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

DOMINIC DULUDE,1,2,3 GABRIEL THÉBERGE-JULIEN,1,3 LÉA BRAKIER-GINGRAS,1,4 and NIKOLAUS HEVEKER1,2,4
1Département de Biochimie, Université de Montréal, Montréal H3T 1J4, Québec, Canada
2Centre de Recherche, Hôpital Sainte-Justine, Montréal H3T 1C5, Québec, Canada

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 Escherichia 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 frameshifting and may permit the development of lead compounds of therapeutic interest.

Keywords: HIV-1 programmed -1 ribosomal frameshift; peptide library

INTRODUCTION
The human immunodeficiency virus of 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 (for review, see Brierley and Pennell 2001; Brierley and 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 and Morrow 1991; Karacostas et al. 1993; Shehu-Xhilaga et al. 2001; Telenti et al. 2002; Dulude et al. 2006). All of 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, UUUUUUA, 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 and 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 and 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) (for review, see Léger et al. 2007).

Kollmus et al. (1996), who replaced the HIV upper stem–loop with IRE, 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 twofold 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, the Rev-response element 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

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 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 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 gray box) to define and isolate events with a decreased GFP/RFP ratio, which corresponds to a reduced frameshift efficiency.
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 (0.1–2 μg/mL), a protein synthesis inhibitor known to increase frameshift efficiency (Thompson et al. 2002), yielded a higher GFP/RFP ratio (up to threefold 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 with 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%–4.8%) observed for HIV-1 frameshift efficiency with a luciferase reporter in bacteria (Léger et al. 2004; Plant and 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 and 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 16 h of peptide induction, a time point that 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 ~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 nontoxic 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, NRRRRRSSRHRHRR. 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 1150 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 1150 clones compared with an average decrease to 87% for a group of unselected clones. Inspection of these 1150 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

<table>
<thead>
<tr>
<th>Process</th>
<th>Number of events</th>
<th>Percent of parental events</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analyzed</td>
<td>2.2 × 10^7</td>
<td>—</td>
</tr>
<tr>
<td>Gated</td>
<td>4.4 × 10^4</td>
<td>0.2</td>
</tr>
<tr>
<td>Growth on agar plates</td>
<td>≈5 × 10^3</td>
<td>11</td>
</tr>
</tbody>
</table>

The peptide library was expressed in bacteria bearing the reporter plasmid and the reporter fluorescence was analyzed by flow cytometry after 16 h 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.
clones decreased the frameshift efficiency by only 13%. The 20 most efficient clones (2% of the clones) decreased the frameshift efficiency by 57%. By 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 1150 clones analyzed, in comparison with unselected peptides. The data show a shift of fluorescence corresponding to a decrease in 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 reassessed 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 with 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 16 h of peptide expression for the 20 most efficient clones compared with a group of nonselected clones is also shown in Figure 4B. The same level of RFP fluorescence is observed for the selected and the nonselected...
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 of 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 of the peptides have a different sequence. Also, residues at positions 7–10 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 7–10 form 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 nonselected 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 toward polyaginine sequences. Rather, single, nonselected sequences with high arginine content (RRR...RRRRRRRRRRRR and RRH...RRRRRRRRRRRR) 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 four other engineered

![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 OD_{600} of 0.4 and 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.]

<table>
<thead>
<tr>
<th>Clone No.</th>
<th>Peptide sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NHRHRRSSR</td>
</tr>
<tr>
<td>2</td>
<td>RRHRHRRSSR</td>
</tr>
<tr>
<td>3</td>
<td>RRRSSRHRPRN</td>
</tr>
<tr>
<td>4</td>
<td>RRNRRSSRHRPRR</td>
</tr>
<tr>
<td>5</td>
<td>RHRSSRHRPRR</td>
</tr>
<tr>
<td>6</td>
<td>RHRSSRHRPRR</td>
</tr>
<tr>
<td>7</td>
<td>RRRSHRRHRPRR</td>
</tr>
<tr>
<td>8</td>
<td>SRRHRHRRRRRRRR</td>
</tr>
<tr>
<td>9</td>
<td>HHRHRRRRRRRRR5</td>
</tr>
<tr>
<td>10</td>
<td>RRRRRHRRTSSR</td>
</tr>
<tr>
<td>11</td>
<td>NRRRRHRRTSSR</td>
</tr>
<tr>
<td>12</td>
<td>RRRRRHRRTSSR</td>
</tr>
<tr>
<td>13</td>
<td>RRSSRRRRRPRR</td>
</tr>
<tr>
<td>14</td>
<td>RRSSRRRRRPRR</td>
</tr>
<tr>
<td>15</td>
<td>SRRRRRRRRRRRNR</td>
</tr>
<tr>
<td>16</td>
<td>SRRRRRRRRRRRNR</td>
</tr>
<tr>
<td>17</td>
<td>HRSSRRRRRHRPRR</td>
</tr>
<tr>
<td>18</td>
<td>HHRHRRRRRRHRPRR</td>
</tr>
<tr>
<td>19</td>
<td>RRRRRRRRRRRRNR</td>
</tr>
<tr>
<td>20</td>
<td>RRRRRRRRRRRRPRR</td>
</tr>
</tbody>
</table>
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 transactivation response element of the bovine immunodeficiency virus. The frame-shift 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 a eukaryote expression and reporter system in HEK293T cells. The scaffold used was derived from Xu et al. (2001), consisting of 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 10 tested peptide

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

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>1 (%)</th>
<th>2 (%)</th>
<th>3 (%)</th>
<th>4 (%)</th>
<th>5 (%)</th>
<th>6 (%)</th>
<th>7 (%)</th>
<th>8 (%)</th>
<th>9 (%)</th>
<th>10 (%)</th>
<th>11 (%)</th>
<th>12 (%)</th>
<th>13 (%)</th>
<th>14 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selected peptides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>65</td>
<td>70</td>
<td>55</td>
<td>75</td>
<td>80</td>
<td>75</td>
<td>65</td>
<td>95</td>
<td>100</td>
<td>60</td>
<td>50</td>
<td>80</td>
<td>75</td>
<td>55</td>
</tr>
<tr>
<td>H</td>
<td>15</td>
<td>25</td>
<td>25</td>
<td>5</td>
<td>20</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>10</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>20</td>
<td>10</td>
<td>5</td>
<td>30</td>
<td>5</td>
<td>0</td>
<td>35</td>
<td>20</td>
<td>15</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Unselected peptides</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>66</td>
<td>52</td>
<td>72</td>
<td>62</td>
<td>69</td>
<td>59</td>
<td>83</td>
<td>72</td>
<td>72</td>
<td>55</td>
<td>69</td>
<td>76</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td>H</td>
<td>10</td>
<td>34</td>
<td>10</td>
<td>17</td>
<td>14</td>
<td>18</td>
<td>28</td>
<td>10</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>3</td>
<td>14</td>
<td>10</td>
</tr>
<tr>
<td>S</td>
<td>21</td>
<td>14</td>
<td>17</td>
<td>21</td>
<td>14</td>
<td>14</td>
<td>3</td>
<td>3</td>
<td>14</td>
<td>14</td>
<td>3</td>
<td>17</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>N</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

Boldface numbers indicate over- and underrepresentation of several residues. (Binomial distribution tests were performed for each position and cumulative probabilities of $P < 0.05$ were seen as significant.)

aFrequency of different amino acid residues at the different positions of 20 active sequences (see Table 2 for peptide sequences).

bFrequency of different amino acid residues at the different positions of 29 random picked sequences of the library.

![Figure 5](https://example.com/figure5.png)

**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 and Narayana 1998), generating derivatives of p3RFP-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.
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 (Biswa et al. 2004; Dulude et al. 2006), in the ribosomes (Léger et al. 2004), or in elongation factors (Dinman and 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 1150 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 nonbasic 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 and 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 nonspecific counterion effect. Indeed, when examining the sequences of randomly chosen nonselected 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 7–10 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 of 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, nonpeptidic 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 GFP coding sequence, such that its expression requires a programmed -1 frameshift. The vector also contains three copies of the 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 pRNALuc2 (Belanger et al. 2002), which 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 pGFPemd-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 rfp promoter of Nsi–XhoI-digested pRNAluc2, generating the pGFP construct. The frameshift region of HIV-1 group M subtype B (nucleotides 2075–2143 of HIV-1NLBR, GenBank accession no. K03455) was introduced in 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 pRNAluc2 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 and Glick 2002) was introduced between the NsiI and XhoI 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 KpnI restriction site at the 5’ and 3’ end of the RFP gene. The KpnI-digested RFP fragment was inserted in the digested KpnI 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 and inserted as a fragment with a Nhel restriction site at the 5’ and 3’ end. This Nhel-digested fragment was then inserted in the digested Nhel 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 HIV-1, the KpnI restriction 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 and inserted as a fragment with a Nhel restriction site at the 5’ and 3’ end. This Nhel-digested fragment was then inserted in the digested Nhel site of pRFP-HIV(-1)-GFP, generating p3RFP-HIV(-1)-GFP.

Plasmid pACYC-Thio was used to express a combinatorial arginin-rich peptide 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’-GACTGACTGTTCCG(MRT)14GTTGCTCTAGCAGT CAG-3’ (Sigma Genosys), containing 14 degenerate codons. For each of these codons, M is a C mixture in a 1:3 ratio, and R is an A or 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’-CAGGGAATTCTGCTGATC-3’), using the Sequana-se 2.0 (USB) as a DNA polymerase. These cassettes were digested with AvaII, ligated into RsrII-digested pACYC-Thio, and used to transform E. coli XL-10 Gold (Strategene).

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 transfectants, 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 arabinine (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 of PBS. Bacterial pellets were resuspended in 250 µL of PBS and transferred into 96-well plates and RFP and GFP

Design of the combinatorial arginin-rich peptide library

Plasmid pACYC-Thio was used to express a combinatorial arginin-rich peptide 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’-GACTGACTGTTCCG(MRT)14GTTGCTCTAGCAGT CAG-3’ (Sigma Genosys), containing 14 degenerate codons. For each of these codons, M is a C mixture in a 1:3 ratio, and R is an A or 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’-CAGGGAATTCTGCTGATC-3’), using the Sequana-se 2.0 (USB) as a DNA polymerase. These cassettes were digested with AvaII, ligated into RsrII-digested pACYC-Thio, and used to transform E. coli XL-10 Gold (Strategene).
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 and XhoI restriction sites of the pLPC expression vector (Serrano et al. 1997). Oligonucleotide cassettes coding for the selected peptides (peptides 1–10 of Table 2) were then inserted between the Bspel and EcoRI restriction sites of this construct. Transfection of the resulting plasmids led to the expression of the peptide 5′-GEFLIVKSG-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 (Grenzmann et al. 1998; Dulude et al. 2006) into 293T cells. The day before transfection, 3 × 10^5 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 cotransfected 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.

**SUPPLEMENTAL DATA**

Supplemental material can be found at http://www.rnajournal.org.

**ACKNOWLEDGMENTS**

This study was supported by a grant from the Canadian Institutes of Health Research (CIHR) to L.B.-G. and N.H. D.D. acknowledges a postdoctoral fellowship from the Groupe de Recherche Universitaire sur le Médicament (GRUM) de l’Université de Montréal and from CIHR and N.H. 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.

Received October 22, 2007; accepted February 4, 2008.

**REFERENCES**


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

Dominic Dulude, Gabriel Théberge-Julien, Léa Brakier-Gingras, et al.

RNA 2008 14: 981-991

Supplemental Material
http://rnajournal.cshlp.org/content/suppl/2008/03/27/rna.887008.DC1

References
This article cites 42 articles, 16 of which can be accessed free at:
http://rnajournal.cshlp.org/content/14/5/981.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to RNA go to:
http://rnajournal.cshlp.org/subscriptions

Copyright © 2008 RNA Society