

SÉBASTIEN LANDRY

**CARACTÉRISATION ET RÉGULATION DE LA
TRANSCRIPTION ANTISENS CHEZ LE VIH-1 ET
LES RÉTROVIRUS HTLVs**

Thèse présentée
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de doctorat en Microbiologie-Immunologie
pour l'obtention du grade de Philosophiae Doctor (Ph.D.)

DÉPARTEMENT DE BIOLOGIE MÉDICALE
FACULTÉ DE MÉDECINE
UNIVERSITÉ LAVAL
QUÉBEC

2009

© Sébastien Landry, 2009

Résumé

On a longtemps cru que, chez les rétrovirus, la production de la totalité des protéines virales dépendait de l'expression d'un transcrit sens de pleine longueur, initié dans le LTR 5'. Alors que la présence de transcription antisens chez les rétrovirus a déjà été suggérée il y a plusieurs années, des preuves concrètes de son rôle biologique ont plus récemment été apportées grâce à une étude qui mettait en évidence l'identification d'une nouvelle protéine virale chez HTLV-1. En effet, cette protéine nommée HBZ est codée par un transcrit antisens épissé de façon alternative et initié dans le LTR 3'. De façon intéressante, la protéine HBZ est impliquée dans la régulation de la transcription virale et est capable d'inhiber la transcription dépendante du transactivateur viral Tax. Ces nouvelles données laissaient donc croire que la transcription antisens pourrait également exister chez d'autres rétrovirus, compte tenu de la présence d'un ORF conservé dans le brin positif de l'ADN proviral de plusieurs d'entre eux. Les principaux objectifs de ce projet étaient d'étudier la régulation du transcrit antisens chez HTLV-1 et d'évaluer la présence de transcription antisens chez le VIH-1 ainsi que chez les rétrovirus humains HTLV-3 et HTLV-4. Premièrement, les résultats obtenus démontrent que la protéine virale Tax est un facteur important dans le contrôle de la transcription antisens chez HTLV-1 et que son action est dépendante de CREB. De plus, nos résultats suggèrent fortement que les sites d'intégration dans le génome de la cellule hôte ont un impact important sur le niveau d'expression du transcrit antisens en absence ou en présence de Tax. De façon intéressante, des analyses de RT-PCR ont permis de démontrer la présence de transcrits antisens épissés dans les cellules transfectées par des clones moléculaires complets d'HTLV-3 et HTLV-4. Des analyses additionnelles ont pu mettre en évidence la localisation majoritairement nucléaire des protéines codées par ces transcrits. Nos résultats démontrent que ces protéines, nommées APH-3 et APH-4, ont également la capacité d'inhiber l'action de Tax. Dans un second temps, l'utilisation d'un protocole de RT-PCR brin-spécifique a permis la détection de transcrits antisens dans plusieurs lignées cellulaires infectées par le VIH-1. La présence de transcription antisens au cours du cycle de réplication viral a pu être clairement établie par l'utilisation d'un virus pseudotypé contenant le gène rapporteur luciférase positionné en

orientation antisens. La protéine antisens du VIH-1 (ASP) a ensuite été exprimée et détectée efficacement dans les cellules d'insectes SF9 ainsi que dans les cellules de mammifères COS-7 et 293T. Ces nouvelles données indiquent que la transcription antisens est une stratégie répandue chez les rétrovirus et suggèrent qu'une analyse approfondie du phénomène est nécessaire à une meilleure compréhension du cycle de réplication rétroviral.

Abstract

Retroviruses have long been thought to lack antisense transcription and to rely on a single sense transcript for viral gene expression. Even though the presence of antisense transcription in retroviruses has been suggested several years ago, convincing data concerning its biological function has only been presented recently, when several studies demonstrated the existence of a new viral protein in HTLV-1. Unexpectedly, this new protein named HBZ was shown to be encoded by the plus strand of the viral genome through the production of an alternatively spliced antisense transcript initiated in the 3' LTR. Interestingly, HBZ has been shown to be involved in viral and cellular transcription regulation and to potentially inhibit Tax-mediated viral sense transcription. Considering that other members of the retrovirus family also possess a conserved ORF on their antisense strand, these data were thus suggesting that this phenomenon could be broadly distributed. The objectives of this study were to characterize the regulation of the HBZ transcript in HTLV-1 and to evaluate the possible existence of antisense transcription in HTLV-3, HTLV-4 and HIV-1. First, our results showed that the viral Tax protein is an important factor controlling the antisense transcription in HTLV-1. Moreover, we demonstrated that the integration site in the host genome importantly affect both the basal and Tax-mediated antisense transcription. Using RT-PCR analysis, we were able to detect the presence of spliced antisense transcripts in cells transfected with HTLV-3 and HTLV-4 molecular clones. Additional experiments demonstrated the predominant nuclear localization of the proteins encoded by these transcripts. Transfection experiments highlighted the Tax-inhibiting capacity of these new proteins, named APH-3 and APH-4. Second, we show that the utilization of a strand-specific RT-PCR protocol allows the specific detection of antisense transcripts in HIV-1 infected cells. The generation of an HIV-1 proviral clone containing the luciferase reporter gene in the antisense orientation clearly demonstrated for the first time the existence of antisense transcription during the replication cycle. We further showed that the antisense-encoded ASP protein could successfully be expressed and detected in SF9 insect cells and in 293T and COS-7 mammalian cells. These data strongly suggest that antisense transcription is common in human retroviruses. Further analysis of

the phenomenon is therefore necessary to better understand its implication in the retroviral life cycle.

Avant-propos

Il m'est d'avis que le milieu de la science en est un plutôt ingrat : la quantité d'énergie investie dans un projet de doctorat est rarement proportionnelle à la quantité de résultats obtenus. Celui qui lit ce document aura tôt fait de constater que les objectifs que nous nous sommes fixés au début de ce projet étaient aussi grands que les risques qui s'y rattachaient. Heureusement, l'impact des fréquents moments de déception et de découragement fut habilement compensé par le support incomparable que j'ai eu la chance de recevoir des gens qui m'ont entouré au cours des dernières années. Je tiens d'abord à remercier mon directeur de recherche, le Dr Benoit Barbeau, pour m'avoir donné l'opportunité de m'attaquer à un projet des plus stimulant, mais également pour sa générosité, sa patience et son sang-froid légendaire. J'aimerais ensuite remercier l'ensemble des étudiants et professionnels de recherche qui ont contribué non seulement au succès de ce projet mais aussi à créer un environnement de travail motivant et agréable. J'adresse des remerciements particuliers à Marie-Hélène Cavanagh et Marilène Halin dont le travail acharné fut un élément essentiel à la réussite de ce projet.

Je voudrais en second lieu remercier mon co-directeur de recherche, le Dr Michel J. Tremblay et toute son équipe, notamment pour leur aide précieuse lors du déménagement de nos laboratoires.

Je suis également reconnaissant envers le Dr Eric A. Cohen, chez qui j'ai pu effectuer plusieurs expériences à un moment crucial.

Finalement, merci à ma copine, ma famille et mes amis, sans qui je n'aurais sûrement pas eu la force de terminer ce doctorat !

Table des matières

Chapitre 1 : La famille des <i>Retroviridae</i>	1
1.1 Généralités.....	1
1.2 Organisation génomique	1
1.3 Classification.....	2
1.4 Cycle réplicatif.....	4
1.5 Structure des rétrovirus	9
1.6 Les rétrovirus HTLVs	10
1.6.1 Le rétrovirus HTLV-1.....	11
1.6.2 La leucémie à cellules T de l'adulte (ATL)	12
1.6.3 La paraparésie spastique tropicale.....	14
1.6.4 Désordres neurologiques associés à HTLV-II	15
1.6.5 Stratégie répllicative des rétrovirus HTLVs.....	16
1.6.6 Les séquences LTR	17
1.6.7 Tax et la régulation transcriptionnelle chez HTLV-1	18
1.6.8 Tax et le cycle cellulaire.....	20
1.6.9 Tax interfère avec les mécanismes de réparation de l'ADN.....	21
1.6.10 La protéine Rex	22
1.6.11 Les protéines accessoires	23
1.7 Le virus de l'immunodéficience humaine de type 1	24
1.7.1 Premières manifestations du SIDA et découverte du VIH	24
1.7.2 Infection par le VIH-1 et syndrome d'immunodéficience acquise (SIDA).....	25
1.7.3 Tropisme et répllication virale	26
1.7.4 Traitements	28
1.7.5 Origines du VIH	29
1.7.6 Épidémiologie et diversité génétique	29
1.7.7 Stratégie répllicative	31
1.7.8 Les régions LTR	33
1.7.9 La protéine Tat	35
1.7.10 La protéine Rev	36
1.7.11 Les protéines accessoires	36
Chapitre 2 : La transcription antisens.....	38
2.1 Transcription chez les eucaryotes : généralités	38
2.1.1 L'initiation de la transcription.....	38
2.1.2 Activation de la transcription.....	41
2.1.3 Traitement des ARNm.....	43
2.1.4 Élongation des transcrits	44
2.1.5 Épissage	45
2.1.6 Export nucléaire des ARNm	46
2.2 La transcription antisens.....	48
2.2.1 Généralités	48
2.2.2 Rôles biologiques des NAT.....	50
2.3 La transcription antisens chez les rétrovirus	52
2.3.1 Transcription antisens chez HTLV-1 et découverte de la protéine HBZ	53

2.3.2	<i>Caractérisation du transcrit antisens chez HTLV-1</i>	54
2.3.3	<i>Rôle de la protéine HBZ dans la réplication virale.</i>	56
2.3.4	<i>Structure de la protéine HBZ</i>	58
2.3.5	<i>Transcription antisens chez HTLV-2 et HTLV-3.</i>	60
2.4	La transcription antisens chez le VIH-1	60
2.5	La transcription antisens chez le VIF	63
Chapitre 3 : Hypothèses et objectifs.....		64
3.1	Hypothèses de recherche.....	64
Chapitre 4 : Régulation positive de la transcription antisens chez HTLV-1 par la protéine virale Tax.....		66
4.1	Résumé	66
4.2	Article.....	67
Chapitre 5 : Les rétrovirus HTLV-3 et HTLV-4 expriment des transcrits antisens permettant l'expression de protéines capables d'inhiber l'action de Tax		91
5.1	Résumé	91
5.2	Article.....	93
Chapitre 6 : Détection, caractérisation et régulation de la transcription antisens chez le VIH-1.....		130
6.1	Résumé	130
6.2	Article.....	132
Chapitre 7 : Expression et détection de la protéine ASP.....		183
7.1	Introduction	183
7.2	Résultats	184
7.2.1	<i>Expression et purification de la protéine ASP dans les cellules d'insectes</i>	184
7.2.2	<i>Détection de la protéine ASP dans les cellules 293T et COS-7</i>	188
7.2.3	<i>Optimisation de la séquence codante d'ASP et détection de la protéine</i>	188
Chapitre 8 : Discussion générale.....		191
8.1	La transcription antisens chez les rétrovirus : un phénomène sous-estimé	191
8.2	La détection de transcrits antisens par des approches de RT-PCR	191
8.3	La régulation de la transcription antisens chez HTLV-1	193
8.4	Rôle de la protéine HBZ dans le développement de l'ATL	196
8.5	Transcription antisens chez HTLV-3 et HTLV-4	198
8.6	Transcription antisens chez le VIH-1	203
8.7	La capacité codante du transcrit antisens chez le VIH.....	204
8.8	Localisation cellulaire de la protéine ASP	207
8.9	L'optimisation des codons permet l'expression efficace de la protéine ASP	209
8.10	ASP : perspectives.....	212
Conclusion.....		214
Bibliographie.....		215

Liste des figures

FIGURE 1 : ORGANISATION GÉNOMIQUE DE QUELQUES RÉTROVIRUS.....	2
FIGURE 2 : LA PHYLOGÉNIE DES RÉTROVIRUS.....	3
FIGURE 3 : LE CYCLE DE RÉPLICATION DES RÉTROVIRUS.....	4
FIGURE 4 : LES ÉTAPES DE LA TRANSCRIPTION INVERSE CHEZ DIFFÉRENTS RÉTROVIRUS.....	6
FIGURE 5 : LES PREMIÈRES ÉTAPES DE L'INFECTION RÉTROVIRALES, DE LA FUSION À L'INTÉGRATION.....	7
FIGURE 6 : LA STRUCTURE DE LA PARTICULE RÉTROVIRALE.....	10
FIGURE 7 : DISTRIBUTION GÉOGRAPHIQUE DE L'INFECTION PAR HTLV-1.....	12
FIGURE 8 : LYMPHOCYTE AU NOYAU POLYLOBÉ CARACTÉRISTIQUE DE L'ATL.....	13
FIGURE 9 : L'ADN PROVIRAL D'HTLV-1.....	16
FIGURE 10 : REPRÉSENTATION SCHÉMATIQUE DU LTR D'HTLV-1.....	17
FIGURE 11 : MODÈLE DE LA TRANSFORMATION CELLULAIRE MÉDIÉE PAR TAX.....	22
FIGURE 12 : ÉVOLUTION DU NOMBRE DE LYMPHOCYTES T CD4+ CIRCULANTS ET DE LA CHARGE VIRALE AU COURS DE L'INFECTION PAR LE VIH-1.....	25
FIGURE 13 : DISTRIBUTION GÉOGRAPHIQUE DE L'ÉPIDÉMIE DU SIDA EN 2007.....	30
FIGURE 14 : LE GÉNOME DU VIH-1 ET LES DIFFÉRENTS ARNm PRODUITS PAR LE VIRUS.....	32
FIGURE 15 : LE LTR DU VIH-1.....	34
FIGURE 16 : PRINCIPALES ÉTAPES DE L'ÉPISSAGE DES ARNm.....	46
FIGURE 17 : LES DIFFÉRENTS TYPES D'ARN ANTISENS.....	49
FIGURE 18 : LES POTENTIELS MÉCANISMES D'ACTION DES ARN ANTISENS NATURELS	51
FIGURE 19 : POSITION DE L'ORF DE LA PROTÉINE HBZ DANS LE GÉNOME D'HTLV-1.....	54
FIGURE 20 : ÉPISSAGE ALTERNATIF DU TRANSCRIT ANTISENS CHEZ HTLV-1.....	56
FIGURE 21 : ACTION DE L'ARNm ET DE LA PROTÉINE HBZ DANS UN CELLULE INFECTÉE PAR HTLV-1.....	58
FIGURE 22 : LES DIFFÉRENTS DOMAINES DE LA PROTÉINE HBZ.....	59
FIGURE 23 : COMPARAISON DE LA SÉQUENCE EN ACIDES AMINÉS DE PLUSIEURS FACTEURS BZIP CODÉS PAR DIFFÉRENTS MEMBRES DES PTLVs.....	60
FIGURE 24 : POSITION DE L'ORF DE LA PROTÉINE ASP DANS LE GÉNOME DU VIH-1	64
FIGURE 25 : EXPRESSION DE LA PROTÉINE ASP DANS LES CELLULES D'INSECTE SF9.....	187
FIGURE 26 : LOCALISATION DE LA PROTÉINE MYC-ASP DANS LES CELLULES 293T ET COS-7.....	189
FIGURE 27 : DÉTECTION DE LA PROTÉINE ASP EXPRIMÉE À L'AIDE D'UN ADNC OPTIMISÉ.....	191
FIGURE 28 : STRATÉGIE DE RT-PCR BRIN-SPÉCIFIQUE UTILISÉE POUR LA DÉTECTION DU TRANSCRIT ANTISENS CHEZ LE VIH-1.....	195
FIGURE 29 : LES ÉTAPES MENANT À L'ONCOGÉNÈSE INDUITE PAR HTLV-1.....	198
FIGURE 30 : REPRÉSENTATION SCHÉMATIQUE D'UN DOMAINE COILED-COIL EN DOUBLE BRINS PARALLÈLES.....	202
FIGURE 31 : STRUCTURE D'UN HOMODIMÈRE BZIP LIÉ À L'ADN.....	203

FIGURE 32 : ALIGNEMENT DES MOTIFS <i>LEUCINE ZIPPER</i> RETROUVÉS DANS LES PROTÉINES ANTISENS EXPRIMÉES PAR LES VIRUS HTLVs.....	204
FIGURE 33 : PROFIL D'HYDROPHOBICITÉ DE LA PROTÉINE ASP.....	207
FIGURE 34 : PRÉDICTION DES HÉLICES TRANSMEMBRANAIRES DE LA PROTÉINE ASP	208
FIGURE 35 : ALIGNEMENT DES SÉQUENCES EN ACIDES AMINÉS DES ORF ANTISENS DU VIH-1 (CLONE NL4.3) ET VISCPZ (CLONE MB66).....	209
FIGURE 36 : POSITION DES SÉQUENCES RRE ET DES ORF ANTISENS CONSERVÉS CHEZ LE VIH-1 ET LE VIF.....	213

Liste des tableaux

TABLEAU 1 : RÔLE DES DIFFÉRENTS FACTEURS DE TRANSCRIPTION GÉNÉRAUX FORMANT LA MACHINERIE TRANSCRIPTIONNELLE.....	39
TABLEAU 2 : POSITION DES ÉLÉMENTS RETROUVÉS DANS LES PROMOTEURS BASAUX DE L'ARN POLYMÉRISE II ET LES FACTEURS TRANSCRIPTIONNELS PAR LESQUELS ILS SONT RECONNUS.....	40

Liste des abréviations

ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complémentaire
ALV	<i>Avian leukemia virus</i>
AP-1	Protéine activatrice 1
APH-3 /APH-4	<i>Antisense protein of HTLV 3/4</i>
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messager
ARNt	Acide ribonucléique de transfert
ASP	<i>Antisense protein</i>
ATL	<i>Adult T-cell leukemia</i>
AZT	Azidothymidine
BRE	<i>TFIIB recognition element</i>
CA	Capside
CAT	Chloramphénicol acétyl-transférase
CDK	<i>Cyclin-dependent kinase</i>
C/EBP	<i>CAAT enhancer binding protein</i>
CPSF	<i>Cleavage and polyadenylation specificity factor</i>
CstF	<i>Cleavage stimulation factor</i>
CTE	<i>Constitutive transport element</i>
CTL	Cellule T cytotoxique
CRE	<i>cyclic AMP responsive element</i>
CREB/ATF	<i>CRE-binding/Activating Transcription Factors</i>

DC	Cellule dendritique
DDR	<i>DNA damage response</i>
DPE	<i>Downstream promoter element</i>
EIAV	<i>Equine infectious anemia virus</i>
ERV	<i>Endogenous Retrovirus</i>
FV	<i>Foamy virus</i>
GFP	<i>Green fluorescent protein</i>
GST	Glutathione-S-transferase
GTF	Facteur de transcription général
GTP	Guanosine 5'-triphosphate
HAART	<i>Highly active antiretroviral therapy</i>
HAT	Histone acétyl-transférase
HBZ	<i>HTLV-1 bZIP factor</i>
HFV	Human Foamy Virus
Inr	Élément initiateur
IST	Inducteur de courts transcrits
LTR	<i>Long terminal repeat</i>
LPS	Lipopolysaccharides
MA	Matrice
MTE	<i>Motif ten element</i>
MPMV	<i>Mason Pfizer Monkey Virus</i>
MVB	<i>Multi-vesicular bodies</i>
NAT	<i>Natural antisense transcript</i>

NC	Nucléocapside
NFAT	<i>Nuclear factor of activated T cells</i>
NF- κ B	<i>Nuclear factor kappa B</i>
NES	<i>Nuclear export signal</i>
NLS	<i>Nuclear localization signal</i>
NPC	<i>Nuclear pore complex</i>
NRE	<i>Negative regulatory region</i>
ORF	<i>Open reading frame</i>
PBS	<i>Primer binding site</i>
PI3K	Phosphatidylinositol-3-kinase
PIC	Complexe de pré-intégration ou complexe de pré-initiation
PPT	<i>Polypurine tract</i>
PRMT6	Protéine arginine méthyl-transférase 6
PTLV	<i>Primate T-cell leukemia virus</i>
Rb	Protéine du rétinoblastome
RCF	<i>Recombinant circulating form</i>
RRE	<i>Rev responsive element</i>
RTC	<i>Reverse Transcriptase complex</i>
RT-PCR	<i>Reverse transcription / Polymerase chain reaction</i>
RxRE	<i>Rex responsive element</i>
SCID/NOD	<i>Severe combined immunodeficiency / Non-obese diabetic</i>
SH3	<i>Src-homology 3</i>
SIDA	Syndrome de l'immunodéficience acquise

snRNP	<i>Small ribonucleoprotein</i>
SRE	<i>Serum responsive element</i>
SRF	<i>Serum responsive factor</i>
STLV-1	<i>Simian T-cell leukemia virus</i>
SU	Protéine de surface
TAR	<i>Transactivation response element</i>
TBP	<i>TATA binding protein</i>
TdT	<i>Terminal deoxynucleodyl transferase</i>
TI	Transcriptase inverse
TM	Protéine transmembranaire
TSP/HAM	<i>Tropical spastic paraparesis / HTLV-1-associated myelopathy</i>
TRE	<i>Tax-responsive element</i>
USF-1	<i>Upstream transcription factor 1</i>
VIF	Virus de l'immunodéficience féline
VIH	Virus de l'immunodéficience humaine
VIS	Virus de l'immunodéficience simienne
VLP	<i>Virus-like particle</i>

Chapitre 1 : La famille des *Retroviridae*

1.1 Généralités

Le terme rétrovirus désigne une grande famille de virus enveloppés à ARN très répandus chez les vertébrés. Ils sont classifiés selon leur structure, leur composition et leurs propriétés répliquatives. La principale caractéristique commune des membres de cette famille consiste en la synthèse d'un ADN double brin à partir du génome d'ARN simple brin de polarité positive et en l'intégration de cet ADN dans le génome de la cellule cible. Les *Retroviridae* ne sont cependant pas les seuls virus capables de transcription inverse puisque d'autres groupes de virus (*Metaviridae*, *Pseudoviridae*, *Hepadnaviridae* et *Caulimoviridae*) sont aussi connus pour utiliser ce type de mécanisme au cours de leur réplication [1, 2].

1.2 Organisation génomique

Les rétrovirus peuvent être divisés en deux grandes catégories, les rétrovirus simples et complexes, et ce, en fonction de leur organisation génomique. Le génome de tous les rétrovirus contient trois gènes de structures qui sont cruciaux pour la production d'une particule virale : le gène *gag* contient l'information nécessaire à la production des protéines de la capsid, nucléocapsid et matrice virale; le gène *pol* permet la synthèse de deux enzymes virales, la transcriptase inverse et l'intégrase; le gène *env* contient les séquences codantes pour les protéines de surface (SU) et transmembranaire (TM) du virus. Le gène *pro* est également commun à tous les rétrovirus et permet la production de la protéase virale. Alors que cette information constitue l'entière composition du génome des rétrovirus simples, les rétrovirus complexes peuvent quant à eux coder pour plusieurs autres protéines virales *via* la production de transcrits d'ARNm multi-épissés (voir figure 1).

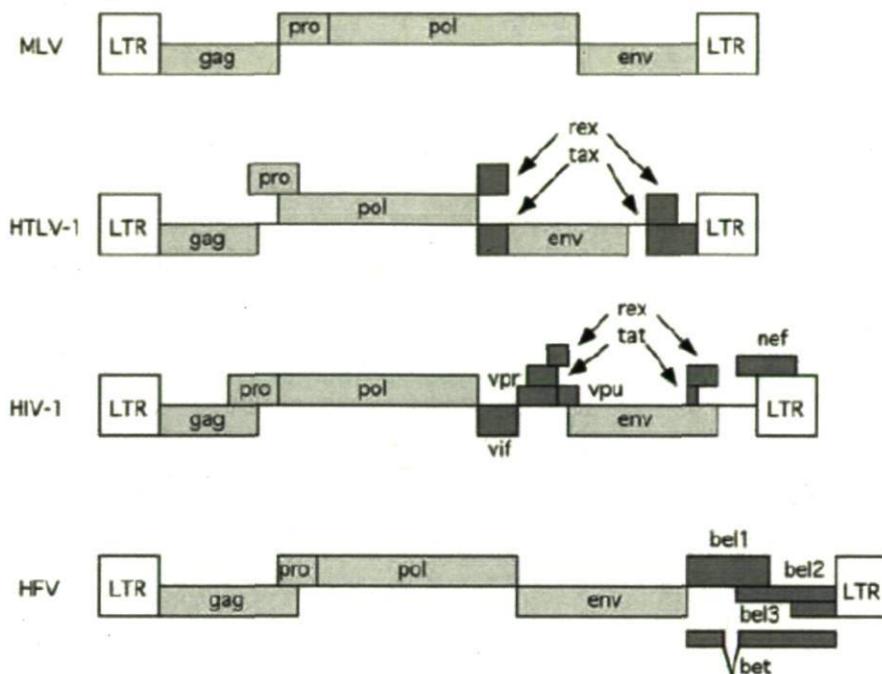


Figure 1 : Organisation génomique de quelques rétrovirus. Les génomes du rétrovirus simple MLV (*Murine Leukemia Virus*) et des rétrovirus complexes HTLV-1, VIH-1 et HFV (*Human Foamy Virus*) sont présentés. Les gènes dits « accessoires » sont présentés en gris foncé.

1.3 Classification

La grande famille des *Retroviridae* est divisée en sept genres qui traduisent la proximité des différents virus d'un point de vue évolutif (voir figure 2). Notons que cinq de ces genres regroupent des virus oncogènes, dont seul le groupe *Deltaretrovirus* est caractérisé par un génome complexe. Les genres non-oncogènes *Lentivirus* et *Spumavirus* regroupent également des rétrovirus complexes. Les membres du groupe *Lentivirus*, dont fait partie notamment le virus de l'immunodéficience humaine de type 1 (VIH-1), sont à l'origine de pathologies à évolution lente, caractérisées par une longue période de latence précédant

l'apparition de signes cliniques. À ce jour, aucune pathologie n'a été associée aux spumavirus [3].

La famille des *Retroviridae* comprend également la vaste sous-famille des rétrovirus endogènes, aussi appelés ERVs (*Endogenous Retrovirus*). Ces séquences sont en fait des vestiges d'intégrations rétrovirales ancestrales qui font toujours partie du génome de plusieurs organismes eucaryotes, puisqu'elles sont transmises de générations en générations selon la loi mendélienne [4]. Bien que la grande majorité des gènes des ERVs soient aujourd'hui défectifs, certains sont toujours capables de produire des protéines ou même, dans de rares cas, des particules virales complètes [5, 6].

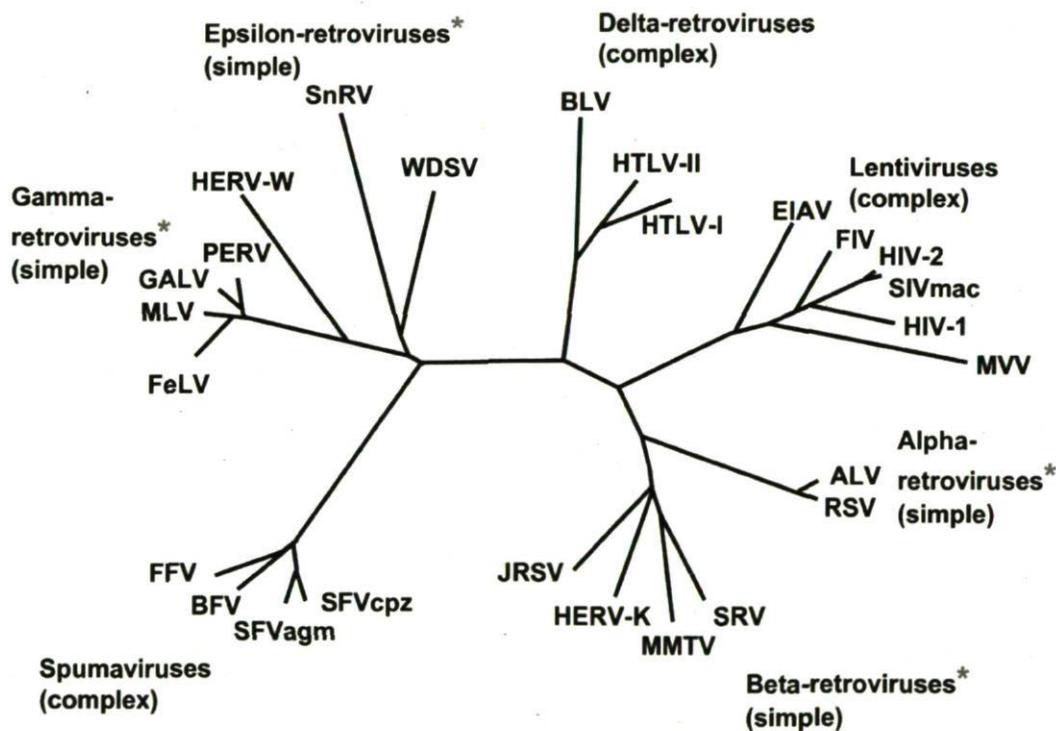


Figure 2 : La phylogénie des rétrovirus. Les astérisques rouges indiquent les genres qui incluent des rétrovirus endogènes. (adapté de [4])

1.4 Cycle répliatif

La première étape du cycle répliatif des rétrovirus consiste en l'attachement de la particule virale à la surface de la cellule cible (figure 3). Cet attachement est possible grâce à l'interaction spécifique entre la glycoprotéine de surface virale (SU) et son récepteur cellulaire. De façon générale, on définit un récepteur rétroviral comme une molécule de la surface cellulaire qui permet l'attachement spécifique de la particule virale et son entrée subséquente dans la cellule [7]. Par contre, l'attachement de la particule virale à la surface de la cellule peut également être promu par d'autres composants de la surface cellulaire. L'entrée du virus nécessite la fusion entre l'enveloppe rétrovirale et la membrane plasmique cellulaire, processus qui implique la participation de la sous-unité transmembranaire (TM) de la protéine d'enveloppe virale.

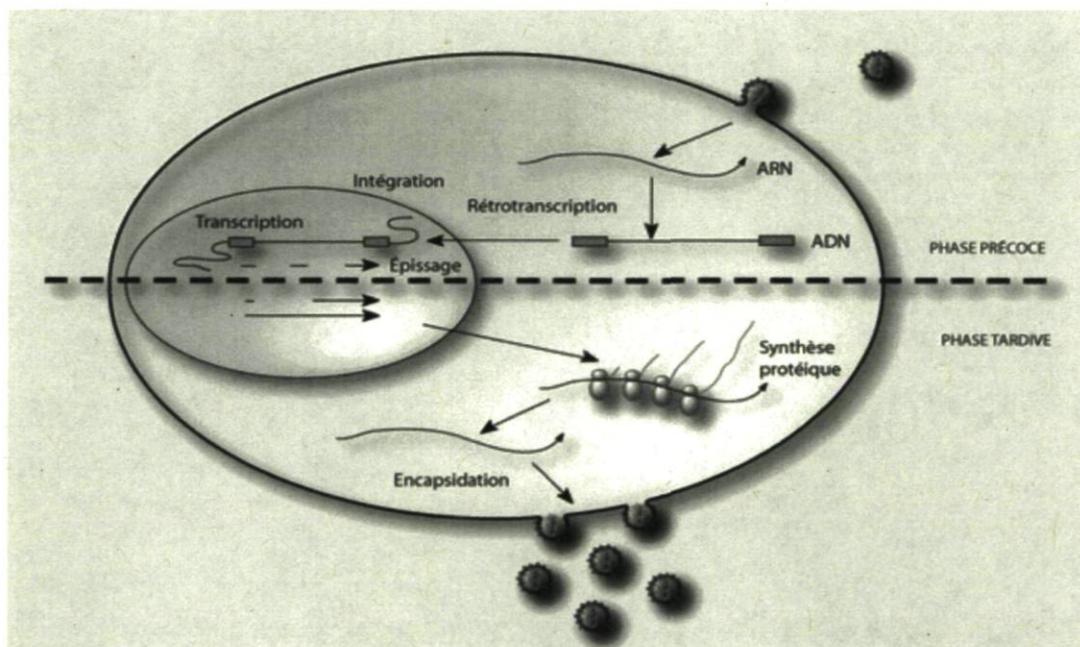


Figure 3 : Le cycle de réplication des rétrovirus. Résumé des différentes étapes de la réplication des rétrovirus, de l'entrée du virus à la production de virions. La phase précoce inclut les étapes d'attachement et de fusion, de synthèse et d'intégration l'ADN proviral ainsi que la transcription des gènes viraux. La phase tardive désigne la production des protéines, l'assemblage et la libération des particules virales.

Suite au relargage de la capsid virale dans le cytoplasme de la cellule s'enclenche un processus de désassemblage ou *uncoating* qui mène à la formation de particules appelées « complexes de transcription inverse » ou RTCs. Bien qu'il semble clair que les RTCs soient formés d'un amalgame de protéines virales et cellulaires, la nature exacte de leurs composants est encore mal connue [8].

C'est au moment du désassemblage que débute l'action de la transcriptase inverse, une enzyme associée au génome viral dont le rôle consiste en la synthèse d'un double brin d'ADN à partir de l'ARN génomique. La transcription inverse est amorcée sur la séquence PBS (*primer binding site*) par un ARN de transfert dont la nature varie selon le rétrovirus. Après une série d'étapes incluant deux sauts de la transcriptase inverse, l'ADN proviral est synthétisé et certaines séquences retrouvées aux extrémités du génome d'ARN sont dupliquées, résultant ainsi en la formation de séquences répétées de part et d'autre de l'ADN viral. Ces séquences, appelées LTRs (*Long Terminal Repeats*) jouent un rôle fondamental dans la régulation de l'expression des gènes viraux. Chez les lentivirus, la transcription inverse implique la synthèse du brin d'ADN positif en deux segments discontinus, puisque cette dernière est amorcée grâce à deux régions riches en purines appelées PPT (*polypurine tracts*). Conséquemment, un fragment d'ADN linéaire doublon appelé *DNA flap* est retrouvé au centre du brin d'ADN positif. Certaines études ont proposé que le *DNA flap* du VIH-1 pourrait être impliqué dans l'import nucléaire du génome viral [9]. Les différentes étapes de la transcription inverse chez les rétrovirus sont présentées à la figure 4 [8].

Les complexes qui contiennent l'ADN double brin rétro-transcrit sont appelés complexes de pré-intégration ou PIC, puisqu'ils sont alors aptes à intégrer le génome de la cellule infectée. Ces complexes doivent tout d'abord migrer jusqu'au noyau, un processus qui implique probablement un transport actif nécessitant des constituants du cytosquelette cellulaire [10].

La majorité des rétrovirus sont incapables d'infecter des cellules dont la membrane nucléaire est intacte : les PIC doivent attendre que l'intégrité de cette dernière soit brisée lors de la mitose cellulaire pour entrer efficacement au noyau. Le VIH-1 fait cependant

exception à cette règle, puisque ses PIC sont capables de traverser au noyau en utilisant les pores nucléaires, ce qui lui donne l'avantage de pouvoir infecter des cellules quiescentes ou différenciées. L'intégrase (IN) du VIH, qui possède une séquence de localisation nucléaire (NLS) dite atypique, semble jouer un rôle majeur dans ce mécanisme [11, 12] bien que d'autres protéines virales soient probablement aussi impliquées [13, 14].

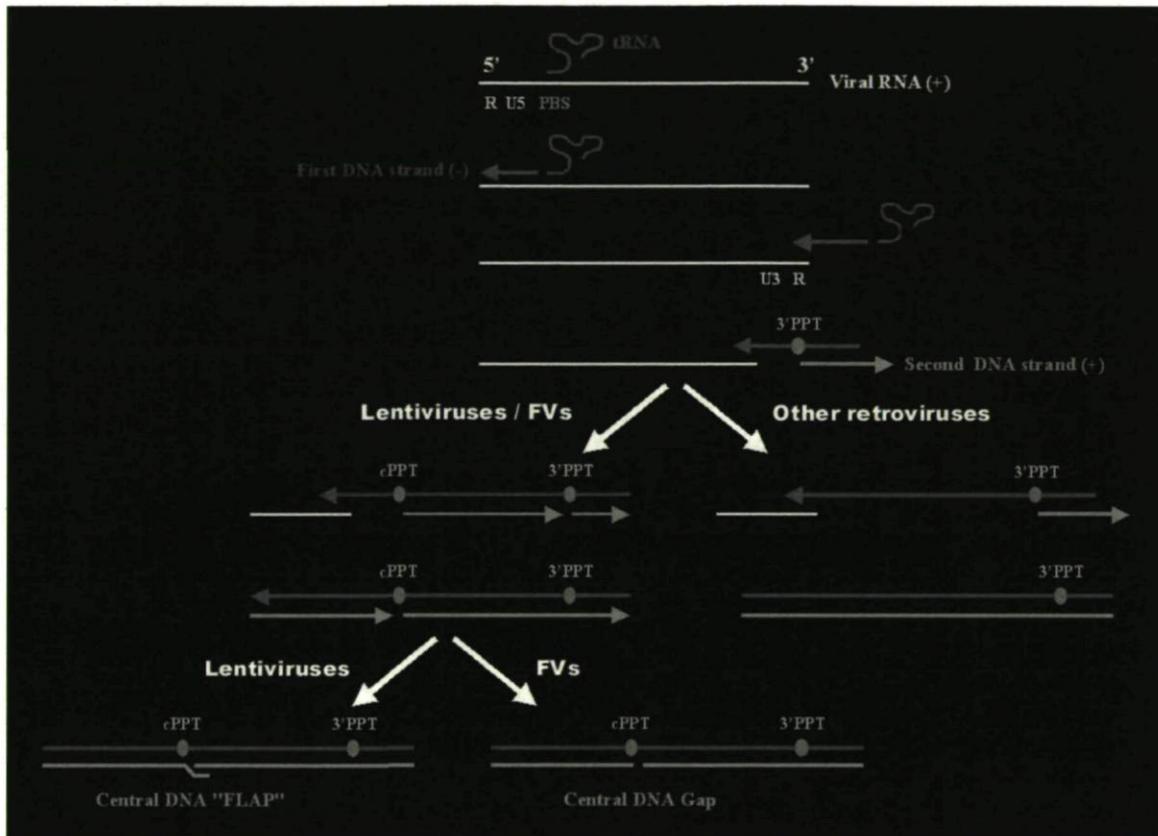


Figure 4 : Les étapes de la transcription inverse chez différents rétrovirus. La réaction de transcription inverse est amorcée par un ARNt spécifique et implique un saut de l'enzyme permettant la duplication des séquences LTRs. L'étape de synthèse du deuxième brin d'ADN chez les *Lentivirus* implique deux séquences riches en purines et conséquemment la présence d'un doublon d'ADN appelé *DNA Flap*. PBS : *primer-binding site*; cPPT : *central polypurine tract*; 3'PPT : *3' polypurine tract*; FVs : *foamy viruses*. (tiré de [8])

Une fois à l'intérieur du noyau, l'ADN viral s'intègre dans l'ADN chromosomique cellulaire grâce à l'action de l'intégrase. Lors de ce processus, les extrémités de l'ADN viral se lient à l'ADN cellulaire à travers plusieurs étapes qui impliquent notamment la

perte de quelques nucléotides aux extrémités de l'ADN viral et la génération de courtes séquences répétées de part et d'autre du site d'intégration. Si le choix de ce dernier semble se faire au hasard, il apparaît par ailleurs clair qu'il puisse être influencé de façon importante par la conformation de la chromatine. En effet, certaines études ont suggéré que les gènes transcriptionnellement actifs pourraient constituer une cible préférentielle de l'intégration rétrovirale [15-18], bien que cela ne semble pas être le cas pour certains rétrovirus [19]. Les principales étapes de l'infection rétrovirale, de la fusion de la particule virale à l'intégration de l'ADN proviral, sont présentées à la figure 5.

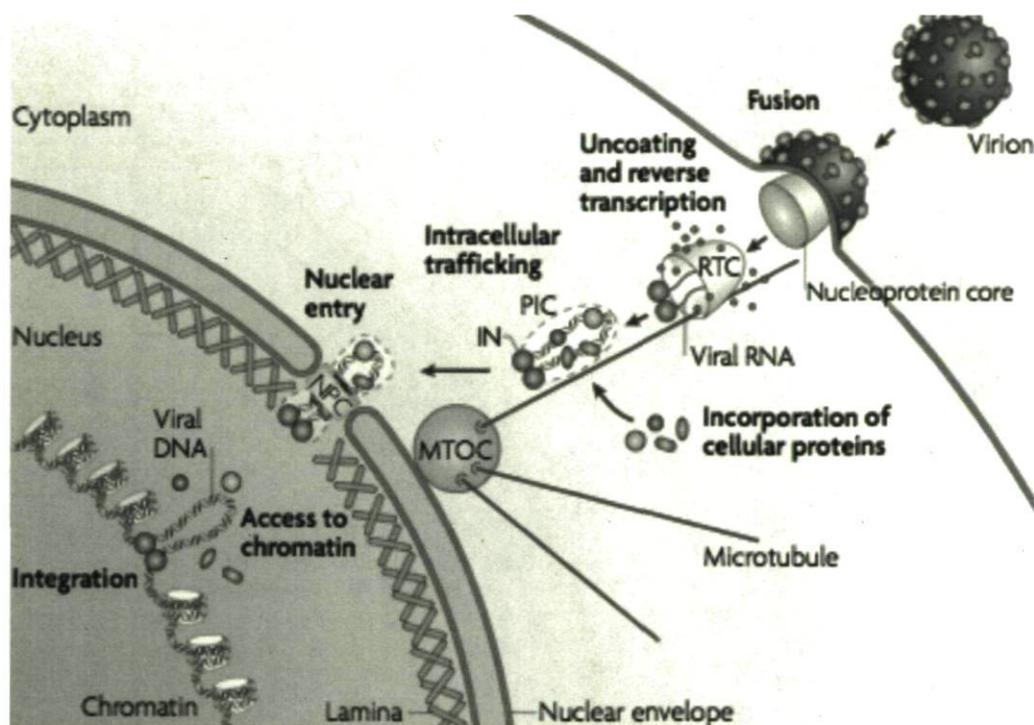


Figure 5 : Les premières étapes de l'infection rétrovirale, de la fusion à l'intégration. Le rétrovirus entre dans la cellule grâce à la fusion des membranes. Le désassemblage de la capsidie mène à la formation des RTCs dans lesquels la transcriptase inverse effectue la synthèse de l'ADN. Cet ADN est associé à des protéines virales et cellulaires dans les PIC, qui atteignent la membrane nucléaire grâce à un transport actif sur les microtubules. (adapté de [10])

L'intégration rétrovirale peut entraîner la dérégulation de l'expression d'un gène cellulaire, ce qui peut parfois mener à l'activation de son potentiel transformant. Ainsi, l'intégration

du provirus peut résulter en l'initiation, l'augmentation ou la terminaison de la transcription d'un gène cellulaire, entraînant conséquemment la production d'une grande quantité de certains messagers d'ARN ou même l'expression de messagers tronqués. Ces mutations insertionnelles sont à l'origine du pouvoir transformant attribué à plusieurs rétrovirus [20]. Le pouvoir transformant des rétrovirus est une caractéristique qui limite l'utilisation de vecteurs rétroviraux en thérapie génique. L'apparition de leucémies liés à l'activation du gène LMO-2 chez certains patients traités pour la correction de l'immunodéficience sévère liés au chromosome X (SCID-X1) à l'aide d'une stratégie reposant sur le transfert de gène via l'utilisation d'un vecteur rétroviral illustre bien cette problématique [21].

Une fois intégré, l'ADN proviral se comporte comme un gène cellulaire : les gènes viraux sont transcrits grâce à l'action de l'ARN polymérase II et l'ADN viral est répliqué à l'aide des enzymes cellulaires, en même temps que le reste de l'ADN génomique. De façon générale, la production des protéines virales résulte de l'expression d'un transcrit d'ARNm pleine longueur initié dans le LTR 5'. L'épissage de ce transcrit permet l'expression des différentes protéines, alors qu'une fraction des transcrits non-épissés qui contiennent un signal d'encapsidation ψ , sera incorporée dans la particule virale.

Si l'expression des protéines d'enveloppe et de la protéase virale est nécessaire à la production de particules virales infectieuses, la seule expression de la protéine Gag permet généralement l'assemblage et le relargage de pseudo-particules virales ou VLPs (*virus-like particles*). L'assemblage des nouveaux virions a lieu à différents endroits selon le type de rétrovirus. Chez les rétrovirus de type C, qui incluent le VIH-1 et les virus HTLVs, l'assemblage a lieu à la membrane plasmique et le bourgeonnement à partir de celle-ci permet au virus d'acquérir son enveloppe, constitué d'une bicouche lipidique empruntée à la cellule hôte. Le bourgeonnement est favorisé par la présence de courts motifs dans la séquence du précurseur Gag connus sous le nom de « *late domains* ». Ces motifs permettent l'interaction avec des protéines cellulaires qui font partie de la machinerie de triage endosomal [22]. L'action de la protéase virale permet de cliver le précurseur protéique Gag en sous-unités structurales matures (matrice (MA), capsid (CA) et nucléocapsid (NC)), une étape essentielle à l'infectivité des virions. La protéase exerce

normalement son action dans la particule virale au niveau de la membrane plasmique, mais peut aussi intervenir lors du bourgeonnement des virus dans la lumière de vésicules intracellulaires appelées *multi-vesicular bodies* ou MVBs. Ces vésicules représentent notamment un lieu de bourgeonnement important dans les macrophages infectés par le VIH-1 (pour une revue, voir [23]).

L'assemblage de plusieurs rétrovirus (incluant le VIH-1) survient au niveau de microdomaines membranaires riches en cholestérol et glycosphingolipides nommés radeaux lipidiques. L'association des protéines structurales avec les radeaux lipidiques peut promouvoir certaines interactions protéines-protéines au cours du bourgeonnement et favoriser l'infectivité des nouveaux virus [24].

1.5 Structure des rétrovirus

Les virions sont d'un diamètre de 80 à 110 nm. Ils contiennent deux copies du génome d'ARN de polarité positive étroitement associées aux protéines de la nucléocapside et aux enzymes virales (protéase, intégrase et transcriptase inverse). Ce complexe ribo-protéique est protégé par une capsid de forme hélicoïdale, elle-même recouverte de la matrice virale. Les virus possèdent également une enveloppe constituée d'une bicouche lipidique dérivée de la membrane plasmique de la cellule hôte. Elle contient les glycoprotéines transmembranaires (TM) et de surface (SU) produites par le virus. La structure de la particule rétrovirale est présentée à la figure 6.

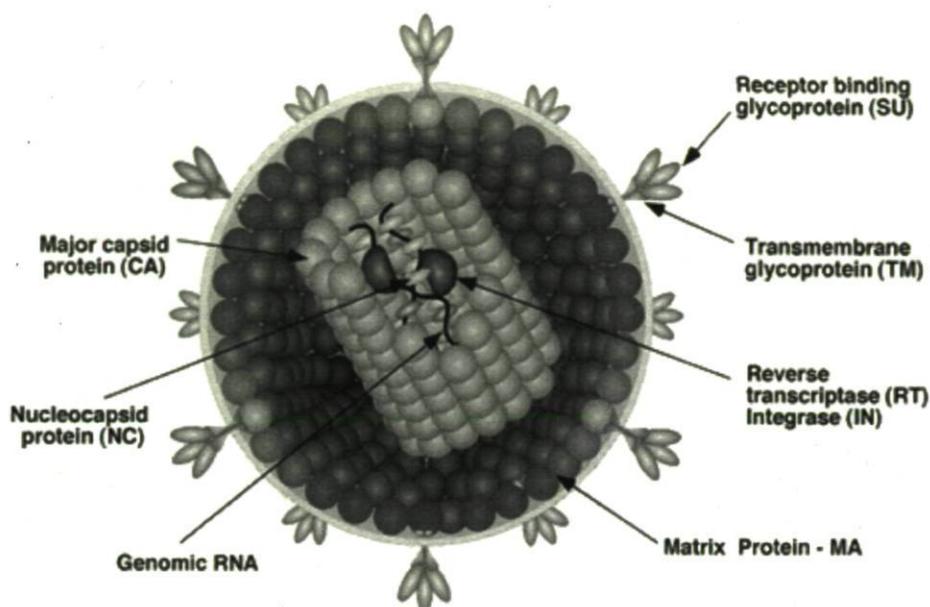


Figure 6 : Structure de la particule rétrovirale. La particule rétrovirale est recouverte d'une bicouche lipidique à la surface de laquelle on retrouve les sous-unités SU et TM de l'enveloppe virale. Cette enveloppe recouvre la matrice (MA) et la capsid virale (CA), au centre de laquelle on retrouve 2 copies du génome d'ARN associées à la nucléocapside et aux enzymes virales. (tiré de [25])

1.6 Les rétrovirus HTLVs

Le rétrovirus HTLV-1 (*Human T-cell Leukemia Virus*) est le premier rétrovirus humain à avoir été identifié. Il fut isolé en 1980 à partir d'une lignée cellulaire dérivée d'un patient atteint de leucémie à cellules T de l'adulte (ATL) [26]. L'année suivante, l'identification de la présence d'anticorps dirigés contre des antigènes viraux dans le sérum des patients atteints de la maladie démontrait pour la première fois l'existence d'un lien entre une pathologie humaine et l'infection par un rétrovirus.

HTLV-1 fait partie du genre *Deltaretrovirus*, qui comprend également le virus HTLV-2 et les virus STLVs (*Simian T-cell Leukemia Viruses*). Ensemble, ils forment un groupe appelé PTLV (*Primate T-cell Leukemia Viruses*). À l'exception de HTLV-2, tous ces virus sont

liés à l'apparition de désordres néoplasiques. Plus récemment, 2 nouveaux membres de la famille des HTLVs ont été découverts : les rétrovirus HTLV-3 et HTLV-4 [27]. Ces nouveaux virus ont tous deux été isolés au Cameroun chez des patients asymptomatiques. Des études phylogénétiques ont permis de démontrer que l'organisation et la séquence génomique du virus HTLV-3 étaient fortement apparentées à celles de son homologue simien STLV-3. Quant à HTLV-4, bien que son identification ne soit basée que sur l'analyse d'une séquence obtenue d'un seul isolat, l'observation d'importantes divergences de séquences avec les autres virus HTLV semble bel et bien indiquer qu'il s'agit d'un virus faisant partie d'une lignée phylogénique unique.

1.6.1 *Le rétrovirus HTLV-1*

HTLV-1 est sans contredit le plus étudié des rétrovirus HTLVs. Il est à l'origine de deux types de pathologie chez l'homme : la leucémie à cellules T de l'adulte (ATL) et la paraparésie spastique tropicale, aussi appelée « myélopathie associée à HTLV-1 » ou HAM/TSP. Vingt millions de personnes sont aujourd'hui infectées par HTLV-1 et certaines régions du globe sont particulièrement touchées. En effet, le Japon, les îles des Caraïbes, l'Amérique du sud, l'Afrique sub-saharienne et l'Iran sont considérés comme des zones endémiques (où la prévalence est supérieure à 2%) alors que le virus est plutôt rare en Amérique du Nord et en Europe (voir figure 7). Il est cependant important de noter que moins de 5% des personnes infectées développeront éventuellement l'une ou l'autre des maladies causées par HTLV-1. Si la majorité des patients infectés sont asymptomatiques, ils sont néanmoins capables de transmettre le virus. Le virus peut se contracter par contacts sanguins, par contacts sexuels ou par le lait maternel. Plusieurs évidences suggèrent que la transmission du virus implique un contact cellule-cellule, au cours duquel le virus profite de la formation d'une jonction cellulaire, appelée synapse virologique, pour passer d'une cellule infectée vers une nouvelle cellule [28]. Plus récemment, certaines études ont pu mettre en évidence la capacité du virus libre à infecter les cellules dendritiques myéloïdes et plasmacytoïdes [29, 30]. Puisque les cellules dendritiques infectées peuvent transférer

efficacement le virus aux cellules CD4+ et entrainer leur transformation, il apparaît qu'elles pourraient jouer un rôle majeur dans la transmission, la dissémination et la persistance du virus *in vivo*.

L'infection par HTLV-1 a aussi été associée à plusieurs autres manifestations pathologiques, comme l'arthrite chronique, les uvéites, les dermatites infectieuses, la polymyosite et la sclérose en plaques [31-33]. Des études plus approfondies sont cependant requises pour clarifier l'existence d'un lien direct entre HTLV-1 et ces maladies.

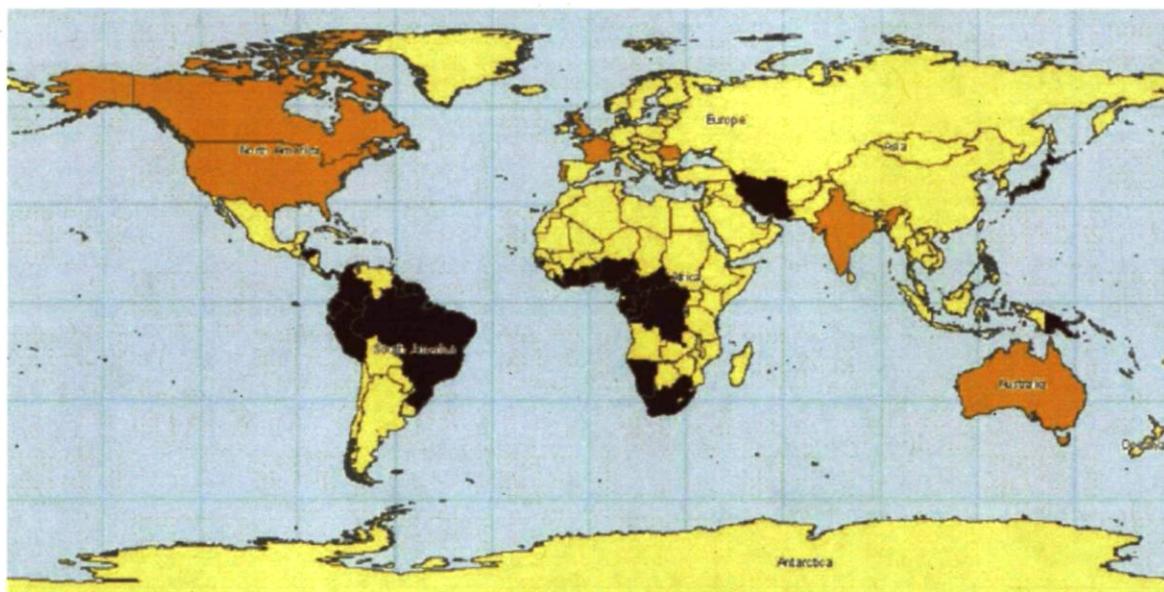


Figure 7 : Distribution géographique de l'infection par HTLV-1. Les pays où l'infection est endémique (prévalence 1-5 %) sont colorés en brun foncé. La couleur brun pâle indique une prévalence inférieure à 1%. (tiré de [34])

1.6.2 La leucémie à cellules T de l'adulte (ATL)

La leucémie à cellules T de l'adulte est un désordre lymphoprolifératif sévère qui peut être classifié en plusieurs sous-types selon certaines caractéristiques cliniques. On distingue les formes aiguës (50%), lymphomateuses (25%), chroniques (20%) et indolentes (5%) [35].

Les formes aiguës et lymphomateuses présentent des évolutions cliniques plus agressives alors que les 2 autres sous-types progressent plus lentement. L'apparition des signes cliniques est généralement précédée d'une très longue période de latence qui peut durer de 20 à 30 ans. Des études épidémiologiques ont démontré que l'ATL se développe surtout chez des patients qui contractent l'infection durant l'enfance.

Cliniquement, l'ATL est caractérisée par une invasion tissulaire par des cellules lymphocytaires T à la morphologie anormale et aux noyaux polylobés, aussi appelées « *flower cells* » (voir figure 8). Les tissus cutanés sont fréquemment atteints. Les patients présentent également un état d'immunosuppression, souvent accompagné d'infections opportunistes. D'un point de vue immunologique, les cellules ATL sont caractérisées par un phénotype CD3+, CD4+, CD8- et CD25+, indiquant qu'elles dérivent de cellules T auxiliaires activées. Par conséquent, il est possible que l'immunodéficience qui accompagne la maladie s'explique en partie par les propriétés immunosuppressives de ces cellules. Par ailleurs, il apparaît que le virus ait aussi la capacité d'interférer avec le processus de thymopoïèse normal *via* l'infection des thymocytes immatures [36].

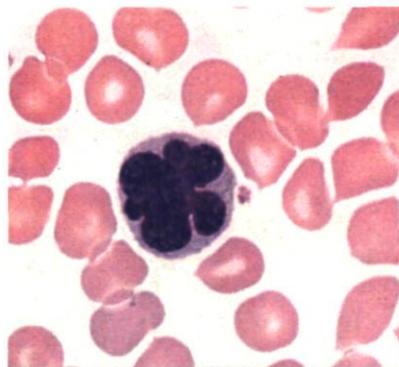


Figure 8 : Lymphocyte au noyau polylobé caractéristique de l'ATL. Les cellules tumorales au noyau polylobé sont des lymphocytes T CD4+ activés infectés par HTLV-1. (tiré de [37])

Le traitement de l'ATL représente un défi de taille pour le clinicien. De façon générale, les patients répondent mal aux traitements de chimiothérapie qui sont efficaces dans le cas de lymphomes agressifs. Les meilleurs résultats sont souvent obtenus en alliant la chimiothérapie classique à un traitement à l'interféron et à la zidovudine (AZT). [38, 39]. Le rôle bénéfique de l'interféron dans cette combinaison pourrait résulter d'une augmentation de la reconnaissance immunologique des cellules transformées ou d'un effet antiprolifératif direct.

1.6.3 *La paraparésie spastique tropicale*

La paraparésie spastique tropicale (TSP) est une maladie neurologique chronique qui apparaît chez environ 1-4% des patients infectés par HTLV-1, avec un risque plus élevé chez la femme que chez l'homme [40]. En général, les symptômes se présentent à l'âge adulte [41], bien que la maladie ait été diagnostiquée chez des enfants de moins de 10 ans [42]. La période de latence est plus courte que dans le cas de l'ATL, la maladie pouvant même se développer dans les quelques années qui suivent la contraction du virus.

La maladie est caractérisée par une démyélinisation qui se traduit par des troubles moteurs des membres inférieurs et des troubles génito-sphinctériens. Son évolution est progressive et est plus rapide chez les patients qui présentent une haute charge provirale. Tout comme dans le cas de l'ATL, on observe la présence de lymphocytes présentant une morphologie atypique. Certaines évidences suggèrent que les dommages au système nerveux central pourraient résulter de la lyse des cellules infectées par le système immunitaire. En effet, le virus HTLV-1 est capable d'infecter les cellules neuronales *in vivo* [43] et on retrouve chez les patients TSP un nombre élevé de lymphocytes T cytotoxiques (CTL) dirigés contre des antigènes viraux [44]. Les résultats obtenus au cours de certaines études ont aussi mené à l'élaboration d'une hypothèse proposant une étiologie auto-immunitaire, basée sur la démonstration que des anticorps dirigés contre des protéines virales auraient la capacité de reconnaître des épitopes cellulaires [45]. Il est par contre peu probable que l'une de ces théories puisse expliquer à elle seule la pathogenèse de la maladie. D'autres groupes de

recherche ont plutôt proposé un modèle dans lequel la production de certaines cytokines induites par la présence de lymphocytes T CD4+ infectés dans le système nerveux central pourrait déclencher le processus de démyélinisation [46].

Malgré la grande quantité d'études effectuées pour tenter de comprendre les désordres neurologiques induits par le virus, aucun traitement efficace n'est disponible pour traiter la TSP. Les stéroïdes sont fréquemment prescrits en association avec un traitement symptomatologique.

1.6.4 Désordres neurologiques associés à HTLV-II

Les séquences nucléotidique des génomes des virus HTLV-1 et HTLV-2 sont homologues à environ 70% [47]. Du point de vue biologique, ils possèdent tous deux un tropisme pour les lymphocytes T CD4+ *in vivo* [48], bien que HTLV-2 semble démontrer une préférence pour les lymphocytes CD8+ [49]. Les voies empruntées par le virus pour la transmission sont les mêmes et le partage de seringues par les utilisateurs de drogues par voie intraveineuse semble jouer un rôle important dans la dissémination du virus HTLV-2 en Europe et en Amérique du Nord [50]. La co-infection avec HTLV-2 et le VIH-1 est d'ailleurs fréquente au sein de ce groupe [51].

Tout comme HTLV-1, le virus HTLV-2 pourrait être à l'origine de désordres neurologiques, mais les évidences sont cependant beaucoup moins nombreuses. En effet, la présence de troubles identiques ou s'apparentant fortement à ceux qui sont observés chez les patients TSP/HAM a été décrite chez un nombre limité de patients infectés par HTLV-2. La présence de co-infection avec le VIH-1 favorise les erreurs de diagnostic et rend difficile la description précise des symptômes associés à ce virus.

1.6.5 Stratégie répliquative des rétrovirus HTLVs

Le génome des virus HTLVs est d'environ 9 kb. En plus des gènes communs aux rétrovirus, il possède une région codante pX permettant la production des protéines régulatrices Tax et Rex, qui jouent un rôle central dans le cycle de répllication virale. Le génome est bordé par les régions LTR qui régulent l'expression et la polyadénylation des ARNm. Les transcrits viraux sont initiés dans le LTR 5' et sont épissés de façon alternative de façon à permettre la production des différentes protéines virales. Le virus utilise aussi le décalage ribosomal (de l'anglais *ribosomal frameshift*), une stratégie qui permet la production des précurseurs protéiques Gag et Gag-Pol à partir du même ARNm grâce à un changement de cadre de lecture. La figure 9 présente les différentes formes d'ARNm produits par HTLV-1.

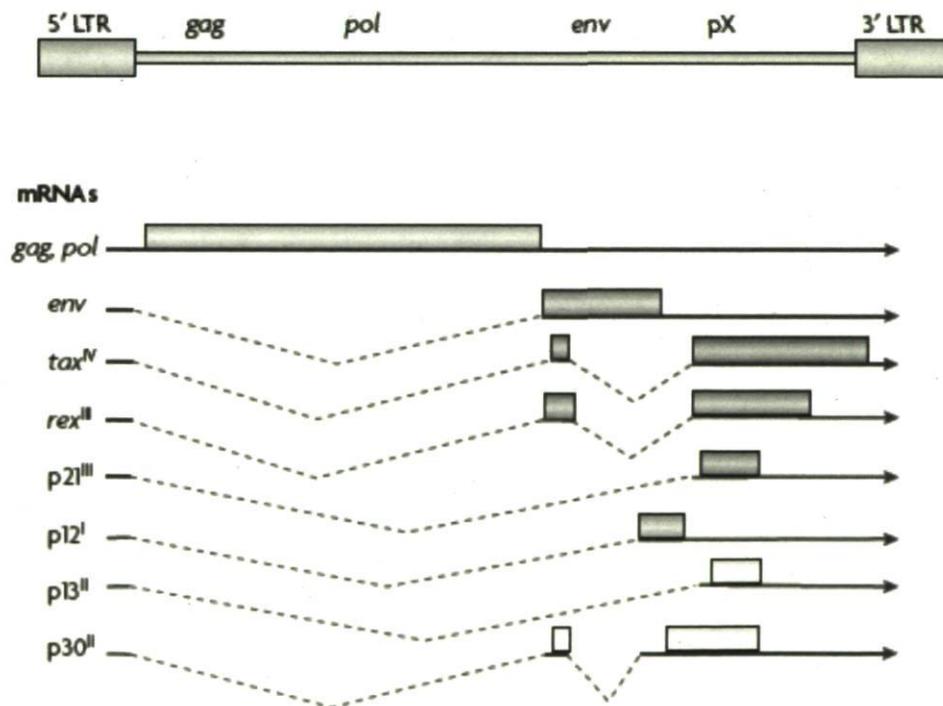


Figure 9 : L'ADN proviral d'HTLV-1. Le génome est bordé de deux régions LTR et contient une région pX qui code pour les protéines Tax et Rex (voir texte) mais également pour plusieurs protéines accessoires : p21, p12, p13 et p30. (tiré de [52])

1.6.6 Les séquences LTR

Les LTR sont constitués de trois régions distinctes, nommées U3 (pour région unique en 3'), R (pour région répétée) et U5 (pour région unique en 5') (voir figure 10). La région U3 contient les éléments qui contrôlent la transcription virale. En plus d'une boîte TATA et d'un signal consensus de polyadénylation, il contient trois éléments répétés de 21 pb : les éléments de réponse à Tax, appelés TRE1 (*Tax Responsive Element*) [53, 54]. Chaque TRE1 contient un motif octamère TGACG(T/A)(C/G)(T/A) bordé par des séquences riches en GC [55]. Ces motifs présentent une forte homologie avec les éléments de réponse à l'AMPc (CRE) et permettent la liaison de facteurs transcriptionnels de la famille CREB/ATF (*CRE-binding/Activating Transcription Factors*). Les éléments TRE1 jouent un rôle important dans la transcription induite par l'action de la protéine virale Tax [56], mais aussi dans le contrôle du niveau basal de transcription initiée par le promoteur viral [57]. Une seconde région, appelée TRE2, est située entre les sites TRE1 central et proximal et contient des éléments de réponse à d'autres facteurs de transcription cellulaires, tels que Ets et c-Myb [58, 59].

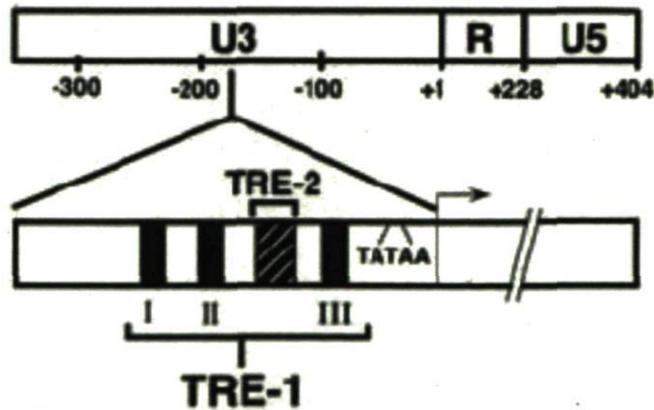


Figure 10 : Représentation schématique du LTR de HTLV-1. La transcription virale est initiée à la jonction des régions U3 et R. La région U3 contient une boîte TATA ainsi que trois éléments de réponse à Tax ou TRE1, essentiels à la transactivation par Tax. (tiré de [60])

1.6.7 Tax et la régulation transcriptionnelle chez HTLV-1

Tax est une protéine de 40 kDa, qui fut initialement décrite comme un activateur de la transcription virale [61]. Le cadre de lecture ouvert de Tax est contenu dans la région du génome nommée pX, qui code également pour la protéine Rex et plusieurs protéines accessoires. Sa séquence en acides aminés contient un signal de localisation nucléaire (NLS) et elle se localise majoritairement au noyau. Par contre, elle possède également un signal d'exportation nucléaire (NES) qui lui permet de faire la navette entre le noyau et le cytoplasme [62].

Tax démontre une affinité faible pour l'ADN et exerce plutôt son action sur le promoteur viral en interagissant avec des facteurs transcriptionnels liés sous forme dimérique aux sites CRE présents dans les éléments TRE1. Plusieurs études ont d'ailleurs démontré sa capacité à interagir avec plusieurs membres de la famille CREB/ATF *in vitro*, dont CREB-1, CREB-2, CREM et XBP-1 [63-66]. Ces protéines possèdent toutes un domaine basique de liaison à l'ADN et un domaine *leucine zipper* ou bZIP, permettant la formation de dimères. Ces homo- ou hétéro-dimérisations régulent l'activité de ces facteurs de transcription en modulant leur capacité à se lier à l'ADN [67]. Tax agit initialement en favorisant la dimérisation des membres de la famille CREB/ATF, augmentant donc leur affinité pour les séquences CRE retrouvées dans le promoteur viral [68-70]. Par la suite, il recrute les histones acétyl-transférase (HATs) CBP/p300 et P/CAF, favorisant ainsi l'initiation de la transcription.

Outre son action positive sur le promoteur viral, Tax est capable de moduler l'expression d'une panoplie de gènes cellulaires [71]. Les cellules infectées par HTLV-1 contiennent notamment un haut niveau de AP-1 (*Activator Protein 1*), un complexe homo- ou hétérodimérique des protéines cellulaires Fos et Jun [72-74]. L'expression de ces protéines est contrôlée par le facteur transcriptionnel SRF (*Serum Responsive Factor*) qui se lie sur les éléments de réponse au sérum (SRE) présents dans les promoteurs des gènes codant pour les protéines Fos et Jun. Tax provoque l'augmentation de l'expression de ces dernières en interagissant directement avec SRF et en favorisant sa liaison au SRE [75].

Tax entraîne également l'activation des facteurs transcriptionnels de la famille NF- κ B, composée de cinq membres : RelA (p65), RelB, c-Rel, NF- κ B1 (p50/p105) et NF- κ B2 (p52/p100). Ces derniers forment différents complexes dimériques qui régulent positivement ou négativement les gènes qui contiennent des éléments de réponse κ B dans leurs séquences promotrices. Ils jouent un rôle important dans la croissance, l'activation, la différenciation et la survie cellulaire [76]. L'activation de NF- κ B survient lorsqu'un stimulus spécifique (cytokines, LPS bactérien, infection virale) entraîne l'activation des sérines kinases IKK α et IKK β . Une fois activées, ces kinases phosphorylent l'inhibiteur de NF- κ B (I κ B), ce qui entraîne son ubiquitination et sa dégradation via le protéasome. Lorsque NF- κ B est libéré de I κ B, il peut transloquer au noyau et activer l'expression des gènes cibles. La protéine Tax est capable d'activer NF- κ B en augmentant l'activité catalytique de IKK α et IKK β *via* son interaction avec un troisième membre du complexe, IKK γ . Il a également été proposé que Tax pourrait agir sur des kinases en amont de IKK, provoquant notamment l'activation de la voie PI3K/AKT [77]. Tax peut aussi entraîner la phosphorylation de la sous-unité p65 et ainsi maximiser l'activation des gènes cibles [78] (pour une revue des différentes voies d'activation de NF- κ B par Tax, voir [79]).

De nombreuses études ont mis en évidence le rôle prépondérant joué par la protéine Tax dans la transformation cellulaire induite par HTLV-1. Son expression est suffisante pour entraîner l'immortalisation des lymphocytes T primaires en culture [80]. En combinaison avec l'oncogène Ras, elle transforme les fibroblastes primaires de rat *in vitro* [81]. Plusieurs modèles de souris transgéniques ont permis de faire un lien direct entre l'expression de Tax et le développement de tumeurs [82]. Malheureusement, aucun de ces modèles ne permet de reproduire les signes cliniques observés chez l'homme dans le développement de l'ATL.

Plus récemment, un nouveau modèle d'infection dans les souris humanisées immunodéficientes SCID/NOD a été présenté [83]. Dans ce modèle très prometteur, les souris sont inoculées avec des cellules souches hématopoïétiques CD34⁺ infectées par HTLV-1 *ex vivo*. De façon intéressante, ce modèle permet de constater chez les souris le développement de lymphomes présentant les signes cliniques de l'ATL et caractérisés par

la prolifération des cellules souches humaines dans la moelle osseuse. Ces observations suggèrent que les cellules souches hématopoïétiques pourraient constituer une cible de l'infection par HTLV-1 *in vivo* et que l'infection de ces cellules pourrait donc jouer un rôle important dans le développement de l'ATL.

1.6.8 Tax et le cycle cellulaire

La progression du cycle cellulaire est un processus hautement régulé qui dépend de l'action des cyclines et de kinases dépendantes des cyclines (CDKs). Les cyclines D et E coopèrent avec CDK4/6 et CDK2 pour médier la progression en phase G1 et la transition en phase G1/S [84]. Les complexes cycline D- CDK4/6 et cyclin E – CDK2 ciblent la protéine du rétinoblastome Rb. Lorsqu'il est hypophosphorylé, Rb se lie au facteur de transcription E2F1. La phosphorylation de Rb résulte en la libération de E2F1 et en l'activation de plusieurs gènes qui régulent la transition de la phase G1 en phase S. Tax interfère avec la progression du cycle cellulaire principalement en agissant sur la transition G1/S et ce, en utilisant plusieurs mécanismes : activation ou répression transcriptionnelle, modifications post-traductionnelles et interactions protéine-protéine. Tax est en effet capable d'interagir avec les cyclines D1, D2 et D3 ainsi qu'avec CDK4 et CDK6 [85, 86]. À travers ces interactions, Tax entraîne la stabilisation du complexe cycline D2-CDK4 et l'augmentation de son activité kinase, ce qui mène à l'hyperphosphorylation de Rb. Il a aussi été démontré que Tax peut contrer l'activité inhibitrice de certaines protéines régulatrices des CDK [87, 88]. Finalement, Tax est capable de se lier directement à Rb pour entraîner sa dégradation *via* le protéasome [89].

1.6.9 Tax interfère avec les mécanismes de réparation de l'ADN

Les dommages à l'ADN induisent une voie de signalisation appelée *DNA damage response* (DDR) au cours des phases S et G2/M du cycle cellulaire. L'activation de cette voie résulte en un délai dans la progression du cycle ou même l'apoptose des cellules endommagées et entraîne l'activation des mécanismes de réparation de l'ADN. Les protéines ATM et ATR ainsi que leurs cibles respectives Chk1 et Chk2 jouent un rôle central dans la voie de signalisation de DDR [90]. Chez les mammifères, Chk1 et Chk2 régulent Cdc25, wee2 et p53 pour mener à l'inactivation des CDK et l'inhibition de la progression du cycle cellulaire. Les cellules qui expriment Tax et les cellules ATL sont caractérisées par la présence de dommages à l'ADN, ce qui laisse croire que la protéine pourrait interférer avec certains mécanismes de réparation de l'ADN [91, 92]. En fait, des études ont mis en évidence la capacité de Tax à lier Rad51 et l'ADN topoisomérase I, deux protéines directement impliquées dans la réparation de l'ADN [93, 94]. De plus, l'expression de Tax entraîne la phosphorylation de Chk2, ce qui provoque une accumulation du nombre de cellules en phases S et G2/M [95, 96]. Par contre, suite à l'exposition des cellules aux rayons gamma, Tax inhibe l'activité kinase de Chk1 et Chk2, ce qui a pour effet d'atténuer l'arrêt du cycle cellulaire en phase G2 et l'apoptose [95, 97]. Il semble donc que Tax puisse réprimer ou activer certains mécanismes de contrôle en fonction des conditions expérimentales. En fait, une étude plus récente démontrait qu'à la suite de dommages induits par irradiation, Tax séquestre Chk2 à l'intérieur de la chromatine, ce qui empêche la dissociation de CDK2 activée, étape nécessaire à la transmission et l'amplification du signal aux substrats solubles [98]. En ciblant plusieurs composants de la machinerie de réparation de l'ADN, Tax favorise donc l'apparition d'anomalies dans le matériel génétique de la cellule.

De plus, Tax inhibe la fonction transactivatrice de la protéine p53, un important suppresseur de tumeurs, permettant ainsi d'éviter l'arrêt du cycle cellulaire et l'apoptose. Il semble donc que les nombreux moyens empruntés par Tax pour permettre la prolifération des cellules infectées provoquent une instabilité génétique pouvant éventuellement mener à la transformation cellulaire (résumé dans la figure 11) [99].

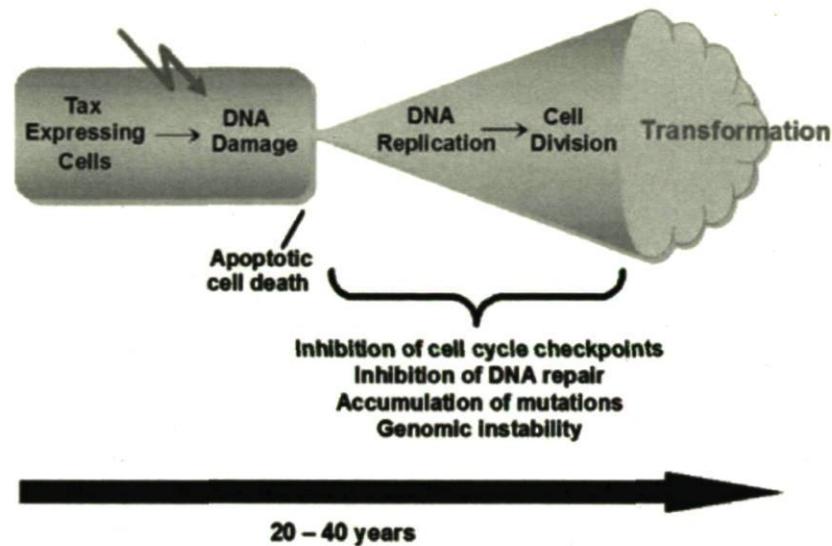


Figure 11 : Modèle de la transformation cellulaire médiée par Tax. Tax interfère avec les mécanismes de réparation de l'ADN ainsi qu'avec la régulation du cycle cellulaire, favorisant donc l'accumulation de mutations et l'instabilité génétique dans les cellules infectées par HTLV-1. Cette situation favorise, après plusieurs années, la sélection de cellules capables de tolérer cette instabilité et donc l'émergence d'une cellule transformée à l'origine de l'ATL. (tiré de [99])

1.6.10 La protéine Rex

Rex est une protéine de 27 kDa qui joue un rôle important dans l'épissage et le transport des ARNm viraux. De façon générale, Rex agit au niveau post-transcriptionnel en favorisant l'exportation au cytoplasme des ARNm mono-épissés (produisant notamment les protéines d'enveloppe) et non épissés (produisant les protéines de structure ou pouvant être encapsidés dans les nouveaux virions en tant que génome). Son expression entraîne une diminution de l'épissage et une plus grande stabilité des transcrits viraux [100]. L'action de Rex nécessite la présence d'une structure d'ARN présente dans tous les transcrits viraux, contenue dans les régions U3 et R du LTR 3' : il s'agit du RxRE ou « élément de réponse à

Rex » [101]. Ce dernier forme une structure secondaire stable et complexe composée de 255 nucléotides organisés en tiges-boucles. Rex exerce son action en se liant spécifiquement au RxRE. En plus de son rôle dans l'exportation des ARNm, le RxRE est nécessaire à la polyadénylation des transcrits viraux, puisqu'il permet de rapprocher le signal consensus de polyadénylation du site de clivage [102].

1.6.11 Les protéines accessoires

En plus des protéines Tax et Rex, la région pX du génome d'HTLV-1 code pour quatre autres protéines accessoires produites grâce à des épissages alternatifs : il s'agit des protéines p12^I, p27^I, p13^{II} et p30^{II}. Alors les premières études avaient laissé croire que ces protéines n'étaient pas nécessaires à la réplication virale [103, 104], un grand nombre de publications ont plus récemment permis de révéler leur rôle clef dans le maintien de la charge virale, l'activation de la cellule cible et la régulation de la transcription [105-118]. Les ARNm codant pour ces protéines peuvent être détectés à partir de cellules transfectées avec un clone moléculaire complet d'HTLV-1, de lignées cellulaires transformées par HTLV-1 et de cellules primaires isolées de patients infectés [119, 120]. Par contre, peu de choses sont connues quant à l'abondance de ces protéines au cours des différentes étapes du cycle répliatif. Des études axées sur la réponse immune à l'infection ont pu apporter des preuves indirectes de l'expression de ces protéines *in vivo* : des anticorps et des lymphocytes T spécifiques dirigés contre des protéines recombinantes ou contre des peptides dérivés de celles-ci ont pu être détectés dans le sérum de patients infectés [121-123] (pour une revue des rôles des protéines accessoires dans la réplication virale, voir [124]).

1.7 Le virus de l'immunodéficience humaine de type 1

1.7.1 Premières manifestations du SIDA et découverte du VIH

Le syndrome d'immunodéficience humaine (SIDA) a été décrit pour la première fois aux États-Unis au cours de l'année 1981, principalement chez de jeunes hommes homosexuels. Ces derniers présentaient des symptômes qui traduisaient une forte immunosuppression et cette maladie alors inconnue leur était fatale. Plusieurs d'entre eux présentaient des lésions cutanées violacées caractéristiques d'une forme rare et agressive de cancer, appelée sarcome de Kaposi. Peu d'options s'offraient alors aux médecins et les tentatives de traiter les différents symptômes n'avaient guère d'influence sur la dégradation progressive de l'état des patients. Puisque la maladie semblait surtout frapper les homosexuels et les utilisateurs de drogues intraveineuses (IDU), plusieurs ont alors tenté d'expliquer ce nouveau mal par des théories qui incluaient des facteurs associés au style de vie : utilisation de certaines drogues, réaction au sperme, mycose inconnue, nombre élevé de partenaires sexuels. Si ces théories parfois loufoques ont contribué au développement de plusieurs préjugés, la description quelques années plus tard des mêmes symptômes chez des patients hétérosexuels a rapidement démontré que cette nouvelle maladie n'épargnerait personne.

D'un point de vue immunologique, l'immunosuppression caractéristique des patients soupçonnés d'être affectés par le SIDA semblait résulter d'une diminution dramatique du nombre de lymphocyte T CD4+. Puisque quelques années auparavant, le rétrovirus HTLV-1 avait été décrit comme un pathogène infectant les cellules T, plusieurs scientifiques ont postulé que le SIDA pourrait être causé par un autre rétrovirus humain. C'est en 1983 que les premières évidences solides ont été apportées par l'équipe du Dr Luc Montagnier de l'Institut Pasteur à Paris, qui a effectivement détecté la présence d'un rétrovirus chez un patient présentant d'importantes lympho-adénopathies [125]. L'année suivante, une équipe américaine confirmait ces résultats, ce qui permettait d'établir un lien plus clair entre le SIDA et l'infection par ce nouveau rétrovirus [126-128]. Ce dernier fut nommé « Virus de l'Immunodéficience Humaine » ou VIH.

1.7.2 Infection par le VIH-1 et syndrome d'immunodéficience acquise (SIDA)

L'infection par le VIH-1 est caractérisée par 3 stades cliniques : une phase de virémie aiguë, une longue phase de latence dont la durée est relativement variable et une phase clinique d'immunodéficience ou phase SIDA (voir figure 12). La primo-infection est suivie d'une forte réplication virale et d'une importante réponse immune de l'hôte [129]. Chez environ 50 à 70 % des patients, on observe l'apparition d'un syndrome rétroviral aigu qui donne lieu à plusieurs manifestations cliniques non-spécifiques [129-131]. Les symptômes apparaissent environ 3 à 6 semaines après l'infection et s'apparentent à ceux d'une mononucléose. Les plus communs sont : fièvre, maux de gorge, fatigue, myalgies et perte de poids. Ils s'accompagnent parfois de lympho-adénopathies et de rash érythémateux. Durant cette période, on assiste à une importante dissémination du virus et à une diminution du nombre de lymphocytes T CD4+, majoritairement attribuable à l'effet cytopathique du virus sur les cellules infectées.

La réponse spécifique de l'hôte permet une diminution progressive du nombre de virus circulants. Elle implique d'abord l'expansion des lymphocytes T cytotoxiques CD8+ et éventuellement une forte réponse humorale. La séroconversion survient environ 1 à 4 semaines après l'infection [132].

Le contrôle de la réplication virale correspond au début d'une longue période de latence clinique au cours de laquelle la quantité de virus dans le sang est souvent sous la limite détectable. Au cours de cette phase asymptomatique, la réplication virale est toujours efficace [133], notamment dans les ganglions lymphatiques. Conséquemment, le taux de lymphocytes T continuera à décroître progressivement, jusqu'à rendre le système immunitaire inefficace. Le patient infecté devient alors la cible d'une foule d'infections opportunistes qui mèneront inévitablement à la mort.

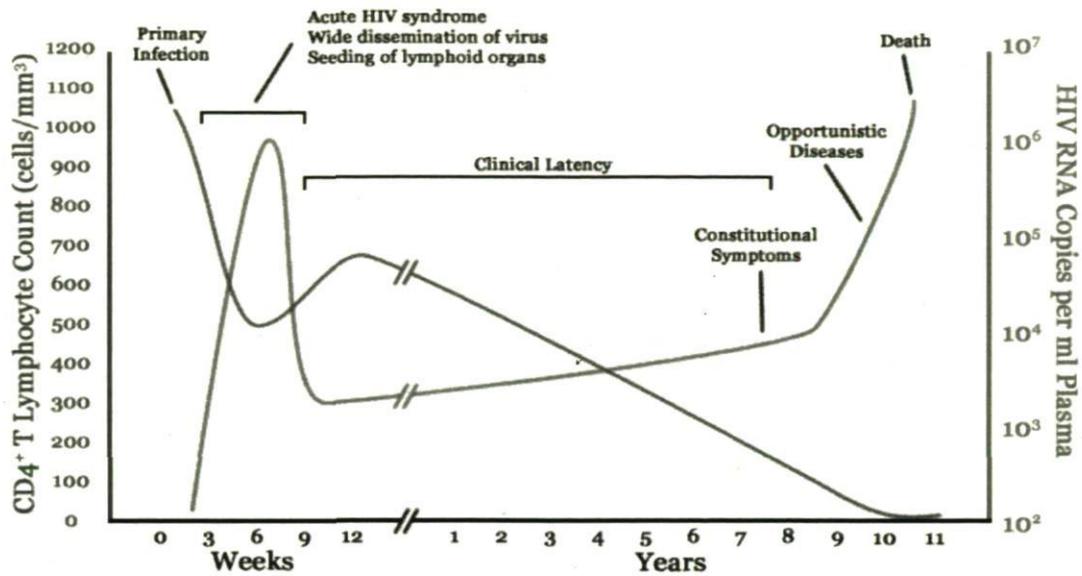


Figure 12 : Évolution du nombre de lymphocytes T CD4+ circulants et de la charge virale au cours de l'infection par le VIH-1. La primo-infection est suivie d'une forte répllication du virus et d'une importante diminution du nombre de lymphocytes T CD4+. La réponse de l'hôte permet un contrôle de l'infection qui mène à une période de latence virale durant laquelle le nombre de cellules CD4 continue de diminuer progressivement. Lorsque ce taux devient inférieur à 200 cellules / mm³, le système immunitaire devient incapable de contrôler les infections opportunistes.

1.7.3 Tropisme et répllication virale

Le tropisme du VIH-1 est déterminé par la présence de protéines cellulaires nécessaires au virus pour permettre l'entrée efficace dans la cellule. Peu de temps après la découverte du virus, des études ont conduit à l'identification de son récepteur : la molécule CD4 [134, 135]. Des recherches additionnelles ont ensuite permis de révéler que les récepteurs de chimiokines CXCR4 et CCR5 agissaient comme co-récepteurs pour le VIH-1 [136, 137]. La distribution de ces différents récepteurs permet au virus d'infecter non seulement les lymphocytes T CD4+, mais également les cellules impliquées dans la présentation des antigènes, comme les macrophages et les cellules dendritiques (DCs). Ce vaste tropisme offre au virus la possibilité d'exister sous différents états réplcatifs dans de nombreux

tissus, stratégie qui joue un rôle important dans sa capacité à échapper à la réponse immune de l'hôte.

Dans les lymphocytes T CD4⁺ activés, la réplication virale est très efficace : il semble que la grande majorité des virus retrouvés dans le plasma proviennent de ces cellules et que leur survie est écourtée par la réplication du virus [138]. Chez les lymphocytes T CD4⁺ quiescents par contre, le virus semble pouvoir exister sous forme latente ou transcriptionnellement inactive. Des analyses effectuées chez des patients sous traitements antiviraux ont pu démontrer que le virus pouvait effectivement persister sous forme latente et que sa réplication pouvait être observée *ex vivo* suite à l'activation de telles cellules [139, 140].

L'infection des macrophages résulte généralement de l'utilisation du co-récepteur CCR5. Il a été démontré que les personnes qui n'expriment pas le CCR5 sont résistantes à l'infection par le virus [141]. Ces données laissent croire que l'infection des macrophages (ou des lymphocytes T CCR5⁺) pourrait jouer un rôle important dans le développement de l'infection. Si la fonction précise des macrophages infectés dans la progression de l'infection est toujours mal définie, de nombreuses hypothèses ont été formulées sur leur rôle potentiel dans la dissémination du virus et dans l'établissement de réservoirs viraux [142-144] (pour une revue des mécanismes spécifiques à l'infection des macrophages par le VIH-1, voir [145]).

Des études plus récentes ont permis de mettre en lumière le rôle primordial des cellules dendritiques dans l'infection par le VIH-1. En plus d'être susceptibles à l'infection, les DCs peuvent capturer les virus pour ensuite les transférer aux cellules CD4⁺ [146, 147]. Cette capacité à lier les particules virales dépend de l'interaction d'une lectine de type C exprimée à la surface des DCs, DC-SIGN, avec la protéine d'enveloppe du virus. L'interaction avec DC-SIGN résulte en l'internalisation des virus dans un compartiment cellulaire non-lysosomal, ce qui permet au virus de rester infectieux et d'infecter les lymphocytes T en *trans* [148].

1.7.4 Traitements

Après plus de 25 ans de recherche sur le VIH, il n'existe toujours aucun traitement curatif pour la maladie qu'il provoque. Il demeure cependant important de mentionner que des progrès significatifs dans le traitement des patients infectés ont été rendus possibles grâce au développement de plusieurs drogues antirétrovirales efficaces à inhiber la réplication virale. Ces drogues ont permis d'améliorer considérablement le pronostic des patients infectés, pour ceux d'entre eux qui ont accès au traitement. Cependant, la capacité du virus à développer des résistances aux drogues affecte grandement les options de traitement et la progression de la maladie. Par conséquent, le développement de drogues ciblant de nouvelles étapes du cycle viral constitue une constante priorité.

La grande facilité du VIH-1 à développer des résistances est associée à sa capacité à évoluer rapidement et donc de générer une grande variabilité génétique. Deux principaux facteurs expliquent cette habileté du virus : la grande quantité d'erreurs produites par la transcriptase inverse virale lors de la synthèse de l'ADN proviral et le haut taux de réplication du virus. [149]. La génération de virus recombinants au cours de l'étape de transcription inverse contribue aussi à l'émergence de clones résistants.

La principale stratégie de traitement utilisée porte le nom de HAART (*highly active antiretroviral therapy*). Elle consiste en la combinaison de 3 ou 4 drogues ciblant 2 enzymes virales. Le plus souvent, 2 classes de drogues visant la transcriptase inverse sont combinées avec un inhibiteur de la protéase virale. Ces cocktails permettent souvent de réduire significativement la réplication virale pour plusieurs années, mais ils sont inefficaces à éradiquer complètement le virus qui semble rester latent dans certains sites souvent désignés comme « réservoirs viraux ».

1.7.5 Origines du VIH

En Afrique, plusieurs espèces de primates sont porteuses de virus qui s'apparentent fortement au VIH. De façon générale, ces virus ne causent pas de maladie lorsqu'ils infectent leurs hôtes naturels [150]. Des études phylogéniques ont permis de démontrer que le VIH de type 1, ou VIH-1 aurait évolué d'une souche de rétrovirus simien infectant une espèce spécifique de chimpanzés appelée *Pan troglodytes troglodytes* [151]. Ce virus, nommé SIVcpz, aurait apparemment franchi la barrière des espèces à trois occasions, donnant naissance à différents groupes de VIH-1, nommés M, N et O. Le VIH de type 2, ou VIH-2, résulterait de l'évolution d'une autre souche virale simienne infectant les singes Mangabey, le VISsm [152].

1.7.6 Épidémiologie et diversité génétique

Le VIH-1 est largement distribué sur la planète et est responsable de la grande majorité des infections par le VIH. Le VIH-2 est quant à lui beaucoup moins répandu et surtout retrouvé dans certains pays d'Afrique de l'Ouest et dans le sud-ouest de l'Inde [153]. Ce confinement à certaines régions géographiques s'explique par le fait que la transmission du VIH-2 est beaucoup moins efficace que celle du VIH-1, probablement en raison d'une production moins grande de particules virales durant la période de latence caractéristique au développement de la maladie.

En 2007, l'application de nouvelles méthodes pour évaluer l'importance de l'épidémie du SIDA a entraîné un changement du nombre estimé de personnes infectées à l'échelle du globe. En effet, en 2007, le nombre de personnes infectées par le VIH a été estimé à 33,2 millions, une diminution de 16% par rapport au nombre publié en 2006 (39,5 millions). Cette différence s'explique notamment grâce à un effort intensifié de l'Inde et de plusieurs pays sud-africains à mieux évaluer et caractériser le problème. Il semble tout de même qu'une partie de cette réduction puisse être attribuée à une diminution de nombre de

nouvelles infections, conséquence probable d'une décroissance des comportements sexuels à risque dans certains de ces pays. Le nombre de nouvelles infections en 2007 s'élève néanmoins à 2,5 millions et le nombre de décès associés à la maladie à 2,1 millions, soit 5700 par jour [154]. Ces chiffres reflètent durement l'importance de la pandémie et nous rappellent que malgré tous les efforts investis dans la prévention, la maladie progresse toujours.

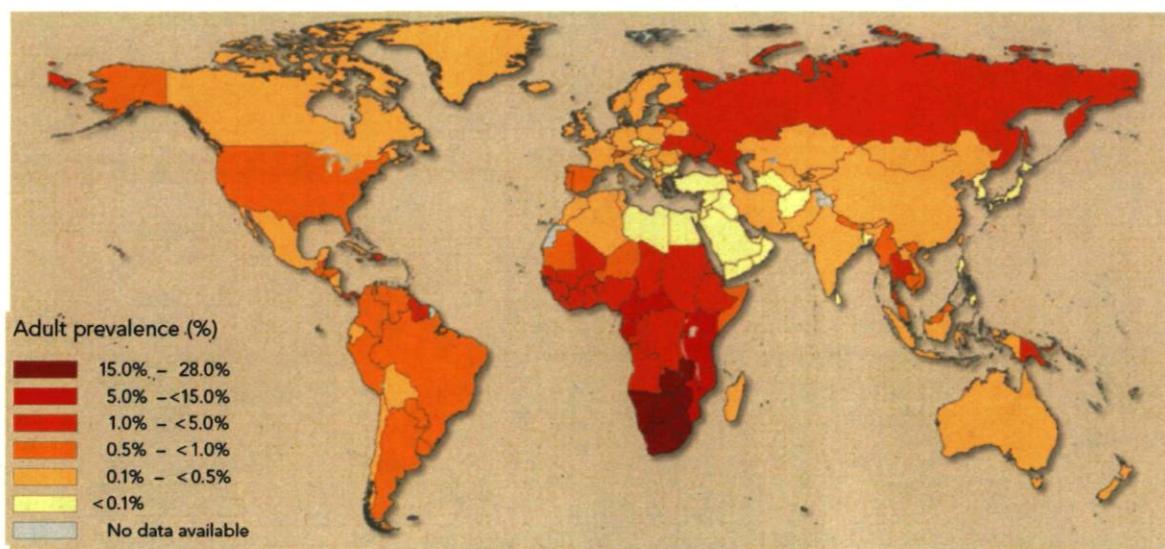


Figure 13 : Distribution géographique de l'épidémie du SIDA en 2007. Les régions les plus touchées par la pandémie sont les pays de l'Afrique subsaharienne ainsi que le continent asiatique (source : UNAIDS).

Les relations sexuelles hétérosexuelles demeurent la principale voie de transmission et sont à l'origine de 85% des infections à l'échelle mondiale. L'Afrique sub-Saharienne, où le taux de nouvelles infections reste très élevé, constitue le principal centre de la pandémie (voir figure 13). L'Asie, où le virus s'est répandu plus récemment, est devenue la deuxième région la plus affectée par ce fléau. Là aussi, les relations sexuelles non-protégées jouent un rôle majeur dans la propagation de la maladie, mais cette dernière est aussi favorisée par un nombre élevé de transmissions résultant de l'utilisation de drogues par voie intraveineuse.

Le VIH-1 est classifié en 3 différents groupes : M (*Major*), O (*Outlier*) et N (*New*). Le groupe M est le plus répandu et est considéré comme le responsable de la pandémie actuelle. Les infections par un virus du groupe O sont plus rares et ne comptent que pour environ 10% des infections par le VIH-1 à l'échelle mondiale [155, 156]. Le groupe M est composé de 11 sous-types ou *clades*, désignés par les lettres A à K, en plus de plusieurs autres formes de virus recombinants appelés *Recombinant Circulating Forms* ou RCF. Ces sous-types sont principalement classés en fonction de la séquence codant pour les protéines d'enveloppe, bien que les séquences d'autres gènes soient aussi utilisées.

La classification en sous-types est un outil important qui permet de mieux comprendre l'évolution de l'épidémie à l'échelle mondiale : chaque région du globe est préférentiellement affectée par différents sous-types. Si l'existence de différences significatives quant à la pathogénicité ou la transmissibilité des différents sous-types n'a toujours pas été démontrée, il semble néanmoins que certains variants répondent mieux au traitement anti-rétroviral. D'autres études ont également mis en évidence la présence de variations dans les séquences LTR des différents sous-types et l'impact de ces variations sur la régulation de la transcription des gènes viraux [157-160].

1.7.7 Stratégie répllicative

Le génome du VIH de type 1 est d'environ 9,8 Kb. Il contient 9 gènes qui codent pour 15 protéines virales. Comme celui des autres rétrovirus, il contient les gènes *gag*, *pol*, et *env* et est bordé par les régions LTR. Il code également pour plusieurs protéines régulatrices et accessoires qui ont toutes été largement étudiées et qui semblent toutes jouer un rôle important dans la persistance du virus *in vivo*. La production des différentes protéines virales est possible grâce à l'épissage alternatif des transcrits viraux qui génère plusieurs espèces d'ARNm (voir figure 14). Ces derniers peuvent être classés en trois catégories : les

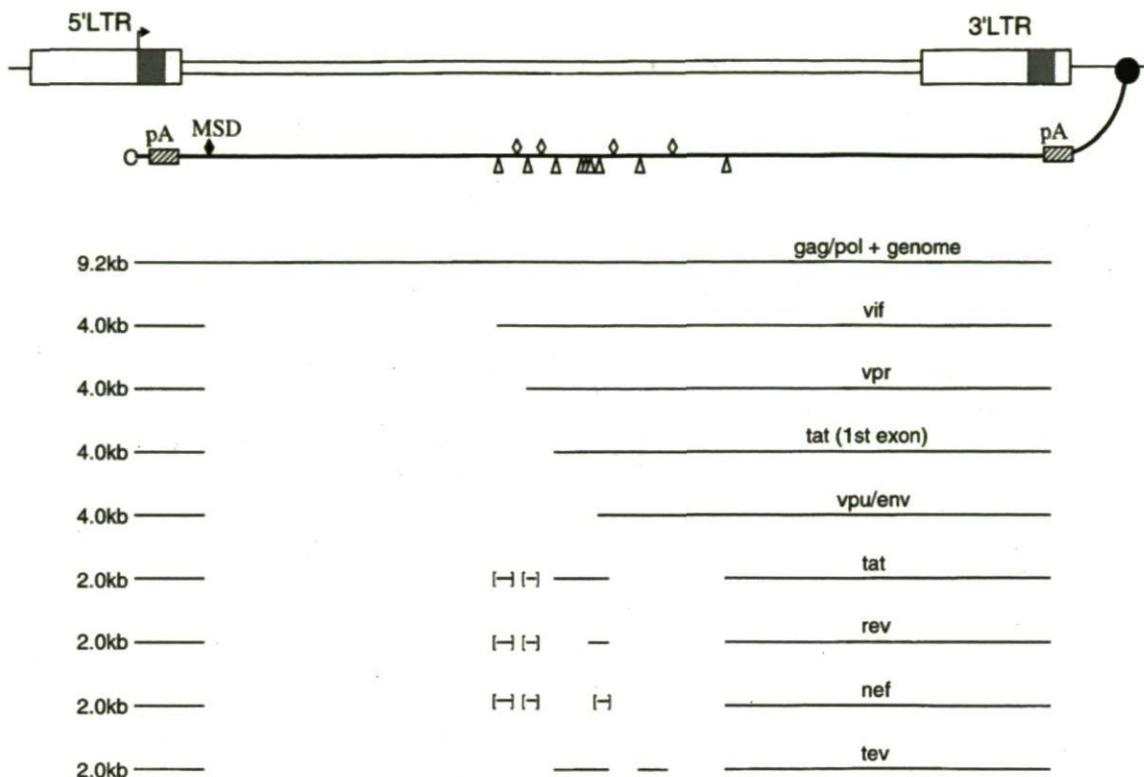


Figure 14 : Le génome du VIH-1 et les différents ARNm produits par le virus. Le virus exprime plusieurs ARNm regroupés en 3 classes : non-épissés (environ 9 kb), mono-épissés (environ 4 kb) et multi-épissés (environ 2 kb). Les triangles représentent les sites donneurs et les losanges les sites accepteurs. MSD : *major splice donor*. pA : *polyadenylation signal* (tiré de [161]).

ARNm de pleine longueur (non-épissés) de 9 kb; les ARNm mono-épissés d'environ 4 kb; les ARNm multi-épissés d'environ 2 kb. Un grand nombre de sites accepteurs et donneurs d'épissage permettent au virus de produire au moins 30 différents types d'ARNm [162]. L'expression de ces messagers est hautement régulée au cours du cycle viral. Dans les premiers stades de la réplication, les ARNm de 2 Kb sont les plus abondants et permettent de produire les protéines Tat, Rev et Nef en grande quantité. Par la suite, il y a une augmentation progressive du nombre d'ARNm non-épissés et mono-épissés, ce qui permet la production des protéines de structure et des protéines accessoires à la fin du cycle [163]. Cette transition est au moins partiellement médiée par la protéine Rev et dépend de l'accumulation de cette dernière jusqu'à l'atteinte d'une concentration critique [164]. Tout

comme HTLV-1, le VIH-1 utilise le décalage du cadre de lecture pour réguler la production des précurseurs protéiques Gag et Gag-Pol.

1.7.8 Les régions LTR

Le LTR 5' du VIH-1 joue le rôle de promoteur viral. Il contient de nombreux sites permettant la liaison de protéines cellulaires capables de réguler positivement ou négativement la transcription virale (voir figure 15, pour une revue, voir [165]). C'est à la jonction des régions U3 et R qu'est initiée la transcription. La région U3 contient une région de régulation négative ou NRE, une région *enhancer* et une région promotrice basale. La région NRE contient plusieurs sites de liaison pour des facteurs transcriptionnels cellulaires tels que USF-1, NFAT1 et AP-1. Alors que certaines études avaient démontré que la suppression de cette région entraînait une faible augmentation du niveau de transcription virale [166], des analyses plus récentes ont plutôt attribué au NRE un rôle de régulation positive.

La région *enhancer* s'étend des nucléotides -82 à -103. Elle contient 2 sites de liaison pour le facteur transcriptionnel NF- κ B qui sont essentiels à la réplication virale. Dans les cellules T, des stimuli tels que certaines protéines virales, les esters de phorbol et plusieurs cytokines peuvent entraîner l'activation de NF- κ B, qui joue un rôle prépondérant dans l'inflammation. Plusieurs récentes études ont également mis en évidence la capacité du facteur transcriptionnel NFAT à se lier à la région *enhancer* du VIH-1 en combinaison avec NF- κ B [167, 168]. L'induction de NFAT favorise la réplication virale dans les lymphocytes primaires [169, 170], ce qui laisse croire que ce facteur transcriptionnel pourrait également jouer un rôle important *in vivo*.

La région promotrice basale contient trois sites de liaison pour le facteur de transcription Sp1. Bien que la suppression de ces sites n'empêche pas la production de particules virales infectieuses, il semble qu'ils jouent un rôle important pour assurer une transactivation

optimale par la protéine virale Tat [171].

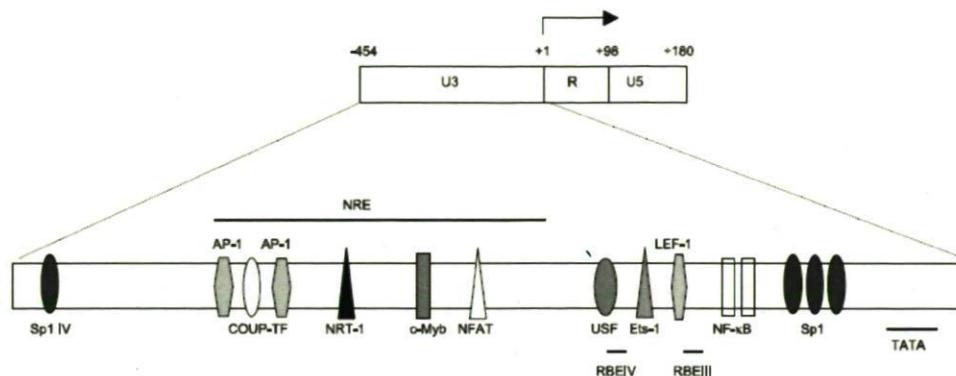


Figure 15 : Le LTR du VIH-1. La région U3 contient les sites de liaison pour les facteurs transcriptionnels qui contrôlent la transcription virale. La région enhancer, qui contient 2 sites de liaison pour NF- κ B, est essentielle à la réplication virale. (adapté de [172])

En aval du site d'initiation de la transcription, se trouvent deux éléments uniquement retrouvés chez certains rétrovirus : l'élément inducteur de courts transcrits (IST) et l'élément de réponse à Tat ou TAR. Ensemble, ils permettent au virus d'adapter son expression génique en fonction des différents environnements cellulaires. L'élément IST s'étend des nucléotides -5 à +82 et induit la synthèse de courts ARN qui s'interrompent prématurément [173]. Il semble aussi stimuler la formation de complexes de pré-initiation incapable d'entraîner une élongation efficace. Les transcrits formés présentent des extrémités 3' hétérogènes et contiennent l'élément TAR. En absence de la protéine Tat, ils constituent la majorité des ARN synthétisés par le promoteur viral. En présence de Tat, la quantité de transcrits de pleine longueur polyadénylés augmente de façon très importante [174] et cette augmentation dépend de l'interaction spécifique de Tat avec l'élément TAR. Ce dernier est un élément de régulation positive unique de par le fait qu'il est efficace seulement sous forme d'ARN. Il forme une structure contenant une tige de nucléotides

pairés, une courte région de 3 nucléotides non-pairés et une boucle riche en guanine de 6 nucléotides.

1.7.9 La protéine Tat

Tat est une protéine de 14 kDa codée par 2 exons : l'exon 1 code pour les acides aminés 1-72 alors que le second exon code pour les acides aminés 73-86 ou 73-101 selon la souche virale. Tat exerce son action en liant l'élément TAR en 5' des transcrits et en recrutant au promoteur viral plusieurs complexes transcriptionnels. Ceux-ci incluent des enzymes possédant des activités histones acétyl-transférase (HAT) qui sont capables de modifier la conformation de la chromatine et le complexe protéique P-TEFb, composé de la cycline T1 et de la kinase CDK9. Ce dernier est capable de phosphoryler l'ARN polymérase II à son extrémité C-terminale, ce qui a pour conséquence l'augmentation de la processivité de la polymérase et donc l'élongation des transcrits. L'action de Tat est dépendante de la présence de l'élément TAR, mais elle nécessite également la présence de co-facteurs transcriptionnels. Tat interagit notamment avec les complexes de pré-initiation qui incluent Sp1, TBP (*TATA-binding protein*), TFIIF, Tip60, et l'ARN polymérase II. Tat est aussi capable de recruter les co-activateurs CBP/p300 and p/CAF [175].

Dans la transactivation du promoteur viral médiée par Tat, le recrutement de nombreuses protéines cellulaires et les modifications de Tat par plusieurs enzymes jouent un rôle crucial. L'acétylation de Tat par les HATs régule sa liaison à l'élément TAR : PCAF, p300/CPB et pGCN5 peuvent toutes acétyler Tat sur différents résidus lysine [175-178]. La protéine Tat acétylée en K28 possède une plus grande affinité pour le complexe cycline T1/CDK9. Par contre, son acétylation en K50 entraîne une baisse d'affinité pour l'élément TAR [179, 180].

Tat interagit également avec la protéine arginine méthyl-transférase 6 (PRMT6), ce qui résulte en la méthylation de Tat sur certains résidus arginine. Cette modification semble

affecter négativement l'action de Tat, puisque la surexpression de PRMT6 entraîne une diminution de la transcription dépendante de Tat à partir du promoteur viral [181].

L'ubiquitine ligase Hdm2 peut lier Tat et promouvoir son ubiquitination sur la Lys71. De façon intéressante, cette modification n'entraîne pas la dégradation de la protéine via le protéasome : elle semble plutôt avoir un effet positif sur le pouvoir transactivateur de la protéine [182].

1.7.10 La protéine Rev

Rev est une protéine de 19 kDa qui régule l'export des transcrits viraux non-épissés et mono-épissés. Elle est codée par deux exons qui sont tous deux essentiels à sa fonction. Elle contient deux domaines fonctionnels : un domaine riche en arginine, requis pour la liaison à l'ARN et la localisation nucléaire, et un domaine hydrophobe riche en leucine important pour l'export nucléaire des ARNm. Les ARNm viraux de pleine longueur (9 kb) et mono-épissés (environ 4 kb) contiennent tous une courte séquence d'ARN d'environ 250 nucléotides qui permet l'action de Rev. Cette séquence est située dans le gène *env* et forme une structure complexe en tige-boucle nommée RRE. En absence de Rev, les ARNm qui contiennent le RRE sont retenus dans le noyau. Lorsqu'elle est exprimée, Rev agit en formant un complexe multimérique sur le RRE, ce qui permet l'interaction avec des constituants de la machinerie d'export nucléaire de la cellule. Rev permet aussi de contrecarrer l'effet de séquences présentes dans les transcrits viraux qui entraînent la rétention de ces derniers dans le noyau *via* l'interaction avec certaines protéines cellulaires.

1.7.11 Les protéines accessoires

Les virus de l'immunodéficience des primates, dont le VIH-1, sont caractérisés par la présence de plusieurs ORF codant pour des protéines accessoires. Les gènes *vif*, *vpr* et *nef* sont présents chez la majorité des isolats du VIH-1, VIH-2 et VIS. En revanche, le gène *vpu*

est retrouvé seulement chez le VIH-1 et certains isolats du VIS. Aucune de ces protéines ne possède une activité enzymatique : elles semblent plutôt toutes jouer un rôle de molécules adaptatrices au cours d'interactions physiques avec d'autres protéines virales ou cellulaires. De façon générale, l'activité de ces protéines semble permettre de contourner plusieurs mécanismes antiviraux de la cellule et de modifier l'environnement cellulaire pour assurer la réplication et la transmission efficaces du virus (pour une revue des fonctions des protéines accessoires du VIH-1, voir [183]).

Chapitre 2 : La transcription antisens

2.1 Transcription chez les eucaryotes : généralités

2.1.1 L'initiation de la transcription

Chez les eucaryotes, on retrouve 3 types d'ARN polymérases, nommées I, II et III. L'ARN polymérase I est localisée dans le nucléole et est impliquée dans la synthèse des ARN ribosomiques. L'ARN polymérase III transcrit les gènes codant pour les ARN de transfert, l'ARN ribosomique 5S et d'autres petits ARN stables. Tous les ARN dont le rôle consiste à coder pour des protéines, les ARNm, sont transcrits par l'ARN polymérase II. Des études ont rapidement démontré que l'action de cette enzyme nécessitait la présence de plusieurs cofacteurs [184, 185]. Ces facteurs ont été identifiés comme étant des facteurs de transcription généraux (GTFs) et ont été nommés TFIIA, TFIIB, TFIID, TFIIE, TFIIF et TFIIH, selon la fraction cellulaire à partir de laquelle ils ont été isolés. Le tableau 1 résume la nature des composantes protéiques de ces facteurs et leurs rôles respectifs dans la machinerie transcriptionnelle [186]. En présence de ces facteurs, l'ARN polymérase II purifiée peut lier spécifiquement une séquence promotrice et initier la transcription *in vitro* [187].

L'étude de la transcription chez les eucaryotes a permis l'identification des éléments constitutifs élémentaires des promoteurs types (voir tableau 2). Ces derniers sont en fait des séquences caractéristiques d'ADN requises pour les fonctions promotrices ainsi que pour l'assemblage et l'orientation adéquats des complexes de pré-initiation (PIC). Tout d'abord, la boîte TATA est une région riche en A/T située environ à 25-30 nucléotides en amont du site d'initiation de la transcription (+1). La reconnaissance de la séquence consensus TATA(A/T)A(A/T)(A/G) par la TBP (*TATA-binding protein*) présente dans le complexe TFIID entraîne la formation du PIC. Alors que les résultats de premières études avaient suggéré que la majorité des promoteurs eucaryotes contenaient des boîtes TATA,

des analyses plus récentes réalisées à partir de bases de données ont démontré qu'elle sont en fait présentes chez seulement 32% des promoteurs [188].

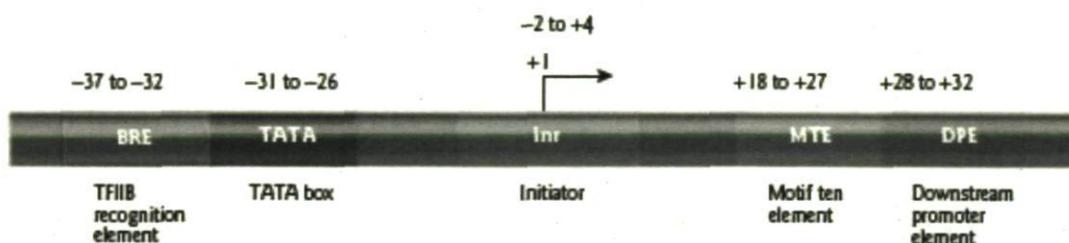
General transcription machinery	RNA pol II (pol II, RNAPII)	RNA polymerase II	Catalysis of RNA synthesis
General transcription factors:			
TFIIB		Transcription factor for RNA polymerase II B	Stabilization of TBP-DNA interactions, recruitment of RNA pol II-TFIIF, start site selection by RNA pol II
TFIID:			
TBP		TATA-binding protein	Core promoter recognition, TFIIB recruitment
TAF		TBP-associated factor	Core promoter recognition/selectivity
TFIIE		Transcription factor for RNA polymerase II E	TFIIF recruitment
TFIIF		Transcription factor for RNA polymerase II F	Recruitment of RNA pol II to promoter DNA-TBP-TFIIB complex
TFIIH		Transcription factor for RNA polymerase II H	Promoter melting, helicase, RNA pol II CTD kinase
Mediator		Mediator	Transduces regulatory information from activator and repressor proteins to RNA pol II

Tableau 1 : Rôle des différents facteurs de transcription généraux formant la machinerie transcriptionnelle. Le tableau présente les abréviations, les noms complets et les fonctions des principaux acteurs de la transcription chez les eucaryotes. (tiré de [188])

Un second élément, l'initiateur (Inr), contient une séquence riche en pyrimidines et borde le site d'initiation de la transcription. L'initiateur est fonctionnellement similaire à la boîte TATA et est capable de diriger l'initiation de façon indépendante. Par contre, chez les promoteurs contenant une boîte TATA, les deux éléments agissent en synergie pour augmenter l'efficacité de l'initiation de la transcription.

Un troisième élément constituant des promoteurs est le *downstream promoter element* (DPE) [189]. On le retrouve environ 30 nucléotides en aval de plusieurs promoteurs dépourvus de boîte TATA. Le DPE est reconnu par TFIID qui pourrait agir en collaboration avec NC2, qui a initialement été caractérisé comme étant un cofacteur agissant négativement sur des promoteurs contenant des boîtes TATA. En plus du DPE, un autre élément est situé en aval du site d'initiation de la transcription: le MTE (*Motif Ten*

Element) (*Downstream Core Element*). Le MTE est normalement retrouvé entre les nucléotides +18 à +27 et fonctionne en association avec l'initiateur pour entraîner une augmentation de la transcription dépendante de l'ARN polymérase II [190]. Il peut également compenser l'absence d'une boîte TATA ou d'un DPE ou même agir de concert avec eux pour augmenter l'activité d'un promoteur de façon Inr-indépendante.



Promoter	Position	Transcription factor	Consensus sequence
Upstream core promoter elements TFIIB recognition element (BRE) TATA box Initiator (Inr)	-37 to -32 -31 to -26 -2 to +4	TFIIB TBP TAF1 (TAF _{II} 250) TAF2 (TAF _{II} 150)	(G/C)(G/C)(G/A)CGCC TATA(A/T)AA(G/A) PyPyA _n N(T/A)PyPy
Downstream core promoter elements Motif ten element (MTE) Downstream promoter element (DPE)	+18 to +27 +28 to +32	TFIID TAF9 (TAF _{II} 40) TAF6 (TAF _{II} 60)	C(G/A)A(A/G)C(G/C) (C/A/G)AACG(G/C) (A/G)G(A/T)(C/T)(G/A/C)
Proximal promoter elements CAAT box GC box	-200 to -70 -200 to -70	CBF, NF1, C/EBP Sp1	CCAAT GGCGG

Tableau 2 : Position des éléments retrouvés dans les promoteurs basaux de l'ARN polymérase II et les facteurs transcriptionnels par lesquels ils sont reconnus. (tiré de [188])

Un autre constituant des promoteurs est reconnu par un cofacteur différent, le TFIIB, à la suite d'une courbure de l'ADN provoquée par la liaison de TBP. Il s'agit du BRE (*TFIIB Recognition Element*) [191]. Cette séquence semble jouer un rôle dans l'orientation des PIC et est retrouvée chez les promoteurs contenant ou non une boîte TATA.

Au cours de la dernière décennie, de nombreuses études ont été entreprises pour analyser les différentes étapes menant à la formation des PIC sur les régions promotrices. Des travaux effectués en utilisant un promoteur adénoviral ont permis d'établir l'ordre d'entrée et la position relative de plusieurs cofacteurs dans le PIC. Ces résultats ont permis de proposer un modèle d'addition séquentielle des différents cofacteurs TFIIA à H pour la formation de ce dernier, dont la première étape consiste en la reconnaissance de la boîte TATA par TFIID. D'autres groupes ont par la suite mis en évidence l'existence d'une autre voie de formation des PIC en découvrant que l'ARN polymérase II pouvait être purifiée sous forme de complexe holoenzyme pré-assemblé ne contenant pas nécessairement les cofacteurs cités ci-haut [192]. Bien que la nature exacte des composants de ce complexe varie selon la méthode de purification, il semble qu'il contienne également d'autres protéines impliquées dans le changement de conformation de la chromatine, la réparation de l'ADN et le traitement de l'ARN. Il est probable que la voie séquentielle et la voie holoenzymatique co-existent *in vivo* et qu'elles soient utilisées dans différents contextes.

2.1.2 Activation de la transcription

Les expériences qui ont mené à une meilleure compréhension des mécanismes d'initiation de la transcription ont souvent été effectuées en utilisant des modèles *in vitro*. Le degré de transcription généré en présence d'ARN polymérase II, des GTF et d'ADN dans de tels modèles est souvent considéré comme étant un niveau basal. Toujours dans ces modèles *in vitro*, il a été démontré que ce niveau basal peut être augmenté par l'addition de protéines capables de lier l'ADN en des sites spécifiques, normalement situés en amont de la séquence promotrice reconnue par l'ARN polymérase II. Ces facteurs sont connus comme étant des activateurs transcriptionnels et sont nécessaires à l'obtention d'un niveau de transcription dite « activée ». Dans la cellule eucaryote, l'ADN est empaqueté à l'intérieur de complexes contenant des octamères d'histones appelés nucléosomes. L'accessibilité à l'ADN dans un tel contexte est très différente de celle associée aux modèles *in vitro*. L'initiation de la transcription *in vivo* est donc un phénomène plus complexe qui nécessite

la participation de facteurs additionnels, dont les activateurs transcriptionnels. En effet, le complexe transcriptionnel contenant l'ARN polymérase II est généralement incapable d'être actif sur une séquence promotrice *in vivo* et le niveau de transcription basale observé *in vitro* souvent inexistant [193]. En fait, plutôt que de suivre la loi du tout ou rien, le processus de transcription *in vivo* semble varier graduellement et de façon très importante en fonction de la conformation de l'ADN, du nombre et du type de facteurs transcriptionnels collaborant avec la polymérase et du nombre de sites de liaisons pour ces facteurs sur l'ADN.

Un promoteur eucaryote type contient souvent plusieurs éléments promoteurs proximaux, localisés environ 70-200 pb en amont du site d'initiation de la transcription. On y retrouve des sites de reconnaissance pour des facteurs transcriptionnels, comme par exemple la boîte CAAT et les boîtes GC. La boîte CAAT est reconnue par la *CAAT-enhancer binding protein* (C/EBP) et les boîtes GC sont des sites de liaison pour le facteur de transcription Sp1. Lorsque ces éléments sont situés près d'un site d'initiation de la transcription, ils entraînent une augmentation de la fréquence de l'initiation de la transcription. Par contre, les facteurs transcriptionnels qui lient les promoteurs proximaux n'entraînent pas nécessairement une augmentation ou une diminution directe de la transcription. En effet, ils peuvent parfois agir en tant que facteur d'attachement pour recruter d'autres éléments régulateurs, tels que les *enhancer*, au niveau du promoteur basal. Les éléments *enhancer* sont normalement plus éloignés du site d'initiation (environ 700-1000 pb). La particularité des éléments *enhancer* consiste au fait qu'ils peuvent être situés en amont, en aval ou même à l'intérieur d'un intron et fonctionner dans n'importe quelle orientation par rapport au promoteur. Un *enhancer* typique est d'environ 500 pb et contient une dizaine de sites de liaison pour des facteurs transcriptionnels. Les *enhancers* régulent positivement l'expression des gènes et peuvent être spécifiques de certains tissus ou de certains stades du développement. Des éléments similaires aux *enhancers* qui régulent négativement l'expression des gènes sont appelés *silencers*.

Les facteurs transcriptionnels qui agissent comme répresseurs exercent leur action en bloquant la machinerie transcriptionnelle, tandis que les facteurs qui activent la transcription peuvent agir [188]:

- 1) En favorisant le recrutement et la liaison des facteurs transcriptionnels généraux et de la RNA polymérase II sur le promoteur basal;
- 2) En entraînant un changement de conformation ou un changement post-traductionnel (ex : phosphorylation) qui stimule l'activité enzymatique de la machinerie transcriptionnelle;
- 3) En interagissant avec des complexes responsables de modifications ou de changements de conformation de la chromatine pour augmenter l'accessibilité des facteurs transcriptionnels ou d'autres activateurs à l'ADN.

2.1.3 *Traitement des ARNm*

Alors que chez les organismes procaryotes les ribosomes ont accès à l'ARNm au moment où le processus de transcription est toujours en cours, le mécanisme est très différent chez les eucaryotes. En effet, chez ces derniers, les ARNm synthétisés dans le noyau doivent être exportés dans le cytoplasme avant de pouvoir servir à la synthèse protéique. De plus, les ARNm initiaux ou pré-ARNm doivent également subir plusieurs modifications. Une première survient immédiatement après l'initiation de la transcription et consiste en l'addition à extrémité 5' de l'ARN synthétisé d'une structure 7-méthylguanosine appelée *cap* ou coiffe. Cette modification débute par l'addition d'un GTP dans l'orientation inverse au dernier nucléotide en 5' du pré-ARNm *via* la génération d'un lien 5'-5', catalysé par la guanylyl transférase. Cette nouvelle guanosine en 5' est par la suite méthylée par une autre enzyme, la guanine méthyltransférase. La coiffe joue un rôle dans l'alignement des ARNm sur le ribosome au cours de la traduction.

Chez les eucaryotes, l'extrémité 3' des ARNm résulte du clivage du transcrit et de l'addition d'une queue de poly-A, par un processus nommé polyadénylation. Le signal

consensus le plus commun est situé environ 10 à 30 nucléotides en amont du site de polyadénylation et est composé des 6 nucléotides AAUAAA. Ce signal est fréquemment suivi 10-30 nucléotides plus loin par une région riche en G et U. Ces deux séquences sont des sites de liaison pour les complexes multi-protéiques CPSF (*cleavage and polyadenylation specificity factor*) et CstF (*cleavage stimulation factor*). Ces complexes permettent le recrutement de la poly-A polymérase qui procédera à l'ajout de la queue de poly-A, qui peut contenir jusqu'à 250 adénosines. En fait, CPSF interagit avec TFIID dès l'initiation de la transcription et permet de mettre en branle le processus de polyadénylation lorsque le signal consensus est rencontré. Plusieurs fonctions ont été attribuées à la queue de polyA et il semblerait qu'elle pourrait jouer un rôle dans la stabilité des transcrits et dans l'initiation de la traduction.

2.1.4 *Élongation des transcrits*

Une fois la machinerie transcriptionnelle mise en branle, l'ARN polymérase II peut fonctionner au rythme de 2000 nucléotides par minute. Cependant, les pré-ARNm de certains gènes peuvent être très longs et faire jusqu'à plus de 2000 kb, en grande partie en raison des nombreux introns qu'ils contiennent. L'extrême longueur des gènes eucaryotes exige donc une grande stabilité des complexes de transcription pour éviter les pauses ou les interruptions du processus. Plusieurs facteurs d'élongation sont capables de s'associer à la polymérase pour assurer cette stabilité [194]. Un second obstacle rencontré par l'ARN polymérase II consiste à poursuivre la transcription sur l'ADN organisé autour des nucléosomes. Il semble encore une fois que de nombreux facteurs d'élongation aient la capacité de modifier la structure de la chromatine de façon à favoriser la transcription [195].

2.1.5 Épissage

Il existe plusieurs types d'introns chez les eucaryotes dont 2 sont retrouvés chez les ARN codants : les introns GU-AG et les introns AU-AC. Le premier type est de loin le plus fréquent. L'épissage des introns s'effectue alors que la transcription est toujours en cours. Il consiste en une succession de 2 réactions de trans-estérification qui permet de joindre les exons et de libérer l'intron sous forme d'une boucle nommée *lariat*. La libération de l'intron nécessite la présence d'une séquence à l'intérieur de ce dernier appelée site de branchement. L'épissage est possible grâce à la participation de plusieurs petits ARN nucléaires ou snARN nommés U1, U2, U4, U5 et U6. Ces derniers s'agencent avec plusieurs sous-unités protéiques pour former les snRNP (*small nuclear ribonucleoprotein*). Les snRNP constituent l'élément de base du splicéosome, un large complexe à l'intérieur duquel s'effectue l'épissage des pre-ARNm [196]. Au cours de la réaction, le site de branchement ainsi que les sites donneurs et accepteurs sont reconnus par les snARN dans une série d'appariements et de réarrangements qui permettent d'assurer un haut niveau de spécificité. Les différents réarrangements entraînent la création d'un site catalytique à l'intérieur du splicéosome. Les ARN U2 et U6 forment éventuellement une structure tridimensionnelle juxtaposée au site donneur et au site de branchement, ce qui permet la première réaction de trans-estérification. Par la suite, la participation de l'ARN U5 permet de joindre les sites donneurs et accepteurs et de faciliter la seconde réaction. Une fois l'intron libéré, le splicéosome subit d'autres changements de conformation de façon à ce qu'il puisse agir dans une nouvelle réaction (les différentes étapes de l'épissage sont présentés dans la figure 16).

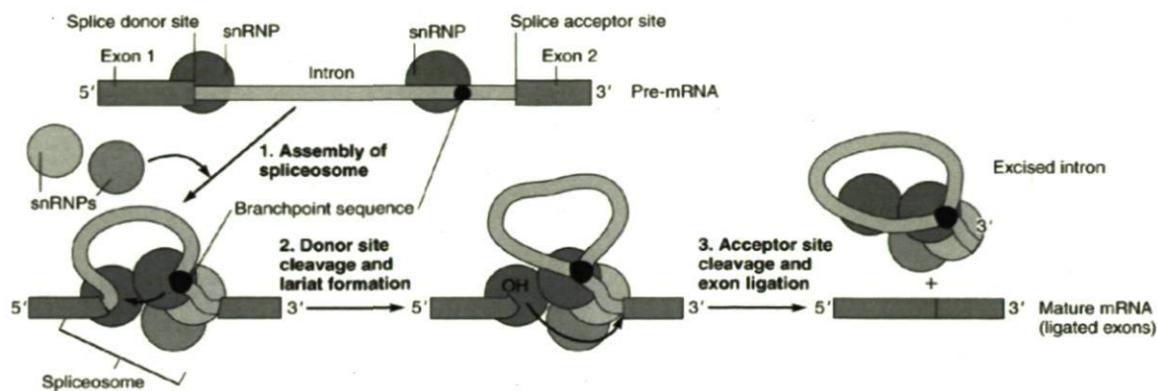


Figure 16 : Principales étapes de l'épissage des ARNm. Les sites sont reconnus par les snRNP, permettant ainsi l'assemblage du splicéosome. La formation d'un site catalytique permet le clivage aux sites donneurs et accepteurs, ce qui résulte en la libération de la séquence intronique et la formation d'un ARNm mature. (tiré de [197])

2.1.6 Export nucléaire des ARNm

Le transport des macromolécules entre le noyau et le cytoplasme est possible grâce à la présence de complexes formant les pores nucléaires ou NPC. Les NPC sont formés par une association de plusieurs nucléoporines et contiennent un canal aqueux pouvant atteindre environ 25 nm de diamètre lors du transport actif. Les NPC au repos demeurent partiellement ouverts (formant des pores d'environ 9 nm de diamètre), ce qui implique que de petites molécules protéiques (≤ 40 kDa) ou d'ARN peuvent passer du noyau au cytoplasme par diffusion passive. Ce mécanisme est cependant plutôt inefficace et même les plus petits ARN qui traversent la membrane nucléaire le font généralement par transport actif [198].

La grande majorité du transport entre le noyau et le cytoplasme est médié par une famille de protéines appelées karyophérines ou importines/exportines. Plusieurs membres de cette famille ont la capacité d'interagir avec des protéines cargo ainsi qu'avec des nucléoporines spécifiques. Leur action dépend de l'activité biologique d'un cofacteur essentiel : la

GTPase Ran (pour une revue voir [199]). De nombreux systèmes ont été utilisés pour identifier les facteurs jouant un rôle important dans l'export d'ARN et l'étude de l'exportation des ARNm rétroviraux fut sans doute l'un des plus efficaces [200]. Des études ont notamment démontré le rôle important de la protéine Crm-1 dans l'export des ARNm viraux chez le VIH-1. La forme de Crm-1 lié au Ran-GTP lie la protéine virale Rev sur son signal d'export nucléaire pour permettre l'export des ARNm contenant le RRE [201, 202]. On sait maintenant que Crm-1 est un joueur majeur dans le mécanisme d'export au cytoplasme des ARN et des protéines cellulaires car il exerce non seulement un rôle central dans l'export de la majorité des protéines qui voyagent entre le noyau et le cytoplasme, mais il est également responsable de l'export des snARN, des ARN ribosomiques ainsi que d'un petit groupe d'ARNm.

Par contre, l'export nucléaire de la majorité des ARNm est indépendant du système impliquant Ran et les karyophérines. Les acteurs clés de l'export des ARNm sont plutôt la protéine ubiquitaire Tap et un petit cofacteur nommé Nxt. L'export nucléaire médié par Tap ne requiert pas d'interaction avec les karyophérines : il implique plutôt deux domaines essentiels de Tap capables de reconnaître directement les NPC [203, 204]. La formation d'hétérodimères de Tap et Nxt favorise leur interaction avec les nucléoporines et éventuellement le transit des ARN vers le cytoplasme.

L'identification du rôle de Tap dans l'export des ARNm est le résultat d'études effectuées sur le rétrovirus MPMV (*Mason Pfizer Monkey Virus*). Comme chez les autres rétrovirus, la réplication de MPMV nécessite l'export d'ARNm non-épissés ou partiellement épissés. Or, ce virus ne produit pas de protéine adaptatrice comme Rev qui peut recruter Crm-1 : il exprime plutôt une séquence d'ARN capable d'entraîner l'export des transcrits indépendamment de la présence des protéines virales. Cette séquence nommée CTE (*constitutive transport element*) permet l'export nucléaire de façon Tap-dépendante [205] (pour une revue des différentes voies d'export de l'ARN chez les eucaryotes, voir [206]).

2.2 La transcription antisens

2.2.1 Généralités

Les transcrits antisens naturels ou NAT (*natural antisense transcripts*) sont des molécules d'ARN endogènes dont la séquence est complémentaire à un autre transcrit. Ils peuvent être divisés en deux groupes : les cis-NAT, qui sont transcrits à partir du brin opposé d'ADN du même locus que le transcrit sens et les trans-NAT, qui sont transcrits à partir d'une région distincte du génome et qui présentent donc une complémentarité partielle ou imparfaite avec un ou plusieurs transcrits sens. En fait, le terme « transcrit sens » est plutôt mal défini et fait souvent référence au transcrit codant, puisqu'un grand nombre de transcrits antisens sont en effet non-codants. Les NAT peuvent être classifiés sur la base de l'orientation de leur hybridation avec le transcrit sens : ils peuvent être divergents (hybridation en 5'), convergents (hybridation en 3') ou complètement chevauchants (un gène étant contenu complètement dans la séquence d'un autre) (voir figure 17). Ces paires de transcrits peuvent être formées par deux transcrits codant pour des protéines, un transcrit codant et un non-codant ou encore deux transcrits non-codants.

La transcription antisens est un phénomène décrit pour la première fois chez les phages [207] et les organismes procaryotes [208]. Chez ces derniers, les cis-NAT sont très fréquents et sont impliqués dans le contrôle de l'expression des gènes à l'origine de fonctions biologiques diverses [209]. L'existence de ce phénomène a par la suite été identifiée chez les eucaryotes [210, 211] et depuis, un nombre croissant de publications sur le sujet laissent croire que les NAT sont beaucoup plus répandus que ce qu'on avait d'abord cru.

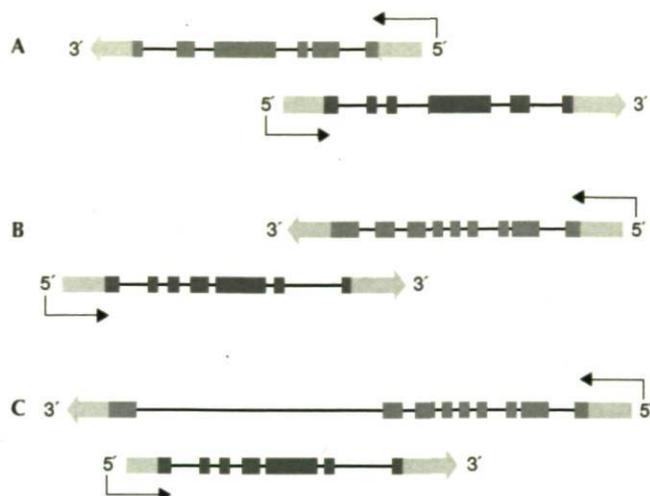


Figure 17: Les différents types d'ARN antisens. (A) Transcrits divergents. (B) Transcrits convergents. (C) Transcrits complètement chevauchants. (tiré de [212])

En effet, selon plusieurs études, la transcription antisens est largement présente dans le génome de nombreuses espèces d'animaux et de plantes, incluant l'homme [213-215], la souris [216-218], le rat, le poulet [215], la drosophile [219] les nématodes [215] le riz [220] *Arabidopsis* [221] et les levures [222]. Chez l'homme, plusieurs études ont rapporté que de 5-29% de toutes les unités transcriptionnelles pourraient être impliquées dans un chevauchement. En utilisant des méthodes de prédiction basées sur la détermination de l'orientation d'un grand nombre d'unités transcriptionnelles (basée notamment sur la présence de signaux de polyadénylation et de queues de polyA), Chen *et al.* ont évalué le nombre d'unités transcriptionnelles formant des paires sens-antisens à près de 22% [213]. La méthode de prédiction développée au cours de cette étude a pu être subséquemment validée par des expériences de RT-PCR brin-spécifique [213]. Ces résultats suggèrent que la transcription antisens est un phénomène bien présent à travers les espèces et que son importance a probablement été sous-estimée.

Il est important de considérer que les analyses effectuées *in silico* sont basées sur une série de paramètres rigoureux qui permettent de déterminer l'orientation d'un transcrit. Ces analyses pourraient donc avoir tendance à sous-évaluer le nombre de transcrits antisens.

D'un autre côté, les analyses par micropuces sont plus sujettes à produire des faux-positifs, qui peuvent résulter de la contamination des échantillons d'ARN, la présence d'hybridation croisée ou d'autre marquage non-spécifique de l'ARN double brin [223, 224]. Une étude récente sur la détection de transcrits antisens par l'utilisation de la méthode *tiling array* (micro-puces) fait état de la présence d'un grand nombre d'artéfacts générant de faux positifs [225]. Ces derniers seraient probablement dus à la synthèse non-spécifique du second brin d'ARN. Bien que le RT-PCR soit une méthode de routine largement employée en biologie moléculaire, elle est souvent appliquée sans l'utilisation des contrôles appropriés pour permettre une détection spécifique d'un brin d'ARN. En effet, dans une réaction conventionnelle de RT-PCR, la spécificité est impossible en raison de l'amorçage endogène non-spécifique durant la phase de transcription inverse, phénomène résultant de la présence inévitable de courts fragments d'ARN ou d'ADN dans l'échantillon de départ. Plusieurs équipes ont d'ailleurs clairement mis en évidence la non-spécificité des méthodes courantes et certaines ont développé divers protocoles qui se sont révélés plus appropriés pour contourner les problèmes associés à l'amorçage endogène. [226-231].

2.2.2 Rôles biologiques des NAT

Bien que quelques groupes aient rapporté la capacité codante de certains NAT [232-234], la grande majorité des transcrits antisens étudiés sont non-codants. En conséquence, les rôles biologiques qui leur sont attribués sont reliés à la régulation transcriptionnelle qui dérive de leurs interactions avec différents transcrits sens. Au moins quatre différents mécanismes ont été proposés (voir figure 18).

1) L'interférence transcriptionnelle : la présence de transcrits chevauchants pourrait bloquer la machinerie transcriptionnelle en raison de la collision des complexes de l'ARN polymérase II. Ce mécanisme a notamment été démontré dans le cas des gènes GAL10 et GAL7 chez la levure, en repositionnant ces deux gènes dans une orientation convergente. [235].

2) Le masquage de l'ARN : la formation d'hydrides ARN-ARN pourrait potentiellement masquer des éléments retrouvés dans l'un ou l'autre des transcrits et donc éventuellement interférer avec des processus qui requièrent des interactions ARN-protéine, tels que l'épissage, le transport, la polyadénylation, la traduction ou la dégradation. Un exemple de ce mécanisme est présenté dans une étude par Hastings *et al.* où l'expression du transcrit antisens au gène *erbA* entraîne une variation de la balance en deux types d'épissage [236].

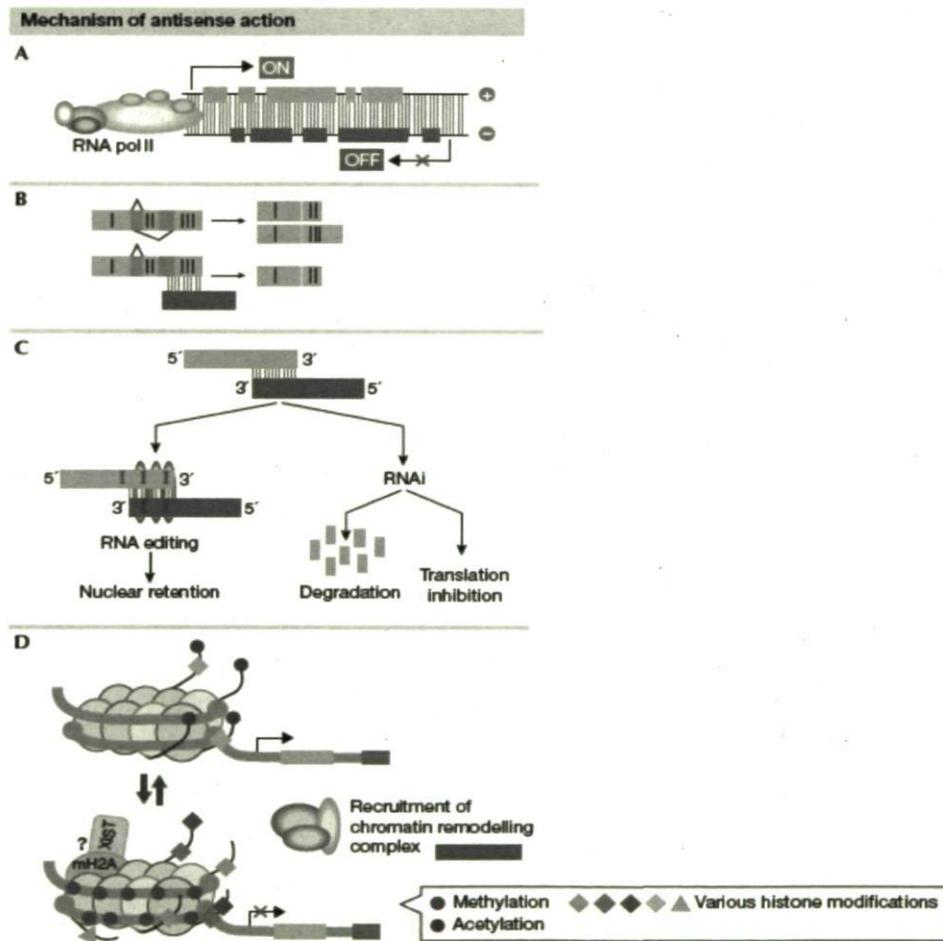


Figure 18 : Les potentiels mécanismes d'action des ARN antisens naturels. (A) L'interférence transcriptionnelle. (B) Le masquage de l'ARN. (C) Mécanismes qui dépendent de la formation d'ARNdb : l'interférence à l'ARN et le *RNA editing*. (D) Méthylation de l'ADN. (tiré de [212])

3) La formation d'ARN double-brin et RNAi : De nombreuses évidences laissent croire que la formation d'hybrides d'ARN pourrait entraîner la mise en marche de mécanismes dépendant de la présence d'ARN double-brin, comme l'interférence à l'ARN ou ARNi ainsi que le *RNA editing*. De tels mécanismes requièrent l'expression simultanée des transcrits sens et antisens, et pourraient donc aller de pair avec l'existence déjà rapportée de la co-expression de plusieurs transcrits sens et antisens [237].

Le *RNA editing* implique la déamination d'adénosines en inosines dans un double-brin d'ARN [238] et pourrait constituer en une stratégie de défense contre l'ARN double-brin. L'hyper-editing de longs duplex d'ARN double-brin peut résulter en leur rétention au noyau ou en leur dégradation au cytoplasme [239, 240].

4) Méthylation et expression mono-allélique : L'ARN double-brin peut induire la méthylation de l'ADN et interférer avec l'expression d'un gène. Par exemple, une forme de thalassémie (chez l'homme) résulte d'une méthylation du gène de l'hémoglobine- β causée par l'expression d'un ARN antisens. [241]. Les ARN antisens peuvent être impliqués dans l'expression mono-allélique, comme par exemple dans l'inactivation sélective du chromosome X [242].

2.3 La transcription antisens chez les rétrovirus

L'existence de la transcription antisens chez les rétrovirus a été suggérée pour la première fois en 1988, dans une étude qui faisait état de la présence d'un ORF conservé de 190 acides aminés dans la séquence complémentaire au gène *env* du VIH-1 [243]. L'analyse de 12 isolats du virus démontrait que la séquence en acides aminés était elle aussi bien conservée, particulièrement en son extrémité N-terminale où 83 des 100 premiers acides aminés sont identiques dans chacun des isolats. La séquence prédite est riche en cystéine, leucine, proline et sérine et présente un profil très hydrophobe, ce qui suggère une association potentielle aux membranes. De façon intéressante, cet ORF est situé au niveau

du site de clivage entre la gp120 et la gp41, une région bien conservée du gène *env*, pourtant très variable dans le reste de sa séquence.

Toujours dans le même article, l'auteur mentionne qu'un ORF antisens est aussi présent chez plusieurs autres rétrovirus et hepadnavirus qui expriment une protéine impliquée dans la transactivation de l'expression des gènes viraux, dont le virus de l'hépatite B, les rétrovirus humains HTLV-1, HTLV-2, VIH-2 ainsi que les rétrovirus Visna, EIAV, et VIS. Selon les analyses effectuées, il apparaît que l'ORF antisens des ancêtres de ces virus était beaucoup plus long et qu'il aurait tendance à raccourcir en raison de l'accumulation de codons stop.

2.3.1 *Transcription antisens chez HTLV-1 et découverte de la protéine HBZ*

Les premières évidences de la présence de transcription antisens chez un rétrovirus furent présentées l'année suivante, alors qu'une étude démontrait la présence d'ARNm de polarité négative chez des cellules chroniquement infectées par le virus HTLV-1 [244]. Dans cet article, les auteurs observent la présence d'un transcrite majeur de 2,5 kb en utilisant des sondes d'ARN situées entre les nucléotides 5159 et 7020 de l'ADN proviral, alors que 2 sondes situées dans le LTR 3' ne permettent pas sa détection. Le transcrite contient un ORF d'environ 200 acides aminés dont la séquence présente une région riche en arginine. Cet ORF se retrouve entre la séquence codante du gène d'enveloppe et de la région pX (voir figure 19). Ces résultats suggéraient donc la possibilité que ce transcrite puisse coder pour une nouvelle protéine virale.

Cette hypothèse n'a été confirmée que de nombreuses années plus tard, alors que Gaudray *et al.* [245] découvraient une nouvelle protéine virale dans une librairie d'ADNc construite à partir des cellules MT2, infectées par HTLV-1. Par la technique du double hybride, les auteurs ont pu identifier l'ADNc de cette protéine dans une série de clones sélectionnés en fonction de leur capacité à interagir avec la région *leucine zipper* (bZIP) du facteur de transcription CREB-2. La séquence en acides aminés de cette protéine présente plusieurs

caractéristiques d'un facteur transcriptionnel de type bZIP et sa capacité à former des complexes avec CREB-2 *in vivo* est confirmée par immunoprécipitation. Conséquemment, les auteurs proposent de nommer cette protéine HBZ (*HTLV-1 bZIP factor*). À l'aide d'anticorps anti-HBZ, un signal spécifique de 31 kDa est détecté à partir d'extraits de cellules C8166-45, confirmant donc l'expression de la protéine *in vivo*.

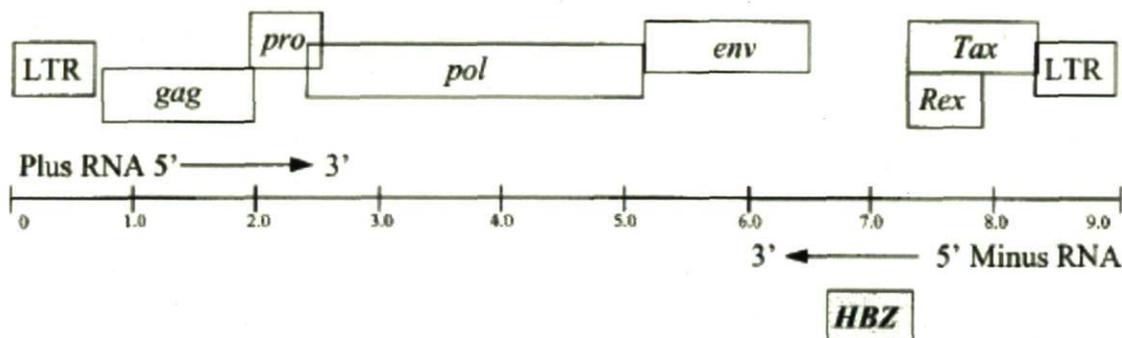


Figure 19 : Position de l'ORF de la protéine HBZ dans le génome d'HTLV-1. La région codante de la protéine HBZ est située entre les ORF de la protéine d'enveloppe (*env*) et des protéines Tax et Rex. (tiré de [245])

Dans cette même étude, des observations intéressantes sont faites quant à la fonction d'HBZ : sa capacité à former des hétérodimères avec CREB-2 entraîne une diminution importante de la transcription virale induite par Tax. Cet effet résulterait de l'incapacité des hétérodimères CREB-2/HBZ à se lier aux éléments TRE1 situés dans le promoteur viral. D'un autre côté, la présence d'un domaine d'activation en N-terminal de la séquence d'HBZ laisse croire qu'il ne s'agit probablement pas de l'unique fonction de cette protéine.

2.3.2 Caractérisation du transcrit antisens chez HTLV-1

À la lumière de ces résultats présentés par Gaudray *et al.*, d'autres groupes s'intéressent ensuite à la structure du transcrit d'HBZ. La présence d'ARN antisens est détectée dans plusieurs lignées infectées par HTLV-1 et chez des cellules dérivées de patients infectés.

De façon surprenante, les résultats présentés prouvent que la transcription antisens est initiée dans le LTR 3' et que le transcrit est épissé de façon alternative [246-248]. Plusieurs sites d'initiation de la transcription sont identifiés dans les régions R et U5 du LTR 3'. L'épissage majeur (SP1) permet le recrutement d'un nouveau codon initiateur ATG dans l'exon 1 du transcrit, ce qui entraîne un changement des premiers acides aminés à l'extrémité N-terminale de la protéine par rapport à la séquence originale identifiée à partir de l'ADN proviral (voir figure 20). Cavanagh *et al.* proposent également l'existence d'un second type d'épissage (SP2) à partir d'un site donneur différent [246]. Cet épissage semble cependant beaucoup moins fréquent que SP1 et ne permet pas de remplacer le codon initiateur original, qui se retrouve dans la séquence de l'intron. Par contre, les auteurs n'excluent pas la possibilité que la protéine HBZ puisse être produite à partir de cet épissage *via* l'utilisation d'un codon initiateur non-ATG, un mécanisme déjà décrit pour la traduction de certaines protéines chez HTLV-1 [249].

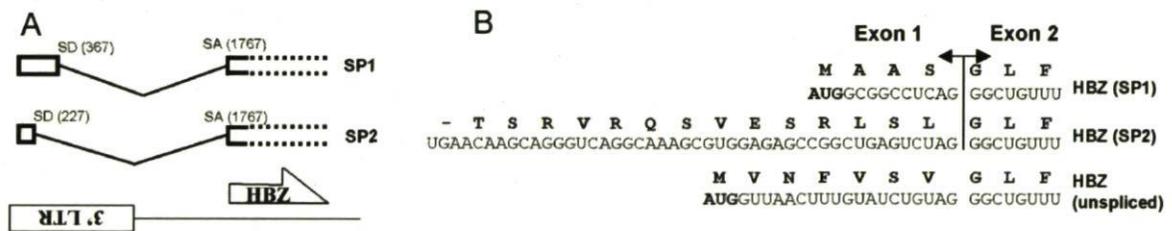


Figure 20 : Épissage alternatif du transcrit antisens chez HTLV-1. (A) Présentation des deux types d'épissage identifiés par Cavanagh *et al.* et position des sites donneurs (SD) et accepteurs (SA). **(B)** Comparaison des séquences nucléotidiques des premiers et deuxièmes exons du transcrit antisens chez HTLV-1 par rapport à la séquence retrouvée dans l'ADN proviral (non-épissé). La figure met en évidence le changement de séquence en acides aminés de la protéine HBZ entraîné par l'épissage majoritaire (SP1). À noter que l'épissage SP2 ne permet pas de recruter un codon initiateur ATG. (adapté de [246])

Fait intéressant, l'importance attribuée à l'épissage du transcrit antisens pour la production de la protéine HBZ varie selon les études. En effet, une première étude suggère l'existence d'un transcrit non-épissé sans pourtant présenter de sites d'initiation de la transcription

pour ce dernier [248]. Toujours selon cette étude, des analyses de RT-PCR en temps réel permettent d'affirmer que la quantité de transcrits non-épissés est environ équivalente à celle des transcrits épissés SP1, ce qui porte à croire que deux isoformes d'HBZ pourraient être exprimés *in vivo*. Cependant, des analyses par immunobuvardage effectuées sur des extraits de cellules ATL démontrent que l'isoforme SP1 d'HBZ semble être exprimé préférentiellement dans ces cellules. Ces résultats sont corroborés par les résultats présentés par Cavanagh *et al.* qui montrent que l'épissage joue un rôle majeur pour permettre la traduction de la protéine, puisque la mutation du codon ATG présent dans l'intron ne semble pas affecter le niveau d'expression d'HBZ observé suite à la transfection d'un ADN proviral. Ces observations sont éventuellement confirmées par une étude plus récente qui rapporte que le transcrit épissé d'HBZ est 4 fois plus présent que le non-épissé dans les cellules de patients infectés [250].

Ces études permettent également l'identification d'un signal consensus de polyadénylation résultant en la production d'un transcrit majeur de 2,4 kb. Ces résultats confirment donc les observations initiales de Larocca *et al.* [244].

2.3.3 Rôle de la protéine HBZ dans la réplication virale.

Ces premiers résultats concernant la transcription antisens chez HTLV-1 s'avèrent efficaces à convaincre la communauté scientifique : de nombreux groupes de recherche se sont intéressés par la suite à HBZ et plusieurs études ont été menées dans le but de mieux comprendre son rôle dans la réplication virale. Ces efforts ont permis de mettre en évidence la capacité d'HBZ à dimériser avec d'autres protéines cellulaires telles que JunB, c-Jun et JunD et à moduler leur activité transcriptionnelle [251, 252]. Il en ressort également que la capacité d'HBZ à inhiber l'action de Tax ne dépend pas uniquement de son interaction avec CREB-2, mais aussi de son association avec CREB [253] et les co-activateurs CBP et p300 [254]. Cette dernière association est possible grâce à la présence de motifs LXXLL présents dans la portion N-terminale d'HBZ et interfère avec la liaison

de Tax avec ces mêmes co-activateurs. Ces résultats démontrent donc que la protéine HBZ utilise plusieurs mécanismes pour réguler négativement l'action de Tax. (la figure 21 résume les rôles connus de la protéine HBZ dans une cellule infectée par HTLV-1).

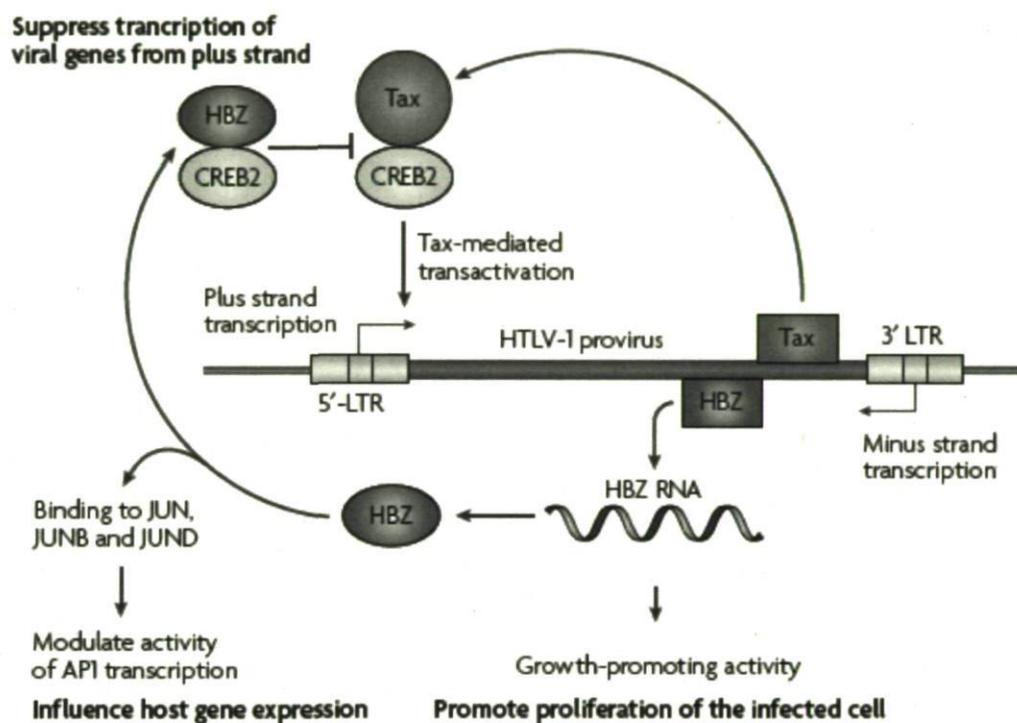


Figure 21 : Action de l'ARNm et de la protéine HBZ dans une cellule infectée par HTLV-1. La protéine HBZ régule négativement la transcription virale en formant des hétérodimères avec les facteurs transcriptionnels CREB et CREB-2. Sa capacité à dimériser avec les membres de la famille AP1 entraîne également une modification de l'expression de gènes cellulaires. L'ARNm d'HBZ pourrait aussi avoir un effet positif sur la prolifération cellulaire. (adapté de [52])

Parallèlement, d'autres groupes s'intéressent à la fonction d'HBZ dans un contexte de réplication *in vivo*. Alors que des essais de prolifération dans les cellules du sang périphérique établissent que le gène d'HBZ n'est pas nécessaire à la réplication virale ou à l'immortalisation induite par le virus, il semble par ailleurs qu'il joue un rôle important dans l'infectivité et la persistance du virus chez le lapin [255]. HBZ pourrait donc réguler la

transcription virale et cellulaire de façon à permettre la survie cellulaire et ultimement la persistance de l'infection. Ces résultats concordent avec ceux d'un autre groupe, qui démontrent que l'inhibition de l'expression d'HBZ par l'utilisation de shRNA diminue la prolifération cellulaire de façon importante [247]. Selon les conclusions de cette dernière étude, il est également possible que le transcrit lui-même soit impliqué dans l'effet prolifératif observé.

Plus récemment, un article venait confirmer l'effet prolifératif d'HBZ dans les lymphocytes et suggérer un rôle de la protéine dans le processus tumorigénique [256]. Cette information est intéressante, compte tenu du fait que contrairement à Tax, HBZ semble être exprimé dans toutes les cellules ATL [248, 250]. Il semble donc que la compréhension des mécanismes menant à la transformation cellulaire par HTLV-1 nécessite des études plus approfondies sur HBZ et son impact sur la régulation de la transcription.

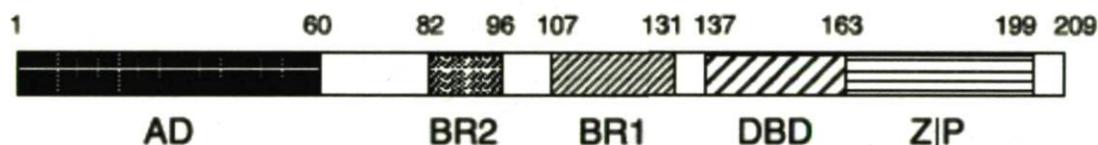


Figure 22 : Les différents domaines de la protéine HBZ. Les domaines basiques BR1 et BR2 ainsi que le domaine de liaison à l'ADN (DBD) sont essentiels à la localisation de la protéine dans les corps nucléaires. ZIP : domaine *leucine zipper* AD : domaine d'activation. (tiré de [257])

2.3.4 Structure de la protéine HBZ

Plusieurs études menées sur la protéine HBZ ont permis de caractériser ses domaines fonctionnels (les différents domaines de la protéine sont présentés à la figure 22). Les expériences initiales effectuées sur la protéine HBZ originale (produite par l'ADNc non-épissé) avaient permis de constater la localisation nucléaire de la protéine [245]. Ces résultats n'étaient pas surprenants, considérant la présence d'un motif *leucine zipper* dans la séquence en acides aminés de la protéine et sa capacité à inhiber l'action de Tax. De façon

intéressante, la distribution de la protéine au noyau est irrégulière et de type granulaire. Des analyses plus poussées ont par la suite permis d'identifier 3 motifs essentiels à la localisation de la protéine : 2 régions riches en acides aminés basiques et une troisième région correspondant au domaine de liaison à l'ADN. Lorsque ces 3 motifs sont présents, la protéine se localise à l'intérieur de corps nucléaires. Des expériences de microscopie électronique ont par la suite permis de constater que la protéine semblait se localiser préférentiellement au niveau de zones correspondant à l'hétérochromatine [257]. Des analyses subséquentes ont permis de démontrer que la localisation de l'isoforme d'HBZ épissé (SP1) semblait différer de celle de la protéine HBZ originale. En effet, en plus de former des structures granulaires, HBZ-SP1 présentent des grandes taches au noyau qui co-localisent avec les nucléoles [248, 258].

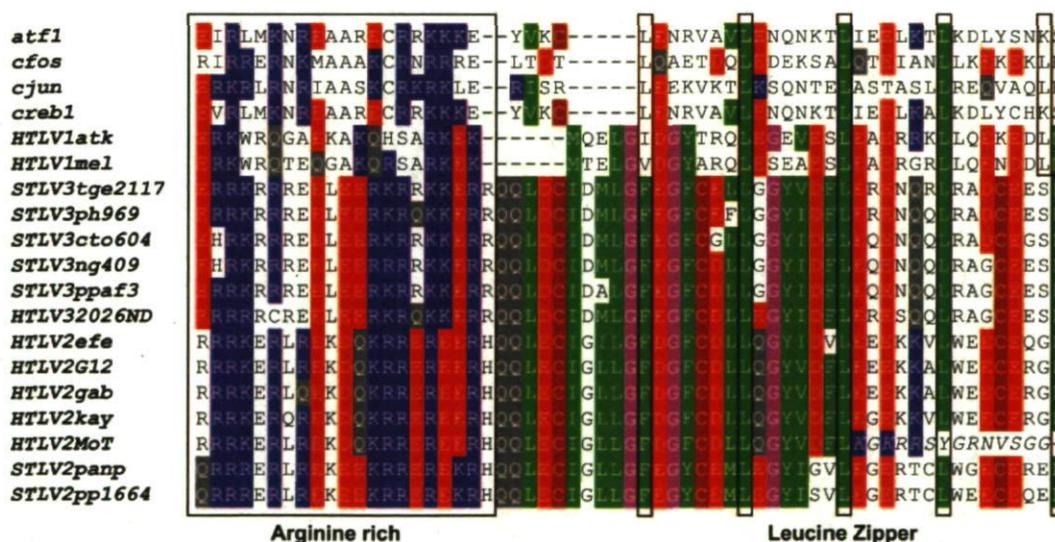


Figure 23 : Comparaison de la séquence en acides aminés de plusieurs facteurs bZIP codée par différents membres des PTLVs. Les domaines *leucine zipper* et les domaines riches en arginine sont encadrés. Le clone HTLV2MoT présente une délétion de 1 nucléotide en position 6823 qui entraîne un changement de cadre de lecture (indiqué en italique). Le motif *leucine zipper* n'est pas conservé dans ce clone. (tiré de [259])

2.3.5 Transcription antisens chez HTLV-2 et HTLV-3

Suite à la découverte du nouveau rétrovirus humain HTLV-3, certaines études ont permis d'établir la séquence complète du virus [260, 261] et ainsi que celle de son homologue simien STLV-3 [262]. Ces analyses ont pu mettre en évidence l'existence d'un ORF antisens dans le génome des virus HTLV-2 et HTLV-3. L'analyse de la séquence en acides aminés démontre que le domaine *leucine zipper* ainsi que la région riche en arginine qui la précède sont conservées chez l'ensemble des représentants de la famille des PTLVs (figure 23).

2.4 La transcription antisens chez le VIH-1

Plusieurs années après avoir été suggérée par Miller [243], la présence de transcription antisens chez le VIH-1 ne fait toujours pas l'objet de consensus. Peu d'équipes se sont intéressées au sujet et les résultats présentés jusqu'à maintenant n'ont vraisemblablement pas convaincu la communauté scientifique de l'importance ou même de l'existence du phénomène.

Peu de temps après la publication de l'article de Miller, un groupe rapportait la détection inattendue de signaux en *Northern blot*, obtenus en hybridant de l'ARN de cellules infectées à l'aide d'une sonde contrôle correspondant au brin positif du VIH-1 [263]. Selon ces résultats, la transcription antisens est présente au cours des premiers stades de la réplication virale et devient indétectable plus tard dans l'infection. Quelques années plus tard, une nouvelle étude plus complète du phénomène paraissait [264]. Cette fois, les transcrits antisens sont détectés par un protocole classique de RT-PCR. À l'aide d'une librairie d'ADNc construite à partir des cellules A3.01 chroniquement infectées par le VIH-1, les auteurs identifient 3 clones correspondant au transcrit antisens. Selon la séquence de ces clones, les sites d'initiation de la transcription se situeraient dans la région R du LTR 3'. Un seul de ces clones contient une queue de poly-A, retrouvée 5 nucléotides en aval du

codon stop de l'ORF antisens décrit par Miller. Malgré l'absence de signal consensus de polyadénylation, les auteurs notent la présence de 2 signaux alternatifs ATTAAA dans la séquence codante de l'ORF antisens (la position de l'ORF d'ASP et du transcript antisens est présentée à la figure 24). De plus, l'analyse du promoteur antisens à l'aide de plusieurs constructions contenant le gène rapporteur CAT permet de conclure que les facteurs de transcription NF- κ B et USF semblent jouer un rôle important dans la régulation positive de la transcription antisens, alors que la protéine Tat pourrait jouer un rôle de régulation négative. Une étude d'un autre groupe attribue également une fonction cruciale à un nouveau site Sp1 situé dans la région U3 du LTR [265]. Le site d'initiation de la transcription est cette fois situé 9 nucléotides en aval de ce site Sp1.

Quelques études ont aussi été menées dans l'objectif de détecter une protéine produite par un potentiel transcrit antisens. Un premier travail, publié en 1994, présente l'efficacité des anticorps de plusieurs patients infectés à immunoprécipiter la protéine antisens synthétisée *in vitro* dans un système de lysats de réticulocytes [266]. Un signal est détecté à 19 kDa, un poids moléculaire qui correspond bien à celui prédit à partir de la séquence en acides aminés de l'ORF contenu dans le transcrit antisens. La protéine est alors nommée ASP (*antisense protein*). Alors que la protéine est reconnue par différents anti-peptides dérivés de régions hydrophiles de la protéine, des expériences de *Western blot* et d'immunoprécipitation s'avèrent inefficaces à détecter ASP dans les cellules infectées ou transfectées. Une fois de plus, la présence de transcription antisens est détectée par des méthodes de RT-PCR classique dans plusieurs lignées chroniquement infectées. Les auteurs concluent que le niveau de transcription antisens est relativement plus faible que celui observé pour la transcription sens, compte tenu notamment de l'absence de signal en *Northern blot*.

En 2002, le même groupe présente de nouveaux résultats quant à la détection et la localisation de la protéine ASP [267]. Cette fois, une protéine de fusion GST-ASP est produite en bactérie et est reconnue efficacement par un sérum anti-ASP. Ce sérum est ensuite utilisé pour déterminer la localisation cellulaire de la protéine dans plusieurs lignées infectées ou surexprimant la protéine. Par microscopie électronique, les auteurs observent

la présence de signaux positifs dans plusieurs compartiments cellulaires, particulièrement au niveau des membranes. De plus, la technique permet de détecter la présence de la protéine à la surface des nouveaux virions, suggérant donc que cette nouvelle protéine pourrait être incorporée dans la particule virale au moment du bourgeonnement.

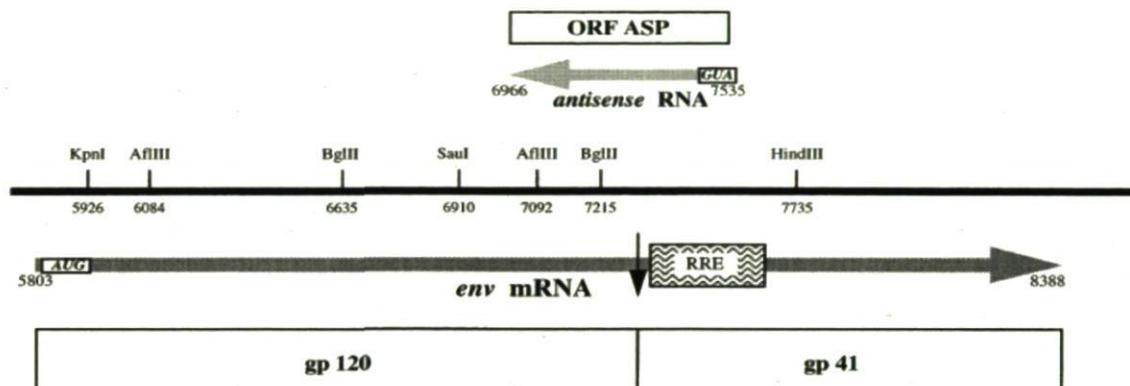


Figure 24 : Position de l'ORF de la protéine ASP dans le génome du VIH-1. (tiré de [268])

Dans le cadre d'une étude plus récente, le LTR du VIH fut étudié de façon plus approfondie pour sa capacité à agir en tant que promoteur du transcrit antisens [172]. Cette étude présente une analyse de promoteur basée sur une série de constructions contenant le gène rapporteur CAT. Les auteurs constatent que l'activité du promoteur antisens est plus faible que celle du sens et qu'elle est régulée positivement par les facteurs de transcription NF- κ B, LEF-1 et Ras. Ces expériences permettent aussi d'attribuer à Tat un rôle de régulation négative, ce qui corrobore les résultats obtenus par Michael *et al.* [264].

Suite à la découverte de la protéine HBZ chez HTLV-1, un nouveau groupe de recherche présente une étude de la transcription antisens chez le VIH-1 [269]. Les auteurs rapportent l'identification d'un élément initiateur (Inr) présentant des similarités avec l'élément initiateur de la *terminal deoxynucleodyl transferase* (TdT). L'orientation de cet élément prédit un rôle susceptible de favoriser l'initiation de la transcription antisens à partir du

LTR. Des expériences d'extension d'amorce dans un modèle breveté de transcription *in vitro* impliquant l'utilisation d'extraits nucléaires de drosophiles permettent de conclure à la fonctionnalité de ce nouvel élément. Contrairement aux précédentes études, les auteurs identifient plusieurs ORF potentiels contenus dans le transcrit antisens du VIH-1. En se basant sur le fait que le virus utilise différentes stratégies pour produire les protéines virales, les auteurs émettent l'hypothèse selon laquelle plusieurs de ces ORF pourraient être traduits *via* l'utilisation de mécanismes spécifiques du virus, tel que le changement de cadre de lecture. En utilisant de nombreuses stratégies, les auteurs identifient, dans les extraits de cellules de patients infectés, plusieurs peptides dérivés de ces ORF. À noter que dans ce même article, il est de plus suggéré que l'élément initiateur identifié pourrait former une boucle d'ARN qui résulterait en la production de micro-ARN.

2.5 La transcription antisens chez le VIF

Dans une étude publiée en 2001, Briquet *et al.* rapportent la présence d'un ORF antisens conservé dans le génome du virus de l'immunodéficience féline (VIF) [270]. Tout comme chez le VIH-1, cet ORF est situé dans le brin complémentaire au gène de l'enveloppe. Cet ORF pourrait coder pour une protéine de 103 acides aminés prédite comme étant fortement hydrophobe. Les résultats de cette étude démontrent que la présence de transcrits antisens chez VIF peut être détectée par RT-PCR lors d'expériences d'infection effectuées *in vivo* et *ex vivo*.

Chapitre 3 : Hypothèses et objectifs.

3.1 Hypothèses de recherche

Au moment de démarrer ce projet dans nos laboratoires, la transcription antisens chez les rétrovirus était toujours un sujet relativement obscur, particulièrement chez le VIH-1. Par contre, les résultats fraîchement obtenus par notre groupe [246] ainsi que par des collaborateurs [245] allaient rapidement conduire à la renaissance de l'intérêt pour le phénomène, particulièrement pour le virus HTLV-1. En effet, l'identification de la protéine HBZ et la mise en évidence de la présence d'épissage alternatif dans le transcrit antisens chez HTLV-1 laissaient peu de doute quant à l'existence et la pertinence biologique de la transcription antisens chez ce rétrovirus.

À la lumière de ces résultats, il semblait logique de croire que la transcription antisens pourrait être une stratégie utilisée par d'autres rétrovirus qu'HTLV-1. Le cas du VIH-1 paraissait particulièrement intéressant, considérant la présence déjà rapportée d'un ORF conservé dans le brin complémentaire du virus. De plus, l'absence de similitude entre l'ORF de la protéine HBZ et l'ORF antisens du VIH-1 laissait croire que ces protéines antisens pourraient jouer des rôles différents dans la réplication virale.

Sur la base de ces prémisses, nous nous sommes donnés différents objectifs spécifiques qui visaient une meilleure compréhension du phénomène de la transcription antisens chez les rétrovirus :

- I. Étudier la régulation de l'expression du transcrit antisens chez HTLV-1 et évaluer l'impact de l'expression de la protéine Tax sur le promoteur antisens
- II. Vérifier l'existence de la transcription antisens chez d'autres virus de la famille HTLV et évaluer la capacité de ces transcrits à coder pour de nouvelles protéines virales.

- III. Confirmer la présence de transcription antisens chez le VIH-1 par l'application de méthodes permettant l'amplification ou la détection spécifique d'un ARN antisens, caractériser ces transcrits et évaluer leur capacité à produire la protéine virale ASP.

Chapitre 4 : Régulation positive de la transcription antisens chez HTLV-1 par la protéine virale Tax.

Ce chapitre a fait l'objet d'une publication dans la revue *The Journal of Virology* (doi:10.1128/JVI.01264-08) et est constitué de la version finale de l'article tel que publié.

4.1 Résumé

Plusieurs études ont récemment démontré l'existence de transcription antisens chez HTLV-1. Ce transcrit permet la synthèse de la nouvelle protéine virale HBZ. Alors que de nombreuses études ont été entreprises dans le but de mieux comprendre le rôle de cette protéine dans le cycle de réplication virale, peu de choses sont connues quant à la régulation de l'expression de ce nouveau gène. Dans cette étude, nous avons utilisé notre construction luciférase K30-3'asLuc pour démontrer que la protéine Tax régule positivement l'expression de l'ARNm d'HBZ. Cette action est dépendante de la présence des éléments TRE1 situés dans le LTR 3'. La génération de lignées stables a permis de mettre en évidence que la réponse à Tax est fortement influencée par le site d'intégration dans le génome de la cellule. L'environnement d'ADN cellulaire pourrait donc affecter le niveau d'expression de la protéine HBZ dans les cellules infectées.

4.2 Article

Upregulation of HTLV-1 antisense transcription by the viral Tax protein

Sébastien Landry¹, Marilène Halin¹, Amandine Vargas¹, Isabelle Lemasson², Jean-Michel Mesnard³ and Benoit Barbeau^{*1}

Université du Québec à Montréal, Département des sciences biologiques, Montréal (Québec) Canada, H2X 3X8¹, East Carolina University, Department of Microbiology and Immunology, Greenville, North Carolina 27834², Centre d'études d'agents Pathogènes et Biotechnologies pour la Santé (CPBS), CNRS/UMI/UM2 UMR 5236, Montpellier, France

3

Running title : Tax activates HTLV-1 antisense transcription

*Corresponding author:

Mailing address:

Université du Québec à Montréal, Département des Sciences Biologiques

Case postale 8888, succursale Centre-ville

Montréal (Québec) H3C 3P8, CANADA

Phone: (514) 987-3000, ext. 4576. Fax: (514) 987-4647

E-mail: barbeau.benoit@uqam.ca

Abstract word count: 116; Text word count : 2274

ABSTRACT

Several studies have recently demonstrated the existence of HTLV-1 antisense transcripts, which allow the synthesis of the newly described HBZ protein. Although previous reports have been aimed at understanding the potential role of the HBZ protein in HTLV-1 pathogenesis, little is known as to how this viral gene is regulated. Here, using our K30-3'asLuc reporter construct, we show that the viral Tax protein upregulates antisense transcription through its action on the TRE sequences located in the 3'LTR. Generation of stable clones in 293T cells demonstrated that Tax-induced HBZ expression is importantly influenced by the integration site in the host genome. The cellular DNA context could thus affect the level of HBZ mRNA expression in infected cells.

HTLV-1 is the etiological agent of Adult T cell Leukemia/Lymphoma (ATLL) and HTLV-1-Associated Myelopathy also known as Tropical Spastic Paraparesis (HAM/TSP) (13, 29, 34, 35, 46), leading to disease development in less than 5% of infected individuals (42). The HTLV-1 Tax protein is generally considered to play a central role in the viral pathogenesis, although its precise role remains to be determined. Tax transactivates the viral promoter, mainly through three Tax Responsive Element (TRE) 1 repeats located in the U3 region of the 5' Long Terminal Repeat (LTR) (14, 15, 17, 19, 44). Each TRE1 contains an imperfect cyclic AMP response element (CRE) and they are crucial for both basal and Tax-induced promoter activity (14, 15, 17, 19, 44). Tax associates with TRE1-bound ATF/CREB transcription factors and recruits the transcriptional co-activator CBP/p300, which results in a strong transcriptional activation (12, 22). Tax also upregulates the expression of cellular genes through the activation of several transcription factors such as CREB-1, NF- κ B and SRF (1, 30, 45).

Recently, we (6) and others (31, 39) have clearly demonstrated the existence of antisense transcripts in HTLV-1, which initiate from the 3'LTR. These transcripts are alternatively spliced and encode for the newly described HTLV-1 bZIP factor (HBZ) through an ATG initiation codon being present in exon 1 located in the 3' LTR. Interestingly, HBZ downregulates the Tax-dependent transactivation of viral gene expression through the formation of heterodimers with CREB-1 and CREB-2. In addition, HBZ interacts with c-Jun, JunB, and JunD and modulates their transcriptional activity (4, 24, 41). Recent studies have shown that HBZ is expressed in all ATL cell lines, suggesting that it might play an unsuspected role in the biological process leading to ATLL development (28, 43).

Despite the increasing number of studies aimed at understanding the role of HBZ in viral pathogenesis, little is known about the promoter region regulating its expression. In order to clarify this issue, we have performed several transfection experiments using our previously described K30-3'asLuc construct (6). In this construct, a cassette containing the luciferase reporter gene and the SV40 polyA signal was cloned in the antisense orientation in a 5' LTR-deleted version of the full length HTLV-1 K30 proviral DNA. More precisely, the luciferase reporter gene lies downstream of the 5' end of HBZ exon 2 while remaining in frame of the HBZ ORF. Jurkat cells were transiently transfected with K30-3'asLuc by electroporation (250V, 950 μ F) and stimulated or left untreated for 8 hours with different T-cell activators. At 16 hour post-transfection, luciferase activity was measured as previously described (20). The antisense promoter was shown to be inducible in stimulated Jurkat cells by PHA and PMA, but more importantly by a combination of the protein tyrosine phosphatase inhibitor bpV[pic] and forskolin (figure 1A) (20). Since bpV[pic]/forskolin have been previously shown to upregulate HTLV-1 sense gene expression by activating CREB transcription factors (32, 36), these results suggested a possible role for this transcription factor in antisense promoter activity. We next transfected various HTLV-1-infected T-cell lines with K30-3'asLuc along with the β -galactosidase-expressing plasmid pRc-actin-LacZ using the same electroporation conditions described above. The β -galactosidase activity was measured using the Galacto-Light™ kit (Applied Biosystems, Bedford, MS). Interestingly, results suggested that antisense promoter activity was higher in HTLV-1-infected cells when compared to non-infected Jurkat cells, with the exception of the Tax non-expressing S1T cell line (23). A strong luciferase signal was obtained for virions-producing cell lines MT-2 and C91/PL, while MJ cells, which produce less viral

particles (unpublished data), showed lower luciferase activity. In addition, non-HTLV-1-producing Tax-expressing C8166-45 cells (5) showed a comparable signal to the one obtained with the MT2 and C91/PL cells. (figure 1B). Hence, these results suggested a possible implication of the Tax protein on the regulation of antisense promoter activity. To test this potential implication of Tax, levels of Tax protein was evaluated. Cell extract preparation and Western blot analysis was conducted as previously described (6). Using our anti-Tax antibody (dilution 1/100) (2) along with an HRP-coupled anti-mouse IgG (1/10000) (Santa Cruz Biotechnology Inc., Santa Cruz CA), Tax was detected in C8166-45, C91/PL and MT-2 cells. In MT-2 cells, a 68 kDa protein was detected and corresponded to a previously identified Tax-Env fusion protein, which has been suggested to be also functionally active (7, 18). No specific band could be detected for the MJ, S1T and Jurkat cell lines. Equal loading of extracts were confirmed with GAPDH using a mouse anti-GAPDH antibody (Santa Cruz Biotechnology Inc.) (1/1000). The absence of a Tax-specific band in MJ cells was most likely due to low expression level since RT-PCR confirmed the presence of Tax mRNA in this cell line (data not shown).

As the presence of Tax correlated with high levels of antisense transcription-driven luciferase activity, we then performed several co-transfection experiments in 293T and Jurkat cells (as outlined in (6)). In both cell types, co-transfection of a Tax expression vector, pTax (25) and K30-3'asLuc resulted in a statistically significant increase in luciferase activity compared to control (figure 2A). Because LTR-driven transcription is differently regulated in a plasmidic context in comparison to the chromatin context (27, 33), the effect of Tax expression on antisense promoter activity was further evaluated in an integrated proviral DNA. As initial RT-PCR experiments indicated the presence of HBZ

spliced mRNA in Tax non-producing S1T cells, S1T cells were thus electroporated (see conditions above) with a Tax-expressing vector (vs. pH β Pr.1neo) and analyzed for the presence of the major HBZ spliced transcript by RT-PCR. PCR amplifications were conducted on oligodT-primed cDNAs with a forward primer spanning the HBZ mRNA splice junction (5'-ATGGCGGCCTCAGGGCTGT-3') and a reverse primer located in the HBZ coding region (5'-TGGAGGGCCCCGTCGCAG-3') and with β -actin-specific primers (5'-CGTGACATTAAGGAGAAGCTG-3' and 5'-CTCAGGAGGAGCAATGATCTT-3'). PCR conditions were as follow: a first step of denaturation at 95°C for 10 seconds, followed by 35 (standard) or 50 (real-time) cycles of denaturation (94°C for 3 seconds), annealing (60°C for 15 seconds) and elongation (72°C for 12 seconds). Real-time RT-PCR experiments were also performed with these conditions using the SYBR^R Premix Ex TaqTM Perfect Real Time kit. The HPRT-1 gene served as the internal control with the following forward and reverse primers: 5'-AAGCTTGCGACCTTGACC-3' and 5'-GACCAGTCAACAGGGGACATAA-3'. Real time and standard semi-quantitative RT-PCR experiments both showed a significant increase in HBZ mRNA level in the Tax-transfected S1T cells, thus confirming the results obtained with our luciferase reporter construct (figure 2 B,C).

Since electroporation of S1T cells resulted in about 5-10% of transfected cells (data not shown), and might account for the low level of induction, 293T cells were next stably transfected with K30-3'asLuc and pCMV-Hyg in a 10:1 ratio using Lipofectamine 2000 (Invitrogen). Individual clones were selected in the presence of 200 μ g/mL Hygromycin B and then tested for luciferase activity. As shown in figure 3 (panel A), we first noted important variations in basal luciferase activities in the different clones (from 1 to 5564

RLU), suggesting that the integration sites importantly influenced the basal activity of the HBZ promoter. Transfection of the 293T cell clones with pTax resulted in an important induction of luciferase activity, reaching up to 2000 fold for certain clones (figure 3B). The observed Tax-inductions were generally inversely correlated with basal luciferase activity levels of the different clones.

Assessment of integrated plasmidic DNA copy number was then conducted by real-time PCR on 100ng of genomic DNA of the different clones using luciferase-specific forward (5'-TGTTGTTCCATTCCATCACG-3') and reverse primers (5'-TGGCGAAGAAGGAGAATAGG-3') (figure 3C). Each DNA sample was normalized with the HPRT-1 control gene amplification. A standard curve was similarly amplified in parallel and consisted of $4,7 \times 10^3$ to $4,7 \times 10^7$ copies of pGL3basic DNA diluted in 100 ng of non-transfected 293T genomic DNA. Results showed variations in the normalized copy number amongst the different clones ranging from 3200 to 14300 copies/ng of genomic DNA, but these variations did not correlate with basal or Tax-induced levels of luciferase activity, indicating that the number of integrated plasmids was not responsible for the observed differences in luciferase levels. In parallel, Western blot analyses were carried out using mouse anti-CREB-1 and anti-CREB-2 mouse antibodies (1/500; Santa Cruz Biotechnology Inc.). Results showed weak variations in the expression of these transcription factors in the different clones (figure 3D).

To better understand this Tax-dependent upregulation of HBZ gene expression, two representative clones (8 and 21) were transfected with vectors expressing various Tax mutants (11, 38). When compared to the Tax-mediated increase in luciferase gene expression, Tax M47 and Tax K88A (deficient for CREB-dependent transcriptional

activation and CBP/p300 recruitment, respectively) resulted in a more limited increase in the activation of luciferase activity. In contrast, NF- κ B activation-defective Tax (Tax M22) or a PDZ binding-motif-mutated Tax (Tax Δ PBM) led to levels of induction comparable to wild-type Tax-transfected cells (figure 3E). Detection of Tax protein levels using our Tax antiserum (1/1000; NIH AIDS Research and Reference Reagent Program, Germantown MD) and HRP-conjugated anti-rabbit IgG (1/10000, Santa Cruz Biotechnology Inc.) demonstrated that the low inductions observed for the M47 and K88A mutants could not be attributed to a lower protein expression (figure 3F). These results thereby suggested that Tax induction of the antisense promoter requires the participation of CREB.

In order to confirm that CREB can directly modulate HBZ promoter activity, expression vectors for wild-type CREB, a constitutively active CREB mutant (CREB-FY) (9) and CREB-2 (10) were transfected in stable 293T cell clones 8 and 21 or pooled stably transfected 293T cells. Results demonstrated that both wild-type CREB and CREB-FY could potently increase luciferase activity, albeit at a lower level than the one observed when Tax is expressed. On the other hand, CREB-2 repressed luciferase gene expression to levels lower than the basal activity (figure 4A). The negative effect of CREB-2 on the antisense promoter is not surprising, as this transcriptional factor has been reported to repress the expression of cellular genes containing CRE elements in their promoter (16). It is likely that CREB-2 on its own downregulates antisense transcription by interacting with CBP/p300, thus preventing it to bind transcription factors positioned on the 3'LTR, although it might have a positive role in the presence of Tax, as previously reported (10).

Based on the above CREB-1-dependent upregulation of luciferase expression in our stably transfected clones, the 3' LTR sequence of K30-3'asLuc was replaced with a LTR mutated

in the TRE1 repeats and unable to bind to CREB (3) thereby generating K30-3'asLuc 21PMD. These constructs were transiently co-transfected in 293T cells with our Tax expression vector (figure 4B). While no major differences were observed in basal luciferase activity between the samples, the K30-3'asLuc 21PMD-transfected cells showed a dramatically reduced response to Tax when compared to the wild-type construct. 293T cells were then stably transfected with the wild-type or TRE1-mutated K30-3'asLuc constructs (figure 4C). Tax expression in both cell populations demonstrated results similar to those presented above, further suggesting that TRE1 repeats were required for optimal Tax-mediated activation of HBZ expression. These results are in agreement with a previous study that demonstrated the nearly equal capacity of the 5'LTR and 3'LTR to bind transcription factors and co-activators in HTLV-1-infected cell lines and ATL cells (21).

To further confirm the role of CREB in Tax-induced HBZ expression in the proviral DNA context, ACH proviral DNA expressing either wild-type Tax (ACH-Tax) or Tax M47 mutant (ACH-Tax M47) were co-transfected in 293T cells along with a pTax or the empty vector (figure 4D). The relative level of HBZ mRNA was then analyzed by real-time RT-PCR. In the absence of pTax, the mutated ACH-Tax M47 proviral DNA showed a significantly reduced level of HBZ mRNA when compared to wild-type ACH-Tax. On the other hand, Tax expression resulted in an important increase in HBZ mRNA levels in both ACH-Tax and ACH-Tax M47 samples, confirming the modulator role of Tax on antisense transcription. No significant difference could be observed between the ACH-Tax and ACH-Tax M47 samples when a Tax-expression vector was co-transfected, most likely because the level of Tax protein level produced from the expression vector is importantly exceeding the one expressed from the proviral DNA.

The results presented in this study bring important information for the understanding of HTLV-1 replication and pathogenesis. First, we provide evidence that Tax can control the expression of its viral repressor HBZ, which is in agreement with a recent study (47). The existence of such a retroaction loop controlling the TAX/HBZ expression balance should prevent exaggerated Tax expression, which could ultimately lead to a strong and unfavorable anti-Tax immune response. In addition, our results show that the Tax-dependent induction of antisense promoter activity importantly depends on the integration site. The more pronounced Tax-dependent induction in 293T cell clones with low basal promoter activity might be consequential to its recently described capacity to induce the disassembly of nucleosomes at the viral promoter (40). In contrast, strong basal HBZ promoter activity would probably be associated with a favorable chromatin context, which would further explain the reduced upregulation mediated by Tax.

Recent studies have shown that HTLV-1 integrates randomly in the host genome (8), although integration in transcriptionally active units could be associated with HAM/TSP development (26). One could equally hypothesize that integration into a transcriptionally active site of the genome would affect the sense/antisense transcriptional balance in favor of HBZ expression, a feature that appears to be characteristic of ATL cells according to another recent study (43). In conclusion, our results demonstrate that Tax plays an important role in regulating HBZ expression. Further analyses are warranted to better understand the regulation of HBZ expression and how this might act upon transformation of HTLV-1-infected cells and consequently ATLL development.

ACKNOWLEDGMENTS

We thank Dr. D. Branch (The Institute of Medical Science, University of Toronto, Toronto, Canada) for the generous gift of the S1T cell line. We are also thankful toward Dr. Lee Ratner (Washington University, St-Louis, MO, USA) for providing us with the pACH Tax WT and pACH M47 vectors. This work was supported by The Cancer Research Society Inc. (B.B.) and by the Centre National de la Recherche Scientifique (CNRS)-Université Montpellier I and the Association pour la Recherche sur le Cancer (ARC n° 3606) (J.M.M.). S.L. was supported by a CIHR Ph.D. scholarship and B.B. holds a Canada Research Chair Tier 2. This work was performed by S.L. in partial fulfillment of a Ph.D. degree from the Microbiology-Immunology Program, Faculty of Medicine, Laval University.

REFERENCES

1. **Armstrong, A. P., A. A. Franklin, M. N. Uittenbogaard, H. A. Giebler, and J. K. Nyborg.** 1993. Pleiotropic effect of the human T-cell leukemia virus Tax protein on the DNA binding activity of eukaryotic transcription factors. *Proc Natl Acad Sci U S A* **90**:7303-7.
2. **Arpin-Andre, C., and J. M. Mesnard.** 2007. The PDZ domain-binding motif of the human T cell leukemia virus type 1 tax protein induces mislocalization of the tumor suppressor hScrib in T cells. *J Biol Chem* **282**:33132-41.
3. **Barnhart, M. K., L. M. Connor, and S. J. Marriott.** 1997. Function of the human T-cell leukemia virus type 1 21-base-pair repeats in basal transcription. *J Virol* **71**:337-44.
4. **Basbous, J., C. Arpin, G. Gaudray, M. Piechaczyk, C. Devaux, and J. M. Mesnard.** 2003. The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity. *J Biol Chem* **278**:43620-7.
5. **Bhat, N. K., Y. Adachi, K. P. Samuel, and D. Derse.** 1993. HTLV-1 gene expression by defective proviruses in an infected T-cell line. *Virology* **196**:15-24.
6. **Cavanagh, M. H., S. Landry, B. Audet, C. Arpin-Andre, P. Hivin, M. E. Pare, J. Thete, E. Wattel, S. J. Marriott, J. M. Mesnard, and B. Barbeau.** 2006. HTLV-I antisense transcripts initiating in the 3'LTR are alternatively spliced and polyadenylated. *Retrovirology* **3**:15.
7. **Datta, S., N. H. Kothari, and H. Fan.** 2000. In vivo genomic footprinting of the human T-cell leukemia virus type 1 (HTLV-1) long terminal repeat enhancer sequences in HTLV-1-infected human T-cell lines with different levels of Tax I activity. *J Virol* **74**:8277-85.
8. **Derse, D., B. Crise, Y. Li, G. Princler, N. Lum, C. Stewart, C. F. McGrath, S. H. Hughes, D. J. Munroe, and X. Wu.** 2007. Human T-cell leukemia virus type 1 integration target sites in the human genome: comparison with those of other retroviruses. *J Virol* **81**:6731-41.

9. **Du, K., H. Asahara, U. S. Jhala, B. L. Wagner, and M. Montminy.** 2000. Characterization of a CREB gain-of-function mutant with constitutive transcriptional activity in vivo. *Mol Cell Biol* **20**:4320-7.
10. **Gachon, F., A. Peleraux, S. Thebault, J. Dick, I. Lemasson, C. Devaux, and J. M. Mesnard.** 1998. CREB-2, a cellular CRE-dependent transcription repressor, functions in association with Tax as an activator of the human T-cell leukemia virus type 1 promoter. *J Virol* **72**:8332-7.
11. **Gachon, F., S. Thebault, A. Peleraux, C. Devaux, and J. M. Mesnard.** 2000. Molecular interactions involved in the transactivation of the human T-cell leukemia virus type 1 promoter mediated by Tax and CREB-2 (ATF-4). *Mol Cell Biol* **20**:3470-81.
12. **Georges, S. A., H. A. Giebler, P. A. Cole, K. Luger, P. J. Laybourn, and J. K. Nyborg.** 2003. Tax recruitment of CBP/p300, via the KIX domain, reveals a potent requirement for acetyltransferase activity that is chromatin dependent and histone tail independent. *Mol Cell Biol* **23**:3392-404.
13. **Gessain, A., F. Barin, J. C. Vernant, O. Gout, L. Maurs, A. Calender, and G. de The.** 1985. Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis. *Lancet* **2**:407-10.
14. **Giebler, H. A., J. E. Loring, K. van Orden, M. A. Colgin, J. E. Garrus, K. W. Escudero, A. Brauweiler, and J. K. Nyborg.** 1997. Anchoring of CREB binding protein to the human T-cell leukemia virus type 1 promoter: a molecular mechanism of Tax transactivation. *Mol Cell Biol* **17**:5156-64.
15. **Harrod, R., Y. Tang, C. Nicot, H. S. Lu, A. Vassilev, Y. Nakatani, and C. Z. Giam.** 1998. An exposed KID-like domain in human T-cell lymphotropic virus type 1 Tax is responsible for the recruitment of coactivators CBP/p300. *Mol Cell Biol* **18**:5052-61.
16. **Karpinski, B. A., G. D. Morle, J. Huggenvik, M. D. Uhler, and J. M. Leiden.** 1992. Molecular cloning of human CREB-2: an ATF/CREB transcription factor that can negatively regulate transcription from the cAMP response element. *Proc Natl Acad Sci U S A* **89**:4820-4.

17. **Kashanchi, F., J. F. Duvall, R. P. Kwok, J. R. Lundblad, R. H. Goodman, and J. N. Brady.** 1998. The coactivator CBP stimulates human T-cell lymphotropic virus type I Tax transactivation in vitro. *J Biol Chem* **273**:34646-52.
18. **Kobayashi, N., N. Yamamoto, Y. Koyanagi, J. Schneider, G. Hunsmann, and M. Hatanaka.** 1984. Translation of HTLV (human T-cell leukemia virus) RNA in a nuclease-treated rabbit reticulocyte system. *EMBO J* **3**:321-5.
19. **Kwok, R. P., M. E. Laurance, J. R. Lundblad, P. S. Goldman, H. Shih, L. M. Connor, S. J. Marriott, and R. H. Goodman.** 1996. Control of cAMP-regulated enhancers by the viral transactivator Tax through CREB and the co-activator CBP. *Nature* **380**:642-6.
20. **Langlois, M., B. Audet, E. Legault, M. E. Pare, M. Ouellet, J. Roy, N. Dumais, J. M. Mesnard, D. M. Rothstein, S. J. Marriott, M. J. Tremblay, and B. Barbeau.** 2004. Activation of HTLV-I gene transcription by protein tyrosine phosphatase inhibitors. *Virology* **329**:395-411.
21. **Lemasson, I., N. J. Polakowski, P. J. Laybourn, and J. K. Nyborg.** 2004. Transcription regulatory complexes bind the human T-cell leukemia virus 5' and 3' long terminal repeats to control gene expression. *Mol Cell Biol* **24**:6117-26.
22. **Lu, H., C. A. Pise-Masison, T. M. Fletcher, R. L. Schiltz, A. K. Nagaich, M. Radonovich, G. Hager, P. A. Cole, and J. N. Brady.** 2002. Acetylation of nucleosomal histones by p300 facilitates transcription from tax-responsive human T-cell leukemia virus type 1 chromatin template. *Mol Cell Biol* **22**:4450-62.
23. **Maeda, M., N. Arima, Y. Daitoku, M. Kashihara, H. Okamoto, T. Uchiyama, K. Shirono, M. Matsuoka, T. Hattori, K. Takatsuki, and et al.** 1987. Evidence for the interleukin-2 dependent expansion of leukemic cells in adult T cell leukemia. *Blood* **70**:1407-11.
24. **Matsumoto, J., T. Ohshima, O. Isono, and K. Shimotohno.** 2005. HTLV-1 HBZ suppresses AP-1 activity by impairing both the DNA-binding ability and the stability of c-Jun protein. *Oncogene* **24**:1001-10.

25. **Matsumoto, K., H. Shibata, J. I. Fujisawa, H. Inoue, A. Hakura, T. Tsukahara, and M. Fujii.** 1997. Human T-cell leukemia virus type 1 Tax protein transforms rat fibroblasts via two distinct pathways. *J Virol* **71**:4445-51.
26. **Meekings, K. N., J. Leipzig, F. D. Bushman, G. P. Taylor, and C. R. Bangham.** 2008. HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP. *PLoS Pathog* **4**:e1000027.
27. **Miller, S. C., A. Taylor, K. Watanabe, K. Mok, and F. M. Torti.** 1997. Regulation of NF-kappaB and HIV-1 LTR activity in mouse L cells by ultraviolet radiation: LTR trans-activation in a nonirradiated genome in heterokaryons. *Exp Cell Res* **230**:9-21.
28. **Miyazaki, M., J. Yasunaga, Y. Taniguchi, S. Tamiya, T. Nakahata, and M. Matsuoka.** 2007. Preferential selection of human T-cell leukemia virus type 1 provirus lacking the 5' long terminal repeat during oncogenesis. *J Virol* **81**:5714-23.
29. **Miyoshi, I., I. Kubonishi, S. Yoshimoto, T. Akagi, Y. Ohtsuki, Y. Shiraishi, K. Nagata, and Y. Hinuma.** 1981. Type C virus particles in a cord T-cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature* **294**:770-1.
30. **Munoz, E., and A. Israel.** 1995. Activation of NF-kappa B by the Tax protein of HTLV-1. *Immunobiology* **193**:128-36.
31. **Murata, K., T. Hayashibara, K. Sugahara, A. Uemura, T. Yamaguchi, H. Harasawa, H. Hasegawa, K. Tsuruda, T. Okazaki, T. Koji, T. Miyanishi, Y. Yamada, and S. Kamihira.** 2006. A novel alternative splicing isoform of human T-cell leukemia virus type 1 bZIP factor (HBZ-SI) targets distinct subnuclear localization. *J Virol* **80**:2495-505.
32. **Nakamura, M., M. Niki, K. Ohtani, and K. Sugamura.** 1989. Differential activation of the 21-base-pair enhancer element of human T-cell leukemia virus type I by its own trans-activator and cyclic AMP. *Nucleic Acids Res* **17**:5207-21.
33. **Okada, M., and K. T. Jeang.** 2002. Differential requirements for activation of integrated and transiently transfected human T-cell leukemia virus type 1 long terminal repeat. *J Virol* **76**:12564-73.

34. **Osame, M., K. Usuku, S. Izumo, N. Ijichi, H. Amitani, A. Igata, M. Matsumoto, and M. Tara.** 1986. HTLV-I associated myelopathy, a new clinical entity. *Lancet* **1**:1031-2.
35. **Poiesz, B. J., F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, and R. C. Gallo.** 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* **77**:7415-9.
36. **Poteat, H. T., P. Kadison, K. McGuire, L. Park, R. E. Park, J. G. Sodroski, and W. A. Haseltine.** 1989. Response of the human T-cell leukemia virus type 1 long terminal repeat to cyclic AMP. *J Virol* **63**:1604-11.
37. **Rodgers-Johnson, P., D. C. Gajdusek, O. S. Morgan, V. Zaninovic, P. S. Sarin, and D. S. Graham.** 1985. HTLV-I and HTLV-III antibodies and tropical spastic paraparesis. *Lancet* **2**:1247-8.
38. **Rousset, R., C. Desbois, F. Bantignies, and P. Jalinot.** 1996. Effects on NF-kappa B1/p105 processing of the interaction between the HTLV-1 transactivator Tax and the proteasome. *Nature* **381**:328-31.
39. **Satou, Y., J. Yasunaga, M. Yoshida, and M. Matsuoka.** 2006. HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci U S A* **103**:720-5.
40. **Sharma, N., and J. K. Nyborg.** 2008. The coactivators CBP/p300 and the histone chaperone NAP1 promote transcription-independent nucleosome eviction at the HTLV-1 promoter. *Proc Natl Acad Sci U S A* **105**:7959-63.
41. **Thebault, S., J. Basbous, P. Hivin, C. Devaux, and J. M. Mesnard.** 2004. HBZ interacts with JunD and stimulates its transcriptional activity. *FEBS Lett* **562**:165-70.
42. **Uchiyama, T.** 1997. Human T cell leukemia virus type I (HTLV-I) and human diseases. *Annu Rev Immunol* **15**:15-37.
43. **Usui, T., K. Yanagihara, K. Tsukasaki, K. Murata, H. Hasegawa, Y. Yamada, and S. Kamihira.** 2008. Characteristic expression of HTLV-1 basic zipper factor (HBZ) transcripts in HTLV-1 provirus-positive cells. *Retrovirology* **5**:34.

44. **Yan, J. P., J. E. Garrus, H. A. Giebler, L. A. Stargell, and J. K. Nyborg.** 1998. Molecular interactions between the coactivator CBP and the human T-cell leukemia virus Tax protein. *J Mol Biol* **281**:395-400.
45. **Yoshida, M.** 1994. Mechanism of transcriptional activation of viral and cellular genes by oncogenic protein of HTLV-1. *Leukemia* **8 Suppl 1**:S51-3.
46. **Yoshida, M., I. Miyoshi, and Y. Hinuma.** 1982. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc Natl Acad Sci U S A* **79**:2031-5.
47. **Yoshida, M., Y. Satou, J. Yasunaga, J. Fujisawa, and M. Matsuoka.** 2008. Transcriptional control of spliced and unspliced human T-cell leukemia virus type 1 bZIP factor (HBZ) gene. *J Virol* **82**:9359-68.

FIGURE LEGENDS

Figure 1. HTLV-1 antisense promoter activity in stimulated Jurkat cells and HTLV-1-infected-cells. (A) Jurkat cells were transfected with 15 μ g of K30-3'asLuc, stimulated for 8h with various T-cell activating agents (PHA 3 μ g/mL, PMA20ng/mL, Ionomycin 1 μ M, bpV[pic] 15 μ M, Forskolin 10 μ M, TNF α 20ng/mL) at 16h post-transfection or left untreated. (NS: no stimulation). Cells were monitored for luciferase activity. Luciferase values are presented as the mean of triplicates \pm SD. Student's T-tests were performed to determine significant differences between samples and the unstimulated control. (* = p-value < 0,05 ; ** = p-value < 0,01) **(B)** Jurkat cells and various HTLV-1-infected cells were transfected with 15 μ g K30-3'asLuc along with 4 μ g pRc-actin-LacZ. Cells were lysed 24h post-transfection and monitored for luciferase and β -galactosidase activity. Luciferase values are presented as the mean of triplicates \pm SD and are expressed as normalized RLU (RLU/ β -galactosidase). **(C)** The amount of Tax protein expressed in each cell line was assessed by Western blot. Detection of GAPDH is shown as a loading control.

Figure 2. Upregulation of antisense transcription by the viral Tax protein. (A) 293T and Jurkat cells were transfected with 500ng K30-3'asLuc, 200ng pRc-actin-LacZ and 300ng of pTax or the empty vector pH β Pr.1neo. Cells were lysed 48h post-transfection and monitored for luciferase and β -galactosidase activity. Luciferase activities represent the mean value of three measured samples \pm SD. and are expressed as normalized RLU (RLU/ β -galactosidase). **(B,C)** S1T cells were electroporated with 15 μ g of pTax or pH β Pr.1neo. Total RNA was isolated 48h post-transfection and the HBZ spliced mRNA

was amplified by standard **(B)** or real time RT-PCR **(C)**, using β -actin or HPRT-1 as an internal control, respectively. These results are representative of three independent experiments and are presented in independently transfected triplicates. Student's T-tests were performed to determine significant differences between samples (* = p-value < 0,05 ; ** = p-value < 0,01).

Figure 3. Tax-induced antisense promoter activity is influenced by the integration site.

(A) 293T cells were stably transfected with 4 μ g K30-3'asLuc and 400ng pCMV-Hyg. Different clones and a pool of clones were measured for luciferase activity. Luciferase activities represent the mean value of three measured samples \pm SD **(B)** Pooled clones and different 293T cell clones were transfected with 800 ng of pTax or the empty vector and monitored for luciferase activity 48h after transfection. Results are presented as fold increase of normalized luciferase activity measured for Tax-expressing cells over the activity of cells transfected with the control vector **(C)** Real-time PCR amplification of the luciferase gene was conducted in order to quantify the integrated plasmid copy number in each clone. Results were normalized with the HPRT-1 gene and are shown as number of copy per ng of genomic DNA. **(D)** Western blot analysis for CREB-1, CREB-2 and GAPDH (loading control) was performed on 50 μ g of total protein extracts from each 293T cell clone. **(E)** Two stable K30-3'asLuc clones were transfected with 800ng pSG-Tax, pSG-Tax M22, pSG-Tax M47, pSG-Tax K88A, pSG-Tax Δ PBM or the empty pSG-5 vector. Cells were lysed 48h post-transfection and monitored for luciferase activity. Luciferase activities represent the mean value of three measured samples \pm SD. **(F)** Tax

levels from each expression vectors was assessed by Western blot using Tax antiserum. GAPDH levels were analyzed as a control.

Figure 4. Positive regulation of the antisense promoter by CREB. (A) Two 293T cells clones stably transfected with K30-3'asLuc and the isolated pool were transfected with 800ng of expression vectors for CREB, CREB-FY, or CREB-2 or the empty vector. Cells were lysed 48hr post-transfection and monitored for luciferase activity. Luciferase activities represent the mean value of three measured samples \pm SD. (B) 293T cells were transfected with 500ng of K30-3'asLuc or K30-3'asLuc 21PMD, 200ng pRc-actin-LacZ and 300ng of pTax (vs pH β Pr.1neo). Cells were lysed 48h post-transfection and monitored for luciferase and β -galactosidase activity. Luciferase activities represent the mean value of three measured samples \pm SD. and are expressed as normalized RLU (RLU/ β -galactosidase). (C) 293T cells stably transfected with K30-3'asLuc or K30-3'asLuc were transfected with 800ng of pTax or pH β Pr.1neo. Cells were lysed 48h post-transfection and monitored for luciferase activity. Luciferase activities represent the mean value of three measured samples \pm SD. (D) Wild-type- or Tax M47-expressing ACH proviral DNA (500 ng) were transfected in 293T cells with 200ng of pRc-actin-LacZ and 300ng of a pTax or pH β Pr.1neo. HBZ mRNA expression levels were quantified by real-time RT-PCR using HPRT-1 gene as a reference control.

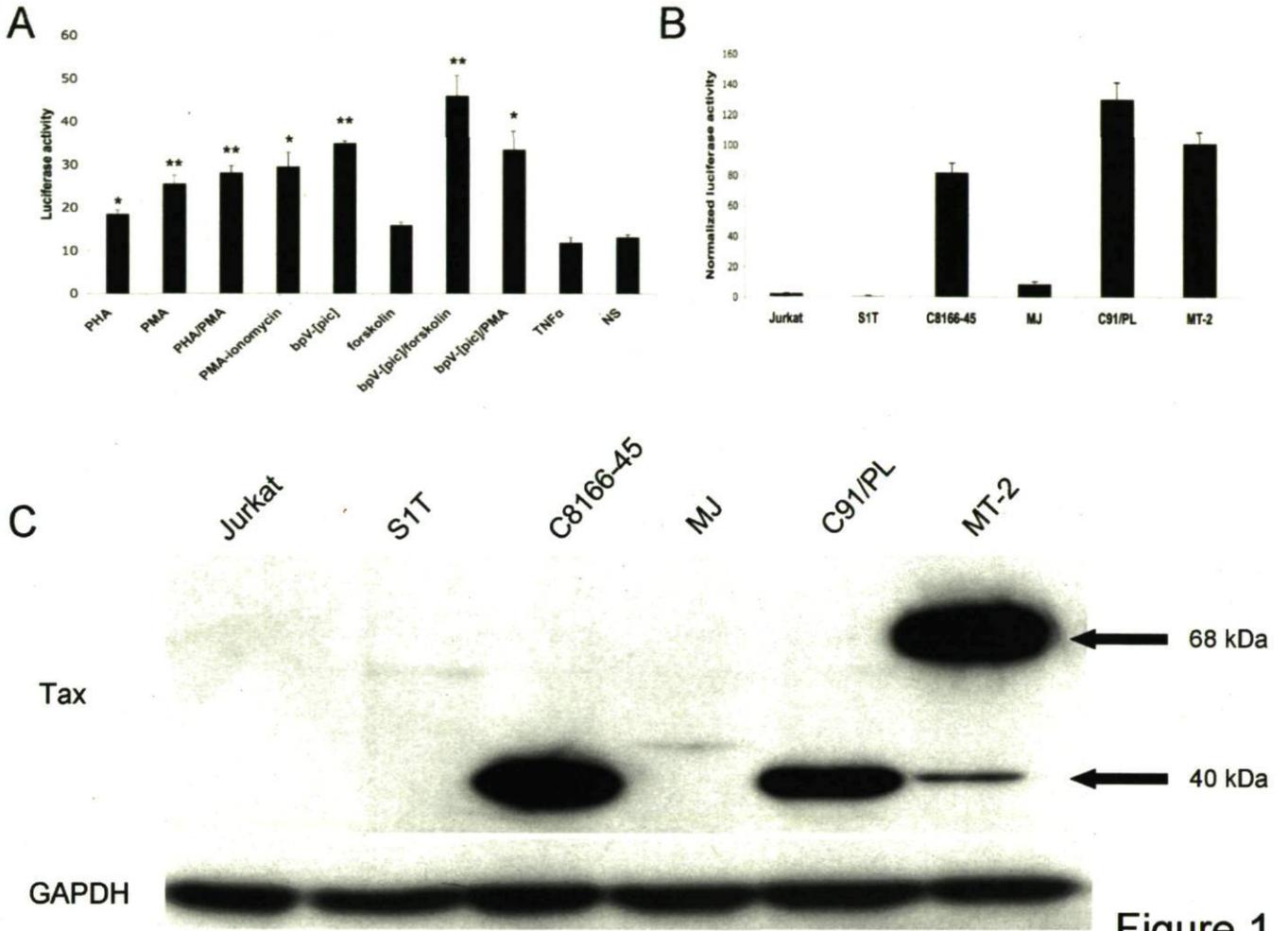


Figure 1

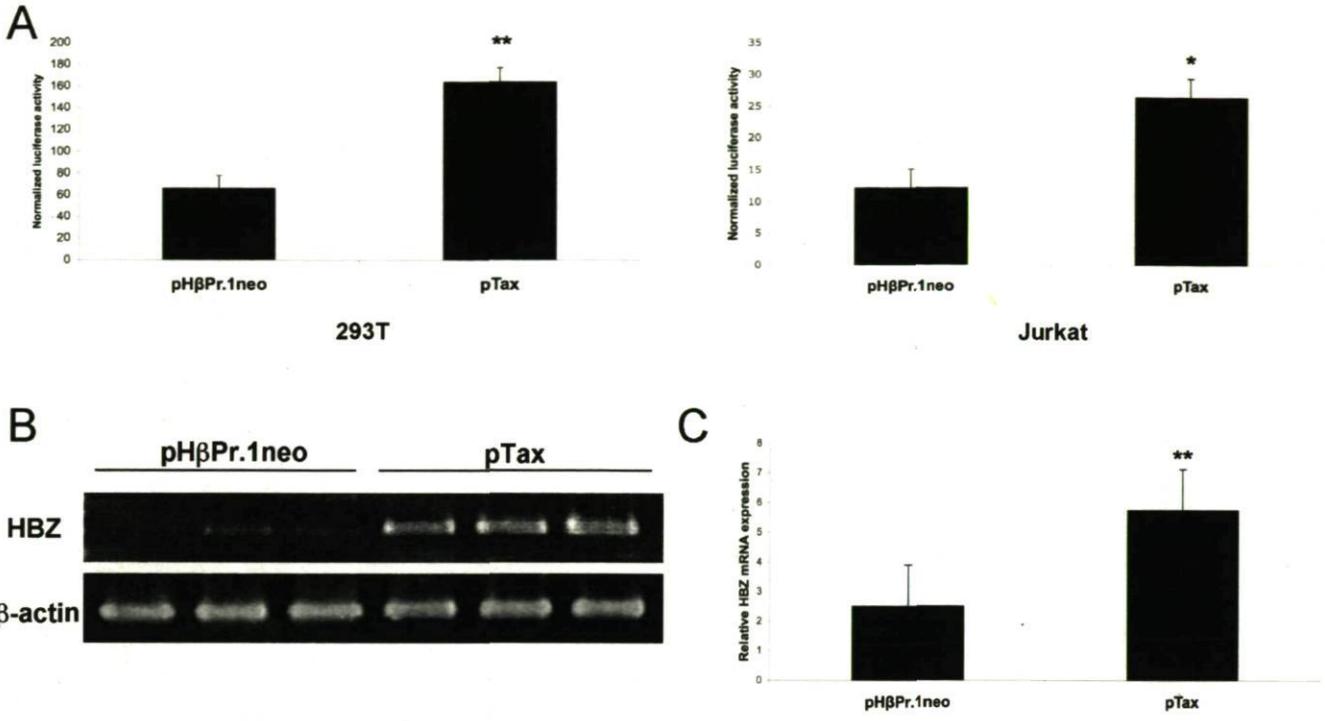


Figure 2

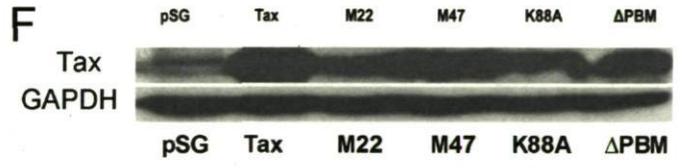
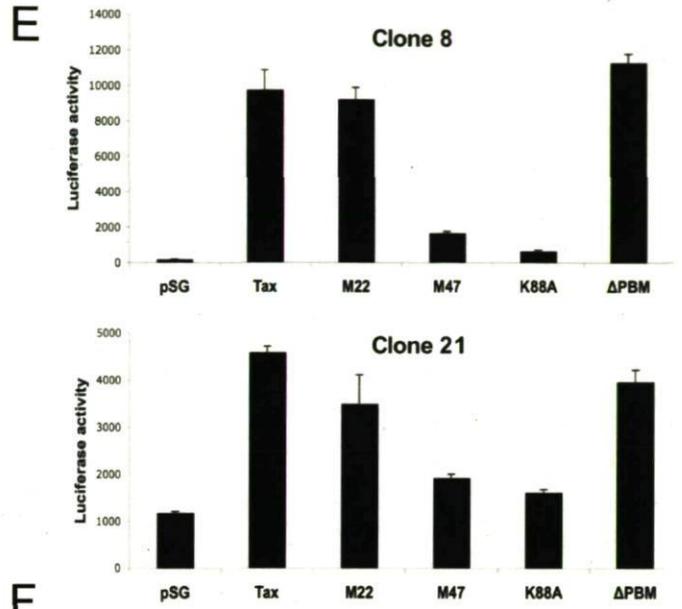
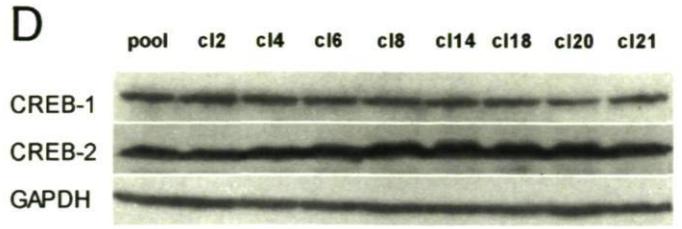
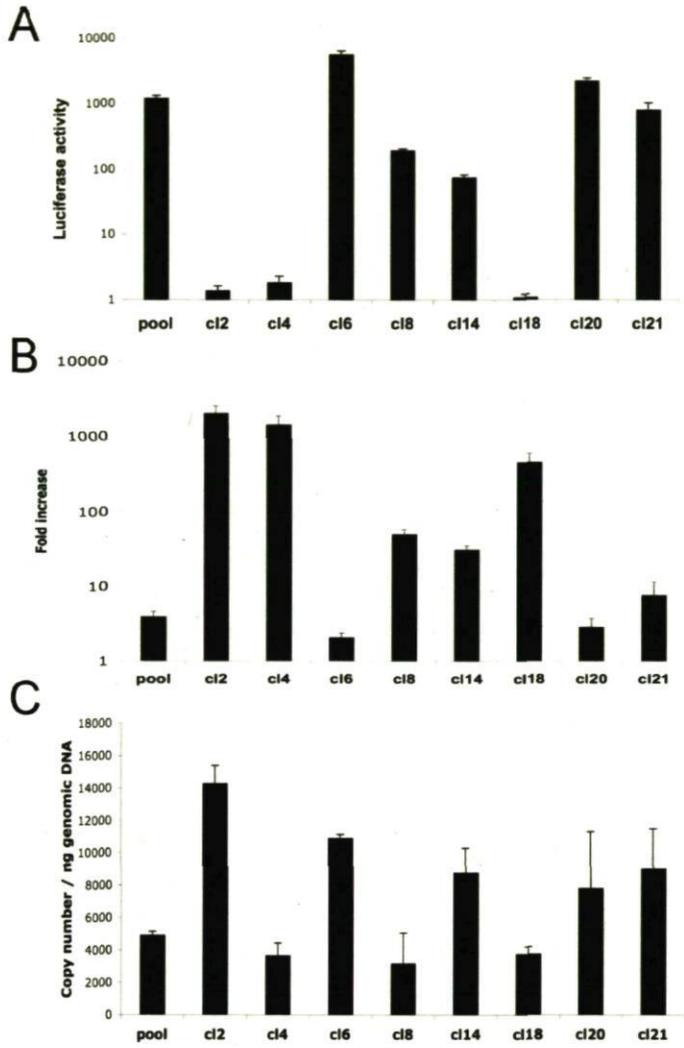


Figure 3

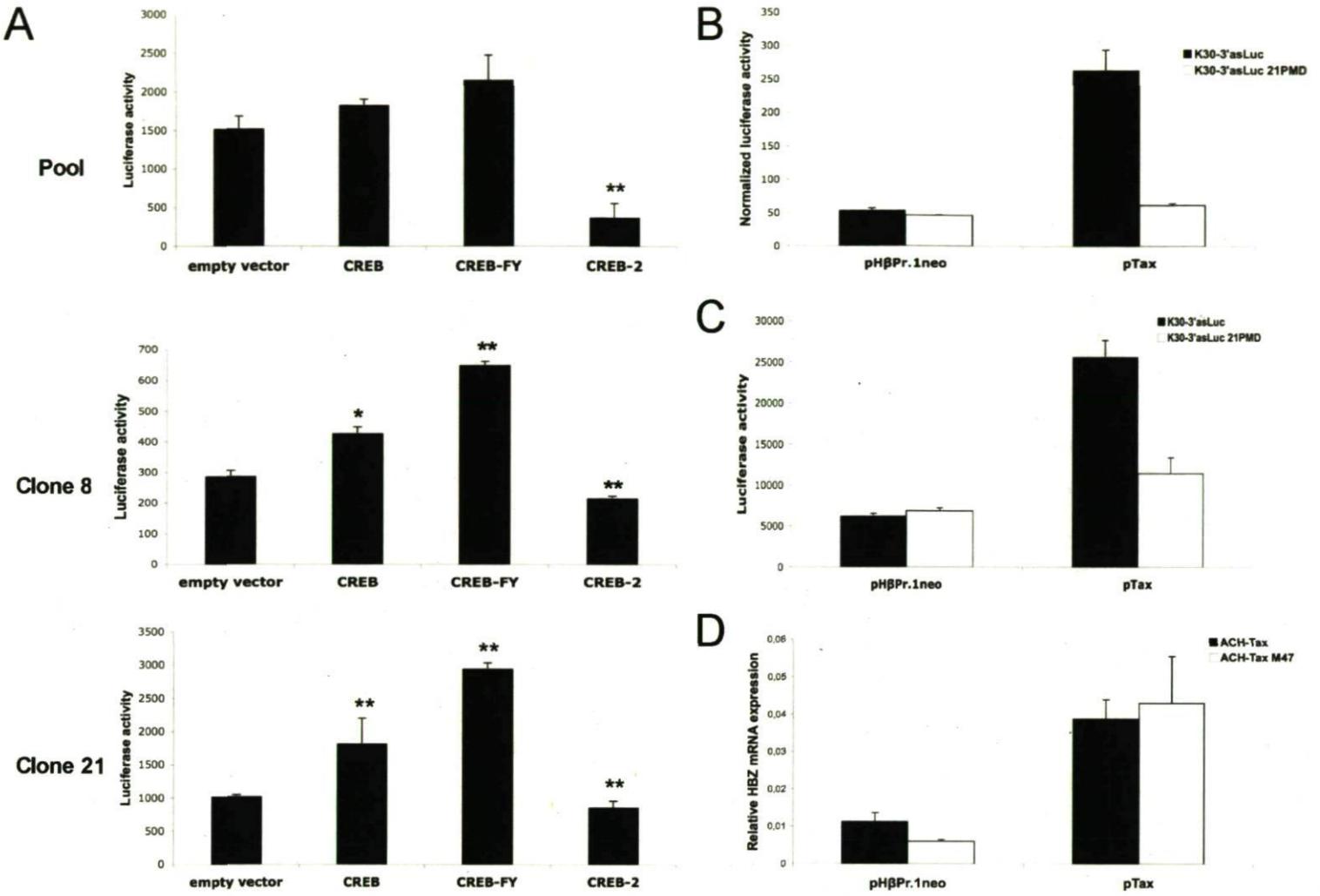


Figure 4

Chapitre 5 : Les rétrovirus HTLV-3 et HTLV-4 expriment des transcrits antisens permettant l'expression de protéines capables d'inhiber l'action de Tax

Ce chapitre est présenté sous forme d'article non-soumis.

5.1 Résumé

Les rétrovirus humains HTLV-3 et HTLV-4, nouvellement découverts, font partie de la famille des PTLVs, comprenant notamment les rétrovirus HTLV-1 et HTLV-2. La démonstration inattendue de l'existence de transcrits antisens chez HTLV-1 et l'identification de la nouvelle protéine HBZ codée par ces transcrits ont récemment attiré l'attention sur un phénomène oublié de la biologie des rétrovirus. Puisque l'analyse de la séquence des rétrovirus HTLV-3 et HTLV-4 a pu mener à l'identification de cadres de lecture ouverts sur le brin complémentaire de leurs génomes, nous étions intéressés à vérifier si ces virus exprimaient des transcrits antisens et si ces derniers pouvaient coder pour des protéines virales. Des analyses initiales effectuées dans les cellules 293T transfectées avec des constructions d'ADN proviral ont permis de détecter la présence de transcrits antisens pour les deux virus. Des analyses de RACE ont démontré que ces transcrits sont épissés et polyadénylés. Les protéines codées par ces transcrits, nommées APH-3 et APH-4, contiennent une région basique, mais n'ont pas de domaine *leucine zipper*. Des analyses en microscopie confocale ont permis de mettre en évidence la localisation nucléaire de la protéine APH-4 et la localisation cytoplasmique et nucléaire de APH-3. Des analyses des régions promotrices ont révélé que l'expression des transcrits est modulée par des agents activateurs des cellules T. De plus, les deux protéines semblent posséder la capacité d'inhiber la transcription virale médiée par Tax1 et Tax3 à partir du

LTR d'HTLV-1. Ces résultats démontrent que la transcription antisens n'est pas un phénomène exclusif à HTLV-1 et suggèrent fortement que les nouvelles protéines APH-3 et APH-4 jouent un rôle important dans la réplication des rétrovirus HTLV-3 et HTLV-4.

5.2 Article

HTLV-3 and HTLV-4 viruses produce antisense transcripts encoding proteins with a Tax-inhibiting function

Marilène Halin^{1*}, Sébastien Landry^{1*}, Émilie Laroque¹, Suzan Marriott², William Switzer³ and Benoit Barbeau^{1†}

¹Département des Sciences Biologiques et Centre de recherche BioMed, Université du Québec à Montréal, 2080 Saint-Urbain, Montréal, Québec, Canada H2X 3X8, ²Department of Molecular Virology and Microbiology, Baylor College of Medicine, Houston TX 77030; ³Laboratory Branch, Division of HIV/AIDS Prevention, National Center for HIV, STD, and TB prevention, Centers for Disease Control and Prevention, Atlanta, GA 30333.

†Correspondence to Benoit Barbeau, Département des Sciences Biologiques, Université du Québec à Montréal, 2080 Saint-Urbain Montréal, Québec, Canada H2X 3X8. Tel: 514 987-3000 ext.4576, E-mail: barbeau.benoit@uqam.ca

* These authors have equally contributed to this work.

Abstract

Human HTLV-3 and HTLV-4 retroviruses have been recently discovered and they are part of the HTLV family composed of the well-known HTLV-1 and HTLV-2 viruses. Antisense transcription in human retroviruses is a concept that has recently emerged and has been conclusively demonstrated in HTLV-1. In this virus, antisense transcription allows the synthesis of a viral protein termed HBZ. In light of the presence of potential ORFs in the antisense strand of HTLV-3 and HTLV-4, we were thus interested in evaluating the occurrence of antisense transcription in these viruses and whether these transcripts could encode a protein. Using proviral DNA constructs, RT-PCR analyses indeed indicated the existence of antisense transcription in transfected 293T cells. Our results showed that these transcripts were spliced and polyadenylated. The proteins encoded by these transcripts, termed APH (Antisense Protein of HTLV)-3 and APH-4, were both devoid of a bZIP domain but contained basic-rich regions. Confocal microscopy and Western blot experiments demonstrated a nuclear-restricted pattern for APH-4, whereas APH-3 could be observed both in the cytoplasm and the nucleus. Promoter analysis revealed that APH-3 and APH-4 expression is modulated by T cells activators. Interestingly, both proteins were shown to be comparable in their capacity to inhibit Tax1- and Tax3-mediated HTLV-1 LTR activation. These results provide further evidence that antisense transcription in human retroviruses is not exclusive to HTLV-1 and that APH-3 and APH-4 are likely to have an impact on HTLV-3 and HTLV-4 replication, respectively.

Introduction

Human T-cell Leukemia Viruses (HTLV) are human deltaretroviruses, which are part of the Primate T-cell Lymphotropic Virus (PTLV) group and are constituted by two well studied members known as HTLV-1, the first retrovirus to be isolated in human (Miyoshi et al., 1981; Poiesz et al., 1980; Poiesz et al., 1981; Yoshida et al., 1982) and HTLV-2. HTLV-1 bears important significance in terms of human health, as this virus is the etiological agent of Adult T-cell Leukemia/Lymphoma (ATLL) and HTLV-1-Associated Myelopathy/Tropical Spastic Paraparesis (HAM/TSP). Unlike HTLV-1, HTLV-2 has only been seriously linked to HAM-like pathologies (and not to leukemia). Recently, two new HTLV viruses, termed HTLV-3 and HTLV-4, have been isolated, the former having been closely compared to HTLV-3 (Calattini et al., 2006; Calattini et al., 2005; Wolfe et al., 2005). Currently, these viruses have been isolated in a relatively low number of individuals from Africa and no pathologies have yet been associated with these viruses.

The discovery of these two retroviruses has led to a series of recent studies comparing them to HTLV-1 and HTLV-2. These studies have indeed demonstrated that both HTLV-3 and HTLV-4 shared similar genomic organization to the former human retroviruses and weak but reproducible cross-reactivity with HTLV-1 and HTLV-2-specific antibodies was observed (Calattini et al., 2006; Calattini et al., 2005; Switzer et al., 2006; Wolfe et al., 2005). Further studies have focussed on the Tax3 protein of HTLV-3 and have highlighted its similarity to HTLV-1 Tax in terms of both intracellular localization and transactivation activity (Calattini et al., 2006; Chevalier et al., 2006). A recent study has also provided

evidence that the HTLV-3 genome, when reconstituted, expresses several different retroviral genes and produces infectious particle (Chevalier et al., 2008). Alike HTLV-1, these reports have also pointed toward the existence of a potential ORF, which could be produced from the antisense strand (Calattini et al., 2006; Chevalier et al., 2008; Switzer et al., 2006).

Former studies had indeed shed light on the existence of a new unexpected retroviral protein named HBZ (HTLV-I bZIP), which is encoded on the antisense strand of the HTLV-1 genome (Gaudray et al., 2002). Typically, two HBZ isoforms are produced with one of them being most abundant and dependent on a spliced transcript (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). Both HBZ isoforms block Tax-induced and basal HTLV-1 transcription and interact with several Jun family members, rendering some of these members inactive through possible sequestration in transcriptionally inactive nuclear bodies (Basbous et al., 2003; Hivin et al., 2007; Matsumoto et al., 2005; Thebault et al., 2004). A number of reports have also demonstrated that HBZ was expressed in cells from ATLL patients and that it could therefore be implicated in ATLL development, partly through its hyperproliferative action on T cells (Arnold et al., 2008; Barbeau and Mesnard, 2007; Matsuoka and Jeang, 2007; Mesnard et al., 2006; Satou et al., 2006).

Although antisense transcription has been suggested in other retroviruses (Arnold et al., 2008; Briquet et al., 2001; Briquet and Vaquero, 2002; Michael et al., 1994; Peeters et al., 1996; Tagieva and Vaquero, 1997; Vanhee-Brossollet et al., 1995), more data are needed to clearly establish that antisense transcription is present in several retroviruses beside HTLV-1. Given that former studies have alluded to the possible existence of antisense transcription in HTLV-3 and HTLV-4, we herein investigated on the presence of antisense

transcripts from these viruses and on its potential encoding capacity. Our results indicate that both viruses produce a spliced and polyadenylated antisense transcript. The proteins encoded by these transcripts show distinct localization, the HTLV-3 antisense protein being both nuclear and cytoplasmic, while the HTLV-4 counterpart is almost exclusively contained in the nucleus. However, both antisense proteins were shown to inhibit Tax-mediated HTLV LTR activation. These results hence provide significant data indicating the importance of antisense proteins in retroviral replication.

Material and methods

Cell lines and antibodies

The human 293T and mouse kidney COS-7 cell lines were maintained in DMEM culture medium supplemented with 10% foetal bovine serum (Hyclone Laboratories, Waltham, MA). The T-cell line Jurkat E6.1 was maintained in RPMI-1640 medium supplemented with 10% foetal bovine serum (Hyclone Laboratories, Logan, UT), 2 mM glutamine, 100 U/ml penicillin G and 100 µg/ml streptomycin. The anti-Myc antibody (9E10) was purchased from Santa Cruz biotechnology while the ECL™ sheep anti-mouse IgG antibody coupled to the horseradish peroxidase was obtained from GE Healthcare.

Plasmids

Both HTLV-3 (2026ND) and HTLV-4 (pUC-HT4v2) proviral DNAs were cloned in pUC18. The pHTLV-3ΔEcoRV construct was derived from HTLV-3 (2026ND) proviral DNA by enzymatic digestion with EcoRV and XhoI, thereby deleting 4359pb of the 5' end of the proviral sequence. The pHTLV-3-as-luc vector was derived from pHTLV-3ΔEcoRV by cloning the luciferase reporter gene from pGL3-basic in frame in the second exon of APH-3 at position 7181 (sense orientation). Briefly, PacI and BstZ171 restriction sites were added to the pHTLV-3ΔEcoRV construct by PCR using Phusion™ Site-Directed Mutagenesis Kit according to manufacturer's instructions using the reverse primer 26-1 (5'P-**ATAGTATAC** TGCAATCCCAGGAACTG-3') containing BstZ171 restriction site (bold) and forward primer 27-1 (5'P-ATATTA**AATTAAT**GTCTCCGGGGCTAGG-3') containing PacI restriction site (bold). Luciferase was amplified using reverse primer 35-1

(5'- **ATATTAATTAAGGAAGACGCCAAAAACATAAAGAA**-3') containing *PacI* restriction site (bold) and forward primer 32-1 (5'- **ATAGTATACTACCACATTTGTAGAGGTTTTAG**-3') containing *BstZ171* restriction site (bold). Luciferase amplicons were digested by *PacI* and *BstZ171* and ligated in the pHTLV-3 Δ EcoRV digested vector. The pHTLV-4 Δ SacI construct was derived from HTLV-4 (pUC-HT4v2) proviral DNA by enzymatic digestion with *SacI* and *HindIII*, thereby deleting 3213pb of the 5' end of the proviral sequence. The pHTLV-4-as-luc vector was obtained by cloning the luciferase reporter gene in frame in the second exon of APH-4 at position 7037 (sense transcript) in pHTLV-4 Δ 6560 (construct containing the last 2183bp of the HTLV-4 genome in pUC18 vector). Briefly, *PacI* and *BstZ171* restriction sites were added to the pHTLV-4 Δ 6560 construct by PCR using PhusionTM Site-Directed Mutagenesis Kit according to manufacturer's instructions using the reverse primer 26-2 (5'-P-**ATAGTATACAGAGGAGATGCCTGGTA**-3') containing *BstZ171* restriction site (bold) and forward primer 31-1 (5'-P-**ATATTAATTAATGGTGTTGAGACCTTCTTTG**-3') containing *PacI* restriction site (bold). Luciferase was amplified using reverse primer 35-1 (5'- **ATATTAATTAAGGAAGACGCCAAAAACATAAAGAA**-3') containing *PacI* restriction site (bold) and forward primer 32-2 (5'- **ATAGTATACTACCACATTTGTAGAGGTTTTAG**-3') containing *BstZ171* restriction site (bold). Luciferase amplicons were digested by *PacI* and *BstZ171* and ligated in the pHTLV-4 Δ 6560 digested vector. The pHTLV-Luc vector contains the HTLV-1 LTR 3' region cloned into the *XhoI*/*HindIII* site of the pGL2-Basic vector (Promega) (Geleziunas et al., 1998). The Tax1 expressing vector has been previously described [295]. The pCMV-Tax3 and pCMV-Tax4 vectors contain the Tax3- and Tax4-coding regions under the

control of the CMV promoter region in the pcDNA3.1Zeo (+). The pMycAPH-3 and pMycAPH-4 were generated by cloning the RT-PCR amplification of the coding regions of APH-3 or APH-4 using a primer that contain a Myc tag in the pcDNA3.1Zeo (+) vector. Briefly, APH-3 was amplified using reverse primer 53-1 (5'-**ATGGAACAAAACTCATCTCAGAAGAGGATCTGATGGCTCGATCCCGAAGC** GG-3') containing a Myc tag (bold) and forward primer 36-1 (5'-**ATATCTAGATTATAACAGATCTGCTACCTCCTGTAG**-3') containing a XbaI restriction site (bold). APH-4 was amplified using reverse primer 59-1 (5'-**ATGGAACAAAACTCATCTCAGAAGAGGATCTGATGGACACTCGAGAATTTT** TTAGGGG-3') containing a Myc tag (bold) and forward primer 37-1 (5'-**ATATCTAGATTATAATAACTCCGCCAATACACCCAAC**-3') containing a XbaI restriction site (bold). Amplified products were digested by XbaI and ligated into pcDNA3.1Zeo (+) digested by XbaI and EcoRV. Plasmids expressing the GFP-APH-3 and GFP-APH-4 fusion proteins were obtained by cloning the amplified coding regions in the pEGFP-N1 vector (Clontech) in the EcoRI/AgeI sites. The pRcActin-LacZ vector contains the β -galactosidase gene under the regulation of the human actin promoter.

Transfection

293T cells were transfected with 10-15 μ g of DNA using the calcium phosphate protocol as previously described [307] or using the LipofectamineTM Reagent (Invitrogen) according to manufacturer's instructions. Jurkat cells were transfected by electroporation with 15 μ g of DNA using the Gene Pulser Xcell system (BioRad, Hercules, CA) (960 μ F, 250V). Transfected 293T cells were lysed 48h post-transfection in lysis buffer (25mM Tris

phosphate, pH 7.8, 2mM DTT, 1% Triton X-100, 10% glycerol) and luciferase activity readout was performed with the MLX microplate luminometer (Dynex Technologies) with a single injection of a luciferase buffer (20 mM tricine, 1.07 mM $(\text{MgCO}_3)_4 \cdot \text{Mg}(\text{OH})_2 \cdot 5\text{H}_2\text{O}$, 2.67 mM MgSO_4 , 0.1 mM EDTA, 220 μM Coenzyme A, 4.7 μM D-Luciferin potassium salt, 530 μM ATP, 33.3 mM DTT). Each sample was co-transfected with a β -gal-expressing vector for transfection efficiency normalisation. The β -galactosidase activity was measured using the Galacto-LightTM kit (Applied Biosystems, Bedford, MS) according to manufacturer's suggestions. Luciferase activities are presented in normalized Relative Light Units (RLU/ β -gal) and represent the calculated mean \pm SD of three transfected samples normalised by the measured β -galactosidase activity. Jurkat cells were resuspended at 2.5×10^7 /ml of complete medium, transfected in bulk and were separated at 16h post-transfection into various treatment groups at a density of 1×10^5 cells/well (100 μl) in 96-well flat-bottom plates. Cells were either left untreated or were treated with phorbol 12-myristate 13-acetate (PMA) at 20 ng/ml (Sigma-Aldrich, Oakville, Canada), phytohemagglutinin (PHA-P) at 3 $\mu\text{g}/\text{ml}$ (Sigma-Aldrich, Oakville, Canada), ionomycin (Sigma-Aldrich, St-Louis, MO) at 1 mM, bpV [pic] (Alexis corp., Farmingdale, NY) at 15 μM , forskolin (Sigma-Aldrich, St-Louis, MO) at 10 μM and $\text{TNF}\alpha$ at 20 ng/ml. Luciferase activity was monitored at 8h post-stimulation.

RT-PCR and 3'RACE analyses

Total RNA was extracted by the TRIzol reagent (Invitrogen) from transfected 293T cells 48 h post-transfection. PolyA⁺ RNA was purified from total RNA using Oligotex mRNA Mini Kit (Qiagen, Mississauga, ON) according to manufacturer's instructions. RT-PCR analyses

were conducted using an oligo(dT) primer (Invitrogen). Briefly, total RNA (5 μ g) was mixed with 1 μ l of 10 μ M of the oligo(dT) primer. The RNA:RT primer mix was heated at 70°C for 5 min, chilled on ice for 5 min and incubated 2 hours at 42°C in the presence of 1X AMV reaction buffer, 1 mM dNTPs, 10U of SUPERase-In RNase inhibitor (Ambion, Austin, TX) and 15 U AMV reverse transcriptase (USB, Cleveland, Ohio). Synthesized cDNAs were then PCR amplified in the presence of 1.25 U Taq DNA polymerase, 1X ThermoPol buffer, 20 μ M dNTPs, 1,5 μ M of each primer using a Tgradient thermocycler (Biometra, Goettingen, Germany). Primers used for PCR reactions were the reverse primer 23-1 (5'-CCACCAAATAGAAGAGGGATGGC-3') and forward primers 19-1 (5'-GACGCCCTGGCCCCAACAG-3') for the amplification of the antisense transcript of HTLV-3 and the reverse primer 21-1 (5'-CGGCGGCGTCTCAACTGATTG-3') and forward primer 23-4 (5'-ACGAGTCCCCCATATGTCCAAA-3') for the amplification of the antisense transcript of HTLV-4. PCR conditions were as follow: a first step of denaturation at 94°C for 5 min followed by 35 cycles of denaturation (94°C for 1 min), annealing (60°C for 1 min) and extension (72°C for 1 min) and a final extension at 72°C for 5 min. RT-PCR amplification were controlled for DNA contamination (RNA samples with no RT step). The polyA signals were identified from isolated polyA⁺ RNA with the FirstChoice RLM-RACE kit from Ambion according to manufacturer's instructions. The cDNA synthesis was performed in the presence of the supplied 3'RACE adapter. PCR amplification was achieved through 3'RACE outer and inner primers and a primer derived from the sequence downstream of the ORFs: 23-2 (5'-GGAGAGGAACCACACTGGATCAT-3') for HTLV-3 antisense transcript and 23-5 (5'-

GAGTCAGGACATGCTCTAGGTCT-3') for HTLV-4 antisense transcript. Amplified products were then cloned in pBlueScript KS and sequenced.

Confocal microscopy

COS-7 cells were seeded onto a Lab-Tek chamber slide (Nunc) for 24 h before transfection and were then transfected using the LipofectamineTM Reagent (Invitrogen) according to manufacturer's instructions. Live cells analysis was performed at 36h post-transfection using a confocal laser-scanning microscope (Bio-Rad MRC1024, CA, USA) equipped with a Nikon Eclipse TE300 camera (Nikon, Tokyo, Japan).

Western blot analysis

Cytoplasmic and nuclear extracts from transfected 293T cells were prepared using NE-PER[®] Nuclear and Cytoplasmic Extraction Reagents (Pierce Biotechnology) according to manufacturer's instructions. Equal quantities of extracts were run on a SDS-10% PAGE and transferred to PVDF membranes (Millipore, Billerica, MA). The blot was next blocked in PBS/5% milk and incubated with a mouse anti-Myc 9E10 antibody (dilution 1:1000). After several washes, signals were revealed by the addition of peroxylase-conjugated goat anti-mouse IgG (dilution 1:10000) antibody and subsequent incubation with the BM Chemiluminescence Blotting Substrate (POD) (Roche, Basel, Switzerland). Membranes were exposed on Amersham hyperfilms ECL (GE Healthcare).

Results

Detection of spliced antisense transcripts in HTLV-3 and HTLV-4-transfected cells.

Previous studies have highlighted the presence of a potential ORF in the antisense strand of HTLV-3 (Calattini et al., 2006; Chevalier et al., 2008; Switzer et al., 2006). Given that, in this retrovirus family, the demonstration of the existence of an antisense protein has been exclusively shown for HTLV-1, we thereby asked whether the newly discovered HTLV-3 and HTLV-4 human retroviruses could express similar antisense proteins.

Our first *in silico* analysis of the ORF present in both previously described HTLV-3 and HTLV-4 proviral DNA (Switzer et al., 2006; Wolfe et al., 2005) indeed revealed the existence of an ORF in the antisense strand between *tax* and *env* genes (Figure 3A and B). The encoded proteins were predicted to contain a maximum of 221 and 158 amino acids (for HTLV-3 and HTLV-4 respectively) from a predicted ATG initiation codon. These ORFs were termed APH (Antisense Protein of HTLV)-3 and APH-4.

To address the existence of an antisense transcript from these retroviruses, we first conducted RT-PCR analyses. Each proviral DNA constructs were first deleted of their 5'LTR region to minimize interference from sense transcription (Figure 1). The resulting pHTLV-3 Δ EcoRV and pHTLV-4 Δ SacI proviral DNA constructs were next transfected in 293T cells and analysed by RT-PCR using forward and reverse primers derived from the 3'LTR and APH-3/APH-4 regions, respectively (see Figure 1A and B). The choice of primers and their positioning were based on the previously described spliced pattern of the antisense HBZ transcript of the related HTLV-1 virus (Cavanagh et al., 2006; Murata et al., 2006; Satou et al., 2006). RT-PCR analyses indeed revealed the presence of a single signal

specific for an antisense transcript (as demonstrated by the absence of a signal in the absence of RT). In addition, its size also lent credence to the occurrence of splicing while no signals with a size expected for unspliced transcripts were amplified (Figure 2A and B). Splice sites of the HTLV-3 antisense transcript were positioned at nt 8556 and 7218 for the splice donor and splice acceptor, respectively, while their corresponding SD and SA sites were at nt 8422 and 7089 for the HTLV-4 antisense transcript. Position of the spliced sites and the length of the intronic region were very well conserved between HTLV-3 and HTLV-4. Similarly to HTLV-1 HBZ, in both APH-3 and APH-4 transcripts, the resulting spliced transcript led to the addition of an *N*-terminal amino acid sequence with a methionine initiation codon to the remaining corresponding APH ORF amino acid (Figure 2C). Interestingly, for APH-3, a total of 8 amino acids were added from exon 1 while APH-4 exon 1 was responsible for the addition of 4 amino acids. An important observation is that unlike HBZ, if the unspliced mRNA could encode for a different APH isoform, in both cases, the resulting encoded APH isoform would initiate at a position downstream of the splice acceptor site thereby leading to a shorter isoform than the protein potentially encoded from the spliced transcript.

These results hence demonstrated that both viruses were capable of producing a spliced antisense transcript with a protein coding capacity. Signals corresponding to an unspliced transcript were not detected in these analyses.

APH-3 and APH-4 transcripts are polyadenylated.

We were next interested in characterizing the 3' end of both retroviral antisense transcripts (Figure 3). Again, proviral DNA-deleted versions of HTLV-3 and HTLV-4 were

transfected in 293T cells and RNA were used to identify the 3' end of the transcripts by 3' RACE analysis. As depicted in Figure 3, both transcripts were cleaved and polyadenylated between a consensus polyA signal and a GT-rich sequence, which is often found in proximity to the polyA tail addition site. Interestingly, although the composition of the targeted dinucleotide for cleavage was different for both retroviruses (AG and CC for HTLV-3 and HTLV-4, respectively), both cleavages occurred at an equal distance from the polyA signal (19 nt).

Amino acid sequence of APH-3 and APH-4 and their cellular localization.

The resulting amino acid sequence of APH-3 and APH-4 were compared to the previously published HBZ sequence (Figure 4). This comparison revealed that the previously described bZIP domains of HBZ was not conserved in both APH-3 and APH-4. Basic-rich regions were however noted at similar positions and could reflect potential NLS sequence for both APH-3 and APH-4. In addition, similar LXXLL and LXXLL-like motifs known to be responsible for an interaction of HBZ with CBP/p300 were identified in both APH-3 and APH-4 proteins.

Based on these analyses, we next determined the cellular localization of APH-3 and APH-4 using expression vectors fused the GFP coding sequence at their amino ends. These vectors were transfected in COS7 cells and analyzed by confocal microscopy. Our first analysis of APH-3 revealed that it localized mainly in the nucleus but showed significant staining in the cytoplasm (Figure 5A). This staining pattern could be observed from 24 to 72 hours post-transfection (data not shown). In addition, punctuated structures were observed in

stained COS7 nuclei. Upon merging, nuclear signals were however shown to poorly co-localize with the nucleolus.

Similar analyses were conducted on 293T and COS7 cells transfected with the GFP-tagged APH-4 expression vector. Interestingly as opposed to APH-3 and more reminiscent of HBZ-SP1, APH-4 was strictly present in the nucleus and did co-localize with the nucleolus (Figure 5A). All these confocal microscopy experiments were retested on cells transfected with amino and carboxyl Myc-tagged APH expression vectors and again led to similar results (data not shown) indicating that the added tag and its position likely did not contribute to the cellular distribution of APH-3 and APH-4.

To confirm these results, Western blot analyses were performed on nuclear vs. cytoplasmic extracts prepared from 293T cells transfected with expression vectors for Myc-tagged HBZ, APH-3 or APH-4 (Figure 5B). Controls consisted of extracts from cells transfected with the empty expression vector. Extracts indeed revealed the expected size for APH-3 (28 kDa) while APH-4 indicated a lower molecular weight than the predicted size being 22 kDa. As previously reported (Arnold et al., 2006; Cavanagh et al., 2006), HBZ showed a slower migration than the expected size of 27 kDa. In addition, as predicted, both APH-4 and HBZ were mainly present in nuclear extracts while showing a weak presence in cytoplasmic extracts. On the other hand, the APH-3 signal was present in both extracts. Interestingly, a higher molecular weight signal was apparent exclusively in the nucleus of APH-3-expressing cells although the signal might be non-specific as judged by the presence of a weaker bands in control transfected cells.

These results hence demonstrated that both APH-3 and APH-4 are differently distributed in transfected cell, which might impact on their function in their respective viruses.

Promoter studies on both HTLV-3 and HTLV-4 antisense transcripts.

To further analyse the *APH-3* and *APH-4* genes, expression levels were analysed in Jurkat T cells using constructs in which a luciferase reporter gene along with the polyA signal was inserted in frame with the amino acid sequence present in exon 2 (Figure 6A). Jurkat cells were subsequently transfected and then either left untreated or stimulated with a series of known T-cell activators. In non-stimulated conditions, luciferase activity was found to be modest but significant showing no strong difference between HTLV-3 and HTLV-4 constructs (data not shown). Stimulation also led to an increase in luciferase activity for both vectors and were more responsive to the addition of the PTP inhibitor, bpV[pic], a strong T-cell activating agent.

Tax-driven HTLV LTR activation is inhibited by APH-3 and APH-4 expression.

Former studies have demonstrated that the HTLV-I-encoded antisense protein HBZ inhibited the Tax-mediated activation of its promoter. In addition, this inhibition was mediated by direct interaction with CREB-2 through its leucine zipper domain. As APH-3 and APH-4 seems to lack such a domain, we asked whether it could nonetheless act upon Tax-mediated HTLV-1 LTR activation. Expression vectors for APH-3, APH-4 and HBZ (in comparison to pcDNA3.1Zeo (+)) were transfected in 293T cells along with expression vectors for Tax 1 or Tax3 and pHTLV-luc. Measurement of normalized luciferase activity demonstrated that, as previously reported, HBZ blocked Tax activation of the HTLV-1 LTR. Importantly, both APH-3 and APH-4 showed a similar capacity to block this Tax-mediated HTLV-1 LTR upregulation, albeit at different levels.

These results therefore indicated that, although APH-3 and APH-4 lack a typical bZIP domain, both proteins can alter the extent of activation of the HTLV-1 LTR mediated by either Tax1 or Tax3 and could therefore play a similar role in HTLV-3 and HTLV-4 replication.

Discussion

HTLV-3 and HTLV-4 retroviruses have been recently isolated from Cameroon individuals, which had no sign of illness (Calattini et al., 2005; Wolfe et al., 2005). Studies have further demonstrated that a number of retroviral genes are shared with the related HTLV-1 and HTLV-2 viruses (Calattini et al., 2006; Chevalier et al., 2006; Switzer et al., 2006). Given that recent evidence points to the importance of the newly discovered antisense-encoded HBZ gene in both HTLV-1 replication and ATL development, the goal of this study was to provide evidence for the existence of antisense transcripts in these viruses and to further assess the possible functional relevance of their encoded proteins.

Our first *in silico* analyses confirmed previous analyses in that a potential ORF was indeed present on the antisense strand of the HTLV-3 and HTLV-4 genomes. Our analyses have demonstrated that these proteins were bZIP-less but contained a certain number of basic-rich regions similar to basic domains identified for HBZ. Although these basic amino acid-rich regions could serve in targeting APH-3 and APH-4 to the nucleus, these domains could also participate in modulating the transcriptional activity of APH-3 and APH-4.

Our RT-PCR analyses and 3' RACE experiments have conclusively indicated that the antisense transcripts of HTLV-3 and HTLV-4 are spliced and polyadenylated. The splicing pattern was similar to that of HBZ with a similar intronic size. Although more analyses are needed, current results argue for low abundance of unspliced transcripts and therefore suggest that no other isoform could be produced from the antisense transcript. In addition, unlike HBZ, translation of APH-3 and APH-4 would initiate with a methionine resulting in the deletion of 15 and 24 amino acids from the NH₂ end, respectively. Such deletions might

cause important change in the protein and might suggest why unspliced transcripts are not abundant and might therefore not lead to synthesis of another isoform.

Sequence analyses of the SD and SA sequence have revealed typical consensus sequence. Consensus polyA signal and GT-rich sequence were also identified in proximity to the polyA addition site for both HTLV-3 and HTLV-4. This is again reminiscent of the situation of the HBZ transcript. In fact positioning of the polyA addition site for *APH-3* and *APH-4* RNA is equivalent to the HBZ gene and likely reveals specific constraints associated with antisense transcription in retroviruses.

The analysis of the amino acid composition of both APH-3 and APH-4 and their comparison to HBZ demonstrated important differences. First and importantly, software analysis of APH-3 and APH-4 did not highlight a LZ domain, a marked difference with HBZ. In addition, APH-3 showed important differences in amino acid sequence. However, it should be stressed that like HBZ, basic regions were identified and LXXLL and LXXLL-like motifs (known to be responsible for binding p300/CBP) were also noted. These similarities might indeed contribute to the similarity in cellular localisation and functional aspects between APH-3, APH-4 and HBZ. Confocal and Western blot experiments have indeed indicated that APH-3 and APH-4 are both nuclear (at least partly for APH-3). The NLS sequence which mediates nuclear targeting remains to be determined but might reasonably involve similar regions to those responsible for nuclear localisation of HBZ. However, unlike HBZ (the SP1 isoform), APH-3 does not co-localize with the nucleolus. In fact, its nuclear distributions are more representative of the formerly published localisation of the HBZ isoform, thought to be produced from unspliced transcripts (Gaudray et al.,

2002). Further experiments will be required to determine how these differences in nuclear distribution could be explained.

Another interesting observation is the cytoplasmic localization of APH-3 observed in COS7 cells. This observation might be indicative of the existence of a potential nuclear export signal. It is also important to note that a higher molecular weight signal was noted in our Western blot analysis in APH-3-expressing cells and that this signal was mainly localised to the nucleus. This opens up the possibility that nuclear localisation for APH-3 (and even for HBZ and APH-4) could be importantly modulated by potential post-translational modifications.

Studies of the APH-3 and APH-4 genes using the luciferase reporter gene have further showed that their expression was stimulated by known T-cell activators. Similar results have been obtained with constructs similar for the HBZ gene. It might be expected that as for HBZ expression, Tax expression would induce its expression *via* the equivalent TRE sequences present in the U3 region (Landry et al., 2008; Yoshida et al., 2008).

Functional analyses have been investigated for both APH-3 and APH-4. HTLV-1 LTR has been previously shown to also be activated by Tax3 (Calattini et al., 2006). Our analyses have indeed revealed that both antisense proteins could indeed inhibit Tax1 and Tax3 activation. This raises the issue, as we have previously demonstrated, that the leucine zipper domain is not the only amino acid segment responsible for Tax inhibition by HBZ (Clerc et al., 2008). Amino acid sequence of both APH-3 and APH-4 highlights possible LXXLL and LXXLL-like motifs which are known to mediate interaction with p300/CBP and are related to the interaction of HBZ and p300/CBP. Hence, the mechanism of Tax inhibition is likely different from HBZ and should further be considered in the context of

the differences in cellular distribution of all three antisense proteins. These differences could also impact on the cellular genes differently modulated by HBZ, APH-3 and APH-4 and might affect their capacity to modify cellular proliferation or transformation.

The current data demonstrate that antisense transcription is a common mode of expression in HTLV viruses (and likely STLV family members). Recent results have also confirmed that a similar antisense transcript and encoded protein was being produced from the HTLV-2 virus (Halin et al., submitted), a virus which has not been related to any haematological malignancies. The similarity between these transcription patterns and the function of their encoded proteins highlight the importance of antisense-encoded HTLV gene for replication. Furthermore, a link between HBZ and ATL development has been suggested. It is not currently known if HTLV-3 and HTLV-4 are also associated with human diseases and future studies will help in assessing this possibility as well as the possible association of viral proteins (including APH-3 and APH-4). In addition, it will be interesting to determine if other human and non-human retroviruses might also be capable of producing antisense transcripts with a protein-coding capacity.

Acknowledgments

This work was supported by a grant to B.B. from the Cancer Research Society (CRS) Inc. M.H. is supported by an institutional Hydro-Quebec scholarship and S.L. holds a CIHR PhD scholarship. B.B. holds a Canada Research Chair of Canada (Tier 2).

References

1. **Arnold, J., B. Yamamoto, M. Li, A. J. Phipps, I. Younis, M. D. Lairmore, and P. L. Green.** 2006. Enhancement of infectivity and persistence in vivo by HBZ, a natural antisense coded protein of HTLV-1. *Blood* **107**:3976-82.
2. **Arnold, J., B. Zimmerman, M. Li, M. D. Lairmore, and P. L. Green.** 2008. Human T-cell Leukemia Virus Type-1 Antisense-encoded Gene, Hbz, Promotes T Lymphocyte Proliferation. *Blood*.
3. **Barbeau, B., and J. M. Mesnard.** 2007. Does the HBZ gene represent a new potential target for the treatment of adult T-cell leukemia? *Int Rev Immunol* **26**:283-304.
4. **Basbous, J., C. Arpin, G. Gaudray, M. Piechaczyk, C. Devaux, and J. M. Mesnard.** 2003. The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity. *J Biol Chem* **278**:43620-7.
5. **Briquet, S., J. Richardson, C. Vanhee-Brossollet, and C. Vaquero.** 2001. Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus. *Gene* **267**:157-64.
6. **Briquet, S., and C. Vaquero.** 2002. Immunolocalization studies of an antisense protein in HIV-1-infected cells and viral particles. *Virology* **292**:177-84.
7. **Calattini, S., S. A. Chevalier, R. Duprez, P. Afonso, A. Froment, A. Gessain, and R. Mahieux.** 2006. Human T-cell lymphotropic virus type 3: complete nucleotide sequence and characterization of the human tax3 protein. *J Virol* **80**:9876-88.
8. **Calattini, S., S. A. Chevalier, R. Duprez, S. Bassot, A. Froment, R. Mahieux, and A. Gessain.** 2005. Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa. *Retrovirology* **2**:30.
9. **Cavanagh, M.-H., S. Landry, B. Audet, C. Arpin, P. Hivin, M.-È. Paré, J. Thête, E. Wattel, S. J. Marriott, J. M. Mesnard, and B. Barbeau.** 2006. HTLV-I antisense transcripts initiating in the 3' LTR are alternatively spliced and polyadenylated. *Retrovirology* (accepted).

10. **Chevalier, S. A., N. L. Ko, S. Calattini, A. Mallet, M. C. Prevost, K. Kehn, J. N. Brady, F. Kashanchi, A. Gessain, and R. Mahieux.** 2008. Construction and characterization of a human T-cell lymphotropic virus type 3 infectious molecular clone. *J Virol* **82**:6747-52.
11. **Chevalier, S. A., L. Meertens, C. Pise-Masison, S. Calattini, H. Park, A. A. Alhaj, M. Zhou, A. Gessain, F. Kashanchi, J. N. Brady, and R. Mahieux.** 2006. The tax protein from the primate T-cell lymphotropic virus type 3 is expressed in vivo and is functionally related to HTLV-1 Tax rather than HTLV-2 Tax. *Oncogene* **25**:4470-82.
12. **Clerc, I., N. Polakowski, C. André-Arpin, P. Cook, B. Barbeau, and I. Lemasson.** 2008. The HTLV-1 bZIP factor downregulates HTLV-i transcription by targeting the transcriptional coactivators p300/CBP. *PLoS Pathogens*:submitted.
13. **Gaudray, G., F. Gachon, J. Basbous, M. Biard-Piechaczyk, C. Devaux, and J. M. Mesnard.** 2002. The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription. *J Virol* **76**:12813-22.
14. **Halin, M., E. Douceron, I. Clerc, C. Journo, N. G. Ko, S. Landry, E. Murphy, J. M. Mesnard, B. Barbeau, and R. Mahieux.** 2008. Human T-cell Leukemia Virus Type 2 produces a spliced antisense transcript encoding a protein that lacks a bZIP domain but inhibits Tax2-mediated transcription. *Blood* **submitted**.
15. **Hivin, P., J. Basbous, F. Raymond, D. Henaff, C. Arpin-Andre, V. Robert-Hebmann, B. Barbeau, and J. M. Mesnard.** 2007. The HBZ-SP1 isoform of human T-cell leukemia virus type I represses JunB activity by sequestration into nuclear bodies. *Retrovirology* **4**:14.
16. **Lamsoul, I., J. Lodewick, S. Lebrun, R. Brasseur, A. Burny, R. B. Gaynor, and F. Bex.** 2005. Exclusive ubiquitination and sumoylation on overlapping lysine residues mediate NF-kappaB activation by the human T-cell leukemia virus tax oncoprotein. *Mol Cell Biol* **25**:10391-406.
17. **Landry, S., M. Halin, A. Vargas, I. Lemasson, and J. M. Mesnard.** 2008. Upregulation of HTLV-1 antisense transcription by the viral Tax protein. *Journal of Virology* **submitted**.

18. **Matsumoto, J., T. Ohshima, O. Isono, and K. Shimotohno.** 2004. HTLV-1 HBZ suppresses AP-1 activity by impairing both the DNA-binding ability and the stability of c-Jun protein. *Oncogene*.
19. **Matsuoka, M., and K. T. Jeang.** 2007. Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation. *Nat Rev Cancer* **7**:270-80.
20. **Mesnard, J. M., B. Barbeau, and C. Devaux.** 2006. HBZ, a new important player in the mystery of adult T-cell leukemia. *Blood* **108**:3979-82.
21. **Michael, N. L., M. T. Vahey, L. d'Arcy, P. K. Ehrenberg, J. D. Mosca, J. Rappaport, and R. R. Redfield.** 1994. Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat. *J Virol* **68**:979-87.
22. **Miyoshi, I., I. Kubonishi, S. Yoshimoto, T. Akagi, Y. Ohtsuki, Y. Shiraishi, K. Nagata, and Y. Hinuma.** 1981. Type C virus particles in a cord T cell line derived by co-cultivating normal human cord leukocytes and human leukaemic T cells. *Nature* **294**:770.
23. **Murata, K., T. Hayashibara, K. Sugahara, A. Uemura, T. Yamaguchi, H. Harasawa, H. Hasegawa, K. Tsuruda, T. Okazaki, T. Koji, T. Miyanishi, Y. Yamada, and S. Kamihira.** 2006. A novel alternative splicing isoform of human T-cell leukemia virus type 1 bZIP factor (HBZ-SI) targets distinct subnuclear localization. *J Virol* **80**:2495-505.
24. **Peeters, A., P. F. Lambert, and N. J. Deacon.** 1996. A fourth Sp1 site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription. *J Virol* **70**:6665-72.
25. **Poiesz, B., F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna, and R. C. Gallo.** 1980. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc. Natl. Acad. Sci. USA* **77**:7415-7419.
26. **Poiesz, B., F. W. Ruscetti, M. S. Reitz, V. S. Kalyanaraman, and R. C. Gallo.** 1981. Isolation of a new type C retrovirus (HTLV) in primary uncultured cells of a patient with Sézary T-cell leukemia. *Nature* **294**:268-271.

27. **Satou, Y., J. I. Yasunaga, M. Yoshida, and M. Matsuoka.** 2006. HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells. *Proc Natl Acad Sci U S A*.
28. **Switzer, W. M., S. H. Qari, N. D. Wolfe, D. S. Burke, T. M. Folks, and W. Heneine.** 2006. Ancient origin and molecular features of the novel human T-lymphotropic virus type 3 revealed by complete genome analysis. *J Virol* **80**:7427-38.
29. **Tagieva, N. E., and C. Vaquero.** 1997. Expression of naturally occurring antisense RNA inhibits human immunodeficiency virus type 1 heterologous strain replication. *J Gen Virol* **78 (Pt 10)**:2503-11.
30. **Thebault, S., J. Basbous, P. Hivin, C. Devaux, and J. M. Mesnard.** 2004. HBZ interacts with JunD and stimulates its transcriptional activity. *FEBS Lett* **562**:165-70.
31. **Vanhee-Brossollet, C., H. Thoreau, N. Serpente, L. D'Auriol, J. P. Levy, and C. Vaquero.** 1995. A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals. *Virology* **206**:196-202.
32. **Wolfe, N. D., W. Heneine, J. K. Carr, A. D. Garcia, V. Shanmugam, U. Tamoufe, J. N. Torimiro, A. T. Prosser, M. Lebreton, E. Mpoudi-Ngole, F. E. McCutchan, D. L. Birx, T. M. Folks, D. S. Burke, and W. M. Switzer.** 2005. Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters. *Proc Natl Acad Sci U S A* **102**:7994-9.
33. **Yoshida, M., I. Miyoshi, and Y. Hinuma.** 1982. Isolation and characterization of retrovirus from cell lines of human adult T-cell leukemia and its implication in the disease. *Proc. Natl. Acad. Sci. USA* **79**:2031-2035.
34. **Yoshida, M., Y. Satou, J. I. Yasunaga, J. I. Fujisawa, and M. Matsuoka.** 2008. Transcriptional control of spliced and unspliced HTLV-1 bZIP factor gene. *J Virol*.

Figure legends

Figure 1 : Position of the antisense ORF in the HTLV-3 and HTLV-4 genomes.

HTLV-3 2026ND (A) and HTLV-4 HT4v2 (B) molecular clones are depicted with all known genes. The positioning of the putative APH-3 and APH-4 ORF from the antisense strand is indicated below the proviral DNA. A 5' deleted version of these vectors termed pHTLV-3 Δ EcoRV and pHTLV-4 Δ SacI are also presented. The size of the proviral DNA (full length or 5' end deleted version) are indicated for each retrovirus. The primers used for the RT-PCR experiments are also represented.

Figure 2 : Detection of HTLV-3 and HTLV-4 spliced antisense transcripts .

293T cells were transfected either with pHTLV-3 Δ EcoRV (A) or pHTLV-4 Δ SacI (B). RNA was extracted and analyzed by RT-PCR for the presence of antisense transcripts using the primer combinations 21-3/19-1 (A) or 21-1/23-2 (B). Deduced spliced transcript for each proviral DNA are depicted and position of the SD and SA sites are indicated below the spliced transcript (nucleotidic positioning from sense strand). Control (-) = no RT in the reaction; M= 100 bp marker; * indicates 600 bp. (C) Amino acid sequence of the spliced APH-3 and APH-4 RNA next to the spliced junction. The amino acid is shown above each sequence. As a reference, the nucleotidic and amino acid sequences predicted for unspliced transcripts are also provided for comparison.

Figure 3 : Identification of a functional polyadenylation site in HTLV-3 and HTLV-4 antisense transcripts.

RNA samples obtained from 293T cells transfected with either pHTLV-3 Δ EcoRV (A) or pHTLV-4 Δ SacI (B) were analyzed by 3' RLM-RACE. Lane (-): PCR amplification in the absence of any cDNA. M: 100 bp marker. The sequences of the *APH* mRNA and of the 3' polyA tail are shown. In the sequence below, the position of the cleavage site and the presence of both polyA signal and GT-rich sequence are highlighted.

Figure 4 : Amino acid sequence of the APH-3 and APH-4 proteins.

The amino acid sequence of APH-3 and APH-4 were deduced from the 2026ND and HT4v2 strains, respectively. Both sequences were compared to the amino acid sequence of the HBZ-SP1 isoform. The previously described BR2, BR1 and DNA-binding domain (DBD) are indicated above the HBZ sequence, while the LZ domain is in a box. LxxLL and LxxLL-like motif are depicted in shaded boxes.

Figure 5 : APH-3 and APH-4 demonstrate different cellular localizations.

(A) pEGFP-C1-APH3 and pEGFP-C1-APH4 expression vectors were transfected into COS7 cells. Thirty six hours post-transfection, cells observed by confocal microscopy. Images are representative of the entire population of transfected cells. Images on the right side of the panels are presented after merging of GFP and transmission signals. (B) Cell

lysates (50 µg) from 293T cells transfected as in (A) or with the pHBZ-Myc vector were first separated in cytoplasmic (Cyt) and nuclear (Nuc) fractions and were subjected to electrophoresis on a 12 % Bis-Tris gel and analyzed by Western blot with anti-Myc, antibodies. pcDNA3.1/Myc parental vector was used as a negative control (CTL). For APH-3, the asterisk points to the major signal with the expected size.

Figure 6 : Antisense expression is upregulated in Jurkat T cells after activation.

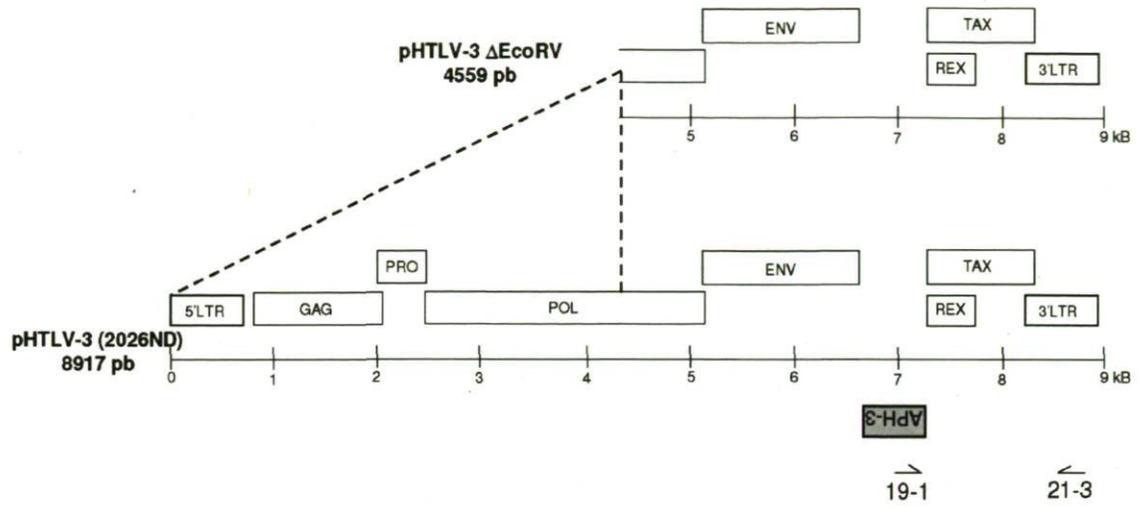
(A) Luciferase reporter gene was inserted in frame with the exon 2 encoding segment of both APH-3 and APH-4 along with a polyA signal cassette at its 3'end generating pHTLV-3-as-luc and pHTLV-4-as-luc. (B) Both luciferase-encoding constructs were transfected in Jurkat T cells and were then activated for 8 hours with PHA, PMA, ionomycin, bpV[pic], forskolin, TNF- α or a combination. Luciferase activity was measured from cell lysates of three independently stimulated samples and are presented as the average fold induction (+/- S.D.) where a value of 1 is given to unstimulated cells. These results are representative of four independent experiments.

Figure 7 : APH-3 and APH-4 repress Tax-1 and Tax-3 dependent transcription from the HTLV-1 LTR.

293T cells were transiently transfected with pHTLV-luc and pCMV-Tax-1 (left panel) or pCMVTax-3 (right panel) together with expression vectors for pMycAPH-3, pMycAPH-4 or HBZ-SP1-Myc (compared to the empty vector pcDNA3.1) along with the pRcActin-lac

Z vector for normalization. The presented luciferase activities are an average of 3 independent transfection experiments and are depicted as the average normalized luciferase activity +/- S.D. These results are representative of three independent experiments

A



B

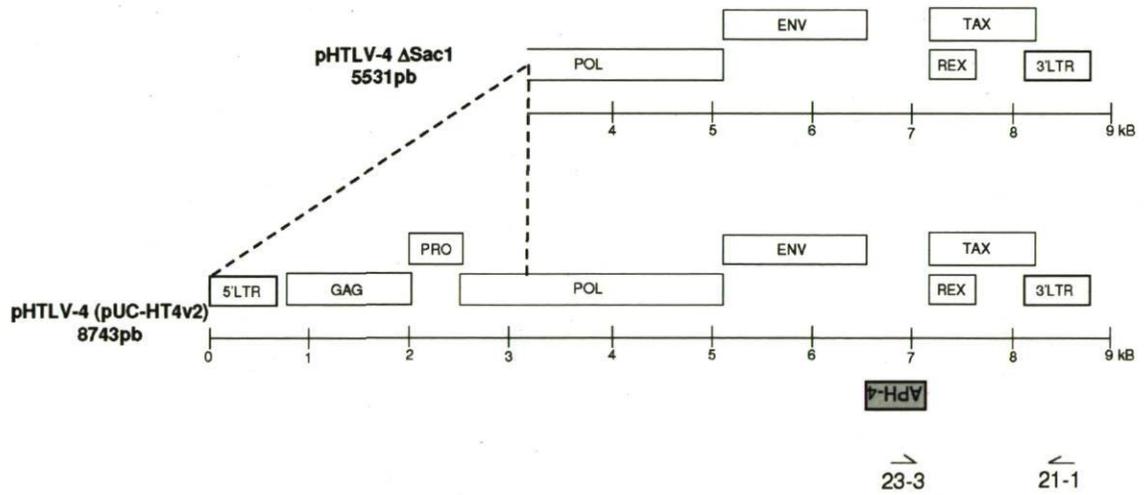


Figure 1

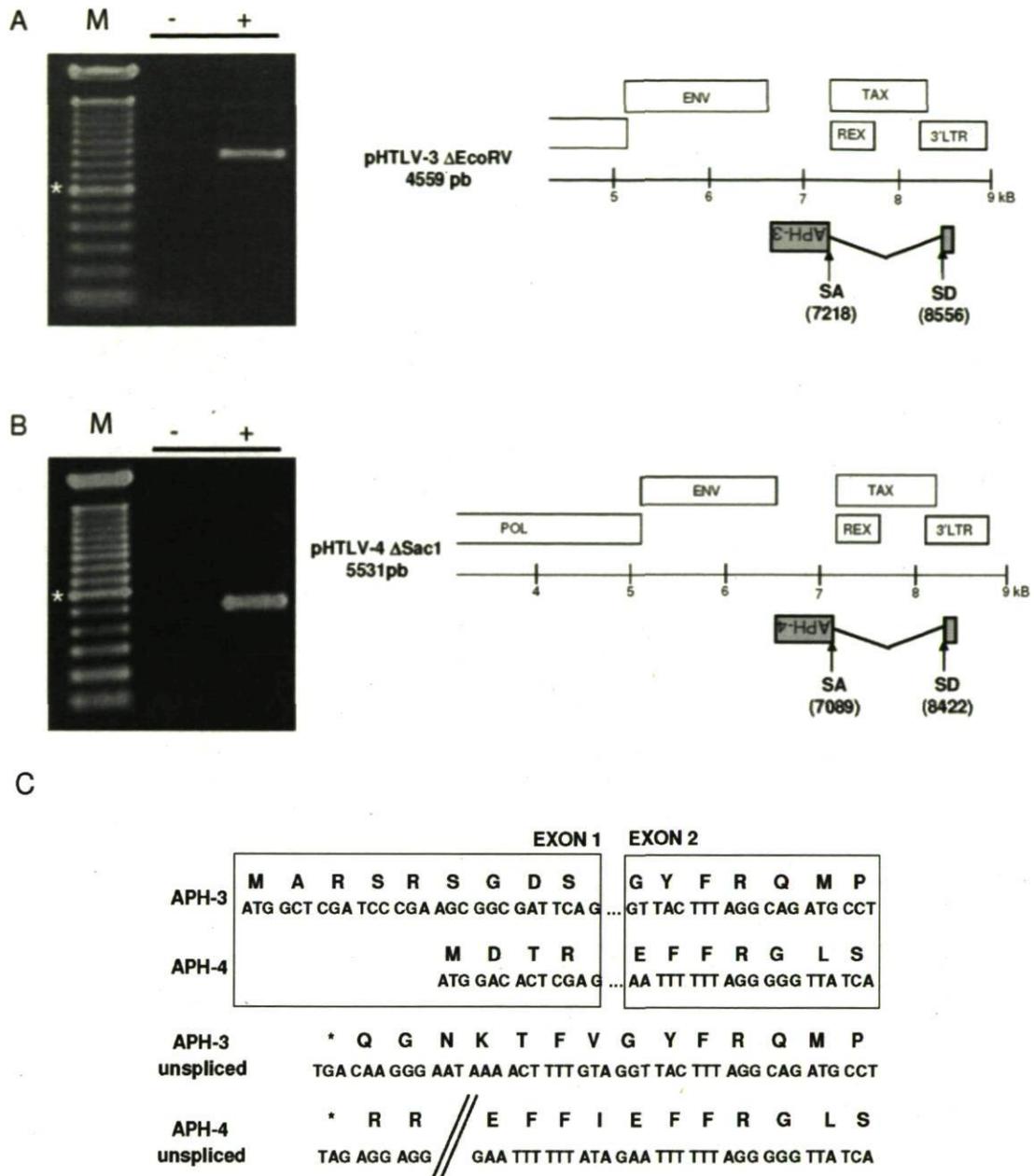


Figure 2

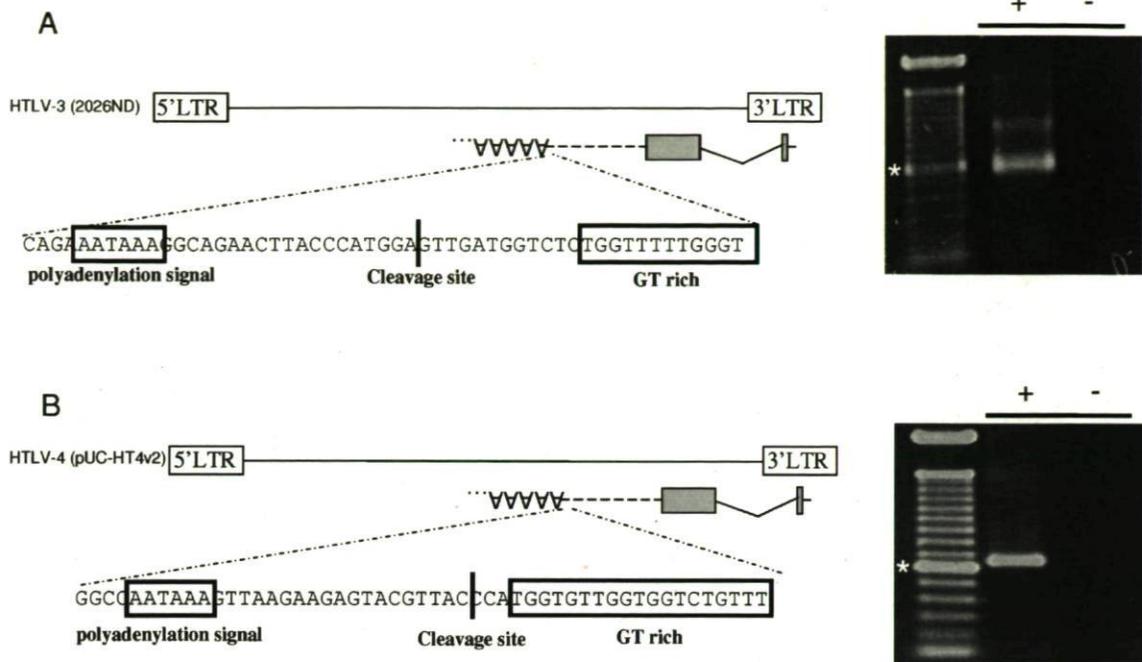


Figure 3

```

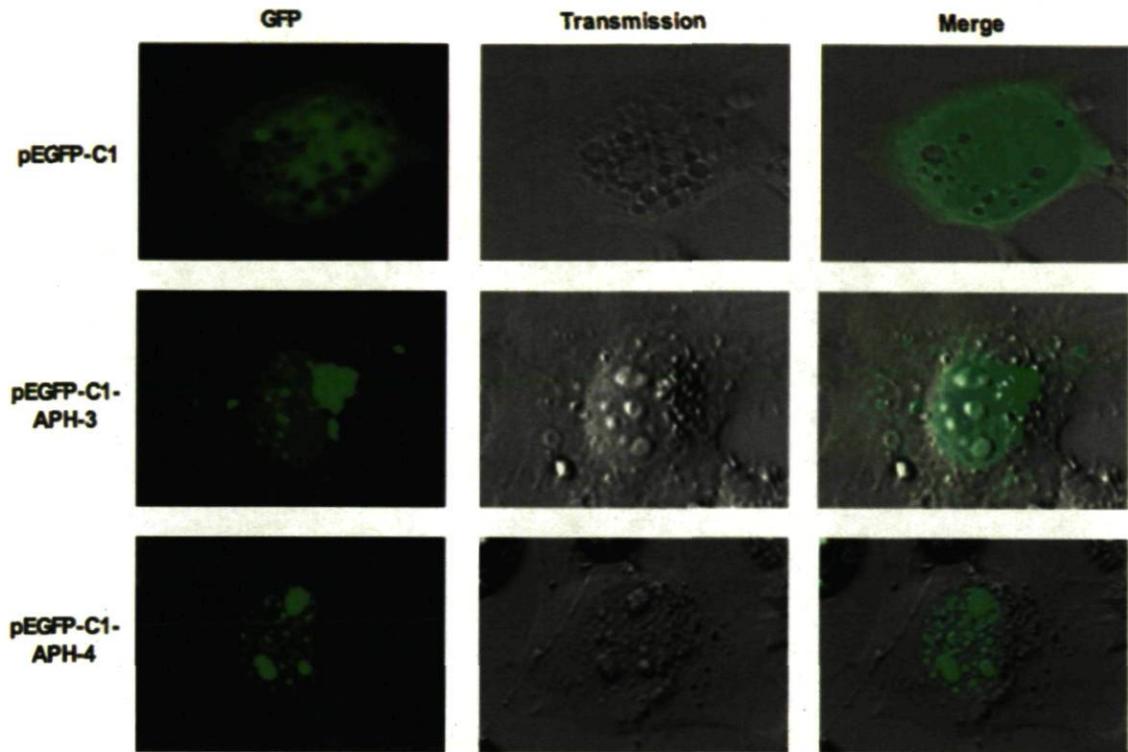
HBZ 1 MAA-----SGLFRCLPVSCPEDLLV-----EELVDGLLSLEEELKDK 37
APH-3 1 MARSRSGDSGYFRQMPVLAPETA|PGTVPVVYSCI STSGGSLTSRSLGPPVLT A 56
APH-4 1 MDT-----REFFRGL----- 10
                                     BR2
HBZ 38 EEEKAVLDGLLSLEEEESRGLRRGPPGEKAPPRGETHRDRQRR A-----EEKRK 86
APH-3 57 EEPLTPVETLLNLCERSGRDQM-----DMLQGNG-KR--PRGRPLGSSSKDLGYWK 104
APH-4 11 -----SSEDRAKKVSTPEEMPGSG-ARRARKGRPLGSGGANDGTRK 50
                                     BR1
HBZ 87 RKKEREKEEEKQIAEYLKRKEEEKARRRRRAEKKAADVARRKQEEQERRERKWRQG 142
APH-3 105 SRLDDLAEQEQQLLSYWRERKKKEERRRRCREEL-----EERKRQKKERRQ 150
APH-4 51 SRLRALVDEQAEIQRLLRQEKEKEARRRERLQER-----ERKRKEREARHQ 96
                                     DBD
HBZ 143 AEKAKQHSARKEKMQLGIDGYTRQLEGEVESLEAERRKLLQEKEDLMGEVNYWQG 198
APH-3 151 QLDC-----IDMLGFEGFCDLLEGYIDFLERESQQLRAGCEESLEEYFWYMD 197
APH-4 97 QRKC-----MDLLGFEGYCDLLQGYIDVLEKERKDL EEECNHNLGELFRLSA 143

HBZ 199 RLE-----AMW-----LQ 206
APH-3 198 CWGQGVDI EGDSSGLC-CTWVSSRGESHSLLLQE VADLL 235
APH-4 144 GLEQTLDI WGDSSGLAECGWVNGRGNHGDLLGV LAELL 182

```

Figure 4

A



B

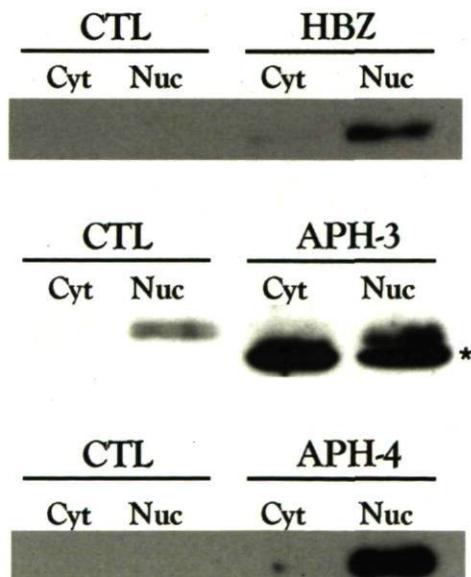


Figure 5

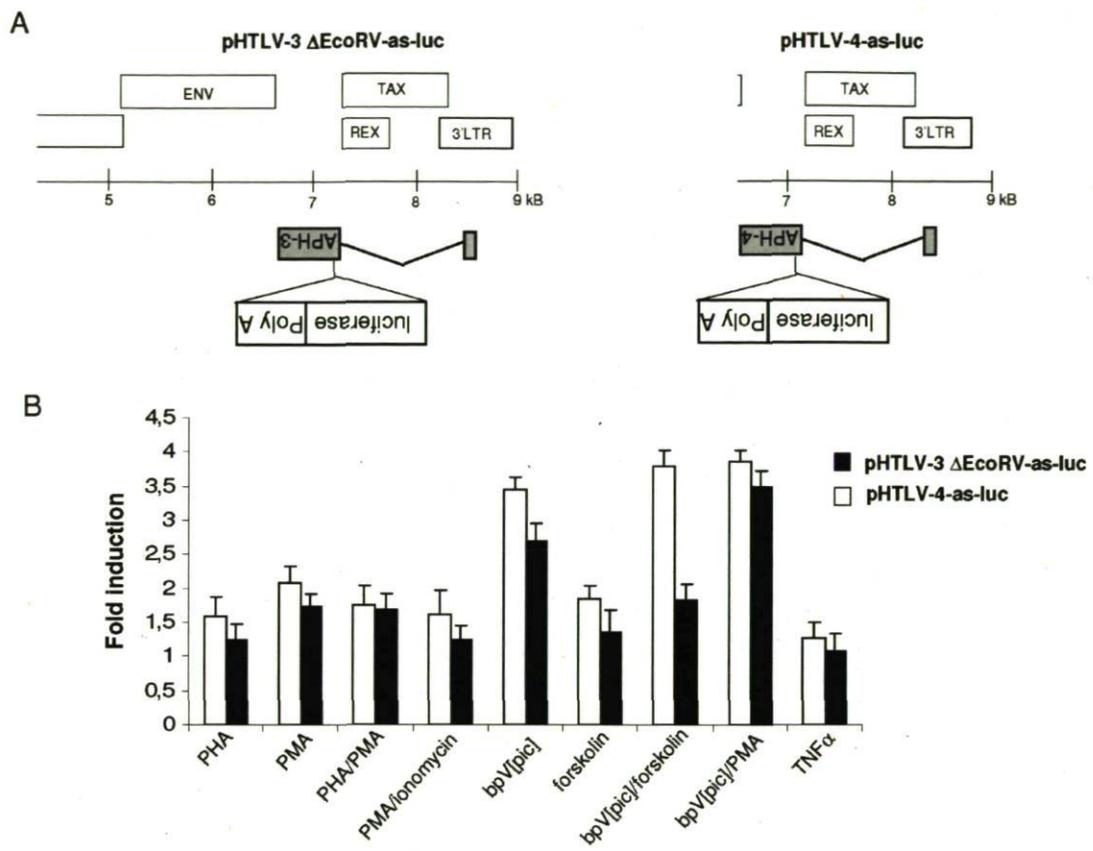
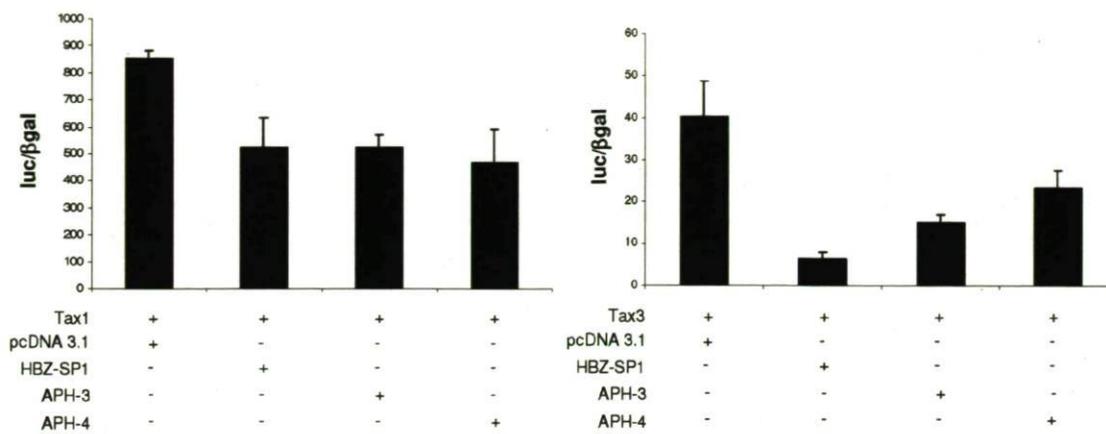


Figure 6

**Figure 7**

Chapitre 6 : Détection, caractérisation et régulation de la transcription antisens chez le VIH-1.

Ce chapitre a fait l'objet d'une publication dans la revue *Retrovirology* (2007, 4:71) et est constitué de la version finale de l'article tel que publié.

6.1 Résumé

Les travaux de plusieurs groupes ont récemment démontré l'existence de transcrits antisens chez HTLV-1. Ces transcrits ont la capacité de coder pour la protéine virale HBZ. L'objectif de la présente étude était de mettre en évidence l'existence de transcription antisens chez le VIH-1 et de mieux caractériser les transcrits. Des expériences initiales effectuées par RT-PCR à partir des cellules chroniquement infectées J1.1 et des cellules 293T transfectées avec un ADN proviral ont permis de confirmer l'existence de transcription antisens. Le développement d'un protocole de RT-PCR permettant d'éviter les artefacts associés à l'amorçage endogène a par la suite permis de détecter la présence de transcrits antisens dans plusieurs lignées infectées. Par des analyses de RACE, plusieurs sites d'initiations de la transcription ont pu être identifiés en bordure du LTR 3'. De façon intéressante, un nouveau signal fonctionnel de polyadénylation présentant la séquence consensus a pu être identifié dans le brin complémentaire du génome viral. Des expériences de transfection effectuées dans les cellules 293T et Jurkat à l'aide d'une construction d'ADN proviral contenant le gène luciférase en orientation antisens ont permis d'observer la présence de signaux luciférase inductibles par des activateurs de cellules T. La

génération de lignée stables avec cette construction a permis de mettre en évidence l'effet positif de Tat sur la transcription antisens. Des expériences d'infection effectuées dans les cellules 293T, Jurkat, SupT1, U937 et CEMT4 en utilisant des virions pseudotypés exprimant le gène rapporteur luciférase à partir du transcrit antisens ont permis d'obtenir des signaux luciférase qui se sont cependant avérés être plus faibles que ceux obtenus à partir des cellules infectées par des virions NL4.3 Luc. Ces résultats démontrent pour la première fois que la transcription antisens existe dans le contexte de l'infection chez le VIH-1. La possibilité que le transcrit antisens du VIH-1 puisse donc exprimer une nouvelle protéine virale devrait donc être réévaluée.

6.2 Article

Detection, characterization and regulation of antisense transcripts in HIV-1

Sébastien Landry¹, Marilène Halin¹, Sylvain Lefort², Brigitte Audet², Catherine Vaquero³
Jean-Michel Mesnard⁴ and Benoit Barbeau^{1*}

¹*Université du Québec à Montréal, Département des sciences biologiques, Montréal (Québec) Canada, H2X 3X8*²; *Centre de Recherche en Infectiologie, Centre Hospitalier Universitaire de Québec, Pavillon CHUL, and Département de Biologie médicale, Faculté de Médecine, Université Laval, Ste-Foy (Québec), Canada G1V 4G2*³; *CHU-Pitié-Salpêtrière, Paris, France*⁴; *Laboratoire Infections Rétrovirales et Signalisation cellulaire, CNRS/UMI UMR 5121/IFR 122, Institut de Biologie, 34960 Cedex 2, Montpellier, France*

*Corresponding author: Mailing address:

Département des sciences biologiques

Université du Québec à Montréal

Rm SB-R860

2080 St-Urbain

Montréal (Québec) Canada

H2X 3X8

Phone: (514) 987-3000 (#4576). Fax: (514) 987-4647

E-MAIL: barbeau.benoit@uqam.ca

E-MAIL addresses of authors: sebastien.landry@crchul.ulaval.ca;

halin.marilene@courrier.uqam.ca; sylvain.lefort@crchul.ulaval.ca;

brigitte.audet@crchul.ulaval.ca; vaquero@ext.jussieu.fr; jean-michel.mesnard@univ-montpl.fr

Abstract

Background: We and others have recently demonstrated that the human retrovirus HTLV-I was producing a spliced antisense transcript, which led to the synthesis of the HBZ protein. The objective of the present study was to demonstrate the existence of antisense transcription in HIV-1 and to provide a better characterization of the transcript and its regulation.

Results: Initial experiments conducted by standard RT-PCR analysis in latently infected J1.1 cell line and pNL4.3-transfected 293T cells confirmed the existence of antisense transcription in HIV-1. A more adapted RT-PCR protocol with limited RT-PCR artefacts also led to a successful detection of antisense transcripts in several infected cell lines. RACE analyses demonstrated the existence of several transcription initiation sites mapping near the 5' border of the 3'LTR (in the antisense strand). Interestingly, a new polyA signal was identified on the antisense strand and harboured the polyA signal consensus sequence. Transfection experiments in 293T and Jurkat cells with an antisense luciferase-expressing NL4.3 proviral DNA showed luciferase reporter gene expression, which was further induced by various T-cell activators. In addition, the viral Tat protein was found to be a positive modulator of antisense transcription by transient and stable transfections of this proviral DNA construct. RT-PCR analyses in 293T cells stably transfected with a pNL4.3-derived construct further confirmed these results. Infection of 293T, Jurkat, SupT1, U937 and CEMT4 cells with pseudotyped virions produced from the antisense luciferase-expressing NL4.3 DNA clone led to the production of an AZT-sensitive luciferase signal, which was however less pronounced than the signal from NL4.3Luc-infected cells.

Conclusions: These results demonstrate for the first time that antisense transcription exists in HIV-1 in the context of infection. Possible translation of the predicted antisense ORF in this transcript should thus be re-examined.

Background

It has been largely accepted that gene expression in retroviruses solely relies on a single transcript, which is in turn either left unspliced, singly or multiply spliced. This transcript is initiated from the 5' LTR region, which harbours in its U3 segment most of the necessary binding sites for important transcription factors regulating the expression of retroviral genes. In addition, for all studied retroviruses, this transcript initiates at a single position and is typically dependent on an upstream TATA box. Few studies have addressed the possible existence of transcripts initiated at other position in the retroviral genome. A number of reports have however provided an interesting and unexpected possibility to retroviral gene expression. Indeed, in a few complex retroviruses including HIV-1, FIV-1 and HTLV-I, it has been suggested that transcripts produced in the antisense direction exist and that these transcripts could have the potential to encode for a protein (1-4). Although these results have been debated and largely contested, new results obtained with the HTLV-I virus have importantly revived the issue over antisense transcription (2, 5-15). Indeed, the HTLV-I retrovirus has been the first retrovirus from which the existence of an antisense transcript has been clearly demonstrated. Recent studies have further highlighted the spliced nature of this transcript (12-14). The HBZ protein encoded from this transcript was shown to have AP-1 and Tax inhibitory activity and to be detected in infected cell lines as well as PBMCs from HTLV-I infected individuals.

The existence of antisense transcription in HIV-1 has been similarly suggested based on the identification of a conserved ORF in the antisense strand of its genome. Hence an initial study by Miller had identified a well conserved ORF of 189 amino acids,

later termed ASP (Antisense Protein) on the antisense strand, which was generally well conserved in all analysed HIV-1 proviral DNA (3). Analysis of the hydrophobic profile of the potentially encoded protein revealed it to be highly hydrophobic and thus to possibly be associated to the membrane. Detection of the ASP protein has only been possible through Western blot analysis of bacterially produced ASP and electron microscopy studies (16). Despite these studies, no functions have yet been assigned to this potential virally encoded protein and its existence remains controversial.

Studies have however been more focussed on the detection of the antisense transcript itself in HIV-1. The existence of the transcript has been previously suggested through Northern blot and RT-PCR approaches (4, 17, 18). Studies based on the identification of the 5' and 3' ends of the antisense transcript have also been performed and suggested that this transcript was initiating next to the 3' LTR border, although no consensus was obtained (4, 19, 20). Promoter analyses have been further conducted by using the isolated 3' LTR positioned in the antisense orientation and T-cell activators were shown to positively modulate promoter activity while Tat had an adverse effect (4, 19, 21). Although these analyses have tended to infer that this pattern of expression was occurring in HIV-1, numerous artefacts and contradictory results have not permitted to unequivocally demonstrate that indeed HIV-1 antisense transcription existed. Therefore reassessment of antisense expression is directly needed to readdress the existence of antisense transcription in HIV-1.

In virtue of the recent results on HTLV-I antisense transcription, the goal of this study was to readdress the existence of antisense transcription in HIV-1. Using an antisense transcription-specific RT-PCR approach (with no non-specific RT priming

artefact) and an HIV-1 proviral DNA construct expressing the luciferase gene in the inverse orientation, we provided for the first time strong evidence demonstrating the presence of HIV-1 antisense transcripts. Our data also highlight the existence of a new polyA signal in the antisense strand and strongly support a positive role for Tat on antisense transcription. These results add new important information, which will likely impact on the understanding of HIV-1 replication.

Results

Detection of the antisense transcript in infected and transfected cells. Previous studies had earlier suggested that antisense transcription could be detected through RT-PCR analyses (4). However, in our hands, these protocols were not suitable for specific detection of antisense transcription as substantial amount of non-specific signals due to endogenous RT priming was apparent. Endogenous RT priming results from priming of RNA by small degraded RNA or DNA fragments present in the extracted RNA pool, which act as primers during the reverse transcriptase step. Given that sense expression in retroviruses has been suggested to be more prominent than antisense transcription, a PCR signal might thus be overwhelmingly derived from cDNA produced from sense mRNA primed by degraded DNA/RNA and not permit to specifically assess the existence of antisense transcripts. Endogenous RT priming is typically controlled by conducting PCR amplification of cDNA produced from RNA in the presence of the reverse transcriptase but in the absence of the RT primer. To detect the antisense transcript, we chose RT and PCR primers in the proviral DNA region located in the ASP ORF. As presented in Figure 1, the ASP ORF is located in the antisense strand in the *env* gene. Several primers were designed to provide signals of different sizes.

Our first RT-PCR analysis was conducted using standard conditions. However, as indicated above, we expected that most of endogenous RT priming artefacts in HIV-1-infected cells might be coming from cDNA synthesis from the sense transcript occurring through the presence of degraded HIV-1 cellular DNA or antisense RNA. To decrease this important source of endogenous RT priming artefact, we first used the J1.1 cell line, which

is latently infected and produces very low amounts of virions when left unstimulated. RNA extracted from this cell line was used for cDNA synthesis with an antisense RNA-specific primer located in the ASP ORF region and PCR was then conducted with two different sets of ASP ORF-derived primers. As presented in Figure 2A, specific signals representing antisense transcription were detected (lanes 6 and 7). The sequencing of these signals confirmed their specificity. Importantly, controls for endogenous RT priming (absence of primer at the RT step) (lanes 4 and 5) and for DNA contamination (no RT step) (lanes 1 to 3) were devoid of any signal further demonstrating that our PCR-amplified fragments were specific to the antisense transcript. Antisense transcription was next tested in 293T cells transfected with HIV-1 proviral DNA. The NL4.3 proviral was thus chosen and, as argued above, a 5' LTR-deleted version termed pNL4.3 Δ NarI was generated to minimize endogenous RT priming on sense transcripts. Wild-type and 5' LTR-deleted pNL4.3 constructs were thus transfected in 293T cells and RT-PCR analyses (as described above) were undertaken on RNA samples from these cells (Figure 2B). Again, a specific signal was easily detected in PCR amplification of cDNA originating from 293T cells transfected with pNL4.3 Δ NarI (lanes 6 and 7) and its specificity was further demonstrated by sequencing. Controls for DNA contamination or endogenous RT priming indicated no contaminating signals (lanes 1 to 5). As opposed to these results, pNL4.3wt-transfected cells clearly demonstrated the presence of endogenous RT priming, which had a comparable intensity to the signal obtained in the presence of the RT primer (compare lanes 10 and 11 versus 12 and 13 respectively).

We were next interested in demonstrating the presence of the antisense transcript in chronically infected cells. As we have demonstrated that sense transcription would likely

be an important source of endogenous RT priming artefact masking the antisense RNA-specific signal, we thus optimized RT-PCR conditions, which would greatly diminish non-specific signals (see Methods). Tested infected cells included OM10.1, ACH-2, J1.1 and U937 HIV-1_{III}B (Figure 3 and data not shown). Although the three first cell lines produces low amounts of infectious particles, the U937-infected cell line is known to be a source of substantial levels of produced infectious particles. The improved RT-PCR protocol consisted of an extraction of mRNA followed by an RT step with a primer containing a 3' end complementary to the antisense transcript in the ASP region and a non-complementary 5' end. To remove the RT primer from the reaction, cDNAs are then purified through the use of a column (DNA cleanup step). PCR is then performed with a forward primer again derived from the ASP region and a reverse primer termed the anchor primer specific to the 5' extremity of the RT primer. This RT-PCR approach hence strongly favoured PCR amplification of RT primer-derived cDNAs. Using this approach, we first tested RNA samples from 293T cells transfected with either pNL4.3wt or pNL4.3ΔNarI. As depicted in Figure 3, this approach permitted the detection of the antisense transcript in both transfected cell lines using two different primer sets (lanes 5 and 6). In these experiments, we have also controlled for the specificity of the forward primer used in PCR amplification (primer 30-20 specific to the non-complementary end of the 24-6F RT primer) by testing cDNAs produced with added 24-6 RT primer devoid of the targeted sequence of this forward primer (lanes 1 and 2). Another crucial control consisted of ensuring that no 24-6F RT primer remained in the cDNA reaction after column purification. Sufficient amount of the RT primer during PCR amplification might allow subsequent amplification with the PCR forward primer 30-20 thereby amplifying any

source of HIV-1 DNA. Hence, purified cDNA prepared from the RT primer 24-6 was incubated in the presence of an aliquot of mock prepared cDNA from non-transfected 293T cells after column purification (lanes 3 and 4). As expected, none of these controls led to a PCR-amplified signal. With this approach, we next tested the infected cell lines listed above. We indeed demonstrate that this protocol allowed us to specifically detect the antisense transcript in these infected cell lines (lanes 5 and 6) while no undesirable contaminating signals were detected (lanes 1 to 4). None of the RNA samples tested through this protocol had contaminating DNA (data not shown).

These results hence demonstrated the existence of an antisense transcript in HIV-1, which included the ASP ORF sequence. The use of HIV-1 proviral DNA clones and of infected cell lines suggested that a wide range of HIV-1 strains are capable of producing this transcript.

Identification of multiple transcription initiation sites for the HIV-1 antisense transcript. Identification of transcription initiation sites of the antisense transcript of HIV-1 was next undertaken. As an important and specific signal was detected by RT-PCR in 293T cells transfected with pNL4.3 Δ NarI, these cells were used as the source of RNA to conduct 5'RACE analyses. In these analyses, PCR amplification was conducted with reverse primers positioned near the 5' end of the ASP ORF region and primers specific to the oligonucleotide ligated to the 5' end of RNAs (as described in Methods). Migration of the 5'RACE products initially indicated the presence of potential multiple transcription initiation sites. Cloning and sequencing of all amplified products generated by 5' RACE (Figure 4) indeed confirmed that numerous transcription initiation sites were identifiable

for the HIV-1 antisense transcript and were positioned near the 5' border of the 3' LTR and at more downstream region (in the antisense strand). The majority of the transcription initiation sites were located in a 462 bp region encompassing the transcription initiation site previously described by Peeters *et al.* (1996) (19). Interestingly, the 3' part of the 462 bp region (containing the previously identified CAP site) presented numerous transcription initiation sites identified repeatedly by numerous sequenced PCR amplicons,

These results hence demonstrated that the HIV-1 antisense transcript initiated next to the 3' LTR at multiple positions (in a region ranging in length from 700 to 1250 bp). This multiplicity of initiation sites might be a consequence of the absence of a TATA box at close distance.

Identification of a new polyA signal in the antisense strand. Previous results had suggested that non-consensus polyA signals existed at the 3' end of the ASP ORF and were likely determinant for the polyadenylation of the antisense transcript (4). We next searched for the 3' end of the antisense transcript by conducting 3'RACE analyses, again using RNA from pNL4.3 Δ NarI-transfected 293T cells. Through this approach, forward primers were initially designed 500 bp upstream of the ASP stop codon in the antisense strand. However, no signals reminiscent of the presence of a polyA tail being added to the antisense transcript at proximity to this region were detected (data not shown). We thus searched for a potentially new polyA signal and found an AATAAA consensus sequence at position 4908 of the pNL4.3 molecular clone (sense strand positioning) in the complementary sequence corresponding to the *pol* gene (Figure 5B). Comparison of this polyA consensus sequence among different HIV-1 strains revealed a high degree of

conservation and further demonstrated its close localisation to a downstream GT-rich sequence, another marker for polyA addition (22) (Figure 5C). We thus repeated our 3' RACE analyses using the RNA sample from the same transfected 293T cells and used a forward primer at a distance closer to this potential polyA signal. After amplification, a specific signal was detected with a size expected for a polyA tail being present near the polyA signal (Figure 5A). Cloning and sequencing of the 3'RACE amplified products further confirmed that the polyA signal was next to the position of the added polyA tail (at a 19 nucleotide distance) (Figure 5C).

These results hence confirmed that the antisense transcript was polyadenylated and a newly identified and well conserved polyA signal located at 2.4 kb distance from the ASP stop codon was likely essential for the addition of the polyA tail.

Modulation of HIV-1 antisense transcription using an antisense luciferase-expressing proviral DNA. Our RT-PCR approach represents an important tool in the detection of a specific signal for antisense transcription. However, the quantification of changes in antisense transcription levels in the proviral DNA context remained an essential element. Sense transcription and HIV-1 infection has been studied by several research groups using proviral DNA constructs containing the luciferase reporter gene inserted in the *nef* gene in frame with its ATG initiation codon (HXB-Luc and NL4.3LucR-E-) (23,24). As these proviral DNA constructs produce virions upon transfection, we used the pNL4.3LucR-E- construct and removed its luciferase reporter gene to reinsert this reporter gene in the same position but in the inverse orientation. The cloning of the luciferase reporter gene in the antisense direction at this position permitted the usage of the transcription initiation sites

identified above (presented in Figure 4). This vector, termed pNL4.3AsLucR-E- was tested in Jurkat and 293T cells and upon transfection, luciferase activity was constantly measured in both cell lines (293T: 893.7 ± 30 RLU and Jurkat: 2.2 ± 0.3 RLU versus 0.2 RLU as a blank value) (Figure 6A-B).

As previous results had suggested that antisense transcription was positively modulated by T-cell activators (4, 19, 21), we next tested our antisense luciferase-expressing proviral DNA clone for its response to T-cell activators upon transfection in Jurkat cells (Figure 6B). Our initial results confirmed previous data in that NF- κ B-activating agents such as PMA and PHA slightly but significantly induced luciferase activity in transfected Jurkat cells. We next tested a version of this construct from which the 5' LTR was deleted to assess if blocking of sense transcription could impact on the level of induced antisense transcription (Figure 6C). This construct termed pNL4.3 Δ BstAsLucR-E- was transfected in Jurkat cells, which were subsequently treated with a wider range of T-cell activators. Measurement of luciferase activity from these transfected cells indeed revealed that the 5'LTR-deleted version was more potent in its response, especially when comparing the responses toward PHA and PMA activation.

These results hence demonstrated that quantification of luciferase activity can be achieved using our pNL4.3AsLucR-E- construct and that, in the context of this proviral DNA, we were able to confirm induction of antisense transcription by T-cell activators. Furthermore, responses toward these T-cell activators seemed to be modulated by sense transcription.

Upregulation of antisense transcription by the viral Tat protein. Previous studies from other groups had hinted on the possible adversary effect of Tat expression on antisense transcription, although these data might have been artefactual (4, 19, 21). We thus readdressed these findings to provide a clearer role of Tat on antisense transcription using our antisense luciferase-expressing NL4.3 clone. As an initial step, we first looked at Tat modulation on a construct containing a complete LTR cloned upstream and in the inverse orientation of a luciferase reporter gene. Co-transfection experiments with this construct and a Tat expression vector (or the empty vector) along with the β -gal expression vector in the Jurkat cell line were performed and normalized luciferase activity was subsequently measured. As demonstrated in Figure 7A, Tat expression importantly reduced luciferase activity as previously shown. It is possible that Tat might importantly induce TAR-dependent sense transcription from the 3'LTR (and thus in the inverse orientation from the luciferase gene), which could lead to interference on antisense transcription. To evaluate this possibility, the antisense pLTRXLuc construct was linearized before transfection and then evaluated for its Tat response. As shown in Figure 7A, luciferase activity in this construct was now modulated positively by the viral Tat protein, although linearization led to an important reduction in basal antisense transcription. These results indeed suggested that Tat was rather a positive modulator of antisense transcription.

We next looked to confirm these data in the context of our luciferase-expressing proviral DNA. In order to determine the basal level of antisense transcription in the absence of Tat protein, we decided to use the 5'LTR-deleted pNL4.3 Δ BstAsLucR-E-, which cannot produce Tat. The luciferase signal from this vector was further compared between a transfected circular form versus its linear form following digestion in the *gag*

gene (thereby further eliminating possible interference from a full-circle sense transcript initiated from the 3'LTR). Upon co-transfection of pActin-LacZ, pNL4.3 Δ BstAsLucR-E and pCMV-tat (or the empty control vector) in 293T cells, normalized luciferase activity was determined. In Figure 7B, results of this transfection experiment are presented and indeed argue for a positive Tat-dependent modulation of antisense transcription in the context of the linearized proviral DNA, whereas a more reduced positive effect of Tat on antisense transcription was noted with the circular form of the plasmid. We then prepared stable 293T cell clones by transfection of linearized pNL4.3AsLucR-E (similarly digested) along with pCMV-Hyg and after selection, pooled clones were co-transfected with pActin-LacZ and pCMV-tat (or the empty control vector) (Figure 7C). Luciferase read-outs confirmed that our proviral DNA construct was still responsive to Tat expression in that Tat could upregulate luciferase gene expression in the antisense strand in the chromatin context by more than 2.5 folds. When isolated clones were similarly transfected, all clones responded positively to Tat expression although at different levels (Figure 7C). Fold induction were observed to range between 2 to 17 folds in their response to Tat expression.

We also tested if Tat modulation of antisense transcription could be demonstrated at the RNA level. We initially transfected linearized pNL4.3 Δ NarI in 239T cells and selected for stable transfectants, from which no HIV-1 sense expression should be detected. The resulting selected pool was then transfected with pCMV-tat (or the control vector) and pActin-LacZ. Cells transfected with comparable efficiency (as determined by β -gal activity) were then analysed by RT-PCR. In the absence of 5'LTR and 3'LTR sense transcription, the RT reaction was initially conducted with random primers (Figure 7D). PCR signals demonstrated that Tat expression in the context of this deleted proviral DNA

construct was positively affecting antisense transcription (compare intensities between lanes 7 and 8). The actin signal was identical in both transfected cells (lanes 5 and 6) while DNA contamination controls presented no signals (lanes 1 to 4). To ascertain of the antisense strand specificity of the signal, cDNAs were also synthesized from this RNA with the 24-6F RT primer and subsequently analysed for antisense transcription using the protocol described in Figure 3 and the combination of two different PCR primer sets. As presented in panel E, both primer sets indicated an important increase in the PCR signal attributed to antisense transcription upon Tat expression (compare lanes 7 and 8 versus 3 and 4, respectively). The presence of the 24-6F primer in the PCR reaction was again controlled for in lanes 1-2 and 5-6, which showed no clear signal indicative of this primer contamination in the PCR reaction.

The body of results presented above illustrates that the HIV-1 Tat protein acts positively on antisense transcription in the context of a 5'LTR-deleted proviral clone.

Detection of luciferase activity following infection of cells by antisense luciferase-expressing virions. Given that the pNL4.3AsLucR-E- vector was constructed by keeping the structure of the parental vector intact, we were thus confident that this new vector should allow us to produce virions, which could be pseudotyped and used to study antisense transcription during infection. Hence, pNL4.3AsLucR-E- and the parental pNL4.3LucR-E- were separately co-transfected with a VSV-G expression vector in 293T cells. Harvested supernatants indicated that high levels of p24 capsid proteins were detectable. Subsequently, an identical quantity of both pseudotyped virions (p24 levels) was used to infect 293T cells (Figure 8 A-B). Luciferase activity was detected for both

types of virions above levels measured in non-infected cells (luciferase activity is indicated above each column). Interestingly, and as expected, luciferase activity was importantly lower in 293T cells infected with NL4.3AsLucR-E virions as compared to cells infected with NL4.3LucR-E- virions. However, it was noted that for both cell lines, a continuous increase in luciferase activity was apparent for both viruses between day 1 and 2 and levelled off at day 3. Further experiments using AZT revealed the specificity of the signal in that pre-treatment of 293T cells with AZT importantly reduced the luciferase signal obtained from infection of either virus (Figure 8B). Additional infection experiments with NL4.3AsLucR-E- virions were equally conducted in other cell lines, which included T-cell lines CEMT4, SupT1 and Jurkat and the monocytic U937 cell line. These infection experiments demonstrated that, albeit at a lower extent, antisense luciferase reporter gene expression was detected in all tested cell lines above levels from uninfected cells, although important differences with NL4.3LucR-E-infected cells remained (Figure 8C). Jurkat cells infected with NL4.3AsLucR-E- virions were also tested for several T-cell activators and, as demonstrated through our transfection experiments (Figure 6), induction of luciferase expression by these agents was demonstrated, being again optimal with the bpV[pic]/PMA combination (Figures 8D).

These results thus indicated that our antisense luciferase-encoding proviral DNA can produce virions allowing quantification of antisense expression by luciferase activity in several cell types and at different time points post-infection.

Discussion

Antisense transcription in retroviruses has remained controversial for several years. We and others have recently demonstrated that this pattern of transcription was existent in HTLV-I and further experiments have convincingly shown its coding potential and the important role played by the HBZ protein (2, 5-15). For HIV-1, antisense transcription has been poorly studied and these studies have not convinced the HIV-1 research community over the existence of both antisense transcripts and the ASP protein (4, 16, 18, 19). In this study, our goal was to readdress antisense transcription in HIV-1. Our results demonstrate for this first time that indeed antisense transcription does exist in HIV-1, likely involves a newly described antisense polyA consensus sequence and is positively modulated by T-cell activators and the viral Tat protein. We further present data on a new proviral DNA construct, which should allow us to quantify antisense expression in different cell types and under different conditions.

We have first set out to detect the antisense transcript in a specific fashion. Although previous studies had presented evidence based on RT-PCR analyses supporting its existence (4, 18), endogenous RT priming has been a major drawback in fully acknowledging the existence of antisense transcription in HIV-1. Our first strategy to avoid endogenous RT priming was to study antisense transcription in the context where sense transcription was greatly reduced, as these transcripts are likely the major source of non-specificity masking antisense transcript-specific signals. J1.1 cells and the pNL4.3 Δ NarI proviral DNA have allowed us to achieve this goal and to detect antisense transcription with a limited presence of endogenous RT priming. Indeed, an important

signal of the expected size was obtained in both conditions and for pNL4.3, the presence of the 5'LTR was, as expected, resulting in endogenous RT priming. Using a more optimized RT-PCR approach, we have been successful in detecting an antisense RNA-derived PCR signal in cell lines in which sense transcription was more active. Our approach uses mRNA to avoid contaminating DNA (also in small fragments) and, through the use of a semi-complementary RT primer, strongly favours amplification of cDNAs synthesized from this primer and not from contaminating small RNA fragments. Such a comparable approach has been previously used successfully for the detection of antisense transcription in the Hepatitis C Virus (25). An interesting observation from our data is that U937-infected cells generated a strong signal. Whether this is only specific to the cell line or representative of infected monocytic cells remain to be determined.

Upon characterization of the antisense transcript, we have demonstrated that multiple transcription initiation sites existed for antisense transcripts. It is however worth noting that these initiation sites mainly occur near the 5' border of the 3' LTR, which differs from what has been demonstrated for HTLV-I. However, one transcription initiation site agrees with the results obtained by Peeters *et al.* (1996) (19), while our current data do not support results from other groups including a recent study (4, 20). Protocols and tested cell lines might explain these discrepancies and further experiments will be needed to confirm our data on the transcription initiation sites. However, we have mapped several transcription initiation sites in this region and have confirmed these data in several experiments using a near full length proviral DNA construct (i.e. pNL4.3 Δ NarI).

In our study, unlike HTLV-I antisense transcripts, no clones from our 5' RACE analyses indicated the presence of splicing events occurring in the 5' end of the antisense

transcript. In addition, numerous RT-PCR experiments with primers designed from proviral DNA spanning the ASP and 3'LTR region have been undertaken and none of these experiments conducted in 293T cells transfected with different wild-type or deleted proviral DNA or HIV-1 infected cell lines has permitted to provide evidence for the existence of splicing at the 5' end of the antisense RNA. Hence, these results might justify the closer positioning of the transcription initiation sites to the ATG initiation codon of the potential ASP ORF (as compared to HTLV-I HBZ initiation codon) in order to permit sufficient translation efficiency.

Our characterization of the transcript has also permitted the identification of a new HIV-1 polyA signal. This signal is different from the previously suggested polyA signal in the study of Michael *et al.* (1994) (4). However, we have confirmed the potential usage of our polyA signal by 3' RACE analysis. In addition, this sequence is conserved among different strains and obeys to the consensus sequence. Furthermore, a GT rich sequence is present in the neighbouring DNA region, again a potent marker for positioning of polyA tail addition in the transcript. We are thus confident that this is the polyA signal used for antisense transcription termination. The 2.4 kb length of the presumed 3'UTR suggest that this region of the transcript might play a role in the stability of the transcript, although it still remains to be determined if splicing events might be occurring at this 3' end. By Northern blot analysis, we have not been able to detect the expected 4.1 kb size of the antisense transcript in HIV-1-infected cell lines or 293T cells transfected with proviral DNA. This is likely reminiscent of low levels of the antisense transcript, which is well illustrated in our infection experiments conducted with NL4.3LucR-E- and NL4.3AsLucR-E virions. Indeed, levels of antisense transcript in these experiments are 30 to 1000 folds

lower than that of sense transcripts. However, based on our 5' and 3' analyses, the Northern blot signals initially attributed to antisense transcription presented by Bukrinsky and Etkin (1990) (17) do not correspond to the expected size and furthermore these results had not subsequently been confirmed by other groups.

An important contribution from our present study is the generation of our pNL4.3AsLucR-E- construct. By transfection experiments, we have indeed demonstrated that luciferase activity was being detected and that this vector permitted the production of infectious pseudotyped virions, which produced a luciferase signal for at least three days. The construct also allowed us to confirm previous results from others in that T-cell activators could upregulate antisense transcription in both transfection and infection experiments. As opposed to these studies, we have not seen any positive effect of TNF α on luciferase activity, although other T-cell lines will need to be tested. Although this might argue that NF- κ B might not be important for induction of antisense transcription, other activators known to activate this transcription factors were capable of upregulating levels of luciferase activity. Transfection experiments with 3'LTR-deleted versions will be required to pinpoint the transcription factors important for this transcriptional modulation.

Our results have also surprisingly demonstrated that Tat could upregulate antisense transcription. We have multiple evidences favouring this conclusion: transient and stable transfection experiments indeed have been conducted and luciferase reporter measurement and RT-PCR analyses all lend credence to a positive modulation by Tat. Previous studies had rather suggested a negative role but as stated in one of these studies (4, 19, 21), these results might be consequential of TAR-dependent sense transcription being strongly induced by Tat and having an interference effect on luciferase transcript initiated from the

inverted 3'LTR. We have indeed demonstrated that this was likely the case by linearizing the vector. Similar conclusions were suggested by Bentley *et al.* who have mutated the 3'LTR TATA box and then observed a low but significant induction of antisense transcription by Tat. The mechanism by which Tat could activate the antisense promoter is unknown. Tat activates the expression of several different cellular genes through different mechanisms including binding to DNA-bound transcription factors (26). Although current model on the action of Tat through TAR binding on sense transcription favours an increase in the processivity of the RNA polII in sense transcription, it is still possible that Tat could also act upon formation of preinitiation complexes (PIC), as it has been recently readdressed (27, 28). Indeed, several reports have indicated that Tat can interact with factors present in the preinitiation complex. However it should also be kept in mind that Tat has been shown to modulate Sp1 binding (29, 30). An important Sp1 site has previously been proposed for antisense transcription (19) and could be the target of Tat-mediated upregulation of antisense transcription. More experiments will be required to determine the mechanism by which Tat activates antisense transcription. However, we are aware that the present effect of Tat on antisense transcription was demonstrated in 5' LTR-deleted proviral DNA constructs and thus this Tat-dependent modulation of antisense transcription remains to be confirmed in the context of a full-length proviral DNA (for example, using a Tat-deficient proviral clone).

Infection experiments with NL4.3 based virions expressing an antisense reporter gene have revealed that this tool could indeed be used to study antisense transcription in the context of infection. The specificity of the luciferase signal was confirmed by the addition of AZT. Interestingly, the extent of luciferase activity was lower in tested T and monocytic

cells than in 293T cells for both sense and antisense transcription. Remarkably, as pointed out above, the results from our infection experiments with the two different luciferase-encoding virions revealed that antisense transcription was importantly lower than sense transcription. A similar situation also exists for antisense transcription in HTLV-I (2, 12). This low abundance might be essential to avoid potential adverse effect of antisense transcript on sense transcripts and thereby virus replication. It is likely that sense transcription, especially upon induction by Tat or T-cell stimulating agents contribute importantly in keeping antisense RNA at low levels. It might be postulated that sense transcription initiated from the 3' LTR could also compete with limited amount of transcription factors and also diminish the extent of PIC formation on the antisense strand. In this regard, part of the function of the antisense RNA might be viewed as a true regulator of viral replication by acting as an antisense as previously suggested (31). As another possibility, the antisense transcript might have a totally different role on HIV-1 replication as recently suggested for the HTLV-I antisense transcript (13).

Another possible function of this transcript is through its encoding potential for the ASP protein. Important information is still lacking and most importantly its detection remains an important issue to solve. The nucleotidic region of this ORF is well conserved and its presumed ATG initiation is also very well preserved. HIV-1 is very heterogenous in the sequence of its different strains, and one might thus argue that the conservation of this ORF is an indication of its encoding capacity. In addition, its position is extremely similar to the HBZ ORF in HTLV-I. Preliminary data from our group has suggested that interrupting its translation by adding an in frame stop codon importantly reduces extracellular p24 levels in transfected cells, confirming previous data from Vaquero's team

(16). We are currently conducting several experiments to address issues related to this potential viral protein.

In conclusion, we have provided strong evidence for the existence of antisense transcription in HIV-1 and our data support a positive modulation of this transcript by the viral Tat protein. We have produced an important tool, which will permit to easily assess antisense transcription following infection or transfection in the context of the proviral DNA and in the chromatin environment. These data are important as they represent an important aspect of HIV-1 replication, which has been left unanswered. In addition, the potentially encoded ASP protein needs to be further studied. Understanding the function of this protein would add to the comprehension of HIV-1 pathogenesis. In addition, its suggested membrane-bound localisation would make it a target for anti-retroviral treatment and vaccine strategies.

Methods

Cell lines – T-cell line used in this study included Jurkat E6.1 (32), SupT1 (33) and CEMT4 (34). Tested HIV-1 infected T-cell lines were the ACH-2 (35) and J1.1 cell lines (36). The monocytic U937 (37), OM10.1 (38) and U937HIV-1IIIB cell lines were also used in the present study, the two latter being infected by two different HIV-1 strains. All these cell lines were maintained in RPMI-1640 medium supplemented with 10% Fetal Bovine Serum (Hyclone Laboratories, Logan, UT), 2mM glutamine, 100U/ml penicillin G, and 100µg/ml streptomycin. The human embryonic kidney 293T fibroblast cell line (39) was grown in similarly supplemented DMEM.

Plasmids and antibodies – The pCMV-tat vector was provided by Dr. O. Schwartz (Unité d'Oncologie Virale, Institut Pasteur, Paris, France) and contains the immediate early enhancer/promoter region of the human cytomegalovirus placed upstream of the viral *tat* gene (40). The control vector pCMV5 was kindly provided by Dr. É. Rassart (Université du Québec à Montréal, Montreal Canada). The pAsLTRXLuc construct was generated by excising the XhoI/HindIII 3'HIV-1 LAI LTR fragment from pLTRXLuc and cloning this LTR fragment after blunting in SmaI-digested pGL3-basic (Promega, Naepean, Canada) in the inverse orientation relative to the luciferase reporter gene. The pNL4.3 proviral DNA clone is a full-length infectious HIV-1 molecular clone provided by the NIH AIDS Research and Reference Reagent Program (Rockville, MD). The pNL4.3LucR-E- vector (containing the luciferase reporter gene and deficient for envelope synthesis) was generously provided by Dr. N. R. Landau (The Salk Institute for Biological Studies, La

Jolla, CA) (41). The pNL4.3 Δ NarI construct was derived from pNL4.3 by enzymatic digestion with NarI and self-ligation, thereby deleting the 5' LTR sequence. The pNL4.3AsLucR-E- vector was constructed by cloning the XbaI (blunted)/XhoI luciferase fragment from pGL3-basic in the NotI (blunted)/XhoI digested pNL4.3LucR-E- vector after excision of the luciferase reporter gene from this vector. The 5'LTR-deleted pNL4.3 Δ BstAsLucR-E- construct was derived from the pNL4.3AsLucR-E- vector by enzymatic digestion with BstZ17I and subsequent self ligation. The pHCMV-G vector expressing the broad host-range Vesicular Stomatitis Virus envelope glycoprotein G (VSV-G) has been described previously (42). The pActin-LacZ vector contains the lacZ gene under the regulation of the actin promoter. The hybridomas producing anti-p24 (clone 31-90-25 and 183H12-5C) antibodies were obtained from the American Type Culture Collection (Manassas VA). Antibodies from these hybridomas were purified with mAbTrap protein G affinity columns according to the manufacturer's instructions (Amersham Biosciences Inc., Uppsala, Sweden).

Transient and stable transfections. 293T cells were transfected in 24-well plates with Lipofectamine 2000 (Invitrogen; Burlington, Canada) according to manufacturer's suggestions. Transfected cells were lysed 48h. post-transfection in a lysis buffer (25mM Tris phosphate, pH 7.8, 2mM DTT, 1% Triton X-100, 10% glycerol) and luciferase activity was measured with the Dynex MLX microplate luminometer (MLX Dynex Technologies, Chantilly, VA) with a single injection of a luciferase buffer (20mM tricine, 1.07mM (MgCO₃)₄·Mg(OH)₂·5H₂O, 2.67mM MgSO₄, 0.1mM EDTA, 220 μ M Coenzyme A, 4.7 μ M D-Luciferin potassium salt, 530 μ M ATP, 33.3mM DTT). Each sample was co-transfected

with the β -gal-expressing vector for normalisation. The β -galactosidase activity was measured using the Galacto-Light™ kit (Applied Biosystems, Bedford, MS) according to manufacturer's suggestions. Luciferase activity is presented in Relative Light Units (RLU) and represents the calculated mean \pm SD of three transfected samples normalized by the measured β -galactosidase activity. Jurkat cells were transfected by electroporation with 15 μ g of DNA using the Gene Pulser Xcell system (Bio-Rad, Hercules CA) (960 μ F, 250V). Cells were resuspended at 37.5×10^6 /ml of complete medium, transfected in bulk and were separated at 16 h post-transfection into various treatment groups at a density of 1×10^5 cells/well (100 μ l) in 96-well flat-bottom plates. Cells were either left untreated or were treated with phorbol 12-myristate 13-acetate (PMA) at 20 ng/mL (Sigma-Aldrich, Oakville, Canada), phytohemagglutinin (PHA-P, Sigma-Aldrich, Oakville, Canada) at 3 μ g/mL, ionomycin (Sigma-Aldrich, Oakville, Canada) at 1 μ M, bpV[pic] (Alexis corp.) at 15 μ M, forskolin (Sigma-Aldrich, Oakville, Canada) at 10 μ M and TNF α at 20ng/ml. Luciferase activity was monitored as described above at 8h. post-stimulation. Stable transfection of 293T cells were performed as follow. Linearized pNL4.3 Δ BstAsLuc and pNL4.3 Δ NarI constructs (upon digestion with BstZ17I and ApaI, respectively) were co-transfected with each linearized plasmid with the pCMV-Hyg vector in the presence of Lipofectamine 2000 according to manufacturer's instructions. After transfection (48 hours), Hygromycin B (Sigma-Aldrich, Oakville, Canada) was added to the cells at a concentration of 200 μ g/ml. Hygromycin-resistant clones were isolated and tested for luciferase activity.

RT-PCR and 5'/3' RACE analyses. Total RNA was extracted by the Trizol reagent (Invitrogen; Burlington, Canada). PolyA+ RNA was purified from lysed cell samples using

the Poly(A)Purist™ Kit (Ambion, Austin TX) and according to manufacturer's instructions. RT-PCR analyses were conducted using the EndoFree RT Kit (Ambion, Austin TX) with RT primer 24-6 (5'-TAAAACAAATTATAAACATGTGGC-3') or the RT primer 24-6F (5'-**ATCATGAAGCATCTAGATTCAAAGTGCTGCTAAAACAAATTATAAACATGTGGC-3'**) (the non-complementary sequence being indicated in bold) according to manufacturer's instructions. Briefly, polyA+ RNA (250 ng) or total RNA (10µg) were mixed with 1µl of 10µM of a specific RT primer. The RNA:RT primer mix was heated at 70°C for 5 min., equilibrated at 50°C for 5 min. and incubated in the presence of the reaction mix containing 1X RT buffer, 1µl of 2.5mM dNTPs, 10 U of SuperAsin RNase inhibitor (Ambion) and 1µl RT. For certain RT reactions, cDNAs were purified through a QIAquick PCR purification kit (Qiagen) before PCR amplification. Synthesized cDNA were then PCR amplified in the presence of 1.25U Taq DNA polymerase, 1X ThermoPol buffer, 20µM dNTPs, 1.5µM of each primer using a Tgradient thermocycler (Biometra, Goettingen, Germany). Primers added to the PCR reactions were the reverse primer 26-6 (5'-AAAGCAATGTATGCCCTCCCATCAG -3') and forward primers 26-5 (5'-TGCTGTTGCGCCTCAATAGCCCTCAG-3'), 25-3 (5'-ATGCCCCAGACTGTGAGTTGCAACA-3'), or the anchor primer 30-20 (5'-ATCATGAAGCATCTAGATTCAAAGTGCTGC-3') (specifically used for the RT reactions performed in the presence of the 24-6F primer). PCR conditions were as follow: a first step of denaturation at 94°C for 5 min followed by 35 cycles of denaturation (94°C for 1 min.), annealing (60°C for 1 min.) and extension (72°C for 1 min.) and a final extension at 72°C for 5 min. RT-PCR amplifications were controlled for DNA

contamination (RNA samples with no RT step), for the presence of the RT floating primer (24-6F) after cDNA purification (mixing a 24-6-synthesized cDNA with an RT reaction performed with a 24-6F floating primer in non-transfected 293T RNA) and endogenous RT priming (cDNA synthesis reaction in the presence of RT with no added specific primer). Extremities of the antisense transcript (5' and 3') were analyzed from isolated total RNA with the FirstChoice RLM-RACE kit from Ambion according to the manufacturer's instructions. For the 5'RACE protocol, RNA was ligated to a supplied anchor at their 5' end and cDNAs were subsequently synthesized with random decamers. Two PCR rounds were conducted with the supplied 5'RACE outer and inner primers and HIV-1-specific primers R1 (5'-AATCACAAGTAGCAATACAGCAGCTAA-3') and R2 (5'-CAATTATTGTCTGATATAGTGCAGCAGCAGAAC-3') successively. For the 3'RACE protocol, cDNA synthesis was performed in the presence of the supplied 3'RACE adapter. PCR amplification was achieved through 3'RACE inner and outer primers and HIV-1-specific forward primer 30-10 (5'-GGCTGGGTTCGGTATTAAGGAA-3'). Amplified products were then cloned in pBlue Script KS+ (Stratagene, La Jolla CA) and sequenced.

Production of virus stocks. Virus particles were produced by calcium phosphate transfection of virus-encoding vectors in 293T cells as previously described (43, 44). Briefly, 293T cells were plated 16h. before transfection to reach a 50-80% confluence at the day of transfection and then transfected with 35-40 μ g through the calcium phosphate protocol as previously described (44). Pseudotyped HIV-1 particles were generated by cotransfecting pNL4.3LucR-E- (25-30 μ g) and the VSV-G envelope expression vector (5 μ g of pHCMV-G). Sixteen hours after transfection, cells were washed twice with PBS and

incubated for 24h. in complete DMEM culture medium. Viruses were collected at this point by filtering the culture media through a 0.22 μ m pore size cellulose acetate membrane. Virus stocks were normalized for virion content using an in-house sensitive double antibody sandwich enzyme-linked immunosorbent assay (ELISA) specific for the major core viral p24 protein as previously described (44). Virus stocks were aliquoted and frozen at -80°C for future use. All virus stocks underwent a single freeze-thaw cycle before use in infection studies.

Infections – 293T cells were plated at 7.5×10^4 cells/well in 24-well plates at 16h before infection. Pseudotyped virions (37.5ng p24) were added to each well and cells were lysed at different time points post-infection with 200 μ l of 1X lysis buffer. Jurkat, U937, CEMT4 and SupT1 cells were seeded in the presence of virions (1ng of p24/10000 cells) in complete medium at 1×10^6 cells/ml and incubated for 5h. at 37°C. Cells were then resuspended at 1×10^6 cells/ml in fresh medium and lysed at 48h. post-infection. Luciferase activity was monitored from 25 μ l aliquot of each cell lysate with the Dynex MLX microplate luminometer (MLX Dynex Technologies, Chantilly, VA) upon addition of 100 μ l of luciferase buffer as described above. Mean luciferase activity \pm standard deviation (SD) was calculated using luciferase activities from three independent and similarly treated cell samples taken from the same experiment.

List of abbreviations

HBZ: HTLV-I bZIP

HIV-1: human immunodeficiency virus type 1

HTLV-I: human T-cell leukemia virus type I

LTR: long terminal repeat

PBMCs: Peripheral Blood Mononuclear Cells

AZT: Azidothymidine

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

S. Landry developed the antisense-specific RT-PCR protocol, designed and carried out most of the RT-PCR analyses, the 5' and 3' RACE analyses, generated most of the DNA constructs and stable cell lines, performed the transfections and infections experiments, conducted the sequence alignments and drafted the manuscript. M Halin has performed RT-PCR and transfection experiments. B. Audet and S. Lefort have conducted several transfection experiments and have constructed certain vectors. C. Vaquero and JM. Mesnard have helped in drafting and finalizing the manuscript and provided important input on the design of the study. B. Barbeau have conceived the study, participated in its coordination, and helped in drafting the manuscript and finalizing the manuscript. All authors read and approved the final manuscript.

Acknowledgments

This work was supported by the Canadian Institutes of Health Research HOP-67257. S.L. was supported by a CIHR Ph.D. scholarship and B.B. holds a Canada Research Chair Tier 2. We are thankful toward Éric Legault and Julie Moreau for excellent technical support.

References

1. Briquet S, Richardson J, Vanhee-Brossollet C, Vaquero C: **Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus.** *Gene* 2001, 267:157-64.
2. Larocca D, Chao LA, Seto MH, Brunck TK: **Human T-cell leukemia virus minus strand transcription in infected T-cells.** *Biochem Biophys Res Commun* 1989, 163:1006-13.
3. Miller RH: **Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand.** *Science* 1988, 239:1420-1422.
4. Michael NL, Vahey MT, d'Arcy L, Ehrenberg PK, Mosca JD, Rappaport J, Redfield RR: **Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat.** *J Virol* 1994, 68:979-87.
5. Gaudray G, Gachon F, Basbous J, Biard-Piechaczyk M, Devaux C, Mesnard JM: **The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription.** *J Virol* 2002, 76:12813-22.
6. Hivin P, Frederic M, Arpin-Andre C, Basbous J, Gay B, Thebault S, Mesnard JM: **Nuclear localization of HTLV-I bZIP factor (HBZ) is mediated by three distinct motifs.** *J Cell Sci* 2005, 118:1355-62.
7. Hivin P, Arpin-Andre C, Clerc I, Barbeau B, Mesnard JM: **A modified version of a Fos-associated cluster in HBZ affects Jun transcriptional potency.** *Nucleic Acids Res* 2006, 34:2761-72.
8. Basbous J, Arpin C, Gaudray G, Piechaczyk M, Devaux C, Mesnard JM: **The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity.** *J Biol Chem* 2003, 278:43620-7.
9. Mesnard JM, Barbeau B, Devaux C: **HBZ, a new important player in the mystery of Adult-T- cell leukemia.** *Blood* 2006.

10. Thebault S, Basbous J, Hivin P, Devaux C, Mesnard JM: **HBZ interacts with JunD and stimulates its transcriptional activity.** FEBS Lett 2004, 562:165-70.
11. Arnold J, Yamamoto B, Li M, Phipps AJ, Younis I, Lairmore MD, Green PL: **Enhancement of infectivity and persistence in vivo by HBZ, a natural antisense coded protein of HTLV-1.** Blood 2006, 107:3976-82.
12. Cavanagh MH, Landry S, Audet B, Arpin-Andre C, Hivin P, Pare ME, Thete J, Wattel E, Marriott SJ, Mesnard JM, et al: **HTLV-I antisense transcripts initiating in the 3'LTR are alternatively spliced and polyadenylated.** Retrovirology 2006, 3:15.
13. Satou Y, Yasunaga J, Yoshida M, Matsuoka M: **HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells.** Proc Natl Acad Sci U S A 2006, 103:720-5.
14. Murata K, Hayashibara T, Sugahara K, Uemura A, Yamaguchi T, Harasawa H, Hasegawa H, Tsuruda K, Okazaki T, Koji T, et al: **A novel alternative splicing isoform of human T-cell leukemia virus type 1 bZIP factor (HBZ-SI) targets distinct subnuclear localization.** J Virol 2006, 80:2495-505.
15. Matsumoto J, Ohshima T, Isono O, Shimotohno K: **HTLV-1 HBZ suppresses AP-1 activity by impairing both the DNA-binding ability and the stability of c-Jun protein.** Oncogene 2005, 24:1001-10.
16. Briquet S, Vaquero C: **Immunolocalization studies of an antisense protein in HIV-1-infected cells and viral particles.** Virology 2002, 292:177-84.
17. Bukrinsky MI, Etkin AF: **Plus strand of the HIV provirus DNA is expressed at early stages of infection.** AIDS Res Hum Retroviruses 1990, 6:425-6.
18. Vanhee-Brossollet C, Thoreau H, Serpente N, D'Auriol L, Levy JP, Vaquero C: **A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals.** Virology 1995, 206:196-202.
19. Peeters A, Lambert PF, Deacon NJ: **A fourth Sp1 site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription.** J Virol 1996, 70:6665-72.

20. Ludwig LB, Ambrus, Jr. JL, Krawczyk KA, Sharma S, Brooks S, Hsiao CB, Schwartz SA: **Human immunodeficiency virus-type 1 LTR DNA contains an intrinsic gene producing antisense RNA and protein products.** *Retrovirology* 2006, 3:80.
21. Bentley K, Deacon N, Sonza S, Zeichner S, Churchill M: **Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription.** *Arch Virol* 2004, 149:2277-94.
22. Zhao J, Hyman L, Moore C: **Formation of mRNA 3' ends in eukaryotes: mechanism, regulation, and interrelationships with other steps in mRNA synthesis.** *Microbiol Mol Biol Rev* 1999, 63:405-45.
23. Chen BK, Saksela K, Andino R, Baltimore D: **Distinct modes of human immunodeficiency virus type 1 proviral latency revealed by superinfection of nonproductively infected cell lines with recombinant luciferase-encoding viruses.** *J. Virol.* 1994, 68:654-660.
24. Connor RI, Chen BK, Choe S, Landau NR: **Vpr is required for efficient replication of human immunodeficiency virus type-1 in mononuclear phagocytes.** *Virology* 1995, 206:935-944.
25. Sangar DV, Carroll AR: **A tale of two strands: reverse-transcriptase polymerase chain reaction detection of hepatitis C virus replication.** *Hepatology* 1998, 28:1173-6.
26. White MK, Gorrill TS, Khalili K: **Reciprocal transactivation between HIV-1 and other human viruses.** *Virology* 2006, 352:1-13.
27. Brady J, Kashanchi F: **Tat gets the "green" light on transcription initiation.** *Retrovirology* 2005, 2:69.
28. Raha T, Cheng SW, Green MR: **HIV-1 Tat stimulates transcription complex assembly through recruitment of TBP in the absence of TAFs.** *PLoS Biol* 2005, 3:e44.
29. Vlach J, Garcia A, Jacque JM, Rodriguez MS, Michelson S, Virelizier JL: **Induction of Sp1 phosphorylation and NF-kappa B-independent HIV promoter domain activity in T lymphocytes stimulated by okadaic acid.** *Virology* 1995, 208:753-61.

30. Chun RF, Semmes OJ, Neuveut C, Jeang KT: **Modulation of Sp1 phosphorylation by human immunodeficiency virus type 1 Tat.** *J Virol* 1998, 72:2615-29.
31. Tagieva NE, Vaquero C: **Expression of naturally occurring antisense RNA inhibits human immunodeficiency virus type 1 heterologous strain replication.** *J Gen Virol* 1997, 78 (Pt 10):2503-11.
32. Weiss A, Imboden J, Shoback D, Strobo J: **Role of T3 surface molecules in human T-cell activation: T3-dependent activation results in an increase in cytoplasmic free calcium.** *Proc. Natl. Acad. Sci. USA* 1984, 81:4169-4173.
33. Smith SD, Shatsky M, Cohen MPS, Warnke R, Link MP, Glader BE: **Monoclonal antibody and enzymatic profiles of human malignant T lymphoid cells and derived cell lines.** *Cancer Res.* 1984, 44:5657-5660.
34. 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.
35. Clouse KA, Powell D, Washington I, Poli G, Strebel K, Farrar W, Barstad P, Kovacs J, Fauci AS, Folks TM: **Monokine regulation of human immunodeficiency virus-1 expression in a chronically infected human T cell clone.** *J. Immunol.* 1989, 142:431-438.
36. Perez VL, Rowe T, Justement JS, Butera ST, June CH, Folks TM: **An HIV-1-infected T cell clone defective in IL-2 production and Ca²⁺ mobilization after CD3 stimulation.** *J. Immunol.* 1991, 147:3145-3148.
37. Sundstrom C, Nilsson K: **Establishment and characterization of a human histiocytic lymphoma cell line (U-937).** *Int. J. Cancer* 1976, 17:565-577.
38. Butera ST, Perez VL, Wu B-Y, Nabel GJ, Folks TM: **Oscillation of the human immunodeficiency virus surface receptor is regulated by the state of viral activation in a CD4⁺ cell model of chronic infection.** *J. Virol.* 1991, 65:4645-4653.
39. Pear WS, Nolan GP, Scott ML, Baltimore D: **Production of high-titer helper-free retroviruses by transient transfection.** *Proc. Natl. Acad. Sci. USA* 1993, 90:8932-8936.

40. Schwartz O, Virelizier J-L, Montagnier L, Hazan U: **A microtransfection method using the luciferase-encoding reporter gene for the assay of human immunodeficiency virus LTR promoter activity.** *Gene* 1990, 88:197-205.
41. Deng H, Liu R, Ellmeier W, Choe S, Unutmaz D, Burkhart M, Di Marzio P, Marmon S, Sutton RE, Mark Hill C, et al: **Identification of a major co-receptor for primary isolates of HIV-1.** *Nature* 1996, 381:661-666.
42. Yee JK, Miyanochara A, Laporte P, Bouic K, Burns JC, Friedmann T: **A general method for the generation of high titer, pantropic retroviral vectors: highly efficient infection of primary hepatocytes.** *Proc. Natl. Acad. Sci. USA* 1994, 92:9564-9568.
43. Cantin R, Fortin J-F, Tremblay M: **The amount of host HLA-DR proteins acquired by HIV-1 is virus strain- and cell type-specific.** *Virology* 1996, 218:372-381.
44. Fortin J-F, Cantin R, Lamontagne G, Tremblay M: **Host-derived ICAM-1 glycoproteins incorporated on human immunodeficiency virus type 1 are biologically active and enhance viral infectivity.** *J. Virol.* 1997, 71:3588-3596.

Figure legends

Figure 1: Positioning of the ASP antisense ORF in the HIV-I proviral DNA. The ASP ORF is located on the antisense strand in the region of the *env* gene. Primers used for RT-PCR experiments and the expected sizes of the amplified signals are indicated below the enlarged ASP ORF. The arrow indicated the antisense transcript.

Figure 2: Specific detection of the antisense transcript in latently infected cells and transfected 293T cells (**A**) RT-PCR analyses were performed on RNA samples from J1.1 cells using the 24-6 RT primer and PCR primer combinations 26-6/26-5 (lanes 1,2,4,6) and 26-6/25-3 (lanes 3,5,7) (expected size of 384 bp and 437 bp, respectively). Samples were tested for DNA contamination (lanes 2-3; no RT and no RT primer) and endogenous RT priming (lanes 4-5; RT with no added RT primer). Lane 1 represents PCR analysis with no added cDNA or RNA. Lanes 6 and 7 show the results of PCR using 2 primers combinations. (**B**) 293T cells were transfected with 5 μ g of pNL4.3 or pNL4.3 Δ Nar1 proviral DNA. RT-PCR analysis of RNA samples from transfected 293 T cells and controls were performed as in **A**. M= Lambda EcoRI/HindIII DNA marker (the asterisk indicates the 564 bp band). The arrows on the right side of panel **A** points to the specific signal.

Figure 3: Specific detection of HIV-1 antisense transcript in infected cells lines and transfected 293T cells. Synthesis of cDNA was performed on polyA⁺ RNA using 24-6 (control RT primer) or 24-6F (floating RT primer) and purified on a PCR-cleanup column. PCR amplifications were performed using the reverse primer 30-20 (anchor) and forward

primers 26-5 (lane 1, 3, 5) or 25-3 (lane 2, 4 and 6) in order to specifically amplify 24-6F-synthesized cDNA. Samples were tested for anchor primer specificity (lanes 1-2) and cDNA cleanup efficiency (lanes 3-4). Lane 5 and 6 show specific amplification of HIV-1 antisense transcript from J1.1 OM10.1, U937 chronically infected cells lines and 293T transfected with complete pNL4.3 proviral DNA or 5'LTR-deleted pNL4 Δ Nar1 construct.

Figure 4: Identification of transcription initiation sites for the antisense transcript. Total RNA from 293T cells transfected with the pNL4 Δ Nar1 construct was used to amplify 5' cDNA ends using the 5' RACE approach. Numerous transcription initiation sites were identified upon sequencing of multiple PCR-amplified signals and are part of a region presented in the enlarged segment encompassing the LTR, and both *nef* and *env* gene regions. Nucleotide numbering corresponds to the sense strand.

Figure 5: Identification of a new functional polyadenylation site for HIV-1 antisense transcript. (A) Total RNA from 293T cells transfected with pNL4.3 Δ NarI was used for 3'RACE analysis. CTL represents PCR amplification conducted in the absence of cDNA and RNA samples. The amplified product was cloned and sequenced. M=100 bp marker (the asterisk indicates the 600 bp band). (B) Position of the newly identified antisense polyA addition site (pA indicated with arrow) in the HIV-1 genome. (C) Sequence alignment of 8 HIV-1 genome showing conservation for the consensus polyA signal and of a GT-rich consensus sequence. Position of the polyA addition site is indicated by an arrow.

Figure 6: Generation of a proviral DNA construct expressing the antisense directed luciferase gene expression. **(A)** 293T cells were transfected with 600ng pNL4.3AsLucR-E- and lysed 48h post-transfection. CTL represents untransfected cells. **(B-C)** Jurkat cells were transfected with 15 μ g pNL4.3AsLucR-E- **(B)** or pNL4.3 Δ BstAsLucR-E- **(C)** and stimulated with different activators for 8h. or left untreated. Lysed cells were then analysed for luciferase activity. Results show the mean luciferase activity values of three measured samples \pm S.D **(A-B)** or fold induction compared to the unstimulated control **(C)**.

Figure 7: Modulatory effect of Tat on antisense transcription. **(A)** Jurkat cells were transfected with 10 μ g of circular or BamHI-digested pAsLTRXLuc plasmid in combination with 5 μ g of pCMV-tat or empty pCMV vector. **(B)** 293T cells were transfected with 400ng of circular or BstZ17I-digested pNL4.3 Δ BstAsLucR-E-, 200ng of pActin-LacZ and 200ng of pCMV-tat or empty vector. **(C)** 293T cell clones and a pooled population stably transfected with pNL4.3 Δ BstAsLucR-E- were transfected with 200ng pCMV-tat or empty vector. Results are presented as fold induction compared to empty vector. **(D)** A pool of 293T cells stably transfected for pNL4.3 Δ NarI was transfected with 5 μ g of pCMV-tat or the empty pCMV5 vector. cDNA synthesis was performed with random primers. PCR amplifications were performed to detect the presence of the HIV antisense transcript (lane 3-4-7-8) using 24-6/25-3 primers. β -actin amplification was performed as control (lane 1-2-5-6). Lanes 1, 2, 3 and 4 represent control for DNA contamination to which RNA was directly added for PCR amplification. **(E)** RNA from the transfected 293T cells described in **D** were also used for amplification of antisense transcripts from 24-6F-synthesized cDNA using 30-20 (anchor) primer in combination with 25-3 (lane 1, 3, 5 and 7) or 26-5

(lanes 2, 4, 6 and 8) primers. Samples were tested for cDNA cleanup efficiency (lanes 1, 2, 5 and 6). Tat expression in transfected cells is indicated above the gel for both latter panels. Luciferase activities in *A*, *B* and *C* represent the mean value of three measured samples \pm S.D.

Figure 8: Expression of HIV-1 antisense transcript during viral infection. (A) 293T cells were infected with 35ng of NL4.3LucR-E- or NL4.3AsLucR-E- virions pseudotyped with Vesicular Stomatitis Virus envelope (VSVg) and were lysed at different time points post-infection and tested for luciferase activity. In (B), 10 μ M or 50 μ M of AZT was added to the culture media 1h before infection of 293T cells and luciferase activity was determined after 48 hr. (C) SupT1, Jurkat, CEMT4 and U937 cells were infected with NL4.3LucR-E- or VSVg-pseudotyped NL4.3AsLucR-E- virions at a ratio of 1ng / 1 x 10⁵ cells. Cells were lysed 48 hr post-infection and tested for luciferase activity. (D) Forty-eight hr post-infection, Jurkat cells were seeded at a concentration of 1 x 10⁶ cells/ml and stimulated with various cell activators for 8h. Luciferase activities represent the mean value of three measured samples \pm S.D. Luciferase activity in panels A, B and C are presented on a logarithmic scale.

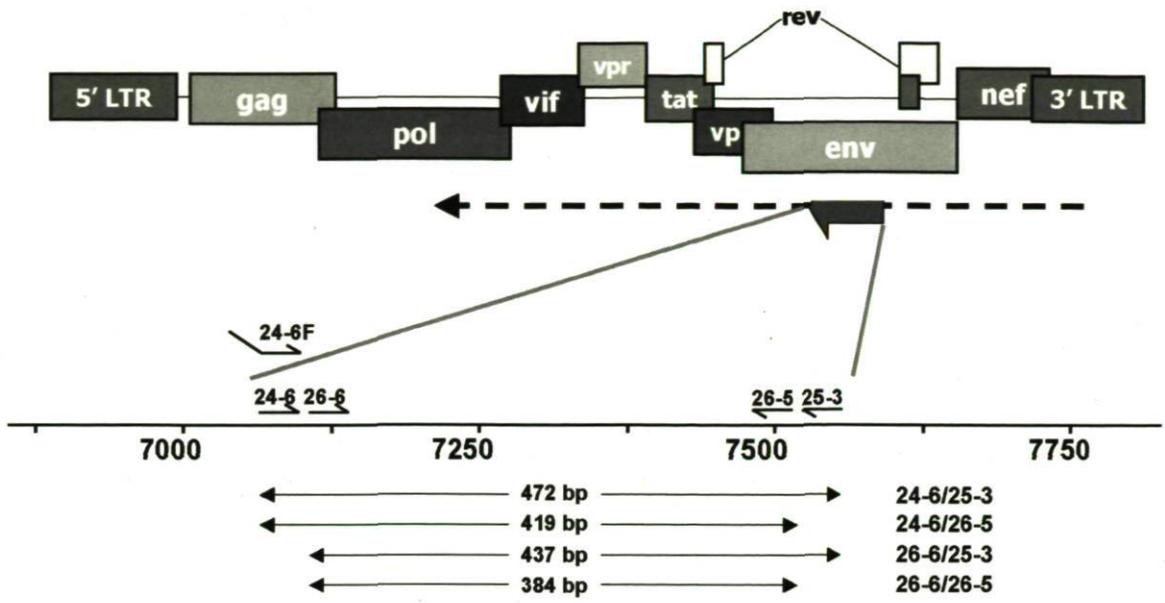


Figure 1

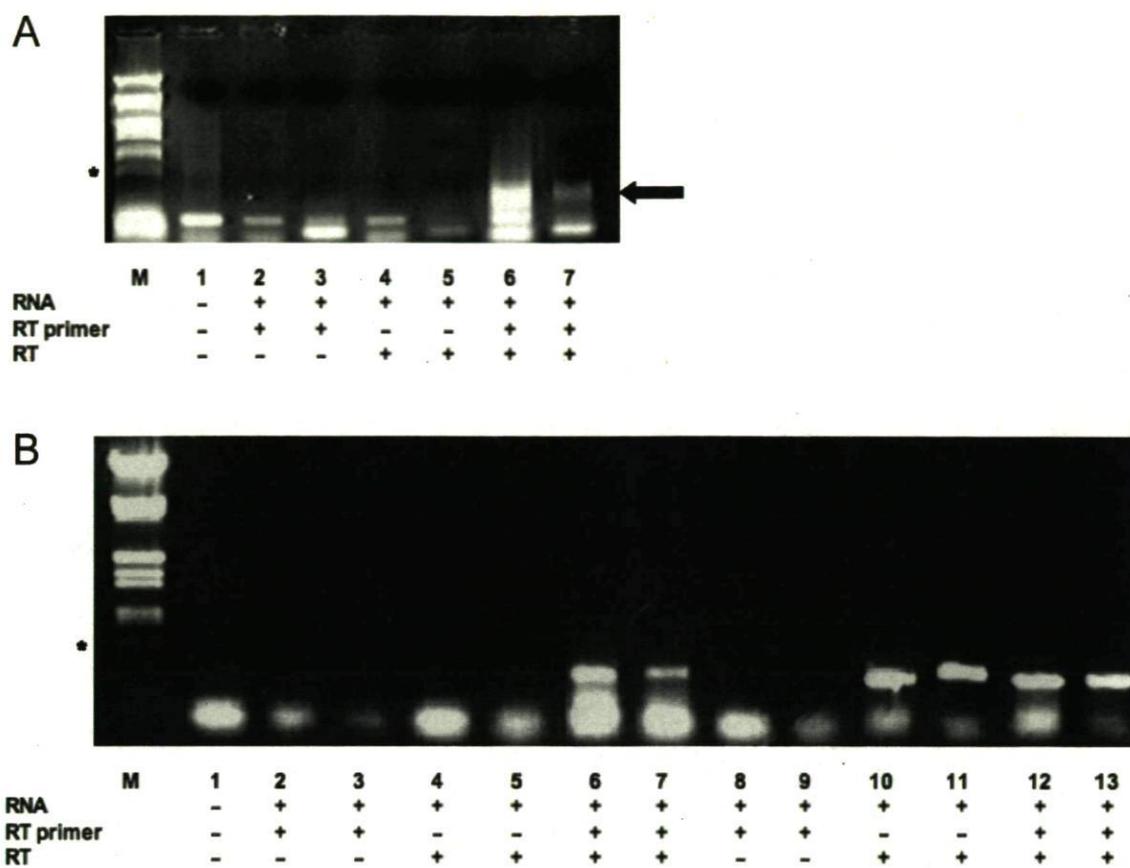


Figure 2

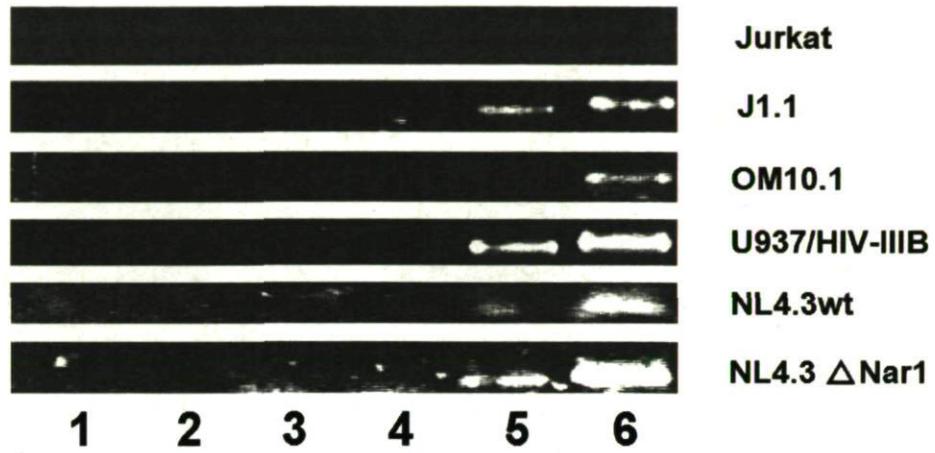


Figure 3

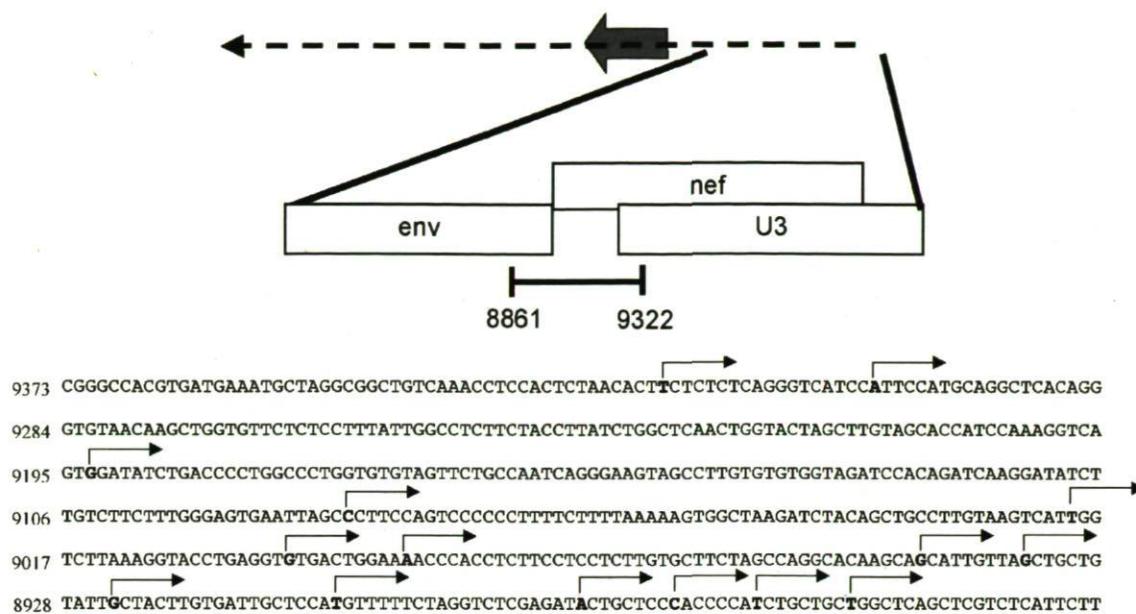


Figure 4

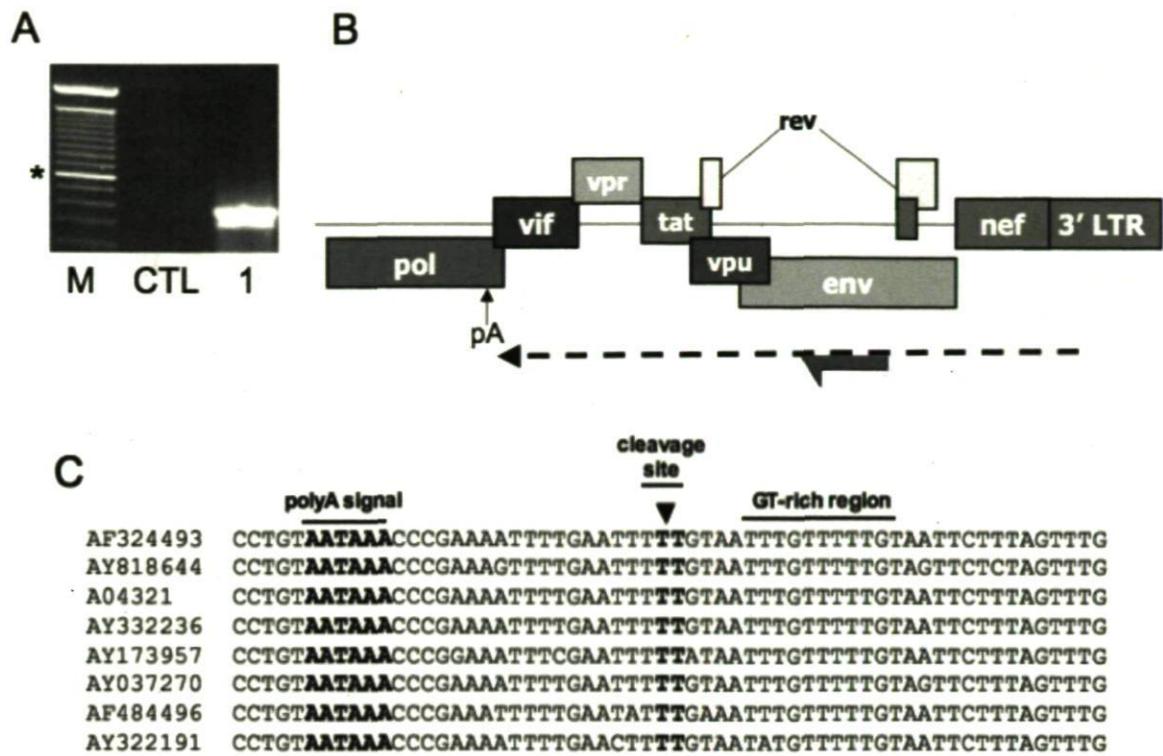


Figure 5

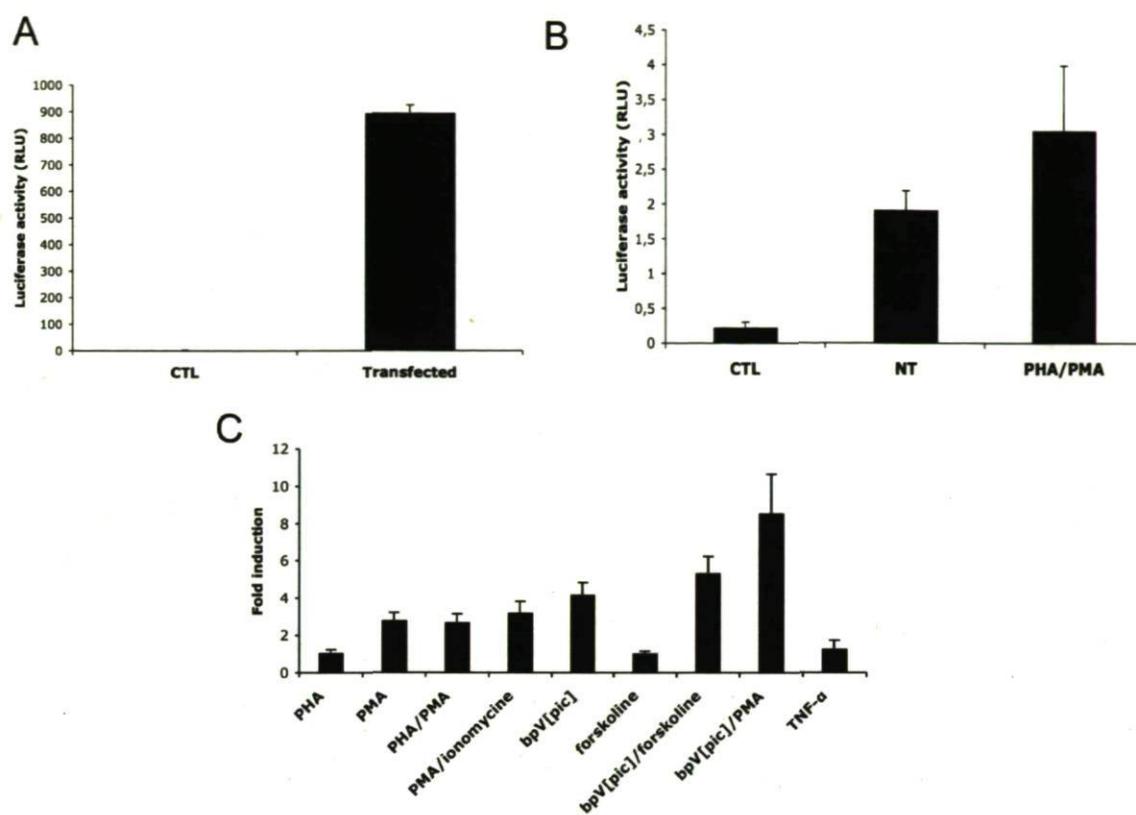


Figure 6

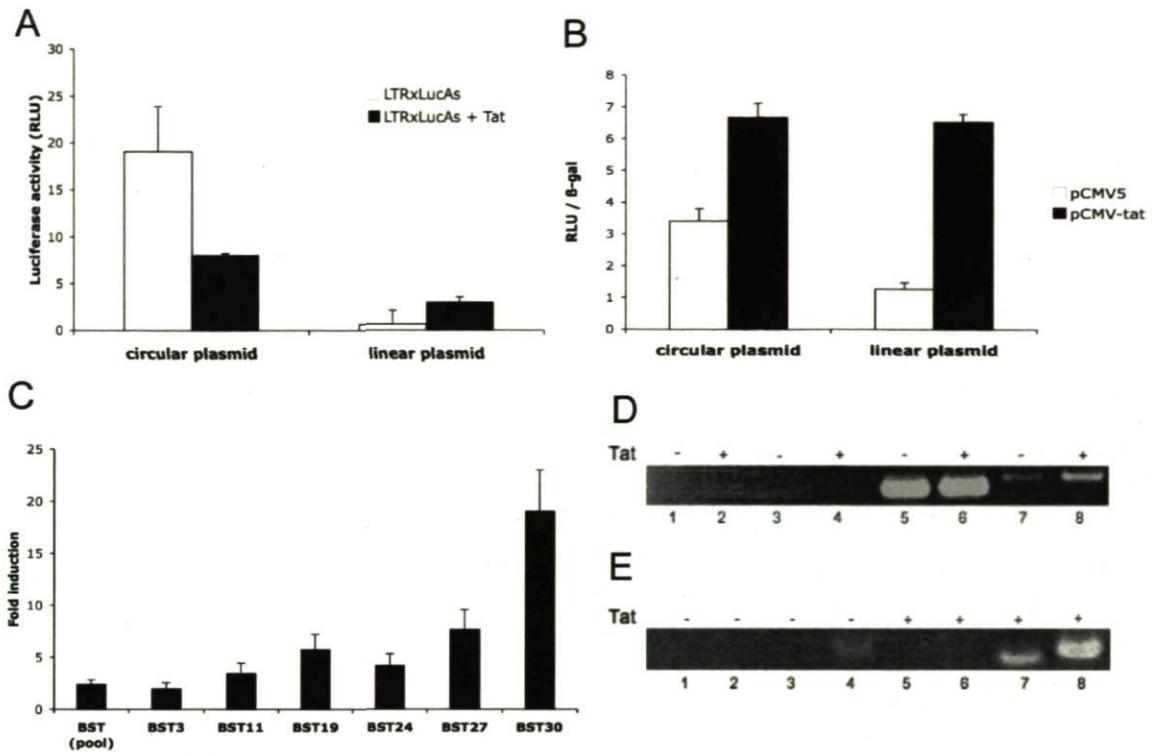


Figure 7

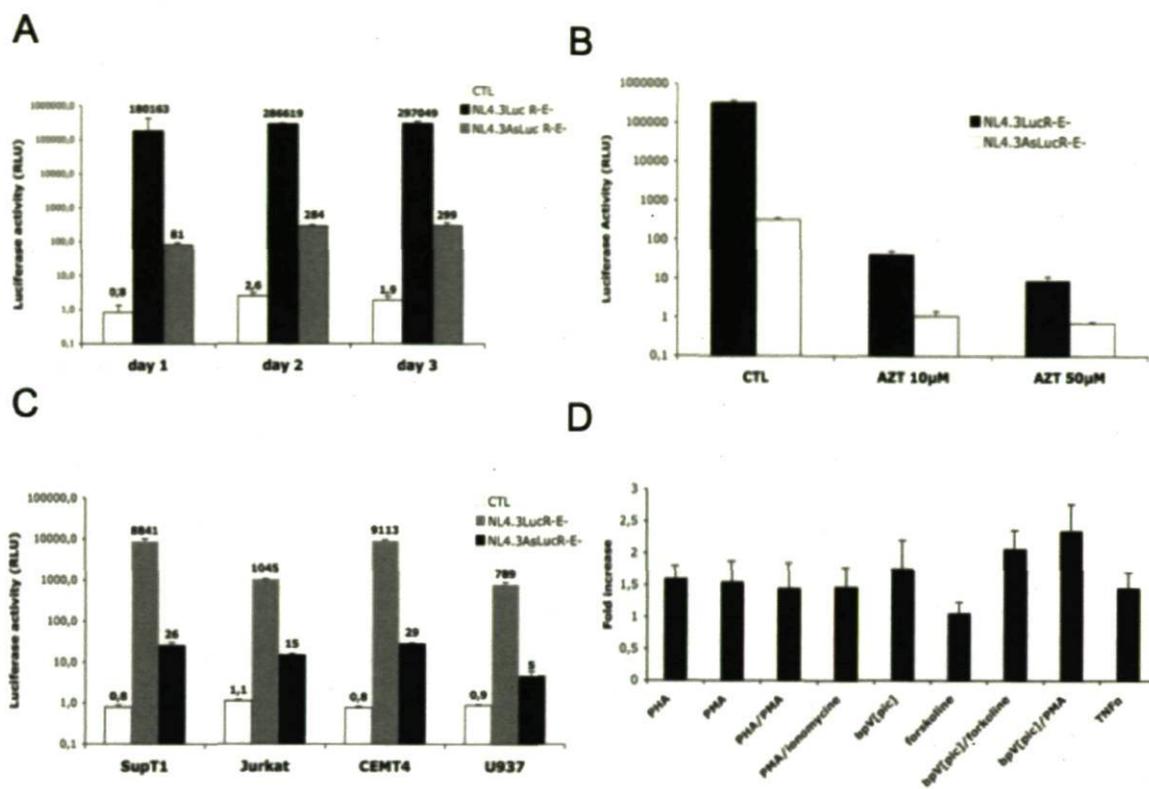


Figure 8

Chapitre 7 : Expression et détection de la protéine ASP

7.1 Introduction

Les résultats présentés dans Landry *et al.* (chapitre 6) démontraient clairement l'existence de la transcription antisens au cours de la réplication virale chez le VIH-1 [271]. À la lumière des résultats obtenus par notre groupe ainsi que par d'autres équipes sur la transcription antisens chez HTLV-1 et la nouvelle protéine virale HBZ, plusieurs expériences ont été entreprises pour tenter d'évaluer la capacité du transcrit antisens du VIH-1 à produire une nouvelle protéine virale. Compte tenu du peu de données disponibles quant à l'existence d'une telle protéine, le schéma expérimental a tout d'abord été conçu dans la simple optique d'identifier et de localiser cette dernière dans un contexte cellulaire.

Cette section est présentée sous forme de résultats non-publiés.

Les étapes de génération des baculovirus recombinants ainsi que la purification de la protéine ASP ont été effectuées dans le laboratoire du Dr Jean-Yves Masson. Cette partie du travail a été accomplie par Yan Coulombe, MSc.

7.2 Résultats

7.2.1 Expression et purification de la protéine ASP dans les cellules d'insectes

Une première étape dans l'étude de la protéine ASP consistait en son expression et sa purification, ceci dans l'objectif d'utiliser la protéine purifiée pour générer des anticorps spécifiques. La séquence codante de la protéine ASP a donc été amplifiée par PCR et clonée dans le vecteur d'expression pFastBac en ajoutant une étiquette Myc à l'extrémité N-terminale et une étiquette 6-His à l'extrémité C-terminale (pour permettre la purification). Ce plasmide donneur permet l'expression d'un gène d'intérêt sous le contrôle d'un promoteur spécifique de baculovirus. La transformation de cette construction dans la souche bactérienne *E. Coli* DH10Bac, qui contient un vecteur navette baculoviral et un plasmide auxiliaire, permet la transposition du vecteur pFasbac et ainsi la génération d'un bacmide recombinant. La transfection de ce dernier dans les cellules d'insectes SF9 permet la production de grandes concentrations de baculovirus recombinants exprimant la protéine d'intérêt. Spécifions que pour ces expériences, la séquence codante de la protéine a été amplifiée à partir d'un vecteur d'expression pSV-III contenant la séquence codante d'une protéine d'enveloppe dérivée d'un isolat clinique du VIH-1 (pSVIII-92HT594.10, NIH AIDS research and reference reagent program). Le choix de cette séquence a été effectué sur la base de la présence d'une similarité optimale avec la séquence consensus de la protéine ASP originalement publiée par Miller [243].

Suite à la sélection de clones positifs, il a été possible d'exprimer et de purifier avec succès la protéine recombinante à partir de cellules SF9 infectées, tel que décrit précédemment [272]. La quantité de protéine obtenue n'a cependant pas été suffisante pour poursuivre nos objectifs premiers de production d'anticorps à l'aide de protocoles animaux. Cependant, considérant le peu d'information disponible concernant la protéine ASP et les résultats peu concluants des premières expériences effectuées dans les cellules eucaryotes, le modèle des

cellules SF9 n'a pas été abandonné. Il présentait en effet l'avantage de permettre un niveau d'expression suffisant pour permettre la détection de la protéine.

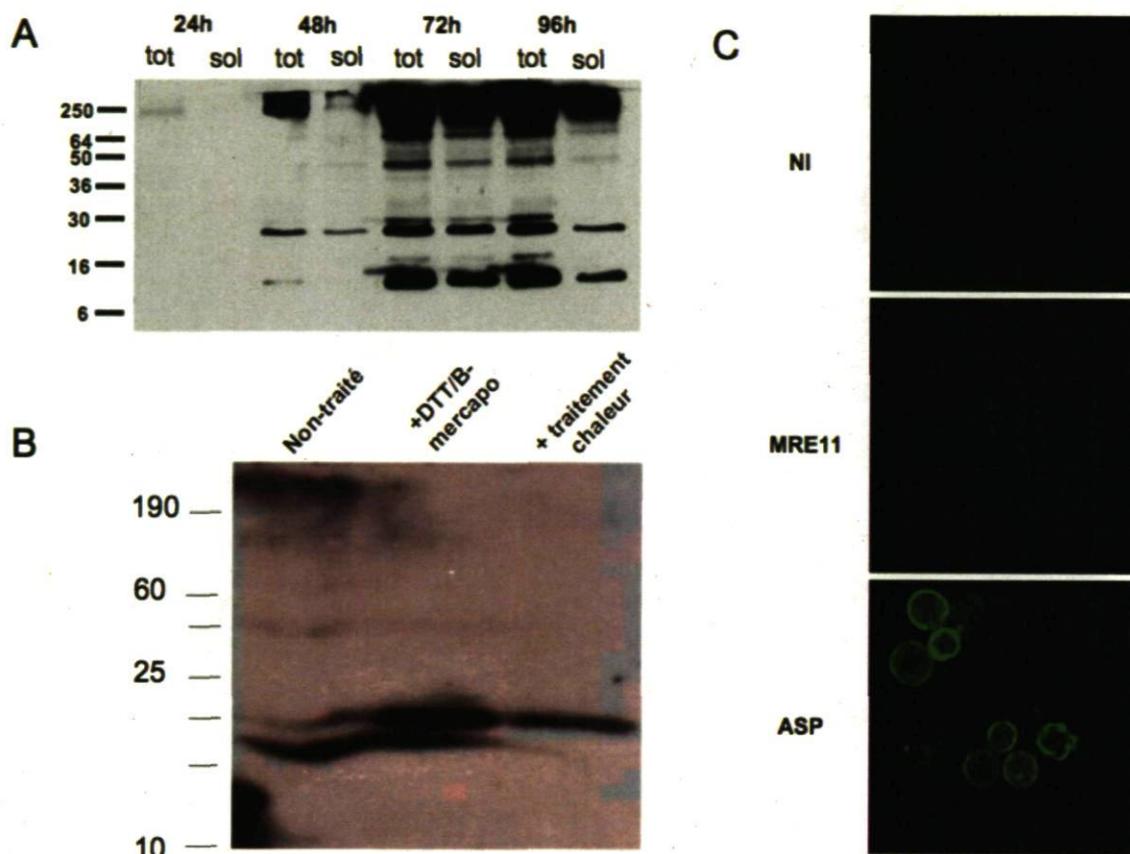


Figure 25 : Expression de la protéine ASP dans les cellules d'insectes SF9. (A) Détection de la protéine myc-ASP à l'aide d'un anticorps anti-myc (Santa Cruz biotechnology) à partir d'extraits totaux ou solubles obtenus de cellules SF9 infectées par un baculovirus recombinant. (B) Détection de la protéine ASP purifiée non-traitée ou suite à un traitement au DTT/ β -mercapto-éthanol combiné ou non avec un traitement à la chaleur. (C) Détection en microscopie confocale de la protéine ASP dans les cellules infectées à 48 h post-infection à l'aide d'un anticorps anti-Myc. Les cellules ont préalablement été fixées à la paraformaldéhyde 4% et perméabilisées au méthanol.

En utilisant un anticorps commercial anti-myc (Santa Cruz biotechnology), des expériences initiales de *Western blot* effectuées à partir d'extraits de cellules infectées ont permis de détecter des bandes spécifiques de la protéine ASP, visibles à partir de 24h post-infection.

De façon intéressante, l'observation de plusieurs bandes à différents poids moléculaires (figure 25A) suggérait fortement la présence de multimères. Cette hypothèse a été appuyée par l'obtention d'une seule bande lorsque les extraits étaient incubés pour 10 minutes à 95°C en présence d'agents réducteurs (DTT/ β -mercapto-éthanol) (résultats non-présentés).

Les mêmes expériences ont subséquemment été répétées en utilisant les fractions purifiées de la protéine ASP et ont permis de confirmer notre hypothèse de départ quant à la capacité de la protéine à former des multimères (figure 25B).

À la lumière de ces résultats, nous étions donc intéressés à potentialiser le modèle des cellules d'insectes qui, bien qu'étant éloigné du contexte des lymphocytes T, permettait pour la première fois la détection efficace de la protéine ASP. Toujours en utilisant un anticorps anti-Myc, nous avons donc entrepris des expériences de microscopie confocale dans le but de déterminer la localisation cellulaire de la protéine. Les cellules infectées par les baculovirus recombinants ont donc été fixées à la paraformaldéhyde et perméabilisées (ou non) au méthanol à différents temps post-infection et marquées selon un protocole décrit précédemment [273] à l'aide d'un anticorps secondaire couplé à un fluorochrome (Alexa 488 anti-mouse, Invitrogen). Les signaux ont été comparés à ceux des échantillons contrôles de cellules non-infectées ou infectées par un baculovirus exprimant une autre protéine (MRE11). À 48 h post-infection, un fort signal positif a pu être observé chez les cellules exprimant ASP, alors que les échantillons contrôles montraient une absence complète de signaux spécifiques. Les cellules positives pour ASP étaient caractérisées par un signal d'apparence circulaire qui semblait délimiter sans interruption la surface cellulaire, caractéristique d'un marquage de la membrane plasmique. Nous avons également pu noter la présence d'un signal d'une intensité plus faible au niveau du cytoplasme. Ces observations corroboraient celles présentées dans une autre étude [267] et semblaient aller de pair avec différentes analyses effectuées *in silico* à partir de la séquence en acides aminés de la protéine. Pour clarifier cette possible localisation membranaire d'ASP, le marquage des cellules non-fixées a été effectué à l'aide de la même combinaison d'anticorps utilisée précédemment et les échantillons ont été analysés par cytométrie de flux (résultats non-présentés). L'obtention d'un taux significatif de cellules positives (30%) (résultats non

montrés) pour l'échantillon ASP supportait alors non seulement la localisation cellulaire de la protéine, mais suggérait également une possible localisation extracellulaire de son extrémité N-terminale, ce qui permettrait de rendre disponible l'étiquette Myc lors du marquage de cellules vivantes. Il est cependant possible que l'effet pathogénique du baculovirus puisse rendre possible la détection d'un épitope intracellulaire même en absence de fixation ou perméabilisation et des contrôles supplémentaires doivent être effectués pour confirmer ces résultats. De plus, des expériences de co-localisation en utilisant un marqueur de la membrane plasmique ou du noyau permettraient sans doute de clarifier la question.

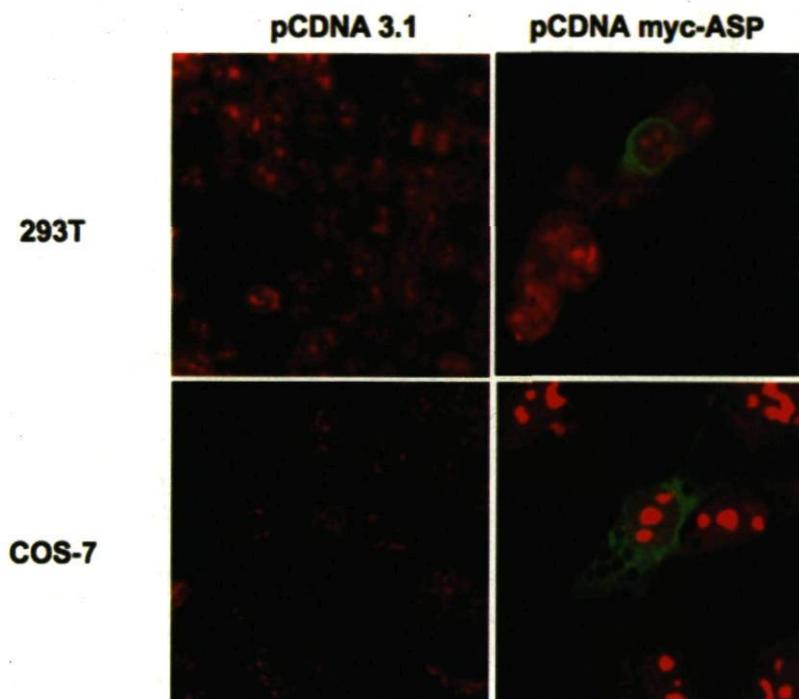


Figure 26 : Localisation de la protéine myc-ASP dans les cellules 293T et COS-7. Les cellules 293T et COS-7 ont été transfectées avec 1 μ g de pcDNA-ASP ou du vecteur vide. Quarante-huit heures post-transfection, les cellules ont été fixées, perméabilisées et marquées en utilisant un anticorps anti-Myc (SantaCruz) et un anticorps secondaire couplé à un fluorochrome (Alexa488). Les noyaux ont été marqués à l'iodure de propidium.

7.2.2 Détection de la protéine ASP dans les cellules 293T et COS-7

Conscients de la nécessité de confirmer nos résultats dans un modèle de cellules humaines, nous avons poursuivi nos expériences de microscopie en utilisant le vecteur d'expression eucaryote pcDNA 3.1 contenant la séquence codante d'ASP. Encore une fois, les cellules ont été marquées à l'aide d'un anticorps primaire anti-Myc et d'un anticorps secondaire couplé à un fluorochrome. L'observation des cellules transfectées par notre vecteur d'expression a permis de constater la présence d'un faible pourcentage de cellules positives, absentes dans les échantillons transfectés avec le vecteur vide (voir figure 26).

Contrairement aux observations faites dans les cellules SF9, le signal positif semblait cette fois-ci se situer majoritairement au cytoplasme. Alors que son absence au noyau ne semblait pas faire de doute, il apparaissait difficile de déterminer si une partie du signal pourrait co-localiser avec la membrane plasmique. Des essais préliminaires de colocalisation effectués à l'aide d'un marqueur spécifique de la membrane plasmique (CellMask, Invitrogen) semblaient démontrer que cela pourrait en effet être le cas (résultats non-présentés).

7.2.3 Optimisation de la séquence codante d'ASP et détection de la protéine

En dépit des résultats obtenus en microscopie confocale, les nombreux essais effectués pour tenter de détecter la protéine ASP en *Western blot* dans les lignées cellulaires humaines se sont tous avérés non-concluants. Ces résultats s'expliquent sans doute par un faible taux d'expression de la protéine, une hypothèse qui va de pair avec le faible pourcentage de cellules positives observées en microscopie. Plusieurs méthodes ont donc été mises de l'avant pour permettre un meilleur taux d'expression de la protéine. Suite à plusieurs essais infructueux, nous nous sommes tournés vers une stratégie déjà employée pour d'autres protéines du VIH-1 : l'optimisation des codons [274-276]. La synthèse de l'ADNc optimisé de la protéine a donc été effectuée de façon à inclure les codons les plus

fréquemment utilisés chez les eucaryotes (Genscript Corp.). L'ADN obtenu a été cloné dans le vecteur d'expression pcDNA3.1 et des expériences de microscopie confocale ont été entreprises pour comparer le taux d'expression obtenu avec cette nouvelle construction avec la construction originale pcDNA-ASP. De façon intéressante, nous avons pu constater que le taux de cellules positives observées pour les cellules 293T ou COS-7 transfectées avec le vecteur pCDNA-ASP-opt était beaucoup plus élevé que celui observé dans les cellules transfectées avec la construction sauvage (Figure 27). Il ne fait donc aucun doute que l'optimisation de la séquence codante de la protéine permet une meilleure expression et rend donc sa détection plus facile.

Puisque nous possédions un nouvel outil permettant une expression efficace de la protéine ASP, de nouvelles tentatives de détection de la protéine ASP à partir de cellules transfectées ont été effectuées par *Western blot*. Alors que les premiers essais exécutés à l'aide d'un protocole standard avaient donné des résultats négatifs, l'utilisation d'un protocole permettant d'enrichir les extraits protéiques en protéines membranaires a finalement permis d'obtenir un signal spécifique au poids attendu de 20 kDa (Figure 27C).

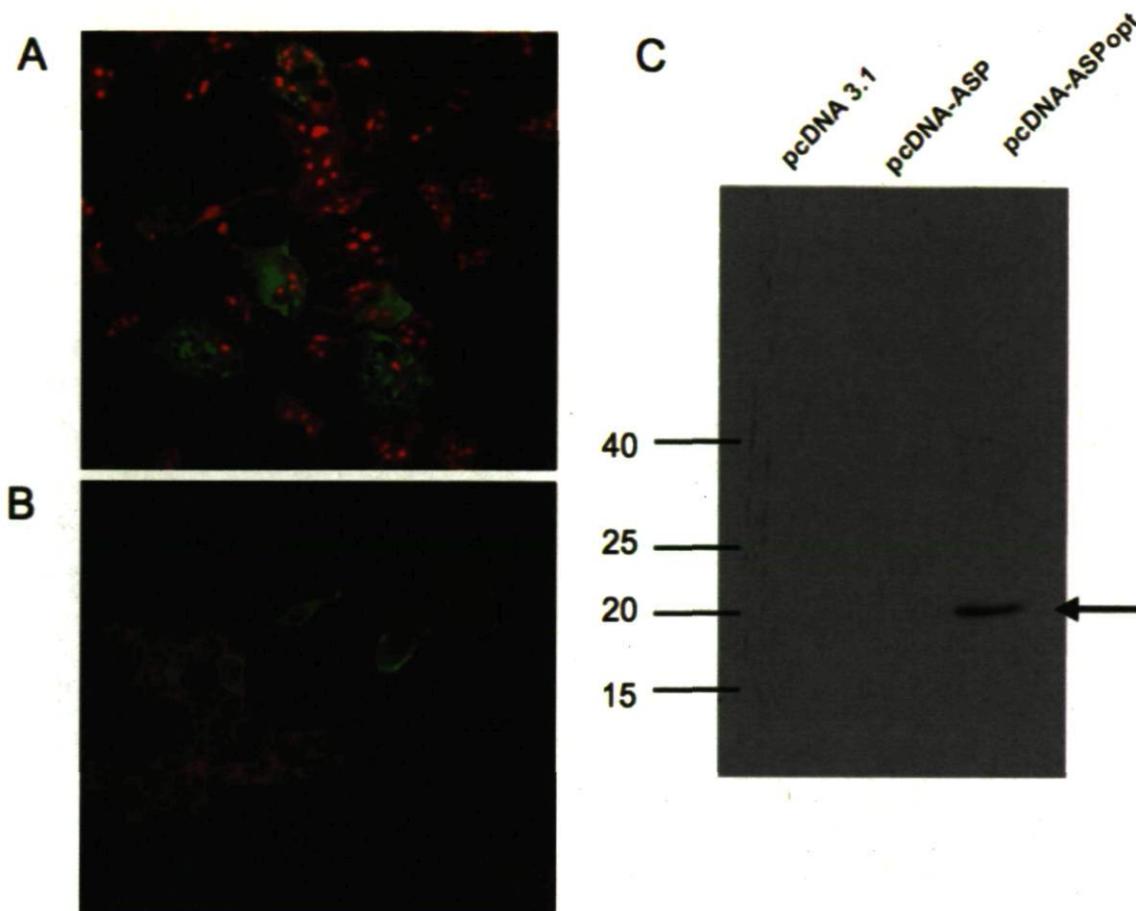


Figure 27 : Détection de la protéine ASP exprimée à l'aide d'un ADNc optimisé. (A et B) Les cellules 293T (A) ou COS-7 (B) ont été transfectées avec 1 μ g de la construction pcDNA-ASP-opt. 48 h post-transfection, les cellules ont été fixées, perméabilisées et marquées en utilisant un anticorps anti-Myc (SantaCruz) et anticorps secondaire marqué (AlexaFluor 488 anti-mouse, Invitrogen). Les noyaux ont été marqués à l'iodure de propidium. (C) Les extraits protéiques de cellules COS-7 transfectées avec 5 μ g de pcDNA3.1, pcDNA-ASP ou pcDNA-ASPOpt ont été obtenus en utilisant la trousse MemPER Mammalian Membrane Protein Extraction (Pierce). Les extraits ont ensuite été migrés sur un gel de polyacrylamide 12% et la détection a été effectuée à l'aide d'un anticorps primaire anti-Myc (SantaCruz) et d'un anticorps secondaire anti-souris couplé à la peroxidase (GE Healthcare).

Chapitre 8 : Discussion générale

8.1 La transcription antisens chez les rétrovirus : un phénomène sous-estimé

Au moment où le projet basé sur l'étude de la transcription antisens se mettait en branle dans nos laboratoires, le sujet constituait toujours un pari risqué : malgré la description récente de la protéine HBZ, qui argumentait fortement pour l'existence et la pertinence biologique du phénomène, rares étaient les rétrovirologistes qui semblaient prêts à lui accorder une réelle importance. Il est en effet surprenant de constater aujourd'hui qu'une protéine virale ait pu demeurer non identifiée pendant plus de 20 ans chez un virus très étudié, qui pourtant n'exprime qu'un nombre très limité de protéines virales. Contre toute attente, il apparaît donc que la transcription antisens est une stratégie répandue chez les rétrovirus, tout comme elle semble l'être chez de nombreux autres organismes. Les observations publiées par Miller étaient donc pertinentes : puisque des ORF conservés semblent également être présents dans le brin complémentaire des génomes de plusieurs autres membres de la famille des rétrovirus, il est possible que d'autres protéines virales codées par des transcrits antisens soient toujours non identifiées.

8.2 La détection de transcrits antisens par des approches de RT-PCR

Un élément majeur à considérer dans l'étude des transcrits antisens consiste en la méthode utilisée pour prouver leur existence. Les approches utilisant les hybridations de type *Northern blot* peuvent s'avérer efficaces et spécifiques, mais leur niveau de sensibilité est souvent insuffisant pour la détection de transcrits peu abondants. L'application d'une approche spécifique de RT-PCR devient alors nécessaire et bien que cette stratégie ait été employée par certains, la seule utilisation d'une amorce spécifique pour amorcer la réaction de transcription inverse n'est pas suffisante pour produire une synthèse d'un ADNc unique. En effet, la présence d'amorçage endogène au cours de cette étape constitue un obstacle

majeur à la détection brin-spécifique et plusieurs méthodes ont été proposées pour éviter l'obtention des faux-positifs associés à ce phénomène, qui résultent souvent de l'amplification non-spécifique du transcrit sens. [227, 229, 231, 277-279]. La méthode qui a été développée au cours de cette étude pour détecter la présence de transcription antisens chez le VIH-1 a déjà été proposée pour la détection brin-spécifique chez d'autres virus, dont notamment le virus de l'hépatite C [280]. Elle consiste essentiellement à utiliser un oligonucléotide flottant pour effectuer la synthèse de l'ADNc à partir du transcrit d'intérêt. Cette amorce flottante comprend une portion non-spécifique à son extrémité 5', permettant donc par la suite d'effectuer une amplification PCR spécifique pour le cDNA synthétisé à partir de l'amorce ajoutée à la réaction (figure 28). Dans l'étude de Landry *et al.*, cette approche est utilisée avec succès pour identifier la présence de transcription dans plusieurs lignées infectées ou transfectées. De plus, nos résultats démontrent clairement que l'utilisation d'une approche classique de RT-PCR est généralement inefficace et que dans une telle situation, l'amplification non-spécifique du transcrit sens est très probable. On serait donc tenté de croire que certains groupes qui se sont à l'origine intéressés à la transcription antisens ont pu être victimes des mêmes artefacts sans nécessairement réaliser l'ampleur du problème. Il apparaît qu'un tel obstacle ait pu nuire à la crédibilité de certains résultats et donc éventuellement à celle du phénomène en entier. Dans le cas du VIH-1 particulièrement, il s'agit en effet d'un point important, considérant l'absence (jusqu'à maintenant) d'épissage dans le transcrit antisens.

La caractérisation des transcrits antisens des rétrovirus humains HTLV est quant à elle beaucoup moins affectée par de tels artefacts : la présence d'un intron permet la détection spécifique et la quantification des transcrits épissés par l'application de méthodes classiques de RT-PCR. Par contre, de grandes précautions devraient tout de même être prises lorsque ces mêmes approches sont utilisées pour étudier l'expression de transcrits non-épissés qui pourraient entraîner la production de différents isoformes des protéines antisens [248]. La présence d'un codon initiateur qui permettrait la production de différents isoformes des protéines antisens en absence d'épissage chez HTLV-1 est en effet intrigante. De plus le fait que les deux isoformes d'HBZ (original vs SP1) présentent une localisation nucléaire distincte laisse croire que ces modifications des acides aminés en N-

terminal pourraient influencer la fonction de ces protéines. Néanmoins, l'évaluation de l'abondance relative des transcrits épissés vs non-épissés nécessitent l'application de protocoles mieux adaptés que ceux qui ont été présentés jusqu'à maintenant.

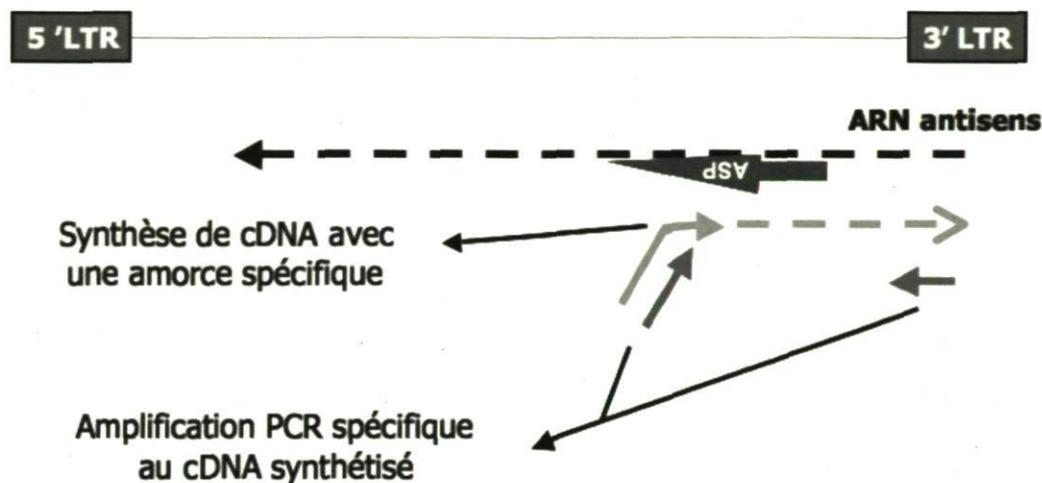


Figure 28 : Stratégie de RT-PCR brin-spécifique utilisé pour la détection du transcrit antisens chez le VIH-1. La stratégie de RT-PCR développée au cours de cette étude repose principalement sur l'utilisation d'une amorce spécifique lors de l'étape de transcription inverse. Cette amorce contient une portion non-spécifique à son extrémité 5' (portion flottante) qui permet d'effectuer une amplification PCR spécifique pour l'ADNc synthétisé à l'aide de cette même amorce.

8.3 La régulation de la transcription antisens chez HTLV-1

Au moment où nous avons caractérisé le transcrit antisens chez HTLV-1 et identifié la présence d'un épissage alternatif, une seconde étude proposait l'existence d'une forme de transcrit non-épissée permettant de produire la protéine HBZ initialement décrite [248]. Selon cette étude, les transcrits non-épissés et les transcrits épissés SP1 sont exprimés à un

niveau similaire dans les cellules primaires ATL ainsi que dans les lignées cellulaires dérivées de cellules ATL. Dans une étude subséquente du même groupe, effectuée sur des cellules de plusieurs patients séropositifs, le transcrit épissé SP1 est plutôt évalué comme étant 4 fois plus abondant que le transcrit non-épissé [250]. Ces variations traduisent sans aucun doute les difficultés associées à l'amplification spécifique du transcrit non-épissé, tel que mentionné ci-haut.

Peu après la soumission de notre étude sur l'effet de la protéine Tax sur la transcription antisens chez HTLV-1 [281], un premier article traitant du même sujet paraissait [282]. Les résultats présentés dans cet article corroborent ceux que nous avons obtenus concernant l'effet positif de Tax sur le promoteur d'HBZ et l'implication des éléments TRE1. Par contre, dans cette étude, les auteurs rapportent également l'existence de plusieurs sites d'initiation de transcription situés dans la région pX. Tout comme celui initialement décrit par Cavanagh *et al.*, ce deuxième promoteur serait sans boîte TATA. Il contiendrait cependant les éléments Inr et DPE et permettrait donc la production de l'isoforme d'HBZ traduit à partir du transcrit non épissé. Suite à plusieurs expériences de transfection, les auteurs concluent que la capacité de cet isoforme à inhiber l'action de Tax sur le promoteur viral est moindre que celle de l'isoforme HBZ-SP1. De plus, Yoshida *et al.* avancent que l'isoforme non-épissé serait exprimé en moins grande quantité à partir de l'ADNc transfecté et serait possiblement moins stable que l'isoforme HBZ-SP1.

Ces études sont basées sur la prémisse selon laquelle le transcrit non-épissé d'HBZ est exprimé dans les cellules infectées. S'il est possible que ce soit effectivement le cas, il n'en reste pas moins que des analyses supplémentaires doivent être effectuées pour clarifier la fréquence de ce transcrit. De plus, d'autres types d'expériences devraient être planifiées pour démontrer l'expression des deux isoformes d'HBZ à partir d'un ADN proviral complet. Par exemple, l'ajout d'étiquettes différentes en amont des codons ATG initiateurs respectifs des deux isoformes pourrait faciliter la détection spécifique de ces protéines en *Western blot*.

Il est intéressant de noter que, selon Yoshida *et al.*, HBZ-S1 est exprimé de façon ubiquitaire, indépendamment de la présence de Tax. Cette conclusion est tirée suite à 1)

l'identification d'un site Sp1 important dans l'activité promotrice basale du promoteur de HBZ-SP1 2) la présomption que la séquence promotrice du transcrit épissé est située en amont du codon initiateur ATG situé dans l'exon 1 d'HBZ SP1. Selon cette théorie, le promoteur du transcrit non-épissé est donc le seul à être modulé de façon importante par l'action de Tax sur les TRE1. Pourtant, il est pour le moins curieux qu'un potentiel rôle des éléments TRE1 dans l'expression du transcrit épissé soit exclu d'emblée sous prétexte que ces derniers sont situés dans l'intron. En effet, autant les séquences introniques que exoniques peuvent avoir un rôle à jouer dans la régulation transcriptionnelle et les descriptions de promoteurs introniques ne sont pas rares [283-285].

Notre analyse de l'action de Tax sur la transcription antisens a été effectuée en grande partie en utilisant notre construction K30-3'asLuc, dans laquelle le gène luciférase a été inséré dans l'exon 2 de HBZ-SP1. Dans cette construction, l'expression du gène luciférase peut donc résulter de l'expression d'un transcrit épissé ou non. Par contre, les expériences effectuées dans les cellules Tax-négatives S1T démontrent bien que l'expression de Tax entraîne une augmentation de la quantité de transcrits épissés. De façon intéressante, les inductions obtenues lorsque la protéine Tax est exprimée sont en général beaucoup plus importantes lorsque les expériences sont effectuées dans le contexte d'un ADN proviral intégré. Les capacités de Tax à recruter les histones acétyl-transférase CBP/p300 [286] et à entraîner l'exclusion des nucléosomes du promoteur viral [287] sont probablement mieux représentées dans notre modèle de cellules stables que lors d'expériences de transfections transitoires impliquant des constructions plasmidiques.

De récentes études ont démontré la capacité de la protéine accessoire p30^{II} à réguler la transcription virale induite par Tax en interagissant directement avec CPB/p300 [110]. Il serait donc intéressant d'analyser l'effet de p30^{II} sur le promoteur d'HBZ. Dans un autre ordre d'idées, il serait pertinent de poursuivre l'analyse du promoteur antisens de façon à pouvoir identifier les éléments qui permettent l'expression d'HBZ en absence de Tax, puisqu'il s'agit d'une situation fréquente dans les cellules ATL. De plus, des expériences effectuées à partir de cellules primaires de patients infectés (malades ou non) pourraient permettre d'établir un lien entre le niveau d'expression d'HBZ et l'apparition de l'ATL.

8.4 Rôle de la protéine HBZ dans le développement de l'ATL

De façon générale, la protéine Tax est considérée comme un des acteurs essentiels à la transformation cellulaire provoquée par HTLV-1. Par contre, il est connu que son expression dans les cellules ATL est souvent supprimée par différents mécanismes, tels que l'hyperméthylation du LTR 5' [288, 289], la présence de mutations dans le gène *tax* [290] ou la suppression du LTR 5' [291]. Puisque la protéine Tax est une cible importante des lymphocytes T cytotoxiques (CTLs) de l'hôte, son expression est probablement désavantageuse pour une cellule infectée : l'émergence d'une cellule transformée qui n'exprime plus Tax est donc favorisée. Il a en effet été proposé que la perte d'expression de Tax survient au cours du processus d'oncogénèse.

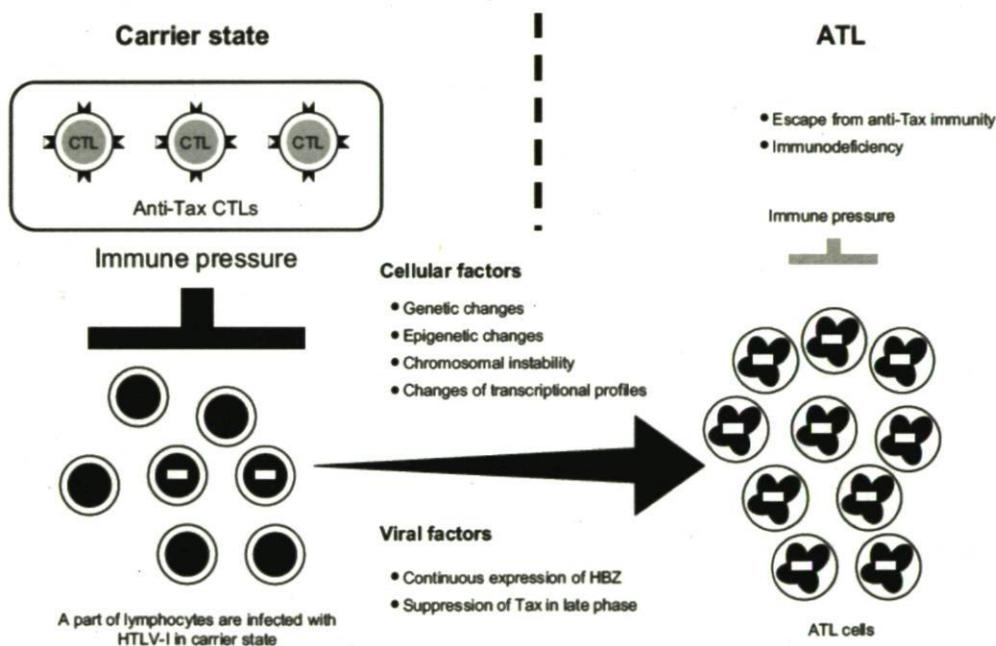


Figure 29 : Les étapes menant à l'oncogénèse induite par HTLV-1. La prolifération clonale des cellules T infectées requiert l'expression de Tax et HBZ. La réponse immune contrôle la prolifération de ces cellules et un équilibre s'établit entre la division cellulaire et l'élimination par les CTLs anti-Tax. Durant la longue phase de latence, plusieurs changements surviennent au niveau cellulaire et certaines cellules transformées ne dépendent plus de l'expression de Tax. La suppression de l'expression de Tax ou la présence d'immunosuppression chez l'hôte conduisent à la prolifération cellulaire incontrôlée et au développement de l'ATL. (tiré de [292])

De façon intéressante, de récentes études ont suggéré que le gène *HBZ* serait exprimé dans toutes les cellules ATL [247, 248, 293]. Il apparaît donc que si l'expression de Tax est associée à la transformation cellulaire, celle d'*HBZ* pourrait quant à elle être requise pour la croissance des cellules ATL. En effet, tant le transcrit que la protéine *HBZ* se sont vu attribuer des rôles dans la prolifération cellulaire [247, 256]. Par contre, des expériences effectuées sur des cellules primaires à l'aide de virus mutants pour l'expression d'*HBZ* ont démontré que la protéine n'est pas nécessaire à l'immortalisation des cellules [255]. La caractérisation du rôle exact du gène *HBZ* dans le processus de transformation cellulaire menant au développement de l'ATL nécessite donc des analyses supplémentaires.

L'identification récente de NF- κ B comme nouveau partenaire d'interaction de la protéine *HBZ* suggère que cette dernière pourrait agir à plusieurs niveaux sur la régulation transcriptionnelle cellulaire et donc possiblement collaborer avec Tax dans le processus d'oncogenèse [294]. Des résultats récemment obtenus par Banerjee *et al.* supportent cette hypothèse : dans un modèle de souris reconstituées SCID/NOD, l'infection par un virus mutant pour *HBZ* entraîne l'apparition de lymphomes présentant un phénotype immature CD4-CD8-CD90+, en opposition avec le phénotype mature CD4+CD25+ CD90- obtenu lors d'infection avec un virus sauvage [83]. La figure 29 présente les différentes étapes de l'oncogenèse induite par HTLV-1.

La démonstration de l'action de Tax sur le promoteur du gène *HBZ* met en évidence l'existence d'une boucle de rétroaction permettant au virus de contrôler l'expression de ses gènes et donc d'éviter une expression exagérée de la protéine Tax. Il s'agit là d'une possible stratégie servant à éviter une forte réponse immune dirigée contre les cellules infectées. De façon intéressante, nos résultats démontrent que le site d'intégration dans le génome de la cellule hôte pourrait influencer de façon importante le niveau basal d'expression du transcrit d'*HBZ* ainsi que l'induction du promoteur antisens lorsque Tax est exprimé. Ces résultats suggèrent que le contexte d'ADN cellulaire pourrait influencer la balance de l'expression entre les gènes *tax* et *HBZ*. Il est tentant de penser que l'intégration dans un site favorisant l'activité du promoteur d'*HBZ* pourrait entraîner la survie cellulaire en prévenant l'expression de Tax. Toujours dans la même optique, on peut penser qu'un

certain rapport entre le niveau d'expression de Tax et HBZ pourrait être optimal pour la sélection à long terme de cellules transformées et donc que certains sites d'intégration pourraient avoir un impact direct sur la probabilité de développer la maladie.

8.5 Transcription antisens chez HTLV-3 et HTLV-4

D'un certain point de vue, la démonstration de l'existence de transcrits antisens chez les rétrovirus HTLV-3 et HTLV-4 n'est pas surprenante : l'analyse de séquence du brin complémentaire de leurs génomes avait rapidement permis d'identifier la présence de cadres de lectures ouverts dont la séquence en acides aminés présentait des similarités avec celle de la protéine HBZ. De plus, des analyses effectuées récemment dans nos laboratoires avaient aussi permis d'identifier la présence d'un transcrit antisens épissé chez le rétrovirus HTLV-2 (Halin *et al.*, soumis). Nos observations suggèrent donc que l'ensemble des virus PTLVs posséderaient la capacité de produire une protéine virale *via* l'expression d'un transcrit antisens. Les expériences qui ont été effectuées par RT-PCR ont ultimement permis de confirmer notre hypothèse de départ et de caractériser l'épissage des transcrits antisens chez les rétrovirus humains HTLV-3 et HTLV-4. Comme chez HTLV-1, le premier exon de ces transcrits est situé dans le LTR 3' et permet de recruter un codon initiateur ATG. Fait intéressant, l'absence de codon initiateur dans les séquences introniques du transcrit antisens des deux virus suggère que l'épissage est essentiel à la production des protéines APH-3 et APH-4. Par contre, la présence de méthionine au début de la séquence des deux protéines (acide aminé 15 pour APH-3 et acide aminé 25 pour APH-4) nous force à considérer la possibilité que des isoformes tronqués de la protéine pourraient être produits par des transcrits non-épissés.

Tel qu'initialement suggéré par Switzer *et al.* [259], ces nouvelles protéines possèdent un domaine basique riche en arginine suivi d'une séquence prédite pour former un domaine *coiled-coil*. Tout comme il a été observé chez HBZ, il est fort possible que les régions basiques soient impliquées dans la localisation de ces protéines au noyau, tel qu'observé en

microscopie confocale. Lors de ces mêmes analyses, nous avons également pu observer que la protéine APH-3 se retrouvait aussi au cytoplasme, sous forme de structures ponctiformes uniformément dispersées. Des analyses additionnelles impliquant la génération de plusieurs mutants pourront clarifier le rôle de différents domaines dans ces localisations distinctes, qui suggèrent donc la présence d'un motif d'exportation nucléaire fonctionnel dans la séquence d'APH-3. De plus, il serait pertinent de s'assurer que la présence de la protéine au cytoplasme ne constitue pas un artéfact de surexpression : la construction d'un ADN proviral complet d'HTLV-3 exprimant la protéine APH-3 en fusion GFP pourrait permettre de visualiser la localisation de la protéine dans un contexte où elle est exprimée à un niveau qui s'approche davantage de celui présent lors de l'infection virale. Dans le cas où la protéine se retrouverait toujours au cytoplasme lorsque son expression dépend du promoteur viral situé dans le LTR 3', il deviendrait alors intéressant d'identifier la nature de ces structures cytoplasmiques (par microscopie confocale) et éventuellement d'étudier le rôle de ces dernières dans la réplication. En effet, cette localisation spécifique de la protéine APH-3 suggère que son activité cellulaire pourrait différer de celle décrite jusqu'à maintenant pour HBZ et donc que des partenaires d'interaction différents pourraient possiblement être identifiés.

Une divergence majeure entre les HBZ et les protéines APH-3 et APH-4 réside dans l'absence d'un motif *leucine zipper* conservé chez ces dernières. Chez HBZ, ce motif est précédé de deux régions basiques riches en arginine d'un domaine de liaison à l'ADN (voir figure 22). Les motifs de dimérisation *leucine zipper* forment des *coiled-coils* parallèles caractérisés par la présence de résidus leucines en position **d** des heptades répétées (**abcdefg**)_n (figure 30). Le motif contient généralement 4 ou 5 heptades répétées, chacune formant deux tours en hélice alpha de 7 acides aminés [295, 296]. Un domaine *coiled-coil* est typiquement composé de deux ou plusieurs hélices alpha capables d'interagir entre elles par l'intermédiaire de forces hydrophobes entre les acides aminés aux positions **a** et **d**. Les domaines *coiled-coil* sont souvent impliqués dans la formation d'homo ou d'hétéro-oligomères [297].

L'alignement de la séquence en acides aminés de la protéine HBZ avec celle des protéines antisens des autres virus HTLVs permet notamment de mettre en évidence des différences au niveau des potentiels motifs bZIP. En effet, chez HTLV-1, le *leucine zipper* possède 5 heptades également retrouvées chez de nombreux facteurs transcriptionnels de type bZIP [298]. Par contre, chez les autres virus, ce motif est incomplet, notamment en raison du décalage d'une leucine en C-terminal (voir figure 32). La figure 31 présente la structure d'un facteur bZIP lié à l'ADN.

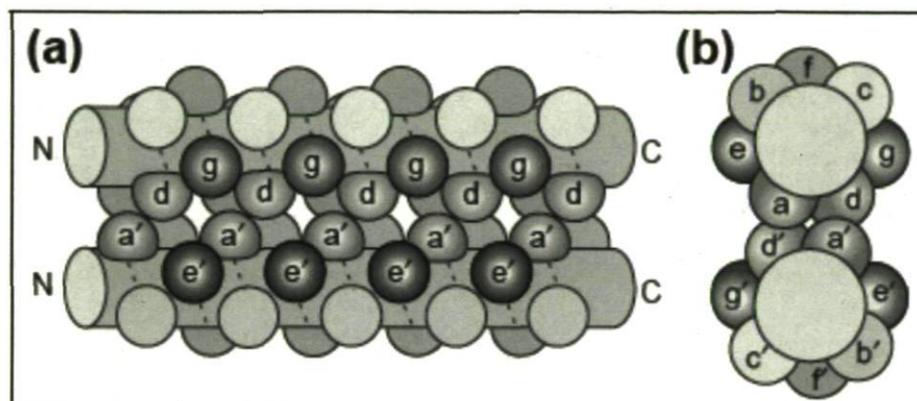


Figure 30 : Représentation schématique d'un domaine *coiled-coil* en double brin parallèles. Dans un domaine *coiled-coil* l'interface entre les deux hélices alpha est formé par les acides aminés en position **a, d, e** et **g**. (a) Vue de côté. (b) Vue de haut. (tiré de [299])

Néanmoins, les séquences des protéines APH-3 et APH-4 sont prédites pour former un domaine *coiled-coil* avec une forte probabilité. (Résultats non-présentés, obtenus à l'aide du serveur web Paircoil2 <http://groups.csail.mit.edu/cb/paircoil2/paircoil2.html> [300]). Ces observations laissent croire que malgré l'absence d'un motif *leucine zipper* typique, ces protéines pourraient contenir un motif *leucine zipper-like* fonctionnel qui permettrait des interactions avec certains facteurs transcriptionnels. Considérant la capacité des protéines APH-3 et APH-4 à inhiber l'effet de Tax sur le promoteur d'HTLV-1, il est logique de penser que ces protéines pourraient donc s'hétérodimériser avec des facteurs transcriptionnels de type bZIP pour accomplir leur action.

Par contre, il est possible que l'absence de certains résidus leucine non conservés entraînent des changements d'affinité ou de spécificité qui influenceraient l'activité des ces nouvelles protéines. Des expériences de double hybride ou d'immunoprécipitation pourraient permettre de clarifier la capacité de ces motifs *leucin zipper-like* à se dimériser avec ceux des facteurs de transcription déjà connus pour interagir avec HBZ. Aussi, la génération de mutants de délétion permettrait d'établir hors de tout doute le rôle de ce domaine dans l'inhibition de l'effet des protéines Tax1 et Tax 3.

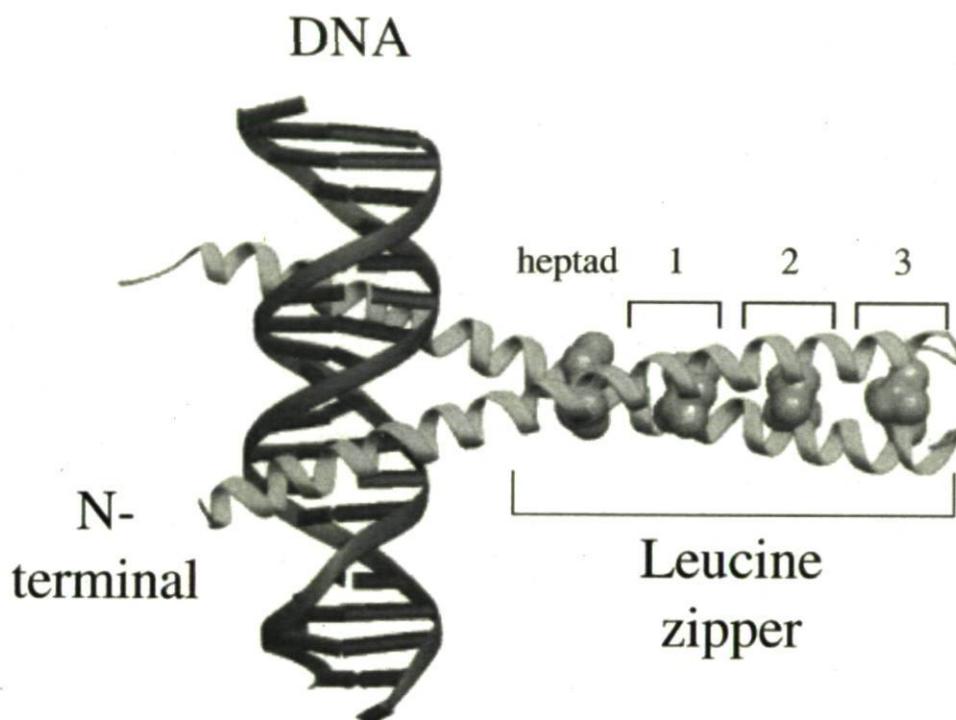


Figure 31 : Structure d'un homodimère bZIP lié à l'ADN. Les leucines conservées présentes à tous les deux tours d'hlice alphasont représentés en gris. (tiré de [298])

Il est possible que les motifs LXXLL présents chez APH-3 et APH-4 soient aussi impliqués dans l'effet observé via une interaction directe avec CPB/p300 et leur importance fonctionnelle devra aussi être considérée. Dans le même ordre d'idée, il serait intéressant d'évaluer la capacité relative des différents protéines antisens à inhiber la transcription dépendante de Tax à partir des promoteurs viraux obtenus de leurs virus respectifs. On

pourrait effectivement émettre l'hypothèse que la présence de variations dans la séquence protéique des protéines APH- 3 APH-4 reflète une adaptation qui vise à réguler efficacement leur propre promoteur.

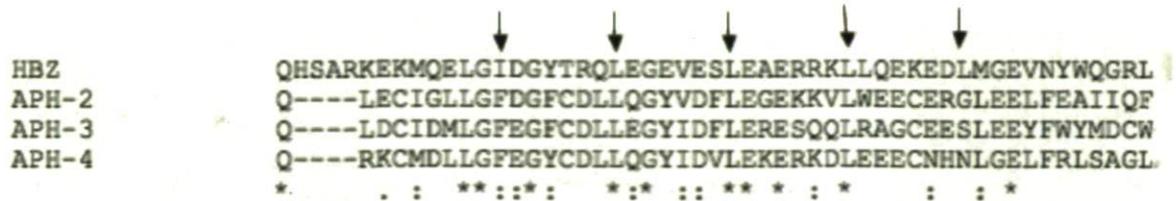


Figure 32 : Alignement des motifs *leucine zipper* des protéines antisens retrouvées chez les rétrovirus HTLVs. Les flèches indiquent les résidus leucine conservés dans la séquence de la protéine HBZ.

8.6 Transcription antisens chez le VIH-1

L'étude de la transcription antisens chez les rétrovirus HTLVs a sans doute été facilitée par deux facteurs non négligeables : 1) l'existence de transcrits épissés et 2) la présence de domaines conservés dans la séquence en acides aminés des ORF présents sur ces transcrits. Le premier a permis d'écarter rapidement la possibilité que la présence de transcrits antisens dans les cellules infectées ou transfectées pourrait résulter d'une fuite de la machinerie transcriptionnelle. Le deuxième a pu donner des pistes qui ont permis d'attribuer des rôles et des partenaires potentiels aux protéines HBZ, APH-3 et APH-4. Ensemble, ces facteurs ont rapidement entraîné l'exclusion de l'hypothèse selon laquelle la transcription antisens chez les virus HTLVs pourrait en fait constituer un artéfact de biologie moléculaire.

Malheureusement pour ceux qui ont osé s'y aventurer, l'étude de la transcription antisens chez le VIH-1 ne présente aucun de ces avantages. Bien que nos méthodes de RT-PCR se soient révélées efficaces pour détecter et caractériser le transcrit, aucune trace d'épissage n'a pu être identifiée. Conséquemment, la question de la pertinence biologique du transcrit antisens chez le VIH-1 ne peut être écartée. L'utilisation de virus pseudotypés capables d'exprimer la luciférase sur le brin complémentaire de leur génome a sans doute permis de répondre partiellement à cette question. En effet, l'expression du gène rapporteur observée lors d'expériences d'infection avec notre clone NL4.3asLuc démontre clairement que le transcrit antisens est exprimé au cours du cycle de réplication virale et qu'il est apte à servir de gabarit pour la traduction de protéine. Il apparaît néanmoins que ce transcrit est peu fréquent : les niveaux d'activité luciférase obtenus en infection sont généralement 100 à 1000 fois moins élevés que ceux obtenus avec un des virus NL4.3Luc. Cette différence n'est pas surprenante, considérant la possibilité que ce transcrit soit responsable de l'expression d'une seule protéine virale. Par contre, il est possible que cette différence soit variable selon la cellule infectée et que le site d'intégration puisse influencer le rapport entre la transcription sens et antisens. L'identification d'un signal de polyadénylation consensus et fonctionnel apporte également un argument de taille pour la fonctionnalité du

transcrit. À noter que ce signal est différent de celui proposé par Michael *et al.* [264] et que notre caractérisation des transcrits ne concorde pas avec la taille des signaux obtenus en *Northern blot* par Bukrinsky et Etkin [263]. Plusieurs facteurs pourraient expliquer ces contradictions et l'existence d'un possible épissage en 3' de l'ORF décrit devrait faire l'objet d'analyses plus approfondies.

De façon intéressante, nos résultats démontrent que la protéine Tat régule positivement la transcription antisens. Ces résultats sont en contradiction avec ceux obtenus précédemment par deux groupes [172, 264]. Notre approche démontre donc que l'utilisation de lignées stables permet d'éviter les problèmes d'interférence qui peuvent être associés à l'étude de promoteurs bidirectionnels à partir de vecteurs plasmidiques circulaires. Par contre, il serait intéressant de confirmer ces résultats dans un contexte où les deux séquences LTR sont présentes. Par exemple, la construction d'un ADN proviral complet contenant deux gènes rapporteurs luciférase (*Firefly* et *Renilla*) permettrait d'évaluer l'action de Tat dans une situation où les deux promoteurs peuvent entrer en compétition. Considérant la capacité de Tat à moduler l'activité du facteur transcriptionnel Sp1 [301], il serait également pertinent de vérifier le rôle du site de liaison Sp1 identifié par Peeters *et al.* dans la transcription antisens induite par Tat.

8.7 La capacité codante du transcrit antisens chez le VIH

D'un point de vue théorique, la présence d'un ORF conservé dans le brin complémentaire du génome du VIH-1 constitue un argument en faveur d'un rôle pour la transcription antisens chez le VIH-1. Par contre, l'analyse de la séquence en acides aminés de cette protéine nommée ASP donne peu d'indices quant à sa fonction potentielle dans la réplication du virus. En effet, de nombreuses analyses *in silico* ont été effectuées pour tenter de mieux comprendre les caractéristiques de cette protéine prédite comme étant particulièrement hydrophobe (voir figure 33).

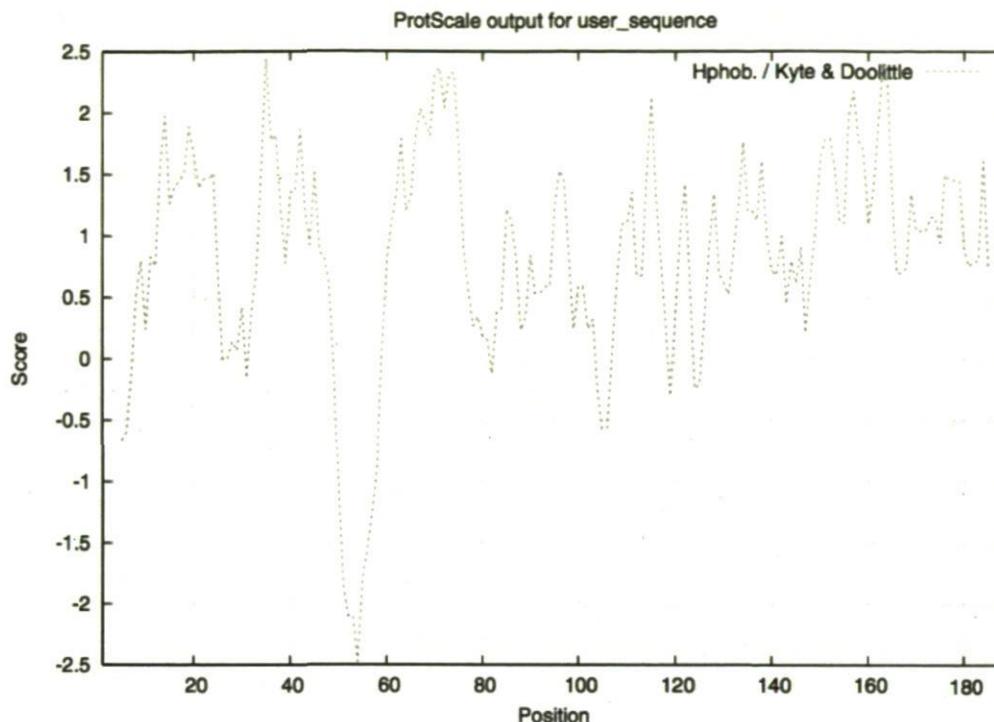


Figure 33 : Profil d'hydrophobicité de la protéine ASP. Prédiction de l'hydrophobicité de la protéine selon l'échelle établie par Kyte et Doolittle [302] en utilisant l'interface web ProtScale (ExPASy).

La protéine posséderait donc une extrémité N-terminale riche en résidus cystéines, un pic hydrophile (situé aux environs des acides aminés 42 à 62) et possiblement 3 domaines transmembranaires hautement hydrophobes. Le nombre de domaines transmembranaires prédits varie selon le logiciel utilisé. En règle générale, 3 domaines sont identifiés, tels qu'initialement proposé par Vaquero *et al.*) (voir figure 34). La région hydrophile est riche en prolines et contient un motif PXXP. Les motifs PXXP sont souvent retrouvés chez les protéines qui interagissent avec des domaines d'homologie SH3 (*Src-homology 3*). Ces domaines d'environ 60 acides aminés sont fréquents dans la séquence de protéines impliquées dans la signalisation cellulaire et de protéines du cytosquelette [303-305]. Les motifs PXXP de la protéine Nef du VIH-1 ont notamment été décrits comme étant nécessaires à l'interaction avec le domaine SH3 des kinases Src [306, 307].

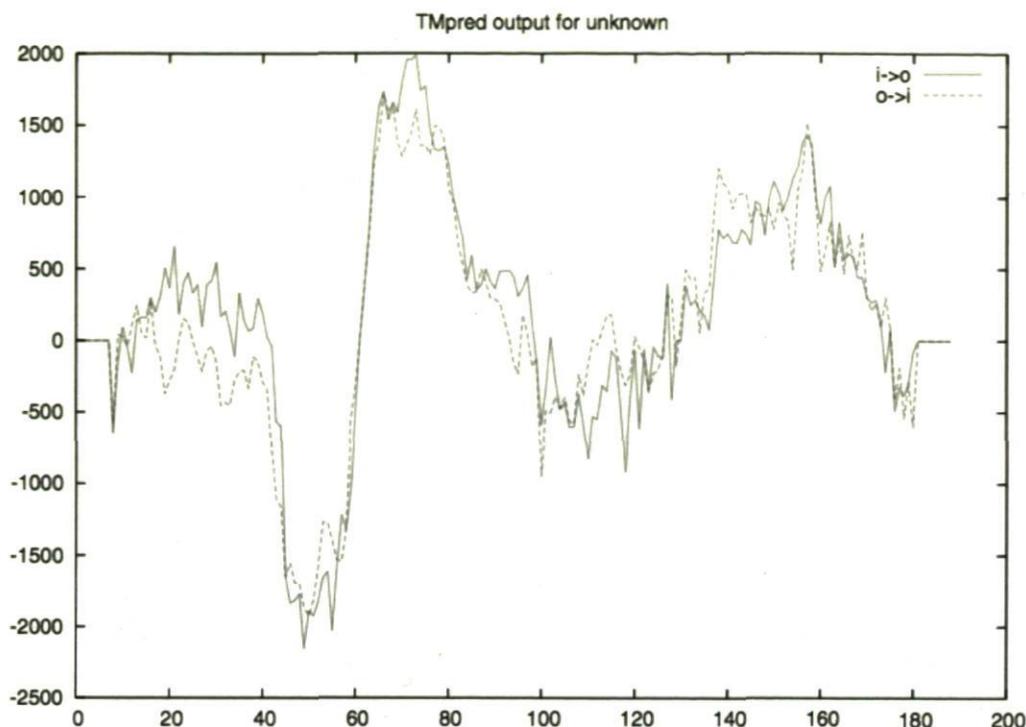


Figure 34 : Prédiction des hélices transmembranaires de la protéine ASP. Selon les modèles utilisés pour effectuer les analyses, 2 ou 3 domaines transmembranaires sont prédits. Le modèle qui prévaut selon le serveur web TMpred (ExPASy) prédit la présence de 3 domaines transmembranaires (présentés ci-haut, valeur plancher de 500) où l'extrémité N-terminale serait intracellulaire.

De façon intéressante, la présence d'un ORF conservé a également été décrite chez d'autres lentivirus, tels que le virus de l'immunodéficience féline (VIF), le virus de l'immunodéficience bovine (VIB), le VIH-2 ainsi que chez l'ancêtre proposé du VIH, le VIS du chimpanzé [243]. Alors que les ORF antisens des trois premiers virus semblent être différents de celui identifié chez le VIH-1, celui du VIScpz présente au contraire d'importantes similitudes (voir figure 35).

```

ASO_HIV_NL4.3  MPQTVSCNRCCCASIALSKLFCCTISDNNCLACTVSVIDVQPIVLPAAAPKNPRNKAPIP
ASO_SIVcpzMB66 MPQTESCNRCCASMAALSRLCCCTIPDSSCLACTVSVSDAAPIVLPAAAPKNPRNKAPNP
*****:***:*:*****.*..***** * . *****

ASO_HIV_NL4.3  T-ALFSLCTLLFALVGATPNGSIFTTLYLYNSLLQLSLISPPPGLKISDPLLLLPPSIV
ASO_SIVcpzMB66 TAALFSLCTVLLFALVGATPKGSILTLYLYNSLLQLFLISPPPTGVNMADLFILSVFCCA
* *****:*****:***:***** *****.*:::* ::* . .

ASO_HIV_NL4.3  NSSPVIFDEHLICPLMGGAYIAFPTSCHMFINCFILHGSVIVSLPSVLFDPVSLQVLLNQ
ASO_SIVcpzMB66 NNIPVIFEIQFIVPLMGGRNIPPLTLCHSLTICLILQGIVIL---SSLVSLKMMEVKSM
*. *****: ::* ***** *.:** ** : *:::* ** : * *.. .:::** .

ASO_HIV_NL4.3  VLLNSCVELQ
ASO_SIVcpzMB66 EVL-----Q
:*          *

```

Figure 35 : Aligement de la séquence en acides aminés des ORF antisens du VIH-1 (clone NL4.3) et VIScpz (clone MB66).

8.8 Localisation cellulaire de la protéine ASP

Une seule étude a jusqu'à maintenant été publiée sur la localisation cellulaire de la protéine ASP [267]. Ces analyses ont été effectuées par microscopie électronique sur des cellules transfectées et infectées. Dans les cellules A3.01 transfectées, les signaux obtenus sont décrits comme étant surtout associés aux membranes internes, particulièrement au niveau des mitochondries. Des signaux positifs sont observés de part et d'autre des membranes plasmiques, nucléaires et mitochondriales. De plus, le marquage des cellules SupT1 infectées permet d'observer la présence de signaux positifs à la surface des nouveaux virions.

Récemment, nous avons tenté d'évaluer la localisation cellulaire de la protéine ASP en utilisant des méthodes de microscopie confocale. Le modèle des cellules SF9 a tout d'abord été utilisé puisque l'infection par un baculovirus recombinant permet un haut niveau d'expression de la protéine. De façon intéressante, nos premières observations semblaient confirmer les résultats présentés par l'équipe de Vaquero : les signaux paraissaient surtout concentrés au niveau de la membrane plasmique, bien que la protéine semble aussi être présente au cytoplasme. La localisation précise de la protéine ASP dans ces cellules devrait

cependant être confirmée par des analyses complémentaires de co-localisation avec des marqueurs spécifiques des membranes plasmiques et du noyau.

Considérant la facilité à produire la protéine en grande quantité dans les cellules SF9, il pourrait être intéressant d'utiliser ce modèle pour évaluer la capacité de la protéine à se retrouver en surface des particules virales. Par exemple, la co-transfection de ces cellules à l'aide de vecteurs exprimant ASP et la protéine de capsid du VIH-1 permettrait la production de VLPs (la production de VLPs à partir des cellules SF9 a déjà été rapportée [308]). Des techniques d'immunocapture ou tout simplement de *Western blot* pourraient par la suite être utilisées pour détecter la présence de la protéine ASP à la surface ou à l'intérieur de ces VLPs.

À la lumière des résultats obtenus dans les SF9, des expériences de microscopie confocale ont été entreprises sur les lignées cellulaires 293T et COS-7. Si les résultats préliminaires laissent croire que la protéine pourrait être majoritairement cytoplasmique dans ces lignées, il apparaît cependant prématuré de tirer des conclusions sans avoir effectué de co-marquages. Des analyses supplémentaires sont donc nécessaires et pourraient permettre d'identifier les compartiments cytoplasmiques préférentiellement ciblés par la protéine. En effet, plusieurs cellules semblent présenter des structures ponctiformes fortement positives. D'un autre côté, le signal est absent du noyau, ce qui concorde avec l'absence de NLS dans la séquence en acides aminés de la protéine.

Les expériences de microscopies ont pu également mettre en évidence le faible taux d'expression obtenu lorsque les cellules 293T ou COS-7 sont transfectées avec notre vecteur d'expression pcDNA-ASP. De façon récurrente, le nombre de cellules positives était évalué visuellement à moins de 1%. Il était alors impossible de déterminer si ces observations résultaient effectivement d'un faible niveau d'expression ou de l'incapacité à détecter efficacement la protéine à l'aide des anticorps utilisés. Ce problème rendait donc les analyses de microscopie laborieuses et laissait souvent planer un doute sur la spécificité des signaux obtenus. Selon toute vraisemblance, le faible pourcentage de cellules positives expliquait notre incapacité à détecter la protéine par *Western blot*.

8.9 L'optimisation des codons permet l'expression efficace de la protéine ASP

Plusieurs hypothèses ont donc été lancées pour expliquer ce phénomène et quelques approches ont été mises de l'avant pour tenter de contourner le problème. L'utilisation de plusieurs protocoles de *Western blot* mieux adaptés à l'extraction de protéines membranaires et l'ajout d'un inhibiteur du protéasome (MG132) se sont d'abord avérés inefficaces. Puis, nous avons opté pour une stratégie qui a déjà été employée pour permettre la surexpression de certaines protéines du VIH-1 : l'optimisation des codons [276, 309].

L'optimisation des codons a été décrite à l'origine comme une stratégie permettant d'augmenter l'efficacité de traduction de certains gènes contenant des codons non optimaux pour les types cellulaires transfectés. Cependant, un nombre croissant d'études démontrent que l'effet positif observé est davantage attribuable à des changements au niveau de la stabilité et l'exportation des ARNm, plutôt qu'à une traduction plus efficace. En effet, Ngumbela *et al.* présentaient récemment une étude dans laquelle la transfection d'ARNm codant pour la protéine Gag est utilisée pour évaluer l'influence réelle de l'optimisation des codons sur l'efficacité de traduction du gène. Les résultats rapportés confirment que l'impact majeur de l'utilisation de cette approche se situe au niveau de la stabilité et de l'export des ARNm [275].

Au cours de la réplication virale, l'expression des protéines Gag et Pol est dépendante de la présence de la protéine Rev, qui permet l'exportation au cytoplasme des transcrits non-épissés. En absence de Rev, il semble que la présence des séquences d'instabilité ou d'éléments répresseurs nuise à l'expression efficace de ces protéines. [310, 311]. Ces régions sont généralement décrites comme étant riches en A et U et la présence de motif AUUUA semble jouer un rôle majeur dans la rétention nucléaire des transcrits non-épissés ou mono-épissés. Alors que ces éléments AU sont le plus souvent situés dans les régions 3'UTR des ARNm cellulaires, ils sont exceptionnellement retrouvés dans les séquences codantes des gènes chez le VIH-1. La mutagénèse de ces régions dans les gènes *gag* et *pol* permet la production de particules virales en absence de Rev [311].

Tel qu'observé dans le cas des protéines de la capsid et de l'intégrase, l'optimisation des codons permet d'augmenter considérablement l'expression de la protéine ASP dans les cellules 293T et COS-7. Bien que le nombre de cellules positives n'ait été évalué que de façon qualitative en microscopie confocale, la différence entre le taux d'expression obtenu avec la construction contenant le gène sauvage est marquée. En utilisant le vecteur d'expression contenant l'ADNc optimisé (pcDNA-ASP-opt), on obtient environ 20-30 % de cellules positives. Il est intéressant de noter que l'ADNc sauvage utilisé contient un motif AUUUA et un pourcentage de codon A/U de 57%. La séquence de l'ADNc optimisé contient plutôt 41% de A/U, ce qui pourrait expliquer (au moins partiellement) l'efficacité de la stratégie.

L'utilisation de notre vecteur d'expression optimisé pourrait s'avérer très utile dans plusieurs aspects de l'étude de la protéine ASP. Dans un premier temps, la possibilité d'exprimer ASP chez des lignées cellulaires humaines devrait faciliter l'étude de sa localisation cellulaire, qui éventuellement pourra être évaluée dans plusieurs lignées de lymphocytes T et de macrophages humains. De plus, il pourrait être intéressant d'exprimer la protéine dans une lignée chroniquement infectée de façon à évaluer si la présence des autres protéines virales peut éventuellement influencer la localisation cellulaire d'ASP.

À la lumière de ces résultats, il serait logique de se questionner sur les mécanismes qui pourraient permettre la production de la protéine ASP *in vivo*. En effet, si l'expression des protéines Gag et Pol en transfection est favorisée par l'optimisation des codons, leur expression efficace durant le cycle de réplication virale est assurée par la présence de la protéine Rev. Or, il est intéressant de noter que l'ORF d'ASP chez le VIH-1 chevauche partiellement la séquence RRE, tout comme c'est le cas chez le VIF (figure 36). Cette séquence ferait donc inévitablement partie du transcrit antisens chez ces deux virus. La structure secondaire du RRE est prédite pour être formée en antisens, mais des études de fonctionnalité du RRE ont démontré que la protéine Rev a une affinité 20 à 100 fois plus importante pour le RRE sens [312, 313]. La possibilité que le RRE antisens puisse jouer un rôle dans l'export des transcrit antisens n'est cependant pas à exclure et des analyses

additionnelles sont requises pour déterminer la capacité de ces derniers à être exportés efficacement au cytoplasme dans un contexte d'infection.

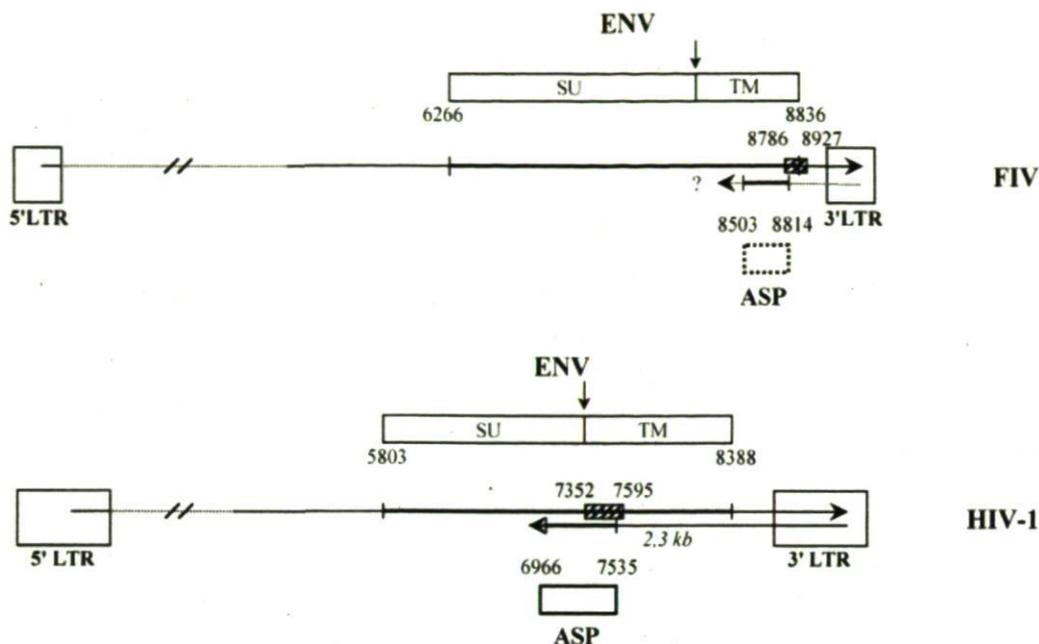


Figure 36 : Position des séquences RRE et des ORF antisens conservés chez le VIH-1 et le VIF. La figure présente un schéma du génome des virus du VIF et du VIH-1. La position des ORF codant pour les protéines d'enveloppe et les protéines antisens est indiquée. Les séquences contenant les éléments RRE sont identifiées par un rectangle hachuré. (adapté de [270])

L'optimisation de la séquence codante de la protéine ASP du VIH-1 a permis sa détection en *Western blot* lors d'expériences de transfections transitoires dans les cellules COS-7. Ceci représente à notre connaissance la première détection de cette protéine dans une lignée cellulaire de mammifère et constitue donc un apport majeur de cette étude. Cependant, les essais effectués à l'aide de protocoles standard d'extraction protéique se sont montrés inefficaces : il semblerait que la détection de la protéine sur gel nécessite une extraction adaptée aux protéines membranaires. Ces observations vont donc de pair avec les résultats de Vaquero *et al.* quant à la localisation de la protéine au niveau des membranes internes de la cellule. Évidemment, ces résultats laissent croire que la détection de la protéine ASP dans un contexte de cellules infectées risque de s'avérer une tâche ardue. Par contre, la

difficulté à détecter la protéine à partir de cellules infectées ne devrait pas permettre de conclure à son inexistence. En effet, certaines protéines accessoires chez HTLV-1 ont fait l'objet de plusieurs études malgré le fait qu'elles n'aient jamais été détectées à partir de cellules infectées [118].

8.10 ASP : perspectives

Convaincre la communauté scientifique de l'existence de la protéine ASP nécessitera 2 étapes cruciales : la démonstration que la protéine peut être exprimée à partir d'un ADN proviral complet ainsi que dans les cellules primaires infectées et l'identification d'un rôle dans la réplication virale. La première est probablement la plus complexe, considérant les difficultés rencontrées pour exprimer la protéine en cellules eucaryotes. La construction d'un clone d'ADN proviral contenant le gène d'ASP optimisé en fusion avec un étiquette Myc ou avec la GFP pourrait dans un premier temps permettre de déterminer si un ORF situé dans la séquence complémentaire au gène de l'enveloppe peut effectivement être traduit efficacement. Des études approfondies de la régulation du transcrit seront également nécessaires, puisqu'elles donneraient probablement des indices importants pour déterminer à quel moment du cycle de réplication le transcrit antisens est exprimé de façon optimale. En effet, on peut penser que le niveau de transcription antisens pourrait être plus élevé dans une cellule où le virus est latent, puisque des études ont déjà démontré que l'activité du LTR 3' du virus est augmentée lorsque le LTR 5' est inactivé [314]. On peut également imaginer que l'expression du transcrit pourrait être supérieure dans un sous-type cellulaire donné. Par exemple, les résultats obtenus en RT-PCR, bien que qualitatifs, suggèrent que le transcrit pourrait être plus abondant chez les lignées de macrophages testées (U937-HIV-IIIb et OM10.1).

L'identification du rôle de la protéine nécessitera dans un premier temps une meilleure évaluation de sa localisation sub-cellulaire. Il sera intéressant de vérifier la localisation d'ASP chez des lymphocytes et des macrophages infectés ou non. Ensuite, l'utilisation de virus mutés incapables d'exprimer la protéine pourrait permettre d'évaluer un possible

impact sur la réplication du virus en culture cellulaire. D'autres méthodes visant à spécifiquement inhiber l'expression du transcrit antisens pourraient aussi être envisagées, comme par exemple la transfection de morpholinos oligonucléotides antisens [315] ou l'utilisation de ribozymes [316].

Conclusion

Alors que les rétrovirus humains ont été jusqu'à aujourd'hui le sujet de nombreuses études, plusieurs aspects de leur stratégie répllicative restent toujours incompris. Les résultats obtenus par notre équipe ainsi que par d'autres groupes de recherche sur la présence de transcription antisens chez plusieurs membres de la famille des rétrovirus laissent croire qu'il s'agit là d'une pièce importante du casse-tête que constitue leur cycle de vie. Il paraît en effet logique de croire que ces virus possédant un très court génome ont évolué de façon à maximiser leur potentiel codant. La surprenante capacité des protéines antisens des rétrovirus HTLVs à agir négativement sur la transcription virale révèle qu'une meilleure compréhension du cycle de réplication de ces virus et des pathologies qui leur sont associées nécessite une analyse approfondie de ce nouveau type de transcrit. De plus, les résultats que nous avons obtenus quant à la présence de transcription antisens au cours du cycle de réplication virale chez le VIH-1 suggère qu'il pourrait s'agir d'un aspect important de la biologie du virus et des analyses additionnelles sont requises pour évaluer la capacité de ce transcrit à produire une nouvelle protéine virale.

Il semble ne faire aucun doute que nos connaissances sur le phénomène de la transcription antisens chez les rétrovirus sont aujourd'hui très fragmentaires : les prochaines années pourraient s'avérer déterminantes dans la compréhension de cette tactique virale et éventuellement dans le développement de nouvelles approches thérapeutiques.

Bibliographie

1. Gorinsek, B., F. Gubensek, and D. Kordis, *Evolutionary genomics of chromoviruses in eukaryotes*. Mol Biol Evol, 2004. **21**(5): p. 781-98.
2. Hull, R. and S.N. Covey, *Genome organization and expression of reverse transcribing elements: variations and a theme*. J Gen Virol, 1986. **67** (Pt 9): p. 1751-8.
3. Murray, S.M. and M.L. Linial, *Foamy virus infection in primates*. J Med Primatol, 2006. **35**(4-5): p. 225-35.
4. Weiss, R.A., *The discovery of endogenous retroviruses*. Retrovirology, 2006. **3**: p. 67.
5. Boller, K., et al., *Human endogenous retrovirus HERV-K113 is capable of producing intact viral particles*. J Gen Virol, 2008. **89**(Pt 2): p. 567-72.
6. Muster, T., et al., *An endogenous retrovirus derived from human melanoma cells*. Cancer Res, 2003. **63**(24): p. 8735-41.
7. Sommerfelt, M.A., *Retrovirus receptors*. J Gen Virol, 1999. **80** (Pt 12): p. 3049-64.
8. Nisole, S. and A. Saib, *Early steps of retrovirus replicative cycle*. Retrovirology, 2004. **1**: p. 9.
9. Zennou, V., et al., *HIV-1 genome nuclear import is mediated by a central DNA flap*. Cell, 2000. **101**(2): p. 173-85.
10. Suzuki, Y. and R. Craigie, *The road to chromatin - nuclear entry of retroviruses*. Nat Rev Microbiol, 2007. **5**(3): p. 187-96.
11. Gallay, P., et al., *HIV-1 infection of nondividing cells: C-terminal tyrosine phosphorylation of the viral matrix protein is a key regulator*. Cell, 1995. **80**(3): p. 379-88.
12. Bouyac-Bertoia, M., et al., *HIV-1 infection requires a functional integrase NLS*. Mol Cell, 2001. **7**(5): p. 1025-35.
13. Haffar, O.K., et al., *Two nuclear localization signals in the HIV-1 matrix protein regulate nuclear import of the HIV-1 pre-integration complex*. J Mol Biol, 2000. **299**(2): p. 359-68.
14. Pryciak, P.M., A. Sil, and H.E. Varmus, *Retroviral integration into minichromosomes in vitro*. EMBO J, 1992. **11**(1): p. 291-303.
15. Meekings, K.N., et al., *HTLV-1 integration into transcriptionally active genomic regions is associated with proviral expression and with HAM/TSP*. PLoS Pathog, 2008. **4**(3): p. e1000027.
16. Craigie, R., *Hotspots and warm spots: integration specificity of retroelements*. Trends Genet, 1992. **8**(6): p. 187-90.
17. Brady, T., et al., *HIV integration site distributions in resting and activated CD4+ T cells infected in culture*. AIDS, 2009. **23**(12): p. 1461-71.
18. Schroder, A.R., et al., *HIV-1 integration in the human genome favors active genes and local hotspots*. Cell, 2002. **110**(4): p. 521-9.
19. Weidhaas, J.B., et al., *Relationship between retroviral DNA integration and gene expression*. J Virol, 2000. **74**(18): p. 8382-9.
20. Jonkers, J. and A. Berns, *Retroviral insertional mutagenesis as a strategy to identify cancer genes*. Biochim Biophys Acta, 1996. **1287**(1): p. 29-57.

21. Hacein-Bey-Abina, S., et al., *LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1*. Science, 2003. **302**(5644): p. 415-9.
22. Freed, E.O. and A.J. Mouland, *The cell biology of HIV-1 and other retroviruses*. Retrovirology, 2006. **3**: p. 77.
23. Cassol, E., et al., *Monocyte-derived macrophages and myeloid cell lines as targets of HIV-1 replication and persistence*. J Leukoc Biol, 2006. **80**(5): p. 1018-30.
24. Briggs, J.A., T. Wilk, and S.D. Fuller, *Do lipid rafts mediate virus assembly and pseudotyping?* J Gen Virol, 2003. **84**(Pt 4): p. 757-68.
25. www.virustaxonomyonline.com. [cited.
26. Poiesz, B.J., et al., *Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma*. Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7415-9.
27. Wolfe, N.D., et al., *Emergence of unique primate T-lymphotropic viruses among central African bushmeat hunters*. Proc Natl Acad Sci U S A, 2005. **102**(22): p. 7994-9.
28. Igakura, T., et al., *Spread of HTLV-I between lymphocytes by virus-induced polarization of the cytoskeleton*. Science, 2003. **299**(5613): p. 1713-6.
29. Jones, K.S., et al., *Cell-free HTLV-1 infects dendritic cells leading to transmission and transformation of CD4(+) T cells*. Nat Med, 2008. **14**(4): p. 429-36.
30. Jain, P., et al., *DC-SIGN mediates cell-free infection and transmission of HTLV-1 by dendritic cells*. J Virol, 2009.
31. Oger, J., *HTLV-1 infection and the viral etiology of multiple sclerosis*. J Neurol Sci, 2007. **262**(1-2): p. 100-4.
32. Murphy, E.L., et al., *Modelling the risk of adult T-cell leukemia/lymphoma in persons infected with human T-lymphotropic virus type I*. Int J Cancer, 1989. **43**(2): p. 250-3.
33. Yamaguchi, K., et al., *A proposal for smoldering adult T-cell leukemia--diversity in clinical pictures of adult T-cell leukemia*. Jpn J Clin Oncol, 1983. **13 Suppl 2**: p. 189-99.
34. Proietti, F.A., et al., *Global epidemiology of HTLV-I infection and associated diseases*. Oncogene, 2005. **24**(39): p. 6058-68.
35. Yamaguchi, K. and K. Takatsuki, *Adult T cell leukaemia-lymphoma*. Baillieres Clin Haematol, 1993. **6**(4): p. 899-915.
36. Feuer, G., et al., *Human T-cell leukemia virus infection of human hematopoietic progenitor cells: maintenance of virus infection during differentiation in vitro and in vivo*. J Virol, 1996. **70**(6): p. 4038-44.
37. Valensi, F., *Leucémie/lymphome T de l'adulte HTLV-1+*. Revue Française des Laboratoires, 2003. **360**.
38. Matutes, E., *Adult T-cell leukaemia/lymphoma*. J Clin Pathol, 2007. **60**(12): p. 1373-7.
39. Gill, P.S., et al., *Treatment of adult T-cell leukemia-lymphoma with a combination of interferon alfa and zidovudine*. N Engl J Med, 1995. **332**(26): p. 1744-8.
40. Orland, J.R., et al., *Prevalence and clinical features of HTLV neurologic disease in the HTLV Outcomes Study*. Neurology, 2003. **61**(11): p. 1588-94.
41. Gessain, A., et al., *Antibodies to human T-lymphotropic virus type-I in patients with tropical spastic paraparesis*. Lancet, 1985. **2**(8452): p. 407-10.

42. Quintas, S., et al., [*Tropical spastic paraparesis and HTLV-I associated myelopathy in infancy. A case report and review of the literature*]. *Rev Neurol*, 2004. **39**(12): p. 1133-6.
43. Yoshida, M., et al., *Viruses detected in HTLV-I-associated myelopathy and adult T-cell leukaemia are identical on DNA blotting*. *Lancet*, 1987. **1**(8541): p. 1085-6.
44. Usuku, K., et al., *Cytotoxic and suppressor activities in patients with HTLV-I-associated myelopathy*. *J Neuroimmunol*, 1991. **33**(3): p. 199-205.
45. Lee, S.M., et al., *Autoantibodies that recognize functional domains of hnRNPA1 implicate molecular mimicry in the pathogenesis of neurological disease*. *Neurosci Lett*, 2006. **401**(1-2): p. 188-93.
46. Ijichi, S., et al., *An autoaggressive process against bystander tissues in HTLV-I-infected individuals: a possible pathomechanism of HAM/TSP*. *Med Hypotheses*, 1993. **41**(6): p. 542-7.
47. Feuer, G. and P.L. Green, *Comparative biology of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2*. *Oncogene*, 2005. **24**(39): p. 5996-6004.
48. Gallo, R.C., *Human retroviruses after 20 years: a perspective from the past and prospects for their future control*. *Immunol Rev*, 2002. **185**: p. 236-65.
49. Ijichi, S., et al., *In vivo cellular tropism of human T cell leukemia virus type II (HTLV-II)*. *J Exp Med*, 1992. **176**(1): p. 293-6.
50. Lowis, G.W., W.A. Sheremata, and A. Minagar, *Epidemiologic features of HTLV-II: serologic and molecular evidence*. *Ann Epidemiol*, 2002. **12**(1): p. 46-66.
51. Toro, C., et al., *Molecular epidemiology of HTLV-2 infection among intravenous drug users in Spain*. *J Clin Virol*, 2005. **33**(1): p. 65-70.
52. Matsuoka, M. and K.T. Jeang, *Human T-cell leukaemia virus type 1 (HTLV-1) infectivity and cellular transformation*. *Nat Rev Cancer*, 2007. **7**(4): p. 270-80.
53. Brady, J., et al., *Identification of p40x-responsive regulatory sequences within the human T-cell leukemia virus type I long terminal repeat*. *J Virol*, 1987. **61**(7): p. 2175-81.
54. Fujisawa, J., et al., *A transcriptional enhancer sequence of HTLV-I is responsible for trans-activation mediated by p40 chi HTLV-I*. *EMBO J*, 1986. **5**(4): p. 713-8.
55. Jeang, K.T., et al., *Characterization of cellular factors that interact with the human T-cell leukemia virus type I p40x-responsive 21-base-pair sequence*. *J Virol*, 1988. **62**(12): p. 4499-509.
56. Kwok, R.P., et al., *Control of cAMP-regulated enhancers by the viral transactivator Tax through CREB and the co-activator CBP*. *Nature*, 1996. **380**(6575): p. 642-6.
57. Barnhart, M.K., L.M. Connor, and S.J. Marriott, *Function of the human T-cell leukemia virus type I 21-base-pair repeats in basal transcription*. *J Virol*, 1997. **71**(1): p. 337-44.
58. Bosselut, R., et al., *The product of the c-ets-1 proto-oncogene and the related Ets2 protein act as transcriptional activators of the long terminal repeat of human T cell leukemia virus HTLV-I*. *EMBO J*, 1990. **9**(10): p. 3137-44.
59. Bosselut, R., et al., *Myb protein binds to multiple sites in the human T cell lymphotropic virus type I long terminal repeat and transactivates LTR-mediated expression*. *Virology*, 1992. **186**(2): p. 764-9.

60. Yao, J. and B. Wigdahl, *Human T cell lymphotropic virus type I genomic expression and impact on intracellular signaling pathways during neurodegenerative disease and leukemia*. *Front Biosci*, 2000. **5**: p. D138-68.
61. Felber, B.K., et al., *The pX protein of HTLV-I is a transcriptional activator of its long terminal repeats*. *Science*, 1985. **229**(4714): p. 675-9.
62. Alefantis, T., et al., *Characterization of a nuclear export signal within the human T cell leukemia virus type I transactivator protein Tax*. *J Biol Chem*, 2003. **278**(24): p. 21814-22.
63. Zhao, L.J. and C.Z. Giam, *Human T-cell lymphotropic virus type I (HTLV-I) transcriptional activator, Tax, enhances CREB binding to HTLV-I 21-base-pair repeats by protein-protein interaction*. *Proc Natl Acad Sci U S A*, 1992. **89**(15): p. 7070-4.
64. Bantignies, F., et al., *Genetic characterization of transactivation of the human T-cell leukemia virus type I promoter: Binding of Tax to Tax-responsive element 1 is mediated by the cyclic AMP-responsive members of the CREB/ATF family of transcription factors*. *Mol Cell Biol*, 1996. **16**(5): p. 2174-82.
65. Reddy, T.R., et al., *Functional interaction of the HTLV-I transactivator Tax with activating transcription factor-4 (ATF4)*. *Oncogene*, 1997. **14**(23): p. 2785-92.
66. Ku, S.C., et al., *XBP-1, a novel human T-lymphotropic virus type 1 (HTLV-1) tax binding protein, activates HTLV-1 basal and tax-activated transcription*. *J Virol*, 2008. **82**(9): p. 4343-53.
67. Hai, T. and M.G. Hartman, *The molecular biology and nomenclature of the activating transcription factor/cAMP responsive element binding family of transcription factors: activating transcription factor proteins and homeostasis*. *Gene*, 2001. **273**(1): p. 1-11.
68. Perini, G., S. Wagner, and M.R. Green, *Recognition of bZIP proteins by the human T-cell leukaemia virus transactivator Tax*. *Nature*, 1995. **376**(6541): p. 602-5.
69. Wagner, S. and M.R. Green, *HTLV-I Tax protein stimulation of DNA binding of bZIP proteins by enhancing dimerization*. *Science*, 1993. **262**(5132): p. 395-9.
70. Yin, M.J. and R.B. Gaynor, *HTLV-1 21 bp repeat sequences facilitate stable association between Tax and CREB to increase CREB binding affinity*. *J Mol Biol*, 1996. **264**(1): p. 20-31.
71. Ng, P.W., et al., *Genome-wide expression changes induced by HTLV-1 Tax: evidence for MLK-3 mixed lineage kinase involvement in Tax-mediated NF-kappaB activation*. *Oncogene*, 2001. **20**(33): p. 4484-96.
72. Fujii, M., et al., *HTLV-I Tax induces expression of various immediate early serum responsive genes*. *Oncogene*, 1991. **6**(6): p. 1023-9.
73. Fujii, M., et al., *Activation of oncogenic transcription factor AP-1 in T cells infected with human T cell leukemia virus type I*. *AIDS Res Hum Retroviruses*, 2000. **16**(16): p. 1603-6.
74. Mori, N., et al., *Constitutive activation of transcription factor AP-1 in primary adult T-cell leukemia cells*. *Blood*, 2000. **95**(12): p. 3915-21.
75. Winter, H.Y. and S.J. Marriott, *Human T-cell leukemia virus type I Tax enhances serum response factor DNA binding and alters site selection*. *J Virol*, 2007. **81**(11): p. 6089-98.

76. Hayden, M.S. and S. Ghosh, *Signaling to NF-kappaB*. Genes Dev, 2004. **18**(18): p. 2195-224.
77. Jeong, S.J., et al., *Activated AKT regulates NF-kappaB activation, p53 inhibition and cell survival in HTLV-1-transformed cells*. Oncogene, 2005. **24**(44): p. 6719-28.
78. O'Mahony, A.M., et al., *Human T-cell lymphotropic virus type 1 tax induction of biologically Active NF-kappaB requires IkappaB kinase-1-mediated phosphorylation of RelA/p65*. J Biol Chem, 2004. **279**(18): p. 18137-45.
79. Peloponese, J.M., M.L. Yeung, and K.T. Jeang, *Modulation of nuclear factor-kappaB by human T cell leukemia virus type 1 Tax protein: implications for oncogenesis and inflammation*. Immunol Res, 2006. **34**(1): p. 1-12.
80. Akagi, T., et al., *Characterization of peripheral blood T-lymphocytes transduced with HTLV-I Tax mutants with different trans-activating phenotypes*. Oncogene, 1997. **14**(17): p. 2071-8.
81. Yamaoka, S., et al., *Constitutive activation of NF-kappa B is essential for transformation of rat fibroblasts by the human T-cell leukemia virus type I Tax protein*. EMBO J, 1996. **15**(4): p. 873-87.
82. Grossman, W.J. and L. Ratner, *Transgenic mouse models for HTLV-I infection*. J Acquir Immune Defic Syndr Hum Retrovirol, 1996. **13 Suppl 1**: p. S162-9.
83. Banerjee, P., Crawford, L , Sieburg, M , Green, P , Beilke, M , Harrington, WJ , Lairmore, MD , Feuer, G. *HTLV-1 Infection of Human Hematopoietic Stem Cells Induces a Novel T Cell Lymphoma in Humanized SCID Mice*. in American Society of Hematology. 2008. San Francisco.
84. Brooks, G., *Cyclins, cyclin-dependent kinases, and cyclin-dependent kinase inhibitors: detection methods and activity measurements*. Methods Mol Biol, 2005. **296**: p. 291-8.
85. Neuveut, C., et al., *Human T-cell leukemia virus type 1 Tax and cell cycle progression: role of cyclin D-cdk and p110Rb*. Mol Cell Biol, 1998. **18**(6): p. 3620-32.
86. Haller, K., et al., *Physical interaction of human T-cell leukemia virus type 1 Tax with cyclin-dependent kinase 4 stimulates the phosphorylation of retinoblastoma protein*. Mol Cell Biol, 2002. **22**(10): p. 3327-38.
87. Suzuki, T., et al., *HTLV-1 Tax protein interacts with cyclin-dependent kinase inhibitor p16INK4A and counteracts its inhibitory activity towards CDK4*. EMBO J, 1996. **15**(7): p. 1607-14.
88. Low, K.G., et al., *Human T-cell leukemia virus type 1 Tax releases cell cycle arrest induced by p16INK4a*. J Virol, 1997. **71**(3): p. 1956-62.
89. Kehn, K., et al., *The HTLV-I Tax oncoprotein targets the retinoblastoma protein for proteasomal degradation*. Oncogene, 2005. **24**(4): p. 525-40.
90. Zhou, B.B. and S.J. Elledge, *The DNA damage response: putting checkpoints in perspective*. Nature, 2000. **408**(6811): p. 433-9.
91. Majone, F. and K.T. Jeang, *Clastogenic effect of the human T-cell leukemia virus type I Tax oncoprotein correlates with unstabilized DNA breaks*. J Biol Chem, 2000. **275**(42): p. 32906-10.

92. Marriott, S.J., F.J. Lemoine, and K.T. Jeang, *Damaged DNA and miscounted chromosomes: human T cell leukemia virus type I tax oncoprotein and genetic lesions in transformed cells*. J Biomed Sci, 2002. **9**(4): p. 292-8.
93. Suzuki, T., et al., *HTLV-1 tax oncoprotein binds to DNA topoisomerase I and inhibits its catalytic activity*. Virology, 2000. **270**(2): p. 291-8.
94. Wu, K., et al., *Protein profile of tax-associated complexes*. J Biol Chem, 2004. **279**(1): p. 495-508.
95. Park, H.U., et al., *Human T-cell leukemia virus type I Tax interacts with Chk1 and attenuates DNA-damage induced G2 arrest mediated by Chk1*. Oncogene, 2004. **23**(29): p. 4966-74.
96. Haoudi, A., et al., *Human T-cell leukemia virus-I tax oncoprotein functionally targets a subnuclear complex involved in cellular DNA damage-response*. J Biol Chem, 2003. **278**(39): p. 37736-44.
97. Park, H.U., et al., *Human T-cell leukemia virus type I Tax attenuates gamma-irradiation-induced apoptosis through physical interaction with Chk2*. Oncogene, 2006. **25**(3): p. 438-47.
98. Li, J. and D.F. Stern, *Regulation of CHK2 by DNA-dependent protein kinase*. J Biol Chem, 2005. **280**(12): p. 12041-50.
99. Marriott, S.J. and O.J. Semmes, *Impact of HTLV-I Tax on cell cycle progression and the cellular DNA damage repair response*. Oncogene, 2005. **24**(39): p. 5986-95.
100. Grone, M., C. Koch, and R. Grassmann, *The HTLV-1 Rex protein induces nuclear accumulation of unspliced viral RNA by avoiding intron excision and degradation*. Virology, 1996. **218**(2): p. 316-25.
101. Toyoshima, H., et al., *Secondary structure of the human T-cell leukemia virus type I rex-responsive element is essential for rex regulation of RNA processing and transport of unspliced RNAs*. J Virol, 1990. **64**(6): p. 2825-32.
102. Ahmed, Y.F., et al., *The HTLV-I Rex response element mediates a novel form of mRNA polyadenylation*. Cell, 1991. **64**(4): p. 727-37.
103. Derse, D., J. Mikovits, and F. Ruscetti, *X-I and X-II open reading frames of HTLV-I are not required for virus replication or for immortalization of primary T-cells in vitro*. Virology, 1997. **237**(1): p. 123-8.
104. Roithmann, S., et al., *The open reading frame I (ORF I)/ORF II part of the human T-cell leukemia virus type I X region is dispensable for p40tax, p27rex, or envelope expression*. J Virol, 1994. **68**(5): p. 3448-51.
105. Nicot, C., et al., *HTLV-1-encoded p30II is a post-transcriptional negative regulator of viral replication*. Nat Med, 2004. **10**(2): p. 197-201.
106. Fukumoto, R., et al., *Inhibition of T-cell receptor signal transduction and viral expression by the linker for activation of T cells-interacting p12(I) protein of human T-cell leukemia/lymphoma virus type 1*. J Virol, 2007. **81**(17): p. 9088-99.
107. Fukumoto, R., et al., *In vivo genetic mutations define predominant functions of the human T-cell leukemia/lymphoma virus p12I protein*. Blood, 2008.
108. Kim, S.J., et al., *A conserved calcineurin-binding motif in human T lymphotropic virus type 1 p12I functions to modulate nuclear factor of activated T cell activation*. J Biol Chem, 2003. **278**(18): p. 15550-7.

109. Ding, W., et al., *Human T-cell lymphotropic virus type 1 p12(I) expression increases cytoplasmic calcium to enhance the activation of nuclear factor of activated T cells*. J Virol, 2002. **76**(20): p. 10374-82.
110. Zhang, W., et al., *Human T-lymphotropic virus type 1 p30(II) regulates gene transcription by binding CREB binding protein/p300*. J Virol, 2001. **75**(20): p. 9885-95.
111. Ding, W., et al., *Endoplasmic reticulum and cis-Golgi localization of human T-lymphotropic virus type 1 p12(I): association with calreticulin and calnexin*. J Virol, 2001. **75**(16): p. 7672-82.
112. Lairmore, M.D., et al., *In vitro and in vivo functional analysis of human T cell lymphotropic virus type 1 pX open reading frames I and II*. AIDS Res Hum Retroviruses, 2000. **16**(16): p. 1757-64.
113. Collins, N.D., et al., *Selective ablation of human T-cell lymphotropic virus type 1 p12I reduces viral infectivity in vivo*. Blood, 1998. **91**(12): p. 4701-7.
114. Collins, N.D., et al., *Proliferation response to interleukin-2 and Jak/Stat activation of T cells immortalized by human T-cell lymphotropic virus type 1 is independent of open reading frame I expression*. J Virol, 1999. **73**(11): p. 9642-9.
115. Bartoe, J.T., et al., *Functional role of pX open reading frame II of human T-lymphotropic virus type 1 in maintenance of viral loads in vivo*. J Virol, 2000. **74**(3): p. 1094-100.
116. Albrecht, B., et al., *Human T-lymphotropic virus type 1 open reading frame I p12(I) is required for efficient viral infectivity in primary lymphocytes*. J Virol, 2000. **74**(21): p. 9828-35.
117. Albrecht, B., et al., *Activation of nuclear factor of activated T cells by human T-lymphotropic virus type 1 accessory protein p12(I)*. J Virol, 2002. **76**(7): p. 3493-501.
118. Albrecht, B. and M.D. Lairmore, *Critical role of human T-lymphotropic virus type 1 accessory proteins in viral replication and pathogenesis*. Microbiol Mol Biol Rev, 2002. **66**(3): p. 396-406, table of contents.
119. Cereseto, A., et al., *Differential expression of alternatively spliced pX mRNAs in HTLV-I-infected cell lines*. Leukemia, 1997. **11**(6): p. 866-70.
120. Berneman, Z.N., et al., *Expression of alternatively spliced human T-lymphotropic virus type 1 pX mRNA in infected cell lines and in primary uncultured cells from patients with adult T-cell leukemia/lymphoma and healthy carriers*. Proc Natl Acad Sci U S A, 1992. **89**(7): p. 3005-9.
121. Dekaban, G.A., et al., *The HTLV-I orfI protein is recognized by serum antibodies from naturally infected humans and experimentally infected rabbits*. Virology, 2000. **274**(1): p. 86-93.
122. Chen, Y.M., et al., *Antibody reactivities to tumor-suppressor protein p53 and HTLV-I Tof, Rex and Tax in HTLV-I-infected people with differing clinical status*. Int J Cancer, 1997. **71**(2): p. 196-202.
123. Pique, C., et al., *Evidence for the chronic in vivo production of human T cell leukemia virus type I Rof and Tof proteins from cytotoxic T lymphocytes directed against viral peptides*. J Exp Med, 2000. **191**(3): p. 567-72.

124. Bindhu, M., A. Nair, and M.D. Lairmore, *Role of accessory proteins of HTLV-1 in viral replication, T cell activation, and cellular gene expression*. Front Biosci, 2004. **9**: p. 2556-76.
125. Barre-Sinoussi, F., et al., *Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)*. Science, 1983. **220**(4599): p. 868-71.
126. Popovic, M., et al., *Detection, isolation, and continuous production of cytopathic retroviruses (HTLV-III) from patients with AIDS and pre-AIDS*. Science, 1984. **224**(4648): p. 497-500.
127. Gallo, R.C., et al., *Frequent detection and isolation of cytopathic retroviruses (HTLV-III) from patients with AIDS and at risk for AIDS*. Science, 1984. **224**(4648): p. 500-3.
128. Schupbach, J., et al., *Serological analysis of a subgroup of human T-lymphotropic retroviruses (HTLV-III) associated with AIDS*. Science, 1984. **224**(4648): p. 503-5.
129. Daar, E.S., et al., *Transient high levels of viremia in patients with primary human immunodeficiency virus type 1 infection*. N Engl J Med, 1991. **324**(14): p. 961-4.
130. Clark, S.J., et al., *High titers of cytopathic virus in plasma of patients with symptomatic primary HIV-1 infection*. N Engl J Med, 1991. **324**(14): p. 954-60.
131. Tindall, B. and D.A. Cooper, *Primary HIV infection: host responses and intervention strategies*. AIDS, 1991. **5**(1): p. 1-14.
132. Levy, J.A., *Pathogenesis of human immunodeficiency virus infection*. Microbiol Rev, 1993. **57**(1): p. 183-289.
133. Pantaleo, G., et al., *HIV infection is active and progressive in lymphoid tissue during the clinically latent stage of disease*. Nature, 1993. **362**(6418): p. 355-8.
134. Dalglish, A.G., et al., *The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus*. Nature, 1984. **312**(5996): p. 763-7.
135. Klatzmann, D., et al., *T-lymphocyte T4 molecule behaves as the receptor for human retrovirus LAV*. Nature, 1984. **312**(5996): p. 767-8.
136. Choe, H., et al., *The beta-chemokine receptors CCR3 and CCR5 facilitate infection by primary HIV-1 isolates*. Cell, 1996. **85**(7): p. 1135-48.
137. Feng, Y., et al., *HIV-1 entry cofactor: functional cDNA cloning of a seven-transmembrane, G protein-coupled receptor*. Science, 1996. **272**(5263): p. 872-7.
138. Wei, X., et al., *Viral dynamics in human immunodeficiency virus type 1 infection*. Nature, 1995. **373**(6510): p. 117-22.
139. Finzi, D., et al., *Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy*. Science, 1997. **278**(5341): p. 1295-300.
140. Chun, T.W., et al., *Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection*. Nature, 1997. **387**(6629): p. 183-8.
141. Dean, M., et al., *Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study*. Science, 1996. **273**(5283): p. 1856-62.
142. Swingler, S., et al., *HIV-1 Nef intersects the macrophage CD40L signalling pathway to promote resting-cell infection*. Nature, 2003. **424**(6945): p. 213-9.
143. Raposo, G., et al., *Human macrophages accumulate HIV-1 particles in MHC II compartments*. Traffic, 2002. **3**(10): p. 718-29.

144. Orenstein, J.M., et al., *Cytoplasmic assembly and accumulation of human immunodeficiency virus types 1 and 2 in recombinant human colony-stimulating factor-1-treated human monocytes: an ultrastructural study*. J Virol, 1988. **62**(8): p. 2578-86.
145. Carter, C.A. and L.S. Ehrlich, *Cell biology of HIV-1 infection of macrophages*. Annu Rev Microbiol, 2008. **62**: p. 425-43.
146. Pope, M., et al., *Conjugates of dendritic cells and memory T lymphocytes from skin facilitate productive infection with HIV-1*. Cell, 1994. **78**(3): p. 389-98.
147. Cameron, P.U., et al., *Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells*. Science, 1992. **257**(5068): p. 383-7.
148. Kwon, D.S., et al., *DC-SIGN-mediated internalization of HIV is required for trans-enhancement of T cell infection*. Immunity, 2002. **16**(1): p. 135-44.
149. Rambaut, A., et al., *The causes and consequences of HIV evolution*. Nat Rev Genet, 2004. **5**(1): p. 52-61.
150. VandeWoude, S. and C. Apetrei, *Going wild: lessons from naturally occurring T-lymphotropic lentiviruses*. Clin Microbiol Rev, 2006. **19**(4): p. 728-62.
151. Keele, B.F., et al., *Chimpanzee reservoirs of pandemic and nonpandemic HIV-1*. Science, 2006. **313**(5786): p. 523-6.
152. Santiago, M.L., et al., *Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercopithecus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2*. J Virol, 2005. **79**(19): p. 12515-27.
153. Schim van der Loeff, M.F. and P. Aaby, *Towards a better understanding of the epidemiology of HIV-2*. AIDS, 1999. **13 Suppl A**: p. S69-84.
154. *UNAIDS WHO epidemic update*. 2007.
155. Yamaguchi, J., et al., *Evaluation of HIV type 1 group O isolates: identification of five phylogenetic clusters*. AIDS Res Hum Retroviruses, 2002. **18**(4): p. 269-82.
156. Yamaguchi, J., et al., *HIV infections in northwestern Cameroon: identification of HIV type 1 group O and dual HIV type 1 group M and group O infections*. AIDS Res Hum Retroviruses, 2004. **20**(9): p. 944-57.
157. Lemieux, A.M., et al., *T-cell activation leads to poor activation of the HIV-1 clade E long terminal repeat and weak association of nuclear factor-kappaB and NFAT with its enhancer region*. J Biol Chem, 2004. **279**(51): p. 52949-60.
158. Montano, M.A., et al., *Divergent transcriptional regulation among expanding human immunodeficiency virus type 1 subtypes*. J Virol, 1997. **71**(11): p. 8657-65.
159. Roof, P., et al., *Differential regulation of HIV-1 clade-specific B, C, and E long terminal repeats by NF-kappaB and the Tat transactivator*. Virology, 2002. **296**(1): p. 77-83.
160. Desfosses, Y., et al., *Regulation of human immunodeficiency virus type 1 gene expression by clade-specific Tat proteins*. J Virol, 2005. **79**(14): p. 9180-91.
161. Ashe, M.P., L.H. Pearson, and N.J. Proudfoot, *The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site*. EMBO J, 1997. **16**(18): p. 5752-63.
162. Schwartz, S., et al., *Cloning and functional analysis of multiply spliced mRNA species of human immunodeficiency virus type 1*. J Virol, 1990. **64**(6): p. 2519-29.

163. Kim, S.Y., et al., *Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: evidence for differential gene expression*. J Virol, 1989. **63**(9): p. 3708-13.
164. Pomerantz, R.J., T. Seshamma, and D. Trono, *Efficient replication of human immunodeficiency virus type 1 requires a threshold level of Rev: potential implications for latency*. J Virol, 1992. **66**(3): p. 1809-13.
165. Pereira, L.A., et al., *A compilation of cellular transcription factor interactions with the HIV-1 LTR promoter*. Nucleic Acids Res, 2000. **28**(3): p. 663-8.
166. Lu, Y., et al., *Effects of long terminal repeat mutations on human immunodeficiency virus type 1 replication*. J Virol, 1989. **63**(9): p. 4115-9.
167. Giffin, M.J., et al., *Structure of NFAT1 bound as a dimer to the HIV-1 LTR kappa B element*. Nat Struct Biol, 2003. **10**(10): p. 800-6.
168. Bates, D.L., et al., *Crystal structure of NFAT bound to the HIV-1 LTR tandem kappaB enhancer element*. Structure, 2008. **16**(5): p. 684-94.
169. Cron, R.Q., et al., *NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells*. Clin Immunol, 2000. **94**(3): p. 179-91.
170. Robichaud, G.A., et al., *Nuclear factor of activated T cells is a driving force for preferential productive HIV-1 infection of CD45RO-expressing CD4+ T cells*. J Biol Chem, 2002. **277**(26): p. 23733-41.
171. Sune, C. and M.A. Garcia-Blanco, *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, 1995. **69**(10): p. 6572-6.
172. Bentley, K., et al., *Mutational analysis of the HIV-1 LTR as a promoter of negative sense transcription*. Arch Virol, 2004. **149**(12): p. 2277-94.
173. Ratnasabapathy, R., et al., *The HIV-1 long terminal repeat contains an unusual element that induces the synthesis of short RNAs from various mRNA and snRNA promoters*. Genes Dev, 1990. **4**(12A): p. 2061-74.
174. Kao, S.Y., et al., *Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product*. Nature, 1987. **330**(6147): p. 489-93.
175. Kiernan, R.E., et al., *HIV-1 tat transcriptional activity is regulated by acetylation*. EMBO J, 1999. **18**(21): p. 6106-18.
176. Kaehlcke, K., et al., *Acetylation of Tat defines a cyclinT1-independent step in HIV transactivation*. Mol Cell, 2003. **12**(1): p. 167-76.
177. Deng, L., et al., *Acetylation of HIV-1 Tat by CBP/P300 increases transcription of integrated HIV-1 genome and enhances binding to core histones*. Virology, 2000. **277**(2): p. 278-95.
178. Col, E., et al., *The histone acetyltransferase, hGCN5, interacts with and acetylates the HIV transactivator, Tat*. J Biol Chem, 2001. **276**(30): p. 28179-84.
179. Bres, V., et al., *Tat acetyl-acceptor lysines are important for human immunodeficiency virus type-1 replication*. J Biol Chem, 2002. **277**(25): p. 22215-21.
180. Bres, V., et al., *Differential acetylation of Tat coordinates its interaction with the co-activators cyclin T1 and PCAF*. EMBO J, 2002. **21**(24): p. 6811-9.
181. Boulanger, M.C., et al., *Methylation of Tat by PRMT6 regulates human immunodeficiency virus type 1 gene expression*. J Virol, 2005. **79**(1): p. 124-31.

182. Bres, V., et al., *A non-proteolytic role for ubiquitin in Tat-mediated transactivation of the HIV-1 promoter*. Nat Cell Biol, 2003. **5**(8): p. 754-61.
183. Malim, M.H. and M. Emerman, *HIV-1 accessory proteins--ensuring viral survival in a hostile environment*. Cell Host Microbe, 2008. **3**(6): p. 388-98.
184. Gerard, M., et al., *Purification and interaction properties of the human RNA polymerase B(II) general transcription factor BTF2*. J Biol Chem, 1991. **266**(31): p. 20940-5.
185. Conaway, J.W., et al., *Transcription initiated by RNA polymerase II and transcription factors from liver. Structure and action of transcription factors epsilon and tau*. J Biol Chem, 1991. **266**(12): p. 7804-11.
186. Kadonaga, J.T., *Regulation of RNA polymerase II transcription by sequence-specific DNA binding factors*. Cell, 2004. **116**(2): p. 247-57.
187. Lee, T.I. and R.A. Young, *Transcription of eukaryotic protein-coding genes*. Annu Rev Genet, 2000. **34**: p. 77-137.
188. Kadonaga, J.T., *Transcription in eucaryotes*, in Cell. 2004.
189. Burke, T.W., et al., *The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box*. Cold Spring Harb Symp Quant Biol, 1998. **63**: p. 75-82.
190. Lim, C.Y., et al., *The MTE, a new core promoter element for transcription by RNA polymerase II*. Genes Dev, 2004. **18**(13): p. 1606-17.
191. Lagrange, T., et al., *New core promoter element in RNA polymerase II-dependent transcription: sequence-specific DNA binding by transcription factor IIB*. Genes Dev, 1998. **12**(1): p. 34-44.
192. Koleske, A.J. and R.A. Young, *An RNA polymerase II holoenzyme responsive to activators*. Nature, 1994. **368**(6470): p. 466-9.
193. Robinson, K.M. and J.T. Kadonaga, *The use of chromatin templates to recreate transcriptional regulatory phenomena in vitro*. Biochim Biophys Acta, 1998. **1378**(1): p. M1-6.
194. Conaway, J.W., et al., *Control of elongation by RNA polymerase II*. Trends Biochem Sci, 2000. **25**(8): p. 375-80.
195. Orphanides, G. and D. Reinberg, *RNA polymerase II elongation through chromatin*. Nature, 2000. **407**(6803): p. 471-5.
196. Smith, C.W. and J. Valcarcel, *Alternative pre-mRNA splicing: the logic of combinatorial control*. Trends Biochem Sci, 2000. **25**(8): p. 381-8.
197. Alberts, B., *Molecular Biology of the Cell*. 2002: Garland Science.
198. Breeuwer, M. and D.S. Goldfarb, *Facilitated nuclear transport of histone H1 and other small nucleophilic proteins*. Cell, 1990. **60**(6): p. 999-1008.
199. Goldberg, M.W., et al., *Ran alters nuclear pore complex conformation*. J Mol Biol, 2000. **300**(3): p. 519-29.
200. Cullen, B.R., *Retroviruses as model systems for the study of nuclear RNA export pathways*. Virology, 1998. **249**(2): p. 203-10.
201. Neville, M., et al., *The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export*. Curr Biol, 1997. **7**(10): p. 767-75.
202. Fornerod, M., et al., *CRMI is an export receptor for leucine-rich nuclear export signals*. Cell, 1997. **90**(6): p. 1051-60.

203. Braun, I.C., et al., *Nuclear export of mRNA by TAP/NXF1 requires two nucleoporin-binding sites but not p15*. *Mol Cell Biol*, 2002. **22**(15): p. 5405-18.
204. Kang, Y. and B.R. Cullen, *The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences*. *Genes Dev*, 1999. **13**(9): p. 1126-39.
205. Gruter, P., et al., *TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus*. *Mol Cell*, 1998. **1**(5): p. 649-59.
206. Cullen, B.R., *Nuclear RNA export*. *J Cell Sci*, 2003. **116**(Pt 4): p. 587-97.
207. Barrell, B.G., G.M. Air, and C.A. Hutchison, 3rd, *Overlapping genes in bacteriophage phiX174*. *Nature*, 1976. **264**(5581): p. 34-41.
208. Tomizawa, J., et al., *Inhibition of ColE1 RNA primer formation by a plasmid-specified small RNA*. *Proc Natl Acad Sci U S A*, 1981. **78**(3): p. 1421-5.
209. Wagner, E.G. and R.W. Simons, *Antisense RNA control in bacteria, phages, and plasmids*. *Annu Rev Microbiol*, 1994. **48**: p. 713-42.
210. Anderson, S., et al., *Sequence and organization of the human mitochondrial genome*. *Nature*, 1981. **290**(5806): p. 457-65.
211. Bibb, M.J., et al., *Sequence and gene organization of mouse mitochondrial DNA*. *Cell*, 1981. **26**(2 Pt 2): p. 167-80.
212. Lapidot, M. and Y. Pilpel, *Genome-wide natural antisense transcription: coupling its regulation to its different regulatory mechanisms*. *EMBO Rep*, 2006. **7**(12): p. 1216-22.
213. Chen, J., et al., *Over 20% of human transcripts might form sense-antisense pairs*. *Nucleic Acids Res*, 2004. **32**(16): p. 4812-20.
214. Yelin, R., et al., *Widespread occurrence of antisense transcription in the human genome*. *Nat Biotechnol*, 2003. **21**(4): p. 379-86.
215. Sun, M., et al., *Evidence for variation in abundance of antisense transcripts between multicellular animals but no relationship between antisense transcription and organismic complexity*. *Genome Res*, 2006. **16**(7): p. 922-33.
216. Kiyosawa, H., et al., *Antisense transcripts with FANTOM2 clone set and their implications for gene regulation*. *Genome Res*, 2003. **13**(6B): p. 1324-34.
217. Carninci, P., et al., *The transcriptional landscape of the mammalian genome*. *Science*, 2005. **309**(5740): p. 1559-63.
218. Katayama, S., et al., *Antisense transcription in the mammalian transcriptome*. *Science*, 2005. **309**(5740): p. 1564-6.
219. Misra, S., et al., *Annotation of the Drosophila melanogaster euchromatic genome: a systematic review*. *Genome Biol*, 2002. **3**(12): p. RESEARCH0083.
220. Osato, N., et al., *Antisense transcripts with rice full-length cDNAs*. *Genome Biol*, 2003. **5**(1): p. R5.
221. Jen, C.H., et al., *Natural antisense transcripts with coding capacity in Arabidopsis may have a regulatory role that is not linked to double-stranded RNA degradation*. *Genome Biol*, 2005. **6**(6): p. R51.
222. David, L., et al., *A high-resolution map of transcription in the yeast genome*. *Proc Natl Acad Sci U S A*, 2006. **103**(14): p. 5320-5.
223. Galante, P.A., et al., *Sense-antisense pairs in mammals: functional and evolutionary considerations*. *Genome Biol*, 2007. **8**(3): p. R40.

224. Johnson, J.M., et al., *Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments*. Trends Genet, 2005. **21**(2): p. 93-102.
225. Perocchi, F., et al., *Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D*. Nucleic Acids Res, 2007. **35**(19): p. e128.
226. Beiter, T., et al., *Sense or antisense? False priming reverse transcription controls are required for determining sequence orientation by reverse transcription-PCR*. Anal Biochem, 2007. **369**(2): p. 258-61.
227. Gu, C., et al., *Plus- and minus-stranded foot-and-mouth disease virus RNA quantified simultaneously using a novel real-time RT-PCR*. Virus Genes, 2007. **34**(3): p. 289-98.
228. Horsington, J. and Z. Zhang, *Analysis of foot-and-mouth disease virus replication using strand-specific quantitative RT-PCR*. J Virol Methods, 2007. **144**(1-2): p. 149-55.
229. Lin, L., J. Fevery, and S. Hiem Yap, *A novel strand-specific RT-PCR for detection of hepatitis C virus negative-strand RNA (replicative intermediate): evidence of absence or very low level of HCV replication in peripheral blood mononuclear cells*. J Virol Methods, 2002. **100**(1-2): p. 97-105.
230. Cummings, M., C. McGurk, and J.R. Masters, *Rapid identification of antisense mRNA-expressing clones using strand-specific RT-PCR*. Antisense Nucleic Acid Drug Dev, 2003. **13**(2): p. 115-7.
231. Gunji, T., et al., *Specific detection of positive and negative stranded hepatitis C viral RNA using chemical RNA modification*. Arch Virol, 1994. **134**(3-4): p. 293-302.
232. Li, A.W., C.K. Too, and P.R. Murphy, *The basic fibroblast growth factor (FGF-2) antisense RNA (GFG) is translated into a MutT-related protein in vivo*. Biochem Biophys Res Commun, 1996. **223**(1): p. 19-23.
233. Doerig, C., L.I. Pizer, and C.L. Wilcox, *Detection of the latency-associated transcript in neuronal cultures during the latent infection with herpes simplex virus type 1*. Virology, 1991. **183**(1): p. 423-6.
234. Lagunoff, M. and B. Roizman, *Expression of a herpes simplex virus 1 open reading frame antisense to the gamma(1)34.5 gene and transcribed by an RNA 3' coterminal with the unspliced latency-associated transcript*. J Virol, 1994. **68**(9): p. 6021-8.
235. Prescott, E.M. and N.J. Proudfoot, *Transcriptional collision between convergent genes in budding yeast*. Proc Natl Acad Sci U S A, 2002. **99**(13): p. 8796-801.
236. Hastings, M.L., et al., *Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels*. Nucleic Acids Res, 1997. **25**(21): p. 4296-300.
237. Chen, J., et al., *Genome-wide analysis of coordinate expression and evolution of human cis-encoded sense-antisense transcripts*. Trends Genet, 2005. **21**(6): p. 326-9.
238. Bass, B.L., *RNA editing by adenosine deaminases that act on RNA*. Annu Rev Biochem, 2002. **71**: p. 817-46.
239. Scadden, A.D. and C.W. Smith, *Specific cleavage of hyper-edited dsRNAs*. EMBO J, 2001. **20**(15): p. 4243-52.

240. Zhang, Z. and G.G. Carmichael, *The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs*. Cell, 2001. **106**(4): p. 465-75.
241. Tufarelli, C., et al., *Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease*. Nat Genet, 2003. **34**(2): p. 157-65.
242. Ogawa, Y. and J.T. Lee, *Antisense regulation in X inactivation and autosomal imprinting*. Cytogenet Genome Res, 2002. **99**(1-4): p. 59-65.
243. Miller, R.H., *Human immunodeficiency virus may encode a novel protein on the genomic DNA plus strand*. Science, 1988. **239**(4846): p. 1420-2.
244. Larocca, D., et al., *Human T-cell leukemia virus minus strand transcription in infected T-cells*. Biochem Biophys Res Commun, 1989. **163**(2): p. 1006-13.
245. Gaudray, G., et al., *The complementary strand of the human T-cell leukemia virus type 1 RNA genome encodes a bZIP transcription factor that down-regulates viral transcription*. J Virol, 2002. **76**(24): p. 12813-22.
246. Cavanagh, M.H., et al., *HTLV-I antisense transcripts initiating in the 3'LTR are alternatively spliced and polyadenylated*. Retrovirology, 2006. **3**: p. 15.
247. Satou, Y., et al., *HTLV-I basic leucine zipper factor gene mRNA supports proliferation of adult T cell leukemia cells*. Proc Natl Acad Sci U S A, 2006. **103**(3): p. 720-5.
248. Murata, K., et al., *A novel alternative splicing isoform of human T-cell leukemia virus type 1 bZIP factor (HBZ-SI) targets distinct subnuclear localization*. J Virol, 2006. **80**(5): p. 2495-505.
249. Corcelette, S., T. Masse, and J.J. Madjar, *Initiation of translation by non-AUG codons in human T-cell lymphotropic virus type I mRNA encoding both Rex and Tax regulatory proteins*. Nucleic Acids Res, 2000. **28**(7): p. 1625-34.
250. Usui, T., et al., *Characteristic expression of HTLV-1 basic zipper factor (HBZ) transcripts in HTLV-1 provirus-positive cells*. Retrovirology, 2008. **5**: p. 34.
251. Basbous, J., et al., *The HBZ factor of human T-cell leukemia virus type I dimerizes with transcription factors JunB and c-Jun and modulates their transcriptional activity*. J Biol Chem, 2003. **278**(44): p. 43620-7.
252. Thebault, S., et al., *HBZ interacts with JunD and stimulates its transcriptional activity*. FEBS Lett, 2004. **562**(1-3): p. 165-70.
253. Lemasson, I., et al., *Human T-cell leukemia virus type 1 (HTLV-1) bZIP protein interacts with the cellular transcription factor CREB to inhibit HTLV-1 transcription*. J Virol, 2007. **81**(4): p. 1543-53.
254. Clerc, I., et al., *An interaction between the human T cell leukemia virus type 1 basic leucine zipper factor (HBZ) and the KIX domain of p300/CBP contributes to the down-regulation of tax-dependent viral transcription by HBZ*. J Biol Chem, 2008. **283**(35): p. 23903-13.
255. Arnold, J., et al., *Enhancement of infectivity and persistence in vivo by HBZ, a natural antisense coded protein of HTLV-1*. Blood, 2006. **107**(10): p. 3976-82.
256. Arnold, J., et al., *Human T-cell leukemia virus type-1 antisense-encoded gene, Hbz, promotes T-lymphocyte proliferation*. Blood, 2008. **112**(9): p. 3788-97.
257. Hivin, P., et al., *Nuclear localization of HTLV-I bZIP factor (HBZ) is mediated by three distinct motifs*. J Cell Sci, 2005. **118**(Pt 7): p. 1355-62.

258. Hivin, P., et al., *The HBZ-SP1 isoform of human T-cell leukemia virus type I represses JunB activity by sequestration into nuclear bodies*. *Retrovirology*, 2007. **4**: p. 14.
259. Switzer, W.M., et al., *Ancient origin and molecular features of the novel human T-lymphotropic virus type 3 revealed by complete genome analysis*. *J Virol*, 2006. **80**(15): p. 7427-38.
260. Chevalier, S.A., et al., *Construction and characterization of a human T-cell lymphotropic virus type 3 infectious molecular clone*. *J Virol*, 2008. **82**(13): p. 6747-52.
261. Calattini, S., et al., *Human T-cell lymphotropic virus type 3: complete nucleotide sequence and characterization of the human tax3 protein*. *J Virol*, 2006. **80**(19): p. 9876-88.
262. Chevalier, S.A., et al., *Construction and characterization of a full-length infectious simian T-cell lymphotropic virus type 3 molecular clone*. *J Virol*, 2007. **81**(12): p. 6276-85.
263. Bukrinsky, M.I. and A.F. Etkin, *Plus strand of the HIV provirus DNA is expressed at early stages of infection*. *AIDS Res Hum Retroviruses*, 1990. **6**(4): p. 425-6.
264. Michael, N.L., et al., *Negative-strand RNA transcripts are produced in human immunodeficiency virus type 1-infected cells and patients by a novel promoter downregulated by Tat*. *J Virol*, 1994. **68**(2): p. 979-87.
265. Peeters, A., P.F. Lambert, and N.J. Deacon, *A fourth Sp1 site in the human immunodeficiency virus type 1 long terminal repeat is essential for negative-sense transcription*. *J Virol*, 1996. **70**(10): p. 6665-72.
266. Vanhee-Brossollet, C., et al., *A natural antisense RNA derived from the HIV-1 env gene encodes a protein which is recognized by circulating antibodies of HIV+ individuals*. *Virology*, 1995. **206**(1): p. 196-202.
267. Briquet, S. and C. Vaquero, *Immunolocalization studies of an antisense protein in HIV-1-infected cells and viral particles*. *Virology*, 2002. **292**(2): p. 177-84.
268. Tagieva, N.E. and C. Vaquero, *Expression of naturally occurring antisense RNA inhibits human immunodeficiency virus type 1 heterologous strain replication*. *J Gen Virol*, 1997. **78 (Pt 10)**: p. 2503-11.
269. Ludwig, L.B., et al., *Human Immunodeficiency Virus-Type 1 LTR DNA contains an intrinsic gene producing antisense RNA and protein products*. *Retrovirology*, 2006. **3**: p. 80.
270. Briquet, S., et al., *Natural antisense transcripts are detected in different cell lines and tissues of cats infected with feline immunodeficiency virus*. *Gene*, 2001. **267**(2): p. 157-64.
271. Landry, S., et al., *Detection, characterization and regulation of antisense transcripts in HIV-1*. *Retrovirology*, 2007. **4**: p. 71.
272. Boisvert, F.M., et al., *Arginine methylation of MRE11 by PRMT1 is required for DNA damage checkpoint control*. *Genes Dev*, 2005. **19**(6): p. 671-6.
273. Masson, J.Y., et al., *Identification and purification of two distinct complexes containing the five RAD51 paralogs*. *Genes Dev*, 2001. **15**(24): p. 3296-307.
274. Holler, T.P., et al., *HIV1 integrase expressed in Escherichia coli from a synthetic gene*. *Gene*, 1993. **136**(1-2): p. 323-8.

275. Ngumbela, K.C., et al., *Quantitative effect of suboptimal codon usage on translational efficiency of mRNA encoding HIV-1 gag in intact T cells*. PLoS ONE, 2008. **3**(6): p. e2356.
276. Nguyen, K.L., et al., *Codon optimization of the HIV-1 vpu and vif genes stabilizes their mRNA and allows for highly efficient Rev-independent expression*. Virology, 2004. **319**(2): p. 163-75.
277. Lanford, R.E., et al., *Lack of detection of negative-strand hepatitis C virus RNA in peripheral blood mononuclear cells and other extrahepatic tissues by the highly strand-specific rTth reverse transcriptase PCR*. J Virol, 1995. **69**(12): p. 8079-83.
278. Huang, X., Y. Li, and C.Y. Zheng, *A novel single-cell quantitative real-time RT-PCR method for quantifying foot-and-mouth disease viral RNA*. J Virol Methods, 2009. **155**(2): p. 150-6.
279. Takyar, S.T., et al., *Specific detection of minus-strand hepatitis C virus RNA by reverse-transcription polymerase chain reaction on PolyA(+)-purified RNA*. Hepatology, 2000. **32**(2): p. 382-7.
280. Komurian-Pradel, F., et al., *Strand specific quantitative real-time PCR to study replication of hepatitis C virus genome*. J Virol Methods, 2004. **116**(1): p. 103-6.
281. Landry, S., et al., *Upregulation of HTLV-1 antisense transcription by the viral Tax protein*. J Virol, 2008.
282. Yoshida, M., et al., *Transcriptional control of spliced and unspliced human T-cell leukemia virus type 1 bZIP factor (HBZ) gene*. J Virol, 2008. **82**(19): p. 9359-68.
283. Zauberman, A., et al., *A functional p53-responsive intronic promoter is contained within the human mdm2 gene*. Nucleic Acids Res, 1995. **23**(14): p. 2584-92.
284. Tiffoche, C., et al., *Novel intronic promoter in the rat ER alpha gene responsible for the transient transcription of a variant receptor*. Endocrinology, 2001. **142**(9): p. 4106-19.
285. Vicini, E. and M. Conti, *Characterization of an intronic promoter of a cyclic adenosine 3',5'-monophosphate (cAMP)-specific phosphodiesterase gene that confers hormone and cAMP inducibility*. Mol Endocrinol, 1997. **11**(7): p. 839-50.
286. Georges, S.A., et al., *Tax recruitment of CBP/p300, via the KIX domain, reveals a potent requirement for acetyltransferase activity that is chromatin dependent and histone tail independent*. Mol Cell Biol, 2003. **23**(10): p. 3392-404.
287. Lemasson, I., et al., *Tax-dependent displacement of nucleosomes during transcriptional activation of human T-cell leukemia virus type 1*. J Biol Chem, 2006. **281**(19): p. 13075-82.
288. Takeda, S., et al., *Genetic and epigenetic inactivation of tax gene in adult T-cell leukemia cells*. Int J Cancer, 2004. **109**(4): p. 559-67.
289. Koiwa, T., et al., *5'-long terminal repeat-selective CpG methylation of latent human T-cell leukemia virus type 1 provirus in vitro and in vivo*. J Virol, 2002. **76**(18): p. 9389-97.
290. Furukawa, Y., et al., *Existence of escape mutant in HTLV-I tax during the development of adult T-cell leukemia*. Blood, 2001. **97**(4): p. 987-93.
291. Tamiya, S., et al., *Two types of defective human T-lymphotropic virus type I provirus in adult T-cell leukemia*. Blood, 1996. **88**(8): p. 3065-73.
292. Yasunaga, J. and M. Matsuoka, *Leukaemogenic mechanism of human T-cell leukaemia virus type I*. Rev Med Virol, 2007. **17**(5): p. 301-11.

293. Miyazaki, M., et al., *Preferential selection of human T-cell leukemia virus type 1 provirus lacking the 5' long terminal repeat during oncogenesis*. J Virol, 2007. **81**(11): p. 5714-23.
294. Zhao, T., et al., *Human T-cell leukemia virus type 1 bZIP factor selectively suppresses the classical pathway of NF- κ B*. Blood, 2008.
295. Rasmussen, R., et al., *X-ray scattering indicates that the leucine zipper is a coiled coil*. Proc Natl Acad Sci U S A, 1991. **88**(2): p. 561-4.
296. O'Shea, E.K., et al., *X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil*. Science, 1991. **254**(5031): p. 539-44.
297. Grigoryan, G. and A.E. Keating, *Structural specificity in coiled-coil interactions*. Curr Opin Struct Biol, 2008. **18**(4): p. 477-83.
298. Vinson, C., et al., *Classification of human B-ZIP proteins based on dimerization properties*. Mol Cell Biol, 2002. **22**(18): p. 6321-35.
299. Fong, J.H., A.E. Keating, and M. Singh, *Predicting specificity in bZIP coiled-coil protein interactions*. Genome Biol, 2004. **5**(2): p. R11.
300. McDonnell, A.V., et al., *Paircoil2: improved prediction of coiled coils from sequence*. Bioinformatics, 2006. **22**(3): p. 356-8.
301. Chun, R.F., et al., *Modulation of Sp1 phosphorylation by human immunodeficiency virus type 1 Tat*. J Virol, 1998. **72**(4): p. 2615-29.
302. Kyte, J. and R.F. Doolittle, *A simple method for displaying the hydropathic character of a protein*. J Mol Biol, 1982. **157**(1): p. 105-32.
303. Kay, B.K., M.P. Williamson, and M. Sudol, *The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains*. FASEB J, 2000. **14**(2): p. 231-41.
304. Sudol, M., *From Src Homology domains to other signaling modules: proposal of the 'protein recognition code'*. Oncogene, 1998. **17**(11 Reviews): p. 1469-74.
305. Dalgarno, D.C., M.C. Botfield, and R.J. Rickles, *SH3 domains and drug design: ligands, structure, and biological function*. Biopolymers, 1997. **43**(5): p. 383-400.
306. Saksela, K., G. Cheng, and D. Baltimore, *Proline-rich (PxxP) motifs in HIV-1 Nef bind to SH3 domains of a subset of Src kinases and are required for the enhanced growth of Nef+ viruses but not for down-regulation of CD4*. EMBO J, 1995. **14**(3): p. 484-91.
307. Lee, C.H., et al., *Crystal structure of the conserved core of HIV-1 Nef complexed with a Src family SH3 domain*. Cell, 1996. **85**(6): p. 931-42.
308. Yamshchikov, G.V., et al., *Assembly of SIV virus-like particles containing envelope proteins using a baculovirus expression system*. Virology, 1995. **214**(1): p. 50-8.
309. Cherepanov, P., et al., *High-level expression of active HIV-1 integrase from a synthetic gene in human cells*. FASEB J, 2000. **14**(10): p. 1389-99.
310. Cochrane, A.W., et al., *Identification and characterization of intragenic sequences which repress human immunodeficiency virus structural gene expression*. J Virol, 1991. **65**(10): p. 5305-13.
311. Schneider, R., et al., *Inactivation of the human immunodeficiency virus type 1 inhibitory elements allows Rev-independent expression of Gag and Gag/protease and particle formation*. J Virol, 1997. **71**(7): p. 4892-903.

312. Heaphy, S., et al., *HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the Rev response element region*. Cell, 1990. **60**(4): p. 685-93.
313. Kjems, J., et al., *Structural analysis of the interaction between the human immunodeficiency virus Rev protein and the Rev response element*. Proc Natl Acad Sci U S A, 1991. **88**(3): p. 683-7.
314. Klaver, B. and B. Berkhout, *Comparison of 5' and 3' long terminal repeat promoter function in human immunodeficiency virus*. J Virol, 1994. **68**(6): p. 3830-40.
315. Moulton, J.D. and Y.L. Yan, *Using Morpholinos to control gene expression*. Curr Protoc Mol Biol, 2008. **Chapter 26**: p. Unit 26 8.
316. Gu, S., et al., *Inhibition of infectious human immunodeficiency virus type 1 virions via lentiviral vector encoded short antisense RNAs*. Oligonucleotides, 2006. **16**(4): p. 287-95.