Université Paris XI

Ecole Doctorale de Cancérologie

Thèse de Doctorat

Pour obtenir le diplôme de

Docteur en Biologie Cellulaire et Moléculaire

Présentée par

Gwenola MANIC

Impacts du design de vecteurs dérivés du VIH-1 et de la machinerie de réparation de l'ADN sur l'expression lentivirale

---

Thèse présentée et soutenue publiquement le 28 septembre 2012

Devant le jury constitué de

Pr. Christian AUCLAIR	Président du jury
Pr. Javier PEREA	Rapporteur
Dr. Stéphane BASMACIOGULLARI	Rapporteur
Pr. Uriel HAZAN	Examinateur
Dr. Stéphanie BURY-MONE	Examinatrice
Dr. Jean-François MOUSCADET	Directeur de thèse

A ILIO

&

A ma famille

REMERCIEMENTS

Je souhaiterai remercier chacune des personnes qui a contribué, à sa manière, de près ou de loin, au sein du laboratoire ou à l'extérieur, à l'aboutissement de ce travail de thèse.

Je tiens à exprimer ma gratitude aux membres du jury pour l'honneur qu'ils m'ont fait en acceptant d'évaluer mon travail de thèse ainsi que pour l'intérêt qu'ils y ont porté.

Je remercie tout particulièrement le Professeur Christian Auclair qui a accepté de présider le jury de cette thèse et de m'accueillir une seconde fois au LBPA pour que j'y réalise ma thèse.

J'adresse mes sincères remerciements au Professeur Javier Perea et au Docteur Stéphane Basmaciogullari pour leurs commentaires enrichissants dont a bénéficié ce mémoire ainsi que pour leur gentillesse.

Je suis reconnaissante envers le Docteur Jean-François Mouscadet d'avoir accepté de faire partie de mon jury de thèse malgré son départ du laboratoire. Je le remercie également de m'avoir m'accueilli au sein de son équipe malgré ma situation délicate et pour son soutien financier lors de ma première année de thèse.

Je tiens à remercier sincèrement le Docteur Stéphanie Bury-Moné pour son encadrement, sa disponibilité, sa rigueur, son enthousiasme et ses idées, ainsi que pour les nombreuses discussions scientifiques et personnelles que nous avons pu avoir tout au long de ces quatre années de thèse.

Je remercie le Professeur Uriel Hazan pour ses nombreux conseils avisés.

Je tiens à remercier l'AFM d'avoir financé mes travaux de thèse, ainsi que les donateurs de cette association qui contribuent au soutien des jeunes chercheurs.

Je remercie le Docteur Malcom Buckle pour son soutien financier aux projets initiés au sein de l'équipe du Docteur Jean-François Mouscadet.

J'adresse mes remerciements au Docteur Yegor Vassetzky et aux membres du laboratoire « Signalisation, noyaux et innovations en cancérologie » de l'institut Gustave Roussy pour leur accueil et leur gentillesse lors de mon apprentissage de la technique FISH 3D. Je souhaite remercier plus particulièrement le Docteur Andrey Pichugin pour m'avoir former aux bases de cette technique avec autant de gentillesse, passion et de générosité.

Je tiens à remercier chaleureusement le Docteur Aurélie Maurin-Marlin de m'avoir transmis son savoir-faire avec gentillesse et générosité, et de m'avoir appris les rudiments de la vie au P3, et c'est peu dire…

Je remercie Fanny Laurent pour sa gentillesse, son humour et ses délicieux macarons…

Je remercie chacun des membres du laboratoire pour leur accueil, leur contribution à l'aboutissement de ce travail et l'ambiance au quotidien, au P3 ou ailleurs. Merci à Manale, Arzu, Gladys, Chiara, Elisa, Corinne, Xiao-Ju, Marie-Christine… Merci à Hervé pour son humour et ses conseils incisifs. Merci à Fred, Olivier et Alex pour l'ambiance au P3 et leurs nombreux conseils. Merci également à Sylvain pour ses conseils en qPCR. Je souhaite aussi remercier Christine pour ses conseils avisés en Western-blot.

Un grand merci à Aurore et Victoria pour leur amitié et leur soutien…

Je tiens à remercier en particulier Soundasse pour son sourire, sa bonne humeur à toute épreuve, son soutien et ses leçons de courage.

Je remercie Maria pour son soutien, sa gentillesse et ses précieux conseils scientifiques et personnels, qui m'ont tant aidé... Merci également à Aména et Didier.

Je tiens à remercier Lorenzo pour ses talents d'écrivain et ses conseils avisés, qui ont permis la publication de notre travail. Merci également à Nicole.

Je remercie mes amis : Rocco, Camille, Fred, Eugenia, Alfredo, il piccolo Alejandro, Charlène, Adrien, Neri, Sylvia, Saverio, Nath, pour leur humour, leur soutien et les inoubliables moments passés ensemble autour de délices des quatre coins du monde, rue Brillat Savarin ou au bout du monde… En espérant qu'ils puissent se renouveler autant que possible…

Je tiens à remercier de tout cœur mes parents et mes frères, sans qui je n'aurai jamais parcouru ce long chemin. Je les remercie pour leur soutien infaillible. Merci également à Sarah et Carole…

Je remercie Ilio, qui m'a donné la force de débuter et poursuivre ma thèse. Je n'aurai pas de mots suffisamment forts pour lui exprimer combien ma formation scientifique et mon bonheur personnel lui doivent… A nous la *dolce vita*…

Enfin, je remercie celles et ceux que j’aurai oublié de nommer dans ces remerciements, et les prie de m’en excuser.

RESUME

Une meilleure connaissance des déterminants viraux et cellulaires impliqués dans la régulation de l'expression lentivirale est essentielle pour (*i*) maitriser l'expression d'un transgène dans le cadre d'un transfert de gène et (*ii*) mieux comprendre l'absence/la perte de l'expression du virus de l'immunodéficience humaine de type 1 (VIH-1) observé en cas de latence virale après infection. Dans ce contexte, les facteurs cellulaires impliqués dans la reconnaissance et la réparation des cassures de l'ADN, activés dès les étapes précoces du cycle viral, pourraient participer au contrôle de l'expression rétrovirale.

Dans la première partie de cette étude, nous avons généré une collection de vecteurs lentiviraux codant le transgène *green fluorescent protein* (*gfp*) avec des *design* variés : promoteur du VIH-1 [*long terminal repeat* (LTRs)] ou d'origine hétérologue [*e.g.*, promoteur viral CMV (*cytomegalovirus*), ou humain PGK (*phosphoglycerate kinase*)], intégrase native ou mutée, LTRs natifs ou *self inactivating* (SIN). En particulier, nous avons caractérisé l'impact de l'insertion de différentes séquences hétérologues au sein des SIN-LTRs sur l'expression du transgène *gfp* au cours du temps dans un contexte compétent ou déficient pour l'intégration. Nous avons mis en évidence un phénomène de modulation du niveau d'expression du transgène (d'un facteur 0,3 à 1,8; comparé aux vecteurs sans insertion) qui est dépendant de la séquence insérée et de son orientation. L'expression transgénique résulterait donc d'une balance coûts/bénéfices associés à l'insertion d'éléments aux extrémités des vecteurs qui pourrait déterminer le niveau d'expression du transgène à partir de vecteurs modifiés.

Dans la seconde partie, nous avons réalisé une analyse systématique et comparative de l'expression lentivirale au sein des cellules humaines de carcinome de côlon HCT 116 déplétées pour le complexe Ku. Ce complexe, qui est essentiel à la survie chez l'homme et est impliqué dans les processus de réparation de l'ADN, a été identifié comme une cible anti-VIH potentielle. Nous avons mis en évidence que la déplétion en Ku induit une diminution de l'expression précoce du VIH-1 de façon spécifique du LTR et indépendante de Tat. Même si l'action de Ku nécessite l'intégration du VIH-1, sa déplétion ne modifie pas l'efficacité d'intégration virale mais agit plutôt au niveau transcriptionnel. Da façon importante, la réactivation de l'expression du transgène après traitement par l'activateur de NF-κB (*Nuclear Factor κB*), le TNFα (*tumor necrosis factor α*) ou un inhibiteur des déacétylases d'histones, la trichostatine A est favorisée

par la déplétion en Ku. A l'inverse, en présence d'un niveau normal en Ku, les cellules exprimant le VIH-1 seraient contre-sélectionnées dans le temps. Ainsi, la déplétion en Ku pourrait promouvoir l'établissement d'un état (reactivable) de latence transcriptionelle associé à une moindre contre-sélection des cellules transduites.

Les résultats issus de ce travail de thèse démontrent que l'expression lentivirale varie en fonction de nombreux paramètres, dont (*i*) le *design* des vecteurs, (*ii*) le type cellulaire transduit et son fond génétique, et (*iii*) le temps écoulé depuis la transduction, reflétant ainsi des interactions différentielles entre le vecteur et son hôte.

ABSTRACT

An improved knowledge of the viral and cellular determinants implicated in the regulation of the lentiviral gene expression is crucial (*i*) for a better control of transgene expression in strategies designed for gene transfer and (*ii*) for uncovering the mechanisms of viral latency observed after infection with the human immunodeficiency virus type 1 (HIV-1) and accounting for the absence/loss of HIV-1 expression. Cellular factors involved in the mechanism of detecting/repairing DNA lesions are largely activated during the initial steps of viral cycle and, thus, may participate in the control of lentiviral expression.

In the first part of this study, we generated a set of lentiviral vectors encoding for the transgene *green fluorescent protein* (*gfp*) with various *design:* the original promoter [*long terminal repeat* (LTRs)] or of heterologuous origins (*e.g.*, the viral *cytomegalovirus*, CMV or the human *phosphoglycerate kinase*, PGK), wild-type or mutated integrase and wild-type or *self inactivating* (SIN) LTRs. By taking advantage of these constructs, we characterized the impact of the insertion of distinct heterologuous sequences within SIN-LTRs on the expression of *gfp* over the time in conditions of proficiency or deficiency for the integration. We put in evidence a phenomenon of modulation of the level of the transgene expression due to the insertion (by a factor of 0.3 to 1.8, as compared to vectors without insert) that was dependant from the nature and/or the orientation of the insert. We speculate that a balance between the costs and the benefits associated to insertion at the extremities of lentiviral vectors may dictates the level expression of transgene from this engineered construct.

In the second part, we performed a systematic and comparative analysis of the lentiviral expression on human HCT 116 colon carcinoma cells depleted from the complex Ku. This complex, which is essential for the survival in humans and has a described role in DNA repair process, has been previously identified as a potential target against HIV-1. Here, we showed that Ku depletion induced a decrease of the HIV-1 early expression in a fashion that was specific for LTR and independent from Tat. Although Ku action needed HIV-1 integration to host genome, its depletion did not modify the viral integration efficiency but rather acted at transcriptional level. Importantly, the reactivation of transgene expression by administering either the NF-κB (*Nuclear Factor κB*) activator, tumor necrosis factor α (TNFα) or the histone deacetylase inhibitor named trichostatin A was favored in a condition of Ku depletion. On the contrary, in

presence of normal level of Ku, cells expressing HIV-1 displayed a high level of counter-selection over the time. Thus, our observations pleased to favor the hypothesis that Ku depletion promotes the establishment of a state of (reactivable) transcriptional latency associated to a lesser counter-selection of transduced cells.

Altogether, the results obtained during this thesis demonstrate that lentiviral expression vary depending on (*i*) specific vector design, (*ii*) the transduced cell line and its genetic backbone, and (*iii*) the time elapse from transduction, as a consequence of modified interactions between the vector and its host.

TABLE DES MATIERES

TABLE DES MATIERES 1

LISTE DES ABREVIATIONS 5

INTRODUCTION 9

I/ LE VIRUS DE L'IMMUNODEFICIENCE HUMAINE DE TYPE I (VIH-1) 10
A/ GENERALITES 10
1/ Historique : découverte du virus de l'immunodéficience humaine 10
2/ Classification 11
3/ Pathologies associées à l'infection rétrovirale 13
a/ Cancers 13
b/ Autres pathologies 13
c/ Le syndrome d'immunodéficience acquise 14
d/ Rétrovirus non-pathogènes 15
4/ Organisation génétique du VIH-1 16
a/ Génome et protéines du VIH-1 16
b/ Epissages des ARNm du VIH-1 18
c/ Eléments *cis* au sein du génome du VIH-1 19
5/ Structure du VIH-1 20
B/ LES BASES DE LA MULTIPLICATION DU VIH-1 21
1/ Entrée : adsorption et fusion 22
2/ La particule virale au sein du cytoplasme de la cellule hôte 23
a/ Facteurs de restriction 23
b/ Décapsidation, transport/trafic cytoplasmique 24
c/ Transcription inverse 25
d/ Formation du complexe de pré-intégration cytoplasmique 26
3/ La traversée active du cytoplasme au noyau : l'import nucléaire 27
4/ Les étapes nucléaires du cycle de multiplication du VIH-1 28
a/ Formation du complexe de pré-intégration nucléaire 28
b/ Processus d'intégration 28
c/ Sites préférentiels d'intégration du VIH-1 30
d/ Autres espèces d'ADN viral générées 32
5/ Expression 33
6/ Assemblage, bourgeonnement et maturation des particules virales 34
7/ Conséquences de l'infection par le VIH-1 35
a/ Mort des cellules infectées 35
b/ Echappement au système immunitaire des cellules infectées 35
c/ Latence des cellules infectées 36

II/ VECTEURS LENTIVIRAUX : LES LENTIVIRUS COMME VECTEURS DE GENES 36
A/ INTERETS DES RETROVIRUS COMME VECTEURS DE GENES 36
1/ Importance des vecteurs rétroviraux dans le transfert de gènes 36
2/ Propriétés des rétrovirus et conséquences dans le *design* de vecteurs dérivés 37
a/ Avantages liés à l'utilisation des rétrovirus comme vecteurs de gènes 37
b/ Limites liées à l'utilisation des rétrovirus comme vecteurs de gènes 39
B/ LA CONVERSION D'UN LENTIVIRUS EN UN LENTIVECTEUR : DEVELOPPEMENT & *DESIGN* 42
1/ Evolution du concept du *design* de lentivecteurs 42
2/ Production de particules lentivirales 43
a/ Principe 43
b/ Cellules productrices 43
c/ Concentration et stockage 44

d/ Titrations 44
3/ Développement de différentes générations de vecteurs lentiviraux 45
4/ Vecteur de transfert minimal 48
5/ Stratégies pour l'amélioration des vecteurs lentiviraux : contrôle et sécurité de l'expression trangénique 50
a/ Ciblage de l'entrée du virus au sein de la cellule hôte 50
b/ Utilisation de vecteurs déficients pour l'intégration 50
c/ Ciblage de l'intégration du transgène 52
d/ Isoler spatialement la cassette d'expression 52
e/ Restreindre l'expression transgénique !de façon spatio-temporelle 53
f/ Améliorer la terminaison de transcription du transgène 53

III/ L'EXPRESSION LENTIVIRALE ET SA REGULATION 53
A/ ORGANISATION STRUCTURALE ET SPATIALE DE LA CHROMATINE CELLULAIRE 53
1/ Organisation de la chromatine cellulaire 54
a/ Différents niveaux de condensation de la chromatine cellulaire 54
b/ Dynamique de la condensation de la chromatine cellulaire 54
c/ Dynamique de la localisation nucléaire de la chromatine cellulaire 55
2/ Exemple de séquences structurantes de la chromatine : les insulateurs 55
B/ L'EXPRESSION DU VIH-1 57
1/ Organisation du 5'LTR, promoteur du génome viral 57
a/ Le 5'LTR, un promoteur-*enhancer* puissant et très optimisé 58
b/ Organisation nucléosomale du 5'LTR 58
2/ Expression du VIH-1 à partir du 5'LTR et régulations mises en jeu 59
a/ Expression du VIH-1 à partir du 5'LTR intégré 59
b/ Régulation de l'expression du VIH-1 par des structures secondaires de l'ADN et/ou de l'ARN viral 63
3/ Expression à partir de l'ADN viral non-intégré 65
4/ La latence du VIH-1 66
a/ Mise en évidence 66
b/ Définition 66
c/ La latence post-intégrative 67
C/ L'EXPRESSION D'UN TRANSGENE AU SEIN D'UN LENTIVECTEUR MODIFIE 69
1/ Promoteurs eucaryotes et régulation différentielle de la transcription 70
a/ Différents types de promoteurs eucaryotes 70
b/ Régulation épigénétique de l'initiation de la transcription 72
2/ L'expression d'un transgène à partir d'un promoteur interne d'origine humaine 74
a/ La kinase phosphoglycérate 74
b/ Organisation du promoteur PGK 74
c/ Régulation de l'expression à partir du promoteur PGK 74
3/ L'expression d'un transgène à partir d'un promoteur interne d'origine virale 75
a/ Le cytomégalovirus 75
b/ Organisation du promoteur CMV 76
c/ Régulation de l'expression à partir du promoteur CMV 77

IV/ ROLES DE LA MACHINERIE DE REPARATION DE L'ADN LORS DU CYCLE LENTIVIRAL 77
A/ STIMULATION DES VOIES DE REPONSE AUX DOMMAGES DE L'ADN LORS DE L'INFECTION RETROVIRALE 78
1/ Signalisation et réparation des dommages de l'ADN 78
2/ L'infection rétrovirale, un stress exogène qui stimule la DDR 80
a/ Différents niveaux de stimulation de la DDR par l'infection rétrovirale 80
b/ Implication des kinases de la DDR après infection par le VIH-1 81
c/ Phosphorylation de H2AX après infection rétrovirale 82
B/ INTERACTIONS DYNAMIQUES ENTRE CYCLE RETROVIRAL ET MACHINERIE DE REPARATION DE L'ADN 83
1/ Etat actuel des connaissances sur les interactions entre cycle rétroviral et machinerie de réparation de l'ADN 83
2/ Implication des voies de réparation des cassures double-brins dans le cycle du VIH-1 85
a/ Les voies de réparation des cassures double-brin 86
b/ Interactions des voies NHEJ et HR avec les différentes étapes du cycle du VIH-1 87
3/ Implication de la machinerie de réparation dans le cycle lentiviral et remodelage chromatinien 91
C/ ROLES DU COMPLEXE KU DANS LE CYCLE RETROVIRAL 92
1/ Le complexe Ku 92

a/ Identification des protéines Ku70 et Ku80 92
b/ Gènes codant le complexe Ku 92
c/ Caractéristiques de l'hétérodimère Ku 93
2/ Rôles de Ku 96
a/ Réparation 96
b/ Télomères 98
c/ Transcription 98
d/ Interactions de Ku avec les virus 99
3/ Rôles de Ku dans le cycle du VIH-1 99
a/ Etapes pré-intégratives 100
b/ Intégration et localisation 100
c/ Expression 101
d/ Comparaison à un autre senseur des lésions de l'ADN : PARP-1 102

PROBLEMATIQUE 104

RESULTATS 107

I/ IMPACTS DU *DESIGN* DE VECTEURS DERIVES DU VIH-1 SUR L'EXPRESSION LENTIVIRALE 108
A/ Contexte de l'étude 108
B/ Article : '3'Self-Inactivating Long Terminal Repeat Inserts for the Modulation of Transgene Expression from Lentiviral Vectors' (publié dans le journal HUMAN GENE THERAPY Methods, 2012, Apr;23(2):84-97) 110

II/ IMPACTS DE LA MACHINERIE DE REPARATION DE L'ADN SUR L'EXPRESSION LENTIVIRALE 133
A/ Contexte de l'étude 133
B/ Article : 'Impact of the Ku complex on HIV-1 expression and consequence of its depletion in the establisment of viral latency' (soumis au journal RETROVIROLOGY) 135

III/ IMPACT DE LA MACHINERIE DE REPARATION DE L'ADN SUR L'EXPRESSION DE VECTEURS LENTIVIRAUX CONTENANT L'INSULATEUR CHS4 AU SEIN DE LEURS SIN-LTRS 170
A/Contexte de l'étude 170
B/ Résultats (non-publiés) 171
C/ Discussion 173

DISCUSSION ET PERSPECTIVES 175

I/ IMPACTS DES DETERMINANTS VIRAUX SUR L'EXPRESSION LENTIVIRALE 176
A/ Impacts des déterminants viraux sur l'expression du VIH-1 176
1/ Perte limitée des cellules transduites par le VIH-1 en absence de Vpr 176
2/ Etat de latence du VIH-1 des cellules transduites par le VIH-1 en absence de Tat 176
B/ Modulation de l'expression lentivirale lors de l'insertion de séquences hétérologues au sein de vecteurs dérivés du VIH-1 177
1/ Impacts de promoteurs internes sur l'expression d'un transgène intégré ou non 177
a/ Le promoteur CMV induit une expression variable qui tend à être éteinte 177
b/ Le promoteur PGK induit une expression homogène et stable 178
2/Modulation de l'expression transgénique lors de l'insertion de séquences hétérologues au sein des SIN-LTRs 178
a/ L'insertion dans les SIN-LTRs influence peu les étapes pré-intégratives du cycle lentiviral 178
b/ Impacts de l'insertion dans les SIN-LTRs sur l'expression transgénique 179
c/ Hypothèses proposées pour expliquer l'atténuation de l'expression du transgène associée à l'insertion dans les SIN-LTRs 180
3/ Augmentation de l'expression du transgène grâce à l'insertion de cHS4 181

II/ IMPACTS DES DETERMINANTS CELLULAIRES SUR L'EXPRESSION LENTIVIRALE 182
A/ Impacts du statut en Ku sur l'expression du VIH-1 182
1/ Impact positif du statut en Ku sur l'expression du VIH-1 à court-terme..... 182
a/ La déplétion en Ku n'affecte pas les étapes pré-intégratives du cycle viral..... 182
b/ Impact positif du statut de Ku sur la transcription du LTR du VIH-1 183
c/ Régulations mutuelles entre Ku et p53 et impact sur l'expression du VIH-1 184
2/ Impact négatif du statut en Ku sur la latence du VIH-1 185
3/ Hypothèses de modes d'action de Ku..... 186
B/ Impact positif de la déplétion en Ku sur l'expression transgénique à partir d'un vecteur lentiviral insulé 187

CONCLUSIONS..... 189

BIBLIOGRAPHIE 192

LISTE DES ABREVIATIONS

AAV, *Associated-Adenovirus Virus*
ADNc, ADN complémentaire
ADNg, ADN génomique
ADNdb, ADN double-brin
ADNni, ADN non-intégré
ADNv, ADN viral
ALV, *Avian Leukosis Virus*
AP1, *activator protein 1*
APOBEC3G, *apolipoprotein B mRNA-editing catalytic polypeptide 3*
ARNi, ARN interférence
ARNm, ARN messager
ARN pol II, ARN polymérase II
ARNr, ARN ribosomaux
ARNsb, ARN simple -brin
ARNsh, ARN *short hairpin*, ou ARN en forme d'épingle à cheveux courte
ARNsi, ARN *small interfering*, ou petits ARN interférents
ARNsn, ARN *small nuclea*, ou petits ARN nucléaires
ARNv, ARN viral
ASV, *Avian Sarcoma Virus*
ATM, *Ataxia Telangiectasia-Mutated*
ATP, *Adenosine tri-phosphate*
ATR, *ATM and Rad3-related*
Att, *attachment sites*
BAF, *Barrier-to-Autointegration Factor*
Bp, paire de bases
BRCA1, *BReast CAncer 1*
3C, *Chromosome conformation capture*
CA, Capside
CBP, *CREB-Binding Protein*
Cdk9, *Cyclin Dependant Kinase 9*
cf., *confer*, se reporter à
CGI, *CpG island*, ou îlot CpG
ChIP, *Chromatin ImmunoPrecipitation*
CHO, *Chinese Hamster Ovary*
cHS4, *chicken hypersensitive site-4*
CMV, Cytomégalovirus.Abréviation désignant le promoteur *immediate early* du CMV (P_{CMV})
cPPT, C*entral PolyPurine Tract*
CTCF, *CCCTC-binding factor*
CTD, *carboxyl-terminal domain*

CTS, *central termination sequence*
DDR, *DNA damage response*
DIS, Signal de dimérisation de l'ARNg
DNA-PK, *DNA-Protein Kinase*
DNA-PKcs, *DNA-Protein Kinase catalytic subunit*
DNase, Désoxyribonucléase
DNTP, *desoxy A/T/C/G tri-phosphate*
DSB, *Double strand break*, ou cassure double-brin
DSIF, *5,6-dichloro-1-*β*-ribofurosuyllbenzimidazole riboside-sentivity-inducing factor*
e.g., *exempli gratia*, par exemple
EIAV, *equine infectious anemia virus*
Env, Enveloppe
ETS-1, *E26 transformation-specific 1*
FISH, *Fluorescent In Situ Hybridation*
Gag, *group-specific antigen*
GFP, *Green Fluorescent Protein*
GRV, gammarétroviral
HAART, *Highly Active Antiretroviral Therapy*
HAT, Histone Acétyl-Transférase
HCMV, Cytomégalovirus humain
HCT, *Human Carcinoma Tissue*
HDAC, Histone DésACétylase
HERV, *Human Endogenous RetroVirus*
HMGA1/2, *High mobility group AT-hook 1/2*
HR, *Homologous Recombinaison*
HS, *hypersensitive sites*
HTLV-1, *Human T-cell Leukemia Virus* 1
IBD, *Integrase Binding Domain*
IDLVs, vecteurs lentiviraux intégrase-déficients
i.e., *id*, c'est-à-dire
IE, *immediate early*
IκBα, *inhibitor of NF-κB α-subunit*
IN, Intégrase
IN1, *INtegrase Interactor* 1 (ou hSNF5)
INR, initiateur de la transcription,
IRES, *Internal Ribosome Entry Site*
KARP-1, *Ku86 Autoantigen Related Protein*-1
Kb, kilobases
KDa, kilodaltons
KBS, *Ku Binding Site*
LEDGF/p75, *Lens Epithelium-Derived Growth Factor p75 splice variant*

LSF, *late SV40 factor*
LTR, *Long terminal repeat*
LV, lentiviral
m.o.i., *multiplicity of infection*
MA, Matrice
MFI, *Mean Fluorescence Intensity* ou intensité (moyenne) de fluorescence
(Mo-)MLV, (*Moloney*) *Murine Leukemia Virus*
NC, Nucléocapside
NELF, *Negative Elongation Factor*
Nef, *Negative factor*
NFAT, *Nuclear Factor of Activated T-cells*
NF-κB, *Nuclear Factor κB*
NHEJ, *Non Homologous End Joining*
NLS, *Nuclear Localization Sequence*
Nm, nanomètres
NRE1, *Negative Regulatory Element 1*
Nt, nucléotides
N-TEF, *Negative transcription Elongation Factor*
OCT-1, *Octamere motif 1*
P53, *tumor suppressor p53*
PARP-1, *Polyadenosine diphosphate(ADP)-ribose polymerase-1*
Pb, paires de bases
PBS, *primer binding site*
PCAF, *p300/CREB-binding protein (CBP)-associated factor*
PFV, Prototype/Primate Foamy Virus
PGK, *PhosphoGlycerate Kinase.* Abréviation désignant le promoteur du gène *pgk* (P_{PGK})
PIC, *PreIntegration Complex* ou complexe de pré-intégration
Poly(A), polyadénylation
PPT, *PolyPurine Tract*
PR, Protéase
P-TEFb, *Positive Transcription Elongation Factor B*
Q-PCR, *Quantitative Polymerase Chain Reaction*
Rev, *regulator of virion*
RRE, *Rev responsive element*
RSV, *Rous sarcoma virus*
RT, *reverse transcriptase*, ou transcriptase inverse
RTC, *reverse transcription complex*, ou complexe de transcription inverse
RT-Q-PCR, *Real-time*-Q-PCR
SA/SD, Site accepteur/donneur d'épissage
SCID, *Severe Combined Immune Deficiency*
SIDA, Syndrome d'ImmunoDéficience Acquise

SIN-LTRs, *Self Inactivating* LTR
SNF-5, *Sucrose Non Fermenting*
SP-1, *SV40-promoter specific factor*
SIV, *simian immunodeficiency virus*, du *sooty mangabey* (sm) ou *chimpanzee* (czp)
TAR, *TransActivation Responsive element*
Tat, *transactivator of transcription*
TFII(D, H), *transcription factor II (D, H)*
TNFα, *Tumor Necrosis Factor*
TBP, *TATA-binding protein*
TRIM5α, *TRIpartite Motif protein* 5α
TSA, Trichostatine A
TSS, *transcription site start*, ou site d'initiation de la transcription
UCOEs, *ubiquitously acting chromatin opening elements*
UNG2, *Uracil DNA Glycosylase 2*
USEs, *upstream enhancer element*
VDJ, Variable, Diversité, Jonction
VIH-1/2, Virus de l'immunnodéficience humaine de type 1/2
Vif, *Virion infectivity factor*
Vpr, *viral protein R*
vs., *versus*
VSV, virus de la Stomatite Vésiculeuse
VSV-G, glycoprotéine G d'enveloppe du VSV
WRN, WeRNer syndrom
WT, *Wild Type*
XLF, *XRCC4 Like Factor*
XRCC4/5/6, *X-Ray Cross-Complementing group 4/5/6*
XMRV, *xenotropic murine leukemia virus-related virus*
YY1, *yin and yang 1*

INTRODUCTION

I/ LE VIRUS DE L'IMMUNODEFICIENCE HUMAINE DE TYPE I

A/ GENERALITES

1/ Historique : découverte du virus de l'immunodéficience humaine

Le syndrome d'immunodéficience acquise (SIDA) fut observé pour la première fois en juin et juillet 1981 dans deux rapports successifs du *Centers for Disease Control and Prevention* (CDC). Ceux-ci décrivaient des groupes de jeunes hommes homosexuels atteints d'infections opportunistes (*e.g.*, pneumonie à *Pneumocystis jiroveci*) et de sarcomes de Kaposi à New York et en Californie. Cette pathologie était caractérisée par une profonde déplétion des cellules de l'immunité associée à des affections jusque-là observées chez des patients immunodéprimés. Des cas similaires furent alors décrits en Europe de l'ouest ainsi que chez des patients hémophiles et/ou transfusés par voie sanguine. Cette maladie pouvait ainsi toucher des personnes jusque-là saines tels des homosexuels de sexe masculin et des toxicomanes (par voie intraveineuse) et impliquait une transmission sanguine ou sexuelle, ce qui permit de suspecter l'implication d'un agent infectieux de type herpes virus (*i.e.*, cytomégalovirus) ou rétrovirus. Le candidat putatif était alors le HTLV (*Human T cell Leukemia Virus*) décrit par le Dr. Gallo et des chercheurs japonais (Gallo et al., 1983).

L'agent causal du SIDA fut découvert en 1983 par le groupe de Luc Montagnier de l'Institut Pasteur (Barre-Sinoussi et al., 1983), récompensé avec Françoise Barré-Sinoussi par le Prix Nobel de Médecine et Physiologie en 2008. Cependant, ce nouveau rétrovirus, alors appelé *LymphAdenopathy Virus* (LAV), ne sera officiellement reconnu être la cause du SIDA qu'en 1984 après la description du même virus par d'autres laboratoires internationaux (Gallo et al., 1984; Levy et al., 1984; Popovic et al., 1984) prenant le nom de virus de l'immunodéficience humaine (VIH) en 1986 (Coffin et al., 1986).

Il existe au moins deux types de VIH de type 1 (VIH-1) et 2 (VIH-2) qui diffèrent de part leur séquence codante [*e.g.*, Vpr (*viral protein R*) et Vpx (*viral protein X*)] et leur origine. VIH-1 est la cause prédominante de l'épidémie de SIDA et a évolué à partir d'un virus, le *Simian immunodeficiency virus* (SIV), qui a franchi la barrière des espèces des chimpanzés (*Pan troglodytes*, SIVczp) à l'homme (Gao et al., 1999). VIH-2, découvert en 1986 (Clavel et al., 1986), est retrouvé principalement en Afrique de l'Ouest et provient du SIV du singe mangabey (dit enfumé, *Cercocebus atys*, du *sooty mangabey*, SIVsm) (Hirsch et al., 1989).

L'isolement du virus causal du SIDA a permis de mettre en œuvre des mesures de prévention rationnelles ainsi que de débuter la recherche d'inhibiteurs viraux efficaces. La

première molécule anti-rétrovirale découverte fut l'azidothymidine (AZT, ou zidovudine) qui est un inhibiteur efficace de la transcription inverse *in vitro* (Nakashima et al., 1986). Cependant, le traitement n'avait cependant pas de bénéfice chez le patient à long terme du fait de l'apparition de résistances [*e.g.*, chez 30 à 80 % des patients traités (Clumeck, 1993)]. Seule une combinaison d'au moins trois inhibiteurs (qui ciblent plusieurs étapes du cycle viral, *cf.*, § I.B), appelée thérapie anti-rétrovirale HAART (*Highly Active Antiretroviral Therapy*), permet depuis 1996 de traiter les patients infectés par le VIH en empêchant leur immunodépression et les infections opportunistes consécutives souvent mortelles (Moir et al., 2011). La HAART permet de maintenir la virémie à un taux faible (*cf.*, § I.A.3.c) mais les malades traités ne guérissent pas : le virus n'est pas éliminé car il est peu/pas exprimé au sein de réservoirs cellulaires(Moir et al., 2011).

Les dernières données épidémiologiques sur l'épidémie de SIDA dans le monde fournies par ONUSIDA estimaient à 33,3 millions de personnes vivant avec le VIH en 2009 dont 2,6 millions nouvellement infectées cette même année. Grâce à une plus large couverture de l'HAART dans le monde (principalement dans les pays développés), le nombre de morts lié à la maladie est en diminution, s'élevant encore malgré tout à 1,8 millions de personnes en 2009. La zone la plus touchée par l'épidémie (en termes de prévalence, nouvelles infections et de mortalité) est l'Afrique sub-saharienne, principalement en Afrique du Sud et au Botswana (Source : http://www.unaids.org/globalreport/Global_report_fr.htm, de ONUSIDA).

2/ <u>Classification</u>

La famille des *Retroviridae* regroupent des virus enveloppés infectant les Vertébrés (groupe VI de la classification de Baltimore). Les rétrovirus possèdent un génome diploïde composé d'ARN simple brin positif non fragmenté qui est coiffé et polyadénylé pouvant mesurer de 7 (*Avian Leukosis Virus*, ALV) à 12,5 (*Walleye Dermal Sarcoma Virus*) kilobases (kb). Leur nom, qui débute par *Retro* signifiant « à reculons », se réfère au fait que le cycle de réplication des rétrovirus est composé des étapes de transcription inverse et d'intégration de l'ADN complémentaire (ADNc) double brin formé lors de la transcription inverse. Ces étapes sont respectivement catalysées par les protéines virales, transcriptase inverse (*reverse transcriptase*, RT) et intégrase (IN), contenues dans la particule virale (Vogt, 1997).

Les *Retroviridae* sont divisés en deux sous-familles et sept genres, comme indiqué dans le **tableau 1**, selon l'homologie de séquence de leur reverse transcriptase, ainsi que leur morphologie (forme de la capside), le site d'assemblage de la capside (membrane plasmique

ou cytoplasme), et la présence/absence de gènes auxiliaires qui distingue les rétrovirus simples des rétrovirus complexes.

Tableau 1 : Classification de La famille des *Retroviridae* (Source : http://talk.ictvonline.org/files/ictv_documents/m/msl/1231.aspx, du *International Committee on Taxonomy of Viruses*, ICTV).

Sous-familles	Genres	Prototype	Espèces à tropisme humain	Structure du génome
Orthoretrovirinae	*Alpharetrovirus*	*Avian Leukosis Virus* (ALV)		SIMPLES
	Betaretrovirus	*Mouse Mammary Tumor Virus*		
	Gammaretrovirus	*Murine Leukemia Virus* (MLV)	*Xenotropic Murine leukemia virus-Related Virus* (in vitro)	
	Epsilonretrovirus	*Walleye Dermal Sarcoma Virus*		COMPLEXES
	Deltaretrovirus	*Bovine Leukemia Virus*	*Human T cell Lymphotropic/ Leukemia/ Lymphoma Virus* HTLV-1, -2, -3 et -4	
	Lentivirus	*Human Immunodeficiency Virus type-1* (HIV-1)	HIV-1 et HIV-2	
Spumaretrovirinae	*Spumavirus*	*Simian Foamy Virus*	*Prototype/Primate foamy virus*	

L'organisation génétique classique du génome des rétrovirus simples comprend les gènes (*group-specific antigen*), *pro*/*pol* (*polymerase*) et *env* (*envelop*) ordonnés successivement de façon invariante. Les rétrovirus complexes possèdent en plus un ou plusieurs gènes auxiliaires codant des protéines dites régulatrices et/ou accessoires quand elles sont dispensables à la multiplication du virus *in vitro*. Ces protéines sont impliquées dans la régulation de la synthèse et maturation des ARN messagers (ARNm) ainsi que l'interaction avec la cellule hôte (Vogt, 1997).

VIH-1 et VIH-2 sont les seuls lentivirus à tropisme humain (**Tableau 1**). Ils ont des structures génétiques similaires mais peuvent avoir des variations de séquences importantes. A titre d'exemple, deux souches de VIH-1 et VIH-2 [comparées dans (Motomura et al., 2008)] contiennent seulement 55 % d'homologie de séquence pour le génome viral (dont respectivement 54 %, 55 %, et 35 % pour leurs gènes *gag*, *pol* et *env*). Le type VIH-1 peut être divisé en trois groupes majeurs, qui reflètent des événements indépendants de transmission à partir de primates (non-humains) (Gao et al., 1999; Keele et al., 2006; Korber et al., 2000). Le groupe M (*major*) a une distribution mondiale alors que les groupes O

(*outlier*) et N (non-M/non-O) sont localisés en Afrique de l'Ouest et centrale, proche de l'habitat naturel du chimpanzé *Pan troglodytes* (Keele et al., 2006; Peeters et al., 1997).

Huit pour cent du génome humain provient de rétrovirus endogènes (*Human Endogenous RetroVirus*, HERVs), le plus souvent dégénérés. Les HERVs sont répartis en 100 familles divisées en 3 classes : classe I, II ou III lorsqu'ils sont, respectivement, proches des gamma-rétrovirus, beta-rétrovirus, ou spumavirus (Bannert and Kurth, 2004). Leurs séquences ont contribué à l'évolution, mais sont source d'instabilité génétique et peuvent ainsi avoir un impact (positif ou négatif) sur la physiologie cellulaire [*e.g.*, placentation, inflammation (Kurth and Bannert, 2010)].

3/ Pathologies associées à l'infection rétrovirale

a/ *Cancers*

Les rétrovirus furent découverts dès 1908 du fait de leur capacité oncogénique [*i.e.*, ALV, identifié comme responsable de leucémie chez le poulet (Ellerman et Bang, 1908)]. De plus, du fait de l'origine cellulaire des oncogènes viraux, les rétrovirus ont permis la découverte de nombreux proto-oncogènes cellulaires [*i.e.*, le RSV (*Rous sarcoma virus*), après transduction de l'oncogène src (Rous, 1911)].

Cinq genres de rétrovirus, les alpha-, beta-, delta-, epsilon- et gamma-rétrovirus, ont été décrits susceptibles de provoquer des cancers chez les organismes qu'ils infectent en utilisant différents mécanismes moléculaires. Le processus de transformation cellulaire induit par l'infection rétrovirale peut être rapide et polyclonale (*e.g.*, sarcomes) lorsqu'il est du à la transduction d'un oncogène viral [*i.e.*, rétrovirus (défectifs) transducteurs; *e.g.*, *mos* par le Mo-MSV (*Moloney Murine Sarcoma Virus*), aidé par le Mo-MLV (*Moloney Murine Leukemia Virus*) (Moloney, 1966)]. La tumorigénèse peut aussi être lente et clonale lorsqu'elle résulte d'une transformation cellulaire indirecte après expression de protéines virales [*i.e.*, rétrovirus *trans*-activateurs; *e.g.*, Tax de HTLV-1 (Tanaka et al., 1990)] ou mutagénèse insertionnelle [*i.e.*, rétrovirus *cis*-activateurs ; *e.g.*, *c-myc* (Steffen, 1984)] au sein de cellules hématopoïétiques (*e.g.*, leucémies, lymphomes).

b/ *Autres pathologies*

Un certain nombre de rétrovirus, même simples, peuvent induire après infection des maladies du système nerveux central [*i.e.*, perte de la fonction neuronale; *e.g.*, encéphalopathie spongiforme due au HTLV-1 (Gessain and Gout, 1992)] ou être la cause

d'immunodéficiences (différentes du SIDA) [*e.g.*, atrophie du thymus chez la souris après infection par Mo-MLV (Wong, 1990)].

Les lentivirus ont été isolés au sein de nombreuses espèces animales (*e.g.*, moutons, chevaux, singes) et, comme le préfixe « lenti » l'indique, induisent des pathologies à évolution lente associées à une longue période d'incubation, à une réplication virale persistante, des troubles neurologiques et la destruction des cellules hématopoïétiques ou immunologiques, entrainant indirectement la croissance de néoplasmes [*i.e.*, lymphomes, sarcome de Kaposi; *e.g.*, *equine infectious anemia virus* (Reagan et al., 1950)].

c/ *Le syndrome d'immunodéficience acquise*

VIH-1 et VIH-2 sont les seuls lentivirus à tropisme humain et causent tous deux la maladie SIDA (avec des symptômes similaires) chez les personnes qu'ils infectent. VIH-1 est hautement pathogène uniquement chez les humains (ce qui n'est pas le cas de son précurseur, le SIVcpz, chez le chimpanzé) comparé à VIH-2, moins virulent, qui dérive du SIVsm (non-pathogène chez le singe mangabey).

L'infection par le VIH-1 est caractérisée (dans les 7 à 21 jours qui suivent) par une phase précoce d'« éclipse » (appelée syndrome VIH aigu) de réplication intense du virus, principalement dans les lymphocytes T $CD4^+$, en absence de réponse immunologique (virémie plasmatique > 10 millions de copies/mL) associée à une déplétion importante de lymphocytes T $CD4^+$ dans le sang périphérique. Le virus est disséminé dans les tissus lymphoïdes (*i.e.*, associés à l'appareil digestif et ganglions lymphatiques) où sont notamment constitués des réservoirs viraux (*e.g.*, lymphocytes T $CD4^+$ au repos). Dans les mois qui suivent l'infection, commence une phase chronique, souvent asymptomatique, d'activation immunitaire soutenue et de réplication virale persistante associée à une virémie en baisse qui devient stable, atteignant les 10 000 copies/mL en absence de traitement. Après plusieurs années (3 à 15 ans), la phase avancée de l'infection par le VIH est caractérisée par une déplétion marquée en lymphocytes T $CD4^+$ qui mène au développement des maladies opportunistes caractéristiques du SIDA (*i.e.*, infections et/ou cancers) [*cf.*, revue (Moir et al., 2011)].

Depuis le développement de la HAART, une prise en charge rapide des personnes infectées par le VIH a considérablement réduit l'incidence des maladies opportunistes caractéristiques du SIDA. Le traitement antirétroviral permet de contrôler la réplication virale et de maintenir la virémie sous le seuil de détection à l'aide des méthodes standards (~1

copies/mL) après un an. Celle-ci chute en quatre phases qui correspondent à la perte de la production virale à partir de différents réservoirs cellulaires aux durées de vie variables (Moir et al., 2011). Cependant, le principal réservoir viral, les lymphocytes T $CD4^+$ quiescents, a une durée de vie illimitée sans activation cellulaire et représente actuellement un obstacle à l'éradication du VIH-1 chez un patient traité par HAART (Pace et al., 2011). De plus, les personnes infectées par le VIH-1 et traitées sur le long terme avec une HAART peuvent être confrontées aux effets secondaires du traitement et/ou de l'infection, tel le risque de développer des maladies cardiovasculaires (Moir et al., 2011). Il est à noter qu'un effet secondaire majeur de la HAART est l'échappement viral caractérisé par l'émergence de virus résistants qui induit un échec thérapeutique chez 7 à 15 % des patients traités (Cortez and Maldarelli, 2011).

Le tableau clinique décrit ci-dessus est le plus fréquent (*i.e.*, 70-80 % des cas). Cependant, l'évolution de l'infection par le VIH-1 et la progression de la maladie SIDA peuvent être influencées par des facteurs viraux et cellulaires. En particulier, le système immunitaire de l'individu infecté permet de ralentir ou, au contraire, d'augmenter l'évolution de l'infection. Il existe ainsi des individus infectés par le VIH-1 qui sont (*i*) des progresseurs rapides ou (*ii*) des non-progresseurs à long-terme, dont (*iii*) des résistants (*elite controllers*), en fonction de leurs caractéristiques génétiques et de l'évolution de l'infection. Les progresseurs rapides représentent 10 % des personnes infectées par le VIH-1 alors que 5-15 % appartiennent à la catégorie des non-progresseurs à long-terme (Sharma et al., 2011). Seul 1 % des individus sont résistants à l'infection par le VIH-1 grâce à différents mécanismes (O'Connell et al., 2009). Ils peuvent, par exemple, possèder un haplotype des HLA (*human leukocyte antigen*) de classe I (*e.g.*, HLA-B57) qui favorise la réponse des lymphocytes T cytotoxiques (en augmentant la présentation des épitopes du VIH-1) ou être hétérozygotes pour le co-récepteur CCR5 muté (CCR5Δ32) (qui rend les cellules cibles réfractaires à l'infection) (O'Connell et al., 2009). L'importance de cette mutation de CCR5 a été très illustrée avec l'annonce d'un premier patient très probablement « guéri » de son infection au VIH après transplantation de cellules souches CCR5Δ32/Δ32 (Allers et al., 2011).

d/ *Rétrovirus non-pathogènes*

Les spumavirus n'ont pas de rôle pathogène connu, quelque soit l'espèce de l'hôte infecté. Il est à noter qu'il n'existe pas de spumavirus humain. L'homme peut cependant être infecté accidentellement par un spumavirus de primate (non-humain), appelé

prototype/primate foamy virus (PFV), sans pour autant qu'une transmission horizontale soit possible (Schweizer et al., 1997).

Il est à noter la description récente du XMRV (*xenotropic murine leukemia virus-related virus*) initialement caractérisé comme un nouveau gammarétrovirus à tropisme humain au sein de tumeurs de la prostate (Urisman et al., 2006). Il s'agit en réalité d'un virus recombinant provenant d'une xénogreffe de cellules humaines tumorales de prostate au sein de souris de laboratoire (Paprotka et al., 2011). Ce phénomène de recombinaison ne serait pas rare puisqu'il a été évalué pouvant survenir avec une fréquence de 25 % (Zhang et al., 2011).

4/ Organisation génétique du VIH-1

a/ *Génome et protéines du VIH-1*

Une fois intégré au sein de la chromatine hôte, le génome du VIH-1 mesure environ 9,7 kb, est flanqué de LTRs (*long terminal repeats*) à chaque extrémité (**Figure 1A**) et permet de coder 17 protéines différentes (**Figure 1B**, **Tableau 2**).

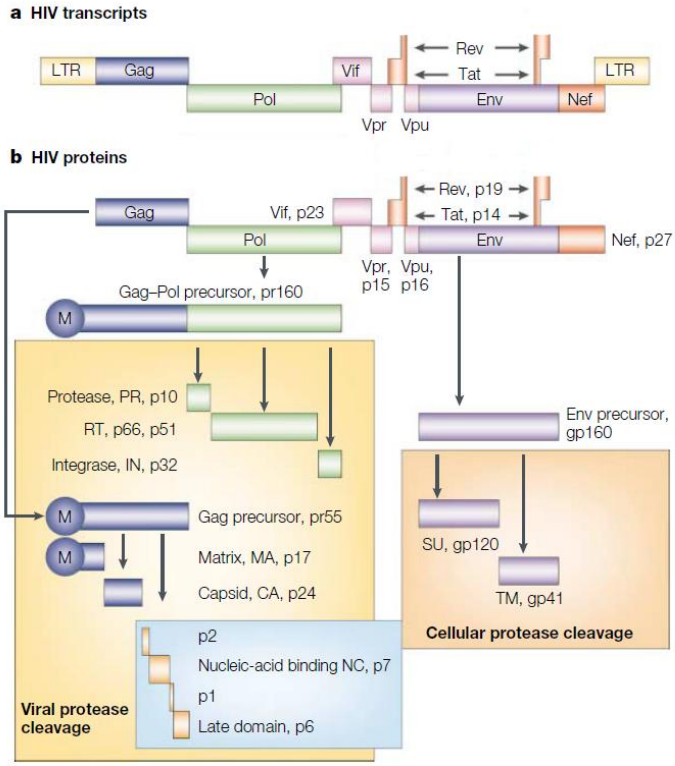


Figure 1 : Génome du VIH-1 sous forme de provirus et protéines synthétisées. (**A**) Le VIH-1 sous forme provirale est bordé des LTRs et contient les gènes *gag*, *pol* et *env*. (**B**) Ils codent les protéines structurales (MA, CA, p2, p7, p1, p6), enzymatiques (PR, RT, IN) et d'enveloppe (SU, TM) après clivage de polyprotéines (Gag, Gag-Pol ou Env) médié par des protéases virales ou cellulaires, comme indiqué. (**A-B**) Les gènes *tat*, *rev*, *vif*, *vpr*, *vpu*, *nef* codent les protéines Tat, Rev, Vif, Vpr, Vpu, Nef [adaptée de (Peterlin and Trono, 2003)].

Les protéines structurales, enzymatiques et d'enveloppe, communes à tous les rétrovirus, sont issues du clivage de polyprotéines codées par les gènes *gag*, *pro*/*pol* et *env*, respectivement. Le génome du VIH-1 contient deux gènes supplémentaires, *tat* et *rev*, codant chacun une protéine régulatrice pouvant se fixer aux ARNs: Tat (*Transactivator of transcription*) et Rev (*Regulator of virion gene expression*), indispensables pour obtenir de hauts niveaux de réplication virale. Enfin, le VIH-1 contient 4 gènes codant les protéines dites accessoires: Vpr (Vpx pour le VIH-2, *Viral protein R/X*), Vpu (*Viral protein U*), Vif (*Virion infectivity factor*) et Nef (*Negative factor*) [**Figure 1**, (Frankel and Young, 1998)].

Tableau 2 : Origine et fonctions des protéines du VIH-1 [adapté de (Frankel and Young, 1998; Pluta and Kacprzak, 2009)].

Gène/Origine	Protéine	Fonction(s)
Protéines structurales		
gag/clivage de la polyprotéine Gag (Pr55) médiée par PR pendant la maturation du virion	Matrice (MA ; p17)	Formation et import nucléaire du PIC (*PreIntegration Complex*), export du complexe Gag-ARN dans le cytosol, attachement à la membrane, formation des virions (oligomérisation de Gag et recrutement de Gag-Pol/Pr160 et Env)
	Capside (CA ; p24)	Attachement à la membrane, formation de la capside
	Nucléocapside (NC ; p7)	Attachement à la membrane, formation de la capside et du PIC, attachement à l'ARN, co-facteur de la RT, sélection de l'ARNg, packaging et dimérisation pendant l'assemblage des virions.
	P1 et P2	Régulation du taux de clivage, attachement à la membrane et formation des virions
	p6	Recrutement de Vpr au virion, motif PTAP impliqué dans les interactions avec les protéines de l'hôte, attachement à la membrane et formation des virions, interaction avec la machinerie ESCRT (*Endosomal Sorting Complex Required for Transport*) qui facilite le bourgeonnement des particules virales immatures
Enzymes		
gag-pol/clivage de la polyprotéine Gag-Pol (Pr160) médiée par PR pendant la maturation du virion	Transcriptase inverse et RNAase H (RT hétérodimère p66/p51)	Synthèse de l'ADNc viral et dégradation de l'ARN matrice
	Protéase (PR ; p10)	Clivage de Pr160 et Pr55, maturation du virion
	Intégrase (IN ; p32)	Import nucléaire du PIC, intégration de l'ADNc viral dans le génome hôte
Glycoprotéines d'enveloppe		
env/clivage de la polyprotéine Env (gp160) par des protéases cellulaires	Gp120 (SU)	Interaction avec le récepteur de la cellule hôte CD4 et co-récepteurs
	Gp41 (TM)	Ancrage du complexe Gp120/Gp41 dans la membrane virale, fusion membranaire avec la cellule hôte
Protéines régulatrices		

tat	Transactivateur (Tat ; p14)	Initiation de la transactivation à partir du LTR viral, élongation de la transcription, remodelage de la chromatine, régulation de l'apoptose, modulateur de la réponse immune (dont la machinerie d'ARN interférence)
rev	Régulateur de l'expression des protéines virales (Rev ; p19)	Stimulation de l'export nucléaire des ARNs du VIH non-et mono-épissés, stimulation de la traduction des ARNs du VIH non-épissés
Protéines accessoires		
nef	*Negative factor* (Nef ; p27)	Diminution de la régulation du récepteur CD4, diminution de la régulation du CMH-I et -II, interférence du signal de transduction, augmentation de l'infectivité (réarrangement du cytosquelette d'actine), impliquée dans la formation de la synapse virologique
vif	*Viral infectivity factor* (Vif ; p23)	Inactivation de la réponse immunitaire de l'hôte par fixation de l'apolipoprotéine B (APOBEC3) et de l'IRF3 (*interferon regulatory factor*)
vpr	*Viral protein R* (Vpr ; p15)	Import nucléaire du PIC, contrôle de la fidélité de la transcription inverse, arrêt du cycle cellulaire (transition G2/M), induction de l'apoptose, transactivation du LTR du HIV et des gènes cellulaires, modulation de la réponse immunitaire de l'hôte, inhibition de l'épissage des pré-ARNm viraux et cellulaires, stimulation de la signalisation des voies de réparation des dommages de l'ADN
vpu	*Viral protein U* (Vpu ; p16)	Dégradation du récepteur CD4, relarguage des particules virales, régulation de l'apoptose

b/ *Epissages des ARNm du VIH-1*

L'ensemble des protéines du VIH-1 est codé par plus de 40 ARNm synthétisés à partir d'épissages alternatifs grâce à la présence de plusieurs sites donneurs/accepteurs d'épissage (**Figure 2**). Chacune des protéines du VIH-1 peut donc être produite à partir d'un ou plusieurs transcrits (**Figure 2**). Les transcrits codant les protéines précoces, le plus souvent régulatrices (Tat, Rev et Nef) sont tous épissés, alors que ceux codant les protéines tardives, composants structuraux et enzymatiques du virion et facteurs permettant d'augmenter l'infectivité, sont mono-épissés ou ne le sont pas (**Figure 2**). Rev permet notamment de réguler la transition entre les phases précoces et tardives de l'expression rétrovirale en favorisant le transport d'ARNm non-épissés ou mono-épissés [(contenant le RRE (*Rev response element*)] du noyau vers le cytoplasme [**Figure 2**, (Stoltzfus and Madsen, 2006)].

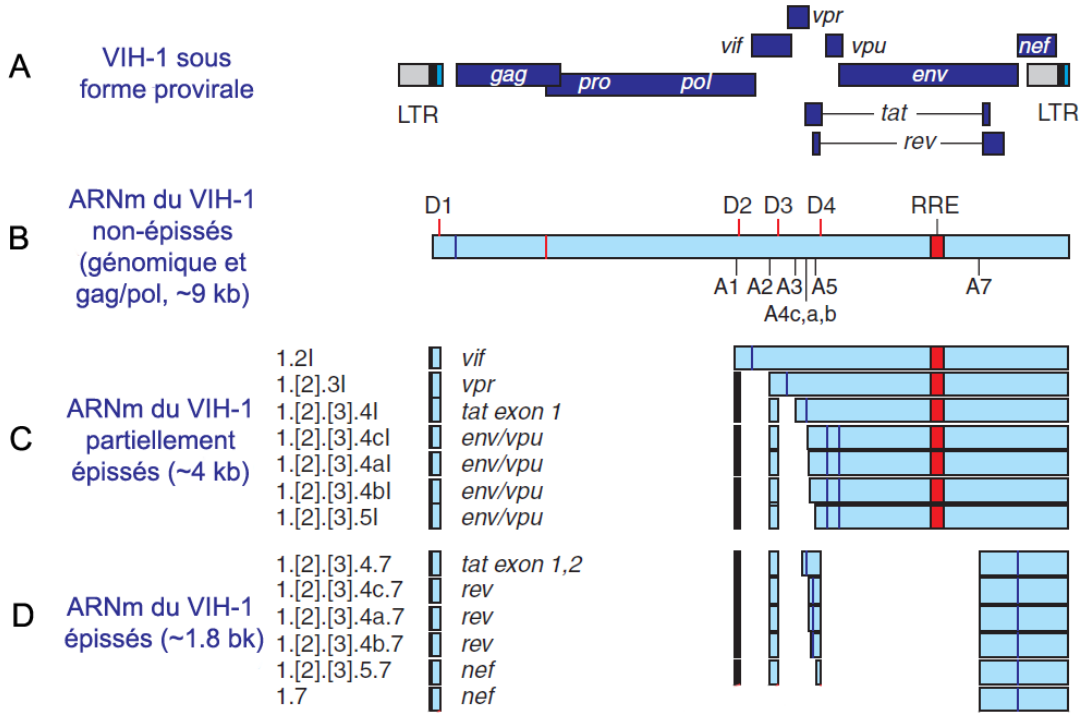


Figure 2 : Génome du VIH-1 sous forme provirale et les ARNm synthétisés. (**A**) Représentation schématique du génome du VIH-1. (**B-D**) Les différents ARNm synthétisés lors de la transcription du LTR sont représentés sous forme (**B**) non-épissée, (C) partiellement ou (**D**) totalement épissée. Il est à noter la présence de différents sites donneurs (D1-4) et accepteurs (A1-7) d'épissage sur la forme non-épissée de l'ARNm qui permettent de générer de nombreux ARNm partiellement ou totalement épissés. Certains exemples d'ARNm partiellement ou totalement épissés sont représentés ici (les sites D/A ainsi que les protéines pour lesquels ces ARNm codent sont précisés) [adaptée de (Karn and Stoltzfus, 2012)].

c/ *Eléments* cis *au sein du génome du VIH-1*

Les éléments *cis* situés au sein du génome du VIH-1 (qui se superposent souvent avec des séquences codantes) optimisent le transfert du génome viral (transcription inverse, intégration, transcription, épissage) une fois la cellule hôte infectée ou sont impliqués dans la production efficace de nouvelles particules virales infectieuses (**Figure 3**).

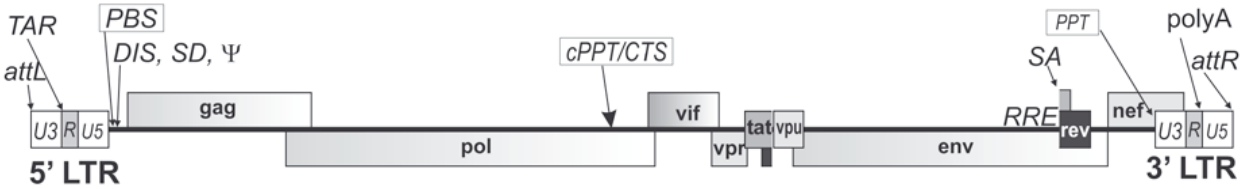


Figure 3 : Eléments *cis* présents au sein du génome du VIH-1. Les éléments *cis* du génome du VIH-1 sont impliqués dans le transfert du génome viral, soient la transcription inverse (PBS, cPPT/CTS, PPT), l'intégration (attL, attR), la transcription (U3, TAR, polyA) et l'épissage (SA, SD) ainsi que dans la production de virions (RRE, DIS, ψ) [issue de (Pluta and Kacprzak, 2009)]. Se reporter au **tableau 3** pour les abréviations.

Les emplacements et fonctions des éléments *cis* du génome du VIH-1 sont détaillés dans le **tableau 3**.

Tableau 3 : Emplacement et fonctions des éléments *cis* inclus dans le génome du VIH-1 sous forme proviral (décrits du LTR 5' au 3').

Eléments *cis*	Fonctions
LTR (*long terminal repeat*)	
attL/R (*attachment sites Left/Right*)	Sites d'attachement pour l'intégration.
U3	Séquence unique non codante contenant les éléments promoteurs et amplificateurs de la transcription du provirus (lorsqu'elle est placée en 5', une fois dupliquée au sein du provirus)
R	Répétition directe à chaque extrémité qui contient un signal de polyadénylation (polyA, arrêt de la transcription en 3') et l'élément TAR.
U5	Région unique, non codante, riche en G/U qui est la première à être rétro-trancrite et permet une reconnaissance efficace du signal polyA contenu dans R (lorsqu'elle est placée en 3', une fois dupliquée au sein du provirus).
5' *gag*	
PBS (*primer binding site*)	Complémentaire de la séquence de l'ARNt. Permet l'initiation de la transcription inverse.
DIS	Signal de dimérisation de l'ARNg
SD	Site donneur d'épissage
Ψ (Psi)	Ensemble de boucles d'ARN servant de signal d'encapsidation de l'ARNg viral. Tête (*leader*) région transcrite non traduite présente à l'extrémité 5' de tous les ARNm viraux.
cPPT (*pol*) PPT (*nef*) CTS (*pol*)	La séquence appelée *central polypurine tract-central termination sequence* (cPPT-CTS) est présente presque exclusivement au sein du génome des lentivirus. Le CTS, localisé en aval du cPPT, est le site auquel la synthèse du brin + d'ADN viral se termine. La présence de cPPT-CTS détermine une discontinuité dans le brin + complet d'ADN viral qui forme une triple hélice appelée DNA flap (cf., § I.6.c) et exerce un effet positif sur l'infectivité virale (Follenzi et al., 2000 ; Sirven et al., 2000), peut-être en favorisant la rétro-transcription
RRE (*env*)	Elément de réponse à Rev qui permet l'export nucléaire des ARNm partiellement ou non-épissés
SA (*rev*)	Site accepteur d'épissage

5/ <u>Structure du VIH-1</u>

Les particules virales du VIH-1 ont un diamètre de 80 à 120 nanomètres (nm) (**Figure 4**). Elles sont recouvertes d'une membrane lipidique contenant des complexes glycoprotéiques constitués de trimères de protéines de surface (SU) Gp120 et transmembranaire (TM) Gp41 liés de façon non-covalente (**Figure 4A**, **B**). Pendant le bourgeonnement des virions à partir de la cellule infectée, le virus peut aussi incorporer des protéines cellulaires (*e.g.*, HLA de classe I et II, cyclophyline A, protéines d'adhésion) (**Figure 4A**, **B**). Une matrice constituée des protéines (MA) p17 est ancrée à la surface interne de l'enveloppe virale (**Figure 4A**, **B**). A l'intérieur, se trouve la capside en forme de trapèze constituée de l'antigène (CA) p24

(**Figure 4B**). Cette dernière contient les deux brins d'ARNg associés à la nucléocapside (constituée des protéines NC, p7) et aux ARNt (qui serviront d'amorce à la transcriptase inverse) (**Figure 4B**). La capside contient également les protéines virales protéase, transcriptase inverse, VPR et intégrase qui permettront d'assurer les étapes précoces du cycle viral indispensables à l'expression, la multiplication et la persistance du virus (**Figure 4B**). Cependant, la particule virales du VIH-1 ne prendra cette forme mature qu'après maturation du virion lors de son bourgeonnement à la surface cellulaire après clivage des polyprotéines Gag et Gag-Pol en MA/CA/NC et PR/RT/IN, respectivement (Pluta and Kacprzak, 2009) (**Figure 4B**).

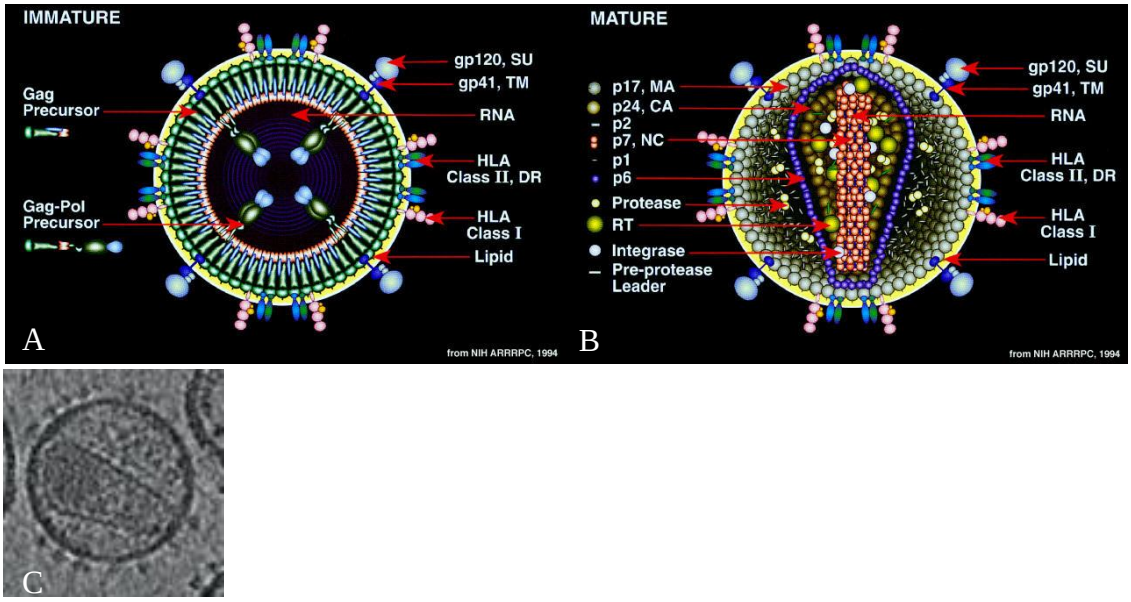


Figure 4 : Structures du VIH-1. Les panels **A** et **B** représentent schématiquement les particules virales immatures (**A**) et matures (**B**) sur lesquelles sont annotées les protéines virales (sous forme de précurseurs ou matures, respectivement) qui les composent ainsi que les protéines cellulaires incorporées (HLA de classe I, II) (Source : http://www.ictvdb.org/WIntkey/Images/ dg_HIV.jpg). Le panel **C** est une image 2D d'un virion mature (de 100 nm de diamètre) prise par microscopie cryo-électronique (Liu et al., 2010).

B/ LES BASES DE LA MULTIPLICATION DU VIH-1

Le cycle de multiplication du VIH-1 est composé des étapes précoces et tardives qui ont lieu, respectivement, avant et après intégration de l'ADN viral (ADNv) au sein de la chromatine de la cellule infectée (**Figure 6**). Certaines étapes du cycle du VIH-1 peuvent être influencées, positivement ou négativement, par les protéines de la cellule infectée. Il a ainsi été récemment identifié ~1400 gènes cellulaires essentiels à l'infection par le VIH-1 à l'aide d'approches génomiques globales (mettant en œuvre l'interférence par les ARNs) (Brass et

al., 2008; Espeseth et al., 2011; Genovesio et al., 2011; Konig et al., 2008; Rato et al., 2010; Waninger et al., 2004; Zhou et al., 2008).

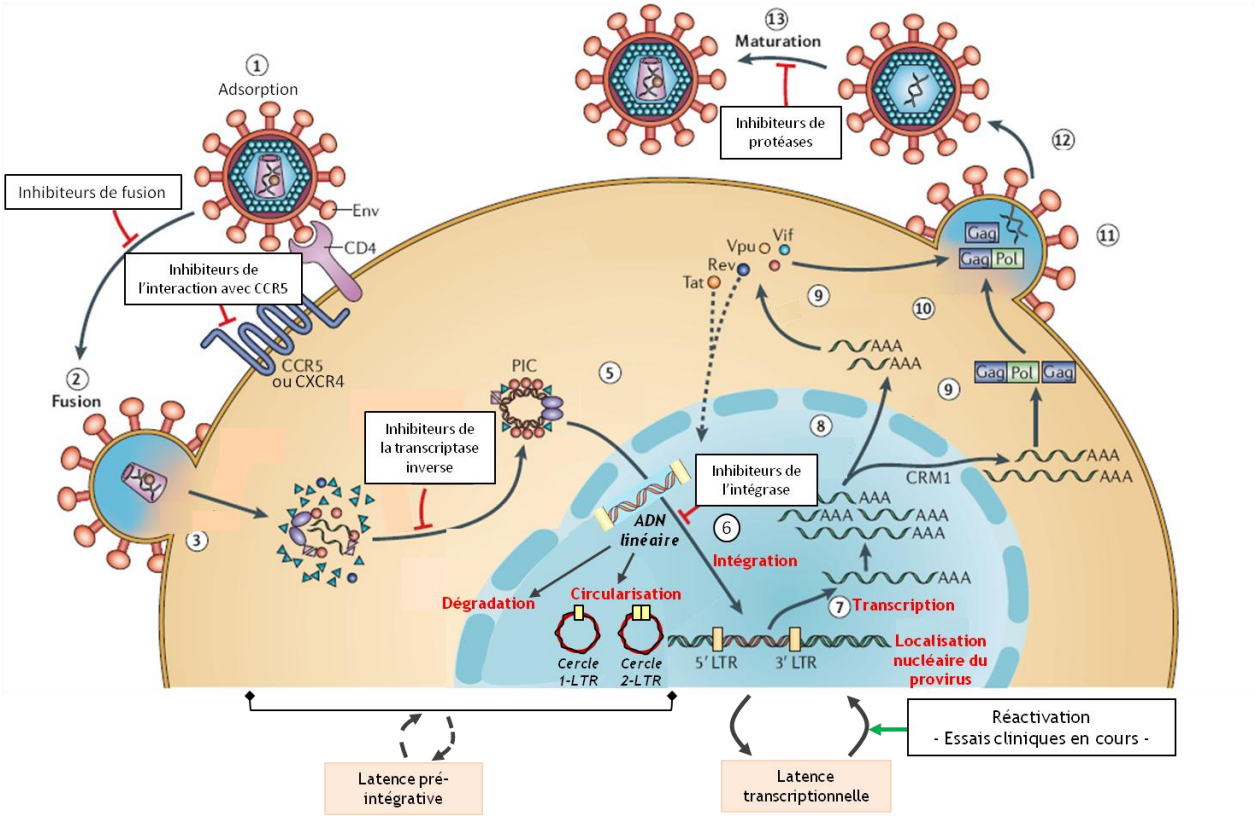


Figure 5 : Cycle de multiplication du VIH-1. (**1**) Le cycle de multiplication du VIH-1 débute par l'adsorption du virus à la surface cellulaire sur le récepteur CD4 et le co-récepteur CCR5 et CXCR4. (**2**) Cette étape est suivie de la fusion des membranes virales et cellulaires et (**3**) la décapsidation du *core* viral (**4**) dont le génome ARN est rétro-transcrit en un ADN double-brin. (**5**) L'ADN viral linéaire, associé aux protéines cellulaires et virales au sein du PIC, est importé au sein du noyau où il peut être dégradé, circularisé en épisomes à 1 ou 2-LTRs ou (**6**) intégré au sein de la chromatine hôte. (**7**) L'ADN viral intégré est transcrit en ARNm aux multiples épissages qui (**8**) sont exportés dans le cytoplame où (**9**) ils synthétisent les protéines régulatrices (qui retournent jouer leur rôle dans le noyau) ou structurales et accessoires qui participent à (**10**) l'assemblage de nouvelles particules virales. (**11**) Celles-ci bourgeonnent à la membrane cellulaire et sont libérées sous forme immature. (**12**) La maturation des virions a ensuite lieu grâce au clivage des polyprotéines grâce à la protéase virale. Comme précisé, plusieurs étapes du cycle de multiplication du VIH-1 peuvent être ciblées lors de la HAART. De plus, sont indiquées les étapes du cycle pouvant aboutir à générer un état de latence pré-intégrative ou transcriptionnelle (étapes 3 à 6, ou transcription, respectivement). Il est à noter que cette dernière peut être réactivée à l'aide de traitements pharmacologiques (*e.g.*, inhibiteurs des déacétylases d'histones) dont certaines sont actuellement en essais cliniques [adaptée de (Engelman and Cherepanov, 2012)].

1/ <u>Entrée : adsorption et fusion</u>

L'attachement du VIH-1 à la surface cellulaire a lieu grâce au complexe d'enveloppe Gp120-Gp41, essentiel à la reconnaissance du virus et à l'entrée au sein des cellules cibles (**Figure 5**). Gp120 fixe une glycoprotéine monomérique transmembranaire, CD4 exprimé à la

surface cellulaire d'environ 60 % des lymphocytes T circulants, des lymphocytes T précurseurs au sein de la moelle osseuse et du thymus, des monocytes et macrophages, des cellules éosinophiles, des cellules dentritiques et des cellules microgliales du système nerveux central. Une fois Gp120 fixée à la protéine CD4, l'enveloppe virale subit une modification structurale lui permettant de se fixer à des récepteurs à chimiokines à la surface cellulaire [*cf.*, revue (Peterlin and Trono, 2003)].

Les plus communs de ces co-récepteurs du VIH-1 sont CXCR4 et CCR5, respectivement exprimés sur de nombreux types cellulaires ou uniquement sur les monocytes/macrophages, les lymphocytes T activés et les cellules dendritiques. L'expression différentielle de ces deux récepteurs à la surface des cellules cibles est le déterminant majeur du tropisme du VIH-1. Ainsi, les souches du VIH-1 qui fixent préférentiellement le co-récepteur CCR5 sont des virus à tropisme M (pour macrophage) ou R5. Les souches se fixant préférentiellement à CXCR4 des lymphocytes T $CD4^+$ primaires sont dites à tropisme T (pour lymphocyte T) ou X4. Les souches du VIH pouvant se fixer sur les deux récepteurs sont des virus X4R5 à tropisme dual [*cf.*, revue (Peterlin and Trono, 2003)]. Il est à noter que ce tropisme double du VIH-1 peut indiquer une évolution du tropisme chez le patient après traitement à l'aide d'inhibiteurs de CCR5 (Mild et al., 2010).

La double fixation de Gp120 à CD4 et à un co-récepteur permet un attachement plus stable du virus, ce qui induit la pénétration de Gp41 au sein de la membrane cellulaire. Cet évènement rapproche les deux membranes qui fusionnent, permettant l'entrée de la capside virale (Peterlin and Trono, 2003).

Il est à noter que le virus VIH-1 et les vecteurs qui en dérivent utilisés au cours de notre étude sont pseudotypés à l'aide de la glycoprotéine G amphotrope d'enveloppe du virus de la stomatite vésiculaire (VSV), ce qui nous permet de nous affranchir de la présence du récepteur CD4 à la surface des types cellulaires utilisés. Mais, l'absence d'interaction du vecteur avec ses récepteurs empêche l'activation de cascades de signalisation pouvant être nécessaires à certaines étapes (notamment précoces) du cycle viral (Aiken, 1997; Yu et al., 2009).

2/ <u>La particule virale au sein du cytoplasme de la cellule hôte</u>

a/ *Facteurs de restriction*

Les facteurs de restriction représentent le système de défense cellulaire anti-virale intrinsèque, ou immunité intrinsèque. Ils ont la capacité de reconnaître directement des

composants viraux, de façon dépendante ou indépendante de leur activité cellulaire et d'agir à différentes étapes du cycle de multiplication du VIH-1 (**Tableau 4**). Leur action résulte en un blocage immédiat de la multiplication virale et peut aussi contribuer à la stimulation de l'immunité innée (*e.g.*, Tétherine). Ces acteurs de la défense anti-virale peuvent également agir sur le cycle de multiplication d'autres (rétro)virus, comme indiqué dans le **Tableau 4**, et/ou être spécifique de l'espèce de l'organisme infecté [*e.g.*, TRIM5α (*tripartite motif protein* 5α); **Tableau 4**] [*cf.*, revues (Wolf and Goff, 2008; Yan and Chen, 2012)].

Tableau 4 : Facteurs de restriction du VIH-1 [adapté de (Yan and Chen, 2012)].

Protéines	Rôle cellulaire	Virus ciblé(s) (spécificité d'espèce)	Fonction(s) dans le cycle viral
APOBEC3G	Déaminase des cytidines	**VIH-1**, SIV, EIAV, MLV, spumavirus, virus de l'hépatite B	Introduit des mutations au sein de l'ADN du VIH-1 (édite C en U dans le brin -). Inhibe la transcription inverse et l'intégration du VIH-1. *Action contrecarrée par Vif.*
TRIM5α	Ubiquitine U3 ligase	**VIH-1** (MR), MLV (H)	Bloque la décapsidation du virus entrant. Promeut la signalisation de l'immunité innée en reconnaissant la capside rétrovirale.
Tétherine		**VIH-1**, MLV, HTLV-1, virus d'Ebola, KSHV	Bloque le relarguage des virus enveloppés. *Action contrecarrée par Vpu.*
SAMHD1	Phosphohydrolase (stimulée par dGTP)	**VIH-1, VIH-2**	Inhibe la réplication du VIH-1 dans les cellules myéloïdes, probablement en régulant le stock de dNTP cellulaires. *Action contrecarrée par Vpx.*
TREX1	3' exonucléase	**VIH-1**	Enlève l'ADN rétro-transcrit abortif cytoplasmique. Inhibe les réponses de l'immunité innée contre le VIH-1

Abréviations : H, Homme ; MR, macaque rhésus ; APOBEC3G, *apolipoprotein B mRNA-editing catalytic polypeptide 3* ; TRIM5α, *tripartite motif protein* 5α ; SAMHD1, *SAM domain and HD domain 1 protein* ; TREX1, *three prime repair exonuclease 1* ; EIAV, *equine infectious anemia virus*; HTLV-1, *human T leukemia virus* de type-1; KHSV, *Kaposi's sarcoma–associated herpes S virus.*

b/ *Décapsidation, transport/trafic cytoplasmique*

Il est généralement admis que la décapsidation du core viral a lieu, au moins partiellement, dès l'entrée du virus dans le cytoplasme de la cellule infectée, libérant ainsi l'ARN viral (ARNv) (Suzuki and Craigie, 2007). Ce réarrangement du core viral est associé la maturation du complexe de transcription inverse (RTC, *reverse transcription complex*) (composé de protéines virales et cellulaire) qui existe sous forme de complexe de pré-initiation de la transcription inverse au sein des virions (Warrilow et al., 2009). Il a notamment été observé qu'une synthèse d'ADNc limitée (à partir du complexe de pré-

initiation de transcription inverse) a toutefois au sein des particules virales avant infection (Lori et al., 1992; Trono, 1992). Le RTC subit ensuite différentes stades de maturation jusqu'à achèvement du processus de transcription inverse (Warrilow et al., 2009).

L'évolution des techniques d'analyses des étapes de pré-intégration du VHI-1, et en particulier de la composition du RTC [illustré dans (Arhel, 2010)], permet actuellement de proposer que le core viral puisse être décapsidé de façon concomitante au processus de transcription inverse pendant le transport de la capside vers le noyau (grâce au réseau de microtubules et d'actine) (McDonald et al., 2002; Nermut and Fassati, 2003). De façon alternative, une autre hypothèse propose que le core viral ne soit désassemblé qu'une fois la transcription inverse achevée (Arhel, 2010). La capside du VIH-1 serait ainsi une structure perméable de maintien du RTC où une concentration importante de transcriptase inverse autour de son substrat, l'ARNg, favoriserait la conversion ARN simple brin-ADN double brin (ARNsb-ADNdb), elle-même facilitée par la fixation de NC. Le core viral favoriserait alors l'import nucléaire (Arhel, 2010).

c/ *Transcription inverse*

La transcription inverse est une étape majeure du cycle rétroviral qui aboutit à la duplication des LTRs. Ce processus a lieu en différentes étapes et permet la synthèse d'un brin d'ADN db à partir d'un ARN sb grâce à l'action de la transcriptase inverse, comme décrit dans la **figure 6**.

L'achèvement de la maturation du RTC et de la transcription inverse permettent de générer le complexe de pré-integration (PIC) également considéré comme la forme la plus mature du RTC (Warrilow et al., 2009).

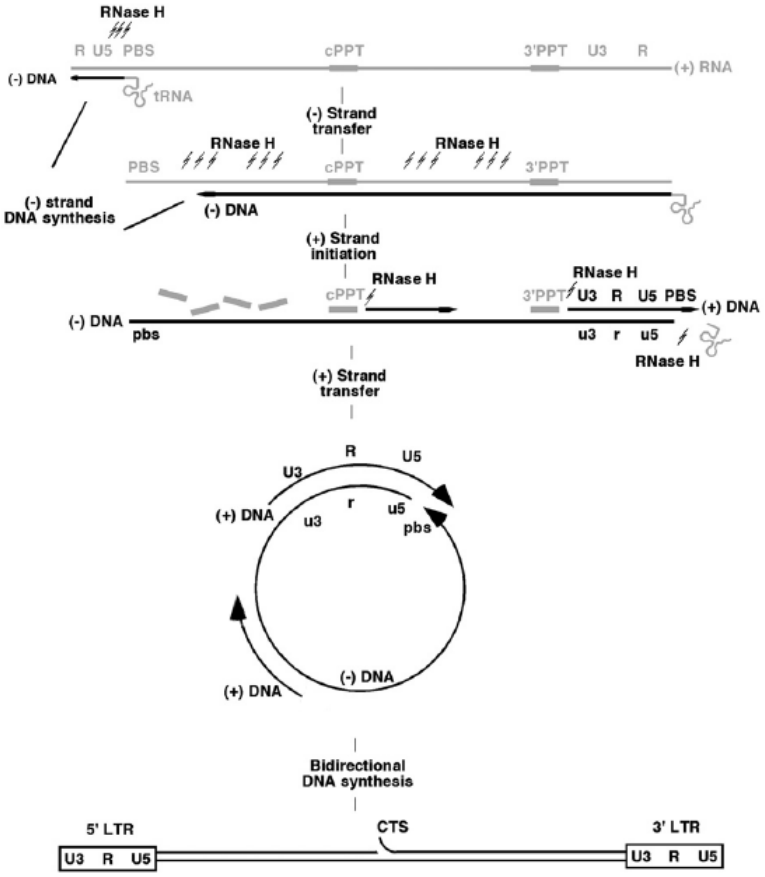


Figure 6 : Les différentes étapes de la transcription inverse. Elle est initiée par la synthèse du brin d'ADN (-) en 5' du brin d'ARN (+) en utilisant un $ARNt_{Lys3}$ hôte qui fixe le site PBS. La synthèse d'ADN s'arrête au R de l'extrémité 5' avant qu'il n'y ait transfert du brin d'ADN (-) néo-synthétisé sur la séquence R du même brin d'ARN ou du second. L'élongation du brin d'ADN (-) se poursuit et est accompagnée de la digestion de la matrice d'ARN (par l'activité RNaseH de la RT) à laquelle résistent partiellement les séquences PPT en position centrale (cPPT) et en 3' (3'PPT) qui servent de sites d'initiation à la synthèse du brin d'ARN (+). De plus, la complémentarité des séquences PBS des brins d'ADN (+) et (-) permet le transfert du brin d'ADN (+) et la synthèse complète du double brin d'ADN viral. Cependant, l'élongation du brin d'ADN (+) à partir de 2 sites d'initiation permet de générer un déplacement du brin en aval sur 100 nucléotides (nt), appelé *DNA flap* central qui se termine au CTS (Pluta and Kacprzak, 2009) [adaptée de (Gotte et al., 2010)].

d/ *Formation du complexe de pré-intégration cytoplasmique*

Le PIC, complexe nucléoprotéique compétent pour l'intégration, contient l'ADNdb viral compact associé à l'IN mais dépourvu de CA (*i.e.*, décapsidation totale) et NC (Bukrinsky et al., 1993; Farnet and Haseltine, 1991; Miller et al., 1997). La présence d'autres protéines virales caryophiles MA, RT, PR et Vpr est débattue (Bukrinsky et al., 1993; Farnet and Haseltine, 1991; Iordanskiy et al., 2006; Karageorgos et al., 1993; Miller et al., 1997).

Il a également été montré que certaines protéines cellulaires peuvent être associées au PIC du VIH-1 *in vitro* et/ou *ex vivo* (**Tableau 5**). Cependant, la composition du PIC évolue

probablement en fonction de sa localisation, *i.e.*, cytoplasme *vs.* noyau. Les protéines cytoplasmiques du PIC pourraient néanmoins favoriser l'import nucléaire.

Tableau 5 : Protéines cellulaires identifiées comme interagissant avec l'intégrase et/ou le PIC [issue de (Raghavendra et al., 2010)].

Protéines cellulaires	**Méthode d'identification**[b]
BAF (*Barrier-to-Autointegration Factor*)	SS, BTD
Gémine 2	IP
HAT p300 (*Histone Acetyl-Transferase p300*)	IP
HMGA1 (*high mobility group AT-hook 1*)	SS
HSP60 (*Heat shock protein 60*)	PD
Protéine EED (*embryonic ectodermal development*)	THS, PD
Importine 7	IP, PD (Zaitseva et al., 2009)
IN1 (*Integrase interactor 1* ou hSNF5)	THS, THS (Rain et al., 2009)
LEDGF/p75 (*Lens Epithelium-Derived Growth Factor p75 splice variant*)	IP, THS (Rain et al., 2009)
UNG2 (*uracil DNA glycosylase 2*)	PD
NUP153 (*nucleoporin 153 kDa*)	PD (Woodward et al., 2009)
TRN-SR2 (*transportin*-SR2 ou TNPO3)	PD (Christ et al., 2008), THS (Rain et al., 2009)
VBP1 (*Von Hippel-Lindau Binding protein 1*)	THS (Rain et al., 2009)
Ku80	PE (Li et al., 2001), THS et PD (Studamire and Goff, 2008)
Np1l1, Rbbp4, An32a, An32e, Calréticuline, Nemo, Nono, LEO1, DDX19A, STAU1, Hnrh-1, Hnrh-3, PairB, Sf3b2, Sfrs3, Snut1, If4g, Dynactine	BTD

[a]En l'absence de précision, les résultats sont issus de (Raghavendra et al., 2010)].
[b]SS (*salt stripped*) : reconstitution de l'activité du PIC altérée par les fortes concentrations en sels ; IP : immunoprécipitation de complexes formés au sein des cellules ; PD (*pull down*) : co-précipitation *in vitro* ; THS (*two hybrid system*) : système double-hybride ; PE (PIC *extraction*) : préparation de PIC et analyse des protéines par Western-blot ; BTD (biotinylated *target DNA*) : identification par spectrométrie de masse des complexes associés à un ADN cible biotinylé.

3/ <u>La traversée active du cytoplasme au noyau : l'import nucléaire</u>

Les lentivirus ont la capacité de se répliquer au sein de cellules quiescentes qui implique le passage de la membrane nucléaire par le PIC. Le PIC du VIH-1 de 56 nm de diamètre (Miller et al., 1997) est transporté activement au travers des pores nucléaires de façon ATP(adénosine triphosphate)-dépendante (Bukrinsky et al., 1992). Ce processus impliquerait, directement ou indirectement, les protéines virales et/ou cellulaires qui

constituent le PIC ainsi que la structure centrale particulière de l'ADNdb viral appelée *DNA flap* (Suzuki and Craigie, 2007). Cependant, le rôle exact de ces facteurs, et notamment du *DNA flap*, reste controversé (Suzuki and Craigie, 2007). Plus récemment, de nouveaux partenaires de l'IN [TRN-SR2 (*transportin*-SR2), l'importine 7 et le composant du pore nucléaire NUP153 (*nucleoporin* 153 kDa)], ont été identifiés comme participant au contrôle de l'accès au noyau (Christ et al., 2008; Woodward et al., 2009; Zaitseva et al., 2009).

4/ Les étapes nucléaires du cycle de multiplication du VIH-1

a/ *Formation du complexe de pré-intégration nucléaire*

Une fois dans le noyau, le PIC incorpore des protéines nucléaires, dont LEDGF/p75 (Llano et al., 2004), HMGA1 (Miller et al., 1997), Ku (Chen and Engelman, 1998 ; Li et al., 2001 ; Lin and Engelman, 2003; Studamire and Goff, 2008), BAF (Lin and Engelman, 2003) et IN1 (Iordanskiy et al., 2006) (**Tableau 5**). LEDGF/p75 a un rôle connu dans l'intégration de l'ADNv, comme développé dans le paragraphe **iii/** de cette partie. Les autres protéines citées ci-dessus pourraient également avoir un impact dans les étapes nucléaires du cycle du VIH-1, *i.e.*, intégration, localisation et expression.

b/ *Processus d'intégration*

Une fois dans le noyau, l'IN, associée à ses partenaires viraux et cellulaires du PIC, circularise l'ADNdb viral et catalyse les deux premières étapes de la réaction d'intégration au sein du génome hôte (**Figure 7**). Ces deux étapes permettent la maturation des extrémités 3' de l'ADN viral (*i.e.*, maturation en 3') et le transfert de celui-ci au sein du génome humain (*i.e.*, transfert de brins) (détails de la réaction **Figure 7**). La troisième et dernière étape du processus d'intégration est la réparation des lésions générées lors du transfert de l'ADNv dans la chromatine de la cellule hôte par la machinerie de réparation de l'ADN (**Figure 8**). L'implication de cette dernière dans le cycle de multiplication du VIH-1 sera décrite dans le chapitre IV de cette thèse.

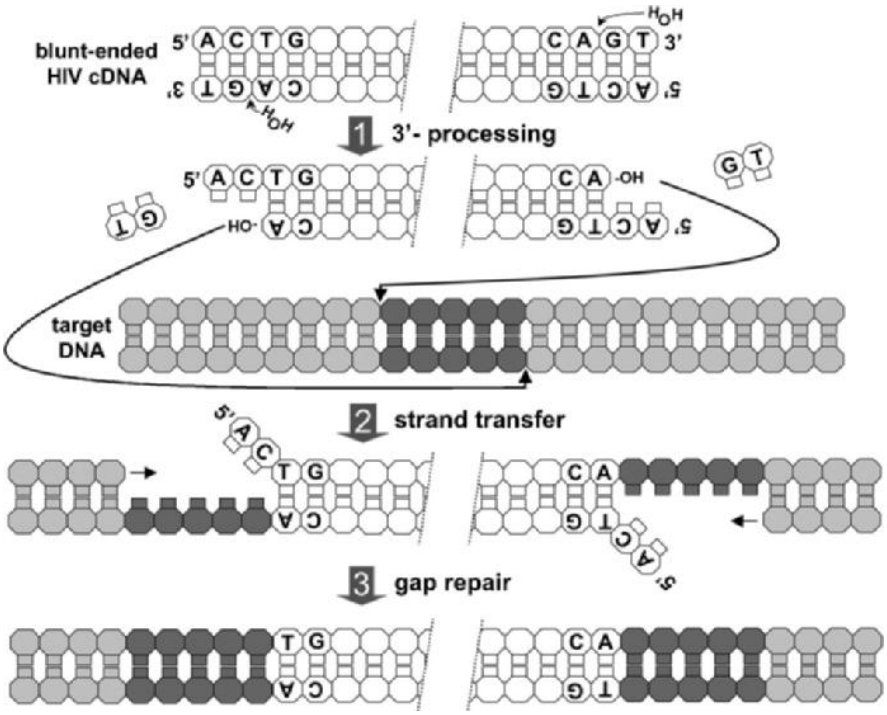


Figure 7 : La réaction d'intégration de l'ADN viral au sein de la chromatine hôte. Le précurseur de l'intégration est une molécule d'ADNc viral (ADNv) à bouts francs associé à l'IN au sein du PIC. (**1**) La première réaction, appelée maturation en 3' (*3' processing*), consiste en un clivage par l'intégrase des 2 nucléotides terminaux de l'ADNv (qui crée des extrémités 3' simple brin CA). Les groupements –OH 3' des extrémités 3' simple brin CA de l'ADNv vont ensuite attaquer les liaisons phosphodiesters des brins opposés (séparées par 5 bases) de l'ADN hôte. (**2**) Cette *trans*-estérification, deuxième étape du processus d'intégration également catalysée par l'intégrase, est appelée transfert de brins (*strand transfer*). (**3**) La troisième étape implique la machinerie de réparation de l'ADN qui répare les cassures générées (*gap repair*) lors de l'intégration et coupent les extrémités 5' qui dépassent de l'ADNv, ce qui créée une duplication des séquences d'ADN de l'hôte entourant le provirus [adaptée de (Poeschla, 2008)].

Il est à noter que la première réaction du processus d'intégration a lieu au sein du PIC avant l'entrée au noyau. Le processus d'intégration se déroule également au sein du PIC en interaction avec les partenaires cellulaires de l'IN. L'un d'entre eux, *i.e.*, LEGDF/p75, influence le positionnement du PIC au sein de la chromatine en interagissant à la fois avec l'IN et certaines séquences de l'ADN hôte cible (Ciuffi, 2008) (**Figure 8**). Il n'est pas exclut que d'autres protéines cellulaires du PIC puissent également influencer le devenir du PIC, *i.e.*, Ku, comme développé dans le chapitre IV de cette thèse.

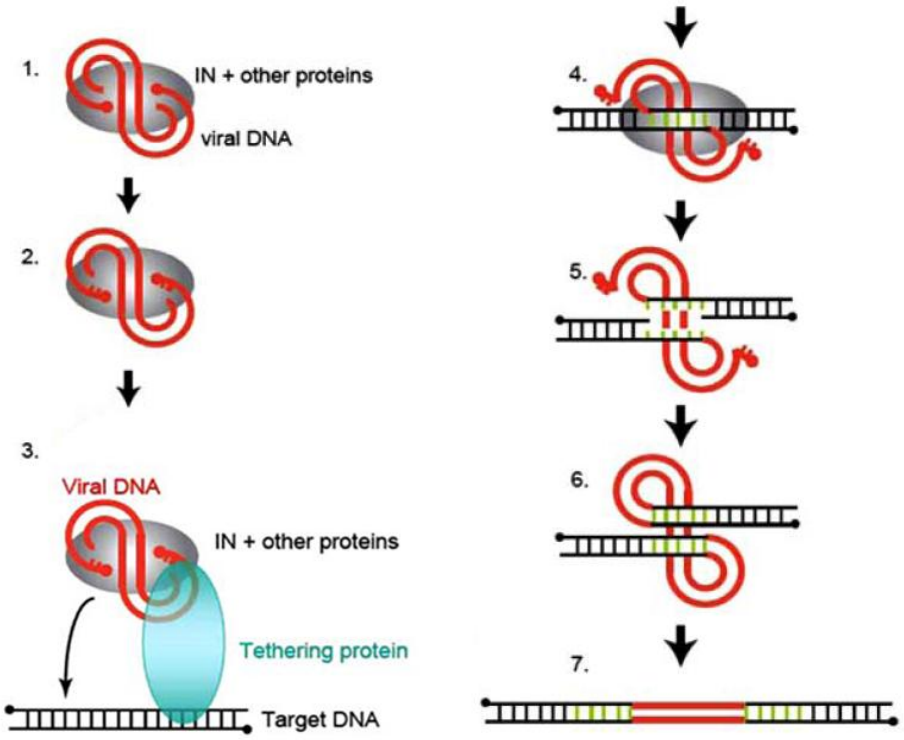


Figure 8: Le processus d'intégration de l'ADN viral dans le génome cellulaire dépend de l'intégrase et des protéines cellulaires du PIC. (**1**) Le processus d'intégration dépend de l'intégrase, qui est associée aux autres protéines virales et cellulaires au sein du PIC (représenté par un ovale gris) et à l'ADNv (ligne rouge en gras, les ronds rouges représentent les extrémités 5' de l'ADN). (**2**) La première étape permet de générer les extrémités simple-brin de l'ADN. (**3**) Le recrutement du PIC au niveau du site d'intégration de l'ADN cible est facilité grâce à l'intervention d'une protéine ou complexe protéique pouvant se fixer à la chromatine (*i.e.*, LEDGF/p75). (**4**) L'IN peut alors simultanément créer des cassures au niveau de l'ADN cible [entre 5 paires de bases (bp)] et les fixer aux extrémités 3' de l'ADNv, (**5**) en laissant un intermédiaire d'intégration avec des bases non-appariées à chacune des jonctions entre ADNs viral et cible. (**6**) La machinerie de réparation de l'ADN de la cellule hôte permet la synthèse des nucléotides manquants, (**7**) ce qui résulte en une duplication des 5 bp qui flanquent l'ADN proviral [adaptée de (Ciuffi, 2008)].

c/ *Sites préférentiels d'intégration du VIH-1*

Des analyses génomiques des sites d'intégration des lentivirus, dont le VIH-1 et les vecteurs dérivés, ont mis en évidence son intégration préférentielle au sein des unités de transcription actives (Bushman et al., 2005; Hacker et al., 2006; Mitchell et al., 2004; Schroder et al., 2002; Wang et al., 2007), plus accessibles aux complexes d'intégration (**Figure 9**).

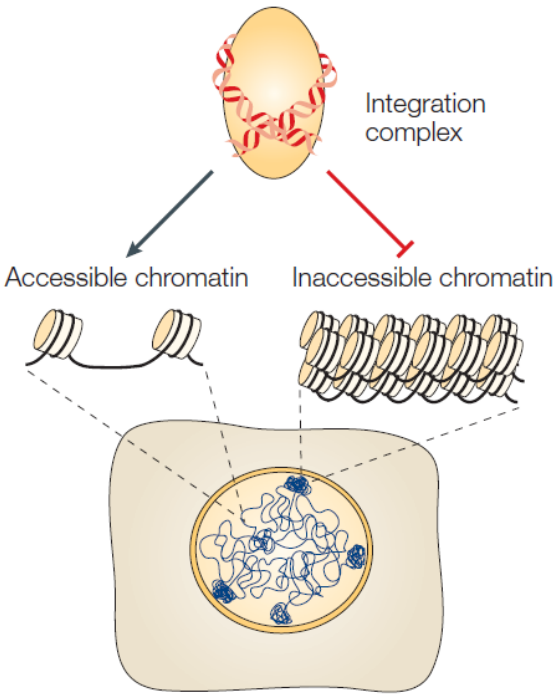


Figure 9 : Régulation de l'intégration par l'accessibilité de l'ADN. L'ADN empaqueté au sein des nucléosomes est relativement inaccessible aux complexes d'intégration. L'exposition de séquences cibles lors de la décompaction de la chromatine promeut ainsi l'intégration [issue de (Bushman et al., 2005)].

Celles-ci contiennent des sites de reconnaissance pour des co-facteurs cellulaires associés à l'IN, tel LEDGF/p75 qui est en outre nécessaire (mais pas essentiel) pour obtenir une intégration et une réplication efficace (Ciuffi et al., 2005; Shun et al., 2007). En effet, LEDGF/p75 permet le recrutement du PIC au sein de la chromatine hôte grâce à sa capacité à fixer directement l'IN du VIH-1 en C-terminal [via son domaine IBD (*integrase binding domain*)] et la chromatine en N-terminal [via son domaine PWWP (*Pro-Trp-Trp-Pro*) et une paire de crochets A/T] (Llano et al., 2006). LEDGF/p75 fixe aussi spécifiquement l'IN d'autres lentivirus (*i.e.*, VIH-2, EIAV), ce qui leur confère une distribution des sites d'intégration semblable à celle du VIH-1 (Hacker et al., 2006; Hematti et al., 2004; Kang et al., 2006; MacNeil et al., 2006). En absence de LEDGF/p75, l'intégration du génome lentiviral ainsi que sa préférence pour les gènes transcriptionellement actifs sont fortement réduites (Ciuffi, 2008). La distribution des sites d'intégration ressemble alors à celle des gammarétrovirus (Wu et al., 2003), *i.e.*, à proximité et au sein des promoteurs ainsi que dans les îlots CpG (Mitchell et al., 2004). Les spumavirus ont la même préférence d'intégration que les gammarétrovirus, dans une proportion cependant moins importante (Nowrouzi et al., 2006; Trobridge et al., 2006).

d/ *Autres espèces d'ADN viral générées*

L'ADNv intégré, appelé provirus, représente cependant uniquement 5 à 10 % de l'ADNv total contenu dans une cellule infectée (Brussel and Sonigo, 2004). La vaste majorité de l'ADNv est donc sous forme non-intégrée. La forme la plus abondante d'ADNv est sous forme linéaire extrachromosomal, produit direct de la transcription inverse de l'ARNv et substrat du processus d'intégration (**Figure 10**). Les autres formes d'ADNv non-intégré sont circulaires et générées à partir de l'ADNv linéaire grâce à différents mécanismes. Ils peuvent être produits par auto-intégration (*i.e.*, « attaque » des brins d'ADNv par les extrémités issus du *3'processing*) sous forme d'auto-intégrants tronqués ou réarrangés, et représenter jusqu'à 20 % de l'ADNv après infection par le Mo-MLV (Shoemaker et al., 1981) (**Figure 10**). D'autres espèces virales circulaires, contenant un ou deux LTRs (respectivement, 30 et 5 % de l'ADNv total) sont formées par recombinaison homologue (*i.e.*, cercles à 1-LTR) et/ou intervention de protéines respectivement impliquées dans les voies de réparation de l'ADN par recombinaison homologue ou non-homologue [*cf.*, revue (Sloan and Wainberg, 2011) et § IV.B.1-2] (**Figure 10**).

Ces différentes espèces virales ont des durées de vie variables. La quantité d'ADN linéaire diminue rapidement s'il n'est pas intégré (durées de demi-vie estimées à 1–2 jours) (Zhou et al., 2005), alors que les cercles à 1- et 2-LTRs s'accumulent dans le noyau à des temps plus tardifs de l'infection et sont éliminés lors des divisions cellulaires. Ces derniers, qui étaient considérées jusqu'à peu comme des produits finis issus d'évènements abortifs (Brown et al., 1989), peuvent exprimer les protéines précoces du VIH-1 (*i.e.*, Nef, Tat et Rev) et ainsi participer à la réplication du VIH-1 (Gelderblom et al., 2008; Kelly et al., 2008) [*cf.*, revue (Sloan and Wainberg, 2011) et § III.B.3]. De plus, il est actuellement proposé que les cercles à 2-LTRs soient des intermédiaires tardifs dans le processus d'intégration médié par le clivage de la jonction 2-LTR via l'IN (Delelis et al., 2007), ce qui permettrait la persistance de l'infection en présence d'inhibiteurs de l'intégrase.

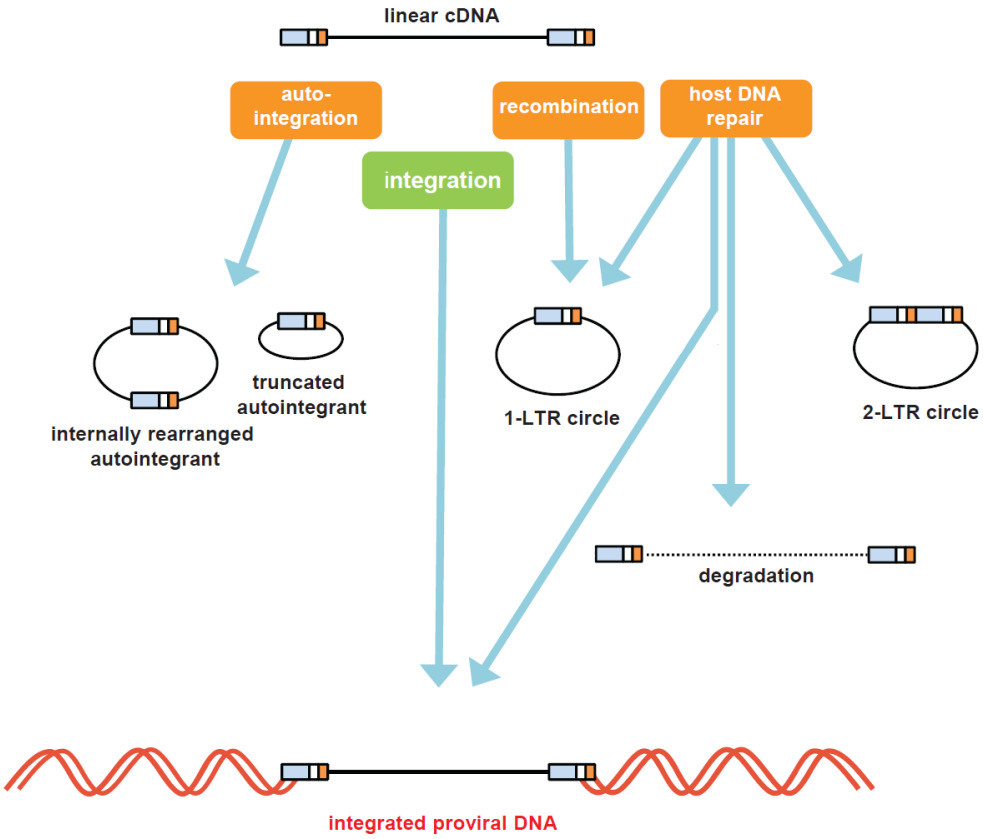


Figure 10 : Les différentes formes d'ADN viral au sein du noyau. Différentes formes d'ADNv sont générées à partir d'ADNv linéaire extra-chromosomale, espèce virale majoritaire, *via* différents processus. Celui-ci peut être intégré, ou circularisé par recombinaison, auto-intégration ou intervention de la machinerie de réparation de l'ADN (qui peut aussi intervenir dans la dégradation de l'ADN linéaire. Ces mécanismes permettent de générer des formes d'ADNv réarrangées/tronquées, ou des cercles à 1-LTR/ 2-LTRs, respectivement [adaptée de (Sloan and Wainberg, 2011)].

5/ Expression

L'ADNv se comporte, une fois intégré, comme n'importe quel gène cellulaire. Les LTRs qui bordent en 5' et 3' le provirus définissent respectivement l'initiation et la terminaison de transcription. L'initiation de la transcription à partir du 5'LTR est régulée par la présence/absence de facteurs de transcription cellulaires qui positionnent l'ARN polymérase II (ARN pol II) et les protéines de la machinerie cellulaire de transcription associées. Cependant, une fois la transcription initiée, ce sont principalement les protéines virales précoces, Tat et Rev, qui permettent d'atteindre de hauts niveaux d'expression en stimulant la transcription, et l'export des ARNm non- et mono-épissés, respectivement. Une fois au sein du cytoplasme, les ARNm viraux sont pris en charge par la machinerie cellulaire de traduction et traduits en protéines correspondantes [*cf.*, revue (Karn and Stoltzfus, 2012)]. Un paragraphe sur l'expression du VIH-1 sera développé dans le chapitre III de cette thèse.

6/ Assemblage, bourgeonnement et maturation des particules virales

Les poly/protéines structurales et enzymatiques virales migrent jusqu'au lieu d'assemblage des particules virales en étant stabilisées par le cytosquelette d'actine et de microtubules (Jolly and Sattentau, 2007). Le lieu d'assemblage prédominant des particules virales pour les différents types cellulaires infectés est la membrane plasmique (Sundquist and Krausslich, 2012).

En effet, dans les lymphocytes T, le VIH s'assemble et bourgeonne au niveau de microdomaines de la membrane plasmique, appelés radeaux lipidiques, plateformes d'interactions protéines-lipides. La polyprotéine Gag/Pr55 joue un rôle central dans l'assemblage du virus puisqu'elle suffit à la production de particules non-infectieuses en l'absence d'autres protéines virales (Gheysen et al., 1989). En particulier, le dimère ARNg est transporté jusqu'à la membrane plasmique en étant complexé à Gag via son domaine NC qui permet son assemblage au sein de la particule virale (Martin-Serrano and Neil, 2011). De plus, le domaine MA de Gag lui permet de cibler et de se fixer la membrane plasmique (Saad et al., 2006), ce qui favorise l'oligomérisation de Gag (Ono et al., 2007) qui recrute Gag-Pol/Pr160 et Env. Le bourgeonnement des particules virales immatures est médiée par le domaine C-terminal de Gag, p6, qui interagit avec des composants de la machinerie ESCRT (*Endosomal Sorting Complex Required for Transport*), facilitant ainsi la dissociation de la membrane plasmique.

Dans les macrophages, l'amélioration des techniques de microscopie a permis de montrer récemment que le VIH s'assemble et bourgeonne au niveau d'invaginations de la membrane plasmique accessibles à la surface (Sundquist and Krausslich, 2012). Ce mécanisme avait initialement laisser supposer que les particules virales de VIH-1 bourgeonnait dans des endosomes ou de structures similaires à des corps multivésiculaires, puis étaient relarguées des cellules par exocytose (Kramer et al., 2005; Nguyen et al., 2003).

Les particules virales sont libérées des cellules productrices sous forme immature. Leur maturation, qui a lieu rapidement après bourgeonnement, repose sur une cascade de réactions de clivage des polyprotéines Gag et Gag-Pro/Pol en protéines matures Gag et Pro/Pol. Le virus mature et infectieux qui en résulte est alors capable d'initier un autre cycle d'infection dans une nouvelle cellule cible. Il a également été observé que les particules virales matures pouvaient être transmises de façon efficace via la formation d'une synapse virologique créée par l'interaction Env-CD4 entre cellule infectée et non-infectée (Puigdomenech et al., 2008; Sato et al., 1992). Cette structure est stable et maintenue par le

recrutement de molécules d'adhésion et des réarrangements du cytosquelette d'actine sans qu'il y ait fusion cellulaire. L'assemblage des particules virales et leur relarguage deviennent polarisés au niveau de la zone de contact, et les virions sont transférés au sein de l'espace intercellulaire avant d'entrer dans la cellule non-infectée (Martin-Serrano and Neil, 2011).

7/ Conséquences de l'infection par le VIH-1

En fonction du contexte cellulaire et des interactions virus/hôte, l'infection peut aboutir à la mort des cellules infectées, souvent associée à une production virale importante, ou à une latence, avec peu ou pas d'expression virale.

a/ *Mort des cellules infectées*

L'infection *in vivo* par le VIH-1 des lymphocytes T $CD4^+$ résulte majoritairement en une apoptose massive des cellules infectées médiée par les protéines virales et/ou de l'hôte en fonction du stade d'infection dans lequel se trouve la cellule infectée et aboutit à une profonde déplétion en lymphocytes T en absence de HAART (Cummins and Badley, 2010). Les cellules infectées peuvent être détruites par le système immunitaire inné lors de la phase aigue d'infection (Cossarizza, 2008). La mort de cellules non-infectées par apoptose peut également être induite par leurs voisines infectées (*i.e.*, effet *bystander*) après stimulation par de facteurs immunitaires cellulaires ou viraux membranaires et/ou solubles lors de la phase chronique d'infection (Cossarizza, 2008). D'autres formes de mort ont été décrites (*i.e.*, nécrose, autophagie) (Cossarizza, 2008). La mort des cellules infectées peut survenir avant qu'il n'y ait production virale (*e.g.*, système immunitaire compétent chez les individus non-progresseurs, quantité létale/non-réparable de dommages de l'ADN; *cf.*, § IV.A.2) ainsi qu'au sein des autres types cellulaires $CD4^+$ [*e.g.*, macrophages; *cf.*, revue (Cummins and Badley, 2010)].

b/ *Echappement au système immunitaire des cellules infectées*

Les cellules infectées par le VIH-1 peuvent cependant échapper au système immunitaire en charge de les éliminer grâce à la mise en œuvre de différentes stratégies. Parmi celles-ci, Nef, localisée au sein de la membrane plasmique, interfère avec la migration et la persistance des molécules CMH de classe I à la surface cellulaire normalement en charge de présenter les peptides viraux nécessaires à la reconnaissance par des lymphocytes T cytotoxiques spécifiques. L'entrée en latence du provirus au sein des lymphocytes T $CD4^+$ au repos peut aussi être considérée comme un échappement au système immunitaire puisqu'il permet la constitution de réservoirs viraux [*cf.*, revue (Peterlin and Trono, 2003)].

c/ *Latence des cellules infectées*

De la même façon qu'il existe deux phases, précoce et tardive, dans le cycle de réplication du VIH, deux types de latence ont été décrites et sont appelées pré-intégrative ou transcriptionnelle en fonction de si elles ont lieu, respectivement, avant ou après intégration du génome viral. L'entrée en latence du virus dépend du type cellulaire infecté et de son état physiologique. Par exemple, un défaut en dNTP (*desoxy A/T/C/G tri-phosphate*) au sein de la cellule (*i.e.*, lymphocytes T $CD4^+$ quiescents) induit une transcription inverse incomplète, ce qui correspond à de la latence pré-intégrative. Par contre, un défaut en facteurs de transcription [*i.e.*, NF-κB (*nuclear factor kappa B*)] au sein de la cellule (*i.e.*, lymphocytes T $CD4^+$ au repos) empêche l'expression virale et induit une latence transcriptionnelle. Les différentes formes de latence du VHI-1, et en particulier, la latence transcriptionnelle, seront décrites dans le chapitre III de cette thèse.

II/ VECTEURS LENTIVIRAUX : LES LENTIVIRUS COMME VECTEURS DE GENES

A/ INTERETS DES RETROVIRUS COMME VECTEURS DE GENES

1/ Importance des vecteurs rétroviraux dans le transfert de gènes

Le transfert de gène basé sur les rétrovirus repose sur le remplacement des gènes viraux par des gènes d'intérêt, pouvant être transférés à des taux importants grâce à une machinerie rétrovirale minimale, dans le cadre d'applications en thérapie génique mais aussi comme outils en recherche fondamentale.

Le concept d'utilisation des rétrovirus comme vecteurs de gènes a émergé dans le début des années 1980 avec la création du premier vecteur gammarétroviral basé sur le Mo-MLV (Tabin et al., 1982). Le développement des premiers lentivecteurs a ensuite été mis au point rapidement après la découverte du VIH-1 dès 1990 et a considérablement bénéficié des avancées déjà acquises dans les recherches sur les vecteurs d'origine gammarétrovirale ainsi que sur les autres familles de virus (Page et al., 1990). Grâce à ce développement précoce, les vecteurs rétroviraux ont été les premiers vecteurs viraux à entrer en essais cliniques en 1989 (Rosenberg et al., 1990). Ils ne représentent cependant en 2012 plus que ~23 % (toutes origines confondues) des vecteurs utilisés en thérapie génique (**Figure 11**). Le concept de thérapie génique est de traiter des maladies par le transfert de matériel génétique. Ainsi, les

vecteurs rétroviraux sont particulièrement adaptés au traitement par transfert de gène des maladies monogéniques pour lesquelles, en théorie, le remplacement stable du gène défectueux/muté homozygote récessif pourrait suffire à traiter la pathologie. Cependant, si les vecteurs rétroviraux sont des outils génétiques de premier choix du fait de leur capacité à s'intégrer stablement dans le génome de leur hôte, la sécurité du transfert de gène doit être garantie, surtout lorsqu'il est à visée thérapeutique, en empêchant notamment la mutagenèse insertionnelle et/ou la production de particules virales.

A l'heure actuelle, la plupart des vecteurs rétroviraux utilisés dérive des lentivirus (VIH), gammarétrovirus (Mo-MLV) et, de façon plus minoritaire, des spumavirus (PFV).

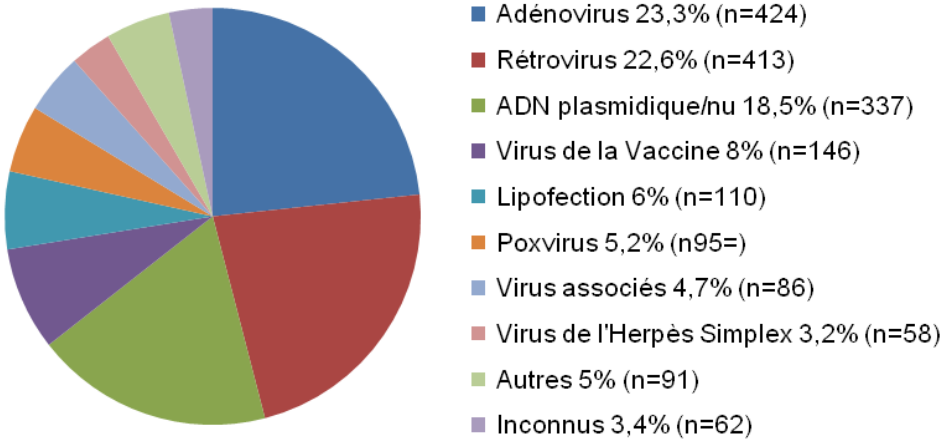


Figure 11 : Vecteurs viraux et non-viraux utilisés en essais cliniques de thérapie génique en 2012 (Source : http://www.abedia.com/wiley/vectors.php, des Editions Wiley).

2/ Propriétés des rétrovirus et conséquences pour le *design* de vecteurs dérivés

Ce paragraphe décrit les propriétés des rétrovirus ainsi que les avantages et inconvénients qu'ils confèrent aux vecteurs rétroviraux. Les aspects concernant l'amélioration du *design* de vecteurs (acquises et en cours d'exploration) seront développés ultérieurement (*cf.*, § II.B.1).

a/ *Avantages liés à l'utilisation des rétrovirus comme vecteurs de gènes*

i/ Intégration de l'ADN viral dans la chromatine hôte

Comme souligné ci-dessus, les vecteurs rétroviraux ont la capacité de pouvoir s'intégrer stablement au sein du génome des cellules qu'ils infectent. Cette propriété unique chez les virus assure l'expression à long terme du transgène au sein des cellules transduites et de leur descendance. Cependant, le site d'intégration de l'ADNv ne peut être ciblé au sein de

la chromatine, ce qui est un inconvénient majeur pouvant être contourné grâce à l'utilisation de vecteurs rétroviraux non-intégratifs (aspects développés dans les § I.B.4 et II.A.2).

ii/ La transcription inverse de l'ARN viral et duplication des LTRs

L'organisation génétique même des rétrovirus facilite le design des vecteurs qui en dérivent puisque le transgène se trouvera bordé de deux LTRs après transcription inverse (qui les dupliquent) et intégration au sein du génome hôte (*cf.*, § I.6.c). La délétion au sein des LTRs des vecteurs de transfert de troisième génération (aspect développé dans le § II.3.c) permet l'inclusion de promoteurs, insulateurs, ou séquences régulatrices de l'expression qui seront donc systématiquement dupliqués lors de la transcription inverse. Cependant, la transcriptase inverse des rétrovirus a un taux d'erreurs important [excepté chez les spumavirus (Rinke et al., 2002)]. Cette source de variabilité génétique chez les rétrovirus peut leur conférer un avantage sélectif mais aussi provoquer des mutations du transgène.

iii/ Import nucléaire des lentivirus

Grâce à leur capacité d'import de l'ADNv au noyau, les vecteurs dérivés des lentivirus sont les seuls parmi les rétrovirus à pouvoir transduire à la fois des cellules se divisant et quiescentes, notamment les cellules souches et progénitrices (Reiser et al., 1996), les hépatocytes primaires (Pfeifer et al., 2001), les neurones (Naldini et al., 1996a), les monocytes et les macrophages (Weinberg et al., 1991). Cette capacité unique permet d'élargir considérablement le tropisme de cellules/tissus disponibles pour le transfert de gènes dans le cas de vecteurs pseudotypés, ce qui a favorisé l'utilisation de lentivecteurs au détriment des vecteurs d'origine gammarétrovirale.

iv/ Pseudotypage des vecteurs rétroviraux

L'utilisation d'enveloppes hétérologues pour la production des vecteurs rétroviraux sans pour autant modifier les caractéristiques des particules virales a été un développement initial majeur acquis dès 1990 pour les lentivecteurs (Page et al., 1990). Le pseudotypage représente un avantage considérable pour ne pas être limité dans le choix des cellules à transduire et/ou cibler un type cellulaire donné. De plus, l'utilisation d'une glycoprotéine d'enveloppe amphotrope (*i.e.*, capable de se répliquer dans différents hôtes) stable [*e.g.*, glycoprotéine G du VSV, VSV-G] permet une concentration par ultracentrifugation des surnageants viraux et leur stockage à long terme (Burns et al., 1993).

v/ Immunogénicité faible lié au transfert de gènes

Les rétrovecteurs (gammarétrovecteurs et lentivecteurs) peuvent déclencher une réaction immunitaire innée et adaptative dirigée contre les particules virales, les cellules transduites et/ou le produit du gène transduit par le vecteur *in vivo* (Matrai et al., 2010). L'immunogénicité des particules rétrovirales est cependant relativement faible si elle comparée à celle provoquée par la transduction d'autres systèmes viraux [*e.g.*, vecteurs adénoviraux et *adeno-associated virus* (AAV)] et est dépendante du *design* du vecteur, de la voie d'administration et du dosage de l'injection (Nayak and Herzog, 2010).

Le *design* actuel des vecteurs lentiviraux (non-réplicatifs) permet également de ne pas transduire de séquences codantes complètes de protéines virales, ce qui empêche donc l'expression de protéines virales au sein de l'hôte après transduction (*cf.*, § II.3.d).

b/ *Limites liées à l'utilisation des rétrovirus comme vecteurs de gènes*

Un certain nombre de limites existe dans le transfert de gène mettant en œuvre les vecteurs lentiviraux qui, bien qu'ayant (au moins en partie) été dépassées, contribue aujourd'hui à réduire l'utilisation des vecteurs lentiviraux en thérapie génique.

i/ Intégration non-ciblée du trangène

L'intégration non-ciblée des rétrovirus au sein du génome de la cellule hôte peut provoquer un phénomène de mutagénèse insertionnelle (**Figure 12**). Celui-ci est caractérié par différents mécanismes dont l'inactivation d'un gène cellulaire et/ou des modifications de l'expression des gènes avoisinant le site d'intégration. Ces dernières impliquent l'activation de proto-oncogènes, l'interférence entre promoteurs, ou encore la transcription au-dela du site de terminaison au sein des gènes hôtes voisins (Nowrouzi et al., 2011). Notamment, il a été récemment décrit que l'intégration d'un transgène au sein de régions transcrites pouvaient induire des épissages aberrants de transcrits chimériques (Moiani et al., 2012). La génération d'une mutagénèse insertionnelle peut aboutir à une dominance clonale voire à une oncogenèse insertionnelle (Pauwels et al., 2009). Un exemple historique est l'activation insertionelle du gène *LMO2* (*LIM domain only 2*, connu comme proto-oncogène) par l'*enhancer* LTR observé après introduction de vecteurs gammarétroviraux, qui a contribué à la leucémogénèse chez cinq patients traités par thérapie génique dans le cas de déficit immunitaire combiné sévère lié à l'X (SCID-X1) dans deux essais cliniques indépendants (Fischer et al., 2011; Hacein-Bey-Abina et al., 2008; Hacein-Bey-Abina et al., 2003; Howe et al., 2008).

Les rétrovirus ont des profils d'intégration différents au sein du génome hôte (*cf.*, § I.B.6). Les préférences de sites d'intégration peuvent ainsi influencer la génotoxicité potentielle des vecteurs (Cattoglio et al., 2007; Nienhuis et al., 2006) et seraient le reflet d'interactions différentielles entre les protéines de l'hôte (*i.e.*, LEDGF/p75 dans le cas des lentivirus; *cf.*, § I.B.6) et la machinerie d'intégration virale (Lewinski and Bushman, 2005). Les lentivirus possèdent donc un profil d'intégration relativement plus sûr que celui des gammarétrovirus (même si l'insertion d'un lentivecteur peut toucher un gène hétérozygote et/ou suppresseur de tumeur) (Pauwels et al., 2009).

Le déterminant majeur de la génotoxicité potentielle du vecteur intégré reste le LTR dont l'inactivation (*i.e.*, *self-inactivating*, SIN) augmente considérablement la sécurité du vecteur transduit, même lorsqu'il est d'origine gammarétrovirale (Montini et al., 2009) (aspect développé dans le § II.B.1).

ii/ Variations de l'expression du transgène

L'altération de l'expression d'un transgène est désignée sous le terme de s*ilencing*. Cependant, ce terme général regroupe différents phénomènes pouvant affecter l'expression d'un transgène après transduction rétrovirale (**Figure 12**) (Ellis, 2005). Il peut y avoir une absence d'expression virale immédiatement après transduction, que nous appelerons s*ilencing* par la suite (**Figure 12**). De plus, le site d'intégration peut impacter positivement ou négativement sur l'expression transgénique. Cet effet de position peut éventuellement conduire à de la variégation qui se caractérise par des différences d'expression au sein d'une population portant le provirus à un même locus (**Figure 12**). Enfin, l'extinction correspond à la diminution progressive de l'expression au cours du temps d'un provirus initialement exprimé (Ellis, 2005).

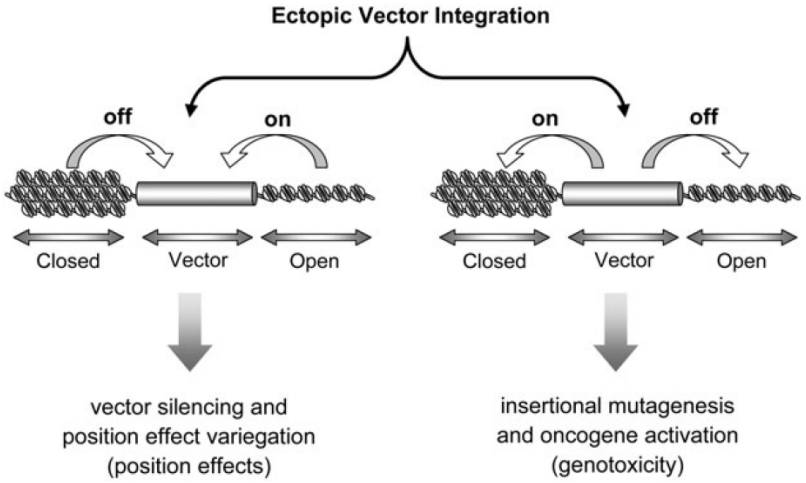


Figure 12 : Conséquences de l'intégration non-ciblée du vecteur sur l'expression transgénique. L'intégration non-ciblée du vecteur a deux impacts majeurs sur son efficacité. (**Gauche**) La structure chromatinienne à proximité du vecteur intégré peut influencer l'expression du transgène, négativement ou positivement. Elle induit ainsi un *silencing* de l'expression et/ou une variégation due à l'effet de position (aussi appelée effet de position du chromosome). (**Droite**) Les séquences du vecteur peuvent aussi influencer l'expression des gènes hôtes, ce qui augmente le risque de mutagénèse insertionnelle et parfois l'activation de proto-oncogènes (aussi appelée génotoxicité médiée par le vecteur) [issue de (Emery, 2011)].

Ainsi, l'expression du transgène peut éventuellement être perdue par *silencing* et/ou extinction après transduction. Il est donc indispensable de la contrôler grâce au choix des promoteurs et l'utilisation de séquences modulatrices de l'expression (*e.g.*, insulateurs, aspects développés dans le § II.B.5) [*cf.*, revue (Emery, 2011)].

iii/ Génération de particules virale recombinantes

Lors de la production de rétrovecteurs incapables de se répliquer, il existe un risque de potentiellement générer et propager un vecteur compétent pour la réplication par recombinaison homologue entre des séquences virales se chevauchant (Bastone and Lochelt, 2004). Ce risque diminue grâce à la séparation des régions codantes pour les protéines structurales des éléments cis associés au gène d'intérêt permettant son transfert (qui augmente de fait le nombre de recombinaisons nécessaires à la génération de vecteurs recombinants). De plus, l'absence de chevauchement entre séquences virales diminue également le risque de mobilisation. La mobilisation est un phénomène de *rescue* de la réplication virale à partir de gènes rétroviraux endogènes (*i.e.*, HERV) ou exogènes lors de thérapie génique effectués sur des patients porteurs du VIH-1 (Levine et al., 2006) [*cf.*, revue (Pauwels et al., 2009)].

iv/ Limites des rétrovecteurs comparés aux autres systèmes

La taille de l'insert (*e.g.*, transgène, séquence modulatrice, promoteur interne) pouvant être introduit au sein des rétrovecteurs est limitée (10 kb) comparée à celle de vecteurs dérivés du virus de l'Herpès simplex (150 kb) ou du virus de la vaccine (50 kb) (Warnock et al., 2011).

Le titre viral maximal obtenu lors de la production de lentivecteurs est faible (*i.e.*, ~10^7 génomes viraux/mL) et difficilement transposable à une production industrielle s'il est comparé à celui obtenu lors d'une production de AAVs (*i.e.*, ~10^{12} génomes viraux/mL) (Warnock et al., 2011). De plus, il ne peut actuellement pas être amélioré grâce à l'utilisation de lignées d'encapsidation qui exprimeraient stablement les (glyco)protéines d'enveloppe (*i.e.*, toxicité) ou d'assemblage (*cf.*, § II.B.5) (Schweizer and Merten, 2010).

B/ LA CONVERSION D'UN LENTIVIRUS EN UN LENTIVECTEUR : DEVELOPPEMENT & *DESIGN*

1/ Evolution du concept du *design* de lentivecteurs

Les premiers vecteurs dérivés du VIH-1 furent décrits dès 1990. Ils étaient simplement basés sur la séparation des séquences codants les protéines structurales et enzymatiques du VIH-1 du gène codant l'enveloppe du virus (pouvant aussi être hétérologue) (Helseth et al., 1990; Page et al., 1990) et ouvraient la voie au pseudotypage des lentivecteurs, déjà existant chez les gammarétrovecteurs, et à l'élargissement de leur tropisme cellulaire.

Le concept du design de vecteurs a ensuite évolué afin d'améliorer l'efficacité et la sécurité des vecteurs en prévenant tout risque de produire des vecteurs recombinants et/ou de mutagenèse insertionnelle en thérapie génique, comme décrit précédemment (*cf.*, § II.A.2). Ces inconvénients seraient également dommageables en recherche puisqu'ils pourraient aboutir à des biais dus à des évènements de recombinaison (*e.g.*, étude de virus non-réplicatifs) et/ou à un changement d'expression au sein des cellules transduites (*e.g.*, transformation cellulaire).

Le développement de rétrovecteurs était principalement basé sur le concept de séparation en différents plasmides des séquences agissant en *cis* essentielles au transfert du génome viral (fournies par le vecteur de transfert/navette) et des éléments *trans* qui codent les protéines virales nécessaires à la formation des particules virales et à leur infectivité (fournis

par les vecteurs d'assemblage et d'enveloppe) (Pluta and Kacprzak, 2009). Ce concept, acquis dans le développement des gammarétrovecteurs, a pour but d'augmenter le nombre de recombinaisons homologues nécessaire à la formation d'une particule virale (*i*) recombinante et/ou (*ii*) issue de la mobilisation de séquences virales déjà présentes (*i.e.*, individus infectés, HERVs), tout en diminuant les homologies des séquences des différents plasmides nécessaires à la production de virions (Pauwels et al., 2009).

L'évolution dans la construction de vecteurs rétroviraux les plus sûrs et performants possibles s'est ensuite caractérisée par le développement successif de différentes générations de navettes avec l'élimination de séquences homologues et/ou ajout de séquences hétérologues pouvant être issues d'autres virus (*e.g.*, tropisme, éléments *cis* améliorant l'infectivité) (Pauwels et al., 2009; Pluta and Kacprzak, 2009).

2/ Production de particules lentivirales

a/ *Principe*

Le protocole de production de particules virales est mis en œuvre à l'aide d'une transfection plasmidique au phosphate de calcium (Chen and Okayama, 1987) ou utilisant des transfectants commerciaux. Au moins trois plasmides codant l'enveloppe, le transgène ou les protéines structurales et enzymatiques sont transfectés de façon transitoire au sein de cellules productrices qui produisent ainsi des particules virales (**Figure 13-haut**).

b/ *Cellules productrices*

La lignée cellulaire HEK (*Human Embryonic Kidney*) 293T est utilisée principalement pour la production de particules virales car la présence de l'antigène T de SV40 (*Simian Virus 40*) lui confère une croissance cellulaire et une efficacité de production virale supérieure à celles des HEK 293 dans les mêmes conditions (Merten et al., 2011). L'utilisation d'autres types cellulaires humains ou de primates (*e.g.*, HeLa, COS-7) réduit également le taux de production virale, à l'exception des cellules COS-1 (Schweizer and Merten, 2010). De façon alternative, des lignées d'encapsidation exprimant stablement ou de façon inductible l'enveloppe (*e.g.*, VSV-G) et/ou les protéines structurales et enzymatiques ont été testées afin de substituer la transfection de plusieurs plasmides qui réduit l'efficacité de production virale (Schweizer and Merten, 2010). Cependant, bien que cette pratique ait été réussie pour la production rétrovirale (*i.e.*, lignée Phoenix) (Coroadinha et al., 2010), elle n'a pas pour le

moment pas été fructueuse pour la production lentivirale du fait de la toxicité des protéines exogènes exprimées (*e.g.*, VSV-G) (Schweizer and Merten, 2010).

c/ *Concentration et stockage*

Les surnageants viraux pseudotypés à l'aide de glycoprotéines d'enveloppe peuvent être concentrés par différentes techniques qui permettent d'augmenter significativement le titre viral des préparations (Kutner et al., 2009). La concentration des productions virales peut se faire par ultracentrifugation (~10^5 *g* pendant 1-4h à 4°C), centrifugation à haute vitesse sur un temps long (~10^4 *g* pendant 8-12h à 4°C), précipitation au phosphate de calcium/polyéthylène glycol ou filtration centrifuge (Sakuma et al., 2012). Il est à noter que l'efficacité de concentration des surnageants viraux peut être affectée par la présence de sérums dans le milieu de culture, notamment en cas de centrifugation à haute vitesse ou de filtration centrifuge (Sakuma et al., 2012). Les productions virales ainsi concentrées peuvent ensuite être aliquotées et stockées à -80°C pour une conservation à long terme.

d/ *Titrations*

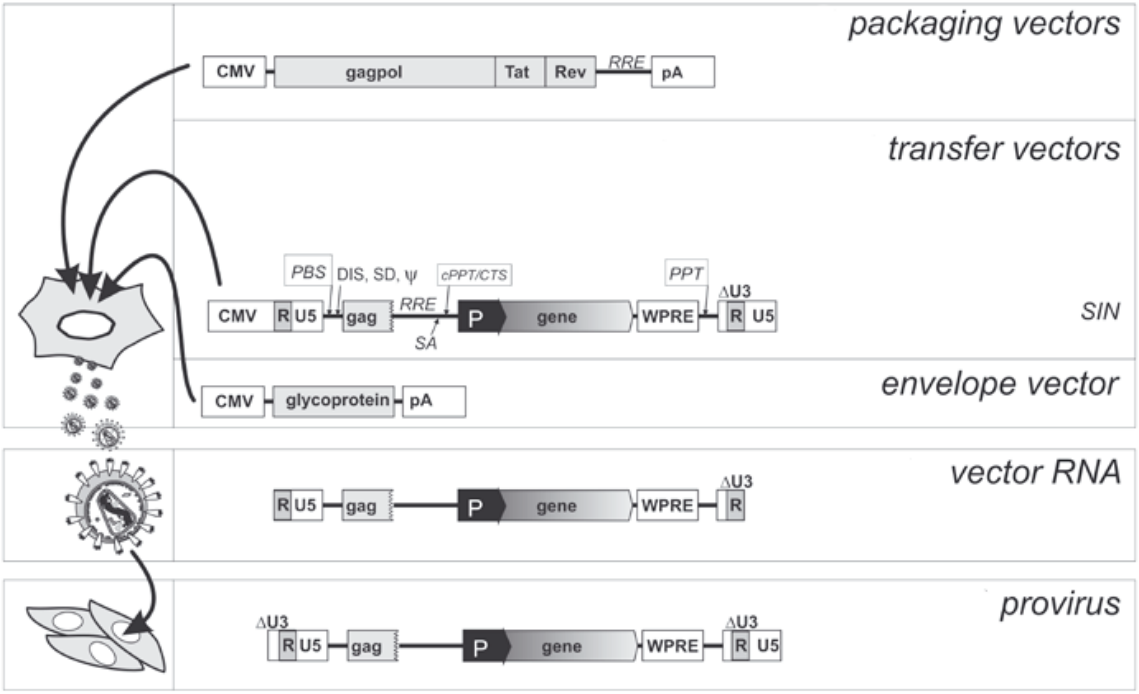


Figure 13 : De la transfection de plasmides à l'intégration d'un vecteur *self-inactivating* (SIN) au sein de la chromatine de la cellule infectée. (**Haut**) Représentation schématique de la méthode de production de particules virales réalisée à l'aide d'une tri-transfection plasmidique. Les virions (contenant deux brins d'ARNg, **Milieu**) ainsi générés sont ensuite transduits (**Bas**) au sein des cellules cibles au sein desquelles l'ARN viral est rétro-transcrit en un ADN qui s'intègre sous la forme de provirus. Il est à noter que la délétion de la séquence U3 du 3'LTR de l'ARN viral est dupliquée lors de la transcription inverse en ADN [adaptée de (Pluta and Kacprzak, 2009)].

Avant transduction des virions (contenant deux brins d'ARNg, **Figure 13-milieu**) au sein de lignées cellulaires ou cellules primaires cibles (**Figure 13-bas**), les surnageants viraux concentrés doivent être titrés à l'aide de différentes méthodes pouvant être complémentaires. Les composants des particules virales peuvent être mesurés en déterminant la concentration de la protéine p24 (par test Elisa), l'activité de la transcriptase inverse, ou le nombre de copies du génome viral (par RT-Q-PCR, *Real-time-Quantitative Polymerase Chain Reaction*) (Sakuma et al., 2012). Le nombre de copies d'ADNv (intégré ou non) et/ou l'expression du transgène des cellules infectées peuvent également être mesurés par Q-PCR ou cytofluorométrie, respectivement (Sakuma et al., 2012).

3/ <u>Développement de différentes générations de vecteurs lentiviraux</u>

La première génération de vecteurs était basée sur la production virale à partir de trois plasmides dont les fonctions d'assemblage étaient fournies par un plasmide codant Env (ou une enveloppe hétérologue, *e.g.*, VSV-G) et un plasmide d'assemblage (*packaging*) exprimant toutes les autres protéines virales (**Figure 14-haut**). L'expression à partir de ces deux plasmides était sous le contrôle d'un promoteur viral (*e.g.*, CMV, RSV) et d'un signal de polyadénylation [poly(A)] hétérologue (*e.g.*, SV40, insuline) (Naldini et al., 1996a; Naldini et al., 1996b). Le vecteur de transfert était composé d'une cassette d'expression encadrée par deux LTRs natifs et portant les séquences nécessaires à l'export des ARN viraux au sein des cellules productrices (RRE, SA et SD, également présentes dans le plasmide d'assemblage), à l'assemblage du génome (Ψ) et à la transcription inverse (PBS, cPPT/CTS, PPT), comme identifié par Parolin *et al.* (Parolin et al., 1994) (description de ces éléments, *cf.*, § I.A.4).

La suppression des gènes viraux accessoires (*vif*, *vpr*, *vpu* et *nef*), importants pour la virulence et la pathogénéicité du VIH-1, du plasmide d'assemblage a ensuite permis de produire les vecteurs de deuxième génération (Zufferey et al., 1997) (**Figure 14-milieu**). La troisième génération fut créée en éliminant le gène *tat* du premier plasmide d'assemblage et en apportant *Rev* à l'aide d'un second plasmide d'assemblage (Dull et al., 1998; Kim et al., 1998) (**Figure 14-bas**). Ce système de production lentivirale a permis d'augmenter la biosécurité des vecteurs générés grâce à la diminution des chevauchements entre régions codantes pour les composants viraux nécessaires à l'assemblage des particules virales et à l'élimination de *tat*.

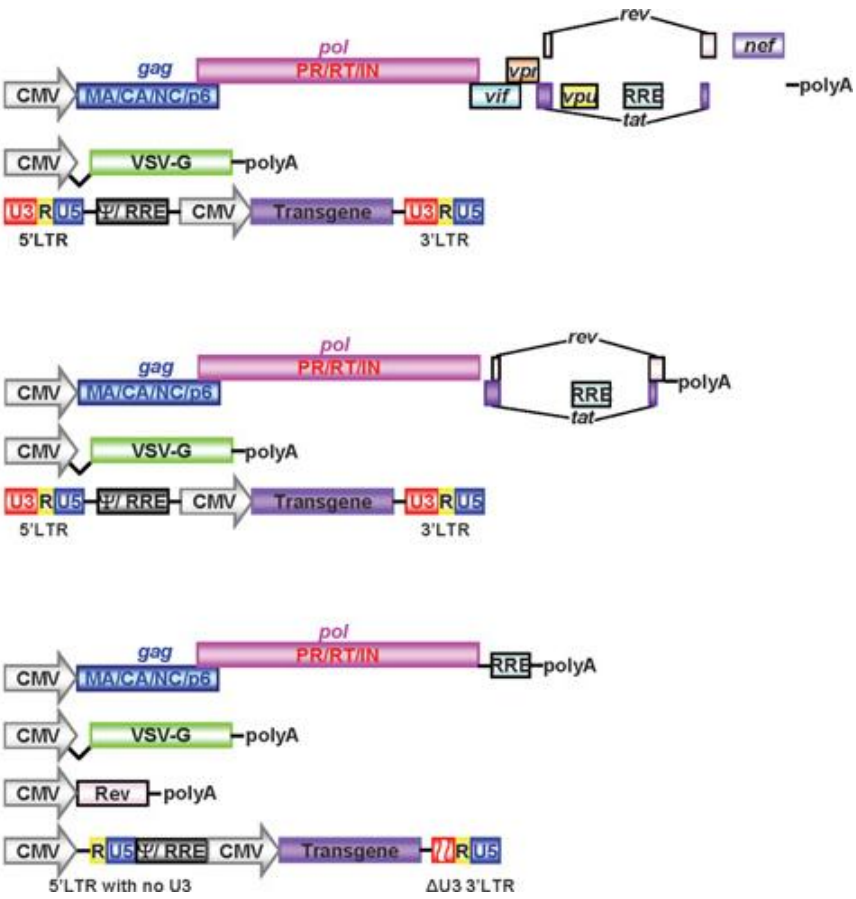


Figure 14 : Représentation schématique des différentes générations de vecteurs dérivés du VIH-1. (**Haut**) La première génération de vecteurs dérivés du VIH-1 inclut toutes les protéines virales, excepté la protéine Env, sur un plasmide d'assemblage. VSV-G est fournit par un plasmide différent, dit d'enveloppe. Le plasmide navette dérivé du VIH-1 contient les LTRs natifs et le transgène sous contrôle d'un promoteur viral fort, le CMV. (**Milieu**) La deuxième génération de vecteurs basés sur le VIH-1 a été générée en excluant toutes les protéines accessoires du plasmide d'assemblage. L'expression de la glycoprotéine d'enveloppe et du transgène ont été fournis par deux plasmides différents identiques à ceux décrits dans le panel du haut. (**Bas**) La troisième génération de vecteurs dérivés du VIH-1 nécessite quatre plasmides différents. En plus des trois plasmides [*i.e.*, plasmides (*i*) d'assemblage, (*ii*) codant l'enveloppe et (*iii*) navette], la protéine Rev est fournit par un plasmide supplémentaire. Le plasmide navette a également été modifié en éliminant complètement la région U3 du 5'LTR remplacée par un promoteur viral fort (*i.e.*, RSV ou CMV) et partiellement la région U3 du 3'LTR [issue de (Sakuma et al., 2012)].

Une stratégie majeure pour l'amélioration de la sécurité des vecteurs lentiviraux a impliqué la délétion des éléments promoteur et *enhancer* localisés dans la région U3 du 3'LTR du vecteur de transfert (Dull et al., 1998; Iwakuma et al., 1999; Miyoshi et al., 1998; Zufferey et al., 1998), déjà appliquée avec succès aux gammarétrovecteurs (Yu et al., 1986) (**Figure 14-bas**). La région U3 du LTR contient des éléments, dont une TATA-*box* et des sites de fixation pour SP-1 (*SV40-promoter specific factor 1*), NF-κB (*Nuclear Factor κB*) et NFAT (*nuclear factor of activated T cells*), essentiels à l'activité du promoteur 5'LTR du

VIH-1 (pour plus de détails, *cf.*, § III.B.1). L'activité promotrice du 5'LTR du vecteur navette permet (*i*) la production d'ARN viral suffisante pour une encapsidation efficace et leur transfert au sein de particules virales et (*ii*) l'expression du vecteur navette intégré au sein des cellules transduites (**Figure 13-milieu**). La délétion partielle de la région U3 du 3'LTR est dupliquée dans le 5'LTR du vecteur intégré (**Figure 13-milieu**) dont l'activité promotrice est considérablement réduite (de 90 %). Les LTRs (et les vecteurs) ainsi générés sont appelés SIN (*self-inactivating*). Toutefois, le promoteur/*enhancer* remplacant le 5'LTR doit être constitutivement actif et suffisamment fort pour assurer une production virale efficace en absence de Tat [*e.g.*, CMV ou promoteur/*enhancer* chimère RSV-HIV du pRRL.SIN.hPGK-gfp, éliminé lors de la transcription inverse (Dull et al., 1998)] (**Figure 14-milieu**).

Les caractéristiques des différentes générations de vecteurs dérivés du VIH-1 sont analysées dans le **tableau 6**.

Tableau 6 : Comparaison des systèmes de productions lentivirales. Il est à noter que le niveau de sécurité lié à l'utilisation des vecteurs lentiviraux *in vivo* est variable en fonction des expériences menées, d'un niveau de sécurité 1 pour le maintien d'animaux inoculés à un niveau 3 pour les animaux greffés avec des cellules humaines transduites. RCL, *replication competent lentivector* [adaptée de (Pauwels et al., 2009)].

	Première génération	Deuxième génération	Troisième génération
Nombre de plasmides transfectés	3	3	4
Délétion dans le 3'LTR (SIN)	Non	Non	Oui
Nombre de plasmides d'assemblage (contenant les gènes du VHI-1)	1	1	2
Gènes accessoires (*vif, vpr, vpu, nef*)	Tous présents	Tous absents	Tous absents
Séquences codantes les protéines Tat et Rev	Présentes sur un même plasmide d'assemblage	Présentes sur un même plasmide d'assemblage	Tat est absente, Rev est codé par un plasmide d'assemblage indépendant
Nombre de recombinaisons pour former un RCL	2	3	4, entre plasmides sans homologie et *trans*-complémentation de la délétion SIN
Niveau de sécurité *in vitro*	3	2	1

D'autres systèmes de productions lentivirales disponibles dans le commerce ont été générés à partir des connaissances issues de l'étude des vecteurs de deuxième et troisième générations. Ils sont notamment basés sur l'utilisation de vecteurs navette non-SIN [Lenti-X™ (Kappes and Wu, 2001)] ou SIN [Translentiviral™ (Kappes and Wu, 2001)] et d'au moins trois plasmides d'assemblage qui séparent les gènes *gag* de *pol* (Wu et al., 2000, Kappes, 2001 #504) ainsi que l'expression de la protéase (Westerman et al., 2007), ce qui limite le chevauchement entre séquences du VIH-1. Toutefois, les efficacités de production sont mises à l'épreuve par le nombre de plasmides nécessaires pour produire la totalité des gènes viraux complémentaires (réduites à partir de quatre plasmides transfectés). La création de vecteurs hybrides associant les composants d'assemblage d'un vecteur de transfert dérivé du VIH-1 à ceux d'un second virus [SIV (White et al., 1999) ou chimères VIH-1/VIH-2 (Sachdeva et al., 2007)] a également été testé afin de réduire la probabilité de recombinaison homologue.

4/ Vecteur de transfert minimal

Le « squelette » du vecteur de transfert ne doit contenir qu'un minimum de séquences virales, ce qui diminue les homologies de séquences tout en permettant l'introduction d'un insert, *i.e.*, transgène et/ou éléments régulateurs (*e.g.*, insulateurs), de taille importante si nécessaire.

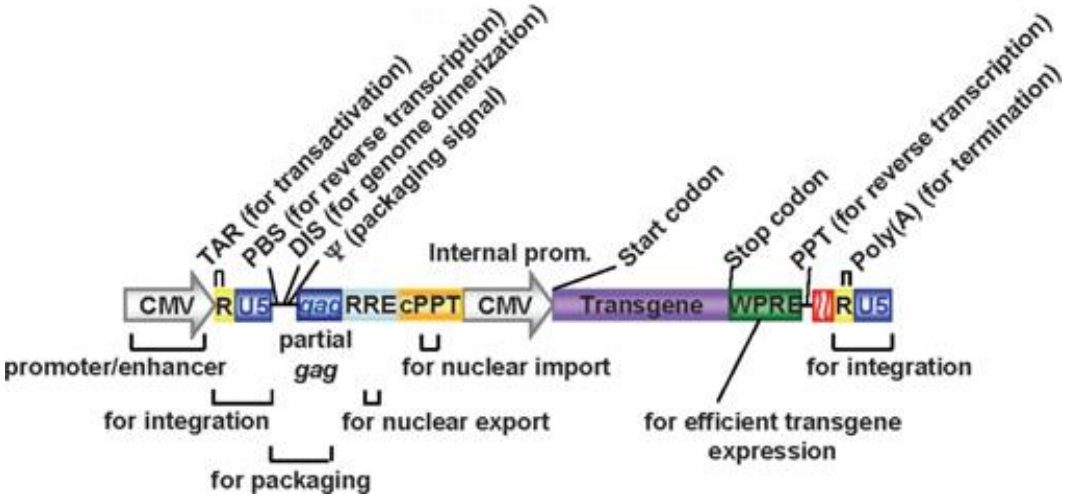


Figure 15 : Eléments *cis* essentiels du vecteur navette. Les position et fonction de chacun des éléments *cis* sont annotées. Ces séquences peuvent être homologues (décrites dans le § I.A.4) ou hétérologues (*i.e.*, promoteur interne CMV, WPRE, décrits dans le § II.B.4). WPRE, WHV (*woodchuck hepatitis virus*) PRE (*post-transcriptional regulatory element*) [issue de (Sakuma et al., 2012)].

Cependant, un certain nombre d'éléments *cis* homologues (décrits dans le § I.A.4) ou hétérologues doivent être présents au sein du vecteur de transfert pour assurer (*i*) une production virale optimale, (*ii*) un transfert efficace du transgène et (*iii*) une expression performante et sûre (**Figure 15**). Ces trois critères, essentiels au *design* d'un vecteur efficace, doivent donc être pris en compte pour assurer un compromis entre la taille du vecteur (dont l'augmentation trop importante peut diminuer le titre viral) et son efficacité. Ces éléments *cis* peuvent provenir du VIH-1 s'ils sont essentiels aux trois paramètres cités ci-dessus (ce qui a pu conduire à leur ré-insertion, *i.e.*, cPPT) ou être remplacés/améliorés par des séquences hétérologues (**Figure 15**, **Tableau 7**).

Tableau 7 : Exemples d'élément *cis* essentiels à un vecteur minimal basé sur le VIH-1.

Eléments *cis* essentiels	Origine	Fonctions (références)
cPPT	VIH-1	Présence de la séquence cPPT-CTS indispensable à l'obtention d'une bonne efficacité de transduction *in vitro* et *in vivo*, qui est fortement stimuler (> 10 fois) [comparé aux vecteurs ne le contenant pas (Follenzi et al., 2000; Park and Kay, 2001; Sirven et al., 2000; Van Maele et al., 2003)].
RRE	VIH-1	Permet l'assemblage de l'ARN (Parolin et al., 1994). Réduction significative du titre viral en son absence [de 5 fois comparé aux systèmes contenant Rev/RRE, (Kotsopoulou et al., 2000)]. Indépendance vis-à-vis de Rev/RRE possible grâce à l'insertion de plusieurs séquences du CTE (*constitutive transport element*) du MPMV (*Mason-Pfizer monkey virus*) (Bray et al., 1994; Oh et al., 2007) dans le vecteur de transfert et/ou le plasmide d'assemblage, pour obtenir des titres viraux équivalents à ceux des systèmes contenant Rev/RRE (Srinivasakumar et al., 1997).
WPRE	WHV	Augmente la quantité d'ARN non-épissés (Zufferey et al., 1999), ce qui permet d'améliorer la performance du transfert de gène (expression du transgène dans les cellules transduites, titre viral) (Oh et al., 2007; Schambach et al., 2007). Contient le gène *x* tronqué impliqué dans l'oncogénese (Kingsman et al., 2005) sous forme mutée (Zanta-Boussif et al., 2009).

Malgré tous les efforts fournis pour obtenir un *design* de vecteurs sûrs et performants, un certain nombre de challenges reste encore à accomplir afin, notamment, de sécuriser l'expression du transgène, en ciblant par exemple son intégration ou l'isolant de façon spatio-temporelle. Toutefois, toutes les connaissances sur la biologie des (rétro)virus ne sont pas acquises et il est actuellement impossible d'anticiper tous les problèmes qui pourront se poser après modification du *design* de vecteur. De plus, il est essentiel d'avoir une meilleure connaissance des interactions vecteur/hôte afin de pouvoir également anticiper les problèmes posés par les interactions du vecteur avec son hôte dans un contexte pathologique.

5/ Stratégies pour l'amélioration des vecteurs lentiviraux : contrôle et sécurité de l'expression trangénique

a/ *Ciblage de l'entrée du virus au sein de la cellule hôte*

Un pseudotypage optimal des lentivecteurs permet de cibler spécifiquement les cellules à transduire afin de limiter la toxicité due à une transduction cellulaire non-appropriée. L'utilisation de l'enveloppe VSV-G, historique pour le pseudotypage des gammarétrovecteurs et lentivecteurs, est non-spécifique (*i.e.*, enveloppe amphotrope), peut être cytotoxique et ne permet pas la transduction efficace des cellules quiescentes [du à un défaut de stimulation des étapes précoces, *e.g.*, transcription inverse (Serafini et al., 2004)]. Une entrée sélective des lentivecteurs nécessite alors un pseudotypage exploitant le tropisme naturel d'autres virus, à l'aide de glycoprotéines d'enveloppes d'autres rétrovirus [*e.g.*, glycoprotéines H et F du virus de la rougeole qui assure l'activation transitoire des cellules B quiescentes (Frecha et al., 2009)] ou d'autres virus enveloppés [*e.g.*, enveloppe neurotropique du virus de la rage (Federici et al., 2009)]. Le pseudotypage peut aussi être créé artificiellement (génétiquement et/ou chimiquement) en utilisant des enveloppes recombinantes/chimères pour le ciblage de molécules de surface cellulaires particulières [domaines de glycoprotéines virales fusionnés entre eux (Zhang et al., 2004) ou avec des domaines d'interactions, *e.g.*, ligands (Froelich et al., 2009) ou anticorps (Ziegler et al., 2008)]. Un pseudotypage alternatif à l'enveloppe VSV-G pourrait cependant générer des vecteurs plus fragiles/instables (Strang et al., 2004).

b/ *Utilisation de vecteurs déficients pour l'intégration*

Une stratégie alternative pour éviter les effets négatifs de l'intégration repose sur l'utilisation de lentivecteurs non-intégratifs (NILVs), aussi appelés lentivecteurs défectifs pour l'intégration (IDLVs).

Ces vecteurs lentiviraux ne peuvent intégrer leur génome au sein de la chromatine hôte du fait de mutation(s) introduite(s) dans l'IN (**Figure 16-gauche**). L'IN a trois domaines fonctionnels (*i.e.*, *core* catalytique et extrémités C- et N-terminales, **Figure 16-gauche**) et les mutations générées se sont pas toutes égales. En effet, elles peuvent affecter uniquement l'intégration mais aussi les autres rôles de l'IN dans le cycle du VIH-1 (*i.e.*, mutations de classe I et II, respectivement) (**Figure 16-droite**). La génération d'IDLVs met en œuvre des mutations de classe I, en particulier au sein du *core* catalytique de l'IN (*e.g.*, résidus D116, D64, E152) (**Figure 16-gauche**). Ces mutations de l'IN peuvent être accompagnées de

modifications des sites *att* du LTR (**Figure 16-droite**) et/ou substituées par le traitement des cellules transduites avec des inhibiteurs de l'intégration (*i.e.*, raltégravir) [*cf.*, revue (Banasik and McCray, 2010)].

En absence d'intégration, les espèces d'ADN virales majoritaires sont les cercles à 1- et 2-LTRs qui, comme décrit dans le paragraphe I.B.4, sont normalement issus d'une intégration abortive (**Figure 16-droite**) et peuvent être le support d'une expression virale. Ces épisomes étant stables, ils peuvent se maintenir et garantir une expression transitoire ou stable au sein de cellules transduites différenciées ou quiescentes, respectivement, ces dernières étant les cibles préférées des transductions par des lentivecteurs (Poon and Chen, 2003; Saenz et al., 2004). Les cercles à 1- et 2-LTRs permettent uniquement une expression transitoire dans les cellules en division desquelles ils sont éliminés par dilution. Cependant, l'introduction d'éléments/signaux de maintien [*i.e.*, SAR/MAR (*scaffold/matrix-attached region*), origine de réplication du SV40] au sein des formes circulaires d'ADNv permettrait de les maintenir dans les cellules qui prolifèrent (Baiker et al., 2000 ; Glover et al., 2005; Vargas et al., 2004) [*cf.*, revue (Wanisch and Yanez-Munoz, 2009)].

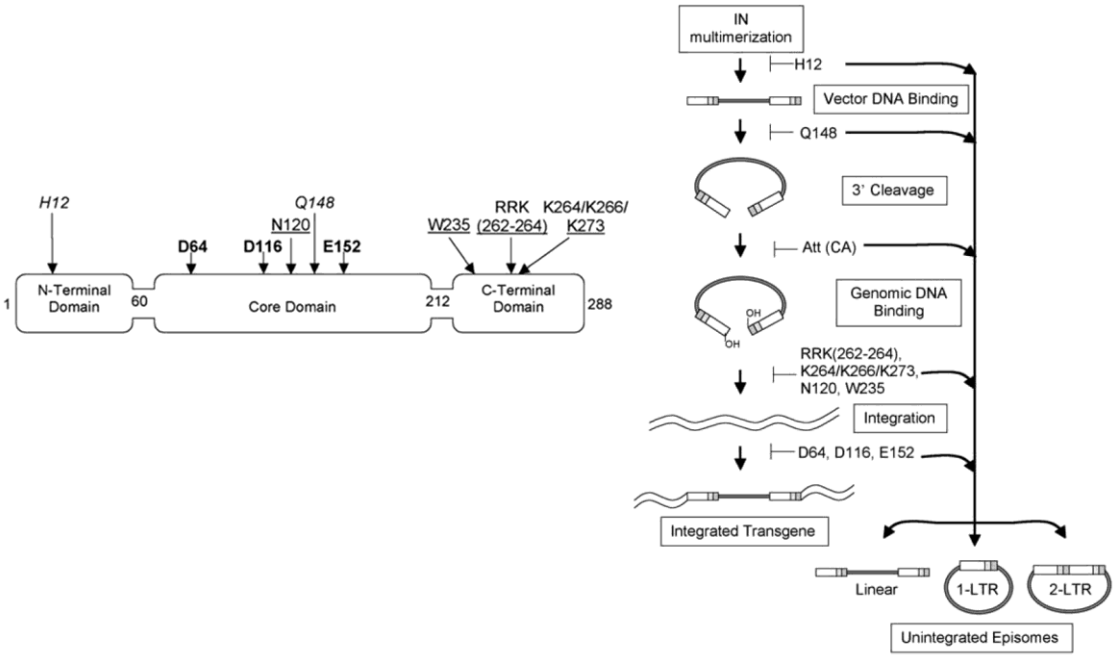


Figure 16 : Mutations de l'IN et conséquences sur les étapes du processus d'intégration. (**Gauche**) Représentation schématique des trois domaines de l'IN et des mutations (indiquées par des flèches) permettant de l'inactiver dans le cadre des approches IDLV. Les mutations affectant l'activité catalytique de l'IN sont indiquées en gras. Les mutations affectant la fixation de l'ADN sont soulignées. Les autres mutations sont en italique. (**Droite**) Les conséquences des différentes mutations affectant les étapes du processus d'intégration (détaillé dans la **figure 7**) sont indiquées [adaptée de (Banasik and McCray, 2010)].

Dans ce contexte, une approche associant IDLVs et ZFNs (*zinc-finger nucleases*), qui reconnaissent et clivent un site génomique unique (Miller et al., 2007), a permis de générer une correction génique parfaite (par recombinaison homologue) en ciblant un site d'insertion du transgène au sein de cellules différenciées et souches (Lombardo et al., 2007). Un *design* expérimental similaire associant IDLVs et méganucléases (nucléases, sauvages ou génétiquement modifiées, de l'ADN qui reconnaissent des séquences rares) a mis en évidence une intégration site-spécifique pouvant être multiple (Izmiryan et al., 2011)

Toutefois, la transduction d'IDLVs peut induire une intégration résiduelle d'ADNv au sein des cellules transduites à un taux généralement très bas (proche de celui observé après transfection plasmidique) de façon intégrase-dépendante et/ou -indépendante. Cependant, des évènements d'intégration, même rares, peuvent conduire à une mutagenèse insertionelle *in vivo* [*cf.*, revue (Wanisch and Yanez-Munoz, 2009)].

c/ *Ciblage de l'intégration* du transgène

Des tentatives de ciblage de l'intégration lentivirale ont été mises en œuvre à l'aide d'IN génétiquement modifiées fusionnées à des méganucléases (Grizot et al., 2010) ou encore, en modifiant les partenaires cellulaires de l'IN [*i.e.*, LEDGF/p75 (Ciuffi et al., 2006; Ferris et al., 2010)], sans succès véritable pour le moment.

d/ *Isoler spatialement la cassette d'expression*

L'utilisation d'éléments génétiques (structurants, *i.e.*, insulateurs et S/MARs) était une stratégie prometteuse pour rendre les vecteurs plus sûrs en isolant l'expression à partir d'un transgène intégré dans l'espace et en la rendant ainsi indépendante du contexte chromatinien hôte (*cf.*, § II.A.2). Les insulateurs agissent comme des frontières génétiques au sein de la chromatine en empêchant les interactions promoteur/*enhancer* (fonction *enhancer blocker*) et/ou bloquant la propagation de l'hétérochromatine (fonction barrière) [*cf.*, § III.A.2 et revue (Emery, 2011)]. Les S/MARs permettent d'ancrer la chromatine à la membrane nucléaire, l'isolant ainsi géographiquement (Heng et al., 2004). Ces éléments peuvent être introduits dans la séquence U3 délétée ou en amont et aval du transgène au sein de vecteurs SIN.

Cependant, l'utilisation de lentivecteurs SIN contenant des insulateurs chromatiniens pourrait ne pas résoudre le problème de mutagénèse insertionnelle. Ainsi, un cas de dominance clonale a été détecté chez un des patients d'un essai clinique soignant la β-thalassémie et utilisant des cellules souches hématopoiétiques transduites par un vecteur SIN

dérivé du VIH-1 contenant l'insulateur cHS4 (*chicken hypersensitive site-4*) (Cavazzana-Calvo et al., 2010). La dominance clonale résulte de l'intégration du vecteur dans le gène *HMGA2* (*high mobility group AT-hook 2*), dont l'activation transcriptionnelle est associée à des tumeurs bénignes et malignes (Cavazzana-Calvo et al., 2010).

e/ *Restreindre l'expression transgénique de façon spatio-temporelle*

Les premiers promoteurs à être utilisés dans un contexte gammarétroviral étaient d'origine virale (*e.g.*, issus des virus CMV, SV40) ou humaine (*e.g.*, issu du gène codant la *phosphoglycérate kinase*, PGK). Cependant, ces promoteurs induisent respectivement une expression importante ou faible en fonction de leurs caractéristiques (*cf.*, § III.C), et ne permettent une expression régulée. La régulation de l'expression peut se faire dans l'espace (au niveau d'un organisme) grâce à l'utilisation de promoteurs tissu-spécifique (*e.g.*, neurones, hépatocytes, etc), ou dans le temps *via* des promoteurs régulés au cours du développement ou inductibles (*i.e.*, avec l'ajout/absence d'agents pharmacologiques). Ces promoteurs peuvent être introduits comme les éléments décrits dans le paragraphe précédent au sein des LTRs ou directement dans le vecteur [*cf.*, revue (Pluta and Kacprzak, 2009)]. Il est à noter que l'introduction d'inserts dans les LTRs est une pratique courante dans le *design* de vecteurs dont l'impact réel n'a cependant jamais été clairement décrit.

f/ *Améliorer la terminaison de transcription du transgène*

L'amélioration de la terminaison de transcription peut à la fois améliorer l'expression du transgène et la biosécurité du vecteur en diminuant notamment la fréquence de transcription au-delà du site de terminaison [due à la délétion dans U3 d'*enhancer* du signal de polyadénylation (Zaiss et al., 2002)]. Des séquences hétérologues augmentant la terminaison de la transcription ont ainsi été apportées dans le vecteur de transfert. Il s'agit d'éléments qui favorisent la reconnaissance du signal de poly(A) [*e.g.*, USE (*upstream enhancer elements*) dérivé du SV40 (Schambach et al., 2007)].

III/ L'EXPRESSION LENTIVIRALE ET SA REGULATION

A/ ORGANISATION STRUCTURALE ET SPATIALE DE LA CHROMATINE CELLULAIRE

L'ADN du VIH-1 s'intègre préférentiellement dans les gènes hôtes transcriptionnellement actifs qui seraient plus accessibles au sein de la chromatine (aspects

développés dans le § I.B.4). L'organisation de la chromatine pourrait ainsi influencer le devenir des différentes espèces d'ADN virales.

1/ Organisation de la chromatine cellulaire

a/ *Différents niveaux de condensation de la chromatine cellulaire*

L'ADN génomique (ADNg) est empaqueté de façon hiérarchisée sous forme de chromatine pour pouvoir être contenu dans le noyau. L'unité structurale et fonctionnelle répétée qui constitue la chromatine est le nucléosome, constitué de 146 bp d'ADN enroulées autour d'un octamère d'histones (*i.e.*, 2 molécules de H2A, H2B, H3 et H4) (Luger et al., 1997). Cette fibre d'un diamètre de 11 nm, appelée '*beads-on-a-string*', représente le premier niveau d'organisation chromatinienne (Luger et al., 1997). La fixation d'une histone *linker* (H1 ou H5) sur cette fibre la condense en une structure de 30 nm de diamètre, second niveau d'organisation de la chromatine (Robinson et al., 2006).

Des niveaux supérieurs de condensation chromatinienne existent et sont caractérisés par des interactions intra-chromosome (*e.g.*, entre *enhancer* et promoteur) et/ou inter-chromosome (*e.g.*, insulateurs et chromatine) pouvant impliquer la formation de boucles chromatiniennes (*i.e.*, *chromatin looping*) de stabilité variable (Deng and Blobel, 2010; Li and Reinberg, 2011).

b/ *Dynamique de la condensation de la chromatine cellulaire*

La dynamique de la condensation de la chromatine est critique dans la régulation de l'expression génique puisqu'elle détermine l'accès à l'ADN aux protéines qui s'y fixent de façon spatio-temporelle. La compaction de la chromatine peut être modulée grâce à un certain nombre de mécanismes, dont l'incorporation de variants d'histone (*e.g.*, H2A.Z), les modifications post-traductionelles covalentes des queues d'histones (*e.g.*, par acétylation, méthylation), la méthylation de l'ADN (associée, le plus souvent, à une inactivation de la transcription), et la fixation de protéines architecturales non-histone (*e.g.*, complexes de remodelage ATP-dépendant) (Kim et al., 2009). Ces modifications réversibles des protéines et/ou de l'ADN au sein de la chromatine sont appelées épigénétiques car elles n'affectent pas la séquence de l'ADN mais sont transmissibles au cours des divisions cellulaires. Un exemple caractéristique est l'hypoacétylation des histones catalysée par des enzymes spécialisées, les déacétylases d'histones (HDACs) et corrélée à une répression transcriptionnelle, alors que

leur hyperacétylation par les acétyltransférases d'histones (HATs) induit une activation de la transcription (Li et al., 2007).

c/ *Dynamique de la localisation nucléaire de la chromatine cellulaire*

Au sein du noyau, on distingue l'euchromatine qui correspond à des régions du génome décondensées généralement associées à des gènes transcriptionellement actifs, de l'hétérochromatine, qui se réfère à des zones très condensées transcriptionnellement inactives [sous une forme facultative (réversible) ou constitutive] (Huisinga et al., 2006). Euchromatine et hétérochromatine sont respectivement localisées au centre et à la périphérie du noyau. Cependant, l'euchromatine est aussi retrouvée associée à la périphérie nucléaire, en particulier aux pores nucléaires (Geyer et al., 2011).

L'organisation spatiale des chromosomes au sein du noyau a une fonction régulatrice essentielle pour l'activation transcriptionnelle. L'initiation de la transcription implique une dynamique de la chromatine qui peut être un mouvement vers des zones du noyau favorables à la transcription, au niveau notamment de corps nucléaires contenant la machinerie nécessaire à l'expression génique (*i.e., transcriptional factories*) (Li and Reinberg, 2011).

2/ Exemple de séquences structurantes de la chromatine : les insulateurs

L'architecture génomique et la régulation des gènes nécessitent des séquences d'ADN particulières, les insulateurs, qui empêchent des interactions inappropriées entre domaines chromatiniens adjacents en participant à la formation et/ou au maintien de frontières physiques et fonctionnelles au sein du génome (Vogelmann et al., 2011).

Deux classes majeures d'insulateurs ont été identifiées des levures à l'homme: (*i*) les insulateurs *enhancer blockers* établissent des domaines qui séparent *enhancers* (et parfois *silencer*) et promoteurs et bloquent ainsi leur interaction, et (*ii*) les insulateurs barrières préviennent le *silencing* des gènes euchromatiniens en créant une barrière contre la propagation de l'hétérochromatine avoisinante, susceptible de donner lieu à de la variégation d'expression (Vogelmann et al., 2011) (**Figure 17**).

L'insulateur le mieux caractérisé chez les Vertébrés à ce jour dérive du site 4 hypersensible à la DNase du LCR (*locus control region*) du cluster de gènes codants la β-globine chez le poulet (appelé cHS4) et est un des rares insulateurs à posséder les deux propriétés décrites ci-dessus (Emery, 2011).

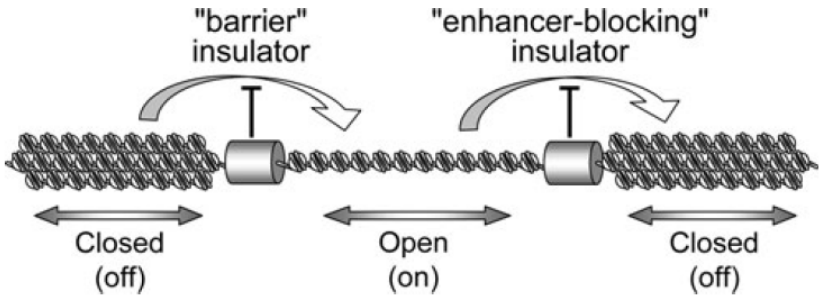


Figure 17 : Propriétés des insulateurs. (**Gauche**) Les insulateurs « barrières » empêchent la propagation des régions d'hétérochromatine à l'euchromatine avoisinante en les séparent physiquement. (**Droite**) Les insulateurs *enhancer*-bloqueurs se positionnent en double-copie entre un *enhancer* et un promoteur et empêchent tout contact entre ces deux éléments [adaptée de (Emery, 2011)].

Les protéines qui fixent les insulateurs sont essentielles à leur activité. Chez les Vertébrés, seul le facteur CTCF (*CCCTC-binding factor*) a été caractérisé comme protéine insulatrice. Cependant, des homologues de protéines insulatrices caractérisées chez la drosophile et la levure ont été récemment identifiés chez l'homme (Vogelmann et al., 2011). CTCF est notamment responsable de l'activité *enhancer blocker* de l'insulateur cSH4 (Recillas-Targa et al., 2002). Chaque protéine insulatrice interagit avec un réseau unique de partenaires (composants de la machinerie de transcription, impliqués dans le remodelage de la chromatine, participant aux modifications d'histones et à l'architecture nucléaire) qui est souvent défini par le contexte chromatinien et la localisation génétique de l'insulateur, ainsi que par la phase du cycle cellulaire et le type cellulaire (Gaszner and Felsenfeld, 2006; Vogelmann et al., 2011).

Les insulateurs peuvent : (*i*) favoriser l'accessibilité à la chromatine, *e.g.*, en interagissant/recrutant avec des facteurs de transcription ou de remodelage pour modifier les nucléosomes/histones avoisinants et/ou en prévenant la méthylation de l'ADN ; (*ii*) permettre l'établissement d'interactions sur de longues distances entre éléments éloignés et/ou le *clustering* de nombreux insulateurs en '*bodies*' ; (*iii*) plus généralement, les insulateurs pourraient être impliqués dans le cloisonnement de la chromatine en interphase pour réguler notamment l'état d'activation des gènes grâce au positionnement de la chromatine [*cf.*, revues (Gaszner and Felsenfeld, 2006; Vogelmann et al., 2011)].

Nous avons décrit dans ce paragraphe le génome de la cellule infectée comme un paysage de domaines intra- et inter-chromosomiques régulés de façon complexe au sein duquel l'ADNv peut s'intégrer et/ou s'exprimer.

B/ L'EXPRESSION DU VIH-1

Parmi les différentes étapes du cycle de réplication du VIH-1, son expression (après intégration) est le principal déterminant du taux de réplication virale qui mène à la progression de la maladie SIDA (Imai and Ochiai, 2011). De plus, l'expression du VIH-1 est dépendante de divers facteurs cellulaires et viraux qui l'initient et la maintiennent en agissant sur le 5'LTR du provirus (Coiras et al., 2009). Notamment, la transcription du VIH-1 est finement régulée au niveau des étapes d'initiation et d'élongation (Karn, 2011). Ainsi, en fonction du contexte cellulaire, l'expression du VIH-1 peut être effective, ou soumise au *silencing*/latence.

1/ Organisation du 5'LTR, promoteur du génome viral

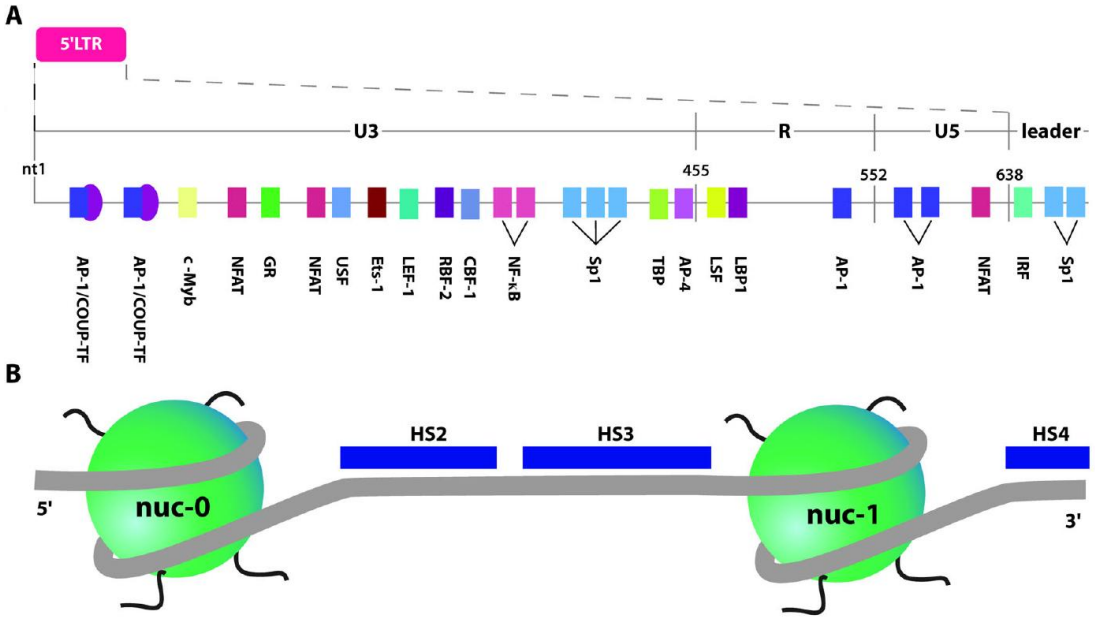


Figure 18 : Organisation du 5'LTR (et de la région *leader*) du VIH-1. Représentations schématiques (**A**) des sites de fixation des principaux facteurs de transcription localisés dans le 5'LTR et au début de la région *leader* du VIH-1 et (**B**) de l'organisation nucléosomale de cette région. (**A**) Les régions U3, R, U5 et *leader* sont indiquées. Le nucléotide 1 (nt1) marque le début de U3 dans le 5'LTR. Le site d'initiation de la transcription (TSS, ou séquence initiatrice) correspond à la jonction entre U3 et R. Le promoteur du VIH-1 est, en outre, constitué des trois sites de fixation à SP-1 et de la *TATA-box*, représentée comme site de fixation à TBP. (**B**) Les sites hypersensibles au traitement à la DNase I (*hypersensitives sites*, HS) n°2, 3 et 4 ainsi que la position des nucléosomes 0 et 1 (nuc-0 et -1) sont indiqués (*cf.*, texte pour les correspondances) [adaptée de (Colin and Van Lint, 2009)].

Une fois intégré dans le génome hôte, Le VIH-1 sous forme provirale se comporte comme un gène cellulaire avec, à ses extrémités 5' et 3', deux LTRs où débute et se termine la transcription, respectivement (Jones and Peterlin, 1994). Ces éléments aux séquences

identiques, jouent différents rôles: le 5'LTR fonctionne comme un promoteur-*enhancer* (**Figure 18**) alors que le 3'LTR contient le site (*i.e.*, la région R) où les transcrits viraux sont polyadénylés (Karn and Stoltzfus, 2012).

a/ *Le 5'LTR, un promoteur*-enhancer *puissant et très optimisé*

Le 5'LTR du VIH-1, constitué d'un promoteur et d'une région *enhancer*, contient de multiples éléments de régulation en amont et aval du site d'initiation de la transcription (*transcription start site*, TSS) qui servent de sites de fixation aux facteurs cellulaires d'initiation de la transcription (**Figure 18A**).

En particulier, la région U3 contient le promoteur composé de trois sites de fixation (en tandem) au facteur de transcription ubiquitaire SP-1 (Jones et al., 1986), une *TATA-box* [-20 à -35 en amont du TSS, (Olsen and Rosen, 1992)] et une séquence initiatrice (Rittner et al., 1995) (**Figure 18A**). En conséquence, le 5'LTR est un promoteur extrêmement efficace capable de supporter de hauts niveaux de transcription, bien plus que le promoteur CMV (Karn and Stoltzfus, 2012).

Le 5'LTR (et le début de la séquence *leader* en aval) possède également une région *enhancer* unique, qui participe à la régulation de l'initiation de la transcription du VIH-1 et contient de nombreux sites de fixation à divers facteurs de transcription, dont deux sites de fixation à NF-κB (Nabel and Baltimore, 1987), et des motifs permettant la fixation de NFAT et ETS-1 (*E26 transformation-specific 1*) (Colin and Van Lint, 2009) (**Figure 18A**). Ces activateurs permettent la réplication virale à un taux important au sein des cellules T activées et des macrophages différenciés (Colin and Van Lint, 2009). Cependant, seuls NF-κB et SP-1 sont indispensables à la réplication du VIH-1 (Perkins et al., 1993; Sune and Garcia-Blanco, 1995). De plus, d'autres sites de fixation ont été identifiés pour NFAT, AP-1 (*activating protein-1*), SP-1 et IRF (*interferon-responsive factor*) en aval du TSS dans la région U5 du 5'LTR et le début de la séquence *leader* en aval (el Kharroubi and Verdin, 1994 ; Van Lint et al., 1997; Verdin, 1991) (**Figure 18A**). Ces sites de fixation peuvent fonctionner indépendamment de, ou en concert avec, les autres séquences du 5'LTR pour activer la transcription du VIH-1 (Van Lint et al., 1997).

b/ *Organisation nucléosomale du 5'LTR*

Après intégration du VIH-1 au sein du génome, en fonction du contexte cellulaire, une structure chromatinienne condensée peut se former au niveau du 5'LTR. Ainsi, dans des

conditions basales (*i.e.*, en absence d'activation de la transcription), deux nucléosomes appelés nuc-0 et nuc-1 sont positionnés de façon précise au niveau du 5'LTR. Ils rendent inaccessible le TSS mais délimitent deux régions dépourvues de nucléosomes correspondant à la région promoteur-*enhancer* [identifiées comme hypersensibles au traitement à la DNase ; *i.e.*, *hypersensitive sites* (HS)-2 et 3] et à la région régulatrice en aval du TSS (*i.e.*, HS4) (Verdin et al., 1993) (**Figure 18B**). De plus, la région en aval du 5'LTR contient trois autres nucléosomes appelés nuc-2, nuc-3 et nuc-4 (Verdin et al., 1993). Ainsi, la structure de la chromatine peut restreindre réversiblement l'accès du LTR ainsi que celui du génome, ce qui contribue à réguler la transcription du VIH-1 (*i.e.*, initiation et élongation). Cependant, le remodelage des nucléosomes associé à l'expression du VIH-1 à partir du 5'LTR peuvent avoir lieu après activation cellulaire (*e.g.*, disponibilité en facteurs de transcription ; *i.e.*, remodelage du nuc-1 grâce à NFAT qui se fixe à la frontière 3' du nuc-1) ou traitement pharmacologique [*e.g.*, utilisation du TNF-α qui permet le recrutement de NF-κB au niveau du LTR (Verdin et al., 1993)].

2/ Expression du VIH-1 à partir du 5'LTR et régulations mises en jeu

L'activité du provirus est grandement influencée par le statut métabolique et l'état d'activation dans lequel se trouve la cellule hôte. Dans un contexte cellulaire propice à la réplication du VIH-1, comme au sein des lymphocytes T CD4 activés, l'expression du VIH-1 peut être régulée à différentes étapes en fonction de la présence/absence de facteurs de transcriptions et/ou protéines virales. Ces différentes étapes sont : (*i*) l'initiation de la transcription et (*ii*) son élongation ainsi que (*iii*) la maturation/export des ARNv transcrits.

a/ *Expression du VIH-1 à partir du 5'LTR intégré*

i/ L'initiation de la transcription du VIH-1

L'activation de la cellule T $CD4^+$ permet la translocation au sein de son noyau de facteurs de transcription indispensables à l'initiation de la transcription du VIH-1 grâce à leur fixation sur le 5'LTR en conformation accessible. C'est le cas de NF-κB, sous forme hétérodimérique p50/p65, libéré de son inhibiteur IκBα (*inhibitor of NF-κB α-subunit*) cytoplamisque (Nabel and Baltimore, 1987), et de NFAT (Okamura et al., 2000). NFAT interagit au sein du 5'LTR au niveau de sites qui chevauchent les sites de fixation de NF-κB dans U3, ce qui suggère des fixations mutuellement exclusives de ces deux facteurs (Cron et al., 2000) (**Figures 18A et 19**).

Ainsi, la transcription est initiée par la présence, au sein du promoteur-*enhancer* 5'LTR, de facteurs de transcription ubiquitaires (*e.g.*, SP-1) et inductibles (après activation/différenciation cellulaire, *e.g.*, NF-κB sous forme p50/p65) qui permettent le recrutement de TFIID (*transcription factor II D*), dont la sous-unité TBP (*TATA-binding protein*) fixe la TATA-*box*, et du complexe de pré-initiation de la transcription, dont l'ARN pol II. L'initiation n'est en réalité effective que lorsque TFIIH (aussi appelé XPB), qui possède des activités ATPase/hélicase/kinase, phosphoryle de l'extrémité CTD (*carboxyl-terminal domain*) de l'ARN pol II et permet ainsi l'ouverture de l'ADN et la libération du complexe de pré-initiation de la transcription du promoteur (Karn and Stoltzfus, 2012). Ce complexe devient alors complexe d'élongation précoce de la transcription.

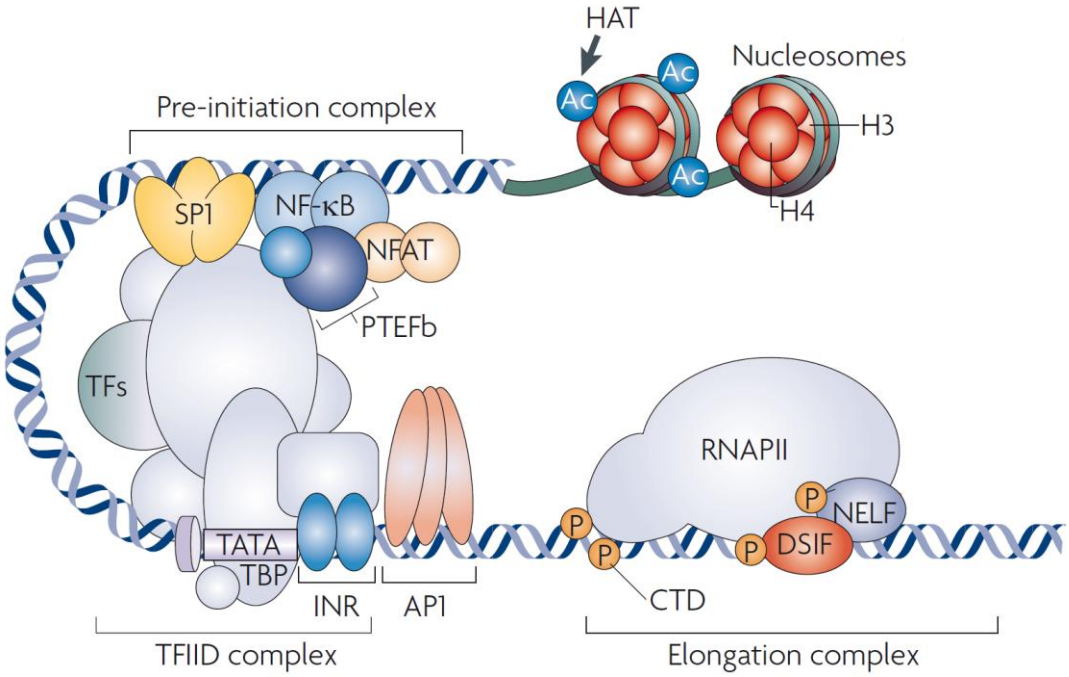


Figure 19: Initiation de la transcription au sein du 5'LTR. La fixation des facteurs de transcription cellulaires (TFs), dont NFAT, NF-κB (la fixation de ces TFs étant mutuellement exclusives), SP-1 et AP-1, associés avec la machinerie transcriptionnelle basale au niveau du 5'LTR forme le complexe de pré-initiation de la transcription. Le recrutement de ces différents partenaires est possible grâce à l'accessibilité accrue du 5'LTR médiée par les HATs qui acétylent (Ac) les queues d'histones (H3, H4). L'initiation de la transcription est effective lors de la phosphorylation de l'extrémité CTD de l'ARN polymérase II (RNAPII) et la libération du promoteur par le complexe d'élongation précoce. Celui-ci est cependant bloqué en absence de Tat et associé aux facteurs répresseurs DSIF et NELF [adaptée de (Coiras et al., 2009)].

ii/ Elongation de la transcription du VIH-1

En absence de Tat, la transcription est initiée mais bloquée à proximité du promoteur à une étape d'élongation précoce (**Figure 20A**). Cette pause du complexe d'élongation est un

mécanisme de régulation de la transcription des gènes eucaryotes nécessitant une réponse rapide. Dans le cas du VIH-1, elle est due à la présence de N-TEF [*negative transcription elongation factor*, composé de NELF (*negative elongation factor*) et DSIF (*DRB-sensitive inducing factor*)] qui fixe l'ARN pol II hypophosphorylée (Ping and Rana, 2001), ainsi qu'à un environnement chromatinien condensé (**Figure 20A**). Ainsi, de courts transcrits de 59 nt s'accumulent dans le cytoplasme (Kao et al., 1987).

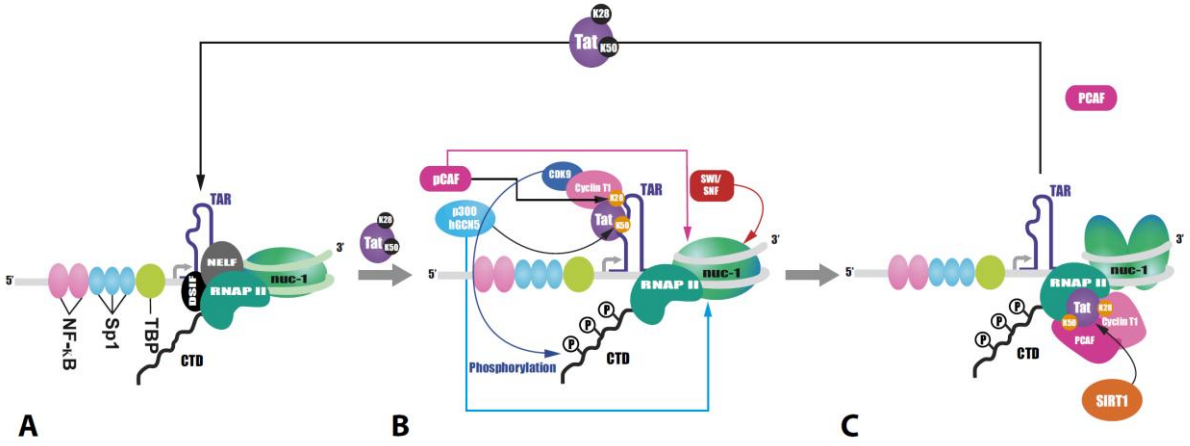


Figure 20: Mécanisme d'activation de l'élongation de la transcription par Tat. (**A**) En absence de Tat, la transcription à partir du 5'LTR produit principalement de courts ARNm contenant la structure TAR à leur extrémité 5'. N-TEF, composé de NELF et DSIF, fixe l'ARN pol II (RNAPII) hypophosphorylée et empêche l'élongation de la transcription. (**B**) Une fois synthétisée, Tat migre au noyau, fixe l'ARN tige-boucle TAR, et active l'élongation virale en recrutant pTEFb, composé de Cdk9 et CyclT1. Ce recrutement est augmenté par l'acétylation de Tat par PCAF sur la lysine (K) 28, localisée dans le domaine de transactivation de la protéine virale. Cdk9 phosphoryle le domaine CTD de l'ARN pol II, ce qui augmente sa processivité et permet la dissociation de N-TEF. De plus, l'acétylation de Tat sur la K50 par p300 et Gcn5 promeut le relarguage de pTEFb, la dissociation de Tat/TAR et son transfert au sein du complexe d'élongation de la transcription. Tat y recrute PCAF et le complexe de remodelage ATP-dépendant SWI/SNF, qui expulse le nucléosome 1 (nuc-1). (**C**) A la fin du processus d'élongation, la déacétylation de Tat par la SIRT1 permet sa dissociation de l'ARN pol II et PCAF, et son recyclage pour initier un nouveau cycle d'activation transcriptionelle [issue de (Colin and Van Lint, 2009), *cf.*, texte pour les abréviations du nom des protéines].

Cependant, de rares transcrits complets peuvent permettre la synthèse de quelques molécules de Tat suffisantes pour stimuler l'élongation de la transcription. Tat se fixe alors sur l'élément ARN TAR en tige-boucle présent à l'extrémité 5' de tous les transcripts viraux naissants (Dingwall et al., 1990) et recrute au niveau de son domaine N-terminal P-TEFb (*positive transcription elongation factor b*), composé la cycline T1 et de la kinase Cdk9 (*cyclin-dependent kinase 9*) (**Figure 20B**) (Garber et al., 1998), jusque-là séquestré par le complexe 7SK snRNP (*small nuclear ribonucleoprotein*). Ce recrutement est augmenté par

l'acétylation de Tat par PCAF [*p300/CREB-binding protein (CBP)-associated factor*] sur la lysine (K) 28 (Kiernan et al., 1999). De plus, Cdk9 phosphoryle le CTD (sur la sérine 2) de l'ARN pol II (Zhou et al., 2000), et N-TEF, permettant la dissociation de ce dernier et d'augmenter la processivité de l'ARN pol II (Fujinaga et al., 2004). L'élongation de la transcription est aussi stimulée par l'interaction de la sous-unité p65 de NF-κB avec pTEFb (Barboric et al., 2001) et Cdk7/TFIIH (Kim et al., 2006), qui phosphoryle la sérine 5 du CTD de l'ARN pol II (Zhou et al., 2000) avant de quitter le complexe d'élongation. L'acétylation de Tat sur la K50 par p300 et Gcn5 (*general control of amino acid synthesis protein 5*) favorise la libération de pTEFb, la dissociation de Tat sur TAR, et son transfert ultérieur au sein du complexe d'élongation de l'ARN pol II (Kaehlcke et al., 2003). Tat peut ensuite recruter PCAF, qui joue un rôle dans le remodelage de la chromatine après activation transcriptionnelle. A la fin du processus d'élongation, la déacetylation de Tat par SIRT1 (*sirtuin 1*), une HDAC (de classe III), permet sa dissociation du complexe ARN pol II-PCAF, et son recyclage pour initier un nouveau cycle d'activation transcriptionnelle (Pagans et al., 2005) (**Figure 20C**). Cette fonction pourrait être importante lorsque la quantité de Tat est un facteur limitant, notamment dans les phases précoces de l'infection (Colin and Van Lint, 2009).

iv/ La terminaison et maturation des pré-mARNs

L'élongation efficace des transcrits viraux permet la synthèse de pré-mARNs qui sont coiffés en 5', maturés en 3' et polyadénylés comme les transcrits primaires eucaryotes (Karn and Stoltzfus, 2012). Le clivage en 3' précède la polyadénylation et a lieu en un site polyadénylé [poly(A)] positionné de façon précise. Cependant, dans le cas du VIH-1, ces séquences, qui sont dupliquées dans les régions R des LTRs 5' et 3', ne sont reconnus qu'en 3' notamment grâce à la présence d'un élément USE et de la structure secondaire du site poly(A) particulièrement propice au clivage en 3' (Karn and Stoltzfus, 2012).

Les ARNm du VIH-1 peuvent arborer plus de 40 épissages alternatifs différents (dus à la présence de plusieurs sites donneurs/accepteurs d'épissage ; aspect développé dans le § I.A.3) (Karn and Stoltzfus, 2012).

Enfin, Tat pourrait aussi être impliqué dans les processus de maturation des pré-mARNs, *i.e.*, addition d'une coiffe (Chiu et al., 2002), polyadénylation (Calzado et al., 2004) et épissage (Berro et al., 2006).

v/ Export et traduction des transcrits du VIH-1

L'expression du VIH-1 se déroule en deux phases : (*i*) la phase précoce, qui correspond à la synthèse des protéines (précoces) Tat, Rev et Nef à partir d'ARNm épissés, et (*ii*) la phase tardive, qui donne lieu à la production des protéines structurales et accessoires à partir d'ARNm partiellement ou non-épissés exportés vers le cytoplasme grâce à Rev (Karn and Stoltzfus, 2012). Ainsi, contrairement aux transcrits eucaryotes partiellement ou non-épissés dégradés au sein du noyau, ceux du VIH-1 s'y accumulent pour être exportés vers le cytoplasme grâce à Rev. Rev s'oligomèrise sur son élément de réponse (*i.e.*, RRE) présent sous forme d'une tige-boucle allongée, au sein des ARNm mono- ou non-épissés et interagit avec Crm1 (aussi appelé Exportin1) pour être transporté de façon énergie-dépendante au sein des pores nucléaires (Karn and Stoltzfus, 2012).

La traduction des ARNm du VIH-1 nécessite la machinerie cellulaire. Cependant, le mécanisme d'initiation de la traduction reste controversé. Ainsi, la présence de nombreuses structures secondaires au sein de l'ARNm viral (exon 1, *e.g.*, TAR, PBS) pourrait interférer avec un *scan* de la séquence initiatrice (AUG) par le ribosome, bien que celui-ci soit expérimentalement possible (Berkhout et al., 2011). Alternativement, l'initiation de la traduction pourrait impliquer l'entrée interne des ribosomes et/ou l'interaction entre des éléments de l'ARNm et des protéines cellulaires (Karn and Stoltzfus, 2012).

b/ *Régulation de l'expression du VIH-1 par des structures secondaires de l'ADN et/ou de l'ARN viral*

i/ Impact des structures secondaires de l'ARN viral sur l'expression du VIH-1

De nombreuses structures secondaires de l'ARNv régulent l'expression du VIH-1 à différentes étapes du processus. C'est notamment le cas des séquences TAR et RRE, comme développé ci-dessus (*cf.*, § III.B.2, **Figure 21A, B**). De plus, comme décrit ci-dessus, l'initiation de la traduction est régulée par les structures secondaires de l'ARNm présentes dans l'exon 1 (**Figure 21A**). Le changement du cadre de lecture (*frameshift*) qui permet la synthèse du précurseur *gag-pro-pol* tous les 20 précurseurs *gag* est régulé (au moins en partie) par une structure secondaire particulière qui permet une pause du ribosome (**Figure 21A**) (Karn and Stoltzfus, 2012). Enfin, l'analyse de la structure d'ARN du génome du VIH-1 par Watts *et al.* a récemment mis en évidence une corrélation entre l'organisation de l'ARN et les séquences codant les boucles inter-domaines des protéines virales, suggérant que la structure

ARN modulerait la vitesse des ribosomes pendant la traduction de façon à promouvoir un repliement protéique correct (Watts et al., 2009).

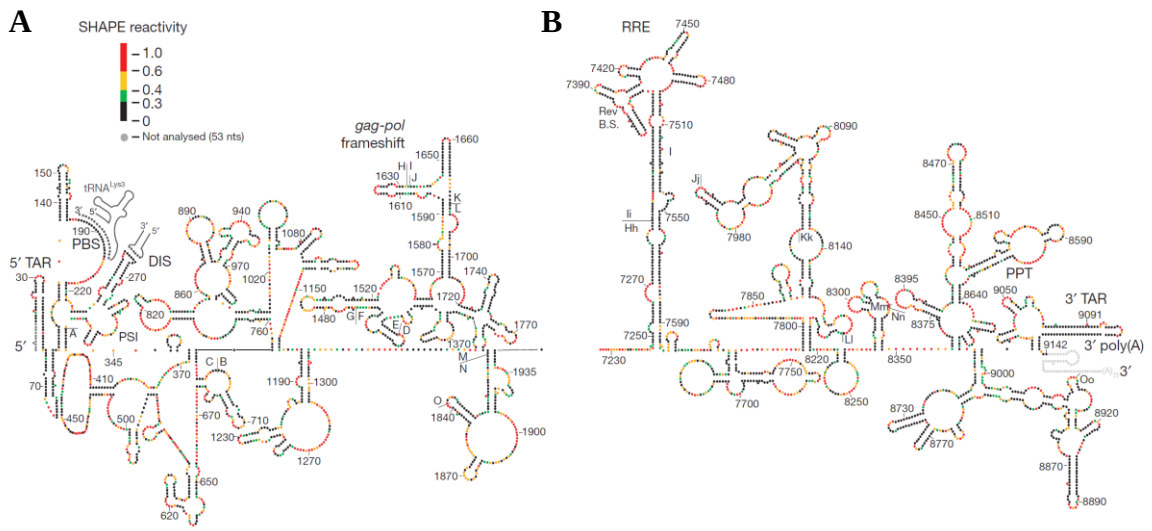


Figure 21 : Structure des extrémités 5' (A) et 3' (B) du génome du VIH-1 sous forme ARN. Ces structures secondaires de l'ARN ont été déterminées par l'analyse SHAPE (*selective 2'-hydroxyl acylation analysed by primer extension*) des ARNs à haut-débit. Chaque nucléotide (nt) est représenté par un rond, dont la couleur indique sa réactivité. L'échelle de réactivité (R) SHAPE représente le niveau de structuration de l'ARN (nt non-structuré : R = 0). Les différentes structures secondaires importantes sont indiquées. Les lettres indiquent les régions codantes des protéines [A/B : début/fin de *gag* (MA), C/D : début/fin de *gag* (CA), E/F : début/fin de *gag* (p2), G/H : début/fin de *gag* (NC), I/K : début/fin de *gag* (p1), J/M : début/fin du peptide en *transframe*, L/O : début/fin de *gag* (p6), N : début de *pol* (PR), Hh : fin de *env* (gp120), Ii/Mm : début/fin de *env* (gp41), Jj : fin de *tat*, Kk : début de l'exon 2 de *tat/rev*, Ll : fin de *rev*, Nn/Oo : début/fin de *nef*] [adaptée de (Watts et al., 2009)].

ii/ Boucles d'ADN/ARN transcription-dépendante

Il a été reporté la formation de boucles d'ADN/ARN transcription-dépendante impliquant des interactions entre promoteur et éléments de terminaison de transcription [poly(A)] chez la levure [*cf.*, revue (Hampsey et al., 2011)], et plus récemment chez l'homme pour le gène *brca1* (Tan-Wong et al., 2008) et le VIH-1 sous forme provirale et plasmidique [**Figure 22A, B** (Perkins et al., 2008)].

Dans le cas du VIH-1, les interactions ont lieu entre le promoteur autologue (*i.e.*, U3) ou hétérologue (*e.g.*, CMV) au sein du 5'LTR et la queue poly(A) située dans la région R du 3'LTR (**Figure 22**). L'expression du génome du VIH-1 serait ainsi régulée par les structures secondaires des séquences d'ADN et d'ARN. Cependant, l'impact de leur modification par insertion/délétion d'éléments hétérologues (dans le cadre du *design* de vecteurs) sur

l'expression transgénique n'est jamais été analysé et fera l'objet d'une partie des résultats de cette thèse.

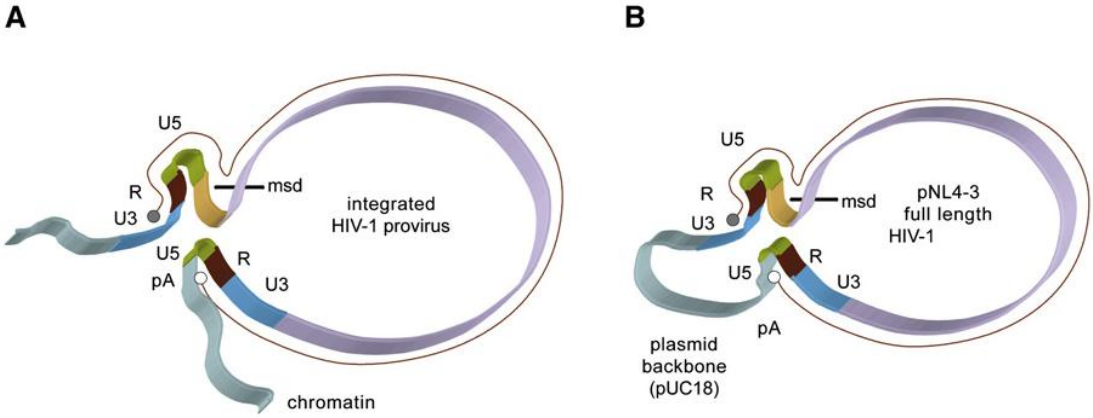


Figure 22: Représentations schématiques des boucles d'ADN-ARN du génome du VIH-1 sous forme provirale (A) et plasmidique (B). Les interactions (potentielles) indiquées entre élèments d'initiation et de terminaison de la transcription ont été analysées par capture 3C (*chromosome conformation capture*). Les lignes fines et épaisses représentent l'ARN et l'ADN, respectivement. Les cercles gris et blanc indiquent les régions transcrites du promoteur (U3) et la queue de polyadénylation (polyA, dans R), respectivement [adaptée de (Perkins et al., 2008)]. MSD, *major site donor*.

3/ <u>Expression à partir de l'ADN viral non-intégré</u>

L'ADNv non-intégré (ADNni), majoritaire après infection, ne permet pas un niveau d'expression nécessaire à la production de virions. L'ADNni existe sous forme d'espèces d'ADNv circulaires (à 1- ou 2-LTRs) stables, et d'ADNv linéaire, rapidement dégradé. Ces trois espèces d'ADNni, et en particulier les cercles à 1-LTR, peuvent servir de substrat pour l'expression du VIH-1 (aspect développé dans le § I.B.4) (Cara et al., 1996). Cependant, l'ADNni peut être sous forme de structure chromatinienne condensée, qui limite l'initiation de la transcription mais qui est réversible par traitement à l'aide d'inhibiteurs de HDACs (Kantor et al., 2009).

Il a été mis en évidence que seuls les transcrits précoces codant Tat, Rev et Nef ainsi que la protéine Nef sont présents dans les cellules infectées par des virus intégrase-défectifs (*i.e.*, contenant uniquement de l'ADNni) (Kelly et al., 2008; Wu and Marsh, 2003). L'expression tardive, qui nécessite une accumulation suffisante de Rev pour permettre l'export des ARNm non-épissés ou mono-épissés, ne peut atteindre un niveau suffisant pour permettre une production virale uniquement à partir d'ADNni (Iyer et al., 2009). De plus, l'expression à partir de l'ADNni a la particularité d'être faible (Gelderblom et al., 2008), mais

pourrait être augmentée par Vpr (Poon and Chen, 2003) et Tat (Cara et al., 1996). Cependant, les quantités de Tat et Nef produites à partir de l'expression d'ADNni sont suffisantes pour l'activation cellulaire et l'augmentation de l'infection productive, suggérant un rôle dans la réplication virale (Wu and Marsh, 2001).

Une hypothèse actuelle propose que l'expression à partir d'ADNni puisse permettre un gain de temps en cas de réinfection de la cellule puisque il y aurait réplication sans intégration du second virus (Sloan and Wainberg, 2011). Ainsi, lors d'une co-infection par (au moins) deux virus (fréquente *in/ex vivo*), il a été montré par Gelderblom *et al.* qu'il y avait (*trans*-)complémentation entre ADNni et ADN intégré, ce qui contribue à la diversification des génomes viraux (Gelderblom et al., 2008).

4/ La latence du VIH-1

a/ *Mise en évidence*

Depuis le milieu des années 1990, les patients infectés par le VIH-1 sont traités à l'aide d'une thérapie antirétrovirale appelée HAART (développé dans le § I.A.3) qui permet de réduire la charge virale en-dessous des limites de détection (par des techniques standards) mais ne permet pas l'éradication du virus (Shen and Siliciano, 2008). En effet, malgré le traitement, des cellules infectées capables de se répliquer persistent au sein de l'organisme. Ces réservoirs viraux correspondent à des sanctuaires anatomiques où la HAART a une faible pénétrance [*e.g.*, système nerveux (Varatharajan and Thomas, 2009)] et/ou à des types cellulaires contenant un provirus latent qui s'exprime peu ou pas (*e.g.*, lymphocytes $CD4^+$ au repos, macrophages) (Colin and Van Lint, 2009).

b/ *Définition*

De la même façon qu'il existe deux phases dans le cycle viral, *i.e.*, avant et après intégration, il existe deux types de latence selon que le virus soit intégré ou non dans le génome de son hôte (Colin and Van Lint, 2009).

La latence pré-intégrative résulte d'un blocage partiel ou complet du cycle viral à une étape précédant l'intégration du virus dans le génome hôte (Colin and Van Lint, 2009). Ce blocage peut être du à une transcription inverse incomplète du fait d'une quantité réduite de dNTP dans les cellules métaboliquement inactives (Gao et al., 1993). Le PIC peut aussi ne pas être importé dans le noyau à cause d'un manque d'ATP (Bukrinsky et al., 1991). La latence

pré-intégrative génère ainsi de l'ADNv (linéaire ou circulaire) susceptible d'être intégré au sein du génome hôte après activation et/ou différenciation cellulaire (Bukrinsky et al., 1991).

La latence post-intégrative, ou transcriptionnelle, a lieu lorsqu'un provirus échoue à exprimer efficacement son génome et est transcriptionnellement inactif (de façon réversible) après intégration dans le génome hôte. Cet état de latence est exceptionnellement stable et uniquement limité par la durée de vie de la cellule infectée et de ses descendants.

c/ *La latence post-intégrative*

La latence des provirus, ou *silencing* transcriptionnel, est médiée par un certain nombre de mécanismes : (*i*) le site et l'orientation de l'intégration, (*ii*) les modifications épigénétiques de l'ADNv, (*iii*) l'absence de Tat, de ses partenaires cellulaires ou d'autres facteurs de l'hôte et (*iv*) l'interférence par les ARNs (ARNi). Ces différents mécanismes, qui régulent positivement/négativement la transcription virale, coopèrent et/ou contribuent à l'établissement et/ou au maintien de la latence post-intégrative du VIH-1.

i/ Site et orientation de l'intégration

Les partenaires (cellulaires, virales) de l'IN et/ou de l'ADNv au sein du PIC pourraient influencer le site d'intégration [*e.g.*, LEDGF/p75 au sein des gènes transcriptionnellement actif, Ku80 en périphérie nucléaire (Masson et al., 2007)] mais aussi permettre le recrutement de facteurs promoteurs/répresseurs de la transcription (aspect développé dans le § I.B.4). Ainsi, IN1, sous-unité du complexe de SWI/SNF, interagit avec l'IN, favorise l'intégration *in vitro* (Lesbats et al., 2011), mais pourrait initier une répression de la transcription du VIH-1 en l'absence de Tat (Boese et al., 2009).

De plus, la majorité des provirus VIH-1 latents est localisée au sein des gènes transcriptionnellement actifs (Han et al., 2004; Lewinski et al., 2005) qui seraient plus accessibles au sein de la chromatine (Schroder et al., 2002). Ainsi, les rares évènements d'intégration au sein de l'hétérochromatine (Carteau et al., 1998) ne permettent pas d'expliquer la latence transcriptionnelle. Le site et l'orientation de l'intégration au sein du génome de l'hôte peuvent cependant générer de l'interférence transcriptionnelle qui jouerait un rôle dans l'établissement de la latence (Colin and Van Lint, 2009). Ce phénomène peut avoir lieu entre des promoteurs actifs divergents (Han et al., 2008) ou convergents (Callen et al., 2004; Crampton et al., 2006). Au contraire, lorsque le provirus est intégré dans la même orientation que le gène hôte, la transcription au-delà de la terminaison (*read-through*) peut

empêcher, par exemple, l'assemblage du complexe de pré-initiation de la transcription au niveau du 5'LTR (Lenasi et al., 2008).

ii/ Modifications épigénétiques

Les modifications post-traductionelles des queues d'histones (*e.g.*, déacétylation, certaines méthylations) modulent les réarrangements chromatiniens et réduisent l'accessibilité des facteurs de transcription et à l'ARN pol II en créant un remodelage des nucléosomes et une structure condensée de la chromatine, qui influence la réplication du VIH-1.

Ainsi, la déacétylation des histones peut être initiée par des facteurs répresseurs de la transcription, tels YY1 (*yin and yang 1*) et LSF (*late SV40 factor*) qui répriment la réplication du VIH-1 dans les cellules $CD4^+$ T infectées (Romerio et al., 1997) en recrutant HDAC1 (Coull et al., 2000). C'est aussi le cas d'autres facteurs de transcription, comme NF-κB sous forme d'homodimères p50 (Williams et al., 2006) ou CBF1 (*C-promoter binding factor 1*) (Tyagi and Karn, 2007) dans de nombreuses lignées cellulaires.

De plus, la méthylation des histones peut être associée à l'activation de la transcription ou à son *silencing* des gènes eucaryotes. Dans le cas du VIH-1, la formation de l'hétérochromatine, associée au recrutement de HP1γ (*heterochromatin protein 1 homologue-γ*), a été corrélée à la triméthylation de l'histone 3 sur K9, médiée par la méthyltransférase SUV39H1 (*histone-lysine N-methyltransferase suppressor of variegation 3−9 homologue 1*) et résulte en un *silencing* transcriptionnel du VIH-1 (du Chene et al., 2007). Le co-répresseur CTIP2 [*chicken ovalbumin upstream promoter-transcription factor* (COUPTF)-*interacting protein 2*] peut aussi recruter HDAC1, HDAC2, SUV39H1 et les protéines HP1 pour établir un environnement hétérochromatinien qui mène au silencing du VIH-1 (Marban et al., 2007). Cet état transcriptionnel latent peut être réactivé par traitement cellulaire à l'aide d'inhibiteurs de HDACs qui induisent une ouverture de la chromatine associée à l'expression du VIH-1.

Cependant, le *silencing* transcriptionnel médié à l'aide de modifications épigénétiques différentes serait graduel et impliquerait une initiation et un maintien de la latence. L'initiation de la latence serait médiée par les modifications des queues d'histones qui, à leur tour, pourraient permettre de recruter les partenaires nécessaires au maintien de la latence par méthylation de l'ADN, comme décrit en 2009 par Blazkova et al. (Blazkova et al., 2009). Cependant, il a été décrit très récemment que le 5' LTR des réservoirs viraux issus de patients traités par HAART est rarement méthylé, suggérant l'intervention d'autres mécanismes dans le maintien de la latence (Blazkova et al., 2012).

iii/ Facteurs de l'hôte

Comme cité ci-dessus, l'expression du VIH-1 est fortement influencée par la présence/absence de facteurs de transcription. Un exemple caractéristique est celui de IκBα, qui séquestre NF-κB sous forme p50/p65 dans le cytoplasme. Sa sur-expression inhibe la réplication du VIH-1 et peut être responsable du maintien de la latence dans les cellules T au repos (Coiras et al., 2007). De plus, le régulateur transcriptionel Myc, qui fixe SP1 au niveau des LTRs, forme un complexe qui recrute HDAC1, inhibe l'expression du VIH-1 et est responsable du maintien de la latence (Jiang et al., 2007).

L'absence/la présence de facteurs associés à Tat peut également être une source de latence transcriptionnelle (Colin and Van Lint, 2009). Ainsi, la méthylation de résidus arginine (Arg52 et 53) de Tat par PRMT6 (*arginine N*-methyltransferase 6) (Boulanger et al., 2005) affecte négativement la formation du complexe Tat/TAR/cycline T1 (Xie et al., 2007).

iv/ Voie d'interférence par les ARN

La voie d'ARNi agit comme une barrière contre la multiplication virale (Corbeau, 2008). Un certain nombre de microARNs (ARNmi) cellulaires semble contrôler la réplication du VIH-1 en ciblant les ARNm épissés ou non-épissés excepté celui codant Nef (Huang et al., 2007). La sur-expression de ces ARNmi dans les cellules T $CD4^+$ au repos inhibe la traduction de la plupart des protéines du VIH-1 dont l'absence contribue à la latence virale (*e.g.*, Tat et Rev) (Han and Siliciano, 2007).

Les connaissances issues de l'étude des mécanismes de la latence du VIH-1 permettent actuellement d'établir des approches thérapeutiques [*e.g.*, combinant des inhibiteurs de HDACs et des cytokines (Colin and Van Lint, 2009)] dont certaines sont testées en essais cliniques dans le but d'éradiquer le virus de l'organisme infecté (Trono et al., 2010).

C/ L'EXPRESSION D'UN TRANSGENE AU SEIN D'UN LENTIVECTEUR MODIFIE

L'expression d'un transgène au sein d'un lentivecteur SIN-LTR est possible grâce à l'utilisation d'un promoteur hétérologue (aspect développé dans le § II.B.2-4). Cependant, le choix de ce promoteur est critique et doit être en adéquation avec l'utilisation qui sera faite du vecteur. Dans ce paragraphe, sont présentées les caractéristiques des promoteurs qui seront utilisés pour les travaux présentés dans la suite de cette thèse.

1/ Promoteurs eucaryotes et régulation différentielle de la transcription

Il existe trois types de polymérases chez les Eucaryotes : (*i*) l'ARN polymérase I produit les pré-ARNs ribosomaux (ARNr) maturés en ARNr 28S, 18S et 8S ; (*ii*) l'ARN polymérase II transcrit les gènes codant les ARNs messagers (ARNm) (codant les protéines) et quatre des cinq petits ARN nucléaires (ARNsn) impliqués dans l'épissage des ARNs ; (*iii*) l'ARN polymérase III produit les ARNs de transfert (ARNt), l'ARNr 5S et un ARNsn. Nous nous intéresserons par la suite à la transcription de l'ARN polymérase II (ARN pol II).

a/ *Différents types de promoteurs eucaryotes*

Le promoteur *core* est défini par la région d'ADN adjacente au TSS qui permet la transcription du gène en ARN. La position de ce dernier est déterminée par les multiples séquences d'ADN spécifiques qui l'entourent, situées principalement en amont mais aussi en aval, et qui varient selon les gènes (Valen and Sandelin, 2011). Ces séquences d'ADN représentent des sites de fixation pour les facteurs de transcription (généraux ou spécifiques) qui régulent le recrutement du complexe de pré-initiation de la transcription (Baumann et al., 2010). Elles sont variables en fonction des différents types de promoteurs existants, *i.e.*, à TATA-*box* ou à îlot CpG, comme décrit ci-dessous.

i/ Les promoteurs de type TATA-*box*

Le TSS peut être positionné de façon précise lorsqu'il est situé (30 nt) en aval d'un *pattern* riche en TA, appelé TATA-*box*, qui recrute la TBP (*TATA-binding protein*) (au niveau de l'initiateur de la transcription, INR). Dix à 35 % des promoteurs possèdent une TATA-*box* (Saxonov et al., 2006). Ils se situent souvent en amont de gènes exprimés spécifiquement dans certains tissus et/ou induits par des signaux particuliers (*e.g.*, stimuli extracellulaires), qui nécessitent une réponse cellulaire adaptée et rapide. L'expression à partir de ces gènes est donc hautement variable mais très forte une fois activée (Valen and Sandelin, 2011).

L'activité promotrice de ces gènes est extrêmement régulée et dépend de nombreuses régions *cis*-régulatrices proches (< 250 bp du TSS) (*i.e.*, promoteur proximal/distal si situé en amont ou en aval du TSS, respectivement) et/ou plus lointaines (constituées d'*enhancers*, *silencers* et/ou insulateurs) (Baumann et al., 2010) (**Figure 23**). Le promoteur core est constitué de la TATA-*box* associée à (et agit de façon synergique avec) l'initiateur de la transcription (Smale and Baltimore, 1989). Les autres séquences sur lesquelles se fixent ou s'associent des composants du complexe de pré-initiation de la transcription qui définissent la

position précise du TSS (Baumann et al., 2010) (**Figure 23**). Par exemple, DPE (*downstream promoter element*) et DCE (*downstream core element*) sont des sites de reconnaissances pour TFIID, alors que BREu et BREd (*TFIIB recognition element upstream* ou *downstream*, respectivement) interagissent avec TFIIB (Baumann et al., 2010). Les sites éloignés pouvant stimuler, inhiber ou bloquer la transcription sont respectivement appelés *enhancers*, *silencers* ou insulateurs, et interagissent avec le promoteur grâce à la formation de boucles (**Figure 23**) (Ong and Corces, 2011). Il est à noter que les promoteurs à TATA-*box* sont particulièrement dépendants de la présence et/ou l'absence de facteurs de transcription spécifiques impliqués dans l'expression tissu-spécifique ou en réponse à un stimulus (Juven-Gershon and Kadonaga, 2010).

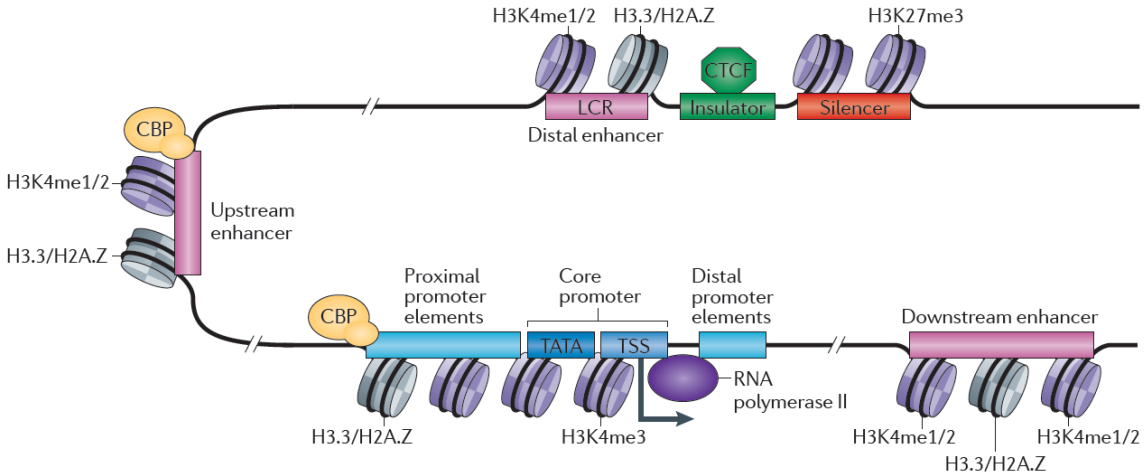


Figure 23: Eléments de régulation d'un promoteur à TATA-*box*. Le promoteur à TATA-*box* est constitué des éléments *core*, proximal et distal. L'initiation de la transcription à partir d'un promoteur à TATA-*box* peut être régulée par de nombreux *enhancers* qui sont localisés à distance et séparés avec des *silencers* et insulateurs [fixés par la protéine régulatrice CTCF (*CCCTC-binding factor*)]. Les *enhancers* sont définis par des caractéristiques chromatiniennes favorables à la transcription (H3K4me1/2, H3.3/H2A.Z) et la fixation de la protéine CBP [(*cyclic AMP-responsive element-binding*(CREB) *binding protein*]. Les *silencers* portent des marques épigénétiques défavorables à la transcription (H3K27me3) [issue de (Ong and Corces, 2011)].

ii/ Les promoteurs de type CGI

Au contraire, le TSS peut être positionné de façon plus large/dispersé (ce qui génère des transcrits aux multiples extrémités 5'). Il se trouve alors au sein d'une large séquence riche en GCs (sur-représentés comparé à la moyenne génomique) contenant une répétition de GCs (localement sur-représentés) de 20-50 nt dans les 100 paires de bases (pb) en amont du TSS. Cette répétition de GCs est appelé îlot CpG (CpG *island*, CGI) et existe dans 40 à 70 %

des promoteurs humains (Illingworth et al., 2010; Saxonov et al., 2006). Ces derniers se trouvent en amont de gènes dit de ménage (*e.g.*, *phosphoglycerate kinase*, PGK) exprimés constitutivement de façon ubiquitaire à un niveau (faible) constant et qui ne nécessitent qu'une régulation simple de leur expression.

Les promoteurs de type CGI ne possède pas les éléments du promoteur de type TATA-*box* mais ont des sites de fixation pour des facteurs de transcription ubiquitaires, tel SP-1, NF-1 (*neurofibromin 1*), E2F, and ETS-1 (Landolin et al., 2010). Ces sites de reconnaissance sont définis par des motifs riches en CpG (Deaton and Bird, 2011). SP-1, par exemple, a été décrit recrutant TBP au niveau de promoteurs ne possédant pas de TATA-*box* (Brandeis et al., 1994; Butler and Kadonaga, 2002).

Les promoteurs de type CGI contiendraient ainsi plusieurs promoteurs *core* faibles, qui définissent plusieurs TSS, alors que celui de type TATA-*box* possèderait un promoteur *core* fort (Baumann et al., 2010).

Il est à noter que les CGIs n'existent que chez les Vertébrés alors que la TATA-*box* est conservée de la levure à l'Homme. Ainsi, les contenus différentiels des promoteurs à TATA ou CpG sont indicatifs de mécanismes de régulation différents. Cependant, les propriétés de ces deux types de promoteurs ne sont pas mutuellement exclusives et il existe de nombreuses exceptions, dont des promoteurs hybrides des deux classes (Valen and Sandelin, 2011).

b/ *Régulation épigénétique de l'initiation de la transcription*

La fixation de l'ARN pol II et l'initiation de la transcription nécessite le recrutement dynamique de la machinerie de transcription basale et des composants régulateurs au niveau des gènes cibles ainsi que des interactions dynamiques entre facteurs de transcription et la chromatine. Cependant, la structure chromatinienne des promoteurs contenant une CGI ou une TATA-*box* définit un accès de l'ADN respectivement favorable ou non à la transcription et impliquent des régulations épigénétiques différentielles.

Ainsi, les promoteurs à CGIs (non-méthylées) sont occupés par des nucléosomes dit instables car faiblement positionnés, ce qui facilite l'accès du promoteur au complexe de pré-initiation de la transcription (**Figure 24**). Ces nucléosomes sont constitués des variants d'histone H3.3 et H2A.Z à la place des histones canoniques H3 et H2A (Jin and Felsenfeld, 2007) qui permettent le recrutement de l'ARN pol II (Valen and Sandelin, 2011) (**Figure 24**). Les queues d'histones peuvent aussi être modifiées, telle l'acétylation des histones H3 et H4

(Birney et al., 2007; Wang et al., 2008b) et la tri-méthylation de l'histone H3 (K4me3) caractéristiques de la chromatine transcriptionnellement active (**Figure 24**). H3K4me3 est une signature des promoteurs à CGIs qui permet le recrutement de TFIID (van Ingen et al., 2008; Vermeulen et al., 2007) et persiste souvent même lorsque le gène associé est inactif (Guenther et al., 2007; Mikkelsen et al., 2007; Thomson et al., 2010). L'expression à partir des promoteurs à CGIs peut cependant être inhibée stablement par méthylation des CGIs pendant le développement normal (ou pathologique dans le cas de cancers) [*cf.*, revues (Deaton and Bird, 2011; Li et al., 2007)]. Il est à noter que les promoteurs de types CGIs sont plutôt régulés lors de l'élongation de la transcription au niveau des nucléosomes se trouvant en aval du TSS (**Figure 24**) (Jones, 2012).

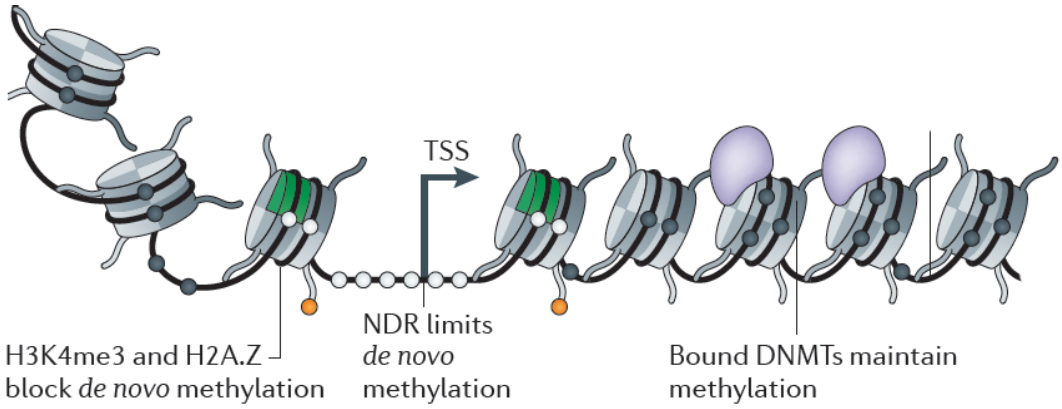


Figure 24: Eléments de régulation transcriptionnelle des promoteurs à CGIs. Les promoteurs à CGIs (non-méthylées) sont occupés par des nucléosomes faiblement positionnés, dit instables, qui facilitent l'accès du TSS aux facteurs de transcription et au complexe de pré-initiation de la transcription. Les nucléosomes entourant le TSS portent la triméthylation de l'histone H3 sur la lysine 4 (H3K4me3, rond orange), qui est associée avec une transcription active, et le variant d'histone H2A.Z (en vert) associé à de l'ADN méthylé (rond blanc). En aval du TSS, l'ADN est ne contient pas de CpG et est principalement méthylé (ronds gris) médié par la fixation de méthyltransférases de l'ADN (DNMTs, en violet), ce qui bloque l'élongation de la transcription [adaptée de (Jones, 2012)].

Les promoteurs à TATA-*box* sont organisés en structure chromatinienne constituée de nucléosomes stables, qui les rend inaccessibles pour le complexe de pré-initiation de la transcription et nécessite l'intervention des complexes de remodelage ATP-dépendants SWI/SNF pour leur activation (Ramirez-Carrozzi et al., 2009). Ces complexes participent à la décondensation de la chromatine et au remodelage local des nucléosomes qui garantissent l'accès au promoteur, caractérisées par des modifications épigénétiques favorables à la transcription. L'occupation stable de ces promoteurs par des nucléosomes fait qu'ils sont pour la plupart transcriptionnellement inactifs par défaut (Valen and Sandelin, 2011).

2/ L'expression transgénique à partir d'un promoteur d'origine humaine

a/ *La kinase phosphoglycérate*

La kinase phosphoglycérate (*phosphoglycerate kinase*, PGK) est une enzyme de la glycolyse qui catalyse le transfert d'un phosphate du 1,3 diphosphoglycérate (1,3-DPG) sur l'adénosine diphosphate (ADP), ce qui génère l'acide 3-phosphoglycérique (3-PGA) et une molécule d'ATP. Cette réaction est la première des deux qui génèrent une molécule d'ATP pendant la glycolyse (Beutler, 2007; Parker and Hoffman, 1967).

Le gène codant cette enzyme est dit « de ménage » car exprimé ubiquitairement au sein de toutes les cellules de l'organisme. Ce gène, appelé *PGK1*, est localisé sur le chromosome X [en position Xq13 (Willard et al., 1985)]. *PGK1*, ainsi que la quasi-intégralité des gènes codés par l'un des deux chromosomes X, peut être transcriptionnellement inactivé au cours de l'embryogénèse chez les mammifères femelles par l'inactivation du chromosome X, qui permet la compensation du dosage des gènes liés à l'X entre les organismes des deux sexes (Monk, 1995). Un défaut en PGK (par mutation du gène) est une des causes relativement rares d'anémie hémolytique congénitale non sphérocytaire, pouvant s'accompagner de retards mentaux et/ou de myopathies (*e.g.*, rhabdomyolyse) (Beutler, 2007).

b/ *Organisation du promoteur PGK*

Le promoteur du gène *PGK1*, appelé promoteur PGK par la suite, ne possède pas de TATA-*box* (ni d'INR)(Singer-Sam et al., 1984) mais est riche en CpGs (dont des CGIs) et contient des sites de fixation potentiels pour des facteurs de transcription ubiquitaires tels SP1, NF1, ATF, C/EBP (déterminés par footprinting ex vivo après traitement à la DNAse (Pfeifer and Riggs, 1991; Pfeifer et al., 1990)) et prédictifs pour NF-κB, EST1, AP2 (déterminés par reconnaissance de sites de fixation consensus sur la séquence du promoteur).

Le promoteur PGK est donc un promoteur de type CGIs.

c/ *Régulation de l'expression à partir du promoteur PGK*

i/ Dans un contexte sain

Le gène *PGK1* est exprimé constitutivement au sein des cellules de l'organisme grâce à la présence de facteurs de transcription ubiquitaires cités ci-dessus. La seule régulation négative que peut subir le promoteur PGK dans un contexte sain est l'inactivation aléatoire de

l'X au cours du développement de l'embryon femelle [comme illustré chez la souris, (Monk, 1995)]. Ce *silencing* est initié par les transcripts spécifiques de l'inactivation de l'X (X-*inactivation specific transcript*, XIST) et est maintenu par un certain nombre de modifications épigénétiques, médiées notamment par le recrutement de complexes protéiques du groupe Polycomb (PcG) et de facteurs de transcription répresseurs de la transcription (*e.g.*, YY1) et la méthylation de l'ADN (Lee, 2011; Wutz, 2011).

ii/ En conditions de stress nutritionnel

En conditions d'hypoxie, la synthèse d'ATP dépend plutôt de la glycolyse ayant lieu dans le cytoplasme à la place de la phosphorylation oxydative qui se passe au sein des mitochondries. Il a ainsi été observé au sein de tumeurs que l'hypoxie augmente l'expression du gène *PGK1* (Daly et al., 2004), de façon dépendant de HIF-1α (*hypoxia-inducible factor* 1α, qui possède des éléments de réponse à l'hypoxie sur le promoteur de ce gène (Kress et al., 1998)], ce qui serait un des mécanismes compensatoires pour assurer la génération d'ATP (Lam et al., 2007).

3/ <u>L'expression transgénique à partir d'un promoteur d'origine virale</u>

a/ *Le cytomégalovirus*

Le promoteur *major immediate early* (MIE) du cytomégalovirus (CMVie, ou CMV), dont l'utilisation est largement répandue dans le domaine du transfert de gènes, est le promoteur-*enhancer* du virus CMV humain (HCMV). Le HCMV appartient à la sous-famille des β-Herpes virus et infecte tous les organes et tissus, causant une infection généralisée avec un large éventail de symptômes (*e.g.*, pneumonies, problèmes hépatiques et gastrointestinaux) principalement chez des individus immunodéprimés ou ayant un système immunologique immature [*cf.*, revue (Stinski and Isomura, 2008)].

Le génome du HCMV [de 230 kb environ (Lakeman and Osborn, 1979; Stinski et al., 1979)] contient deux promoteurs divergents séparés par une région *enhancer*. En amont de la région *enhancer* MIE, se trouvent le promoteur et les gènes très précoces (*immediate early*, IE) qui permettent une réplication productive au sein des cellules permissives. Le promoteur et les gènes UL127 ne sont quant à eux pas essentiels à la réplication en culture cellulaire. Il est à noter que la transcription simultanée des deux gènes n'est pas possible [*cf.*, revue (Stinski and Isomura, 2008)].

b/ *Organisation du promoteur CMV*

Le promoteur-*enhancer* CMVie est constitué de la région enhancer MIE (distale et proximale) et du promoteur des gènes IE. Le promoteur IE contient une *TATA-box* entre -28 et -22, une séquence crs (*cis-repression*) entre -13 and +1, et une INR entre +1 et +7 (Macias et al., 1996) (**Figure 25**). Crs est essentiel à la phase précoce de la transcription des gènes IE, et donc à l'infection.

La région *enhancer* MIE contient de nombreux sites de fixation, pouvant être répétés, pour des facteurs activateurs (RAR, ETS, SP-1, ERF, CREB/ATF et NF-κB) et répresseurs de la transcription (YY1) (**Figure 25**). Le promoteur-*enhancer* CMVie utilisé comme outil pour les travaux de cette étude comprend la région -522 à +55 (Scharfmann et al., 1991). La région *enhancer* MIE distale peut fonctionner dans les deux orientations et n'est pas essentielle à forte m.o.i.. Cependant, sa délétion ralentie l'expression des gènes IE à faible m.o.i. (Meier and Pruessner, 2000). La région enhancer MIE proximale influence directement le niveau de transcription des gènes IE (Isomura et al., 2004), contient une séquence *enhancer* minimale que sont les sites de fixation à SP-1 en -75 et -55 (Isomura et al., 2005) (**Figure 25**).

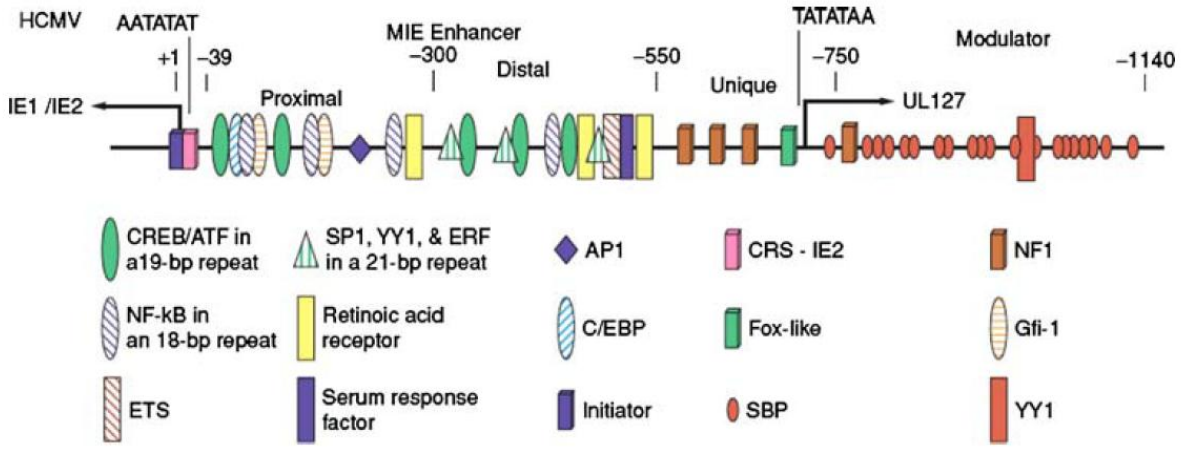


Figure 25: La région *enhancer* MIE du CMV, entourée de deux promoteurs divergents. Les TSS des gènes viraux (IE1/2 et UL127) sont représentés par une flèche. La région *enhancer* MIE est composée de deux parties, proximal et distal. Une région appelée « unique » se trouve entre l'*enhancer* distal et le promoteur UL127. Elle contient des sites de fixation pour des facteurs répresseurs de la transcription et aurait un rôle d'insulateur. En aval du promoteur UL127, se trouve une région appelée « modulateur ». Les sites de fixation pour des facteurs activateurs/répresseurs de la transcription sont représentés par les formes et couleurs indiquées [adaptée de (Stinski and Isomura, 2008)].

c/ *Régulation de l'expression à partir du promoteur CMV*

i/ Expression « normale » à partir du promoteur CMV

Les différents éléments *cis*-régulateurs peuvent agir de façon indépendante et coopérer pour recruter l'ARN pol II et augmenter la transcription des gènes MIE (Meier et al., 2002). Les protéines IE permettent ensuite de stimuler l'expression des gènes précoces (*Early*), dont les produits induisent à leur tour l'expression des gènes tardifs (*Late*) (Wathen and Stinski, 1982). Cependant, les protéines régulent également négativement les promoteurs à partir desquels elles sont exprimées. Ainsi, la région Crs adjacente au TSS régule l'expression des gènes IE au cours de l'infection puisque la fixation de la protéine IE86 sur Csr inhibe la transcription à partir du promoteur IE (Jupp et al., 1993).

ii/ Latence du promoteur CMV

Le HCMV a un large tropisme cellulaire mais ne peut se répliquer qu'au sein des cellules différenciées (*e.g.*, macrophages, cellules denditriques, cellules endothéliales, fibroblastes). Il est donc incapable de se répliquer au sein de cellules peu différenciées (*e.g.*, cellules progénitrices de la moelle osseuse, monocytes) dans lesquelles il se maintient dans un état latent réactivable (Sinclair and Sissons, 1996). Le HCMV peut ainsi établir une infection latente chez des individus immunocompétents qui est réactivée après traitement immunosuppresseur (*i.e.*, production de virus) (Glenn, 1981).

Le promoteur CMV latent est caractérisé par un certain nombre de modifications épigénétiques, telles que des histones hypoacétylées, l'association avec des HDACs (*e.g.*, HDAC1) et HP1 qui fixe des histones méthylées (Murphy et al., 2002; Reeves et al., 2005b). L'activation du promoteur IE latent peut avoir lieu par traitement à la TSA (Murphy et al., 2002) ou après différentiation cellulaire (Reeves et al., 2005a; Reeves et al., 2005b; Soderberg-Naucler et al., 2001). Ainsi, la stimulation de l'expression des gènes IE, facteur limitant de la réplication du HCMV, permet la réactivation du virus latent [*cf.*, revue (Stinski and Isomura, 2008)].

IV/ ROLES DE LA MACHINERIE DE REPARATION DE L'ADN DANS LE CYCLE LENTIVIRAL

Les interactions hôte/virus sont inéluctables dans un contexte où le virus dépend des fonctions cellulaires de son hôte pour se répliquer. Les stratégies de détournement cellulaire

varient en fonction des virus mais sont toujours conduites dans un même but : assurer la réplication virale. Cependant, les protéines de l'hôte, dont la machinerie de réparation de l'ADN, peuvent avoir un impact positif ou négatif sur le cycle viral afin de limiter l'infection virale tout en préservant l'intégrité génomique de la cellule [*cf.*, revue (Weitzman et al., 2010)]. Les protéines/structures virales peuvent ainsi stimuler les voies de réponse aux dommages de l'ADN qui peuvent intervenir à différentes étapes du cycle viral.

A/ STIMULATION DES VOIES DE REPONSE AUX DOMMAGES DE L'ADN LORS DE L'INFECTION RETROVIRALE

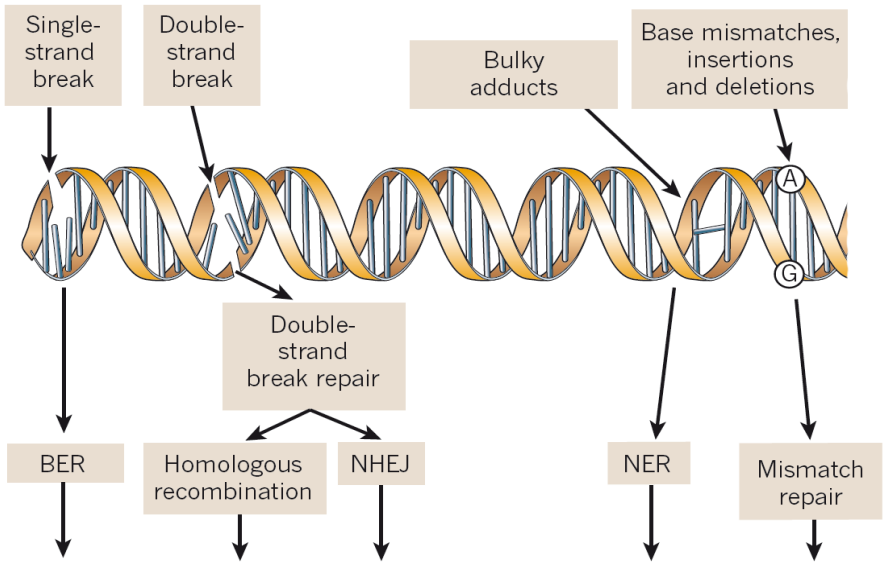
1/ Signalisation et réparation des dommages de l'ADN

L'intégrité du génome humain est continuellement mise à l'épreuve par des agents exogènes (de source physique ou chimique, *e.g.*, lumière UV) et endogènes (*e.g.*, espèces oxygène réactives issues du métabolisme) qui génèrent différents types de lésions de l'ADN (détaillés dans la **Figure 26**). Nos cellules ont évolué un système sophistiqué de surveillance et de réparation pour détecter l'ADN endommagé, signaler sa présence et le réparer afin de prévenir la transmission de mutations génétiques [*cf.*, revues (Ciccia and Elledge, 2010; Jackson and Bartek, 2009)].

Ainsi, la réponse aux dommages de l'ADN (*DNA damage response*, DDR) a lieu par le recrutement successif des différents partenaires : senseurs, transducteurs, médiateurs, effecteurs, qui détectent la lésion, transduisent et amplifient des signaux d'activation/de recrutement (Polo and Jackson, 2011) (**Figure 26**). Ces signaux sont transmissibles par modifications post-traductionnelles réversibles des protéines cibles (*e.g.*, phosphorylation, ubiquitination) pour obtenir rapidement une/des réponse(s) adaptée(s) (non-exclusives) au type de lésion et à son intensité (*e.g.*, apoptose, arrêt de la transcription, arrêt du cycle cellulaire, remodelage de la chromatine, etc) (Polo and Jackson, 2011).

La DDR permet le recrutement des protéines qui réparent la lésion détectée en permettant notamment la maturation des extrémités (avec résection ou non, *e.g.*, endo/exo-nucléase), la synthèse d'ADN (*i.e.*, polymérase) et la ligature des extrémités générées (*i.e.*, ligase). La signalisation en réponse à une cassure de l'ADN n'est pas dissociable de la réparation physique puisque les deux agissent collectivement grâce à des protéines aux activités (de signalisation et réparation) non-exclusives [*e.g.*, Ku fixe les extrémités de l'ADN et recrute la sous-unité catalytique de la DNA-PK (*DNA-dependent protein kinase*) pour former le complexe DNA-PK] (Ciccia and Elledge, 2010). La réparation des dommages de

l'ADN inclut différentes voies qui sont activées en fonction du dommage détecté et du contexte cellulaire. Ainsi, une cassure double-brin de l'ADN peut activer la voie de réparation non-homologue NHEJ (*non-homologous DNA end joining*) ou la voie de recombinaison homologue HR (*homologous recombination*) en fonction de la phase du cycle cellulaire dans laquelle se trouve la cellule endommagée (et donc de son état de différenciation, *i.e.*, cellule souche qui se divise *vs.* différenciée) (**Figure 26**) (Ciccia and Elledge, 2010).



senseurs	PARP-1/2	MRN, PARP-1/2	Ku80/Ku70	RPA	MSH2/6
transducteurs		ATM	DNA-PK	ATR ATRIP	
Médiateurs		Claspin, BRCA1/2, 53BP1			
Effecteurs		Chk1, Chk2, p53			
Effecteurs de la réparation	glycosylase (UNG2), endonucléase(FEN-1), polymérase β, DNase, ligase I/III, Rad18	RPA, Rad51, Rad52, polymérase	ARTEMIS, XRCC4, Ligase IV	Rad17, Hus1, TFIIH (XPB/XPD, polymérase κ	MLH1, PSM2, PCNA, polymérase δ/ε, ligase I

Figure 26 : Types de lésions générés et principales voies de signalisation/réparation mises en jeu. La figure représente les types de lésions de l'ADN pouvant être générés au sein du génome humain et les voies de réparation activées en conséquence. Les protéines majeures de ces voies sont annotées et classées en fonction de leur rôle (*i.e.*, senseur, transducteur, médiateur, ou effecteur) dans la signalisation des voies de réponse aux dommages de l'ADN. Les protéines effectrices interviennent dans l'activation/le recrutement de facteurs impliqués dans la réparation physique de la lésion (des exemples sont indiqués) ainsi que dans le remodelage de la chromatine et/ou l'arrêt du cycle cellulaire [*cf.*, revues (Ciccia and Elledge, 2010; Polo and Jackson, 2011)] [adaptée de (Lord and Ashworth, 2012)].

Le terme « machinerie de réparation de l'ADN » utilisé ultérieurement inclut les protéines impliquées dans la signalisation et/ou la réparation des dommages de l'ADN, sans pour autant qu'une voie de réparation entière soit activée.

2/ L'infection rétrovirale, un stress exogène qui stimule la DDR

a/ *Différents niveaux de stimulation de la DDR par l'infection rétrovirale*

Lors du cycle rétroviral, il existe au moins trois niveaux de stimulation des voies de signalisation des dommages de l'ADN : (*i*) l'ADNv linéaire non-intégré, (*ii*) les protéines virales telle que Vpr et (*iii*) les lésions de l'ADN générées lors de l'intégration (**Figure 27**). En fonction de la nature et de l'intensité de cette stimulation, la cellule peut mourir ou survivre.

i/ l'ADNv linéaire non-intégré

Il a été mis en évidence que les cellules déficientes pour la voie NHEJ (*i.e.*, Ku, DNA-PK ou Ligase IV) meurent par apoptose après infection par le VIH-1 à forte m.o.i. (*multiplicity of infection*) de cellules de rongeurs (murines ou de hamster, *i.e.*, Ku, DNA-PK) et humaines (*i.e.*, Ligase IV) (Daniel et al., 2004; Daniel et al., 1999; Li et al., 2001). Cette réponse nécessiterait l'intégration de l'ADNv (Daniel et al., 2004; Daniel et al., 1999), ou au moins la transcription inverse qui le forme (Li et al., 2001). En absence de facteurs de la voie NHEJ, la présence d'extrémités doubles sur l'ADNv linéaire non-intégré favoriserait la mort de la cellule infectée par une forte m.o.i. (Daniel et al., 2004; Daniel et al., 1999; Li et al., 2001). Celle-ci peut cependant avoir lieu de façon indépendante de l'accumulation d'ADNv linéaire en présence des facteurs NHEJ (Bergeron and Sodroski, 1992 ; Laurent-Crawford and Hovanessian, 1993) qui le circularisent en cercles à 2-LTRs (Li et al., 2001). Ces résultats restent malgré tout controversés et seront développés ultérieurement dans le paragraphe IV.B.2.

ii/ Vpr

L'expression de la protéine virale Vpr induit un arrêt du cycle cellulaire en G2/M après activation de ATR [*Ataxia-telangeictasia mutated* (*ATM*)*-and Rad3-related* (ATR); accompagnée de la phosphorylation de CHK1 (*Checkpoint kinase 1*) et dépendante de Rad17 et Hus1] et la formation de foyers (*foci*) nucléaires de 53BP1 (*p53-binding protein 1*), γH2AX (phosphorylée sur la sérine 139) et BRCA1 (*Breast Cancer Gene 1*) (Andersen et al., 2005;

Lai et al., 2005; Zimmerman et al., 2004). Vpr cause ainsi la mort (par apoptose) des cellules infectées (Roshal et al., 2003), après fixation sur la chromatine (Lai et al., 2005; Tachiwana et al., 2006) qui pourrait être médiée par DDB1 (*damage-specific DNA-binding protein 1*) (Schrofelbauer et al., 2007) et interférée avec la réplication de l'ADN (Lai et al., 2005; Zimmerman et al., 2006) et/ou induire des cassures double-brins (*double-strand breaks*, DSB) de l'ADN (Tachiwana et al., 2006). Il est à noter que cette propriété de Vpr pourrait être mise à profit dans des approches anti-cancéreuses (Muthumani et al., 2009).

iii/ Lésions de l'ADN générées lors de l'intégration

Les lésions de l'ADN générées lors de la troisième étape du processus d'intégration de l'ADNv dans le génome hôte nécessitent une réparation médiée par la machinerie de réparation de l'ADN de la cellule hôte (*cf.*, § I.B.4). Il ressort des multiples études menées pour identifier les acteurs essentiels à la réparation de ces lésions exogènes (reportées dans le **tableau 8**) qu'aucune des voie de réparation cellulaire ne serait essentielle. En effet, seules des fonctions de réparation données seraient nécessaires pour une réparation correcte [*e.g.*, polymérase, nucléase et ligase (Yoder and Bushman, 2000)] du fait de la redondance des protéines cellulaires ayant une même fonction lors du processus de réparation physique.

Cependant, les protéines impliquées dans les étapes initiales (dont détection et signalisation) des voies de réparation des DSBs NHEJ (*e.g.*, Ku, DNA-PK) ou HR (*e.g.*, MRE11-NBS1, PARP-1, ATM) semblent importantes dans le cycle du VIH-1 (**Tableau 8**, aspects développés dans le § IV.B.2).

b/ *Implication des kinases de la DDR après infection par le VIH-1*

Les kinases ATM, ATR, et DNA-PK de la famille PI3K (*phosphoinositide-3-kinase*) jouent un rôle essentiel dans la transduction des signaux en aval de la détection d'une lésion de l'ADN. Il a été montré l'implication, le recrutement et/ou l'activation de ces trois kinases en réponse à une infection cellulaire par le VIH-1 (**Tableau 8, Figure 27**). En effet, ce sont des cibles privilégiées du virus car elles possèdent de nombreuses cibles et induisent une réponse rapide en les phosphorylant de façon réversible (Zimmerman et al., 2004). Il y a cependant un certain nombre de controverses sur les rôles exacts de ces protéines (**Tableau 8**, aspects développés dans le § IV.B.2).

c/ *Phosphorylation de H2AX après infection rétrovirale*

En absence d'infection, les kinases ATM, ATR et DNA-PK activées phosphorylent le variant d'histone H2A, γH2AX au niveau du site de la lésion de l'ADN (Ciccia and Elledge, 2010). γH2AX a été observé précocement, à partir de 4 h après infection par l'ASV, lors des premiers évènements d'intégration sans pour autant être nécessaire au processus de réparation (Daniel et al., 2004). Cette caractéristique majeure d'un dommage de l'ADN n'a cependant été observée que tardivement (20 h, 48 h) après infection par le VIH-1 et résulterait de l'expression de Vpr [*cf.*, § IV.A.2, (Zimmerman et al., 2004)] (**Figure 27**).

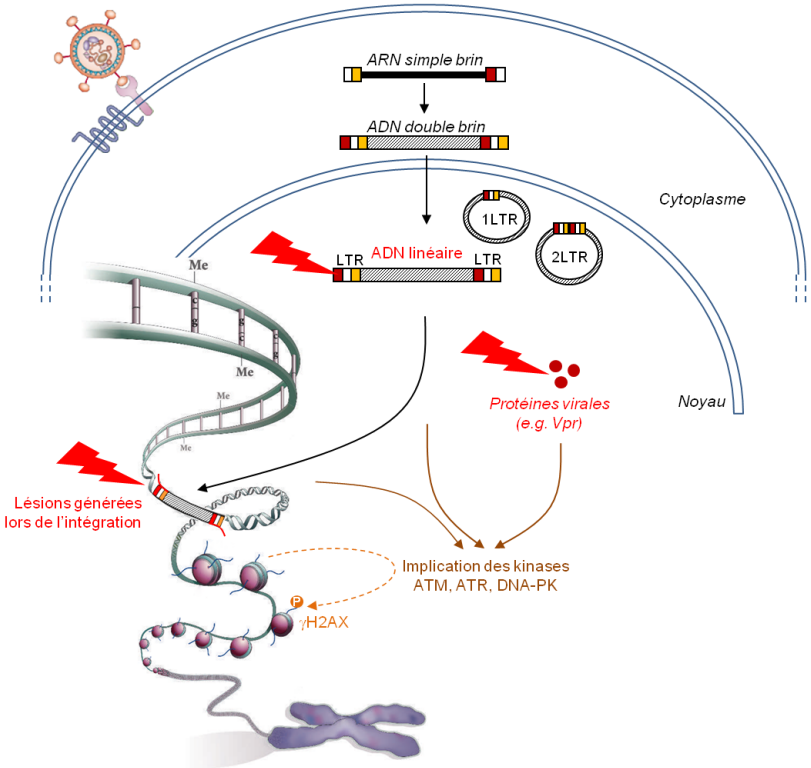


Figure 27 : Stimulation des voies de réponse aux dommages de l'ADN lors de l'infection rétrovirale. Trois niveaux de stimulation des voies de signalisation des dommages de l'ADN peuvent être activées pendant le cycle rétroviral, il existe: (*i*) l'ADNv linéaire non-intégré, (*ii*) les protéines virales (*e.g.*, Vpr), et (*iii*) les lésions de l'ADN générées lors de l'intégration. Ces protéines/intermédiaires viraux peuvent impliqués/activés/recrutés les kinases ATM, ATR ou DNA-PK qui participent à la transduction de la signalisation, pouvant faire intervenir la phosphorylation du variant d'histone H2AX (*i.e.*, γH2AX).

B/ INTERACTIONS DYNAMIQUES ENTRE LE CYCLE RETROVIRAL ET LA MACHINERIE DE REPARATION DE L'ADN

1/ Etat actuel des connaissances sur les interactions entre cycle rétroviral et machinerie de réparation de l'ADN

En réponse à la stimulation de la DDR, et/ou de par leur rôle physiologique, un certain nombre de protéines impliquées dans la machinerie de réparation de l'ADN peuvent avoir un rôle positif ou négatif sur différentes étapes de l'infection rétrovirale (**Tableau 8**).

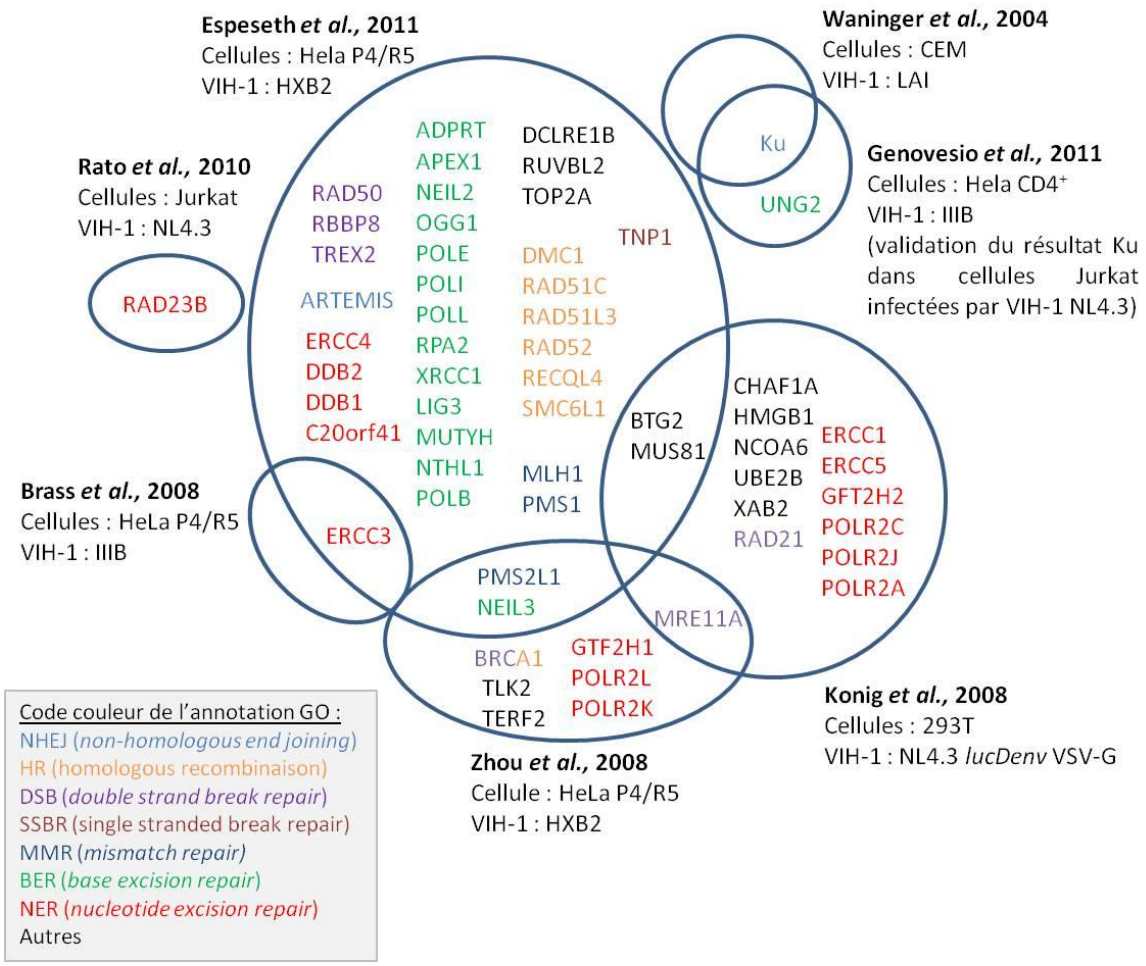


Figure 28 : Protéines de la réparation de l'ADN identifiées comme nécessaires à la multiplication du VIH lors d'approches globales. Les études de criblage du génome basées sur l'utilisation de l'ARNi ont mis en œuvre des lignées cellulaires ou des virus/vecteurs VIH-1 qui peuvent varier, comme indiqué. Les voies de réparation auxquelles appartiennent les protéines identifiées sont également annotées à l'aide d'un code couleur.

Dans ce contexte, de récentes analyses génomiques par ARNi des gènes cellulaires impliqués dans le cycle du VIH-1 ont reporté ceux codant des protéines de la machinerie de réparation de l'ADN parmi les ~1400 nécessaires à la réplication du VIH-1 (Brass et al., 2008; Espeseth et al., 2011; Genovesio et al., 2011; Konig et al., 2008; Rato et al., 2010;

Waninger et al., 2004; Zhou et al., 2008) [études ayant globalement un faible chevauchement, comme indiqué dans (Bushman et al., 2009)].

Ces interactions dynamiques entre hôte et virus sont organisées de façon spatio-temporelle en différents niveaux, dépendants de la localisation cellulaire et de l'étape du cycle viral (**Tableau 8**). Il est à noter que les controverses indiquées dans le **tableau 8** ainsi que le peu de chevauchement entre protéines de la réparation identifiées lors d'approches globales (**Figure 28**) pourraient s'expliquer par la mise en œuvre de protocoles expérimentaux différents (*e.g.*, espèce ou type cellulaire utilisé, virus réplicatif *vs.* vecteur défectif pseudotypés), comme décrit dans le § IV.B.2).

Tableau 8 : Interactions dynamiques entre le cycle rétroviral et la machinerie de réparation de l'ADN. Sont indiqués le rôle positif (+) ou négatif (-) des protéines de la machinerie de réparation de l'ADN avec l'étape du cycle rétroviral concerné (cytoplasmique ou nucléaire). Les protéines impliquées dans les voies HR et NHEJ sont indiquées en vert et orange, respectivement.

Cellule	Etape du cycle rétroviral	Rôle	Protéines (et références)
Cytoplasme	Apoptose	+	**ATR** stimulée par Vpr (Roshal et al., 2003)
		-	**ATM** [(Daniel et al., 2001), intégration-dépendante (Lau et al., 2005)], **ATR** [via intégration (Daniel et al., 2003)], **DNA-PK** [(Daniel et al., 2001), via intégration (Daniel et al., 1999), controversé (Baekelandt et al., 2000)] , **Ku** [(Li et al., 2001), controversé (Baekelandt et al., 2000)], **Ligase IV** [intégration-dépendante (Daniel et al., 2004), transcription inverse-dépendante (Li et al., 2001)], **NBS1** [via ATM et intégration (Smith et al., 2008)]
	Dégradation de l'ADNv	-	**Rad18** (Lloyd et al., 2006), **XPB** et **XPD** [(Yoder et al., 2006), au sein du noyau (Yoder et al., 2011)], **UNG2** (Priet et al., 2005)
	Insertions de mutations dans l'ADNv	+	**PMS2** [MLV, à un taux faible (Tang and Zhang, 2000)], **UNG2** [via encapsidation dépendante de Vpr (Chen et al., 2004)]
		-	**UNG2** [(Priet et al., 2005), controversé (Kaiser and Emerman, 2006)]
Noyau	Maturation et protection des extrémités		**ATM**, **Mre11**, **NBS1**, **Artemis** (Sakurai et al., 2009), TREX1 [MoMLV (Beck-Engeser et al., 2009)]
	Circularisation des 2-LTRs	+	**Ku** [(Li et al., 2001), dont Ku80 (Jeanson et al., 2002; Masson et al., 2007)], **Ligase IV** (Kilzer et al., 2003; Li et al., 2001), **XRCC4** (Li et al., 2001)
		-	**Rad 52** [compétition avec Ku pour fixer le LTR (Lau et al., 2004)]
	Circularisation des 1-LTR	+	**MRE11** [(Kilzer et al., 2003), controversé (Sakurai et al., 2009)], **NBS1** (Kilzer et al., 2003)

	Intégration au sein du génome	+	**ATM** [(Lau et al., 2005; Nunnari et al., 2005), controversé (Ariumi et al., 2005)], **ATR** [(Daniel et al., 2003; Nunnari et al., 2005; Smith et al., 2007), controversé (Ariumi et al., 2005; Dehart et al., 2005)], **DNA-PK** [(Daniel et al., 1999), controversé (Ariumi et al., 2005; Baekelandt et al., 2000)], **Ku80** [(Jeanson et al., 2002; Waninger et al., 2004; Zheng et al., 2011), controversé (Li et al., 2001; Masson et al., 2007)], **NBS1** (Smith et al., 2008), **PARP-1** [(Ha et al., 2001; Kameoka et al., 2005), controversé (Ariumi et al., 2005; Baekelandt et al., 2000; Siva and Bushman, 2002)], **WRN** [recrutement de PCAF/P-TEFb (Sharma et al., 2007)], **voie BER** (Espeseth et al., 2011; Yoder et al., 2011) dont **FEN-1** (Brin et al., 2000; Yoder and Bushman, 2000) et **complexe SET** [inhibe autointégration (Yan et al., 2009)] dont Metnase (Williamson et al., 2008)
		-	**Rad52** [compétition avec Ku sur fixation LTR (Lau et al., 2004)], **Rad51** (intéragit avec IN (Cosnefroy et al., 2012))
	Localisation		**ATM** (Sakurai et al., 2009), **DNA-PK** [FLV, (Hasegawa et al., 2005)], **Ku80** [LV-SIN en périphérie nucléaire (Masson et al., 2007)], **PARP-1** [centromères (Kameoka et al., 2005)]
	Expression	+	**ATM** (via Rev, (Ariumi and Trono, 2006)), **DNA-PK** (interaction avec pol II (Tyagi et al., 2011)), **Ku80** ((Waninger et al., 2004), **PARP-1** (activation LTR intégré (Kameoka et al., 2004)), **Rad51** (interagit avec C/EBPb (Chipitsyna et al., 2006))
		-	**Ku80** (MMTV, via DNA-PK (Giffin et al., 1999), KBS (Jeanson and Mouscadet, 2002 ; Masson et al., 2007)), **PARP-1** (compétition avec Tat ou le complexe Tat/P-TEFb pour TAR *in vitro* (Parent et al., 2005)), **UNG2** (Fenard et al., 2009)

Abréviations : BER, *Base Excision Repair* [qui inclut Ogg1, Myh, Neil1 et Polß dans (Yoder et al., 2011), et MUTYH, NTHL1, NEIL3, et POLB dans (Espeseth et al., 2011)]; UNG2, *uracil DNA glycosylase* ; MMR, *Mispathch repair* ; MMTV, *mouse mammary tumor virus* ; FEN-1, *Flap endonucléase 1* ; Complexe SET, contient les endonucleases APE1 et NM23-H1, la 5'-3' exonuclease TREX1 et Metnase ; FLV, *Friend leukemia virus* ; PARP-1, *Poly(ADP-ribose) polymerase-1*.

Il ressort des études citées dans le **tableau 8** que les membres des voies de réparation des DSBs NHEJ et HR sont largement représentés dans les interactions hôte/virus, et ce des étapes pré-intégratives du cycle viral jusqu'à l'étape cruciale d'expression du génome viral intégré.

2/ Implications des voies de réparation des cassures double-brins dans le cycle du VIH-1

La sur-représentation des voies de réparation des DSBs dans les interactions hôte/virus peut s'expliquer par la capacité des protéines senseurs de ces voies à s'associer aux extrémités double-brin de l'ADN et à ensuite recruter ses partenaires. Une autre hypothèse (non-exclusive) repose sur le fait que les lésions simple-brin de l'ADN générées lors de l'intégration de l'ADNv sont transformées en DSB au cours de la réplication si elles ne sont

pas réparées (Smith and Daniel, 2006). Elles doivent cependant être réparées en privilégiant la voie de réparation associée au moins de délétions possibles des séquences virale/hôte initiales.

a/ *Les voies de réparation des cassures double-brin*

i/ Principales voies de réparation des cassures double-brin

Les voies de réparation des DSBs sont divisées en deux voies majeures, appelées voie de réparation non-homologue NHEJ et voie de recombinaison homologue HR. Deux voies mineures de réparation des DSBs qui nécessitent une homologie de séquence entre les régions se situant de part et d'autre de la cassure double-brin ont été décrites. Il s'agit des voies de réparation médiées par l'hybridation des simples brins (*single-strand annealing*, SSA), ou de micro-homologies (*micro-homology-mediated end joining*, MMEJ ; aussi appelé *alternative*-NHEJ) (Ciccia and Elledge, 2010; Mladenov and Iliakis, 2011).

ii/ Choix de la voie de réparation

Le choix de la voie de réparation des DSBs est régulé par de nombreux facteurs, dont la nature de la lésion et la phase du cycle cellulaire dans laquelle se trouve la cellule au matériel génétique endommagé. En particulier, la régulation du choix entre NHEJ et HR se fait par la résection des extrémités double-brin de l'ADN dont l'activation dépend des kinases cycline-dépendantes (CDK) (Ciccia and Elledge, 2010).

C'est la voie NHEJ qui est principalement utilisée chez les vertébrés et peut être activée tout au long du cycle cellulaire. Elle met en jeu la ligature des extrémités double-brin de l'ADN lésé qui n'ont, au préalable, été que peu ou pas maturées et protégées par Ku (**Figure 29A**). Cette voie de réparation est considérée comme peu fidèle mais est corrélée à une réparation sans mutation/insertion/délétion si les deux extrémités des cassures ne sont ni maturées, ni endommagées (Ciccia and Elledge, 2010).

Les autres voies de réparation nécessitent une résection 5'-3' des extrémités de l'ADN lésé. La voie HR est une réparation fidèle qui nécessite un substrat homologue, principalement la chromatide sœur du chromosome portant la lésion présente après la duplication de l'ADN, pendant la fin de la phase S et la phase G2 du cycle cellulaire (**Figure 29B**). La voie MMEJ nécessite des micro-homologies au niveau des extrémités de la DSB (**Figure 29C**). La voie SSA a lieu par recombinaison homologue entre les séquences homologues se situant de part et d'autre de la DSB (**Figure 29D**). Ces deux dernières voies de

réparation aboutissent, respectivement, à des délétions/insertions minimes ou importantes (Ciccia and Elledge, 2010).

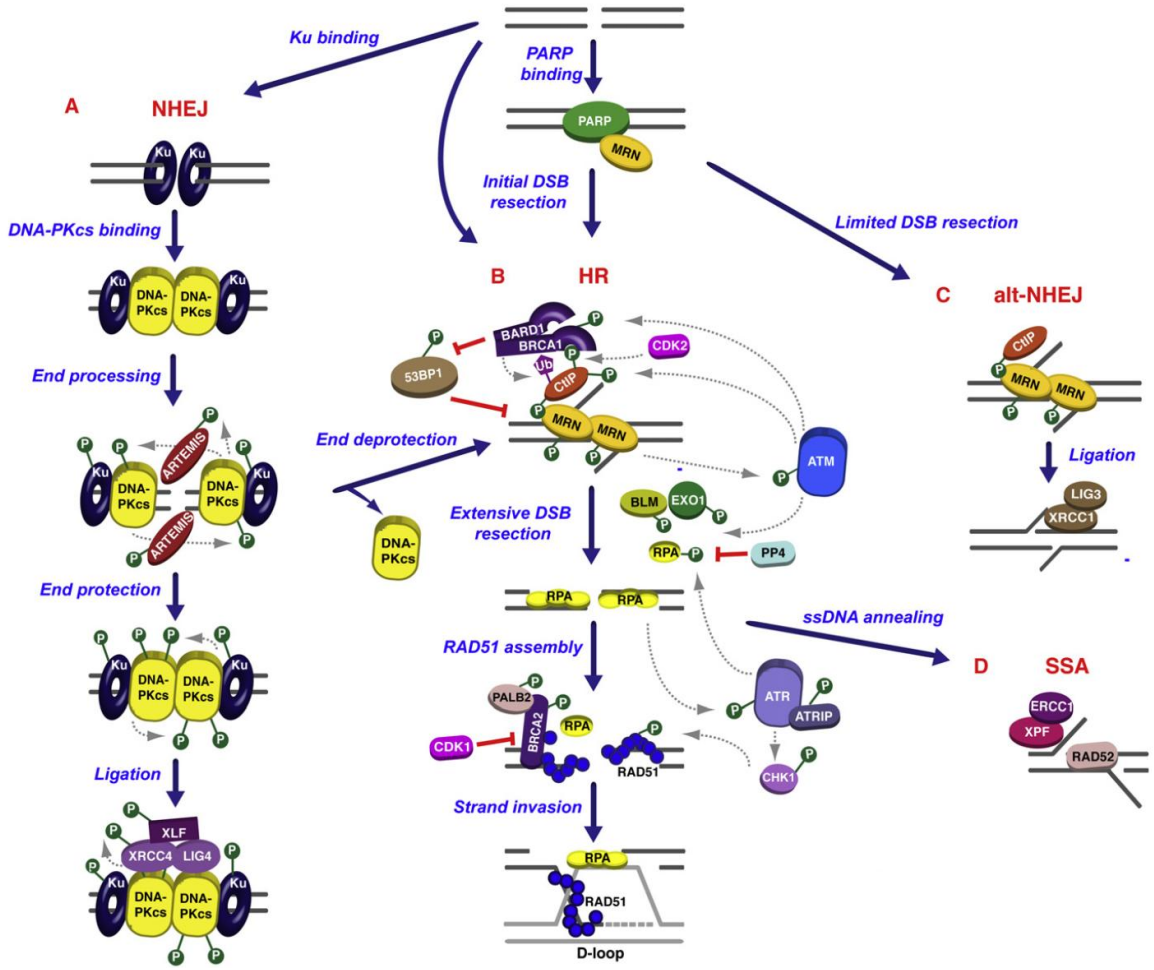


Figure 29 : Voies de réparation des cassures double-brins de l'ADN. Le choix de réparation des cassures double-brins a lieu par fixation des extrémités double-brin par Ku (**A**) ou PARP-1 (**B**) qui induisent respectivement les voies NHEJ ou HR [ou alt-NHEJ/MMEJ (**C**), ou SSA (**D**)]. Chacune des voies activées induit un processus de réparation qui diffère de par les partenaires recrutés, aboutissant *in fine* à la réparation de la lésion prise en charge [adaptée de (Ciccia and Elledge, 2010)].

b/ *Interactions des voies NHEJ et HR avec les différentes étapes du cycle du VIH-1*

i/ Importance de la voie NHEJ dans le cycle du VIH-1

Il a été mis en évidence que les protéines de la voie NHEJ (*i.e.*, Ku, DNA-PK, Ligase IV) ont un effet positif sur la réplication du VIH au sein de cellules de rongeurs (*i.e.*, Ku, DNA-PK) et humaines (*i.e.*, Ligase IV). Elles agissent en limitant l'apoptose (Daniel et al., 2001; Daniel et al., 1999) et en favorisant l'intégration de l'ADNv (Daniel et al., 2004; Daniel et al., 1999) ou au moins la transcription inverse (Li et al., 2001). Cependant, ces résultats

sont controversés et ne sont visibles qu'après une infection à forte m.o.i. (> 1) (Ariumi et al., 2005; Baekelandt et al., 2000).

Les protéines de la voie NHEJ seraient également impliquées dans la circularisation des espèces circulaires d'ADNv à 2-LTRs, comme reporté pour Ku (à faible et forte m.o.i) dans des cellules humaines infectées par un virus VIH-1 réplicatif (Jeanson et al., 2002; Zheng et al., 2011) ou des cellules de rongeurs transduites par un vecteur dérivé du VIH-1 (*i.e.*, non-réplicatif) (Li et al., 2001; Masson et al., 2007). De plus, Ligase IV/XRCC4 a été décrit influencer la génération des cercles à 2-LTRs à forte m.o.i. dans des cellules humaines transduites par un vecteur VIH-1 (Li et al., 2001).

L'implication de DNA-PK dans l'intégration du VIH-1 au sein du génome hôte, décrite à forte m.o.i. au sein de cellules de rongeurs (Daniel et al., 1999), reste controversée (Ariumi et al., 2005; Baekelandt et al., 2000). Il en est de même pour Ku80 qui participerait à ce processus à faible m.o.i. uniquement dans les cellules humaines dépourvues en Ku80 infectées par un virus VIH-1 réplicatif (Jeanson et al., 2002 ; Zheng et al., 2011) ou transduites par un vecteur dérivé du VIH-1 (Waninger et al., 2004).

Cependant, DNA-PK et/ou Ku80 interviendraient dans le ciblage rétroviral dans la chromatine [du FLV (Hasegawa et al., 2005) et d'un vecteur SIN dérivé du VIH-1 en périphérie nucléaire (Masson et al., 2007)] au sein de cellules de rongeurs. Ces protéines ont également été décrites impliquées dans la régulation positive [DNA-PK (Tyagi et al., 2011), Ku80 (Waninger et al., 2004)] et/ou négative [Ku80 (Jeanson and Mouscadet, 2002 ; Masson et al., 2007)] de l'expression du VIH-1 au sein de cellules humaines (Jeanson and Mouscadet, 2002; Tyagi et al., 2011; Waninger et al., 2004), et de rongeurs (Masson et al., 2007) (aspects concernant Ku développés dans le § IV.C.3).

Les controverses liées à ces résultats proviendraient de la mise en œuvre de *designs* expérimentaux différents, comme décrit dans le paragraphe **iv/** ci-dessous (*cf.*, IV.B.2).

ii/ Importance de la voie HR dans le cycle du VIH-1

La voie HR, via ATM, a également été décrite comme ayant un rôle positif dans la réplication du VIH-1. ATM agit en réduisant l'apoptose induite par l'infection du VIH-1 (Daniel et al., 2001) à faible m.o.i. de façon dépendante de NBS1 et de l'intégration de l'ADNv (Lau et al., 2005; Smith et al., 2008) dans des cellules humaines (Daniel et al., 2001; Lau et al., 2005; Smith et al., 2008) et murines (Daniel et al., 2001; Lau et al., 2005). La voie HR serait aussi impliquée dans la maturation et protection des extrémités de l'ADNv (Sakurai

et al., 2009) et la circularisation des cercles à 1-LTR à forte m.o.i. (Kilzer et al., 2003) au sein des cellules humaines transduites par un vecteur dérivé du VIH-1 (Kilzer et al., 2003; Sakurai et al., 2009). La voie HR a également été décrite comme inhibant la circularisation des cercles à 2-LTRs dans les cellules murines infectées par le VIH-1 réplicatif (Lau et al., 2004).

ATM et PARP-1 agiraient positivement sur l'intégration [(Ha et al., 2001; Kameoka et al., 2005; Lau et al., 2005; Nunnari et al., 2005), résultat controversé (Ariumi et al., 2005; Baekelandt et al., 2000)] au sein de cellules humaines (Baekelandt et al., 2000; Kameoka et al., 2005; Lau et al., 2005; Nunnari et al., 2005) et de rongeurs (Ariumi et al., 2005; Baekelandt et al., 2000; Ha et al., 2001; Kameoka et al., 2005) transduites par un vecteur dérivé du VIH-1 (Ariumi et al., 2005; Baekelandt et al., 2000; Ha et al., 2001; Kameoka et al., 2005; Lau et al., 2005) ou infectées par un virus VIH-1 réplicatif (Baekelandt et al., 2000; Nunnari et al., 2005).

Les protéines ATM et PARP-1 de la voie HR interviendraient dans le ciblage (partiel) du VIH-1 au sein du génome hôte [ATM (Sakurai et al., 2009), PARP-1 (Kameoka et al., 2005)] au sein de cellules humaines (Kameoka et al., 2005; Sakurai et al., 2009) et de rongeurs (Kameoka et al., 2005) transduites par un vecteur dérivé du VIH-1. Elles auraient également un rôle dans la régulation de l'expression virale de façon positive [ATM (Ariumi and Trono, 2006), PARP-1 (Kameoka et al., 2004)] et/ou négative [PARP-1 *in vitro* (Parent et al., 2005)] au sein de cellules humaines (Ariumi and Trono, 2006; Kameoka et al., 2004) et de rongeurs (Ariumi and Trono, 2006) transduites par un vecteur VIH-1 (Ariumi and Trono, 2006; Kameoka et al., 2004) ou infectées par le VIH-1 réplicatif (Ariumi and Trono, 2006) (aspects concernant PARP-1 développés dans le § IV.C.3).

Cependant, des acteurs de la voie HR se trouvant en aval de ATM, Rad51 et Rad52, défavoriseraient l'intégration du VIH-1 au sein du génome hôte (Cosnefroy et al., 2012), ce dernier pouvant être en compétition avec Ku pour fixer le LTR (Lau et al., 2004).

iii/ Importance des voies de réparation lors d'une infection dans un contexte indépendant de l'intégration

Il a très récemment été mis en évidence que l'infection par le VIH-1 dans un contexte déficient pour l'intégration est favorisée dans des conditions de dommages de l'ADN (*i.e.*, pré-traitement des cellules par irradiation ou au peroxyde d'hydrogène avant infection) et corrélée à une expression virale moindre (comparé au VIH-1 compétent pour l'intégration) (Ebina et al., 2012). De plus, l'analyse des jonctions LTR-chromatine hôte a montré qu'elles

portent des délétions pour le LTR ou des insertions > 50 bp, résultat corrélé à une insertion fréquente du VIH-1 dans les répétitions minisatellites-*like* (Ebina et al., 2012). Cette étude suggère une prise en charge de l'ADNv par la voie NHEJ et/ou HR qui pourrait influencer le site d'intégration du VIH-1 et être associée à la constitution de réservoirs viraux (Ebina et al., 2012).

iv/ Controverses

Comme décrit ci-dessus, il existe un certain nombre de controverses sur l'implication de la machinerie de réparation de l'ADN sur le cycle rétroviral qui repose sur la mise en œuvre de *designs* expérimentaux différents. Ils varient quant à l'utilisation du type cellulaire (hôte naturel du VIH-1, ou autre pouvant avoir des voies de réparation privilégiées), du virus/vecteur (*e.g.*, VIH-1 réplicatif, pseudotypé, etc), de la m.o.i. [comme discuté dans (Sinclair et al., 2006) et illustré dans (Yang et al., 2010)], du temps d'analyse post-transduction, de molécules pharmacologiques [dont la spécificité est concentration- et type cellulaire-dépendante (Yang et al., 2010)] et, enfin, des techniques (qui deviennent plus précises, *e.g.*, espèces d'ADNv circulaires analysées sur gel d'agarose après PCR *vs.* Q-PCR actuellement).

Notamment, un certain nombre d'études a pu être mené au sein de cellules de rongeurs [*i.e.*, souris (*e.g.*, *mouse embryonic fibroblasts*, MEFs) ou hamster (*e.g.*, *Chinese Hamster Ovary*, CHO)] qui peuvent être aisément déplétées en facteurs de la réparation de l'ADN. De plus, les protéines impliquées dans la réparation de l'ADN n'ont pas la même abondance dans les cellules humaines et de rongeurs. C'est ains le cas de DNA-PK, très abondante dans les cellules humaines (Ting et al., 1999). Cette différence sous-tend des rôles attribués aux protéines impliquées dans la réparation de l'ADN qui pourraient varier entre espèces (chez les mammifères) et être spécifique, au sein d'une espèce, d'interaction(s) avec un agent infectieux donné (*e.g.*, Ku avec le VIH-1, aspect développé dans le § IV.C.3).

v/ NHEJ *versus* HR

La circularisation des espèces d'ADNv non-intégrées à 1- et 2-LTRs par les voies HR et NHEJ est cohérente avec leur fonction de réparation par recombinaison homologue et ligature, respectivement.

Il a été proposé que la voie NHEJ soit utilisée préférentiellement dans la réparation post-intégration de l'ADNv dans le génome hôte du fait de son activation tout au long du

cycle cellulaire et du substrat que représente l'ADNv avec des extrémités maturées mais non-endommagées (Smith and Daniel, 2006). Cependant, la voie HR (ainsi que les voies de réparation des cassures simple-brin) semble avoir aussi un rôle non négligeable dans ce processus (**Tableau 8**). Le choix de la voie de réparation mise en jeu dans la réparation post-intégration de l'ADNv intégré serait multifactoriel et/ou non-exclusif, surtout en cas d'infection multiple.

Enfin, l'implication des protéines des voies NHEJ et HR dans la localisation du provirus intégré et/ou la régulation de son expression pourraient être des conséquences secondaires dues à leur rôle dans la réparation post-intégration, le résultat d'une prise en charge précoce de l'ADNv, et/ou indépendant de leur fonction/recrutement lors du processus de réparation.

3/ Implication de la machinerie de réparation dans le cycle lentiviral et remodelage chromatinien

Certaines des conséquences secondaires de l'implication de la machinerie de réparation dans le cycle rétroviral ont déjà été évoquées dans les paragraphes ci-dessus (*e.g.*, apoptose, arrêt du cycle cellulaire, etc). Cependant, l'intervention de la machinerie de réparation dans un contexte « sain » nécessite un remodelage de la chromatine associé à des modifications épigénétiques qui permettent d'assurer l'accès à la machinerie de réparation de l'ADN à la zone chromatinienne lésée, de la même façon que lors de l'activation de l'initiation de la transcription (*e.g.*, code histone spécifique de la réparation *vs.* transcription, aspects développés dans le § III.A.1).

Dans ce contexte, l'infection par le VIH-1 pourrait induire un remodelage local de la chromatine en réponse à un recrutement initié par la DDR, ou un changement global de l'organisation chromatinienne, en fonction du degré d'intensité de stress auquel est corrélée l'infection (Lilley et al., 2010). Un remodelage de la chromatine global ou local au niveau du site d'intégration en réponse à l'infection du VIH-1, médiée par le recrutement de la machinerie de réparation de l'ADN et/ou la mise en place de modifications épigénétiques spécifiques, pourrait ainsi avoir un impact sur l'expression virale qui pourrait être variable au cours du temps.

A titre d'exemple, comme décrit précédemment, l'infection par l'ASV peut induire la phosphorylation de H2AX (Daniel et al., 2004). De plus, Metnase, impliquée dans la méthylation de l'histone H3 après activation de la voie NHEJ (Lee et al., 2005), a été décrit

avoir un rôle dans l'intégration lentivirale au sein du génome hôte (Williamson et al., 2008). Néanmoins, il n'a pas été montré de lien direct entre remodelage de la chromatine, réparation de l'ADN et infection par le VIH-1.

C/ ROLES DU COMPLEXE KU DANS LE CYCLE RETROVIRAL

Au cours de notre étude, nous nous sommes plus précisément intéressés aux rôles du complexe Ku dans le cycle de réplication du VIH-1. Ainsi, ce facteur a été décrit participer à de nombreuses étapes du cycle rétroviral (**Tableau 8**) mais ses fonctions restent malgré tout controversées.

1/ Le complexe Ku

a/ *Identification des protéines Ku70 et Ku80*

Le complexe protéique Ku est composé de deux sous-unités : Ku80 (de 86 kDa, aussi appelée Ku86) et Ku70 (de 70 kDa). Cet hétérodimère a été identifié en 1981 dans le sang de patients atteints de syndrome de chevauchement sclérodermie-polymyosite (*i.e.*, maladie autoimmune caractérisée par deux types d'atteintes cliniques/biologiques) dont la moitié avaient des anticorps anti-Ku (nom issu des initiales du premier patient identifié) (Mimori et al., 1981). Il a ensuite été retrouvé dans le sang de patients atteints d'autres maladies autoimmunes (*e.g.*, lupus érythrémateux systémique) (Francoeur et al., 1986; Mimori et al., 1986; Reeves, 1985).

b/ *Gènes codant le complexe Ku*

Ku80 est codé par le gène *XRCC5* tandis que Ku70 est codé par *XRCC6* (*X-ray repair complementing defective repair in Chinese hamster cells 5* et *6*, respectivement), en référence aux études menées chez le rongeur qui ont permis d'identifier Ku comme un facteur important dans la réparation des lésions de l'ADN (Errami et al., 1996; Singleton et al., 1997). Ces gène sont situés sur les chromosome humains 2 et 22 (positions 2q35 et 22q13, respectivement) (Koike, 2002).

Chez les eucaryotes supérieurs, il existe un autre membre au sein de la famille Ku, appelé KARP-1 (*Ku80-autoantigen-related protein-1*) qui est transcrit à partir du locus *XRCC5* en utilisant un promoteur en aval et des exons supplémentaires (Myung et al., 1997). La protéine qui en résulte possède une extension en N-terminal de ~9 kDa comparée à Ku80

(Myung et al., 1997). KARP-1 ne pourrait cependant pas remplacer Ku80 lors de la réparation des DSBs malgré le fait qu'elle puisse s'hétérodimériser avec Ku70, comme décrit récemment (Koike et al., 2011).

c/ *Caractéristiques de l'hétérodimère Ku*

i/ Conservation de Ku au cours de l'évolution

Chez les eucaryotes, des homologues de Ku ont été retrouvés chez les singes, les amphibiens (*i.e.*, *Xenopus laevis*), les insectes (*i.e.*, *Drosophilia melanogaster*) et les rongeurs (*i.e.*, souris, hamster). Il est cependant à noter, comme cité précédemment, que la concentration en Ku est nettement moins importante dans les cellules de rongeurs comparées à celles d'origine humaine (Tuteja and Tuteja, 2000). Cependant, la majorité des eucaryotes inférieurs ne possèdent ni la DNA-PKcs, ni le domaine d'interaction avec cette protéine au sein de l'homologue de Ku80. La capacité de Ku80 à recruter la DNA-PKcs est ainsi spécifique des eucaryotes supérieurs (Downs and Jackson, 2004). La conservation structurale et fonctionnelle de Ku chez les eucaryotes suggère cependant l'existence de fonctions de Ku indépendantes de la réparation des DSBs chez les eucaryotes supérieurs, ainsi que de nouvelles fonctions dépendantes de la réparation de l'ADN.

Enfin, des homologues de Ku, codés par un gène, ont été découverts chez les bactéries et archées (Aravind and Koonin, 2001; Doherty et al., 2001).

ii/ Propriétés structurales du complexe Ku

Les sous-unités Ku70 et Ku80 du complexe Ku présentent seulement 14 % d'identité de séquences mais un repliement quasi-identique, ce qui suggère qu'elles proviendraient d'un même gène ancestral (Walker et al., 2001). Elles ont une organisation structurale proche puisque toutes deux sont constituées de trois régions : le domaine N-terminal (term) en hélices α et feuillets β, un domaine central en tonneau β et de domaines C-term divergents (**Figure 30A, B, C**).

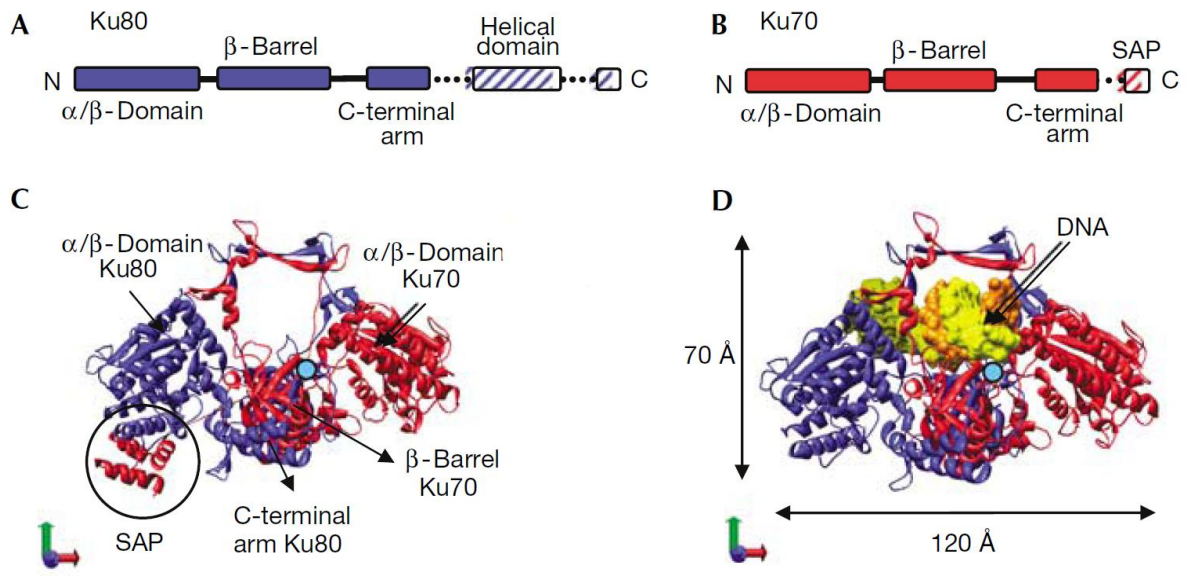


Figure 30 : **Organisation du complexe Ku.** (**A-B**) Représentations schématiques de l'organisation des domaines des sous-unités Ku80 (**A**) et Ku70 (**B**). Elles sont constituées des régions suivantes : le domaine N-terminal (term) α/β , un domaine central en tonneau β et de domaines C-term divergents. Ku80 possède un long domaine C-term en hélices, comparé au domaine SAP (*SAF-A/B, Acinus and PIAS*) en C-term de Ku70. (**C-D**) Structures tridimensionnelles du complexe Ku seul (**C**) ou en interaction avec l'ADN (**D**). Ku80 et Ku70 sont représentées en bleu et rouge, respectivement. Sont indiqués les domaines décrits ci-dessus. Le rond bleu ciel indique le dernier acide aminé du domaine C-term de Ku80 [adaptée de (Rivera-Calzada et al., 2007)].

Les domaines aux extrémités ont la capacité de fixer des protéines alors que le domaine central intéragit avec l'ADN de façon non-spécique (**Figure 30D**). Ku80 possède un long domaine C-term lui permettant d'intéragir avec la sous-unité catalytique DNA-PKcs pour former la DNA-PK. Quant à Ku70, il possède en C-term un domaine SAP (*SAF-A/B, Acinus and PIAS*) qui lui permettrait de se lier à l'ADN. De plus, le domaine C-term de chacune des deux sous-unités lui permet d'interagir avec le domaine central de l'autre sous-unité (**Figure 30 C, D**) (Gullo et al., 2006; Rivera-Calzada et al., 2007).

iii/ Stabilité du complexe Ku

Le complexe hétérodimérique Ku70-Ku80 est extrêmement stable *in vivo*. Cette stabilité lui est conférée par l'hétérodimérisation des deux sous-unités ainsi que par le recrutement/activation de la DNA-PK. Ainsi, la déplétion stable en Ku80 dans les cellules de hamster *xrs-6* résulte en la perte de Ku70 (Singleton et al., 1997). De plus, la déplétion (partielle stable ou transitoire) en Ku80 dans les cellules humaines HCT 116 est corrélée à la diminution de Ku70, et *vice-versa* [(Li et al., 2002), observations personnelles]. De la même façon, la *trans*-complémentation (transitoire) de Ku80 au sein de cellules de rongeurs ou

humaines déplétées pour cette protéine permet de rétablir les niveaux protéiques des deux sous-unités (Errami et al., 1996; Ghosh et al., 2007).

Cependant, Ku80 et Ku70 auraient chacune des fonctions qui leur sont propres (Ochem et al., 1997). La stabilité des sous-unités de Ku dépendrait donc de leur hétérodimérisation mais aussi de leur dégradation sous forme monomérique (Gullo et al., 2006).

iv/ Localisation du complexe Ku

Les protéines Ku sont impliquées dans de nombreux processus nucléaires et cytoplasmiques, ce qui suggère qu'elles puissent être actives à de nombreux emplacements au sein de la cellule. Elles sont cependant majoritairement nucléaires, et localisées au sein des régions transcriptionnellement actives du noyau (Gullo et al., 2006). Les sous-unités du complexe Ku peuvent aussi être localisées dans le cytoplasme, notamment au cours de la mitose au sein des cellules humaines (Koike, 2002). La translocation nucléaire de chacune des deux sous-unités pendant la phase G1 se ferait ensuite de façon indépendante et le retour de Ku70 précèderait alors celui de Ku80 (Koike, 2002). De plus, le transport au noyau de Ku80 et Ku70 impliquerait des régulateurs différents, comme en témoignent les signaux de localisation nucléaire qui diffèrent entre ces deux protéines (Koike, 2002). Les protéines Ku ont également été décrites être présentes à la membrane cellulaire de différents types cellulaires (Prabhakar et al., 1990), où elles pourraient jouer le rôle de molécules d'adhésion (Teoh et al., 1998).

Enfin, des changements de localisation subcellulaire de Ku (*i.e.*, translocation nucléaire) peuvent être contrôlés par des stimuli initiés lors de contacts entre cellules, ou de stress produits par la privation en facteurs de croissance (*i.e.*, apportés par le sérum) ou le traitement à l'hydroxyurée (Fewell and Kuff, 1996).

v/ Activités de fixation des nucléotides et protéines par Ku

Ku a la capacité d'interagir avec les nucléotides (*i.e.*, ADN, ARN). Les premières études menées sur les capacités de fixation de l'ADN de Ku ont mis en évidence qu'elles se produisent avec une haute affinité aux extrémités (endommagées ou non) de l'ADN de façon non-spécifique, indépendante de leurs structures/séquences (Tuteja and Tuteja, 2000). Cependant, Ku peut aussi se fixer de façon séquence-spécifique au niveau d'éléments de régulation de la transcription de certains gènes viraux, tel le NRE (*negative regulatory*

element) au sein du LTR du MMTV (Giffin et al., 1999) et du HTLVI (Okumura et al., 1994), et de nombreux gènes humains (Tuteja and Tuteja, 2000).

Ku peut également fixer sélectivement des structures ARN, tel l'élément TAR du VIH-1 *in vitro* (Kaczmarski and Khan, 1993).

Il a également été mis en évidence que Ku interagit directement (de façon indépendante de DNA-PK ou non) avec un certain nombre de protéines, dont la télomérase (Ting et al., 2005), PARP-1 (Galande and Kohwi-Shigematsu, 1999), YY1 (Sucharov et al., 2004) et l'histone H2AX (Heo et al., 2008).

vi/ Sous-unité de la DNA-PK

L'activité sérine-thréonine kinase de DNA-PK est portée par sa sous-unité catalytique, DNA-PKcs. L'activité kinase de cette dernière est stimulée lors de son recrutement à l'ADN par Ku (qui y est déjà fixé) pour ainsi former le complexe DNA-PK. Il a été démontré *in vitro* la capacité de DNA-PK à phosphoryler de nombreuses protéines nucléaires régulatrices pouvant se fixer à l'ADN, dont p53, SP-1, l'ARN pol II, TFIID, Artemis, Ku et elle-même, ce qui souligne son implication dans un certain nombre de processus (Tuteja and Tuteja, 2000).

2/ Rôles de Ku

a/ *Réparation*

Le complexe Ku est essentiel à la réparation des DSBs médiée par la voie NHEJ puisque son absence réduit l'activité NHEJ dans les cellules de rongeurs (Errami et al., 1996), et au moins l'activité DNA-PK chez l'homme (Li et al., 2002).

Brièvement, Ku détecte les cassures double-brin de l'ADN et, *via* sa fixation, permet de (*i*) choisir la voie NHEJ, (*ii*) protéger et rapprocher les extrémités de l'ADN et (*iii*) recruter les autres facteurs de la voie NHEJ (**Figure 31A**). La première à être ensuite recrutée est la DNA-PKcs qui forme avec Ku le complexe DNA-PK (**Figure 31B**). Les (auto-)phosphorylations séquentielles de DNA-PK et Artemis favorisent la maturation (*e.g.*, résection) des extrémités par Artemis et leur protection par DNA-PK (**Figure 31C**), ce qui permet un rapprochement des extrémités (**Figure 31D**) avant ligation par XRCC4/Ligase IV/XLF (**Figure 31E**) (Ciccia and Elledge, 2010). Il est à noter que le détachement de Ku de la zone d'ADN réparée nécessite son ubiquitinilation préalable (Postow et al., 2008).

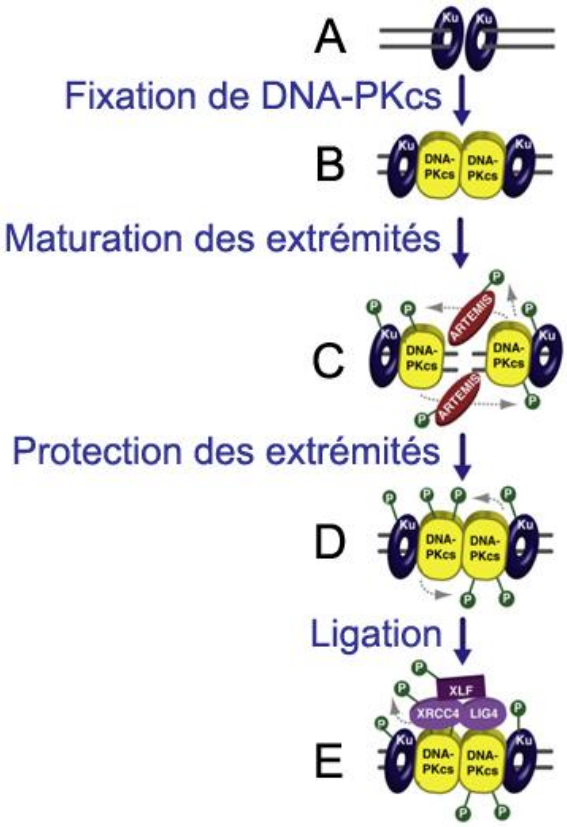


Figure 31 : Processus de réparation d'une cassure double-brin de l'ADN impliquant la voie non-homologue NHEJ. La fixation des protéines Ku70/Ku80 aux extrémités double-brin de la cassure de l'ADN (**A**) permet le recrutement de la DNA-PKcs (**B**). Le complexe DNA-PK ainsi formé et Artemis (auto)phosphorylés favorisent la maturation des extrémités de l'ADN (**C**) qui sont protégés par la DNA-PK (**D**) puis liguées par XRCC4/Ligase IV/XLF (**E**)[adaptée de (Ciccia and Elledge, 2010)].

Cependant, l'accès à la lésion, permettant l'intervention des facteurs de la réparation, nécessite un remodelage de la chromatine pouvant faire intervenir des protéines *polycomb* [*e.g.*, PHF1, (Hong et al., 2008)] et/ou des méthyltransférases [*e.g.*, Metnase (Lee et al., 2005)]. De la même façon, la chromatine doit être ré-assemblée après réparation. Ce processus pourrait faire intervenir Ku, qui interagit *in vitro* avec CAF-1 (*Chromatin assembly factor 1*), responsable de l'assemblage *de novo* des histones H3 et H4 (Hoek et al., 2011).

De plus, la voie NHEJ, dont Ku, participe à la recombinaison V(D)J (variable-diversity-joining) impliquée dans le réarrangement des gènes d'immunoglobulines générant ainsi de la diversité. En particulier, la voie NHEJ, grâce au processus décrit ci-dessus, permet la maturation et la réparation des structures en boucles dont la formation est médiée par les protéines RAG1 et RAG2 (*recombination-activating gene*) (Downs and Jackson, 2004).

Enfin, Ku inhiberait l'apoptose médiée par Apaf-1 après lésion de l'ADN en interagissant directement avec le promoteur du gène codant ce facteur pro-apoptotique et en régulant négativement sa transcription (De Zio et al., 2011).

b/ *Télomères*

Les télomères des chromosomes représentent une source perpétuelle d'extrémités de l'ADN pour les cellules eucaryotes. Néanmoins, la longueur de ces extrémités varient en fonction des espèces. Elles sont ainsi plutôt longues chez les rongeurs mais relativement courtes chez l'homme. Leur taille est cependant critique pour la survie cellulaire (Espejel et al., 2002).

Les évidences de l'implication de Ku dans le maintien des télomères proviennent principalement des études menées chez *Saccharomyces cerevisiae*. Ku y a été décrit participer à la localisation des télomères en périphérie nucléaire et être nécessaire à l'établissement d'un état chromatinien trancriptionnellement inactif (Downs and Jackson, 2004).

Chez d'autres organismes, Ku a été mis en évidence dans la régulation des télomères. Ainsi, Ku s'associe avec les télomères au sein des cellules humaines et de rongeurs, et en particulier avec la télomérase chez l'homme (Ting et al., 2005). Notamment, Ku permet de maintenir la longueur des télomères et protéger les extrémités des chromosomes (Riha et al., 2006). Ku est donc essentielle à la survie des cellules humaines de par son rôle dans le maintien des télomères (Li et al., 2002; Wang et al., 2009).

c/ *Transcription*

Un certain nombre de gènes possède des élèments de régulation pouvant être fixés par Ku qui peut ainsi réguler leur expression (Tuteja and Tuteja, 2000). Le complexe a notamment été décrit se fixer spécifiquement à des séquences d'ADN riches en purines et en AT au niveau des gènes humains codant le récepteur à la transferrine (Roberts et al., 1994), le récepteur aux glucocorticoïdes (Warriar et al., 1996), la xanthine oxidoréductase (Xu et al., 2004), le récepteur aux androgènes (Mayeur et al., 2005), l'*IL-2* (Shi et al., 2007) *et c-jun* (Jiang et al., 2008) ainsi que des gènes viraux, dont ceux du rétrovirus murin MMTV (Giffin et al., 1996). Le complexe Ku peut ainsi avoir un effet répresseur [*e.g.*, du promoteur du MMTV (Giffin et al., 1996)] ou activateur de la transcription [*e.g.*, promoteur du gène codant le récepteur aux glucocorticoïdes (Warriar et al., 1996) ou *c-jun* (Jiang et al., 2008)].

De plus, Ku a été retrouvé impliqué dans la régulation de l'expression médiée par les ARN pol I, II et III (Downs and Jackson, 2004).

Un exemple caractéristique est la stimulation de la transcription du gène *IL-2* par Ku et NF90 au sein de cellules T activées. Dans ces cellules au repos, Ku(70/80) et NF90 interagissent avec le promoteur *IL-2 in vivo* et participerait à créer/maintenir un environnement défavorable à la transcription. Lors de l'activation cellulaire, le remodelage de la chromatine modifierait les fixations protéines-ADN en diminuant la fixation de Ku70 et augmentant les interactions Ku80/NF90 avec le promoteur *IL-2* qui favoriserait alors l'initiation de la transcription (Shi et al., 2007).

Enfin, il a été très récemment mis en évidence que DNA-PK provoque une inhibition de la transcription par l'ARN pol II en présence des DSBs, et son éviction de l'ADN, de manière à favoriser la réparation des lésions (Pankotai et al., 2012).

Ainsi, Ku, de part ses caractéristiques particulières, peut être impliqué dans la régulation de la transcription de façon dépendante ou indépendante du processus de réparation.

d/ *Interactions de Ku avec les virus*

Il a été mis en évidence que DNA-PK/Ku régulerait négativement la réplication du KSHV (*Kaposi's sarcoma-associated herpesvirus*) sous forme latente en phosphorylant une protéine clé de la latence appelée LANA (*latency-associated nuclear antigen*), impliquée notamment dans la persistance et la réplication des épisomes viraux (Cha et al., 2010).

De plus, DNA-PK/Ku favoriserait la réplication du AAV *in vitro* (Choi et al., 2010). Parallèlement, le ciblage de gène médié par un AAV est augmenté si la voie NHEJ, dont Ku70, est inhibée (Fattah et al., 2008; Paulk et al., 2012).

Enfin, Ku80 permettrait d'augmenter la transcription lytique de l'EBV (*Epstein-Barr virus*) en interagissant avec son facteur de transcription Zta au niveau du promoteur du gène de la phase lytique *BZLF1* (Chen et al., 2011).

3/ <u>Rôles de Ku dans le cycle du VIH-1</u>

Ku interagit à différentes étapes du cycle du VIH-1 (aspect développé dans le § IV.B.2). Cependant, son rôle est controversé du fait de l'utilisation de *designs* expérimentaux qui peuvent varier (*cf.*, § IV.B.2). Le complexe Ku a malgré tout été identifié comme cible

cellulaire qui affecte la réplication du VIH-1 par deux analyses génomiques globales (Genovesio et al., 2011; Waninger et al., 2004). Dans la suite de ce paragraphe, seront distingués les rôles décrits à l'une ou l'autre des sous-unités du complexe Ku, ou du complexe Ku lorsque les deux sous-unités sont impliquées.

a/ *Etapes pré-intégratives*

Il a été mis en évidence que Ku70 interagit avec l'IN du VIH-1 [ainsi qu'avec celle du MLV (Li et al., 2001; Studamire and Goff, 2008)] au sein du PIC *in vitro* (Li et al., 2001) et *in vivo* (de façon indépendante de Ku80) au sein de cellules humaines (Zheng et al., 2011). De plus, grâce à cette interaction, Ku70 pourrait être encapsidé au sein des particules virales lors de l'assemblage des virions et protégé l'IN d'une dégradation par le protéasome lors de l'infection (Zheng et al., 2011). Ku est également impliqué dans la circularisation des espèces circulaires d'ADNv à 2-LTRs à faible et forte m.o.i. dans des cellules humaines infectées par un virus VIH-1 réplicatif (Jeanson et al., 2002; Zheng et al., 2011). } ou des cellules de rongeurs transduites par un vecteur dérivé du VIH-1 (*i.e.*, non-réplicatif) (Li et al., 2001; Masson et al., 2007). Cependant, Ku n'interviendrait pas directement dans la stimulation de la réaction d'intégration puisqu'elle ne participe ni au clivage des extrémités de l'ADNv, ni au transfert de brin au sein du génome hôte (Li et al., 2001).

b/ *Intégration et localisation*

Ku80 a été décrit participer au processus d'intégration du VIH-1 uniquement dans les cellules humaines déplétées (transitoirement ou stablement) en Ku80 infectées par le VIH-1 (Jeanson et al., 2002; Zheng et al., 2011) ou transduites par un vecteur dérivé (Waninger et al., 2004). Le rôle de Ku dans l'intégration virale semble donc spécifique de l'espèce humaine puisqu'aucun effet de la déplétion en Ku n'a été observé sur l'intégration du VIH-1 dans des cellules de rongeur dépourvues en Ku (Baekelandt et al., 2000; Li et al., 2001; Masson et al., 2007).

Enfin, notre laboratoire a décrit que Ku80 pourrait participer à la localisation d'un vecteur SIN dérivé du VIH-1 en périphérie nucléaire de cellules de rongeur dépourvues en Ku et influencer négativement l'expression virale dans le temps (Masson et al., 2007). Cependant, dans cette étude, il n'a pas été mis en évidence si Ku80 participait au ciblage de l'intégration ou à la relocalisation du provirus en périphérie nucléaire après transduction (Masson et al., 2007).

c/ *Expression*

i/ Régulation de l'expression du VIH-1 et controverses

Il a été reporté que Ku80 peut influencer positivement (Waninger et al., 2004) ou négativement (Jeanson and Mouscadet, 2002; Masson et al., 2007) l'expression du VIH-1. De plus, DNA-PK régule de façon négative la transcription du MMTV grâce à la fixation de Ku sur un élément de régulation spécifique présent au sein de l'élément NRE du LTR du MMTV (Giffin et al., 1999). Ce site de fixation potentiel pour Ku (*Ku binding site*, KBS) a également été identifié dans le NRE du LTR du VIH-1 (Jeanson and Mouscadet, 2002) et pourrait participer à la répression de la transcription du VIH-1 au sein des cellules U1 contenant deux provirus pseudolatents [*i.e.*, qui porte une mutation au niveau du gène *tat* (Emiliani et al., 1998)] (Jeanson and Mouscadet, 2002). Il a notamment observé dans ces cellules que la déplétion transitoire de Ku80 stimule la production virale induite par le traitement au TNFα (Jeanson and Mouscadet, 2002). De plus, l'expression (précoce) du VIH-1 et des vecteurs dérivés est augmentée au sein des cellules de rongeurs déplétées en Ku80 (Jeanson and Mouscadet, 2002; Masson et al., 2007).

Cependant, de façon contradictoire, Waninger *et al.* ont mis en évidence que la déplétion en Ku80 (médiée par ARNi) diminue l'activation (précoce) du LTR du VIH-1 médiée par Tat (Waninger et al., 2004). Ce résultat est à rapprocher de l'implication positive de DNA-PK (fonctionnelle) dans la régulation de l'expression précoce à partir du LTR du VIH-1 (ou du promoteur CMV, observé à forte m.o.i.) et tardive lors de la réactivation de provirus latent, compétent pour la réplication (Tyagi et al., 2011).

ii/ Explications possibles

Un certain nombre d'études analysant les effets de Ku sur le cycle du VIH-1 ont été réalisés à l'aide de virus réplicatifs (Jeanson and Mouscadet, 2002; Jeanson et al., 2002; Tyagi et al., 2011; Waninger et al., 2004; Zheng et al., 2011). Cependant, l'utilisation d'un virus réplicatif peut conduire à la superposition des effets dus à l'absence d'une protéine sur les étapes pré-intégratives, l'expression et les étapes de production virale (*e.g.*, assemblage) puisque les cellules cibles sont aussi les cellules productrices. Dans ce contexte, leur usage dans les études précédemment citées a pu aboutir à masquer l'impact réel de Ku sur l'expression du VIH-1 (Tyagi et al., 2011). Par exemple, une diminution de l'expression pourrait réduire la quantité de particules virales produites qui se traduirait par un défaut visible d'intégration après ré-infection des cellules productrices (Jeanson et al., 2002).

De plus, le temps d'analyse, *i.e.*, à court- ou moyen/long-terme, après infection/transduction pourrait expliquer les contradictions issues des études réalisées sur l'impact de Ku sur l'expression du VIH-1 (et les étapes pré-intégratives du cycle viral). Une partie des travaux présentés dans cette thèse vise notamment à éclaircir le rôle de Ku dans la régulation de l'expression lentivirale dans le temps.

Enfin, des différences entre les caractéristiques des protéines impliquées dans la réparation de l'ADN (*e.g.*, abondance, rôles spécifiques et/ou essentiels) au sein des cellules humaines et de rongeurs pourraient expliquer des contributions différentielles lors du cycle viral, notamment lors de l'intégration du VIH-1 au sein de la chromatine hôte.

d/ *Comparaison à un autre senseur des lésions de l'ADN : PARP-1*

PARP-1 est une enzyme qui fixe les cassures simple- et double-brin de l'ADN [avec une haute affinité (D'Silva et al., 1999)] et catalyse la modification post-traductionnelle de nombreuses protéines à l'aide de polymères d'ADP-ribose (Luo and Kraus, 2012). Cette protéine serait en compétition avec Ku pour la fixation des extrémités double-brin de l'ADN lors d'une lésion double-brin de l'ADN et favoriserait l'intervention de la voie HR (Mladenov and Iliakis, 2011). Cependant, cette compétition serait limitée par la régulation de la résection des extrémités de l'ADN qui est dépendante du cycle cellulaire [*cf.*, § IV.B.2, (Mladenov and Iliakis, 2011)].

Comme décrit dans le paragraphe IV.B.2, il a été montré (de façon controversé) que PARP-1 a un rôle dans l'intégration du VIH-1 (Ha et al., 2001; Kameoka et al., 2005), au sein de cellules humaines (Kameoka et al., 2005) et de rongeurs (Ha et al., 2001; Kameoka et al., 2005) transduites par un vecteur dérivé du VIH-1 (Ha et al., 2001; Kameoka et al., 2005). Néanmoins, PARP-1 pourrait participer à la régulation de l'expression du VIH-1 puisqu'elle active le LTR intégré d'un vecteur VIH-1 au sein de cellules humaines (Kameoka et al., 2004). De plus, PARP-1 a été décrit être un co-facteur dans l'activation des gènes dépendants de NF-κB (Hassa and Hottiger, 1999), qui pourrait dépendre de sa capacité à maintenir l'expression des HATs (Ota et al., 2003). Cependant, il a été mis en évidence *in vitro* que PARP-1 est en compétition avec le complexe Tat/p-TEFb pour la fixation de TAR (Parent et al., 2005). D'autres investigations seront nécessaires pour déterminer le(s) rôle(s) exact(s) de PARP-1 dans le cycle du VIH-1, et en particulier, si sa participation dans la régulation de l'expression virale évolue au cours du temps.

Ainsi, les protéines de la machinerie de la réparation qui interviennent très tôt dans la détection des lésions de l'ADN, et en particulier Ku, pourraient participer à différentes étapes du cycle viral, dont la régulation de l'expression virale dans le temps. Ces capacités qui leur sont conférées ne dépendraient pas uniquement de leur rôle dans la signalisation et la réparation des dommages de l'ADN, mais feraient intervenir leurs autres fonctions de façon non-exclusive. Cependant, il est très difficile de mettre clairement en évidence les rôles exacts des protéines étudiées, du fait notamment de phénomènes de compensation entre voies de réparation au sein des lignées étudiées. Enfin, dans le cas de notre protéine d'intérêt Ku, se pose le problème de son abondance et du fait qu'elle est essentielle à la survie des cellules humaines.

Les impacts du complexe Ku sur le cycle du VIH-1 ont été évalués au cours de ce travail de thèse.

PROBLEMATIQUE

Après infection de sa cellule hôte, le VIH-1 peut se répliquer uniquement s'il s'exprime à un niveau suffisant pour produire les protéines/ARNs nécessaires à la formation de nouveaux virions. L'expression nécessaire à la multiplication du VIH-1 requiert l'intégration de son génome au sein de la chromatine hôte. Cependant, après intégration, l'expression virale peut être sujette au *silencing* transcriptionnel. Cette forme de latence post-intégrative contribue à la formation de réservoirs viraux qui constituent un obstacle à l'éradication du virus chez les patients séropositifs traités par HAART.

D'autre part, le transfert de gène basé sur l'utilisation de vecteurs rétroviraux nécessite le développement d'outils sûrs et efficaces basés sur un meilleur contrôle du site d'intégration du transgène et de son expression dans le temps. En effet, l'intégration non-ciblée d'un transgène peut générer une/des mutagénèse(s) insertionnelle(s), pouvant aboutir au processus d'oncogénèse. De plus, l'expression du transgène intégré peut être influencée par le contexte chromatinien environnant et atténuée et/ou perdue rapidement ou au cours du temps après transduction.

L'expression du VIH-1 ainsi que celle d'un transgène d'origine lentivirale, sont donc soumises à une régulation stricte de la part de déterminants viraux et cellulaires. Une meilleure caractérisation de ces déterminants permettrait de contrôler l'établissement et/ou le maintien d'un *silencing* post-intégratif en vu de (*i*) développer des traitement curatifs du VIH-1 et (*ii*) maîtriser/moduler l'expression transgénique dans le cadre de transfert de gènes.

Le but de cette thèse a donc été de caractériser les impacts de certains éléments viraux (*e.g.*, promoteur natif ou hétérologue) et cellulaires (*e.g.*, protéine de la réparation de l'ADN) qui agissent spécifiquement sur l'expression lentivirale dans un contexte de transduction par le VIH-1 ou des vecteurs qui en dérivent.

Dans un premier temps, nous avons étudié l'impact de l'insertion de séquences hétérologues au sein des SIN-LTRs de vecteurs dérivés du VIH-1 sur l'expression lentivirale. En effet, cette approche est particulièrement prisée lors du transfert de gène utilisant les vecteurs rétroviraux du fait de la duplication des LTRs lors de la transcription inverse. Ainsi, l'insertion d'un insulateur au sein des SIN-LTRs permettrait d'isoler le transgène au sein de la chromatine hôte et d'(au moins) maintenir son expression dans le temps. Cependant, l'impact réel de l'ajout de séquences au sein des SIN-LTRs sur l'expression transgénique reste mal caractérisé. De plus, celui-ci pourrait varier en fonction du promoteur interne qui contrôle l'expression du transgène et de la fonctionnalité de l'intégrase. La première partie des travaux

expérimentaux de cette thèse avait donc pour objectif de caractériser l'impact de l'insertion au sein des SIN-LTRs sur l'expression lentivirale en fonction du *design* des vecteurs utilisés.

Dans un second temps, nous avons analysé les rôles de la machinerie de réparation de l'ADN dans le cycle lentiviral. En effet, les facteurs impliqués dans la réparation de l'ADN peuvent avoir un impact positif ou négatif sur le cycle rétroviral, en fonction de la protéine et/ou de l'étape du cycle considéré, dont l'issue aboutit à une multiplication virale productive ou abortive. Cependant, les études menées peuvent être contradictoires en fonction du *design* expérimental. En particulier, le complexe Ku, impliqué dans la voie NHEJ de réparation de l'ADN, aurait des fonctions dans le cycle du VIH-1 qui restent controversées. Il a notamment été décrit que cet hétérodimère pouvait réguler positivement ou négativement l'expression lentivirale. La deuxième partie des résultats de la présente étude visait à éclaircir les rôles de Ku dans le cycle du VIH-1, et plus particulièrement, dans la régulation de son expression au sein de cellules humaines.

Enfin, nous nous sommes intéressés à l'impact de l'insulation des vecteurs lentiviraux sur l'expression transgénique dans un contexte défectueux pour la réparation de l'ADN. Ainsi, les facteurs de la réparation de l'ADN sont décrits comme nécessaires à la réparation post-intégration mais pourraient participer plus précocement à la prise en charge de l'ADN viral, puis influencer son site d'intégration, et contrôler son expression dans le temps. Dans ce contexte, le laboratoire a décrit que Ku participait à la localisation (pré- et/ou post-intégrative) en périphérie nucléaire de vecteurs dérivés du VIH-1. La troisième partie de cette étude a permis d'analyser l'influence de l'haplodéplétion en Ku sur l'expression lentivirale *a priori* isolée du contexte chromatinien hôte.

RESULTATS

I/ IMPACTS DU DESIGN DE VECTEURS DERIVES DU VIH-1 SUR L'EXPRESSION LENTIVIRALE

A/ CONTEXTE DE L'ETUDE

Le transfert de gènes pour la recherche ou la thérapie génique nécessite le développement d'outils sûrs et efficaces. Environ 23% des essais cliniques actuellement menés en thérapie génique utilisent des vecteurs de transfert dérivés de rétrovirus (http://www.wiley.co.uk/genmed/clinical/). Ceux-ci sont particulièrement adaptés au transfert de gènes, notamment pour leur capacité à s'intégrer de façon quasi-aléatoire au sein de la chromatine hôte. Cette caractéristique assure potentiellement une expression transgénique à long-terme au sein des cellules transduites (Pauwels et al., 2009).

Cependant, il existe des problèmes de toxicités, pouvant être médiés par le vecteur et/ou le transgène, qui proviennent des interactions entre les séquences du vecteur et l'environnement génomique entourant les sites d'intégration [*cf.*, revues (Emery, 2011; Pauwels et al., 2009)].

Ainsi, l'intégration non-ciblée peut mener à l'inactivation d'un gène cellulaire ou des modifications d'expression des gènes hôtes situés à proximité du provirus. Ces évènements peuvent conférer un avantage sélectif aux cellules transduites, ce qui augmente le risque de dominance clonale, ou même d'oncogénèse (Pauwels et al., 2009). Dans ce contexte, le développement de vecteurs SIN, qui ne possèdent plus les séquences *enhancer*-promoteur de la région U3 des LTRs, a considérablement contribué à réduire la génotoxicité des vecteurs rétroviraux, en particulier s'ils contiennent un promoteur interne modérément actif (Maruggi et al., 2009; Montini et al., 2009; Montini et al., 2006; Zhou et al., 2010; Zychlinski et al., 2008). Depuis peu, le développement de IDLVs émerge également comme une approche alternative (Banasik and McCray, 2010).

Toutefois, l'efficacité du transfert de gènes médiée par les vecteurs rétroviraux peut être influencée par la chromatine à proximité du provirus et être variable, en particulier à cause du *silencing* que peut subir le transgène. Celui-ci peut avoir lieu rapidement après transduction ou s'établir progressivement au cours du temps (*i.e.*, extinction) et/ou conduire à de la variégation entre cellules sœurs qui exprimeraient différemment un même provirus (Ellis, 2005). Diverses stratégies sont actuellement menées pour optimiser l'expression transgénique [*i.e.*, utilisation de UCOEs (*ubiquitously acting chromatin opening elements*) optimisés (Knight et al., 2012)]. L'une d'entre elles repose sur la configuration des vecteurs

SIN qui permet l'insertion de séquences au sein du 3' LTR (qui sont dupliquées dans le 5' LTR au cours de la transcription inverse). Ces éléments incorporés dans les SIN-LTR incluent les USEs (*upstream polyadenylation enhancers*) qui augmentent l'efficacité de terminaison de la transcription, et les insulateurs (Schambach et al., 2009). L'utilisation de séquences insulatrices, douées d'activités barrière et/ou *enhancer*-bloqueur, pourrait limiter le *silencing* du transgène et/ou prévenir l'activation des gènes endogènes par l'*enhancer* hétérologue, respectivement [*cf.*, revues (Emery, 2011; Schambach et al., 2009)]. L'importance de cette dernière a récemment été illustrée par l'introduction de courts insulateurs synthétiques au sein des SIN-LTRs d'un vecteur gammarétroviral (Gaussin et al., 2012).

L'insulateur cHS4 (de 1,2 kb) possède les activités *enhancer*-bloqueur et barrière, dont la première nécessite la fixation du facteur CTCF (Aker et al., 2007; Arumugam et al., 2009; Chung et al., 1993). La présence de cHS4 au sein de vecteurs rétroviraux a été décrite comme réduisant le risque d'oncogénèse insertionelle (Evans-Galea et al., 2007; Ryu et al., 2007), bien qu'un cas de dominance clonale ait été récemment reporté pendant un essai clinique utilisant un vecteur SIN contenant cHS4 (Cavazzana-Calvo et al., 2010). De manière frappante, l'activité barrière de cHS4 au sein de vecteurs rétroviraux varie grandement en fonction du *design* des vecteurs étudiés (Emery et al., 2000; Nielsen et al., 2009; Uchida et al., 2011; Yannaki et al., 2002). L'insertion de cHS4 dans le 3' LTR a notamment été décrite réduire le titre viral fonctionel (Arumugam et al., 2009; Hanawa et al., 2009; Nielsen et al., 2009; Uchida et al., 2011), du à une moindre transcription inverse (Arumugam et al., 2009; Hanawa et al., 2009) et/ou à une intégration de séquences virales incomplètes (Arumugam et al., 2009; Nielsen et al., 2009). L'insertion de séquences autres que cHS4 au sein des SIN-LTRs a aussi été explorée, telle celle de MAR, sans succès pour le moment (Grandchamp et al., 2011).

L'impact réel de l'insertion de séquences hétérologues au sein des SIN-LTRs reste donc à élucider. Cette tâche peut se révéler complexe de part les propriétés intrinsèques des séquences clonées. Etant donné les nouvelles caractéristiques que pouvaient conférer l'insertion de fragments exogènes au sein des SIN-LTRs des vecteurs rétroviraux, nous avons analysé les titres viraux, l'intégration, l'efficacité de transduction ainsi que l'expression du transgène d'un set de vecteurs lentiviraux intégrase-compétent ou -déficient. Ceux-ci contiennent, au sein des SIN-LTRs, des inserts d'origine aviaire (cHS4), humaine (*i.e.*, insulateur putatif et séquence « déserte ») ou bactérienne (*spacer*).

Nous avons observé que ces insertions induisent le plus souvent une diminution de la transcription du transgène. Dans un contexte lentiviral et gammarétroviral, ce phénomène d'attenuation est dépendant de l'orientation, stable à long terme (au moins 3 mois) et n'augmente pas l'extinction ou la variégation. Les vecteurs générés au cours de cette étude permettent la modulation de la transcription du transgène d'un facteur 0,3 à 1,8 (comparé au vecteur parental) de façon stable ou transitoire (en fonction du statut de l'intégrase). Ils pourraient ainsi constituer des outils génétiques intéressants pour réguler finement l'expression du transgène.

B/ ARTICLE : '3'SELF-INACTIVATING LONG TERMINAL REPEAT INSERTS FOR THE MODULATION OF TRANSGENE EXPRESSION FROM LENTIVIRAL VECTORS'

(PUBLIE DANS LE JOURNAL HUMAN GENE THERAPY METHODS, 2012 APR;23(2):84-97)

HUMAN GENE THERAPY Part B: Methods 23:1– (April 2012)
© Mary Ann Liebert, Inc.
DOI: 10.1089/hgtb.2011.154

3′ Self-Inactivating Long Terminal Repeat Inserts for the Modulation of Transgene Expression from Lentiviral Vectors

Gwenola Manic,[1] Aurélie Maurin-Marlin,[1] Lorenzo Galluzzi,[2,3] Frédéric Subra,[1] Jean-François Mouscadet,[1,4] and Stéphanie Bury-Moné[1]

Abstract

Gene transfer for research or gene therapy requires the design of vectors that allow for adequate and safe transgene expression. Current methods to modulate the safety and expression profile of retroviral vectors can involve the insertion of insulators or scaffold/matrix-attachment regions in self-inactivating long terminal repeats (SIN-LTRs). Here, we generated a set of lentiviral vectors (with internal CMV or PGK promoter) in which we inserted (at the level of SIN-LTRs) sequences of avian (i.e., chicken hypersensitive site-4, cHS4), human (i.e., putative insulator and desert sequence), or bacterial origin. We characterized them with respect to viral titer, integration, transduction efficiency and transgene expression levels, in both integrase-proficient and -deficient contexts. We found that the cHS4 insulator enhanced transgene expression by a factor of 1.5 only when cloned in the antisense orientation. On the other hand, cHS4 in the sense orientation as well as all other inserts decreased transgene expression. This attenuation phenomenon persisted over long periods of time and did not correspond to extinction or variegation. Decreased transgene expression was associated with lower mRNA levels, yet RNA stability was not affected. Insertions within the SIN-LTRs may negatively affect transgene transcription in a direct fashion through topological rearrangements. The lentiviral vectors that we generated constitute valuable genetic tools for manipulating the level of transgene expression. Moreover, this study demonstrates that SIN-LTR inserts can decrease transgene expression, a phenomenon that might be overcome by modifying insert orientation, thereby highlighting the importance of careful vector design for gene therapy.

Introduction

Gene transfer for research or gene therapy requires the development of safe and efficient delivery tools. Approximately 20% of current clinical trials involving gene therapy rely on retroviral transgenesis (http://www.wiley.co.uk/genmed/clinical/). Lentiviral (LV) or gammaretroviral (GRV) vectors, which are derived from human immunodeficiency virus type 1 (HIV-1) or Moloney leukemia virus (MLV), respectively, represent suitable tools for gene transfer for several reasons. First, their genome is simple; second, these vectors have been extensively characterized; third, they can integrate in the host cell genome, potentially ensuring long-term transgene expression (Pauwels *et al.*, 2009). Of note, LV and GRV vectors are characterized by different infection and integration patterns (Mitchell *et al.*, 2004; Yamashita and Emerman, 2006).

However, there are toxicity issues (related to the vector and/or the transgene) that must be carefully considered for gene therapy approaches (reviewed in Chang and Sadelain, 2007; Pauwels *et al.*, 2009). Unspecific integration can lead to the inactivation of a cellular gene or to changes in the expression of host genes that reside in the vicinity of the provirus (e.g., after *trans*-activation, promoter interference, or read-through into neighboring genes). These events can confer a selective advantage to transduced cells, increasing the risk of clonal dominance or even oncogenesis (Pauwels *et al.*, 2009). Thus, the design of safe and efficient vectors for transgene expression remains a great challenge, and integrase-deficient lentivectors (IDLVs) are emerging as an alternative approach (Banasik and McCray, 2010).

In this context, the use of self-inactivating (SIN) vectors, which lack enhancer/promoter sequences in the U3 region of

[1]Laboratoire de Biologie et de Pharmacologie Appliquée, UMR 8113 CNRS, Ecole Normale Supérieure de Cachan, FR-94230 Cachan, France.
[2]INSERM, U848, FR-94800 Villejuif, France.
[3]Université Paris Descartes, Sarbonne Paris Cité, FR-75006, Paris, France.
[4]Present address: Bio-Rad Laboratories, 92430 Marnes-la-Coquette, France.

the long terminal repeats (LTRs), reportedly reduces genotoxicity, in particular if they contain an internal promoter that is moderately active (Montini *et al.*, 2006, 2009; Zychlinski *et al.*, 2008; Maruggi *et al.*, 2009; Zhou *et al.*, 2010). The configuration of SIN vectors allows for the insertion of sequences within the 3′ LTR; these sequences are duplicated in the 5′ LTR during retrotranscription (Pauwels *et al.*, 2009). The elements most commonly incorporated into SIN-LTRs include (1) transgenes (e.g., short hairpin RNAs); (2) upstream polyadenylation enhancer elements (or upstream sequence elements, USEs), which limit read-through into neighboring (onco)genes; and (3) insulators, which either block the interaction between the enhancer and promoter (so-called "enhancer-blockers") or stop the propagation of heterochromatin (so-called "barrier insulators") (reviewed in Pauwels *et al.*, 2009; Schambach *et al.*, 2009), as illustrated by the introduction of short synthetic insulators within SIN-LTRs of a GRV vector (Gaussin *et al.*, 2012).

The canonical chicken hypersensitive site-4 (cHS4) insulator (a 1.2-kb fragment) possesses both enhancer-blocking and barrier activities, of which the former (and perhaps also the latter) requires the binding of the CCCTC-binding factor (CTCF) (Chung *et al.*, 1993; Aker *et al.*, 2007; Arumugam *et al.*, 2009). The presence of cHS4 within retroviral vectors appears to reduce the risk of insertional oncogenesis (Evans-Galea *et al.*, 2007; Ryu *et al.*, 2007), although one case of clonal dominance has been reported during a clinical trial based on a cHS4-containing SIN vector (Cavazzana-Calvo *et al.*, 2010). Strikingly, the cHS4 barrier activity in retroviral vectors greatly varies with vector design (Emery *et al.*, 2000; Yannaki *et al.*, 2002; Nielsen *et al.*, 2009; Uchida *et al.*, 2011) and insertion of cHS4 into the 3′ LTR reportedly reduces the functional titer (Hanawa *et al.*, 2009; Nielsen *et al.*, 2009; Urbinati *et al.*, 2009; Uchida *et al.*, 2011), due to impaired reverse transcription (Hanawa *et al.*, 2009; Urbinati *et al.*, 2009) and/or integration of incomplete viral sequences (Nielsen *et al.*, 2009; Urbinati *et al.*, 2009). The insertion of sequences other than cHS4 has also been explored. Bos and colleagues demonstrated that large double-copy vectors that contain different promoter–transgene cassettes within their SIN-LTRs are functional, yet exhibit a size-dependent decline in transduction efficiency (Bos *et al.*, 2010). The inclusion of CTCF-and CTF/NF1-binding sites was also shown to have a little effect on titers obtained from a GRV vector (Gaussin *et al.*, 2012).

Nevertheless, the real impact of cloning exogenous sequences within SIN-LTRs remains to be elucidated, a task that is complicated by the intrinsic properties of cloned sequences. Because the insertion of exogenous fragments within the LTR may confer novel properties to retroviral vectors, we analyzed viral titer, integration, transduction efficiency, and transgene expression of a set of integrase-proficient or -deficient LV vectors containing, within the SIN-LTRs, inserts of avian (cHS4), human (i.e., putative insulator and desert sequence), or bacterial ("spacer") origin. These insertions most often led to decreased transgene transcription. In the LV and GRV context, this attenuation phenomenon was orientation dependent, was stable over long periods of time, and did not increase extinction or variegation. The vectors that we generated allow for the modulation of transgene transcription from 0.3- to 1.5-fold (compared with the parental vector), and may constitute valuable genetic tools for the refined control of transgene expression.

Materials and Methods

Vector constructs

The genetic organization of LV and GRV vectors (see Supplementary Table S1; supplementary data are available online at www.liebertonline.com/hgtb) is represented in Fig. 1. All LV constructs were derived from RRLSIN.cPPT.hPGK-gfp (shortened to "LV-P_{PGK}"), an SIN LV vector carrying the enhanced *gfp* gene under the control of the human internal phosphoglycerate kinase (PGK) promoter (P_{PGK}). RRLSIN.cPPT .hPGK-gfp was modified from a previously described vector (Zufferey *et al.*, 1998) (a gift from the University of Turin Medical School, Turin, Italy) by deletion of the woodchuck hepatitis posttranscriptional regulatory element (WPRE). All GRV constructs were derived from MLV and the murine sarcoma virus (MuSV)-based LNCX vector (Clontech/Takara Bio, Mountain View, CA), into which a multiple cloning site (including an *Xho*I restriction site) was introduced between the *Hin*dIII and *Cla*I sites ("LNCX*"). The vector lacking both the transgene and the internal promoter ("GRV-LTR-empty1") was generated by deletion of the *Bcl*I–*Bgl*II fragment of LNCX*, containing the neomycin resistance gene and the cytomegalovirus (CMV) promoter. A module containing the *GFP* transgene under the control of the PGK promoter (*Eco*RV–*Sal*I fragment from LV-P_{PGK}) was introduced between the compatible *Bsa*BI and *Sal*I sites of GRV-LTR-empty1. GRV-P_{PGK}, a self-inactivating version of the resulting GRV-LTR-P_{PGK} vector, was generated by deleting 177 bp from the U3 3′ LTR as follows: the product of a PCR amplification of LNCX* performed with OBM95 and OBM96 primers (Supplementary Table S2) was digested with *Xho*I and *Xba*I and used to replace the *Xho*I–*Xba*I region of GRV-LTR-P_{PGK}. For the insertions of fragments in LV vectors, a single *Xba*I site was introduced into the 3′ SIN-LTR of LV-P_{PGK} by directed PCR mutagenesis, done with a QuikChange XL site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) and OBM164 and OBM165 primers (Supplementary Table S2). The cHS4 1.2-kb insulator was obtained by digestion of the pJC13-1 plasmid kindly provided by G. Felsenfeld (National Institutes of Health, Bethesda, MD) (Chung *et al.*, 1993). The 1.2-kb cHS4 insulator *Xba*I fragment was inserted into the *Xba*I site within the 3′ SIN-LTR of LV-P_{PGK}. The 1.07-kb Ins023′ fragment (chr2: 220214795-220215860, Hg18 coordinates) was obtained by amplifying human genomic DNA (from CEM-T4 cells) with the OBM176 and OBM178 primers (Supplementary Table S2). The resulting amplicon was digested with *Xba*I and introduced into the *Xba*I site within the 3′ SIN-LTR of LV-P_{PGK}. The 1.59-kb DS18 fragment (chr2: 220332055-220333640, Hg18 coordinates) was obtained by amplifying human genomic DNA (from HeLa cells) with the OBM29 and OBM30 primers (Supplementary Table S2) followed by digestion with *Bam*HI and the introduction of the resulting fragment into the 3′ SIN-LTR at the *Bbs*I site of LV-P_{PGK} or at the *Xba*I site of GRV-P_{PGK} (with blunting before ligation). The 1.54-kb fragment of the *Escherichia coli lacZ* gene was obtained by amplification of pHR′-CMV-LacZ (Naldini *et al.*, 1996) with the OBM179 and OBM180 primers (Supplementary Table S2), followed by digestion with *Bam*HI or *Xba*I and introduction of the resulting fragment into the *Bam*HI site within the 3′ SIN-LTR of LV-P_{PGK}. All insertions (selected for the sense or antisense orientation) were checked by plasmid sequencing. Deleted

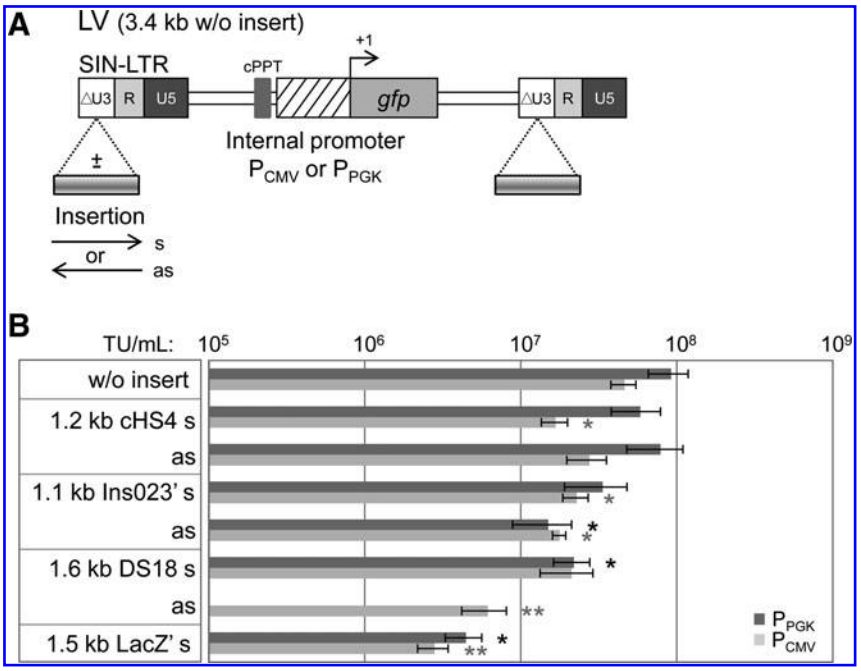


FIG. 1. Retroviral vector design and impact on functional titers. **(A)** Schematic representation of the lentiviral (LV) vectors (in the proviral, integrated form) used in this study. These vectors encode the green fluorescent protein transgene (*gfp*) under the control of either the immediate-early human cytomegalovirus promoter (P_{CMV}) or human phosphoglycerate kinase promoter (P_{PGK}) and contain a self-inactivating long terminal repeat (SIN-LTR). Within the SIN-LTRs, inserts were cloned in either the sense (s) or antisense (as) orientation (except for the LacZ′ sequence, which was introduced only in the sense orientation). cPPT, central polypurine tract; Ψ, encapsidation signal. **(B)** Human cervical carcinoma HeLa cells were transduced with serial dilutions of GFP-encoding vectors based on the indicated promoter and containing within their SIN-LTRs the indicated inserts (insert size is reported in kilobases) in the sense (s) or antisense (as) orientation and, 48 hr later, functional titers were assessed. Bars report functional titers as determined in at least 11 independent titrations of at least 4 productions made with 2 different shuttle plasmid preparations (means ± SEM, $^*p < 0.05$ and $^{**}p < 0.01$ compared with parental vectors based on the same internal promoter). Statistically significant differences were not observed between P_{PGK}- and P_{CMV}-based vectors that contained the same inserts ($p > 0.05$). TU, transduction units.

forms of DS18 were generated by double-digestion of the vectors with appropriate enzymes (*Bbs*I–*Afe*I, *Bbs*I–*Xba*I, or *Sgr*AI–*Afe*I; as indicated in Fig. 6), followed by blunting and ligation. On the other hand, deletion of the mammalian-wide interspersed repeat (MIR) element was achieved by directed mutagenic PCR (as described previously) of DS18-containing vectors, performed with the OBM174 and OBM175 primers (Supplementary Table S2). Replacement of the CMV promoter with the PGK promoter in LV vectors was done by substituting the *Cla*I–*Bam*HI region (PGK promoter) with the *Cla*I–*Bam*HI fragment (CMV promoter) of the pHR′-CMV-eGFP vector (Miyoshi *et al.*, 1997). All constructs were maintained and stored in the Stratagene SURE 2 *E. coli* host strain (Agilent Technologies).

LV and GRV vectors were produced with the psPAX2 or pCMVR8.74 packaging plasmid (kindly given by the D. Trono laboratory, Lausanne, Switzerland; Addgene plasmids 12260 and 22036, respectively) encoding HIV-1 Gag-Pol, Tat, and Rev proteins; and the Stratagene pVPack-GP packaging plasmid (Agilent Technologies) encoding MLV Gag-Pol proteins, respectively. IDLVs were produced with the D116A pCMV-R8.74 packaging plasmid encoding a mutated (D116A) integrase (Drelich *et al.*, 1992). This mutation was introduced into the pCMV-R8.74 packaging plasmid by directed PCR mutagenesis (as described above) and OBM156 and OBM157 primers (Supplementary Table S2). All vectors were pseudotyped with the vesicular stomatitis virus envelope glycoprotein (VSV-G) encoded by pMD.G (Ory *et al.*, 1996).

Cell culture and treatment

Human cervix adenocarcinoma epithelial HeLa cells and human embryonic kidney (HEK) 293T cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA; CCL-2 and CRL-11268, respectively) and maintained in Dulbecco's modified Eagle's medium. Human T lymphoblastoid CEM-T4 cells (a naturally isolated subclone of the CEM line that displays high levels of surface CD4 expression; obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy and Infectious Diseases, National Institutes of Health: CEM-T4 from J.P. Jacobs) (Foley *et al.*, 1965) were cultured in RPMI 1640. All media were supplemented with 10% fetal bovine serum (FBS; PAA Laboratories, Pasching, Austria) and 1% penicillin–streptomycin (P/S; 100 units/ml). Wild-type human colorectal carcinoma epithelial HCT 116 cells (kindly provided by E.E. Hendrickson) and Chinese hamster ovary (CHO) cells (CRL-9618; ATCC) were maintained in McCoy's 5A and minimal essential medium (MEM), respectively. Media were supplemented with 10% FBS, 1% P/S, and 10% essential amino acids for MEM. All cell lines used in this study were incubated at 37°C, under an atmosphere

containing 5% CO_2. Unless otherwise indicated, media and supplements for cell culture were purchased from Gibco/Life Technologies (Carlsbad, CA) and BD Falcon plasticware was obtained from BD Discovery Labware (Bedford, MA).

For the analysis of RNA stability, cells were treated with actinomycin D (Sigma-Aldrich, St. Louis, MO) at a concentration of 5 μM (Leclerc *et al.*, 2002) in 25-cm^2 flasks. Six hours after treatment, 10^5 cells were harvested, centrifuged for 5 min at $270 \times g$, washed once with phosphate-buffered saline (PBS), and stored at −80°C for total RNA extraction.

Vector production

Retroviral vectors were produced by calcium phosphate-mediated transfection of 293T cells as previously described (Chen and Okayama, 1987; Masson *et al.* 2007) with the following modifications. Briefly, 10^6 cells were plated in 75-cm^2 flasks and maintained in culture for 3 days. Two hours before the transfection, medium was replaced with fresh medium. Meanwhile, a mixture of plasmids composed of 6 μg of shuttle vector plasmid, 4 μg of packaging plasmid (pVPack-GP) (psPAX2, pCMV-R8.74 or its derivate encoding a mutated integrase), and 2 μg of pMD.G per flask was prepared. A precipitate was formed by mixing the plasmids with 1 ml of 0.25 *M* $CaCl_2$ and adding the resulting solution mix to 1 ml of HEPES-buffered saline (280 m*M* NaCl, 10 m*M* KCl, 1.5 m*M* Na_2HPO_4, 50 m*M* HEPES; pH 7.05) while vortexing. The mixture was incubated for 20 min at room temperature and then added to the cells. The next day, the medium was replaced with fresh medium. Supernatants were then harvested 48 and 72 hr after transfection, centrifuged for 5 min at $270 \times g$, following by centrifugation at $1730 \times g$ for 20 min and then ultracentrifugation for 2 hr at 20,000 rpm, at a temperature of 4°C, in an Optima L-80 XP ultracentrifuge (Beckman Coulter, Brea, CA) with an SW 28 Ti rotor. Vectors were finally resuspended in PBS at a concentration ratio of 200 for lentiviral vectors and of 150 for gammaretroviral vectors, and aliquots were stored at −80°C. No difference between viral titers was observed when the psPAX2 or pCMV-R8.74 packaging plasmid was employed for vector production.

Transduction procedure and analysis

The day before transduction, 10^5 cells per well were seeded in duplicate for each set of conditions, in 12-well plates. Where indicated, 25 μM 3′-azido-3′-deoxythymidine (AZT; Sigma-Aldrich) or 400 n*M* raltegravir (median inhibitory concentration [IC_{50}], 20 n*M*; Prestwick Chemical, Washington, DC) was added 2 hr before transduction. Two, 3, 6, and 10 days after transduction, the percentage and geometric mean fluorescence intensity (MFI) of green fluorescent protein-positive (GFP^+) cells were assessed by flow cytometry on a FACSCalibur cytofluorometer (BD Biosciences, San Jose, CA). The multiplicity of infection (MOI) was calculated from the percentage of GFP^+ cells during the linear phase of infection of HeLa cells. For experiments in which we explored the expression of transgene in a GFP^+ population of transduced cells, GFP^+ cells were sorted by fluorescence-activated cell sorting (FACS) 1 week after transduction at low MOI by cytofluorometry and cultured for 3 months. Sorted cells were subjected to expression analysis every week, as described previously.

Quantification of viral DNA species

Twenty-four and 72 hr after transduction, cells were recollected, centrifuged for 5 min at $270 \times g$, and washed once with PBS. Genomic DNA (gDNA) was isolated with a QIAamp blood DNA minikit (Qiagen, Hilden, Germany) and quantified with a NanoVue spectrophotometer (GE Healthcare, Chalfont St. Giles, UK). Primers and probes (Supplementary Table S2) for total lentiviral DNA amplification (MH531/MH532 primers and MH-FL/MH-FC probes for HR′ and SIN vectors) and integrated viral DNA nested PCR (L-M667, Alu1 and Alu2 primers for the first PCR; Lambda-T/AA55M primers and LTR-FL/LTR-FC probes for the second PCR) have been described elsewhere (Brussel and Sonigo, 2003). Quantitative PCR (qPCR) with these primers and probes was carried out with 2 μl (between 40 and 500 ng) of gDNA in a final volume of 20 μl of LightCycler FastStart DNA Master HybProbe mix, according to the manufacturer's recommendations (Roche Diagnostics, Basel, Switzerland). Amplification was carried out on a LightCycler (Roche Diagnostics) as previously described (Brussel and Sonigo, 2003). Each sample was examined in duplicate with each set of primers. The cell equivalents in sample DNA were calculated on the basis of the amplification of the β-globin gene (two copies per diploid cell) with the LightCycler instrument and commercially available materials (control kit DNA; Roche Diagnostics). Copy numbers of both total and integrated viral DNA were determined with reference to two DNA standard curves obtained by amplification of 10^1 to 10^4 HeLa cells containing one pR7-Neo Δ*env* provirus per cell (Brussel and Sonigo, 2003) and 10^2 to 10^6 copies of linearized LV-P_{PGK} plasmid, as recommended by Hou and colleagues (2010). Viral DNA titers were determined on the basis of total viral DNA quantified 24 hr posttransduction, a time point at which viral DNA is maximal (Brussel and Sonigo, 2003).

Quantitative RT-PCR

Total RNA was extracted from 10^5 cells with an RNeasy mini kit (Qiagen) and resuspended in 50 μl of diethylpyrocarbonate (DEPC)-treated water. Genomic DNA contamination was removed by treating 25 μl of all mRNA samples with 20 U of RNase-free DNase (Roche Diagnostics) for 1 hr at 37°C, according to the manufacturer's instructions. RNA was repurified on RNeasy mini kit columns (Qiagen). The quantity and quality of the isolated RNA were determined with a NanoVue spectrophotometer (GE Healthcare). DNA-free RNA (100–150 ng) was reverse-transcribed in a 20-μl final reaction volume with PrimeScript reverse transcriptase (Takara Bio, Shiga, Japan) for 1 hr at 42°C, with 2 pmol of OBM13 and OBM242 primers (Supplementary Table S2). To confirm the specificity of RT-qPCR, samples without template or without reverse transcriptase were used as negative controls. The OBM12–OBM13 and OBM241–OBM242 primer pairs (Supplementary Table S2) were used to amplify the *gfp* and *GAPDH* genes, respectively, from 2 μl (between 10 and 20 ng) of complementary DNA in a final volume of 20 μl of LightCycler DNA Master SYBR Green I mix, according to the manufacturer's recommendations (Roche Diagnostics). Each sample was analyzed in duplicate for each set of primers. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as an invariant internal control (Zheng *et al.*, 2008). Fold-change values were calculated by comparative C_t

analysis after normalizing for the quantity of *GAPDH* mRNA in the samples concerned (Pfaffl, 2001).

Statistical procedures

Unless otherwise specified, experiments were performed in triplicate parallel instances and independently repeated at least twice. Data were analyzed with Microsoft Excel (Microsoft, Redmond, WA) and statistical significance was assessed by means of two-tailed Student *t* tests. Unless otherwise indicated, asterisks depict statistical significance ($*p < 0.05$, $**p < 0.01$).

Results

Retroviral vector design

We took advantage of a *gfp*-encoding HIV-1-based vector (Zufferey *et al.*, 1998) and modified it at the promoter level as well as by inserting different elements, in sense or antisense orientation, within its SIN-LTRs (Fig. 1A, and Supplementary Table S1). To overcome possible promoter-specific effects and to determine the impact of insertions on transcription from both moderate and strong promoters, we used the human PGK and immediate-early human CMV promoters (P_{PGK} and P_{CMV}, respectively). The elements inserted within the SIN-LTRs were as follows: (1) the 1.2-kb avian cHS4 insulator, (2) a 1.1-kb modified version of the human Ins023 sequence (DHS3) (Xi *et al.*, 2007) (which we called Ins023′), (3) a 1.6-kb human sequence from a regulatory factor-binding desert (which we named DS18), and (4) a 1.5-kb spacer sequence from the *Escherichia coli lacZ* gene (which we called LacZ′).

Using the University of California at Santa Cruz (UCSC) Genome Browser (http://genome.ucsc.edu/), we selected Ins023 because, similar to cHS4, it binds CTCF (Kim *et al.*, 2007). In cell culture-based insulator assays, Ins023 displays enhancer-blocking activity three times stronger than that of cHS4 (Xi *et al.*, 2007). Moreover, Ins023 is located on chromosome 2 at the frontier between an island and a regulatory factor-binding desert (Zhang *et al.*, 2007), suggesting that this Ins023 may also have barrier properties. To avoid any possible interference from the endogenous retroviral element (ERV) contained in Ins023, we removed it by PCR amplification, thereby generating Ins023′. DS18 corresponds to a sequence on chromosome 2 located in the regulatory factor-binding desert downstream of Ins023 (Zhang *et al.*, 2007), which presents heterochromatin marks (e.g., H3K27me3). In contrast to Ins023′ and cHS4, DS18 does not bind CTCF and contains a CAT repetition as well as a mammalian-wide interspersed repeat (MIR) of the subfamily 2, also known as MIRb (Smit and Riggs, 1995). The 1.54-kb fragment from *E. coli lacZ* (LacZ′) was cloned into the SIN-LTRs not in frame, to serve as an exonic spacer control.

By this approach, we wanted to elucidate the properties conferred to the vector by SIN-LTR inserts that bind CTCF and display enhancer-blocking and/or barrier activity (i.e., cHS4, Ins023′), while discriminating these from effects provided by nonspecific sequences (i.e., DS18, LacZ′).

Impact of vector design on lentiviral titers

Functional viral titers (which reflect the percentage of cells expressing GFP for given quantity of virus) were estimated by transducing human cervical carcinoma HeLa cells with serial dilutions of vector stocks followed by cytofluorometric quantification of GFP-expressing cells (48 hr posttransduction). The linear phase of transduction (approximately 3–30% GFP^+ cells) was used to calculate the functional titer. To precisely discriminate between GFP transduction and pseudotransduction (i.e., the presence of proteins carried on by viral particles) (Haas *et al.*, 2000), we performed a parallel analysis of the percentage of GFP^+ cells in the presence or absence of AZT, an inhibitor of reverse transcription (Ruprecht *et al.*, 1986). This led us to exclude the LV-P_{PGK}-DS18as vector from further analysis, as it was associated with consistent pseudotransduction (Supplementary Fig. S1). The production of infectious viral particles from parental vectors was not significantly affected by the internal promoter (Fig. 1B). With the exception of cHS4 in the LV-P_{PGK} backbone, SIN-LTR modifications tended to decrease functional viral titers in a pattern that did not correlate with CTCF-binding properties or sequence origin (Fig. 1B). Decrease in viral titer due to insert size has been reported in previous studies (Ramezani *et al.*, 2003; Hanawa *et al.*, 2009; Nielsen *et al.*, 2009; Urbinati *et al.*, 2009; Bos *et al.*, 2010; Uchida *et al.*, 2011). This effect was minimal in our study, perhaps because of vector size (rather small, i.e., from 3.4 to 5 kb, in our study) or because of the nature of SIN-LTR inserts (which in our study always lacked promoter activity). We observed a sequence- and orientation-dependent, rather than size-dependent, effect on functional titers, with

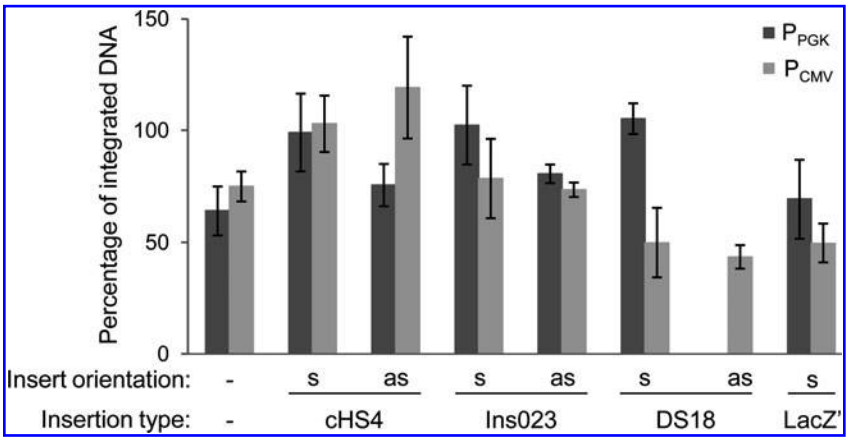


FIG. 2. Insertions within SIN-LTRs do not affect integration efficiency. Human cervical carcinoma HeLa cells were transduced with GFP-encoding, P_{PGK}- or P_{CMV}-based vectors that contained within SIN-LTRs the indicated inserts in either the sense (s) or antisense (as) orientation. Seventy-two hours later, the levels of both total and integrated viral DNA were quantified by qPCR. Columns report the percentage of integrated (out of total) viral DNA (means ± SEM; results are from three independent experiments).

the LacZ′ insertion being more deleterious than the DS18s insertion (Fig. 1B). Of note, the viral DNA titer was decreased only slightly by the presence of inserts within the SIN-LTRs of P_{CMV}-based vectors, except in the case of the LacZ′ insertion (Supplementary Fig. S2). This may be because large inserts (e.g., LacZ′) within the 3′ LTR might affect reverse transcription, as previously reported (Urbinati *et al.*, 2009). Altogether, these data demonstrate that the production of infectious viral particles is a complex phenomenon that is affected by multiple parameters, not only by vector length.

Therefore, we tested our set of vectors for integration efficiency. The insertion of cHS4 into the 3′ LTR has been suggested to facilitate genetic instability, leading to the loss of the transgene expression cassette on recombination between LTRs (Nielsen *et al.*, 2009). Because this phenomenon may introduce biases into the PCR-based determination of viral integration efficiency, we specifically quantified integrated viral DNA with probes and primers recognizing human *Alu* repeats and viral LTR sequences, as previously described (Brussel and Sonigo, 2003). DNA levels were determined 24 and 72 hr after transduction at an MOI of 0.1 (average percentage of GFP^+ cells, 10% at 48 hr, detected by cytofluorometry). Total viral DNA was greatly reduced (10- to 25-fold) 72 hr posttransduction (as compared with 24 hr after transduction), corresponding to 0.04 ± 0.01 copy per cell of (mainly integrated) vector DNA (Fig. 2, and Supplementary Fig. S3). Minimal variations were observed with the various inserts (Supplementary Fig. S3). Of note, the efficiency of integration was not affected by the insert size and composition (Fig. 2). These data indicate that a defect at the integration step does not explain the functional titer reduction observed in Fig. 1. We therefore investigated whether such a reduction might be linked to transgene expression in transduced cells.

Impact of insertions within SIN-LTRs on transgene expression early after transduction

By analyzing *gfp* transduction and the expression levels of GFP by cytofluorometry, we studied how the internal promoter and the genetic elements cloned within SIN-LTRs influence transgene expression. We decided to study these aspects on the basis of an MOI less than 0.5 (i.e., to 3 to 50% GFP^+ cells) so as to work in conditions in which cells were transduced in the linear phase of functional titration. We analyzed GFP positivity as well as the geometric mean fluorescence intensity (MFI) of the GFP signal for up to 10 days posttransduction, according to standardized, currently accepted procedures (Buceta *et al.*, 2011). To elucidate the influence of integrated versus unintegrated DNA forms on transgene expression, the same experiments were performed with integrase-proficient vectors in the presence or absence of the integrase inhibitor raltegravir (Hazuda *et al.*, 2000), or with IDLVs (encoding a functionally inactive integrase variant). Efficient integrase inhibition was assessed by the absence of highly fluorescent cells 48 hr posttransduction for LV-P_{PGK} vectors (Supplementary Fig. S4) and by the rapid loss of GFP expression on mid-term culture (Supplementary Fig. S5).

Forty-eight hours posttransduction, at comparable percentages of GFP^+ cells, we found that the cHS4 insulator cloned in the antisense orientation leads to a 1.5-fold increase in early transgene expression from the strong P_{CMV} internal promoter (Fig. 3A). Consistent with previous observations (Emery *et al.*, 2000; Yannaki *et al.*, 2002; Hanawa *et al.*, 2009; Uchida *et al.*, 2011), this does not hold true when the same sequence is cloned in reverse (sense) orientation, or in the presence of the moderately active P_{PGK} promoter. These results corroborate the notion that P_{CMV} is more susceptible to short-term silencing than P_{PGK} (Gerolami *et al.*, 2000) and suggest that transcription from P_{CMV} is facilitated by the

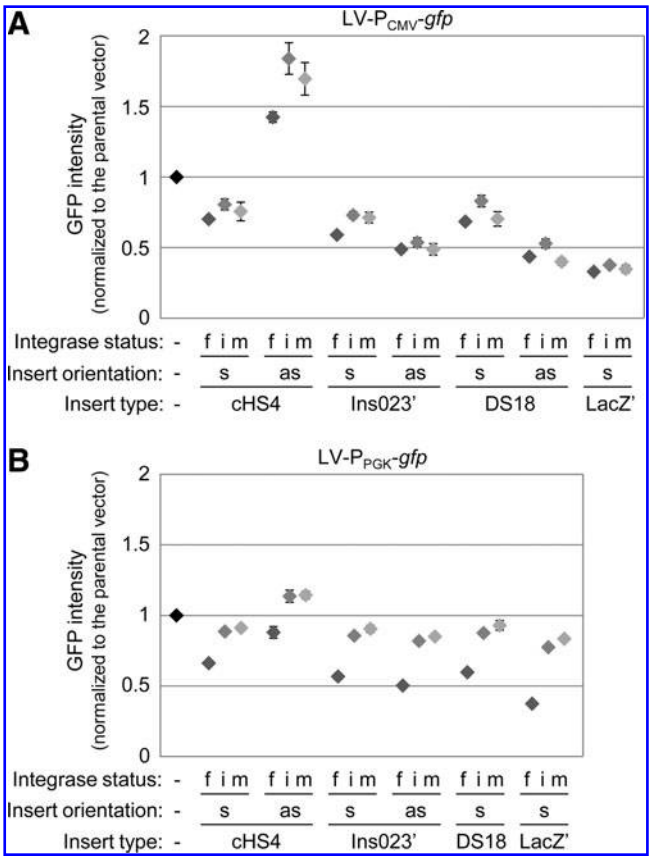


FIG. 3. Insertions within SIN-LTRs modulate early transgene expression from integrase-proficient or -deficient vectors in an orientation-dependent fashion. Human cervical carcinoma HeLa cells were transduced with GFP-encoding vectors based either on the **(A)** CMV promoter or **(B)** PGK promoter, and containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. Forty-eight hours later, cytofluorometry was employed to determine the percentage and intensity of GFP expression. Cell populations that exhibited comparable transduction efficiency (percentage of GFP^+ cells) were employed to estimate GFP expression levels. Here, the intensity of GFP expression from insert-containing vectors is normalized to that of the corresponding parental vectors (means ± SEM; values were calculated from at least 6 MOI obtained from at least six independent experiments). Note that, 72 hr from transduction, the percentage of integrated viral DNA (determined by qPCR experiments) is significantly reduced (50-fold) in cells transduced with integrase-deficient vectors in comparison with cells transduced with integrase-proficient vectors. f, functional, wild-type integrase; i, integrase inhibition with raltegravir; m, mutated integrase.

insulator capacity of cHS4. Such a positive impact of the antisense cHS4 insert on P_{CMV}-driven transgene expression is independent of integration (Fig. 3A) and declines with time (Fig. 4A). This suggests that cHS4 inserted in the antisense orientation influences mostly early expression and/or transcription from unintegrated DNA. Perhaps, the strong P_{CMV}-driven expression from unintegrated circular DNA is improved in the presence of the transcriptional termination that is associated with antisense cHS4 (Hanawa *et al.*, 2009). On the other hand, the antisense cHS4 insert positively modulates P_{PGK}-driven transgene expression only on integration and mid-term culture (Fig. 4B). Taken together, our results suggest that the antisense cHS4 enhances transgene expression in a promoter- and integration-dependent fashion.

Surprisingly, 48 hr after transduction, the insertion within SIN-LTRs of all sequences studied but cHS4 in the antisense orientation led to a significant decrease in GFP expression (Fig. 3). Notably, for DS18-containing and (although to a lesser extent) Ins023′-containing vectors, we recorded a difference in transgene expression that was promoter independent but orientation dependent. Thus, DS18 in the antisense orientation led to a more relevant decrease in GFP signal than the same sequence cloned in the sense direction (Fig. 3A). This effect depended on integration for the LV-P_{PGK}, but not the LV-P_{CMV}, backbone (Fig. 3A and B), perhaps because P_{CMV} is stronger than P_{PGK}, and efficient transgene expression from the latter might require integration (Supplementary Fig. S1).

As vectors that contain moderate promoters present an optimal safety profile (Pauwels *et al.*, 2009), we investigated further the reduction of transgene expression associated with cloning within the SIN-LTRs of P_{PGK}-based constructs. To explore putative cell type-dependent effects (Rivella *et al.*, 2000; Yao *et al.*, 2003), we transduced P_{PGK}-based vectors in human colorectal carcinoma HCT 116 cells, human T lymphoblastoid leukemia CEM-T4 cells, and rodent ovarian carcinoma CHO cells. Two days after transduction, vectors containing insertions within their SIN-LTRs were significantly less efficient in expressing GFP than the parental P_{PGK}-based vector, in all tested cell lines (Fig. 5). Such a decrease persisted over 6 days whereas, at this time point, antisense cHS4 had no impact (in CEM-T4 and CHO cells) or a positive impact (in HCT 116 cells) on transgene expression. Thus, vector design affects transgene expression more than cellular determinants. Similarly, the insertion of DS18 in the sense orientation within the SIN-LTRs of an MLV-derived, GFP-encoding, P_{PGK}-based GRV vector decreased transgene expression by a factor of 0.6 ± 0.1 (data not shown), demonstrating that DS18 attenuates transgene expression independently of a vector-specific effect.

We observed a slight tendency for large inserts (i.e., DS18 in the antisense orientation, LacZ′) to decrease transgene expression more relevantly than small inserts (i.e., cHS4 and Ins023′, both in the sense direction) (Fig. 3). However, the specific nature and orientation of the insert played an even more important role, as demonstrated by the activating behavior of cHS4 as well as by the fact that GFP expression levels on transduction with DS18-containing vectors varied from 0.6 ± 0.1 to undetectable, and from 0.7 ± 0.1 to 0.5 ± 0.1 (in the context of the P_{PGK} and P_{CMV} promoter, respectively), depending on insert orientation (Fig. 3).

Intrigued by the expression profile of DS18-containing vectors, we generated four subclones of this sequence in the sense orientation and carried out a systematic analysis to determine which part of DS18 (from within SIN-LTRs) is responsible for reducing transgene expression. Surprisingly, the effect of sense DS18 on GFP expression was not modified by the deletion of a 0.5-kb fragment upstream of the *Xba*I restriction site (sΔBX subclone), by the deletion of a 1.1-kb fragment upstream of the *Afe*I restriction site (sΔBA subclone), or by the concomitant removal of the fragment upstream of the *Xba*I restriction site and of the fragment downstream of the *Afe*I restriction site (sΔBXΔAS subclone) (Fig. 6). As the two latter subclones have no sequence in common, it appears that no specific region of sense DS18 is responsible for its inhibitory effect on transgene expression. Conversely, we found that this effect is exacerbated by deletion of the MIRb element in the sense orientation (Fig. 6). Interestingly, inserts of equal size did not necessarily have similar effects on transgene expression. When the same experiments were performed with antisense DS18, GFP

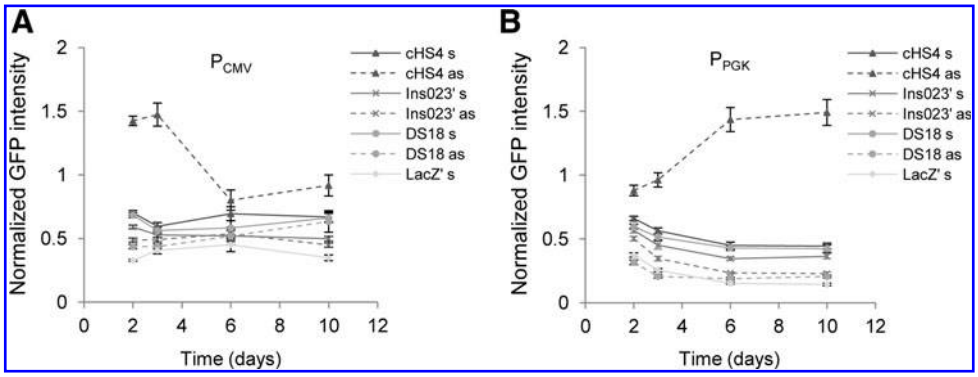


FIG. 4. Kinetics of GFP expression from P_{CMV}- and P_{PGK}-based vectors bearing insertions within their SIN-LTRs. Human cervical carcinoma HeLa cells were transduced with GFP-encoding vectors based either on the **(A)** CMV promoter or **(B)** PGK promoter, and containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. GFP expression was assessed by cytofluorometry 2, 3, 6, and 10 days after transduction, as described in Fig. 3. The intensity of GFP expression from insert-containing vectors is normalized to that of the corresponding parental vectors (means ± SEM; results are from at least 8 MOI obtained from at least eight independent experiments).

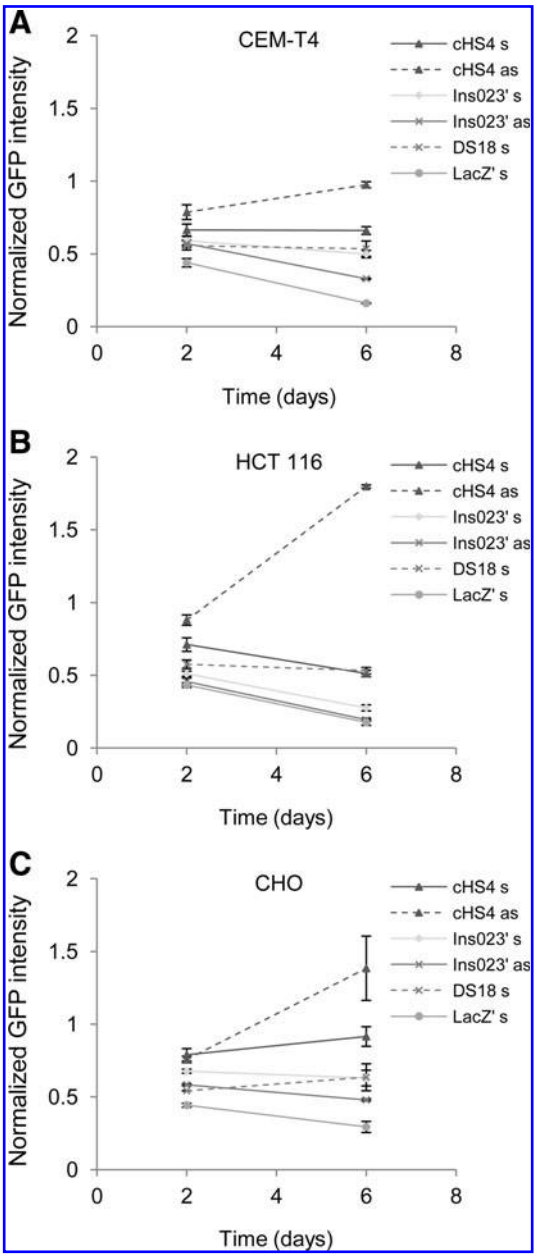


FIG. 5. Attenuation of GFP expression due to insertions within SIN-LTRs is cell type independent. **(A)** Human lymphoblastic CEM-T4 cells, **(B)** human colorectal carcinoma HCT 116 cells, or **(C)** Chinese hamster ovary CHO cells were transduced (at an MOI < 0.3) with GFP-encoding, integrase-proficient, P_{PGK}-based vectors containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. The GFP expression was analyzed by cytofluorometry (as described in Fig. 3), 2 and 6 days after transduction. Here, the intensity of GFP expression from insert-containing vectors is normalized to that of the corresponding parental vectors (means ± SEM; values were calculated from at least 3 MOI obtained from at least three independent experiments).

expression never exceeded the detection threshold imposed by pseudotransduction (data not shown). Taken together, these results and the fact that MIRb elements are described to adopt tRNA-like structures (Smit and Riggs, 1995) suggest that sequence folding at the proviral level, rather than any functional activity provided by the insertion, affects transgene expression. This phenomenon appears to be associated with the presence of an insert within SIN-LTRs and to be dependent on the orientation of the insert, rather than its size.

Impact of SIN-LTR insertions on RNA stability

Because SIN-LTR insertions that are located upstream of the poly(A) signal of the R region (Fig. 1) are transcribed and constitute the 3′ end of the transgene mRNA, they may affect its folding and/or stability. To explore this issue, we transduced CEM-T4 cells (at a low MOI, < 0.15) with P_{PGK}-based vectors containing within their SIN-LTRs either no insert (control vector), cHS4 (in both orientations), or sense DS18 or its derivative sΔBXΔAS (as the minimal sequence presenting an interesting profile of modifications on transgene expression; see previously). We sorted GFP^+ cells by FACS and, 5 weeks later, we analyzed them for *gfp* mRNA content and stability.

We found that cHS4 in the antisense orientation favors mid-term transgene expression from P_{PGK}-based vectors, in contrast to cHS4 cloned in the sense direction as well as DS18 (Fig. 7). Similar to HeLa cells (Fig. 4), CEM-T4 cells manifested increased GFP expression when transduced with a P_{PGK}-based vector containing cHS4 in the antisense orientation within the SIN-LTRs over time (starting from 2 weeks posttransduction) (Fig. 7A). We also observed a correlation of normalized *gfp* mRNA levels and GFP expression with all inserts but cHS4 in the antisense orientation (Fig. 7A), suggesting that decreased transgene expression from sense cHS4- and DS18-containing P_{PGK}-based vectors reflects lowered mRNA levels. Intriguingly, GFP expression from the vector containing cHS4 in the antisense orientation was higher than expected from mRNA levels, pointing to the existence of a posttranscriptional mechanism (e.g., polyadenylation enhancement) (Hanawa *et al.*, 2009) whereby antisense cHS4 improves transgene expression.

To obtain more insights into these phenomena, we quantified *gfp* mRNA by RT-qPCR in cells before and after 6 hr of treatment with actinomycin D (a transcriptional inhibitor). Notably, we failed to detect any difference in the turnover of *gfp* mRNAs transcribed from P_{PGK}-based vectors containing cHS4 and DS18 within their SIN-LTRs (compared with the insert-free vector) (Fig. 7B). Thus, mRNA stability appears not to be involved in the decrease in transgene expression that results from the insertion of sense cHS4 and DS18. Moreover, the *gfp* mRNA species transcribed from the vector that contained cHS4 in the antisense orientation was not more stable than its counterparts (Fig. 7B), indicating that there must be a posttranscriptional mechanism not linked to mRNA stability accounting for increased transgene expression in this context.

Long-term decreases in transgene expression can result from (1) *extinction*, the progressive silencing of an initially expressed provirus during long-term culture or differentiation and/or (2) *variegation*, whereby genetically identical sister cells inheriting the same provirus may express it or not (Ellis, 2005). To formally exclude that DS18-containing P_{PGK}-based vectors might be preferentially associated with these phenomena, we analyzed GFP expression in

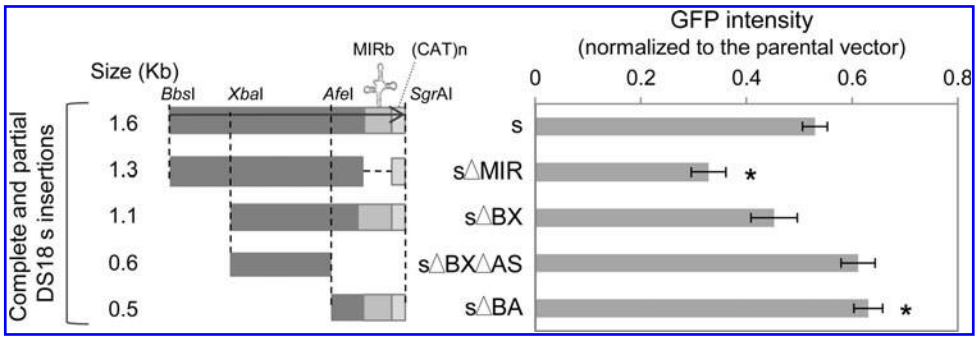


FIG. 6. Effects of the DS18 insert on early transgene expression are not specific to sequence. Human cervical carcinoma HeLa cells were transduced with <0.5 MOI of a GFP-encoding, P_{PGK}-based vector that contained within its SIN-LTRs no insert, full-length DS18 in the sense orientation, or the depicted DS18 restriction fragments. Forty-eight hours later, cells were analyzed as described in Fig. 3. *Left:* Sense DS18 and its restriction fragments are depicted. *Right:* Bars depict GFP expression from insert-containing vectors on normalization to that of the P_{PGK}-based parental vector (means ± SEM; $*p < 0.05$ compared with the vector containing full-length DS18; results are from at least 3 MOI obtained from at least three independent experiments). All the constructs tested were significantly different from the vector without insert. ΔBA, deletion of the region upstream of the *Afe*I site; ΔBXΔAS, deletion of the region upstream of the *Xba*I site and of the region downstream of the *Afe*I site; ΔBX, deletion of the region upstream of the *Xba*I site; ΔMIR, deletion of the MIRb element.

HeLa cells transduced with a P_{PGK}-based vector containing either no insert (control vector) or DS18 in the sense orientation, sorted by FACS for GFP positivity, and kept the cells in culture for 3 months. Attenuation of transgene expression was stable over this period (Fig. 8A and B). Three months after sorting, 97.6% of the cells transduced with the empty P_{PGK}-based vector expressed GFP, some at high levels (Fig. 8B). Such highly fluorescent cells were not present in the 96.6% GFP^+ cell population transduced with the DS18-containing vector, which had a more homogeneous (and less intense) pattern of GFP expression (Fig. 8A and B).

We then determined the percentage of cells expressing GFP from LV-P_{PGK}-based, LV-P_{CMV}-based, and GRV-P_{PGK}-based vectors (containing, or not containing, DS18) at multiple time points after transduction and sorting for GFP positivity. Notably, GFP^+ cells decreased with time, following a perfectly overlapping pattern in the case of P_{PGK}-based vectors irrespective of the vector type and DS18 presence (Fig. 8C). Initially, GFP^+ cells decreased in a relatively pronounced fashion when transduced with P_{CMV}-based vectors (compared with their P_{PGK}-based counterparts) (Fig. 8C), perhaps linked to a higher propensity for silencing (immediate extinction) of the strong CMV promoter (Fig. 8D, and Supplementary Fig. S5) (Gerolami *et al.*, 2000), associated with the fact that the presence of the DS18 insert reduces transgene expression below the threshold of detection. However, after 40 days of cell culture, the difference between LV-P_{CMV}-based vectors containing, or not containing, DS18 remained constant. In conclusion, decreased expression from DS18-containing vectors is distinct from transgene extinction.

Decreased transgene expression by the SIN-LTR insertion of DS18 persisted for at least 3 months and did not increase transgene extinction or variegation

Last, on transduction and sorting, we isolated 7 to 10 (per transduction condition) HeLa cell clones and maintained them in culture for a further 6 weeks. Most of the clones randomly sorted from cells transduced with the DS18-containing vector were characterized by a lower GFP signal compared with those transduced with the control insert-free vector, consistent with the stable nature of the attenuation phenomenon (Supplementary Fig. S6A and B). Moreover, we did not observe any significant difference among the clones relative to the frequency of variegation, which we assessed in terms of (1) a multimodal distribution of GFP expression and (2) the proportion of GFP-negative cells (Fig. 8D, and Supplementary Fig. S7). All together, these results demonstrate that the attenuation phenomenon associated with the insertion of DS18 (in the sense orientation) within SIN-LTRs is stable over long periods of time and does not correspond to transgene extinction or variegation.

Discussion

Gene transfer requires the development of safe, efficient, and flexible genetic tools. The ability of retroviral vectors to efficiently integrate in the host cell genome and express the transgene of interest can be influenced by the presence of exogenous sequences within their (SIN) LTRs. Driven by these considerations, we investigated the impact on retrotransgenesis of large (1.1- to 1.6-kb) sequences of avian (cHS4), human (Ins023′, DS18), or bacterial (LacZ′) origin cloned (in either the sense or antisense orientation) within the SIN-LTRs of (integrase-proficient or -deficient) LV and GRV vectors that express GFP under the control of the strong CMV or the moderate PGK promoter.

Kumar and colleagues have shown that viral titers decrease semilogarithmically with increasing vector size (Kumar *et al.*, 2001). Our results showed a similar tendency: large inserts led to a more remarkable decrease in viral titers than did small inserts (Fig. 1B), irrespective of internal promoter, insert orientation, and CTCF-binding abilities. The reduction of viral titers may be linked to less efficient preintegrative steps (including production and reverse transcription) (Hanawa *et al.*, 2009), genetic instability (Nielsen *et al.*, 2009; Urbinati *et al.*, 2009), and expression level in the target cell.

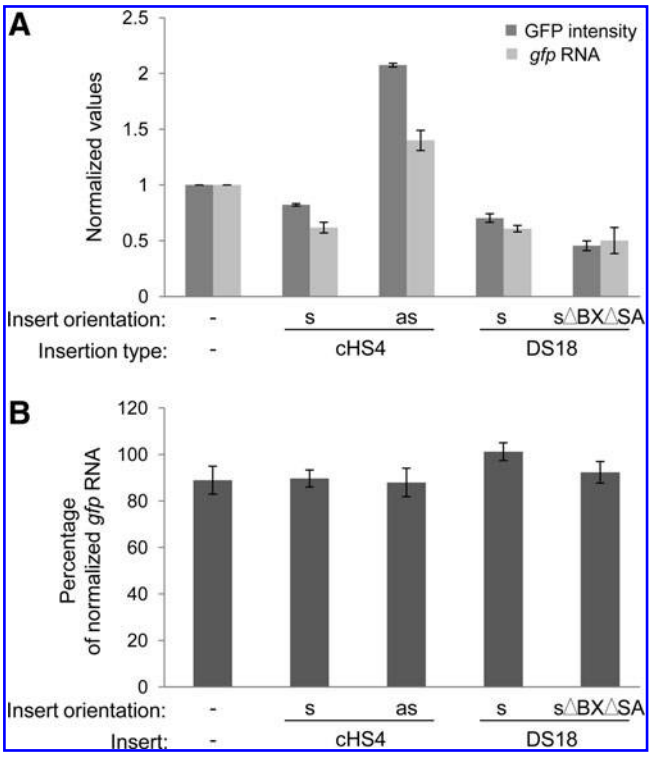


FIG. 7. Impact of insertions within the SIN-LTRs on mRNA levels and stability. Human T lymphoblastoid CEM-T4 cells were transduced with a GFP-encoding, P_{PGK}-based vector that contained within its SIN-LTRs no insert or the indicated sequence in the sense (s) or antisense (as) orientation. Two weeks later, GFP^+ cells (containing 1 ± 0.1 provirus per cell) were sorted and maintained in culture for 7 weeks, after which GFP protein expression was analyzed by cytofluorometry and *gfp* mRNA levels were quantified (before and after a 6-hr-long treatment with actinomycin D) by RT-qPCR. **(A)** Dark gray columns report the intensity of GFP expression from insert-containing vectors on normalization to that of the corresponding parental vector. Light gray columns illustrate the levels of *gfp* mRNA in cells transduced with insert-containing vectors (normalized to the *gapdh* mRNA content as well as to the *gfp* mRNA levels recorded in cells transduced with the parental vector). **(B)** Columns depict the levels of *gfp* mRNA after actinomycin D treatment normalized to the *gapdh* mRNA content as well as to the levels of *gfp* mRNA recorded before the administration of actinomycin D thus illustrating RNA stability. Data represent mean values from two independent experiments performed in duplicate ± SEM. Significant differences were not observed ($p > 0.05$).

This said, the influence of cHS4 on retroviral titers was marginal, in line with results from previous studies pointing to negligible (Yannaki *et al.*, 2002) or even positive (Emery *et al.*, 2000) effects of cHS4.

The vectors used in this study did not exhibit significant integration defects (Fig. 2), yet all inserts, in at least one orientation, significantly decreased early GFP expression in LV-transduced cells (Fig. 3). cHS4 cloned within SIN-LTRs has previously been shown to decrease transgene expression from internal promoters (Yannaki *et al.*, 2002; Ramezani *et al.*, 2003; Hanawa *et al.*, 2009; Uchida *et al.*, 2011), but this phenomenon was ascribed to the ability of cHS4 to block the 3′ LTR enhancer (Yannaki *et al.*, 2002) and/or to topological constraints (Yannaki *et al.*, 2002; Ramezani *et al.*, 2003). Our findings indicate that this effect is not specifically linked to the capacity of cHS4 to recruit CTCF, as it persists when sequences that do not bind CTCF are cloned within SIN-LTRs (Fig. 3). Aker and colleagues previously reported that the insertion of putatively inert "spacer" sequences of exogenous origin slightly decreases transgene expression (Aker *et al.*, 2007). We observed a similar phenomenon, because both DS18 and LacZ′, which were selected as control sequences (because they do not contain CTCF-binding sites), relevantly attenuated transgene expression. Altogether, these observations argue against the existence of totally "inert" sequences that can be used as spacer controls for insertion within retroviral vectors.

Previous studied indicate that the efficiency of transduction with vectors containing large inserts (of various nature) within the SIN-LTR decreases along with insert size (Hanawa *et al.*, 2009; Nielsen *et al.*, 2009; Urbinati *et al.*, 2009; Bos *et al.*, 2010). We also observed a tendency for the most pronounced attenuation of transgene expression to occur with large inserts (Fig. 3). However, further exploration with deletion derivatives of DS18 did not confirm this trend (Fig. 6). Moreover, insert orientation as well as the deletion of some sequences (e.g., MIRb in sense DS18) significantly influenced transgene expression attenuation, at least in some cases (Figs. 3 and 6). Taken together, our results indicate that insert size is not the only element influencing transgene expression from LV vectors.

Several phenomena might explain the attenuation of transgene expression resulting from the insertion of sequences within SIN-LTRs: (1) the recruitment or exclusion of cellular factors in the vicinity of insertions, which may influence the activity of the internal promoter; (2) changes in LTR DNA topology, which may prevent transcription factors from gaining access to the internal promoter; and (3) changes in RNA stability and/or (4) changes in RNA topology, which may affect transcription/translation. The attenuation of transgene expression that we observed was sequence and cell type independent, *de facto* excluding the first of these possibilities. The three-dimensional organization of the proviral DNA might be disturbed by insertions, and features favoring transcription, such as loops that bring the SIN-LTRs close to each other, might be destabilized. The LV provirus reportedly forms a transcription-dependent gene loop structure between the 5′ LTR or P_{CMV} and the 3′ LTR poly(A) signal, suggesting that pre-mRNA processing signals may indeed affect gene transcription (Perkins *et al.*, 2008). We failed to detect significant changes in RNA stability (Fig. 7B), suggesting that the attenuation of transgene expression that we observed derives from topological rearrangements at the DNA and/or the RNA level. The U3 region of the LTR is included in the 3′ UTR of the mRNA, and hence is likely to affect folding in an orientation-dependent manner, in particular when inserts are large. We propose that these RNA-folding modifications directly modify the level of transgene transcription from the retroviral loop structure. The impact of distinct inserts within the LV-P_{CMV} backbone even in the absence of integration favors the hypothesis of a direct action through topological rearrangements (Fig. 3A). Further studies will be needed to better elucidate this issue and possibly identify

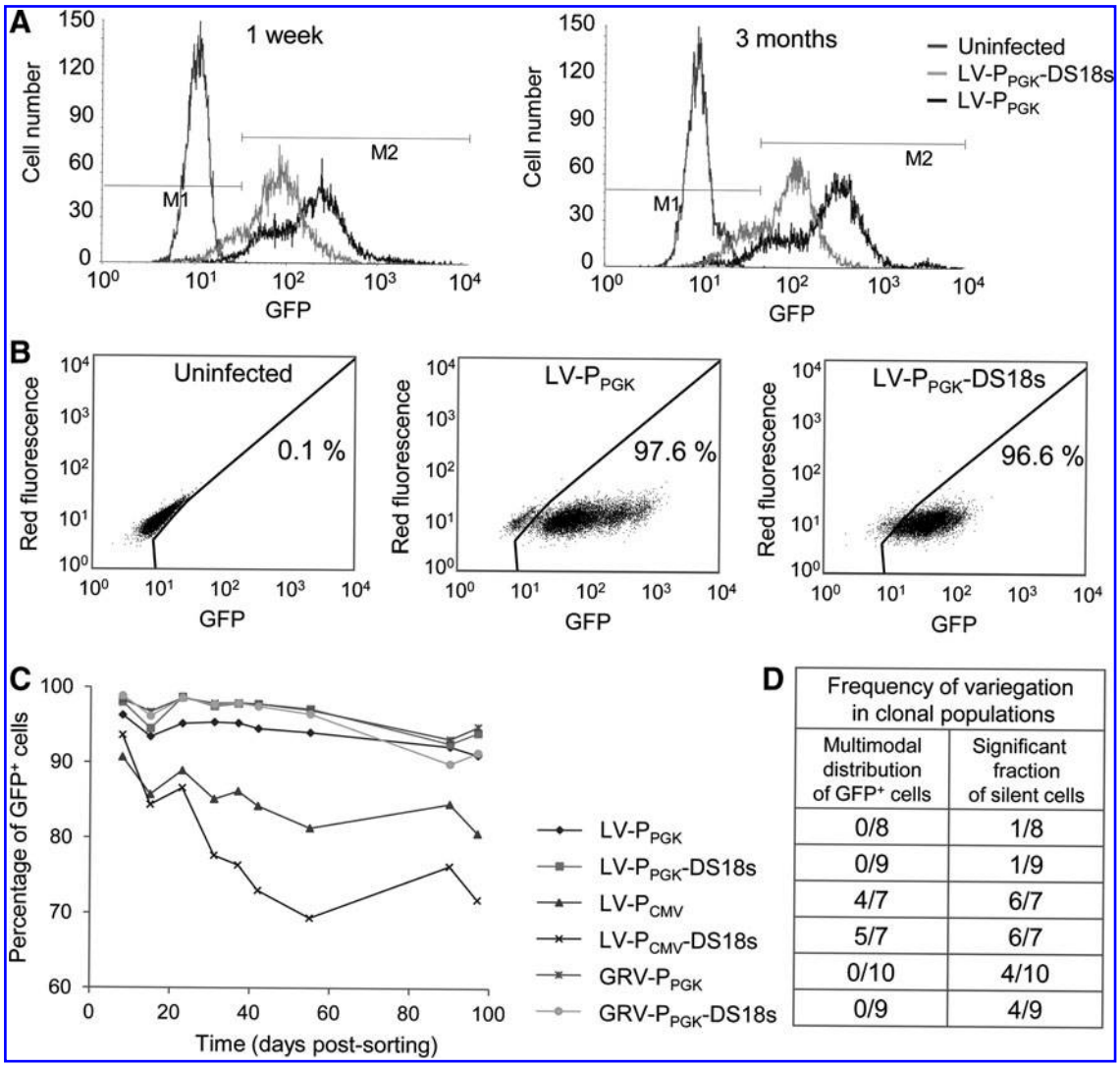


FIG. 8. Reduction of transgene expression by the insertion within SIN-LTRs of sense DS18 is stable and does not involve extinction or variegation. **(A** and **B)** GFP expression from DS18-containing P_{PGK}-based vectors is stable. Human cervical carcinoma HeLa cells were left uninfected or were transduced with <0.2 MOI of a GFP-encoding, P_{PGK}-based vector that contained within its SIN-LTR no insert or full-length DS18 in the sense (s) orientation. One week later, cells were sorted by FACS and GFP^+ cells were maintained in culture for an additional 3 months. At the indicated time points, GFP expression was evaluated by cytofluorometry. **(A)** GFP expression profiles (M1 and M2 define GFP-negative and -positive cell populations, respectively). **(B)** GFP expression is plotted as a function of red autofluorescence and percentages indicate the incidence of long-term GFP positivity. Results are representative of two independent experiments. **(C)** Changes in the long-term percentage of GFP^+ cells on transduction with vectors containing, or not containing, the DS18 element in their SIN-LTRs. Human cervical carcinoma HeLa cells were transduced with less than 0.2 MOI of lentiviral (LV) or gammaretroviral (GRV) GFP-encoding, P_{CMV}- or P_{PGK}-based vectors containing, or not containing, within their SIN-LTRs the DS18 element in the sense orientation. One week later, cells were sorted by FACS and GFP^+ cells were maintained in culture for an additional 3 months. The percentage of GFP-expressing cells is reported as a function of time. Results are from one representative experiment in which the sorted populations of LV-P_{PGK}-, LV-P_{CMV}-, and GRV-P_{PGK}-based vectors (containing, or not containing, sense DS18) had, respectively, 1 ± 0.3 vector copy number per cell (determined by qPCR experiments performed at various times after cell sorting). **(D)** Frequency of variegation. Seven to 10 clones were sorted from HeLa cells transduced as in **(C)** and maintained in culture for 6 weeks. The number of clones displaying a variegated GFP expression pattern (clear multimodal distribution of GFP expression or significantly higher fraction of completely silenced cells) among the total number of clones analyzed for each vector is reported. Detailed profiles of GFP expression for each clone are reported in Supplementary Fig. 7.

which topological elements are optimal for transgene transcription. Hanawa and colleagues detected increased amounts of polyadenylated *gfp* mRNA resulting from the insertion within SIN-LTRs of the 0.25-kb core insulator (in one given orientation) (Hanawa *et al.*, 2009). In line with this, we observed increased *gfp* mRNA levels in cells transduced with P_{PGK}-based vectors containing within SIN-LTRs cHS4 in the antisense (but not the sense) orientation (Fig. 7A). Taken together, these observations suggest that multiple elements influence transgene expression from a given vector, including

(but perhaps not limited to) vector backbone and size, insert size and orientation, internal promoter, and the transgene itself. In our study, cHS4 increased transgene expression only when cloned within SIN-LTRs in the antisense configuration (Figs. 3, 4, 5, and 7A). The importance of insert orientation for the modulation of transgene expression from retroviral vectors has been previously reported (Emery *et al.*, 2000; Yannaki *et al.*, 2002; Hanawa *et al.*, 2009; Uchida *et al.*, 2011). In this context, attenuation may minimize the positive effects of some SIN-LTR inserts, and this may explain partially contradictory results about the efficiency of cHS4 as an insulating element in retroviral vectors (Nielsen *et al.*, 2009; Uchida *et al.*, 2011). Here, we demonstrate that SIN-LTR inserts can decrease transgene expression, a phenomenon that sometimes can be overcome by modifying insert orientation, thus highlighting the importance of careful vector design for gene therapy.

Our results reveal that cHS4 in the antisense orientation can improve transgene expression, although in a promoter-dependent fashion, from unintegrated DNA forms (Fig. 3A), bringing new perspectives to approaches of IDLV-based gene transfer. Conversely, increased transgene expression from P_{PGK}-based constructs bearing antisense cHS4 insert appeared over time, on integrated DNA forms (Figs. 4B and 5). These observations emphasize the fact that cHS4 might act through several mechanisms (including termination improvement, enhancer-blocker and barrier effects).

We observed no increased transgene expression from vectors including the Ins023′ sequence, which contains a CTCF-binding site associated with blocker-enhancer activity three times stronger than that of cHS4 (Xi *et al.*, 2007). We cannot rule out the possibility that Ins023′ is not functional in the retroviral context. As an alternative, the insertion-dependent attenuation of transgene expression might mask any enhancer activity by Ins023′, or the binding of CTCF might not be sufficient to ensure increased transgene expression (Yao *et al.*, 2003; Arumugam *et al.*, 2009).

Is attenuation of transgene expression advantageous for gene transfer? The level of transgene expression influences gene transfer efficiency and safety. Most gene therapy strategies that are currently under evaluation for the cure of inherited diseases are based on complementation approaches, in which the transfer of a functional copy of the mutant gene leads to phenotypic correction of the genetic defect. For such an approach to be successful, transgene expression should be stable over time and appropriate (in terms of expression levels and localization) for the physiological context (reviewed in Chang and Sadelain, 2007). Moreover, a balance must be found between the benefits and toxicity of gene expression, in particular relative to the possible recognition, neutralization, or destruction of transduced targets by the immune system (in immune-proficient hosts). Thus, depending on the disease, the expression of the therapeutic transgene ranges from low (e.g., X-linked severe combined immunodeficiency syndrome or adenosine deaminase deficiency), to moderate (e.g., Wiskott-Aldrich syndrome), to high (e.g., β-thalassemia and sickle cell disease, hemophilia) (Chang and Sadelain, 2007). In this context, or to establish well-defined cellular models, it would be advantageous to use vectors whose transgene expression levels can be finely modulated by SIN-LTR inserts, such as those presented in this study.

Acknowledgments

The authors thank Sylvain Thierry, Olivier Delelis, and Hervé Leh for stimulating discussions and Yann Lécluse for technical support. G. Manic and A. Maurin-Marlin held fellowships from the Association Française contre les Myopathies and the Ecole Normale Supérieure de Cachan, respectively.

Conflict of Interest Statement

The authors declare no conflict of interest.

References

Aker, M., Tubb, J., Groth, A.C., *et al.* (2007). Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. Hum. Gene Ther. 18, 333–343.

Arumugam, P.I., Urbinati, F., Velu, C.S., *et al.* (2009). The 3′ region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. PLoS One 4, e6995.

Banasik, M.B., and McCray, P.B., Jr. (2010). Integrase-defective lentiviral vectors: Progress and applications. Gene Ther. 17, 150–157.

Bos, T.J., De Bruyne, E., Van Lint, S., *et al.* (2010). Large double copy vectors are functional but show a size-dependent decline in transduction efficiency. J. Biotechnol. 150, 37–40.

Brussel, A., and Sonigo, P. (2003). Analysis of early human immunodeficiency virus type 1 DNA synthesis by use of a new sensitive assay for quantifying integrated provirus. J. Virol. 77, 10119–10124.

Buceta, M., Galbete, J.L., Kostic, C., *et al.* (2011). Use of human MAR elements to improve retroviral vector production. Gene Ther. 18, 7–13.

Cavazzana-Calvo, M., Payen, E., Negre, O., *et al.* (2010). Transfusion independence and HMGA2 activation after gene therapy of human β-thalassaemia. Nature 467, 318–322.

Chang, A.H., and Sadelain, M. (2007). The genetic engineering of hematopoietic stem cells: The rise of lentiviral vectors, the conundrum of the LTR, and the promise of lineage-restricted vectors. Mol. Ther. 15, 445–456.

Chen, C., and Okayama, H. (1987). High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. 7, 2745–2752.

Chung, J.H., Whiteley, M., and Felsenfeld, G. (1993). A 5′ element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell 74, 505–514.

Drelich, M., Wilhelm, R., and Mous, J. (1992). Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein *in vitro*. Virology 188, 459–468.

Ellis, J. (2005). Silencing and variegation of gammaretrovirus and lentivirus vectors. Hum. Gene Ther. 16, 1241–1246.

Emery, D.W., Yannaki, E., Tubb, J., and Stamatoyannopoulos, G. (2000). A chromatin insulator protects retrovirus vectors from

chromosomal position effects. Proc. Natl. Acad. Sci. U.S.A. 97, 9150–9155.

Evans-Galea, M.V., Wielgosz, M.M., Hanawa, H., *et al.* (2007). Suppression of clonal dominance in cultured human lymphoid cells by addition of the cHS4 insulator to a lentiviral vector. Mol. Ther. 15, 801–809.

Foley, G.E., Lazarus, H., Farber, S., *et al.* (1965). Continuous culture of human lymphoblasts from peripheral blood of a child with acute leukemia. Cancer 18, 522–529.

Gaussin, A., Modlich, U., Bauche, C., *et al.* (2012). CTF/NF1 transcription factors act as potent genetic insulators for integrating gene transfer vectors. Gene Ther. 19, 15–24.

Gerolami, R., Uch, R., Jordier, F., *et al.* (2000). Gene transfer to hepatocellular carcinoma: Transduction efficacy and transgene expression kinetics by using retroviral and lentiviral vectors. Cancer Gene Ther. 7, 1286–1292.

Haas, D.L., Case, S.S., Crooks, G.M., and Kohn, D.B. (2000). Critical factors influencing stable transduction of human $CD34^+$ cells with HIV-1-derived lentiviral vectors. Mol. Ther. 2, 71–80.

Hanawa, H., Yamamoto, M., Zhao, H., *et al.* (2009). Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken β-globin locus HS4 insulator element. Mol. Ther. 17, 667–674.

Hazuda, D.J., Felock, P., Witmer, M., *et al.* (2000). Inhibitors of strand transfer that prevent integration and inhibit HIV-1 replication in cells. Science 287, 646–650.

Hou, Y., Zhang, H., Miranda, L., and Lin, S. (2010). Serious overestimation in quantitative PCR by circular (supercoiled) plasmid standard: Microalgal PCNA as the model gene. PLoS One 5, e9545.

Kim, T.H., Abdullaev, Z.K., Smith, A.D., *et al.* (2007). Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245.

Kumar, M., Keller, B., Makalou, N., and Sutton, R.E. (2001). Systematic determination of the packaging limit of lentiviral vectors. Hum. Gene Ther. 12, 1893–1905.

Leclerc, G.J., Leclerc, G.M., and Barredo, J.C. (2002). Real-time RT-PCR analysis of mRNA decay: Half-life of β-actin mRNA in human leukemia CCRF-CEM and Nalm-6 cell lines. Cancer Cell Int. 2, 1.

Maruggi, G., Porcellini, S., Facchini, G., *et al.* (2009). Transcriptional enhancers induce insertional gene deregulation independently from the vector type and design. Mol. Ther. 17, 851–856.

Masson, C., Bury-Moné, S., Guiot, E., Saez-Cirion, A., Schoëvaërt-Brossault, D., Brachet-Ducos, C., Delelis, O., Subra, F., Jeanson-Leh, L., Mouscadet, JF. (2007). Ku80 participates in the targeting of retroviral transgenes to the chromatin of CHO cells. J Virol. 81(15):7924–7932. Epub 2007 May 16. PubMed PMID: 17507472; PubMed Central PMCID: PMC1951289.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., *et al.* (2004). Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. PLoS Biol. 2, E234.

Miyoshi, H., Takahashi, M., Gage, F.H., and Verma, I.M. (1997). Stable and efficient gene transfer into the retina using an HIV-based lentiviral vector. Proc. Natl. Acad. Sci. U.S.A. 94, 10319–10323.

Montini, E., Cesana, D., Schmidt, M., *et al.* (2006). Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. Nat. Biotechnol. 24, 687–696.

Montini, E., Cesana, D., Schmidt, M., *et al.* (2009). The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. J. Clin. Invest. 119, 964–975.

Naldini, L., Blomer, U., Gallay, P., *et al.* (1996). *In vivo* gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science 272, 263–267.

Nielsen, T.T., Jakobsson, J., Rosenqvist, N., and Lundberg, C. (2009). Incorporating double copies of a chromatin insulator into lentiviral vectors results in less viral integrants. BMC Biotechnol. 9, 13.

Ory, D.S., Neugeboren, B.A., and Mulligan, R.C. (1996). A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. Proc. Natl. Acad. Sci. U.S.A. 93, 11400–11406.

Pauwels, K., Gijsbers, R., Toelen, J., *et al.* (2009). State-of-the-art lentiviral vectors for research use: Risk assessment and biosafety recommendations. Curr. Gene Ther. 9, 459–474.

Perkins, K.J., Lusic, M., Mitar, I., *et al.* (2008). Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. Mol. Cell 29, 56–68.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29, e45.

Ramezani, A., Hawley, T.S., and Hawley, R.G. (2003). Performance- and safety-enhanced lentiviral vectors containing the human interferon-β scaffold attachment region and the chicken β-globin insulator. Blood 101, 4717–4724.

Rivella, S., Callegari, J.A., May, C., *et al.* (2000). The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. J. Virol. 74, 4679–4687.

Ruprecht, R.M., O'Brien, L.G., Rossoni, L.D., and Nusinoff-Lehrman, S. (1986). Suppression of mouse viraemia and retroviral disease by 3′-azido-3′-deoxythymidine. Nature 323, 467–469.

Ryu, B.Y., Persons, D.A., Evans-Galea, M.V., *et al.* (2007). A chromatin insulator blocks interactions between globin regulatory elements and cellular promoters in erythroid cells. Blood Cells Mol. Dis. 39, 221–228.

Schambach, A., Swaney, W.P., and van der Loo, J.C. (2009). Design and production of retro- and lentiviral vectors for gene expression in hematopoietic cells. Methods Mol. Biol. 506, 191–205.

Smit, A.F., and Riggs, A.D. (1995). MIRs are classic, tRNA-derived SINEs that amplified before the mammalian radiation. Nucleic Acids Res. 23, 98–102.

Uchida, N., Washington, K.N., Lap, C.J., *et al.* (2011). Chicken HS4 insulators have minimal barrier function among progeny of human hematopoietic cells transduced with an HIV1-based lentiviral vector. Mol. Ther. 19, 133–139.

Urbinati, F., Arumugam, P., Higashimoto, T., *et al.* (2009). Mechanism of reduction in titers from lentivirus vectors carrying large inserts in the 3′LTR. Mol. Ther. 17, 1527–1536.

Xi, H., Shulha, H.P., Lin, J.M., *et al.* (2007). Identification and characterization of cell type-specific and ubiquitous chromatin regulatory structures in the human genome. PLoS Genet. 3, e136.

Yamashita, M., and Emerman, M. (2006). Retroviral infection of non-dividing cells: Old and new perspectives. Virology 344, 88–93.

Yannaki, E., Tubb, J., Aker, M., *et al.* (2002). Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. Mol. Ther. 5, 589–598.

Yao, S., Osborne, C.S., Bharadwaj, R.R., *et al.* (2003). Retrovirus silencer blocking by the cHS4 insulator is CTCF independent. Nucleic Acids Res. 31, 5317–5323.

Zhang, Z.D., Paccanaro, A., Fu, Y., *et al.* (2007). Statistical analysis of the genomic distribution and correlation of regulatory elements in the ENCODE regions. Genome Res. 17, 787–797.

Zheng, J., Shan, Y., Lambrecht, R.W., *et al.* (2008). Differential regulation of human ALAS1 mRNA and protein levels by heme and cobalt protoporphyrin. Mol. Cell. Biochem. 319, 153–161.

Zhou, S., Mody, D., Deravin, S.S., *et al.* (2010). A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate *LMO2* expression in human T cells. Blood 116, 900–908.

Zufferey, R., Dull, T., Mandel, R.J., *et al.* (1998). Self-inactivating lentivirus vector for safe and efficient *in vivo* gene delivery. J. Virol. 72, 9873–9880.

Zychlinski, D., Schambach, A., Modlich, U., *et al.* (2008). Physiological promoters reduce the genotoxic risk of integrating gene vectors. Mol. Ther. 16, 718–725.

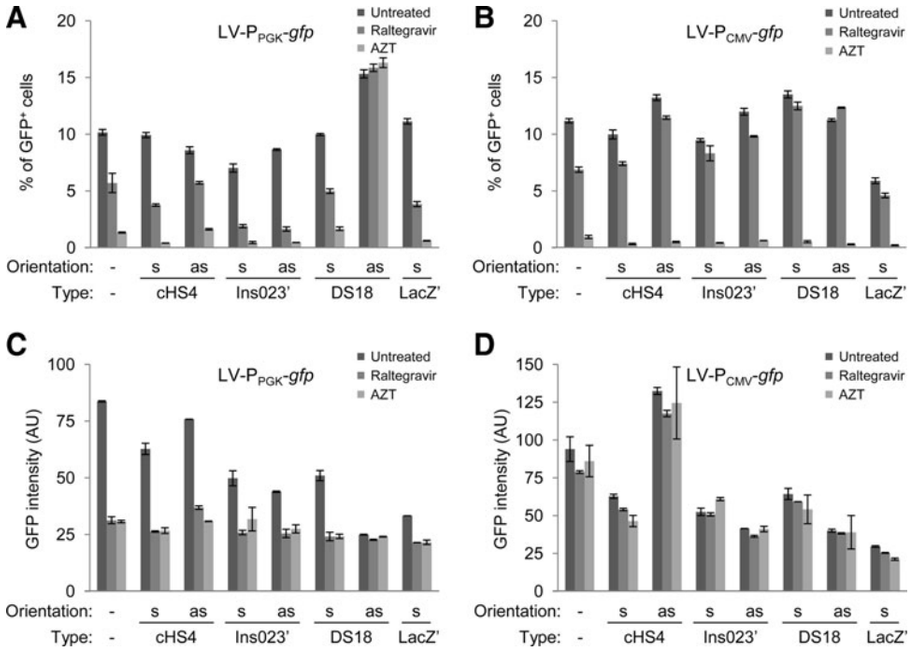
Address correspondence to:
Dr. Stéphanie Bury-Moné
LBPA, CNRS
Ecole Normale Supérieure (ENS) de Cachan
61 avenue du Président Wilson
94235 Cachan
France

E-mail: stephanie.bury-mone@lbpa.ens-cachan.fr

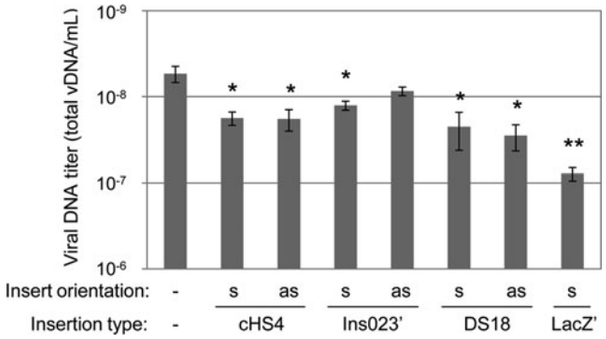
Received for publication August 21, 2011;
accepted after revision February 27, 2012.

Published online: March 2, 2012.

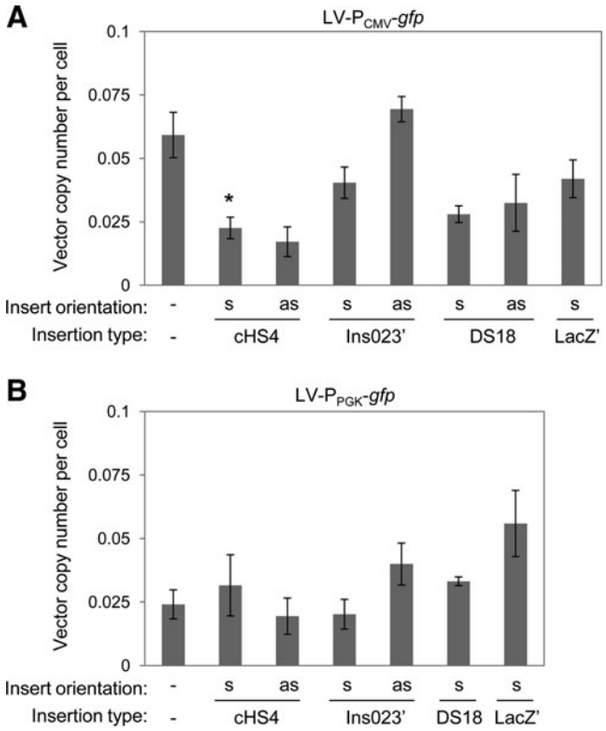
Supplementary Data



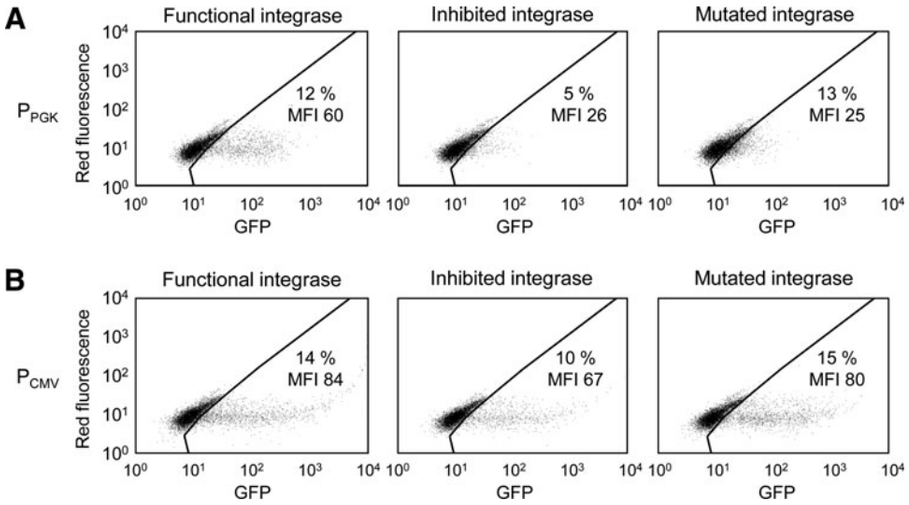
SUPPLEMENTARY FIG. S1. Cytofluorometric analysis of AZT- or raltegravir-treated cells transduced with parental vectors or its derivatives bearing insertions within SIN-LTRs. Untreated cells, AZT-treated cells, and raltegravir-treated HeLa cells were transduced with <0.4 MOI of GFP-encoding vectors based on the PGK promoter **(A** and **C)** or CMV promoter **(B** and **D)**, and containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. Forty-eight hours later, cytofluorometry was employed to determine the **(A** and **B)** percentage and **(C** and **D)** intensity of GFP expression (which is expressed in arbitrary units. Data are from representative experiments performed at least four times in duplicate (means ± SEM). "Orientation" and "type" indicate the orientation and type of the insert contained within the SIN-LTR of the transduced vector. AU, arbitrary units.



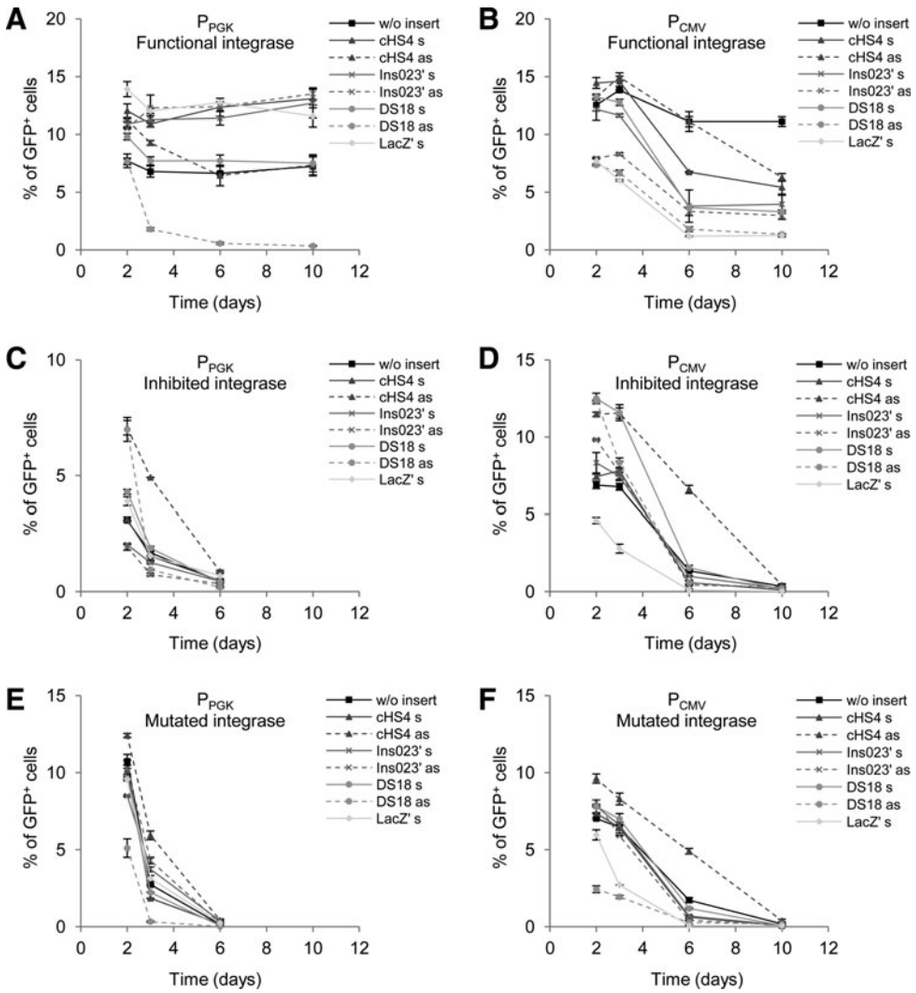
SUPPLEMENTARY FIG. S2. Viral DNA titers of P_{CMV}-based vectors. Human cervical carcinoma HeLa cells were transduced with serial dilutions of GFP-encoding, P_{CMV}-based vectors bearing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. Twenty-four hours later, total viral DNA was quantified by qPCR (on cell populations with an average GFP positivity of 10 ± 3%). Viral DNA titers were calculated by dividing the quantity of total DNA (values per cell × number of transduced cells) by the volume of viral preparation used in each experiment. Columns report mean values ± SEM (*$p<0.05$, **$p<0.01$ compared with parental vector) obtained from seven independent experiments. vDNA, viral DNA.



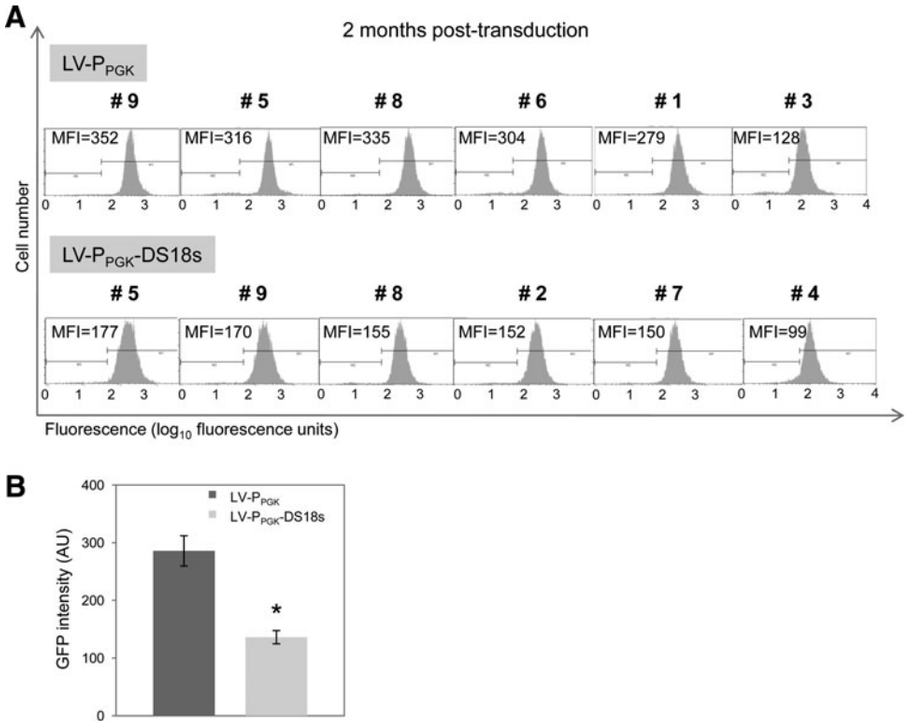
SUPPLEMENTARY FIG. S3. Vector copy number per cell transduced with parental vectors or its derivatives containing insertions within their SIN-LTRs. Human cervical carcinoma HeLa cells were transduced with serial dilutions of GFP-encoding vector based on the **(A)** PGK promoter or **(B)** CMV promoter and containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation. Seventy-two hours later, total viral DNA and cell number were quantified by qPCR (on cell populations with an average GFP positivity of $10 \pm 3\%$). Vector copy number per cell was calculated by reporting the total viral DNA to cell number. Columns report mean values $\pm$ SEM (*$p < 0.05$ compared with parental vector) obtained from three independent experiments. Of note, the rapid reduction in total viral DNA (0.04 ± 0.01 copy per cell, 72 hr posttransduction as compared with 0.7 ± 0.1 copy per cell 24 hr posttransduction) that we observed suggests that a fraction of GFP fluorescence detected 48 hr after transduction comes from either GFP synthesized from unintegrated DNA or from stable GFP protein/mRNA species synthesized at earlier time points.



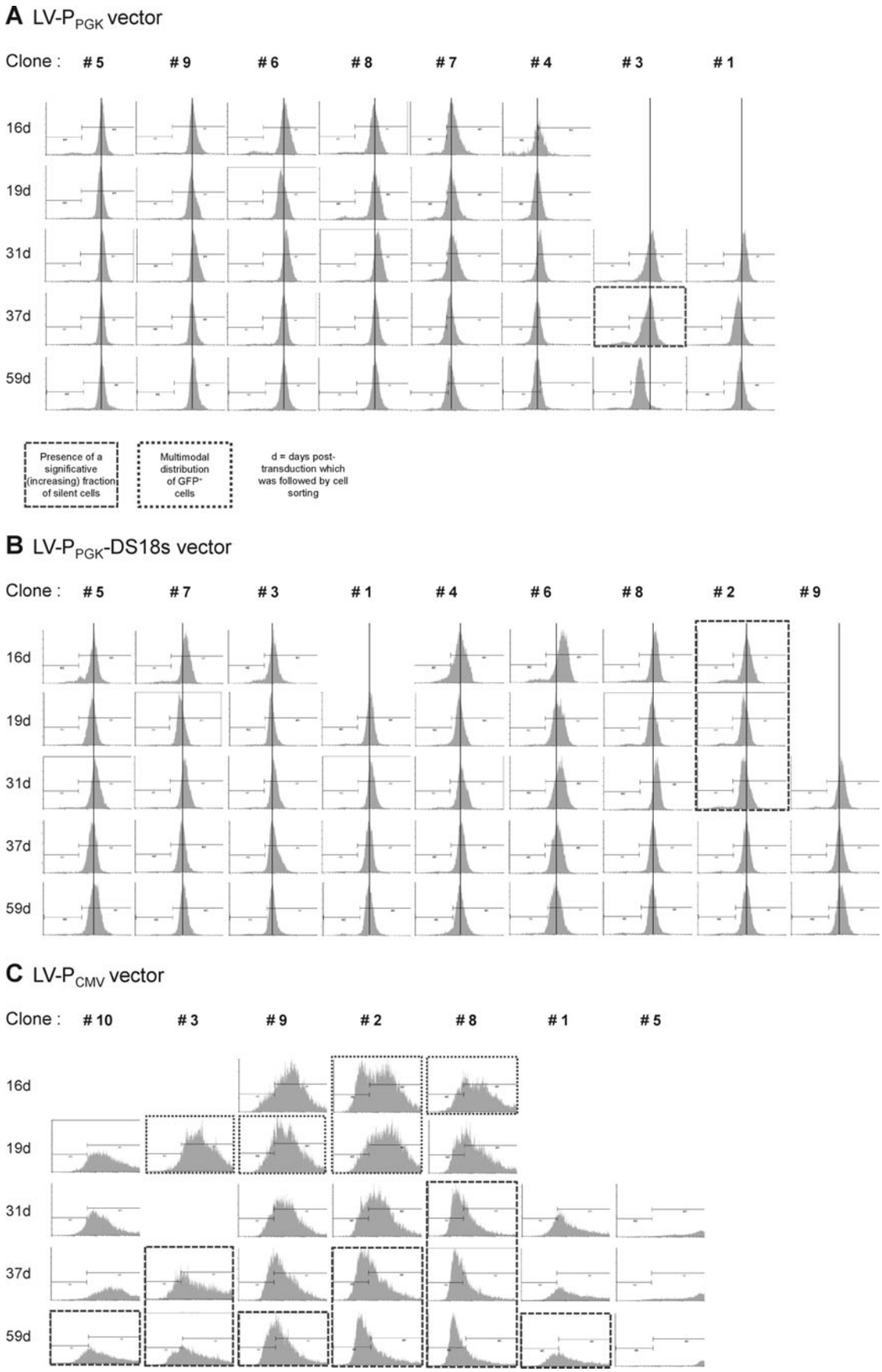
SUPPLEMENTARY FIG. S4. GFP expression profiles of cells transduced with integrase-proficient or -deficient vectors. Human cervical carcinoma HeLa cells were transduced with less than 0.2 MOI of GFP-encoding vectors based on the **(A)** PGK promoter or **(B)** CMV promoter. The vectors used carried a functional integrase that was inhibited by treatment of the cells with a specific integrase inhibitor (i.e., raltegravir) before transduction. Alternatively, these experiments were performed with vectors carrying a mutated integrase. Forty-eight hours later, cytofluorometry was employed to determine the percentage and geometric mean fluorescence intensity (MFI) of GFP expression. Profiles are representative of at least six independent experiments.



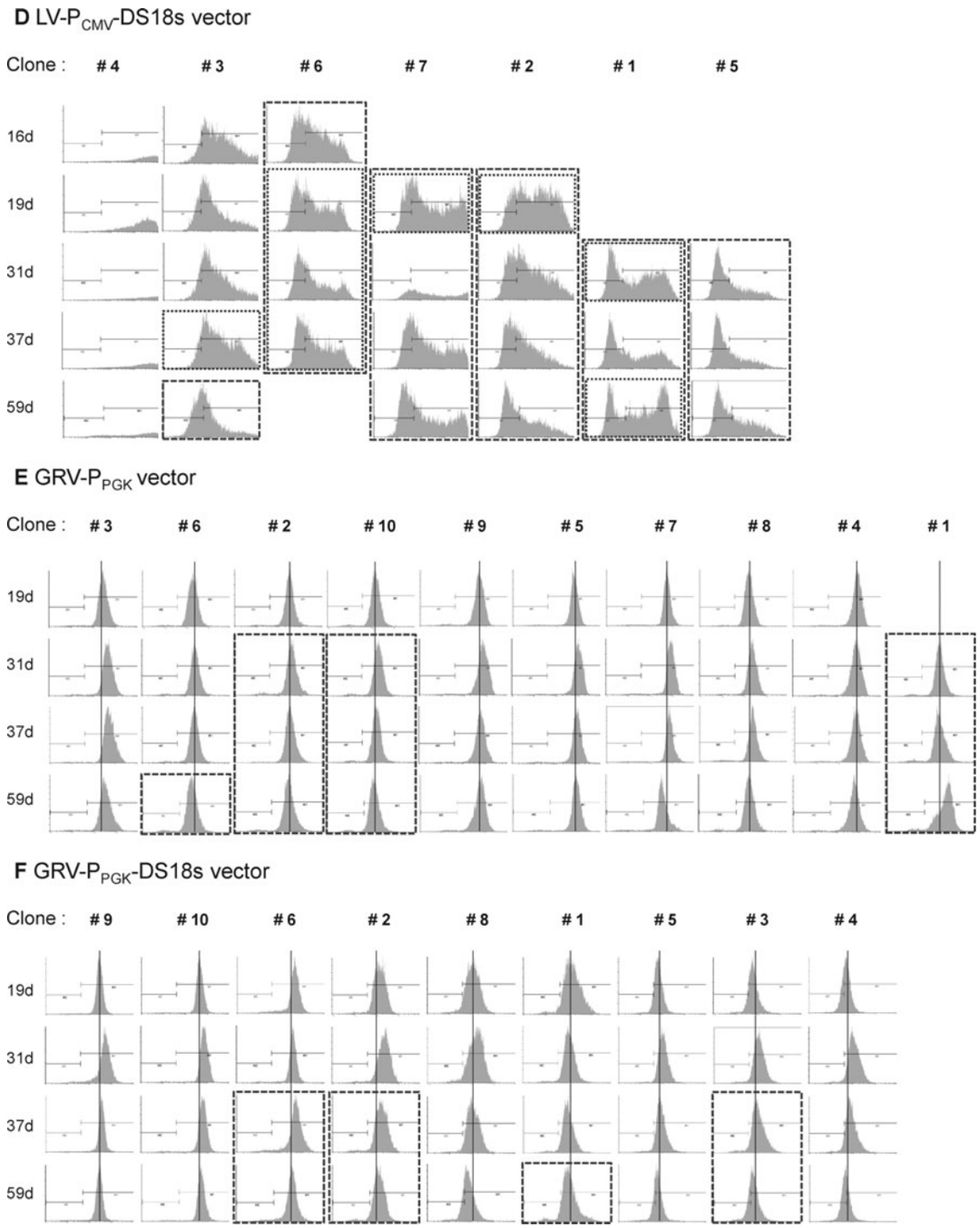
SUPPLEMENTARY FIG. S5. Kinetics of GFP expression from P_{CMV}- and P_{PGK}-based vectors bearing insertions within their SIN-LTRs. Human cervical carcinoma HeLa cells were transduced with <0.4 MOI of GFP-encoding vectors based on the **(A, C,** and **E)** PGK promoter or **(B, D,** and **F)** CMV promoter and containing within their SIN-LTRs the indicated inserts in the sense (s) or antisense (as) orientation, as described in Supplementary Fig. S4. GFP expression was assessed by cytofluorometry 2, 3, 6, and 10 days after transduction in **(A)**, **(B)**, **(D)**, and **(F)**, as described in Fig. 3. For conditions represented in **(C)** and **(E)**, cytofluorometric analysis were performed until 6 days, because GFP expression was lost at this time point. Kinetics are representative of at least six independent experiments.



SUPPLEMENTARY FIG. S6. GFP expression analysis of HeLa cell clones 2 months after transduction with a GFP-encoding, P_{PGK}-based vector containing within its SIN-LTR DS18 in the sense orientation. Human cervical carcinoma HeLa cells were transduced as detailed in Fig. 8C and 7 to 10 clones were sorted by cytofluorometry 1 week after transduction. Seven weeks later, these clones were characterized by cytofluorometry for GFP expression. **(A)** Representative GFP expression profiles. **(B)** Quantitative data are reported (mean GFP expression ± SEM; $*p < 0.05$ compared with the parental vector).



SUPPLEMENTARY FIG. S7. GFP expression profiles of HeLa cell clones transduced with lentiviral (LV) or gammaretroviral (GRV) GFP-encoding, P_{CMV}- or P_{PGK}-based vectors bearing within their SIN-LTRs the DS18 element in the sense orientation and maintained in culture for 1.5 months. Representative expression profiles at multiple time points are reported.



SUPPLEMENTARY FIG. S7. (Continued).

Supplementary Table S1. Vectors Used in Study

Type	*Promoter*	*Insertion*	*Insert size (kb)*	*Orientation*	*Name*
LV	PGK	—	—	—	LV-P_{PGK}
		cHS4	1.2	s	LV-P_{PGK}-cHS4 s
				as	LV-P_{PGK}-cHS4 as
		Ins023′	1.1	s	LV-P_{PGK}-Ins023′ s
				as	LV-P_{PGK}-Ins023′ as
		DS18	1.6	s	LV-P_{PGK}-DS18 s
				as	LV-P_{PGK}-DS18 as
		LacZ′	1.5	s	LV-P_{PGK}-LacZ′ s
LV	CMV	—	—	—	LV-P_{CMV}
		cHS4	1.2	s	LV-P_{CMV}-cHS4 s
				as	LV-P_{CMV}-cHS4 as
		Ins023′	1.1	s	LV-P_{CMV}-Ins023′ s
				as	LV-P_{CMV}-Ins023′ as
		DS18	1.6	s	LV-P_{CMV}-DS18 s
				as	LV-P_{CMV}-DS18 as
		LacZ′	1.5	s	LV-P_{CMV}-LacZ′ s
GRV	PGK	—	—	—	GRV-P_{PGK}
		DS18	1.6	s	GRV-P_{PGK}-DS18 s
				as	GRV-P_{PGK}-DS18 as

as, antisense; cHS4, chicken hypersensitive site-4; CMV, cytomegalovirus; GRV, gammaretroviral; LV, lentiviral; PGK, phosphoglycerate kinase; s, sense.

Supplementary Table S2. Primers Used for This Study

Name	*Target*	*Sequence*[a]
Primers used for vector cloning		
OBM29	DS18	GGATCCTTGGGGAATAATTTTACTCTTTTG
OBM30	DS18	GGATCCTGATGGTGATGAAATAAAGTTG
OBM95	LNCX 3′ region	CTCGAGCACGTGATGCATCA
OBM96	LNCX U3	TCTAGAGTTAACCTGTTCCATCTGTTCCTGAC
OBM156	IND116A-sense	GTACATACAGCCAATGGCAGCAATTTCACCAG
OBM157	IND116A-antisens	TGCCATTGGCTGTATGTACTGTTTTTACTGG
OBM164	3′LV SIN-LTR	CAACGATCTAGACAAGATCTGCTTTTTGC
OBM165	3′LV SIN-LTR	GATCTTGTCTAGATCGTTGGGAGTGAATTAGCC
OBM174	DS18	CACTGGTTCTCTATGATAGTATCATCTTCACGATCATGATCATCATCACCAGGGCC
OBM175	DS18	ATCATGATCGTGAAGATGATACTATCATAGAGAACCAGTGGCCCTCTGAG
OBM176	Ins023	CGGCGCGCCGGTACCTCTAGAAGAGATGTGCCTGGAACCACTCCTGATG
OBM178	Ins023	CGGCGCGCCGGTACCTCTAGAGCACATTGGGGAGCAAAATTGGAAAAAG
OBM179	*lacZ*	CGGCGCGCCTCTAGAGGATCCGAGCTCGGCGTTAACTCGGCGTTTCATCTGTGGTGC
OBM180	*lacZ*	CGGCGCGCCTCTAGAGGATCCGAGCTCGTTATCGCTATGACGGAACAGG
Primers used for quantitative PCR		
OBM12	*eGfp*	AGAACGGCATCAAGGTGAAC
OBM13	*eGfp*	ACTGGGTGCTCAGGTAGTGG
OBM241	Human *gapdh*	CGCTCTCTGCTCCTCCTGTT
OBM242	Human *gapdh*	CCATGGTGTCTGAGCGATGT
Alu1	Human *Alu*	TCCCAGCTACTGGGGAGGCTGAGG
Alu2	Human *Alu*	GCCTCCCAAAGTGCTGGGATTACAG
L-M667	LV	ATGCCACGTAAGCGAAACTCTGGCTAACTAGGGAACCCACTG
Lambda-T	L-M667	ATGCCACGTAAGCGAAACT
AA55M	LV	GCTAGAGATTTTCCACACTGACTAA
MH 531	LV	TGTGTGCCCGTCTGTTGTGT
MH 532	LV	GAGTCCTGCGTCGAGAGA**G**C
Probes used for quantitative PCR		
LTR-FL	LV	CACAACAGACGGGCACACACTACTTGA (3′ fluorescein)
LTR-FC	LV	CACTCAAGGCAAGCTTTATTGAGGC (5′ LC red 640 dye and 3′-P)
MH-FL	LV	CCCTCAGACCCTTTTAGTCAGTGTGGAA (3′ fluorescein)
MH-LC	LV	TCTCTAGCAGTGGCGCCCGAACAG (5′ LC red 640 dye and 3′-P)

[a]Restriction sites are underlined.

II/ IMPACTS DE LA MACHINERIE DE REPARATION DE L'ADN SUR L'EXPRESSION LENTIVIRALE

A/ CONTEXTE DE L'ETUDE

Les facteurs impliqués dans la réparation de l'ADN des cellules infectées ont été décrits comme des acteurs clés dans les interactions hôte/rétrovirus ainsi que dans la réponse à leur infection (Skalka and Katz, 2005; Smith and Daniel, 2006). Les signaux déclenchés lors des premières étapes de l'infection rétrovirale (*e.g.*, lésions créées lors l'intégration) peuvent stimuler les voies de réponses aux dommages de l'ADN en activant et/ou recrutant les kinases ATM et ATR (Lau et al., 2005; Roshal et al., 2003; Smith et al., 2008). Cette stimulation peut alors induire la mort (Andersen et al., 2005) ou la survie (Daniel et al., 1999; Lau et al., 2005; Smith et al., 2008) de la cellule infectée.

Les protéines des différentes voies de réparation de l'ADN interviennent directement dans plusieurs étapes du cycle rétroviral en stimulant ou inhibant la multiplication virale (*cf.*, Introduction, Tableau 8). Cependant, la littérature concernant certaines de ces protéines est contradictoire. Notre laboratoire a notamment mis en évidence un rôle négatif de Ku dans la régulation de l'expression précoce du VIH-1 au sein de cellules de rongeurs (Jeanson and Mouscadet, 2002; Masson et al., 2007) ainsi que dans la latence virale au sein des cellules humaines (Jeanson and Mouscadet, 2002). A l'inverse, Waninger *et al.* ont démontré que la déplétion en Ku80 diminue l'activation (précoce) du LTR du VIH-1 (Waninger et al., 2004).

Ku70 et Ku80 forment l'hétérodimère Ku qui recrute la DNA-PK et active le complexe du même nom lors de l'activation de la voie de réparation non-homologue (NHEJ) des cassures double-brins de l'ADN. De part ses capacités à fixer les acides nucléiques, Ku est important dans un certain nombre d'autres processus nucléaires, dont le maintien des télomères et la transcription (Gullo et al., 2006).

Récemment, deux analyses génomiques indépendantes par ARNi des gènes cellulaires nécessaires à la réplication du VIH-1 ont identifié Ku comme cible cellulaire qui affecte la multiplication du VIH-1 (Genovesio et al., 2011; Waninger et al., 2004), soulignant son importance dans le cycle du VIH-1. En accord avec ces données, Ku agirait à différentes étapes du cycle du VIH-1. Ku70 a été décrit interagir avec l'IN du VIH-1 (Studamire and Goff, 2008) au sein du PIC (Li et al., 2001; Zheng et al., 2011), ce qui pourrait influencer les étapes pré-intégratives. De plus, grâce à cette interaction, Ku70 pourrait être encapsidé au sein des particules virales (Zheng et al., 2011). Ku serait également impliqué dans la circularisation des espèces d'ADNv à 2-LTRs (Jeanson et al., 2002; Li et al., 2001; Masson et

al., 2007; Zheng et al., 2011). Ku a aussi été décrit participer au processus d'intégration du VIH-1 après transduction de cellules humaines (Jeanson et al., 2002; Waninger et al., 2004; Zheng et al., 2011). Enfin, l'action de Ku sur l'expression du VIH-1 pourrait être médiée par la fixation de la KBS identifiée dans le LTR du VIH-1 (Jeanson and Mouscadet, 2002) et/ou le ciblage/relocalisation du site d'intégration lentivirale en périphérie nucléaire au sein de cellules de rongeurs (Masson et al., 2007).

Afin de clarifier le rôle du complexe Ku dans le cycle du VIH-1, nous avons réalisé une étude de l'expression lentivirale au cours du temps à partir d'un set de vecteurs dérivés du VIH-1 non-réplicatifs transduisant le transgène *gfp* sous le contrôle du promoteur du VIH-1 (LTRs natifs) ou hétérologue (*i.e.*, d'origine virale, le CMV ou humaine, le PGK). Etant donné le caractère essentiel de Ku à la survie des cellules humaines (Wang et al., 2009), nous avons utilisé le modèle cellulaire humain HCT 116 issu d'un carcinome colorectal stablement haplodéplété en Ku80. Ce modèle a été obtenu par inactivation ciblée du locus *XRCC5* (Li et al., 2002) et est l'un des seuls modèles de déplétion en Ku à être caractérisé phénotypiquement (Ghosh et al., 2007; Li et al., 2002). Les cellules HCT 116 *Ku80*$^{+/-}$ ont notamment un niveau élevé en p53 (TP53) WT (*wild-type*) (comparés aux cellules WT) (Ghosh et al., 2007; Li et al., 2002), protéine clé dans la réponse aux dommages de l'ADN qui est non-fonctionnelle dans les autres lignées cellulaires utilisées pour l'étude du VIH-1. Ce modèle a aussi l'avantage de nous affranchir d'une approche ARNi qui peut être non-spécifique, abolie lors d'une infection par le VIH-1 (Bennasser et al., 2005; Bennasser et al., 2006; Qian et al., 2009) et/ou d'une sur-expression de sh/siRNA (Grimm et al., 2006).

Cette étude nous a permis de démontrer que le statut en Ku joue un rôle double en réponse à la transduction par le VIH-1 dans les cellules HCT 116, en stimulant spécifiquement la transcription précoce du LTR du VIH-1 et en limitant la quantité de cellules contenant un provirus latent. La déplétion en Ku n'a pas d'impact sur les étapes pré-intégratives du VIH-1 mais diminue l'expression précoce du VIH-1, en agissant au niveau transcriptionnel. De plus, cet effet peut être partiellement aboli lors de la déplétion en p53 des cellules HCT 116 *Ku80*$^{+/-}$ qui rétablit un niveau de Ku proche à celui des cellules WT. Enfin, nous avons mis en évidence que la déplétion en Ku augmente la proportion de cellules contenant un provirus peu ou pas exprimé ainsi que la (re-)activation de l'expression virale observée après traitement par la TSA ou au TNFα. Notre étude souligne ainsi les conséquences potentielles du ciblage de protéines cellulaires impliquées dans la régulation de l'expression du VIH-1 sur la constitution de cellules réservoirs contenant un provirus latent.

B/ ARTICLE : 'IMPACT OF THE KU COMPLEX ON HIV-1 EXPRESSION AND CONSEQUENCE OF ITS DEPLETION IN THE ESTABLISHMENT OF VIRAL LATENCY' (SOUMIS AU JOURNAL RETROVIROLOGY)

Impact of the Ku complex on HIV-1 expression and consequence of its depletion in the establishment of viral latency

Gwenola Manic[1,*], Aurélie Maurin-Marlin[1,*], Fanny Laurent[1], Ilio Vitale[2,3], Sylvain Thierry[1], Philippe Dessen[2,4], Michèle Vincendeau[5], Christine Leib-Mösch[5,6], Uriel Hazan[1], Jean-François Mouscadet[1,#], Stéphanie Bury-Moné[1,§]

[1]LBPA, CNRS, Ecole Normale Supérieure de Cachan, FR-94230 Cachan, France; [2]Institut Gustave Roussy, FR-94800 Villejuif, France; [3]INSERM, U848, FR-94800 Villejuif, France; [4]INSERM, U985, FR-94800 Villejuif, France; [5]Institute of Virology, Helmholtz Zentrum München, German Research Center for Environmental Health, DE-85764 Neuherberg, Germany; [6]Department of Hematology and Oncology, Mannheim Medical Center, University of Heidelberg, DE-68135 Mannheim, Germany. *These authors contributed equally to the work.

Abstract

Background : Ku, a cellular complex required for human cell survival and involved in double strand break (DSB) DNA repair and multiple other cellular processes, may modulate retroviral multiplication, although the precise mechanism through which it acts is still controversial. Recently, two global approaches identified Ku as a possible anti-human immunodeficiency virus type 1 (HIV-1) target in human cells. Here we investigated the role of Ku on HIV-1 replication cycle by analyzing the expression level of a panel of non-replicative lentiviral vectors expressing the *green fluorescent protein* (*gfp*) on human colorectal carcinoma HCT 116 cells, stably or transiently depleted from Ku.

Results : We found that the depletion of Ku did not affect the efficiency of pre-integrative steps but decreased the early HIV-1 expression by acting at the transcriptional level. This negative effect was specific for the HIV-1 promoter, required the obligatory step of viral DNA integration and was reversed by the partial restoration of Ku by means of p53 transient depletion. In presence of a normal level of Ku, HIV-1 expression was gradually lost over time, likely due to the counter-selection of HIV-1-expressing cells. On the contrary, the reactivation of transgene expression from HIV-1 by means of trichostatin A (TSA) or tumor necrosis factor α (TNFα) administration, was enhanced in condition of Ku depletion, suggesting a phenomenon of provirus latency.

Conclusions : These observations pleased to favor the hypothesis that Ku has an impact on HIV-1 expression at short- and middle-time after integration, as its depletion reduces the level of early transgene expression from HIV-1 promoter, decreases the counter-selection of HIV-1-expressing cells over the time and promotes the establishment of a viral latency.

Background

The signals triggered by retroviral infection have been shown to promote either the survival [1-4] or the death [5] of the infected cell, and this depending on their nature and intensity.

The components of the machinery responsible for DNA repair in host cells seem to play a critical role in the interactions with retroviruses, thus potentially influencing the outcome of viral infections (for a more detailed description, refer to [6, 7]). Indeed, the DNA damage response pathways are stimulated in the initial steps of retroviral infection - and in particular during virus integration in host genome - of multiple retrovirus [7, 8]. For instance, the activation and/or the recruitment to viral DNA (vDNA) of the kinases ataxia telangiectasia mutated (ATM) [3 , 4] and ATM-and Rad3-related (ATR) [9] have been observed following infection by the lentiviral (LV) human immunodeficiency virus type 1 (HIV-1). In addition, either avian sarcoma virus (ASV) or HIV-1 infections induced the generation of phosphorylated H2AX (γ-H2AX) foci in host cells [10, 11]. DNA repair response signaling in HIV-1-infected cells may be triggered by distinct signals, including viral proteins, (*e.g.*, Vpr [5]), viral genomes presented in a linear unintegrated form [2] and DNA lesions generated by the vDNA integration [1, 12].

So, it appears that the DNA repair machinery is necessary to complete viral integration [8], but may also constitute one among the first lines of defense against viral infection put in place by the cell [13].

Indeed, DNA repair proteins could also directly interfere with HIV-1 cycle, namely by influencing the stability of vDNA [*e.g.*, XPB/XPD [14]], the integration of the virus [*e.g.*, RAD51 [15]] and/or the expression of viral genes [*e.g.*, DNA-dependant protein kinase (DNA-PK) [16]]. As a consequence, retroviral replication in host cells may be either stimulated or inhibited. Nevertheless, for some among these proteins, the results present in literature are sometimes controversial. For instance, in two previous studies performed by our and another group, the Ku complex (also known as Ku) has been reported to affect proviral gene expression in a negative [17, 18] or a positive [19] fashion, respectively.

Ku is a heterodimer complex composed by two DNA-binding subunits, Ku70 and Ku80 (official gene names *xrcc6* and *xrcc5*, respectively). In presence of double strand-breaks (DSBs), Ku recruits the kinase subunit DNA-PKcs to generate the DNA-PK complex, which is involved in the non-homologous end-joining (NHEJ) pathway of DSB repair [20]. Ku also

contributes to additional nuclear processes, including the maintenance of telomere, VDJ recombination, and transcription (for a comprehensive description on Ku roles, see [21]).

Recently, two independent genome-wide screens assays identified Ku complex as a target against HIV-1 replication, thus underlining a potential role of this complex on HIV-1 cycle [19, 22]. Accordingly, several studies reported that Ku might act at different stages of the HIV-1 cycle. For instance, we and others previously demonstrated that Ku had a role in the formation of 2-LTR circular forms in human and rodent cells [2, 18, 23, 24]. Moreover, Ku might be involved in provirus integration in human cells [19, 23, 24]. However, in contrast with these findings, a dispensable function of Ku and/or the DNA PK complex in HIV-1 integration into human [25, 26], horse [26] and rodent [2, 18, 26, 27] cells have been also described. In addition, Ku was shown to interact directly with the HIV-1 integrase (IN), thus potentially influencing the pre-integrative steps [2, 24, 28]. Finally, we recently described that the Ku level may modulate provirus localization in chinese hamster ovary (CHO) cells [18]. Nevertheless, the precise impact of Ku on both HIV-1 integration and HIV-1 expression remains unclear and needs further examinations.

To gain insights on Ku functions on LV cycle, we performed a systematic analysis on viral expression from a collection of defective LV vectors transducing the *green fluorescent protein* (*gfp*) transgene under the control of viral or cellular promoters. As Ku is essential for human cell survival [29], the studies were carried out in *Ku80*$^{+/-}$ human colorectal carcinoma HCT 116 cells that have previously been subjected to targeted disruption of the *ku80* locus [30]. The results have been confirmed by parallel studies on HCT 116 cells transiently depleted from Ku70 or Ku80 through RNA interference approaches.

With this strategy, we demonstrated that Ku played a dual function in response to HIV-1 infection, namely to promote the early transcription from the HIV-1 promoter and to limit the constitution of viral latency. Ku depletion had no impact on the pre-integrative steps, but rather decreased HIV-1 expression by acting at the transcriptional level. Accordingly, p53 knockdown in Ku80-haplodepleted cells enhanced both Ku protein levels and HIV-1 expression. Finally, we showed that the depletion of Ku limited the counter-selection of HIV-1-expressing cells over the time and favored the establishment of cells containing latent viruses, which may be readily reactivated by the administration of the histone deacetylase inhibitor trichostatine A (TSA) or the nuclear factor κB (NF-κB) activator tumor necrosis factor α (TNFα).

Results

The level of Ku influences the efficiency of HIV-1 transduction

To investigate the effects of Ku on the HIV-1 replication cycle, we took advantage of heterozygous *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells [30]. These cells represent a valid model for Ku depletion for at least two reasons: (*i*) they are characterized by an expression of Ku80 (at both the mRNA and protein levels) that is half that of parental *Ku80*$^{+/+}$ (wild type, WT) cells ([30] and **Figure S1A**, **B**) and (*ii*) as the binding to Ku80 stabilizes Ku70, they also manifest a marked (approximately 50 %) depletion of Ku70 ([30] and **Figure S1B**).

We first evaluated the effect of Ku depletion on HIV-1 transduction efficiency. For this, WT and *Ku80*$^{+/-}$ HCT 116 cells were transduced with XCD3 [an *env*-defective HIV-1 construct in which *nef* was replaced by a *gfp* transgene under the control of the native HIV-1 LTR and an internal ribosome binding site (IRES) (**Figure 1A**)] followed by the cytofluorometry-mediated assessment of GFP expression 48 h later. When performing this analysis at a low multiplicity of infection (m.o.i. of 0.3), we observed that the percentage of GFP-positive (GFP^{+}) cells among *Ku80*$^{+/-}$ HCT 116 cells was approximately half that of their WT counterparts (~15 % *vs.* ~30 %) (**Figure 2A**, **B**). Moreover, as compared to *Ku80*$^{+/+}$ cells, transduced *Ku80*$^{+/-}$ HCT 116 cells displayed lower GFP expression levels, as monitored by the geometric mean fluorescence intensity (MFI) (**Figure 2C**, **D**). At high m.o.i., the percentage of GFP^{+} cells among the Ku80-haploinsufficient population was nearly equal to that observed among WT cells, perhaps due to a phenomenon of saturation of the number of cell expressing the transgene (**Figure 2A**, **B**). However, the difference in GFP expression levels was conserved (**Figure 2C**, **D**), indicating that Ku depletion affected transgene expression even at high m.o.i. XCD3 transduction had no significant effect on proliferation/viability in either WT or *Ku80*$^{+/-}$ HCT 116 cells, as evaluated by a colorimetric assay performed 48 h after transduction (data not shown). This finding indicates that the difference in GFP expression levels that we observed (**Figure 2**) cannot be ascribed to the loss of transduced cells.

To confirm these results, we performed additional experiments in which WT and *Ku80*$^{+/-}$ HCT 116 cells were transiently depleted of Ku by means of transfection with small interfering (si) RNAs directed against either Ku80 or Ku70 (**Figure 3A**). Seventy-two hours after transfection, cells were transduced with XCD3 for additional 48 h, and then analyzed by cytofluorometry for transgene expression. As shown in **Figure 3B**, the knockdown of Ku

significantly decreased HIV-1 expression levels in WT cells. Apparently, in *Ku80*$^{+/-}$ HCT 116 cells transgene expression was not altered by the siRNAs further depleting Ku (**Figure 3B**), suggesting that a ~50 % depletion of Ku was already sufficient to affect HIV-1 transcription.

Taken together, these observations demonstrate that either the prolonged (**Figure 2**) or the transient (**Figure 3**) depletion of Ku negatively affects GFP expression from the HIV-1 promoter.

Ku and p53 may cooperate for the modulation of HIV-1 expression

In accordance to what previously reported by [30, 31], we observed that the basal levels of p53 were higher in Ku80-haploinsufficient cells than in their WT counterparts (**Figure 3C**, **Figure S1C**). Intrigued by this finding, we thus decided to investigate the effects of a p53 specific siRNA (**Figure 3B**, **C**).

First, we analyzed the effect of p53 knockdown on the levels of the Ku complex, finding that p53 depletion in *Ku80*$^{+/-}$ cells was associated with increased amounts of Ku70 (not shown) and Ku80 (**Figure 3C**). Indeed, in this condition, Ku80 reached level close to that exhibited by WT HCT 116 cells (**Figure 3C**).

Strikingly, while the knockdown of p53 in WT HCT 116 cells was associated to a decreased in transgene expression from XCD3, p53 depletion in *Ku80*$^{+/-}$ HCT 116 cells increased HIV-1 expression to a level that was intermediate between those observed in WT and *Ku80*$^{+/-}$ HCT 116 cells (**Figure 3B**). Although we could not exclude a direct role for p53 on HIV-1 expression, we speculated that this finding might be due to the increase in Ku protein levels as observed in *Ku80*$^{+/-}$ cells depleted of p53. To confirm this hypothesis, we performed experiments in which we transiently or stably *trans*-complemented Ku in *Ku80*$^{+/-}$ HCT 116 cells by means of Ku80-expressing plasmids or viral vectors (either alone or combined with the fluorescent protein DsRed). Thereafter, the polyclonal cell population as well as ~30 DsRed^{+} cell clones isolated by cytofluorometry-assisted sorting were assessed for Ku protein levels. In addition, we took advantage of a commercially available Ku *trans*-complemented *Ku80*$^{+/-}$ HCT 116 clone (named pKu80 *Ku80*$^{+/-}$ clone). In none of these experiments, which we repeated at least 2 times yielding similar results, we observed increased Ku expression upon *trans*-complementation (data not shown and **Figure S2** for pKu80 *Ku80*$^{+/-}$ clone). This observation is in contrast with previous findings [31] and suggests that Ku is tightly controlled in *Ku80*$^{+/-}$ HCT 116 cells, perhaps by a mechanism directly

controlling protein levels or negatively affecting the survival/proliferation of cells expressing excessive amounts of Ku.

In conclusion, we provided experimental evidences on a process linking in a tight fashion the expression levels of Ku and p53 that may be relevant for HIV-1 expression.

Ku depletion negatively affects the HIV-1 expression *via* a transcriptional mechanism

For elucidating the impact of the Ku complex on the pre-integrative steps of the HIV-1 cycle, we quantified by quantitative PCR (Q-PCR) all vDNA species - namely, total vDNA, 2-LTR circular viral forms, and integrated proviral DNA - in human colon carcinoma (WT and *Ku80*$^{+/-}$ HCT 116) and in T lymphoid leukemia (Jurkat and CEM-T4) cells previously transduced with XCD3.

As determined in parallel cytofluorometric assessments, the early GFP expression was similar in WT HCT 116, Jurkat and CEM-T4 cells, being in all cases higher than in *Ku80*$^{+/-}$ HCT 116 cells (**Figure 4A**). In line with previous results [32], transgene expression rapidly decreased over time, in all analyzed cell lines (**Figure 4A**). As demonstrated by Q-PCR analyses, XCD3 efficiently integrated in the genome of both *Ku80*$^{+/+}$ and *Ku80*$^{+/-}$ HCT 116 cells. Indeed, ~50 % of total vDNA was integrated in the cellular genome 24 h post-transduction (**Figure 4B**), a percentage that is consistent with previous reports [33]. Similar levels of vDNA integration were also observed in both CEM-T4 and Jurkat cells (**Figure 4B**). In our hands, the decrease in integrated or total vDNA was more rapid in HCT 116 and CEM-T4 cells than in Jurkat cells (**Figure 4C**). Moreover, both the levels of integrated DNA and the proportion of vDNA present as 2-LTR circular forms at an early time (from 2 to 72 h) following XCD3 transduction were similar in *Ku80*$^{+/+}$ and *Ku80*$^{+/-}$ HCT 116 cells (**Figure 4C**, **D**). Intriguingly, the quantity of total and integrated vDNA per cell 10 days after transduction was higher in *Ku80*$^{+/-}$ HCT 116 cells than in their parental WT counterparts (**Figure 4B**).

These results demonstrate that the depletion of Ku affects neither the capacity of the vDNA to form 2-LTR circles nor that to integrate in the cellular genome.

Importantly, lower levels of GFP-encoding mRNAs were detected in *Ku80*$^{+/-}$ HCT 116 cells 2 days after transduction, than in their WT counterparts (**Figure S3**). Indeed, in two distinct experiments, the quantitative difference between GFP-encoding mRNAs detected in *Ku80*$^{+/+}$ *vs. Ku80*$^{+/-}$ correlated with that observed in GFP expression levels (**Figure S3**).

Altogether, these results indicate that Ku depletion has no impact on the HIV-1 integration step but rather affects HIV-1 expression at the transcriptional level.

HIV-1-driven expression is stimulated by Ku status

Given that Ku has been implicated in the transcriptional regulation of various genes (*e.g.*, *IL2* [34]), we wondered whether this complex also controls transcription from the HIV-1 promoter, which directs transgene expression.

For this, we conducted transduction experiments with LV vectors in which *gfp* was placed under the control of the cytomegalovirus immediate early (CMV) (*i.e.*, the HR'-CMV vector, **Figure 1B**) or the human phosphoglycerate kinase (PGK) internal promoter [*i.e.*, the self-inactivating (SIN)-PGK vector, **Figure 1C**]. As opposed to what observed for XCD3 (harboring the HIV-1 promoter, **Figure 1A**), the percentage of GFP^+ cells after transduction with SIN-PGK or HR'-CMV at an m.o.i. of 0.1 was comparable in WT and *Ku80*$^{+/-}$ HCT 116 cell lines (**Figure 5A**), and this independently from the m.o.i. used (data not shown). So, Ku appeared to have no significant impact on the GFP expression of SIN-PGK and HR'-CMV LV vectors, pointing to a specific role for Ku related to the HIV-1 promoter.

We then evaluated the impact of HIV-1 accessory proteins on this phenomenon. To this aim, we took advantage of XCD24, a vector containing the *gfp* under the control of HIV-1 promoter and deleted from all HIV-1 coding sequences (**Figure 1D**). Consistent with our observations on XCD3, GFP expression from XCD24 was reduced in Ku-depleted cells as compared to their WT counterparts, at different m.o.i. (**Figure 5A**). Similar results were obtained with an XCD3 variant lacking Vpr (**Figure S4A**, **B**), *de facto* excluding that HIV-1 accessory proteins are required for the effects of Ku on HIV-1 expression. Vpr is reported to induce the G_2/M arrest of cells infected by HIV-1 [9]. Nevertheless, the cytofluorometry-assisted cell cycle distribution analyses on cells stained with the DNA dye propidium iodure (PI) put in evidence no significant difference between the cell cycle profiles of WT and *Ku80*$^{+/-}$ cell or transduced with XCD3 or XCD3 lacking Vpr (data not shown). Of note, these cell lines displayed a similar cell cycle profile also in control conditions (**Figure S1D**).

To rule out the possible contribution of Tat proteins coming from packaging cells, we employed a XCD24 variant carrying the C37T mutation in the *trans*activation-responsive (TAR) element [XCD24 TAR(C37T), **Figure 1D**], which affects Tat responsiveness [35]. Both WT and *Ku80*$^{+/-}$ HCT 116 cells were less efficiently transduced by this vector, as compared to XCD3 (**Figure 5A**), owing to a defective virus production (for more details refer to [35]). Nonetheless, this decrease was more pronounced in *Ku80*$^{+/-}$ cells than in their WT counterparts, confirming that the effect of Ku on the expression of the HIV-1-based constructs is independent from Tat. Similar results were also recorded with the XCD24 $Sc^{-216-196}$ vector, carrying a scrambled sequence in place of a previously identified Ku-binding sequence

(KBS), extending from nucleotides -216 to -196 in the U3 LTR [17] (**Figure 1D**), suggesting that Ku influenced transgene expression without binding the putative KBS (**Figure 5A**).

Altogether, these findings indicate that the effect of Ku on HIV-1 expression is specific for the HIV-1 promoter, is abolished by the presence of an internal heterologous promoter (*e.g.*, PGK or CMV), is independent from Tat and does not require Ku to bind the HIV-1 $KBS^{-216-196}$ sequence.

Ku specifically modulates the expression from exogenous and integrated viral DNA

The re-expression of silent retroviral sequences in the human genome has been reported during oncogenesis as well as upon UVB-irradiation, two conditions that activate the DNA damage response [36, 37]. As HIV-1-driven transcription appeared to be modulated by the Ku level, we investigated whether Ku would exert a similar effect on the expression of human endogenous retroviruses (HERV). To this aim, we performed a comprehensive microarray-based analysis of the transcriptional activity of HERVs in WT and *Ku80*$^{+/-}$ HCT 116 cells (**Figure S5**, **Table S1**), as previously described [38]. Only 4 HERV taxons (HML-1, HML-3, HML-5 and HML-9) among 57 (24 groups of HERVs related to β-retrovirus, γ-retroviruses and spumaviruses) probed by RetroArrays exhibited a small but significant variation of expression in WT *vs.* Ku80-haplodeficient cells (non parametric p value < 0.01, intensity threshold fixed at ± 30 % of variation between the two cell lines, **Table S1**; for more details, see also Methods). However, subsequent RT-Q-PCR experiments performed for the HERV groups exhibiting the most robust difference (namely, HML-1 and HML-5) did not confirm this result, which is probably due to a limited sensibility of the RetroArrays and the fact that the primers used for RT-Q-PCR do not match exactly the same subset of HERV sequences as the capture probes spotted on the RetroArrays [38] (**Table S1**).

Irrespective to this discrepancy, our results suggest that the Ku level has a mild (if any) impact on HERV expression in HCT 116 cells, further confirming that Ku specifically influences HIV-1 expression.

To corroborate once more this issue, we assessed transgene expression in WT and *Ku80*$^{+/-}$ HCT 116 cancer cells upon transient transfection with pSIN-PGK, pHR'-CMV, pXCD3 or pXCD24 shuttle plasmids. We observed that the percentages of GFP^+ cells were similar between the two cell lines (data not shown), while the MFI of GFP were even higher in *Ku80*$^{+/-}$ than in WT HCT 116 cells (**Figure 5B**). This result is apparently at odds to what we observed with viral transduction (**Figure 2**) yet is consistent with similar findings obtained in a rodent model of Ku depletion [17, 18]. This finding suggests that Ku modulation of HIV-

1 expression needs viral integration. In line with this hypothesis, Ku had no effect on the basal levels of transgene expression from unintegrated vDNA forms in cells transduced with XCD3 harboring the IN D116A mutation [XCD3 IN(D116A)] (**Figure 1A**, **Figure 5C**).

In conclusion, the Ku complex appears to specifically promote the expression of the integrated form of the HIV-1 genome.

The reactivation of latent HIV-1 is favored in Ku80-haplodepleted cells

Finally, we monitored the role of Ku in the events that are associated with the establishment of the provirus latency and with the reactivation of latent proviruses.

To address this issue, transgene expression was evaluated in WT and *Ku80*$^{+/-}$ HCT 116 cells upon transduction with XCD3 followed by exposure to either the TNFα (which promotes HIV-1 LTR-driven transcription by activating the transcription factor NF-κB [39]), or the histone deacetylase inhibitor TSA (which stimulates transgene expression by promoting histone acetylation [40]) (**Figure 6A**). This approach represents a standard experimental procedure for monitoring viral latency and reactivation [32, 41].

As expected**,** the administration of TNFα or TSA to both WT and *Ku80*$^{+/-}$ HCT 116 cells transduced 9 days earlier with XCD3 induced a significant increase in GFP expression (**Figure 6A**, **B**). However, even in response to TNFα or TSA, the percentage of GFP^{+} cells and the GFP MFI were higher in *Ku80*$^{+/-}$ than in WT cells (**Figure 6B**). Similar results were observed with an alternative strategy, in which WT and *Ku80*$^{+/-}$ HCT 116 cells were transduced with XCD3 at low or middle m.o.i. (< 0.6), and treated one day later with TSA for additional 24 h (**Figure 6A**). Indeed, TSA increased GFP expression in WT and *Ku80*$^{+/-}$ HCT 116 cells, *de facto* cancelling the difference between these two cell lines in this respect (**Figure 6C**).

This observation suggests that a pool of TSA-responsive GFP^{-} cells is present at early and middle time upon XCD3 transduction in both cell lines and that this fraction of cells harboring a silent provirus in a reactivable state is higher after the depletion of Ku.

In line with this hypothesis, the Ku level altered neither the kinetic of transgene extinction nor the percentage of GFP reactivation in GFP-positive HCT 116 cells isolated 2 days after transduction and incubated or not for 24 h with TSA 2 or 4 weeks later (**Figure S6**).

To give further confirmation to our hypothesis, we repeated the same experiments by using the Tat-deficient vector XCD24, a condition that reportedly induces virus pseudo-latency [42]. Transgene expression from the XCD24 vector was subjected to a dramatic increase by TNFα and TSA in both WT and *Ku80*$^{+/-}$ HCT 116 cells (**Figure 6D**). However,

the depletion of Ku favored transgene re-expression from this vector. Indeed, the ratio between the percentage of GFP^+ cells in presence or in absence of TSA or TNFα 10 days after transduction with XCD24 was almost 1.5 fold higher in Ku-depleted than in WT cells, a value that was similar to that observed for XCD3 (**Figure 6E**). Of note, as opposed to what observed for XCD3, WT and *Ku80*$^{+/-}$ cells transduced with XCD24 (**Figure 6D**) or with the *Vpr*$^-$ XCD3 vector (**Figure S4A**, **C**) failed to exhibit or presented a limited decrease in transgene expression over time. Hence, it appears that the decrease in transgene expression from XCD3 may be due to the potential cytostatic and/or cytotoxic effects of HIV-1 accessory proteins [*e.g.*, Vpr, as described in [5]], which may promote the counter-selection/elimination of cells expressing the integrated virus. As expected, the administration of reactivating agents did not modify transgene expression from SIN-PGK in both WT and Ku80-haplodeficient cells (**Figure 6D, E**). Of note, the percentage of GFP expression from this latter vector was stable over the time (**Figure 6D**).

Altogether, these results suggest that the depletion of Ku, which has a negative impact on early HIV-1-driven transgene expression, may favor the establishment of a pool of cells bearing an integrated, latent and readily re-activable provirus.

Discussion

Ku depletion decreases HIV-1 expression

Here, we investigated the impact of Ku on LV cycle by transducing distinct non-replicative LV vectors encoding for the *gfp* transgene. We put in evidence a two-fold reduction of early (48/72 h) HIV-1 expression in human Ku80-haplodeficent HCT 116 colon carcinoma cells as compared to their Ku80-proficient counterparts (**Figure 2**). Similar results were obtained on WT HCT 116 cells transfected with Ku70- or Ku80-depleting siRNAs (**Figure 3B**). Although the positive effect of Ku on GFP expression from HIV-1 required the obligatory step of vDNA integration in host genome (**Figure 5C**), Ku had no impact in HIV-1 pre-integrative stages. This finding was demonstrated by comparing the capacity to form 2-LTR circles and the integration efficiency of either WT *vs.* Ku80-haplodepleted HCT 116 (**Figure 4B**, **D**) cells or WT HCT 116 cells transfected with an unrelated *vs.* a Ku80- or a Ku70-directed siRNAs (data not shown). Accordingly, GFP expression from HIV-1-based vectors containing an internal promoter (*e.g.*, HR'-CMV and SIN-PGK) were as efficient in WT as in *Ku80*$^{+/-}$ HCT 116 cells (**Figure 5A**).

As detailed in the introduction section, the impact of Ku complex in provirus integration is debated. This controversy may be due to the intrinsic differences in the cellular

model, which may display a variable requirement of DNA damage signaling pathways during retroviral infection, as well as to the characteristic of the virus (*i.e.*, replicative *vs.* non-replicative form). Moreover, our findings seem to exclude, at least in this experimental model, a role not only for the endogenous but also for the exogenous Ku in pre-integrative steps of LV replication cycle. Indeed, the level of HIV-1 expression in HCT 116 cells was similar when these cells were transduction with HIV-1 vectors produced in human embryonic kidney (HEK) 293T cells transiently depleted or not from Ku (data not shown).

We found that, in *Ku80*$^{+/-}$ HCT 116 cells transduced with the HIV-1 vector XCD3, the decrease in transgene expression was correlated to a reduced level of *gfp* mRNA as compared to their WT counterparts (**Figure 2**, **Figure S2**), suggesting that Ku rather acts at a transcriptional level. Reportedly, Ku displayed both a positive and a negative impact on the transcription of multiple genes, and this through the direct interaction with their promoters [*e.g.*, [34, 43]]. Concerning HIV-1 expression, previous reports ascribed opposite effects to Ku. In one study, we showed that Ku negatively affect HIV-1 re-expression in human promonocytic CD4^{+} U1 cells pseudo-latently infected with a mutated form of HIV-1 [17]. We hypothesized that this impact could be due to the binding of Ku80 to a putative KBS element in the HIV-1 promoter that we identified [17]. In striking opposition to this observation, the silencing of Ku70 impaired HIV-1 replication in different human cell lines including cervical carcinoma HeLa cells and Jurkat cells [19, 22] Accordingly, Waninger and colleagues reported a negative effect of Ku80 depletion on HIV-1 expression on HeLa and CEM cells [19]. This latter impact was explained by a low level of *de novo* Tat synthesis in *Ku80*$^{+/-}$ cells, resulting in the decrease of Tat *trans*-activation of HIV-1 LTR. Finally, Ku has been reported to bind the TAR RNA sequence of HIV-1 [44].

Here, we demonstrated that Ku has a positive effect on HIV-1 short-term expression in HCT 116 cells that was (*i*) specific to the HIV-1 promoter (**Figure 5A**, **Figure S5**); (*ii*) independent from HIV-1 accessory proteins (**Figure 5A**, **Figure S4**); and (*iii*) not correlated to the binding to the putative KBS of HIV-1, as proven by the low transgene expression from XCD24 Sc$^{-216-196}$ in Ku80-haplodepleted cells as compared to their WT counterparts (**Figure 5A**). Importantly, as GFP expression from XCD24 TAR(C37T) was also affected by Ku depletion, we ruled out that the impact of Ku depended on Tat as it was previously suggested [19].

Although to date the exact contribution of Ku to HIV-1 expression is not precisely known, we propose two alternative and non-mutually exclusive hypotheses. Theoretically, Ku may promote the early transcription from the LTR by (contributing to the) targeting of the

virus to specific host chromatin region characterized by high level of transcription. Alternatively, as Ku80 interacts with multiple proteins that are positively implicated in HIV-1 transcription - including Sp1, RNA polymerase II, NF-κB, Ets-1 and DNA-PK [16, 45-48] - it may participate to the establishment of a local environment favorable for the early transcription from the LTR. The obligatory requirement of virus integration shown in this study may corroborate this hypothesis. Nevertheless, so far, distinct - and a more direct - role of Ku on HIV-1 expression could not be excluded. Further analyses are needed in order to elucidate this issue.

Cross-regulation between Ku and p53

We also put in evidence a cross-regulation between the expression levels of Ku and p53, a finding that prompt us for hypothesizing the existence of an unknown compensatory mechanism that increases the level of one protein when the other is low expressed (**Figure 3C**). These observations confirm the complex and mutual regulation between Ku80 and p53 anticipated by [31]. Previous studies reported that either human *Ku80*$^{+/-}$ or murine *Ku80*$^{-/-}$ cells display a high level of spontaneous DNA lesions and/or DNA damage response [30, 31, 49]. So, an appealing hypothesis would be that this mechanism could participate to limit the potential deleterious effets due to this endogenous DNA damages [30, 31].

Intriguingly, here, HIV-1 expression appeared to be affected by this Ku80/p53 interlinked control (**Figure 3B**). These preliminary results, which may suggest both a direct (*e.g.*, by acting on HIV-1-LTR) or indirect (*e.g.*, by modulating the expression of Ku) contribution of p53 in the regulation of HIV-1 transcription, need further investigations. Nevertheless, these findings point to the evidence that both DNA repair machinery and p53 level may influence HIV-1-driven expression, an observation that has to be taken into account for the choice of election of the cellular model used in the study of HIV-1 cycle. As exemples, p53 is mutated in CEM, Jurkat and U1 cell lines (database: http://www-p53.iarc.fr) and p53 functions are altered by the human papillomavirus E6 protein in HeLa cells [50] or by the human T leukemia virus type I (HTLV-I) Tax protein in C8166 cells [24]. So, given that the human HCT 116 cells express a wild-type p53, and that the complete absence of Ku is lethal in human, *Ku80*$^{+/-}$ HCT 116 cells represent a suitable model to evaluate Ku functions on HIV-1 cycle.

Ku depletion promotes viral latency

Finally, we demonstrated that Ku might promote viral latency. Indeed, the depletion of Ku negatively affected early transgene expression (**Figure 2**, **Figure 5**), did not influence the

kinetic of transgene extinction (**Figure S6**) but induced a marked reactivation of *gfp* expression in presence of the reactivating agents TNFα and TSA (**Figure 6**). This latter effect was observed either at early (2 days) or middle time (10 days) upon transduction, in agreements with our previous observations made in Ku80-depleted U1 cells, a model of virus pseudo-latency [17].

Driven by these observations, we propose the following model for describing the action of Ku in response to HIV-1 transduction (**Figure 7**). When normally expressed in target cells, Ku may favor the early HIV-1-driven transcription (2 days post-transduction) by direct and/or indirect interactions with the provirus (promoter). This positive effect is associated to the limitation of viral latency and is followed by the decrease in transgene expression occurring at middle time (10 days) post-transduction and mainly due to the counter-selection of transduced cells (**Figure 7A**). Conversely, in absence of Ku, the HIV-1-driven transcription is negatively affected at early time upon transduction, when a higher proportion of cells harbors silenced or poorly expressed proviruses (**Figure 7B**). However, these cells with an integrated provirus are less subjected to counter-selection over time and display an elevated capacity to be reactivated by TNFα or TSA administration (**Figures 7B**).

Conclusions

The present study demonstrates that the efficiency of HIV-1 transduction and expression in host human cells are affected by multiple parameters, including the design of vector, the genetic/phenotypic background of target cells and the time elapsed from cell transduction. In particular, among the DNA repair factors that have been proposed as promising targets for antiretroviral therapy [51], here, we found that Ku had roles to sustain HIV-1 expression at the transcriptional level after proviral integration and to prevent provirus latency. The precise mechanisms by which Ku acts are of great interest and need further investigations.

Methods

Plasmids and vectors. The genetic organization of the LV vectors is presented in **Figure 1**. RRLSIN.cPPT.hPGK-gfp (shortened to "SIN-PGK", with a supplementary "p" for the plasmid form) is a SIN-LV vector containing the *enhanced gfp* gene under the control of the human PGK internal promoter. SIN-PGK was modified from a previously described vector [52] (gift from the University of Turin Medical School, Italy) by deletion of the woodchuck hepatitis post-transcriptional regulatory element (WPRE). pXCD3 is derived from pNL4.3 containing the HIV-1 genome (GenBank: AF324493.1) into which the *Nhe*I-*Bsa*BI region of the *env* gene was deleted and the *Hpa*I-*Xho*I fragment (corresponding to the *nef* gene) was replaced by the compatible

IRES-humanized *Renilla* GFP (hrGFP) fragment from the pIRES-hrGFP 1a plasmid (Clontech Laboratories, Inc., Mountain View, CA, US). pXCD3 IN(D116A) is an integrase-defective version of pXCD3 encoding a mutated (D116A) integrase. This mutation was introduced into pXCD3 by directed PCR mutagenesis made by using the QuikChange XL Site-Directed Mutagenesis Kit (Agilent Technologies, Inc., Santa Clara, CA, US) according to manufacturer's indications and with OBM156/OBM157 primers (**Table S2**). HR'-CMV-gfp (shortened to "HR'-CMV") is a LV vector containing the *gfp* gene under the control of both the CMV and U3 LTR promoters [53]. pXCD24 was derived from pHR'-CMV by deleting the *Cla*I-*Bam*HI fragment corresponding to the CMV promoter. The pXCD24 TAR(C37T) construct was generated by replacing the original *Kpn*I-*Xba*I 3'LTR fragment of pXCD24 by a TAR(C37T) mutated version. This latter was obtained after insertion of the original *Kpn*I-*Xba*I 3'LTR fragment into the pGem-T-easy vector (Promega Corporation, Madison, WI, US) and directed mutagenesis (as described above) with the OBM200/OBM201 primers (**Table S2**). The pXCD24 $Sc^{-216-196}$ construct was obtained by replacing the putative KBS (CGGAGAGAGAAGTATTAGAGT) located between -216 and -196 nucleotides of the *Kpn*I-*Xba*I U3 3' LTR region of pXCD24 by a scrambled sequence (ACTCTCCTAGGTCTACTTGAC). This latter was generated after insertion of the *Kpn*I-*Xba*I 3'LTR region of pXCD24 into the pGem-T-easy vector (Promega) followed by fusion PCR. In this aim, two separate PCR were performed on pXCD24 by using the OBM101/OBM102 primers (**Table S2**) for the first PCR, and the OBM103/OBM104 primers (**Table S2**) for the second one following the manufacturer's recommendations for *Taq* Platinium High Fidelity amplification reactions (Invitrogen, Life Technologies Co., Grand Island, NY, US). Then, 1/50 of each PCR products were mixed to perform a third PCR without primers for the 10 first cycles (98°C for 10s/45°C for 10s/72°C for 30s), and with the OBM101/OBM104 primers (**Table S2**) for the 20 last amplification cycles (98°C for 10 s/60°C for 10 s/72°C for 30 s). The resulting PCR product was inserted into the pGem-T-easy vector (Promega) according to the manufacturer's recommendations, thereby generating the *Kpn*I-*Xba*I 3'LTR mutated fragment containing the scrambled sequence. The *Kpn*I-*Xba*I 3'LTR mutated fragment containing the scrambled sequence in place of the putative KBS was reintroduced into pXCD24 to replace the original *Kpn*I-*Xba*I region, thereby generating the pXCD24 $Sc^{-216-196}$ construct. All constructs were maintained and stored in the Sure2 *E. coli* host strain (Agilent Technologies).

LV vectors (except XCD3 and its derivatives) were produced with the psPAX2 or the pCMV-R8.74 packaging plasmid (kindly given by Didier Trono's laboratory, Lausanne, Switzerland - Addgene plasmids 12260 and 22036) encoding HIV-1 Gag-Pol, Tat and Rev proteins. All vectors were pseudotyped with the vesicular stomatitis virus envelope glycoprotein (VSV-G) encoded by pMD.G [54].

Cell culture and treatment. Wild type and *Ku80*$^{+/-}$ 70.32 human colorectal carcinoma epithelial HCT 116 cells (kindly provided by Dr E. A. Hendrickson) [30] were maintained in McCoy's 5A medium. *Ku80*$^{+/-}$ HCT 116 derivatives containing either a neomycin (Neo) or a hygromycin-Ku80 expression pcDNA3.1 plasmid (named A6 clone) (shortened as 'pCtl' and 'pKu80', respectively) were obtained from Horizon Discovery Ltd (Cambridge, UK) and maintained in McCoy's 5A medium (under a 10 µg/ml hygromycin selection for pKu80 *Ku80*$^{+/-}$ HCT 116 cells). HEK 293T cells were maintained in Dulbecco's modified Eagle medium. Human T $CD4^+$ lymphoid leukemia Jurkat cells [55] and human T $CD4^+$ lymphoblastoid leukemia CEM-T4 cells (a naturally isolated subclone of the CEM line which displays high levels of surface CD4 expression; obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, National Institute of Allergy

and Infectious Diseases, National Institutes of Health, from J.P. Jacobs) [56] were cultured in RPMI 1640. All media were supplemented with 10 % fetal bovine serum (FBS; PAA Laboratories GmbH, Pasching, Austria) and 1 % penicillin/streptomycin (p/s; 100 units/mL). All cell lines used in this study were incubated at 37°C, under an atmosphere containing 5 % CO_2. Unless otherwise indicated, media and supplements for cell culture were purchased from Gibco (Life technologies) and plastic ware was obtained from BD Falcon (BD Biosciences, Franklin Lakes, NJ, US).

Vector production. LV vectors were produced by calcium phosphate-mediated transfection of 293T cells as previously described [57, 58] with the following modifications. Briefly, cells were transfected by a mixture of plasmids constitued of 6 µg of shuttle vector plasmid, 2 µg of envelop plasmid (pMD.G) and 4 µg of packaging plasmid (psPAX2 or pCMV-R8.74, except for XCD3 production). After supernatant harvesting and ultracentrifugation, vector pellets were finally resuspended in PBS (Gibco) at a final concentration of 200 (as compared to the initial uncentrifuged supernatant), and aliquots were stored at -80°C.

Plasmid transfection and RNA interference. For plasmid transfection, WT and *Ku80*$^{+/-}$ HCT 116 cells were seeded in 12-well plates at a density of 10^5 cells per well the day before. Then, cells were transfected at 80 % confluence with 1 µg of shuttle vector plasmids (pSIN-PGK, pHR'-CMV, pXCD3, or pXCD24), in triplicate for each set of conditions, in 1.5 µl of Attracten® transfection reagent (QIAGEN GmbH, Hilden, Germany) dissolved in a final volume of 80 µl of free-McCoy's 5A as recommended by the manufacturer. After complexing time of 20 min at RT, transfection complexes were added to the cultures overnight. Transfection efficiency was determined after 48 h by cytofluorometric analysis, as described below.

Custom-designed siRNAs duplexes (**Table S2**) targeting Ku80, Ku70, p53 encoding mRNA or a sequence unrelated to human genome (Ctl) were purchased from Eurogentec (Eurogentec S.A., Seraing, Belgium). SiRNA transfection was achieved by plating HCT 116 cells in 12-well plates at a density of 10^5 cells per well (30–40 % confluence) at least 6 h before transfection using the HiPerFect® transfection reagent (QIAGEN) following the manufacter's recommendations. Briefly, 2.8 µl of HiPerFect® transfection reagent were added to 20 pmol of each siRNA dissolved in 100 µl of free-McCoy's 5A for a final concentration of 23.5 µM. After 10 min at RT, transfection complexes were added to the cultures overnight. After 72 h, transfection efficiency was determined by immunoblotting or transduced by vector, as described below.

Transduction procedure and analysis. The day before transduction, 10^5 cells per well were plated in triplicate for each set of conditions, in 12-well plates. Problems due to differences in plating and cell growth between cell lines were overcome by estimating the confluence of each cell line on the day of transduction and modifying the volume of the viral suspension added, in accordance with this estimation. Two days after transduction, the percentage and MFI of GFP^+ cells were determined by flow cytometry analysis as described below. The m.o.i. was calculated from the percentages of GFP^+ cells observed during the linear phase of transduction of WT HCT 116 cells. Alternatively, cell viability was measured by a MTT assay. For experiments exploring the reactivation of transgene expression in a mixed population of GFP^- and GFP^+ cells, cells were transduced and maintained in a 25 cm² flask. One week after transduction, 10^5 cells per well were plated in duplicate for each set of conditions, in 12-well plates. One day later, cells were treated for 24 h with 200 nM of TSA (Sigma-Aldrich Co., St Louis, MO, US) or 10 ng/ml of TNFα (BD Biosciences, San Jose, CA, US) before the analysis of transgene expression by flow cytometry. DMSO was used as a negative control for TSA treatment. For experiments exploring the reactivation of transgene expression in a pure population of transduced cells, GFP^+ cells were sorted two or three

days after transduction at low m.o.i. (< 0.1). Cells were then cultured for six weeks and subjected every week to a TSA treatment of 24 h before GFP expression analysis, as described above.

Cytofluorometric studies. To assess cell cycle distribution, cells were collected, fixed by gentle vortexing in ice-cold 80 % (v/v) ethanol and stained with 50 mg/ml propidium iodide (Sigma-Aldrich) in 0.1 % (w/v) D-glucose in PBS (Gibco) supplemented with 1 mg/ml (w/v) RNAse A (Sigma-Aldrich). The analysis of transfection or transduction (*i.e.*, percentage and MFI of GFP^+ cells) was estimated by flow cytometry on a FACSCalibur™ cytofluorometer (BD Biosciences). Statistical analysis was carried out by using the CellQuestTM software (BD Biosciences), upon gating on the events characterized by normal forward scatter and side scatter parameters.

Immunoblotting. Cells were washed with ice-cold PBS, resuspended in 100 µL of Mammalian Lysis Buffer from Qproteome Mammalian Protein Prep Kit (QIAGEN) and incubated on a rotary shaker for 5 min at 4°C. Protein concentration in cellular extracts was determined using a standard bicinchoninic acid assay (Thermo Fisher Scientific Inc., Waltham, MA, US). Thereafter, protein extracts (50 µg per lane) were separated according to their molecular weight on 12 % SDS-PAGE gels, followed by electrotransfer to nitrocellulose membranes (Amersham Biosciences, GE Healthcare, Piscataway, NJ, US) and immunoblotting with specific antibodies for Ku80, Ku70, p53 [all from Santa Cruz Biotechnology Inc., Santa Cruz, CA, US; references (ref.) H-300, N3H10, D0-1, respectively]. Primary antibody that specifically recognizes β-actin (from Santa Cruz Biotechnology Inc., ref.I-19) was used as loading control. Finally, membranes were incubated with appropriate secondary antibodies conjugated to the horseradish peroxidase (Santa Cruz Biotechnology), followed by chemiluminescence detection using the ECL Western Blotting Detection Reagents and photographic films (both from GE Healthcare).

Quantification of viral DNA species. Two hours 8 h, 24 h, 72 h and 240 h after transduction, cells were recollected, centrifuged at 270 g and washed once with PBS. Thereafter, genomic DNA (gDNA) samples were prepared and quantitative PCR (Q-PCR) were performed as previously described [58]. Primers and probes (**Table S2**) used for integrated vDNA nested PCR as well as LV total vDNA and 2-LTR amplification [carried out on LightCycler 1.5 or II 480 (Roche Diagnostics, Basel, Switzerland)] have been described in [33]. The cell equivalents in sample DNA as well as copy number of total vDNA and integrated vDNA were calculated as previously described by [58]. Alternatively, copy number of 2-LTR circular vDNA forms were determined with reference to10^2 to 10^6 copies of linearized pXCD3 plasmid or pGem-T-easy vector containing XCD3 2-LTR sequence.

RetroArrays experiments and analysis. Total RNA from 4.10^6 cells was extracted with the RNAeasy Mini kit (QIAGEN) and resuspended in 50 µl of diethylpyrocarbonate (DEPC)-treated water (Gibco). Genomic DNA contamination was removed by treating 30 µg of all mRNA samples with 2 U of Turbo DNase I (Ambion, from Life technologies) for 30 min at 37°C, according to the manufacturer's instructions. RNA was repurified on RNAeasy Mini kit columns (QIAGEN). The quantity and quality of the isolated RNA were determined by analysis on gel (1 % TBE 0.5X) and with a Nanovue spectrophotometer (GE HealthCare). The absence of genomic DNA was controlled by classic PCR on 50 ng of RNA with *Taq* Platinium enzyme (Invitrogen, from Life technologies) and mixed oligonucleotide primers (MOP) [38] with the following amplification program : 94°C 2 min ; 30X (94°C 30S, 50°C 2 min, 72°C 2 min); 72°C 7 min. Reverse transcription of mRNA, hybridization target synthesis and labeling MOP PCR as well as microarray preparation and hybridization and

post-processing of retrovirus-specific microarrays were performed as previously described [38]. Two independent PCR labeling and microarrays experiments, with positive (mix of RNA) and negative (H_2O) controls, were performed for each sample. RetroArray contains 57 HERV groups, each in triplicate. Images saved in the Bitmap file format were used to present the results by false color mapping. One result (representative of the triplicate) has been aligned for each microarray and each condition (**Figure S5**). Images saved in TIF file format were processed with ImaGene™ 8.0 software tool package (BioDiscovery Inc., Los Angeles, CA, US). Local background was subtracted, generating "Norm1" set of data. Data were then normalized to the total fluorescence of each microarray (data "Norm2"). Ratios of mean intensity of each locus in WT over *Ku80*$^{+/-}$ HCT 116 condition were calculated, as well as *p* values associated to a two-paired Student's *t* test. Loci presenting a ratio superior or equal to 1.3 or inferior or equal to 0.7, and a *p* value inferior to 0.01 are presented in **Table S1**. In a second step, the distribution of log-intensity microarray spots for each microarray was used to normalize hybridization efficiencies of microarrays, bringing them all to an equivalent distribution (data "Norm3"). Given the relatively low number of spots per microarray, the following analyses were performed on "Norm1" and "Norm3" data. A statistical non parametric analysis in BRB (Biometric Research Branch-ArrayTools, Dr. Richard Simon and BRB-ArrayTools Development Team) was conducted. The *p* value was set to the robust value of 0.01 in order to highlight loci differentially expressed. Candidate genes for differential expression between the conditions studied were selected from theses analyses (**Table S1**).

Quantitative RT-PCR. DNA-free RNA samples were prepared and RT-Q-PCR analyses were performed as previously described [58]. The *actin*, *glyceraldehyde-3-phosphate dehydrogenase* (*gapdh*), *hr-gfp*, or *xrcc5* genes or HERV loci were amplified by means of the corresponding primer pairs (**Table S2**), from 2 µl (between 10 and 20 ng) of complementary DNA in a final volume of 20 µl of LightCycler DNA Master SYBR Green I mix, according to the manufacturer's recommendations (Roche Diagnostics). For HERV loci expression analysis, the amplification was: 95°C 10 min; 40-45 x [95°C 1"; 50°C 5"; 72°C 10-12"]; 1 x [95°C 0"; 65°C 15"; 95°C 0" 0.1°C/sec). For other genes, the amplification program was the same with the following modifications: the hybridization temperature was 60°C, the elongation times were 6" (*gapdh*), 10" (*hr-gfp*, *actin*), or 15" (*xrcc5*). For Ku80 encoding mRNA quantification and HERV transcriptome analysis, fold-change values were calculated by comparative Ct analysis after normalizing for the quantity of *gapdh* or *actin* mRNA in the samples concerned [59]. For analysis of (HR-)GFP expression, copy numbers of *hr-gfp* and *gapdh* mRNAs were determined with reference to a standard curve obtained by amplification 10^2 to 10^7 copies of linearized cloned DNA with matching sequences, and used to calculate the normalized ratios.

Statistical procedures. Unless otherwise specified, experiments were performed in triplicate parallel instances and independently repeated at least twice. Data were analyzed with Microsoft Excel (Microsoft Co., Redmond, USA) and statistical significance was assessed by means of two-tailed Student's *t* tests. Unless otherwise indicated, asterisks depict statistical significance (*$p < 0.05$, ** $p < 0.01$).

Competing interests.

The authors declare no conflict of interest.

Author's contributions.

GM and AMM have carried out clonings; designed, performed, analyzed and interpreted Q-PCR/cellular experiments, and drafted the manuscript. MV and FL performed RNA extraction and RetroArrays analysis. IV

was involved in siRNA experiment design and cell cycle experiment analyses, and participated in writing the manuscript. PD performed the statistical analysis. ST contributed to analysis and interpretation of data on HIV-1 expression and cell counter-selection at middle term. CLM participated in the design of the study of endogenous retroviruses expression. UH revised the manuscript. JFM participated in the study design and helped to draft the manuscript. SBM conceived the study, participated in its design/coordination, performed Q-PCR analysis, and wrote the manuscript. All authors read and approved the final manuscript.

Acknowledgments.

We thank Hervé Leh, Frédéric Subra and Olivier Delelis for stimulating discussions and Yann Lécluse for technical support. We acknowledge Wolfgang Seifarth, from Medical Faculty Mannheim, for discussions about the microarray data. We specially thank Arzu Demir and Antonin Demange for their technical help and enthusiasm. The research leading to these results was funded by the *Agence Nationale de Recherches sur le SIDA et les Hépatites* (ANRS) and the ANR (grant PCV-0015I IMFOVIR). GM holds a fellowship from the *Association Française contre les Myopathies* (AFM). AMM holds a fellowship from the *Ecole Normale Supérieure* of Cachan.

References

1. Daniel R, Katz RA, Skalka AM: A role for DNA-PK in retroviral DNA integration. *Science* 1999, 284:644-647.
2. Li L, Olvera JM, Yoder KE, Mitchell RS, Butler SL, Lieber M, Martin SL, Bushman FD: Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *Embo J* 2001, 20:3272-3281.
3. Lau A, Swinbank KM, Ahmed PS, Taylor DL, Jackson SP, Smith GC, O'Connor MJ: Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase. *Nat Cell Biol* 2005, 7:493-500.
4. Smith JA, Wang FX, Zhang H, Wu KJ, Williams KJ, Daniel R: Evidence that the Nijmegen breakage syndrome protein, an early sensor of double-strand DNA breaks (DSB), is involved in HIV-1 post-integration repair by recruiting the ataxia telangiectasia-mutated kinase in a process similar to, but distinct from, cellular DSB repair. *Virol J* 2008, 5:11.
5. Andersen JL, Zimmerman ES, DeHart JL, Murala S, Ardon O, Blackett J, Chen J, Planelles V: ATR and GADD45alpha mediate HIV-1 Vpr-induced apoptosis. *Cell Death Differ* 2005, 12:326-334.
6. Skalka AM, Katz RA: Retroviral DNA integration and the DNA damage response. *Cell Death Differ* 2005, 12 Suppl 1:971-978.
7. Smith JA, Daniel R: Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses. *ACS Chem Biol* 2006, 1:217-226.
8. Craigie R, Bushman FD: HIV DNA Integration. *Cold Spring Harb Perspect Med* 2012, 2:a006890.
9. Roshal M, Kim B, Zhu Y, Nghiem P, Planelles V: Activation of the ATR-mediated DNA damage response by the HIV-1 viral protein R. *J Biol Chem* 2003, 278:25879-25886.
10. Daniel R, Ramcharan J, Rogakou E, Taganov KD, Greger JG, Bonner W, Nussenzweig A, Katz RA, Skalka AM: Histone H2AX is phosphorylated at sites of retroviral DNA integration but is dispensable for postintegration repair. *J Biol Chem* 2004, 279:45810-45814.
11. Zimmerman ES, Chen J, Andersen JL, Ardon O, Dehart JL, Blackett J, Choudhary SK, Camerini D, Nghiem P, Planelles V: Human immunodeficiency virus type 1 Vpr-mediated G2 arrest requires Rad17 and Hus1 and induces nuclear BRCA1 and gamma-H2AX focus formation. *Mol Cell Biol* 2004, 24:9286-9294.
12. Daniel R, Greger JG, Katz RA, Taganov KD, Wu X, Kappes JC, Skalka AM: Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the nonhomologous end joining repair pathway. *J Virol* 2004, 78:8573-8581.
13. Lilley CE, Schwartz RA, Weitzman MD: Using or abusing: viruses and the cellular DNA damage response. *Trends Microbiol* 2007, 15:119-126.
14. Yoder K, Sarasin A, Kraemer K, McIlhatton M, Bushman F, Fishel R: The DNA repair genes XPB and XPD defend cells from retroviral infection. *Proc Natl Acad Sci U S A* 2006, 103:4622-4627.

15. Cosnefroy O, Tocco A, Lesbats P, Thierry S, Calmels C, Wiktorowicz T, Reigadas S, Kwon Y, De Cian A, Desfarges S, et al: Stimulation of the human RAD51 nucleofilament restricts HIV-1 integration in vitro and in infected cells. *J Virol* 2012, 86:513-526.
16. Tyagi S, Ochem A, Tyagi M: DNA-dependent protein kinase interacts functionally with the RNA polymerase II complex recruited at the human immunodeficiency virus (HIV) long terminal repeat and plays an important role in HIV gene expression. *J Gen Virol* 2011, 92:1710-1720.
17. Jeanson L, Mouscadet JF: Ku represses the HIV-1 transcription: identification of a putative Ku binding site homologous to the mouse mammary tumor virus NRE1 sequence in the HIV-1 long terminal repeat. *J Biol Chem* 2002, 277:4918-4924.
18. Masson C, Bury-Mone S, Guiot E, Saez-Cirion A, Schoevaert-Brossault D, Brachet-Ducos C, Delelis O, Subra F, Jeanson-Leh L, Mouscadet JF: Ku80 participates in the targeting of retroviral transgenes to the chromatin of CHO cells. *J Virol* 2007, 81:7924-7932.
19. Waninger S, Kuhen K, Hu X, Chatterton JE, Wong-Staal F, Tang H: Identification of cellular cofactors for human immunodeficiency virus replication via a ribozyme-based genomics approach. *J Virol* 2004, 78:12829-12837.
20. Lieber MR: The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem* 2010, 79:181-211.
21. Gullo C, Au M, Feng G, Teoh G: The biology of Ku and its potential oncogenic role in cancer. *Biochim Biophys Acta* 2006, 1765:223-234.
22. Genovesio A, Kwon YJ, Windisch MP, Kim NY, Choi SY, Kim HC, Jung S, Mammano F, Perrin V, Boese AS, et al: Automated genome-wide visual profiling of cellular proteins involved in HIV infection. *J Biomol Screen* 2011, 16:945-958.
23. Jeanson L, Subra F, Vaganay S, Hervy M, Marangoni E, Bourhis J, Mouscadet JF: Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells. *Virology* 2002, 300:100-108.
24. Zheng Y, Ao Z, Wang B, Jayappa KD, Yao X: Host protein Ku70 binds and protects HIV-1 integrase from proteasomal degradation and is required for HIV replication. *J Biol Chem* 2011, 286:17722-17735.
25. Ariumi Y, Turelli P, Masutani M, Trono D: DNA damage sensors ATM, ATR, DNA-PKcs, and PARP-1 are dispensable for human immunodeficiency virus type 1 integration. *J Virol* 2005, 79:2973-2978.
26. Kilzer JM, Stracker T, Beitzel B, Meek K, Weitzman M, Bushman FD: Roles of host cell factors in circularization of retroviral dna. *Virology* 2003, 314:460-467.
27. Baekelandt V, Claeys A, Cherepanov P, De Clercq E, De Strooper B, Nuttin B, Debyser Z: DNA-Dependent protein kinase is not required for efficient lentivirus integration. *J Virol* 2000, 74:11278-11285.
28. Studamire B, Goff SP: Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology* 2008, 5:48.
29. Wang Y, Ghosh G, Hendrickson EA: Ku86 represses lethal telomere deletion events in human somatic cells. *Proc Natl Acad Sci U S A* 2009, 106:12430-12435.
30. Li G, Nelsen C, Hendrickson EA: Ku86 is essential in human somatic cells. *Proc Natl Acad Sci U S A* 2002, 99:832-837.
31. Ghosh G, Li G, Myung K, Hendrickson EA: The lethality of Ku86 (XRCC5) loss-of-function mutations in human cells is independent of p53 (TP53). *Radiat Res* 2007, 167:66-79.
32. Duverger A, Jones J, May J, Bibollet-Ruche F, Wagner FA, Cron RQ, Kutsch O: Determinants of the establishment of human immunodeficiency virus type 1 latency. *J Virol* 2009, 83:3078-3093.
33. Brussel A, Sonigo P: Analysis of early human immunodeficiency virus type 1 DNA synthesis by use of a new sensitive assay for quantifying integrated provirus. *J Virol* 2003, 77:10119-10124.
34. Shi L, Qiu D, Zhao G, Corthesy B, Lees-Miller S, Reeves WH, Kao PN: Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. *Nucleic Acids Res* 2007, 35:2302-2310.
35. Emiliani S, Van Lint C, Fischle W, Paras P, Jr., Ott M, Brady J, Verdin E: A point mutation in the HIV-1 Tat responsive element is associated with postintegration latency. *Proc Natl Acad Sci U S A* 1996, 93:6377-6381.
36. Buscher K, Trefzer U, Hofmann M, Sterry W, Kurth R, Denner J: Expression of human endogenous retrovirus K in melanomas and melanoma cell lines. *Cancer Res* 2005, 65:4172-4180.
37. Hohenadl C, Germaier H, Walchner M, Hagenhofer M, Herrmann M, Sturzl M, Kind P, Hehlmann R, Erfle V, Leib-Mosch C: Transcriptional activation of endogenous retroviral sequences in human epidermal keratinocytes by UVB irradiation. *J Invest Dermatol* 1999, 113:587-594.
38. Seifarth W, Frank O, Zeilfelder U, Spiess B, Greenwood AD, Hehlmann R, Leib-Mosch C: Comprehensive analysis of human endogenous retrovirus transcriptional activity in human tissues with a retrovirus-specific microarray. *J Virol* 2005, 79:341-352.

39. Israel N, Hazan U, Alcami J, Munier A, Arenzana-Seisdedos F, Bachelerie F, Israel A, Virelizier JL: Tumor necrosis factor stimulates transcription of HIV-1 in human T lymphocytes, independently and synergistically with mitogens. *J Immunol* 1989, 143:3956-3960.
40. Van Lint C, Emiliani S, Ott M, Verdin E: Transcriptional activation and chromatin remodeling of the HIV-1 promoter in response to histone acetylation. *Embo J* 1996, 15:1112-1120.
41. Pearson R, Kim YK, Hokello J, Lassen K, Friedman J, Tyagi M, Karn J: Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. *J Virol* 2008, 82:12291-12303.
42. Emiliani S, Fischle W, Ott M, Van Lint C, Amella CA, Verdin E: Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. *J Virol* 1998, 72:1666-1670.
43. Giffin W, Gong W, Schild-Poulter C, Hache RJ: Ku antigen-DNA conformation determines the activation of DNA-dependent protein kinase and DNA sequence-directed repression of mouse mammary tumor virus transcription. *Mol Cell Biol* 1999, 19:4065-4078.
44. Kaczmarski W, Khan SA: Lupus autoantigen Ku protein binds HIV-1 TAR RNA in vitro. *Biochem Biophys Res Commun* 1993, 196:935-942.
45. Bertinato J, Tomlinson JJ, Schild-Poulter C, Hache RJ: Evidence implicating Ku antigen as a structural factor in RNA polymerase II-mediated transcription. *Gene* 2003, 302:53-64.
46. Choul-li S, Drobecq H, Aumercier M: DNA-dependent protein kinase is a novel interaction partner for Ets-1 isoforms. *Biochem Biophys Res Commun* 2009, 390:839-844.
47. Kim H: DNA repair Ku proteins in gastric cancer cells and pancreatic acinar cells. *Amino Acids* 2008, 34:195-202.
48. Tuteja R, Tuteja N: Ku autoantigen: a multifunctional DNA-binding protein. *Crit Rev Biochem Mol Biol* 2000, 35:1-33.
49. Holcomb VB, Rodier F, Choi Y, Busuttil RA, Vogel H, Vijg J, Campisi J, Hasty P: Ku80 deletion suppresses spontaneous tumors and induces a p53-mediated DNA damage response. *Cancer Res* 2008, 68:9497-9502.
50. Kessis TD, Slebos RJ, Nelson WG, Kastan MB, Plunkett BS, Han SM, Lorincz AT, Hedrick L, Cho KR: Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc Natl Acad Sci U S A* 1993, 90:3988-3992.
51. Daniel R: DNA repair in HIV-1 infection: a case for inhibitors of cellular co-factors? *Curr HIV Res* 2006, 4:411-421.
52. Zufferey R, Dull T, Mandel RJ, Bukovsky A, Quiroz D, Naldini L, Trono D: Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol* 1998, 72:9873-9880.
53. Miyoshi H, Takahashi M, Gage FH, Verma IM: Stable and efficient gene transfer into the retina using an HIV-based lentiviral vector. *Proc Natl Acad Sci U S A* 1997, 94:10319-10323.
54. Ory DS, Neugeboren BA, Mulligan RC: A stable human-derived packaging cell line for production of high titer retrovirus/vesicular stomatitis virus G pseudotypes. *Proc Natl Acad Sci U S A* 1996, 93:11400-11406.
55. Schneider U, Schwenk HU, Bornkamm G: Characterization of EBV-genome negative "null" and "T" cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. *Int J Cancer* 1977, 19:621-626.
56. Foley GE, Lazarus H, Farber S, Uzman BG, Boone BA, McCarthy RE: Continuous Culture of Human Lymphoblasts from Peripheral Blood of a Child with Acute Leukemia. *Cancer* 1965, 18:522-529.
57. Chen C, Okayama H: High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol* 1987, 7:2745-2752.
58. Manic G, Maurin-Marlin A, Galluzzi L, Subra F, Mouscadet JF, Bury-Mone S: 3' Self-Inactivating Long Terminal Repeat Inserts for the Modulation of Transgene Expression from Lentiviral Vectors. *Hum Gene Ther Methods* 2012.
59. Pfaffl MW: A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001, 29:e45.
60. Nimura Y, Kawata T, Uzawa K, Okamura J, Liu C, Saito M, Shimada H, Seki N, Nakagawara A, Ito H, et al: Silencing Ku80 using small interfering RNA enhanced radiation sensitivity in vitro and in vivo. *Int J Oncol* 2007, 30:1477-1484.
61. Ayene IS, Ford LP, Koch CJ: Ku protein targeting by Ku70 small interfering RNA enhances human cancer cell response to topoisomerase II inhibitor and gamma radiation. *Mol Cancer Ther* 2005, 4:529-536.
62. Vitale I, Senovilla L, Jemaa M, Michaud M, Galluzzi L, Kepp O, Nanty L, Criollo A, Rello-Varona S, Manic G, et al: Multipolar mitosis of tetraploid cells: inhibition by p53 and dependency on Mos. *Embo J* 2010, 29:1272-1284.

Figures

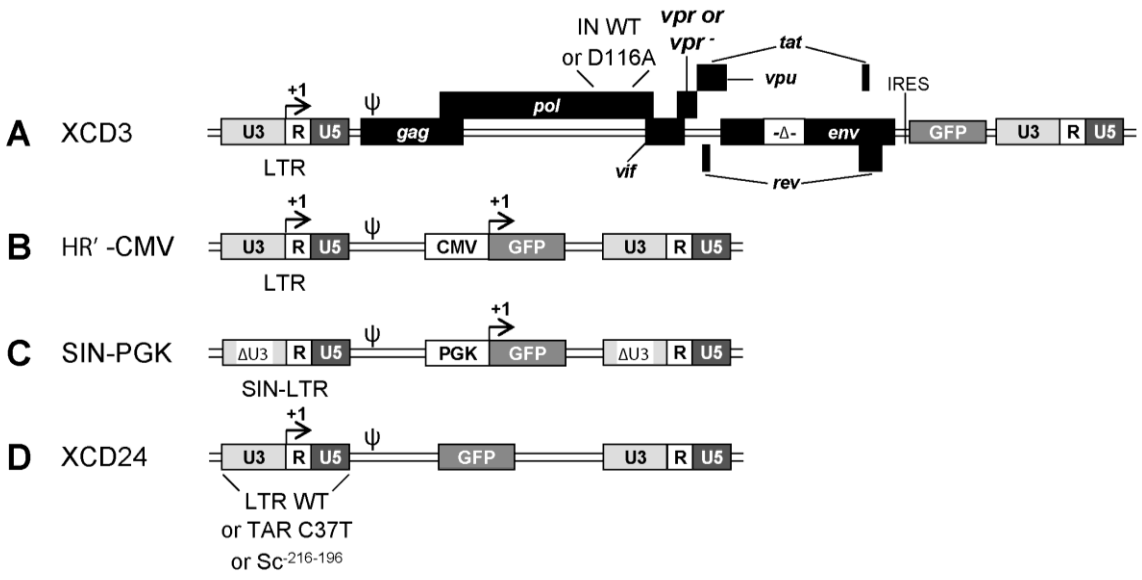


Figure 1: Design of lentiviral vectors. (**A**-**D**) Schematic representation of the lentiviral (LV) vectors (in the proviral, integrated form) used in this study. (**A**) XCD3 is a human immunodeficiency virus type 1 (HIV-1) construct lacking the *env* and *nef* genes and carrying the *green fluorescent protein* (*gfp)* transgene under the control of the native HIV-1 long terminal repeat (LTR) and an internal ribosome entry site (IRES) (also referred as HIV-1 *env*$^-$ *nef*$^-$ IRES-*gfp*). Two derived versions of this vector have been engineered, the first presenting the deletion of Vpr (*vpr*$^-$) and the second the D116A mutation in the integrase (IN) coding sequence [named XCD3 *vpr*$^-$ and XCD3 IN(D116A), respectively]. (**B**) HR'-CMV is a LV vector carrying the *gfp* transgene under the control of both the immediate early human cytomegalovirus (CMV) and HIV-1 promoters. (**C**) SIN-PGK is a self-inactivating (SIN, *i.e.*, U3-deleted, ΔU3) LV vector containing the *gfp* under the control of the human phosphoglycerate kinase (PGK) internal promoter. (**D**) XCD24 is a LV vector lacking all HIV-1 genes and carrying the *gfp* under the control of the WT HIV-1 promoter. Two variants of this vector have been derived, the first mutated in the putative Ku binding site [KBS, XCD24 Sc$^{-216-196}$] and the second in the *trans*-activation response (TAR) element [XCD24 TAR(C37T)]. ψ, encapsidation signal.

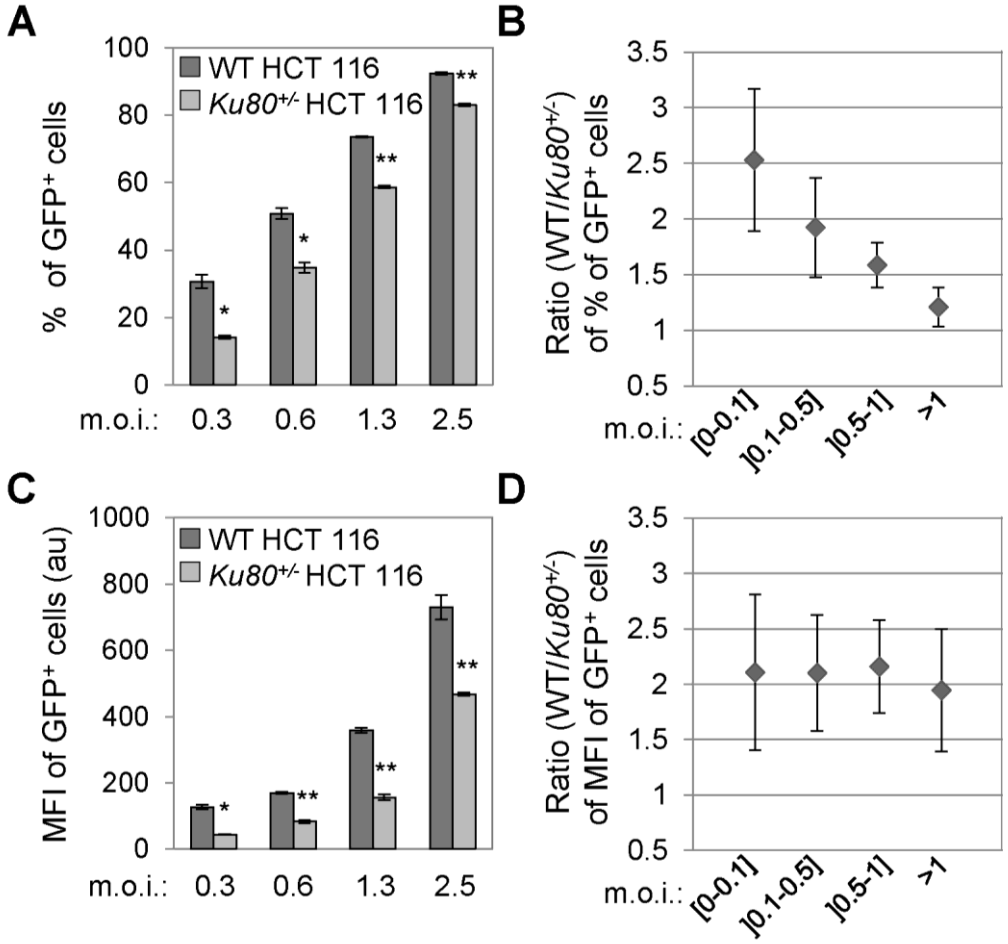


Figure 2: Ku80 haplodepletion reduces HIV-1-driven GFP expression. (**A**-**D**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transduced with serial dilutions of XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*) as indicated and, 48 hours later, analyzed by means of cytofluorimetry for the percentage and the geometric mean fluorescence intensity (MFI) of GFP-positive (GFP^{+}) cells. Panels (**A**) and (**C**) show the results (mean of duplicates ± SD; *, $p < 0.05$; **, $p < 0.01$) of the percentage and the MFI of GFP^{+} cells, respectively, obtained in one representative experiment (out of three independent ones that yielded similar results) at the multiplicity of infection (m.o.i.) illustrated. The ratios between WT and *Ku80*$^{+/-}$ HCT 116 cells for the percentage and the MFI of GFP^{+} cells at the depicted m.o.i. are illustrated in panels (**B**) and (**D**), respectively (mean ± SD, n = 3). AU, arbitrary units.

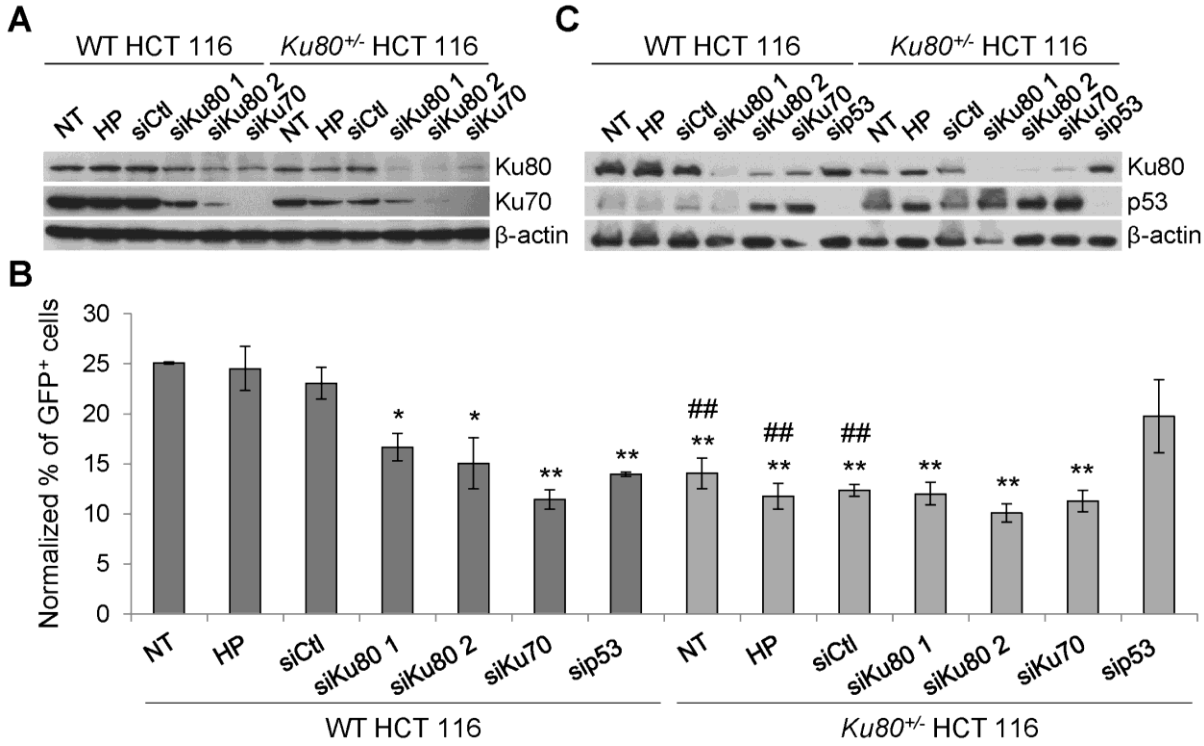


Figure 3: HIV-1-driven GFP expression in WT HCT 116 cells is decreased by Ku transient depletion. (**A**-**C**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transfected with an unrelated sequence (siCtl) or with siRNAs directed against Ku80 (siKu80 1 or 2), Ku70 (siKu70) or p53 (sip53). Seventy-two hours later, the protein levels of Ku80 and Ku70 (**A**) or Ku80 and p53 (**C**) were assessed by Western-blot analysis. Actin levels were monitored to ensure equal loading of lanes. Alternatively, 72 h upon transfection, cells were transduced with XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*) at low m.o.i. (< 0.3) for additional 48 h followed by cytofluorimetric assessment of GFP expression (**B**). NT (non-transfected) and HP (Hiperfect®) correspond to cells kept in control condition or exposed to the same amount of liposomes used for transfection but in absence of siRNA, respectively. To avoid bias linked to cell counting, in (**B**) the percentages of GFP^{+} cells were normalized to the total viral DNA content per cell as determined by quantitative PCR (Q-PCR) 24 h after transduction. The results obtained from at least 3 independent experiments were then normalized to a GFP expression of 25 % (of note, the level of GFP expression of NT, HP or siCtl WT cells before normalization were always comprised between 20 to 30 %). Results are expressed as mean ± SEM. *, $p < 0.05$ and **, $p < 0.01$ as compared to HP and siCtl-transfected WT cells; $^{\#}$, $p < 0.05$ and $^{\#\#}$, $p < 0.01$ as compared to WT cells subjected to the same transfection condition.

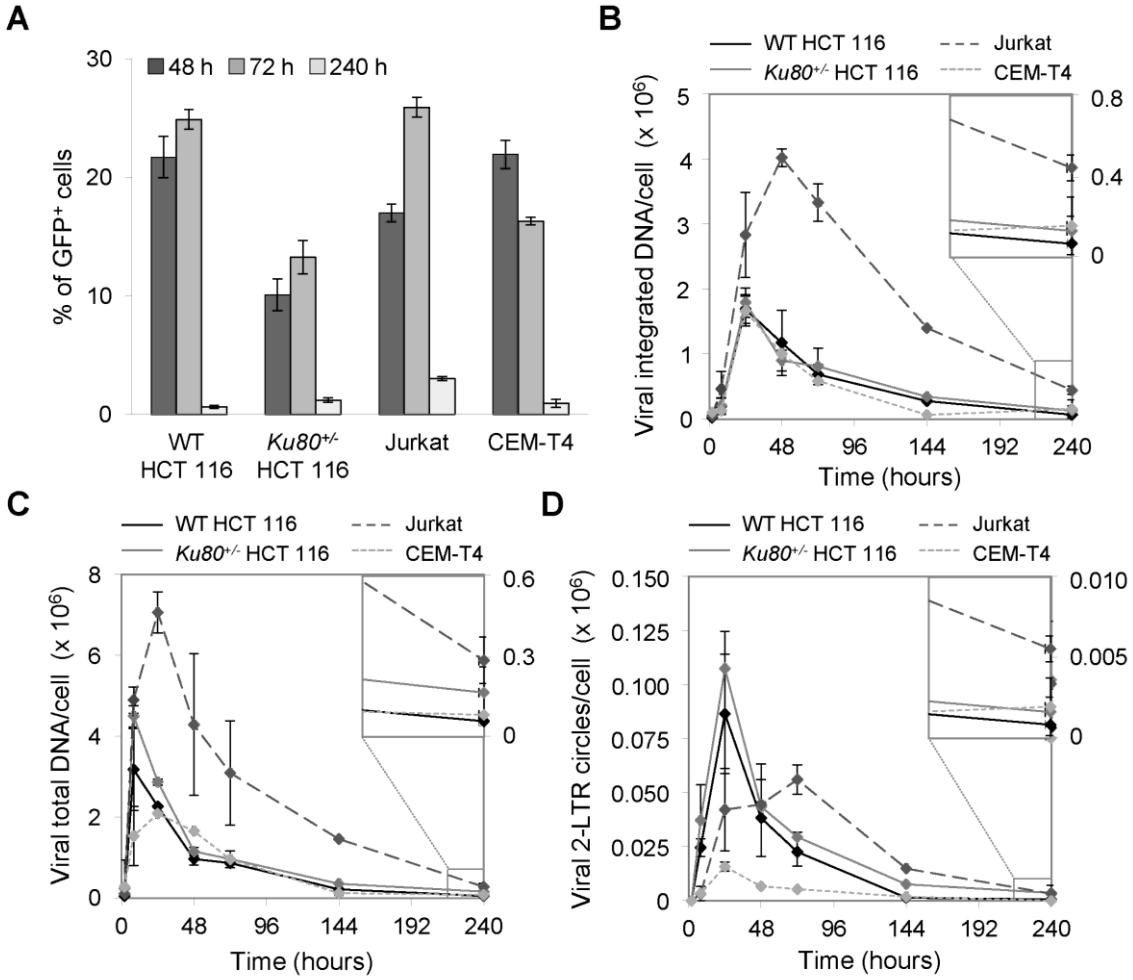


Figure 4: Ku80 haplodepletion does not affect HIV-1 integration efficiency and 2-LTR circle formation. (**A-D**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells as well as the human T lymphoid leukemia Jurkat and CEM-T4 cell lines were transduced with the same volume of XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*) at a m.o.i < 0.2. Upon the depicted time from transduction, cells were subjected either to cytofluorimetric analysis, to measure GFP expression (**A**) or to Q-PCR assays, to quantify viral DNA species (**B-D**). In panel (**A**), results obtained in one representative experiment (out of 2 independent ones that yielded similar results) are shown (mean of duplicates ± SD). In panels (**B-D**), the copy number of the integrated viral DNA (**B**), the total viral DNA (**C**) and the 2-LTR circular forms of viral DNA (**D**) per 1x10^{6} cells for each cell line are reported (mean ± SEM; n = 2). The last data point (240 h from the transduction) is shown at higher magnification.

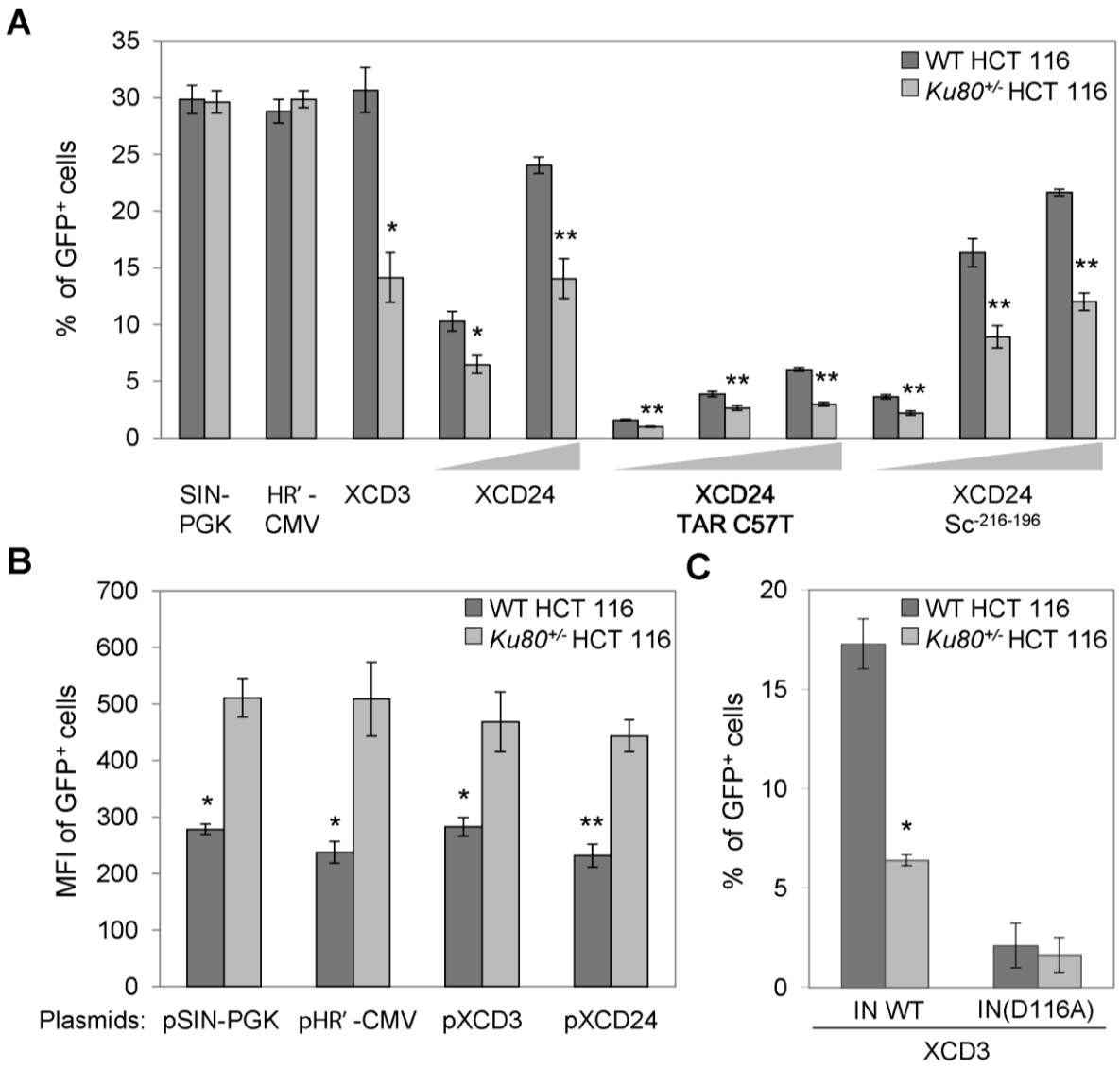


Figure 5: Ku stimulates HIV-1-driven expression independently from Tat or the putative KBS$^{-216-196}$ sequence but requires provirus integration. (**A**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transduced for 48 h either with SIN-PGK, HR'-CMV, XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*), XCD24 (at a m.o.i. 0.1 and 0.25) or its derivatives in which the putative KBS (XCD24 Sc$^{-216-196}$) or TAR element [XCD24 TAR(C37T)] was mutated (at a m.o.i. 0.02 to 0.06 or 0.04 to 0.22, respectively) (for more detail on these vectors, refer to **Figure 1**). Thereafter, the percentage of fluorescent (GFP^{+}) cells was measured by flow cytometry. Of note, the volume of viral suspension was adjusted to the number of cells to be transduced for both cell lines. (**B**) Alternatively, cells were transfected with pXCD3, pXCD24, pHR'-CMV and pSIN-PGK shuttle plasmids, followed by cytofluorimetry-mediated determination of the MFI of fluorescent cells. (**C**) WT and *Ku80*$^{+/-}$ HCT 116 cells were transduced for 48 h either with XCD3 or its integrase-defective derivative, XCD3 IN(D116A) (at a m.o.i. 0.15), followed by cytofluorometry-mediated assessment of the percentage of GFP^{+} cells. In all panels, results are expressed as mean ± SD (*, $p < 0.05$; **, $p < 0.01$) and refer to the values obtained in one representative experiment conducted in duplicate out of (at least) 2 independent observations yielding similar results (**A**) or coming from at least 3 experiments (**B**,**C**). When not specified, the m.o.i. was 0.3.

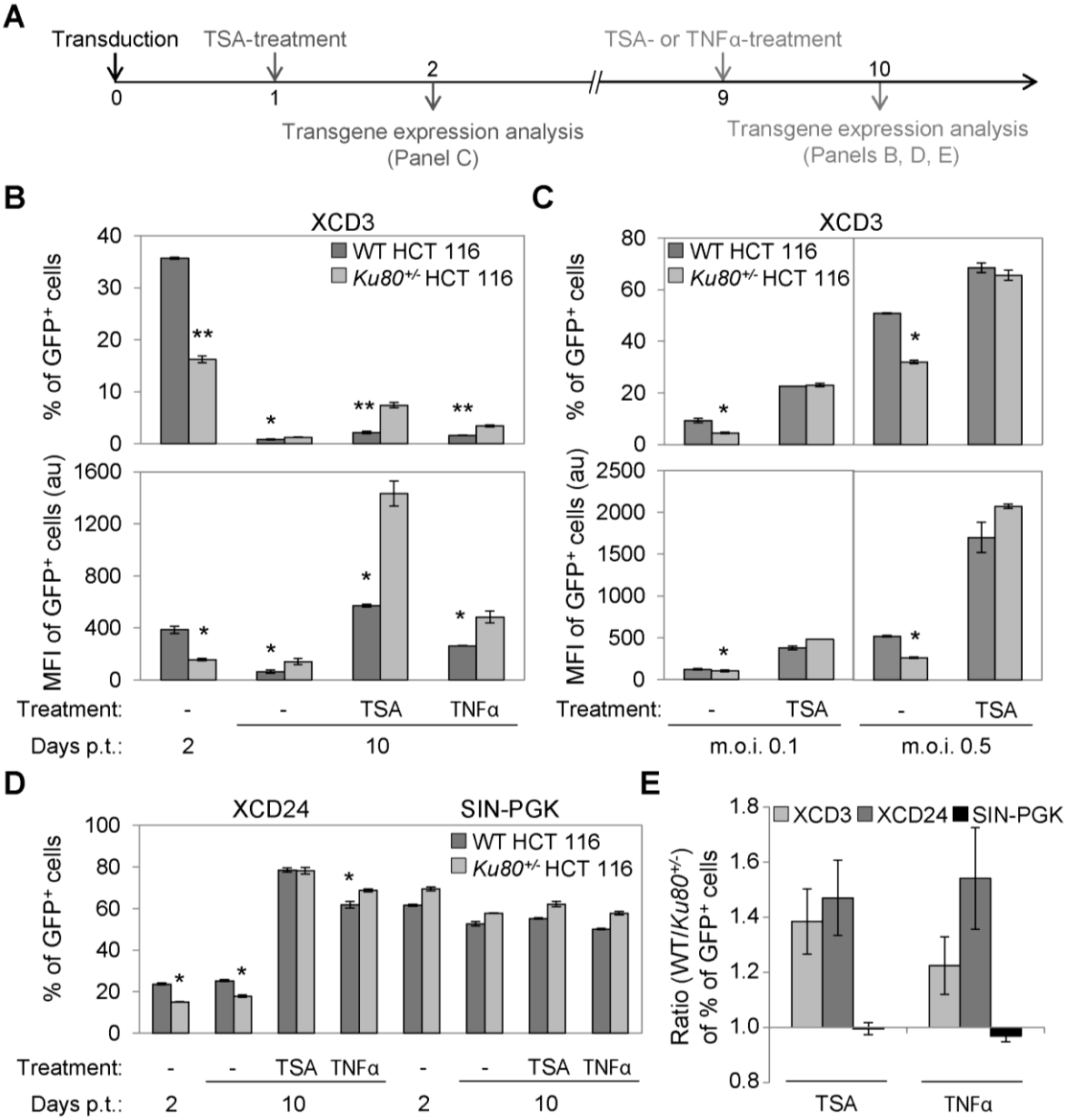


Figure 6: Ku80 haplodepletion increases the level of TNFα- or TSA-induced HIV-1 reactivation. (**A-C**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells transduced with XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*) at an m.o.i. < 0.6 were incubated for 24 h in absence or presence of 200 nM tricostatin A (TSA) or 10 ng/ml tumor necrosis factor α (TNFα) upon cultivation for 9 days in standard condition (**A**,**B**), or - alternatively - upon 24 h post-transduction (**A**,**C**). Thereafter, the percentage and the MFI of GFP^{+} cells were estimated by flow cytometry 2 (**B**,**C**) and 10 (**B**) days post-transduction (p.t.). Results from one representative experiment (out of 3 three independent ones yielding similar results) performed in duplicate are presented as mean ± SD (*, $p < 0.05$; **, $p < 0.01$). (**D**) WT and *Ku80*$^{+/-}$ HCT 116 cells transduced with XCD24 or SIN-PGK were treated 9 days later with TSA or TNFα as described in (**A**,**B**). The panel reports data for one representative observation (out of 3 three independent ones yielding similar results) performed in duplicate (mean ± SD; *, $p < 0.05$; **, $p < 0.01$). (**E**) Ratio between the percentage of GFP^{+} cells of *Ku80*$^{+/-}$ and WT HCT 116 cells in presence or absence of TSA or TNFα, as determined in (**B**) and (**D**) for each LV vector (mean ± SD, n = 3).

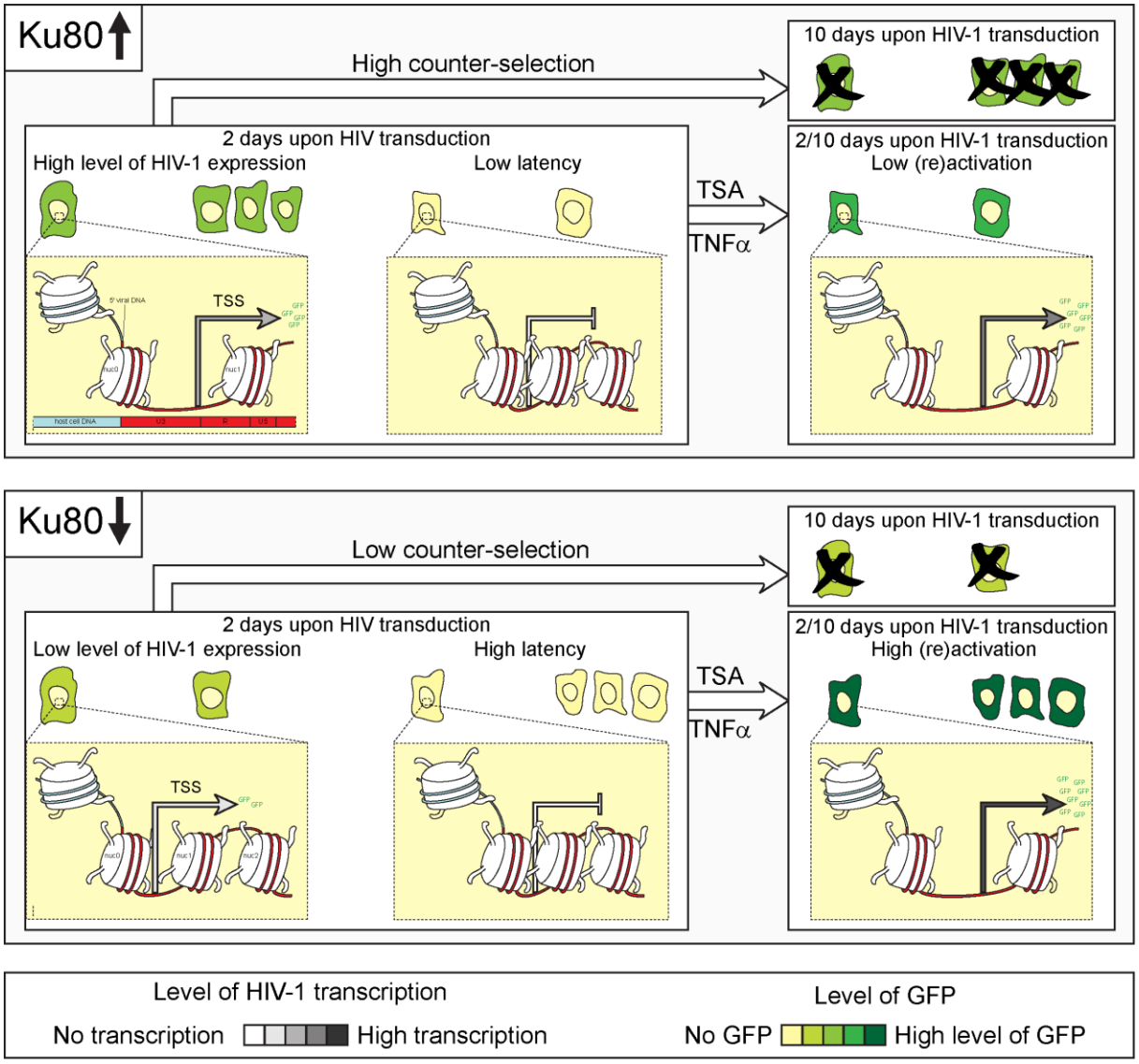
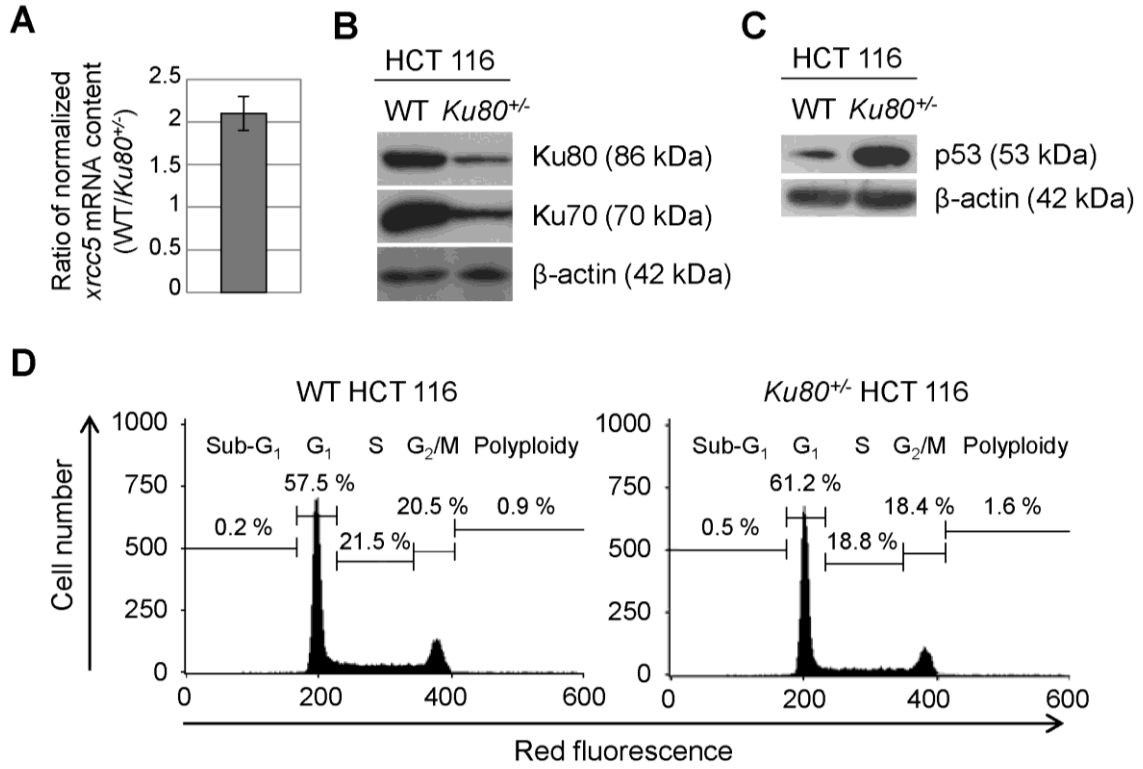
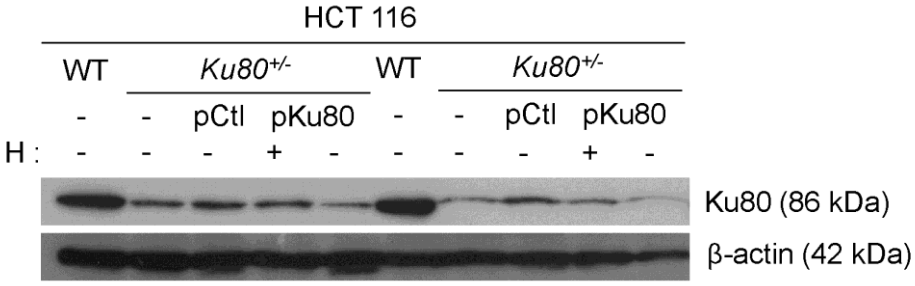


Figure 7: Hypothetical model for the role of Ku on HIV-1 expression and latency at early and middle-time upon transduction. (**A**) When expressed at normal level (as in Ku80-proficient cells), Ku stimulates the early expression of HIV-1, a condition that is accompanied by a low level of viral latency and in turn leads to a high counter-selection of transduced cells over the time. This latter effect is probably due to Vpr-induced toxicity (see text for more details). (**B**) On the contrary, low level of Ku (as in *Ku80*$^{+/-}$ cells or in WT cells transfected with Ku-depleting siRNAs) negatively affects the expression of HIV-1, increases the proportion of cells with latent viruses and reduces the counter-selection phenomenon. Accordingly, the reactivation of HIV-1 upon TSA- or TNFα-treatment at 2 or 10 days after transduction is more efficient in *Ku80*$^{+/-}$ cells that in their WT counterparts. Hence, we hypothesize that the depletion of Ku may favor the appearance of cells harboring integrated proviruses in a latent and readily reactivable state starting from the early phases from HIV-1 transduction.

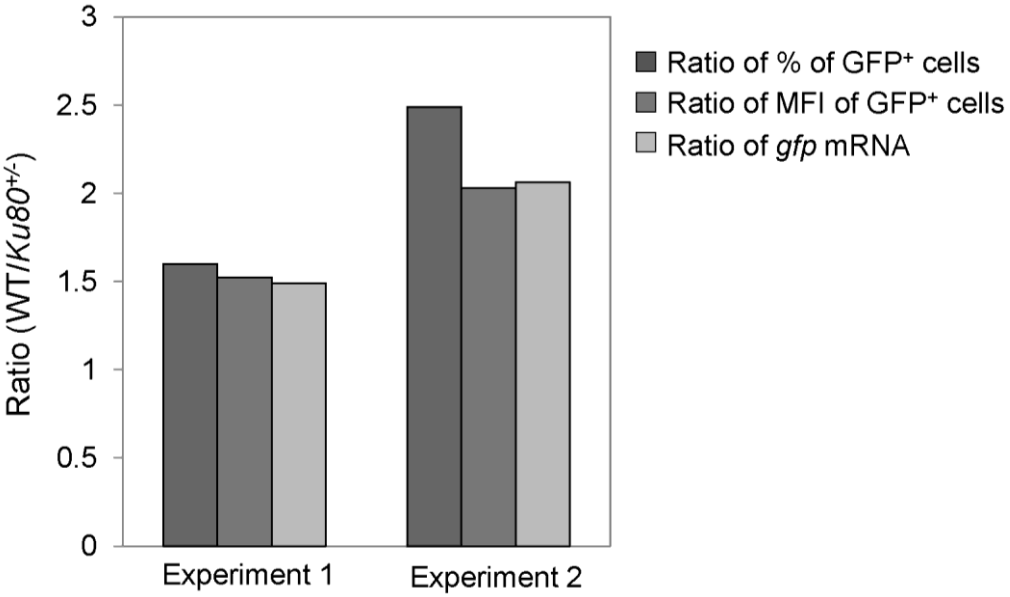
Supplementary Figures



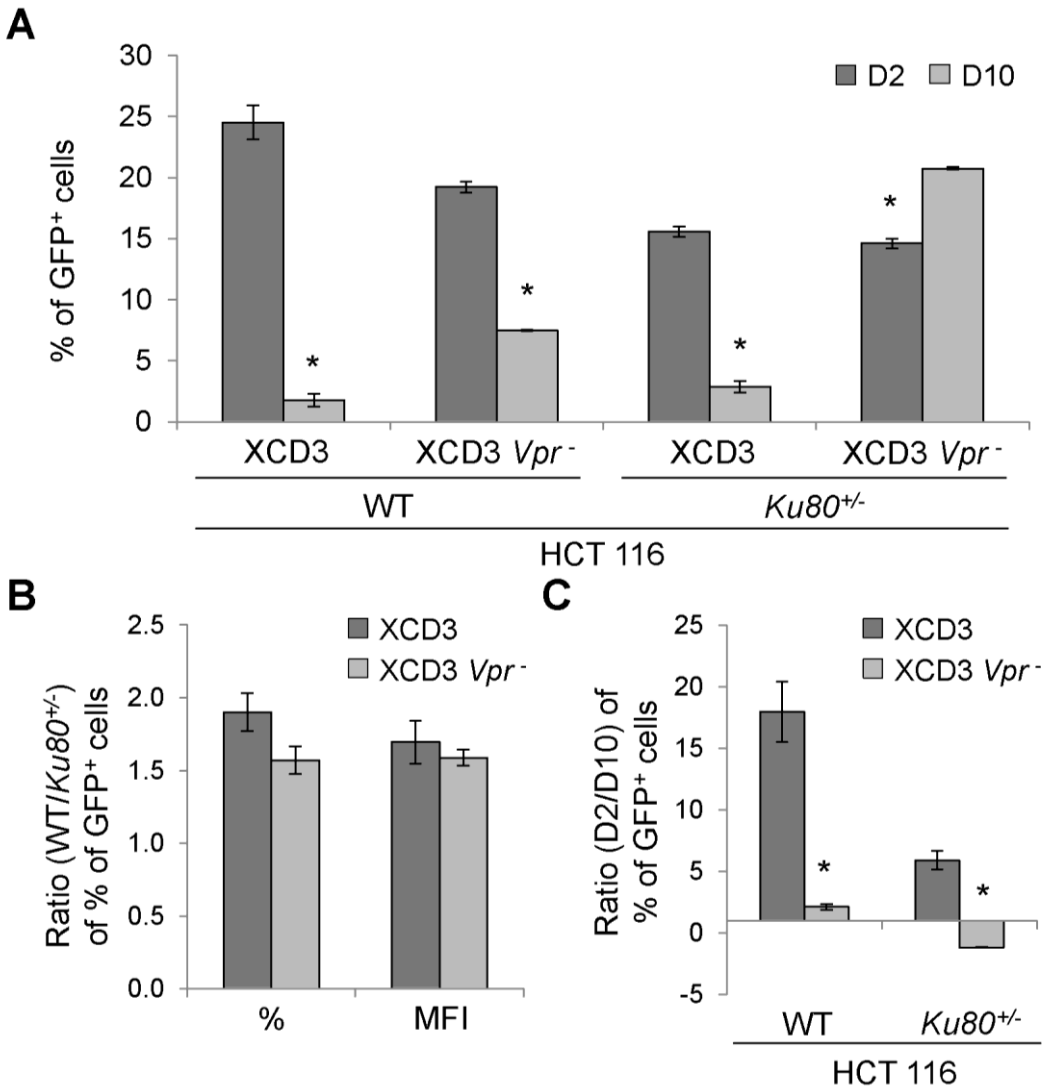
Supplementary Figure 1: Characterization of WT and *Ku80*$^{+/-}$ HCT 116 cells. (**A**,**C**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were subjected to RT-Q-PCR analysis (**A**), to evaluate the level of Ku80 encoding *xrcc5* mRNA and western-blot assay (**B-C**), to assess Ku80, Ku70 ad p53 protein contents. In panel (**A**), *xrcc5* mRNA content (normalized to that of *β-actin* mRNA) of WT HCT 116 cells was normalized to that of *Ku80*$^{+/-}$ cells (mean ± SD, n=3). In panels **B**,**C**, actin levels were monitored to ensure equal loading of lanes. (**D**) WT and *Ku80*$^{+/-}$ HCT 116 cells were stained with propidium iodide for the cytofluorometric assessment of cell cycle progression. Panel **D** reports cell cycle distributions and quantitative data for both cell lines of one representative observation out of 3 three independent ones yielding similar results.



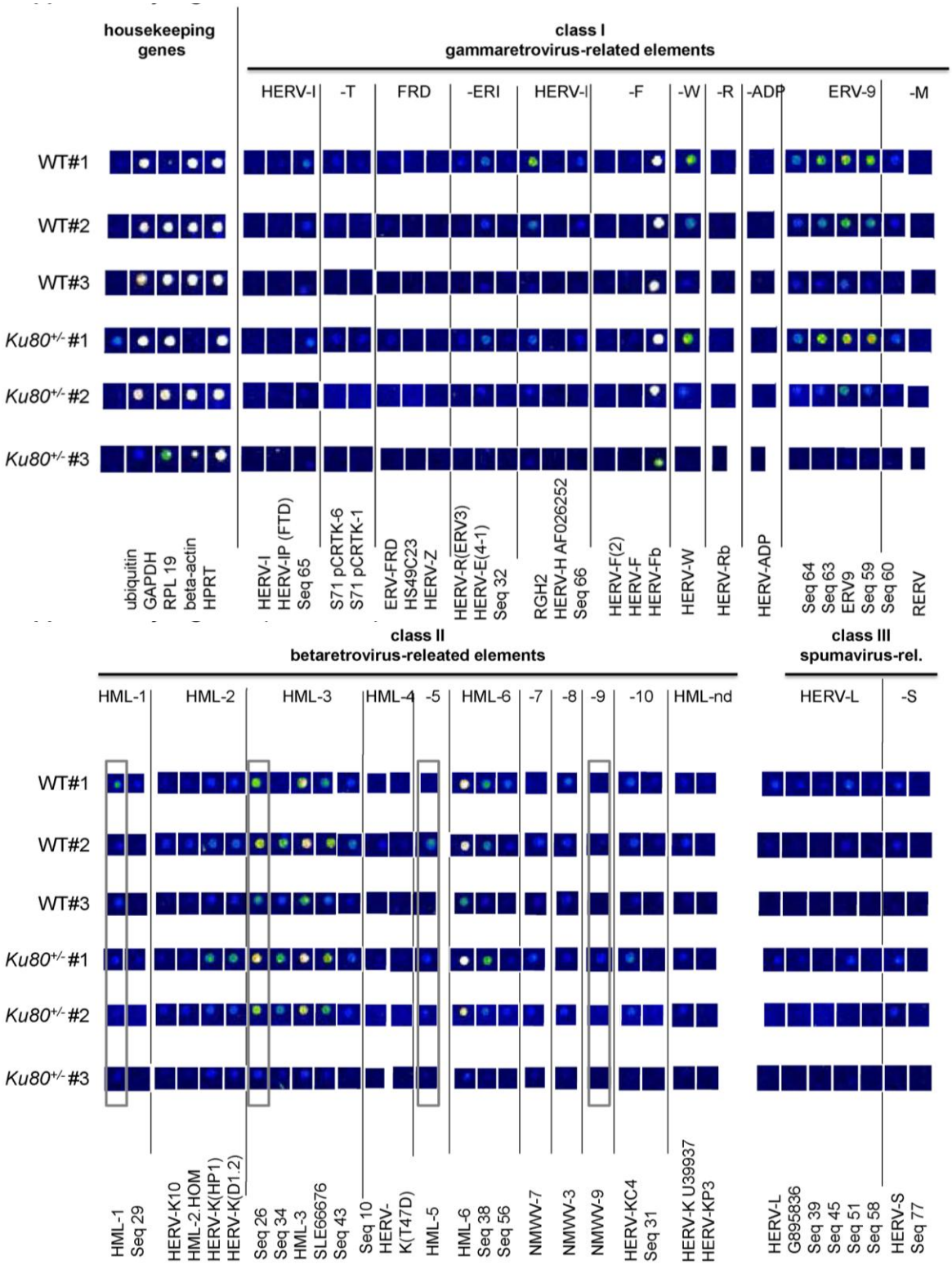
Supplementary Figure 2: Characterization of a Ku *trans*-complemented *Ku80*$^{+/-}$ HCT 116 clone. Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells as well as derivatives containing either a neomycin (Neo) or a hygromycin-Ku80 expression pcDNA3.1 plasmid (shortened as 'pCtl' and 'pKu80', respectively) were subjected to western-blot assessment of Ku80 protein contents. Actin level was monitored to ensure equal loading of lanes. Please note that the pKu80 *Ku80*$^{+/-}$ HCT 116 clone was maintained or not under a 10 µg/ml hygromycin selection (noted as 'H').



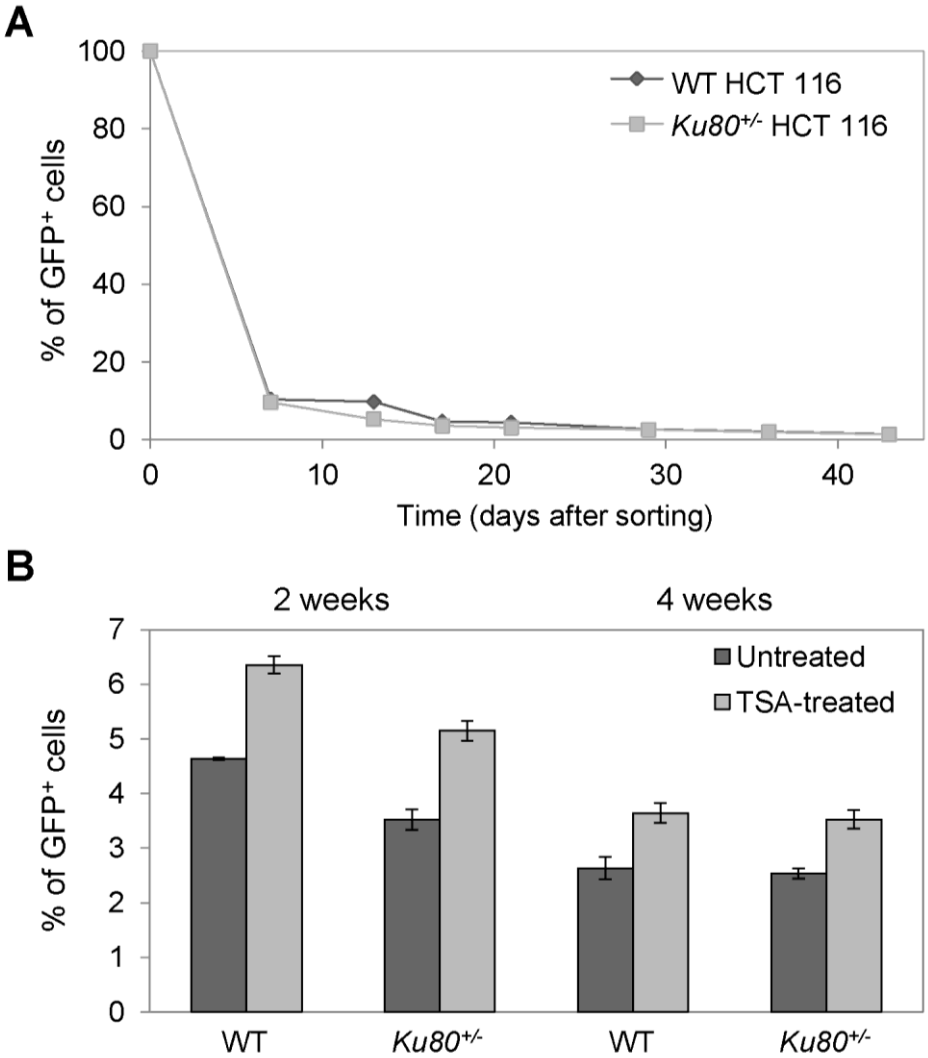
Supplementary Figure 3: GFP expression level correlates to *gfp* mRNA content in both WT and *Ku80*$^{+/-}$ HCT 116 cells. Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transduced at middle m.o.i. (< 0.5) with XCD3 (HIV-1 *env*$^{-}$ *nef*$^{-}$ IRES-*gfp*) and subjected 48 h later to either flow cytometric analyses, to measure the percentage and the MFI of GFP-positive cells or RT-Q-PCR assays, to quantify the *gfp* mRNA content (normalized to *gapdh* mRNA content). The ratios between WT and *Ku80*$^{+/-}$ cells HCT 116 cells for the indicated parameters are illustrated. Results correspond to two representative and independent experiments. Of note, the percentages of GFP^{+} WT cells in experiments 1 and 2 were of 15 % and 25 %, respectively.



Supplementary Figure 4: The decrease of HIV-1-driven expression over time is reduced in the absence of Vpr. (**A-C**) Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transduced with XCD3 expressing *Vpr* (XCD3), or its derivative lacking *Vpr* (XCD3 *Vpr*$^{-}$, refer to **Figure 1A** for more details on these vectors). Upon 2 or 10 days from transduction, the percentage of GFP^{+} cells was determined by flow cytometry (D2 and D10, respectively). (**A**) One representative experiment performed in triplicates (out of four independent ones yielding similar results) is shown (mean ± SD; *, $p < 0.05$). In panel **B**, the ratios of GFP-positive cells between WT and *Ku80*$^{+/-}$ HCT 116 cells 2 days after transduction with XCD3 or XCD3 *Vpr*$^{-}$ [as obtained in (A)] are illustrated (mean ± SD, n = 4). In panel **C**, the ratios of GFP-positive cells between 2 and 10 days after transduction with XCD3 or XCD3 *Vpr*$^{-}$, as obtained in (**A**) in cells with the depicted genetic background are represented (mean ± SD; *, $p < 0.05$; n = 3).



Supplementary Figure 5: HERVs show the same transcription profile in both WT and *Ku80*$^{+/-}$ HCT 116 cells. Three independent pairs of total RNA from WT and Ku80$^{+/-}$ human colon carcinoma HCT 116 were compared by microarray hybridization. Images were saved in the Bitmap file format and were used to present the results by false color mapping. Each hybridization result is representative of a triplicate yielding similar results. Grey boxes indicate the four HERV groups that were identified by Biometric Research Branch-ArrayTools(BRB)-mediated analyses (for more detail, please refer to Methods section).



Supplementary Figure 6: Ku80 haplodepletion does not increase the HIV-1 reactivation level in GFP$^+$ transduced cells. (**A**,**B**). Wild-type (WT) and *Ku80*$^{+/-}$ human colon carcinoma HCT 116 cells were transduced for 2 days with XCD3 (HIV-1 *env*$^-$ *nef*$^-$ IRES-*gfp*) at low m.o.i. (< 0.1; *i.e.*, one provirus per cell), followed by the isolation of GFP$^+$ cells by means of cytofluorometry-mediated cell sorting. Thereafter, GFP$^+$ isolated cells were maintained in culture for 45 days and the percentages of GFP-expressing were routinely monitored by flow cytometry-mediated analysis, as indicated (**A**). Alternatively, 2 and 4 weeks after cell sorting, transduced cells were left untreated or exposed for 24 h to 200 nM TSA (**B**). Columns in (**B**) represent percentages of GFP$^+$ cells in each condition (mean ± SD). Results in panels (**A**) and (**B**) are from one representative experiment performed at least 2 independent times.

Table S1: HERV groups analyzed by RT-Q-PCR

Condition: WT HCT 116 versus *Ku80*$^{+/-}$ HCT 116						
Locus	**Class**	**Parametric analysis on Norm2 data** (Ratio, Student's *t* test)		**BrB analysis**		**RT-Q-PCR results:** Ratio WT/*Ku80*$^{+/-}$ (± SD) 2 independent experiments realized on each RNA extract
		Ratio WT/*Ku80*$^{+/-}$ (Norm1 data)	***P* value**	**Norm1 Data** (parametric *p* value or «- » if p>0,01)	**Norm3 data** (parametric *p* value or «- » if p>0,01)	
ERV-FRD	γ	0,7	0,00002	-	-	0,8 (± 0,2)
ERV9 (seq59)	γ	0,7	0,009	-	-	nd
HML-2 HERV-K (HP1)	β	0,7	0,003	-	-	0,8 (± 0,3)
HERV-S	spuma	0,6	0,005	-	-	nd
ERI HERV-R (ERV3)	γ	0,7	0,001	-	-	nd
HML-7 MMWV-7	β	0,6	0,0008	-	-	nd
HML-9 MMWV-9	β	**0,6**	**0,000001**	**0,001**	**0,006**	ns
HML-4 (Seq10)	β	0,7	0,0000003	-	-	nd
HML-5	β	**0,4**	**0,00053**	**0,009**	**0,003**	0,7 (± 0,3)
HML-1	β	**1,6**	**0,008**	**0,007**	**-**	0,8 (± 0,3)
HML-3 (Seq26)	β	**0,8**	**0,006**	**-**	**0,09**	nd
Actin						1,0 (± 0,2)

nd (not determined)
ns (non specific signal of RT-Q-PCR)

Table S2: Sequence of primers, probes and siRNA used for this study

Name	Target	Sequence*	Reference
Primers used for cloning			
OBM101	LV LTR	GGTACCTTTAAGACCAATG	This study
OBM102		GTCAAGTAGACCTAGGAGAGTGGTCATCCATCCCATGCAGGCT	
OBM103		ACTCTCCTAGGTCTACTTGACGGAGGTTTGACAGCCGCCTAG	
OBM104		TCTAGATGCTGCTAGAGATTTTC	
OBM156	HIV *pol*	GTACATACAGCCAATGGCAGCAATTTCACCAG	[58]
OBM157		TGCCATTGGCTGTATGTACTGTTTTTACTGG	
OBM200	HIV *tar* sequence	AGAGAACTCCCAGGCTCAGATCT	This study
OBM201		GGGAGTTCTCTGGCTAACTAGG	
Primers used for quantitative RT-PCR			
OBM209	human *XRCC5* gene (encoding Ku80)	ACCAAAGCGCCTGAGGAC	This study
OBM210		CTCATGGTAAAGCCCACGTC	
OBM257	*hr-gfp*	AAACCTGGCCCTGTCTTCTT	This study
OBM258		GAGGAACTGCTTCCTTCACG	
OBM24	Human *β-actin*	AAATCTGGCACCACACCTTC	This study
R		TCACCGGAGTCCATCACGAT	
OBM241	human *gapdh*	CGCTCTCTGCTCCTCCTGTT	[58]
OBM242		CCATGGTGTCTGAGCGATGT	
OBM247	ERV-FRD	AAAAAGGAAGAAGTTAACAGC	[38]
OBM248		ATATAAAGACTTAGGTCCTGC	
OBM249	HML2-ERV-K-HP1	GGCCATCAGAGTCTAAACCAC	From Leib-Mösch's lab
OBM250		CTGACTTTCTGGGGGTGGCCG	
OBM251	HML-5	TGAAAGGCCAGCTTGCTG	[38]
OBM252		CAATTAGGAAATTCTTTTCTAC	
OBM253	HML-1	AAAATCAGGGAAATGGAGAATG	This study
OBM254		CTTGGGCAGCTAAAGGAATG	
Primers used for quantitative PCR			
Alu1	Human Alu	TCCCAGCTACTGGGGAGGCTGAGG	[33]
Alu2		GCCTCCCAAAGTGCTGGGATTACAG	
L-M667	LV	ATGCCACGTAAGCGAAACTCTGGCTAACTAGGGAACCCACTG	
Lambda-T	L-M667	ATGCCACGTAAGCGAAACT	
AA55M	LV	GCTAGAGATTTTCCACACTGACTAA	

MH 531		TGTGTGCCCGTCTGTTGTGT	
MH 532-UNIV		CCGAGTCCTGCGTCGAGAGA	This study
HIV F	2-LTR LV junction	GTGCCCGTCTGTTGTGTGACT	[33]
HIV R1		ACTGGTACTAGCTTGTAGCACCATCCA	
BGLOB sense	β-globin	CAACTTCATCCACGTTCACC	Control kit DNA; Roche
BGLOB antisense		ACACAACTGTGTTCACTAGC	
Probes used for quantitative PCR			
LTR-FL	LV	CACAACAGACGGGCACACACTACTTGA (3' fluoresceine)	[33]
LTR-FC		CACTCAAGGCAAGCTTTATTGAGGC (5' LC red 640 dye and 3'-P)	
MH-FL		CCCTCAGACCCTTTTAGTCAGTGTGGAA (3' fluoresceine)	
MH-LC		TCTCTAGCAGTGGCGCCCGAACAG (5' LC red 640 dye and 3'-P)	
HIV FL	2-LTR LV junction	CCACACACAAGGCTACTTCCCTGA (3' fluoresceine)	
HIV LC		TGGCAGAACTACACACCAGGGC (5' LC red 640 dye et 3' phosphorylée)	
siRNA (5'-3'sense sequence indicated)			
Ku80 2	*xrcc5* mRNA encoding Ku80)	GCGAGUAACCAGCUCAUAAdTdT	[60]
Ku80 1		GAGCUAAUCCUCAAGUCGGdTdT	[19]
Ku70	*xrcc6* mRNA encoding Ku70)	GAUGCCCUUUACUGAAAAAdTdT	[61]
P53	*tp53* mRNA	GUGAGCGCUUCGAGAUGUUdTdT	[62]
Ctl	Scramble sequence unrelated to human genome	GCCGGUAUGCCGGUUAAGUdTdT	[62]

III/ IMPACT DE LA MACHINERIE DE REPARATION DE L'ADN SUR L'EXPRESSION DE VECTEURS LENTIVIRAUX CONTENANT L'INSULATEUR CHS4 AU SEIN DE LEURS SIN-LTRS

A/ CONTEXTE DE L'ETUDE

Au cours de notre première étude, nous avons mis en évidence l'action positive orientée de l'insulateur cHS4 (en antisens) sur l'expression du transgène (Manic et al., 2012). Cependant, nous avons observé que cet effet est dépendant du promoteur qui contrôle l'expression de la *gfp* (Manic et al., 2012). Ainsi, l'augmentation de l'expression transgénique en présence de cHS4 antisens est transitoire ou révélée au cours du temps lorsque la *gfp* est sous contrôle du promoteur CMV ou PGK, respectivement (Manic et al., 2012). De plus, le bénéfice apporté par l'insulateur cHS4 en antisens sur l'expression transgénique contrôlée par le P_{PGK} est variable en fonction du type cellulaire transduit (*e.g.*, HCT 116 *vs* CEM-T4) (Manic et al., 2012). Il est à noter, qu'en accord avec la littérature, nous avons caractérisé l'expression à partir du P_{PGK} comme stable au cours du temps alors que celle contrôlée par le P_{CMV} est sujette à extinction (Manic et al., 2012).

Lors de notre deuxième étude, nous n'avons pas observé d'impact de la déplétion en Ku sur l'expression transgénique précoce de vecteurs dérivés du VIH-1 contenant un promoteur interne CMV ou PGK (également à moyen-terme pour ce dernier) au sein des cellules humaines HCT 116 (Manic *et al.*, en préparation ; *cf.*, § II). Cependant, nous n'avons pas testé l'évolution de l'expression de la *gfp* contrôlée par le P_{CMV} dans un contexte déplété en Ku. En effet, Ku pourrait influencer l'expression du transgène au cours du temps en intervenant, par exemple, sur sa localisation, comme récemment décrit au sein des cellules de rongeurs (Masson et al., 2007).

Au cours de cette étude, nous avons souhaité évaluer si la déplétion en Ku pouvait influencer l'extinction de l'expression transgénique contrôlée par le P_{CMV}. De plus, nous avons voulu tester un éventuel impact du statut en Ku sur l'expression du transgène *a priori* isolé de son environnement (*i.e.*, bordé de l'insulateur cHS4). Pour répondre à cette problèmatique, nous avons réalisé des expériences de transduction au sein des cellules HCT 116 WT et *Ku80*$^{+/-}$ avec des vecteurs dérivés du VIH-1 codant la *gfp* sous contrôle du P_{CMV} ou P_{PGK} et contenant l'insulateur cHS4 au sein de leur SIN-LTRs.

Nous avons ainsi mis en évidence que le promoteur interne (CMV ou PGK) ou le fond génétique transduit (WT ou *Ku80*$^{+/-}$) n'influence pas la perte du pourcentage des cellules

GFP^+ transduites par les vecteurs contenant cHS4 orienté en antisens. De plus, l'insertion de cHS4 en antisens au sein des SIN-LTRs induit une augmentation spécifique de l'intensité de l'expression de la *gfp* normalisée. Enfin, de façon étonnante, l'expression de la *gfp* sous contrôle du P_{PGK} augmentée par la présence de cHS4 en antisens est favorisée lors de la déplétion en Ku. Ces données suggèrent une régulation différentielle de l'activité de cHS4 en antisens en fonction du statut en Ku dans notre modèle.

B/ RESULTATS (NON-PUBLIES)

Les cellules humaines HCT 116 WT et *Ku80*$^{+/-}$ (issues d'un carcinome colorectal) ont été transduites avec des vecteurs codant la *gfp* sous contrôle du promoteur CMV ou PGK, et contenant au sein de leur SIN-LTRs l'insulateur cHS4 dans l'orientation sens ou antisens. Le pourcentage des cellules GFP^+ et l'intensité de l'expression (*i.e.*, MFI) de la GFP ont été estimés par cytofluorométrie 2, 6 et 10 jours après transduction.

Nous avons tout d'abord mis en évidence que la perte du pourcentage des cellules GFP^+ a lieu principalement entre 2 et 6 jours post-transduction de façon identique pour chacun de vecteurs transduits (vide, ou contenant cHS4 orienté en sens ou en antisens) (**Figure 32A-C**), comme observé précédemment (**Figure S5**, #1). Nous avons toutefois noté que cette chute du pourcentage des cellules GFP^+ semble supérieure pour les cellules transduites par les vecteurs contenant le promoteur CMV, comparé à celles transduites avec les constructions SIN-P_{PGK} (**Figure 32A-C**). Ces données pourraient s'expliquer par la capacité du P_{CMV} à s'exprimer à partir des formes d'ADN virales non-intégrées qui sont perdues lors des divisions cellulaires, en accord avec nos données précédentes (**Figure S4**, #1). Cependant, nous n'avons pas observé de différence significative de la perte du pourcentage de cellules GFP^+ entre les deux fonds génétiques HCT 116 transduits (*i.e.*, WT et *Ku80*$^{+/-}$, **Figure 32A-C**).

La perte du pourcentage des cellules WT ou *Ku80*$^{+/-}$ GFP^+ transduites par les vecteurs SIN-P_{PGK} ou -P_{CMV} contenant cHS4 (en antisens) de 2 à 10 jours post-transduction est donc similaire quelque soit le promoteur interne contrôlant l'expression de la *gfp* et le fond génétique transduit (**Figure 32**).

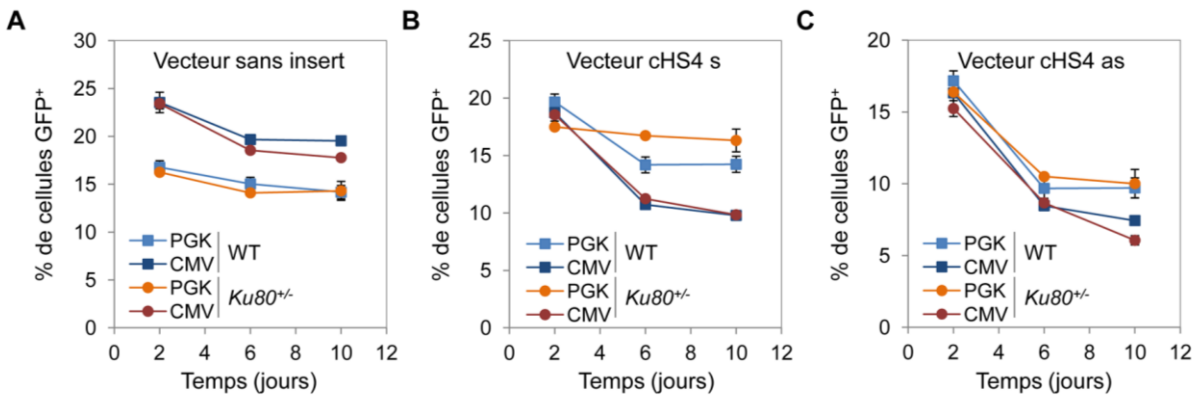


Figure 32 : Suivi des cellules HCT 116 WT et *Ku80*$^{+/-}$ GFP^{+} transduites par des vecteurs lentiviraux contenant cHS4 au sein de leur SIN-LTRs. Les cellules humaines HCT 116 WT et Ku80$^{+/-}$ (issues d'un carcinome colorectal) ont été transduites avec des vecteurs codant la *gfp* sous contrôle du promoteur CMV ou PGK, et contenant au sein de leur SIN-LTRs l'insulateur cHS4 dans l'orientation sens (s) ou antisens (as). Le pourcentage des cellules GFP^{+} a été estimé par cytofluorométrie 2, 6 et 10 jours après transduction. (**A**) Cinétique du pourcentage des cellules WT ou Ku80$^{+/-}$ GFP^{+} transduites par les vecteurs SIN-PGK ou -CMV sans insert (**A**) ou contenant cHS4 orienté en antisens (**B**) ou en sens (**C**) (moyenne ± SD, résultats représentatifs d'au moins deux expériences indépendantes).

L'intensité de l'expression de la GFP à partir des vecteurs contenant cHS4 (orienté en sens ou en antisens) a ensuite été normalisée sur celle du vecteur parental correspondant au sein de populations cellulaires ayant des efficacités de transduction comparables. L'analyse de l'intensité de l'expression de la GFP normalisée pendant 10 jours post-transduction a mis en évidence une réduction de l'expression de la GFP des cellules WT ou *Ku80*$^{+/-}$ transduites par les vecteurs contenant cHS4 (en sens) quelque soit le promoteur interne (**Figure 33A, B**).

A l'inverse, l'expression transgénique contrôlée par le P_{PGK} est augmentée au sein des cellules WT ou *Ku80*$^{+/-}$ transduites par les vecteurs contenant cHS4 en antisens (**Figure 33A, B**), de façon cohérente avec nos observations précédentes au sein de différentes lignées cellulaires (**Figures 4** et **5**, #1). Cependant, l'expression de la *gfp* contrôlée par le promoteur CMV à partir des vecteurs contenant cHS4 en antisens est également augmentée à 10 jours post-transduction (**Figure 33B, C**). Ces données sont en désaccord avec ce que nous avons observé précédemment après transduction des cellules HeLa (**Figure 4**, #1).

Enfin, de façon étonnante, nous avons mis en évidence que l'augmentation de l'intensité de l'expression de la *gfp* des vecteurs contenant cHS4 (en antisens) est supérieure lorsque la *gfp* était sous contrôle du promoteur PGK dans un fond génétique haplodéplété en Ku80 (comparé au promoteur CMV et aux cellules WT, **Figure 33A-C**).

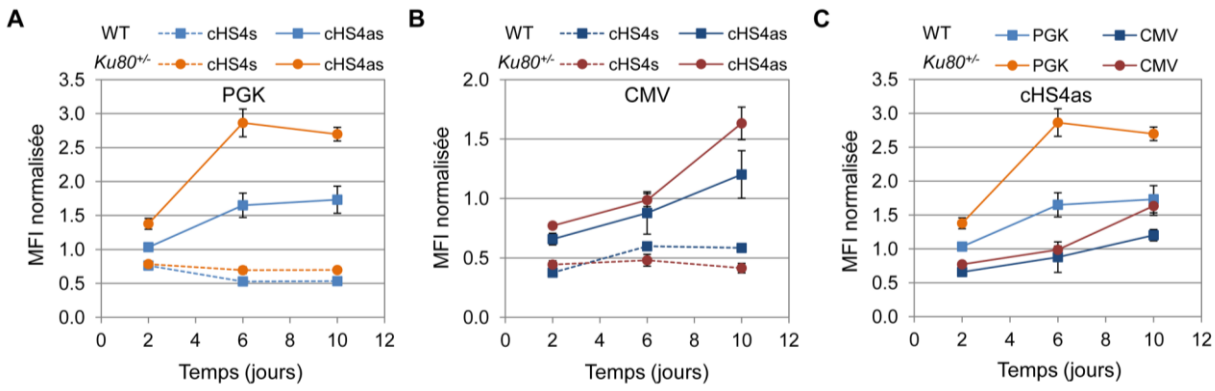


Figure 33 : L'expression de la *gfp* à partir d'un vecteur lentiviral est augmentée lorsque cHS4 est inséré en antisens dans les SIN-LTRs au sein des cellules *Ku80*$^{+/-}$. (**A-C**) Les cellules humaines HCT 116 WT et Ku80$^{+/-}$ (issues d'un carcinome colorectal) ont été transduites avec des vecteurs codant la *gfp* sous contrôle du promoteur CMV ou PGK, et contenant au sein de leur SIN-LTRs l'insulateur cHS4 dans l'orientation sense (s) ou antisense (as). Le pourcentage et l'intensité de l'expression de la GFP ont été estimés par cytofluorométrie 2, 6 et 10 jours après transduction. Ici, l'intensité de l'expression de la GFP des vecteurs contenant cHS4 (en s ou as) a été normalisée sur celle du vecteur parental correspondant. (**A-B**) Cinétique de l'intensité de l'expression de la GFP normalisée des cellules WT ou Ku80$^{+/-}$ transduites par les vecteurs (**A**) SIN-PGK ou (**B**) SIN-CMV contenant cHS4 (en s ou as) (moyenne ± SEM, résultats représentatifs d'au moins deux expériences indépendantes). (**C**) Cinétique de l'intensité de l'expression de la GFP normalisée des cellules WT ou Ku80$^{+/-}$ transduites par les vecteurs SIN-PGK ou -CMV contenant cHS4as (moyenne ± SEM, valeurs issues de trois expériences indépendantes).

C/ Discussion

Nous avons mis en évidence que la présence de cHS4 (en antisens) favorise l'expression de la *gfp* à partir du P_{PGK} (dès 2 jours post-transduction) et du P_{CMV} (à partir de 10 jours post-transduction) au sein des cellules HCT 116. Ces données sont cohérentes avec nos observations précédentes dans le cas du P_{PGK} (**Figures 4** et **5**, #1) mais pas dans le cas du P_{CMV} pour lequel nous avons observé un impact positif transitoire de cHS4 (en antisens) sur l'expression transgénique au sein des cellules HeLa (**Figure 4**, #1). Ces contradictions pourraient s'expliquer par le fait que (*i*) l'expression contrôlée par le P_{CMV} (à partir des formes intégrées et non-intégrées de l'ADNv) soit plus sensible au contexte cellulaire au sein duquel le P_{CMV} agit (comparé au P_{PGK}) et/ou (*ii*) que l'action positive sur l'expression de la séquence cHS4 en antisens puisse être dépendante du type cellulaire transduit (*e.g.*, HCT 116 *vs.* HeLa).

De façon intéressante, nous avons également montré que la déplétion en Ku favorise spécifiquement l'expression de la *gfp* à partir du P_{PGK} en présence de cHS4 (en antisens) au sein des cellules HCT 116. Cet effet pourrait avoir lieu grâce à l'augmentation de l'activité

barrière et/ou de terminaison de transcription de cHS4 (en antisens) par (*i*) un ciblage dans des zones transcriptionnellement actives par défaut en absence de Ku, comme hypothétisé au sein des cellules de rongeurs, où l'action de cHS4 se révèlerait davantage, et/ou (*ii*) l'intervention, par défaut ou favorisée par la déplétion de Ku, d'une protéine impliquée dans la régulation de l'activité barrière de cHS4 (*i.e.*, PARP-1). Un environnement favorable à l'expression qui serait améliorée/créé par cHS4 (en antisens) devrait davantage favoriser le P_{CMV}, plus susceptible au *silencing*. Cependant, aucune différence significative de l'expression du transgène contrôlée par le P_{CMV} n'a été observée entre le fond génétique WT ou *Ku80*$^{+/-}$ (**Figure 33B, C**). Nous ne pouvons exclure cette notion de ciblage mais privilégions l'intervention de facteurs propices à l'expression transgénique en présence de cHS4 (en antisens) favorisée par la déplétion en Ku. Cette hypothèse inclut la possibilité d'une prise en charge précoce de l'ADNv linéaire par Ku, ou par PARP-1 lorsque les quantités de Ku sont moindres, qui se fixerait de façon non-spécifique aux extrémités de l'ADNv (*cf.*, discussion § II.B.2 pour plus de détails).

DISCUSSION ET PERSPECTIVES

Ces travaux de thèse visaient à étudier les impacts de déterminants viraux et cellulaires sur la régulation de l'expression du VIH-1 et de vecteurs qui en dérivent. Les figures des articles de la section Résultats seront citées dans la section Discussion en utilisant les annotations #1 pour (Manic et al., 2012) et #2 pour Manic *et al.* (soumis à *Retrovirology*).

I/ IMPACTS DES DETERMINANTS VIRAUX SUR L'EXPRESSION LENTIVIRALE

Nous avons mis en évidence que l'absence de certaines protéines virales au sein du VIH-1 (**partie A**) ou la présence de séquences hétérologues dans les vecteurs dérivés (**partie B**) peut influencer l'expression lentivirale et le devenir des cellules transduites.

A/ IMPACTS DES DETERMINANTS VIRAUX SUR L'EXPRESSION DU VIH-1

1/ Perte limitée des cellules transduites par le VIH-1 en absence de Vpr

Nous avons décrit une chute importante de l'expression du VIH-1 à moyen-terme (10 jours) après transduction de cellules tumorales épithéliales (HCT 116 WT) ou lymphocytaires (CEM-T4, Jurkat) (**Figure 4A** #2). Celle-ci était corrélée à une forte diminution de la quantité d'ADN viral par cellule (**Figure 4C, B** #2), ce qui indique qu'il y a une contre-sélection massive des cellules GFP^+ qui expriment les protéines du VIH-1. A l'inverse, la perte de l'expression du VIH-1 au cours du temps est limitée ou nulle en absence de Vpr (**Figure S4** #2) ou de l'ensemble des protéines auxiliaires du VIH-1 (**Figure 6C** #2), respectivement, après transduction des cellules HCT 116.

Ces observations suggèrent que l'expression des protéines auxiliaires, dont Vpr, induirait une toxicité et/ou, au moins, une contre-sélection des cellules transduites par le VIH-1 dans notre modèle. Ces résultats sont cohérents avec le rôle double décrit de Vpr qui, néo-synthétisée après infection, induit un arrêt du cycle cellulaire en G_2/M, ce qui favorise l'expression du LTR (Goh et al., 1998) mais provoque une mort par apoptose (Andersen et al., 2008). Nos données devront toutefois être confirmées par des expériences de Q-PCR.

2/ Etat de latence des cellules transduites par le VIH-1 en absence de Tat

Nous avons observé que l'expression du VIH-1 est réduite en absence de protéines auxiliaires (dont Tat) après transduction des cellules HCT 116 (données non-montrées). Ce résultat est corrélé à une absence de contre-sélection au cours du temps (comme décrit ci-

dessus, § I.A.1) et à une (ré)activation massive après traitement à l'aide d'un inhibiteur des HDACs, la TSA ou d'un activateur de NF-κB, le TNFα (**Figure 6C** #2).

Ces données indiquent que l'absence des protéines auxiliaires, dont Tat, diminue l'expression précoce du VIH-1 et induit une latence post-intégrative du VIH-1 au sein des cellules HCT 116. Nos observations sont en accord avec le rôle de Tat dans la *trans*-activation du LTR (Colin and Van Lint, 2009) ainsi que l'état de pseudo-latence du VIH-1 décrit lorsque Tat n'est pas fonctionnelle, comme au sein des cellules U1 (Emiliani et al., 1998), ou de latence transcriptionnelle en absence des co-facteurs cellulaires de Tat (Colin and Van Lint, 2009). Tat aurait donc un rôle dans l'inhibition de la mise en place d'un *silencing* transcriptionnel du VIH-1, comme récemment décrit par (Donahue et al., 2012). Nos résultats doivent cependant être confirmés par une *trans*-complémentation en Tat.

*L'étude des impacts des déterminants viraux sur l'expression du VIH-1 nous a permis de décrire que (*i*) l'absence de Vpr limiterait la contre-sélection des cellules transduites, et (*ii*) l'absence des autres protéines auxiliaires virales, dont Tat, réduirait considérablement l'expression des provirus, qui se trouveraient dans un état latent (re)activable.*

B/ MODULATION DE L'EXPRESSION LENTIVIRALE LORS DE L'INSERTION DE SEQUENCES HETEROLOGUES AU SEIN DE VECTEURS DERIVES DU VIH-1

1/ Impacts de promoteurs internes sur l'expression d'un transgène intégré ou non

a/ *Le promoteur CMV induit une expression variable qui tend à être éteinte*

Nous avons mis en évidence que le promoteur viral CMV (P_{CMV}) induit une expression hétérogène du transgène (**Figure S4B** #1) qui est cependant équivalente dans un contexte compétent ou déficient pour l'intégration (**Figure S1B, D** #1). Ces données suggèrent que les formes d'ADNv intégrées et non-intégrées s'expriment (précocement) de la même façon à partir du P_{CMV}. L'expression transgénique contrôlée par le P_{CMV} à partir des vecteurs non-intégratifs a toutefois été décrite être inférieure à celle issue de vecteurs intégratifs (Bayer et al., 2008). Malgré cela, le P_{CMV} reste un promoteur de choix dans les approches IDLV (Wanisch and Yanez-Munoz, 2009).

De plus, l'expression de la *gfp* contrôlée par le P_{CMV} est (*i*) sujette à un *silencing* précoce, comme observé après traitement à la TSA 2 jours post-transduction au sein d'une population mixte (*i.e.*, GFP^- et GFP^+, données non-montrées), (*ii*) soumise à extinction au sein

des cellules HeLa GFP^+ isolées après transduction (**Figure 8C** #1), et (*iii*) associée à de la variégation au sein de populations clonales (**Figure 8D** #1). Ces résultats sont en accord avec la description de *silencing* de l'expression régulée par le P_{CMV} observé après transduction de lentivecteurs (Gerolami et al., 2000) ou infection par le CMV (Stinski and Isomura, 2008).

b/ *Le promoteur humain PGK induit une expression homogène et stable*

Le promoteur cellulaire PGK (P_{PGK}) induit une expression homogène du transgène (**Figure S4A** #1) de façon indépendante du type cellulaire transduit (données non-montrées), en accord avec (Qin et al., 2010). De plus, cette expression est moindre en absence d'intégrase fonctionnelle (**Figures S1 et S4B, D** #1), ce qui suggère que le transgène doit être intégré pour que le P_{PGK} soit fonctionnel et/ou être localisé au sein d'un environnement chromatinien favorable. Enfin, l'expression de la *gfp* contrôlée par P_{PGK} est stable dans le temps, comme observé au sein d'une population transduite mixte (**Figure S5A** #1) ou GFP^+ (au moins lorsque les vecteurs sont d'origine lentivirale, **Figure 8C** #1).

L'ensemble de ces données concernant la régulation de l'expression par les promoteurs CMV et PGK sont en accord avec leurs caractéristiques intrinsèques précédemment décrites (*cf.*, Introduction § III.C.2). Nos observations sont également cohérentes avec les études comparant l'expression transgénique contrôlée par ces promoteurs à partir de vecteurs rétroviraux (Norrman et al., 2010; Wang et al., 2008a; Xia et al., 2007).

2/ <u>Modulation de l'expression transgénique lors de l'insertion de séquences hétérologues au sein des SIN-LTRs</u>

a/ *L'insertion dans les SIN-LTRs influence peu les étapes pré-intégratives du cycle lentiviral*

Une étude précédente de Kumar *et al.* a mis en évidence une réduction des titres viraux de façon semi-logarithmique avec l'augmentation de la taille du vecteur (Kumar et al., 2001). En accord avec cette étude, nos résultats montrent une tendance des larges inserts à induire une diminution plus importante des titres viraux comparé aux inserts plus petits (**Figure 1B** #1), quelque soit le promoteur interne, l'orientation de l'insert et sa capacité à fixer CTCF. Nous avons toutefois observé que la réduction des titres viraux semblait plus dépendante de la séquence insérée dans les SIN-LTRs et de son orientation, que de sa taille (*i.e.*, LacZ' sens, **Figure 1B** #1). Cette diminution des titres viraux pourrait être liée à des étapes pré-intégratives moins efficaces (dont la transcription inverse dans le cas de LacZ',

Figure S2 #1) (Arumugam et al., 2009; Hanawa et al., 2009) et/ou au niveau d'expression au sein de la cellule cible, comme détaillé dans le paragraphe **b/** ci-dessous. Il est à noter que l'influence de l'insulateur cHS4 (cloné en orientation sens ou antisens) sur les titres lentiviraux est marginale, en accord avec les résultats d'études publiées qui montraient un effet négligable (Yannaki et al., 2002) ou même positif (Emery et al., 2000) de cHS4.

Nous avons observé une efficacité d'intégration similaire des vecteurs utilisés dans cette étude, quelque soit la séquence insérée dans les SIN-LTRs et sa taille (**Figure 2** #1). Ces donnés nous ont donc permis d'exclure une instabilité génétique des inserts (qui induirait une perte du transgène après recombinaison homologue entre inserts) comme cause de la diminution des titres viraux proposée par (Arumugam et al., 2009; Nielsen et al., 2009).

b/ *Impacts de l'insertion dans les SIN-LTRs sur l'expression transgénique*

Comme suggéré ci-dessus, la diminution des titres viraux pourrait provenir du niveau d'expression au sein des cellules transduites puisque tous les inserts, excepté l'insulateur cHS4 inséré en orientation antisens, diminuent significativement l'expression précoce de la *gfp* après transduction (**Figure 3** #1). En accord avec nos données, cHS4 cloné au sein des (SIN-)LTRs a précédemment été décrit comme diminuant l'expression transgénique à partir de promoteurs internes (Hanawa et al., 2009; Ramezani et al., 2003; Uchida et al., 2011; Yannaki et al., 2002) mais ce phénomène avait été attribué à sa capacité de bloquer l'*enhancer* du 3' LTR (Yannaki et al., 2002)et/ou à des contraintes topologiques (Ramezani et al., 2003; Yannaki et al., 2002). Nos résultats indiquent que l'atténuation de l'expression du transgènique en présence de cHS4 inséré en orientation sens n'est pas spécifiquement lié à la capacité de cHS4 à recruter CTCF puisqu'il persiste en présence de séquences ne fixant pas CTCF au sein des SIN-LTRs (*i.e.*, DS18 et LacZ', **Figure 3** #1). Il a été décrit que l'insertion de séquences *spacer* potentiellement inertes d'origine exogène diminuait légèrement l'expression du transgène (Aker et al., 2007). Nous avons observé un phénomène similaire : DS18 et LacZ', qui ont été sélectionnées comme des séquences contrôles, atténuent significativement l'expression du transgène.

Ces observations soulèvent la question de l'existence de séquences totalement inertes pouvant être utilisées comme contrôles lors de l'insertion au sein des vecteurs viraux.

Des études précédentes avaient montré que l'efficacité de transduction des vecteurs contenant de larges inserts au sein des SIN-LTRs diminue en même temps qu'augmente la taille de l'insert (Arumugam et al., 2009; Bos et al., 2010; Hanawa et al., 2009; Nielsen et al., 2009). Nous avons également observé une tendance de la majorité des atténuations de

l'expression du transgène les plus prononcées à se produire avec les larges inserts (**Figure 3** #1). Cependant, une exploration approfondie de ce phénomène à l'aide de séquences dérivées de DS18 n'a pas confirmé cette tendance (**Figure 6** #1). De plus, l'orientation de l'insert ainsi que la délétion de certains éléments (*e.g.*, MIRb de DS18) influencent significativement l'atténuation de l'expression transgènique (au moins dans certains cas, **Figures 3 et 6** #1).

L'ensemble de nos résultats indique que la taille de l'insert n'est pas le seul paramètre qui influence l'expression du transgène à partir des vecteurs lentiviraux.

Il est à noter que l'atténuation de l'expression du transgène associée à l'insertion dans les SIN-LTRs est visible en absence d'intégration de l'ADNv uniquement lorsque l'expression est sous contrôle du P_{CMV} (**Figure 3A** #1), de façon cohérente avec le fait que l'expression à partir du P_{CMV} soit indépendant de l'intégration de l'ADNv (*cf.*, ci-dessus § a). Enfin, l'atténuation de l'expression du transgène associé à l'insertion dans les SIN-LTRs que nous décrivons est stable à long-terme, et ne modifie pas la variégation (**Figure 8** #1).

c/ *Hypothèses proposées pour expliquer l'atténuation de l'expression du transgène associée à l'insertion dans les SIN-LTRs*

Plusieurs phénomènes pourraient expliquer l'atténuation de l'expression du transgène qui résulte de l'insertion de séquences au sein des SIN-LTRs. Premièrement, le recrutement et/ou l'exclusion de facteurs cellulaires à proximité des insertions pourrait influencer l'activité du promoteur interne. Cependant, l'atténuation de l'expression du transgène que nous avons observé est indépendante de la séquence (**Figures 3 et 4** #1) et du type cellulaire (**Figure 5** #1), excluant *de facto* cette première possibilité. Deuxièmement, des changements de la topologie ADN du LTR pourraient empêcher l'accès des facteurs de transcription au promoteur interne. De plus, l'organisation 3D de l'ADNv pourrait être perturbée par la présence d'inserts, et les caractères favorisant la transcription, comme la formation de boucles qui provoquent un rapprochement des SIN-LTRs, pourraient être déstabilisées. En effet, le provirus lentiviral a été décrit former une structure dépendante de la transcription en forme de boucle entre le 5'LTR ou le P_{CMV} et le signal poly(A) du 3'LTR, ce qui suggère que les signaux de maturation du pré-ARNm pourraient affecter la transcription génique (Perkins et al., 2008). Nous avons observé une corrélation entre l'expression de la *gfp* et les niveaux d'ARNm sans détecter de changements significatifs de la stabilité des ARNs (**Figure 7** #1), ce qui exclut des modifications de la stabilité de l'ARN. Enfin, des changements de la topologie de l'ARN pourraient également affecter la transcription/traduction. La région U3 du LTR est inclut dans le 3'UTR de l'ARNm et pourrait donc affecter son repliement, de façon

dépendante de l'orientation, en particulier quand les inserts sont larges. Nous proposons que ces modifications du repliement de l'ARN, ainsi que celui de l'ADN matrice, puissent influencer directement le niveau de transcription du transgène. Des études supplémentaires seront nécessaires pour mieux comprendre ce phénomène et peut-être identifier les éléments topologiques nécessaires à une transcription optimale du transgène, grâce notamment à la technique 3C (Perkins et al., 2008).

3/ Augmentation de l'expression du transgène grâce à l'insertion de cHS4

Nous avons cependant mis en évidence que l'insertion de l'insulateur cHS4 dans les SIN-LTRs peut augmenter l'expression du transgène uniquement lorsqu'il est cloné dans l'orientation antisens, en accord avec (Emery et al., 2000; Hanawa et al., 2009; Uchida et al., 2011; Yannaki et al., 2002). Cette augmentation de l'expression du transgène en présence de cHS4 en antisens est transitoire et/ou visible uniquement à partir des formes d'ADNni lorsque l'expression est contrôlée par le P_{CMV} (**Figures 3A et 4B** #1). A l'inverse, elle apparaît au cours du temps à partir des formes d'ADN intégré lors de la transduction de contructions contenant le P_{PGK} (**Figures 4A et 7A** #1).

Ces observation suggèrent que l'expression à partir du P_{PGK} est facilitée par l'activité barrière et/ou la terminaison de la transcription que possède la séquence cHS4 insérée en antisens (Hanawa et al., 2009). En accord avec (Hanawa et al., 2009), nous avons observé des quantités plus importantes d'ARNm *gfp* grâce à l'insertion au sein des SIN-LTRs de l'insulateur orienté en antisens (**Figure 7A** #1). En particulier, l'insertion de cHS4 orienté en antisens pourrait favoriser la mise en place d'une structure ADN/ARN favorable à la transcription à partir d'un promoteur faible comme le P_{PGK} en augmentant la terminaison de la transcription. Cette hypothèse est supportée par les études menées sur l'importance des éléments d'initiation et de terminaison de la transcription dans la formation d'une boucle ADN/ARN qui favorise l'expression du VIH-1 (Perkins et al., 2008) et du gène *BRCA-1* (Tan-Wong et al., 2008). Dans le cas d'une expression régulée par le P_{CMV}, cette structure ne serait que transitoire et/ou observée uniquement à partir des formes non-intégrées de l'ADNv. Ces données mettent l'accent sur le fait que cHS4 peut agir par plusieurs mécanismes, dont l'amélioration de la terminaison, et/ou ses activités *enhancer-blocker* et barrière. Ces mécanismes pourraient être non-exclusifs, comme décrit dans le cas de l'activité barrière de cHS4 (en antisens) et de sa capacité à augmenter la terminaison de la transcription (Hanawa et al., 2009).

L'ensemble de ces observations suggère que de multiples éléments influencent l'expression du transgène à partir d'un vecteur donné, dont (mais peut-être pas limité à) le type de vecteur et sa taille, la taille de l'insert et son orientation, le promoteur interne, la fonctionnalité de l'intégrase et le transgène lui-même.

Nos travaux mettent en lumière l'hypothèse de l'existence d'une structure formée entre l'ADNv et son transcrit dont la modification par insertion dans les SIN-LTRs pourrait influencer le niveau de transcription à partir des formes virales intégrées ou non. L'impact de l'insertion au sein des SIN-LTRs sur l'expression serait alors négatif pour la majorité des séquences insérées (au moins dans une orientation) ou positif dans le cas d'un élément spécifique comme l'insulateur cHS4 (en antisens). La topologie des éléments hétérologues insérés dans les vecteurs viraux nécessiterait d'être davantage explorer dans le *design* de vecteurs afin d'optimiser l'expression du transgène.

Notre étude a néanmoins permis de générer des vecteurs capables de moduler d'un facteur 0,3 à 1,8 l'expression de façon stable ou transitoire (en fonction de l'intégration ou non du transgène). L'utilisation des séquences que nous avons décrites réduire l'expression du transgène pourrait être envisagée pour limiter une éventuelle toxicité de l'expression transgénique, en complément de promoteurs tissu-spécifique ou inductibles par exemple.

II/ IMPACTS DES DETERMINANTS CELLULAIRES SUR L'EXPRESSION LENTIVIRALE

Les composants de la signalisation et la réparation des dommages de l'ADN, dont le complexe Ku, interagissent étroitement avec les rétrovirus à de nombreuses étapes de leur cycle de multiplication (*cf.*, Introduction, § IV.A/B). Nous avons donc analysé les impacts du complexe Ku sur l'expression du VIH-1 (**partie A**) et de vecteurs dérivés (**partie B**).

A/ IMPACTS DU STATUT EN KU SUR L'EXPRESSION DU VIH-1

1/ Impact positif du statut en Ku sur l'expression du VIH-1 à court-terme

a/ *La déplétion en Ku n'affecte pas les étapes pré-intégratives du cycle viral*

La déplétion stable ou transitoire en Ku au sein des cellules HCT 116 *Ku80*$^{+/-}$ n'a pas d'impact ni sur la formation des cercles à 2-LTRs, ni sur l'intégration du VIH-1 (comparé aux cellules WT, **Figure 4B, D** #2). De plus, en désaccord avec l'étude de (Zheng et al., 2011), nous avons observé une même efficacité de transduction de vecteurs lentiviraux (*i.e.*, XCD3

et SIN-PGK) produits en conditions proficientes ou déficientes en Ku au sein des cellules cibles HCT 116 WT et *Ku80*$^{+/-}$ (résultats non-montrés). Ces données indiquent que la présence de Ku, d'origine endogène ou exogène, n'a pas d'impact sur les étapes pré-intégratives du cycle lentiviral dans notre modèle.

L'ensemble de ces résultats est cohérent avec le rôle dispensable de Ku dans les étapes pré-intégratives du cycle du VIH-1 décrit dans les cellules humaines (Ariumi et al., 2005; Kilzer et al., 2003) ou issues d'autres mammifères [cheval (Kilzer et al., 2003) et rongeur (Baekelandt et al., 2000; Kilzer et al., 2003)]. Ils sont cependant en désaccord avec des observations précédentes sur le rôle positif de Ku sur la circularisation des espèces d'ADNv à 2-LTRs au sein de cellules de rongeurs [CHO (Li et al., 2001; Masson et al., 2007)] et humaines [CEM (Jeanson et al., 2002)]. Nos résultats sont également en contradiction avec le rôle décrit de Ku sur l'intégration du VIH-1 dans les cellules humaines [CEM (Jeanson et al., 2002; Waninger et al., 2004); C8166 (Zheng et al., 2011)].

Ces controverses pourraient s'expliquer par le fait qu'une moindre intégration du VIH-1 mise en évidence en absence de Ku puisse être le résultat de l'utilisation de virus réplicatifs, comme dans (Jeanson et al., 2002; Waninger et al., 2004; Zheng et al., 2011), qui s'exprimeraient peu et ne se multiplieraient donc pas avec la même efficacité au sein de cellules déplétées en Ku. Il est à noter que l'utilisation de vecteurs VIH-1 pseudotypés avec l'enveloppe VSV-G [dans notre étude et (Li et al., 2001; Masson et al., 2007)] pourrait également influencer le cycle de multiplication du VIH-1 (comparé aux virus réplicatifs). Toutefois, nous ne pouvons pas exclure une spécificité de chacun des types cellulaires étudiés à nécessiter le recrutement et/ou l'activation d'une voie de réponse aux dommages de l'ADN lors de l'infection lentivirale. Dans certains types cellulaires, l'intervention de celle-ci et/ou d'un facteur spécifique en réponse aux signaux d'infection pourrait aussi être compensée via l'utilisation de protéines pouvant le substituer.

b/ *Impact positif du statut de Ku sur la transcription du LTR du VIH-1*

Nous avons mis en évidence que la déplétion en Ku au sein des cellules HCT 116 a un impact négatif sur la transcription du VIH-1 48/72 h post-transduction (**Figures 2**, **3 et S3** #2). Cet effet n'est toutefois pas visible lorque l'expression de la *gfp* est sous contôle d'un promoteur hétérologue interne CMV ou PGK (**Figure 5A** #2). L'impact positif du statut en Ku sur l'expression du VIH-1 est donc spécifique du promoteur LTR. Ces résultats sont en accord avec l'étude de Waninger *et al.* qui décrivait une activation transcriptionnelle Tat-dépendante limitée du LTR du VIH-1 en absence de Ku80 (Waninger et al., 2004).

Cependant, dans notre étude, le statut en Ku a un effet stimulateur sur le LTR du VIH-1 indépendamment de la présence de protéines auxiliaires du VIH-1, dont Tat, et de la présence de TAR non-muté (**Figure 5A** #2). Ce résultat exclut un effet direct de Ku sur Tat ou indirect médié par la fixation de Ku sur TAR, comme suggéré dans (Kaczmarski and Khan, 1993; Waninger et al., 2004).

Nous avons montré que l'absence d'impact de la mutation de la KBS, séquence potentielle de fixation de Ku au sein du LTR du VIH-1 (Jeanson and Mouscadet, 2002), au sein d'un vecteur dérivé du VIH-1 sur l'effet positif de Ku sur l'expression du VIH-1 (**Figure 5A** #2). Cette donnée suggère une action positive indirecte de Ku sur le LTR sans exlure le possible recrutement de Ku (de façon non-spécifique) au niveau des extrémités de l'ADNv qui pourrait aussi expliquer la spécificité de Ku pour le LTR (*vs.* promoteurs internes). De plus, une prise en charge de l'ADNv par Ku pourrait influencer le site d'intégration et/ou médié des interactions avec d'autres facteurs qui régulent positivement la transcription, dont SP1, l'ARN pol II, Ets-1 ou DNA-PK (Bertinato et al., 2003; Choul-li et al., 2009; Tuteja and Tuteja, 2000; Tyagi et al., 2011). En particulier, il a récemment été décrit le recrutement de DNA-PK au niveau du LTR du VIH-1 qui interagit avec l'ARN pol II lors de la transcription du VIH-1 (Tyagi et al., 2011). D'autres facteurs qui favorisent l'activité du promoteur LTR et sont régulés positivement par Ku pourraient aussi être impliqués [*e.g.*, NF-κB (Kim, 2008)].

Enfin, l'impact positif du statut en Ku sur l'expression précoce du VIH-1 nécessite l'intégration de l'ADNv (**Figure 5C** #2), ce qui suggère que le processus d'intégration est, au moins dans notre modèle, intimement lié au ciblage de l'ADNv dans un environnement local qui favorise la transcription précoce du VIH-1 et/ou à l'établissement d'un tel environnement.

c/ *Régulations mutuelles entre Ku et p53 et impact sur l'expression du VIH-1*

Nous avons observé une régulation mutuelle des contenus en Ku et p53 (**Figures 3C et S1C** #2), comme précédemment décrit (Ghosh et al., 2007; Li et al., 2002). Celle-ci pourrait permettre de limiter les effets délétères causés par les dommages endogènes auxquels sont confrontées les cellules HCT 116 *Ku80*$^{+/-}$ (Li et al., 2002). Cette hypothèse est supportée par l'observation d'une augmentation de la quantité de lésions et de l'activation de la DDR médiée par p53 dans les cellules murines *Ku80*$^{-/-}$ (Holcomb et al., 2008).

De plus, la déplétion de p53 dans les cellules HCT 116 *Ku80*$^{+/-}$ permet de rétablir une efficacité de transduction du VIH-1 intermédiaire à celles observées dans les cellules WT et *Ku80*$^{+/-}$ (**Figure 3B** #2). Cependant, p53 a un effet positif sur la transduction du VIH-1 dans

un contexte WT (**Figure 3B** #2), en accord avec des études précédentes (Mukerjee et al., 2008; Pauls et al., 2006). L'efficacité de transduction du VIH-1 dans les cellules HCT 116 *Ku80*$^{+/-}$ déplétées en p53 résulterait donc plutôt d'une balance entre effet négatif de la déplétion en p53 et effet positif de l'augmentation en Ku sur la transduction du VIH-1 dans ces cellules. Cependant, nous ne pouvons pas exclure qu'il n'y ait pas d'impact d'une déplétion en Ku dans les cellules HCT 116 *Ku80*$^{+/-}$ du fait d'une mortalité accrue des cellules fortement déplétées en Ku (observable au microscope) et/ou du requièrement du peu de protéines Ku restantes pour le maintien de ses fonctions essentielles.

p53 est un effecteur majeur dans la signalisation en réponse aux dommages de l'ADN. De ce fait, la régulation réciproque des contenus en p53 et Ku et son impact sur l'expression du VIH-1 suggère l'importance d'une protéine p53 fonctionnelle dans les modèles d'études des rôles de la machinerie de réparation de l'ADN sur l'infection du VIH-1. Cependant, les autres études ayant identifié Ku comme facteur clé de la réplication du VIH-1 (Genovesio et al., 2011) ou, au moins, de son expression (Waninger et al., 2004) ont été menées au sein de cellules possédant une protéine p53 non-fonctionnelle. Notamment, les fonctions de p53 ont été inactivées par mutation(s), comme au sein de lignées lymphocytaires (*e.g.*, CEM, Jurkat; http://www-p53.iarc.fr) ou sont altérées par la présence d'une protéine virale [*e.g.*, E6 du papillomavirus au sein des cellules HeLa (Kessis et al., 1993)].

L'ensemble de nos résultats favoriserait donc plutôt un impact direct de Ku sur la transcription du VIH-1 au sein des cellules HCT 116, sans exclure une possible contribution de p53. Des expériences supplémentaires devront être conduites pour conclure quant au rôle de p53 dans la régulation de l'expression du VIH-1 en absence de Ku grâce notamment à la transduction du VIH-1 dans la lignée HCT 116 déficiente en p53 déplétée ou non en Ku.

2/ <u>Impact négatif du statut en Ku sur l'établissement de la latence du VIH-1</u>

A court-terme post-transduction, la diminution de l'expression du VIH-1 visible lors de l'haplodéplétion de Ku80 au sein des cellules HCT 116 est réversible par traitement cellulaire à la TSA (**Figure 6B** #2). Ainsi, lorsque le statut de Ku atteint une valeur critique (*i.e.*, inférieure ou égale à la moitié de la valeur normale), au moins la moitié des cellules HCT 116 transduites possèdent un provirus peu ou pas exprimé qui pourrait être associé à une ou des HDAC(s) (dont l'action est inhibée par la TSA).

L'impact négatif de la déplétion en Ku sur l'expression précoce du VIH-1 est corrélé à une quantité deux fois supérieure d'ADNv à moyen terme (10 jours) post-transduction au sein

des cellules *Ku80*$^{+/-}$ (**Figure 4B, C** #2). En accord avec les éléments discutés dans le paragraphe I.A.1, une moindre expression du VIH-1 dans les cellules déplétées en Ku diminuerait la proportion de cellules transduites contre-sélectionnées à moyen-terme. De plus, l'haplodéplétion en Ku est associée à une réactivation 1,5 fois plus importante des provirus latents (après traitement à la TSA ou au TNFα) dans une population mixte GFP^{+} et GFP^{-} de cellules transduites (**Figure 6A, D** #2).

Ces résultats sont cohérents avec nos observations précédentes montrant une stimulation de l'expression du VIH-1 après traitement au TNFα des cellules U1 qui était augmentée après déplétion en Ku (Jeanson and Mouscadet, 2002). Cet état de latence dépendrait donc de la disponibilité en NF-κB, comme décrit par (Duverger et al., 2009), mais est (ré)activable après traitement adapté (*e.g.*, TSA ou TNFα), en accord avec (Pearson et al., 2008). Nos données sont en accord avec la description du LTR du provirus latent interagissant (indirectement) avec des HDACs de classe I (*i.e.*, HDAC-1, -2 et -3) qui induiraient la formation d'une structure de la chromatine empêchant l'initiation de la transcription provirale (Colin and Van Lint, 2009).

3/ <u>Hypothèses de modes d'action de Ku</u>

L'ensemble de nos résultats supporte l'hypothèse que des interactions directes/indirectes entre l'ADN (pro)viral et Ku favoriseraient la transcription précoce du LTR du VIH-1 lorsque celui-ci est intégré. Ku pourrait stimuler la transcription du LTR, de façon indépendante de Tat, en interagissant directement avec le promoteur du VIH-1 [au niveau d'une séquence autre que la KBS putative décrite par (Jeanson and Mouscadet, 2002)] ou indirectement en recrutant des facteurs cellulaires qui régulent positivement l'expression du VIH-1 (*e.g.*, NF-κB) (*cf.*, § II.A.1 ci-dessus). Cette stimulation précoce induirait cependant une contre-sélection plus importante des cellules transduites à moyen terme post-transduction (**Figure 7** #2).

A l'inverse, la déplétion en Ku favoriserait la mise en place précoce d'une latence post-intégrative du VIH-1. Celle-ci pourrait être établit et/ou maintenue par (*i*) une intégration par défaut dans des régions peu exprimées du génome, soit une intégration silencieuse comme suggéré par (Jordan et al., 2003); (*ii*) un recrutement (par défaut) de facteurs cellulaires régulant négativement l'expression du VIH-1, dont des HDACs; (*iii*) un niveau insuffisant en facteurs stimulant la transcription du LTR (*e.g.*, NF-κB sous forme p65/p50); et/ou (*iv*) un *silencing* épigénétique causé, par exemple, par la déacétylation des histones associées au

LTR. *In fine*, ce *silencing* transcriptionnel du VIH-1 résulterait en une proportion plus importante de cellules contenant un provirus peu ou pas exprimé (mais réactivable), qui seraient moins sujettes au phénomène de contre-sélection (**Figures 6C et 7** #2).

Nos résultats ne nous ont pas permis d'identifier si l'action de Ku pouvait être engagée à une étape précédant l'expression virale et, si oui, laquelle. Son absence ne modifie pas l'efficacité d'intégration du VIH-1 dans notre modèle. Nous ne pouvons cependant pas exclure un rôle de Ku dans la prise en charge de l'ADNv et/ou dans le ciblage de son intégration. Les expériences de FISH (*Fluorescence In Situ Hybridization*) 3D que nous avons mené, en collaboration avec le laboratoire du Dr. Y. Vassetzky (Villejuif, France), pour tenter de répondre à cette question n'ont pas été fructueuses, faute d'optimisation expérimentale. Il serait cependant intéressant de les poursuivre, de façon couplée avec le pyroséquençage des sites d'intégration du VIH-1 en condition compétente ou déplétée en Ku. Ces approches permettraient de conclure quant à un éventuel rôle de Ku dans le ciblage de l'intégration du VIH-1 dans des régions chromatiniennes transcriptionnellement actives des cellules humaines. Elles pourraient être complétées par des expériences de ChIP (*Chromatin ImmunoPrecipitation*) qui permettraient d'identifier les partenaires de Ku et/ou interagissant avec le LTR dans un contexte haplodéplété en Ku.

L'ensemble de ces donnés a permis de démontrer que la déplétion de Ku qui stimule la transcription précoce du LTR du VIH-1 favoriserait la mise en place d'une latence post-intégrative au sein de cellules transduites moins susceptibles d'être contre-sélectionnées au cours du temps.

Notre étude pointe les éventuels effets secondaires d'approches actuellement menées contre les protéines de l'hôte impliquées dans la multiplication du VIH-1.

B/ IMPACT POSITIF DE LA DEPLETION EN KU SUR L'EXPRESSION TRANSGENIQUE A PARTIR D'UN VECTEUR LENTIVIRAL INSULE

Nous avons mis en évidence que la déplétion en Ku favorise spécifiquement l'expression de la *gfp* à partir du P_{PGK} déjà augmentée par la présence de cHS4 en antisens au sein des cellules HCT 116 *Ku80*$^{+/-}$(**Figure 34**).

Comme anticipé dans la discussion du paragraphe III.B de la section Résultats, nous proposons que PARP-1 puisse être impliquée positivement dans le contrôle de l'activité

barrière de cHS4 en antisens (Aker et al., 2010). En effet, Ku et PARP-1 sont en compétition pour la fixation des cassures double-brin de l'ADN (Ciccia and Elledge, 2010). De plus, la fixation de l'une ou l'autre de ces protéines détermine le choix de la voie de réparation activée [*i.e.*, HR ou NHEJ (Mladenov and Iliakis, 2011)]. Une hypothèse intéressante repose sur l'idée que le contexte haplodéplété en Ku pourrait favoriser la fixation de PARP-1 au niveau de l'insulateur cHS4 qui augmenterait l'activité barrière de ce dernier et, par conséquent, l'expression de la *gfp*. Il n'est cependant pas exclut une prise en charge des extrémités double-brin de l'ADNv par défaut par PARP-1 en condition de déplétion en Ku qui serait non-spécifique et pourrait expliquer l'augmentation progressive de l'expression de la *gfp*.

Cette hypothèse sera évaluée par la transduction des vecteurs SIN-PGK contenant ou non cHS4 en antisens, couplées à des expériences de ChIP en conditions compétentes ou déficientes pour PARP-1 au sein des cellules HCT 116 *Ku80*$^{+/-}$.

Ces observations suggèrent que le contexte cellulaire, dont la déplétion d'un facteur de la réparation de l'ADN, pourrait influencer l'impact de séquences insulatrices comme cHS4, ce qui est à prendre en considération lors du transfert de gène au sein de patients atteints de cancers ou de maladies génétiques spécifiques.

CONCLUSIONS

L'ensemble de ces travaux de thèse a permis de mettre en évidence que l'expression lentivirale est régulée par de nombreux paramètres, dont (*i*) le *design* du vecteur (*e.g.*, LTR *vs.* promoteur interne, intégrase fonctionnelle *vs.* mutée, absence *vs.* présence de séquences hétérologues au sein des SIN-LTRs et leur orientation), (*ii*) le type cellulaire transduit et son fond génétique (*e.g.*, niveaux d'expression en Ku et p53), et (*iii*) le temps écoulé depuis la transduction, comme résumé dans le **tableau 9** ci-dessous.

Tableau 9 : Bilan des différents paramètres du *design* des vecteurs ou du contexte cellulaire qui influencent l'expression trangénique en fonction du vecteur utilisé. Il est à noter que « court et moyen terme » se réfèrent à 2 et 10 jours post-transduction dans notre étude. Abréviations : as, antisens ; (+), effet positif ; (-), effet négatif ; Nd, non-déterminé.

		Vecteur SIN-PGK	Vecteur SIN-CMV	Vecteur LTR VIH-1
Caractéristiques du vecteur		Promoteur interne de type CGI, SIN-LTRs	Promoteur interne de type TATA-*box*, SIN-LTRs	Promoteur LTR à TATA-*box*, Expression des protéines du VIH-1
Promoeteur	Extinction/contre-sélection	Non	Oui (extinction)	Oui (contre-sélection associée à l'expression de Vpr)
	Réactivation	Non	Nd	Oui
	Expression à partir des formes d'ADNv non-intégré	Faible	Oui (équivalente à celle à partir d'ADN intégré)	Oui
	Impact du type cellulaire sur l'expression	Non	Nd	Oui
Design des SIN-LTRs	Impact du clonage de larges inserts (*e.g.*, DS18) au sein des SIN-LTRs sur l'expression	-	-	Nd
	Impact de l'insulateur cHS4 as sur l'expression du provirus	+	+	Nd
	Impact de l'insulateur cHS4 as sur l'expression des formes d'ADNv non intégré	-	+	Nd
Contexte cellulaire	Impact de la déplétion en Ku (en absence de cHS4 as) à court et moyen terme	Pas d'impact	Pas d'impact	- sur l'expression (et la contre-sélection), + sur la latence
	Impact de la déplétion en Ku (en présence de cHS4 as) sur l'expression	+	Pas d'impact	Nd
	Impact de la déplétion en p53 sur la transduction	Nd	Nd	- en condition WT, + en condition *Ku80*$^{+/-}$

L'anayse du *design* de vecteurs dérivés du VIH-1 sur l'expression transgénique nous a permis de montrer que l'expression à partir de promoteurs viraux fortement régulés, comme le CMV, est variable en fonction du type cellulaire transduit et soumise à extinction. A l'inverse, un promoteur décrit comme constitutif et ubiquitaire tel que le P_{PGK} a une activité indépendante du contexte cellulaire transduit et stable au cours du temps. En absence d'intégration de l'ADNv, la force du P_{CMV} lui confère la capacité d'être activé alors que le P_{PGK} doit être intégré pour être fonctionnel.

Néanmoins, la présence de séquences hétérologues dans les SIN-LTRs atténue l'expression transgénique contrôlée par le P_{PGK} ainsi que le P_{CMV}. Ce type d'insertion pourrait affecter la formation d'une structure ADN/ARN décrite comme favorisant la transcription. Toutefois, le clonage de séquences spécifiques dans les SIN-LTRs, comme l'insulteur cHS4 en orientation antisens, peut influencer positivement l'expression à partir du P_{PGK} et du P_{CMV} d'un transgène intégré (et non-intégré dans le cas de ce dernier). Cet effet est visible à court- et moyen-terme de façon dépendante du type cellulaire transduit et, dans le cas du P_{PGK}, de la présence de facteurs cellulaires comme Ku.

Nous avons également démontré que le contexte cellulaire de transduction, et en particulier, le statut de la machinerie de réparation de l'ADN, influence l'expression du LTR du VIH-1. Ainsi, la déplétion de Ku au sein des cellules HCT 116 diminue l'expression précoce du VIH-1 intégré de façon spécifique du LTR et indépendante de Tat. De plus, la déplétion en Ku dans ces cellules favoriserait un état de latence transcriptionnelle (re)activable du provirus qui limiterait la contre-sélection des cellules transduites. A l'inverse, en présence d'un niveau normal en Ku, les cellules exprimant le VIH-1 seraient contre-sélectionnées au cours du temps de façon dépendante de l'expression de Vpr.

Notre étude confirme l'importance de certains déterminants viraux et cellulaires dans la régulation de l'expression lentivirale. Elle met notamment l'accent sur la nécessité d'étudier de façon approfondie les éléments topologiques hétérologues indispensables à une expression optimale en vu de l'amélioration du *design* des vecteurs lentiviraux. De plus, nos travaux suggèrent que le contexte cellulaire pourrait influencer l'impact de séquences insulatrices comme cHS4, ce qui est à prendre en considération lors du transfert de gène au sein de patients atteints de cancers ou de maladies génétiques spécifiques. Enfin, nos données mettent en avant les éventuels effets secondaires que pourraient avoir les approches actuellement menées visant à cibler les protéines de l'hôte impliquées dans l'expression du VIH-1.

BIBLIOGRAPHIE

Aiken, C. (1997) Pseudotyping human immunodeficiency virus type 1 (HIV-1) by the glycoprotein of vesicular stomatitis virus targets HIV-1 entry to an endocytic pathway and suppresses both the requirement for Nef and the sensitivity to cyclosporin A. *J Virol*, **71**, 5871-5877.

Aker, M., Bomsztyk, K. and Emery, D.W. (2010) Poly(ADP-ribose) polymerase-1 (PARP-1) contributes to the barrier function of a vertebrate chromatin insulator. *J Biol Chem*, **285**, 37589-37597.

Aker, M., Tubb, J., Groth, A.C., Bukovsky, A.A., Bell, A.C., Felsenfeld, G., Kiem, H.P., Stamatoyannopoulos, G. and Emery, D.W. (2007) Extended core sequences from the cHS4 insulator are necessary for protecting retroviral vectors from silencing position effects. *Hum Gene Ther*, **18**, 333-343.

Allers, K., Hutter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E. and Schneider, T. (2011) Evidence for the cure of HIV infection by CCR5Delta32/Delta32 stem cell transplantation. *Blood*, **117**, 2791-2799.

Andersen, J.L., Le Rouzic, E. and Planelles, V. (2008) HIV-1 Vpr: mechanisms of G2 arrest and apoptosis. *Exp Mol Pathol*, **85**, 2-10.

Andersen, J.L., Zimmerman, E.S., DeHart, J.L., Murala, S., Ardon, O., Blackett, J., Chen, J. and Planelles, V. (2005) ATR and GADD45alpha mediate HIV-1 Vpr-induced apoptosis. *Cell Death Differ*, **12**, 326-334.

Aravind, L. and Koonin, E.V. (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res*, **11**, 1365-1374.

Arhel, N. (2010) Revisiting HIV-1 uncoating. *Retrovirology*, **7**, 96.

Ariumi, Y. and Trono, D. (2006) Ataxia-telangiectasia-mutated (ATM) protein can enhance human immunodeficiency virus type 1 replication by stimulating Rev function. *J Virol*, **80**, 2445-2452.

Ariumi, Y., Turelli, P., Masutani, M. and Trono, D. (2005) DNA damage sensors ATM, ATR, DNA-PKcs, and PARP-1 are dispensable for human immunodeficiency virus type 1 integration. *J Virol*, **79**, 2973-2978.

Arumugam, P.I., Urbinati, F., Velu, C.S., Higashimoto, T., Grimes, H.L. and Malik, P. (2009) The 3' region of the chicken hypersensitive site-4 insulator has properties similar to its core and is required for full insulator activity. *PLoS One*, **4**, e6995.

Baekelandt, V., Claeys, A., Cherepanov, P., De Clercq, E., De Strooper, B., Nuttin, B. and Debyser, Z. (2000) DNA-Dependent protein kinase is not required for efficient lentivirus integration. *J Virol*, **74**, 11278-11285.

Baiker, A., Maercker, C., Piechaczek, C., Schmidt, S.B., Bode, J., Benham, C. and Lipps, H.J. (2000) Mitotic stability of an episomal vector containing a human scaffold/matrix-attached region is provided by association with nuclear matrix. *Nat Cell Biol*, **2**, 182-184.

Banasik, M.B. and McCray, P.B., Jr. (2010) Integrase-defective lentiviral vectors: progress and applications. *Gene Ther*, **17**, 150-157.

Bannert, N. and Kurth, R. (2004) Retroelements and the human genome: new perspectives on an old relation. *Proc Natl Acad Sci U S A*, **101 Suppl 2**, 14572-14579.

Barboric, M., Nissen, R.M., Kanazawa, S., Jabrane-Ferrat, N. and Peterlin, B.M. (2001) NF-kappaB binds P-TEFb to stimulate transcriptional elongation by RNA polymerase II. *Mol Cell*, **8**, 327-337.

Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vezinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, **220**, 868-871.

Bastone, P. and Lochelt, M. (2004) Kinetics and characteristics of replication-competent revertants derived from self-inactivating foamy virus vectors. *Gene Ther*, **11**, 465-473.

Baumann, M., Pontiller, J. and Ernst, W. (2010) Structure and basal transcription complex of RNA polymerase II core promoters in the mammalian genome: an overview. *Mol Biotechnol*, **45**, 241-247.

Bayer, M., Kantor, B., Cockrell, A., Ma, H., Zeithaml, B., Li, X., McCown, T. and Kafri, T. (2008) A large U3 deletion causes increased in vivo expression from a nonintegrating lentiviral vector. *Mol Ther*, **16**, 1968-1976.

Beck-Engeser, G.B., Eilat, D., Harrer, T., Jack, H.M. and Wabl, M. (2009) Early onset of autoimmune disease by the retroviral integrase inhibitor raltegravir. *Proc Natl Acad Sci U S A*, **106**, 20865-20870.

Bennasser, Y., Le, S.Y., Benkirane, M. and Jeang, K.T. (2005) Evidence that HIV-1 encodes an siRNA and a suppressor of RNA silencing. *Immunity*, **22**, 607-619.

Bennasser, Y., Yeung, M.L. and Jeang, K.T. (2006) HIV-1 TAR RNA subverts RNA interference in transfected cells through sequestration of TAR RNA-binding protein, TRBP. *J Biol Chem*, **281**, 27674-27678.

Bergeron, L. and Sodroski, J. (1992) Dissociation of unintegrated viral DNA accumulation from single-cell lysis induced by human immunodeficiency virus type 1. *J Virol*, **66**, 5777-5787.

Berkhout, B., Arts, K. and Abbink, T.E. (2011) Ribosomal scanning on the 5'-untranslated region of the human immunodeficiency virus RNA genome. *Nucleic Acids Res*, **39**, 5232-5244.

Berro, R., Kehn, K., de la Fuente, C., Pumfery, A., Adair, R., Wade, J., Colberg-Poley, A.M., Hiscott, J. and Kashanchi, F. (2006) Acetylated Tat regulates human immunodeficiency virus type 1 splicing through its interaction with the splicing regulator p32. *J Virol*, **80**, 3189-3204.

Bertinato, J., Tomlinson, J.J., Schild-Poulter, C. and Hache, R.J. (2003) Evidence implicating Ku antigen as a structural factor in RNA polymerase II-mediated transcription. *Gene*, **302**, 53-64.

Beutler, E. (2007) PGK deficiency. *Br J Haematol*, **136**, 3-11.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., Kuehn, M.S., Taylor, C.M., Neph, S., Koch, C.M., Asthana, S., Malhotra, A., Adzhubei, I., Greenbaum, J.A., Andrews, R.M., Flicek, P., Boyle, P.J., Cao, H., Carter, N.P., Clelland, G.K., Davis, S., Day, N., Dhami, P., Dillon, S.C., Dorschner, M.O., Fiegler, H., Giresi, P.G., Goldy, J., Hawrylycz, M., Haydock, A., Humbert, R., James, K.D., Johnson, B.E., Johnson, E.M., Frum, T.T., Rosenzweig, E.R., Karnani, N., Lee, K., Lefebvre, G.C., Navas, P.A., Neri, F., Parker, S.C., Sabo, P.J., Sandstrom, R., Shafer, A., Vetrie, D., Weaver, M., Wilcox, S., Yu, M., Collins, F.S., Dekker, J., Lieb, J.D., Tullius, T.D., Crawford, G.E., Sunyaev, S., Noble, W.S., Dunham, I., Denoeud, F., Reymond, A., Kapranov, P., Rozowsky, J., Zheng, D., Castelo, R., Frankish, A., Harrow, J., Ghosh, S., Sandelin, A., Hofacker, I.L., Baertsch, R., Keefe, D., Dike, S., Cheng, J., Hirsch, H.A., Sekinger, E.A., Lagarde, J., Abril, J.F., Shahab, A., Flamm, C., Fried, C., Hackermuller, J., Hertel, J., Lindemeyer, M., Missal, K., Tanzer, A., Washietl, S., Korbel, J., Emanuelsson, O., Pedersen, J.S., Holroyd, N., Taylor, R., Swarbreck, D., Matthews, N., Dickson, M.C., Thomas, D.J., Weirauch, M.T., Gilbert, J., Drenkow, J., Bell, I., Zhao, X., Srinivasan, K.G., Sung, W.K., Ooi, H.S., Chiu, K.P., Foissac, S., Alioto, T., Brent, M., Pachter, L., Tress, M.L., Valencia, A., Choo, S.W., Choo, C.Y., Ucla, C., Manzano, C., Wyss, C., Cheung, E., Clark, T.G., Brown, J.B., Ganesh, M., Patel, S., Tammana, H., Chrast, J., Henrichsen, C.N., Kai, C., Kawai, J., Nagalakshmi, U., Wu, J., Lian, Z., Lian, J., Newburger, P., Zhang, X., Bickel, P., Mattick, J.S., Carninci, P., Hayashizaki, Y., Weissman, S., Hubbard, T., Myers, R.M., Rogers, J., Stadler, P.F., Lowe, T.M., Wei, C.L., Ruan, Y., Struhl, K., Gerstein, M., Antonarakis, S.E., Fu, Y., Green, E.D., Karaoz, U., Siepel, A., Taylor, J., Liefer, L.A., Wetterstrand, K.A., Good, P.J., Feingold, E.A., Guyer, M.S., Cooper, G.M., Asimenos, G., Dewey, C.N., Hou, M., Nikolaev, S., Montoya-Burgos, J.I., Loytynoja, A., Whelan, S., Pardi, F., Massingham, T., Huang, H., Zhang, N.R., Holmes, I., Mullikin, J.C., Ureta-Vidal, A., Paten, B., Seringhaus, M., Church, D., Rosenbloom, K., Kent, W.J., Stone, E.A., Batzoglou, S., Goldman, N., Hardison, R.C., Haussler, D., Miller, W., Sidow, A., Trinklein, N.D., Zhang, Z.D., Barrera, L., Stuart, R., King, D.C., Ameur, A., Enroth, S., Bieda, M.C., Kim, J., Bhinge, A.A., Jiang, N., Liu, J., Yao, F., Vega, V.B., Lee, C.W., Ng, P., Shahab, A., Yang, A., Moqtaderi, Z., Zhu, Z., Xu, X., Squazzo, S., Oberley, M.J., Inman, D., Singer, M.A., Richmond, T.A., Munn, K.J., Rada-Iglesias, A., Wallerman, O., Komorowski, J., Fowler, J.C., Couttet, P., Bruce, A.W., Dovey, O.M., Ellis, P.D., Langford, C.F., Nix, D.A., Euskirchen, G., Hartman, S., Urban, A.E., Kraus, P., Van Calcar, S., Heintzman, N., Kim, T.H., Wang, K., Qu, C., Hon, G., Luna, R., Glass, C.K., Rosenfeld, M.G., Aldred, S.F., Cooper, S.J., Halees, A., Lin, J.M., Shulha,

H.P., Zhang, X., Xu, M., Haidar, J.N., Yu, Y., Ruan, Y., Iyer, V.R., Green, R.D., Wadelius, C., Farnham, P.J., Ren, B., Harte, R.A., Hinrichs, A.S., Trumbower, H., Clawson, H., Hillman-Jackson, J., Zweig, A.S., Smith, K., Thakkapallayil, A., Barber, G., Kuhn, R.M., Karolchik, D., Armengol, L., Bird, C.P., de Bakker, P.I., Kern, A.D., Lopez-Bigas, N., Martin, J.D., Stranger, B.E., Woodroffe, A., Davydov, E., Dimas, A., Eyras, E., Hallgrimsdottir, I.B., Huppert, J., Zody, M.C., Abecasis, G.R., Estivill, X., Bouffard, G.G., Guan, X., Hansen, N.F., Idol, J.R., Maduro, V.V., Maskeri, B., McDowell, J.C., Park, M., Thomas, P.J., Young, A.C., Blakesley, R.W., Muzny, D.M., Sodergren, E., Wheeler, D.A., Worley, K.C., Jiang, H., Weinstock, G.M., Gibbs, R.A., Graves, T., Fulton, R., Mardis, E.R., Wilson, R.K., Clamp, M., Cuff, J., Gnerre, S., Jaffe, D.B., Chang, J.L., Lindblad-Toh, K., Lander, E.S., Koriabine, M., Nefedov, M., Osoegawa, K., Yoshinaga, Y., Zhu, B. and de Jong, P.J. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799-816.

Blazkova, J., Murray, D., Justement, J.S., Funk, E.K., Nelson, A.K., Moir, S., Chun, T.W. and Fauci, A.S. (2012) Paucity of HIV DNA methylation in latently infected, resting CD4+ T cells from infected individuals receiving antiretroviral therapy. *J Virol*, **86**, 5390-5392.

Blazkova, J., Trejbalova, K., Gondois-Rey, F., Halfon, P., Philibert, P., Guiguen, A., Verdin, E., Olive, D., Van Lint, C., Hejnar, J. and Hirsch, I. (2009) CpG methylation controls reactivation of HIV from latency. *PLoS Pathog*, **5**, e1000554.

Boese, A., Sommer, P., Holzer, D., Maier, R. and Nehrbass, U. (2009) Integrase interactor 1 (Ini1/hSNF5) is a repressor of basal human immunodeficiency virus type 1 promoter activity. *J Gen Virol*, **90**, 2503-2512.

Bos, T.J., De Bruyne, E., Van Lint, S., Heirman, C. and Vanderkerken, K. (2010) Large double copy vectors are functional but show a size-dependent decline in transduction efficiency. *J Biotechnol*, **150**, 37-40.

Boulanger, M.C., Liang, C., Russell, R.S., Lin, R., Bedford, M.T., Wainberg, M.A. and Richard, S. (2005) Methylation of Tat by PRMT6 regulates human immunodeficiency virus type 1 gene expression. *J Virol*, **79**, 124-131.

Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A. and Cedar, H. (1994) Sp1 elements protect a CpG island from de novo methylation. *Nature*, **371**, 435-438.

Brass, A.L., Dykxhoorn, D.M., Benita, Y., Yan, N., Engelman, A., Xavier, R.J., Lieberman, J. and Elledge, S.J. (2008) Identification of host proteins required for HIV infection through a functional genomic screen. *Science*, **319**, 921-926.

Bray, M., Prasad, S., Dubay, J.W., Hunter, E., Jeang, K.T., Rekosh, D. and Hammarskjold, M.L. (1994) A small element from the Mason-Pfizer monkey virus genome makes human immunodeficiency virus type 1 expression and replication Rev-independent. *Proc Natl Acad Sci U S A*, **91**, 1256-1260.

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

Brown, P.O., Bowerman, B., Varmus, H.E. and Bishop, J.M. (1989) Retroviral integration: structure of the initial covalent product and its precursor, and a role for the viral IN protein. *Proc Natl Acad Sci U S A*, **86**, 2525-2529.

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

Bukrinsky, M.I., Sharova, N., Dempsey, M.P., Stanwick, T.L., Bukrinskaya, A.G., Haggerty, S. and Stevenson, M. (1992) Active nuclear import of human immunodeficiency virus type 1 preintegration complexes. *Proc Natl Acad Sci U S A*, **89**, 6580-6584.

Bukrinsky, M.I., Sharova, N., McDonald, T.L., Pushkarskaya, T., Tarpley, W.G. and Stevenson, M. (1993) Association of integrase, matrix, and reverse transcriptase antigens of human immunodeficiency virus type 1 with viral nucleic acids following acute infection. *Proc Natl Acad Sci U S A*, **90**, 6125-6129.

Bukrinsky, M.I., Stanwick, T.L., Dempsey, M.P. and Stevenson, M. (1991) Quiescent T lymphocytes as an inducible virus reservoir in HIV-1 infection. *Science*, **254**, 423-427.

Burns, J.C., Friedmann, T., Driever, W., Burrascano, M. and Yee, J.K. (1993) Vesicular stomatitis virus G glycoprotein pseudotyped retroviral vectors: concentration to very high titer and efficient gene transfer into mammalian and nonmammalian cells. *Proc Natl Acad Sci U S A*, **90**, 8033-8037.

Bushman, F., Lewinski, M., Ciuffi, A., Barr, S., Leipzig, J., Hannenhalli, S. and Hoffmann, C. (2005) Genome-wide analysis of retroviral DNA integration. *Nat Rev Microbiol*, **3**, 848-858.

Bushman, F.D., Malani, N., Fernandes, J., D'Orso, I., Cagney, G., Diamond, T.L., Zhou, H., Hazuda, D.J., Espeseth, A.S., Konig, R., Bandyopadhyay, S., Ideker, T., Goff, S.P., Krogan, N.J., Frankel, A.D., Young, J.A. and Chanda, S.K. (2009) Host cell factors in HIV replication: meta-analysis of genome-wide studies. *PLoS Pathog*, **5**, e1000437.

Butler, J.E. and Kadonaga, J.T. (2002) The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes Dev*, **16**, 2583-2592.

Callen, B.P., Shearwin, K.E. and Egan, J.B. (2004) Transcriptional interference between convergent promoters caused by elongation over the promoter. *Mol Cell*, **14**, 647-656.

Calzado, M.A., Sancho, R. and Munoz, E. (2004) Human immunodeficiency virus type 1 Tat increases the expression of cleavage and polyadenylation specificity factor 73-kilodalton subunit modulating cellular and viral expression. *J Virol*, **78**, 6846-6854.

Cara, A., Cereseto, A., Lori, F. and Reitz, M.S., Jr. (1996) HIV-1 protein expression from synthetic circles of DNA mimicking the extrachromosomal forms of viral DNA. *J Biol Chem*, **271**, 5393-5397.

Carteau, S., Hoffmann, C. and Bushman, F. (1998) Chromosome structure and human immunodeficiency virus type 1 cDNA integration: centromeric alphoid repeats are a disfavored target. *J Virol*, **72**, 4005-4014.

Cattoglio, C., Facchini, G., Sartori, D., Antonelli, A., Miccio, A., Cassani, B., Schmidt, M., von Kalle, C., Howe, S., Thrasher, A.J., Aiuti, A., Ferrari, G., Recchia, A. and Mavilio, F. (2007) Hot spots of retroviral integration in human CD34+ hematopoietic cells. *Blood*, **110**, 1770-1778.

Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., Cavallesco, R., Gillet-Legrand, B., Caccavelli, L., Sgarra, R., Maouche-Chretien, L., Bernaudin, F., Girot, R., Dorazio, R., Mulder, G.J., Polack, A., Bank, A., Soulier, J., Larghero, J., Kabbara, N., Dalle, B., Gourmel, B., Socie, G., Chretien, S., Cartier, N., Aubourg, P., Fischer, A., Cornetta, K., Galacteros, F., Beuzard, Y., Gluckman, E., Bushman, F., Hacein-Bey-Abina, S. and Leboulch, P. (2010) Transfusion independence and HMGA2 activation after gene therapy of human beta-thalassaemia. *Nature*, **467**, 318-322.

Cha, S., Lim, C., Lee, J.Y., Song, Y.J., Park, J., Choe, J. and Seo, T. (2010) DNA-PK/Ku complex binds to latency-associated nuclear antigen and negatively regulates Kaposi's sarcoma-associated herpesvirus latent replication. *Biochem Biophys Res Commun*, **394**, 934-939.

Chen, C. and Okayama, H. (1987) High-efficiency transformation of mammalian cells by plasmid DNA. *Mol Cell Biol*, **7**, 2745-2752.

Chen, C.C., Yang, Y.C., Wang, W.H., Chen, C.S. and Chang, L.K. (2011) Enhancement of Zta-activated lytic transcription of Epstein-Barr virus by Ku80. *J Gen Virol*, **92**, 661-668.

Chen, H. and Engelman, A. (1998) The barrier-to-autointegration protein is a host factor for HIV type 1 integration. *Proc Natl Acad Sci U S A*, **95**, 15270-15274.

Chen, R., Le Rouzic, E., Kearney, J.A., Mansky, L.M. and Benichou, S. (2004) Vpr-mediated incorporation of UNG2 into HIV-1 particles is required to modulate the virus mutation rate and for replication in macrophages. *J Biol Chem*, **279**, 28419-28425.

Chipitsyna, G., Sawaya, B.E., Khalili, K. and Amini, S. (2006) Cooperativity between Rad51 and C/EBP family transcription factors modulates basal and Tat-induced activation of the HIV-1 LTR in astrocytes. *J Cell Physiol*, **207**, 605-613.

Chiu, Y.L., Ho, C.K., Saha, N., Schwer, B., Shuman, S. and Rana, T.M. (2002) Tat stimulates cotranscriptional capping of HIV mRNA. *Mol Cell*, **10**, 585-597.

Choi, Y.K., Nash, K., Byrne, B.J., Muzyczka, N. and Song, S. (2010) The effect of DNA-dependent protein kinase on adeno-associated virus replication. *PLoS One*, **5**, e15073.

Choul-li, S., Drobecq, H. and Aumercier, M. (2009) DNA-dependent protein kinase is a novel interaction partner for Ets-1 isoforms. *Biochem Biophys Res Commun*, **390**, 839-844.

Christ, F., Thys, W., De Rijck, J., Gijsbers, R., Albanese, A., Arosio, D., Emiliani, S., Rain, J.C., Benarous, R., Cereseto, A. and Debyser, Z. (2008) Transportin-SR2 imports HIV into the nucleus. *Curr Biol*, **18**, 1192-1202.

Chung, J.H., Whiteley, M. and Felsenfeld, G. (1993) A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. *Cell*, **74**, 505-514.

Ciccia, A. and Elledge, S.J. (2010) The DNA damage response: making it safe to play with knives. *Mol Cell*, **40**, 179-204.

Ciuffi, A. (2008) Mechanisms governing lentivirus integration site selection. *Curr Gene Ther*, **8**, 419-429.

Ciuffi, A., Diamond, T.L., Hwang, Y., Marshall, H.M. and Bushman, F.D. (2006) Modulating target site selection during human immunodeficiency virus DNA integration in vitro with an engineered tethering factor. *Hum Gene Ther*, **17**, 960-967.

Ciuffi, A., Llano, M., Poeschla, E., Hoffmann, C., Leipzig, J., Shinn, P., Ecker, J.R. and Bushman, F. (2005) A role for LEDGF/p75 in targeting HIV DNA integration. *Nat Med*, **11**, 1287-1289.

Clavel, F., Guetard, D., Brun-Vezinet, F., Chamaret, S., Rey, M.A., Santos-Ferreira, M.O., Laurent, A.G., Dauguet, C., Katlama, C., Rouzioux, C. and et al. (1986) Isolation of a new human retrovirus from West African patients with AIDS. *Science*, **233**, 343-346.

Clumeck, N. (1993) Current use of anti-HIV drugs in AIDS. *J Antimicrob Chemother*, **32 Suppl A**, 133-138.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H., Vogt, P. and et al. (1986) What to call the AIDS virus? *Nature*, **321**, 10.

Coiras, M., Lopez-Huertas, M.R., Perez-Olmeda, M. and Alcami, J. (2009) Understanding HIV-1 latency provides clues for the eradication of long-term reservoirs. *Nat Rev Microbiol*, **7**, 798-812.

Coiras, M., Lopez-Huertas, M.R., Rullas, J., Mittelbrunn, M. and Alcami, J. (2007) Basal shuttle of NF-kappaB/I kappaB alpha in resting T lymphocytes regulates HIV-1 LTR dependent expression. *Retrovirology*, **4**, 56.

Colin, L. and Van Lint, C. (2009) Molecular control of HIV-1 postintegration latency: implications for the development of new therapeutic strategies. *Retrovirology*, **6**, 111.

Corbeau, P. (2008) Interfering RNA and HIV: reciprocal interferences. *PLoS Pathog*, **4**, e1000162.

Coroadinha, A.S., Gama-Norton, L., Amaral, A.I., Hauser, H., Alves, P.M. and Cruz, P.E. (2010) Production of retroviral vectors: review. *Curr Gene Ther*, **10**, 456-473.

Cortez, K.J. and Maldarelli, F. (2011) Clinical management of HIV drug resistance. *Viruses*, **3**, 347-378.

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, E.L., Rethwilm, A., Fleury, H., Connell, P.P., Sung, P., Delelis, O., Andreola, M.L. and Parissi, V. (2012) Stimulation of the human RAD51 nucleofilament restricts HIV-1 integration in vitro and in infected cells. *J Virol*, **86**, 513-526.

Cossarizza, A. (2008) Apoptosis and HIV infection: about molecules and genes. *Curr Pharm Des*, **14**, 237-244.

Coull, J.J., Romerio, F., Sun, J.M., Volker, J.L., Galvin, K.M., Davie, J.R., Shi, Y., Hansen, U. and Margolis, D.M. (2000) The human factors YY1 and LSF repress the human immunodeficiency virus type 1 long terminal repeat via recruitment of histone deacetylase 1. *J Virol*, **74**, 6790-6799.

Crampton, N., Bonass, W.A., Kirkham, J., Rivetti, C. and Thomson, N.H. (2006) Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. *Nucleic Acids Res*, **34**, 5416-5425.

Cron, R.Q., Bartz, S.R., Clausell, A., Bort, S.J., Klebanoff, S.J. and Lewis, D.B. (2000) NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells. *Clin Immunol*, **94**, 179-191.

Cummins, N.W. and Badley, A.D. (2010) Mechanisms of HIV-associated lymphocyte apoptosis: 2010. *Cell Death Dis*, **1**, e99.

D'Silva, I., Pelletier, J.D., Lagueux, J., D'Amours, D., Chaudhry, M.A., Weinfeld, M., Lees-Miller, S.P. and Poirier, G.G. (1999) Relative affinities of poly(ADP-ribose) polymerase and DNA-dependent protein kinase for DNA strand interruptions. *Biochim Biophys Acta*, **1430**, 119-126.

Daly, E.B., Wind, T., Jiang, X.M., Sun, L. and Hogg, P.J. (2004) Secretion of phosphoglycerate kinase from tumour cells is controlled by oxygen-sensing hydroxylases. *Biochim Biophys Acta*, **1691**, 17-22.

Daniel, R., Greger, J.G., Katz, R.A., Taganov, K.D., Wu, X., Kappes, J.C. and Skalka, A.M. (2004) Evidence that stable retroviral transduction and cell survival following DNA integration depend on components of the nonhomologous end joining repair pathway. *J Virol*, **78**, 8573-8581.

Daniel, R., Kao, G., Taganov, K., Greger, J.G., Favorova, O., Merkel, G., Yen, T.J., Katz, R.A. and Skalka, A.M. (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.

Daniel, R., Katz, R.A., Merkel, G., Hittle, J.C., Yen, T.J. and Skalka, A.M. (2001) Wortmannin potentiates integrase-mediated killing of lymphocytes and reduces the efficiency of stable transduction by retroviruses. *Mol Cell Biol*, **21**, 1164-1172.

Daniel, R., Katz, R.A. and Skalka, A.M. (1999) A role for DNA-PK in retroviral DNA integration. *Science*, **284**, 644-647.

De Zio, D., Bordi, M., Tino, E., Lanzuolo, C., Ferraro, E., Mora, E., Ciccosanti, F., Fimia, G.M., Orlando, V. and Cecconi, F. (2011) The DNA repair complex Ku70/86 modulates Apaf1 expression upon DNA damage. *Cell Death Differ*, **18**, 516-527.

Deaton, A.M. and Bird, A. (2011) CpG islands and the regulation of transcription. *Genes Dev*, **25**, 1010-1022.

Dehart, J.L., Andersen, J.L., Zimmerman, E.S., Ardon, O., An, D.S., Blackett, J., Kim, B. and Planelles, V. (2005) The ataxia telangiectasia-mutated and Rad3-related protein is dispensable for retroviral integration. *J Virol*, **79**, 1389-1396.

Delelis, O., Parissi, V., Leh, H., Mbemba, G., Petit, C., Sonigo, P., Deprez, E. and Mouscadet, J.F. (2007) Efficient and specific internal cleavage of a retroviral palindromic DNA sequence by tetrameric HIV-1 integrase. *PLoS One*, **2**, e608.

Deng, W. and Blobel, G.A. (2010) Do chromatin loops provide epigenetic gene expression states? *Curr Opin Genet Dev*, **20**, 548-554.

Dingwall, C., Ernberg, I., Gait, M.J., Green, S.M., Heaphy, S., Karn, J., Lowe, A.D., Singh, M. and Skinner, M.A. (1990) HIV-1 tat protein stimulates transcription by binding to a U-rich bulge in the stem of the TAR RNA structure. *Embo J*, **9**, 4145-4153.

Doherty, A.J., Jackson, S.P. and Weller, G.R. (2001) Identification of bacterial homologues of the Ku DNA repair proteins. *FEBS Lett*, **500**, 186-188.

Donahue, D.A., Kuhl, B.D., Sloan, R.D. and Wainberg, M.A. (2012) The viral protein Tat can inhibit the establishment of HIV-1 latency. *J Virol*, **86**, 3253-3263.

Downs, J.A. and Jackson, S.P. (2004) A means to a DNA end: the many roles of Ku. *Nat Rev Mol Cell Biol*, **5**, 367-378.

du Chene, I., Basyuk, E., Lin, Y.L., Triboulet, R., Knezevich, A., Chable-Bessia, C., Mettling, C., Baillat, V., Reynes, J., Corbeau, P., Bertrand, E., Marcello, A., Emiliani, S., Kiernan, R. and Benkirane, M. (2007) Suv39H1 and HP1gamma are responsible for chromatin-mediated HIV-1 transcriptional silencing and post-integration latency. *Embo J*, **26**, 424-435.

Dull, T., Zufferey, R., Kelly, M., Mandel, R.J., Nguyen, M., Trono, D. and Naldini, L. (1998) A third-generation lentivirus vector with a conditional packaging system. *J Virol*, **72**, 8463-8471.

Duverger, A., Jones, J., May, J., Bibollet-Ruche, F., Wagner, F.A., Cron, R.Q. and Kutsch, O. (2009) Determinants of the establishment of human immunodeficiency virus type 1 latency. *J Virol*, **83**, 3078-3093.

Ebina, H., Kanemura, Y., Suzuki, Y., Urata, K., Misawa, N. and Koyanagi, Y. (2012) Integrase-independent HIV-1 infection is augmented under conditions of DNA damage and produces a viral reservoir. *Virology*, **427**, 44-50.

el Kharroubi, A. and Verdin, E. (1994) Protein-DNA interactions within DNase I-hypersensitive sites located downstream of the HIV-1 promoter. *J Biol Chem*, **269**, 19916-19924.

Ellerman V., Bang O. Experimentell leukamie bei huhnern. Zentralbl. Bakteriol. 1908;46:595–609

Ellis, J. (2005) Silencing and variegation of gammaretrovirus and lentivirus vectors. *Hum Gene Ther*, **16**, 1241-1246.
Emery, D.W. (2011) The use of chromatin insulators to improve the expression and safety of integrating gene transfer vectors. *Hum Gene Ther*, **22**, 761-774.
Emery, D.W., Yannaki, E., Tubb, J. and Stamatoyannopoulos, G. (2000) A chromatin insulator protects retrovirus vectors from chromosomal position effects. *Proc Natl Acad Sci U S A*, **97**, 9150-9155.
Emiliani, S., Fischle, W., Ott, M., Van Lint, C., Amella, C.A. and Verdin, E. (1998) Mutations in the tat gene are responsible for human immunodeficiency virus type 1 postintegration latency in the U1 cell line. *J Virol*, **72**, 1666-1670.
Engelman, A. and Cherepanov, P. (2012) The structural biology of HIV-1: mechanistic and therapeutic insights. *Nat Rev Microbiol*, **10**, 279-290.
Errami, A., Smider, V., Rathmell, W.K., He, D.M., Hendrickson, E.A., Zdzienicka, M.Z. and Chu, G. (1996) Ku86 defines the genetic defect and restores X-ray resistance and V(D)J recombination to complementation group 5 hamster cell mutants. *Mol Cell Biol*, **16**, 1519-1526.
Espejel, S., Franco, S., Rodriguez-Perales, S., Bouffler, S.D., Cigudosa, J.C. and Blasco, M.A. (2002) Mammalian Ku86 mediates chromosomal fusions and apoptosis caused by critically short telomeres. *Embo J*, **21**, 2207-2219.
Espeseth, A.S., Fishel, R., Hazuda, D., Huang, Q., Xu, M., Yoder, K. and Zhou, H. (2011) siRNA screening of a targeted library of DNA repair factors in HIV infection reveals a role for base excision repair in HIV integration. *PLoS One*, **6**, e17612.
Evans-Galea, M.V., Wielgosz, M.M., Hanawa, H., Srivastava, D.K. and Nienhuis, A.W. (2007) Suppression of clonal dominance in cultured human lymphoid cells by addition of the cHS4 insulator to a lentiviral vector. *Mol Ther*, **15**, 801-809.
Farnet, C.M. and Haseltine, W.A. (1991) Determination of viral proteins present in the human immunodeficiency virus type 1 preintegration complex. *J Virol*, **65**, 1910-1915.
Fattah, F.J., Lichter, N.F., Fattah, K.R., Oh, S. and Hendrickson, E.A. (2008) Ku70, an essential gene, modulates the frequency of rAAV-mediated gene targeting in human somatic cells. *Proc Natl Acad Sci U S A*, **105**, 8703-8708.
Federici, T., Kutner, R., Zhang, X.Y., Kuroda, H., Tordo, N., Boulis, N.M. and Reiser, J. (2009) Comparative analysis of HIV-1-based lentiviral vectors bearing lyssavirus glycoproteins for neuronal gene transfer. *Genet Vaccines Ther*, **7**, 1.
Fenard, D., Houzet, L., Bernard, E., Tupin, A., Brun, S., Mougel, M., Devaux, C., Chazal, N. and Briant, L. (2009) Uracil DNA Glycosylase 2 negatively regulates HIV-1 LTR transcription. *Nucleic Acids Res*, **37**, 6008-6018.
Ferris, A.L., Wu, X., Hughes, C.M., Stewart, C., Smith, S.J., Milne, T.A., Wang, G.G., Shun, M.C., Allis, C.D., Engelman, A. and Hughes, S.H. (2010) Lens epithelium-derived growth factor fusion proteins redirect HIV-1 DNA integration. *Proc Natl Acad Sci U S A*, **107**, 3135-3140.
Fewell, J.W. and Kuff, E.L. (1996) Intracellular redistribution of Ku immunoreactivity in response to cell-cell contact and growth modulating components in the medium. *J Cell Sci*, **109 (Pt 7)**, 1937-1946.
Fischer, A., Hacein-Bey-Abina, S. and Cavazzana-Calvo, M. (2011) Gene therapy for primary adaptive immune deficiencies. *J Allergy Clin Immunol*, **127**, 1356-1359.
Follenzi, A., Ailles, L.E., Bakovic, S., Geuna, M. and Naldini, L. (2000) Gene transfer by lentiviral vectors is limited by nuclear translocation and rescued by HIV-1 pol sequences. *Nat Genet*, **25**, 217-222.
Francoeur, A.M., Peebles, C.L., Gompper, P.T. and Tan, E.M. (1986) Identification of Ki (Ku, p70/p80) autoantigens and analysis of anti-Ki autoantibody reactivity. *J Immunol*, **136**, 1648-1653.
Frankel, A.D. and Young, J.A. (1998) HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem*, **67**, 1-25.
Frecha, C., Costa, C., Levy, C., Negre, D., Russell, S.J., Maisner, A., Salles, G., Peng, K.W., Cosset, F.L. and Verhoeyen, E. (2009) Efficient and stable transduction of resting B lymphocytes and primary chronic lymphocyte leukemia cells using measles virus gp displaying lentiviral vectors. *Blood*, **114**, 3173-3180.

Froelich, S., Ziegler, L., Stroup, K. and Wang, P. (2009) Targeted gene delivery to CD117-expressing cells in vivo with lentiviral vectors co-displaying stem cell factor and a fusogenic molecule. *Biotechnol Bioeng*, **104**, 206-215.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T. and Peterlin, B.M. (2004) Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. *Mol Cell Biol*, **24**, 787-795.

Galande, S. and Kohwi-Shigematsu, T. (1999) Poly(ADP-ribose) polymerase and Ku autoantigen form a complex and synergistically bind to matrix attachment sequences. *J Biol Chem*, **274**, 20521-20528.

Gallo, R.C., Salahuddin, S.Z., Popovic, M., Shearer, G.M., Kaplan, M., Haynes, B.F., Palker, T.J., Redfield, R., Oleske, J., Safai, B. and et al. (1984) Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS. *Science*, **224**, 500-503.

Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., Leibowitch, J. and Popovic, M. (1983) Isolation of human T-cell leukemia virus in acquired immune deficiency syndrome (AIDS). *Science*, **220**, 865-867.

Gao, F., Bailes, E., Robertson, D.L., Chen, Y., Rodenburg, C.M., Michael, S.F., Cummins, L.B., Arthur, L.O., Peeters, M., Shaw, G.M., Sharp, P.M. and Hahn, B.H. (1999) Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature*, **397**, 436-441.

Gao, W.Y., Cara, A., Gallo, R.C. and Lori, F. (1993) Low levels of deoxynucleotides in peripheral blood lymphocytes: a strategy to inhibit human immunodeficiency virus type 1 replication. *Proc Natl Acad Sci U S A*, **90**, 8925-8928.

Garber, M.E., Wei, P. and Jones, K.A. (1998) HIV-1 Tat interacts with cyclin T1 to direct the P-TEFb CTD kinase complex to TAR RNA. *Cold Spring Harb Symp Quant Biol*, **63**, 371-380.

Gaszner, M. and Felsenfeld, G. (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat Rev Genet*, **7**, 703-713.

Gaussin, A., Modlich, U., Bauche, C., Niederlander, N.J., Schambach, A., Duros, C., Artus, A., Baum, C., Cohen-Haguenauer, O. and Mermod, N. (2012) CTF/NF1 transcription factors act as potent genetic insulators for integrating gene transfer vectors. *Gene Ther*, **19**, 15-24.

Gelderblom, H.C., Vatakis, D.N., Burke, S.A., Lawrie, S.D., Bristol, G.C. and Levy, D.N. (2008) Viral complementation allows HIV-1 replication without integration. *Retrovirology*, **5**, 60.

Genovesio, A., Kwon, Y.J., Windisch, M.P., Kim, N.Y., Choi, S.Y., Kim, H.C., Jung, S., Mammano, F., Perrin, V., Boese, A.S., Casartelli, N., Schwartz, O., Nehrbass, U. and Emans, N. (2011) Automated genome-wide visual profiling of cellular proteins involved in HIV infection. *J Biomol Screen*, **16**, 945-958.

Gerolami, R., Uch, R., Jordier, F., Chapel, S., Bagnis, C., Brechot, C. and Mannoni, P. (2000) Gene transfer to hepatocellular carcinoma: transduction efficacy and transgene expression kinetics by using retroviral and lentiviral vectors. *Cancer Gene Ther*, **7**, 1286-1292.

Gessain, A. and Gout, O. (1992) Chronic myelopathy associated with human T-lymphotropic virus type I (HTLV-I). *Ann Intern Med*, **117**, 933-946.

Geyer, P.K., Vitalini, M.W. and Wallrath, L.L. (2011) Nuclear organization: taking a position on gene expression. *Curr Opin Cell Biol*, **23**, 354-359.

Gheysen, D., Jacobs, E., de Foresta, F., Thiriart, C., Francotte, M., Thines, D. and De Wilde, M. (1989) Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell*, **59**, 103-112.

Ghosh, G., Li, G., Myung, K. and Hendrickson, E.A. (2007) The lethality of Ku86 (XRCC5) loss-of-function mutations in human cells is independent of p53 (TP53). *Radiat Res*, **167**, 66-79.

Giffin, W., Gong, W., Schild-Poulter, C. and Hache, R.J. (1999) Ku antigen-DNA conformation determines the activation of DNA-dependent protein kinase and DNA sequence-directed repression of mouse mammary tumor virus transcription. *Mol Cell Biol*, **19**, 4065-4078.

Giffin, W., Torrance, H., Rodda, D.J., Prefontaine, G.G., Pope, L. and Hache, R.J. (1996) Sequence-specific DNA binding by Ku autoantigen and its effects on transcription. *Nature*, **380**, 265-268.

Glenn, J. (1981) Cytomegalovirus infections following renal transplantation. *Rev Infect Dis*, **3**, 1151-1178.

Glover, D.J., Lipps, H.J. and Jans, D.A. (2005) Towards safe, non-viral therapeutic gene expression in humans. *Nat Rev Genet*, **6**, 299-310.

Goh, W.C., Rogel, M.E., Kinsey, C.M., Michael, S.F., Fultz, P.N., Nowak, M.A., Hahn, B.H. and Emerman, M. (1998) HIV-1 Vpr increases viral expression by manipulation of the cell cycle: a mechanism for selection of Vpr in vivo. *Nat Med*, **4**, 65-71.

Gotte, M., Rausch, J.W., Marchand, B., Sarafianos, S. and Le Grice, S.F. (2010) Reverse transcriptase in motion: conformational dynamics of enzyme-substrate interactions. *Biochim Biophys Acta*, **1804**, 1202-1212.

Grandchamp, N., Henriot, D., Philippe, S., Amar, L., Ursulet, S., Serguera, C., Mallet, J. and Sarkis, C. (2011) Influence of insulators on transgene expression from integrating and non-integrating lentiviral vectors. *Genet Vaccines Ther*, **9**, 1.

Grimm, D., Streetz, K.L., Jopling, C.L., Storm, T.A., Pandey, K., Davis, C.R., Marion, P., Salazar, F. and Kay, M.A. (2006) Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*, **441**, 537-541.

Grizot, S., Epinat, J.C., Thomas, S., Duclert, A., Rolland, S., Paques, F. and Duchateau, P. (2010) Generation of redesigned homing endonucleases comprising DNA-binding domains derived from two different scaffolds. *Nucleic Acids Res*, **38**, 2006-2018.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. and Young, R.A. (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell*, **130**, 77-88.

Gullo, C., Au, M., Feng, G. and Teoh, G. (2006) The biology of Ku and its potential oncogenic role in cancer. *Biochim Biophys Acta*, **1765**, 223-234.

Ha, H.C., Juluri, K., Zhou, Y., Leung, S., Hermankova, M. and Snyder, S.H. (2001) Poly(ADP-ribose) polymerase-1 is required for efficient HIV-1 integration. *Proc Natl Acad Sci U S A*, **98**, 3364-3368.

Hacein-Bey-Abina, S., Garrigue, A., Wang, G.P., Soulier, J., Lim, A., Morillon, E., Clappier, E., Caccavelli, L., Delabesse, E., Beldjord, K., Asnafi, V., MacIntyre, E., Dal Cortivo, L., Radford, I., Brousse, N., Sigaux, F., Moshous, D., Hauer, J., Borkhardt, A., Belohradsky, B.H., Wintergerst, U., Velez, M.C., Leiva, L., Sorensen, R., Wulffraat, N., Blanche, S., Bushman, F.D., Fischer, A. and Cavazzana-Calvo, M. (2008) Insertional oncogenesis in 4 patients after retrovirus-mediated gene therapy of SCID-X1. *J Clin Invest*, **118**, 3132-3142.

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Leboulch, P., Lim, A., Osborne, C.S., Pawliuk, R., Morillon, E., Sorensen, R., Forster, A., Fraser, P., Cohen, J.I., de Saint Basile, G., Alexander, I., Wintergerst, U., Frebourg, T., Aurias, A., Stoppa-Lyonnet, D., Romana, S., Radford-Weiss, I., Gross, F., Valensi, F., Delabesse, E., Macintyre, E., Sigaux, F., Soulier, J., Leiva, L.E., Wissler, M., Prinz, C., Rabbitts, T.H., Le Deist, F., Fischer, A. and Cavazzana-Calvo, M. (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, **302**, 415-419.

Hacker, C.V., Vink, C.A., Wardell, T.W., Lee, S., Treasure, P., Kingsman, S.M., Mitrophanous, K.A. and Miskin, J.E. (2006) The integration profile of EIAV-based vectors. *Mol Ther*, **14**, 536-545.

Hampsey, M., Singh, B.N., Ansari, A., Laine, J.P. and Krishnamurthy, S. (2011) Control of eukaryotic gene expression: gene loops and transcriptional memory. *Adv Enzyme Regul*, **51**, 118-125.

Han, Y., Lassen, K., Monie, D., Sedaghat, A.R., Shimoji, S., Liu, X., Pierson, T.C., Margolick, J.B., Siliciano, R.F. and Siliciano, J.D. (2004) Resting CD4+ T cells from human immunodeficiency virus type 1 (HIV-1)-infected individuals carry integrated HIV-1 genomes within actively transcribed host genes. *J Virol*, **78**, 6122-6133.

Han, Y., Lin, Y.B., An, W., Xu, J., Yang, H.C., O'Connell, K., Dordai, D., Boeke, J.D., Siliciano, J.D. and Siliciano, R.F. (2008) Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host Microbe*, **4**, 134-146.

Han, Y. and Siliciano, R.F. (2007) Keeping quiet: microRNAs in HIV-1 latency. *Nat Med*, **13**, 1138-1140.

Hanawa, H., Yamamoto, M., Zhao, H., Shimada, T. and Persons, D.A. (2009) Optimized lentiviral vector design improves titer and transgene expression of vectors containing the chicken beta-globin locus HS4 insulator element. *Mol Ther*, **17**, 667-674.

Hasegawa, M., Yamaguchi, S., Aizawa, S., Ikeda, H., Tatsumi, K., Noda, Y., Hirokawa, K. and Kitagawa, M. (2005) Resistance against Friend leukemia virus-induced leukemogenesis in DNA-dependent protein kinase (DNA-PK)-deficient scid mice associated with defective viral integration at the Spi-1 and Fli-1 site. *Leuk Res*, **29**, 933-942.

Hassa, P.O. and Hottiger, M.O. (1999) A role of poly (ADP-ribose) polymerase in NF-kappaB transcriptional activation. *Biol Chem*, **380**, 953-959.

Helseth, E., Kowalski, M., Gabuzda, D., Olshevsky, U., Haseltine, W. and Sodroski, J. (1990) Rapid complementation assays measuring replicative potential of human immunodeficiency virus type 1 envelope glycoprotein mutants. *J Virol*, **64**, 2416-2420.

Hematti, P., Hong, B.K., Ferguson, C., Adler, R., Hanawa, H., Sellers, S., Holt, I.E., Eckfeldt, C.E., Sharma, Y., Schmidt, M., von Kalle, C., Persons, D.A., Billings, E.M., Verfaillie, C.M., Nienhuis, A.W., Wolfsberg, T.G., Dunbar, C.E. and Calmels, B. (2004) Distinct genomic integration of MLV and SIV vectors in primate hematopoietic stem and progenitor cells. *PLoS Biol*, **2**, e423.

Heng, H.H., Goetze, S., Ye, C.J., Liu, G., Stevens, J.B., Bremer, S.W., Wykes, S.M., Bode, J. and Krawetz, S.A. (2004) Chromatin loops are selectively anchored using scaffold/matrix-attachment regions. *J Cell Sci*, **117**, 999-1008.

Heo, K., Kim, H., Choi, S.H., Choi, J., Kim, K., Gu, J., Lieber, M.R., Yang, A.S. and An, W. (2008) FACT-mediated exchange of histone variant H2AX regulated by phosphorylation of H2AX and ADP-ribosylation of Spt16. *Mol Cell*, **30**, 86-97.

Hirsch, V.M., Olmsted, R.A., Murphey-Corb, M., Purcell, R.H. and Johnson, P.R. (1989) An African primate lentivirus (SIVsm) closely related to HIV-2. *Nature*, **339**, 389-392.

Hoek, M., Myers, M.P. and Stillman, B. (2011) An analysis of CAF-1-interacting proteins reveals dynamic and direct interactions with the KU complex and 14-3-3 proteins. *J Biol Chem*, **286**, 10876-10887.

Holcomb, V.B., Rodier, F., Choi, Y., Busuttil, R.A., Vogel, H., Vijg, J., Campisi, J. and Hasty, P. (2008) Ku80 deletion suppresses spontaneous tumors and induces a p53-mediated DNA damage response. *Cancer Res*, **68**, 9497-9502.

Hong, Z., Jiang, J., Lan, L., Nakajima, S., Kanno, S., Koseki, H. and Yasui, A. (2008) A polycomb group protein, PHF1, is involved in the response to DNA double-strand breaks in human cell. *Nucleic Acids Res*, **36**, 2939-2947.

Howe, S.J., Mansour, M.R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M.H., Pike-Overzet, K., Chatters, S.J., de Ridder, D., Gilmour, K.C., Adams, S., Thornhill, S.I., Parsley, K.L., Staal, F.J., Gale, R.E., Linch, D.C., Bayford, J., Brown, L., Quaye, M., Kinnon, C., Ancliff, P., Webb, D.K., Schmidt, M., von Kalle, C., Gaspar, H.B. and Thrasher, A.J. (2008) Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *J Clin Invest*, **118**, 3143-3150.

Huang, J., Wang, F., Argyris, E., Chen, K., Liang, Z., Tian, H., Huang, W., Squires, K., Verlinghieri, G. and Zhang, H. (2007) Cellular microRNAs contribute to HIV-1 latency in resting primary CD4+ T lymphocytes. *Nat Med*, **13**, 1241-1247.

Huisinga, K.L., Brower-Toland, B. and Elgin, S.C. (2006) The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma*, **115**, 110-122.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R. and Bird, A.P. (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS Genet*, **6**.

Imai, K. and Ochiai, K. (2011) Role of histone modification on transcriptional regulation and HIV-1 gene expression: possible mechanisms of periodontal diseases in AIDS progression. *J Oral Sci*, **53**, 1-13.

Iordanskiy, S., Berro, R., Altieri, M., Kashanchi, F. and Bukrinsky, M. (2006) Intracytoplasmic maturation of the human immunodeficiency virus type 1 reverse transcription complexes determines their capacity to integrate into chromatin. *Retrovirology*, **3**, 4.

Isomura, H., Stinski, M.F., Kudoh, A., Daikoku, T., Shirata, N. and Tsurumi, T. (2005) Two Sp1/Sp3 binding sites in the major immediate-early proximal enhancer of human cytomegalovirus have a significant role in viral replication. *J Virol*, **79**, 9597-9607.

Isomura, H., Tsurumi, T. and Stinski, M.F. (2004) Role of the proximal enhancer of the major immediate-early promoter in human cytomegalovirus replication. *J Virol*, **78**, 12788-12799.

Iwakuma, T., Cui, Y. and Chang, L.J. (1999) Self-inactivating lentiviral vectors with U3 and U5 modifications. *Virology*, **261**, 120-132.

Iyer, S.R., Yu, D., Biancotto, A., Margolis, L.B. and Wu, Y. (2009) Measurement of human immunodeficiency virus type 1 preintegration transcription by using Rev-dependent Rev-CEM cells reveals a sizable transcribing DNA population comparable to that from proviral templates. *J Virol*, **83**, 8662-8673.

Izmiryan, A., Basmaciogullari, S., Henry, A., Paques, F. and Danos, O. (2011) Efficient gene targeting mediated by a lentiviral vector-associated meganuclease. *Nucleic Acids Res*, **39**, 7610-7619.

Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071-1078.

Jeanson, L. and Mouscadet, J.F. (2002) Ku represses the HIV-1 transcription: identification of a putative Ku binding site homologous to the mouse mammary tumor virus NRE1 sequence in the HIV-1 long terminal repeat. *J Biol Chem*, **277**, 4918-4924.

Jeanson, L., Subra, F., Vaganay, S., Hervy, M., Marangoni, E., Bourhis, J. and Mouscadet, J.F. (2002) Effect of Ku80 depletion on the preintegrative steps of HIV-1 replication in human cells. *Virology*, **300**, 100-108.

Jiang, D., Zhou, Y., Moxley, R.A. and Jarrett, H.W. (2008) Purification and identification of positive regulators binding to a novel element in the c-Jun promoter. *Biochemistry*, **47**, 9318-9334.

Jiang, G., Espeseth, A., Hazuda, D.J. and Margolis, D.M. (2007) c-Myc and Sp1 contribute to proviral latency by recruiting histone deacetylase 1 to the human immunodeficiency virus type 1 promoter. *J Virol*, **81**, 10914-10923.

Jin, C. and Felsenfeld, G. (2007) Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes Dev*, **21**, 1519-1529.

Jolly, C. and Sattentau, Q.J. (2007) Regulated secretion from CD4+ T cells. *Trends Immunol*, **28**, 474-481.

Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R. (1986) Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. *Science*, **232**, 755-759.

Jones, K.A. and Peterlin, B.M. (1994) Control of RNA initiation and elongation at the HIV-1 promoter. *Annu Rev Biochem*, **63**, 717-743.

Jones, P.A. (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*, **13**, 484-492.

Jordan, A., Bisgrove, D. and Verdin, E. (2003) HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *Embo J*, **22**, 1868-1877.

Jupp, R., Hoffmann, S., Depto, A., Stenberg, R.M., Ghazal, P. and Nelson, J.A. (1993) Direct interaction of the human cytomegalovirus IE86 protein with the cis repression signal does not preclude TBP from binding to the TATA box. *J Virol*, **67**, 5595-5604.

Juven-Gershon, T. and Kadonaga, J.T. (2010) Regulation of gene expression via the core promoter and the basal transcriptional machinery. *Dev Biol*, **339**, 225-229.

Kaczmarski, W. and Khan, S.A. (1993) Lupus autoantigen Ku protein binds HIV-1 TAR RNA in vitro. *Biochem Biophys Res Commun*, **196**, 935-942.

Kaehlcke, K., Dorr, A., Hetzer-Egger, C., Kiermer, V., Henklein, P., Schnoelzer, M., Loret, E., Cole, P.A., Verdin, E. and Ott, M. (2003) Acetylation of Tat defines a cyclinT1-independent step in HIV transactivation. *Mol Cell*, **12**, 167-176.

Kaiser, S.M. and Emerman, M. (2006) Uracil DNA glycosylase is dispensable for human immunodeficiency virus type 1 replication and does not contribute to the antiviral effects of the cytidine deaminase Apobec3G. *J Virol*, **80**, 875-882.

Kameoka, M., Nukuzuma, S., Itaya, A., Tanaka, Y., Ota, K., Ikuta, K. and Yoshihara, K. (2004) RNA interference directed against Poly(ADP-Ribose) polymerase 1 efficiently suppresses human immunodeficiency virus type 1 replication in human cells. *J Virol*, **78**, 8931-8934.

Kameoka, M., Nukuzuma, S., Itaya, A., Tanaka, Y., Ota, K., Inada, Y., Ikuta, K. and Yoshihara, K. (2005) Poly(ADP-ribose)polymerase-1 is required for integration of the human immunodeficiency virus type 1 genome near centromeric alphoid DNA in human and murine cells. *Biochem Biophys Res Commun*, **334**, 412-417.

Kang, Y., Moressi, C.J., Scheetz, T.E., Xie, L., Tran, D.T., Casavant, T.L., Ak, P., Benham, C.J., Davidson, B.L. and McCray, P.B., Jr. (2006) Integration site choice of a feline immunodeficiency virus vector. *J Virol*, **80**, 8820-8823.

Kantor, B., Ma, H., Webster-Cyriaque, J., Monahan, P.E. and Kafri, T. (2009) Epigenetic activation of unintegrated HIV-1 genomes by gut-associated short chain fatty acids and its implications for HIV infection. *Proc Natl Acad Sci U S A*, **106**, 18786-18791.

Kao, S.Y., Calman, A.F., Luciw, P.A. and Peterlin, B.M. (1987) Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product. *Nature*, **330**, 489-493.

Kappes, J.C. and Wu, X. (2001) Safety considerations in vector development. *Somat Cell Mol Genet*, **26**, 147-158.

Karageorgos, L., Li, P. and Burrell, C. (1993) Characterization of HIV replication complexes early after cell-to-cell infection. *AIDS Res Hum Retroviruses*, **9**, 817-823.

Karn, J. (2011) The molecular biology of HIV latency: breaking and restoring the Tat-dependent transcriptional circuit. *Curr Opin HIV AIDS*, **6**, 4-11.

Karn, J. and Stoltzfus, C.M. (2012) Transcriptional and Posttranscriptional Regulation of HIV-1 Gene Expression. *Cold Spring Harb Perspect Med*, **2**, a006916.

Keele, B.F., Van Heuverswyn, F., Li, Y., Bailes, E., Takehisa, J., Santiago, M.L., Bibollet-Ruche, F., Chen, Y., Wain, L.V., Liegeois, F., Loul, S., Ngole, E.M., Bienvenue, Y., Delaporte, E., Brookfield, J.F., Sharp, P.M., Shaw, G.M., Peeters, M. and Hahn, B.H. (2006) Chimpanzee reservoirs of pandemic and nonpandemic HIV-1. *Science*, **313**, 523-526.

Kelly, J., Beddall, M.H., Yu, D., Iyer, S.R., Marsh, J.W. and Wu, Y. (2008) Human macrophages support persistent transcription from unintegrated HIV-1 DNA. *Virology*, **372**, 300-312.

Kessis, T.D., Slebos, R.J., Nelson, W.G., Kastan, M.B., Plunkett, B.S., Han, S.M., Lorincz, A.T., Hedrick, L. and Cho, K.R. (1993) Human papillomavirus 16 E6 expression disrupts the p53-mediated cellular response to DNA damage. *Proc Natl Acad Sci U S A*, **90**, 3988-3992.

Kiernan, R.E., Vanhulle, C., Schiltz, L., Adam, E., Xiao, H., Maudoux, F., Calomme, C., Burny, A., Nakatani, Y., Jeang, K.T., Benkirane, M. and Van Lint, C. (1999) HIV-1 tat transcriptional activity is regulated by acetylation. *Embo J*, **18**, 6106-6118.

Kilzer, J.M., Stracker, T., Beitzel, B., Meek, K., Weitzman, M. and Bushman, F.D. (2003) Roles of host cell factors in circularization of retroviral dna. *Virology*, **314**, 460-467.

Kim, H. (2008) DNA repair Ku proteins in gastric cancer cells and pancreatic acinar cells. *Amino Acids*, **34**, 195-202.

Kim, J.K., Samaranayake, M. and Pradhan, S. (2009) Epigenetic mechanisms in mammals. *Cell Mol Life Sci*, **66**, 596-612.

Kim, V.N., Mitrophanous, K., Kingsman, S.M. and Kingsman, A.J. (1998) Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *J Virol*, **72**, 811-816.

Kim, Y.K., Bourgeois, C.F., Pearson, R., Tyagi, M., West, M.J., Wong, J., Wu, S.Y., Chiang, C.M. and Karn, J. (2006) Recruitment of TFIIH to the HIV LTR is a rate-limiting step in the emergence of HIV from latency. *Embo J*, **25**, 3596-3604.

Kingsman, S.M., Mitrophanous, K. and Olsen, J.C. (2005) Potential oncogene activity of the woodchuck hepatitis post-transcriptional regulatory element (WPRE). *Gene Ther*, **12**, 3-4.

Knight, S., Zhang, F., Mueller-Kuller, U., Bokhoven, M., Gupta, A., Broughton, T., Sha, S., Antoniou, M.N., Brendel, C., Grez, M., Thrasher, A.J., Collins, M. and Takeuchi, Y. (2012) Safer, silencing-resistant lentiviral vectors: Optimization of ubiquitous chromatin opening element (UCOE) through elimination of aberrant splicing. *J Virol*.

Koike, M. (2002) Dimerization, translocation and localization of Ku70 and Ku80 proteins. *J Radiat Res*, **43**, 223-236.

Koike, M., Yutoku, Y. and Koike, A. (2011) KARP-1 works as a heterodimer with Ku70, but the function of KARP-1 cannot perfectly replace that of Ku80 in DSB repair. *Exp Cell Res*, **317**, 2267-2275.

Konig, R., Zhou, Y., Elleder, D., Diamond, T.L., Bonamy, G.M., Irelan, J.T., Chiang, C.Y., Tu, B.P., De Jesus, P.D., Lilley, C.E., Seidel, S., Opaluch, A.M., Caldwell, J.S., Weitzman, M.D., Kuhen, K.L., Bandyopadhyay, S., Ideker, T., Orth, A.P., Miraglia, L.J., Bushman, F.D., Young, J.A. and Chanda, S.K. (2008) Global analysis of host-pathogen interactions that regulate early-stage HIV-1 replication. *Cell*, **135**, 49-60.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S. and Bhattacharya, T. (2000) Timing the ancestor of the HIV-1 pandemic strains. *Science*, **288**, 1789-1796.

Kotsopoulou, E., Kim, V.N., Kingsman, A.J., Kingsman, S.M. and Mitrophanous, K.A. (2000) A Rev-independent human immunodeficiency virus type 1 (HIV-1)-based vector that exploits a codon-optimized HIV-1 gag-pol gene. *J Virol*, **74**, 4839-4852.

Kramer, B., Pelchen-Matthews, A., Deneka, M., Garcia, E., Piguet, V. and Marsh, M. (2005) HIV interaction with endosomes in macrophages and dendritic cells. *Blood Cells Mol Dis*, **35**, 136-142.

Kress, S., Stein, A., Maurer, P., Weber, B., Reichert, J., Buchmann, A., Huppert, P. and Schwarz, M. (1998) Expression of hypoxia-inducible genes in tumor cells. *J Cancer Res Clin Oncol*, **124**, 315-320.

Kumar, M., Keller, B., Makalou, N. and Sutton, R.E. (2001) Systematic determination of the packaging limit of lentiviral vectors. *Hum Gene Ther*, **12**, 1893-1905.

Kurth, R. and Bannert, N. (2010) Beneficial and detrimental effects of human endogenous retroviruses. *Int J Cancer*, **126**, 306-314.

Kutner, R.H., Zhang, X.Y. and Reiser, J. (2009) Production, concentration and titration of pseudotyped HIV-1-based lentiviral vectors. *Nat Protoc*, **4**, 495-505.

Lai, M., Zimmerman, E.S., Planelles, V. and Chen, J. (2005) Activation of the ATR pathway by human immunodeficiency virus type 1 Vpr involves its direct binding to chromatin in vivo. *J Virol*, **79**, 15443-15451.

Lakeman, A.D. and Osborn, J.E. (1979) Size of infectious DNA from human and murine cytomegaloviruses. *J Virol*, **30**, 414-416.

Lam, W., Leung, C.H., Bussom, S. and Cheng, Y.C. (2007) The impact of hypoxic treatment on the expression of phosphoglycerate kinase and the cytotoxicity of troxacitabine and gemcitabine. *Mol Pharmacol*, **72**, 536-544.

Landolin, J.M., Johnson, D.S., Trinklein, N.D., Aldred, S.F., Medina, C., Shulha, H., Weng, Z. and Myers, R.M. (2010) Sequence features that drive human promoter function and tissue specificity. *Genome Res*, **20**, 890-898.

Lau, A., Kanaar, R., Jackson, S.P. and O'Connor, M.J. (2004) Suppression of retroviral infection by the RAD52 DNA repair protein. *Embo J*, **23**, 3421-3429.

Lau, A., Swinbank, K.M., Ahmed, P.S., Taylor, D.L., Jackson, S.P., Smith, G.C. and O'Connor, M.J. (2005) Suppression of HIV-1 infection by a small molecule inhibitor of the ATM kinase. *Nat Cell Biol*, **7**, 493-500.

Laurent-Crawford, A.G. and Hovanessian, A.G. (1993) The cytopathic effect of human immunodeficiency virus is independent of high levels of unintegrated viral DNA accumulated in response to superinfection of cells. *J Gen Virol*, **74 (Pt 12)**, 2619-2628.

Lee, J.T. (2011) Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nat Rev Mol Cell Biol*, **12**, 815-826.

Lee, S.H., Oshige, M., Durant, S.T., Rasila, K.K., Williamson, E.A., Ramsey, H., Kwan, L., Nickoloff, J.A. and Hromas, R. (2005) The SET domain protein Metnase mediates foreign DNA integration and links integration to nonhomologous end-joining repair. *Proc Natl Acad Sci U S A*, **102**, 18075-18080.

Lenasi, T., Contreras, X. and Peterlin, B.M. (2008) Transcriptional interference antagonizes proviral gene expression to promote HIV latency. *Cell Host Microbe*, **4**, 123-133.

Lesbats, P., Botbol, Y., Chevereau, G., Vaillant, C., Calmels, C., Arneodo, A., Andreola, M.L., Lavigne, M. and 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.

Levine, B.L., Humeau, L.M., Boyer, J., MacGregor, R.R., Rebello, T., Lu, X., Binder, G.K., Slepushkin, V., Lemiale, F., Mascola, J.R., Bushman, F.D., Dropulic, B. and June, C.H. (2006) Gene transfer in humans using a conditionally replicating lentiviral vector. *Proc Natl Acad Sci U S A*, **103**, 17372-17377.

Levy, J.A., Hoffman, A.D., Kramer, S.M., Landis, J.A., Shimabukuro, J.M. and Oshiro, L.S. (1984) Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS. *Science*, **225**, 840-842.

Lewinski, M.K., Bisgrove, D., Shinn, P., Chen, H., Hoffmann, C., Hannenhalli, S., Verdin, E., Berry, C.C., Ecker, J.R. and Bushman, F.D. (2005) Genome-wide analysis of chromosomal features repressing human immunodeficiency virus transcription. *J Virol*, **79**, 6610-6619.

Lewinski, M.K. and Bushman, F.D. (2005) Retroviral DNA integration--mechanism and consequences. *Adv Genet*, **55**, 147-181.

Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707-719.

Li, G., Nelsen, C. and Hendrickson, E.A. (2002) Ku86 is essential in human somatic cells. *Proc Natl Acad Sci U S A*, **99**, 832-837.

Li, G. and Reinberg, D. (2011) Chromatin higher-order structures and gene regulation. *Curr Opin Genet Dev*, **21**, 175-186.

Li, L., Olvera, J.M., Yoder, K.E., Mitchell, R.S., Butler, S.L., Lieber, M., Martin, S.L. and Bushman, F.D. (2001) Role of the non-homologous DNA end joining pathway in the early steps of retroviral infection. *Embo J*, **20**, 3272-3281.

Lilley, C.E., Chaurushiya, M.S. and Weitzman, M.D. (2010) Chromatin at the intersection of viral infection and DNA damage. *Biochim Biophys Acta*, **1799**, 319-327.

Lin, C.W. and 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.

Liu, J., Wright, E.R. and Winkler, H. (2010) 3D visualization of HIV virions by cryoelectron tomography. *Methods Enzymol*, **483**, 267-290.

Llano, M., Vanegas, M., Fregoso, O., Saenz, D., Chung, S., Peretz, M. and Poeschla, E.M. (2004) LEDGF/p75 determines cellular trafficking of diverse lentiviral but not murine oncoretroviral integrase proteins and is a component of functional lentiviral preintegration complexes. *J Virol*, **78**, 9524-9537.

Llano, M., Vanegas, M., Hutchins, N., Thompson, D., Delgado, S. and Poeschla, E.M. (2006) Identification and characterization of the chromatin-binding domains of the HIV-1 integrase interactor LEDGF/p75. *J Mol Biol*, **360**, 760-773.

Lloyd, A.G., Tateishi, S., Bieniasz, P.D., Muesing, M.A., Yamaizumi, M. and Mulder, L.C. (2006) Effect of DNA repair protein Rad18 on viral infection. *PLoS Pathog*, **2**, e40.

Lombardo, A., Genovese, P., Beausejour, C.M., Colleoni, S., Lee, Y.L., Kim, K.A., Ando, D., Urnov, F.D., Galli, C., Gregory, P.D., Holmes, M.C. and Naldini, L. (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*, **25**, 1298-1306.

Lord, C.J. and Ashworth, A. (2012) The DNA damage response and cancer therapy. *Nature*, **481**, 287-294.

Lori, F., di Marzo Veronese, F., de Vico, A.L., Lusso, P., Reitz, M.S., Jr. and Gallo, R.C. (1992) Viral DNA carried by human immunodeficiency virus type 1 virions. *J Virol*, **66**, 5067-5074.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251-260.

Luo, X. and Kraus, W.L. (2012) On PAR with PARP: cellular stress signaling through poly(ADP-ribose) and PARP-1. *Genes Dev*, **26**, 417-432.

Macias, M.P., Huang, L., Lashmit, P.E. and Stinski, M.F. (1996) Cellular or viral protein binding to a cytomegalovirus promoter transcription initiation site: effects on transcription. *J Virol*, **70**, 3628-3635.

MacNeil, A., Sankale, J.L., Meloni, S.T., Sarr, A.D., Mboup, S. and Kanki, P. (2006) Genomic sites of human immunodeficiency virus type 2 (HIV-2) integration: similarities to HIV-1 in vitro and possible differences in vivo. *J Virol*, **80**, 7316-7321.

Manic, G., Maurin-Marlin, A., Galluzzi, L., Subra, F., Mouscadet, J.F. and Bury-Mone, S. (2012) 3' self-inactivating long terminal repeat inserts for the modulation of transgene expression from lentiviral vectors. *Hum Gene Ther Methods*, **23**, 84-97.

Marban, C., Suzanne, S., Dequiedt, F., de Walque, S., Redel, L., Van Lint, C., Aunis, D. and Rohr, O. (2007) Recruitment of chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. *Embo J*, **26**, 412-423.

Martin-Serrano, J. and Neil, S.J. (2011) Host factors involved in retroviral budding and release. *Nat Rev Microbiol*, **9**, 519-531.

Maruggi, G., Porcellini, S., Facchini, G., Perna, S.K., Cattoglio, C., Sartori, D., Ambrosi, A., Schambach, A., Baum, C., Bonini, C., Bovolenta, C., Mavilio, F. and Recchia, A. (2009) Transcriptional enhancers induce insertional gene deregulation independently from the vector type and design. *Mol Ther*, **17**, 851-856.

Masson, C., Bury-Mone, S., Guiot, E., Saez-Cirion, A., Schoevaert-Brossault, D., Brachet-Ducos, C., Delelis, O., Subra, F., Jeanson-Leh, L. and Mouscadet, J.F. (2007) Ku80 participates in the targeting of retroviral transgenes to the chromatin of CHO cells. *J Virol*, **81**, 7924-7932.

Matrai, J., Chuah, M.K. and VandenDriessche, T. (2010) Recent advances in lentiviral vector development and applications. *Mol Ther*, **18**, 477-490.

Mayeur, G.L., Kung, W.J., Martinez, A., Izumiya, C., Chen, D.J. and Kung, H.J. (2005) Ku is a novel transcriptional recycling coactivator of the androgen receptor in prostate cancer cells. *J Biol Chem*, **280**, 10827-10833.

McDonald, D., Vodicka, M.A., Lucero, G., Svitkina, T.M., Borisy, G.G., Emerman, M. and Hope, T.J. (2002) Visualization of the intracellular behavior of HIV in living cells. *J Cell Biol*, **159**, 441-452.

Meier, J.L., Keller, M.J. and McCoy, J.J. (2002) Requirement of multiple cis-acting elements in the human cytomegalovirus major immediate-early distal enhancer for viral gene expression and replication. *J Virol*, **76**, 313-326.

Meier, J.L. and Pruessner, J.A. (2000) The human cytomegalovirus major immediate-early distal enhancer region is required for efficient viral replication and immediate-early gene expression. *J Virol*, **74**, 1602-1613.

Merten, O.W., Charrier, S., Laroudie, N., Fauchille, S., Dugue, C., Jenny, C., Audit, M., Zanta-Boussif, M.A., Chautard, H., Radrizzani, M., Vallanti, G., Naldini, L., Noguiez-Hellin, P. and Galy, A. (2011) Large-scale manufacture and characterization of a lentiviral vector produced for clinical ex vivo gene therapy application. *Hum Gene Ther*, **22**, 343-356.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P., Lee, W., Mendenhall, E., O'Donovan, A., Presser, A., Russ, C., Xie, X., Meissner, A., Wernig, M., Jaenisch, R., Nusbaum, C., Lander, E.S. and Bernstein, B.E. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553-560.

Mild, M., Kvist, A., Esbjornsson, J., Karlsson, I., Fenyo, E.M. and Medstrand, P. (2010) Differences in molecular evolution between switch (R5 to R5X4/X4-tropic) and non-switch (R5-tropic only) HIV-1 populations during infection. *Infect Genet Evol*, **10**, 356-364.

Miller, J.C., Holmes, M.C., Wang, J., Guschin, D.Y., Lee, Y.L., Rupniewski, I., Beausejour, C.M., Waite, A.J., Wang, N.S., Kim, K.A., Gregory, P.D., Pabo, C.O. and Rebar, E.J. (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*, **25**, 778-785.

Miller, M.D., Farnet, C.M. and Bushman, F.D. (1997) Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition. *J Virol*, **71**, 5382-5390.

Mimori, T., Akizuki, M., Yamagata, H., Inada, S., Yoshida, S. and Homma, M. (1981) Characterization of a high molecular weight acidic nuclear protein recognized by autoantibodies in sera from patients with polymyositis-scleroderma overlap. *J Clin Invest*, **68**, 611-620.

Mimori, T., Hardin, J.A. and Steitz, J.A. (1986) Characterization of the DNA-binding protein antigen Ku recognized by autoantibodies from patients with rheumatic disorders. *J Biol Chem*, **261**, 2274-2278.

Mitchell, R.S., Beitzel, B.F., Schroder, A.R., Shinn, P., Chen, H., Berry, C.C., Ecker, J.R. and Bushman, F.D. (2004) Retroviral DNA integration: ASLV, HIV, and MLV show distinct target site preferences. *PLoS Biol*, **2**, E234.

Miyoshi, H., Blomer, U., Takahashi, M., Gage, F.H. and Verma, I.M. (1998) Development of a self-inactivating lentivirus vector. *J Virol*, **72**, 8150-8157.

Mladenov, E. and Iliakis, G. (2011) Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways. *Mutat Res*, **711**, 61-72.

Moiani, A., Paleari, Y., Sartori, D., Mezzadra, R., Miccio, A., Cattoglio, C., Cocchiarella, F., Lidonnici, M.R., Ferrari, G. and Mavilio, F. (2012) Lentiviral vector integration in the human genome induces alternative splicing and generates aberrant transcripts. *J Clin Invest*, **122**, 1653-1666.

Moir, S., Chun, T.W. and Fauci, A.S. (2011) Pathogenic mechanisms of HIV disease. *Annu Rev Pathol*, **6**, 223-248.

Moloney, J.B. (1966) A virus-induced rhabdomyosarcoma of mice. *Natl Cancer Inst Monogr*, **22**, 139-142.

Monk, M. (1995) Epigenetic programming of differential gene expression in development and evolution. *Dev Genet*, **17**, 188-197.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Bartholomae, C.C., Ranzani, M., Benedicenti, F., Sergi, L.S., Ambrosi, A., Ponzoni, M., Doglioni, C., Di Serio, C., von Kalle, C. and Naldini, L. (2009) The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J Clin Invest*, **119**, 964-975.

Montini, E., Cesana, D., Schmidt, M., Sanvito, F., Ponzoni, M., Bartholomae, C., Sergi Sergi, L., Benedicenti, F., Ambrosi, A., Di Serio, C., Doglioni, C., von Kalle, C. and Naldini, L. (2006) Hematopoietic stem cell gene transfer in a tumor-prone mouse model uncovers low genotoxicity of lentiviral vector integration. *Nat Biotechnol*, **24**, 687-696.

Motomura, K., Chen, J. and Hu, W.S. (2008) Genetic recombination between human immunodeficiency virus type 1 (HIV-1) and HIV-2, two distinct human lentiviruses. *J Virol*, **82**, 1923-1933.

Mukerjee, R., Deshmane, S.L., Fan, S., Del Valle, L., White, M.K., Khalili, K., Amini, S. and Sawaya, B.E. (2008) Involvement of the p53 and p73 transcription factors in neuroAIDS. *Cell Cycle*, **7**, 2682-2690.

Murphy, J.C., Fischle, W., Verdin, E. and Sinclair, J.H. (2002) Control of cytomegalovirus lytic gene expression by histone acetylation. *Embo J*, **21**, 1112-1120.

Muthumani, K., Lambert, V.M., Sardesai, N.Y., Kim, J.J., Heller, R., Weiner, D.B. and Ugen, K.E. (2009) Analysis of the potential for HIV-1 Vpr as an anti-cancer agent. *Curr HIV Res*, **7**, 144-152.

Myung, K., He, D.M., Lee, S.E. and Hendrickson, E.A. (1997) KARP-1: a novel leucine zipper protein expressed from the Ku86 autoantigen locus is implicated in the control of DNA-dependent protein kinase activity. *Embo J*, **16**, 3172-3184.

Nabel, G. and Baltimore, D. (1987) An inducible transcription factor activates expression of human immunodeficiency virus in T cells. *Nature*, **326**, 711-713.

Nakashima, H., Matsui, T., Harada, S., Kobayashi, N., Matsuda, A., Ueda, T. and Yamamoto, N. (1986) Inhibition of replication and cytopathic effect of human T cell lymphotropic virus type III/lymphadenopathy-associated virus by 3'-azido-3'-deoxythymidine in vitro. *Antimicrob Agents Chemother*, **30**, 933-937.

Naldini, L., Blomer, U., Gage, F.H., Trono, D. and Verma, I.M. (1996a) Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector. *Proc Natl Acad Sci U S A*, **93**, 11382-11388.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F.H., Verma, I.M. and Trono, D. (1996b) In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science*, **272**, 263-267.

Nayak, S. and Herzog, R.W. (2010) Progress and prospects: immune responses to viral vectors. *Gene Ther*, **17**, 295-304.

Nermut, M.V. and Fassati, A. (2003) Structural analyses of purified human immunodeficiency virus type 1 intracellular reverse transcription complexes. *J Virol*, **77**, 8196-8206.
Nguyen, D.G., Booth, A., Gould, S.J. and Hildreth, J.E. (2003) Evidence that HIV budding in primary macrophages occurs through the exosome release pathway. *J Biol Chem*, **278**, 52347-52354.
Nielsen, T.T., Jakobsson, J., Rosenqvist, N. and Lundberg, C. (2009) Incorporating double copies of a chromatin insulator into lentiviral vectors results in less viral integrants. *BMC Biotechnol*, **9**, 13.
Nienhuis, A.W., Dunbar, C.E. and Sorrentino, B.P. (2006) Genotoxicity of retroviral integration in hematopoietic cells. *Mol Ther*, **13**, 1031-1049.
Norrman, K., Fischer, Y., Bonnamy, B., Wolfhagen Sand, F., Ravassard, P. and Semb, H. (2010) Quantitative comparison of constitutive promoters in human ES cells. *PLoS One*, **5**, e12413.
Nowrouzi, A., Dittrich, M., Klanke, C., Heinkelein, M., Rammling, M., Dandekar, T., von Kalle, C. and Rethwilm, A. (2006) Genome-wide mapping of foamy virus vector integrations into a human cell line. *J Gen Virol*, **87**, 1339-1347.
Nowrouzi, A., Glimm, H., von Kalle, C. and Schmidt, M. (2011) Retroviral vectors: post entry events and genomic alterations. *Viruses*, **3**, 429-455.
Nunnari, G., Argyris, E., Fang, J., Mehlman, K.E., Pomerantz, R.J. and Daniel, R. (2005) Inhibition of HIV-1 replication by caffeine and caffeine-related methylxanthines. *Virology*, **335**, 177-184.
O'Connell, K.A., Bailey, J.R. and Blankson, J.N. (2009) Elucidating the elite: mechanisms of control in HIV-1 infection. *Trends Pharmacol Sci*, **30**, 631-637.
Ochem, A.E., Skopac, D., Costa, M., Rabilloud, T., Vuillard, L., Simoncsits, A., Giacca, M. and Falaschi, A. (1997) Functional properties of the separate subunits of human DNA helicase II/Ku autoantigen. *J Biol Chem*, **272**, 29919-29926.
Oh, T., Bajwa, A., Jia, G. and Park, F. (2007) Lentiviral vector design using alternative RNA export elements. *Retrovirology*, **4**, 38.
Okamura, H., Aramburu, J., Garcia-Rodriguez, C., Viola, J.P., Raghavan, A., Tahiliani, M., Zhang, X., Qin, J., Hogan, P.G. and Rao, A. (2000) Concerted dephosphorylation of the transcription factor NFAT1 induces a conformational switch that regulates transcriptional activity. *Mol Cell*, **6**, 539-550.
Okumura, K., Takagi, S., Sakaguchi, G., Naito, K., Minoura-Tada, N., Kobayashi, H., Mimori, T., Hinuma, Y. and Igarashi, H. (1994) Autoantigen Ku protein is involved in DNA binding proteins which recognize the U5 repressive element of human T-cell leukemia virus type I long terminal repeat. *FEBS Lett*, **356**, 94-100.
Olsen, H.S. and Rosen, C.A. (1992) Contribution of the TATA motif to Tat-mediated transcriptional activation of human immunodeficiency virus gene expression. *J Virol*, **66**, 5594-5597.
Ong, C.T. and Corces, V.G. (2011) Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nat Rev Genet*, **12**, 283-293.
Ono, A., Waheed, A.A. and Freed, E.O. (2007) Depletion of cellular cholesterol inhibits membrane binding and higher-order multimerization of human immunodeficiency virus type 1 Gag. *Virology*, **360**, 27-35.
Ota, K., Kameoka, M., Tanaka, Y., Itaya, A. and Yoshihara, K. (2003) Expression of histone acetyltransferases was down-regulated in poly(ADP-ribose) polymerase-1-deficient murine cells. *Biochem Biophys Res Commun*, **310**, 312-317.
Pace, M.J., Agosto, L., Graf, E.H. and O'Doherty, U. (2011) HIV reservoirs and latency models. *Virology*, **411**, 344-354.
Pagans, S., Pedal, A., North, B.J., Kaehlcke, K., Marshall, B.L., Dorr, A., Hetzer-Egger, C., Henklein, P., Frye, R., McBurney, M.W., Hruby, H., Jung, M., Verdin, E. and Ott, M. (2005) SIRT1 regulates HIV transcription via Tat deacetylation. *PLoS Biol*, **3**, e41.
Page, K.A., Landau, N.R. and Littman, D.R. (1990) Construction and use of a human immunodeficiency virus vector for analysis of virus infectivity. *J Virol*, **64**, 5270-5276.
Pankotai, T., Bonhomme, C., Chen, D. and Soutoglou, E. (2012) DNAPKcs-dependent arrest of RNA polymerase II transcription in the presence of DNA breaks. *Nat Struct Mol Biol*, **19**, 276-282.
Paprotka, T., Delviks-Frankenberry, K.A., Cingoz, O., Martinez, A., Kung, H.J., Tepper, C.G., Hu, W.S., Fivash, M.J., Jr., Coffin, J.M. and Pathak, V.K. (2011) Recombinant origin of the retrovirus XMRV. *Science*, **333**, 97-101.

Parent, M., Yung, T.M., Rancourt, A., Ho, E.L., Vispe, S., Suzuki-Matsuda, F., Uehara, A., Wada, T., Handa, H. and Satoh, M.S. (2005) Poly(ADP-ribose) polymerase-1 is a negative regulator of HIV-1 transcription through competitive binding to TAR RNA with Tat.positive transcription elongation factor b (p-TEFb) complex. *J Biol Chem*, **280**, 448-457.

Park, F. and Kay, M.A. (2001) Modified HIV-1 based lentiviral vectors have an effect on viral transduction efficiency and gene expression in vitro and in vivo. *Mol Ther*, **4**, 164-173.

Parker, J.C. and Hoffman, J.F. (1967) The role of membrane phosphoglycerate kinase in the control of glycolytic rate by active cation transport in human red blood cells. *J Gen Physiol*, **50**, 893-916.

Parolin, C., Dorfman, T., Palu, G., Gottlinger, H. and Sodroski, J. (1994) Analysis in human immunodeficiency virus type 1 vectors of cis-acting sequences that affect gene transfer into human lymphocytes. *J Virol*, **68**, 3888-3895.

Paulk, N.K., Marquez Loza, L., Finegold, M. and Grompe, M. (2012) AAV-mediated gene targeting is significantly enhance by transient inhibition of NHEJ or the proteasome in vivo.. *Hum Gene Ther*.

Pauls, E., Senserrich, J., Clotet, B. and Este, J.A. (2006) Inhibition of HIV-1 replication by RNA interference of p53 expression. *J Leukoc Biol*, **80**, 659-667.

Pauwels, K., Gijsbers, R., Toelen, J., Schambach, A., Willard-Gallo, K., Verheust, C., Debyser, Z. and Herman, P. (2009) State-of-the-art lentiviral vectors for research use: risk assessment and biosafety recommendations. *Curr Gene Ther*, **9**, 459-474.

Pearson, R., Kim, Y.K., Hokello, J., Lassen, K., Friedman, J., Tyagi, M. and Karn, J. (2008) Epigenetic silencing of human immunodeficiency virus (HIV) transcription by formation of restrictive chromatin structures at the viral long terminal repeat drives the progressive entry of HIV into latency. *J Virol*, **82**, 12291-12303.

Peeters, M., Gueye, A., Mboup, S., Bibollet-Ruche, F., Ekaza, E., Mulanga, C., Ouedrago, R., Gandji, R., Mpele, P., Dibanga, G., Koumare, B., Saidou, M., Esu-Williams, E., Lombart, J.P., Badombena, W., Luo, N., Vanden Haesevelde, M. and Delaporte, E. (1997) Geographical distribution of HIV-1 group O viruses in Africa. *Aids*, **11**, 493-498.

Perkins, K.J., Lusic, M., Mitar, I., Giacca, M. and Proudfoot, N.J. (2008) Transcription-dependent gene looping of the HIV-1 provirus is dictated by recognition of pre-mRNA processing signals. *Mol Cell*, **29**, 56-68.

Perkins, N.D., Edwards, N.L., Duckett, C.S., Agranoff, A.B., Schmid, R.M. and Nabel, G.J. (1993) A cooperative interaction between NF-kappa B and Sp1 is required for HIV-1 enhancer activation. *Embo J*, **12**, 3551-3558.

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

Pfeifer, A., Kessler, T., Yang, M., Baranov, E., Kootstra, N., Cheresh, D.A., Hoffman, R.M. and Verma, I.M. (2001) Transduction of liver cells by lentiviral vectors: analysis in living animals by fluorescence imaging. *Mol Ther*, **3**, 319-322.

Pfeifer, G.P. and Riggs, A.D. (1991) Chromatin differences between active and inactive X chromosomes revealed by genomic footprinting of permeabilized cells using DNase I and ligation-mediated PCR. *Genes Dev*, **5**, 1102-1113.

Pfeifer, G.P., Tanguay, R.L., Steigerwald, S.D. and Riggs, A.D. (1990) In vivo footprint and methylation analysis by PCR-aided genomic sequencing: comparison of active and inactive X chromosomal DNA at the CpG island and promoter of human PGK-1. *Genes Dev*, **4**, 1277-1287.

Ping, Y.H. and Rana, T.M. (2001) DSIF and NELF interact with RNA polymerase II elongation complex and HIV-1 Tat stimulates P-TEFb-mediated phosphorylation of RNA polymerase II and DSIF during transcription elongation. *J Biol Chem*, **276**, 12951-12958.

Pluta, K. and Kacprzak, M.M. (2009) Use of HIV as a gene transfer vector. *Acta Biochim Pol*, **56**, 531-595.

Poeschla, E.M. (2008) Integrase, LEDGF/p75 and HIV replication. *Cell Mol Life Sci*, **65**, 1403-1424.

Polo, S.E. and Jackson, S.P. (2011) Dynamics of DNA damage response proteins at DNA breaks: a focus on protein modifications. *Genes Dev*, **25**, 409-433.

Poon, B. and Chen, I.S. (2003) Human immunodeficiency virus type 1 (HIV-1) Vpr enhances expression from unintegrated HIV-1 DNA. *J Virol*, **77**, 3962-3972.

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

Postow, L., Ghenoiu, C., Woo, E.M., Krutchinsky, A.N., Chait, B.T. and Funabiki, H. (2008) Ku80 removal from DNA through double strand break-induced ubiquitylation. *J Cell Biol*, **182**, 467-479.

Prabhakar, B.S., Allaway, G.P., Srinivasappa, J. and Notkins, A.L. (1990) Cell surface expression of the 70-kD component of Ku, a DNA-binding nuclear autoantigen. *J Clin Invest*, **86**, 1301-1305.

Priet, S., Gros, N., Navarro, J.M., Boretto, J., Canard, B., Querat, G. and Sire, J. (2005) HIV-1-associated uracil DNA glycosylase activity controls dUTP misincorporation in viral DNA and is essential to the HIV-1 life cycle. *Mol Cell*, **17**, 479-490.

Puigdomenech, I., Massanella, M., Izquierdo-Useros, N., Ruiz-Hernandez, R., Curriu, M., Bofill, M., Martinez-Picado, J., Juan, M., Clotet, B. and Blanco, J. (2008) HIV transfer between CD4 T cells does not require LFA-1 binding to ICAM-1 and is governed by the interaction of HIV envelope glycoprotein with CD4. *Retrovirology*, **5**, 32.

Qian, S., Zhong, X., Yu, L., Ding, B., de Haan, P. and Boris-Lawrie, K. (2009) HIV-1 Tat RNA silencing suppressor activity is conserved across kingdoms and counteracts translational repression of HIV-1. *Proc Natl Acad Sci U S A*, **106**, 605-610.

Qin, J.Y., Zhang, L., Clift, K.L., Hulur, I., Xiang, A.P., Ren, B.Z. and Lahn, B.T. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One*, **5**, e10611.

Raghavendra, N.K., Shkriabai, N., Graham, R., Hess, S., Kvaratskhelia, M. and Wu, L. (2010) Identification of host proteins associated with HIV-1 preintegration complexes isolated from infected CD4+ cells. *Retrovirology*, **7**, 66.

Rain, J.C., Cribier, A., Gerard, A., Emiliani, S. and Benarous, R. (2009) Yeast two-hybrid detection of integrase-host factor interactions. *Methods*, **47**, 291-297.

Ramezani, A., Hawley, T.S. and Hawley, R.G. (2003) Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood*, **101**, 4717-4724.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M. and Smale, S.T. (2009) A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell*, **138**, 114-128.

Rato, S., Maia, S., Brito, P.M., Resende, L., Pereira, C.F., Moita, C., Freitas, R.P., Moniz-Pereira, J., Hacohen, N., Moita, L.F. and Goncalves, J. (2010) Novel HIV-1 knockdown targets identified by an enriched kinases/phosphatases shRNA library using a long-term iterative screen in Jurkat T-cells. *PLoS One*, **5**, e9276.

Reagan, R.L., Lillie, M.G., Hickman, J.W. and Brueckner, A.L. (1950) Studies of the virus of equine infectious anemia. *Am J Vet Res*, **11**, 157-158.

Recillas-Targa, F., Pikaart, M.J., Burgess-Beusse, B., Bell, A.C., Litt, M.D., West, A.G., Gaszner, M. and Felsenfeld, G. (2002) Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A*, **99**, 6883-6888.

Reeves, M.B., Lehner, P.J., Sissons, J.G. and Sinclair, J.H. (2005a) An in vitro model for the regulation of human cytomegalovirus latency and reactivation in dendritic cells by chromatin remodelling. *J Gen Virol*, **86**, 2949-2954.

Reeves, M.B., MacAry, P.A., Lehner, P.J., Sissons, J.G. and Sinclair, J.H. (2005b) Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. *Proc Natl Acad Sci U S A*, **102**, 4140-4145.

Reeves, W.H. (1985) Use of monoclonal antibodies for the characterization of novel DNA-binding proteins recognized by human autoimmune sera. *J Exp Med*, **161**, 18-39.

Reiser, J., Harmison, G., Kluepfel-Stahl, S., Brady, R.O., Karlsson, S. and Schubert, M. (1996) Transduction of nondividing cells using pseudotyped defective high-titer HIV type 1 particles. *Proc Natl Acad Sci U S A*, **93**, 15266-15271.

Riha, K., Heacock, M.L. and Shippen, D.E. (2006) The role of the nonhomologous end-joining DNA double-strand break repair pathway in telomere biology. *Annu Rev Genet*, **40**, 237-277.

Rinke, C.S., Boyer, P.L., Sullivan, M.D., Hughes, S.H. and Linial, M.L. (2002) Mutation of the catalytic domain of the foamy virus reverse transcriptase leads to loss of processivity and infectivity. *J Virol*, **76**, 7560-7570.

Rittner, K., Churcher, M.J., Gait, M.J. and Karn, J. (1995) The human immunodeficiency virus long terminal repeat includes a specialised initiator element which is required for Tat-responsive transcription. *J Mol Biol*, **248**, 562-580.

Rivera-Calzada, A., Spagnolo, L., Pearl, L.H. and Llorca, O. (2007) Structural model of full-length human Ku70-Ku80 heterodimer and its recognition of DNA and DNA-PKcs. *EMBO Rep*, **8**, 56-62.

Roberts, M.R., Han, Y., Fienberg, A., Hunihan, L. and Ruddle, F.H. (1994) A DNA-binding activity, TRAC, specific for the TRA element of the transferrin receptor gene copurifies with the Ku autoantigen. *Proc Natl Acad Sci U S A*, **91**, 6354-6358.

Robinson, P.J., Fairall, L., Huynh, V.A. and Rhodes, D. (2006) EM measurements define the dimensions of the "30-nm" chromatin fiber: evidence for a compact, interdigitated structure. *Proc Natl Acad Sci U S A*, **103**, 6506-6511.

Romerio, F., Gabriel, M.N. and Margolis, D.M. (1997) Repression of human immunodeficiency virus type 1 through the novel cooperation of human factors YY1 and LSF. *J Virol*, **71**, 9375-9382.

Rosenberg, S.A., Aebersold, P., Cornetta, K., Kasid, A., Morgan, R.A., Moen, R., Karson, E.M., Lotze, M.T., Yang, J.C., Topalian, S.L. and et al. (1990) Gene transfer into humans--immunotherapy of patients with advanced melanoma, using tumor-infiltrating lymphocytes modified by retroviral gene transduction. *N Engl J Med*, **323**, 570-578.

Roshal, M., Kim, B., Zhu, Y., Nghiem, P. and Planelles, V. (2003) Activation of the ATR-mediated DNA damage response by the HIV-1 viral protein R. *J Biol Chem*, **278**, 25879-25886.

Rous, P. (1911) A Sarcoma of the Fowl Transmissible by an Agent Separable from the Tumor Cells. *J Exp Med*, **13**, 397-411.

Ryu, B.Y., Persons, D.A., Evans-Galea, M.V., Gray, J.T. and Nienhuis, A.W. (2007) A chromatin insulator blocks interactions between globin regulatory elements and cellular promoters in erythroid cells. *Blood Cells Mol Dis*, **39**, 221-228.

Saad, J.S., Miller, J., Tai, J., Kim, A., Ghanam, R.H. and Summers, M.F. (2006) Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly. *Proc Natl Acad Sci U S A*, **103**, 11364-11369.

Sachdeva, G., D'Costa, J., Cho, J.E., Kachapati, K., Choudhry, V. and Arya, S.K. (2007) Chimeric HIV-1 and HIV-2 lentiviral vectors with added safety insurance. *J Med Virol*, **79**, 118-126.

Saenz, D.T., Loewen, N., Peretz, M., Whitwam, T., Barraza, R., Howell, K.G., Holmes, J.M., Good, M. and Poeschla, E.M. (2004) Unintegrated lentivirus DNA persistence and accessibility to expression in nondividing cells: analysis with class I integrase mutants. *J Virol*, **78**, 2906-2920.

Sakuma, T., Barry, M.A. and Ikeda, Y. (2012) Lentiviral vectors: basic to translational. *Biochem J*, **443**, 603-618.

Sakurai, Y., Komatsu, K., Agematsu, K. and Matsuoka, M. (2009) DNA double strand break repair enzymes function at multiple steps in retroviral infection. *Retrovirology*, **6**, 114.

Sato, H., Orenstein, J., Dimitrov, D. and Martin, M. (1992) Cell-to-cell spread of HIV-1 occurs within minutes and may not involve the participation of virus particles. *Virology*, **186**, 712-724.

Saxonov, S., Berg, P. and Brutlag, D.L. (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A*, **103**, 1412-1417.

Schambach, A., Galla, M., Maetzig, T., Loew, R. and Baum, C. (2007) Improving transcriptional termination of self-inactivating gamma-retroviral and lentiviral vectors. *Mol Ther*, **15**, 1167-1173.

Schambach, A., Swaney, W.P. and van der Loo, J.C. (2009) Design and production of retro- and lentiviral vectors for gene expression in hematopoietic cells. *Methods Mol Biol*, **506**, 191-205.

Scharfmann, R., Axelrod, J.H. and Verma, I.M. (1991) Long-term in vivo expression of retrovirus-mediated gene transfer in mouse fibroblast implants. *Proc Natl Acad Sci U S A*, **88**, 4626-4630.

Schroder, A.R., Shinn, P., Chen, H., Berry, C., Ecker, J.R. and Bushman, F. (2002) HIV-1 integration in the human genome favors active genes and local hotspots. *Cell*, **110**, 521-529.
Schrofelbauer, B., Hakata, Y. and Landau, N.R. (2007) HIV-1 Vpr function is mediated by interaction with the damage-specific DNA-binding protein DDB1. *Proc Natl Acad Sci U S A*, **104**, 4130-4135.
Schweizer, M., Falcone, V., Gange, J., Turek, R. and Neumann-Haefelin, D. (1997) Simian foamy virus isolated from an accidentally infected human individual. *J Virol*, **71**, 4821-4824.
Schweizer, M. and Merten, O.W. (2010) Large-scale production means for the manufacturing of lentiviral vectors. *Curr Gene Ther*, **10**, 474-486.
Serafini, M., Naldini, L. and Introna, M. (2004) Molecular evidence of inefficient transduction of proliferating human B lymphocytes by VSV-pseudotyped HIV-1-derived lentivectors. *Virology*, **325**, 413-424.
Sharma, A., Awasthi, S., Harrod, C.K., Matlock, E.F., Khan, S., Xu, L., Chan, S., Yang, H., Thammavaram, C.K., Rasor, R.A., Burns, D.K., Skiest, D.J., Van Lint, C., Girard, A.M., McGee, M., Monnat, R.J., Jr. and Harrod, R. (2007) The Werner syndrome helicase is a cofactor for HIV-1 long terminal repeat transactivation and retroviral replication. *J Biol Chem*, **282**, 12048-12057.
Sharma, G., Kaur, G. and Mehra, N. (2011) Genetic correlates influencing immunopathogenesis of HIV infection. *Indian J Med Res*, **134**, 749-768.
Shen, L. and Siliciano, R.F. (2008) Viral reservoirs, residual viremia, and the potential of highly active antiretroviral therapy to eradicate HIV infection. *J Allergy Clin Immunol*, **122**, 22-28.
Shi, L., Qiu, D., Zhao, G., Corthesy, B., Lees-Miller, S., Reeves, W.H. and Kao, P.N. (2007) Dynamic binding of Ku80, Ku70 and NF90 to the IL-2 promoter in vivo in activated T-cells. *Nucleic Acids Res*, **35**, 2302-2310.
Shoemaker, C., Hoffman, J., Goff, S.P. and Baltimore, D. (1981) Intramolecular integration within Moloney murine leukemia virus DNA. *J Virol*, **40**, 164-172.
Shun, M.C., Raghavendra, N.K., Vandegraaff, N., Daigle, J.E., Hughes, S., Kellam, P., Cherepanov, P. and Engelman, A. (2007) LEDGF/p75 functions downstream from preintegration complex formation to effect gene-specific HIV-1 integration. *Genes Dev*, **21**, 1767-1778.
Sinclair, A., Yarranton, S. and Schelcher, C. (2006) DNA-damage response pathways triggered by viral replication. *Expert Rev Mol Med*, **8**, 1-11.
Sinclair, J. and Sissons, P. (1996) Latent and persistent infections of monocytes and macrophages. *Intervirology*, **39**, 293-301.
Singer-Sam, J., Keith, D.H., Tani, K., Simmer, R.L., Shively, L., Lindsay, S., Yoshida, A. and Riggs, A.D. (1984) Sequence of the promoter region of the gene for human X-linked 3-phosphoglycerate kinase. *Gene*, **32**, 409-417.
Singleton, B.K., Priestley, A., Steingrimsdottir, H., Gell, D., Blunt, T., Jackson, S.P., Lehmann, A.R. and Jeggo, P.A. (1997) Molecular and biochemical characterization of xrs mutants defective in Ku80. *Mol Cell Biol*, **17**, 1264-1273.
Sirven, A., Pflumio, F., Zennou, V., Titeux, M., Vainchenker, W., Coulombel, L., Dubart-Kupperschmitt, A. and Charneau, P. (2000) The human immunodeficiency virus type-1 central DNA flap is a crucial determinant for lentiviral vector nuclear import and gene transduction of human hematopoietic stem cells. *Blood*, **96**, 4103-4110.
Siva, A.C. and Bushman, F. (2002) Poly(ADP-ribose) polymerase 1 is not strictly required for infection of murine cells by retroviruses. *J Virol*, **76**, 11904-11910.
Skalka, A.M. and Katz, R.A. (2005) Retroviral DNA integration and the DNA damage response. *Cell Death Differ*, **12 Suppl 1**, 971-978.
Sloan, R.D. and Wainberg, M.A. (2011) The role of unintegrated DNA in HIV infection. *Retrovirology*, **8**, 52.
Smale, S.T. and Baltimore, D. (1989) The "initiator" as a transcription control element. *Cell*, **57**, 103-113.
Smith, J.A. and Daniel, R. (2006) Following the path of the virus: the exploitation of host DNA repair mechanisms by retroviruses. *ACS Chem Biol*, **1**, 217-226.
Smith, J.A., Nunnari, G., Preuss, M., Pomerantz, R.J. and Daniel, R. (2007) Pentoxifylline suppresses transduction by HIV-1-based vectors. *Intervirology*, **50**, 377-386.

Smith, J.A., Wang, F.X., Zhang, H., Wu, K.J., Williams, K.J. and Daniel, R. (2008) Evidence that the Nijmegen breakage syndrome protein, an early sensor of double-strand DNA breaks (DSB), is involved in HIV-1 post-integration repair by recruiting the ataxia telangiectasia-mutated kinase in a process similar to, but distinct from, cellular DSB repair. *Virol J*, **5**, 11.

Soderberg-Naucler, C., Streblow, D.N., Fish, K.N., Allan-Yorke, J., Smith, P.P. and Nelson, J.A. (2001) Reactivation of latent human cytomegalovirus in CD14(+) monocytes is differentiation dependent. *J Virol*, **75**, 7543-7554.

Srinivasakumar, N., Chazal, N., Helga-Maria, C., Prasad, S., Hammarskjold, M.L. and Rekosh, D. (1997) The effect of viral regulatory protein expression on gene delivery by human immunodeficiency virus type 1 vectors produced in stable packaging cell lines. *J Virol*, **71**, 5841-5848.

Steffen, D. (1984) Proviruses are adjacent to c-myc in some murine leukemia virus-induced lymphomas. *Proc Natl Acad Sci U S A*, **81**, 2097-2101.

Stinski, M.F. and Isomura, H. (2008) Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. *Med Microbiol Immunol*, **197**, 223-231.

Stinski, M.F., Mocarski, E.S. and Thomsen, D.R. (1979) DNA of human cytomegalovirus: size heterogeneity and defectiveness resulting from serial undiluted passage. *J Virol*, **31**, 231-239.

Stoltzfus, C.M. and Madsen, J.M. (2006) Role of viral splicing elements and cellular RNA binding proteins in regulation of HIV-1 alternative RNA splicing. *Curr HIV Res*, **4**, 43-55.

Strang, B.L., Ikeda, Y., Cosset, F.L., Collins, M.K. and Takeuchi, Y. (2004) Characterization of HIV-1 vectors with gammaretrovirus envelope glycoproteins produced from stable packaging cells. *Gene Ther*, **11**, 591-598.

Studamire, B. and Goff, S.P. (2008) Host proteins interacting with the Moloney murine leukemia virus integrase: multiple transcriptional regulators and chromatin binding factors. *Retrovirology*, **5**, 48.

Sucharov, C.C., Helmke, S.M., Langer, S.J., Perryman, M.B., Bristow, M. and Leinwand, L. (2004) The Ku protein complex interacts with YY1, is up-regulated in human heart failure, and represses alpha myosin heavy-chain gene expression. *Mol Cell Biol*, **24**, 8705-8715.

Sundquist, W.I. and Krausslich, H.G. (2012) HIV-1 Assembly, Budding, and Maturation. *Cold Spring Harb Perspect Med*, **2**.

Sune, C. and Garcia-Blanco, M.A. (1995) Sp1 transcription factor is required for in vitro basal and Tat-activated transcription from the human immunodeficiency virus type 1 long terminal repeat. *J Virol*, **69**, 6572-6576.

Suzuki, Y. and Craigie, R. (2007) The road to chromatin - nuclear entry of retroviruses. *Nat Rev Microbiol*, **5**, 187-196.

Tabin, C.J., Hoffmann, J.W., Goff, S.P. and Weinberg, R.A. (1982) Adaptation of a retrovirus as a eucaryotic vector transmitting the herpes simplex virus thymidine kinase gene. *Mol Cell Biol*, **2**, 426-436.

Tachiwana, H., Shimura, M., Nakai-Murakami, C., Tokunaga, K., Takizawa, Y., Sata, T., Kurumizaka, H. and Ishizaka, Y. (2006) HIV-1 Vpr induces DNA double-strand breaks. *Cancer Res*, **66**, 627-631.

Tan-Wong, S.M., French, J.D., Proudfoot, N.J. and Brown, M.A. (2008) Dynamic interactions between the promoter and terminator regions of the mammalian BRCA1 gene. *Proc Natl Acad Sci U S A*, **105**, 5160-5165.

Tanaka, A., Takahashi, C., Yamaoka, S., Nosaka, T., Maki, M. and Hatanaka, M. (1990) Oncogenic transformation by the tax gene of human T-cell leukemia virus type I in vitro. *Proc Natl Acad Sci U S A*, **87**, 1071-1075.

Tang, L.Y. and Zhang, J. (2000) The cellular mismatch repair system is able to repair mismatches within MLV retroviral double-stranded DNA at a low frequency. *Nucleic Acids Res*, **28**, 2302-2306.

Teoh, G., Urashima, M., Greenfield, E.A., Nguyen, K.A., Lee, J.F., Chauhan, D., Ogata, A., Treon, S.P. and Anderson, K.C. (1998) The 86-kD subunit of Ku autoantigen mediates homotypic and heterotypic adhesion of multiple myeloma cells. *J Clin Invest*, **101**, 1379-1388.

Thomson, J.P., Skene, P.J., Selfridge, J., Clouaire, T., Guy, J., Webb, S., Kerr, A.R., Deaton, A., Andrews, R., James, K.D., Turner, D.J., Illingworth, R. and Bird, A. (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. *Nature*, **464**, 1082-1086.

Ting, N.S., Chan, D.W., Lintott, L.G., Allalunis-Turner, J. and Lees-Miller, S.P. (1999) Protein-DNA complexes containing DNA-dependent protein kinase in crude extracts from human and rodent cells. *Radiat Res*, **151**, 414-422.

Ting, N.S., Yu, Y., Pohorelic, B., Lees-Miller, S.P. and Beattie, T.L. (2005) Human Ku70/80 interacts directly with hTR, the RNA component of human telomerase. *Nucleic Acids Res*, **33**, 2090-2098.

Trobridge, G.D., Miller, D.G., Jacobs, M.A., Allen, J.M., Kiem, H.P., Kaul, R. and Russell, D.W. (2006) Foamy virus vector integration sites in normal human cells. *Proc Natl Acad Sci U S A*, **103**, 1498-1503.

Trono, D. (1992) Partial reverse transcripts in virions from human immunodeficiency and murine leukemia viruses. *J Virol*, **66**, 4893-4900.

Trono, D., Van Lint, C., Rouzioux, C., Verdin, E., Barre-Sinoussi, F., Chun, T.W. and Chomont, N. (2010) HIV persistence and the prospect of long-term drug-free remissions for HIV-infected individuals. *Science*, **329**, 174-180.

Tuteja, R. and Tuteja, N. (2000) Ku autoantigen: a multifunctional DNA-binding protein. *Crit Rev Biochem Mol Biol*, **35**, 1-33.

Tyagi, M. and Karn, J. (2007) CBF-1 promotes transcriptional silencing during the establishment of HIV-1 latency. *Embo J*, **26**, 4985-4995.

Tyagi, S., Ochem, A. and Tyagi, M. (2011) DNA-dependent protein kinase interacts functionally with the RNA polymerase II complex recruited at the human immunodeficiency virus (HIV) long terminal repeat and plays an important role in HIV gene expression. *J Gen Virol*, **92**, 1710-1720.

Uchida, N., Washington, K.N., Lap, C.J., Hsieh, M.M. and Tisdale, J.F. (2011) Chicken HS4 insulators have minimal barrier function among progeny of human hematopoietic cells transduced with an HIV1-based lentiviral vector. *Mol Ther*, **19**, 133-139.

Urisman, A., Molinaro, R.J., Fischer, N., Plummer, S.J., Casey, G., Klein, E.A., Malathi, K., Magi-Galluzzi, C., Tubbs, R.R., Ganem, D., Silverman, R.H. and DeRisi, J.L. (2006) Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS Pathog*, **2**, e25.

Valen, E. and Sandelin, A. (2011) Genomic and chromatin signals underlying transcription start-site selection. *Trends Genet*, **27**, 475-485.

van Ingen, H., van Schaik, F.M., Wienk, H., Ballering, J., Rehmann, H., Dechesne, A.C., Kruijzer, J.A., Liskamp, R.M., Timmers, H.T. and Boelens, R. (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure*, **16**, 1245-1256.

Van Lint, C., Amella, C.A., Emiliani, S., John, M., Jie, T. and Verdin, E. (1997) Transcription factor binding sites downstream of the human immunodeficiency virus type 1 transcription start site are important for virus infectivity. *J Virol*, **71**, 6113-6127.

Van Maele, B., De Rijck, J., De Clercq, E. and Debyser, Z. (2003) Impact of the central polypurine tract on the kinetics of human immunodeficiency virus type 1 vector transduction. *J Virol*, **77**, 4685-4694.

Varatharajan, L. and Thomas, S.A. (2009) The transport of anti-HIV drugs across blood-CNS interfaces: summary of current knowledge and recommendations for further research. *Antiviral Res*, **82**, A99-109.

Vargas, J., Jr., Gusella, G.L., Najfeld, V., Klotman, M.E. and Cara, A. (2004) Novel integrase-defective lentiviral episomal vectors for gene transfer. *Hum Gene Ther*, **15**, 361-372.

Verdin, E. (1991) DNase I-hypersensitive sites are associated with both long terminal repeats and with the intragenic enhancer of integrated human immunodeficiency virus type 1. *J Virol*, **65**, 6790-6799.

Verdin, E., Paras, P., Jr. and Van Lint, C. (1993) Chromatin disruption in the promoter of human immunodeficiency virus type 1 during transcriptional activation. *Embo J*, **12**, 3249-3259.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M. and Timmers, H.T. (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell*, **131**, 58-69.

Vogelmann, J., Valeri, A., Guillou, E., Cuvier, O. and Nollmann, M. (2011) Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. *Nucleus*, **2**, 358-369.

Vogt, V.M. (1997) Retroviral Virions and Genomes.

Walker, J.R., Corpina, R.A. and Goldberg, J. (2001) Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature*, **412**, 607-614.

Wang, G.P., Ciuffi, A., Leipzig, J., Berry, C.C. and Bushman, F.D. (2007) HIV integration site selection: analysis by massively parallel pyrosequencing reveals association with epigenetic modifications. *Genome Res*, **17**, 1186-1194.

Wang, R., Liang, J., Jiang, H., Qin, L.J. and Yang, H.T. (2008a) Promoter-dependent EGFP expression during embryonic stem cell propagation and differentiation. *Stem Cells Dev*, **17**, 279-289.

Wang, Y., Ghosh, G. and Hendrickson, E.A. (2009) Ku86 represses lethal telomere deletion events in human somatic cells. *Proc Natl Acad Sci U S A*, **106**, 12430-12435.

Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q. and Zhao, K. (2008b) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*, **40**, 897-903.

Waninger, S., Kuhen, K., Hu, X., Chatterton, J.E., Wong-Staal, F. and Tang, H. (2004) Identification of cellular cofactors for human immunodeficiency virus replication via a ribozyme-based genomics approach. *J Virol*, **78**, 12829-12837.

Wanisch, K. and Yanez-Munoz, R.J. (2009) Integration-deficient lentiviral vectors: a slow coming of age. *Mol Ther*, **17**, 1316-1332.

Warnock, J.N., Daigre, C. and Al-Rubeai, M. (2011) Introduction to viral vectors. *Methods Mol Biol*, **737**, 1-25.

Warriar, N., Page, N. and Govindan, M.V. (1996) Expression of human glucocorticoid receptor gene and interaction of nuclear proteins with the transcriptional control element. *J Biol Chem*, **271**, 18662-18671.

Warrilow, D., Tachedjian, G. and Harrich, D. (2009) Maturation of the HIV reverse transcription complex: putting the jigsaw together. *Rev Med Virol*, **19**, 324-337.

Wathen, M.W. and Stinski, M.F. (1982) Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *J Virol*, **41**, 462-477.

Watts, J.M., Dang, K.K., Gorelick, R.J., Leonard, C.W., Bess, J.W., Jr., Swanstrom, R., Burch, C.L. and Weeks, K.M. (2009) Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature*, **460**, 711-716.

Weinberg, J.B., Matthews, T.J., Cullen, B.R. and Malim, M.H. (1991) Productive human immunodeficiency virus type 1 (HIV-1) infection of nonproliferating human monocytes. *J Exp Med*, **174**, 1477-1482.

Weitzman, M.D., Lilley, C.E. and Chaurushiya, M.S. (2010) Genomes in conflict: maintaining genome integrity during virus infection. *Annu Rev Microbiol*, **64**, 61-81.

Westerman, K.A., Ao, Z., Cohen, E.A. and Leboulch, P. (2007) Design of a trans protease lentiviral packaging system that produces high titer virus. *Retrovirology*, **4**, 96.

White, S.M., Renda, M., Nam, N.Y., Klimatcheva, E., Zhu, Y., Fisk, J., Halterman, M., Rimel, B.J., Federoff, H., Pandya, S., Rosenblatt, J.D. and Planelles, V. (1999) Lentivirus vectors using human and simian immunodeficiency virus elements. *J Virol*, **73**, 2832-2840.

Willard, H.F., Goss, S.J., Holmes, M.T. and Munroe, D.L. (1985) Regional localization of the phosphoglycerate kinase gene and pseudogene on the human X chromosome and assignment of a related DNA sequence to chromosome 19. *Hum Genet*, **71**, 138-143.

Williams, S.A., Chen, L.F., Kwon, H., Ruiz-Jarabo, C.M., Verdin, E. and Greene, W.C. (2006) NF-kappaB p50 promotes HIV latency through HDAC recruitment and repression of transcriptional initiation. *Embo J*, **25**, 139-149.

Williamson, E.A., Farrington, J., Martinez, L., Ness, S., O'Rourke, J., Lee, S.H., Nickoloff, J. and Hromas, R. (2008) Expression levels of the human DNA repair protein metnase influence lentiviral genomic integration. *Biochimie*, **90**, 1422-1426.

Wolf, D. and Goff, S.P. (2008) Host restriction factors blocking retroviral replication. *Annu Rev Genet*, **42**, 143-163.

Wong, P.K. (1990) Moloney murine leukemia virus temperature-sensitive mutants: a model for retrovirus-induced neurologic disorders. *Curr Top Microbiol Immunol*, **160**, 29-60.

Woodward, C.L., Prakobwanakit, S., Mosessian, S. and Chow, S.A. (2009) Integrase interacts with nucleoporin NUP153 to mediate the nuclear import of human immunodeficiency virus type 1. *J Virol*, **83**, 6522-6533.

Wu, X., Li, Y., Crise, B. and Burgess, S.M. (2003) Transcription start regions in the human genome are favored targets for MLV integration. *Science*, **300**, 1749-1751.

Wu, X., Wakefield, J.K., Liu, H., Xiao, H., Kralovics, R., Prchal, J.T. and Kappes, J.C. (2000) Development of a novel trans-lentiviral vector that affords predictable safety. *Mol Ther*, **2**, 47-55.

Wu, Y. and Marsh, J.W. (2001) Selective transcription and modulation of resting T cell activity by preintegrated HIV DNA. *Science*, **293**, 1503-1506.

Wu, Y. and Marsh, J.W. (2003) Early transcription from nonintegrated DNA in human immunodeficiency virus infection. *J Virol*, **77**, 10376-10382.

Wutz, A. (2011) Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet*, **12**, 542-553.

Xia, X., Zhang, Y., Zieth, C.R. and Zhang, S.C. (2007) Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. *Stem Cells Dev*, **16**, 167-176.

Xie, B., Invernizzi, C.F., Richard, S. and Wainberg, M.A. (2007) Arginine methylation of the human immunodeficiency virus type 1 Tat protein by PRMT6 negatively affects Tat Interactions with both cyclin T1 and the Tat transactivation region. *J Virol*, **81**, 4226-4234.

Xu, P., LaVallee, P.A., Lin, J.J. and Hoidal, J.R. (2004) Characterization of proteins binding to E-box/Ku86 sites and function of Ku86 in transcriptional regulation of the human xanthine oxidoreductase gene. *J Biol Chem*, **279**, 16057-16063.

Yan, N. and Chen, Z.J. (2012) Intrinsic antiviral immunity. *Nat Immunol*, **13**, 214-222.

Yan, N., Cherepanov, P., Daigle, J.E., Engelman, A. and Lieberman, J. (2009) The SET complex acts as a barrier to autointegration of HIV-1. *PLoS Pathog*, **5**, e1000327.

Yang, Y.X., Guen, V., Richard, J., Cohen, E.A. and Berthoux, L. (2010) Cell context-dependent involvement of ATR in early stages of retroviral replication. *Virology*, **396**, 272-279.

Yannaki, E., Tubb, J., Aker, M., Stamatoyannopoulos, G. and Emery, D.W. (2002) Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. *Mol Ther*, **5**, 589-598.

Yoder, K., Sarasin, A., Kraemer, K., McIlhatton, M., Bushman, F. and Fishel, R. (2006) The DNA repair genes XPB and XPD defend cells from retroviral infection. *Proc Natl Acad Sci U S A*, **103**, 4622-4627.

Yoder, K.E. and Bushman, F.D. (2000) Repair of gaps in retroviral DNA integration intermediates. *J Virol*, **74**, 11191-11200.

Yoder, K.E., Roddick, W., Hoellerbauer, P. and Fishel, R. (2011) XPB mediated retroviral cDNA degradation coincides with entry to the nucleus. *Virology*, **410**, 291-298.

Yu, D., Wang, W., Yoder, A., Spear, M. and Wu, Y. (2009) The HIV envelope but not VSV glycoprotein is capable of mediating HIV latent infection of resting CD4 T cells. *PLoS Pathog*, **5**, e1000633.

Yu, S.F., von Ruden, T., Kantoff, P.W., Garber, C., Seiberg, M., Ruther, U., Anderson, W.F., Wagner, E.F. and Gilboa, E. (1986) Self-inactivating retroviral vectors designed for transfer of whole genes into mammalian cells. *Proc Natl Acad Sci U S A*, **83**, 3194-3198.

Zaiss, A.K., Son, S. and Chang, L.J. (2002) RNA 3' readthrough of oncoretrovirus and lentivirus: implications for vector safety and efficacy. *J Virol*, **76**, 7209-7219.

Zaitseva, L., Cherepanov, P., Leyens, L., Wilson, S.J., Rasaiyaah, J. and Fassati, A. (2009) HIV-1 exploits importin 7 to maximize nuclear import of its DNA genome. *Retrovirology*, **6**, 11.

Zanta-Boussif, M.A., Charrier, S., Brice-Ouzet, A., Martin, S., Opolon, P., Thrasher, A.J., Hope, T.J. and Galy, A. (2009) Validation of a mutated PRE sequence allowing high and sustained transgene expression while abrogating WHV-X protein synthesis: application to the gene therapy of WAS. *Gene Ther*, **16**, 605-619.

Zhang, X.Y., La Russa, V.F. and Reiser, J. (2004) Transduction of bone-marrow-derived mesenchymal stem cells by using lentivirus vectors pseudotyped with modified RD114 envelope glycoproteins. *J Virol*, **78**, 1219-1229.

Zhang, Y.A., Maitra, A., Hsieh, J.T., Rudin, C.M., Peacock, C.D., Karikari, C., Brekken, R.A., Stastny, V., Gao, B., Girard, L., Wistuba, I., Frenkel, E., Minna, J.D. and Gazdar, A.F. (2011) Frequent detection of infectious xenotropic murine leukemia virus (XMLV) in human cultures established from mouse xenografts. *Cancer Biol Ther*, **12**, 617-628.

Zheng, Y., Ao, Z., Wang, B., Jayappa, K.D. and Yao, X. (2011) Host protein Ku70 binds and protects HIV-1 integrase from proteasomal degradation and is required for HIV replication. *J Biol Chem*, **286**, 17722-17735.

Zhou, H., Xu, M., Huang, Q., Gates, A.T., Zhang, X.D., Castle, J.C., Stec, E., Ferrer, M., Strulovici, B., Hazuda, D.J. and Espeseth, A.S. (2008) Genome-scale RNAi screen for host factors required for HIV replication. *Cell Host Microbe*, **4**, 495-504.

Zhou, M., Halanski, M.A., Radonovich, M.F., Kashanchi, F., Peng, J., Price, D.H. and Brady, J.N. (2000) Tat modifies the activity of CDK9 to phosphorylate serine 5 of the RNA polymerase II carboxyl-terminal domain during human immunodeficiency virus type 1 transcription. *Mol Cell Biol*, **20**, 5077-5086.

Zhou, S., Mody, D., DeRavin, S.S., Hauer, J., Lu, T., Ma, Z., Hacein-Bey Abina, S., Gray, J.T., Greene, M.R., Cavazzana-Calvo, M., Malech, H.L. and Sorrentino, B.P. (2010) A self-inactivating lentiviral vector for SCID-X1 gene therapy that does not activate LMO2 expression in human T cells. *Blood*, **116**, 900-908.

Zhou, Y., Zhang, H., Siliciano, J.D. and Siliciano, R.F. (2005) Kinetics of human immunodeficiency virus type 1 decay following entry into resting CD4+ T cells. *J Virol*, **79**, 2199-2210.

Ziegler, L., Yang, L., Joo, K., Yang, H., Baltimore, D. and Wang, P. (2008) Targeting lentiviral vectors to antigen-specific immunoglobulins. *Hum Gene Ther*, **19**, 861-872.

Zimmerman, E.S., Chen, J., Andersen, J.L., Ardon, O., Dehart, J.L., Blackett, J., Choudhary, S.K., Camerini, D., Nghiem, P. and Planelles, V. (2004) Human immunodeficiency virus type 1 Vpr-mediated G2 arrest requires Rad17 and Hus1 and induces nuclear BRCA1 and gamma-H2AX focus formation. *Mol Cell Biol*, **24**, 9286-9294.

Zimmerman, E.S., Sherman, M.P., Blackett, J.L., Neidleman, J.A., Kreis, C., Mundt, P., Williams, S.A., Warmerdam, M., Kahn, J., Hecht, F.M., Grant, R.M., de Noronha, C.M., Weyrich, A.S., Greene, W.C. and Planelles, V. (2006) Human immunodeficiency virus type 1 Vpr induces DNA replication stress in vitro and in vivo. *J Virol*, **80**, 10407-10418.

Zufferey, R., Donello, J.E., Trono, D. and Hope, T.J. (1999) Woodchuck hepatitis virus posttranscriptional regulatory element enhances expression of transgenes delivered by retroviral vectors. *J Virol*, **73**, 2886-2892.

Zufferey, R., Dull, T., Mandel, R.J., Bukovsky, A., Quiroz, D., Naldini, L. and Trono, D. (1998) Self-inactivating lentivirus vector for safe and efficient in vivo gene delivery. *J Virol*, **72**, 9873-9880.

Zufferey, R., Nagy, D., Mandel, R.J., Naldini, L. and Trono, D. (1997) Multiply attenuated lentiviral vector achieves efficient gene delivery in vivo. *Nat Biotechnol*, **15**, 871-875.

Zychlinski, D., Schambach, A., Modlich, U., Maetzig, T., Meyer, J., Grassman, E., Mishra, A. and Baum, C. (2008) Physiological promoters reduce the genotoxic risk of integrating gene vectors. *Mol Ther*, **16**, 718-725.

RESUME

Une meilleure connaissance des déterminants viraux et cellulaires impliqués dans la régulation de l'expression lentivirale est essentielle pour (*i*) maitriser l'expression d'un transgène dans le cadre d'un transfert de gène et (*ii*) mieux comprendre l'absence/la perte de l'expression du virus de l'immunodéficience humaine de type 1 (VIH-1) observé en cas de latence virale après infection. Dans ce contexte, les facteurs cellulaires impliqués dans la reconnaissance et la réparation des cassures de l'ADN, activés dès les étapes précoces du cycle viral, pourraient participer au contrôle de l'expression rétrovirale.

Dans la première partie de cette étude, nous avons généré une collection de vecteurs lentiviraux codant le transgène *green fluorescent protein* (*gfp*) avec des *design* variés : promoteur du VIH-1 [*long terminal repeat* (LTRs)] ou d'origine hétérologue [*e.g.*, promoteur viral CMV (*cytomegalovirus*), ou humain PGK (*phosphoglycerate kinase*)], intégrase native ou mutée, LTRs natifs ou *self inactivating* (SIN). En particulier, nous avons caractérisé l'impact de l'insertion de différentes séquences hétérologues au sein des SIN-LTRs sur l'expression du transgène *gfp* au cours du temps dans un contexte compétent ou déficient pour l'intégration. Nous avons mis en évidence un phénomène de modulation du niveau d'expression du transgène (d'un facteur 0,3 à 1,8; comparé aux vecteurs sans insertion) qui est dépendant de la séquence insérée et de son orientation. L'expression transgénique résulterait donc d'une balance coûts/bénéfices associés à l'insertion d'éléments aux extrémités des vecteurs qui pourrait déterminer le niveau d'expression du transgène à partir de vecteurs modifiés.

Dans la seconde partie, nous avons réalisé une analyse systématique et comparative de l'expression lentivirale au sein des cellules humaines de carcinome de côlon HCT 116 déplétées pour le complexe Ku. Ce complexe, qui est essentiel à la survie chez l'homme et est impliqué dans les processus de réparation de l'ADN, a été identifié comme une cible anti-VIH potentielle. Nous avons mis en évidence que la déplétion en Ku induit une diminution de l'expression précoce du VIH-1 de façon spécifique du LTR et indépendante de Tat. Même si l'action de Ku nécessite l'intégration du VIH-1, sa déplétion ne modifie pas l'efficacité d'intégration virale mais agit plutôt au niveau transcriptionnel. Da façon importante, la réactivation de l'expression du transgène après traitement par l'activateur de NF-κB (*Nuclear Factor κB*), le TNFα (*tumor necrosis factor α*) ou un inhibiteur des déacétylases d'histones, la trichostatine A est favorisée par la déplétion en Ku. A l'inverse, en présence d'un niveau normal en Ku, les cellules exprimant le VIH-1 seraient contre-sélectionnées dans le temps. Ainsi, la déplétion en Ku pourrait promouvoir l'établissement d'un état (reactivable) de latence transcriptionelle associé à une moindre contre-sélection des cellules transduites.

Les résultats issus de ce travail de thèse démontrent que l'expression lentivirale varie en fonction de nombreux paramètres, dont (*i*) le *design* des vecteurs, (*ii*) le type cellulaire transduit et son fond génétique, et (*iii*) le temps écoulé depuis la transduction, reflétant ainsi des interactions différentielles entre le vecteur et son hôte.