

A-1 frameshift in the HIV-1 env gene is enhanced by arginine deficiency via a hungry codon mechanism

By: Babatunde Olubajo and [Ethan Will Taylor](#)

Babatunde Olubajo and Ethan Will Taylor. A-1 frameshift in the HIV-1 env gene is enhanced by arginine deficiency via a hungry codon mechanism. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*. Volume 579, Issues 1–2, 11 November 2005, Pages 125-132. <https://doi.org/10.1016/j.mrfmmm.2005.02.018>



This work is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#).

***© 2005 Elsevier B.V. Reprinted with permission. This version of the document is not the version of record. ***

Abstract:

Ribosomal frameshifting is used by various organisms to maximize protein coding potential of genomic sequences. It is commonly exploited by RNA viruses to overcome the constraint of their limited genome size. Frameshifting requires specific RNA structural features, such as a suitable heptanucleotide “slippery” sequence and an RNA pseudoknot. Previous genomic analysis of HIV-1 indicated the potential for several hidden genes encoded through frameshifting; one of these, overlapping the envelope gene, has an RNA pseudoknot just downstream from a slippery sequence, AAAAAGA that features an adenine quadruplet prior to a potential hungry arginine codon (AGA). This *env*-frameshift (*env-fs*) gene has been shown to encode a truncated glutathione peroxidase homologue, with both antioxidant and anti-apoptotic activities in transfected cells. Using a dual reporter cell-based frameshift assay, we demonstrate that the *env-fs* frameshift sequence is active in vitro. Furthermore, in arginine deficient media, *env-fs* frameshifting increased over 100% ($p < 0.005$), consistent with the hypothesized hungry codon mechanism. As a response to arginine deficiency, increased expression of the antioxidant viral GPx gene (*env-fs*) by upregulation of frameshifting could be protective to HIV-infected cells, as a countermeasure to the increased oxidative stress induced by arginine deficiency (because NO is a known scavenger of hydroxyl radical).

Keywords: –1 Frameshift | *env*-Frameshift | HIV-1 | Hungry codon | Arginine deficiency

Article:

1. Introduction

Several potential virally encoded selenoproteins have been identified in the HIV-1 genome by computational analysis [1]. One of the proposed genes is located overlapping the gp41 coding region of the envelope (*env*) gene of HIV-1; hence, the putative selenoprotein it encodes would be expressed as a-1 frameshift fusion protein variant of the HIV-1 *env* gp120 protein, yielding an initial product of about 130 kDa. Just downstream of the predicted frameshift site, a non-ideal

HIV-1 protease cleavage site was identified at the sequence level, which suggested that this selenoprotein module might be expressed both as a small independent 9 kDa protein, as well as a component of the 130 kDa fusion protein of gp120 [2], [3], [4].

The hypothetical protein encoded by this gene, which was named *env*-frameshift (*env-fs*) was shown to have highly significant sequence similarity to the active site regions of selenium-dependent mammalian glutathione peroxidases [2], [4]. The similarity (i.e., multiple alignment score) of the *env-fs* protein sequence to an aligned set of mammalian glutathione peroxidases was 6.3 standard deviations higher than the average score of randomly shuffled *env-fs* sequences [4].

This sequence similarity provided a basis for molecular modeling of the proposed *env-fs* encoded protein as a viral glutathione peroxidase (GPx), which was modeled based on the X-ray crystal structure of bovine GPx. Some outer structural regions of the mammalian GPx were not intrinsic to the viral GPx, but the important core structural and catalytic elements, as well as the orientation of the active site, were nonetheless evident in the viral protein model [4].

The proposed *env-fs* protein was later cloned and expressed in mammalian cell lines. Two cell lines were stably transfected with an *env-fs* construct; and both exhibited a significantly increase in GPx activity relative to control cell lines [4]. Furthermore, the HIV-1 GPx protein was found to possess cytoprotective and anti-apoptotic properties in transfected cells exposed to oxidant stimuli [5]. These results suggest that the virally encoded GPx gene may serve to protect HIV infected cells from apoptosis in order to create conditions that facilitate viral replication and survival.

The idea of a-1 frameshift in the HIV-1 *env* gene that would produce a viral GPx module is supported by the presence of a pseudoknot and slippery sequence [2], [3] two RNA structural elements that are generally required for ribosomal frameshifting to occur (Fig. 1). Although frameshifting is not completely understood, current opinion is that a frameshift is facilitated when a ribosome encounters a pseudoknot in the mRNA during translation. The process of the ribosome unraveling the pseudoknot leads to a pause in translation, which in turn facilitates slippage of the ribosome at a suitable slippery sequence just upstream of the pseudoknot. The ribosome can slip either forward (+ change) or backward (- change), leading to the continuation of protein translation in one of the two overlapping reading frames.

Frameshifting can be regulated by various factors; one of the more interesting of these is the presence of a “hungry codon”, which can enhance frameshifting. Because they are associated with rare tRNA isoacceptors, hungry codons can induce a pause in translation when they are encountered by the ribosome, an effect which would be increased by a deficiency of, their cognate amino acid. Hungry codons have been shown to induce or enhance frameshifting, and an amino acid effect has been demonstrated [6], [7], [8].

The -1 frameshift in the *env* gene of HIV was predicted to involve an arginine hungry codon mechanism [3]. The predicted slippery sequence, (A AAA AGA in the zero frame, slipping to AAA AAG A in the -1 frame), would dictate P-site slippage on the run of four A bases, facilitated by a pause at the hungry arginine codon (AGA) in the zero frame [3]. The implications of a hungry arginine codon can be related to the role of arginine in oxidative stress.

Arginine deficiency leads to a decrease in nitric oxide, which in turn causes an increase in hydroxy radical oxidative stress [9]. Thus, production of a defensive antioxidant viral GPx protein could be controlled by arginine deficiency through its frameshifting mechanism. In this study, the proposed HIV-1 *env*-frameshift sequence will be examined in vitro, as well as the potential for hungry codon induced frameshifting.

2. Materials and methods

2.1. Frameshift assay constructs

The luciferase gene was cloned from the T7-luc vector (Promega) by polymerase chain reaction, adding Sal I and Not I restriction sites to the 5' and 3' ends of the cloned gene respectively. Two versions of the luciferase gene were cloned and they differed in that two extra bases were added to the 5' end of one of the clones to embed the luciferase gene in a -1 frameshift conformation. The second luciferase gene is cloned in a 0 frame.

PCI vector (Promega) and the cloned luciferase genes were digested with Sal I and Not I. The luciferase genes were ligated into the digested PCI vector to give PCL (+) and PCL (-) vector. PCL (-) is the -1 luciferase vector while PCL (+) is the 0 frame luciferase vector.

The green fluorescent protein gene (GFP) was digested from PEGFP-CI (Clontech) with Nhe I and Sal I. The digested GFP gene was incorporated into the PCL (+) and PCL (-) vectors digested with Nhe I and Sal I to form PGL (+) and PGL (-) vectors.

The envelope frameshift sequence was cloned from the PBH10 clone of the HIV-1 genome (NIH AIDS Reagent Program) by polymerase chain reaction, with the addition of Xho I and Sal I sites at the 5' and 3' ends of the cloned *env*-frameshift sequence. A stop codon TAA was engineered directly after the frameshift sequence in the zero frame of GFP to truncate translation of unframeshifted protein.

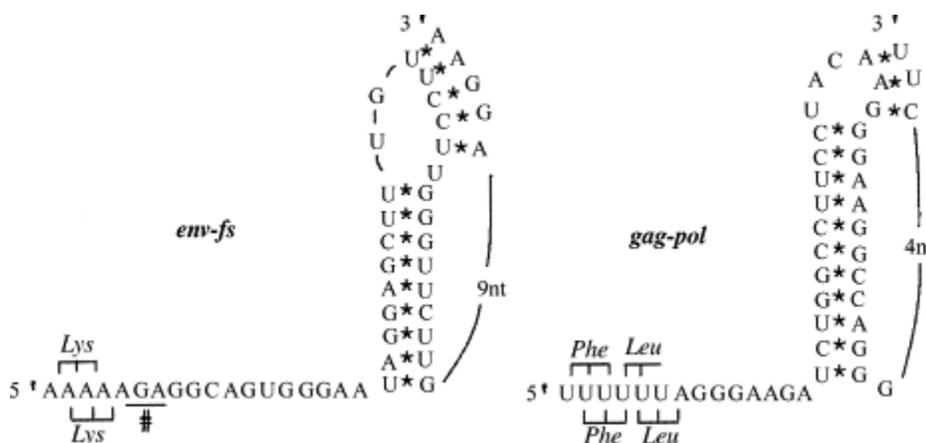


Fig. 1. Frameshift sequences in the envelope (proposed) and gag-pol [10] of HIV-1. The underlined AGA arginine codon (indicated by #) in the *env-fs* sequence is a potential hungry codon [3].

The PGL (+) and PGL (-) vectors along with the envelope frameshift sequence were digested with Xho I and Sal I and ligated to give the assaying vectors PGenVL (+) and PGenVL (-). An identical process was also used to make PGgpL (+) and PGgpL (-) vectors which contain the known gag-pol frameshift sequence of HIV [10]. The sequence of all vectors was confirmed by dideoxy sequencing (Fig. 1, Fig. 2).

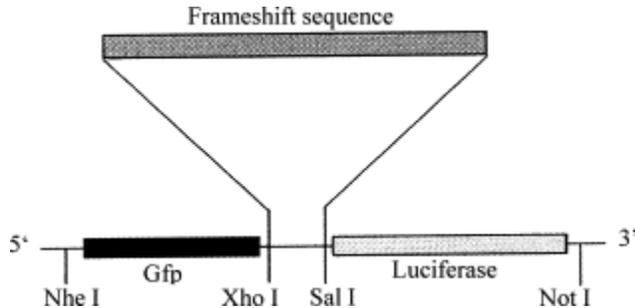


Fig. 2. Map of frameshift assay construct.

2.2. Transfection of cell lines

Two hundred and ninety-three T cells were grown in Dubelco minimal essential medium with 2 mM l-glutamine, 1.5 g/L sodium bicarbonate and 1.0 mM sodium pyruvate. The final media volume was adjusted to contain 10% fetal bovine serum. Cell lines were grown in 35 mm diameter, 6-well plates to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 2 μ g of appropriate vector for 6 h. Cells were transiently transfected with PGenVL (+) and PGenVL (-). PGgpL (+) and PGgpL (-) were used to transiently transfect cell lines as controls. The cells were incubated at 37 °C with 5% CO₂ for 48 h.

2.3. Cell culture and l-arginine supplementation

Two hundred and ninety-three T cells were grown in Dubelco minimal essential medium, which deviated in a few components, depending on the study. The cells were grown in 24-well plates. For arginine deficiency studies, 293-T cells were grown in DMEM lacking l-Arginine, and l-lysine, and supplemented appropriately as the experiments dictate. The media for the l-arginine studies, was supplemented with 10% volume FBS and 0.8 mM l-lysine. The samples were starved of l-arginine for 24 h. The samples were then supplemented accordingly with either 0.1 or 1 mM of l-arginine. Cells were incubated for 48 h at 37 °C with 5% CO₂ with media changes and l-arginine supplementation every 24 h. Cell lines were grown to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 0.5 μ g of PGenVL (-). The transfected cells are incubated at 37 °C with 5% CO₂ for 48 h with l-arginine supplementation at 24 h.

2.4. Cell based frameshift assay

Transfected cell lines were lysed by incubating with 1ml of glo-lysis buffer (Promega) per well for 5 min. Crude extracts were centrifuged and supernatant was recovered. The supernatant was assayed for reporter gene activity.

Luciferase activity was quantitated in a Turner Biosystems TD 20/20 luminometer or a FLUOstar Optima Galaxy plate reader. For the TD20/20 luminometer, luciferase assay involves adding 50 μ l of the cell lysates to 100 μ g of Promega luciferase assay buffer, which contains luciferin as the active luminescence agent. The reaction mixture was placed in the luminometer and luminescence was read immediately. Relative luminescence intensity was read for ten seconds following a programmed 2-s delay. Background luminescence is normally 0.000–0.020 relative light units.

As mentioned earlier, a BMG Labtechnologies FLUOstar Optima plate reader was also used to measure luminescence. Sample volumes of 20 μ l were pipetted into an opaque/white 96-well luminescence plate. The machine was programmed to dispense 40 μ l of luciferase assay reagent into the well of each sample to be tested. Luminescence of each sample is measured over a 12 s period.

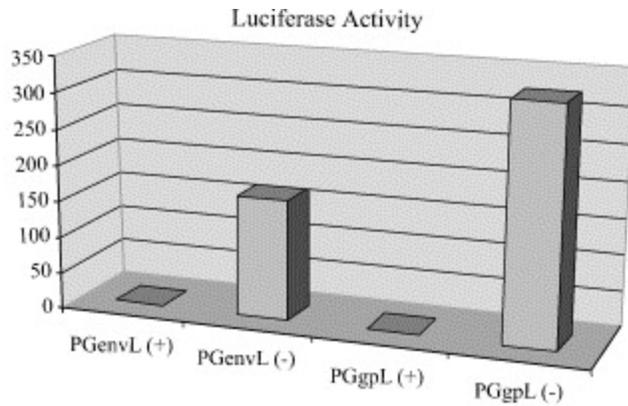
Fluorescence of GFP is measured using a Shimadzu RF-5301 or FLUOstar Optima Galaxy fluorescent plate reader. Cell lysate samples were excited at 488 nm wavelength and the relative intensity of emission at 511 nm was recorded.

Comparing fluorescence of GFP and luminescence of luciferase quantitates frameshifting. Luciferase translation measures frameshifting while GFP translation standardizes the assay. Therefore, frameshifting is represented by the amount of luciferase luminescence in relation to GFP fluorescence. We define the *frameshifting factor* as the ratio of (Luciferase activity)/(GFP fluorescence).

3. Results and discussion

3.1. Frameshift assays: comparison of env-fs to known gag-pol frameshift site

The *env*-frameshift sequence exhibited frameshifting activity upon analysis with the frameshift assay. There was significant luminescence emitted from cell lysates expressing the *env*-frameshift sequence test vector PGenV (-), as compared to mock-transfected cells with no vector DNA. The mock-transfected cells continuously showed a good baseline, giving relative luminescence values of between 0.000 and 0.007 relative light units in comparison with 10,000-fold increase in luminescence of assay vector cell lysates. Lysates of cells transfected with the *env*-frameshift assay vector PGenV (-) also presented a 10,000-fold increase in relative luminescence in comparison with cells transfected with the control vector PGenV (+) as shown in Fig. 3. Cell lysates of cells transfected with the PGenV (+) construct, which contains the test *env* sequence and the luciferase gene in a 0 frame to the upstream GFP gene, had relative luminescence values similar to the mock transfected cells. Cells expressing the control vector PGgpL (-), containing the known *gag-pol* frameshift sequence also showed a 10,000-fold increase in luciferase activity over mock transfected cells, and 1000-fold increase over cells expressing the PGgpL (+) vector. The PGgpL (+) vector showed almost no luciferase activity. Luminescence quantitated from PGenV (+) expressing vectors was almost identical to mock transfected cells and was practically inexistent.



Construct	Luciferase activity in Relative light Units Average +/- Standard Deviation
PGenV (+)	0.011 +/- 0.004
PGenV (-)	166 +/- 20.4
PGgpL (+)	0.61 +/- 0.16
PGgpL (-)	324 +/- 65.0

Fig. 3. Average luminescence from cell lysates expressing frameshift constructs measured on a TD 20/20 luminometer ($P < 0.001$ for PGenV (+) and PGenV (-), $P < 0.001$ for PGgpL (+) and PGgpL (-)).

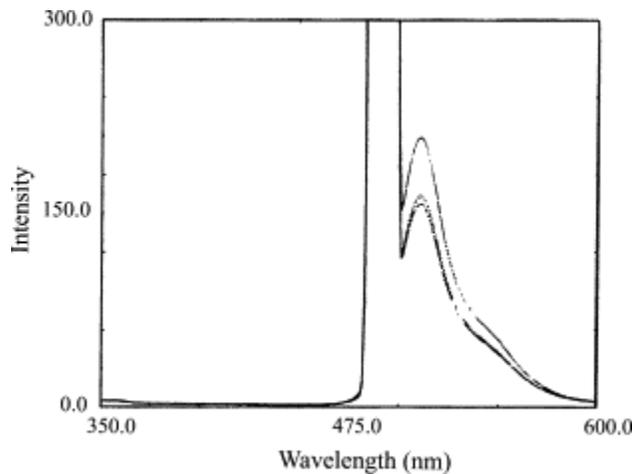


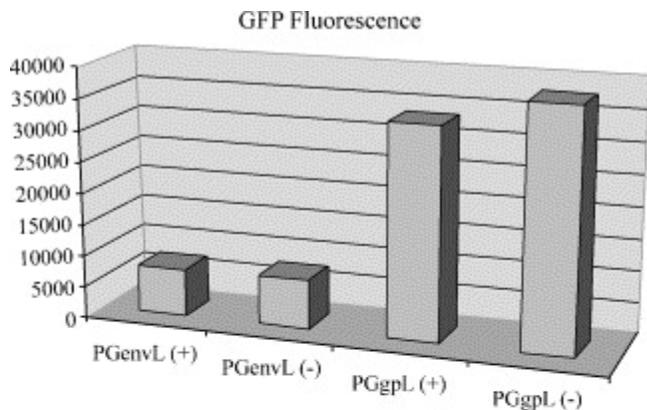
Fig. 4. Emission spectra of GFP fluorescence from representative triplicate cell lysate samples of PGgpL (-) excited at 488 nm measured on the Shimadzu RF-5301. Emission peak was recorded at 511 nm.

GFP fluorescence was measured in the Shimadzu RF-5301 and the POLARstar OPTIMA plate reader with a 485 nm excitation filter and 510 nm emission filter. The peak of relative GFP

fluorescence for assayed cell lysates was at 511 nm when an emission spectrum was measured as shown in Fig. 4.

The relative fluorescence of GFP in cell lysates of transfected cells was fairly strong with a 10-fold increase in fluorescence over auto fluorescence of mock-transfected cell lysates. GFP expression appeared to be similar in the assayed lysates of cells transfected with vectors of similar test frameshift sequences. PGenvL (+) and PGenvL (-) had similar fluorescence profiles, which were much less, than the GFP fluorescence of PGgpL (+) and PGgpL (-). The lower but significant GFP fluorescence from PGenvL (+) and PGenvL (-) vectors could be attributed to low transfection efficiency for those vectors.

The relative fluorescence of GFP in cell lysates of transfected cells was fairly strong with a 10-fold increase in fluorescence over auto fluorescence of mock-transfected cell lysates. GFP expression appeared to be similar in the assayed lysates of cells transfected with vectors of similar test frameshift sequences as shown in Fig. 5. PGenvL (+) and PGenvL (-) had similar fluorescence profiles, which were much less than the GFP fluorescence of PGgpL (+) and PGgpL (-). The lower but significant GFP fluorescence from PGenvL (+) and PGenvL (-) vectors could be attributed to low transfection efficiency for that vector.



Vector	GFP Fluorescence
	Average +/- Standard Deviation
PGenvL (+)	7370 +/- 449
PGenvL (-)	7800 +/- 979
PGgpL (+)	33500 +/- 4980
PGgpL (-)	38000 +/- 11800

Fig. 5. Average GFP fluorescence measured for frameshift assay constructs.

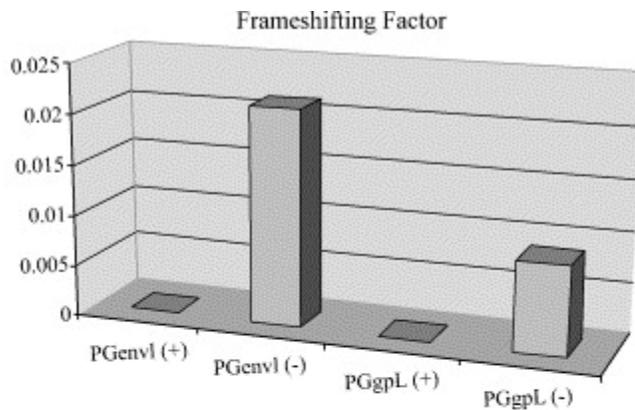
The frameshift factor of the frameshift assay constructs, was calculated as previously described. PGenvL (+) and PGgpL (+) showed virtually no frameshifting due to relatively inexistent luciferase translation. The frameshift factor of PGenvL (-) is consistently two times higher than for the known frameshift sequence PGgpL (-).

3.2. Effects of l-arginine deficiency on frameshifting

The effect of arginine deficient media on frameshifting facilitated by the *env*-frameshift sequence was assessed. Wells supplemented with mM media concentration of arginine are compared with wells supplemented with 0.1 mM arginine. It was found that with the PGenVL (-) construct there was a significant increase in frameshifting under arginine deficient conditions ($p < 0.005$). In comparison the two concentrations of arginine had no significant effects on frameshifting on the PGgpL (-) frameshift sequence (Fig. 7).

4. Conclusions

The presence of hidden genes expressed by frameshifting in genomes of viruses and other organisms is a complex topic that requires careful analysis. Whether embedded overlapping genes arise from recombinant events or a gradual random and convergent evolutionary process, the elucidation of their mechanisms of translation, which may not follow conventional knowledge, requires a cautious but unprejudiced approach. Because of the inherent inefficiency of frameshifting, the protein products of such genes are certain to be less abundant and therefore harder to detect than conventional protein products.



Construct	Frameshift Factor
PGenVL (+)	1.48×10^{-5}
PGenVL (-)	0.02
PGgpL (+)	1.8×10^{-3}
PGgpL (-)	0.01

Fig. 6. Frameshift factor of frameshift assay constructs. The values represent averages of triplicate experiments that have been repeated several times. ($P < 0.001$ for PGenVL (+) and PGenVL (-), $P < 0.001$ for PGgpL (+) and PGgpL (-)).

The *env-fs* gene of HIV is such a hidden gene with an even more elusive protein product, because of the compounded inefficiency of expressing a selenoprotein, which requires a stop codon, UGA, to be reprogrammed as a sense codon for Sec, a process which can be as or more

inefficient than ribosomal frameshifting. We have taken the genomic approach of cloning the gene and showing that the encoded product is functional [4], and, in the present study, demonstrating that the frameshift required for its expression is active in vitro.

We have shown that the viral sequence spanning the predicted pseudoknot and slippery sequence of the *env-fs* gene are capable of inducing a change in reading frame during translation, as evidenced by expression of the downstream luciferase gene engineered in the PGenVL (-) construct (Fig. 6). Furthermore, it is of considerable functional significance that the expression of the putative HIV-1 GPx protein encoded by *env-fs* is under the translational regulation of an amino acid, arginine, whose codon was previously predicted to serve as a hungry codon in the *env-fs* frameshift sequence [3] (Fig. 7).

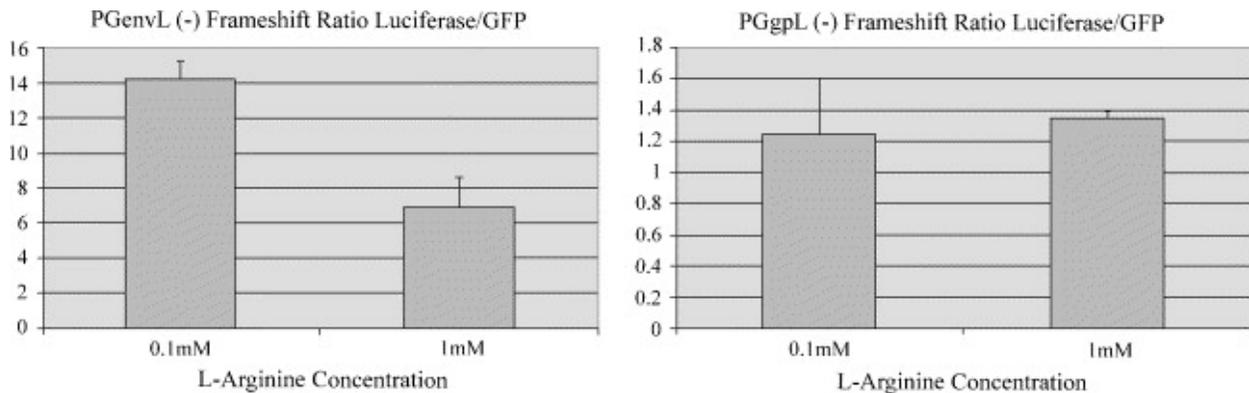


Fig. 7. Frameshift factor of PGenVL (-) and PGgpL, PGenVL (-) shows an increase in frameshift factor with arginine deficient media (left panel) while the comparable construct (right panel) for the known HIV-1 *gal-pol* frameshift site, PGgpL (-), shows no significant change in frameshift factor in response to differences in arginine concentration ($P < 0.005$ for PGenVL (-), $P > 0.1$ for PGgpL (-)).

In this case, the increase in frameshifting in response to arginine deficiency that we have demonstrated would be expected to lead to increased expression of a viral gene, *env-fs*, which has been previously shown to encode antioxidant (GPx) activity, and which confers protection against pro-apoptotic oxidant stimuli in transfected cells [4], [5]. One can easily envision the utility of such a system for the protection of viral replication and survival. Because NO is a known scavenger of hydroxyl radical, arginine deficiency has been associated with oxidative stress via increased production of reactive oxygen species [9]. Hence, enhanced expression of the HIV GPx under those conditions should be protective to HIV-infected cells, as a countermeasure to the increased oxidative stress induced by arginine deficiency. Such a role would be consistent with the previously demonstrated antioxidant and cytoprotective effects of the HIV GPx in living cells [5]. The significance of these observations for HIV pathogenesis in relation to dietary selenium status remains to be elucidated.

While definitive evidence that the HIV-1 GPx protein is expressed in vivo has yet to be obtained, the data of Gladyshev et al. in regard to ^{75}Se labeling of HIV-infected cells are highly consistent with the expression of the expected isoforms of the HIV-1 GPx, at the predicted molecular masses. In HIV-infected cells, Gladyshev et al. observed a decline in levels of cellular selenoproteins, and an increase in “low molecular mass” Se compounds [11]. The latter

corresponded to a wide band partially overlapping and higher in mass than a 6 kDa band seen in uninfected cells. These “low molecular mass” compounds were therefore in the correct mass range to correspond at least in part to the expected location of the predicted protease-processed isoform of the HIV-1 GPx, which is 9 kDa. Furthermore, a slower moving faint ⁷⁵Se-labeled band unique to the HIV-infected cells was in the correct mass range to correspond to the initial *env-fs* gene product produced by frameshifting, in which the GPx module would be fused to the HIV-1 *env* gp120 protein (hence, expected to be about 130 kDa). Thus, the existing experimental evidence is highly consistent with the expression of the HIV-1 GPx in vivo, at least in infected cells. In regard to human data, the results of Cohen et al. showing that the HIV-1 GPx reading frame is more conserved in long-term non-progressors, but has more “loss of function” mutations (i.e., premature terminations) in those who have progressed to AIDS, suggests that the gene is functionally significant in regard to clinical HIV disease [5]. This also suggests that the viral GPx gene will not be useful as a conventional “target” for antiviral therapy, if it is actually a “survival marker” in HIV infection.

The present demonstration that frameshifting at the predicted HIV-1 *env-fs* site occurs in vitro, and is up regulated under arginine deficiency, is to the best of our knowledge the first time that a hungry codon frameshift mechanism has been demonstrated involving an arginine codon. Combined with previous demonstrations that this putative gene of HIV-1 encodes functional GPx activity in transfected cells [4], which furthermore is cytoprotective via an antioxidant anti-apoptotic mechanism [5], these results strongly support the hypothesis that HIV-1 encodes a functionally significant GPx module.

Acknowledgments

This work was supported by PHS Grant DA13561 from the National Institute of Drug Abuse (NIDA) and a training grant from the Sloan Foundation.

References

- [1] E.W. Taylor, C.S. Ramanathan, R.K. Jalluri, R.G. Nadimpalli, A basis for new approaches to the chemotherapy of AIDS: novel genes in HIV-1 potentially encode selenoproteins expressed by ribosomal frameshifting and termination suppression, *J. Med. Chem.* 37 (1994) 2637–2654.
- [2] E.W. Taylor, A. Bhat, R.G. Nadimpalli, W. Zhang, J. Kececioglu, HIV-1 encodes a sequence overlapping *env* gp41 with highly significant similarity to selenium-dependent glutathione peroxidases, *J. Acquir. Immune Defic. Syndr. Hum. Retrovirol.* 15 (1997) 393–394.
- [3] E.W. Taylor, C.S. Ramanathan, R.G. Nadimpalli, A general method for predicting potential new genes in nucleic acid sequences: application to the human immunodeficiency virus, in: M. Witten (Ed.), *Computational Medicine, Public Health and Biotechnology*, World Scientific, Singapore, *Ser Math Biol Med* 5, 1996, pp. 285–309.
- [4] L.J. Zhao, A.G. Cox, J.A. Ruzicka, A.A. Bhat, W.Q. Zhang, E.W. Taylor, Molecular modeling and in vitro activity of an HIV-1-encoded glutathione peroxidase, *Proc. Natl. Acad. Sci. U.S.A.* 97 (2000) 6356–6361.

- [5] I. Cohen, P. Boya, L. Zhao, D. Metivier, K. Andreau, J.L. Perfettini, J.G. Weaver, A. Badley, E.W. Taylor, G. Kroemer, Anti-apoptotic activity of the glutathione peroxidase homologue encoded by HIV-1, *Apoptosis* 9 (2004) 181–192.
- [6] J.A. Gallant, D. Lindsley, Leftward ribosome frameshifting at a hungry codon, *J. Mol. Biol.* 223 (1992) 31–40.
- [7] J. Gallant, D. Lindsley, Ribosome, Frameshifting at hungry codons—sequence rules, directional specificity and possible relationship to mobile element behavior, *Biochem. Soc. Trans.* 21 (1993) 817–821.
- [8] K. Kolor, D. Lindsley, J.A. Gallant, On the role of the P-site in leftward ribosome frameshifting at a hungry codon, *J. Mol. Biol.* 230 (1993) 1–5.
- [9] K. Aoyagi, Inhibition of arginine synthesis by urea: a mechanism for arginine deficiency in renal failure which leads to increased hydroxyl radical generation, *Mol. Cell Biochem.* 244 (2003) 11–15.
- [10] T. Jacks, M.D. Power, F.R. Masiarz, P.A. Luciw, P.J. Barr, H.E. Varmus, Characterization of ribosomal frameshifting in HIV-1 gag-pol expression, *Nature* 331 (1988) 280–283.
- [11] V.N. Gladyshev, T.C. Stadtman, D.L. Hatfield, K.-T. Jeang, Levels of major selenoproteins in T cells decrease during HIV infection and low molecular mass selenium compounds increase, *Proc. Natl. Acad. Sci. U.S.A.* 96 (1999) 835–839.