
Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

doi:10.1016/S0022-2836(03)00784-8 J. Mol. Biol. (2003) 331, 571–583

JMB

Available online at www.sciencedirect.com

SCIENCE DIRECT®

# The Frameshift Stimulatory Signal of Human Immunodeficiency Virus Type 1 Group O is a Pseudoknot

**Martin Baril, Dominic Dulude, Sergey V. Steinberg and Léa Brakier-Gingras***

*Département de Biochimie Université de Montréal, 2900 Boulevard Edouard-Montpetit Montréal, Qué., Canada H3T 1J4*

*Corresponding author

Human immunodeficiency virus type 1 (HIV-1) requires a programmed −1 ribosomal frameshift to produce Gag–Pol, the precursor of its enzymatic activities. This frameshift occurs at a slippery sequence on the viral messenger RNA and is stimulated by a specific structure, downstream of the shift site. While in group M, the most abundant HIV-1 group, the frameshift stimulatory signal is an extended bulged stem-loop, we show here, using a combination of mutagenesis and probing studies, that it is a pseudoknot in group O. The mutagenesis and probing studies coupled to an *in silico* analysis show that group O pseudoknot is a hairpin-type pseudoknot with two coaxially stacked stems of eight base-pairs (stem 1 and stem 2), connected by single-stranded loops of 2 nt (loop 1) and 20 nt (loop 2). Mutations impairing formation of stem 1 or stem 2 of the pseudoknot reduce frameshift efficiency, whereas compensatory changes that allow re-formation of these stems restore the frameshift efficiency to near wild-type level. The difference between the frameshift stimulatory signal of group O and group M supports the hypothesis that these groups originate from a different monkey to human transmission.



*Keywords:* human immunodeficiency virus type 1; HIV-1 group O; ribosomal frameshifting; RNA structure; RNA pseudoknot

## Introduction

Human immunodeficiency virus type 1 (HIV-1) is classified into three groups: M (main), that accounts for more than 99% of the worldwide pandemic, N (new or non-M/non-O), that is not clearly defined yet, and O (outlier), that has a genomic sequence differing by 60% from group M. Each group is furthermore divided into subtypes based on differences in their *env*, *gag* or *pol* gene sequences.[1] Although much less abundant than group M, group O viruses are spread widely in Cameroon and Equatorial Guinea.[2] Their resistance to non-nucleoside reverse transcriptase inhibitors active against group M[3] as well as the rapid rate at which they develop resistance to protease inhibitors indicate that they could become an important threat.[4–6]

HIV-1 uses a programmed −1 ribosomal frameshift during translation of its full-length messenger RNA (mRNA) to produce the Gag–Pol polyprotein, the precursor of the viral enzymes. This programmed event allows a minority of ribosomes to shift the reading frame at a specific sequence before they encounter the stop codon of *gag*, and to extend the translation to the *pol* gene, while conventional translation of the same mRNA by the majority of ribosomes produces the Gag polyprotein, the precursor of the viral structural proteins.[7,8] The efficiency of the frameshift event controls the Gag–Pol to Gag ratio, which is critical for RNA dimerization, particle assembly, replication and viral infectivity.[9–11] Therefore, increasing or decreasing the frameshifting efficiency can interfere with the formation of infectious viral particles.

Abbreviations used: BWYV, beet western yellow virus; CMV, cytomegalovirus; HIV-1, human immunodeficiency virus type 1; IBV, infectious bronchitis virus; LUC, firefly luciferase; MMTV, mouse mammary tumor virus; PCR, polymerase chain reaction; RRL, rabbit reticulocyte lysate; RSV, Rous sarcoma virus; SRV-1, simian retrovirus-1; TFP, Gag–Pol transframe protein.

E-mail address of the corresponding author: lea.brakier.gingras@umontreal.ca

The programmed −1 ribosomal frameshift requires two *cis*-acting elements in the viral mRNA: a heptameric slippery sequence where the frameshift takes place (U UUU UUA in HIV-1) and a downstream stimulatory signal. In group M, the stimulatory signal was previously assumed to be a simple 11-bp stem capped by a 4-nt loop,[7,12] although an alternative structure was proposed but without any experimental support.[13] It was recently shown independently by our group[14] and by Dinman *et al*.[15] that the frameshift stimulatory signal is longer than believed and that the sequence downstream of the classic stem-loop contributes to the frameshifting event. Using a combination of site-directed mutagenesis and probing experiments, we demonstrated that this signal is an extended bulged stem-loop for subtype B of group M,[14] the subtype that prevails in North America and Europe, and also that the same extended bulged stem-loop structure is found in all subtypes of group M (our unpublished results). In this signal, the upper portion of the stem corresponds to the classic stem-loop previously considered as the frameshift stimulatory signal, while the lower portion is formed by a base-pairing between the spacer separating the slippery sequence from the classic stem-loop and a complementary pyrimidine-rich sequence downstream of this stem-loop (Figure 1). This structure thus differs from most frameshift stimulatory signals, which were shown to be pseudoknots.[8,16] Dinman *et al*.[15] proposed an alternative structure for the frameshift signal of HIV-1 group M. In their model, the pyrimidine-rich sequence downstream of the classic stem-loop interacts with three bases in the loop capping this stem through Watson–Crick interactions and forms a triplex with 4 bp on top of this stem. However, as discussed by Dulude *et al*.,[14] a careful analysis of this model shows that it is highly improbable, a major criticism being that the pyrimidine-rich strand that forms the triplex is in an antiparallel orientation relative to the purine-rich strand of the duplex with which it interacts, an orientation that is sterically not favored.

Here, the structure of the frameshift stimulatory signal of HIV-1 group O was investigated with a combination of mutagenesis and probing studies. The HIV-1 group O frameshift region or its mutants was inserted at the beginning of the coding sequence of a luciferase reporter gene, such that luciferase expression depends on a −1 frameshift. The luciferase activity was then assessed in 293T cultured cells and *in vitro*, in a rabbit reticulocyte lysate (RRL). The structure of the frameshift stimulatory signal of HIV-1 group O was also analyzed by probing the RNA fragment encompassing this region with RNase $V_1$. The proposed structure was then assessed by computer modeling. Group O is subdivided into two subtypes: MVP5180 and ANT70, for which there are seven and one complete *pol* sequences, respectively. Our results show that the frameshift stimulatory signal of HIV-1 group O MVP5180 consists of a pseudoknot, and also, that this pseudoknot promotes a higher frameshift efficiency (and thus a higher Gag–Pol to Gag ratio) than the group M frameshift stimulatory signal. However, sequence differences in subtype ANT70 prevent the frameshift stimulatory signal from adopting the same pseudoknot structure as in subtype MVP5180.

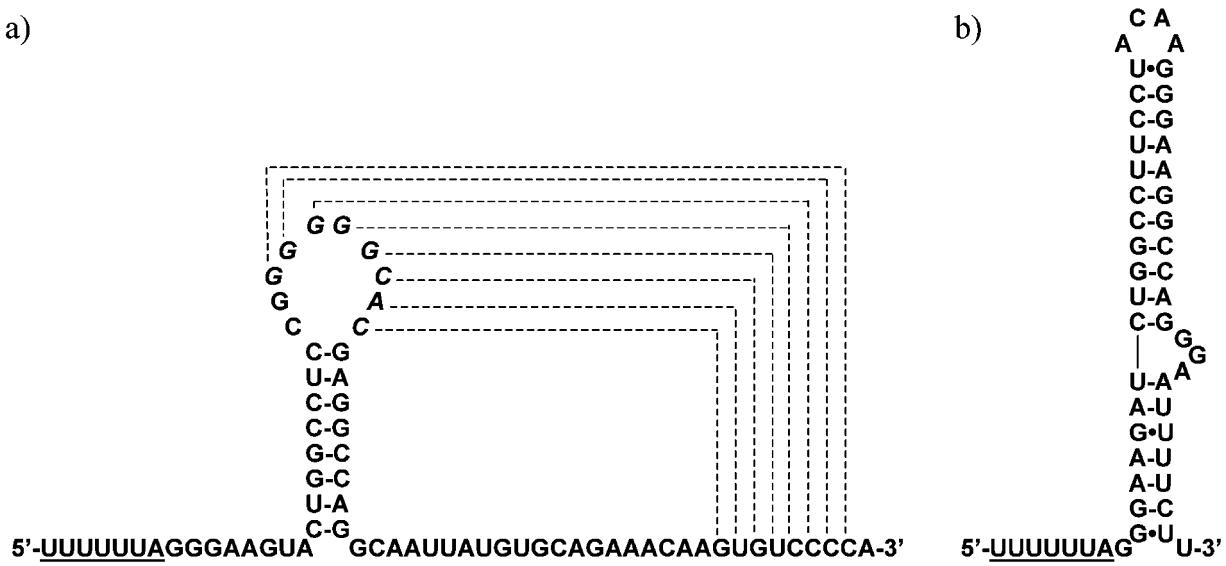


**Figure 1**. Predicted structure for the frameshift region of subtype MVP5180 of HIV-1 group O. (a) The slippery sequence (underlined) is followed by an 8-nt spacer and an 8-bp stem, capped by a 10-nt loop. Eight nucleotides of this loop could base-pair with a complementary region downstream of the stem-loop (proposed pairings are represented by broken lines). Such an interaction results in a pseudoknot with an 8-nt stem 1, a 2-nt loop 1, an 8-nt stem 2 and a 20-nt loop 2. (b) Structure of the frameshift region of subtype B of HIV-1 group M, as determined by Dulude *et al*.[14]

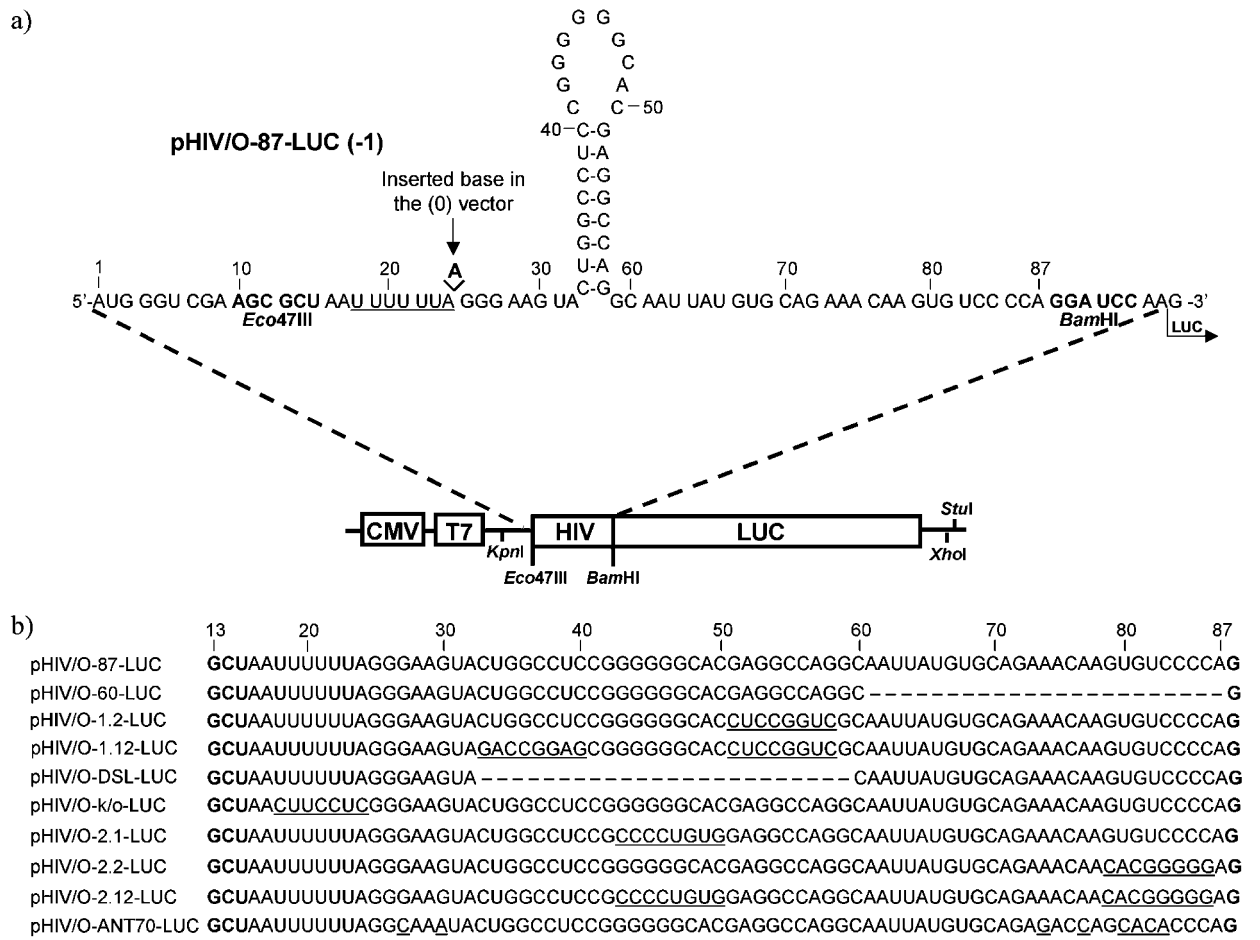


**Figure 2**. Description of the luciferase expression vectors used for the study of the programmed −1 ribosomal frameshift of HIV-1 group O *in vitro* and in cultured cells. (a) The HIV-1 frameshift region of subtype MVP5180 of group O was inserted upstream of the coding sequence of the luciferase reporter gene, generating construct pHIV/O-87-LUC. The slippery site is UUUUUUA (underlined). All mutants of subtype MVP5180 were cloned by inserting between the *Kpn* I-*Bam* HI sites of the vector, the PCR product bearing the mutation investigated. For subtype ANT70 of group O, the corresponding vector was constructed by exchanging the *Eco*47III-*Bam* HI fragment with an appropriate oligonucleotide cassette. For the (−1) constructs, the luciferase sequence is in the −1 reading frame relative to the AUG initiation codon, so that a −1 frameshift is required to produce luciferase. An adenine was added immediately after the slippery sequence (at position 25) for the (0) constructs, so that luciferase is expressed by ribosomes that do not shift the reading frame. (b) Sequences of the frameshift region of all constructs used in this study. Nucleotides substituted or deleted compared to subtype MVP5180 of group O are underlined or represented by broken lines, respectively.

## Results

### Sequence analysis of HIV-1 group O frameshift region

While the totality of studies on the HIV-1 frameshift have been carried out using group M, the frameshift region of group O has not been examined so far. We decided to investigate whether the frameshift region of HIV-1 group O encompasses the same slippery site and stimulatory signal as group M. As mentioned above, there are seven complete *pol* sequences for subtype MVP5180 and one complete *pol* sequence for subtype ANT70 (Los Alamos National Laboratory HIV Sequence Database).[17] We started the analysis of the frameshift region with MVP5180, the most represented subtype of group O. Examination of MVP5180 sequences suggests that the slippery sequence in this group O subtype is the same as in group M, while the frameshift stimulatory signal could be predicted to be an 8-bp stem (the thermodynamic stability calculated with *m*-fold[18] is −14.9 kcal/mol), capped by a 10-nt loop and separated from the slippery sequence by an 8-nt spacer. This stem-loop is different from that initially proposed as a frameshift stimulatory signal for group M. Moreover, for group O, because of the absence of base complementarity between the spacer region preceding the stem-loop and the sequence downstream of this stem-loop, the extended bulged stem-loop structure demonstrated for group M[14] cannot form (Figure 1). However, examination of the frameshift region sequence in MVP5180 suggests that its frameshift stimulatory signal could form a

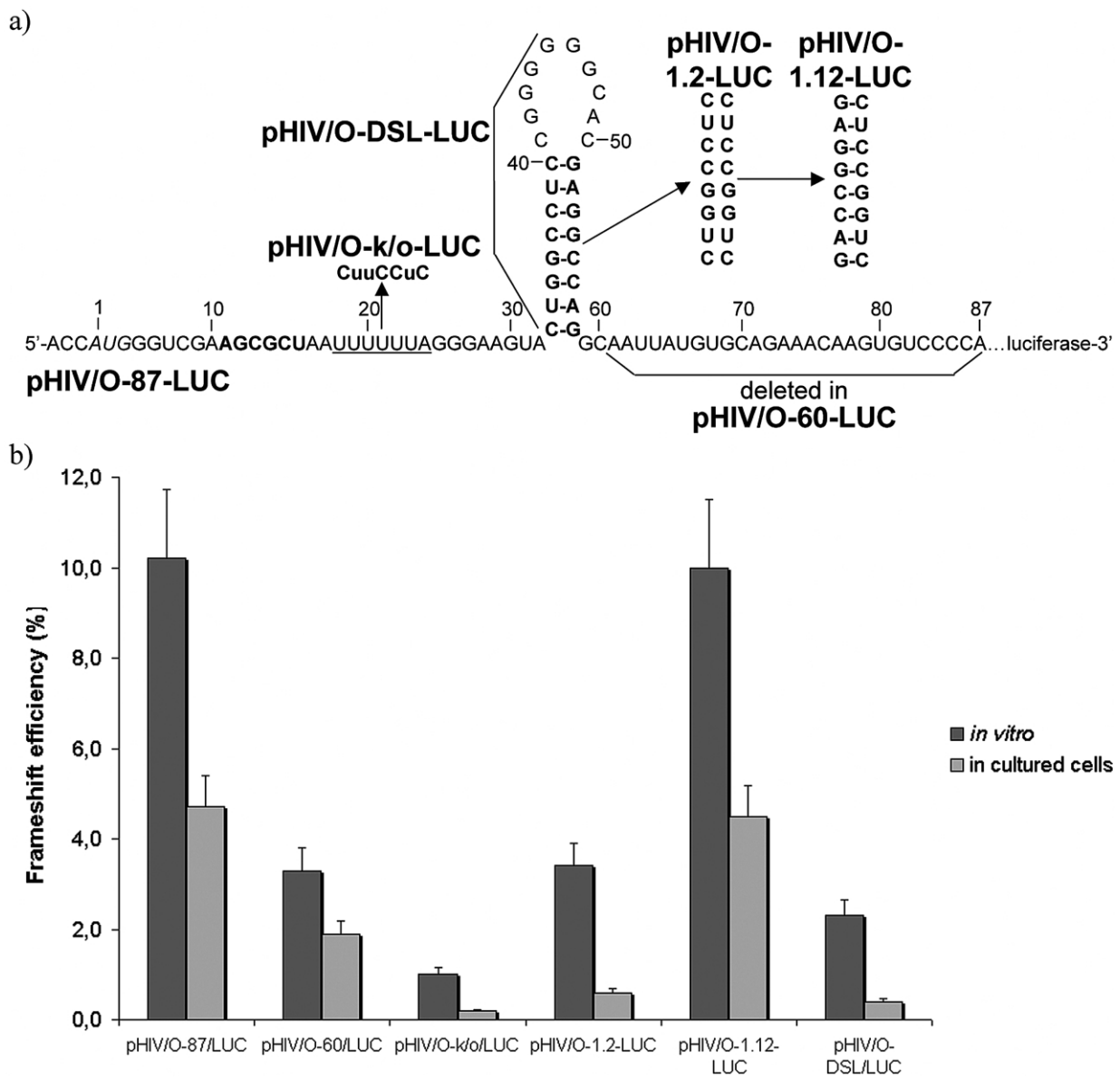


**Figure 3**. Effect of different mutations in the frameshift region of subtype MVP5180 of HIV-1 group O on the frameshift efficiency. (a) A series of mutations were made within the frameshift region of pHIV/O-87-LUC (the dots correspond to the *Bam*HI linker connecting the frameshift region to the luciferase coding sequence): a slippery site mutant, pHIV/O-k/o-LUC, where the slippery sequence (underlined) is mutated (bases that are changed are in capital letters); deletion mutant pHIV/O-DSL-LUC, where the region encompassing stem 1 and its capping loop is deleted (deletion of bases 33 to 60); deletion mutant pHIV/O-60-LUC, where the region 3′ to stem 1 is deleted; substitution mutants where the 3′ strand of stem 1 is altered, impairing formation of stem 1 (pHIV/O-1.2-LUC), and where the 3′ and 5′ strands of stem 1 are simultaneously altered, allowing re-formation of stem 1 (pHIV/O-1.12-LUC). (b) Frameshift efficiency *in vitro* and in cultured cells for the pHIV/O-LUC constructs described above. *In vitro* translation experiments were made in 25 μl of RRL with 0.2 μg of mRNAs transcribed from the *Stu* I-digested pHIV/O-LUC constructs. Assays in cultured cells were made by co-transfecting 293T cells with 3 μg of pHIV/O-LUC and 1.25 μg of pcDNA3.1/Hygro(+)/lacZ. Frameshift efficiencies were calculated as described in the text. Each value represents the mean ± standard error of five to six independent experiments. The bars indicate the standard error on the mean.

pseudoknot structure by base-pairing between 8 nt of the loop capping the 8-bp stem and a complementary downstream sequence (Figure 1).

### Characterization of the frameshift region of subtype MVP5180 of HIV-1 group O

To investigate whether the frameshift stimulatory signal of HIV-1 group O could adopt the proposed pseudoknot structure, we made a reporter construct for the frameshift region of subtype MVP5180 of HIV-1 group O, using a modified pcDNA3.1/Hygro(+) plasmid, in which the firefly luciferase (*luc*) reporter gene had been introduced under control of a CMV and a T7 promoter. The frameshift region of subtype MVP5180 or its

**Table 1.** Frameshift efficiency for different mutants of group O of HIV-1

| Construct | Description | Frameshift efficiency (%) | |
|---|---|---|---|
| | | *In vitro* | In cultured cells |
| *Frameshift efficiency for regions of different length of group O* | | | |
| pHIV/O-87-LUC | Long frameshift region | 10.2 (100) | 4.7 (100) |
| pHIV/O-60-LUC | Short frameshift region | 3.3 (32) | 1.9 (40) |
| *Frameshift efficiency for mutants of group O* | | | |
| pHIV/O-k/o-LUC | Altered slippery sequence | 1.0 (10) | 0.2 (4) |
| pHIV/O-1.2-LUC | Substitution (3′ strand of stem 1) | 3.4 (33) | 0.6 (13) |
| pHIV/O-1.12-LUC | Substitution (stem 1 re-formed) | 10.0 (98) | 4.5 (96) |
| pHIV/O-DSL-LUC | Deletion (minus stem 1 and its capping loop) | 2.3 (23) | 0.4 (9) |
| *Frameshift efficiency for mutations impairing or re-forming the pseudoknot* | | | |
| pHIV/O-2.1-LUC | Substitution (loop capping stem 1) | 4.4 (43) | 1.6 (34) |
| pHIV/O-2.2-LUC | Substitution (downstream segment) | 3.9 (38) | 2.1 (45) |
| pHIV/O-2.12-LUC | Compensatory (pseudoknot re-formed) | 10.3 (101) | 3.8 (81) |
| pHIV/O-ANT70-LUC | Long frameshift region for subtype ANT70 of group O | 4.0 (39) | 2.2 (47) |

All pHIV/O-LUC constructs contain the HIV-1 *gag/pol* frameshift region of group O MVP5180 (except for construct pHIV/O-ANT70-LUC). Mutants of the group O frameshift region are further identified by a short description recalling their characteristics (see details in Figures 3(a) and 4(a)). For each (−1) construct, an in-frame (0) control was made to monitor the frameshift efficiency. The numbers between brackets represent the frameshift efficiency of each construct relative to pHIV/O-87-LUC, which is arbitrarily set at 100%. Results are the means of five to six independent experiments. Standard error on the mean was inferior or equal to 10%.

derivatives was inserted between the first and second codon of *luc*, generating a series of pHIV/O-LUC (−1) constructs (Figure 2). These insertions were such that a −1 frameshift was required to express the LUC reporter protein. An in-frame (0) control construct, in which ribosomes synthesize LUC through conventional translation, was also made for each (−1) frameshift construct examined, by adding an adenine immediately after the slippery sequence. The frameshift efficiency for each construct was measured by dividing the luciferase activity of the (−1) construct by the sum of that of the (0) and (−1) constructs, assuming that the same level of frameshift occurs in both (0) and (−1) constructs. The frameshift efficiency of these constructs was assessed in cultured cells and *in vitro*. For assays in cultured cells, the different pHIV/O-LUC (0) and (−1) constructs were transfected into 293T cells, and luciferase activity was measured in cell extracts 48 hours later. For *in vitro* assays, luciferase activity was measured after translation in RRL of the HIV/O-LUC mRNAs generated by transcription with T7 RNA polymerase of the *Stu* I-linearized plasmids.

Figure 3 shows the characteristics of the structure of the different constructs and their frameshift efficiency, and the results are summarized in Table 1. Consistent with what was systematically observed in previous reports, the frameshift efficiency was higher *in vitro* than in cultured cells (see Ref. 14 and references therein), and was less sensitive *in vitro* to the presence of the stimulatory signal, compared to the situation in cultured cells,[19,20] which was proposed to result from differences between the rate of translation *in vitro* and in cultured cells. As a reference construct for the study of the frameshift region of HIV-1 group O, we used pHIV/O-87-LUC. This construct bears the frameshift region of subtype MVP5180 of group O, encompassing the putative slippery site, the predicted stem-loop and the downstream region, with the sequence proposed to interact with the loop, thus forming a pseudoknot. The frameshift efficiency of pHIV/O-87-LUC was 4.7% in cultured cells and 10.2% *in vitro*. Using construct pHIV/O-k/o-LUC, in which the predicted slippery sequence UUUUUUA of pHIV/O-87-LUC was replaced by CUUCCUC, we first verified that the slippery sequence is the same in group O as in group M. The frameshift in pHIV/O-k/o-LUC was abolished or dramatically decreased in cultured cells and *in vitro*, respectively, confirming that group O viruses use the same slippery sequence as group M viruses. The influence on frameshift efficiency of the region downstream of the predicted 8-bp stem-loop was assessed with construct pHIV/O-60-LUC, bearing the slippery site and the 8-bp stem-loop, but lacking this downstream region. The frameshift efficiency of pHIV/O-60-LUC decreased two to threefold in cultured cells and *in vitro*, compared to pHIV/O-87-LUC. This indicates the presence of an element in the region downstream of the putative 8-bp stem-loop of group O that stimulates the frameshift efficiency, although it does not prove the existence of a pseudoknot structure.

Next, the importance of the proposed 8-bp stem-loop for frameshift was demonstrated using three constructs: pHIV/O-1.2-LUC, pHIV/O-1.12-LUC and pHIV/O-DSL-LUC, all derivatives of pHIV-87-LUC. In pHIV/O-1.2-LUC, the 8 nt of the 3′-strand of the stem-loop were replaced with the 8 nt of the 5′-strand, thus impairing the formation of the stem. In pHIV-1.12-LUC, the 3′ and the 5′-strands of the stem were exchanged, thus restoring the capacity to form the stem. Finally, in pHIV/O-DSL-LUC, the stem-loop was deleted. For the two constructs where the stem-loop was either

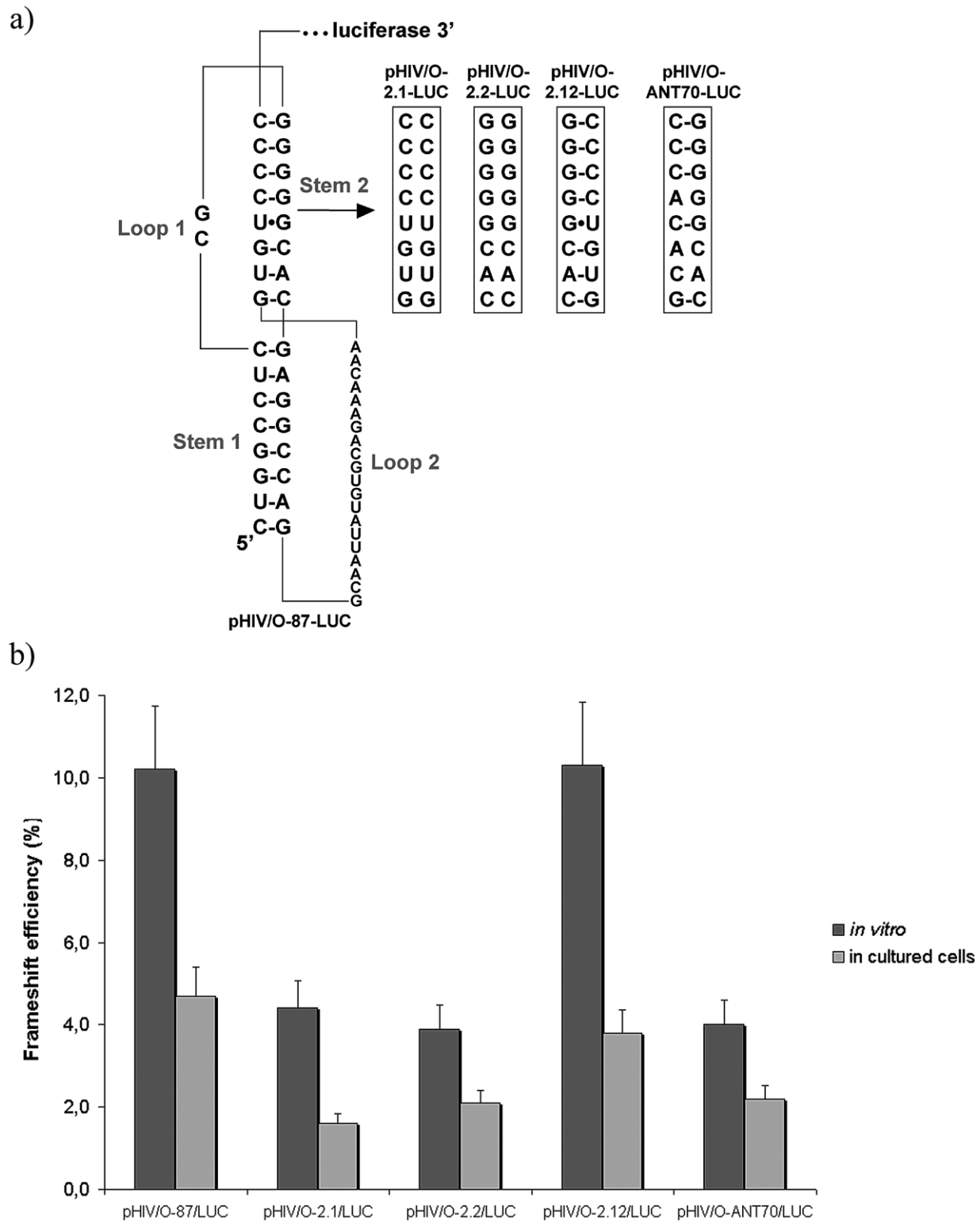


**Figure 4**. Effect on the frameshift efficiency of mutations impairing formation of stem 2 of the pseudoknot of subtype MVP5180 of HIV-1 group O. (a) Description of mutations made within the *gag/pol* frameshift region of pHIV/O-87-LUC (the dots correspond to the *Bam* HI linker connecting the frameshift region to the luciferase coding sequence). Mutations were introduced either in the loop capping stem 1 or in the complementary downstream region, impairing formation of stem 2 (pHIV/O-2.1-LUC and pHIV/O-2.2-LUC, respectively), or allowing re-formation of this stem (pHIV/O-2.12-LUC). Subtype ANT70 of group O, where sequence differences compared to subtype MVP5180 impair formation of stem 2, was also assessed. (b) Frameshift efficiency *in vitro* and in cultured cells with the HIV/O-LUC constructs described above. Assays were as described in the legend to Figure 3.

destabilized or deleted, the frameshift efficiency was reduced by about tenfold in cultured cells and three to fivefold *in vitro*, compared to pHIV/O-87-LUC. However, in pHIV/O-1.12-LUC, where the re-formation of the stem-loop was possible, the frameshift efficiency was restored to the wild-type level *in vitro* and in cultured cells. These results confirm the existence of the predicted 8-bp stem-loop structure and its involvement in the frameshift process. Since the region downstream of the stem-loop increases frameshifting in the presence of this stem-loop, as shown in Figure 3, these results also support our suggestion that the frameshift stimulatory signal in MVP5180 is more complex than a simple stem-loop and could correspond to a pseudoknot structure.

### Characterization of the pseudoknot in the frameshift region of subtype MVP5180 of HIV-1 group O

Figure 4(a) shows a standard representation of the hypothetical pseudoknot that we proposed to act as a frameshift stimulatory signal in HIV-1 group O. To demonstrate the existence of this pseudoknot, we made a series of mutations in the loop capping the 8-bp stem, named stem 1, as well as in the downstream complementary sequence predicted to interact with this capping loop and to form stem 2 of the pseudoknot. Individually, these mutations impair the formation of stem 2, but when combined, should allow it to re-form (Figure 4). The frameshift efficiencies of the constructs containing these mutations were assessed *in vitro* and in cultured cells and are presented in Figure 4(b) (see also Table 1). The mutations consisted of replacing 8 nt of the loop capping stem 1, with the downstream set of complementary nt (pHIV/O-2.1-LUC). The frameshift efficiency of this construct was reduced by about two to threefold *in vitro* and in cultured cells, compared to the wild-type construct pHIV/O-87-LUC. This level of frameshift efficiency is similar to that obtained with the 8-bp stem-loop alone (pHIV/O-60-LUC). Similarly, the frameshift efficiency of pHIV/O-2.2-LUC, in which the eight bases downstream of stem 1 were replaced with the eight complementary bases of the loop, was decreased by about two to threefold *in vitro* and in cultured cells. However, combining both mutations (pHIV/O-2.12-LUC), so as to allow re-formation of stem 2 of the pseudoknot, restored the frameshift efficiency to the wild-type level *in vitro* and in cultured cells. Therefore, mutagenesis studies fully support our hypothesis that the stimulatory signal of subtype MVP5180 of HIV-1 group O adopts a pseudoknot structure.

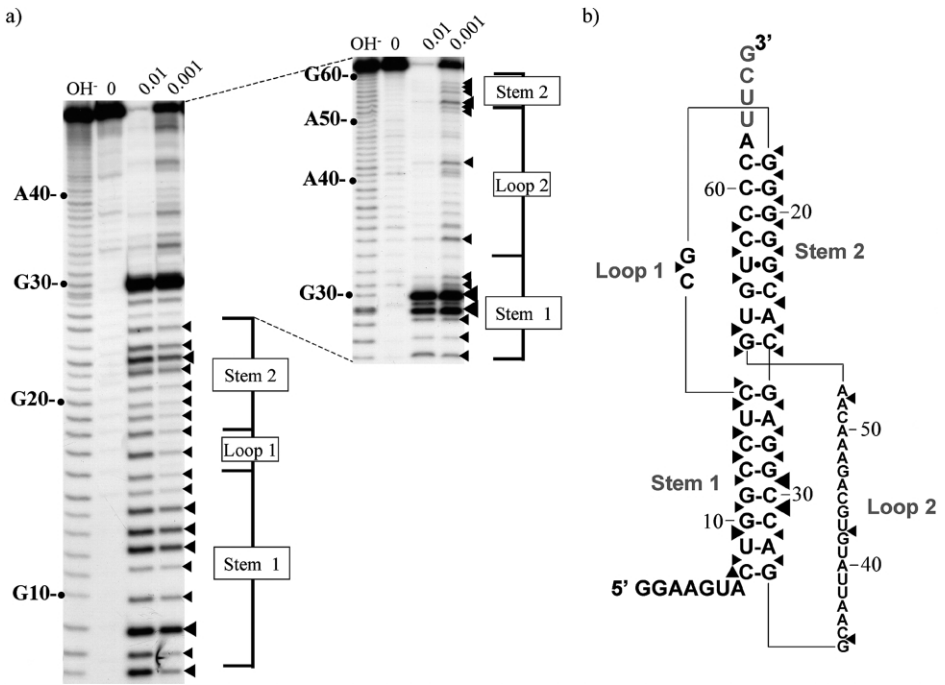


**Figure 5**. Probing of the structure proposed for the frameshift stimulatory signal of subtype MVP5180 of HIV-1 group O. (a) Structure probing of the frameshift stimulatory signal by RNase $V_1$ attack. An RNA transcript encompassing the *gag/pol* frameshift region was 5′ end-labeled with γ-$^{32}$P and digested with RNase $V_1$. Digestion products were analyzed on a 20% (left) and a 10% (right) acrylamide 7 M urea gel. The sites of cleavage were identified by comparison with a ladder of bands created by limited alkaline hydrolysis of the RNA ($OH^-$). The nucleotides that were cleaved were identified by the absence of cleavage in the untreated control lane (0). The amount of units of enzyme added to each reaction is also indicated. (b) Summary of the RNase $V_1$ attacks in the pseudoknot structure of the frameshift stimulatory signal of subtype MVP5180 of HIV-1 group O. The sensitivity of nucleotides to RNase $V_1$ is shown by arrows of different size, where the size is approximately proportional to the intensity of the cleavage at that site. Bases 1 to 62 originate from subtype MVP5180, while bases 63 to 66 (in gray) originate from the vector.

### RNA structure probing of the pseudoknot of subtype MVP5180 of HIV-1 group O

To provide an independent support for the pseudoknot structure in subtype MVP5180 of HIV-1 group O frameshift stimulatory signal, we made a structural probing analysis of an RNA fragment encompassing the *gag/pol* frameshift region of this subtype. Enzymatic probing with RNase $V_1$, an enzyme that cleaves RNA in helical conformation, showed that the regions corresponding to stems 1 and 2 were attacked by the enzyme, with the strongest cleavage sites in stem 1 (Figure 5). Stem 2 was also sensitive to RNase $V_1$, contrasting with loop 2, which was only weakly attacked. The stacking of the two nucleotides in loop 1 likely accounts for its sensitivity to RNase $V_1$. Altogether, probing experiments are in perfect agreement with the mutagenesis studies, showing that the frameshift stimulatory signal of HIV-1 group O MVP5180 is a pseudoknot.

### Computer modeling of the pseudoknot of subtype MVP5180 of HIV-1 group O

Structural analysis of frameshifting pseudoknots and other frameshift stimulatory signals can lead to the identification of conserved motifs and also to a prediction of the mechanism through which frameshift is stimulated.[21] Based on the characteristics of their structure, two major classes of frameshifting pseudoknots can be proposed. In the first class, pseudoknots have an unpaired residue (usually an adenine) at the junction of the two stems, allowing a characteristic bent conformation that was proposed to be the conserved motif promoting frameshift stimulation.[22] The pseudoknot of the mouse mammary tumor virus (MMTV), resolved by NMR,[23] and that of the beet western yellow virus (BWYV), resolved by X-ray crystallography,[24] belong to this class. The frameshifting pseudoknots of the second class do not have an unpaired residue at the junction of the coaxially stacked stems. Among pseudoknots of this class are those of the simian retrovirus-1 (SRV-1), resolved by NMR,[25] and of the infectious bronchitis virus (IBV), which has been extensively studied by mutagenesis and probing studies.[26–28] No additional structural data on frameshifting pseudoknots with stacked stems are available, making identification of a conserved motif difficult. In addition to these two classes of pseudoknots, the Rous sarcoma virus (RSV) has an unusual pseudoknot, where the loop capping stem 1 is partly involved in the formation of stem 2 but contains two additional small sub-stem elements.[29]

Combining our mutagenesis and probing results

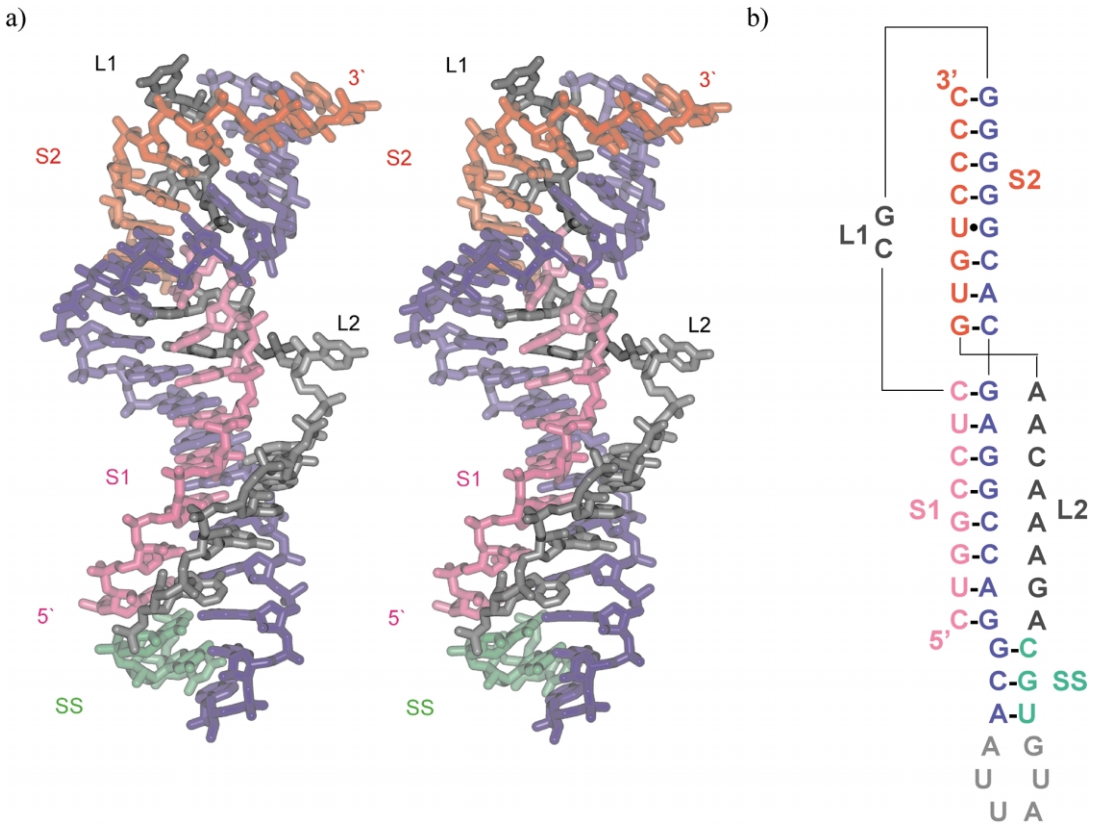


**Figure 6**. Stereo view of the computer modeled structure for the pseudoknot of subtype MVP5180 of HIV-1 group O. Stems (S1 and S2) and loops (L1 and L2) are represented in different colors. (a) Structure of the MVP5180 pseudoknot, with a putative sub-stem (SS) in loop 2 (see the text). Nucleotides in the loop of this sub-stem were not included in our modeling. (b) Schematic representation of the pseudoknot.

with the structural information on −1 frameshifting pseudoknots recently provided by crystallographic and NMR studies, we made an *in silico* modeling of the HIV-1 group O pseudoknot (Figure 6). As shown above, this pseudoknot was predicted to be a classic hairpin-type pseudoknot (e.g. with base-pairing in a hairpin loop). It has two 8-bp stems (stem 1 and stem 2) connected by loops of 2 nt (loop 1) and 20 nt (loop 2). Although the lengths of the stems and loops make the RNA pseudoknot of group O different from other pseudoknots promoting frameshift, these lengths still fit to the general steric requirements that allow the coaxial arrangement of the two stems.[30] Among the pseudoknots whose tertiary structure has been resolved, that of the SRV-1 RNA[25] is the best to be used as a starting conformation to model the structure of the group O RNA pseudoknot. In both pseudoknots, there is no intervening nucleotide between the two stems, so that one can presume that, like in the SRV-1 RNA pseudoknot, there is no kink between the two stems of the group O pseudoknot. Also, the lengths of stems 1 and 2 in both molecules differ by only two base-pairs while the length of loop 1 differs by only one nucleotide, which necessitates only minimal rearrangements. The only region that is notably different in both pseudoknots is loop 2, which is 8-nt longer in the group O RNA than in the SRV-1 RNA, and is much longer than needed for connecting the two stems, which suggests the existence of additional elements in the secondary structure of the pseudoknot. Indeed, visual analysis of the nucleotide sequence of loop 2 revealed the possibility to form an additional 3-bp stem in this loop, immediately after the junction with stem 1 (stem SS in Figure 6). If this sub-stem is positioned coaxially to stem 1, it would effectively reduce the number of nucleotides in loop 2 down to only eight. Modeling such a structure showed that formation of this additional substructure and its coaxial position with respect to stem 1 does not create any problem for the connection of stem 1 and stem 2 with loop 2. Formation of this sub-stem would provide an additional stabilizing effect for the whole structure, where most nucleotides of the connector region become involved in H-bonding with the minor groove of stem 1, forming a triplex interaction. Although disruption of the putative sub-stem by mutagenesis did not affect frameshift efficiency in our assays (data not shown), it is still possible that it could affect frameshifting with the full-length viral RNA, and not with a reporter system. A similar 4-bp stem in loop 2 can also be proposed for the IBV pseudoknot, although a deletion in loop 2, impairing the formation of this putative stem, was found not to affect the frameshift efficiency.[31] The triplex interaction between loop 2 and stem 1 in group O pseudoknot involves an AACAA sequence. This is reminiscent of the situation observed in BWYV[24] and SRV-1,[25] where an AACAA sequence present in loop 2 is implicated in loop-helix triplex interactions. Furthermore, as in SRV-1, the cytosine of the AACAA sequence bulges out in our modeling of group O pseudoknot. Su *et al*.[24] and Michiels *et al*.[25] from their studies with BWYV and SRV-1, respectively, suggested that the formation of a triplex structure in a pseudoknot that promotes frameshift could be a signal to frameshift recognized by the translating ribosome. Our results with HIV-1 group O support this suggestion.

### Subtype ANT70 of HIV-1 group O does not contain the same frameshift stimulatory signal as subtype MVP5180

As mentioned above, there is a second subtype for HIV-1 group O, ANT70,[1] for which only one complete *pol* sequence is available.[17] Sequence analysis suggests that subtype ANT70 uses the same slippery sequence, followed by the same 8-bp stem-loop as in subtype MVP5180. However, sequence differences in the region downstream of this stem-loop prevent formation in ANT70 of stem 2 of the pseudoknot present in MVP5180 (see Figure 4(a)). The frameshift efficiency of subtype ANT70 was assessed with construct pHIV/O-ANT70-LUC containing the frameshift region from subtype ANT70. It was found to be about half of that of pHIV/O-87-LUC both *in vitro* and in cultured cells (Figure 4(b)), which is consistent with the incapacity for this frameshift region to form the pseudoknot found in subtype MVP5180. This does not exclude however the possibility that the frameshift region of subtype ANT70 can form another pseudoknot (see Discussion).

## Discussion

Our results show that the frameshift stimulatory signal of subtype MVP5180 of HIV-1 group O is a pseudoknot, where 8 nt of a 10-nt loop capping an 8-bp stem base-pair with a downstream complementary sequence. The presence of this pseudoknot was demonstrated by site-directed mutagenesis and enzymatic probing of the frameshift stimulatory region. Among the two distinct classes of −1 frameshifting pseudoknots described above, the group O pseudoknot falls in the same class as those of SRV-1[25] and IBV,[26] characterized by the absence of any intervening nucleotide between the coaxial stems. Interestingly, the frameshift efficiency of the group O pseudoknot is two and four times lower than that of the SRV-1 and IBV pseudoknots, respectively. This can result from the difference in the length of stem 1, since shortening this stem from 11 to 10 bp caused an 85% loss of the frameshift efficiency for the IBV pseudoknot.[28] The lower frameshift efficiency observed for the HIV-1 group O pseudoknot can also relate to the absence of a guanosine-rich region in the 5′ arm of stem 1, which was found to contribute significantly to stimulation of frameshift in IBV,[28] SRV-1[32] and MMTV.[33] Formation of stem 2

in group O stimulates frameshift by two to three-fold compared to the effect of a simple stem-loop (pHIV/O-87-LUC *versus* pHIV/O-60-LUC). A similar situation is encountered in the IBV[28] and RSV pseudoknots,[29] where destabilization of stem 2 has a less drastic effect than in SRV-1[34] and MMTV,[33] for which the frameshift efficiency is reduced by eight to tenfold under these conditions.

The use of a pseudoknot as a frameshift stimulatory signal makes HIV-1 group O distinct from group M, where this signal is an extended bulged stem-loop.[14] This observation supports previous phylogenetic analyses according to which the group O and group M clusters of HIV-1 originated from different monkey to human transmissions.[35,36] Since evolution led to the selection of a stimulatory signal in group O different from group M, it likely corresponds to a need specific to group O. Here, we found that the frameshift efficiency in cultured cells was nearly 5% for group O MVP5180. When the frameshift region of group O was replaced with the frameshift region of group M subtypes, under conditions ensuring that any difference observed between the frameshift efficiency of group O and that of group M results from the difference in the structure of the stimulatory signal, we found that the frameshift efficiency of group M subtypes was about 2%, half of that of group O (our unpublished results). We can thus propose that the protease of HIV-1 group O viruses could be less active than group M protease, and that group O viruses require a Gag–Pol to Gag ratio higher than that of group M members, so as to incorporate more Gag–Pol in their virions for an optimal fitness during infection. Although there are no data available on the protease activity of group O, this protease is known to contain several secondary mutations that are found together with primary mutations causing resistance to protease inhibitors in variants of group M.[4–6] Moreover, mutations implicated in resistance to protease inhibitors usually reduce the activity of the enzyme in the absence of inhibitors. Another possibility to account for the higher frameshift efficiency of group O is that the sequence variations in the frameshift region of group O, compared to group M, reduce the activity of the three peptides encoded by this region, the spacer peptide, p1, the transframe peptide, TFP, and p6. These peptides are important for the formation of infectious viral particles,[37–39] and a higher production of Gag–Pol in group O viruses could compensate for a possible decrease in the activity of these peptides.

Surprisingly, the frameshift region of the other subtype of group O, ANT70, also contains an 8-bp stem capped by a 10-nt loop, but cannot form the same pseudoknot as MVP5180, since variations in the sequence downstream this stem-loop impair the formation of stem 2. The level of frameshift efficiency in this subtype, as measured in our reporter system, is about twofold less than in MVP5180. We cannot, however, exclude the possibility that the loop capping stem 1 interacts with another region further downstream to form a pseudoknot that would promote frameshift to a level close to that in MVP5180, and such complementary regions can indeed be found in the viral RNA sequence. Loop 2 in these putative pseudoknots would be very large, but other viruses, such as the human coronavirus, are known to use a frameshift stimulatory pseudoknot where loop 2 contains more than 160 nt.[40]

Finally, ribosomal frameshifting in HIV-1 represents an interesting novel target for an antiviral treatment. It is particularly important to develop new anti-HIV agents directed against group O viruses, since these viruses are naturally resistant to non-nucleoside reverse transcriptase inhibitors[3] and bear mutations that could lead to faster resistance against protease inhibitors.[4–6] The presence of a pseudoknot as a frameshift stimulatory signal in HIV-1 group O suggests that future anti-frameshift agents targeted against this signal must be developed independently for group M and group O members.

## Materials and Methods

### Construction of plasmids

All the plasmids used in this study were derived from pcDNA3.1-LUC, originating from pcDNA3.1/Hygro(+) (Invitrogen) where the luciferase gene was inserted under control of a CMV and a T7 promoter.[14] An oligonucleotide cassette containing a 5′UTR region of 55 nt, an initiator AUG and, two codons downstream, an *Eco*47III restriction site, was inserted between restriction sites *Kpn*I and *Bam*HI of pcDNA3.1-LUC, generating pcDNA3.1-5′UTR-LUC, where the insertion precedes the second codon of luciferase. The frameshift region of HIV-1 group O subtypes MVP5180 and ANT70 (GenBank accession nos. L20571 and L20587, respectively) was inserted between the *Eco*47III and *Bam*HI sites, creating pHIV/O-87-LUC (−1) and pHIV/O-ANT70-LUC (−1) constructs (Figure 2). These constructs are such that the luciferase coding sequence is in the −1 reading frame relative to the initiator codon, so that only ribosomes that make a −1 frameshift produce luciferase, while ribosomes translating in the 0 reading frame terminate translation at a stop codon overlapping codon six of the *luc* sequence. For each construct, an in-frame control plasmid, the (0) construct, was created by inserting an additional adenine just after the slippery sequence of the (−1) construct. Derivatives of pHIV/O-87-LUC (deletion and substitution mutants) were created by PCR, by first amplifying mutated DNA fragments from pHIV/O-87-LUC construct with two primers for deletion mutants or mutants with substitutions 3′ of the stem-loop of the stimulatory signal, and with four primers for mutants with substitutions in the slippery sequence and in the stem-loop, as described by Ho *et al*.[41] The amplified DNA fragments were then subcloned between the *Kpn*I and *Bam*HI sites of pcDNA3.1-LUC, and all constructs were verified by sequencing the entire insert.

### Transient transfections and luciferase assays

Transfections of HIV-1 group O plasmids in 293T cells were carried using a standard calcium-phosphate precipitation method,[42] with 3 μg of a pHIV/O-LUC construct and 1.25 μg of pcDNA3.1/Hygro(+)/lacZ coding for β-galactosidase, as described.[14] For luciferase assays, 1.5 μl of cell extract (600 μl) was added to 50 μl of Luciferase Assay Reagent (Promega) and the light emitted was measured with a Berthold Lumat LB 9507 luminometer. Frameshift efficiencies were calculated by dividing the luciferase activity of the (−1) construct by the sum of the luciferase activity of the (0) and (−1) constructs. The β-galactosidase activity was measured with the chlorophenolred-β-galactopyranoside substrate (Calbiochem), as described,[43] with aliquots of 10 μl of cell extracts, and used to normalize luciferase activities for variations in transfection efficiency.

### *In vitro* transcription and translation

*In vitro* transcriptions were carried out essentially as described,[14] using *Stu*I-linearized pHIV/O-LUC constructs. These RNA transcripts (0.2 μg) were translated in 25 μl of RRL (Promega) at 30 °C for 15 minutes, a reaction time for which the translation system functions at its maximal rate. The reaction was stopped by addition of EDTA to a final concentration of 6 mM. Luciferase activity was monitored as mentioned above, with 2.5 μl of the translation mixture. Frameshift efficiencies were calculated as described above.

### Enzymatic probing of RNA structure

Enzymatic probing of the structure of an RNA fragment encompassing the frameshift region of HIV-1 was performed as described with minor modifications.[14,29] An oligonucleotide cassette containing a T7 promoter followed by the HIV-1 group O *gag/pol* frameshift region (bases 26 to 87, Figure 2) was cloned between the *Nae*I and *Bsp*119I sites of pGEM®-7Zf(−) (Promega), generating the recombinant plasmid pGEM-HIV/O. The RNA transcript, produced by *in vitro* transcription of the *Bsp*119I-linearized plasmid, was 5′-end-labeled with γ-$^{32}$P using a standard dephosphorylation–rephosphorylation method,[44] purified from a 10% (w/v) acrylamide–7 M urea gel and dissolved in 500 mM $NH_4OAc$, 10 mM $Mg(OAc)_2$, 1 mM EDTA and 0.1% (w/v) SDS. Probing with RNase $V_1$ (in a total volume of 10 μl containing $10^5$ cpm of 5′-end-labeled RNA supplemented with 1 μg of yeast tRNA) was done at 25 °C for 15 minutes in 10 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 100 mM KCl and 0 to 0.01 units of enzyme (Ambion). The reaction was stopped by adding an equal volume of formamide gel loading buffer, the sample was heated two minutes at 95 °C and immediately loaded for analysis on a 10% and a 20% polyacrylamide–7 M urea gel.

### Computer modeling

Preliminary modeling was done interactively, using InsightII/Discover package (Version 2000, Accelrys Inc., San Diego, CA). The solution structure of the SRV-1 RNA pseudoknot[25] (PDB identificator 1E95), having six base-pairs in each of the two stems and one and 12 nucleotides in loops 1 and 2, respectively, was used as a starting conformation. The nucleotide sequence of the molecule was changed and additional base-pairs were appended to both stems according to the suggested model of the secondary structure for the RNA pseudoknot of group O subtype MVP5180. The unpaired nucleotides of loops 1 and 2 were arranged as in the structure of these regions in the SRV-1 RNA pseudoknot, providing a reasonable system of H-bonds and base–base stacking interactions. The model was submitted to unrestrained energy minimization using the AMBER forcefield[30] until an energy minimum was reached. Visualizations were done in a Silicon Graphics O2 computer. The PDB file is available on request.

## Acknowledgements

This study was supported by a grant from the Canadian Institutes of Health Research to L.B.-G. and Guy Lemay and a grant from the Canadian Institutes for Health Research to S.V.S. S.V.S. acknowledges fellowships from the Canadian Institutes of Health Research and from the Fonds de la Recherche en Santé du Québec. We thank Guy Lemay for his interest in this work. We are grateful to Pascal Chartrand, Gerardo Ferbeyre and Hugo Soudeyns for stimulating discussions and critical reading of this manuscript.

## References

1. Yamaguchi, J., Vallari, A. S., Swanson, P., Bodelle, P., Kaptue, L., Ngansop, C. *et al.* (2002). Evaluation of HIV type 1 group O isolates: identification of five phylogenetic clusters. *AIDS Res. Hum. Retroviruses*, **18**, 269–282.
2. Peeters, M., Gueye, A., Mboup, S., Bibollet-Ruche, F., Ekaza, E., Mulanga, C. *et al.* (1997). Geographical distribution of HIV-1 group O viruses in Africa. *AIDS*, **11**, 493–498.
3. Descamps, D., Collin, G., Letourneur, F., Apetrei, C., Damond, F., Loussert-Ajaka, I. *et al.* (1997). Susceptibility of human immunodeficiency virus type 1 group O isolates to antiretroviral agents: *in vitro* phenotypic and genotypic analyses. *J. Virol.* **71**, 8893–8898.
4. Yang, C., Gao, F., Funjungo, P. N., Zekeng, L., van der Groen, G., Pieniazek, D. *et al.* (2000). Phylogenetic analysis of protease and transmembrane region of HIV type 1 group O. *AIDS Res. Hum. Retroviruses*, **16**, 1075–1081.
5. de Baar, M. P., Janssens, W., de Ronde, A., Fransen, K., Colebunders, R., Kestens, L. *et al.* (2000). Natural residues versus antiretroviral drug-selected mutations in HIV type 1 group O reverse transcriptase and protease related to virological drug failure *in vivo*. *AIDS Res. Hum. Retroviruses*, **16**, 1385–1394.
6. Luk, K. C., Kaptue, L., Zekeng, L., Soriano, V., Gurtler, L., Devare, S. G. *et al.* (2001). Naturally occurring sequence polymorphisms within HIV type 1 group O protease. *AIDS Res. Hum. Retroviruses*, **17**, 1555–1561.
7. Jacks, T., Power, M. D., Masiarz, F. R., Luciw, P. A., Barr, P. J. & Varmus, H. E. (1988). Characterization of ribosomal frameshifting in HIV-1 gag–pol expression. *Nature*, **331**, 280–283.

8. Brierley, I. & Pennel, S. (2001). Structure and function of the stimulatory RNAs involved in programmed eukaryotic −1 ribosomal frameshifting. *The Ribosome*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY pp. 233–248.
9. Karacostas, V., Wolffe, E. J., Nagashima, K., Gonda, M. A. & 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.
10. Park, J. & Morrow, C. D. (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.
11. Shehu-Xhilaga, M., Crowe, S. M. & 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.
12. Kang, H. (1998). Direct structural evidence for formation of a stem-loop structure involved in ribosomal frameshifting in human immunodeficiency virus type 1. *Biochim. Biophys. Acta*, **1397**, 73–78.
13. Du, Z., Giedroc, D. P. & Hoffman, D. W. (1996). Structure of the autoregulatory pseudoknot within the gene 32 messenger RNA of bacteriophages T2 and T6: a model for a possible family of structurally related RNA pseudoknots. *Biochemistry*, **35**, 4187–4198.
14. 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.
15. Dinman, J. D., Richter, S., Plant, E. P., Taylor, R. C., Hammell, A. B. & Rana, T. M. (2002). The frameshift signal of HIV-1 involves a potential intramolecular triplex RNA structure. *Proc. Natl Acad. Sci.* **99**, 5331–5336.
16. Giedroc, D. P., Theimer, C. A. & Nixon, P. L. (2000). Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting. *J. Mol. Biol.* **298**, 167–185.
17. Los Alamos National Laboratory (2002). A compilation and analysis of nucleic acid and amino acid sequences. *Human Retroviruses and AIDS edit*, Los Alamos National Laboratory.
18. Mathews, D. H., Sabina, J., Zuker, M. & Turner, D. H. (1999). Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *J. Mol. Biol.* **288**, 911–940.
19. Reil, H., Kollmus, H., Weidle, U. H. & Hauser, H. (1993). A heptanucleotide sequence mediates ribosomal frameshifting in mammalian cells. *J. Virol.* **67**, 5579–5584.
20. Parkin, N. T., Chamorro, M. & Varmus, H. E. (1992). Human immunodeficiency virus type 1 gag–pol frameshifting is dependent on downstream mRNA secondary structure: demonstration by expression *in vivo*. *J. Virol.* **66**, 5147–5151.
21. Wang, Y., Wills, N. M., Du, Z., Rangan, A., Atkins, J. F., Gesteland, R. F. & Hoffman, D. W. (2002). Comparative studies of frameshifting and non-frameshifting RNA pseudoknots: a mutational and NMR investigation of pseudoknots derived from the bacteriophage T2 gene 32 mRNA and the retroviral gag–pro frameshift site. *RNA*, **8**, 981–996.
22. Chen, X., Kang, H., Shen, L. X., Chamorro, M., Varmus, H. E. & Tinoco, I., Jr (1996). A characteristic bent conformation of RNA pseudoknots promotes −1 frameshifting during translation of retroviral RNA. *J. Mol. Biol.* **260**, 479–483.
23. Shen, L. X. & Tinoco, I., Jr (1995). The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol.* **247**, 963–978.
24. Su, L., Chen, L., Egli, M., Berger, J. M. & Rich, A. (1999). Minor groove RNA triplex in the crystal structure of a ribosomal frameshifting viral pseudoknot. *Nature Struct. Biol.* **6**, 285–292.
25. Michiels, P. J., Versleijen, A. A., Verlaan, P. W., Pleij, C. W., Hilbers, C. W. & Heus, H. A. (2001). Solution structure of the pseudoknot of SRV-1 RNA, involved in ribosomal frameshifting. *J. Mol. Biol.* **310**, 1109–1123.
26. Brierley, I., Digard, P. & Inglis, S. C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell*, **57**, 537–547.
27. Liphardt, J., Napthine, S., Kontos, H. & Brierley, I. (1999). Evidence for an RNA pseudoknot loop-helix interaction essential for efficient −1 ribosomal frameshifting. *J. Mol. Biol.* **288**, 321–335.
28. Napthine, S., Liphardt, J., Bloys, A., Routledge, S. & Brierley, I. (1999). The role of RNA pseudoknot stem 1 length in the promotion of efficient −1 ribosomal frameshifting. *J. Mol. Biol.* **288**, 305–320.
29. Marczinke, B., Fisher, R., Vidakovic, M., Bloys, A. J. & Brierley, I. (1998). Secondary structure and mutational analysis of the ribosomal frameshift signal of Rous sarcoma virus. *J. Mol. Biol.* **284**, 205–225.
30. Pearlman, D. A. (1995). *AMBER 4.1*, University of California, San-Francisco.
31. Brierley, I., Rolley, N. J., Jenner, A. J. & Inglis, S. C. (1991). Mutational analysis of the RNA pseudoknot component of a coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* **220**, 889–902.
32. ten Dam, E. B., Verlaan, P. W. & Pleij, C. W. (1995). Analysis of the role of the pseudoknot component in the SRV-1 gag–pro ribosomal frameshift signal: loop lengths and stability of the stem regions. *RNA*, **1**, 146–154.
33. Chamorro, M., Parkin, N. & Varmus, H. E. (1992). An RNA pseudoknot and an optimal heptameric shift site are required for highly efficient ribosomal frameshifting on a retroviral messenger RNA. *Proc. Natl Acad. Sci.* **89**, 713–717.
34. ten Dam, E., Brierley, I., Inglis, S. & Pleij, C. (1994). Identification and analysis of the pseudoknot-containing gag–pro ribosomal frameshift signal of simian retrovirus-1. *Nucleic Acids Res.* **22**, 2304–2310.
35. Gao, F., Bailes, E., Robertson, D. L., Chen, Y., Rodenburg, C. M., Michael, S. F. *et al.* (1999). Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. *Nature*, **397**, 436–441.
36. Simon, F., Mauclere, P., Roques, P., Loussert-Ajaka, I., Muller-Trutwin, M. C., Saragosti, S. *et al.* (1998). Identification of a new human immunodeficiency virus type 1 distinct from group M and group O. *Nature Med.* **4**, 1032–1037.
37. Hill, M. K., Shehu-Xhilaga, M., Crowe, S. M. & 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.
38. Louis, J. M., Clore, G. M. & Gronenborn, A. M.

(1999). Autoprocessing of HIV-1 protease is tightly coupled to protein folding. *Nature Struct. Biol.* **6**, 868–875.
39. Demirov, D. G., Orenstein, J. M. & Freed, E. O. (2002). The late domain of human immunodeficiency virus type 1 p6 promotes virus release in a cell type-dependent manner. *J. Virol.* **76**, 105–117.
40. Brierley, I. (1995). Ribosomal frameshifting viral RNAs. *J. Gen. Virol.* **76**, 1885–1892.
41. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K. & Pease, L. R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene*, **77**, 51–59.
42. Jordan, M., Schallhorn, A. & Wurm, F. M. (1996). Transfecting mammalian cells: optimization of critical parameters affecting calcium-phosphate precipitate formation. *Nucleic Acids Res.* **24**, 596–601.
43. Eustice, D. C., Feldman, P. A., Colberg-Poley, A. M., Buckery, R. M. & Neubauer, R. H. (1991). A sensitive method for the detection of beta-galactosidase in transfected mammalian cells. *Biotechniques*, **11**, 739–740, see also 742–743.
44. Knapp, G. (1989). Enzymatic approaches to probing of RNA secondary and tertiary structure. *Methods Enzymol.* **180**, 192–212.

***Edited by D. E. Draper***

*(Received 25 April 2003; received in revised form 11 June 2003; accepted 12 June 2003)*