

Direction des bibliothèques

AVIS

Ce document a été numérisé par la Division de la gestion des documents et des archives de l'Université de Montréal.

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

This document was digitized by the Records Management & Archives Division of Université de Montréal.

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

**Sélection de peptides altérant le changement de cadre de
lecture -1 programmé du VIH-1**

par

Gabriel Théberge-Julien

Département de Biochimie

Faculté de Médecine

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de maîtrise
en biochimie

Mars 2008

© Gabriel Théberge-Julien, 2008



Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Sélection de peptides interférant avec un changement de cadre de lecture ribosomique -1
programmé chez le virus de l'immunodéficience humaine de type 1.

présenté par :

Gabriel Théberge-Julien

a été évalué par un jury composé des personnes suivantes :

Pascal Chartrand, président-rapporteur
Léa Brakier-Gingras, directeur de recherche
Nikolaus Heveker, co-directeur
Luc DesGroseillers, membre du jury

Résumé

Le virus de l'immunodéficience humaine de type 1 (VIH-1) utilise un changement de cadre de lecture (*frameshift*) ribosomique -1 programmé pour produire le précurseur de ses enzymes. Ce *frameshift* se produit dans une région spécifique de l'ARN messager (ARNm) pleine longueur du virus, au niveau d'une séquence glissante, un heptanucléotide. Cette séquence est suivie d'une structure secondaire particulière, le signal stimulateur du *frameshift*, une tige-boucle irrégulière qui contrôle l'efficacité de *frameshift* du VIH-1. Des variations de l'efficacité du *frameshift* diminuent la réplication du virus. Notre but était de sélectionner à partir d'une librairie de peptides des candidats qui perturbent l'efficacité de *frameshift* du VIH-1. Nous avons utilisé un système d'études bactérien, sachant que le *frameshift* du VIH-1 peut être reproduit dans des bactéries et que le ribosome bactérien répond comme le ribosome eucaryote à des modifications du signal stimulateur du VIH-1. Nous avons utilisé un vecteur codant pour deux rapporteurs : la protéine fluorescente verte (GFP) et la protéine fluorescente rouge (RFP). La séquence de *frameshift* du VIH-1 était insérée au début de la séquence codante de la GFP qui était donc produite par les ribosomes effectuant un *frameshift* -1 alors que la RFP était produite par traduction conventionnelle. Une librairie de peptides de 14 acides aminés de long, composés d'arginine, de sérine, d'histidine et d'asparagine, fut introduite dans le site actif de la thiorédoxine de *E. coli*. Des bactéries *E. coli* Top10 ont été transformées avec le vecteur exprimant les deux rapporteurs et le plasmide exprimant la librairie de peptides. Elles furent ensuite triées par cytométrie de flux pour sélectionner les clones exprimant un peptide qui perturbe l'efficacité de *frameshift*, ce qui se manifeste par une variation dans le rapport GFP/RFP. Nous avons identifié des peptides diminuant l'efficacité de *frameshift* de plus de 50% sans affecter la traduction conventionnelle. Ces peptides se sont aussi révélés actifs avec un système rapporteur où le signal stimulateur du VIH-1 était remplacé par d'autres signaux stimulateurs de *frameshift*. Ce résultat suggère que les peptides se lient non au signal stimulateur de *frameshift*, mais au ribosome. Dix des peptides actifs en système bactérien ont été examinés dans un système eucaryote avec un système rapporteur à double luciférase mesurant l'efficacité de *frameshift* du VIH-1 et deux d'entre eux ont conservé leur capacité

à diminuer le *frameshift*. Les peptides actifs pourront servir de base pour le développement d'agents anti-rétroviraux ciblant le *frameshift* ribosomique -1 programmé. De plus, la détermination du site de liaison des peptides actifs ainsi que de leur mode d'action permettrait de mieux caractériser le mécanisme du *frameshift* programmé du VIH-1.

Mots-clés : *frameshift* ribosomique -1 programmé, librairie de peptides, VIH-1, interaction ARN-protéine

Abstract

The human immunodeficiency virus type 1 (HIV-1) uses a programmed -1 ribosomal frameshift to produce the precursor of its enzymes. This frameshift occurs at a specific region of the full length viral messenger RNA (mRNA) on a slippery sequence, a heptanucleotide. This sequence is followed by a peculiar secondary structure, the frameshift stimulatory signal, an irregular stem-loop which controls the frameshift efficiency of HIV-1. A change in the frameshift efficiency perturbs the replication of the virus. Our goal was to select candidates that reduced the frameshift efficiency of HIV-1 from a combinatorial peptide library. We used a bacterial system for our study, knowing that HIV-1 frameshift can be recapitulated in bacteria and that the bacterial ribosome responds like the eukaryotic ribosome to modifications in the frameshift stimulatory signal of HIV-1. We used a reporter system consisting of the green fluorescent protein (GFP) and of the red fluorescent protein (RFP). The HIV-1 frameshift region was inserted at the beginning of the GFP coding sequence, so that GFP is produced by ribosomes that perform a programmed -1 frameshift whereas RFP is produced by conventional translation. A 14-amino acid long peptide library, made of arginine, serine, histidine and asparagine was inserted in the active center of *E. coli* thioredoxin. *E. coli* Top10 bacteria were transformed with the vector that expresses the reporter system and the plasmid that expresses the peptide library. These bacteria were then screened by flow cytometry, in order to select clones expressing a peptide that perturbs the frameshift efficiency, which corresponds to changes in the GFP/RFP ratio. We identified peptides that reduce frameshift efficiency by over 50% without affecting conventional translation. These peptides were also active with reporter systems where the HIV-1 stimulatory signal was replaced by another frameshift stimulatory signal. That result suggests that the peptides did not bind to the frameshift stimulatory signal, but to the ribosome. Ten peptides that were active in the bacterial system were assayed in a eukaryotic system with a dual-luciferase reporter system which assessed the frameshift efficiency of HIV-1. Two of them conserved their capacity to decrease the frameshift. These active peptides constitute a basis for the development of anti-retroviral compounds that target the programmed -1 ribosomal frameshift. The characterization of the

the binding site and the mode of action of active peptides will allow a better characterization of the mechanism accounting for the frameshift.

Keywords : -1 programmed ribosomal frameshift, peptide library, HIV-1, RNA-protein interaction

Remerciements

De nombreuses personnes ont rendu la réalisation de cette maîtrise possible et agréable. Tout d'abord, d'immenses remerciements à ma directrice, Dr Léa Brakier-Gingras, pour son soutien, tant pédagogique que moral, pour son encadrement exemplaire, même dans la maladie, et pour sa patience. Je vous remercie sincèrement de m'avoir donné cette merveilleuse opportunité d'études et de m'avoir guidé avec justesse et compréhension à travers mon parcours. Je remercie également mon co-directeur, le Dr Nikolaus Heveker, pour son apport critique précieux ainsi que pour les nombreux cours d'histoire qui auront su nous instruire et nous divertir.

J'aimerais aussi démontrer toute ma reconnaissance aux deux personnes qui auront été mes mentors lors de mon séjour au laboratoire. Premièrement, François, qui m'a suivi lorsque j'y faisais mon entrée en tant que stagiaire. Merci pour ton enseignement et pour ces heures de plaisir à citer François Pérusse. Deuxièmement, Dominic qui, pendant près de trois ans, aura été un soutien technique hors pair, un bien patient professeur pour un élève imparfait et un ami qui gagne à être connu. Ce fut une réelle chance d'avoir comme superviseurs des gens aussi expérimentés et amicaux que vous.

Je ne peux omettre de mentionner notre trio de filles au laboratoire, Mélissa, Karine et Johanie, avec qui ce fut un réel plaisir d'échanger, autant sur le plan scientifique qu'humain. Je n'oublierai pas de sitôt nos fameuses conversations du midi et tout ce temps passé ensemble lors de congrès. Je remercie aussi Martin que j'aurai brièvement côtoyé au laboratoire du temps où j'étais stagiaire, mais qui n'est jamais parti bien loin. L'ambiance amicale au laboratoire qui se reflétait dans nos nombreuses réunions du vendredi soir aura su faire passer bien rapidement ces trois dernières années.

Finalement, je tiens à remercier mes parents pour leurs encouragements tout au long de mon parcours scolaire ainsi que pour leur support financier. Je remercie aussi mes deux âmes sœurs, ma tendre moitié Nancy, pour son soutien inconditionnel, et mon grand-père Gabriel Théberge (et oui!), pour m'avoir donné le goût des sciences et la soif de connaissances. Merci à ceux qui liront ce mémoire, en espérant que vous l'apprécierez.

En souvenir de toi, grand-papa.

Table des matières

Liste des tableaux	xi
Liste des figures	xii
Liste des abréviations	xiii
Chapitre I : Introduction.....	1
1.1 La problématique.....	2
1.2 Le virus de l'immunodéficience humaine de type 1 (VIH-1).....	4
1.2.1 Son génome	4
1.2.1.1 Les éléments d'ARN et d'ADN.....	6
1.2.1.2 Les protéines virales.....	9
1.2.2 La réplication du VIH-1	13
1.2.2.1 Les étapes précoces du cycle réplcatif viral.....	13
1.2.2.2 Les étapes tardives du cycle réplcatif.....	15
1.3 Les agents anti-rétroviraux connus à ce jour.....	17
1.3.1 Les agents ciblant la protéase virale.....	17
1.3.2 Les agents ciblant la transcriptase inverse.....	18
1.3.3 Les agents ciblant la fusion des virions aux cellules.....	19
1.3.4 Les agents en développement contre de nouvelles cibles	19
1.4 La traduction conventionnelle et le <i>frameshift</i>	22
1.4.1 Les <i>frameshifts</i> programmés chez les eucaryotes et les procaryotes	26
1.4.2 Le <i>frameshift</i> ribosomique -1 programmé chez le VIH-1	28
1.4.2.1 Les éléments contrôlant l'efficacité de <i>frameshift</i> chez le VIH-1.....	29
1.4.2.2 Le mécanisme du <i>frameshift</i> -1 chez le VIH-1	34
1.4.2.3 L'importance de l'efficacité de <i>frameshift</i> chez le VIH-1.....	38
1.4.2.4 Le <i>frameshift</i> ribosomique -1 programmé du VIH-1 chez les bactéries	39
1.5 Les interactions ARN-protéine.....	40
1.5.1 La tendance à être présent à l'interface des acides aminés	40
1.5.2 La structure de peptides interagissant avec l'ARN.....	42

1.5.3 L'utilisation de bibliothèques combinatoires de peptides actifs <i>in vivo</i>	45
1.5.3.1 Sélection de peptides conférant une résistance au Taxol	46
1.5.3.2 Sélection de peptides liant la structure RRE du VIH-1	47
1.6 Le ribosome est-il une cible potentielle?	50
1.7 Le but de ce projet	51
Chapitre II : <i>Selection of peptides interfering with a ribosomal frameshift in the human immunodeficiency virus type 1</i>	53
Chapitre III : Discussion	95
3.1 Rappel de la problématique.....	96
3.2 La cytométrie de flux comme outil pour le criblage de peptides affectant le frameshift.	99
3.3 Étude de l'activité des peptides sélectionnés sur l'efficacité du <i>frameshift</i> et sur la traduction conventionnelle.....	101
3.4 Analyse des séquences des peptides les plus actifs.....	102
3.5 Étude de la sélectivité des peptides sélectionnés pour le VIH-1 groupe M sous-type B.	104
3.6 Étude de l'effet des peptides sélectionnés chez les bactéries dans un système eucaryote.	107
Conclusion et perspectives.....	109
Bibliographie.....	I

Liste des tableaux

Tableau 1. Les principales interactions moléculaires dans les interactions ARN-protéine. .43	
Tableau 2. L'activité de 20 peptides actifs sur des régions de frameshift avec différents signaux stimulateurs de frameshift. 105	

Liste des figures

Figure 1. La prévalence du VIH dans le monde.	3
Figure 2. Distribution des sous-types prédominants du groupe M du VIH-1 dans le monde.	5
Figure 3. Le génome du VIH de type 1.	7
Figure 4. Le cycle de réplication du VIH-1.	14
Figure 5. L'entrée d'un virion dans une cellule.	16
Figure 6. La traduction conventionnelle.	23
Figure 7. Les modèles du signal stimulateur de frameshift du sous-type B du VIH-1.	32
Figure 8. Le signal stimulateur de frameshift du sous-type B du VIH-1.	33
Figure 9. Les modèles mécanistiques du frameshift -1.	35
Figure 10. La tendance des acides aminés à être présents à l'interface d'une interaction ARN-protéine.	41
Figure 11. Structures adoptées par les protéines interagissant avec l'ARN.	44
Figure 12. Les spectres de fluorescence de la RFP et de la GFP.	98
Figure 13. La cytométrie de flux.	100

Liste des abréviations

30S : petite sous-unité ribosomique bactérienne

40S : petite sous-unité ribosomique eucaryote

43S : complexe de préinitiation eucaryote comprenant la 40S associée à différents facteurs d'initiation et à l'ARN de transfert initiateur chargé de méthionine

50S : grande sous-unité ribosomique bactérienne

60S : grande sous-unité ribosomique eucaryote

Ψ : signal d'empaquetage de l'ARNg

λ N : protéine d'antiterminaison du phage λ

A : adénine

A (site) : aminoacyl

Aa : acide aminé

ADN : acide désoxyribonucléique

ADNg : ADN génomique

ARN : acide ribonucléique

ARNg : ARN génomique

ARNm : ARN messenger

ARNr : ARN ribosomique

ARNt : ARN de transfert

C : cytidine

CA : capsid

CCR5 : récepteur de chimiokine 5

CD4 : groupe de différenciation 4

Cryo-EM : microscopie cryoélectronique

CTD : domaine carboxy-terminal

CTL : lymphocyte T cytotoxique

Cyp A : cyclophiline A

CXCR4 : récepteur de chimiokine 4

DEN : domaine d'export nucléaire

DIS : site d'initiation de la dimérisation

DLN : domaine de localisation nucléaire

E (site) : sortie

eIF2 : facteur d'initiation eucaryote 2

Env : enveloppe

FACS : *fluorescence activated cell sorting*

Frameshift : changement de cadre de lecture

G : guanine ou glycine (selon le contexte)

Gag : précurseur de protéines de structure virales

GFP : protéine fluorescente verte

GTP : guanine triphosphate

H : histidine

IN : intégrase

IRE : élément de réponse au fer

IRES : site d'entrée interne du ribosome

IST : inducteur de courts transcrits

KanR : gène de résistance à la kanamycine

LacZ : gène de la β -galactosidase

LTR : longues répétitions terminales

MA : matrice

N : asparagine

NC : nucléocapside

Nef : protéine virale à nombreuses fonctions effectives

NRE : élément régulateur négatif

P (site) : peptidyle

PBS : site de liaison de l'amorce de la transcriptase inverse

PKR : protéine kinase répresseur

Pol : précurseur d'enzymes virales

PR : protéase

R : arginine

RE : réticulum endoplasmique

Rev : protéine virale régulatrice de l'expression des virions

RFP : protéine fluorescente rouge

RISC : *RNA induced silencing complex*

RRE : élément de réponse à Rev

RT : transcriptase inverse

S : sérine

SIDA : syndrome d'immunodéficience acquise

SU : protéine de surface

T : thymidine

T (site) : entrée

TAK : kinase associée à Tat

TAR : élément de réponse d'activation par Tat

Tat : protéine régulatrice de transactivation

TM : protéine trans-membranaire

U : uracile

VIB : virus de l'immunodéficience bovine

Vif : facteur d'infectivité des virions

VIH : virus de l'immunodéficience humaine

Vpr : protéine virale R

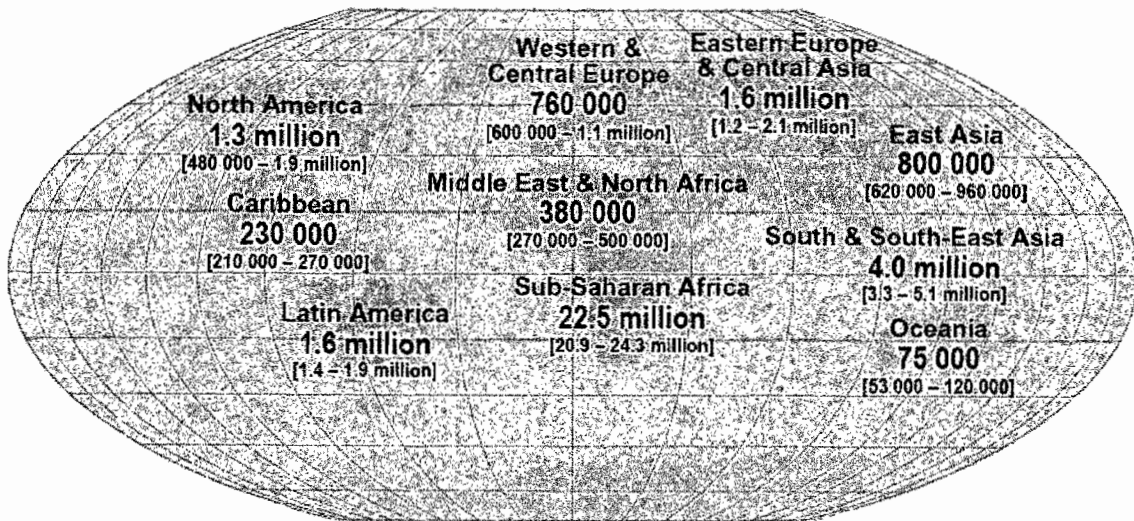
Vpu : protéine virale U

Chapitre I : Introduction

1.1 La problématique

Le virus de l'immunodéficience humaine de type 1 (VIH-1) est l'agent causatif du syndrome de l'immunodéficience acquise (SIDA). Ce virus est à l'origine d'une des plus graves pandémies du dernier siècle, ayant causé plus de 25 millions de décès depuis 1981. En 2007, environ 33,2 millions de personnes étaient infectées par le VIH-1 (figure 1) et 2,1 millions de personnes infectées ont perdu la vie (détails sur le site www.unaids.org). Pour faire face à ce fléau, de nombreux agents anti-rétroviraux ont été développés et sont utilisés en clinique. À ce jour, la trithérapie est le principal traitement administré aux patients atteints par le VIH-1. Elle se base sur l'utilisation de trois inhibiteurs de deux enzymes virales : deux contre la transcriptase inverse et un contre la protéase (Mathe, 1997). Même si le développement récent de comprimés « trois-en-un » a simplifié la consommation de médicaments par les patients et réduit partiellement leur coût (Laurent et al., 2004), le traitement comporte encore plusieurs inconvénients. Outre les coûts importants qui y sont toujours reliés et des effets secondaires sérieux, le principal problème avec les agents anti-rétroviraux couramment utilisés provient de l'émergence de variants viraux résistants à ces agents (Bartlett et al., 2001). Ceci est attribuable au haut degré de mutations du VIH-1, qui est causé par la faible fidélité de la transcriptase inverse jumelée aux pressions sélectives des médicaments utilisés. Il est donc urgent que de nouveaux agents anti-rétroviraux agissant contre de nouvelles cibles du cycle répliatif du VIH-1 soient développés. À court terme, une avenue prometteuse est l'inhibition de la fusion du virus aux cellules et un inhibiteur de la fusion est maintenant utilisé pour traiter les patients infectés (Temesgen et al., 2006). Une autre avenue est l'inhibition d'une autre enzyme virale, l'intégrase. Des molécules inhibitrices de cette enzyme sont présentement en essais cliniques. Notre équipe s'intéresse à une autre cible potentielle : le changement de cadre de lecture (*frameshift*) -1 programmé du VIH-1.

Figure 1. La prévalence du VIH dans le monde. Nombre de personnes infectées par le virus de l'immunodéficience humaine de type-1 selon les régions du monde. (UNAIDS, 2006)

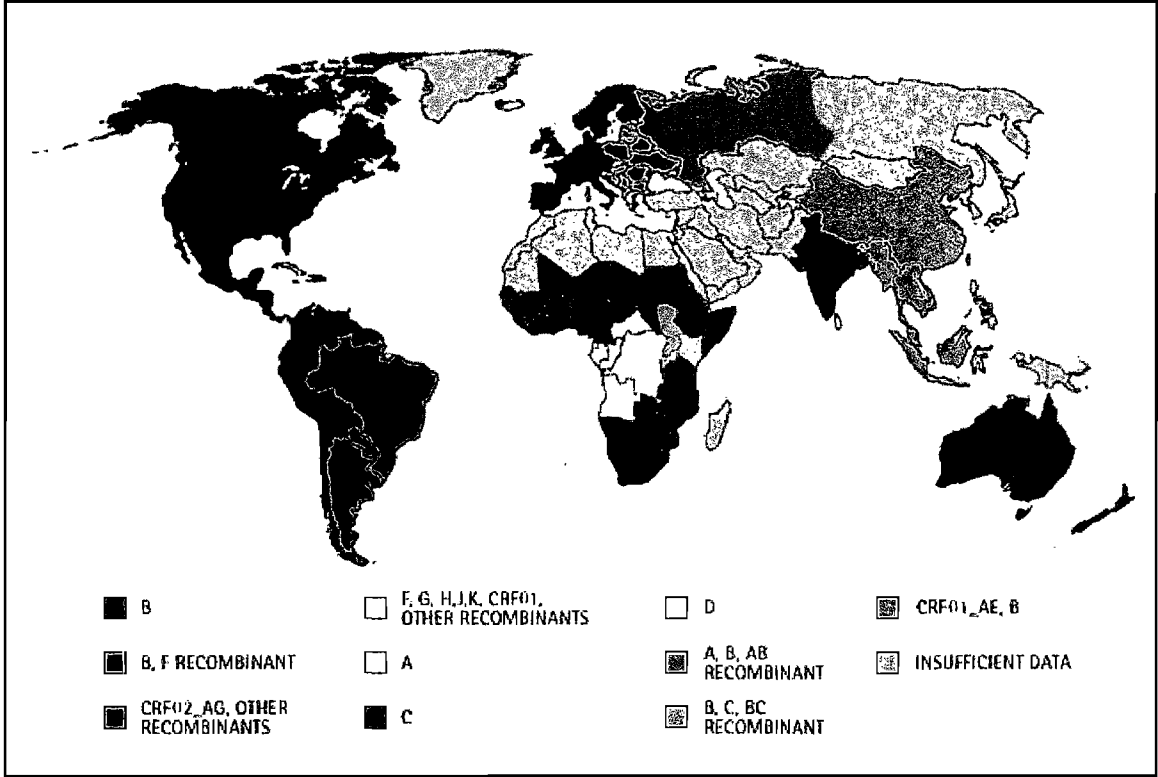


1.2 Le virus de l'immunodéficience humaine de type 1 (VIH-1).

1.2.1 Son génome

Le VIH est un virus à génome d'ARN d'une taille de 9.8 kb. Il existe deux types de VIH : le VIH de type 1, responsable de la pandémie, et le VIH de type 2 (VIH-2), qui est beaucoup moins infectieux que le type 1 (Gilbert et al., 2003) et qui ne sera pas examiné dans cet ouvrage. Le VIH-1 peut quant à lui être divisé en trois groupes : le groupe M pour *main*, le groupe O pour *outlier* et le groupe N pour *new* ou *non-M non O*. Le groupe M est responsable de plus de 99% des infections à travers le monde et est lui-même divisé en plusieurs sous-types nommés A, B, C, D, F, G, H, J et K (figure 2). Notons qu'il existe un grand nombre de variétés de virus recombinants entre les divers sous-types (Burke, 1997). Ces classifications entre groupes et sous-types sont effectuées selon les variations dans les séquences nucléotidiques des précurseurs protéiques de l'enveloppe (Env), des protéines de structure (Gag) et des enzymes virales (Pol) (Robertson et al., 2000). Les sous-types A, B et C du groupe M sont les plus répandus mondialement. Le sous-type C est responsable de près de 50% des infections répertoriées, se retrouvant surtout dans les pays d'Asie et d'Afrique les plus lourdement touchés par la pandémie tels que l'Afrique du Sud et l'Inde. Le sous-type A quant à lui se retrouve surtout dans les pays du centre et de l'est de l'Afrique, ainsi qu'en Europe de l'est, dans les anciens pays soviétiques. Pour ce qui est du sous-type B, il sévit dans la plupart des pays occidentaux, en Amérique du Nord, en Europe centrale et de l'ouest ainsi qu'en Australie (Buonaguro et al., 2007). Notre étude porte sur le sous-type B du groupe M du VIH-1.

Figure 2. Distribution des sous-types prédominants du groupe M du VIH-1 dans le monde. (PBS media, 2003)

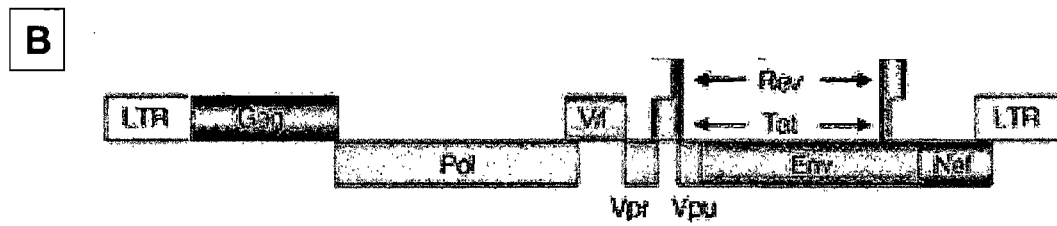
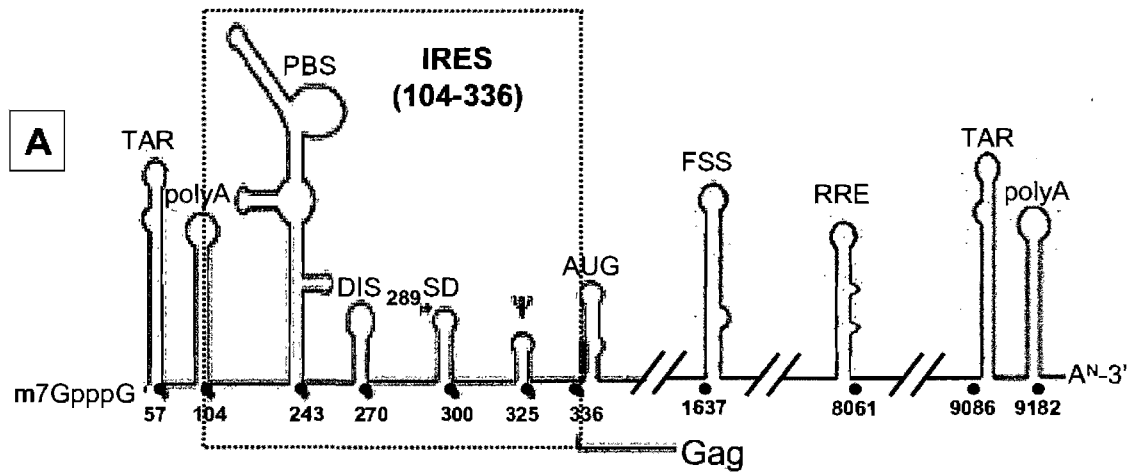


1.2.1.1 Les éléments d'ARN et d'ADN

L'ARN génomique (ARNg) du VIH-1 possède neuf cadres de lecture codant au total pour quinze protéines (les informations contenues dans la section 1.2.1.1 et 1.2.1.2 se réfèrent majoritairement aux revues suivantes : Cullen, 1998; Frankel & Young, 1998; Tang et al., 1999; Freed & Mouland, 2006). L'ARNg est incorporé dans les virions sous forme de dimères où deux copies d'ARNg sont unies par des liens non-covalents. Il existe plusieurs régions régulatrices dans l'ARNg et les ARN messagers (ARNm) du VIH-1, qui jouent des rôles importants dans la réplication du virus. L'ARNg possède à ses extrémités des régions appelées LTR (*long terminal repeats*). Une fois recopié en ADN, le génome viral est introduit dans le génome de la cellule hôte. Plusieurs domaines fonctionnels du 5'-LTR, qui s'étend du début du provirus au codon initiateur de Gag, accomplissent différents rôles dans la transcription des ARNm viraux. Une portion appelée NRE (*negative regulatory element*) diminue la transcription des ARNm viraux en interagissant avec une panoplie de protéines cellulaires (telles que GRE, AP-1 et NF-AT), alors qu'une région activatrice (*enhancer*) ainsi que la région promotrice comprenant une boîte TATA activent la transcription des ARNm viraux. De plus, une région nommée IST (*inducer of short transcripts*) induit la production de transcrits terminés prématurément.

Plusieurs autres régions du génome du VIH-1 comprises dans les LTR jouent des rôles importants dans la réplication du virus lorsqu'elles sont transcrites (figure 3A). Une cellule infectée génère, à partir du provirus, différents ARNm multi-épissés, mono-épissés et non-épissés. Une région en 3' du 5'-LTR, située immédiatement en aval du site d'initiation de la transcription, forme la tige-boucle TAR (*tat-activation response element*) lorsqu'elle est transcrite en ARN. TAR interagit avec la protéine virale Tat (*transactivating regulatory protein*) et promeut la transcription des ARNm naissants via une augmentation de la processivité de l'ARN polymérase II. Cette augmentation est induite par une kinase appelée CDK2 (*cycline dependant kinase*) recrutée par Tat (Nekhai et al., 2002). CDK2 hyperphosphoryle le domaine CTD (*carboxyl terminal domain*) de l'ARN polymérase II, ce

Figure 3. Le génome du VIH de type 1. A) La région 5' non-traduite de l'ARNg du VIH-1 contient plusieurs domaines fonctionnels : la tige-boucle TAR, la tige-boucle poly-A, le site de liaison de l'amorce pour la transcriptase inverse (PBS), le site de dimérisation de l'ARN (DIS), le site donneur majeur de l'épissage (SD) et le signal d'emballage de l'ARNg dans les virions (Ψ). Le site d'initiation interne pour la traduction coiffe-indépendante (IRES) s'étend entre les positions 104 et 336. Le signal stimulateur de *frameshift* et l'élément de réponse à Rev (RRE) sont deux autres structures de l'ARNm ayant des fonctions importantes (tiré et adapté de Brasey et al., 2003). B) Le génome viral contient neuf cadres de lecture encodant quinze protéines virales. La région 5' LTR du génome viral contient les régions régulant la transcription lorsque le génome viral est intégré à celui de l'hôte sous forme de provirus. Un LTR se retrouve aussi en 3' de l'ADNg viral, dû à la duplication du LTR par la transcriptase inverse (http://upload.wikimedia.org/wikipedia/en/e/e8/HIV_genome.png).



qui libère du complexe de pré-initiation localisé au promoteur et lui permet de passer du mode initiation au mode élongation. Cette activité de Tat est régulée par les transcrits terminés prématurément produits par IST qui possèdent TAR. TAR joue aussi un rôle dans la régulation de l'activité de la protéine cellulaire PKR, une protéine qui peut phosphoryler le facteur d'initiation eIF2 et inhiber l'initiation de la traduction (Zhang et al., 2001). Lorsqu'il est présent en grande quantité, TAR inhibe PKR en empêchant sa dimérisation, ce qui stimule la synthèse protéique. Par opposition, lorsque TAR est présent en faible quantité, il active PKR qui phosphoryle alors le facteur d'initiation de la traduction eIF2, ce qui inhibe l'initiation de la traduction coiffe-dépendante. Cette inhibition de l'initiation de la traduction pourrait avoir une incidence sur l'efficacité du *frameshift* ribosomique -1 programmé (voir section 1.4.2.1) et serait associée à une phase latente d'infection dans la cellule hôte (Tang et al., 1999). Une autre structure fonctionnelle des ARNm viraux, la tige-boucle poly-A, se retrouve aux extrémités des ARNm de par la duplication des LTR lors de la transcription inverse de l'ARNg viral pendant la phase précoce du cycle de réplication (voir section 1.2.2.1). Cette structure, lorsqu'elle est située en 5' des ARNm, ne peut entraîner la polyadénylation de l'ARN dû à la présence de la protéine U1 snRNP qui intervient aussi dans l'épissage (Ashe et al., 1997). Toutefois, lorsqu'elle est située à l'extrémité 3', la tige-boucle poly-A entraîne la polyadénylation des ARNm viraux. Le site de liaison de l'amorce pour la transcriptase inverse (PBS) est, comme son nom l'indique, la région où la transcriptase inverse initie la transcription de l'ARNg viral en ADN. Une portion qu'on appelle site de dimérisation de l'ARN (DIS) permet à deux molécules d'ARNg d'interagir ensemble, ce qui facilite l'incorporation de deux copies d'ARNg par virion. Le site donneur majeur de l'épissage (SD), quant à lui, est utilisé dans la production de tous les ARNm épissés. La région appelée Ψ (Psi) chevauche le codon initiateur du précurseur protéique Gag. Formé de quatre tiges-boucles, Ψ est le signal d'emballage des ARNg qui seront transportés à la membrane plasmique via une interaction avec le précurseur protéique Gag. La région comprise entre les positions 100 et 336 de l'ARNm contient l'IRES (*internal ribosomal entry site*) du VIH-1 (Brasey et al., 2003). On ignore toutefois quelle est la structure adoptée par l'ARNm lorsque l'IRES est fonctionnel. La présence d'un IRES permet une initiation de la traduction coiffe-indépendante des ARNm.

Bien que son utilisation et son importance dans la réplication virale soient controversées, cet IRES pourrait permettre au virus de synthétiser ses protéines en situation d'inhibition de la traduction coiffe-dépendante, comme en phase G2 du cycle cellulaire où les cellules infectées sont arrêtées en présence de la protéine virale Vpr. Près de l'extrémité 3' de la région codante de Gag se trouve la région de *frameshift* où se produit un changement de cadre de lecture programmé et auquel nous nous intéressons particulièrement. Cette région comprend deux éléments responsables du *frameshift* ribosomique -1 programmé : la séquence glissante où se produit le *frameshift* et le signal stimulateur de *frameshift* une structure secondaire particulière qui contrôle l'efficacité de *frameshift*. Le *frameshift* ribosomique -1 programmé ainsi que l'implication de ces deux éléments seront décrits en détail plus loin (section 1.4.2.1). Entre les positions 7710 et 8061 de l'ARNm pleine longueur (non-épissé) se retrouve le RRE (*rev response element*). Cette structure, lorsqu'elle est reconnue par la protéine virale Rev, permet l'exportation des ARNm non-épissés ou mono-épissés qui seraient normalement retenus dans le noyau, vu leur épissage incomplet. Cette exportation sera traitée de façon plus détaillée ultérieurement.

1.2.1.2 Les protéines virales

Les protéines du VIH-1 (figure 3B) peuvent être divisées en trois groupes fonctionnels : les protéines de structure, les enzymes ainsi que les protéines régulatrices et accessoires. Les protéines de structure sont traduites sous forme d'un précurseur appelé Gag (Pr55 Gag). Ce précurseur comprend, de l'extrémité N-terminale vers l'extrémité C-terminale, les protéines suivantes : la matrice (MA ou p27), la capside (CA ou p24), le peptide p2, la nucléocapside (NC ou p7), le peptide p1 et la protéine p6^{Gag}.

Gag est produit suite à la traduction de l'ARNm non-épissé (dit pleine longueur) du VIH-1. C'est suite au clivage séquentiel de Gag par la protéase virale que les protéines adopteront leur forme mature. Des six protéines résultant de Gag, MA, CA et NC sont celles qu'on qualifie de protéines de structure. MA possède une double fonction de localisation intracellulaire. Durant l'assemblage des virions, l'addition d'un groupement myristyle hydrophobe en N-terminal de MA permet son ciblage et son ancrage à la

membrane cellulaire (Tang et al., 2004). Ce groupement myristyle est exposé lorsque MA est présent sous forme de multimère. Suite à la maturation de Gag dans les virions, MA se retrouvera davantage sous forme de monomère et son groupement myristyle sera internalisé, assurant ainsi que, après leur entrée dans une cellule hôte, les virions ne soient pas retenus à la membrane plasmique. De plus, MA possède un domaine de localisation nucléaire (DLN) utilisé après le désassemblage des virions dans les cellules et qui participe à l'importation au noyau du domaine de pré-intégration (celui-ci sera examiné ultérieurement). La protéine CA possède aussi deux fonctions dans la réplication virale. La première est la multimérisation des précurseurs Gag. Le tiers C-terminal de CA possède un domaine d'homologie qui lui permet de se lier à ses homologues pour favoriser la multimérisation des précurseurs protéiques (Franke et al., 1994a). Le second rôle de CA est l'incorporation de la cyclophiline A (CypA) dans les virions (Franke et al., 1994b). Cette molécule est une chaperonne cellulaire qui est impliquée dans le processus de désassemblage des virions suite à l'infection d'une cellule (Endrich et al., 1999). NC joue son rôle via son aptitude à lier l'ARN. En plus de lier spécifiquement le signal d'empaquetage Ψ pour assurer l'incorporation d'ARNg dans les virions, NC lie l'ARN de manière non spécifique, ce qui lui permet de recouvrir l'ARNg pour le protéger des nucléases cellulaires (Aldovini & Young, 1990; Luban & Goff, 1994). En ce qui concerne les trois plus courtes protéines produites à partir du précurseur Gag, le clivage au niveau du peptide p2 régule la vitesse avec laquelle la protéase virale procède à la maturation des virions (Pettit et al., 1994), le peptide p1 semble stabiliser les dimères d'ARNg (Hill et al., 2007) alors que p6^{Gag} intervient dans l'importation de la protéine virale Vpr durant l'assemblage des virions (Kondo & Gottlinger, 1996), le relâchement des virions de la membrane plasmique lors du bourgeonnement (Gottlinger et al., 1991) et participerait aussi à la régulation de l'activité de la protéase virale (Huang et al., 1995).

Le précurseur des enzymes virales, Pol, est produit sous forme de protéine de fusion Gag-Pol lorsqu'une partie des ribosomes effectuent un *frameshift* -1 programmé à un endroit précis de l'ARNm pleine longueur du VIH-1, vers la fin de la région codante de Gag. Gag-Pol génère, suite à la maturation des précurseurs par la protéase virale, cinq

autres protéines (de la partie N-terminale de Pol vers l'extrémité C-terminale) : la protéine *trans-frame* (TRF), p6^{Pol}, la protéase (PR), la transcriptase inverse (RT ou p66/p51) et l'intégrase (IN ou p31). Suite au *frameshift* -1, p1 et p6^{Gag} sont remplacés par TRF et p6^{Pol}, deux protéines qui contrôlent l'activité de PR (Hill et al., 2002). PR, tout comme RT et IN, a été exhaustivement étudiée par plusieurs groupes et ciblée pour le développement d'agents anti-rétroviraux. PR est une protéase aspartique qui est active sous forme de dimère (Navia & McKeever, 1990). La concentration de Gag-Pol, et par conséquent le taux de *frameshift* ribosomique -1 programmé, sont déterminants pour le contrôle de l'activité de la protéase et dans l'infectivité des virions. D'une part, une surexpression de Gag-Pol mène à une activation prématurée de PR ce qui diminue l'infectivité virale (Sheng et al., 1997). En effet, l'activation de PR dans le cytoplasme nuit à l'assemblage des virions à la membrane et est toxique pour la cellule hôte (Krausslich, 1991). D'autre part, une baisse de l'expression de Gag-Pol diminue aussi l'infectivité virale. Cette baisse d'infectivité peut être attribuable à une activation trop lente de PR qui mène à des défauts de maturation ainsi qu'à une incorporation insuffisante d'ARNt^{Lys} et de RT dans les virions pour une infectivité optimale (Mak et al., 1994). Tel qu'il a été précédemment mentionné, le rôle principal de PR est de cliver séquentiellement les précurseurs Gag et Gag-Pol pour générer les protéines matures décrites ci-dessus. La RT accomplit deux fonctions via deux domaines. La première fonction est la synthèse d'ADN à partir d'une matrice d'ARN ou d'ADN, qui permet la transcription inverse de l'ARNg viral simple brin en ADN double brin. La seconde est une activité RNase qui digère l'ARNg suite à sa transcription en ADN. La RT existe sous deux formes différentes : p66, une forme complète possédant les deux domaines fonctionnels, et p51, une forme dépourvue du domaine RNase. L'enzyme n'est active que sous forme d'hétérodimère p66/p51. En ce qui concerne IN, elle permet l'intégration de l'ADN génomique viral généré par la RT dans le génome de la cellule hôte.

La protéine Env est un précurseur qui est traduit à partir d'un transcrit mono-épissé du génome viral. Une fois Env clivé par une protéase cellulaire (Freed, 2001), les protéines matures SU (gp120) et TM (gp41) sont générées. SU et TM peuvent être classées dans la catégorie des protéines de structure avec MA, CA et NC. SU médie l'entrée du virus dans

une cellule cible grâce à une interaction de haute affinité avec les récepteurs cellulaires CD4. Cette liaison entre SU et CD4 induit un changement de conformation de SU qui entraîne la formation d'un site de liaison pour un co-récepteur CCR5 ou CXCR4. C'est l'identité du corécepteur membranaire de SU qui détermine le tropisme des virions. Une fois SU lié à CD4 et à son co-récepteur, TM se charge de promouvoir la fusion du virion à la cellule hôte pour permettre son entrée.

Une panoplie de protéines virales auxiliaires provenant de transcrits mono-épissés (Vpr, Vpu, Vif) et multi-épissés (Tat, Rev, Nef) augmente l'infectivité du VIH-1. Comme expliqué précédemment, Tat promeut la transcription du génome viral via son interaction avec la tige-boucle d'ARN TAR. De plus, en compétitionnant avec PKR pour le facteur d'initiation eIF2, Tat inhibe l'action de PKR sur l'initiation de la traduction. En liant RRE, Rev permet l'exportation d'ARNm viraux non-épissés ou mono-épissés, qui ne peuvent pas être exportés via la machinerie nucléaire dû à un épissage incomplet. La protéine Nef (*numerous effector functions*) possède plusieurs fonctions, comme son nom l'indique. En premier lieu, Nef régule négativement les récepteurs CD4 de la surface cellulaire en provoquant leur internalisation et leur dégradation. Nef entraîne aussi l'internalisation et la dégradation des récepteurs de surface MHC I, récepteurs requis pour la présentation de peptides viraux aux lymphocytes T cytotoxiques (CTL), inhibant ainsi la réponse immunitaire médiée par ces CTL (Greenway et al., 2003). Nef serait aussi impliqué dans différents processus visant à moduler l'état d'activation de la cellules hôte ou l'infectivité des virions via son aptitude à recruter certaines kinases, mais ces fonctions sont peu caractérisées. Vpu agit de concert avec Nef dans la régulation négative des récepteurs CD4 en empêchant leur transport du réticulum endoplasmique (RE) à la membrane plasmique. Vpu est aussi actif dans le processus de bourgeonnement des virions. Vpu agit en activant l'exocytose des virions à la surface des cellules pour ensuite faciliter leur relâche de la membrane plasmique. Vif (*virion infectivity factor*) est une protéine virale dont le rôle a été cerné récemment. Ce facteur inhibe l'effet d'APOBEC3G, une cytidine désaminase cellulaire qui est empaquetée dans les virions et qui altère le génome viral lors

de la transcription inverse suite à l'entrée dans une cellule (Malim, 2006). Notons que Vif est requis pour la réplication virale uniquement dans les lignées cellulaires dites non-permissives où APOBEC3G est exprimé. La protéine accessoire Vpr est empaquetée en proportions équimolaires à Gag dans les virions. Une première fonction de Vpr est l'importation au noyau du complexe de pré-intégration du VIH-1 (composé de MA, IN, ADN_v viral, Vpr) via son DLN. Sa seconde fonction consiste à arrêter en phase G2 les cellules infectées. Il a été observé que les promoteurs LTR du VIH-1 sont plus actifs en phase G2. De plus, la traduction coiffe-dépendante étant inhibée en phase G2, le VIH-1 utiliserait la machinerie traductionnelle à son profit via son IRES.

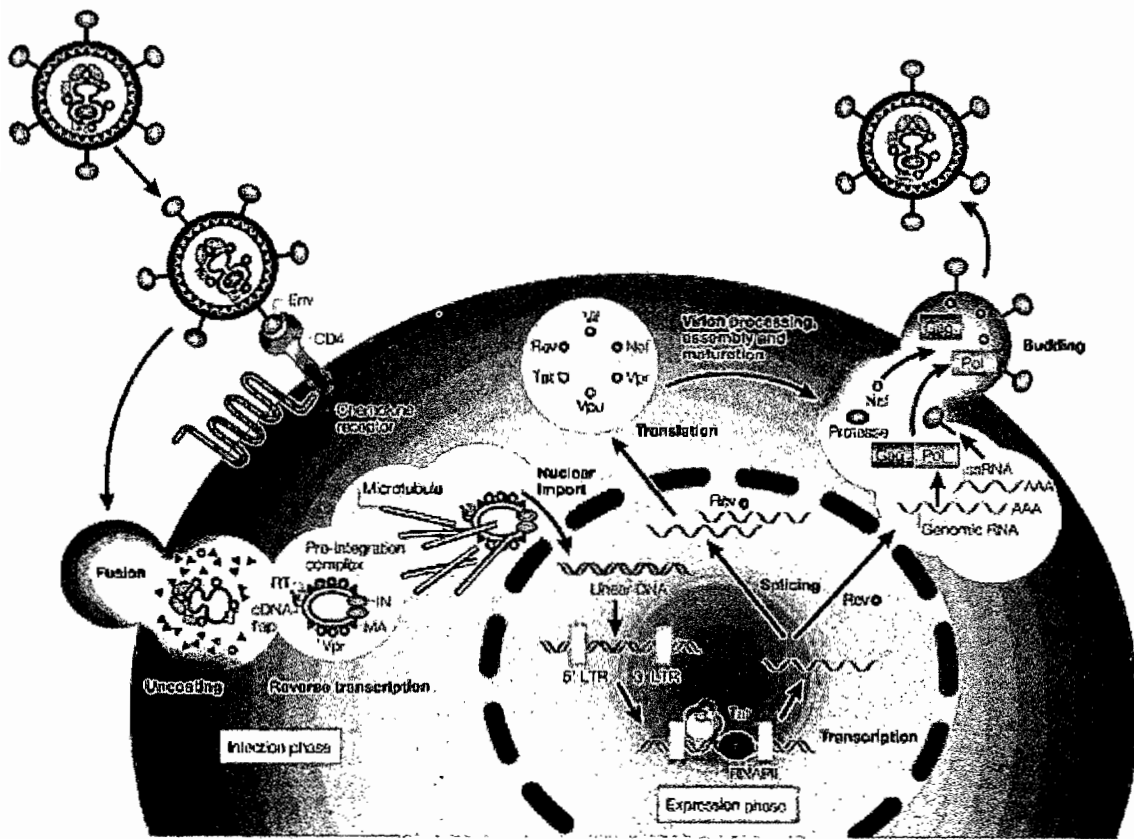
1.2.2 La réplication du VIH-1

Le cycle de réplication du VIH-1 (figure 4) comporte plusieurs étapes clés qui peuvent être regroupées en deux phases : les étapes précoces et les étapes tardives du cycle répliatif (revu dans Tang et al., 1999).

1.2.2.1 Les étapes précoces du cycle répliatif viral

Le cycle de réplication viral débute avec l'entrée du virion dans une cellule cible. Pour ce faire, le virion doit se lier à cette cellule via l'interaction entre SU et les récepteurs/corécepteurs membranaires pour ensuite fusionner à la cellule grâce à l'action de TM (figure 5). Une fois à l'intérieur de son hôte, le virus procède à sa décapsidation médiée par la cyclophiline A (CypA), une chaperonne cellulaire présente dans le virion. Cette décapsidation libère dans le cytoplasme les protéines de structure, les enzymes et l'ARN_v du virus. La RT procède ensuite à la transcription inverse de l'ARN_v simple brin en ADN double brin. Une fois synthétisé, l'ADN_v viral formera avec IN, MA et Vpr le complexe de pré-intégration qui sera activement transporté dans le noyau de la cellule infectée. Après l'importation du complexe dans le noyau, IN catalysera l'intégration de l'ADN_v viral dans le génome de la cellule hôte. Lorsque le génome du virus est inséré dans celui de la cellule, il est appelé provirus. Le provirus est transcrit à un niveau basal par la machinerie de

Figure 4. Le cycle de réplication du VIH-1. Un virion se lie à la surface d'une cellule en interagissant avec les récepteurs CD4 et CCR5/CXCR4. Le virus entre dans la cellule par fusion des membranes lipidiques. Le virion est désassemblé à l'intérieur de la cellule et la transcriptase inverse convertit le génome viral d'ARN simple brin en ADN double brin. L'ADNg viral est acheminé au noyau et l'intégrase catalyse son incorporation dans le génome de l'hôte. Le provirus est transcrit et divers ARNm sont générés via différents épissages. Ces ARNm sont exportés dans le cytoplasme et traduits en protéines. L'ARNg est incorporé avec des précurseurs protéiques dans des nouveaux virions. Cet assemblage se fait à la membrane plasmique. Il y a ensuite bourgeonnement et maturation de nouveaux virions.



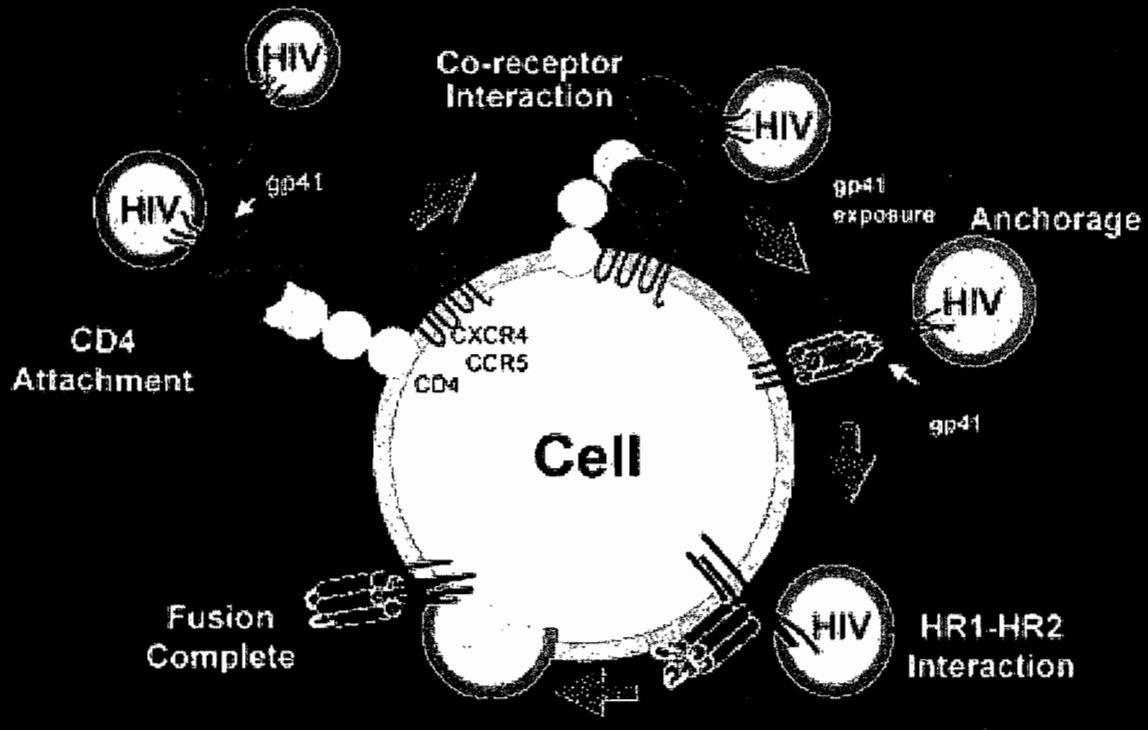
transcription nucléaire, les transcrits produits sont épissés et passent dans le cytoplasme, ce qui permet la production des protéines virales précoces Tat, Rev et Nef. Ces protéines pourront accomplir leurs rôles respectifs tels que décrits précédemment, ce qui marquera le début des étapes tardives du cycle de réplication.

1.2.2.2 Les étapes tardives du cycle répliatif

Une fois Tat synthétisé, la transcription du provirus sera grandement augmentée, ce qui permettra la génération des nombreux transcrits du VIH-1. La protéine Rev permettra l'exportation de transcrits incomplètement épissés (non-épissés ou mono-épissés) qui codent pour Vif, Vpr, Vpu, Env et Gag/Gag-Pol ou qui seront empaquetés en tant qu'ARNg dans les virions naissants dans le cas d'ARNm pleine longueur. Les précurseurs Gag et Gag-Pol s'assemblent dans le cytoplasme pour former des multimères via le domaine d'homologie de CA et le domaine de liaison à l'ARN de NC qui regroupe plusieurs exemplaires des précurseurs autour d'un ARNg. Cette multimérisation entraîne l'exposition du groupement myristyle de Gag, ce qui cible les multimères à la membrane plasmique (Gottlinger, 2001). Il y a alors assemblage des virions immatures de forme sphérique à la membrane plasmique (Jouvenet et al., 2006), suivi du bourgeonnement et de la relâche de virions immatures entourés d'une bicouche de phospholipides cellulaires. La protéase virale procédera enfin à la maturation des précurseurs protéiques pour générer les protéines virales matures et rendre les virions infectieux, leur conférant une forme conique (Gottlinger, 2001). Il est à noter que plusieurs facteurs cellulaires sont utilisés par le virus dans diverses étapes du cycle de réplication du VIH-1, mais que, par souci de simplicité, nous n'en parlerons pas dans cet ouvrage (revu par Tang et al., 1999).

Figure 5. L'entrée d'un virion dans une cellule. La glycoprotéine virale SU (gp120) interagit d'abord avec le récepteur cellulaire CD4. Cette liaison induit un changement de conformation de SU permettant son interaction avec un co-récepteur CCR5 ou CXCR4, selon le tropisme du virion. Cette seconde liaison permet le positionnement adéquat de TM (gp41) qui pourra alors induire la fusion des membranes lipidiques de la cellule et du virion afin de permettre l'entrée du virion dans la cellule.

HIV interaction with CD4 cell



1.3 Les agents anti-rétroviraux connus à ce jour

Afin de combattre le VIH-1, de nombreux composés anti-rétroviraux ont été développés (Temesgen et al., 2006). Les agents anti-rétroviraux couramment utilisés ciblent la protéase virale, la transcriptase inverse ou la fusion des virions aux cellules. Malheureusement, l'émergence de variants résistants à ces agents rend urgent le développement de nouveaux composés contre de nouvelles cibles, ce à quoi de nombreux groupes s'affairent.

1.3.1 Les agents ciblant la protéase virale

La protéase virale est essentielle à la maturation des précurseurs protéiques. Une inhibition de cette enzyme entraîne le relâchement de particules virales non-infectieuses ayant une structure désorganisée. De nombreux inhibiteurs de la protéase virale ont été développés et sont utilisés pour traiter les patients atteints du VIH-1. Ces molécules sont principalement des inhibiteurs compétitifs réversibles qui inhibent l'activité de la protéase virale en se liant directement au site actif de l'enzyme. Les premiers inhibiteurs de protéase virale utilisés sont qualifiés d'inhibiteurs de première génération. L'Atazanavir, le Fosamprenavir, l'Indinavir, le Lopinavir, le Nelfinavir, le Ritonavir et le Saquinavir en font partie. Afin de contrer les problèmes de résistance aux inhibiteurs de première génération et d'optimiser leur action inhibitrice, des modifications structurales ont été apportées pour obtenir des inhibiteurs de deuxième génération (Boggetto & Reboud-Ravaux, 2002). Ceux-ci comprennent entre autre le Tipranavir, le Darunavir et le KNI-764 (Vega et al., 2004). De plus, une autre catégorie d'inhibiteurs de protéase est en développement. Plutôt que de cibler le site actif de l'enzyme, ces nouveaux agents empêchent la formation de l'enzyme dimérique active (Boggetto & Reboud-Ravaux, 2002; Bannwarth & Reboud-Ravaux, 2007).

De nombreuses études ont permis de démontrer diverses propriétés autres que l'activité anti-rétrovirale de plusieurs agents ciblant la protéase. En effet, certains agents

comme le Ritonavir auraient des propriétés pro-apoptotiques chez des cellules cancéreuses (Ikezoe et al., 2004). Aussi, la co-administration de plusieurs inhibiteurs de protéase virale aurait un effet synergique en entraînant l'inhibition des enzymes hépatiques CYP3A responsables du métabolisme de ces médicaments (Ernest et al., 2005). Par contre, d'autres effets indésirables ont aussi été répertoriés tels une altération du métabolisme lipidique (Ranganathan & Kern, 2002) et le développement d'une résistance à l'insuline (Schutt et al., 2004).

1.3.2 Les agents ciblant la transcriptase inverse

La transcriptase inverse est responsable de la conversion de l'ARNg viral simple brin en ADN double brin qui est inséré dans le génome de l'hôte. Deux classes d'inhibiteurs existent contre cette enzyme. Une première classe comprend les analogues de nucléotides. Ces molécules sont structurellement similaires aux nucléotides utilisés dans la polymérisation de l'ADN, mais le groupement hydroxyle en 3' a été remplacé par un groupement différent afin d'empêcher la formation du lien phosphodiester nécessaire à l'élongation de l'ADN. Lorsque la transcriptase inverse inclut ces analogues à la chaîne d'ADN en croissance, elle produit un transcrit arrêté prématurément. Parmi ces analogues on trouve l'Abacavir, le Lamivudine, le Didanosine, l'Emtricitabine, le Tenofovir, le Stavudine et le Zidovudine. Le Racivir est le dernier analogue nucléotidique à avoir vu le jour et est présentement en phase I/II d'essais cliniques. Les effets secondaires les plus souvent rencontrés suite à un traitement avec les analogues de nucléotides sont relatifs à la toxicité mitochondriale. En effet, l'inhibition de la synthèse de l'ADN mitochondrial par ces médicaments peut conduire au développement de myopathies, de neuropathies et de pancréatites (Brinkman et al., 1998). Une deuxième classe d'inhibiteurs de la RT comprend les molécules non analogues de nucléotides. Ces inhibiteurs se lient directement à la RT de manière non compétitive. Leur liaison à un site différent du site actif de l'enzyme entraîne un changement de conformation du site actif perturbant ainsi son activité catalytique. L'Efavirenz, le Dalevirdine et le Nevirapine sont trois agents anti-rétroviraux faisant partie de cette catégorie d'inhibiteurs. Trois nouveaux agents de cette classe d'inhibiteur sont en

essais cliniques : l'Etravirine, le Rilpivirine et le BILR-355 BS. Les effets indésirables des inhibiteurs de RT non analogues de nucléotides sont moins graves que pour les analogues de nucléotides. Les plus communément rencontrés sont des défaillances hépatiques et des éruptions cutanées (Kappelhoff et al., 2005).

1.3.3 Les agents ciblant la fusion des virions aux cellules

Trois événements sont essentiels à l'entrée d'un virion dans une cellule cible. D'abord il y a interaction entre la protéine virale SU et le récepteur cellulaire de surface CD4. Cette liaison permet qu'une seconde interaction entre SU et un corécepteur CCR5 ou CXCR4 (selon le tropisme du virus) ait lieu. Finalement, TM procède à la fusion des membranes phospholipidiques entourant le virion et la cellule (Tang et al., 1999). L'Enfuvirtide et le Maraviroc (Fadel & Temesgen, 2007) sont à ce jour les seuls agents anti-rétroviraux ciblant l'entrée des virions dans les cellules qui ont été approuvés par la *Food and Drugs Administration* (FDA). L'Enfuvirtide est un court peptide qui agit au niveau de TM en l'empêchant de promouvoir la fusion des membranes phospholipidiques. Cet inhibiteur possède peu d'effets secondaires connus, la plupart se limitant à des réactions inflammatoires au site d'injection du peptide (Ball & Kinchelov, 2003). Le Maraviroc est une molécule chimique se liant au récepteur CCR5, compétitionnant ainsi avec les virions pour ce récepteur membranaire. Cet agent augmenterait toutefois le risque de maladies cardiovasculaires et démontre une toxicité hépatique chez certains patients. Deux autres antagonistes de CCR5, le PRO-140 et le Vicriviroc, et un inhibiteur de la liaison entre SU et CD4, le TNX-355, sont présentement à divers stades d'essais cliniques et pourraient être les premiers agents approuvés contre ces deux nouvelles cibles.

1.3.4 Les agents en développement contre de nouvelles cibles

Vu l'émergence de variants viraux résistants aux médicaments utilisés, plusieurs groupes se sont tournés vers de nouvelles cibles pour le développement de nouveaux agents anti-rétroviraux. Une des cibles visées est la maturation des virions suite à leur relâche. Le Bevirimat, un agent présentement en essais cliniques, empêche la maturation de CA. Si le

mécanisme d'action de ce composé reste encore à déterminer, l'hypothèse la plus plausible veut qu'il lie directement le site de clivage entre CA et p1 et inhibe ainsi l'action de la protéase virale (Li et al., 2003). La maturation incomplète de CA entraîne la production de particules virales non-infectieuses.

L'intégrase, la seule enzyme virale pour qui aucun inhibiteur approuvé n'existe, reçoit beaucoup d'attention. IN accomplit trois rôles distincts dans le processus d'intégration du génome viral dans celui de l'hôte. Tout d'abord, IN clive l'ADNg de l'hôte et agit comme exonucléase et retire deux nucléotides à chaque extrémité 3' de l'ADNg viral. Par la suite, il y a insertion de l'ADN viral dans le génome cellulaire où une ligation des extrémités 3' de l'ADNg viral avec le génome de l'hôte est effectuée. Finalement, l'enzyme retire les deux nucléotides non-pairés aux extrémités 5' de l'ADNg viral, puis procède à la ligation de ces extrémités dans le génome de l'hôte. C'est au niveau de l'étape d'insertion qu'un nouvel agent ciblant IN, le MK-0518, inhibe l'action de cette enzyme virale. Le MK-0518, un dérivé d'acide β -dikétoïque, est en phase II d'essais cliniques. D'autres dérivés d'acides β -dikétoïques sont aussi à des phases précoces de développement et pourraient agir sur la première fonction 3'-exonucléase de IN (Nair & Chi, 2007).

Une autre avenue innovatrice est celle de l'interférence de petits ARN (*siRNA*) contre l'ARN viral. De courts ARN pouvant se lier aux ARN viraux entraînent la dégradation de ces derniers par l'intermédiaire du complexe cellulaire RISC (Jacque et al., 2002). Une voie intéressante apparentée aux *siRNA* est celle des oligonucléotides antisens avec un squelette ribose-phosphate modifié, qui ne déclenchent pas d'action de la RNase H et qui démontrent du potentiel contre la réplication du VIH-1. Ces oligonucléotides antisens n'entraînent pas la dégradation d'ARNm cellulaires dont la séquence est voisine de la séquence virale ciblée, mais parviennent, par encombrement stérique, à bloquer la traduction d'ARNm viraux (Turner et al., 2006). Le GEM 91 est l'oligonucléotide antisens ayant démontré les meilleurs résultats jusqu'à maintenant. Celui-ci se lie à une région comprenant le codon initiateur AUG de Gag. Par contre, les siRNA comme les oligonucléotides antisens possèdent le désavantage d'avoir une faible demi-vie *in vivo* et leur administration pose des problèmes.

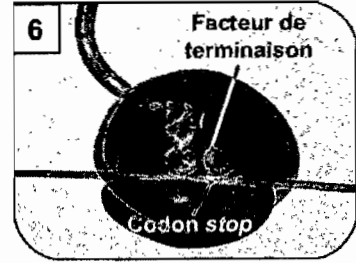
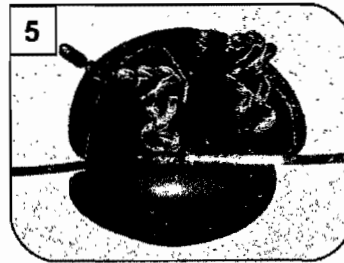
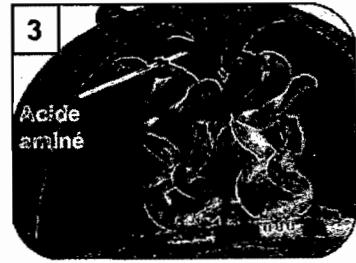
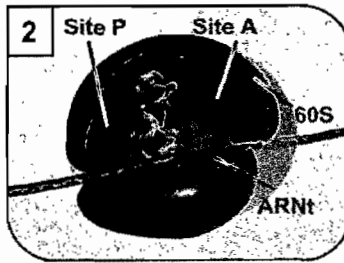
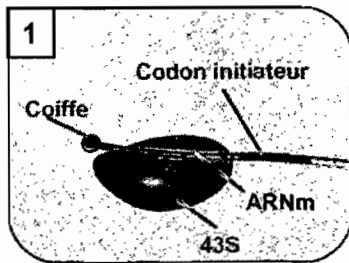
Récemment, la protéine Vif a été cernée comme cible pour le développement d'agents contre le VIH-1. Il a été montré que de courts peptides liant Vif qui empêchent sa multimérisation diminuaient l'infectivité virale. En effet, le rôle de Vif dans la dégradation d'APOBEC3G chez les lignées cellulaires non-permissives nécessite que Vif se retrouve sous forme de multimère (Miller et al., 2007).

Finalement, le *frameshift* ribosomique -1 programmé est une autre cible qui attire l'attention. En 1998, le criblage d'une librairie de molécules chimiques permit d'identifier un composé augmentant ce *frameshift* programmé (Hung et al., 1998). Nommée RG501, cette molécule à la structure s'apparentant considérablement à la spermidine, une polyamine, augmente le taux de *frameshift* d'un facteur deux. Des essais de réplication virale ont démontré que le RG501 diminuait de 50% la réplication du virus chez les lymphocytes T. Toutefois, le composé est toxique à faible concentration chez cette lignée cellulaire et aucun résultat sur le RG501 ou un de ses dérivés ne fut publié depuis. En 2007, une molécule qui lie le signal stimulateur de *frameshift* du VIH-1 a été sélectionnée à partir d'une librairie dynamique de composés chimiques liés à une résine (*resin-bound dynamic combinatorial chemistry* ou RBDCC). Toutefois, son effet sur la réplication virale n'a pas encore été étudié (McNaughton et al., 2007).

1.4 La traduction conventionnelle et le *frameshift*

Le ribosome est un grand complexe ribonucléoprotéique responsable de la traduction des ARNm en protéines (revue de la traduction dans Ramakrishnan, 2002; Kapp & Lorsch, 2004). Le ribosome eucaryote possède deux sous-unités nommées d'après leur coefficient de sédimentation : la 40S ou petite sous-unité et la 60S ou grande sous-unité. Chez les procaryotes, les deux sous-unités sont plus petites et ont un coefficient de sédimentation de 30S et 50S respectivement. Lorsque le ribosome effectue la traduction, l'ARNm transite à travers un tunnel dans le ribosome appelé tunnel de l'ARNm. À l'entrée de ce tunnel se retrouve l'activité hélicase ribosomique chargée de défaire les structures secondaires dans l'ARNm qui s'opposent au déplacement du ribosome lors de la traduction (Takyar et al., 2005). La traduction (figure 6) s'effectue en trois étapes distinctes, soient l'initiation, l'élongation et la terminaison. L'élongation et la terminaison sont semblables chez les eucaryotes et les procaryotes, alors que l'initiation est différente. Durant l'initiation chez les procaryotes, la sous-unité 30S est associée aux facteurs d'initiations IF1, IF2, IF3, à l'ARNt^{Met} chargé d'une méthionine formylée ainsi qu'à une molécule de GTP. Sur l'ARNm, une séquence riche en purines appelée Shine-Dalgarno à proximité d'un codon initiateur AUG est reconnue par la portion complémentaire dite anti-Shine-Dalgarno située à l'extrémité 3' de l'ARNr 16S. Cette interaction permet le recrutement du complexe de préinitiation aux codons initiateurs de l'ARNm suivi du recrutement de la sous-unité 50S du ribosome et du début de l'étape d'élongation. Chez les eucaryotes, la sous-unité ribosomique 40S s'associe avec un ARNt^{Met} chargé d'une méthionine ainsi qu'avec le facteur d'initiation eIF2 et une molécule de GTP pour former le complexe de pré-initiation 43S. Via le facteur eIF4F, lui-même lié à la coiffe à l'extrémité 5' de l'ARNm, le complexe 43S sera recruté à l'ARNm. C'est la traduction coiffe-dépendante. Une fois recruté, le complexe 43S effectuera un balayage en direction 3' de l'ARNm jusqu'à ce qu'il rencontre un codon initiateur. Il y a alors départ des facteurs d'initiation et recrutement de la sous-unité 60S entraînant la formation du ribosome fonctionnel et le début de l'étape

Figure 6. La traduction conventionnelle chez les eucaryotes. L'initiation commence par la liaison, via plusieurs facteurs cellulaires non-illustrés, du complexe 43S à l'extrémité 5' de l'ARNm dans une interaction impliquant la coiffe (1). Le complexe 43S effectue ensuite un balayage en direction 3' de l'ARNm jusqu'à ce qu'il rencontre un codon initiateur AUG complémentaire à l'anticodon de l'ARNt^{Met} chargé de méthionine lié au site P. Il y a alors recrutement de la grande sous-unité ribosomique et formation du ribosome complet fonctionnel (2). L'élongation s'amorce alors et un ARNt-aa complémentaire au codon du messager au site A se lie au site hybride A/T, puis est accommodé au site A (3). Il y a formation du lien peptidique, où la chaîne en croissance sur l'ARNt au site P est transférée sur l'acide aminé présent sur l'ARNt au site A. Le ribosome effectue une translocation de sorte que l'ARNt désaminoacylé se retrouve au site E, l'ARNt avec la chaîne peptidique au site P et que le site A est vacant (4). L'ARNt désaminoacylé est expulsé lors de l'accommodation d'un nouvel ARNt-aa au site A (5). Ce processus se répète jusqu'à ce qu'un codon stop soit rencontré et que les facteurs de terminaison entraînent l'arrêt de la traduction et le relâchement du peptide (6). (<http://vcell.ndsu.nodak.edu/animations/translation/advanced.htm>)



d'élongation. Toutefois, certains ARNm possèdent une région hautement structurée appelée IRES (site d'entrée interne du ribosome) qui permet le recrutement du complexe de pré-initiation à proximité d'un codon AUG initiateur de l'ARNm. Il s'agit alors d'une initiation de la traduction coiffe-indépendante.

Suite à l'initiation, les codons sont décodés un à un pour former la chaîne polypeptidique. C'est l'étape d'élongation qui se subdivise en trois sous-étapes : l'accommodation d'un aminoacyl-ARNt correspondant au codon de l'ARNm, la formation du lien peptidique contrôlée par le ribosome et la translocation des ARNt. Les ARNt font la liaison entre l'information contenue dans l'ARNm et la croissance de la chaîne peptidique. Les ARNt possèdent une boucle anticodon pouvant s'apparier avec un codon complémentaire de l'ARNm ainsi qu'une tige acceptrice porteuse d'un acide aminé correspondant à l'identité de la boucle anticodon. Pour qu'un anticodon soit complémentaire à un codon, les deux premières bases doivent respecter un appariement Watson-Crick (A-U, G-C) alors que la troisième position, appelée *wobble*, accepte certains autres appariements moins stables tels que G-U. Il existe trois sites classiques sur le ribosome où les ARNt peuvent se lier : le site A où se trouvent l'ARNt aminoacylé, le site P où se trouve l'ARNt porteur de la chaîne peptidique en croissance et le site E, occupé par l'ARNt désaminoacylé qui s'apprête à quitter le ribosome. En plus, il y a un site supplémentaire exclusivement sur la grande sous-unité, le site T, avec lequel les ARNt aminoacylés interagissent avant l'accommodation. Lors de la traduction, les ARNt occupent successivement les sites de liaison avec le ribosome appelés A/T, A, P et E. La tige-boucle anticodon et la tige acceptrice d'un ARNt peuvent occuper des portions correspondant à un même site sur leur sous-unité ribosomique respective, soit les portions du site A, du site P ou du site E. Il s'agit alors de l'occupation de sites classiques. La tige-boucle anticodon et la tige acceptrice d'un ARNt peuvent aussi occuper des sites différents lors de l'accommodation ou lorsque la portion anticodon se déplace avant la tige acceptrice durant la translocation. Un ARNt qui chevauche deux sites différents est dit occupant un site hybride tel que A/T, P/A et E/P (où la première lettre désigne la position de l'ARNt par rapport à la petite sous-unité et la seconde lettre la position par rapport à la grande). Les

ARNt se lieut sous forme de complexe ternaire avec le facteur d'élongation eEF1A (EF-Tu chez les procaryotes) et une molécule de GTP au site hybride A/T. Il y a accommodation de l'ARNt au site A si sa boucle anticodon est complémentaire au codon de l'ARNm situé au site A. Il y a ensuite formation d'un lien peptidique, avec transfert de la chaîne peptidique en croissance de l'ARNt au site P à l'ARNt aminoacylé au site A. Les tiges acceptrices de l'ARNt désaminoacylé et de l'ARNt porteur de la chaîne peptidique, aux sites P et A respectivement, se déplacent alors de sorte que les ARNt occupent respectivement les sites P/E et A/P. Ensuite, via l'action d'eEF2 (EF-G chez les procaryotes) les tige-boucles anticodons de ces ARNt effectuent à leur tour une translocation et ces ARNt occupent les sites classiques E et P respectivement. L'ARNt désaminoacylé au site E sera finalement éjecté lorsqu'un nouvel ARNt aminoacylé sera accommodé au site A.

L'étape de terminaison survient lorsque le ribosome rencontre un codon UGA, UAG ou UAA appelé codon de terminaison ou codon *stop*. Il y a libération de la chaîne protéique et le ribosome se dissocie de l'ARNm grâce à des facteurs de terminaison. Chez les eucaryotes, la circularisation des ARNm par des facteurs cellulaires permet le recyclage rapide des ribosomes ayant terminé la traduction dans un nouveau cycle traductionnel.

Le ribosome traduit les ARNm avec une fidélité élevée. Toutefois, il arrive que le ribosome effectue des erreurs lors de la traduction comme l'incorporation d'acides aminés ne correspondant pas à la séquence de l'ARNm, la poursuite de la traduction malgré la rencontre d'un codon *stop* ou un changement de cadre de lecture. Un *frameshift* survient lorsque le ribosome glisse d'une base en direction 5' (*frameshift* -1) ou 3' (*frameshift* +1). Un tel événement survient de manière accidentelle à une fréquence approximative de 0,003% par codon (Farabaugh, 1996). Toutefois, plusieurs *frameshifts* non accidentels, nommés *frameshifts* programmés, de fréquence plus élevée existent. Ils confèrent à ce qui est habituellement une erreur de lecture du ribosome un rôle régulateur dans l'expression d'un gène. Tel que mentionné plus haut, les événements de *frameshift* programmé sont contrôlés par deux éléments présents sur l'ARNm : une séquence glissante où le changement de cadre de lecture se produit et un signal stimulateur de *frameshift*, une séquence ou une structure secondaire dans l'ARNm qui contrôle l'efficacité de *frameshift*.

1.4.1 Les *frameshifts* programmés chez les eucaryotes et les procaryotes

Il existe plusieurs cas connus où l'expression d'une protéine est régulée par l'entremise du *frameshift* (revu dans Baranov et al., 2002; Namy et al., 2004). Un premier exemple concerne le facteur de terminaison de la traduction RF2 chez la plupart des eubactéries. Ce facteur est nécessaire pour que la terminaison de la traduction ait lieu lorsqu'un codon *stop* UGA ou UAG est présent au site de décodage. Le début de la protéine RF2 est encodé dans le cadre 0 de l'ARNm, mais la rencontre d'un codon UGA génère une protéine tronquée non-fonctionnelle lorsque la traduction s'y termine. La portion C-terminale de la RF2 est encodée dans le cadre de lecture +1 et est traduite lorsque des ribosomes rencontrant le codon *stop* UGA positionné au site A effectuent un *frameshift* +1. Deux éléments stimulateurs interviennent dans le cas de ce *frameshift* programmé. Le premier est le codon *stop* lui-même, responsable d'une pause effectuée par le ribosome, et le second est une séquence riche en purines, semblable à un motif Shine-Dalgarno, en 5' du codon UGA qui interagit avec la région 3' de l'ARNr 16S (Weiss et al., 1988) et stimule le *frameshift* en induisant une tension dans l'ARNm. L'expression de RF2 est auto-régulatoire : si ce facteur de terminaison est abondant, les ribosomes termineront la traduction au codon UGA du cadre 0, alors que si ce facteur est présent en faible quantité, des ribosomes seront immobilisés au codon UGA. Au cours de cette pause, le ribosome et l'ARNt^{Leu} au site P effectueront un glissement d'une base en direction 3' (la séquence où se fait le glissement est CUU UGA C et le codon avec lequel interagit l'ARNt^{Leu} est souligné). Le ribosome est alors placé dans le cadre +1 (C UUU GAC), la tension dans l'ARNm est libérée et la traduction se poursuit dans ce nouveau cadre de lecture.

Un autre exemple de *frameshift* programmé chez *E. coli* implique la traduction de l'ARNm codant pour l'ADN polymérase III. Cette enzyme est formée de dix sous-unités différentes. L'expression de deux d'entre elles, gamma(γ) et tau(τ), dépend d'un *frameshift* ribosomique -1. Il a été montré que, dans les régions de *frameshift* -1 programmé, la séquence glissante possède la séquence consensus suivante : X XXY YYZ. Des études ont

démontré que le *frameshift* est possible lorsque X est n'importe quel nucléotide, alors que Y doit être A ou U pour le permettre (Brierley et al., 1992). Si Z = G, le *frameshift* est défavorisé chez les mammifères et la levure, contrairement à ce que l'on observe chez les bactéries. Chez le gène codant pour l'ADN polymérase III de *E. coli*, une séquence glissante A AAA AAG (AAA au site P, AAG au site A) se retrouve aux deux tiers de l'ARNm codant pour les deux sous-unités et est suivie cinq nucléotides en aval par une tige boucle irrégulière agissant en tant que signal stimulateur de *frameshift* (Tsuchihashi & Brown, 1992). Une séquence ressemblant à la séquence Shine-Dalgarno est aussi présente en amont de la séquence glissante et constitue un second signal stimulateur de *frameshift* en interagissant avec l'extrémité 3' de l'ARNr 16S. Dans ce cas-ci, la distance sur l'ARNm entre la séquence où s'effectue le *frameshift* et la séquence Shine-Dalgarno est beaucoup plus grande que dans l'exemple avec RF2. Dans ces conditions, l'interaction entre le ribosome et la séquence Shine-Dalgarno favorise le *frameshift* -1 d'une façon encore inexpliquée. Alors que τ est produit lorsque l'ARNm est traduit de manière conventionnelle, jusqu'à 50% des ribosomes vont effectuer un *frameshift* -1 au niveau de la séquence glissante pour ensuite rapidement rencontrer un codon stop et produire γ . Ce haut taux de *frameshift* est possible du fait qu'il n'existe qu'un type d'ARNt^{Lys} dont l'anticodon est UUU, l'ARNt^{Lys} dont l'anticodon est UUC n'étant pas exprimé chez *E. coli*. De plus, cet ARNt^{Lys} (UUU) possède une uridine modifiée (5-méthylaminométhyl-2-thiouridine) en position 3 de son anticodon (UUU^{mod}) qui a une faible affinité pour un G par comparaison à une uridine non-modifiée. De ce fait, un ARNt^{Lys} (UUU^{Mod}) se lie préférentiellement à un codon AAA que AAG, ce qui favorise le *frameshift* -1 (cadre 0 : A AAA AAG; cadre -1 : AAA AAA G (le codon avec lequel l'ARNt^{Lys} se lie est souligné)). Il est à noter que l'expression d'un ARNt^{Lys} (UUC) chez *E. coli* diminue significativement le taux de *frameshift* chez le gène de l'ADN polymérase III.

Les eucaryotes possèdent aussi leur lot de *frameshifts* régulateurs. Un des cas les plus connus est celui de l'antizyme chez les mammifères. Cette enzyme contrôle le niveau de polyamines cellulaires. Pour produire la forme fonctionnelle de l'antizyme, les ribosomes doivent effectuer un *frameshift* +1 à une séquence glissante comprenant le codon

stop du cadre de lecture 0 (UCC UGA U) (Ivanov & Atkins, 2007). Cette séquence glissante est suivie d'un signal stimulateur, ici un pseudonoeud, une structure secondaire dans l'ARNm composée de deux tiges-boucles où la boucle de l'une des tiges participe à la formation d'une seconde tige. D'une manière encore mal définie, la présence de polyamines augmente le taux de ce *frameshift* +1. Cette voie est donc auto-régulatoire, puisqu'une faible concentration d'antizyme permet l'accumulation de polyamines, ce qui entraîne une augmentation du taux de *frameshift* +1 et, par le fait même, une production accrue d'antizyme qui réduira alors la concentration de polyamines.

Chez l'homme, deux gènes utilisent un *frameshift* ribosomique -1 programmé. Notons que ces deux gènes semblent être d'origine rétrovirale et qu'aucun gène propre à l'humain utilisant un tel mécanisme n'a jusqu'à maintenant été découvert. Le premier, PEG10, possède une région de *frameshift* s'apparentant aux rétrovirus où une séquence glissante G GGA AAC suivie d'un pseudonoeud agissant en tant que signal stimulateur assurent l'efficacité du *frameshift* (Shigemoto et al., 2001; Manktelow et al., 2005). Le produit du *frameshift* de PEG10 est une protéase aspartique comme celle retrouvée chez le VIH-1. Les produits de PEG10 sont exprimés durant l'embryogenèse et jouent un rôle dans le développement, mais ne sont pas normalement exprimés chez les tissus adultes (Clark et al., 2007). Notons que s'il est anormalement exprimé chez l'adulte, PEG10 posséderait des effets oncogènes en interagissant avec SIAH1 et SIAH2, deux supprimeurs tumoraux, et serait impliqué dans l'apparition de plusieurs cancers, particulièrement hépatiques (Okabe et al., 2003; Clark et al., 2007). Le second cas de *frameshift* programmé chez l'homme concerne le gène *Ma3* (Wills et al., 2006). *Ma3* fait partie d'une famille de gènes dont les produits sont la cible de réponses auto-immunes liées à des troubles paranéoplasiques associés à une tumeur. Les rôles des protéines produites par *Ma3* sont toujours inconnus et on ignore si le produit du *frameshift* a une fonction.

1.4.2 Le *frameshift* ribosomique -1 programmé chez le VIH-1

Un grand nombre de virus utilisent le *frameshift* -1 programmé pour produire le précurseur de leurs enzymes. De ce nombre, mentionnons le virus jaune occidental de la

betterave (BWYV), le virus de la bronchite infectieuse (IBV), le coronavirus causant le syndrome respiratoire aigu sévère (SARS-CoV) et, évidemment, le VIH-1. Le *frameshift* chez le VIH-1 a été l'objet de plusieurs études, ce qui a permis de cerner les éléments qui le contrôlent, de démontrer son importance pour le virus et aussi de formuler des mécanismes qui expliquent cet événement.

1.4.2.1 Les éléments contrôlant l'efficacité de *frameshift* chez le VIH-1

Le *frameshift* ribosomique -1 programmé du VIH-1 se produit lors de la traduction de l'ARNm pleine longueur du virus. Lorsqu'ils traduisent le précurseur Gag, les ribosomes rencontrent la région de *frameshift* vers la fin de la séquence codante de Gag. Alors que la majorité des ribosomes vont continuer la traduction de Gag dans le cadre 0 jusqu'à un codon *stop* à proximité pour produire ce précurseur, une minorité de ribosomes effectuent un *frameshift* -1, poursuivent la traduction dans ce cadre de lecture et synthétisent le précurseur Pol sous forme d'une protéine de fusion Gag-Pol. Deux éléments en cis présents dans la région de *frameshift* contrôlent l'efficacité du *frameshift* -1 programmé chez le VIH-1. Le premier élément, la séquence glissante, a comme séquence UUUU UUA, respectant le consensus X XXY YYZ rencontré dans les régions de *frameshift* -1 programmé tel que mentionné dans la section 1.4.1.

Il a par la suite été démontré que les codons flanquant la séquence glissante étaient importants pour l'efficacité de *frameshift*. D'une part, la présence de purines dans le codon en 3' pourrait diminuer l'efficacité du *frameshift* (Bertrand et al., 2002). Cette influence serait due à la stabilisation de l'interaction codon-anticodon par des purines, mais non par des pyrimidines, en 3' par effet d'empilement (*base stacking*).

Kim *et al.* ont démontré en 2001 que l'identité de la région immédiatement en 5' de la séquence glissante pouvait altérer l'efficacité de *frameshift* en remplaçant celle du VIH-1 par celle d'un autre virus. De plus, des études de mutagenèse du codon au site E du VIH-1 lors du *frameshift* de notre laboratoire ont ensuite démontré que le codon en 5' de la séquence glissante est impliqué dans le *frameshift* (Leger et al., 2007). Appelons ce codon

BCX et la base qui le précède A (A BCX XXY YYZ dans le cadre 0). Les variations dans l'efficacité de *frameshift* ont été observées, dépendamment de la nature du codon BCX. Ces variations seraient reliées à l'identité de l'ARNt correspondant à BCX (voir plus loin dans la section 1.4.2.2).

En absence du signal stimulateur de *frameshift*, la séquence glissante du VIH-1 induit un taux de *frameshift* -1 basal de 0,2 à 0,4% selon la lignée cellulaire (Reil et al., 1993; Dulude et al., 2002). La faible expression de la luciférase dans les lymphocytes T, cible naturelle du VIH-1, n'a pu permettre de déterminer le niveau basal de *frameshift* qui y est produit par la séquence glissante (Reil et al., 1993).

Le second élément de l'ARNm pleine longueur du VIH-1 contrôlant l'efficacité de *frameshift*, le signal stimulateur, est une structure secondaire particulière dans l'ARNm située immédiatement en aval de la séquence glissante. Celui-ci augmente l'efficacité de *frameshift* d'un facteur d'environ dix à vingt selon les résultats rapportés. L'efficacité de *frameshift* chez les lymphocytes T mesurée par notre laboratoire avec un système à double luciférase est d'environ 5% (Gendron et al., 2007). Dans ce système rapporteur, la luciférase de *Renilla* et la luciférase de luciole sont exprimées sur le même ARNm. La région de *frameshift* du VIH-1 a été insérée entre les deux luciférases, de sorte que la luciférase de luciole soit produite par les ribosomes qui effectuent un *frameshift* -1. La structure du signal stimulateur du sous-type B du groupe M du VIH-1 a pendant plusieurs années été controversée. La structure initialement proposée pour ce signal était une tige-boucle de onze paires de bases coiffée d'une tétraboucle ACAA qui fut nommée signal stimulateur classique. Il a ensuite été proposé que la structure du signal stimulateur du sous-type B était plus complexe. Deux nouveaux modèles ont alors été proposés (figure 7). Le premier, proposé par notre groupe, est celui d'une hélice à deux tiges coiffée d'une tétraboucle ACAA où les deux tiges sont interrompues par une excroissance de purines (Dulude et al., 2002). Dans ce modèle, la tige-boucle supérieure correspond au signal classique alors que la tige inférieure résulte de l'appariement des séquences nucléotidiques flanquant le signal classique. Le second modèle est un pseudonoeud impliqué dans un triplex intramoléculaire (Dinman et al., 2002). Les groupes de Butcher et Fourmy ont

tranché le débat en 2005 par des études de RMN sur le signal stimulateur (figure 8) et ont confirmé le modèle de l'hélice à deux tiges présenté par Dulude *et al.* (Gaudin et al., 2005; Staple & Butcher, 2005). La tige supérieure est celle ayant la plus grande importance dans la promotion du *frameshift*. Elle stimule le *frameshift* d'un facteur d'environ dix. La délétion de la tige supérieure en ne laissant que la tige inférieure diminue l'efficacité de *frameshift* au niveau basal de la séquence glissante (Dulude et al., 2002). La tige inférieure quant à elle a une plus faible importance dans l'efficacité de *frameshift*. En effet, une déstabilisation de cette dernière diminue approximativement d'un facteur deux l'efficacité de *frameshift*. Il est à noter qu'améliorer la stabilité de la tige inférieure n'augmente pas l'efficacité de *frameshift*. Le rôle de cette tige serait de positionner la tige supérieure afin de favoriser une interaction initiale avec le ribosome (voir section 1.4.2.2).

Une étude de notre laboratoire a permis de déterminer que l'inhibition de l'initiation de la traduction augmente l'efficacité de *frameshift* (Gendron et al., 2007). Ce phénomène s'explique par la plus grande distance entre les ribosomes traduisant les ARNm lorsque l'initiation de la traduction est inhibée, situation rencontrée par exemple lorsque PKR phosphoryle le facteur d'initiation eIF2 (section 1.2.1.1). Cet écart accru entre les ribosomes permet au signal stimulateur de *frameshift* de se reformer après qu'un ribosome l'ait déroulé suite à sa rencontre. À l'inverse, une augmentation de l'initiation de la traduction diminue le taux de *frameshift* en réduisant l'espacement entre les ribosomes qui effectuent la traduction d'un même ARNm et donc la probabilité que les ribosomes rencontrent un signal stimulateur reformé.

Figure 7. Les modèles du signal stimulateur de frameshift du sous-type B du VIH-1. À gauche, le modèle démontré par Dulude et al. (2002). Celui-ci est une hélice à deux tiges interrompues par une excroissance de purines. À droite, le modèle proposé par Dinman *et al.* (2003). Ce dernier consiste en un pseudonoeud avec un triplex intramoléculaire entre la tige-boucle et des nucléotides situés en 3' de la tige (Dulude et al., 2002).

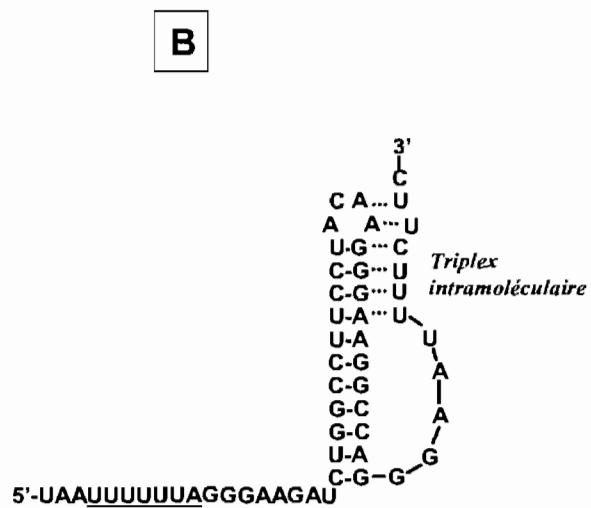
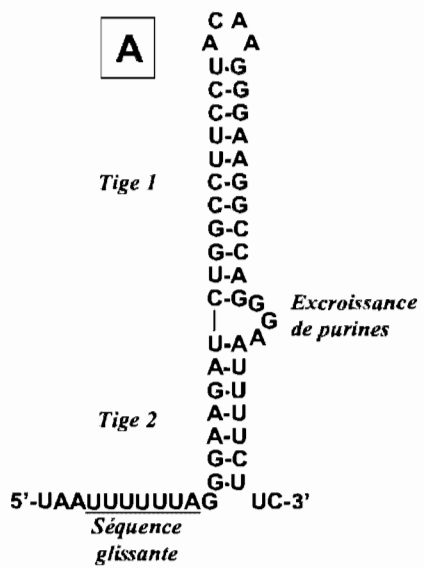


Figure 8. Le signal stimulateur de frameshift du sous-type B du VIH-1. Structure du signal stimulateur de *frameshift* obtenue par résonance magnétique nucléaire (Staple & Butcher, 2005).



1.4.2.2 Le mécanisme du *frameshift* -1 chez le VIH-1

La première étape menant au *frameshift* serait la rencontre du signal stimulateur complet par le ribosome (Leger et al., 2007). À ce moment, la séquence glissante est située en 3' du centre de décodage du ribosome. La tige inférieure du signal stimulateur est ensuite déroulée par le ribosome qui poursuit sa traduction jusqu'à ce qu'il rencontre la tige supérieure à l'entrée du tunnel de l'ARNm. La tige supérieure, de par la stabilité de sa structure secondaire, résiste à son déroulement par l'activité hélicase ribosomique et force le ribosome à faire une pause. La séquence glissante (X XXY YYZ) est à ce moment au centre de décodage, le site P étant occupé par le codon XXY et le site A par le codon YYZ. La pause favorise le *frameshift*, mais n'est pas suffisante pour expliquer l'efficacité de *frameshift* (Kontos et al., 2001). Il est suggéré qu'une interaction spécifique entre le signal stimulateur et le ribosome est responsable de cette efficacité accrue.

Plusieurs modèles ont été suggérés pour expliquer le mécanisme du *frameshift* ribosomique -1 chez le VIH-1 (figure 9I) (revus dans Leger et al., 2007). Le premier modèle proposé, provenant du groupe de Varmus (Jacks et al., 1988), a été celui du glissement simultané. Dans ce modèle, l'ARNt portant la chaîne peptidique au site P (ARNt-peptidyle) et l'ARNt aminoacylé au site A (ARNt-aa) sont aux sites P et A respectivement quand la séquence glissante occupe le centre de décodage du ribosome. Les deux ARNt se dissocient ensuite simultanément de l'ARNm pour effectuer, avec le ribosome, un mouvement d'un nucléotide en direction 5' de l'ARNm. Les deux ARNt se réappariaient ensuite à l'ARNm dans le cadre de lecture -1 et le ribosome poursuit alors la traduction dans ce cadre de lecture. Ce modèle a été critiqué parce qu'il ne tenait pas compte de la très grande rapidité à laquelle se produit la formation du lien peptidique (Rodnina & Wintermeyer, 2001).

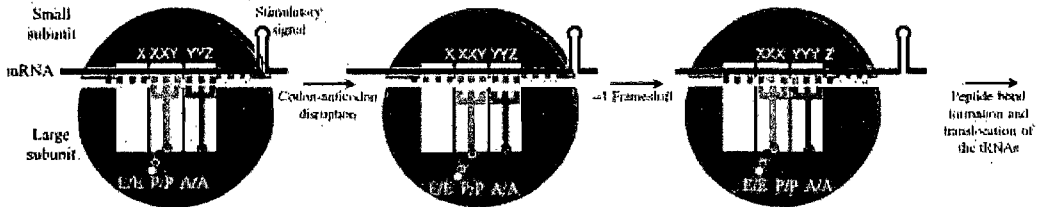
En conséquence, le temps disponible pour que le *frameshift* ait lieu suite à l'accommodation de l'ARNt-aa au site A est extrêmement limité.

Figure 9. Les modèles mécanistiques du frameshift -1. I. En jaune, de gauche à droite, les sites E, P et A. L'ARNt désaminoacylé est en vert, l'ARNt porteur de la chaîne peptidique en rose et celui chargé d'un acide aminé en bleu. Les triplets de même couleur de l'ARNm représentent les codons du cadre 0. A) Modèle proposé par Jacks et al. (1988) et développé par Plant et al. (2003). Le glissement simultané des ARNt s'effectue alors que ceux-ci sont aux sites P et A avant la formation du lien peptidique. B) Modèle proposé par Weiss et al. (1989). Le *frameshift* survient suite à la formation du lien peptidique alors que les ARNt se retrouvent aux sites hybrides P/E et A/P. C) Modèle suggéré par Farabaugh et appuyé par Léger et al. (2004). Le *frameshift* se produit avant la formation du lien peptidique, alors qu'un ARNt-aa n'est pas encore accommodé au site A. L'ARNt-peptidyle et l'ARNt-aa se situent respectivement aux sites P et A/T.

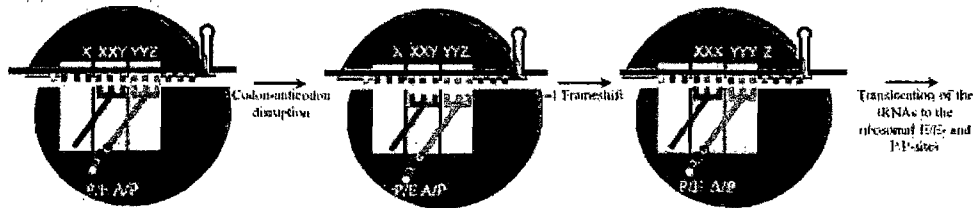
II. Modèle proposé par Léger et al. (2007). Suite à la formation du lien peptidique, il y a translocation des ARNt désaminoacylé et peptidyle vers les sites E et P respectivement. Toutefois, la résistance au déroulement du signal stimulateur à l'entrée du tunnel de l'ARNm du ribosome bloque ces ARNt aux sites E*/E* et P*/P*, où « * » indique un état de transition translocationnel. Le ribosome sélectionne ensuite un nouvel ARNt-aa au site A*/T*, aussi bloqué dans un état de transition. Successivement, à commencer avec l'ARNt désaminoacylé pour terminer avec l'ARNt-aa, les ARNt vont se désappairer de l'ARNm pour glisser vers leur site classique et se réappairer dans le cadre -1. (tirés de Leger et al., 2007)

I

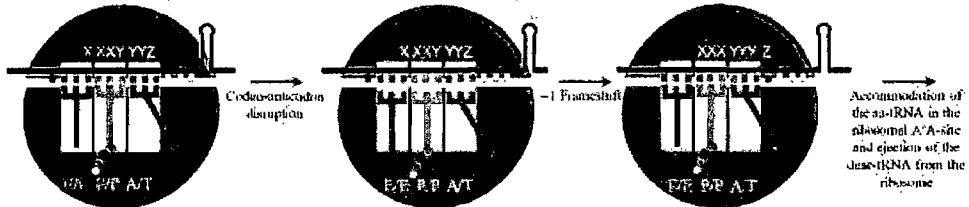
A The two-tRNA simultaneous slippage model proposed by Jacks et al. (19) and refined by Plant et al. (23)



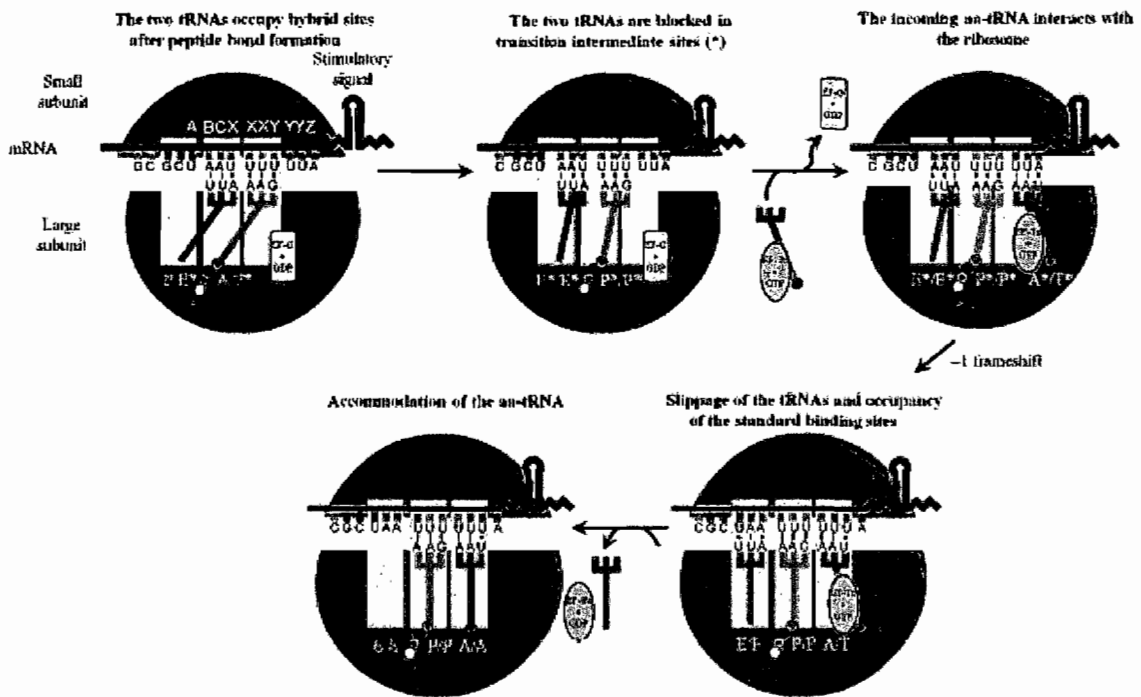
B Weiss et al. (26) proposed that the frameshift occurs during translocation



C Farabaugh et al. (38) and Léger et al. (35) proposed that the frameshift occurs when the aminoacyl-tRNA occupies the A/T entry site



II



Le groupe de Dinman a par la suite voulu raffiner ce modèle en suggérant qu'un mouvement de 9 Å de la boucle anticodon de l'ARNt-aa suite à son accommodation au site A créait une tension dans l'ARNm. Cette tension, due à la résistance au déroulement du signal stimulateur, inciterait les deux ARNt à se dissocier de l'ARNm, ce qui permettrait à l'ARNm de se déplacer d'un nucléotide en direction 3' afin d'éliminer la tension dans l'ARNm. Les deux ARNt se réassocieraient ensuite avec l'ARNm dans le cadre de lecture -1. Toutefois, ce mouvement de 9 Å de la boucle anticodon n'est pas appuyé par les analyses de structure et de dynamique moléculaire du ribosome.

Un second modèle provient du groupe d'Atkins (Weiss et al., 1989). Celui-ci propose que le *frameshift* survient suite à la translocation des ARNt aux sites P et A vers les sites hybrides P/E et A/P respectivement (décrit dans la section 1.4). Ce déplacement des tiges acceptrices des ARNt se produit après que la chaîne peptidique naissante soit transférée sur l'ARNt-aa. Normalement, les boucles anticodons des deux ARNt se déplaceraient ensuite aux sites E et P pour que l'ARNt désaminoacylé soit au site E et que le nouvel ARNt-peptidyle se retrouve au site P. Par contre, en présence du signal stimulateur, ce déplacement serait perturbé par la résistance au déroulement par le ribosome de cette structure secondaire de l'ARNm. Le modèle propose que, lorsqu'elles sont aux sites P/E et A/P, les boucles anticodons des ARNt peuvent se dissocier de l'ARNm et se déplacer avec le ribosome pour se réappairer avec l'ARNm dans le cadre de lecture -1. Il y aurait ensuite déplacement des ARNt vers les états E et P dans le nouveau cadre de lecture. Des analyses par Cryo-EM de complexes ribosome-ARNm-signal stimulateur bloqués au niveau de la région de *frameshift* appuient l'hypothèse que la translocation des ARNt est impliquée dans le *frameshift* (Namy et al., 2006). Toutefois, des études de mutagenèse chez les bactéries sur l'ARNr 16S (Leger et al., 2004) et chez la levure sur le facteur d'élongation eEF1A altérant l'accommodation des ARNt vers l'état A (Harger et al., 2002) démontrent que l'accommodation peut affecter l'efficacité de *frameshift*, ce qui n'est pas expliqué par le modèle de Weiss *et al.* ni par celui de Jacks *et al.* Ces études de mutagenèse ont conduit à un troisième modèle initialement formulé par Farabaugh en 1997, puis appuyé par des études de notre groupe en 2004 (Leger et al., 2004). Ce modèle suggère que le

frameshift a lieu lorsque l'ARNt peptidyle est au site P et l'ARNt-aa au site A/T, avant l'accommodation au site A (rappelons que la première lettre est le site de liaison avec la petite sous-unité du ribosome et la seconde lettre le site de liaison avec la grande sous-unité).

Enfin, un dernier modèle s'apparentant au troisième a été développé par notre groupe en 2007 (figure 9II) (Leger et al., 2007) qui intègre l'ensemble des observations portant sur le *frameshift*. Ce modèle montre comment des mutations en 5' de la séquence glissante (au site E) influencent le *frameshift*. La séquence glissante du VIH-1 rallongée d'un codon en 5' pour tenir compte du site E est la suivante : U AAU UUU UUA (les espaces déterminent le cadre de lecture 0 de Gag). Le *frameshift* est amorcé lorsque le codon AAU de la séquence glissante allongée occupe le site P et que le codon UUU occupe le site A. Suite à la formation du lien peptidique, les ARNt et l'ARNm procèdent à la translocation. Lors de la translocation des ARNt pendant l'élongation, les ARNt passent par un état de transition lorsqu'ils transitent d'un site du ribosome au suivant. D'abord, les tiges acceptrices se déplacent vers le site suivant, de sorte que les ARNt préalablement aux sites P et A se retrouvent dans des états intermédiaires P/E* et A/P* respectivement (l'astérisque représente l'état de transition). Ensuite, les boucles anticodons des ARNt et l'ARNm tentent de poursuivre la translocation. Toutefois, dû à la résistance au déroulement du signal stimulateur à l'entrée du tunnel de l'ARNm du ribosome, l'ARNm ne se déplace que de deux bases. L'ARNt désaminoacylé et l'ARNt peptidyle se retrouvent alors coincés dans des états de transitions E*/E* et P*/P* respectivement. Un nouvel ARNt-aa est alors sélectionné par le ribosome au site A et se retrouve lui aussi dans un état de transition A*/T*. Puisque les interactions codon-anticodon sont dynamiques et peuvent se briser et se reformer, la tendance des ARNt à occuper un site classique plutôt qu'un site de transition sur le ribosome favorise le glissement dans le cadre -1 des ARNt. Ce processus se fait de manière successive, à commencer par l'ARNt désaminoacylé au site E, puis l'ARNt-peptidyl au site P pour terminer avec l'ARNt-aa au site A/T. Une fois le *frameshift* -1 effectué, la traduction conventionnelle serait poursuivie dans ce cadre de lecture. Il faut mentionner ici que l'interaction codon-anticodon n'a jamais été démontrée au site E. Si elle

devait exister, il est supposé que le site E serait beaucoup plus tolérant que les sites P et A en ce qui a trait aux interactions codon-anticodon non-conventionnelles, ce qui rendrait possible le glissement de l'anticodon UUA vers le codon UAA dans le cadre -1 malgré les interactions U-U et A-A défavorables. Tel que mentionné précédemment, changer le codon de l'ARNm au site E peut affecter l'efficacité de *frameshift*. On sait aussi que cette remarque s'applique aux codons aux sites P et A. Le dernier modèle explique ces observations par des variations dans la tendance des ARNt à glisser d'un état de transition à un état classique, dépendamment de leur identité (Leger et al., 2007).

1.4.2.3 L'importance de l'efficacité de *frameshift* chez le VIH-1

L'efficacité du *frameshift* ribosomique -1 programmé contrôle la production du précurseur protéique Gag-Pol chez les cellules infectées par le VIH. La traduction conventionnelle du même messenger conduit à la synthèse du précurseur protéique Gag. Il a été montré que le rapport Gag-Pol : Gag chez les cellules infectées était critique pour la formation de particules virales infectieuses. En effet, une augmentation de vingt fois du rapport Gag-Pol : Gag en surexprimant Gag-Pol à partir d'un vecteur transfecté réduit l'infectivité des virions d'un facteur 1000 et une augmentation de ce rapport de trois fois diminue l'infectivité des virions de dix fois (Shehu-Xhilaga et al., 2001). Cet effet sur la réplication virale serait attribué à la formation de virions anormaux dans le cytoplasme ainsi qu'à une diminution de la stabilité des dimères d'ARNg (Shehu-Xhilaga et al., 2002a; Shehu-Xhilaga et al., 2002b). Des changements plus subtils dans le rapport Gag : Gag-Pol se sont aussi révélés délétères pour la réplication virale. Une étude réalisée par notre groupe en 2006 rapporte par contre que des mutations dans le signal stimulateur entraînant une diminution de l'efficacité de *frameshift* de seulement 30% handicape sévèrement la réplication virale (Dulude et al., 2006). La sensibilité de la réplication virale aux variations dans l'efficacité de *frameshift* du VIH-1 fait donc du *frameshift* ribosomique -1 programmé une cible intéressante dans la lutte contre ce virus.

1.4.2.4 Le *frameshift* ribosomique -1 programmé du VIH-1 chez les bactéries

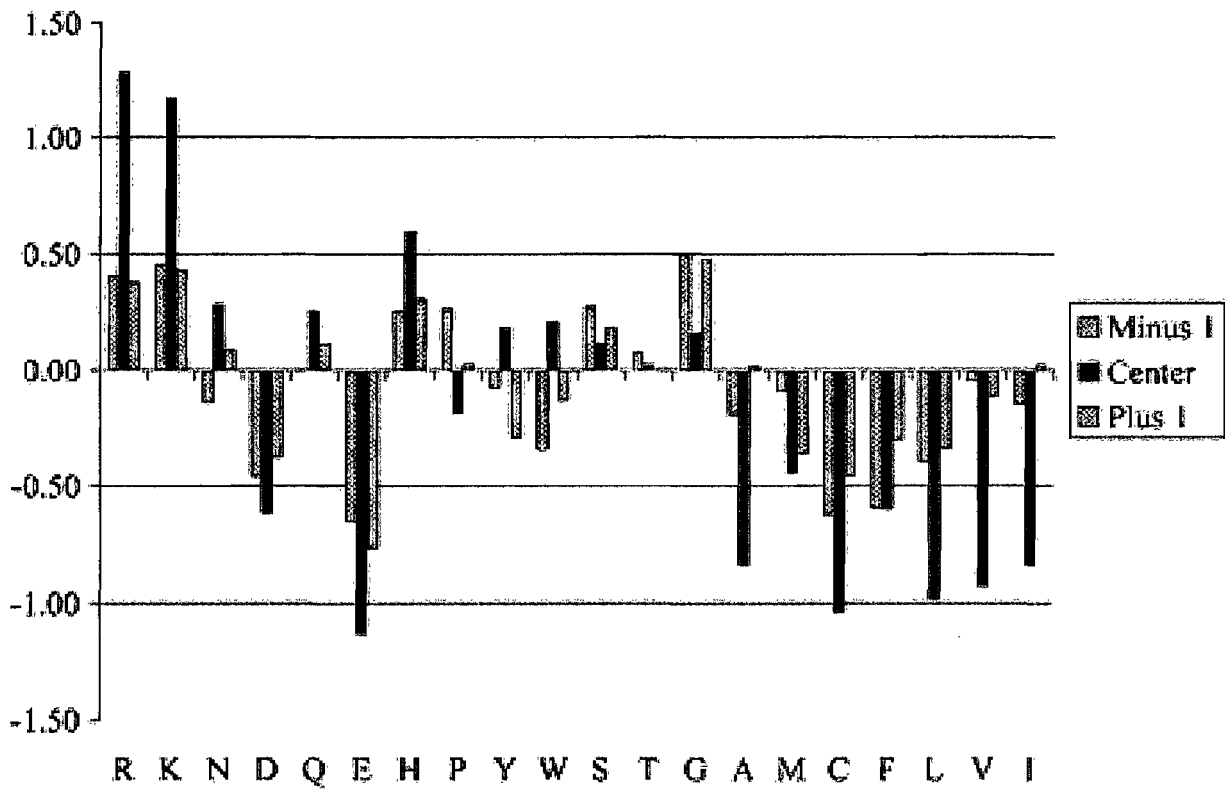
Une étude de notre groupe a permis de démontrer que le *frameshift* programmé du VIH-1 peut être reproduit dans un système bactérien (Leger et al., 2004). Cette étude a aussi montré que le ribosome bactérien répondait exactement de la même façon que le ribosome eucaryote aux mutations dans le signal stimulateur. En effet, les mutations étudiées précédemment en cellules eucaryotes (Dulude et al., 2002) avaient le même effet sur le *frameshift* lorsqu'elles étaient testées chez les bactéries (Leger et al., 2004). Sur la base de ces observations, nous avons décidé d'utiliser un système bactérien dans le cadre du projet ici présenté. Ce choix nous permet de profiter d'un système facile à utiliser et la rapidité de reproduction des bactéries facilite et accélère l'obtention de résultats.

1.5 Les interactions ARN-protéine

1.5.1 La tendance à être présent à l'interface des acides aminés

Les interactions ARN-protéine ont une importance cruciale dans un bon nombre de processus biologiques, tels que la synthèse protéique, la régulation de l'expression des gènes et l'assemblage de virions. Si les sites protéiques de liaison à l'ARN sont tout aussi variés que les séquences à lier, plusieurs biais dans les acides aminés présents à l'interface des interactions ARN-protéine existent (Terribilini et al., 2006). Le plus connu de ces biais concerne sans aucun doute les motifs riches en arginine abondants dans les sites de liaison de protéines à l'ARN (Weiss & Narayana, 1998). En effet, la chaîne latérale des arginines avec son groupement guanidium rend cet acide aminé unique, lui conférant la possibilité d'effectuer une multitude d'interactions ioniques, de ponts hydrogènes et d'interactions de van der Waals. La lysine, un autre acide aminé basique, possède des propriétés semblables et est aussi souvent impliquée dans les interactions ARN-protéine. D'autres acides aminés chargés ou polaires tels que l'histidine et la sérine se retrouvent fréquemment dans de telles interactions. D'ailleurs, une étude statistique a été effectuée sur un échantillonnage de plus de cent protéines liant l'ARN afin de déterminer la tendance de chacun des vingt acides aminés à être impliqué dans une interface ARN-protéine (Terribilini et al., 2006). Cette analyse a aussi déterminé la tendance de chaque acide aminé à se trouver à proximité immédiate d'un acide aminé effectuant une interaction d'interface avec l'ARN (figure 10). Il se dégage de cette étude que des résidus comme l'arginine, la lysine et l'histidine sont très présents dans les interfaces ARN-protéine alors que d'autres tels que la sérine et la glycine ont une haute tendance à être des résidus flanquants. La plupart des résidus non-polaires comme la valine et la phénylalanine ont une très faible tendance à être présents à l'interface, mais leur tendance à être un résidu flanquant est variable. Si les acides aminés chargés et

Figure 10. La tendance des acides aminés à être présents à l'interface d'une interaction ARN-protéine. La barre centrale foncée représente la tendance d'un acide aminé à établir un contact direct à l'interface lors d'une interaction ARN-protéine. Les barres latérales représentent la tendance d'un acide aminé à flanquer un acide aminé effectuant un contact d'interface lors d'une interaction ARN-protéine. Les barres de gauche et de droite correspondent respectivement à la tendance de l'acide aminé à se situer à une position du côté N-terminal ou à une position du côté C-terminal. Les acides aminés sont placés par ordre croissant d'hydrophobicité. La valeur sur l'axe des Y représente la tendance de présence à la position donnée en \log_2 . (tiré de Terribilini et al., 2006)



polaires ont une importance particulière dans le contact entre un site de liaison à l'ARN et sa cible, les acides aminés non polaires jouent aussi des rôles importants, spécialement dans la reconnaissance spatiale du site à lier par une protéine. Par exemple, il existe plusieurs motifs connus où des glycines (interaction Tat-TAR) ou des tryptophanes (interaction Rev-RRE) sont impliqués dans la reconnaissance tridimensionnelle du site à lier (Weiss & Narayana, 1998). Le tableau 1 présente les principaux résidus ou groupements protéiques interagissant avec les différentes cibles de l'ARN. On note encore une fois la prédominance de l'arginine dans les interactions ARN-protéine et que le squelette protéique, de par ses groupements carbonyles et amides, peut aussi participer à un grand nombre d'interactions.

1.5.2 La structure de peptides interagissant avec l'ARN

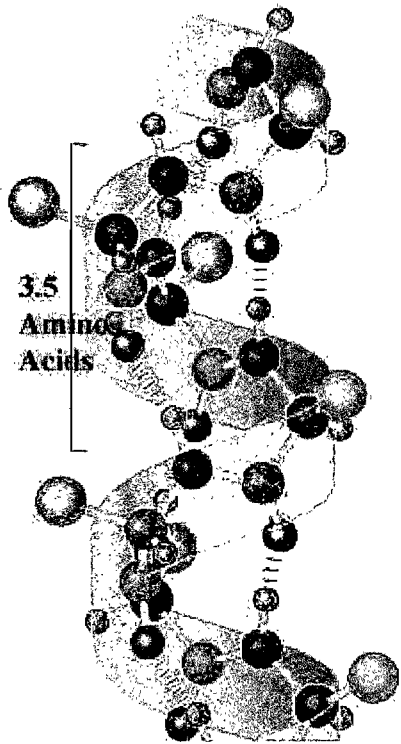
Il existe deux modèles principaux de peptides liant l'ARN (figure 11) (revu par Draper, 1999). Le premier modèle est celui de peptides ou de protéines liant un sillon de l'ARN, le plus souvent une hélice α protéique interagissant avec le sillon majeur d'une hélice A d'un ARN. La forme A d'hélices d'ARN possède un sillon majeur et un sillon mineur, tout comme les hélices d'ADN de forme B. Par contre, le sillon majeur d'hélices de forme A d'ARN n'est pas accessible aux structures secondaires protéiques contrairement aux hélices d'ADN. Deux facteurs viennent nuancer l'impossibilité pour les structures secondaires de protéines d'interagir avec le sillon majeur d'hélices d'ARN. Premièrement, il est rare de retrouver un tour complet d'hélice, ce qui fait que les sillons majeurs sont accessibles via les extrémités des portions hélicoïdales. Deuxièmement, les nombreux mésappariements et excroissances présents sur les structures secondaires d'ARN causent des distorsions qui entraînent des élargissements suffisants au niveau du sillon majeur pour permettre à des protéines d'y interagir. Il n'existe que peu de cas connus où une protéine lie l'ARN en interagissant avec son sillon mineur. Parmi ceux-ci, la protéine ribosomique L11, dont le site de liaison à l'ARN est composé de trois hélices α , qui interagit avec une portion distordue du sillon mineur de l'ARNr de sorte que cette région présente une surface relativement plane. Sinon, les structures protéiques, le plus souvent des hélices α , se lient au sillon majeur de

Tableau 1. Les principales interactions moléculaires dans les interactions ARN-protéine. Les acides aminés ainsi que le squelette d'un peptide peuvent interagir avec le squelette ribose-phosphate ainsi qu'avec les bases azotées de l'ARN. Ce tableau présente les principaux acides aminés et groupements protéiques impliqués dans des interactions avec les différentes cibles de l'ARN. Un « * » réfère au squelette peptidique alors que lorsque les acides aminés sont nommés, c'est leur chaîne latérale qui effectue l'interaction avec la cible.

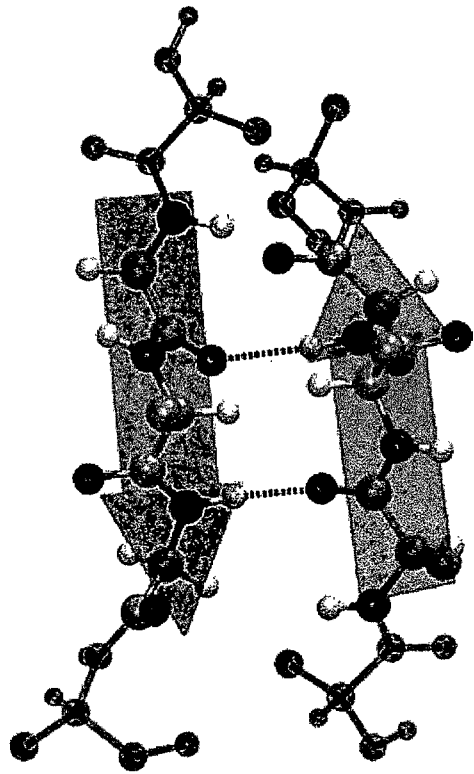
Tableau 1. Les principales interactions moléculaires dans les interactions ARN-protéine.

Type d'interaction	Cible	Résidus et groupements protéiques pouvant effectuer une interaction avec la cible
Ionique / Polaire	Avec le squelette de l'ARN (sucre)	Arginine, lysine, groupements amide* et carbonyle*
	Avec le squelette de l'ARN (phosphate)	Arginine, lysine, groupement amide*, sérine, thréonine, tyrosine
	Adénine	Arginine, groupement carbonyle*
	Guanine	Arginine, groupement amide*, lysine, acide aspartique, acide glutamique
	Cytosine	Arginine, groupement amide* et carbonyle*
	Uracile	Arginine, groupement amide et carbonyle, acide aspartique, acide glutamique, asparagine, glutamine
Empilement de cycles aromatiques	Adénine, guanine, cytosine, uracile	Phénylalanine (préférence pour l'adénine), tyrosine, tryptophane, histidine, arginine

Figure 11. Structures adoptées par les protéines interagissant avec l'ARN. À gauche, un modèle d'une structure protéique en hélice α . Cette structure est rencontrée chez les protéines interagissant avec le sillon majeur d'une hélice A d'ARN. À droite, le modèle d'une structure en feuillet β . Cette conformation est présente chez les protéines interagissant avec des résidus non-appariés ou faiblement structurés de l'ARN (tirés de <http://creationwiki.org/>).



Hélice α



Feuillet β

l'ARN. La protéine virale Rev du VIH-1 et la protéine d'antiterminaison λ N du phage λ font partie de cette catégorie. Le second modèle est celui de feuilletts β reconnaissant spécifiquement certaines bases non appariées de l'ARN. Ces bases non-appariées se retrouvent fréquemment dans les structures d'ARN dû à la présence d'excroissances et de mésappariements. De plus, dans le cas d'ARN simple brin, certaines bases peuvent être faiblement empilées (*base stacking*) et aussi être disponibles pour interagir avec les protéines liant l'ARN possédant une structure en feuilletts β . Ces feuilletts β ont des encavements appelés « pochettes de liaison » où les bases libres sont reconnues. Ce type de liaison à l'ARN ignore le squelette ribose-phosphate de la molécule d'ARN. Un motif protéique connu communément appelé domaine RNP (*ribonucleoprotein*) ou domaine RRM (*RNA recognition motif*) correspond à ce type de modèle de liaison à l'ARN. Par exemple, la protéine humaine U1A, qui est impliquée dans l'épissage, possède un tel motif.

1.5.3 L'utilisation de bibliothèques combinatoires de peptides actifs *in vivo*

Les bibliothèques combinatoires de peptides constituent une avenue prometteuse pour la découverte de molécules pouvant lier des cibles d'ARN. Ces molécules peuvent être utilisées pour compétitionner avec un ligand naturel ou pour perturber un mécanisme *in vivo*. Les peptides sont généralement des molécules thérapeutiques moins intéressantes vu leur faible stabilité *in vivo*, leur faible biodisponibilité lorsqu'ils sont consommés oralement et leur potentielle immunogénicité (Patch & Barron, 2002). Toutefois, des molécules actives peptidomimétiques peuvent être synthétisées en modifiant la nature du peptide, soit en changeant son squelette ou en introduisant des groupements différents. Ces molécules modifiées sont moins vulnérables aux protéases, peuvent être administrées oralement avec une bonne biodisponibilité et sont moins immunogènes. Par exemple, le groupe de Rana a développé des homologues de la protéine Tat du VIH-1 aptes à lier la structure d'ARN TAR des ARNm viraux en utilisant un squelette d'oligocarbamate ou d'oligourée

(Tamilarasu et al., 2001). Un autre problème rencontré dans l'administration de peptides comme agents thérapeutiques est leur incapacité à pénétrer dans les cellules. Pour remédier à cette caractéristique, plusieurs groupements peuvent être greffés à une extrémité des peptides pour médier leur entrée dans les cellules. Parmi ceux-ci, notons le domaine de transduction protéique de Tat (YGRKKRRQRRR) (Schwarze & Dowdy, 2000) et un dérivé synthétique de ce dernier (YARAAARQARA) (Ho et al., 2001).

Les bibliothèques combinatoires sont habituellement exprimées à l'intérieur d'un échafaud (*scaffold*) protéique visant à prévenir la dégradation des peptides générés et à leur imposer une conformation rigide. La nucléase inactivée du staphylocoque, la thiorédoxine et la GFP sont des exemples de protéines échafauds permettant l'expression de bibliothèques de peptides *in vivo* (Norman et al., 1999; Skerra, 2000). Par exemple, la thiorédoxine a déjà fait ses preuves en tant qu'échafaud dans la sélection de peptides inhibiteurs de la thymidylate synthase (Blum et al., 2000). Certaines bibliothèques sont par contre exprimées à une extrémité d'une protéine, de sorte que le peptide se retrouve sous forme flottante, sans contrainte structurale. La section suivante présente deux exemples de projets où des bibliothèques combinatoires de peptides furent utilisées pour sélectionner des peptides actifs *in vivo*. Notons que, dans ces deux études, les peptides utilisés sont produits *in vivo* via transfection ou transformation d'un plasmide et non pas introduits sous forme de peptides synthétiques.

1.5.3.1 Sélection de peptides conférant une résistance au Taxol

En 2001, le groupe de Nolan publiait un article décrivant le crible d'une bibliothèque de peptides sélectionnés selon leur aptitude à induire la survie de cellules eucaryotes en présence d'un agent entraînant la mort cellulaire : le taxol (Xu et al., 2001). Souvent utilisé dans les traitements de chimiothérapie, le taxol empêche le mouvement des microtubules, causant l'arrêt mitotique et la mort cellulaire (Orr et al., 2003). L'objectif de la sélection de peptides inhibant l'effet chimiothérapeutique du taxol était d'identifier des voies de résistance pouvant être empruntées par les cellules cancéreuses. Les peptides criblés provenaient d'une bibliothèque codant pour des peptides de 18 acides aminés de long où les

vingt acides aminés sont représentés. Un échafaud protéique artificiel fut utilisé. Celui-ci est composé de deux bras aptes à dimériser qui flanquent les peptides, ces derniers étant alors exprimés dans une boucle. Des cellules HeLa ont été infectées avec des rétrovirus recombinant contenant la librairie de peptides et les peptides actifs ont été sélectionnés en soumettant les cellules à une dose de 200 nM de taxol pendant 48 h. Suite à l'élimination de faux positifs, les séquences des peptides actifs ont été déterminées et des tests de dose-réponse au taxol effectués pour chaque peptide actif afin de déterminer le plus efficace. Le RGP8.5 s'avéra être le clone le plus actif.

L'investigation de la voie par laquelle le RGP8.5 détient son effet anti-taxol a permis de découvrir que le peptide augmentait le niveau cellulaire de la protéine ABCB1. La surexpression d'ABCB1 est souvent associée avec la résistance à de multiples drogues (MDR pour *multiple drug resistance*). ABCB1 est une protéine transmembranaire qui effectue le transport actif de certains métabolites et protège les cellules de substances cytotoxiques. Une étude approfondie du mécanisme d'action de RGP8.5 a révélé que le peptide augmente le niveau d'ARNm codant pour ABCB1 via l'inhibition du protéasome cellulaire. Le lien direct entre l'inhibition du protéasome et la protection contre la dégradation des ARNm codant pour ABCB1 reste à déterminer, mais il pourrait s'agir de l'inhibition de l'activité ribonucléase de la sous-unité PSMA5 du protéasome. Même si, dans ce cas-ci, la cible des peptides n'est pas l'ARN, la grande diversité de la librairie utilisée ainsi que la possibilité de cribler des peptides actifs sans cible déterminée démontrent le potentiel des bibliothèques combinatoires dans l'identification de molécules actives contre des processus complexes, comme le développement de résistances aux médicaments.

1.5.3.2 Sélection de peptides liant la structure RRE du VIH-1

Dans la course au développement d'agent anti-rétroviraux, la structure RRE présente dans certains ARNm du VIH-1 a attiré l'attention de plusieurs groupes. Le groupe de Frankel, en 1996, a ciblé le RRE avec deux bibliothèques combinatoires de peptides en utilisant un système d'étude bactérien (Harada et al., 1996). La première librairie utilisée

codait pour des peptides de 14 acides aminés de long où seuls quatre acides aminés sont représentés : l'arginine, la sérine, l'histidine et l'asparagine (RSHN). Les pourcentages théoriques de représentation étaient respectivement de 56,25%, 18,75%, 18,75% et 6,25%. Ces acides aminés ont une haute tendance à être présents à l'interface et/ou à être des résidus flanquants dans des interactions ARN-protéines. La seconde librairie codait pour des peptides de 19 acides aminés de long. De ces 19 acides aminés, les neuf acides aminés du centre étaient aléatoirement composés d'arginine, de sérine et de glycine (RSG), avec des pourcentages théoriques de représentation de 33,3% chacun, et étaient flanqués de cinq arginines. La glycine possède une très faible tendance de présence à l'interface dans les interactions ARN-protéine, mais une forte tendance à être un résidu flanquant. Dans les deux librairies, l'abondance d'arginines accroît la probabilité de retrouver des motifs riches en arginine qui augmentent l'affinité des peptides générés pour l'ARN. Les librairies de peptides furent fusionnées à l'extrémité N-terminale de la protéine d'antitermination λ N du phage λ dans un premier plasmide. Dans un second plasmide, RRE fut insérée au début de la séquence codante du gène rapporteur β -galactosidase (LacZ). Des sites terminateurs entre la région codant pour RRE et celle codant pour LacZ faisaient en sorte que le rapporteur ne pouvait être transcrit que si λ N était recruté à la portion RRE des ARNm naissants afin d'assurer l'antitermination de la transcription. Un criblage colorimétrique via LacZ de 600 000 colonies possédant la librairie RSHN et de 200 000 colonies possédant la librairie RSG fut effectué. Suite à l'élimination de faux positifs, une cinquantaine de clones furent séquencés et quatre séquences peptidiques uniques provenant de la librairie RSHN et trois provenant de la librairie RSG furent obtenues. Deux des peptides provenant de la librairie RSHN possédaient une séquence semblable à celle de Rev et liaient RRE avec une affinité similaire, alors que les cinq autres séquences liaient RRE avec une affinité moindre. Des mesures de dichroïsme circulaire ont montré que les peptides provenant de la librairie RSHN avaient une portion hélicoïdale ($\approx 10\%$ de la structure) alors que les peptides provenant de la librairie RSG étaient plutôt non-structurés.

Une autre étude du groupe de Frankel parue en 2003 fait état d'un nouveau crible contre RRE (Peled-Zehavi et al., 2003). Cette fois-ci, la librairie utilisée codait pour des

peptides de quinze acides aminés où les dix premiers résidus N-terminaux étaient constitués de douze acide aminés polaires ou à courte chaîne latérale, excluant les acides aminés avec une longue chaîne latérale hydrophobe. Les cinq acides aminés en C-terminal des peptides étaient des arginines et l'utilisation d'oligonucléotides dégénérés permettait d'insérer un biais pour l'arginine aux dix premières positions de sorte que l'arginine y ait une représentation théorique de 58,33%. La librairie était encore fusionnée à l'extrémité N-terminale de λ N. Le système de sélection utilisé impliquait toujours un procédé d'antiterminaison, mais cette fois-ci deux gènes rapporteurs ont été testés : LacZ et le gène de résistance à la kanamycine (néomycine phosphotransférase ou KanR). Des rondes successives à concentrations croissantes de kanamycine avec le rapporteur KanR ainsi qu'un crible colorimétrique avec LacZ ont permis de sélectionner plusieurs peptides. Des tests comparatifs d'antiterminaison où la protéine virale Rev était fusionnée en N-terminal de λ N ont révélé que certains peptides liaient RRE avec une affinité plus grande que Rev. De plus, la spécificité des peptides pour RRE fut démontrée en remplaçant la portion RRE du plasmide rapporteur par la tige-boucle TAR du virus d'immunodéficience bovin (BIV).

L'obtention de peptides liant spécifiquement RRE avec une affinité égale, voire supérieure, à Rev confirme le potentiel de bibliothèques de peptides générant des motifs riches en arginine dans la sélection de peptides ciblant des structures d'ARN.

1.6 Le ribosome est-il une cible potentielle?

Tel qu'expliqué en détails dans la section 1.7, le but de notre projet est de perturber l'efficacité de *frameshift* du VIH-1 en utilisant une librairie de peptides. Si le signal stimulateur est jugé comme une cible de choix dans la sélection de peptides perturbant le *frameshift* ribosomique -1 programmé, le ribosome est une autre alternative incontournable. Il ne faut pas oublier que le ribosome est un ribozyme et que des peptides liant l'ARNr pourraient influencer l'efficacité de *frameshift*. Bon nombre de molécules ciblant le ribosome et la synthèse protéique sont connues telles que l'anisomycine, la sparsomycine, la puromycine et la cycloheximide (Grollman, 1967; Southworth & Green, 2003). Certaines de ces molécules, comme la sparsomycine, ont démontré du potentiel en tant qu'agent antitumoral. De plus, deux autres inhibiteurs de la synthèse protéique, le bouvardin et l'aplidine, se sont avérés efficaces chez des patients atteints de la leucémie ou de cancers endocriniens (Faivre et al., 2005).

Certains inhibiteurs ciblant le ribosome ont aussi démontré de l'effet sur le *frameshift*. En effet, il a été rapportée que, chez la levure, l'anisomycine et la sparsomycine entraînaient respectivement une diminution et une augmentation du *frameshift* ribosomique -1 programmé du virus L-A de la levure, sans avoir d'effet sur les *frameshifts* +1 (Dinman et al., 1997). Chez *E. coli*, le chloramphénicol, un inhibiteur de la formation du lien peptidique, augmente le *frameshift* -1 du VIH-1 (Brunelle et al., 1999). De telles molécules qui agissent aussi sur la traduction conventionnelle possèdent une fenêtre thérapeutique restreinte et ne peuvent être considérés comme des candidats potentiels en vue d'un traitement contre le VIH-1 ciblant le *frameshift*. Pour être thérapeutiquement attrayants, des peptides actifs sélectionnés pour un effet sur le *frameshift* en liant le ribosome devraient remplir la condition suivante : posséder une fenêtre thérapeutique raisonnable où la concentration requise pour la perturbation de l'efficacité de *frameshift* est significativement plus faible que la concentration affectant la traduction conventionnelle jugée toxique ou, mieux encore, ne pas affecter la traduction conventionnelle.

1.7 Le but de ce projet

Tel que mentionné précédemment, il est impératif de développer de nouveaux agents anti-rétroviraux contre de nouvelles cibles afin de contrer l'émergence de variants du VIH-1 résistants aux agents couramment utilisés. Le VIH-1 est particulièrement sensible aux variations dans l'efficacité de *frameshift* ribosomique -1 programmé, comme en témoignent les résultats présentés dans la section 1.4.2.3, qui montrent qu'une réduction de l'efficacité de *frameshift* de 30% est suffisante pour handicaper sévèrement l'infectivité virale. L'événement du *frameshift* chez le VIH-1 est une cible d'intérêt dans la lutte contre ce virus.

Notre objectif était de sélectionner des peptides aptes à perturber l'efficacité de *frameshift* ribosomique -1 programmé du VIH-1 en interagissant avec le signal stimulateur de *frameshift* ou en liant des régions du ribosome impliquées dans le *frameshift*. Pour ce faire, nous avons décidé d'employer la librairie RSHN utilisée avec succès contre RRE par le groupe de Frankel (Harada et al., 1996). Cette librairie génère des peptides de 14 acides aminés, où le biais pour l'arginine qu'on y retrouve favorise la présence de motifs riches en arginine qui sont connus pour lier efficacement l'ARN. Comme il a été montré que le *frameshift* -1 du VIH peut être reproduit chez *E. coli* (Leger et al., 2004), nous avons opté pour la rapidité et la facilité d'utilisation d'un système bactérien pour effectuer le criblage de la librairie. Nous avons décidé d'utiliser la cytométrie de flux comme outil de criblage, en nous servant d'un système rapporteur fluorescent. Cette méthode a comme avantage de pouvoir cribler rapidement un très grand nombre de clones. Dans ces clones, l'expression de la protéine fluorescente verte (GFP) dépend du *frameshift* par insertion de la région de *frameshift* du sous-type B du groupe M du VIH-1 au début de la région codante de la GFP. L'expression de la protéine fluorescente rouge (RFP) dépend de la traduction conventionnelle et sert à normaliser l'expression de la GFP. Nous avons examiné l'effet des peptides sélectionnés sur la traduction conventionnelle ainsi que leur activité sur des régions de *frameshift* où le signal stimulateur du sous-type B était substitué par un autre signal stimulateur. De plus, nous avons procédé à l'analyse des séquences des peptides

actifs. Finalement, nous avons examiné l'effet des peptides sélectionnés sur le *frameshift* du VIH-1 dans les cellules eucaryotes HEK 293T. L'identification de peptides qui perturbent l'efficacité de *frameshift* -1 du VIH-1 pourra servir de base pour la conception de nouveaux agents anti-rétroviraux ciblant le *frameshift*.

Chapitre II :

Selection of peptides interfering with a ribosomal frameshift in the human immunodeficiency virus type 1

Selection of peptides interfering with a ribosomal frameshift in the human immunodeficiency virus type 1

Dominic Dulude,^{1,2†} Gabriel Théberge-Julien,^{1†} Léa Brakier-Gingras,^{1*} and Nikolaus Heveker,^{1,2*}

¹Département de Biochimie, Université de Montréal, Montréal H3T 1J4, Québec, Canada

²Centre de Recherche/Hôpital Sainte-Justine, Montréal H3T 1C5, Québec, Canada

Article sous presse dans RNA, 2008.

Dominique Dulude et Gabriel Théberge-Julien ont contribué également à toutes les expériences décrites et ont participé à la rédaction du manuscrit.

Léa Brakier-Gingras et Nikolaus Heveker ont supervisé le travail expérimental et la rédaction du manuscrit.

† equal contribution

* joint senior authorship

Correspondence to:

Nikolaus Heveker, Centre de Recherche, #6737 Hôpital Sainte-Justine, 3175 Chemin de la Côte Sainte-Catherine, Montréal, Québec, H3T 1C5, Canada, Tel: (514) 345-4931 ext. 4190, Fax : (514) 345-4801, email :
[REDACTED]

And

Léa Brakier-Gingras, Département de Biochimie, #D-353, Université de Montréal, 2900, boul. Édouard-Montpetit, Montréal, Québec, H3T 1J4, Canada, Tél. : (514) 343-6316, Fax : (514) 343-2210, email :
[REDACTED]

Abstract

The human immunodeficiency virus of type 1 (HIV-1) uses a programmed -1 ribosomal frameshift to produce the precursor of its enzymes, and changes in frameshift efficiency reduce replicative fitness of the virus. We used a fluorescent two-reporter system to screen for peptides that reduce HIV-1 frameshift in bacteria, knowing that the frameshift can be reproduced in *E. coli*. Expression of one reporter, the green fluorescent protein GFP, requires the HIV-1 frameshift whereas the second reporter, the red fluorescent protein RFP, is used to assess normal translation. A peptide library biased for RNA binding was inserted into the sequence of the protein thioredoxin and expressed in reporter-containing bacteria, which were then screened by fluorescence-activated cell sorting (FACS). We identified peptide sequences that reduce frameshift efficiency by over 50% without altering normal translation. The identified sequences are also active against different frameshift stimulatory signals, suggesting that they bind a target important for frameshifting in general, probably the ribosome. Successful transfer of active sequences to a different scaffold in a eukaryotic test system demonstrates that the anti-frameshift activity of the peptides is neither due to scaffold-dependent conformation nor effects of the scaffold protein itself on frameshifting. The method we describe identifies peptides that will provide useful tools to further study the mechanism of frameshift and may permit the development of lead compounds of therapeutic interest.

Introduction

The human immunodeficiency virus type 1 (HIV-1), the causative agent of AIDS, uses a programmed -1 ribosomal frameshift to express its enzymes. The HIV-1 full length viral messenger encodes the structural and enzymatic components from the overlapping *gag* and *pol* open reading frames, respectively. Gag is synthesized by translation of the viral mRNA according to conventional rules by the ribosomes in the infected cells, whereas Pol is produced as a Gag-Pol precursor *via* a programmed -1 ribosomal frameshift that occurs near the 3' end of Gag (reviewed in Brierley & Pennell, 2001; Brierley & Dos Ramos, 2006). The Gag-Pol to Gag ratio is critical for virus replication and infectivity, and changes in this ratio severely decrease viral infectivity and replication (Park & Morrow, 1991; Karacostas et al., 1993; Shehu-Xhilaga et al., 2001; Telenti et al., 2002; Dulude et al., 2006). All these observations stress the potential of targeting the frameshift event for the development of novel antiretroviral agents.

The HIV-1 frameshift event requires two *cis*-acting elements in the viral mRNA: a heptameric slippery sequence, U UUU UUA, where the shift of frame occurs, and a stimulatory signal located immediately downstream, which is an irregular two-stem helix structure (Dulude et al., 2002; Gaudin et al., 2005; Staple & Butcher, 2005) that controls the frameshift efficiency. It is assumed that the ribosome makes a first encounter with the complete frameshift stimulatory signal before its decoding center occupies the slippery site and that the presence of the lower stem favors an interaction between the ribosome and the upper stem-loop of the signal. The ribosome then unfolds the lower stem and progresses along the mRNA, but pauses when its decoding center occupies the slippery site. The upper stem-loop is then at the entrance of the mRNA channel and the pause would be caused by resistance to unwinding of this stem-loop (Léger et al., 2004; Staple & Butcher, 2005). The ribosome unwinds the secondary structures it encounters with its helicase activity made of proteins lining the

entrance channel (S3 and S4 in bacteria corresponding to RPS3 and RPS9 in human ribosomes) (Yusupova et al., 2001; Takyar et al., 2005; Namy et al., 2006). The mechanism of frameshifting is still unclear, and various models have been proposed (Jacks et al., 1988; Weiss et al., 1989; Plant et al., 2003; Léger et al., 2004; Namy et al., 2006; reviewed in Léger et al., 2007).

Kollmus et al. (1996), who replaced the HIV upper stem-loop with the iron-responsive element from ferritin mRNA, a stem-loop structure that binds iron regulatory proteins depending upon the iron status of the cells, found that the frameshift efficiency could be increased about two-fold upon binding of the regulatory proteins. This observation demonstrates that HIV frameshifting can be regulated by RNA-binding proteins. In this study, we decided to search for peptides that could interfere with the HIV-1 frameshift by binding to the frameshift stimulatory signal or to regions of the ribosome that control the frameshift. In a previous study, we had shown that the bacterial ribosome responds exactly like the eukaryote ribosome to the frameshift stimulatory signal of HIV-1 (Léger et al., 2004). We therefore used a bacterial system for the selection of peptides, taking advantage of the speed and simplicity associated with studies using bacteria. The group of Frankel (Harada et al., 1996) succeeded in selecting peptides that bind to RRE of HIV-1, using a bacterial transcriptional antitermination reporter to screen a 14-residue peptide library composed of arginine, serine, asparagine and histidine at each position and with a codon bias for arginine. We developed a bacterial fluorescent frameshift reporter system for the screening of such a library with arginine-rich peptides that bind RNA. This fluorescent reporter system allowed us to isolate bacteria containing peptides that alter HIV-1 frameshift efficiency by fluorescence-activated cell sorting (FACS).

Results

Establishment of the GFP/RFP screening reporter

An expression vector was constructed that expresses the green fluorescent protein (GFP) and its red-shifted variant, RFP, in a specialized bacterial ribosome system in which the ribosome binding site for the reporter messengers and the messenger binding site in 16S rRNA encoded by the same plasmid bear complementary mutations (Lee et al., 1996; 1997). As a result, the reporters are exclusively translated by ribosomes containing the plasmid-encoded 16S rRNA. This structure also ensures that a large fraction of ribosomes are devoted to the translation of the reporters. The red and green fluorescence of the reporters are monitored by flow cytometry or fluorescence plate reading. For the frameshift reporter plasmid, p3RFP-HIV(-1)-GFP (Fig. 1A), the HIV-1 programmed -1 frameshift region, encompassing the slippery site and the complete stimulatory signal from HIV-1 group M subtype B, the group and subtype that prevail in North America and Western Europe, was used. This frameshift region was inserted downstream of the AUG start codon of GFP, such that GFP translation depends on a -1 frameshift. The RFP translation serves as a control for conventional protein synthesis. With this system, changes in the frameshift efficiency are reflected by changes in the GFP/RFP ratio. The GFP/RFP fluorescence intensity ratio was stable over a wide range of densities of bacterial cultures and was set to be 100% frameshift efficiency. Pilot experiments in which bacteria transformed with the reporter plasmid were exposed to different doses of chloramphenicol (0.1 to 2 $\mu\text{g/ml}$), a protein synthesis inhibitor known to increase frameshift efficiency (Thompson et al., 2002), yielded a higher GFP/RFP ratio (up to 3 fold increase, data not shown), demonstrating the suitability of the p3RFP-HIV(-1)-GFP reporter for monitoring changes in frameshift efficiency. Also, the fluorescence intensity of GFP under control of HIV-1 frameshift region was compared to a conventionally translated GFP control to directly measure the frameshift efficiency. It was found to be 3.7%, which is within the range of values (2.4 to 4.8%) observed for

HIV-1 frameshift efficiency with a luciferase reporter in bacteria (Léger et al., 2004; Plant & Dinman, 2006).

Screening of an arginine-rich peptide library

Using the RFP/GFP reporter, we set out to identify peptide sequences that specifically affect frameshift efficiency, by interacting either with the frameshift stimulatory signal or with regions of the ribosome involved in frameshifting, without affecting conventional translation. We designed an arginine-rich peptide library containing arginine, serine, asparagine and histidine residues presented by thioredoxin as a protein scaffold (Fig. 1B). Such library has previously been shown to contain specific binders of RNA motifs (Harada et al., 1996; 1997). However, whereas the library of Harada et al.(1996) was fused to the N-terminal portion of protein N from bacteriophage λ , we chose thioredoxin as a carrier protein for random peptides. Thioredoxin is stable and peptides presented in this scaffold are conformationally constrained (LaVallie et al., 1993), which may increase the affinity of a peptide for its target (Corey & Corey, 1996; Lu et al., 1998). Based on the number of bacterial transformants obtained, we expressed a diversity of 10^7 peptides, which represents only a small fraction of the theoretical diversity (3×10^8) of the library, suggesting that there should be no duplication of sequences in this library. The library was expressed in bacteria transformed with the p3RFP-HIV(-1)-GFP reporter plasmid. The fluorescence of the reporters was analyzed by flow cytometry after 16h of peptide induction, a time point which we had established as optimal to reveal the effect of the peptides on reporter expression. Bacteria with a decreased GFP/RFP ratio, which corresponds to reduced frameshift efficiency, were selected by flow cytometry (Fig. 1C) and plated on agar. The sorting of 2.2×10^7 events resulted in the selection of 4.5×10^4 cells, which corresponds to 0.2% of the total. Of these, only approximately 11% (5×10^3) resulted in colonies when transferred to

agar plates (see Table 1 for a summary of the sorting). We suggest that the discrepancy between the number of plated bacterial clones and the number of growing colonies results from the fact that RNA-binding peptides could be generally toxic to cells; for example, targeting the ribosome and thus interfering with normal translation. It is likely that only clones expressing non-toxic peptide sequences form colonies in our selection procedure.

We investigated whether reiterative rounds of sorting/culture would increase effective selection of active peptides by our screening method. A preliminary experiment revealed that, while the proportion of active clones identified in each round of selection increased, their effect was not. Also, sequence analysis showed that a loss of diversity in sequences accompanied this enrichment, converging to one sequence, NRRRRRRSRHRRRR. We therefore decided to investigate clones obtained after only one round of cell sorting and the results presented below correspond to one single round of selection.

Quantification of the effect of selected peptides

We analyzed about 1,150 colonies obtained on agar plates after the flow cytometry selection to further monitor the effect on the frameshift efficiency of these clones. Clones were grown with or without induction of the peptides and the GFP and RFP fluorescence was measured with a plate reader. As shown in Figure 2A, the relative frameshift efficiency was decreased to an average of 76% of the value in the absence of peptides for the group of 1,150 clones compared to an average decrease to 87% for a group of unselected clones. Inspection of these 1,150 clones led to the identification of the best 160 clones (corresponding to 14% of the clones) that decreased the frameshift efficiency by an average of 49 % whereas the best 14% of the unselected clones decreased the frameshift efficiency by only 13 %. The 20 most efficient clones (2% of the clones) decreased the frameshift efficiency by

57 %. In comparison, the best 2% of unselected bacterial clones only decreased frameshifting by 26%. Figure 2B illustrates the effect of the induction of the selected peptides on the GFP and RFP fluorescence for the 1,150 clones analyzed, in comparison to unselected peptides. The data show a shift of fluorescence corresponding to a decrease of GFP expression, consistent with the fluorescence profile of bacteria defined for the selection of these clones by flow cytometry.

The plasmids isolated from the best 160 clones were retransformed into bacteria and re-assessed for the effect of the peptide they code for by induction and fluorescence plate reading. As shown in Figure 3, the average frameshift decrease caused by induction of the peptides was reproducible. Indeed, inspection of the data for the individual clones revealed that the phenotype was maintained for 145 clones out of 160, and in only 15 clones, the effect on frameshift disappeared, suggesting that about 10% of the clones were false-positive.

The selected peptides affect frameshift-dependent translation, but not conventional translation

As it cannot be excluded *a priori* that peptides that modulate frameshift also affect conventional translation, we investigated the effect of the selected clones on frameshift-dependent GFP expression and on conventional RFP translation. We retransformed 10 randomly chosen clones among the 20 most efficient selected clones into bacteria and measured the fluorescence of both RFP and GFP as a function of culture density after peptide induction. As shown in Figure 4A, GFP fluorescence is inhibited when peptide expression is induced, whereas RFP expression is not affected, as compared to bacteria in which the peptide-bearing plasmid originates from unselected clones. This indicates that conventional protein translation is indeed not affected by the peptides selected as frameshift modulators. Moreover, the GFP and RFP fluorescence measured after 16h of peptide expression for the 20 most efficient clones compared to a group of non-selected clones is also shown in Figure 4B.

The same level of RFP fluorescence is observed for the selected and the non-selected clones whereas the level of GFP fluorescence is decreased for the selected clones when the peptides are induced. These results confirm that the selected peptides do not affect conventional translation and support our suggestion that toxic peptides, among which those affecting normal translation, were eliminated in the process of selection.

Analysis of the sequences of peptides that affect HIV-1 frameshift efficiency

The next step consisted in comparing the sequences of the selected peptides to gain insight into the sequence determinants that are important for the effect of the peptides. The sequence of the 20 most efficient peptides was determined. As shown in Table 2, all the peptides have a different sequence. Also, residues at positions 7 to 10 of the selected peptides tend to be different from unselected peptides (see Table 3), suggesting selection for these features, whereas the other positions show no statistically significant variations. As positions 7 to 10 forms the turn in the peptide loop presented by the thioredoxin scaffold, they may be the most solvent-exposed and therefore available to form intermolecular interactions. An arginine residue is always present at position 9 and predominates at position 8. Serine residues are also more represented in positions 7 and 10. In some cases, other amino acids such as threonine and proline that apparently arise from changes of serine or arginine codons are found at position 11, but the occurrence of these amino acids was equally represented in the selected and non-selected sequences. The appearance of such mutant peptides is frequently observed with peptide libraries (Harada et al., 1996; Xu et al., 2001). The absence of a clear consensus sequence of active peptides was expected, as it is possible that active peptides bind to different targets or to the same target in different ways. There is no selection towards polyarginine sequences. Rather, single, non-selected sequences with high arginine content (RRRSRRRRRRRRRS

and RHRSSRRRRRRRPHRH) displayed only the expected library background activity (a frameshift efficiency decrease of less than 15%, data not shown), indicating that non-specific counterion effects are not the basis of the observed activity of the selected peptides.

The activity of the selected clones is not specific for the HIV-1 group M frameshift stimulatory signal

In order to establish whether the effect observed with the selected peptides was specific or not for the HIV-1 group M frameshift stimulatory signal, we assessed the activity of the 20 most active clones against 4 other engineered frameshift systems, where the frameshift stimulatory signal of HIV-1 group M subtype B was replaced with another frameshift stimulatory signal. As depicted in Figure 5, we used the HIV-1 group O frameshift stimulatory signal (Baril et al., 2003) and artificial stimulatory signals such as the human IRE (Kollmus et al., 1996, see Introduction), HIV-1 RRE and BIV TAR. The frameshift efficiency measured with IRE and HIV-1 group O stimulatory signal was about 5.9% and 4.6%, respectively, which is higher than HIV-1 group M subtype B (3.7%). Lower frameshift efficiency values were observed when the RRE and TAR structures were used as a stimulatory signal (1.8% and 3.4%, respectively). For the selectivity assays, the peptides were expressed in the presence of each of the five different reporters with or without induction of the peptides and changes in the GFP/RFP ratio were measured with a plate reader. As shown in Figure 6, no specificity for the HIV-1 group M frameshift stimulatory signal is observed, although the activity profiles with the different reporters slightly differ between active peptides.

Scaffold-independent effects of peptides on frameshift in eukaryote cells

In the bacterial screening system we used, the peptide sequences are expressed as conformationally constrained loops by thioredoxin as a scaffold. To explore whether the activity of the candidate peptides was brought about by a scaffold-dependent peptide conformation, and to exclude that residues derived from the scaffold contribute crucially to their activity, we tested the activity of several peptides in a different scaffold, this time using an eukaryote expression and reporter system in HEK293T cells. The scaffold used was derived from Xu et al. (2001), consisting in short dimerizing protein sequences, which lead to the presentation of the inserted peptide as a loop, similar to cyclic peptides. The eukaryote reporter expressed *Renilla* and firefly luciferase on the same mRNA, such that *Renilla* luciferase expression follows conventional rules of translation and firefly luciferase expression depends on -1 frameshifting brought about by the HIV-1 frameshift region (Grentzmann et al., 1998; Dulude et al., 2006). As shown in Figure 7, two out of ten tested peptide candidates reduced frameshift efficiency by 35-40%. This last selection step, assessing scaffold independence and activity on frameshifting by the eukaryote ribosome, thus revealed that a number of the candidate sequences identified in the bacterial screening system retain activity when presented in eukaryotes by a different scaffold. Our results demonstrate that a rapid screening procedure in bacteria permits efficient identification of a large number of active peptides and is suitable for the identification of candidates that display activity in a relevant, eukaryotic context.

Discussion

We described here a bacterial system that enables the selection of peptides that reduce HIV-1 frameshift efficiency, using a two-fluorescent reporter system. In this system, the expression of one reporter, GFP, depends upon the frameshift whereas the expression of the other reporter, RFP, depends upon normal translation. The peptides were selected from a combinatorial arginine-rich

peptide library adapted from Harada et al. (1996). Several studies have reported that frameshifting can be modulated by mutations in the slippery sequence (Biswas et al., 2004; Dulude et al., 2006), in the frameshift stimulatory signal (Telenti et al., 2002; Dulude et al., 2006), in the ribosomes (Léger et al., 2004) or in elongation factors (Dinman & Kinzy, 1997) and by drugs acting on the ribosome (Dinman et al., 1997). The screening of a library of chemical compounds enabled the selection of one compound that increased HIV-1 frameshift efficiency and interfered with HIV-1 replication (Hung et al., 1998). Like the active peptide sequences that we report here, this compound was found to be active against a variety of frameshift stimulators. Recently, a small ligand binding to the upper stem-loop of HIV-1 frameshift stimulatory signal was selected from a resin-bound dynamic combinatorial library (McNaughton et al., 2007), but the effect of this compound on frameshift efficiency has not yet been reported. Our study presents the first report of the selection of frameshift modulators using a combinatorial peptide library approach. The use of fluorescent reporters to monitor the changes in frameshifting allowed us to sort promising candidates by flow cytometry. Out of 1,150 clones selected by flow cytometry as containing peptides that decrease HIV-1 frameshift efficiency, 20 clones that decreased frameshift efficiency by about 50% were further analyzed.

Binding affinity of peptides for RNA is largely contributed by charge-charge interactions between basic amino acids and the phosphate groups of the acid backbone of nucleic acids, and polybasic stretches are found in many RNA binding proteins. Binding specificity in turn is brought about by non-basic amino acids that specifically interact with a given motif in the RNA, or with a specific RNA structure. The rationale for the use of arginine-based peptide sequences is therefore to obtain sufficient binding affinity to observe biological effects of a given interaction involving RNA structures, as is the case in frameshifting. However, binding specificity must not depend on the basic nature of the peptide (see Weiss & Narayana, 1998). No consensus sequence was seen in the

sequence of the peptides we selected, suggesting that the peptides bind different targets or the same target in a different way. However, it is clear that the activity of the selected peptides does not result from a non-specific counterion effect. Indeed, when examining the sequences of randomly chosen non-selected peptides, we found examples of sequences that consist almost entirely of arginine but had only a marginal effect on frameshift efficiency (data not shown). Interestingly, positions 7 to 10 present a characteristic pattern in a subset of the active sequences, and could be critical for recognition of the target. None of these peptides was selective for the HIV-1 group M subtype B frameshift stimulatory signal. We cannot exclude the possibility that these peptides could interact with each of the investigated signals, with different affinities that are responsible for the observed differences in activity against the different reporters. Another explanation is that they act on the ribosome, in a way that does not equally affect the frameshift induced by the different stimulatory signals. We mentioned above that our selection procedure excluded peptides that bind to the ribosome and affect normal protein synthesis. However, the selected peptides could interfere with ribosome-dependent steps that are critical for the frameshift. For example, an effect of the selected peptides on the ribosome helicase is an attractive hypothesis. Indeed, an increase of helicase activity would not affect normal translation whereas the frameshift event, which depends upon the pausing caused by the resistance to unwinding of the signal, would be expected to be sensitive to variations in the helicase activity. Such a change of helicase activity should only marginally affect global translation, since initiation, and not elongation, is the rate-limiting step of translation (Laursen et al., 2005). Alternatively, interference of the peptides with the tRNA slippage required for the frameshift could be their mechanism of action.

The proportion of the peptides that retain their activity independently of the presenting scaffold, and are active in both bacterial and eukaryote reporter systems is fair enough to constitute the basis

for further affinity maturation steps, be it by targeted mutagenesis or subsequent library screening. Loss of activity during the change of scaffold and expression system may be due to the differences in the scaffold, and differences between prokaryotic and eukaryotic ribosomes. Indeed, although prokaryote and eukaryote ribosomes are very similar, an extensive number of additional proteins in the eukaryote ribosome (Spahn et al., 2004) could occlude potential interaction sites for the peptides that are present in the bacterial ribosome. The fact that not all the selected peptides are active in a eukaryote system further supports that the decrease in frameshift efficiency observed in the presence of these peptides results from a specific effect, and not from an unspecific counterion effect. Our results suggest that synthetic peptides derived from the active sequences could represent leads for the development of smaller, non-peptidic anti-frameshift compounds. In addition, to our best knowledge, this is also the first report that peptides targeting RNA structures identified in bacterial screening systems can retain their activity also in eukaryotes. The assessment of their activity in HIV-infected cells may result in leads for the development of compounds that interfere with HIV-1 frameshifting and, consequently, its replication. Moreover, the identification of the target to which the peptides bind, ribosomal components or the frameshift stimulatory signal, will permit further investigation of the mechanism of frameshift.

Materials and Methods

Reporter and expression plasmids

The p3RFP-HIV(-1)-GFP reporter plasmid was constructed to monitor changes in the HIV-1 frameshift efficiency in the presence of peptides generated from a library. Briefly, this bacterial vector contains the HIV-1 group M subtype B frameshift region (slippery site and the stimulatory

signal) inserted in the beginning of the green fluorescent protein (GFP) coding sequence such that its expression requires a programmed -1 frameshift. The vector also contains three copies of the red fluorescent protein (RFP) coding sequence that is used as an internal reporter to assess conventional translation. Three copies were used to obtain a sufficiently high level of fluorescence to be monitored by flow cytometry. The GFP and RFP transcripts are both constitutively expressed from a *trp* promoter. This reporter plasmid was derived from pRNA_{luc2} (Belanger et al., 2002), that contains, under control of a *lac* promoter, an IPTG inducible ribosomal RNA operon in which the 16S anti Shine-Dalgarno sequence is mutated and a reporter luciferase gene that was eliminated in our reporter. The RFP and GFP reporter genes have a mutated ribosome-binding site that is complementary to the mutated anti Shine-Dalgarno sequence of 16S rRNA, such that only ribosomes containing the plasmid-encoded 16S rRNA translate the reporter mRNAs. This specificity of the plasmid ensures a high expression of the reporter genes, and no interference with translation of bacterial mRNAs.

The p3RFP-HIV(-1)-GFP reporter plasmid was created by first amplifying the GFP coding sequence by PCR from pGFP_{emd-N1}, a GFP fusion protein expression vector (a generous gift from Dr M. Bouvier, Université de Montréal). The A and B primers (see supporting data for these primer sequences and all the primers used in this study) used for this amplification introduced an *NsiI* and an *XhoI* restriction site at the 5' and 3' end of the gene, respectively. The amplified GFP gene was inserted downstream of the *trp* promoter of *NsiI-XhoI*-digested pRNA_{luc2}, generating the pGFP construct. The frameshift region of HIV-1 group M subtype B (nucleotides 2075 to 2143 of HIV-1_{HXB2}, GenBank accession number K03455) was introduced at the beginning of the GFP coding sequence by inserting an appropriate annealed and phosphorylated 72 bp oligonucleotide cassette (purchased from Sigma Genosys), in the digested *NsiI* site of pGFP. This generated pHIV(-1)-GFP.

An in-frame control plasmid, pHIV(0)-GFP, was derived from pGFP(-1) by inserting an adenine immediately after the slippery sequence in the HIV-1 frameshift region, such that GFP is synthesized according to the conventional rules of translation. A pRFP construct was derived from pRNA_{luc2} using the same approach that was used for creating pGFP. The RFP gene amplified from pDsRed-T4, a generous gift from Dr. B.S. Glick, University of Chicago (Bevis & Glick, 2002) was introduced between the *Nsi*I and *Xho*I restriction sites using the primers C and D. The pRFP construct was used as a template to PCR-amplify the RFP coding sequence and its promoter with primers E and F that introduced a *Kpn*I restriction site at the 5' and 3' end of the RFP gene. The *Kpn*I-digested RFP fragment was inserted in the digested *Kpn*I site of pHIV(-1)GFP and pRFP to generate pRFP-HIV(-1)-GFP and p2RFP, respectively. Finally, the two RFP copies contained in plasmid p2RFP were amplified with primers G and H as a fragment with a *Nhe*I restriction site at the 5' and 3' end. This *Nhe*I digested fragment was then inserted in the digested *Nhe*I site of pRFP-HIV(-1)GFP, generating p3RFP-HIV(-1)-GFP.

Derivatives of p3RFP-HIV(-1)-GFP, where the HIV-1 group M frameshift stimulatory signal was replaced with another frameshift stimulatory signal, were also constructed. The frameshift stimulatory signal of HIV-1 group O, the RRE secondary structure of HIV-1, the TAR secondary structure of BIV or the human IRE were used as stimulatory signals. To this end, a 2.4kB *Nde*I-*Mfe*I fragment of p3RFP-HIV(-1)-GFP containing the frameshift region was inserted between the equivalent restriction sites of pGFP_{emd-N1}, the plasmid from which was derived the GFP coding sequence we use (see above), generating plasmid pHIV-1-GFP_{emd-N1}. This plasmid was digested with the *Bgl*II and *Mfe*I restriction enzymes to remove the HIV-1 group M frameshift stimulatory signal and an oligonucleotide cassette containing the sequences of the pseudoknot from HIV-1 group O, or the irregular stem-loops corresponding to BIV-1 TAR, HIV-1 RRE or human IRE was inserted.

The 2.4Kb NdeI-MfeI fragment of the intermediates generated with these insertions was then reintroduced into p3RFP-HIV(-1)-GFP to create the p3RFP-HIV(-1)grO-GFP, p3RFP-RRE(-1)-GFP, p3RFP-TAR(-1)-GFP and p3RFP-IRE(-1)-GFP derivatives.

The plasmid used for the expression of the peptide library was derived from the pBAD-Thio-E (Invitrogen), which contains the *E. coli* thioredoxin (*trxA*) gene under control of an arabinose-inducible promoter. This plasmid was modified into pACYC-Thio by replacing its ampicillin resistance gene and its origin of replication with a kanamycin resistance gene and a p15a origin, so as to be compatible with p3RFP-HIV(-1)-GFP. To this end, pBAD-Thio-E was digested with *Bsr*BI, the 2.4 Kb fragment containing the *trxA* gene with the upstream arabinose-inducible promoter was isolated and ligated with a 3.2 Kb fragment from *Dra*I-digested pACYC177 (New England Biolabs) that contains the p15a origin and the kanamycin resistance gene.

Design of the combinatorial arginin-rich peptide library

Plasmid pACYC-Thio was used to express a combinatorial library coding for 14-residue peptides embedded within the thioredoxin coding sequence. The library was made following a procedure adapted from Harada et al. (1996), using a degenerate oligonucleotide, 5'-GACTGACTGGTCCG(MRT)₁₄GGTCCTCAGCAGTCAG-3' (Sigma Genosys), containing 14 degenerate codons. For each of these codons, M is a A:C mixture in a 1:3 ratio, and R is an A:G mixture in a 3:1 ratio. Each randomized MRT codon encodes arginine, serine, histidine and asparagine in a proportion of 56.25%, 18.75%, 18.75% and 6.25%, respectively. Double-stranded DNAs were synthesized by elongation of an annealed primer (5'-CAGGGGATTTGCTGCATTC-3'), using the Sequanase 2.0 (USB) as a DNA polymerase. These cassettes were digested with *Ava*II, ligated into *Rsr*II-digested pACYC-Thio and used to transform *E. coli* XL-10 Gold (Stratagene).

Selection of peptides altering HIV-1 frameshift efficiency

E. coli Top10 competent cells containing the p3RFP-HIV(-1)-GFP reporter plasmid were used for the expression of the peptide library. For the flow cytometry screening, 300 ng of pThio-LIB DNA were electroporated into these Top10 competent cells, generating about 1.5×10^7 transformants, and grown in 1 ml of LB for 1 h at 37°C without antibiotic. The culture was then diluted with LB to an O.D. at 600 nm of about 0.4 and 100 µg/ml ampicillin and 40 µg/ml kanamycin were added. IPTG (1 mM) and arabinose (0.2% w/v) were also added to the culture medium to allow expression of RFP and GFP from the p3RFP-HIV(-1)-GFP reporter and expression of the arginine-rich peptides from the pThio-LIB plasmid, respectively. Bacteria were allowed to grow for an additional 16 h at 37°C before being sorted by a FACSVantage SE cell sorter (BD Biosciences). For the sorting, bacteria were diluted in PBS at a density of 0.05 O.D. at 600 nm. Fluorescent proteins were excited at 488 nm and the fluorescence emission for the GFP and RFP was recorded at 530 and 586 nm, respectively. Bacteria with the desired profile of fluorescent protein expression were collected and plated on agar.

Fluorescence measurements

E. coli Top10 competent cells containing the p3RFP-HIV(-1)-GFP reporter plasmid or one of its derivatives, p3RFP-HIV(-1)grO-GFP, p3RFP-RRE(-1)-GFP, p3RFP-TAR(-1)-GFP or p3RFP-IRE(-1)-GFP, were prepared using a standard CaCl₂ procedure and heat-shock transformed with the DNA encoding for the selected peptide (1 ng for 40 µl of competent cells) before being plated on ampicillin/kanamycin-containing agar plates. For the fluorescence readings, overnight cultures of individual clones grown in ampicillin/kanamycin LB medium were diluted to 0.1 O.D. at 600 nm and grown in 96-well plates (2 ml/well) containing 1 ml of LB with 0.2% arabinose, 100 µg/ml

ampicillin and 40 µg/ml kanamycin for 1 h at 37°C at a rotating speed of 250 rpm. IPTG (1 mM) was added and the culture grown for an additional 16 h. Bacteria were then spun down and washed with 500 µl PBS. Bacterial pellets were resuspended in 250 µl of PBS and transferred into 96-well plates and RFP and GFP fluorescence was measured with a Fusion α-FP (Packard) plate reader at 580/15 nm and 535/20 nm, respectively, following an excitation at 485/20 nm.

Eukaryotic frameshift assays

The peptides were flanked by two short dimerizing sequences for induced conformational constraint (Xu et al., 2001). To this end, an oligonucleotide cassette containing the dimerizing sequence was first introduced between the BstXI-XhoI restriction sites of the pLPC expression vector (Serrano et al. 1997). Oligonucleotide cassettes coding for the selected peptides sequences (peptides 1 to 10 of Table 2) were then inserted between the BspEI-EcoRI restriction sites of this construct. Transfection of the resulting plasmids led to the expression of the peptide 5'-GEFLIVKSGVSKAH-X₁₄-GEFLIVKSGPP-3', where X is the 14 amino acid arginine-rich inserted sequence.

The effect of the selected peptides on the frameshift efficiency was measured by cotransfection with a dual-luciferase reporter (Grentzmann et al., 1998; Dulude et al., 2006) into 293T cells. The day before transfection, 3x10⁵ cells/well were seeded in 6-well plates and maintained in DMEM (10% FBS, Wisent). The derivatives of the pLPC vector containing the peptides (2.5 µg) were co-transfected with 0.1 µg of pDual-HIV(-1) into 293T cells, using polyethyleneimine at 3 µg per µg of DNA, and cells were grown for 48 h before being harvested. The firefly versus the *Renilla* luciferase activities were measured with a Lumat LB9507 luminometer (Berthold) with a Dual luciferase® reporter assay kit (Promega), following the manufacturer's instructions.

Statistical analysis

To assess the significance of accumulation of amino acid residues at given positions in selected peptides, binomial distribution tests were performed, and cumulative probabilities of $p < 0.05$ were seen as significant. For these tests, means from unselected peptides were used as expected frequencies.

Acknowledgements

This study was supported by a grant from the Canadian Institutes of Health Research (CIHR) to Léa Brakier-Gingras and Nikolaus Heveker. Dominic Dulude acknowledges a postdoctoral fellowship from CIHR and Nikolaus Heveker acknowledges a CIHR New Investigator Fellowship. We are grateful to Gerardo Ferbeyre and Stephen Michnick for critical reading of the manuscript and for stimulating discussions.

References

- Baril M, Dulude D, Steinberg SV, Brakier-Gingras L. 2003. The frameshift stimulatory signal of human immunodeficiency virus type 1 group O is a pseudoknot. *J Mol Biol* 331:571-583.
- Belanger F, Leger M, Saraiya AA, Cunningham PR, Brakier-Gingras L. 2002. Functional studies of the 900 tetraloop capping helix 27 of 16S ribosomal RNA. *J Mol Biol* 320:979-989.
- Bevis BJ, Glick BS. 2002. Rapidly maturing variants of the *Discosoma* red fluorescent protein (DsRed). *Nat Biotechnol* 20:83-87.

- Biswas P, Jiang X, Pacchia AL, Dougherty JP, Peltz SW. 2004. The human immunodeficiency virus type 1 ribosomal frameshifting site is an invariant sequence determinant and an important target for antiviral therapy. *J Virol* 78:2082-2087.
- Brierley I, Dos Ramos FJ. 2006. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res* 119:29-42.
- Brierley I, Pennell S. 2001. Structure and function of the stimulatory RNAs involved in programmed eukaryotic-1 ribosomal frameshifting. *Cold Spring Harb Symp Quant Biol* 66:233-248.
- Corey MJ, Corey E. 1996. On the failure of de novo-designed peptides as biocatalysts. *Proc Natl Acad Sci U S A* 93:11428-11434.
- Dinman JD, Kinzy TG. 1997. Translational misreading: mutations in translation elongation factor 1alpha differentially affect programmed ribosomal frameshifting and drug sensitivity. *RNA* 3:870-881.
- Dinman JD, Ruiz-Echevarria MJ, Czaplinski K, Peltz SW. 1997. Peptidyl-transferase inhibitors have antiviral properties by altering programmed -1 ribosomal frameshifting efficiencies: development of model systems. *Proc Natl Acad Sci U S A* 94:6606-6611.
- Dulude D, Baril M, Brakier-Gingras L. 2002. Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res* 30:5094-5102.
- Dulude D, Berchiche YA, Gendron K, Brakier-Gingras L, Heveker N. 2006. Decreasing the frameshift efficiency translates into an equivalent reduction of the replication of the human immunodeficiency virus type 1. *Virology* 345:127-136.
- Gaudin C, Mazauric MH, Traikia M, Guittet E, Yoshizawa S, Fourmy D. 2005. Structure of the RNA signal essential for translational frameshifting in HIV-1. *J Mol Biol* 349:1024-1035.
- Grentzmann G, Ingram JA, Kelly PJ, Gesteland RF, Atkins JF. 1998. A dual-luciferase reporter system for studying recoding signals. *RNA* 4:479-486.
- Harada K, Martin SS, Frankel AD. 1996. Selection of RNA-binding peptides in vivo. *Nature* 380:175-179.

- Harada K, Martin SS, Tan R, Frankel AD. 1997. Molding a peptide into an RNA site by in vivo peptide evolution. *Proc Natl Acad Sci U S A* 94:11887-11892.
- Hung M, Patel P, Davis S, Green SR. 1998. Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication. *J Virol* 72:4819-4824.
- Jacks T, Madhani HD, Masiarz FR, Varmus HE. 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell* 55:447-458.
- Karacostas V, Wolffe EJ, Nagashima K, Gonda MA, Moss B. 1993. Overexpression of the HIV-1 gag-pol polyprotein results in intracellular activation of HIV-1 protease and inhibition of assembly and budding of virus-like particles. *Virology* 193:661-671.
- Kollmus H, Hentze MW, Hauser H. 1996. Regulated ribosomal frameshifting by an RNA-protein interaction. *RNA* 2:316-323.
- Laursen BS, Sorensen HP, Mortensen KK, Sperling-Petersen HU. 2005. Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* 69:101-123.
- LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM. 1993. A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Biotechnology (N Y)* 11:187-193.
- Lee K, Holland-Staley CA, Cunningham PR. 1996. Genetic analysis of the Shine-Dalgarno interaction: selection of alternative functional mRNA-rRNA combinations. *RNA* 2:1270-1285.
- Lee K, Varma S, SantaLucia J, Jr., Cunningham PR. 1997. In vivo determination of RNA structure-function relationships: analysis of the 790 loop in ribosomal RNA. *J Mol Biol* 269:732-743.
- Léger M, Dulude D, Steinberg SV, Brakier-Gingras L. 2007. The three transfer RNAs occupying the A, P and E sites on the ribosome are involved in viral programmed -1 ribosomal frameshift. *Nucleic Acids Res*, 35:5581-5592.
- Léger M, Sidani S, Brakier-Gingras L. 2004. A reassessment of the response of the bacterial ribosome to the frameshift stimulatory signal of the human immunodeficiency virus type 1. *RNA* 10:1225-1235.

- Lu Z, Tripp BC, McCoy JM. 1998. Displaying libraries of conformationally constrained peptides on the surface of *Escherichia coli* as flagellin fusions. *Methods Mol Biol* 87:265-280.
- McNaughton BR, Gareiss PC, Miller BL. 2007. Identification of a Selective Small-Molecule Ligand for HIV-1 Frameshift-Inducing Stem-Loop RNA from an 11,325 Member Resin Bound Dynamic Combinatorial Library. *J Am Chem Soc*.
- Namy O, Moran SJ, Stuart DI, Gilbert RJ, Brierley I. 2006. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* 441:244-247.
- Park J, Morrow CD. 1991. Overexpression of the gag-pol precursor from human immunodeficiency virus type 1 proviral genomes results in efficient proteolytic processing in the absence of virion production. *J Virol* 65:5111-5117.
- Plant EP, Dinman JD. 2006. Comparative study of the effects of heptameric slippery site composition on -1 frameshifting among different eukaryotic systems. *RNA* 12:666-673.
- Plant EP, Jacobs KL, Harger JW, Meskauskas A, Jacobs JL, Baxter JL, Petrov AN, Dinman JD. 2003. The 9-A solution: how mRNA pseudoknots promote efficient programmed -1 ribosomal frameshifting. *RNA* 9:168-174.
- Shehu-Xhilaga M, Crowe SM, Mak J. 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J Virol* 75:1834-1841.
- Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. 2004. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* 118:465-475.
- Staple DW, Butcher SE. 2005. Solution structure and thermodynamic investigation of the HIV-1 frameshift inducing element. *J Mol Biol* 349:1011-1023.
- Takyar S, Hickerson RP, Noller HF. 2005. mRNA helicase activity of the ribosome. *Cell* 120:49-58.
- Telenti A, Martinez R, Munoz M, Bleiber G, Greub G, Sanglard D, Peters S. 2002. Analysis of natural variants of the human immunodeficiency virus type 1 gag-pol frameshift stem-loop structure. *J Virol* 76:7868-7873.

- Thompson J, O'Connor M, Mills JA, Dahlberg AE. 2002. The protein synthesis inhibitors, oxazolidinones and chloramphenicol, cause extensive translational inaccuracy in vivo. *J Mol Biol* 322:273-279.
- Weiss MA, Narayana N. 1998. RNA recognition by arginine-rich peptide motifs. *Biopolymers* 48:167-180.
- Weiss RB, Dunn DM, Shuh M, Atkins JF, Gesteland RF. 1989. E. coli ribosomes rephase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol* 1:159-169.
- Xu X, Leo C, Jang Y, Chan E, Padilla D, Huang BC, Lin T, Gururaja T, Hitoshi Y, Lorens JB, Anderson DC, Sikic B, Luo Y, Payan DG, Nolan GP. 2001. Dominant effector genetics in mammalian cells. *Nat Genet* 27:23-29.
- Yusupova GZ, Yusupov MM, Cate JH, Noller HF. 2001. The path of messenger RNA through the ribosome. *Cell* 106:233-241.

Table 1. Summary of the process of selection of peptides by flow cytometry.

	Number of events	% of parental events
Analyzed	2.2×10^7	-
Gated	4.4×10^4	0.2
Growth on agar plates	$\approx 5 \times 10^3$	11

The peptide library was expressed in bacteria bearing the reporter plasmid and the reporter fluorescence was analyzed by flow cytometry after 16h of peptide induction. Gates were set to define the events with a decreased GFP/RFP ratio. (Fig. 1C). Bacteria detected within these gates were isolated by cell sorting and plated on agar. The table shows the number of events and the percentage of the parental events for each step of the selection process.

Table 2. Sequence of the 20 most efficient selected peptides reducing HIV-1 group M frameshift efficiency.

Clone No.	Peptide sequence
1	NHRRHSRRSSRRS
2	RRRHRSRRRRRS
3	RRRSRRRRRPRNS
4	RRNRSSRRSRRRH
5	RHHSRRRRRSRRRR
6	RHSRSRSRRRPRRS
7	RRRSRHRRRSRRHR
8	SRHSRRHRRRRRRR
9	HHRRRRRRRRSSRS
10	RRRRRRRRRRRTSSR
11	NRRRRHRRRRSRRS
12	RRRRHSRRRPRRR
13	RRSSNRRRRSRRRR
14	RSHRRRRRRRRSNR
15	SRRRSRRRRPRRR
16	SRRRRSRRSRRRR
17	HRRRRRRRRRSRHH
18	HRRRRRRRRHRRRR
19	RRRRRRRRRSNRR
20	RRSRRRRRRRPRRH

Table 3. Distribution of the amino acids in the selected and unselected peptides.

A) Selected peptides														
Amino acid	Amino acid position													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R	65%	70%	55%	75%	80%	75%	65%	95%	100%	60%	50%	80%	75%	55%
H	15%	25%	25%	5%	5%	20%	5%	0%	0%	5%	0%	0%	10%	15%
S	10%	5%	15%	20%	10%	5%	30%	5%	0%	35%	20%	15%	5%	30%
N	10%	0%	5%	0%	5%	0%	0%	0%	0%	0%	0%	5%	10%	0%

B) Unselected peptides														
Amino acid	Amino acid position													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R	66%	52%	72%	62%	69%	59%	83%	72%	72%	72%	55%	69%	76%	69%
H	10%	34%	10%	17%	14%	28%	10%	21%	14%	14%	3%	14%	10%	17%
S	21%	14%	17%	21%	14%	14%	3%	3%	14%	14%	3%	17%	10%	14%
N	3%	0%	0%	0%	3%	0%	3%	3%	0%	0%	7%	0%	3%	0%

Frequency of different amino acid residues at the different positions of A) 20 active sequences (see Table for peptide sequences) and B) 29 random picked sequences of the library. Shaded numbers indicate over and under representation of several residues (binomial distribution tests were performed for each position and cumulative probabilities of $p < 0.05$ were seen as significant).

Figure legends

Figure 1. The two-plasmid reporter system used for the selection of arginine-rich peptides affecting HIV-1 frameshift. A) Frameshift reporter plasmid p3RFP-HIV(-1)-GFP. In this construct, the HIV-1 programmed -1 frameshift region encompassing the slippery site and the complete stimulatory signal was inserted downstream from the initiator AUG codon of GFP, such that GFP translation depends on a -1 frameshift. RFP translation was used as a control for conventional protein synthesis. With this system, changes in the frameshift efficiency are reflected by changes in the GFP/RFP ratio. B) Expression vector for the arginine-rich peptide library. Peptides are displayed in the active site of the thioredoxin used as a scaffold for peptide presentation. The expression of the peptide library is induced with arabinose. C) Fluorescence of the p3RGFP-HIV(-1) reporter, as analyzed by flow cytometry. The display shows the fluorescence obtained with an excitation at 488 nm and recorded at 530 nm and 585 nm for a bacterial culture expressing the reporter plasmid. RFP fluorescence is shown on the Y axis and the GFP fluorescence on the X axis. Gates for sorting were set (depicted by the red box) to define and isolate events with a decreased GFP/RFP ratio, which corresponds to a reduced frameshift efficiency.

Figure 2. Analysis of the selected peptides. A) Effect of the peptides selected by flow cytometry on the frameshift efficiency. The GFP/RFP ratio in the absence of peptides (no induction) is set as 100%. The frameshift efficiency of the 1,150 selected clones, of the most efficient 160 and 20 selected clones (14% and 2% of 1,150, respectively) are shown (black bars) and compared, respectively, to 350 unselected clones and to the best 14% and 2% of these unselected clones (gray bars). Results are the means of three assays. The bars indicate the standard error of the means. *** indicates statistical significance of the differences between selected and unselected clones ($p < 0.05$). B) Effect of peptide

induction on the reporter fluorescence. The absolute fluorescence of GFP (x axis) and RFP (y axis) measured for the 1,150 selected clones is shown (red dots) and compared to unselected clones (green dots) and to bacteria without peptide induction (blue dots).

Figure 3. Activity of the 160 most efficient selected peptides before and after re-transformation.

The peptide encoding-plasmids of the 160 most efficient clones were retrieved and re-transformed into bacteria expressing the p3RFP-HIV(-1)-GFP reporter. Single colonies from each re-transformation were analyzed to confirm the effect of the peptide on the frameshift efficiency. Changes in frameshift efficiency are assessed by comparing the GFP/RFP ratio when expression of peptides is induced to the GFP/RFP ratio without induction (set to 100%). The frameshift efficiency of 20 unselected clones and of the 160 most efficient clones before and after re-transformation is shown. Results are the means of three experiments. The bars indicate the standard error of the means.

Figure 4. Effect of the selected peptides on conventional and on frameshift-dependent translation. A) Fluorescence of RFP (measuring conventional translation) and GFP (measuring frameshift-dependent translation) as a function of the bacterial culture density. Ten randomly chosen clones from the selected peptides (circles) and three unselected peptides (squares) are shown. Green symbols represent GFP expression and red symbols RFP expression. Peptide expression was induced in growing bacterial cultures and fluorescence was measured at several growth time points (between O.D.₆₀₀ of 0.4 to 1.1). B) Effect of selected and unselected clones on conventional and frameshift-dependent translation. The 20 most efficient selected and 80 unselected clones were grown for 12 h with induction of the peptides, and the GFP and RFP fluorescence were measured. Results are the

means of three experiments. The bars indicate the standard error of the means for the GFP and RFP fluorescence.

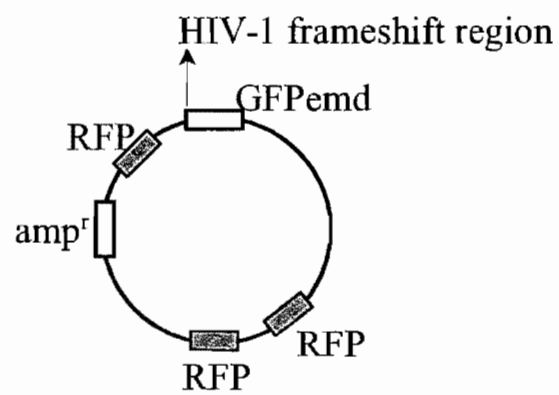
Figure 5. Frameshift stimulatory signals used to assess the selectivity of the peptides. The HIV-1 group M subtype B frameshift stimulatory signal used for the selection of frameshift-altering peptides was replaced with either the HIV-1 group O frameshift stimulatory signal (Baril et al., 2003) or artificial stimulatory signals such as the human IRE (Kollmus et al., 1996), the RRE of HIV-1 and TAR of BIV (structures taken from Weiss & Narayana, 1998), generating derivatives of p3RGF-HIV(-1)-GFP. Human IRE, BIV TAR, HIV-1 RRE and HIV-1 group M stimulatory signals are irregular stem-loops whereas the HIV-1 group O stimulatory signal folds into a pseudoknot.

Figure 6. Activity of selected peptides toward different frameshift stimulatory signals. The 20 most efficient clones presented in Table 2 were tested for their effect in frameshift systems containing the HIV-1 group M subtype B (HIV-1 stB), the human IRE, the HIV-1 Group O (HIV-1 grO), the HIV-1 rev-responsive element (HIV-1 RRE) or the BIV TAR structure acting as the frameshift stimulatory signal. Changes in the frameshift efficiency are obtained by dividing the GFP/RFP ratio obtained in the presence of peptides by the GFP/RFP ratio obtained without peptide. Selectivity of peptides 12, 14, 16 and 19 is shown in A) through D) respectively, which are representative examples for results obtained with all other peptides. Results are the means of at least three independent experiments. The bars indicate the standard error of the means.

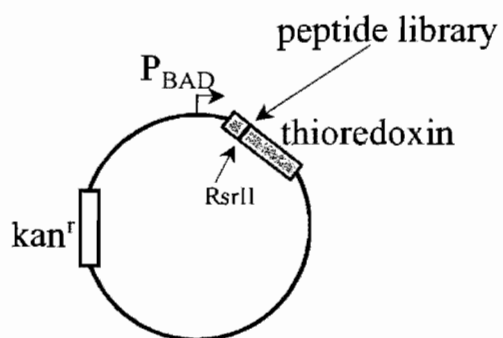
Figure 7. Effect of selected peptides on frameshift efficiency in a eukaryotic system. DNA sequences corresponding to peptides 3 and 10 selected with the bacterial system (see sequences in

Table 2) were cloned into a eukaryotic expression vector and co-transfected with the pDual-HIV(-1) reporter into 293T cells. Frameshift efficiencies are represented relative to the empty vector, which is set to be 100%. Ctrl is a random peptide sequence (RSRSRRRHRRRRRR) that did not affect the frameshift in the bacterial system. Results are the means of three independent experiments performed in duplicates. The bars indicate the standard error of the means.

A)



B)



C)

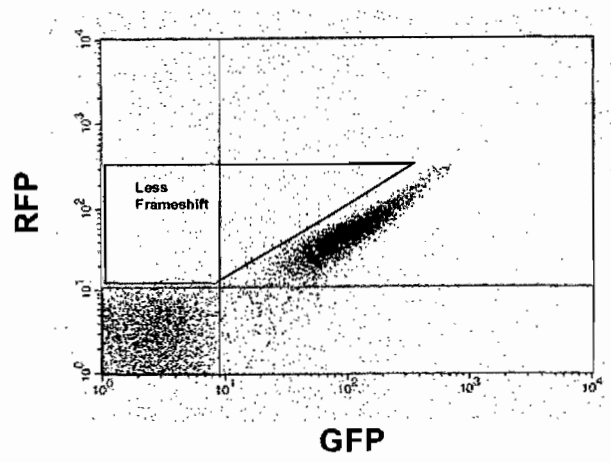
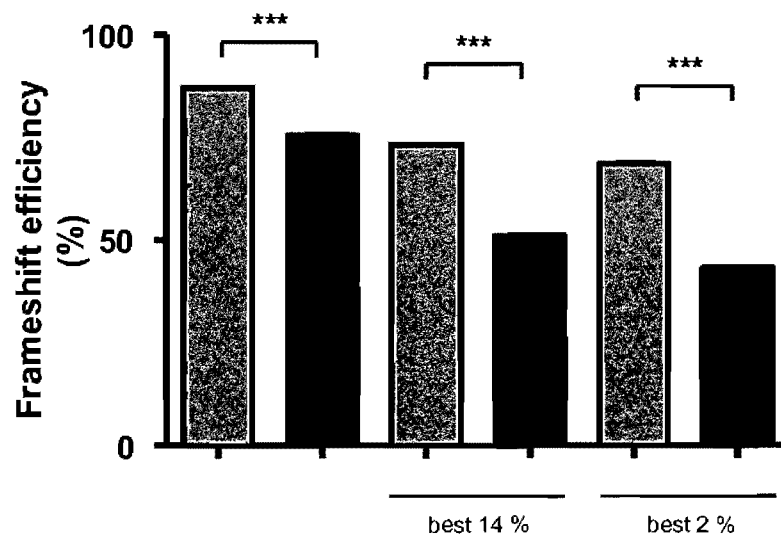


Figure 1.

A)



B)

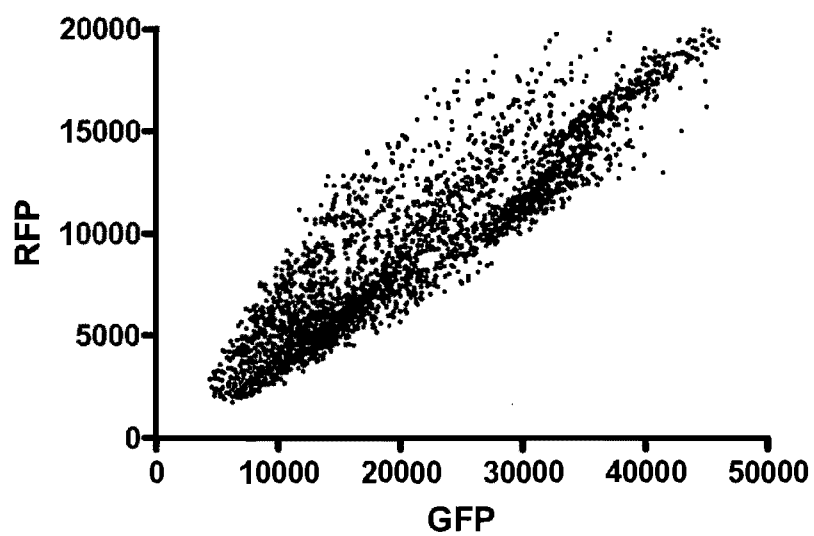


Figure 2.

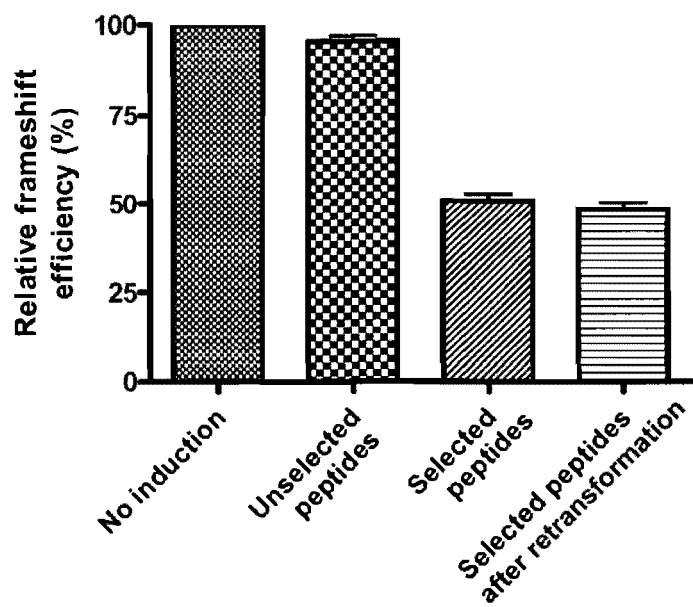
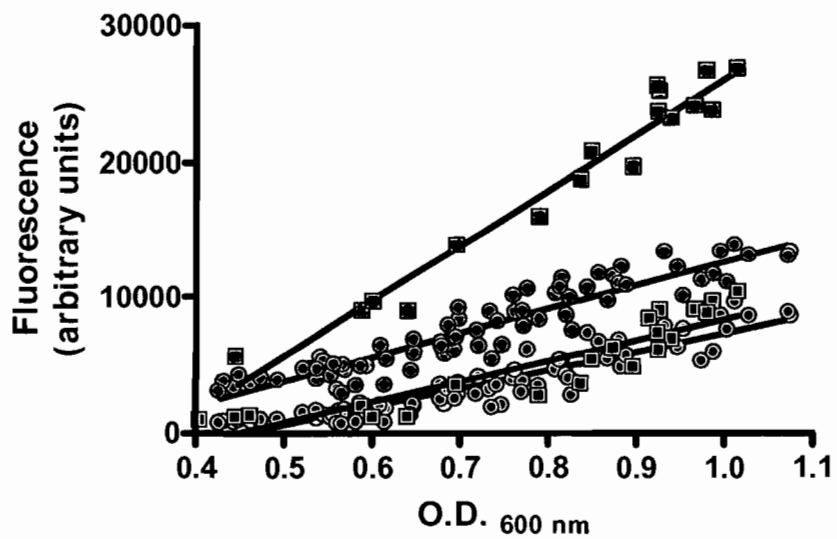


Figure 3.

A)



B)

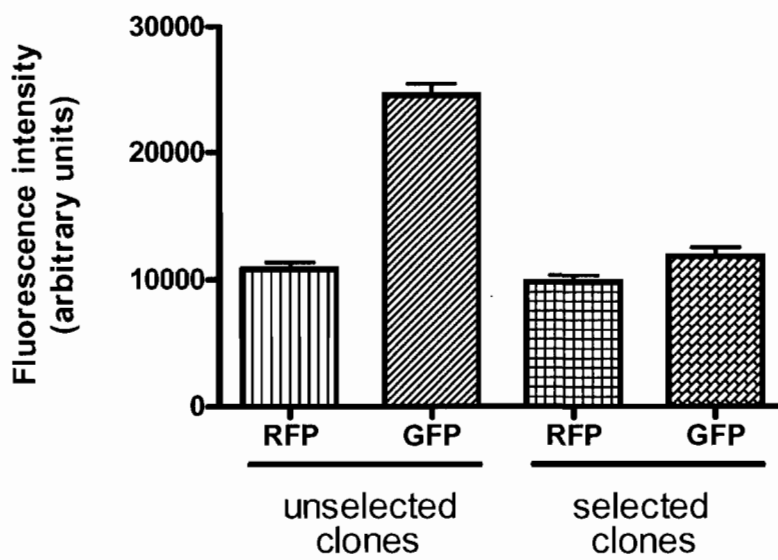


Figure 4.

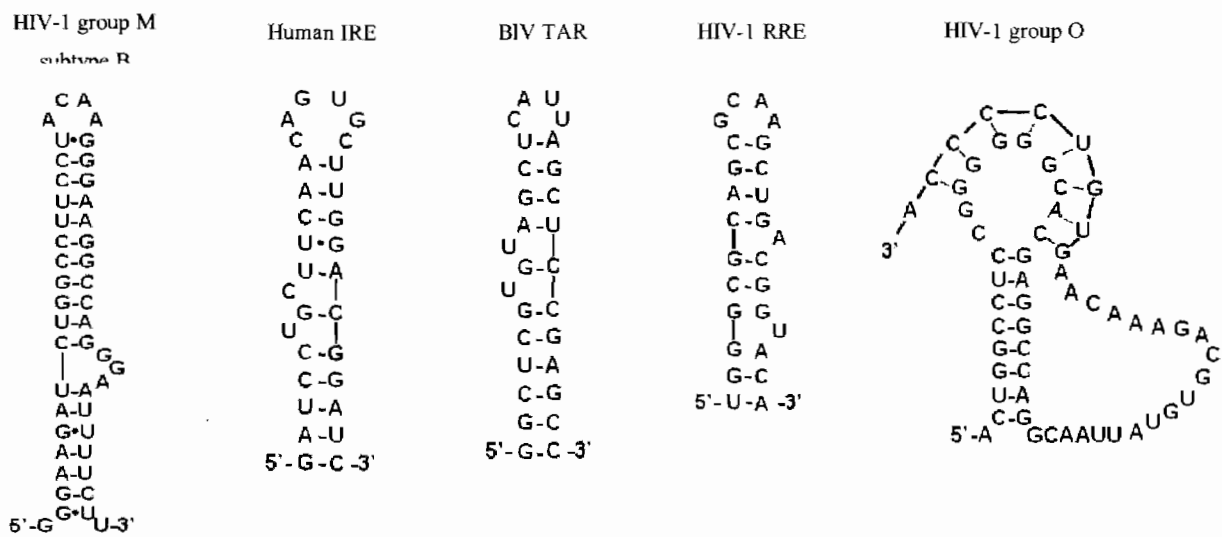


Figure 5.

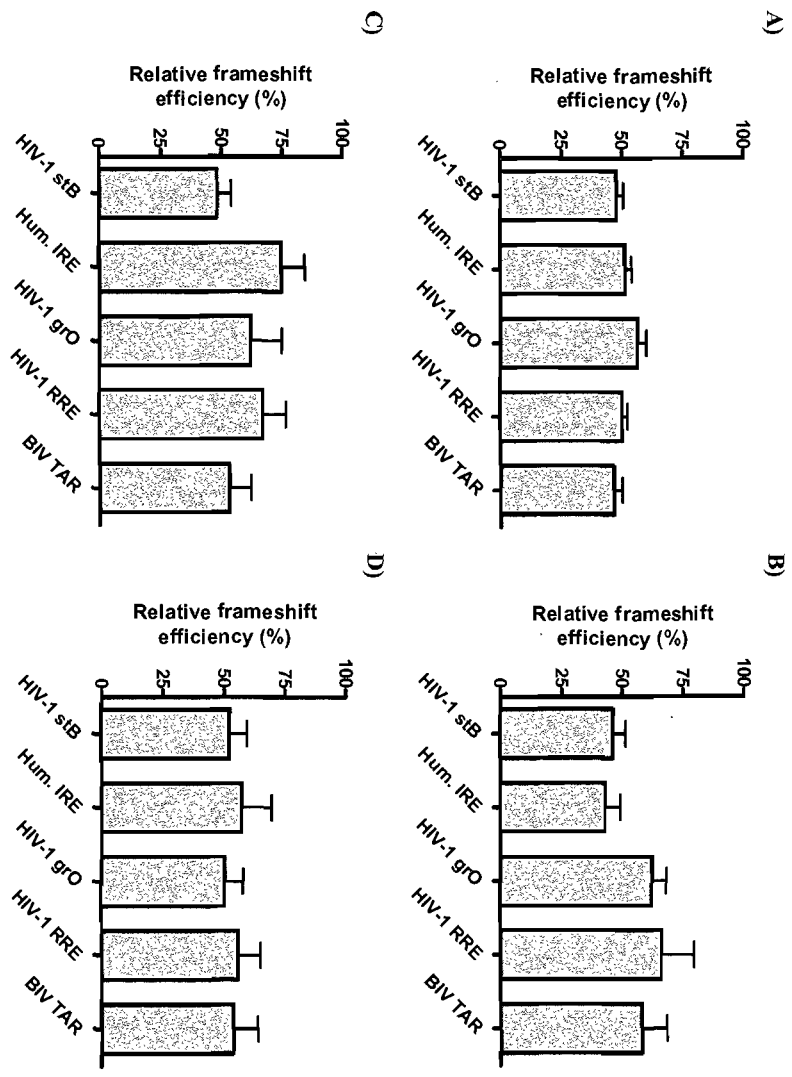


Figure 6.

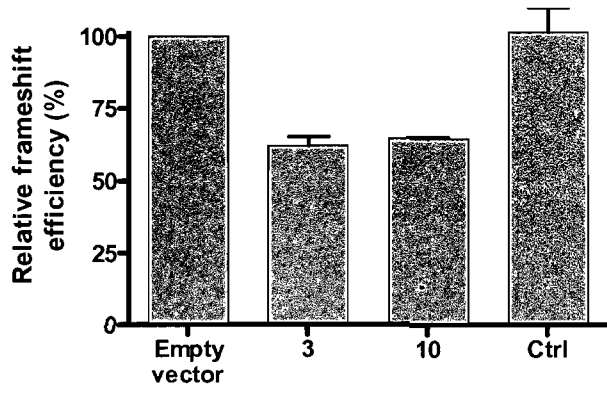


Figure 7.

Chapitre III : Discussion

3.1 Rappel de la problématique

Le virus de l'immunodéficience humaine de type 1 est à la source d'une importante pandémie touchant plus de 33 millions de personnes à travers le monde. Le développement d'inhibiteurs utilisés dans la trithérapie a permis d'offrir les premiers traitements efficaces contre ce virus. Malheureusement, l'émergence de variants résistants aux inhibiteurs de protéase et de transcriptase inverse utilisés dans cette thérapie demande le développement de nouveaux agents antirétroviraux ayant de nouvelles cibles.

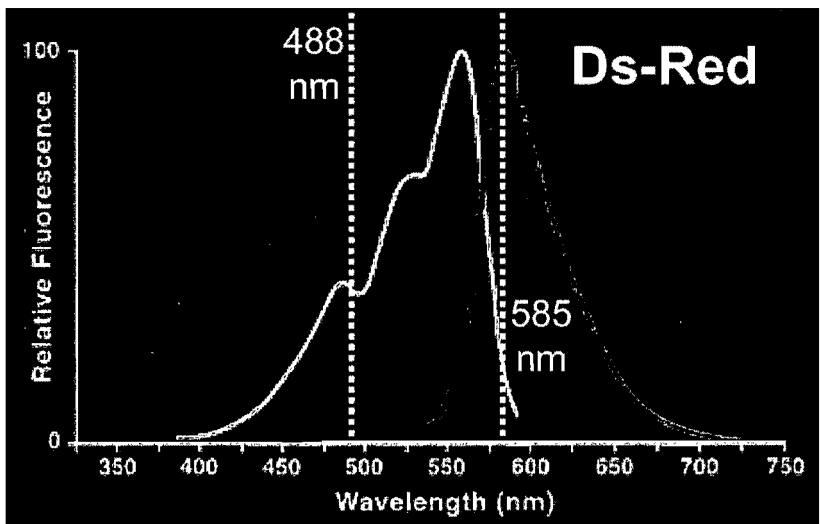
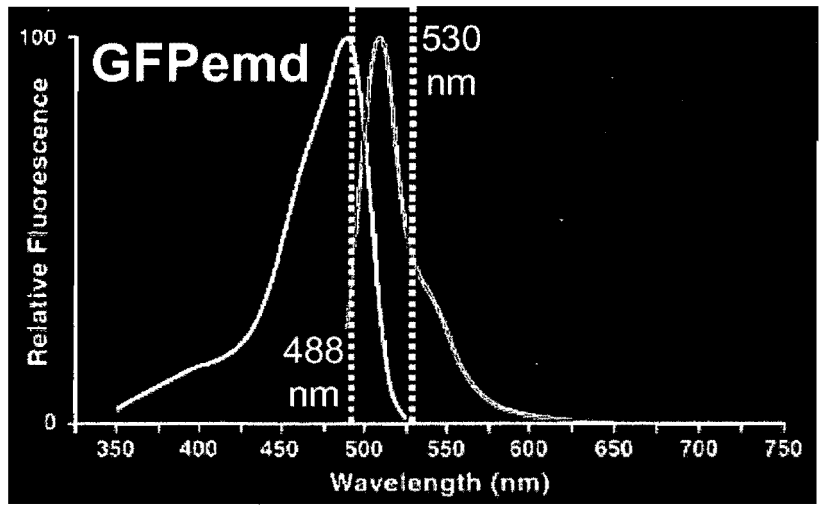
C'est dans cette optique que notre groupe s'intéresse depuis plusieurs années au *frameshift* ribosomique -1 programmé du VIH-1 que ce virus utilise pour produire Gag-Pol, le précurseur de ses enzymes virales. Le sous-type B du groupe M du VIH-1 étant celui qui est le plus répandu en Occident, nos recherches portent sur ce sous-type. Notre équipe a dans le passé caractérisé la structure du signal stimulateur de *frameshift*. Il s'agit d'une hélice à deux tiges interrompues par une excroissance de purines. Notre groupe a aussi démontré par mutagenèse que des perturbations de l'efficacité de *frameshift* aussi faibles que 30% handicapaient sévèrement la réplication du virus. À présent, nous voulons cibler l'événement du *frameshift* dans un objectif thérapeutique. Pour ce faire, nous nous sommes tournés vers de courts peptides riches en arginines aptes à interagir efficacement avec l'ARN afin de sélectionner des candidats qui pourraient interférer avec l'efficacité de *frameshift* en se liant au signal stimulateur, ou encore à des régions du ribosome impliquées dans le mécanisme du *frameshift* -1 programmé.

La librairie de peptides RSHN adaptée du groupe de Frankel (Harada et al., 1996) que nous utilisons a été insérée au site actif de la thiorédoxine de *E. coli*. Cette dernière agit donc en tant que protéine échafaud pour contraindre la structure des peptides générés et empêcher leur dégradation. Les peptides de 14 acides aminés de long sont formés d'arginine, de sérine, d'histidine et d'asparagine dans des proportions statistiques théoriques de 56,25%, 18,75%, 18,75% et 6,25%, respectivement. Le biais pour l'arginine

permet la formation de motifs riches en arginine retrouvés dans de nombreuses protéines liant l'ARN. Un système rapporteur à deux protéines fluorescentes a été développé. Dans ce système, la RFP est produite par traduction conventionnelle, alors que la séquence de *frameshift* du VIH-1 a été introduite au début de la séquence codante de la GFP, de sorte que seuls les ribosomes effectuant un *frameshift* -1 produisent la GFP. Le plasmide codant pour ce système rapporteur contient trois copies de la RFP. Il avait été montré lors de tests préliminaires que ce nombre était requis pour l'obtention d'un signal suffisamment intense de RFP dans les conditions de criblage de la librairie. Ceci est dû au fait qu'à la longueur d'onde d'excitation utilisée lors de mesures d'émission de fluorescence (488 nm), la GFP est excitée à 100% alors que la RFP ne l'est qu'à 40% et que la maturation de la RFP est plus lente que celle de la GFP (figure 12). Le rapport GFP/RFP mesuré nous reflète donc l'efficacité de *frameshift* dans les cultures, la RFP étant utilisée pour la normalisation des mesures de GFP.

Des bactéries *E. coli* Top10 préalablement transformées avec le système rapporteur ont été électroporées en présence du plasmide codant pour la librairie de peptides. Suite à une induction de 16 heures de l'expression des rapporteurs et de la librairie de peptides, les bactéries furent analysées par cytométrie de flux pour sélectionner les clones exprimant un peptide qui affecte l'efficacité de *frameshift*. Le promoteur utilisé pour contrôler l'expression des peptides est P_{BAD}, un promoteur inductible à l'arabinose. Ce promoteur a un patron d'expression « tout ou rien » (Khlebnikov et al., 2002). En l'absence d'arabinose, la répression du promoteur est totale, mais en présence d'arabinose, l'expression entraînée par le promoteur est maximale. Dans des conditions où l'induction des cultures cellulaires est faite avec une quantité suffisante d'arabinose (0,2% m/v) afin d'éviter des patrons d'expression hétérogènes d'une cellule à l'autre, ce promoteur a l'avantage d'assurer une expression maximale et uniforme des différents peptides. Les divers phénotypes mesurés en présence de peptides ne sont donc pas dus à des variations dans la concentration de peptides exprimés, mais à des variations dans l'activité des différentes séquences peptidiques.

Figure 12. Les spectres de fluorescence de la RFP et de la GFP. La protéine fluorescente rouge utilisée dans cette étude est la DsRed-T4, dérivée de la DsRed du corail *Discosoma genus*, alors que la verte est la GFP *Emerald* (GFPemd), dérivée de la GFP provenant de la méduse *Aequorea victoria*. La longueur d'onde d'excitation utilisée lors de la cytométrie de flux et des tests de fluorescence en plaque est de 488 nm, où la DsRed-T4 est excitée à 40% de son plein potentiel et la GFPemd à pratiquement 100%. Les longueurs d'onde d'émission lues sont 530 nm dans le cas de la GFPemd, ce qui correspond à environ 50% de l'intensité maximale, et 585 nm dans le cas de la DsRed, ce qui correspond à l'intensité maximale de fluorescence. Dans les deux graphiques, la courbe de gauche correspond au spectre d'excitation et celle de droite au spectre d'émission. L'axe des Y présente le pourcentage d'excitation ou d'émission selon la longueur d'onde par rapport à l'intensité maximale pouvant être obtenue. (tiré de Patterson et al., 2001)

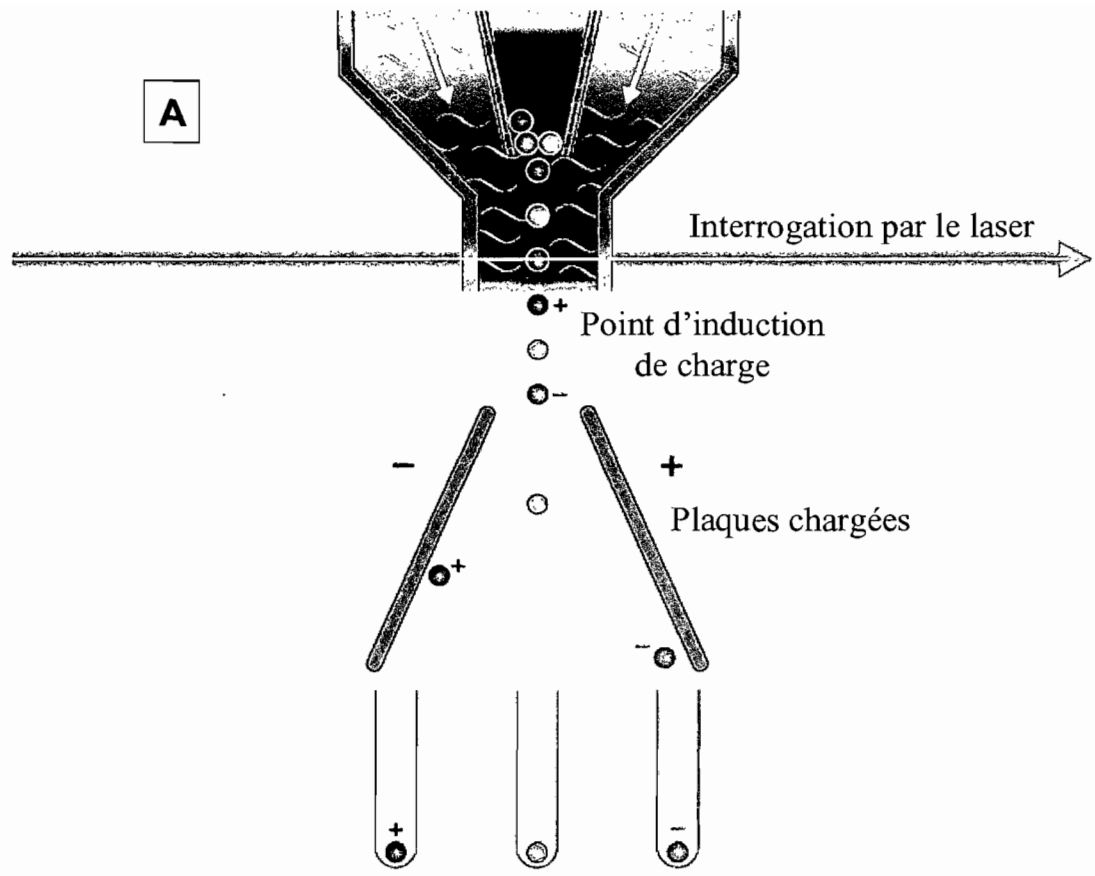


3.2 La cytométrie de flux comme outil pour le criblage de peptides affectant le frameshift.

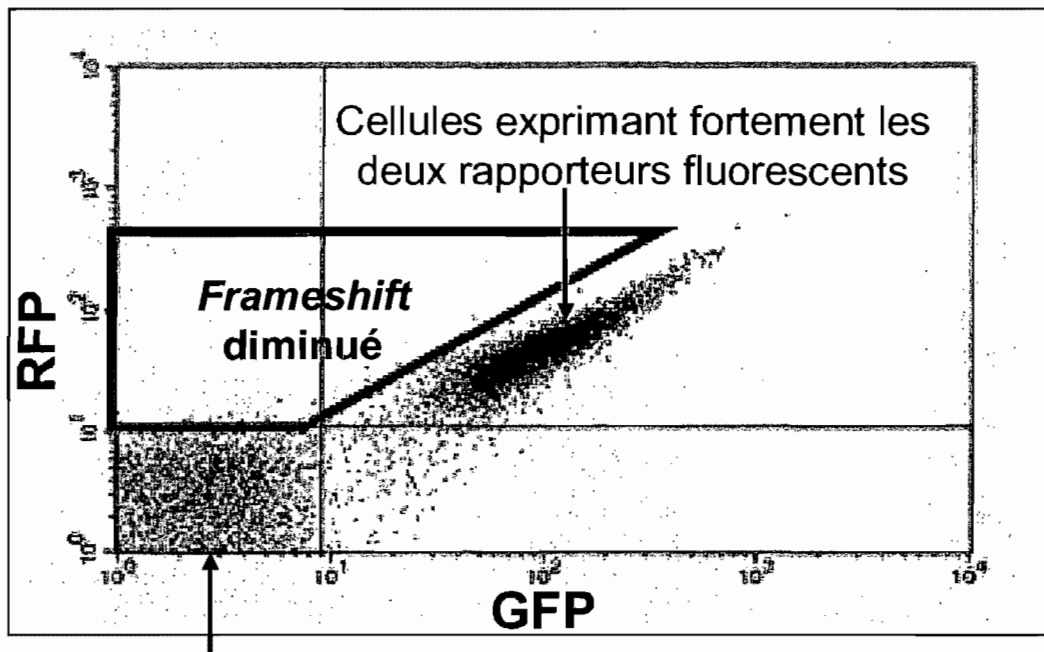
La librairie RSHN criblée possède une diversité théorique de séquences de $2,68 \times 10^8$. Suite à la ligation de la librairie à la thiorédoxine, le nombre de transformant se chiffrait à approximativement cinq millions, soit 1,9% de la diversité totale. Notre projet visait à valider la méthode de sélection plutôt que d'effectuer un crible exhaustif de la librairie. Nous avons donc jugé que cette proportion de transformants était suffisante. Afin de nous assurer que la totalité des séquences disponibles serait analysée par cytométrie de flux (figure 13A), 22 millions de clones exprimant un peptide ainsi que le système rapporteur RFP-GFP furent criblés. Des tests préliminaires ont démontré qu'une grande majorité des clones semblait réduire l'efficacité de *frameshift* alors qu'un très petit nombre de clones semblait l'augmenter. Nous avons alors décidé de ne sélectionner que des clones diminuant l'efficacité de *frameshift* (figure 13B). Quatre rondes successives de criblage ont démontré que la technique permettait un enrichissement significatif de la population de clones sélectionnés dans la même fenêtre. En effet, alors qu'à la ronde initiale seulement 0,4% des clones criblés répondaient aux critères de sélection, les trois rondes suivantes ont respectivement démontré une sélection de 7%, 44% et 69% des clones totaux. Toutefois, le séquençage d'un échantillon de clones a montré que l'enrichissement se faisait au détriment de la variété des séquences retrouvées et que les peptides retrouvés après quatre rondes d'enrichissement n'étaient pas plus actifs que ceux retrouvés en première ronde. Notons aussi que les peptides retrouvés après quatre rondes de cytométrie de flux ne sont pas plus riches en arginine que ceux sélectionnés après une seule ronde. Nous avons donc décidé d'analyser que les clones provenant d'une seule ronde de cytométrie de flux et les résultats que nous présentons correspondent à cette analyse.

Figure 13. La cytométrie de flux. La cytométrie de flux permet le tri de cellules démontrant différents patrons de fluorescence. Ce type d'application est communément appelé FACS (*fluorescence activated cell sorting*). **A)** L'échantillon de cellules à trier est d'abord concentré en un fil très fin de sorte à ne présenter qu'une cellule à la fois au rayonnement laser. L'étape où les cellules sont présentées au laser est l'interrogation. À ce moment, divers paramètres peuvent être déterminés. Le paramètre FSC (*forward scattering*) fournit de l'information sur la taille de la cellule alors que le paramètre SSC (*side scattering*) nous informe sur la granulosité du milieu intracellulaire. De plus, dans le cas où des fluorophores sont utilisés, divers patrons de fluorescence peuvent être obtenus selon les fluorophores et la longueur d'onde d'excitation du laser. Dans notre étude, le laser émet à une longueur d'onde de 488 nm et les mesures d'intensité de fluorescence émise sont prises à 530 nm pour la protéine fluorescente verte (GFP) et 588 nm pour la protéine fluorescente rouge (RFP) (figure 12). Suite à l'interrogation, selon les paramètres décidés à l'ordinateur pour que les cellules soient sélectionnées ou non, les cellules répondant aux critères de sélection sont chargées positivement ou négativement et acheminées latéralement dans un contenant de récolte. Les débris et cellules non-sélectionnées poursuivent leur chemin au centre vers une section recueillant les déchets (<http://www.abdirect.com/uploads/Flow-Cytometry.pdf>). **B)** Les cellules n'exprimant ni la RFP, ni la GFP constituent le bruit de fond et se retrouvent dans le cadran en bas à gauche. Les cellules qui expriment la RFP et qui expriment peu ou pas la GFP se retrouvent dans le cadran en haut à gauche. Inversement, les cellules où la GFP est fortement exprimée, mais pas la RFP se retrouvent dans le cadran en bas à droite. Finalement, les cellules où les deux rapporteurs sont fortement exprimés se retrouvent dans le cadran en haut à droite. Le quadrilatère rouge représente la fenêtre utilisée lors de la sélection de peptides diminuant le *frameshift*. Les bactéries sélectionnées expriment un peptide qui diminue l'expression de la GFP sans affecter celle de la RFP.

A



B



Cellules n'exprimant pas de rapporteurs fluorescents

3.3 Étude de l'activité des peptides sélectionnés sur l'efficacité du *frameshift* et sur la traduction conventionnelle.

Tel que mentionné dans la section précédente, 22 millions de bactéries furent criblées pour sélectionner les clones exprimant un peptide qui altère l'efficacité de *frameshift*. Au total, 45 000 des 22 millions (0,2%) d'événements analysés furent sélectionnés par cytométrie de flux. Ces clones récoltés furent étalés sur boîtes de Pétri, où seulement environ 5000 survécurent et formèrent des colonies. Il est probable que les autres clones exprimaient un peptide toxique à long terme pour la cellule ou que le stress causé par le passage en cytométrie de flux ait été fatal pour ces bactéries. Afin de cerner un échantillonnage de peptides perturbant le *frameshift*, 1150 colonies ont été repiquées et testées individuellement. Des mesures de fluorescence avec un lecteur de plaques Fusion™ α -FP (Packard) permirent d'identifier, à partir des 1150 clones testés, 160 clones qui diminuaient le plus l'efficacité de *frameshift*. Ces 160 peptides furent retransformés dans des bactéries *E. coli* Top10 possédant aussi le système rapporteur RFP-GFP afin de discriminer les clones où la variation dans le rapport GFP/RFP était due à un phénotype cellulaire plutôt qu'à l'action d'un peptide. Sur les 160 clones, 15 furent ainsi éliminés.

Les clones les plus actifs furent ensuite utilisés, parallèlement avec quelques clones non-sélectionnés, pour examiner s'ils pouvaient aussi interférer avec la traduction conventionnelle. Pour ce faire, dix des vingt clones les plus actifs et trois clones non sélectionnés furent aléatoirement choisis et soumis à des essais d'émission de fluorescence en fonction du temps de croissance des bactéries. Les résultats obtenus démontrent qu'en présence de peptides actifs, l'expression de la GFP est inhibée, mais qu'en présence d'un peptide non-sélectionné, l'expression de la GFP augmente de manière constante dans le temps, proportionnellement à la croissance de la culture bactérienne. Pour ce qui est de

l'expression de la RFP, les niveaux de fluorescence sont les mêmes autant avec un peptide sélectionné qu'avec un peptide non-sélectionné et ne sont pas perturbés. Ces résultats nous permettent de conclure que les peptides actifs agissant sur le *frameshift* n'ont pas d'effet délétère sur la traduction conventionnelle. Il est fort probable qu'un peptide agissant sur la traduction conventionnelle soit toxique et que les bactéries contenant de tels peptides n'aient pu survivre.

3.4 Analyse des séquences des peptides les plus actifs.

Les vingt peptides diminuant le plus l'efficacité de *frameshift* et trente peptides non-sélectionnés furent ensuite séquencés afin de comparer la distribution en acides aminés pour chacune des 14 positions. Nous avons privilégié la comparaison des clones sélectionnés avec un échantillonnage de clones non-sélectionnés plutôt que de simplement effectuer une comparaison avec la distribution théorique de R, S, H et N de la librairie afin de tenir compte de possibles biais de fabrication de la librairie lors de sa synthèse. Alors que, pour la plupart des positions, le séquençage de clones non-sélectionnés fournit des données correspondant à la distribution théorique, pour quelques positions comme 7 et 11, il semble y avoir un biais de fabrication en faveur de codons codant pour l'arginine. De plus, un autre biais fut décelé au codon codant pour le 11^e acide aminé, où la seconde base du codon qui ne devrait être que G ou A dans des proportions respectives de 75% et 25% contient plutôt des G à 50%, des A à 20% et des C à 30%, ce qui entraîne la génération de codons CCT (proline) et ACT (thréonine).

La comparaison des séquences des vingt clones les plus actifs avec celles de clones non-sélectionnés permet de cerner un consensus parmi les peptides, aux positions 7 à 10. En effet, alors que pour les positions 8 et 9, une arginine est quasiment toujours présente (dans 95% et 100% des cas respectivement), on observe une surreprésentation de sérines aux positions 7 et 10. Ce motif SRRS se retrouve d'ailleurs dans trois des vingt clones les plus

actifs, alors que six autres clones ont XRRS ou SRRX aux positions 7 à 10 (où X est R ou H). Les positions 7 à 10 se trouvant dans la portion centrale du peptide exposé, il est concevable que ces positions soient impliquées dans une interaction ARN-protéine et que le motif SRRS joue un rôle important dans cette interaction. Ce motif s'accorde d'ailleurs avec les tendances à occuper l'interface des acides aminés ainsi que des résidus flanquants discutées dans la section 1.5.1, où R a une forte tendance à être à l'interface et S une haute probabilité à être un résidu flanquant. Notons par ailleurs que les résidus d'asparagine aussi peu représentés dans les clones non-sélectionnés que dans les plus actifs, se retrouvent pratiquement tous (à 86%) dans les trois positions à chaque extrémité des peptides. Il est donc peu probable que cet acide aminé soit appelé à participer directement à une interaction ARN-protéine dans le contexte de ce criblage, et encore moins indirectement, vu sa tendance négligeable à être un résidu flanquant. Par ailleurs, il est possible de conclure qu'il y a une contre-sélection pour les histidines dans les séquences de peptides actifs. Cet acide aminé à haute tendance à être présent à l'interface lors d'interaction ARN-protéine se fait probablement voler ce rôle par les résidus d'arginine surreprésentés.

Nous avons aussi voulu savoir si l'activité des peptides sélectionnés pouvait être le résultat d'un effet non-spécifique de charge résultant des résidus arginine. L'examen des séquences de peptides non-sélectionnés nous a permis d'identifier deux séquences (RRRSRRRRRRRRRS et RHRSRRRRRRPHRH) particulièrement riches en arginine au niveau de leur portion centrale, portion où nous avons identifié le consensus « SRRS » chez les vingt peptides les plus actifs. Ces deux clones non-sélectionnés diminuent le *frameshift* dans des proportions de moins de 15%, bien moins que la diminution d'environ 50% obtenue avec les vingt peptides les plus actifs. Nous en déduisons que les peptides sélectionnés ne diminuent pas l'efficacité de *frameshift* à cause de leur charge positive qui interagirait de manière non-spécifique avec le ribosome.

3.5 Étude de la sélectivité des peptides sélectionnés pour le VIH-1 groupe M sous-type B.

La prochaine étape a consisté à déterminer si l'activité des peptides sélectionnés était sélective pour le sous-type B du groupe M du VIH-1. Pour ce faire, le système rapporteur RFP-GFP fut de nouveau utilisé. Quatre nouvelles constructions furent générées en substituant le signal stimulateur de *frameshift* du sous-type B par une structure secondaire d'ARN différente. Les structures secondaire utilisées furent la tige boucle TAR du virus de l'immunodéficience bovine (BIV), le RRE du VIH-1, l'IRE humain (*iron response element*), qui ne sont pas des signaux stimulateurs de *frameshift*, ou un autre signal stimulateur de *frameshift*, celui du groupe O du VIH-1, qui est un pseudonoeud (tableau 2). Même si TAR, RRE et IRE ne sont pas des signaux stimulateurs de *frameshift* naturels, leur présence conjointement avec la séquence glissante sauvage du sous-type B entraîne des taux de *frameshift* mesurables. Alors que l'efficacité de *frameshift* mesurée avec le système rapporteur fluorescent était de 3,7% avec le signal stimulateur sauvage du sous-type B, les constructions avec IRE, le signal stimulateur du groupe O, TAR et RRE présentèrent des efficacités de *frameshift* respectives de 5,9%, 4,6%, 3,4% et 1,8%. Les vingt peptides les plus actifs contre le sous-type B furent testés avec les nouveaux rapporteurs pour déterminer leur activité sur le *frameshift* en présence d'un signal stimulateur différent. Aucun des peptides ne démontra de spécificité pour le sous-type B du VIH-1. Globalement, les peptides diminuent l'efficacité de *frameshift* de tous les systèmes rapporteurs dans des proportions variables. De plus, plusieurs peptides ont des effets différents d'un signal stimulateur à l'autre, ce qui suggère que les peptides lient probablement différentes cibles ou la même cible, de manière différente. Il est donc peu probable, vu les divergences entre les structures agissant en tant que signal stimulateur, que les vingt peptides sélectionnés agissent sur le *frameshift* en se liant au signal stimulateur.

Tableau 2. L'activité de 20 peptides actifs sur des régions de frameshift avec différents signaux stimulateurs de frameshift. Quatre constructions dérivées du système rapporteur RFP-GFP furent générées en remplaçant le signal stimulateur de *frameshift* du sous-type B du groupe M du VIH-1 par une autre structure secondaire d'ARN. Les quatre structures remplaçantes sont l'IRE humain, le signal stimulateur du groupe O (un pseudonoeud), le RRE du VIH-1 et la tige-boucle TAR du VIB. Les valeurs indiquées sont les moyennes (%) de l'efficacité du *frameshift* mesurée en présence du peptide (où l'absence de peptide actif = 100% et où n = 3 à 5).

Tableau 2. Sélectivité de 20 peptides sélectionnés pour le sous-type B du VIH-1.

# Clones	Séquence du peptide	Signal stimulateur de <i>frameshift</i>				
		VIH-1 grM	Humain IRE	VIH-1 grO	VIH-1 RRE	VIB TAR
1	NHRRHSRRSSRRS	53	55	52	67	57
2	RRRHRSRRRRRS	66	80	74	59	63
3	RRRSRRRRRPRNS	51	63	57	54	54
4	RRNRRRSRRSRRH	52	66	62	56	53
5	RHHSRRRRRSRRRR	53	63	60	54	51
6	RHSRSRSRRRPRRS	64	78	71	67	59
7	RRRSRHRRRSRRHR	65	77	78	71	64
8	SRHSRRHRRRRRRR	52	58	55	66	56
9	HHRRRRRRRRSSRS	59	57	53	63	53
10	RRRRRRRRRRRTSSR	55	61	61	62	54
11	NRRRRHRRRRSRRS	57	63	64	54	57
12	RRRRHSRRRPRRR	53	57	63	55	52
13	RRSSNRRRRSRRRR	59	73	68	57	62
14	RSHRRRRRRRRSNR	51	48	69	61	64
15	SRRRSRRRRRPRRR	58	64	74	60	48
16	SRRRRSRRSRRRR	66	85	64	60	59
17	HRRRRRRRRRSRHH	51	56	71	64	49
18	HRRRRRRRRHRRRR	60	63	76	65	60
19	RRRRRRRRRSRNRR	54	83	69	74	59
20	RRSRRRRRRRPRRH	58	64	56	62	59
moyenne		57	66	65	62	57

Une cible plus crédible pour les peptides serait le ribosome. Toutefois, comme il a été montré précédemment, les peptides affectant l'efficacité de *frameshift* n'altèrent pas la traduction conventionnelle. L'interaction des peptides avec le ribosome devrait n'affecter que le *frameshift*. Une première hypothèse serait que les peptides ciblent l'activité hélicase du ribosome. Celle-ci s'exerce à l'entrée du tunnel des ARNm du ribosome. Une étude chez *E. coli* démontre que les protéines ribosomiques S3, S4 et S5 projettent des acides aminés vers l'entrée de ce tunnel et suggère que ces trois protéines interviennent dans l'activité hélicase. Ce serait le mouvement de la tête par rapport au corps du ribosome durant la traduction qui fournirait l'énergie mécanique suffisante pour rompre les interactions formant la structure secondaire dans l'ARNm en éloignant S3 et S4 l'un de l'autre (Yusupova et al., 2001). Les peptides actifs contre le *frameshift* pourraient, de manière allostérique en interagissant avec des régions spécifiques de l'ARNr, augmenter l'activité hélicase du ribosome. Cet effet permettrait un déroulement plus rapide du signal stimulateur. La pause du ribosome sur la séquence glissante serait donc raccourcie, ce qui diminuerait la probabilité qu'un *frameshift* -1 programmé se produise. Le *frameshift* ribosomique dépend d'une manière critique de cette pause, alors que les pauses au niveau de structures secondaires n'affectent guère la traduction conventionnelle, l'initiation étant l'étape limitante de la cinétique traductionnelle (Laursen et al., 2005).

Une seconde hypothèse est que la liaison des peptides pourrait influencer spécifiquement le glissement des ARNt associés au *frameshift*. Cette hypothèse se réfère au mécanisme du *frameshift* de Léger *et al.* 2007 tel que présenté dans la section 1.4.2.2. Ce mécanisme implique que, lors de la translocation, les ARNt présents sont coincés dans un état de transition dû à la résistance au déroulement du signal stimulateur. Dans cette situation, le *frameshift* -1 serait promu par la propension des ARNt à quitter leur état de transition pour occuper des sites classiques du ribosome après s'être détachés du messager. Les peptides actifs pourraient, toujours en se liant à certaines régions de l'ARNr, stabiliser

les ARNt impliqués dans le *frameshift* dans leur état de transition et diminuer leur propension à effectuer un changement de cadre de lecture pour occuper des sites classiques. Ceci pourrait accorder davantage de temps à l'activité hélicase du ribosome pour dérouler le signal stimulateur de *frameshift*, les ARNt pourraient alors occuper les sites classiques en entraînant le messager et poursuivre la traduction dans le cadre 0.

3.6 Étude de l'effet des peptides sélectionnés chez les bactéries dans un système eucaryote.

Dix peptides furent choisis aléatoirement parmi les vingt peptides sélectionnés comme les plus actifs pour examiner leur effet sur le *frameshift* chez les eucaryotes. Nous avons choisi d'étudier l'activité des peptides dans des cellules HEK293T en culture. Les dix clones furent exprimés dans un échafaud différent afin d'exclure toute dépendance vis à vis de la thiorédoxine en tant qu'échafaud. L'échafaud employé chez les eucaryotes consistait en deux « bras » flanquants d'une quinzaine d'acides aminés aptes à dimériser ensemble. De la sorte, les peptides sont exposés dans une boucle. Un système rapporteur à deux luciférase fut utilisé, ce qui élimine les peptides qui auraient pu agir sur l'expression ou l'activité de la GFP sans affecter l'efficacité de *frameshift*. Dans ce système rapporteur, la luciférase de *Renilla* et la luciférase de la luciole sont encodées par le même ARNm. La région de *frameshift* du VIH-1 a été insérée entre les deux de sorte que la luciférase de *Renilla* soit produite par traduction conventionnelle alors que seuls les ribosomes qui effectuent un *frameshift* -1 programmé synthétisent la luciférase de luciole (Grentzmann et al., 1998).

Parmi les dix clones testés chez les eucaryotes, deux diminuent l'efficacité de *frameshift* d'environ 35 à 40%. La perte d'activité chez les eucaryotes d'une partie des peptides sélectionnés chez les bactéries pourrait être attribuable au changement d'échafaud,

mais surtout aux différences entre le ribosome bactérien et le ribosome eucaryote en admettant que la cible des peptides est le ribosome. En effet, même s'il y a une grande conservation entre les ribosomes procaryotes et eucaryotes, plusieurs protéines ribosomiques présentes seulement chez les eucaryotes (Spahn et al., 2004) pourraient masquer des sites de l'ARNr avec lesquels les peptides actifs interagissent.

Nous avons aussi examiné si l'expression des deux peptides actifs chez les eucaryotes sous forme flottante plutôt que contrainte dans un échafaud pouvait affecter leur activité. Pour ce faire, nous avons introduit un codon *stop* immédiatement en aval de la séquence codant pour les peptides actifs, qui sont alors exprimé à l'extrémité du bras en 5' et non dans une boucle. Dans ces conditions, aucun des deux peptides ne démontra d'activité sur l'efficacité de *frameshift*, ce qui suggère que les contraintes structurales imposées par l'échafaud sont favorables à l'interaction des peptides avec leur cible. Alternativement, en absence d'échafaud, ces peptides non-structurés sont peut-être très rapidement dégradés.

Conclusion et perspectives

Le but de notre projet était de sélectionner des peptides perturbant l'efficacité de *frameshift* ribosomique -1 programmé du sous-type B du groupe M du VIH-1 à partir d'une librairie combinatoire de peptides codant pour des peptides de 14 acides aminés composés d'arginine, de sérine, d'histidine et d'asparagine avec un biais pour l'arginine. La sélection des peptides était effectuée par cytométrie de flux dans des bactéries *E. coli* Top10 avec un système rapporteur utilisant la GFP et la RFP. Notre étude a démontré que l'analyse des clones obtenus après une ronde de FACS était préférable à l'analyse des clones recueillis après de multiples rondes de FACS, afin de conserver un maximum de séquences actives. Les clones sélectionnés comme les plus actifs diminuent l'efficacité de *frameshift* du VIH-1 de plus de 50% chez *E. coli*.

Il a été démontré que les vingt clones les plus actifs diminuaient l'efficacité de *frameshift* sans altérer la traduction conventionnelle. Une séquence consensus « SRRS » aux positions 7 à 10 de ces vingt clones a été observée lors de l'analyse des séquences de peptides diminuant le *frameshift*. L'utilisation de quatre dérivés du système rapporteur RFP-GFP où le signal stimulateur de *frameshift* du sous-type B du groupe M du VIH-1 était remplacé par une autre structure secondaire a permis de conclure qu'aucun des vingt peptides sélectionnés les plus actifs n'était spécifique pour ce signal stimulateur. Étant donné la grande divergence entre les structures secondaires agissant en tant que signal stimulateur de *frameshift*, il est peu probable que les peptides se lient au signal. Nous proposons que les peptides actifs interagissent avec le ribosome et affectent des fonctions qui n'influencent pas la traduction conventionnelle, mais qui sont critiques pour le *frameshift* comme décrit dans la section 3.5.

La transposition dans un système eucaryote de dix des vingt clones les plus actifs chez les bactéries permit d'identifier deux séquences diminuant l'efficacité de *frameshift* de 35 à 40%. Donc, 20% de clones sélectionnés dans un système bactérien qui conservent leur

activité chez les eucaryotes. Cette proportion peut être considérée comme raisonnable dans la mesure où nous avons pu bénéficier lors du processus de sélection de la rapidité, la facilité et des faibles coûts associés à un système bactérien. De plus, puisque nous n'avons criblé qu'environ 3% de la diversité totale de la librairie, nous pouvons supposer qu'un criblage exhaustif de la librairie RSHN pourrait permettre de sélectionner jusqu'à soixante clones actifs chez les eucaryotes. Vu le changement d'échafaud protéique et de système rapporteur lors de la transposition des clones dans le système eucaryote, nous pouvons exclure que l'échafaud joue un rôle crucial dans l'activité des deux clones qui ont conservé leur activité contre le *frameshift*.

Notre étude rapporte, à notre connaissance, le premier cas de peptides ciblant l'ARN criblés dans un système procaryote qui conservent leur activité chez les eucaryotes. Nous comptons, dans un futur rapproché, tester dans des cellules eucaryotes l'activité sur le *frameshift* des séquences peptidiques actives synthétiques fusionnées au domaine de transduction de Tat. Il sera ensuite important de vérifier l'activité de ces peptides sur la réplication virale. S'ils s'avèrent efficaces pour handicaper la réplication virale, il faudra aussi examiner si cet effet est dû à l'action des peptides sur le *frameshift* en mesurant les concentrations de RT (du précurseur Pol) et de CA (du précurseur Gag) de cellules infectées en présence de peptides actifs. Si les résultats s'avéraient positifs, ces peptides pourraient servir de base dans le développement de nouveaux agents anti-rétroviraux ciblant le *frameshift* ribosomique -1 programmé.

L'analyse approfondie du mécanisme d'action des peptides actifs en identifiant leur site d'interaction avec le ribosome, par exemple par pontage, pourra permettre d'explorer plus en profondeur le mécanisme du *frameshift*. L'activité hélicase du ribosome pourra être mesurée en présence des peptides actifs (selon le procédé décrit dans Takyar et al., 2005) afin de déterminer si les peptides altèrent cette fonction du ribosome. Une autre avenue intéressante pour poursuivre ce projet à long terme serait le crible d'une librairie de peptides plus diversifiée dans le choix des acides aminés, telle que celle utilisée par Xu et

al., 2001 (section 1.5.3.1). Ceci pourrait permettre l'identification d'autres peptides perturbant l'efficacité de *frameshift* en liant le signal stimulateur de *frameshift* ou d'autres régions du ribosome.

Bibliographie

- Aldovini A, Young RA. 1990. Mutations of RNA and protein sequences involved in human immunodeficiency virus type 1 packaging result in production of noninfectious virus. *J Virol* **64**:1920-1926.
- Ashe MP, Pearson LH, Proudfoot NJ. 1997. The HIV-1 5' LTR poly(A) site is inactivated by U1 snRNP interaction with the downstream major splice donor site. *Embo J* **16**:5752-5763.
- Ball RA, Kinchelov T. 2003. Injection site reactions with the HIV-1 fusion inhibitor enfuvirtide. *J Am Acad Dermatol* **49**:826-831.
- Bannwarth L, Reboud-Ravaux M. 2007. An alternative strategy for inhibiting multidrug-resistant mutants of the dimeric HIV-1 protease by targeting the subunit interface. *Biochem Soc Trans* **35**:551-554.
- Baranov PV, Gesteland RF, Atkins JF. 2002. Recoding: translational bifurcations in gene expression. *Gene* **286**:187-201.
- Bartlett JA, DeMasi R, Quinn J, Moxham C, Rousseau F. 2001. Overview of the effectiveness of triple combination therapy in antiretroviral-naive HIV-1 infected adults. *Aids* **15**:1369-1377.
- Bertrand C, Prere MF, Gesteland RF, Atkins JF, Fayet O. 2002. Influence of the stacking potential of the base 3' of tandem shift codons on -1 ribosomal frameshifting used for gene expression. *Rna* **8**:16-28.
- Blum JH, Dove SL, Hochschild A, Mekalanos JJ. 2000. Isolation of peptide aptamers that inhibit intracellular processes. *Proc Natl Acad Sci U S A* **97**:2241-2246.
- Boggetto N, Reboud-Ravaux M. 2002. Dimerization inhibitors of HIV-1 protease. *Biol Chem* **383**:1321-1324.
- Brasey A, Lopez-Lastra M, Ohlmann T, Beerens N, Berkhout B, Darlix JL, Sonenberg N. 2003. The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. *J Virol* **77**:3939-3949.
- Brierley I, Jenner AJ, Inglis SC. 1992. Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. *J Mol Biol* **227**:463-479.
- Brinkman K, ter Hofstede HJ, Burger DM, Smeitink JA, Koopmans PP. 1998. Adverse effects of reverse transcriptase inhibitors: mitochondrial toxicity as common pathway. *Aids* **12**:1735-1744.
- Brunelle MN, Payant C, Lemay G, Brakier-Gingras L. 1999. Expression of the human immunodeficiency virus frameshift signal in a bacterial cell-free system: influence of an interaction between the ribosome and a stem-loop structure downstream from the slippery site. *Nucleic Acids Res* **27**:4783-4791.

- Buonaguro L, Tornesello ML, Buonaguro FM. 2007. Human immunodeficiency virus type 1 subtype distribution in the worldwide epidemic: pathogenetic and therapeutic implications. *J Virol* **81**:10209-10219.
- Burke DS. 1997. Recombination in HIV: an important viral evolutionary strategy. *Emerg Infect Dis* **3**:253-259.
- Clark MB, Janicke M, Gottesbuhren U, Kleffmann T, Legge M, Poole ES, Tate WP. 2007. A mammalian gene *peg10* expresses two reading frames by high efficiency -1 frameshifting in embryonic-associated tissues. *J Biol Chem*.
- Cullen BR. 1998. HIV-1 auxiliary proteins: making connections in a dying cell. *Cell* **93**:685-692.
- Dinman JD, Richter S, Plant EP, Taylor RC, Hammell AB, Rana TM. 2002. The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc Natl Acad Sci U S A* **99**:5331-5336.
- Dinman JD, Ruiz-Echevarria MJ, Czaplinski K, Peltz SW. 1997. Peptidyl-transferase inhibitors have antiviral properties by altering programmed -1 ribosomal frameshifting efficiencies: development of model systems. *Proc Natl Acad Sci U S A* **94**:6606-6611.
- Draper DE. 1999. Themes in RNA-protein recognition. *J Mol Biol* **293**:255-270.
- Dulude D, Baril M, Brakier-Gingras L. 2002. Characterization of the frameshift stimulatory signal controlling a programmed -1 ribosomal frameshift in the human immunodeficiency virus type 1. *Nucleic Acids Res* **30**:5094-5102.
- Dulude D, Berchiche YA, Gendron K, Brakier-Gingras L, Heveker N. 2006. Decreasing the frameshift efficiency translates into an equivalent reduction of the replication of the human immunodeficiency virus type 1. *Virology* **345**:127-136.
- Endrich MM, Gehrig P, Gehring H. 1999. Maturation-induced conformational changes of HIV-1 capsid protein and identification of two high affinity sites for cyclophilins in the C-terminal domain. *J Biol Chem* **274**:5326-5332.
- Ernest CS, 2nd, Hall SD, Jones DR. 2005. Mechanism-based inactivation of CYP3A by HIV protease inhibitors. *J Pharmacol Exp Ther* **312**:583-591.
- Fadel H, Temesgen Z. 2007. Maraviroc. *Drugs Today (Barc)* **43**:749-758.
- Faivre S, Chieze S, Delbaldo C, Ady-Vago N, Guzman C, Lopez-Lazaro L, Lozahic S, Jimeno J, Pico F, Armand JP, Martin JA, Raymond E. 2005. Phase I and pharmacokinetic study of aplidine, a new marine cyclodepsipeptide in patients with advanced malignancies. *J Clin Oncol* **23**:7871-7880.
- Farabaugh PJ. 1996. Programmed translational frameshifting. *Annu Rev Genet* **30**:507-528.
- Franke EK, Yuan HE, Bossolt KL, Goff SP, Luban J. 1994a. Specificity and sequence requirements for interactions between various retroviral Gag proteins. *J Virol* **68**:5300-5305.
- Franke EK, Yuan HE, Luban J. 1994b. Specific incorporation of cyclophilin A into HIV-1 virions. *Nature* **372**:359-362.
- Frankel AD, Young JA. 1998. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem* **67**:1-25.
- Freed EO. 2001. HIV-1 replication. *Somat Cell Mol Genet* **26**:13-33.

- Freed EO, Mouland AJ. 2006. The cell biology of HIV-1 and other retroviruses. *Retrovirology* **3**:77.
- Gaudin C, Mazauric MH, Traikia M, Guittet E, Yoshizawa S, Fourmy D. 2005. Structure of the RNA signal essential for translational frameshifting in HIV-1. *J Mol Biol* **349**:1024-1035.
- Gendron K, Charbonneau J, Dulude D, Heveker N, Ferbeyre G, Brakier-Gingras L. 2007. The presence of the TAR RNA structure alters the programmed -1 ribosomal frameshift efficiency of the human immunodeficiency virus type 1 (HIV-1) by modifying the rate of translation initiation. *Nucleic Acids Res*.
- Gilbert PB, McKeague IW, Eisen G, Mullins C, Gueye NA, Mboup S, Kanki PJ. 2003. Comparison of HIV-1 and HIV-2 infectivity from a prospective cohort study in Senegal. *Stat Med* **22**:573-593.
- Gottlinger HG. 2001. The HIV-1 assembly machine. *Aids* **15 Suppl 5**:S13-20.
- Gottlinger HG, Dorfman T, Sodroski JG, Haseltine WA. 1991. Effect of mutations affecting the p6 gag protein on human immunodeficiency virus particle release. *Proc Natl Acad Sci U S A* **88**:3195-3199.
- Greenway AL, Holloway G, McPhee DA, Ellis P, Cornall A, Lidman M. 2003. HIV-1 Nef control of cell signalling molecules: multiple strategies to promote virus replication. *J Biosci* **28**:323-335.
- Grentzmann G, Ingram JA, Kelly PJ, Gesteland RF, Atkins JF. 1998. A dual-luciferase reporter system for studying recoding signals. *Rna* **4**:479-486.
- Grollman AP. 1967. Inhibitors of protein biosynthesis. II. Mode of action of anisomycin. *J Biol Chem* **242**:3226-3233.
- Harada K, Martin SS, Frankel AD. 1996. Selection of RNA-binding peptides in vivo. *Nature* **380**:175-179.
- Harger JW, Meskauskas A, Dinman JD. 2002. An "integrated model" of programmed ribosomal frameshifting. *Trends Biochem Sci* **27**:448-454.
- Hill MK, Bellamy-McIntyre A, Vella LJ, Campbell SM, Marshall JA, Tachedjian G, Mak J. 2007. Alteration of the proline at position 7 of the HIV-1 spacer peptide p1 suppresses viral infectivity in a strain dependent manner. *Curr HIV Res* **5**:69-78.
- Hill MK, Shehu-Xhilaga M, Crowe SM, Mak J. 2002. Proline residues within spacer peptide p1 are important for human immunodeficiency virus type 1 infectivity, protein processing, and genomic RNA dimer stability. *J Virol* **76**:11245-11253.
- Ho A, Schwarze SR, Mermelstein SJ, Waksman G, Dowdy SF. 2001. Synthetic protein transduction domains: enhanced transduction potential in vitro and in vivo. *Cancer Res* **61**:474-477.
- Huang M, Orenstein JM, Martin MA, Freed EO. 1995. p6Gag is required for particle production from full-length human immunodeficiency virus type 1 molecular clones expressing protease. *J Virol* **69**:6810-6818.
- Hung M, Patel P, Davis S, Green SR. 1998. Importance of ribosomal frameshifting for human immunodeficiency virus type 1 particle assembly and replication. *J Virol* **72**:4819-4824.
- Ikezoe T, Saito T, Bandobashi K, Yang Y, Koeffler HP, Taguchi H. 2004. HIV-1 protease inhibitor induces growth arrest and apoptosis of human multiple myeloma cells via

- inactivation of signal transducer and activator of transcription 3 and extracellular signal-regulated kinase 1/2. *Mol Cancer Ther* **3**:473-479.
- Ivanov IP, Atkins JF. 2007. Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans: close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res* **35**:1842-1858.
- Jacks T, Power MD, Masiarz FR, Luciw PA, Barr PJ, Varmus HE. 1988. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**:280-283.
- Jacque JM, Triques K, Stevenson M. 2002. Modulation of HIV-1 replication by RNA interference. *Nature* **418**:435-438.
- Jouvenet N, Neil SJ, Bess C, Johnson MC, Virgen CA, Simon SM, Bieniasz PD. 2006. Plasma membrane is the site of productive HIV-1 particle assembly. *PLoS Biol* **4**:e435.
- Kapp LD, Lorsch JR. 2004. The molecular mechanics of eukaryotic translation. *Annu Rev Biochem* **73**:657-704.
- Kappelhoff BS, Huitema AD, Beijnen JH. 2005. Should non-nucleoside reverse transcriptase inhibitors be combined? *Drugs R D* **6**:61-69.
- Khlebnikov A, Skaug T, Keasling JD. 2002. Modulation of gene expression from the arabinose-inducible araBAD promoter. *J Ind Microbiol Biotechnol* **29**:34-37.
- Kondo E, Gottlinger HG. 1996. A conserved LXXLF sequence is the major determinant in p6gag required for the incorporation of human immunodeficiency virus type 1 Vpr. *J Virol* **70**:159-164.
- Kontos H, Naphthine S, Brierley I. 2001. Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. *Mol Cell Biol* **21**:8657-8670.
- Krausslich HG. 1991. Human immunodeficiency virus proteinase dimer as component of the viral polyprotein prevents particle assembly and viral infectivity. *Proc Natl Acad Sci USA* **88**:3213-3217.
- Laurent C, Kouanfack C, Koulla-Shiro S, Nkoue N, Bourgeois A, Calmy A, Lactuock B, Nzeusseu V, Mougnotou R, Peytavin G, Liegeois F, Nerrienet E, Tardy M, Peeters M, Andrieux-Meyer I, Zekeng L, Kazatchkine M, Mpoudi-Ngole E, Delaporte E. 2004. Effectiveness and safety of a generic fixed-dose combination of nevirapine, stavudine, and lamivudine in HIV-1-infected adults in Cameroon: open-label multicentre trial. *Lancet* **364**:29-34.
- Laursen BS, Sorensen HP, Mortensen KK, Sperling-Petersen HU. 2005. Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev* **69**:101-123.
- Leger M, Dulude D, Steinberg SV, Brakier-Gingras L. 2007. The three transfer RNAs occupying the A, P and E sites on the ribosome are involved in viral programmed -1 ribosomal frameshift. *Nucleic Acids Res* **35**:5581-5592.
- Leger M, Sidani S, Brakier-Gingras L. 2004. A reassessment of the response of the bacterial ribosome to the frameshift stimulatory signal of the human immunodeficiency virus type 1. *Rna* **10**:1225-1235.
- Li F, Goila-Gaur R, Salzwedel K, Kilgore NR, Reddick M, Matallana C, Castillo A, Zoumplis D, Martin DE, Orenstein JM, Allaway GP, Freed EO, Wild CT. 2003.

- PA-457: a potent HIV inhibitor that disrupts core condensation by targeting a late step in Gag processing. *Proc Natl Acad Sci U S A* **100**:13555-13560.
- Luban J, Goff SP. 1994. Mutational analysis of cis-acting packaging signals in human immunodeficiency virus type 1 RNA. *J Virol* **68**:3784-3793.
- Mak J, Jiang M, Wainberg MA, Hammarskjold ML, Rekosh D, Kleiman L. 1994. Role of Pr160gag-pol in mediating the selective incorporation of tRNA(Lys) into human immunodeficiency virus type 1 particles. *J Virol* **68**:2065-2072.
- Malim MH. 2006. Natural resistance to HIV infection: The Vif-APOBEC interaction. *C R Biol* **329**:871-875.
- Manktelow E, Shigemoto K, Brierley I. 2005. Characterization of the frameshift signal of Edr, a mammalian example of programmed -1 ribosomal frameshifting. *Nucleic Acids Res* **33**:1553-1563.
- Mathe G. 1997. Hypothetical reasons of the HIV1-AIDS "tritherapy" failure. A challenging model. *Biomed Pharmacother* **51**:413-416.
- McNaughton BR, Gareiss PC, Miller BL. 2007. Identification of a selective small-molecule ligand for HIV-1 frameshift-inducing stem-loop RNA from an 11,325 member resin bound dynamic combinatorial library. *J Am Chem Soc* **129**:11306-11307.
- Miller JH, Presnyak V, Smith HC. 2007. The dimerization domain of HIV-1 viral infectivity factor Vif is required to block APOBEC3G incorporation with virions. *Retrovirology* **4**:81.
- Nair V, Chi G. 2007. HIV integrase inhibitors as therapeutic agents in AIDS. *Rev Med Virol* **17**:277-295.
- Namy O, Moran SJ, Stuart DI, Gilbert RJ, Brierley I. 2006. A mechanical explanation of RNA pseudoknot function in programmed ribosomal frameshifting. *Nature* **441**:244-247.
- Namy O, Rousset JP, Napthine S, Brierley I. 2004. Reprogrammed genetic decoding in cellular gene expression. *Mol Cell* **13**:157-168.
- Navia MA, McKeever BM. 1990. A role for the aspartyl protease from the human immunodeficiency virus type 1 (HIV-1) in the orchestration of virus assembly. *Ann N Y Acad Sci* **616**:73-85.
- Nekhai S, Zhou M, Fernandez A, Lane WS, Lamb NJ, Brady J, Kumar A. 2002. HIV-1 Tat-associated RNA polymerase C-terminal domain kinase, CDK2, phosphorylates CDK7 and stimulates Tat-mediated transcription. *Biochem J* **364**:649-657.
- Norman TC, Smith DL, Sorger PK, Drees BL, O'Rourke SM, Hughes TR, Roberts CJ, Friend SH, Fields S, Murray AW. 1999. Genetic selection of peptide inhibitors of biological pathways. *Science* **285**:591-595.
- Okabe H, Satoh S, Furukawa Y, Kato T, Hasegawa S, Nakajima Y, Yamaoka Y, Nakamura Y. 2003. Involvement of PEG10 in human hepatocellular carcinogenesis through interaction with SIAH1. *Cancer Res* **63**:3043-3048.
- Orr GA, Verdier-Pinard P, McDaid H, Horwitz SB. 2003. Mechanisms of Taxol resistance related to microtubules. *Oncogene* **22**:7280-7295.
- Patterson G, Day RN, Piston D. 2001. Fluorescent protein spectra. *J Cell Sci* **114**:837-838.
- Patch JA, Barron AE. 2002. Mimicry of bioactive peptides via non-natural, sequence-specific peptidomimetic oligomers. *Curr Opin Chem Biol* **6**:872-877.

- Peled-Zehavi H, Horiya S, Das C, Harada K, Frankel AD. 2003. Selection of RRE RNA binding peptides using a kanamycin antitermination assay. *Rna* **9**:252-261.
- Pettit SC, Moody MD, Wehbie RS, Kaplan AH, Nantermet PV, Klein CA, Swanstrom R. 1994. The p2 domain of human immunodeficiency virus type 1 Gag regulates sequential proteolytic processing and is required to produce fully infectious virions. *J Virol* **68**:8017-8027.
- Ramakrishnan V. 2002. Ribosome structure and the mechanism of translation. *Cell* **108**:557-572.
- Ranganathan S, Kern PA. 2002. The HIV protease inhibitor saquinavir impairs lipid metabolism and glucose transport in cultured adipocytes. *J Endocrinol* **172**:155-162.
- Reil H, Kollmus H, Weidle UH, Hauser H. 1993. A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J Virol* **67**:5579-5584.
- Robertson DL, Anderson JP, Bradac JA, Carr JK, Foley B, Funkhouser RK, Gao F, Hahn BH, Kalish ML, Kuiken C, Learn GH, Leitner T, McCutchan F, Osmanov S, Peeters M, Pieniazek D, Salminen M, Sharp PM, Wolinsky S, Korber B. 2000. HIV-1 nomenclature proposal. *Science* **288**:55-56.
- Rodnina MV, Wintermeyer W. 2001. Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem* **70**:415-435.
- Schutt M, Zhou J, Meier M, Klein HH. 2004. Long-term effects of HIV-1 protease inhibitors on insulin secretion and insulin signaling in INS-1 beta cells. *J Endocrinol* **183**:445-454.
- Schwarze SR, Dowdy SF. 2000. In vivo protein transduction: intracellular delivery of biologically active proteins, compounds and DNA. *Trends Pharmacol Sci* **21**:45-48.
- Shehu-Xhilaga M, Crowe SM, Mak J. 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. *J Virol* **75**:1834-1841.
- Shehu-Xhilaga M, Hill M, Marshall JA, Kappes J, Crowe SM, Mak J. 2002a. The conformation of the mature dimeric human immunodeficiency virus type 1 RNA genome requires packaging of pol protein. *J Virol* **76**:4331-4340.
- Shehu-Xhilaga M, Lee JY, Campbell S, Marshall JA, Crowe SM, Mak J. 2002b. Overexpression and incorporation of GagPol precursor does not impede packaging of HIV-1 tRNA(Lys3) but promotes intracellular budding of virus-like particles. *J Biomed Sci* **9**:697-705.
- Sheng N, Pettit SC, Tritch RJ, Ozturk DH, Rayner MM, Swanstrom R, Erickson-Viitanen S. 1997. Determinants of the human immunodeficiency virus type 1 p15NC-RNA interaction that affect enhanced cleavage by the viral protease. *J Virol* **71**:5723-5732.
- Shigemoto K, Brennan J, Walls E, Watson CJ, Stott D, Rigby PW, Reith AD. 2001. Identification and characterisation of a developmentally regulated mammalian gene that utilises -1 programmed ribosomal frameshifting. *Nucleic Acids Res* **29**:4079-4088.
- Skerra A. 2000. Engineered protein scaffolds for molecular recognition. *J Mol Recognit* **13**:167-187.

- Southworth DR, Green R. 2003. Ribosomal translocation: sparsomycin pushes the button. *Curr Biol* **13**:R652-654.
- Spahn CM, Jan E, Mulder A, Grassucci RA, Sarnow P, Frank J. 2004. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. *Cell* **118**:465-475.
- Staple DW, Butcher SE. 2005. Solution structure and thermodynamic investigation of the HIV-1 frameshift inducing element. *J Mol Biol* **349**:1011-1023.
- Takyar S, Hickerson RP, Noller HF. 2005. mRNA helicase activity of the ribosome. *Cell* **120**:49-58.
- Tamilarasu N, Huq I, Rana TM. 2001. Targeting RNA with peptidomimetic oligomers in human cells. *Bioorg Med Chem Lett* **11**:505-507.
- Tang C, Loeliger E, Luncsford P, Kinde I, Beckett D, Summers MF. 2004. Entropic switch regulates myristate exposure in the HIV-1 matrix protein. *Proc Natl Acad Sci U S A* **101**:517-522.
- Tang H, Kuhen KL, Wong-Staal F. 1999. Lentivirus replication and regulation. *Annu Rev Genet* **33**:133-170.
- Temesgen Z, Warnke D, Kasten MJ. 2006. Current status of antiretroviral therapy. *Expert Opin Pharmacother* **7**:1541-1554.
- Terribilini M, Lee JH, Yan C, Jernigan RL, Honavar V, Dobbs D. 2006. Prediction of RNA binding sites in proteins from amino acid sequence. *Rna* **12**:1450-1462.
- Tsuchihashi Z, Brown PO. 1992. Sequence requirements for efficient translational frameshifting in the Escherichia coli dnaX gene and the role of an unstable interaction between tRNA(Lys) and an AAG lysine codon. *Genes Dev* **6**:511-519.
- Turner JJ, Fabani M, Arzumánov AA, Ivanova G, Gait MJ. 2006. Targeting the HIV-1 RNA leader sequence with synthetic oligonucleotides and siRNA: chemistry and cell delivery. *Biochim Biophys Acta* **1758**:290-300.
- Vega S, Kang LW, Velazquez-Campoy A, Kiso Y, Amzel LM, Freire E. 2004. A structural and thermodynamic escape mechanism from a drug resistant mutation of the HIV-1 protease. *Proteins* **55**:594-602.
- Weiss MA, Narayana N. 1998. RNA recognition by arginine-rich peptide motifs. *Biopolymers* **48**:167-180.
- Weiss RB, Dunn DM, Dahlberg AE, Atkins JF, Gesteland RF. 1988. Reading frame switch caused by base-pair formation between the 3' end of 16S rRNA and the mRNA during elongation of protein synthesis in Escherichia coli. *Embo J* **7**:1503-1507.
- Weiss RB, Dunn DM, Shuh M, Atkins JF, Gesteland RF. 1989. E. coli ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol* **1**:159-169.
- Wills NM, Moore B, Hammer A, Gesteland RF, Atkins JF. 2006. A functional -1 ribosomal frameshift signal in the human paraneoplastic Ma3 gene. *J Biol Chem* **281**:7082-7088.
- Xu X, Leo C, Jang Y, Chan E, Padilla D, Huang BC, Lin T, Gururaja T, Hitoshi Y, Lorens JB, Anderson DC, Sikic B, Luo Y, Payan DG, Nolan GP. 2001. Dominant effector genetics in mammalian cells. *Nat Genet* **27**:23-29.

- Yusupova GZ, Yusupov MM, Cate JH, Noller HF. 2001. The path of messenger RNA through the ribosome. *Cell* **106**:233-241.
- Zhang F, Romano PR, Nagamura-Inoue T, Tian B, Dever TE, Mathews MB, Ozato K, Hinnebusch AG. 2001. Binding of double-stranded RNA to protein kinase PKR is required for dimerization and promotes critical autophosphorylation events in the activation loop. *J Biol Chem* **276**:24946-24958.