FRAMESHIFTING IN HIV GENES AND CHEMOTAXIS OF NEF PROTEIN ISOFORMS

by

BABATUNDE OLUBAJO

(Under the Direction of Ethan Will Taylor)

ABSTRACT

Ribosomal frameshifting is a translational mechanism commonly exploited by RNA viruses in an effort to overcome the constraints of their limited genome size. Frameshifting requires a suitable heptanucleotide “slippery” sequence and a RNA pseudoknot, and the process allows viruses to translate genes hidden in any of the three frames that exist for all genes.

Previous analysis of the HIV-1 genome in our laboratory indicates the potential for viral selenoproteins to be encoded by frameshifting through re-coding of UGA stop codons as selenocysteine. One of the main goals of this study was to test the validity of predicted frameshift sequences in the envelope, protease and nef sequences of HIV-1. Furthermore, there was an interest in determining the role intracellular selenium and sulfur concentrations might play in frameshifting when tRNA-lys is involved, as well as the effects of arginine on a proposed hungry arginine codon in the envelope frameshift sequence. A frameshift assay was developed to test the validity of hypothesized frameshift sequences. Green fluorescent protein (GFP) and luciferase genes were cloned and arranged in a eukaryotic expression vector, in a manner that allowed putative frameshift sequences to be inserted into the vector between the two reporter

genes, with luciferase in a –1 frame of an upstream GFP gene. The vectors were expressed in 293-T cells and the ratio of reporter protein activity was assessed to determine the extent and efficiency of frameshifting. The envelope and protease genes had significant frameshift activity, which was found to decrease in the presence of selenium and sulfur donor compounds. The hungry arginine codon in the envelope gene also exhibited an increase in frameshifting in arginine deficient media.

Previous analysis of the nef frameshift resultant reading frame in our laboratory, indicated amino acid similarity to gamma chemokines. Furthermore, several conserved UGA codons were evident in this reading frame with potential for selenocysteine incorporation. To test this hypothesis, several mutant constructs were made that induced the proposed frameshift reading frames. When these vectors were expressed in 293-T cells, the lysates of recombinant nef and one of the frameshift induced nef mutants displayed chemoattractant ability for Jurkat cells.

INDEX WORDS: Frameshifting, envelope, gag-pol, protease, nef, chemotaxis, green fluorescent protein, luciferase, β-galactosidase, selenoprotein, HIV

FRAMESHIFTING IN HIV GENES AND CHEMOTAXIS OF NEF PROTEIN ISOFORMS

by

BABATUNDE OLUBAJO

B.S., University of Georgia, 1998

B.S.A., University of Georgia, 2000

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2004

FRAMESHIFTING IN HIV GENES AND CHEMOTAXIS OF NEF PROTEIN ISOFORMS

by

BABATUNDE OLUBAJO

Major Professor: Ethan Will Taylor

Committee: Warren Beach
Arthur Grider
Diane Hartle
Bi Cheng Wang

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2004

DEDICATION

I would like to dedicate this dissertation to my family: Rongbe, Kemi, Femi, Ronke and Oyinkan Olubajo for all their support. To Kim for all the support, friendship and love during the tough times in graduate school.

ACKNOWLEDGEMENTS

I would like to thank my parents Kemi and Rongbe for their love, support and pushing me to succeed through the years. My brother and sisters, Femi, Ronke and Oyinkan for their friendship, love and support. To Kim for helping me through graduate school with encouragement, love and support.

I would like to acknowledge my major advisor Dr. Taylor for all his support and for giving us a “think outside the box” approach to scientific problems to follow. I would like to acknowledge my committee members: Dr. Wang for his support and allowing me to work in his lab as an undergraduate where I gained my foundation in molecular biology. Dr. Hartle who allowed me to use equipment in her lab for which I am very grateful. To Dr. Beach and Dr. Grider for support and constructive criticisms. Cummings, Momany and Bruckner labs for allowing me to use equipment in their lab for which I am very grateful. My lab mates and departmental graduate students for their friendship. I would like to acknowledge the Alfred P. Sloan Foundation for financial support.

Last but not least I would like to thank God for giving me the strength to survive the tough times that life brings.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

BACKGROUND AND SIGNIFICANCE

Overview

This chapter contains a review of the current knowledge on the human immunodeficiency virus (HIV). Statistics on current infection rates are presented to emphasize the importance of HIV research. The function of HIV proteins is discussed highlighting current HIV antiviral therapeutic targets. Furthermore, regions of the HIV genome that are necessary for transcription and translation of HIV genes are discussed along with the life cycle of the virus. The chapter ends with an introduction to the studies presented in the subsequent chapters of this dissertation.

Introduction and Review

The human immunodeficiency virus was determined to be the causative agent of the acquired immune deficiency syndrome in the early 1980s (Barre-Sinoussi et al. 1983, Gallo et al. 1983, Levy et al. 1984). HIV and its associated disease, acquired immune deficiency syndrome (AIDS), still pose one of the most difficult therapeutic tasks for biomedical researchers. HIV/AIDS is a disease that has led to a current day epidemic that is of an alarming magnitude. According to the National Institute of Allergy and Infectious Diseases, at the end of 2003, approximately 40 million people were found to be living with HIV/AIDS worldwide. There were also 5 million new cases of HIV reported in the same year. There were 3 million deaths to HIV/AIDS worldwide in the year 2003.

The majority of HIV/AIDS cases are in sub-Saharan Africa, Asia and the Pacific. In some countries like Botswana and South Africa, 30% of the population is infected with HIV. Since the beginning of the epidemic, there have been approximately 25 million deaths to the disease.

The Centers for Disease Control and Prevention estimate that by the end of 2003 between 850,000 and 950,000 people were living with HIV infection in the United States. The CDC also estimates that 40,000 people are infected with HIV in the Unites states annually. The last report on HIV/AIDS deaths in the United States from the year 2002 shows that there were greater than 16,000 deaths in that year alone. By the end of 2002, there were over 500,000 deaths to HIV/AIDS in the United States.

These statistics alone highlight the importance of HIV/AIDS research. Though decades of substantial funding for HIV/AIDS research in an effort to combat the disease have led to notable advances, no actual cure for HIV infection has been discovered. Some HIV therapeutics have attempted to, and have been somewhat successful in, slowing progression of the disease. Preventative measures, like condoms and efficient HIV diagnostic tests, have been useful in controlling the spread of HIV, though methods of fully eradicating HIV infection are still non-existent. In light of these facts, understanding HIV's various mechanisms of survival should be a critical component of biomedical research, and it is imperative that scientists find new ways of attacking the virus. It is important that scientists continue to approach a full understanding of HIV in order to address a serious and growing epidemic worldwide.

Understanding HIV's mechanisms of survival can also be useful to increasing the knowledge base of general science. A key component of HIV's virulence is its ability to

mutate rapidly giving the virus more adaptable genes and proteins. Investigating this process can serve to aid scientists in understanding evolution. Studies into the mechanisms of HIV replication, infection, virulence and defense mechanisms can give researchers more insights into retroviruses.

Furthermore, the study of viruses like HIV can lead to useful scientific and therapeutic tools, as is evident in the utilization of viruses for gene delivery. Research aimed at delivering therapeutic copies of genes to specific cell types, has involved retroviruses and adenoviruses as carrier agents in the delivery of genes. Such methods are only possible with a good comprehension of the carrier virus in order to understand what segments are important for needed function and what portions must be removed in order to prevent detriment to the recipient host.

HIV is very successful at reproduction, which appears to be the only purpose for its survival. HIV infected patients can produce up to 10 billion new viruses in a 24-hour period. The virus attacks and replicates in immune cells, hence attacking the very defense system against viruses, bacteria and toxins. HIV's invasion of immune cells is one of the factors that governs the difficult dynamics of treating HIV infection.

Another attribute that plays a role in the complexity of medical intervention in HIV infection is HIV's error prone nucleotide polymerase known as reverse transcriptase (Bebenek et al. 1989). The diminished fidelity of the HIV polymerase results in the equivalent of an accelerated evolutionary process, which produces mutant genes and proteins at an accelerated rate. The constant production of variant proteins introduces more efficient ways for the virus to attack the host and elude therapeutic agents. These factors combine to make HIV a difficult disease to manage, and most AIDS patients, as

a result of a compromised immune system, succumb to opportunistic infections or cancer.

HIV mainly attacks and depletes CD4+ T lymphocytes, but has been found in macrophages, Langerhans and other immune cells (Cameron et al. 1992, Zambruno et al. 1995). CD4+ T-lymphocyte depletion is a hallmark of HIV infection and has been a means of monitoring the progression of AIDS (Levy et al. 1993). T lymphocytes play a major role in specific immunity against intracellular microbes. Loss of T lymphocytes results in an increased susceptibility to mycobacteria and pneumocystis carinii infection. There is also a higher incidence of Kaposi's sarcoma and B-cell lymphomas by the Epstein Barr virus.

HIV as a Lentivirus

HIV is a member of the lentivirus group of the *retroviridae* family. The word lentivirus contains the Latin prefix *lenti,* which translates to slow virus, and describes the pro-longed incubation period of most viral members of the group. Other viruses in this group include the encephalitis virus, simian and feline immunodeficiency viruses.

Previous Lentiviral studies have given researchers a slight advantage in the comprehension of HIV's biological mechanisms. Lentiviruses tend to have a positive, single stranded, linear RNA genome that is translated to DNA by reverse transcriptase and incorporated into the host genome for processing with the host organism's transcriptional and translational machinery. Viruses in this group normally have an icosahedral capsid with 7 particle polypeptides and are approximately 100nm in size. These viruses have been shown to infect humans and other animals. The lentivirus infections are normally transmitted through fluid exchange and are prone to produce

immunodeficiency disorders. Lentiviruses were thought to be primarily exogenous, but studies show homology of some HIV and human genes, indicating an evolutionary relationship (Veljkovic et al. 1990).

The HIV Genome

The HIV genome is approximately 9.8 kilobases in length. Nine main proteins have been identified as direct or, early by-products of translation. These proteins are gag, pol, env, vif, vpr, rev, vpu, env, tat and nef. There are additional HIV encoded proteins that undergo post-translational modifications such as reverse transcriptase, protease and integrase. A stretch of repeated adjacent sequences flank the 5' and 3' ends of the HIV genome and are known as the Long Terminal Repeats (LTR). The LTR also contains promoters controlled by host and viral proteins that are important for regulation of HIV gene expression. These components of the HIV genome will be discussed in further detail later in the chapter.

HIV genes and proteins

LTR

The 640bp HIV LTR is involved in transcription regulation of HIV genes (Starcich et al. 1985). The LTR of HIV flanks the 5' and 3' ends of the HIV genome. The LTR is important in reverse transcription of DNA from the RNA genome, and integration into the host genome. The 5' LTR contains the promoter region that regulates mRNA transcript production from the HIV genome. The LTRs are divided into the U3, R and U5 regions (Starcich et al. 1985). The U3 of the LTR includes a TATA box, the TAR element and promoters for transcription factors such as NFκB and SP1.

NFκB is the most studied LTR activating transcription factor. NFκB has been found to play a key role in HIV genome transcription (Nabel et al. 1987). The NFκB promoter region of the LTR is the most conserved sequence of the HIV-1 genome. Under normal physiologic conditions, NFκB is bound to Iκ B in a dormant state and this complex is localized in the cytoplasm. Upon exposure of a cell to environmental or physiological stimuli, activation of Iκ B through phosphorylation occurs, causing a dissociation of Iκ B from NFκB. NFκB is then localized to the nucleus where it binds to the NFκB promoter region to induce transcription of an associated gene. In the case of HIV, NFκB binds to a homologous NFκB promoter in the LTR leading to transcription of the HIV genome (Antoni et al. 1994, Gaynor et al. 1992).

There are several other cellular transcription factors that bind the LTR and affect HIV production. The LTR contains 2 promoters for the SP1 gene, which encodes a constitutively expressed eukaryotic protein that is involved in regulation of transcription (Pugh et al. 1990). Studies indicate that SP1 is required for TAT activation of transcription (Roebuck et al. 1999). AP-1 was found to bind the U5 region of the LTR and affect HIV LTR activity (Roebuck et al. 1993, Rabbi et al. 1995) and a promoter for NF-AT is a component of the LTR-U5 region and has been found to control HIV gene expression (Cron et al. 2000).

The TAR element

The TAR element plays an important role in HIV transcription. It consists of a region of RNA approximately 60 base pairs in length and is found adjacent to the LTR in the 5' end of HIV and next to the nef gene in the 3' region (Churcher et al. 1993). The TAR element is a stem-loop structure with a 3-nucleotide bulge. The TAR element is

believed to associate with the tat protein, one of the earliest HIV proteins to be expressed. An arginine rich region of the tat protein is believed to complex with a uridine nucleotide at position 23 of the TAR element (Churcher et al. 1993). The TAR-tat complex is essential for effective production of HIV transcripts and gene expression (Churcher et al. 1993 Roy et al. 1990).

The importance of the TAR-tat complex to HIV replication has resulted in therapeutic research focused on disruption of the complex. Compounds designed to inhibit the interaction between tat and TAR have been effective *in vitro* (Hamy et al. 1998). Anti-sense RNA technology is another strategy that has proven to be successful at inhibiting the tat-TAR interaction (Hamma et al. 2003).

Tat

The tat protein is one of the earliest proteins made from the HIV genome and is involved in upregulation of several cellular genes. Tat stands for transcription activator, and is a protein that is required for efficient transcription of the host-chromosome-integrated HIV genome (Kao et al. 1987, Feinberg et al.1991).

Transcription of the host incorporated HIV genome proceeds with RNA polymerase binding to a promoter in the U3 region of the HIV LTR. It has been found that transcription of the HIV genome by RNA polymerase alone yields short transcripts that result from premature transcription termination. Tat facilitates efficient transcriptional elongation by binding the RNA element TAR in the 5' region of the HIV genome (Kao et al. 1987, Feinberg et al. 1991). Recent studies also suggest that tat binds HIV-nef to enhance HIV gene expression (Joseph et al. 2003).

Gag

The gag gene yields a precursor protein of approximately 55 KD that is proteolitically cleaved to give proteins that comprise the core of the virion particle. HIV protease aids in cleavage and maturation of the gag precursor polypeptide, producing several by-products, which include the p24 capsid protein, nucleocapsid protein p7 and the matrix protein p17 (Gottlinger et al. 1989). The p24 protein is a structural protein of HIV that forms the HIV capsid, while P17 is important to virion particle stability, and nucleus-directed translocation of viral DNA for processing by cellular proteins (Lewis et al. 1992, Gallay et al. 1995). P7 forms the nucelocapsid structural element of the virus and is involved in packaging of the viral RNA genome (Harrison et al. 1992).

Pol

The pol gene is responsible for encoding a precursor protein that is processed by HIV protease in the virion particle. Proteolytic cleavage of the pol precursor protein yields the integrase, protease, RNAse H and reverse transcriptase proteins of HIV. The pol protein is encoded as a fusion to the gag precursor protein as a result of –1 frameshifting, which accounts for 5% of translation product (Jacks et al. 1988). Frameshifting regulates the 20:1 ratio between the gag and gag-pol precursor protein (Jacks et al. 1988), a ratio that is important to HIV infectivity (Shehu-Zhilaga et al. 2000).

Integrase

Integrase is the HIV protein responsible for incorporation of the HIV genome into the host chromosome for subsequent processing by cellular transcription and translation machinery. Integrase facilitates incorporation of the HIV genome into the host

chromosome by cleaving nucleotides off a double stranded HIV DNA genome (Bushman et al. 1990). A process called strand transfer is required for incorporation of the viral genome into the host chromosome at the beginning of transcription. This process involves cleavage of the host DNA to give two 5' ends (Bushman et al. 1990). The 5' ends of the host DNA and the 3' ends of the viral genome are then ligated (Bushman et al. 1990).

Envelope

The envelope gene of HIV produces gp161, a 161 KD glycoprotein that serves as the first translation product of the envelope gene. Gp161 is proteolytically cleaved to form gp41 and gp120 and both proteins combine to form a dimer that is found on the surface of the virion particle. Gp41 anchors the dimer on the membrane of the HIV virion particle, while gp120 resides on the surface of HIV. The gp120 protein is responsible for the attachment of HIV to chemokine receptors on T-cells, the action of which precedes cell-entry through fusion of the virus with the host cell membrane (Berger et al. 1999). Gp120 is found to bind both CD4 molecules and CCR5 receptors on the surface of T-lymphocytes (Landau et al. 1988, Berger et al. 1999).

Protease

The protease enzyme of HIV-1 plays a critical role in HIV production. The protease enzyme of HIV is a dimerized aspartyl protease (Navia et al. 1989) .The aspartyl-proteases generally have a sequence specific catalytic site represented by the sequence Asp-Thr-Gly (Pearl et al. 1987). Ribosomal frameshifting in the HIV genome produces a transcript that combines a portion of the 3' end of the gag gene with the 5' end of the pol gene and protease is a by-product of translation of this new transcript

(Jacks et al. 1988). Frameshifting in the gag-pol gene results in a gag-pol precursor protein, which undergoes a catalytic process involving protease that leads to a separation of gag, pol and the protease enzyme from the precursor protein (Farmerie et al. 1987). HIV protease is also involved in processing the integrase, RNAse H, and reverse transcriptase proteins from the gag-pol precursor polyprotein. The importance of protease to HIV infection has presented the protein as a successful target for treatment therapy of HIV infection, by arresting production of necessary HIV proteins.

Reverse transcriptase

Reverse transcriptase is encoded as a component of HIV-1 pol gene translation. The pol gene is encoded as a fusion to the gag protein as a result of a –1 frameshift and subsequently processed by HIV protease. Reverse transcriptase is one of 4 enzymes produced from cleavage of the pol precursor protein. Reverse transcriptase, which is found in the virion particle, is necessary for the reverse transcription of DNA from the single stranded HIV genome. Reverse transcription produces a double stranded DNA copy of the HIV genome and occurs in the cytosol of the host cell. Nucleoside drugs, which target reverse transcriptase, were the first successful AIDS drugs, and were effective inhibitors of HIV genome transcription.

Nef

Nef is the largest of HIV's accessory proteins and is expressed in significant quantities early in HIV infection (Kim et al. 1989 Virology). Nef, which stands for negative factor, was named due to early reports indicating that nef caused a decrease in HIV transcription. Current opinions suggest that nef is important for several different aspects of viral replication.

Cells infected with the HIV virus have been shown to contain nef molecules localized in the nucleus, cytoplasm and in extracellular space (Fujii et al. 1996, Murti et al. 1993). Nef has also been reported as a monomer, and in multimeric forms (Fujii et al. 1996). Nef is found in virion particles and is cleaved by HIV protease in the virion particle (Pandori et al. 1996)

Nef associates with several cellular kinases and can be serine or threonine phosphorylated (Arnold et al. 1997, Sawai et al. 1994). Nef has been shown to induce T-cell activation and also to block IL-2 expression (Sawai et al. 1994, Schrager et al. 1999). Studies have demonstrated that HIV-1 Nef acts within infected cells to increase the infectivity of progeny viruses (Spina et al. 1994, Miller et al. 1994).

Nef has other functions that aid HIV's control of the host cell. Nef down regulates the expression of a chemokine transcription factor NFκB (Luria et al. 1991). Nef down-modulates MHC class 1 expression on the surface of T-cells (Schwartz et al. 1996). Nef also leads to a decrease in the number of CD4 molecules on immune cells (Garcia et al. 1992).

The molecular weight of nef has been cited in scientific literature to range between 25kd and 35kd. It is difficult to determine the exact basis for the disparity of nef's molecular weight in different studies, but this underscores the complexity of characterizing nef and understanding its multitude of functions.

The Life Cycle of HIV

HIV is believed to attach and gain entry to immune cells by means of the CD4 and chemokine (CCR5) receptors. The GP120/GP41 protein on the surface of HIV associates with the CD4 protein on the membrane of host cells (Dagleish et al. 1984).

Upon GP120/GP41 subsequent association with CCR5 receptor, a conformational change occurs that exposes GP41 and begins fusion of the virus with the cell membrane of the host cell (Chan et al. 1998). The contents of the virion particle are released into the cytoplasm of the cell following fusion with the cell membrane. Subsequently, the capsid coating of the virus core disintegrates to release the contents of the core. The HIV RNA genome, reverse transcriptase and tRNA-lys are released into the host cell and reverse transcription proceeds (Karageorgos et al. 1993). Reverse transcriptase is a polymerase involved in transcription of DNA from RNA and tRNA-lys serves as the primer for the reverse transcription process. The preintegration complex (PIC), which includes the newly synthesized double stranded DNA complex, integrase and other associated proteins, is transported into the nucleus of the cell (Miller et al. 1997). While in the nucleus, integration of the DNA genome into the host chromosome occurs by action of proteins in the preintegration complex (Miller et al. 1997). New viral particles are made in the nucleus by host cellular transcriptional and translational machinery, which produce genome transcripts and proteins necessary for assembly of the virus. Assembled virion particles eventually bud off the host cell, to infect new cells and continue the viral replication cycle.

Objectives of the dissertation

The objectives of this study was to examine potential hidden mechanisms of HIV's survival, through alternate translation mechanisms and host protein mimicry. The main studies documented in this dissertation were to understand 3 predicted frameshift sequences in the envelope, nef and protease regions of HIV. Frameshift sequences consist of a pseudoknot and slippery heptanucleotide sequence, and both

elements cause slippage of the ribosome to a different reading frame during translation. These frameshift sequences were predicted by an extensive computational analysis of the HIV genome (Taylor et al. 1994,1996). Furthermore, the effects of cysteine, sodium selenite and arginine supplementation on frameshifting is presented to understand the role sulfur and selenium may play on frameshifting, and the effects of arginine on a predicted hungry arginine codon in the envelope frameshift sequence. A study of the chemoattractant ability of resultant frameshift proteins from the nef gene of HIV for Jurkat cells is described, and engineering of constructs that will be useful in the study of selenocysteine incorporation into nef gene was also carried out.

References

Antoni, B. A., S. B. Stein, et al. (1994). "Regulation of Human-Immunodeficiency-Virus Infection - Implications for Pathogenesis." Advances in Virus Research, Vol 43 43: 53-145.

Arnold, S., P. Franken, et al. (1997). "The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling." Structure 5(10): 1361-72.

Barre-Sinoussi, et al. (1983). “Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS).” Science 220, 868-871.

Bebenek, K., J. Abbotts, et al. (1989). "Specificity and mechanism of error-prone replication by human immunodeficiency virus-1 reverse transcriptase." J Biol Chem 264(28): 16948-56.

Berger, E. A., P. M. Murphy, et al. (1999). "Chemokine receptors as HIV-1 coreceptors: roles in viral entry, tropism, and disease." Annu Rev Immunol 17: 657-700.

Bushman, F. D., T. Fujiwara, et al. (1990). "Retroviral DNA integration directed by HIV integration protein in vitro." Science 249(4976): 1555-8.

Cameron, P. U., P. S. Freudenthal, et al. (1992). "Dendritic cells exposed to human immunodeficiency virus type-1 transmit a vigorous cytopathic infection to CD4+ T cells." Science 257(5068): 383-7.

Chan, D. C. and P. S. Kim (1998). "HIV entry and its inhibition." Cell 93(5): 681-4.

Churcher, M. J., C. Lamont, et al. (1993). "High affinity binding of TAR RNA by the human immunodeficiency virus type-1 tat protein requires base-pairs in the RNA stem and amino acid residues flanking the basic region." J Mol Biol 230(1): 90-110.

Cron, R. Q., S. R. Bartz, et al. (2000). "NFAT1 enhances HIV-1 gene expression in primary human CD4 T cells." Clin Immunol 94(3): 179-91.

Dagleish, A, Beverly, P., et al. (1984) “The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature 312:763-764

Farmerie, W. G., D. D. Loeb, et al. (1987). "Expression and processing of the AIDS virus reverse transcriptase in Escherichia coli." Science 236(4799): 305-8.
Feinberg, M. B., D. Baltimore, et al. (1991). "The role of Tat in the human immunodeficiency virus life cycle indicates a primary effect on transcriptional elongation." Proc Natl Acad Sci U S A 88(9): 4045-9.

Fujii, Y., K. Otake, et al. (1996). "Clustered localization of oligomeric Nef protein of human immunodeficiency virus type 1 on the cell surface." FEBS Lett 395(2-3): 257-61.

Garcia, J. V. and A. D. Miller (1992). "Downregulation of cell surface CD4 by nef." Res Virol 143(1): 52-5.

Gallay, P., S. Swingler, et al. (1995). "HIV nuclear import is governed by the phosphotyrosine-mediated binding of matrix to the core domain of integrase." Cell 83(4): 569-76.

Gallo RC, Sarin PS, et al. (1983) “Isolation of Human T- Cell Leukemia Virus in Acquired Immune Deficiency Syndrome (AIDS).” Science 220:865-867.

Gaynor, R. (1992). "Cellular transcription factors involved in the regulation of HIV-1 gene expression." Aids 6(4): 347-63.

Gottlinger, H. G., J. G. Sodroski, et al. (1989). "Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1." Proc Natl Acad Sci U S A 86(15): 5781-5.

Hamma, T., A. Saleh, et al. (2003). "Inhibition of HIV tat-TAR interactions by an antisense oligo-2'-O-methylribonucleoside methylphosphonate." Bioorg Med Chem Lett 13(11): 1845-8.

Hamy, F., V. Brondani, et al. (1998). "A new class of HIV-1 Tat antagonist acting through Tat-TAR inhibition." Biochemistry 37(15): 5086-95.

Harrison GP, Lever AM. The human immunodeficiency virus type 1 packaging signal and major splice donor region have a conserved stable secondary structure. J Virol 1992;66:4144-4153.

Jacks, T., M. D. Power, et al. (1988). "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression." Nature 331(6153): 280-3.

Joseph, A. M., J. S. Ladha, et al. (2003). "Human immunodeficiency virus-1 Nef protein interacts with Tat and enhances HIV-1 gene expression." FEBS Lett 548(1-3): 37-42.

Kao, S. Y., A. F. Calman, et al. (1987). "Anti-termination of transcription within the long terminal repeat of HIV-1 by tat gene product." Nature 330(6147): 489-93.

Karageorgos, L., P. Li, et al. (1993). "Characterization of HIV replication complexes early after cell-to-cell infection." AIDS Res Hum Retroviruses 9(9): 817-23.

Kim, S. Y., Byrn. R et al (1989). “Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: Evidence for differential gene expression.” Journal of Virology 63: 3708-3713

Landau, N. R., M. Warton, et al. (1988). "The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4." Nature 334(6178): 159-62.

Levy, J.A. et al. (1984). “Isolation of lymphocytopathic retroviruses from San Francisco patients with AIDS.” Science 225: 840-842.

Levy, J. A. (1993). "Pathogenesis of human immunodeficiency virus infection." Microbiol Rev 57(1): 183-289.

Lewis, P., M. Hensel, et al. (1992). "Human immunodeficiency virus infection of cells arrested in the cell cycle." Embo J 11(8): 3053-8.

Luria, S., I. Chambers, et al. (1991). "Expression of the type 1 human immunodeficiency virus Nef protein in T cells prevents antigen receptor-mediated induction of interleukin 2 mRNA." Proc Natl Acad Sci U S A 88(12): 5326-30.

Miller, M. D., M. T. Warmerdam, et al. (1994). "The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages." J Exp Med 179(1): 101-13.

Miller, M. D., C. M. Farnet, et al. (1997). "Human immunodeficiency virus type 1 preintegration complexes: studies of organization and composition." J Virol 71(7): 5382-90.

Murti, K. G., P. S. Brown, et al. (1993). "Highly localized tracks of human immunodeficiency virus type 1 Nef in the nucleus of cells of a human CD4+ T-cell line." Proc Natl Acad Sci U S A 90(24): 11895-9.

Nabel, G. and D. Baltimore (1987). "An inducible transcription factor activates expression of human immunodeficiency virus in T cells." Nature 326(6114): 711-3.

Navia, M. A., P. M. Fitzgerald, et al. (1989). "Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1." Nature 337(6208): 615-20.

Pandori, M. W., N. J. Fitch, et al. (1996). "Producer-cell modification of human immunodeficiency virus type 1: Nef is a virion protein." J Virol 70(7): 4283-90.

Pearl, L. H. and W. R. Taylor (1987). "Sequence specificity of retroviral proteases." Nature 328(6130): 482.

Pugh, B. F. and R. Tjian (1990). "Mechanism of Transcriptional Activation by Sp1 - Evidence for Coactivators." Cell 61(7): 1187-1197.

Rabbi, M, Saifuddin. M, et al (1997). “U5 region of the human immunodeficiency virus type 1 long terminal repeat contains TRE-like cAMP-responsive elements that bind both AP-1 and CREB/ATF proteins.” Journal of Virology 233:235-245.

Roebuck, K. A., D. A. Brenner, et al. (1993). "Identification of c-fos-responsive elements downstream of TAR in the long terminal repeat of human immunodeficiency virus type-1." J Clin Invest 92(3): 1336-48.

Roebuck, K. A. and M. Saifuddin (1999). "Regulation of HIV-1 transcription." Gene Expression 8(2): 67-84.

Roy, S., U. Delling, et al. (1990). "A bulge structure in HIV-1 TAR RNA is required for Tat binding and Tat-mediated trans-activation." Genes Dev 4(8): 1365-73.

Sawai, E.T. , Baur, A. et al (1994) “Human immunodeficiency virus type 1 Nef associates with a cellular serine kinase in T lymphocytes.” Proc. Natl Acad. Sci. 91: 1539-1543

Schrager, J. and Marsh, J. (1999) “HIV-1 Nef increases T cell activation in a stimulus-dependent manner “ Proc. Natl. Acad. Sci. 96: 8167-8172.

Schwartz, O., V. Marechal, et al. (1996). "Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein." Nat Med 2(3): 338-42.

Shehu-Xhilaga, M, Crowe, S. et al (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.

Spina, C., Kwoh, T. (1994) “The importance of nef in the induction of human immunodeficiency virus type 1 replication from primary quiescent CD4 lymphocytes” Journal of Exp. Med., 179 (1), 115-123

Starcich, B., L. Ratner, et al. (1985). "Characterization of Long Terminal Repeat Sequences of Htlv-Iii." Science 227(4686): 538-540.

Taylor, E. W., Ramanathan, C.S., et al. (1994). "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(17): 2637-54.

Taylor, E.W., Ramanathan, C.S., and Nadimpalli, R.G. (1996) “A general method for predicting potential new genes in nucleic acid sequences: application to the human

immunodeficiency virus.“ In: Computational Medicine, Public Health and Biotechnology, M. Witten, Ed. World Scientific, Singapore. Ser. Math. Biol. Med. 5: 285-309.

Veljkovic, V. and R. Metlas (1990). "Sequence similarity between human immunodeficiency virus type 1 envelope protein (gp120) and human proteins: a new hypothesis on protective antibody production." Immunol Lett 26(2): 193-5.

Zambruno, G., A. Giannetti, et al. (1995). "Langerhans cells and HIV infection." Immunol Today 16(11): 520-4.

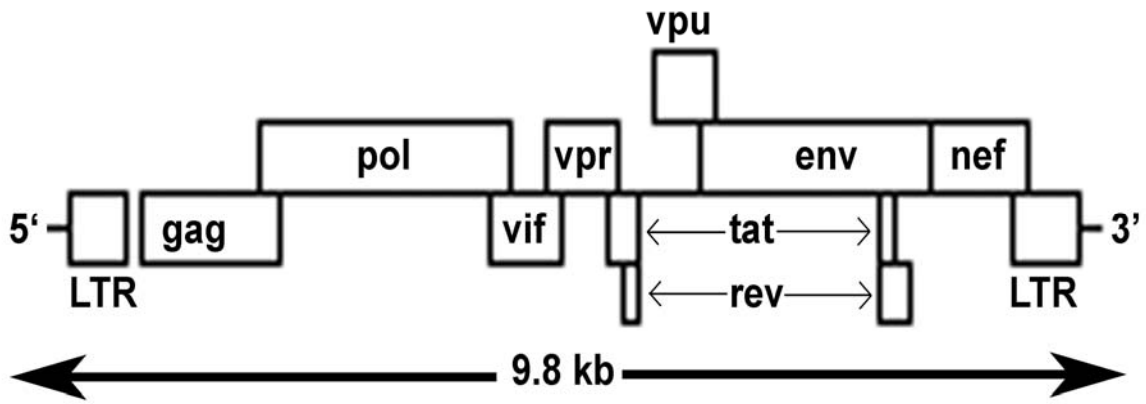


Figure 1-1

HIV-1 genome. Schematic of HIV-1 genome illustrating compact organization of genes for efficient utilization of genome.

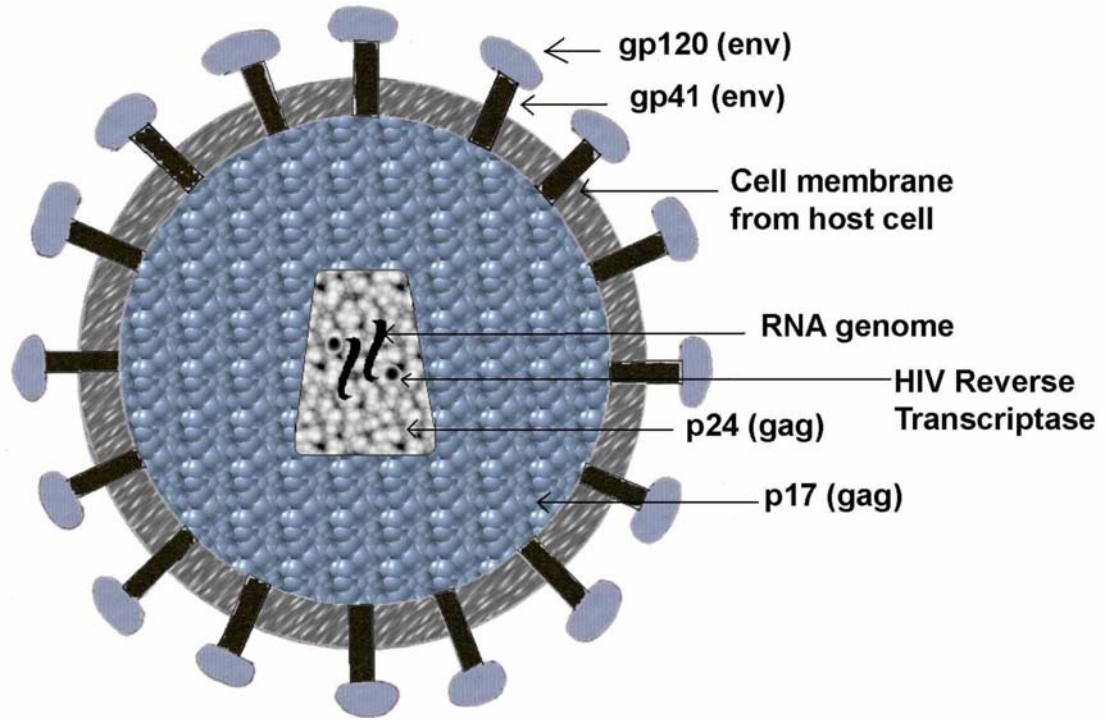


Figure 1-2

Cross-sectional schematic of the HIV virion particle.

CHAPTER 2

RIBOSOMAL FRAMESHIFTING IN THE PROTEASE GENE OF HIV-1[1]

[1]Olubajo, B and Taylor, E.W. To be submitted to Journal of RNA

Overview

This chapter contains a review of ribosomal frameshifting and the function and properties of HIV-protease. Furthermore, the development of a GFP-luciferase, dual-reporter frameshift assay is discussed and frameshifting of the proposed protease frameshift sequence is assessed. The effects of cysteine and selenium on frameshifting in the protease region of HIV are also assessed. The main objectives of the experiments conducted in this study was to: 1. Determine the efficiency of the predicted frameshift sequence in the protease region of HIV. 2. Determine the effects of sodium selenite and cysteine supplementation on the predicted protease frameshift sequence through enhancement of sulfur or selenium hypermodified base synthesis and incorporation into tRNA-lys.

Introduction and Review

Viruses commonly employ alternative methods in encoding proteins from DNA sequences. All genes have three possible reading frames owing to translation proceeding by 3-nucleotide codons encoding for each amino acid. This occurrence, coupled with deletion and insertion of portions of a gene, is exploited by RNA and protein translational machinery in arriving at different transcripts and hence different proteins. One such alternative method of encoding genes is ribosomal frameshifting.

Ribosomal frameshifting is a means by which viruses maximize utility of their genomes or gene sequences. Ribosomal frameshifting involves slippage of the ribosome during protein translation. Slippage of the ribosome can either be in a forward or backward direction, and in most cases, by one or two bases. The transcripts that result from such an event are termed negative (-) for backward slippage or (+) for

forward slippage. A transcript encoded as a result of the ribosome slipping back one base can be described as a – 1 frameshift product. When the ribosome slips forward by two bases, the resultant transcript is referred to as a + 2 frameshift product. From the point of slippage, translation continues on a totally different reading frame hereby allowing for a new protein to be encoded. Frameshifting events that occur midway through a gene lead to insertion or deletion of DNA sequences that can add or delete functions to a protein that is translated as a result.

Frameshifting requires a suitable heptanucleotide “slippery” sequence and an RNA pseudoknot. Slippery sequences and pseudoknots have been shown to be involved in frameshifting in several cases either as a pair, or individually. Slippery sequences are heptanucleotide RNA sequences, which in most cases have a less efficient interaction with the anticodon loop of the corresponding tRNAs in the ribosome during protein translation. Slippery sequences generally follow the pattern X-XXY-YYZ where X, Y or Z represent a different purine or pyrimidine base, though deviations to this rule have been reported.

Pseudoknots are tertiary RNA structures that are formed by base pairing between a secondary loop structure and complement bases outside the loop. The result is an RNA structure composed of two loops and two stems. The two stems coaxially stack to form a quasi-continuous RNA helix. The two loops span the major and minor grooves of the helix.

Though frameshifting is not completely understood, current opinions suggest that a frameshift begins when a ribosome encounters a pseudoknot in the mRNA to be encoded. The ribosome unraveling the pseudoknot leads to a pause in translation,

which in turn causes slippage of the ribosome at a suitable slippery sequence. The ribosome can slip either forward (+change) or backward (-change) and a different sequence of codons is interpreted.

The protease enzyme of HIV-1 plays a critical role in HIV production. The protease enzyme of HIV assumes a dimerized aspartyl form (Navia et al. 1989). The aspartyl-proteases generally have a sequence specific catalytic site represented by the sequence Asp-Thr-Gly (Pearl et al. 1987). Protease of HIV is encoded through ribosomal frameshifting, the process of which produces a gene that combines a portion of the 3' end of the gag gene with the 5' end of the pol gene (Jacks et al. 1988). Frameshifting in the gag-pol gene results in a gag-pol precursor protein, which undergoes an autocatalytic process involving HIV protease (Farmerie et al. 1987). This catalysis leads to a separation of gag from pol and release of the protease enzyme from the precursor protein (Farmerie et al. 1987). HIV protease is also involved in cleaving the integrase, RNAse H, and reverse transcriptase proteins (Farmerie et al. 1987).

Several studies have been aimed at inhibiting HIV protease because of its importance to critical steps in HIV production. Crystal structures of the protease gene have led to a better understanding of protease functions in HIV. Structure determination studies of HIV protease complexed with potential inhibitor compounds has led to current protease inhibitors that are today used in the treatment of HIV infected patients (Molla et al. 2003, Turner et al. 1999).

Computational analysis of the HIV-1 genome indicates the potential for frameshifting events (Taylor et al. 1994,1996). Frameshift sequences were detected in the *gag, pro, rt, int, pro, nef* genes of HIV. These frameshifts would occur in addition to

the known gag-pol –1 frameshift responsible for bypassing the gag stop codon and encoding the pol gene (Jacks et al. 1988). The viability of the frameshift sequences is supported by potential RNA pseudoknot structures and ideal heptameric slippery sequences that were found in the HIV genome, both of which are required for frameshifting.

The frameshift sequence found in the protease gene of HIV-I contains a heptanucleotide sequence that encodes lys-glu, followed by a pseudoknot comprising of two 5 base pair stems and 2 loops 4 and 9 nucleotides in length as shown in Figure 2-1.

The frameshift sequence was tested with a cell based frameshift assay developed in our lab. The frameshift assay was developed to test the validity of the hypothesized frameshift sequence. The frameshift sequence was studied with a dual protein reporter system. Green fluorescent protein and Luciferase genes were cloned and arranged in a eukaryotic expression vector respectively. Test frameshift sequences were cloned into the vector between the two reporter genes with luciferase in a –1 frame. The new frameshift assaying vector was expressed in Jurkat, MDCK, T293 cells all of which had similar results, but T-293 cells were the most sensitive. Fluorescence of GFP and luminescence of luciferase were detected from cell lysates. Their ratios were compared to determine the extent and efficiency of frameshifting in the protease frameshift sequence.

Comparing the fluorescence of gfp and luminescence of luciferase quantitates frameshifting. Luciferase translation measures frameshifting while gfp translation standardizes the assay. Therefore frameshifting is represented by luciferase luminescence as a factor of gfp fluorescence. Frameshifting efficiency is calculated by

comparing the reporter ratios of a test vector with the reporter ratios of a 100% fusion protein expression vector (PGL).

Materials and Methods

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. 2 versions of the luciferase gene were cloned and they differed in that 2 extra bases were added to the 5' end of one of the clones to embed the luciferase gene in a –1 frameshift conformation, while the 2nd 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 where PCL (-) denotes the –1 luciferase vector and PCL (+) is the 0 frame luciferase vector.

The GFP gene was digested from PEGFP-CI (Clontech) with Nhe I and Sal I restriction endonucleases. The digested GFP gene was ligated with T4 DNA ligase into, previously Nhe I and Sal I digested PCL (+) and PCL (-) vectors to form PGL (+) and PGL (-) vectors.

The protease frameshift sequence was cloned from PBH10 (HIV genome from NIH AIDS Reagent Program) by polymerase chain reaction, and in the process, Xho I and Sal I restriction sites were added to the 5' and 3' ends of the cloned protease frameshift sequence. A stop codon (TAA) was engineered to directly follow the frameshift sequence in the zero frame of GFP to truncate unframeshifted fusion protein.

The PGL (+), PGL (-) vectors and the protease frameshift sequence were digested with Xho I and Sal I and subsequently ligated to give the assaying vectors PGproL (+) and PGproL (-). An identical process was used to make PGgpL (+) and PGgpL (-) vectors which contain the known gag-pol frameshift sequence of HIV (Jacks et al. 1988). The sequence of all vectors was confirmed by dideoxy sequencing. The result of these experiments is shown in Figures 2-2 – 2-9.

Transfection of cell lines

293-T cells were grown in Dubelco minimal essential medium with 2mM L-glutamine, 1.5g/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 35mm diameter, 6 well plates to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 2μg of appropriate vector for 6 hours. Cells were transiently transfected with PGproL (+) and PGproL (-). PGL, PGgpL (+) and PGgpL (-) were used to transiently transfect cell lines as controls. The cells were incubated at 37º C with 5% CO_2 for 48 hours.

Cell-based Frameshift Assay

Transfected cell samples were collected and lysed by incubating with 1ml of glo-lysis buffer (Promega) per well for 5 minutes. 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 TD 20/20 luminometer, the luciferase assay involves adding 50μl of the cell lysates to 100μg 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-second delay. Background luminescence is normally 0.000 to 0.020 relative light units.

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

Fluorescence of GFP is measured using a Shimadzu RF-5301 or FLUOstar Optima Galaxy fluorescent plate reader. Cell lysate samples were excited at 488nm wavelength and the relative intensity of emission at 511nm 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 to GFP Fluorescence and Frameshifting Efficiency is the frameshift factor of a sequence as a percentage of the PGL frameshift factor.

Results

The protease frameshift sequence exhibited frameshifting capabilities when studied with the frameshift assay. There was significant luminescence emitted from cell lysates expressing the protease frameshift sequence test vector PGproL (-), over mock-transfected cells with no vector DNA. The mock transfected cells continuously showed a

good baseline giving relative luminescence values of between 0.000 to 0.007 relative light units in comparison with 10,000 fold increase in luminescence of assay vector cell lysates. Luminescence of samples tested with the TD 20/20 luminometer was recorded at a decreased sensitivity in order to obtain readings in the luminometers detectable range. Lysates of cells transfected with the protease frameshift assay vector PGproL (-) also presented a 10,000-fold increase in relative luminescence in comparison with cells transfected with the control vector PGproL (+). Cell lysates of cells transfected with the PGproL (+) construct, which contains the test protease 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 vectors PGgpL (-), containing the known gag-pol frameshift sequence also showed a 10,000 fold increase in luciferase activity over mock transfected cells, and 10,000 fold increase over cells expressing the PGproL (+) vector. The PGgpL (+) vector showed almost no luciferase activity. Luminescence quantitated from PGproL (+) expressing vectors was almost identical to mock-transfected cells and was practically inexistent. The protease frameshift sequence displayed luciferase activity that was about half the activity of the known gag-pol frameshift sequence. The positive control vector PGL which served as a 100% fusion protein expression construct, was consistently a million fold higher than mock transfected and less than 100 fold higher than the protease and gag-pol frameshift test vector when these vectors are expressed in 293-T cells. These results are shown in Figures 2-10.

GFP fluorescence was measured by excitation of cell lysates at 488 nm with a 5nm slit width and fluorescence emission was quantitated over a spectrum of

wavelengths. The peak of relative GFP fluorescence for assayed cell lysates was at 511nm. The relative fluorescence of GFP in cell lysates of transfected cells was fairly strong with a 10-fold increase in fluorescence over autofluorescence 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. PGproL (+) and PGproL (-) had similar fluorescence profiles, which were more than the gfp fluorescence of PGgpL (+) and PGgpL (-). GFP fluorescence quantitated from PGL transfected cells was significantly less than fluorescence of PGgpL (+) and PGgpL (-), PGproL (+) and PGproL (-). The decrease in GFP fluorescence of PGL in comparison with the other assay vectors may have been due to translation efforts of the cell involved in expression of mostly the GFP-Luciferase fusion protein in comparison with the PGproL (+), PGproL (-), PGgpL (+), and PGgpL (-) which would experience translation termination at the end of the GFP gene. These results are illustrated in Figure 2-11 – 2-16.

The frameshift factor was calculated by determining the ratio of luciferase activity to GFP fluorescence. Frameshifting efficiency was calculated by comparing test vector ratio to PGL ratio respectively. The Fluostar Optima plate reader was used to obtain the best quantitation of GFP fluorescence, and the ratios of fluorescence between test constructs were identical to relative fluorescence comparisons between test constructs with the single cell fluorometer. These results are illustrated in Figure 2-17 and 2-18.

The protease frameshift vector PGproL (-) was 10,000 fold higher than the control vector PGproL (+). The PGproL (+) and PGgpL (+) vectors had virtually no frameshifting based on the frameshifting factor and efficiency. The protease frameshift test vector, PGproL (-) had about 2-fold lower frameshifting efficiency than the known

gag-pol frameshift sequence (PGgpL (-)). The frameshifting efficiency of the PGproL (-) ranges between 0.8% and 1.2% while the frameshifting efficiency of PGgpL (-) is between 2% and 3.5% in the assay.

Selenium and Sulfur containing Hypermodified Bases in tRNA-lys

The variety of hypermodified bases that exist in transfer RNA serve a variety of functions (Watts et al. 1978). Selenium and sulfur have been widely found in hypermodified bases of tRNA in different bases and at different substituent positions (Sprinzi et al. 1998). These elements have also been found to exist in compounds components of tRNAs in several different conformations (Ching et al. 1984, Sprinzi et al. 1998). Selenium and sulfur containing nucleotides have been found in tRNA as selenouridine or thiouridine (Sprinzi et at 1998). These elements along with oxygen, which are group VIA members of the periodic table, are interchangeable in hypermodified tRNA bases, and tend to exist at position 2 or 4 of uridine in most cases.

The strong preference for oxygen in nucleotides may not have always been the case, and may have happened as a result of the gradual predominance of oxygen in the atmosphere during evolution. Natural selection for highly functional selenium and sulfur hypermodified bases therefore may have kept the few sulfur and selenium containing tRNAs that we still see in nature today. Hypermodified selenium and sulfur containing tRNAs have been found in mice, rat, *E. coli*, *B. subtillus* and *S. typhimurium*. Identification of sulfur and selenium containing tRNAs in several different species indicate that these hypermodified bases may exist in all species but at low levels.

Studies suggest that having sulfur and selenium in place of oxygen in these tRNA-lys modified bases may have a regulatory role (Agris et al. 1997, Durant et al.

1999, Sundaram et al. 2000). NMR structural studies of the canonical U-turn of the anticodon loop of sulfur containing ecoli tRNA-lys indicate a regulatory advantage (Durant et al. 1999). The modified uridine containing nucleotide is the first base of the anticodon loop of tRNA –lys. When this nucleotide contains an unmodified base, it has been shown that the anti-codon loop is unstructured (Durant et al.1999). The thiolated uridine of the tRNA-lys was shown to be involved in stabilizing the anti-codon loop leading to a more rigid loop that may induce a stronger association between the tRNA-lys anticodon loop and lysine codon during translation (Sundaram et al. 2000).

TRNA-lys is also the known primer for reverse transcription of HIV's RNA genome (Ratner et al. 1985, Marquet et al. 1995). The 18 bases of the 3' end of tRNA-lys are complementary to the primer-binding site of the HIV RNA genome (Ratner et al. 1985, Benas et al. 2000). There is an important interaction between the U rich region of the anticodon loop of tRNA-lys and the A rich region of the primer binding site of HIV (Isel et al. 1993, 1995). Similar to situations previously described, sulfur is found in the 2 position of uridine 34 of the tRNA, which is the first base of the anticodon loop. It has been shown that a thiolated uridine base at the first position of the anticodon loop is important for recognition of the A-rich region of the primer-binding site, and for efficient initiation of transcription of the HIV genome (Benas et al. 2000, Isel et al. 1993,1995). These attributes are also facilitated by the rigidity conferred to the anticodon/codon loop interaction by the thiolated uridine 34 and a thiolated A37 (the base that follows the anticodon loop) as is the case in the NMR structural studies of the similarly substituted *E. coli* tRNA previously described (Agris et al. 1997, Sundaram et al. 2000).

The AAA codon is known to serve as the beginning of several suitable heptanucleotide slippery sequences. The AAA codon is also the complement of the tRNA-lys anticodon loop. Previous studies have suggested a role for hypermodified uridine in the anticodon loop of tRNA-lys. Some suggestions indicate that a thiolated uridine will lead to higher propensity for frameshifting in bacterial and mammalian translation systems (Agris et al. 1997). Other studies suggest a decrease in frameshifting due to a more rigid anticodon loop that interacts more readily with its corresponding codon (Sundaram et al. 2000). To give some insight to this issue the role selenium and sulfur play in frameshifting was studied. The frameshift assay developed is useful in determining a role that selenium or sulfur may play in frameshifting. To determine such a role, the frameshift assay was carried out as previously described under incubation with varying amounts of a selenium donor compound and a sulfur donor compound. Upon literature review, it was evident that the sulfur donor for incorporation of sulfur in several biological systems is L-cysteine. A candidate commonly used to introduce selenium in selenobiology experiments is sodium selenite. Hence the effects of L-cysteine and sodium selenite on the frameshift assay were quantitated to give some understanding of the role sulfur and selenium play in frameshifting and in HIV's mechanisms of survival as a whole.

Materials and Methods

Frameshift Assay Constructs

The frameshift assay vectors were constructed as described in earlier sections. All test frameshift sequences were cloned and placed upstream of a cloned firefly luciferase gene (Promega) and downstream of a green fluorescent protein gene

(Clontech). These were arranged accordingly in a PCI vector (Promega) and expression was assessed.

Cell Culture and Transfection of Cell Lines

293-T cells were grown in Dubelco minimal essential medium and deviated in a few components depending on the element to be tested. The cells were grown in 24 well plates. For L-cysteine studies 293-T cells were grown in DMEM lacking L-cysteine, L-methionine, L-glutamine, sodium bicarbonate and sodium pyruvate. The media in L-cysteine studied sample wells was supplemented with 10% Defined fetal bovine serum (Hyclone), which contains approximately 1.09μM L-cysteine, 2mM L-glutamine, 1.5g/L sodium bicarbonate and 1.0 mM sodium pyruvate and 1mM methionine. The methionine was decreased to limit possible interference of sulfur donation by L-methionine, though this had to be balanced with regular methionine translational requirements. The sample wells were supplemented accordingly with L-cysteine ranging in concentration from 0 to 3mM. Cells are starved of cysteine for 24 hours and subsequently incubated for 48 hours at 37º C with 5% CO_2 with media changes and L-cysteine supplementation every 24 hours.

For sodium selenite studies 293-T cells were grown in DMEM lacking a selenium source. The media was supplemented with Defined Fetal Bovine Serum (Hyclone), which supplements the media with 0.05uM sodium selenite. The sample wells were supplemented accordingly with sodium selenite ranging in concentration from 0 to 3uM. Cells are starved for 24 hours and subsequently incubated for 48 hours at 37º C with 5% CO_2 with media changes and sodium selenite supplementation every 24 hours.

Cell lines were grown to 95% confluency. Lipofectamine (invitrogen) was used to transfect the cell lines in each well with 0.5μg of PGproL (-) DNA in Opti-MEM (Invitrogen) reduced serum media. Following a 24-hour transfection period, the media of transfected cells is replaced with DMEM and supplemented with L-cysteine or sodium selenite and the cells are incubated at 37º C with 5% CO_2 for the 48-hour post transfection period.

Reporter Gene Assay

The transfected cells were lysed by incubating the cells with 200μl of glo-lysis buffer (Promega) per well for 5 minutes. Crude extracts are centrifuged and supernatant is recovered. The supernatant is assayed for reporter gene activity.

Luciferase activity and Green fluorescent protein fluorescence is quantitated in a Fluostar Optima Galaxy plate reader. 10μl of the cell lysate is added to 40μl of Promega luciferase assay buffer, which contains luciferin as the active luminescent agent. Cell lysates are pipetted into a white luminescent reading plate. The luciferase assay buffer is dispensed by the pump into the 96 well plate containing cell lysates and luminescence is recorded over a 12 second interval. Background luminescence is normally 3 to 4 fold lower than test samples. Fluorescence of green fluorescent protein is measured using a Fluostar Optima Galaxy, or Polarstar Optima plate reader. Cell lysate samples are excited at 488nm wavelength and relative intensity of emission at approximately 507nm is recorded.

Comparing fluorescence of green fluorescent protein and luminescence of luciferase quantitates Frameshifting. Luciferase translation measures frameshifting while green fluorescent protein translation standardizes the assay. Therefore

frameshifting is represented by luciferase luminescence as a factor of green fluorescent protein fluorescence.

Results

Effects of Sulfur on Frameshifting

To determine the effects of sulfur on frameshifting, the media of PGproL (-) transfected 293-T cell sample wells was supplemented with cysteine as previously mentioned. 293-T cells were chosen based on the high transfection efficiency of the cell line and literature recommendation that the kidney has a significant number of sulfurtransferases, and a suitable number of available sulfur transfer sites available in tRNA relative to the number of sulfurtransferases (Harris et al. 1978). The recommended concentration of cysteine in cell culture media is approximately 0.2mM L-cysteine. Increasing the concentration of L-cysteine caused the ratio of frameshifting to decrease as shown in Figure 21. This decrease cannot be attributed to effects of cysteine on translation, because several cases exist where the rate of luciferase activity and GFP fluorescence was actually highest for the samples supplemented with the highest amounts of cysteine, but their ratios of frameshifting still proved to be lower than for samples supplemented with lower concentrations of cysteine.

Effects of selenium on frameshifting

To determine the effects of selenium on frameshifting, the media of PGproL (-) transfected 293-T cells were supplemented with sodium selenite as previously mentioned. Sodium selenite is toxic to most cell lines in high doses, but is necessary for efficient cell culture. The recommended concentration of sodium selenite in cell culture media is approximately 0.05 uM. The selenium component of the DMEM

media is from the Fetal Bovine Serum added to the media, which makes up 10% of the total media volume. In increasing the concentration of sodium selenite, the ratio of frameshifting was found to decrease. This decrease could not be attributed to effects of sodium selenite on translation, because several cases exist where the rate of luciferase activity and GFP fluorescence were actually highest for the samples supplemented with the highest amounts of sodium selenite, but their ratios of frameshifting still proved to be lower than for samples supplemented with lower concentrations of cysteine. These results are illustrated in Figure 2-22

References

Agris, P. F., R. Guenther, et al.. (1997). "Unconventional structure of tRNA(Lys)SUU anticodon explains tRNA's role in bacterial and mammalian ribosomal frameshifting and primer selection by HIV-1." RNA 3(4): 420-8.

Benas, P., G. Bec, et al. (2000). "The crystal structure of HIV reverse-transcription primer tRNA(Lys,3) shows a canonical anticodon loop." Rna 6(10): 1347-55.

Ching, W., (1984) “Occurrence of selenium-containing tRNAs in mouse leukemia cells” Proc. Natl. Acad. Sci. 81:3010-3013

Durant P.C., Davis D.R. (1999).”Stabilization of the anticodon stem-loop of tRNA(LyS,3) by an A(+)-C base-pair and by pseudouridine” Journal of Molecular Biology 285 (1): 115-131

Farmerie, W. G., D. D. Loeb, et al. (1987). "Expression and processing of the AIDS virus reverse transcriptase in Escherichia coli." Science 236(4799): 305-8.

Harris, C.,(1978). “Mammalian tRNA sulfurtransferases: properties of the enzyme in rat liver” Nucl. Acids Research 5(2)599-613

Isel, C., R. Marquet, et al. (1993). "Modified nucleotides of tRNA(3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription." J Biol Chem 268(34): 25269-72.

Isel, C., C. Ehresmann, et al. (1995). "Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA(3Lys) (template/primer)." J Mol Biol 247(2): 236-50.

Isel, C., J. M. Lanchy, et al. (1996). "Specific initiation and switch to elongation of human immunodeficiency virus type 1 reverse transcription require the post-transcriptional modifications of primer tRNA3Lys." Embo J 15(4): 917-24.

Isel, C., E. Westhof, et al. (1999). "Structural basis for the specificity of the initiation of HIV-1 reverse transcription." Embo J 18(4): 1038-48.

Jacks, T., M. D. Power, et al. (1988). "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression." Nature 331(6153): 280-3.

Marquet, R., C. Isel, et al. (1995). "tRNAs as primer of reverse transcriptases." Biochimie 77(1-2): 113-24.

Molla, A. and W. Kohlbrenner (2003). "Evolving therapeutic paradigms for HIV and HCV." Curr Opin Biotechnol 14(6): 634-40.

Navia, M. A., P. M. Fitzgerald, et al. (1989). "Three-dimensional structure of aspartyl protease from human immunodeficiency virus HIV-1." Nature 337(6208): 615-20.

Pearl, L. H. and W. R. Taylor (1987). "Sequence specificity of retroviral proteases." Nature 328(6130): 482.

Ratner, L., W. Haseltine, et al. (1985). "Complete nucleotide sequence of the AIDS virus, HTLV-III." Nature 313(6000): 277-84.

Sprinzi, M., Horn, C. et al. (1998). “Compilation of tRNA sequences and sequences of tRNA genes.” Nucleic Acids Research, 26(1): 148-153

Sundaram M, Durant PC et al (2000) “Hypermodified nucleosides in the anticodon of tRNA(Lys) stabilize a canonical U-turn structure” Biochemistry 39 (50): 15652-15652

Taylor, E. W., Ramanathan, C.S., et al. (1994). "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(17): 2637-54.

Taylor, E.W., Ramanathan, C.S., and Nadimpalli, R.G. (1996) “A general method for predicting potential new genes in nucleic acid sequences: application to the human immunodeficiency virus.“ In: Computational Medicine, Public Health and Biotechnology, M. Witten, Ed. World Scientific, Singapore. Ser. Math. Biol. Med. 5: 285-309.

Watts, M. T. and I. Tinoco, Jr. (1978). "Role of hypermodified bases in transfer RNA. Solution properties of dinucleoside monophosphates." Biochemistry 17(12): 2455-63.

Zhang W., Ramanathan C.S., et al (1999) “Selenium-dependent glutathione peroxidase modules encoded by RNA viruses” Biol. Trac. Elem. Res. 70 (2): 97-116

Turner, B. G. and M. F. Summers (1999). "Structural biology of HIV." J Mol Biol 285(1): 1-32.

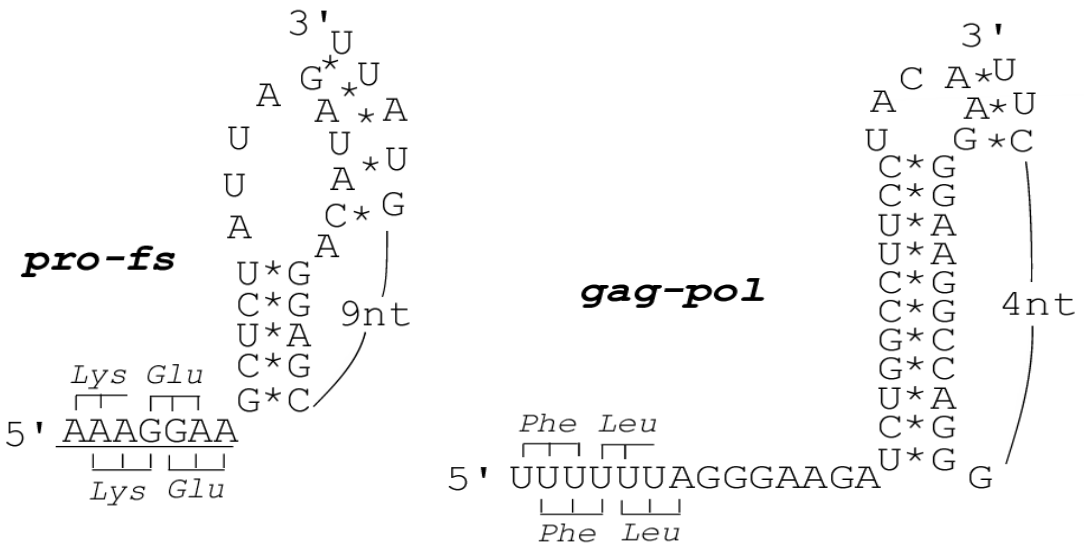


Figure 2-1

Protease and gag-pol frameshift sequences. Slippery sequence and pseudoknot in the protease (proposed) and gag-pol (Jacks et al 1988) genes of HIV-1. Proposed pro-fs slippery sequence contains a run of lysine and glutamine codons in the 0 and –1 frame, which would suit slippage of the ribosome during frameshifting.

Table 2-1

Primers designed for construction of frameshift assay vectors. These primers were designed for PCR cloning of the luciferase gene as well as frameshift sequence. They were also used to incorporate restriction sites, and stop codons at the end of frameshift sequences.

Oligos designed for PCR

All primers are presented as synthesized in the 5' → 3' direction

Pos-Luc-Sal I
TTG GCC GTC GAC ATG GAA GAC GCC AAA AAC ATA AAG
Sal I

Neg-Luc-Sal I
TTG GCC GTC GAC AAA TGG AAG ACG CCA AAA ACA TAA AG
Sal I

Luc-Not I
AAC CGG GCG GCC GCT TTA CAA TTT GGA CTT TCC GCC CT
Not I

GP-fs-Fw

TTG GCC TCT CGA GAG GCT AAT TTT TTA GGG AAG ATC
Xho I

GP-fs-Fw
TTC CGG GTC GAC TTA GCT CTG AAG AAA ATT CCC AGG
Sal I

Pro-fs-Fw
TTG GCC TCT CGA GGG CAA CTA AAG GAA GCT CTA TTA G
Xho I

Pro-fs-Rv
TTC CGG GTC GAC TTA TTC TTC TAA TAC TGT ATC ATC
Sal I

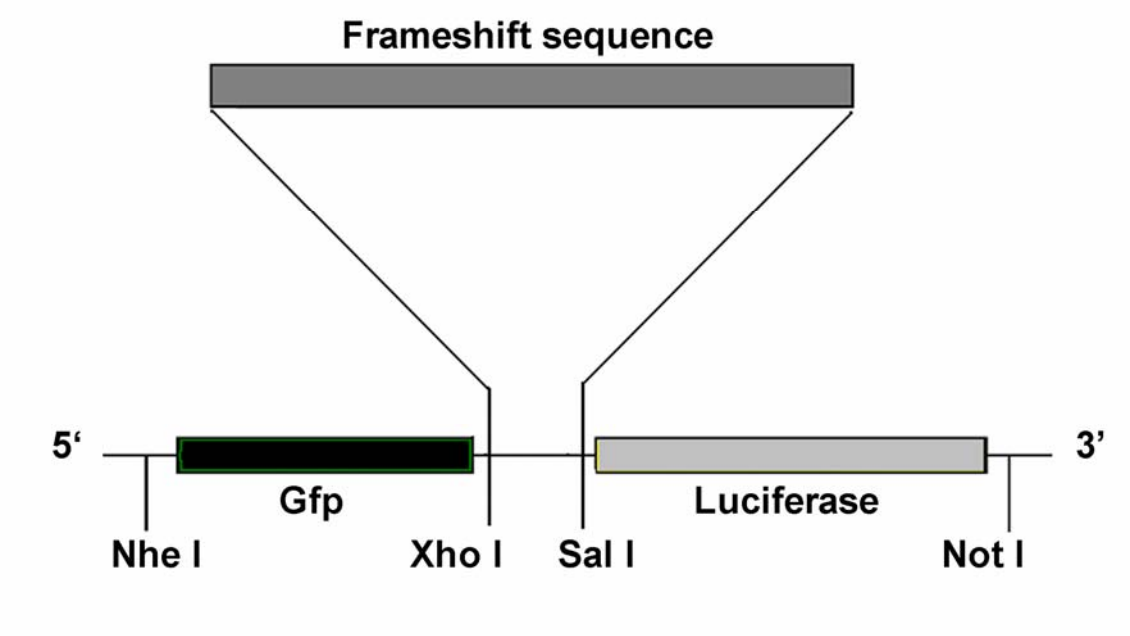


Figure 2-2

A map of the frameshift assay construct used in the frameshift assay. The map describes the organization of gene components for all (+) and (-) assaying vectors**.**

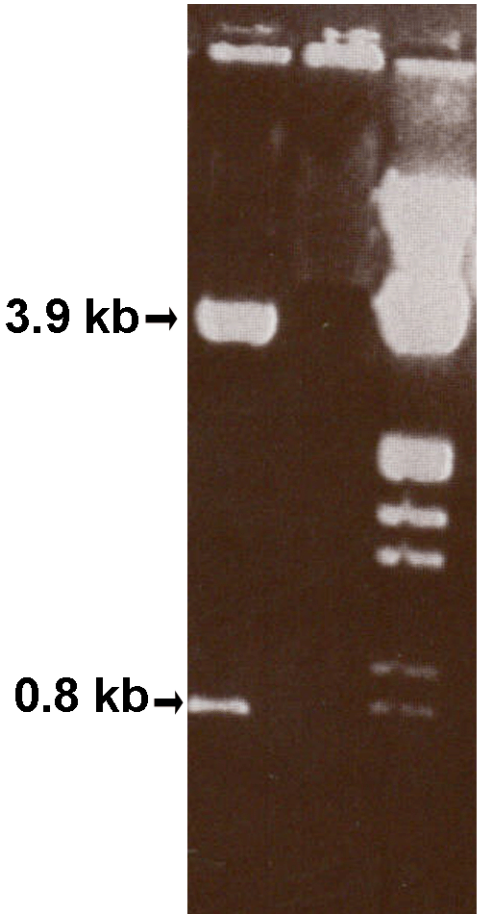


Figure 2-3

Restriction digest of PEGFP-C1. Agarose gel showing PEGFP-C1 digested with NheI and Bgl II in 1st lane and DNA ladder in 3rd lane. Bands were detected at 3.8 KB and 0.8 KB by ethidium bromide and UV detection and correspond to the GFP gene and remainder of the PEGFP-C1 vector respectively.

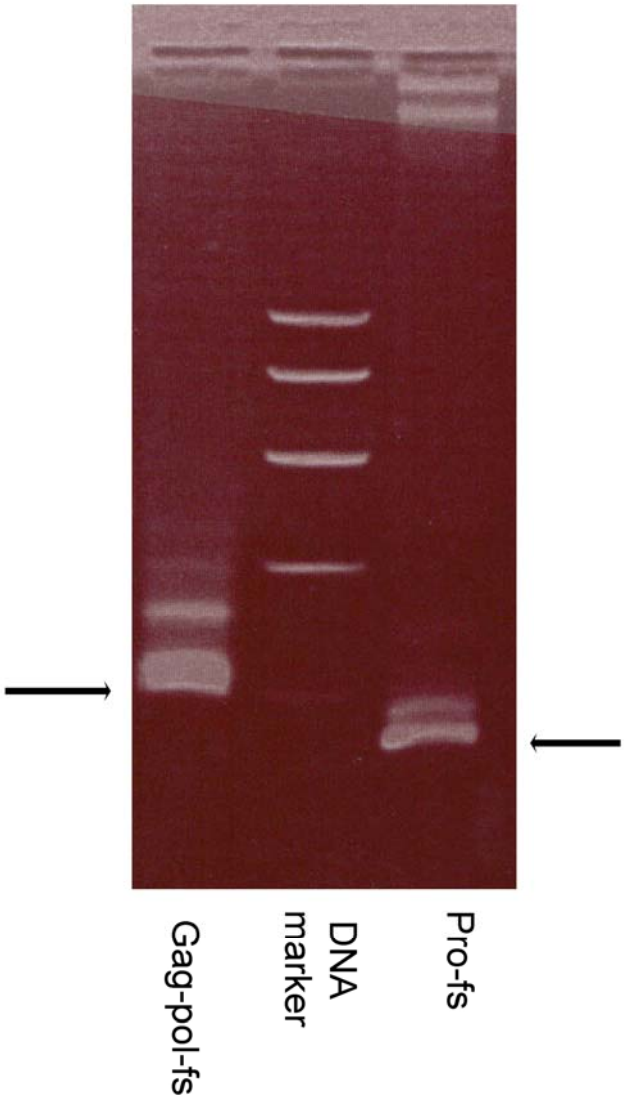


Figure 2-4

PCR amplification of gag-pol and protease frameshift sequences. Agarose gel with frameshift sequence bands indicated by the arrows.

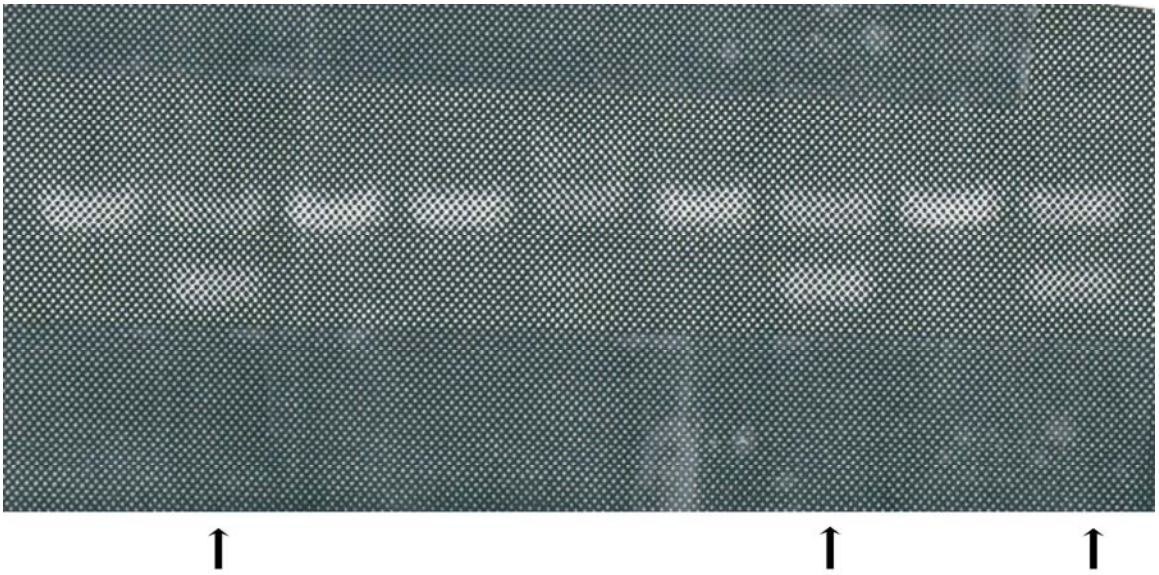


Figure 2-5

PCR identification of positive clones. The Agarose gel shows identification of positive clones with the primers GP-fs-Fw, GP-fs-Rv and Luc-Not I. PGgpL (-) clones are in lanes 1 – 4 and 6 - 8. Lane 5 electrophoresis of DNA marker and lane 9 is positive control with T7-Luc vector (Promega) with Pos-Luc-Sal I and Neg-Luc-Sal I primers. 2 positive PGgpL (-) clones found in (lanes 2,7).

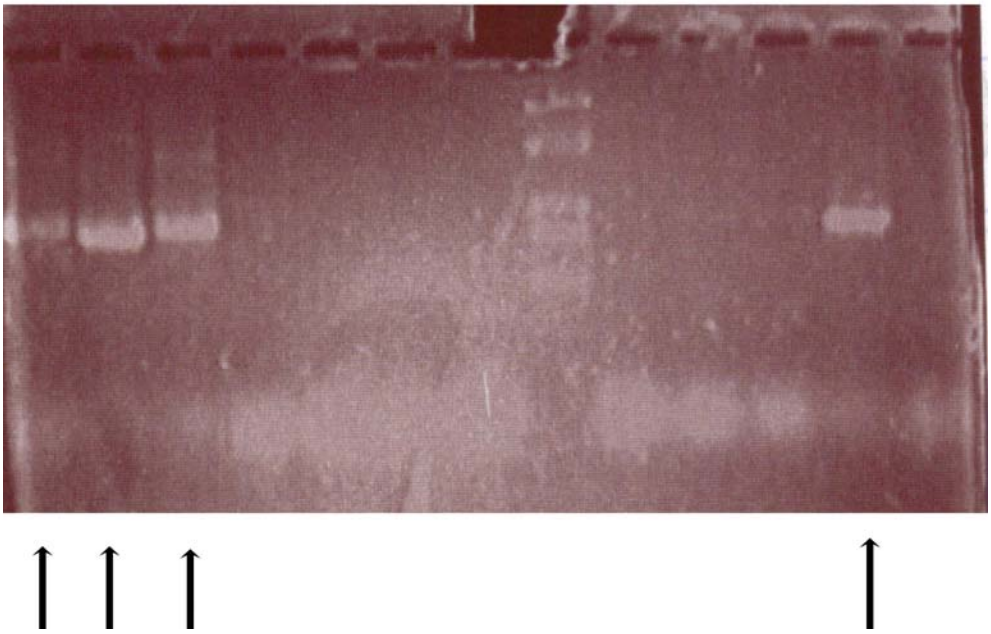


Figure 2-6

PCR identification of positive clones. Agarose gel shows PCR identification of positive clones with primers GP-fs-Fw, and Luc-Not I. PGgpL (+) clones are in lanes 1 – 7 and 10 - 11. Lane 12 is PCR of T7-Luc with Pos-Luc-Sal I and Neg-Luc-Sal I as positive control. DNA marker in lane 8. 3 positive PGgpL (+) clones found in lanes 1,2 and 3.

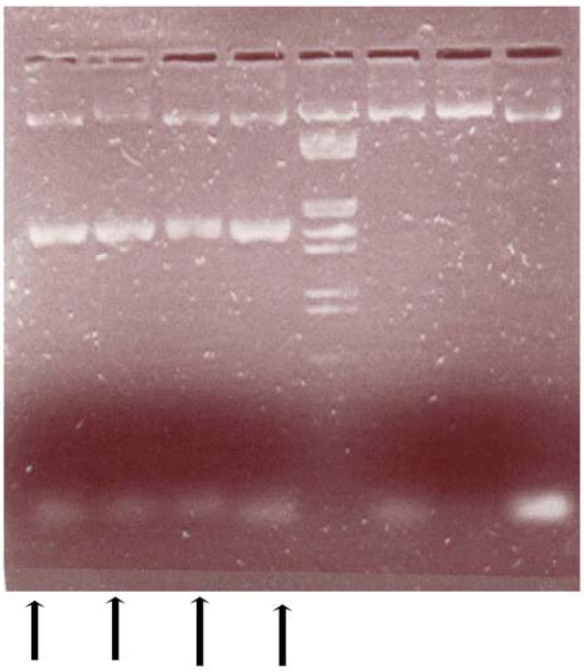


Figure 2-7

PCR identification of PGproL (+) positive clones. Clones tested were in Lanes 1 – 4 and 6 – 8. Lanes 1 – 4 show positive clones identified.

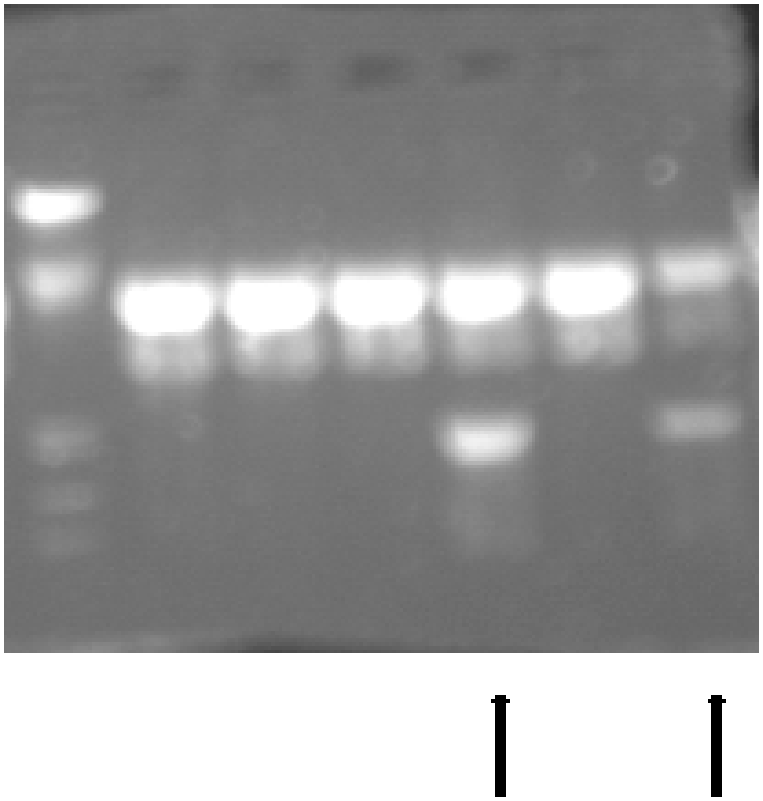


Figure 2-8

PCR identification of PGproL (-) clones. PGproL (-) clones were in lanes 2 - 6, while lane 7 was the positive control. Clone in lane 5 was positive.

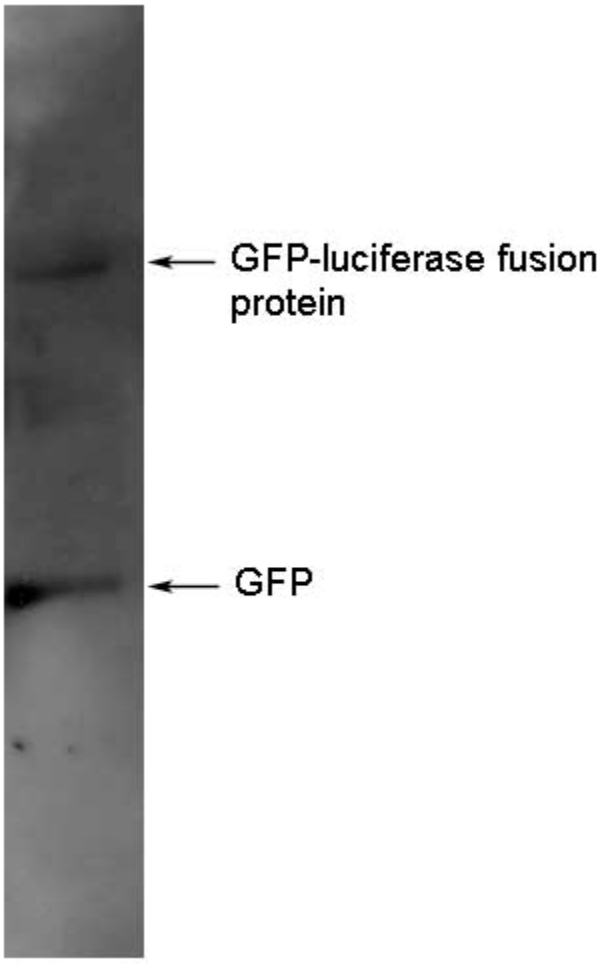
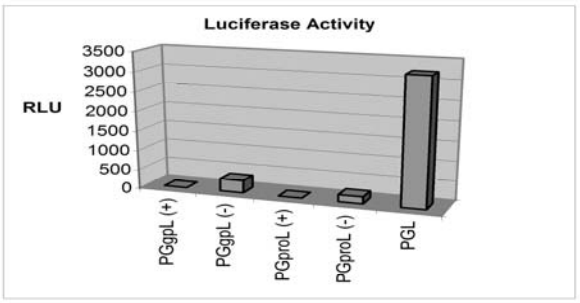
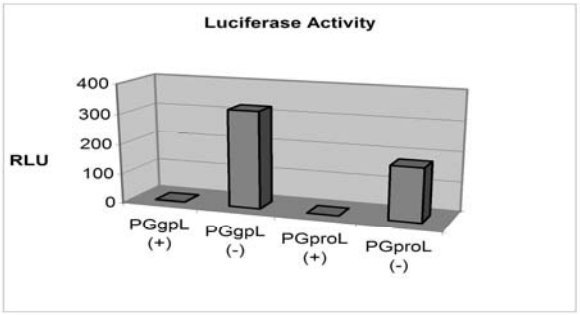


Figure 2-9

Western blot detection of GFP. Western blot analysis of representative construct identified 2 bands corresponding to the GFP-luciferase fusion protein and unframeshifted GFP.





Construct	Luciferase Activity in Relative Light Units Average +/- Standard Deviation
PGproL (+)	0.02 +/- 0.002
PGproL (-)	180 +/- 21
PGgpL (+)	0.6 +/- 0.2
PGgpL (-)	320 +/- 65
PGL	3200 +/- 260

Figure 2-10

Luminescence from cell lysates expressing frameshift assay constructs.

Luminescence was measured on a TD 20/20 luminometer. Cell lysates of PGproL (+) and PGgpL (+) expressing cells show virtually no luminescence ($P<0.001$ for PGproL (+) and PGproL (-), $P<0.0001$ for PGproL (+) and PGL, $P<0.001$ for PGgpL (+) and PGgpL (-), $P<0.0001$ for PGgpL (+) and PGL).

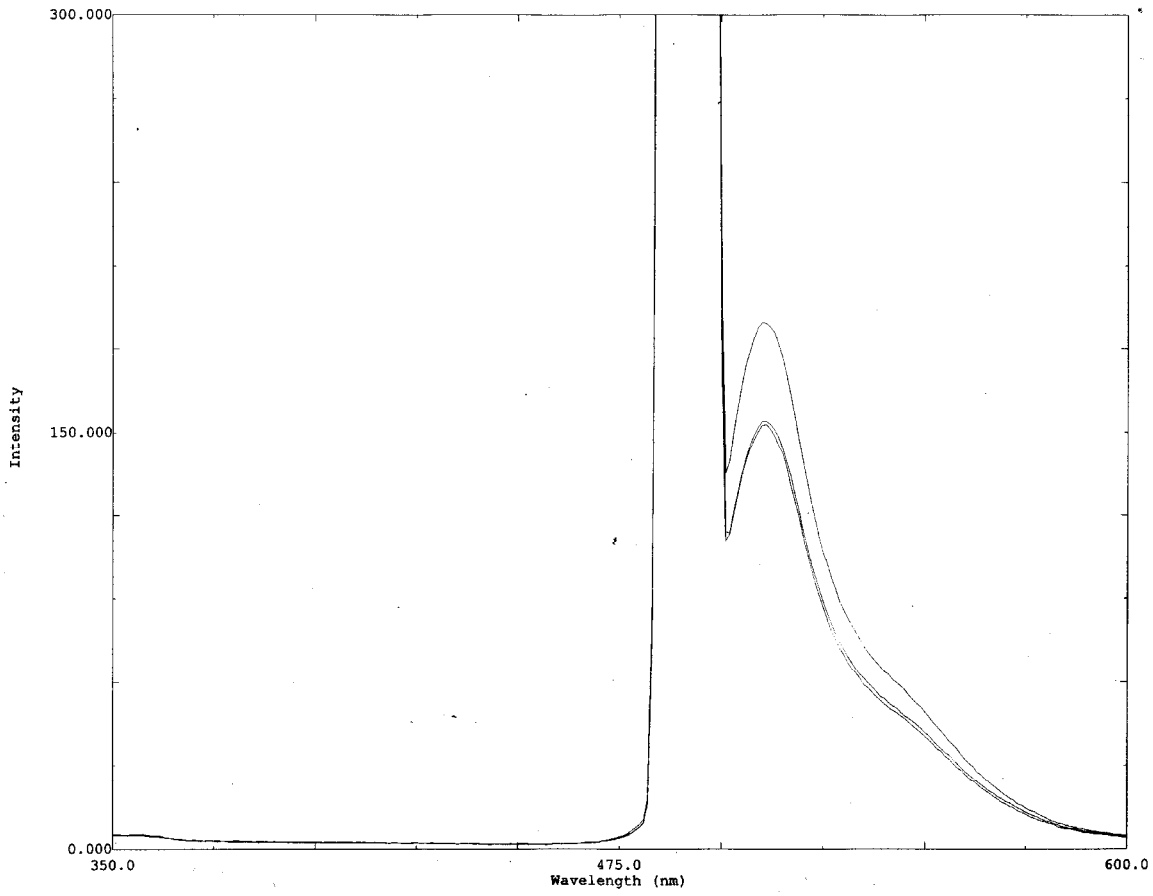


Figure 2-11

GFP fluorescence emission spectra of PGproL (+) lysates. Fluorescence spectra from three PGproL (+) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

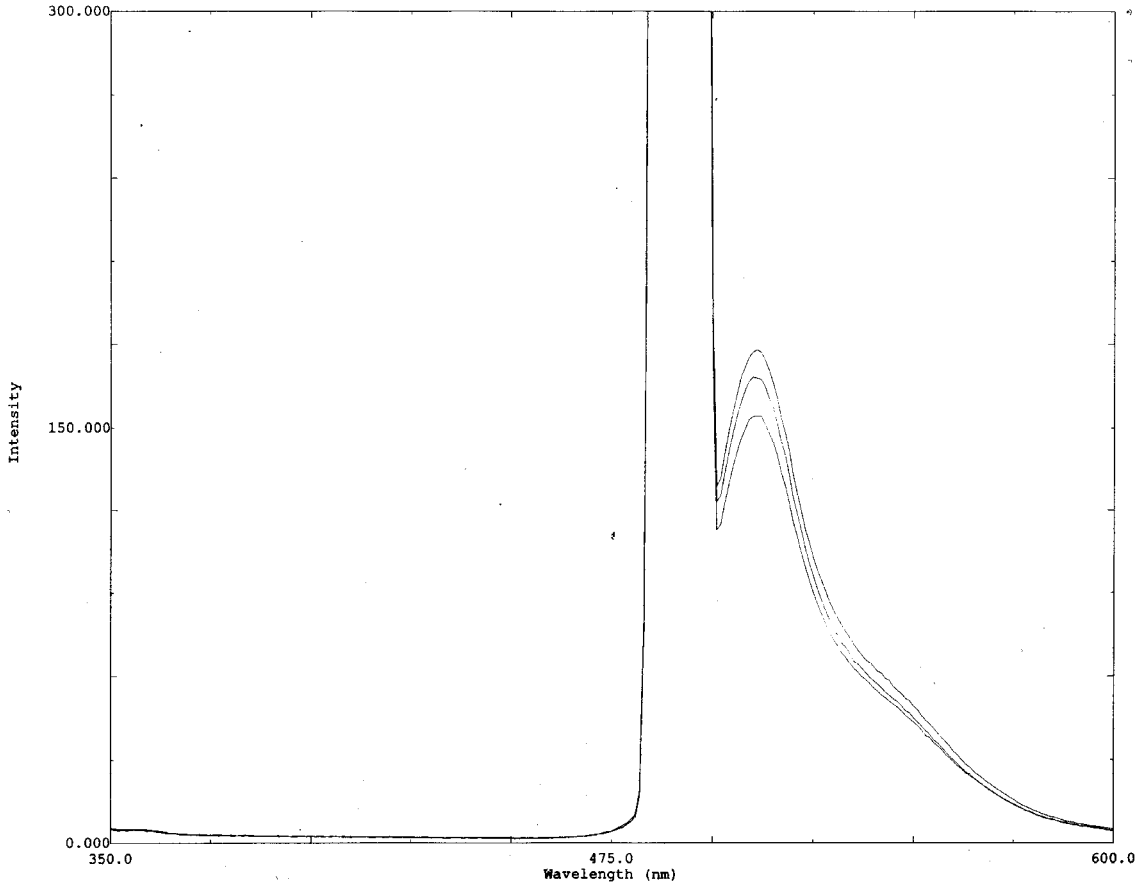


Figure 2-12

GFP fluorescence emission spectra of PGproL (-) lysates. Fluorescence spectra from three PGproL (-) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

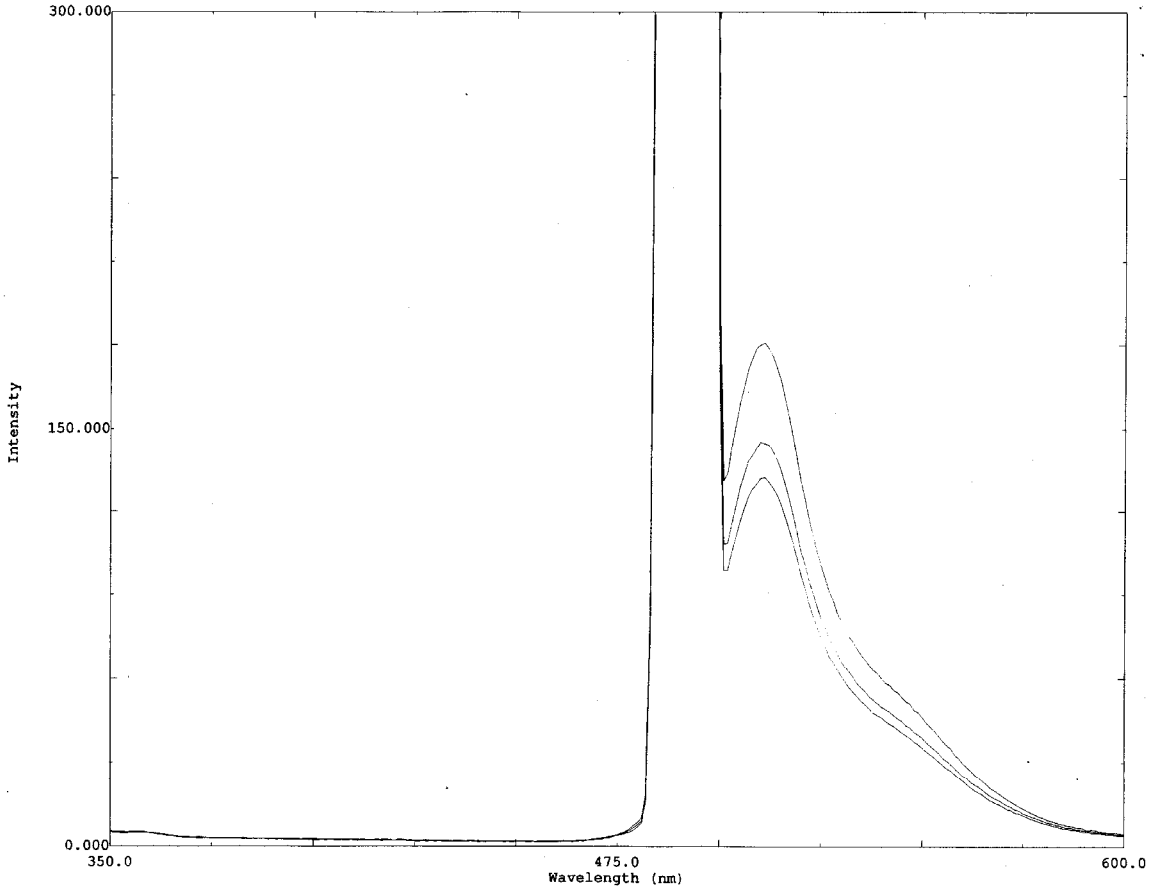


Figure 2-13

GFP fluorescence emission spectra of PGgpL (+) lysates. Fluorescence spectra from three PGgpL (+) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

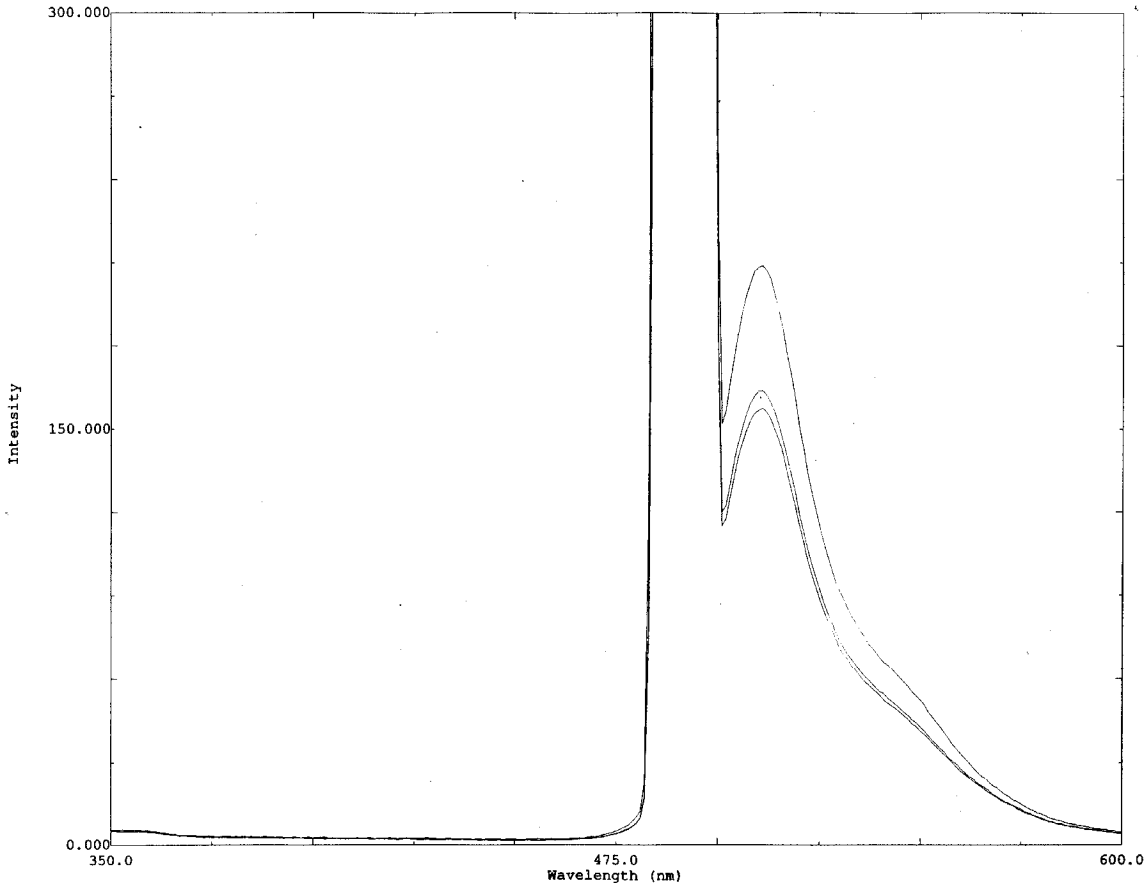


Figure 2-14

GFP fluorescence emission spectra of PGgpL (-) lysates. Fluorescence spectra from three PGgpL (-) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

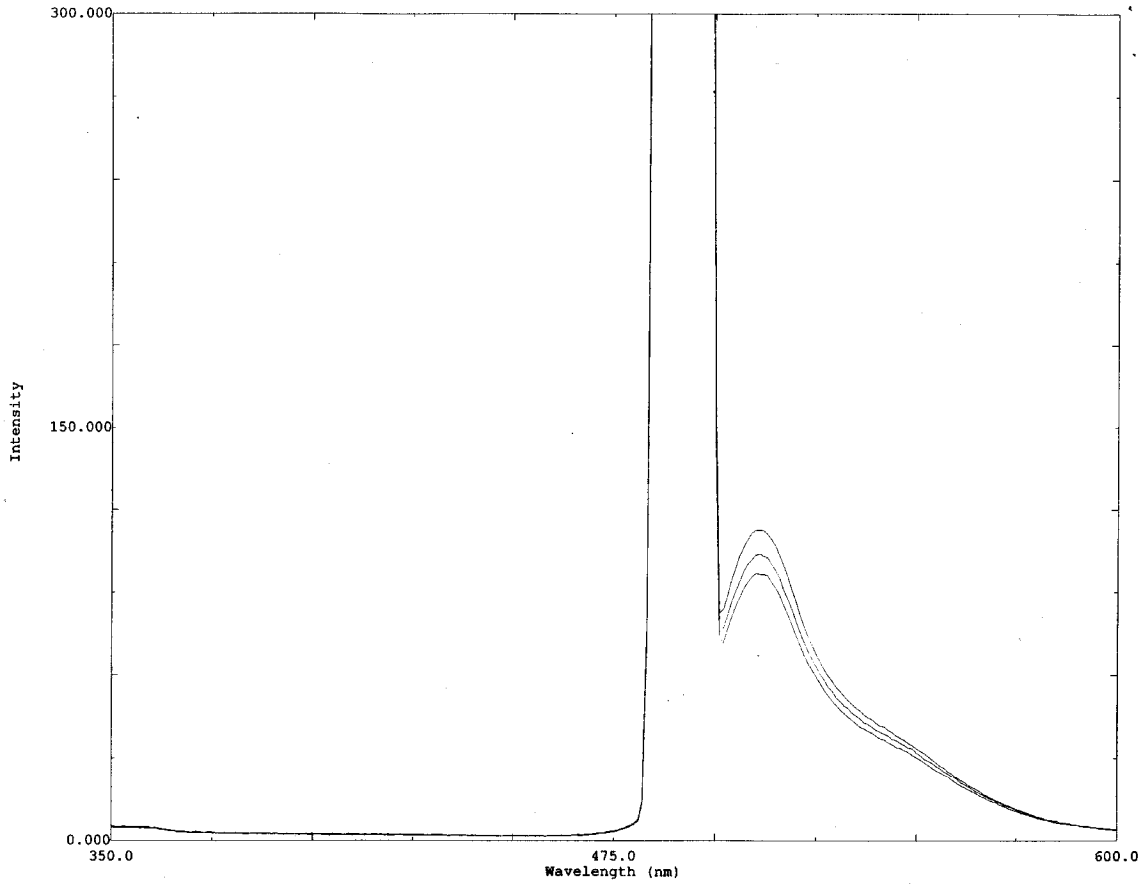
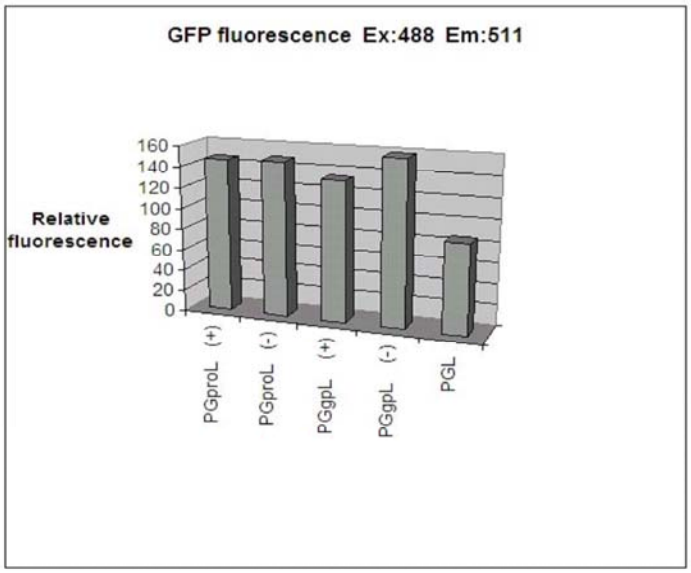


Figure 2-15

GFP fluorescence emission spectra of PGL lysates. Fluorescence spectra from three PGL expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

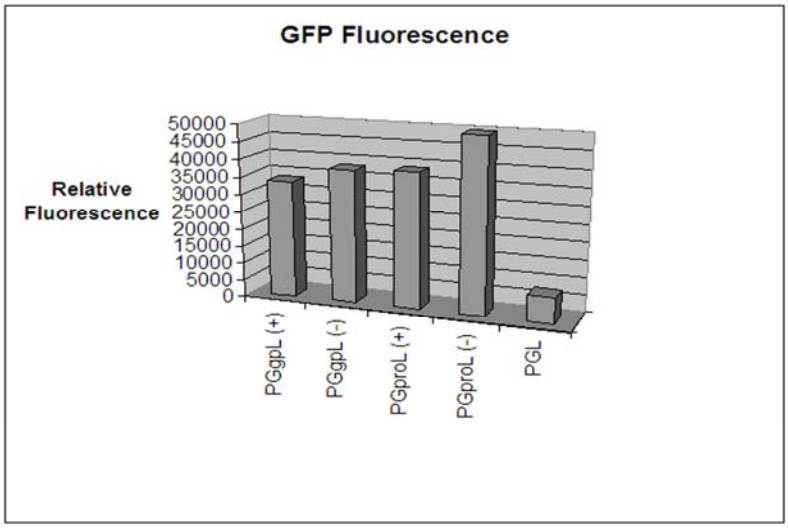


Construct	GFP Relative Fluorescence Excitation: 488 Emission: 511 Average +/- Standard Deviation
PGproL (+)	146 +/- 21
PGproL (-)	148 +/- 12
PGgpL (+)	135 +/- 25
PGgpL (-)	159 +/- 28
PGL	86 +/- 8

Figure 2-16

Average Relative fluorescence of frameshift construct expressed cell lysates.

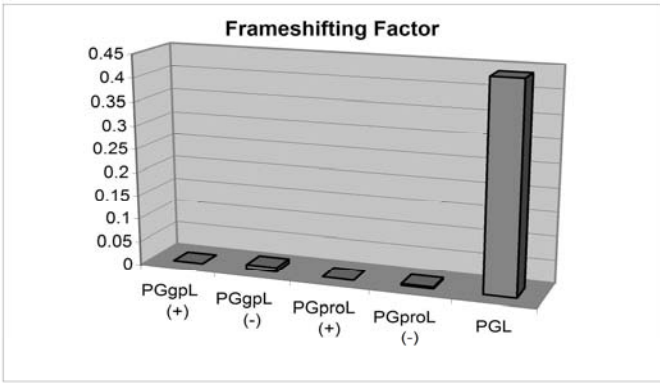
GFP fluorescence was similar for constructs with similar frameshift sequences.

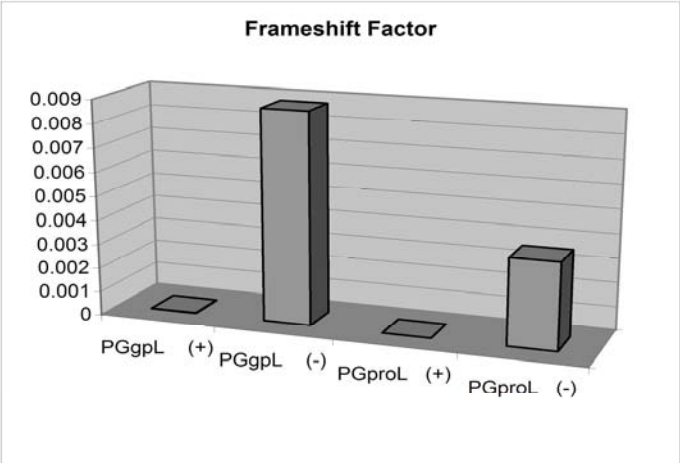


Vector	GFP Fluorescence Average +/- Standard Deviation
PGproL (+)	39000 +/- 16000
PGproL (-)	50000 +/- 8100
PGgpL (+)	34000 +/- 5000
PGgpL (-)	38000 +/- 11800
PGL	7400 +/- 720

Figure 2-17

GFP fluorescence measured on FLUOstar Optima plate reader. GFP fluorescence was measured on a FLUOstar Optima plate reader for enhanced quantitation of GFP fluorescence and calculation of frameshift efficiency.





Construct	Frameshift Factor	Frameshift Efficiency
PGproL (+)	4.6×10^{-7}	0.0001%
PGproL (-)	0.004	1%
PGgpL (+)	1.8×10^{-5}	0.004%
PGgpL (-)	0.009	2%
PGL	0.44	100%

Figure 2-18

Frameshift Factor of constructs. The frameshift factor is calculated by determining the ratio of luciferase activity to GFP fluorescence. PGproL (+) and PGgpL (+) showed virtually no frameshifting. Frameshifting efficiency is a constructs' frameshift factor as a percentage of PGL frameshift factor.

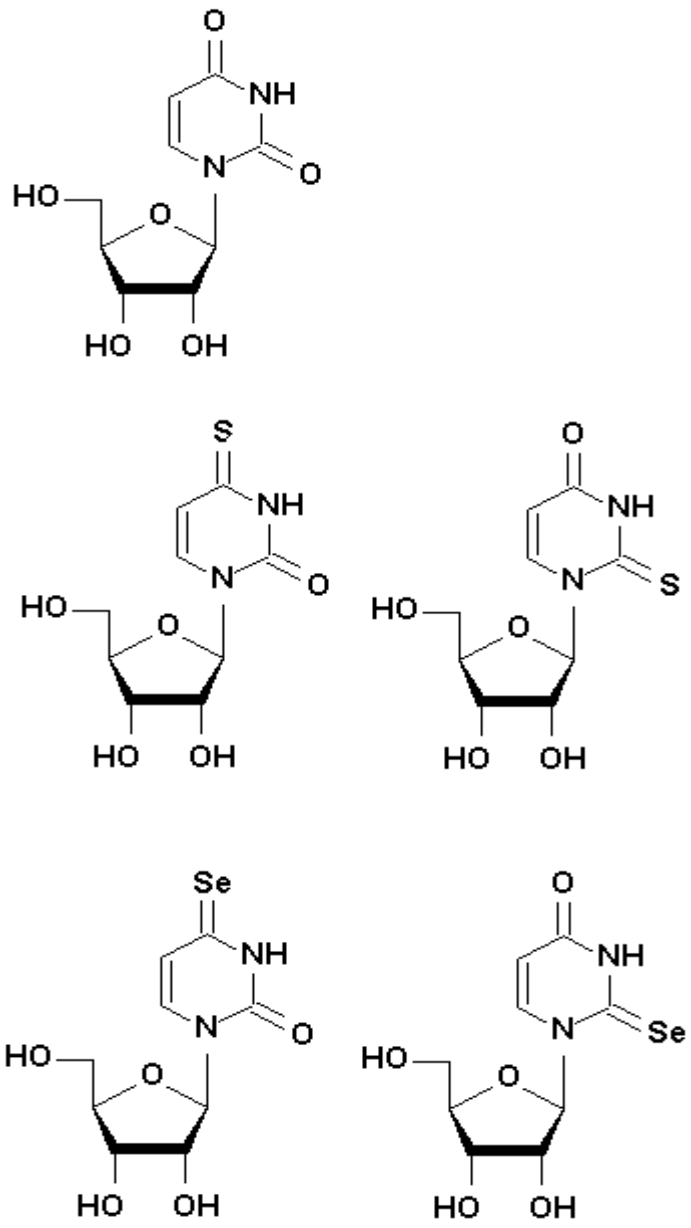


Figure 2-19

Selenium and sulfur substituted uridine. Hypermodified uridine is commonly found in nature to contain sulfur or selenium in place of oxygen in the 2’ or 5’ positions of the base.

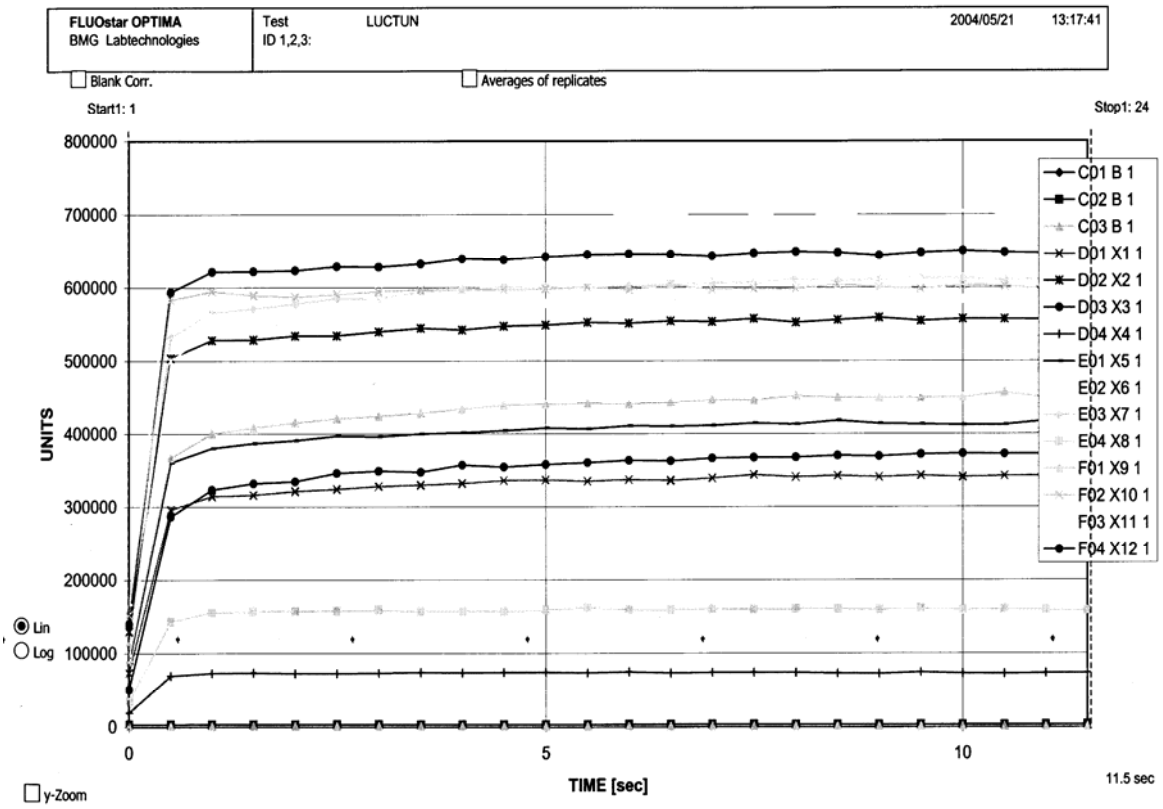


Figure 2-20

Luminescence of constructs expressing cell lysates. Luminescence of cell lysates is quantitated in a FLUOstar Optima plate reader read over a 12 second period.

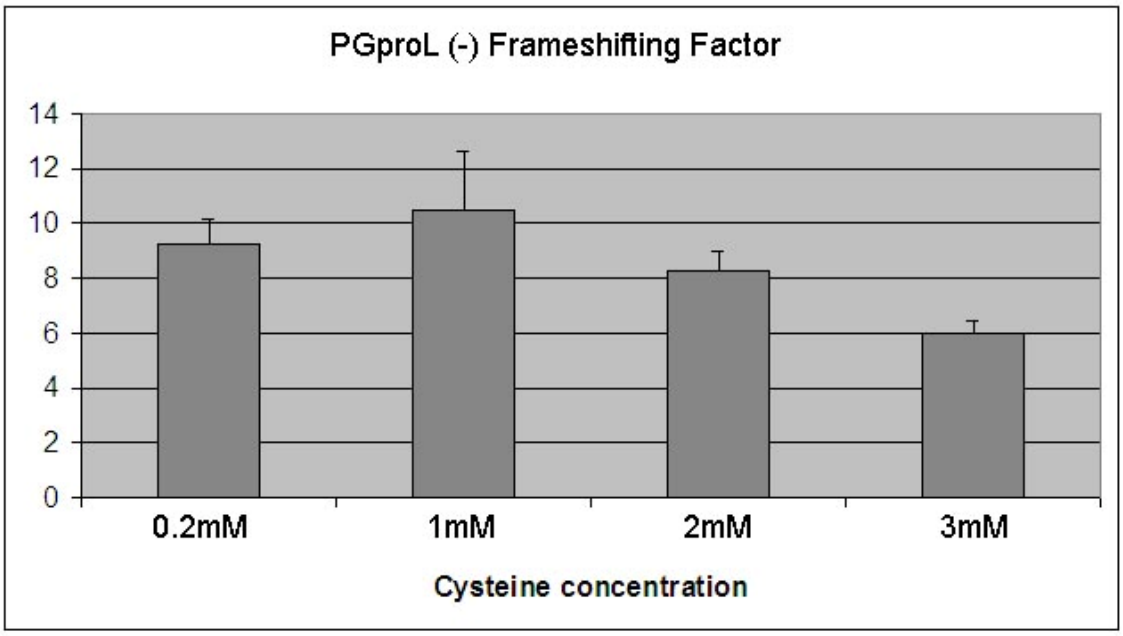


Figure 2-21

The Frameshift Factor of PGproL (-) with cysteine supplementation. Sample wells were supplemented with cysteine concentrations ranging from 0.2mM to 3mM. The Frameshift Factor of PGproL (-) was found to decrease with an increase in cysteine concentrations.

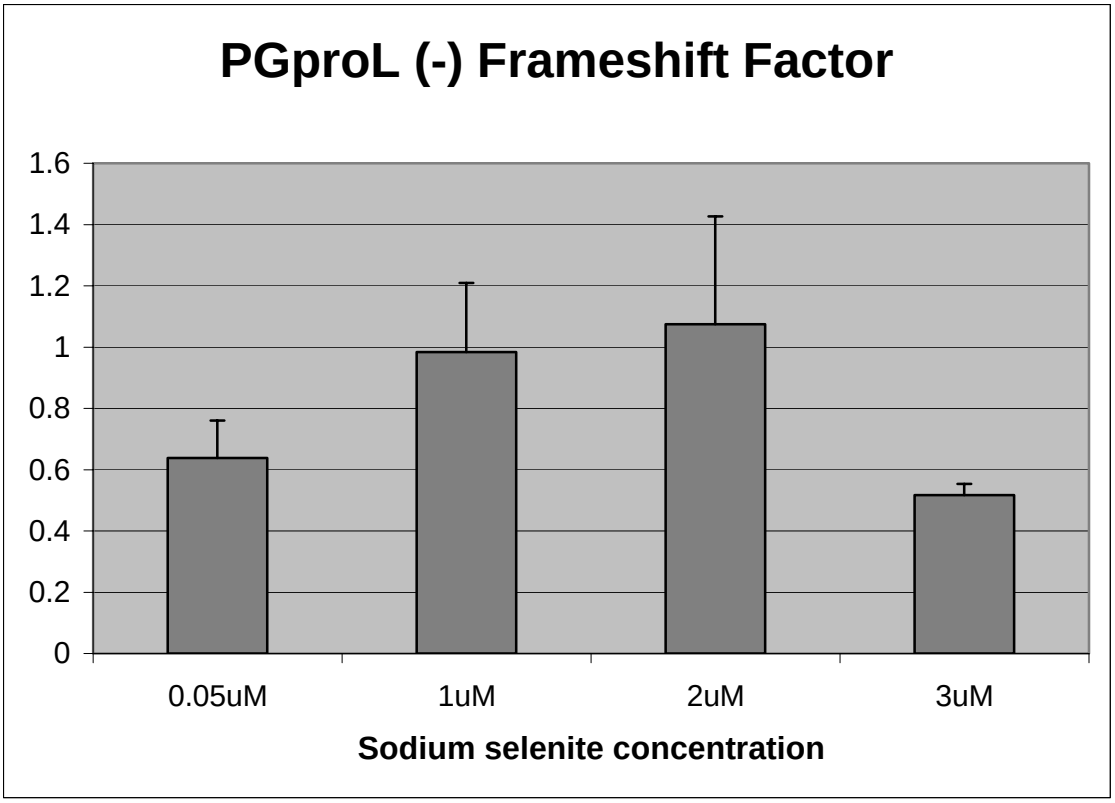


Figure 2-22

The Frameshift factor of PGproL (-) with sodium selenite supplementation.

PGproL (-) cell samples were supplemented with sodium selenite concentrations from 0.05uM to 3uM. The Frameshift Factor of PGproL (-) was found to decrease with an increase in cysteine concentrations.

CHAPTER 3

RIBOSOMAL FRAMESHIFTING IN THE ENVELOPE GENE OF HIV-1[1]

---

[1]Olubajo, B and Taylor, E.W. To be submitted to Journal of Nucleic Acid Research

Overview

This chapter contains a review of ribosomal frameshifting and the properties and function of proteins expressed from the envelope gene of HIV. Furthermore, the development of a GFP-luciferase, dual-reporter frameshift assay is discussed, and frameshifting of a proposed envelope frameshift sequence is assessed. The effects of cysteine, selenium and arginine on frameshifting in the envelope region of HIV are also assessed. The main objectives of the experiments conducted in this study was to: 1. Determine the efficiency of the predicted frameshift sequence in the protease region of HIV. 2. Determine the effects of sodium selenite and cysteine supplementation on the predicted protease frameshift sequence through enhancement of sulfur or selenium hypermodified base synthesis and incorporation into tRNA-lys. 3. Determine the effects of L-arginine on a proposed hungry arginine codon in the frameshift sequence of the envelope gene.

Introduction and Review

The proteins that form the external structural elements of the HIV virion particle are encoded from the envelope gene. The envelope gene of HIV-1 is approximately 1700 base pairs and encodes gp160, a HIV glycoprotein precursor. Proteolytic cleavage of gp160 results in the gp120 and gp41, and these 2 proteins form a glycoprotein heterodimer, which resides on the surface of the HIV virion particle. Gp41 serves as a transmembrane anchor of the dimer, and gp120 is found on the surface of the virus. The gp120 protein is thought to be responsible for the attachment of HIV to the chemokine receptors on T-cells, followed by fusion of the HIV particle with the host cell membrane (Berger et al. 1999). Gp120 is found to bind both CD4 molecules and CCR5

receptors on the surface of T-lymphocytes, and both interactions are important to HIV cell evasion (Landau et al. 1988, Berger et al. 1999).

Computational analysis of the HIV-1 genome identified several potential selenoproteins in the HIV-1 genome with conserved UGA codons (Taylor et al. 1994, 1996). Selenoproteins contain selenium in the form of selenocysteine, an amino acid that is encoded by the UGA codon. Since UGA was initially known as the codon responsible for translation termination, selenoprotein reading frames are difficult to detect. These highly conserved UGA codons discovered in the putative selenoprotein reading frames of the HIV-1 genome could not be attributed simply to translation termination of known HIV-1 proteins.

One of the proposed reading frames was discovered in the gp41coding region of envelope gene of HIV-1. The reading frame was thought to encode a selenoprotein encoded as a –1 frameshift fusion protein from the *env* gene. A putative envelope cleavage site upstream of this reading frame was discovered which further supports the validity of the reading frame (Taylor et al. 1994). Sequence analysis and homology studies of the gene, which was named *env-fs,* was found to have high amino acid similarity to several known selenium dependent mammalian glutathione peroxidase proteins (Zhao et al. 2000). The similarity of the *env-fs* gene to an aligned set of selenocysteine containing mammalian glutathione peroxidase scored 6.3 standard deviations higher than randomized env-fs sequences (Zhao et al. 2000).

The high similarity produced a basis for molecular modeling of a proposed *env-fs* encoded protein. The viral glutathione peroxidase was modeled based on x-ray crystal structure of bovine glutathione peroxidase. Some outer structural regions of the

mammalian glutathione peroxidase were not intrinsic to the viral glutathione peroxidase, but the important core structural elements as well as the orientation of the catalytic site were still evident in the viral protein model (Zhao, et al. 2000). 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 consistently exhibited increased glutathione peroxidase activity over control cell lines (Zhao et al. 2000).

The idea of a –1 frameshift in the env-fs gene that would produce an HIV glutathione peroxidase is supported by a pseudoknot and slippery sequence. The two are RNA elements required for ribosomal frameshifting to occur.

Ribosomal frameshifting is a method used by several organisms to maximize protein output from DNA sequences. It is a process that is very common in viruses, and is important for maximum utilization of viral genomes. Frameshifting requires a suitable heptanucleotide “slippery” sequence and an RNA pseudoknot.

Slippery sequences are heptanucleotide RNA sequences, which in most cases have a less efficient interaction with the anticodon loop of the corresponding tRNAs in the ribosome during protein translation. Slippery sequences generally follow the pattern X-XXY-YYZ where X, Y or Z represent a different purine or pyrimidine base, though deviations to this rule have been reported.

Pseudoknots are tertiary RNA structures that are formed by base pairing between a secondary loop structure and complement bases outside the loop. The result is an RNA structure composed of two loops and two stems. The two stems coaxially stack to form a quasi-continuous RNA helix, while the two loops span the major and minor grooves of the helix.

Though frameshifting is not completely understood, current studies suggest that a frameshift begins when a ribosome encounters a pseudoknot in the mRNA to be encoded. The process of the ribosome unraveling the pseudoknot leads to a pause in translation, which in turn causes slippage of the ribosome at a suitable slippery sequence. The ribosome can slip either forward (+change) or backward (- change) and a different sequence of codons is interpreted.

The frameshift sequence found in the envelope gene of HIV-1 was one of several frameshift sequences found in the HIV genome and serve as a major component of a proposed viral selenoprotein theory (Taylor et al. 1994,1997). The pseudoknot involved in frameshifting of the env-fs gene was found upstream of the proposed reading frame. The slippery sequence involved in env-fs frameshifting was found upstream of the pseudoknot and involves two lysine codons in the –1 and 0 frames at the initiation point of frameshifting (Taylor et al. 1994,1997). Computational analysis of the envelope gene of HIV using software such as mFOLD led to the discovery of this frameshift sequence (Taylor et al. 1994). The program was used to search for stable RNA hairpin structures in the HIV genome. Regions of no more than 20 base pairs from the base of the hairpin structure were searched for complement bases to the loop of the RNA hairpin structure (Taylor et al. 1994).

The slippery sequence of the env-fs frameshift sequence is 11 bases upstream of the stem loop structure as shown in Figure 1. The heptanucleotide sequence AAAAAGA was determined to potentiate p-site –1 frameshifting and translation of a hidden glutathione peroxidase gene.

A cell based frameshift assay using a dual protein reporter system was developed to understand the functionality of the env-fs sequence. GFP and Luciferase genes were cloned and arranged respectively into a eukaryotic expression vector. The putative frameshift sequence was inserted into the vector between the two reporter genes with luciferase in a –1 frame. The new frameshift assaying vector was expressed in different cell lines and 293-T produced the best results. Fluorescence of GFP and luminescence of luciferase are then detected and quantitated from cell lysates. The ratio of fluorescence to luminescence of a sample is compared to determine the extent and efficiency of frameshifting.

Comparing fluorescence of green fluorescent protein and luminescence of luciferase quantitates frameshifting capability of a frameshift sequence. Luciferase translation measures frameshifting while green fluorescent protein translation standardizes the assay. Therefore frameshifting is represented by luciferase luminescence as a factor of GFP fluorescence. Frameshifting efficiency is calculated by comparing the reporter ratios of a test vector with the reporter ratios of a 100% fusion protein expression vector (PGL).

Materials and Methods

Frameshift Assay Constructs

Frameshift construct map and the results of vector construction are shown in Figures 3-2 – 3-9.

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. 2 versions of the luciferase gene were cloned and they

differed in that 2 extra bases were added to the 5' end of one of the clones to embed the luciferase gene in a –1 frameshift conformation, while the 2nd 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 where PCL (-) denotes the –1 luciferase vector and PCL (+) is the 0 frame luciferase vector.

The GFP gene was digested from PEGFP-CI (Clontech) with Nhe I and Sal I restriction endonucleases. The digested GFP gene was ligated with T4 DNA ligase into, previously Nhe I and Sal I digested PCL (+) and PCL (-) vectors to form PGL (+) and PGL (-) vectors.

The protease frameshift sequence was cloned from PBH10 (HIV genome from NIH AIDS Reagent Program) by polymerase chain reaction, and in the process, Xho I and Sal I restriction sites were added to the 5' and 3' ends of the cloned envelope frameshift sequence. A stop codon (TAA) was engineered to directly follow the frameshift sequence in the zero frame of GFP to truncate unframeshifted fusion protein.

The PGL (+), PGL (-) vectors and the envelope frameshift sequence were digested with Xho I and Sal I and subsequently ligated to give the assaying vectors PGenvL (+) and PGenvL (-). An identical process was used to make PGgpL (+) and PGgpL (-) vectors which contain the known gag-pol frameshift sequence of HIV (Jacks et al. 1988). The sequence of all vectors was confirmed by dideoxy sequencing.

Transfection of Cell Lines

293-T cells were grown in Dubelco minimal essential medium with 2mM

L-glutamine, 1.5g/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 35mm diameter, 6 well plates to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 2μg of appropriate vector for 6 hours. Cells were transiently transfected with PGenvL (+) and PGenvL (-). PGL, PGgpL (+) and PGgpL (-) were used to transiently transfect cell lines as controls. The cells were incubated at 37º C with 5% CO_2 for 48 hours.

Cell-based Frameshift Assay

Transfected cell samples were collected and lysed by incubating with 1ml of glo-lysis buffer (Promega) per well for 5 minutes. 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 TD 20/20 luminometer, the luciferase assay involves adding 50μl of the cell lysates to 100μg 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-second delay. Background luminescence is normally 0.000 to 0.020 relative light units.

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

Fluorescence of GFP is measured using a Shimadzu RF-5301 or FLUOstar Optima Galaxy fluorescent plate reader. Cell lysate samples were excited at 488nm wavelength and the relative intensity of emission at 511nm 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) and Frameshifting Efficiency is the frameshift factor of a sequence as a percentage of the PGL frameshift factor.

Results and Discussion

Luciferase Activity of Test Constructs

The envelope frameshift sequence exhibited frameshifting capabilities when studied with the frameshift assay. There was significant luminescence emitted from cell lysates expressing the envelope frameshift sequence test vector PGenvL (-), over mock-transfected cells with no vector DNA. The mock transfected cells continuously showed a good baseline giving relative luminescence values of between 0.000 to 0.007 relative light units in comparison with 10,000 fold increase in luminescence of assay vector cell lysates. Luminescence was recorded at a decreased sensitivity in order to obtain readings in the luminometers detectable range. Lysates of cells transfected with the envelope frameshift assay vector PGenvL (-) also presented a 10,000-fold increase in relative luminescence in comparison with cells transfected with the control vector PGenvL (+). Cell lysates of cells transfected with the PGenvL (+) construct, which

contains the test envelope sequence and the luciferase gene in a 0 frame of the upstream GFP gene had relative luminescence values similar to the mock transfected cells. Cells expressing the control vectors PGgpL (-), containing the known gagpol frameshift sequence also showed a 10,000 fold increase in luciferase activity over mock transfected cells, and 100 fold increase over cells expressing the PGenvL (+) vector. The PGgpL (+) vector showed almost no luciferase activity. Luminescence quantitated from PGenvL (+) expressing vectors was almost identical to mock transfected cells and was practically inexistent. The envelope frameshift sequence displayed luciferase activity that was consistently 2 fold higher than the known gag-pol frameshift sequence. The positive control vector PGL which served as a 100% fusion protein expression construct, was consistently a million fold higher than mock transfected and about 100 fold higher than the envelope and gag-pol frameshift test vector when these vectors are expressed in 293-T cells. These results are shown in Figures 3-10.

GFP Fluorescence

GFP fluorescence was measured by excitation of cell lysates at 488 nm with a 5nm slit width and quantitation of fluorescence emission over a spectrum of emission wavelengths. The peak of relative GFP fluorescence for assayed cell lysates was at 511nm. 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 (-). GFP fluorescence quantitated from PGL

transfected cells was significantly less than fluorescence of PGgpL (+) and PGgpL (-), but about 4 times higher than GFP fluorescence of PGenvL (+) and PGenvL (-). The decrease in GFP fluorescence of PGL in comparison with the PGgpL (+) and PGgpL (-) vectors may have been due to translation efforts of the cell involved in expression of mostly the GFP-Luciferase fusion protein in comparison with the PGgpL (+) and PGgpL (-) which would experience more translation termination at the end of the GFP gene. The lower but significant gfp fluorescence from PGenvL (+) and PGenvL (-) vectors could be attributed to low transfection efficiency for that vector. These results are illustrated in Figures 3-11 – 3-16.

Frameshift Factor and Efficiency

The frameshift factor was calculated by determining the ratio of luciferase activity to GFP fluorescence. Frameshifting efficiency was calculated by comparing test vector ratio to PGL ratio respectively. The Fluostar Optima plate reader was used to obtain the best quantitation of GFP fluorescence, and the ratios of fluorescence between test constructs was identical to relative fluorescence comparisons between test constructs with the single cell fluorometer.

The envelope frameshift vector PGenvL (-) was 10,000 fold higher than the control vector PGenvL (+). The PGenvL (+) and PGgpL (+) vectors had virtually no frameshifting based on the frameshifting factor and efficiency. The envelope frameshift test vector, PGenvL (-) had about 2-fold higher frameshifting than the known gag-pol frameshift sequence (PGgpL (-)). The frameshifting efficiency of the PGenvL (-) ranges between 5% and 7% while the frameshifting efficiency of PGgpL (-) is between 2% and 3.5% in the assay. These results are illustrated in Figures 18.

Selenium and Sulfur containing Hypermodified Bases in tRNA-lys

The variety of hypermodified bases that exist in transfer RNA serve a variety of functions (Watts et al. 1978). Selenium and sulfur have been widely found in hypermodified bases of tRNA in different bases and at different substituent positions (Sprinzi et al. 1998). These elements have also been found to exist in compounds components of tRNAs in several different conformations (Ching et al. 1984, Sprinzi et al. 1998). Selenium and sulfur containing nucleotides have been found in tRNA as selenouridine or thiouridine (Sprinzi et at 1998). These elements along with oxygen, which are group VIA members of the periodic table, are interchangeable in hypermodified tRNA bases, and tend to exist at position 2 or 4 of uridine in most cases.

The strong preference for oxygen in nucleotides may not have always been the case, and may have happened as a result of the gradual predominance of oxygen in the atmosphere during evolution. Natural selection for highly functional selenium and sulfur hypermodified bases therefore may have kept the few sulfur and selenium containing tRNAs that we still see in nature today. Hypermodified selenium and sulfur containing tRNAs have been found in mice, rat, *E. coli*, *B. subtillus* and *S. typhimurium*. Identification of sulfur and selenium containing tRNAs in several different species indicate that these hypermodified bases may exist in all species but at low levels.

Studies suggest that having sulfur and selenium in place of oxygen in these tRNA-lys modified bases may have a regulatory role (Agris et al. 1997, Durant et al. 1999, Sundaram et al. 2000). NMR structural studies of the canonical U-turn of the anticodon loop of sulfur containing ecoli tRNA-lys indicate a regulatory advantage (Durant et al. 1999). The modified uridine containing nucleotide is the first base of the

anticodon loop of tRNA –lys. When this nucleotide contains an unmodified base, it has been shown that the anti-codon loop is unstructured (Durant et al.1999). The thiolated uridine of the tRNA-lys was shown to be involved in stabilizing the anti-codon loop leading to a more rigid loop that may induce a stronger association between the tRNA-lys anticodon loop and lysine codon during translation (Sundaram et al. 2000).

TRNA-lys is also the known primer for reverse transcription of HIV's RNA genome (Ratner et al. 1985, Marquet et al. 1995). The 18 bases of the 3' end of tRNA-lys are complementary to the primer-binding site of the HIV RNA genome (Ratner et al. 1985, Benas et al. 2000). There is an important interaction between the U rich region of the anticodon loop of tRNA-lys and the A rich region of the primer binding site of HIV (Isel et al. 1993, 1995). Similar to situations previously described, sulfur is found in the 2 position of uridine 34 of the tRNA, which is the first base of the anticodon loop. It has been shown that a thiolated uridine base at the first position of the anticodon loop is important for recognition of the A-rich region of the primer-binding site, and for efficient initiation of transcription of the HIV genome (Benas et al. 2000, Isel et al. 1993,1995). These attributes are also facilitated by the rigidity conferred to the anticodon/codon loop interaction by the thiolated uridine 34 and a thiolated A37 (the base that follows the anticodon loop) as is the case in the NMR structural studies of the similarly substituted *E. coli* tRNA previously described (Agris et al. 1997, Sundaram et al. 2000).

The AAA codon is known to serve as the beginning of several suitable heptanucleotide slippery sequences. The AAA codon is also the complement of the tRNA-lys anticodon loop. Previous studies have suggested a role for hypermodified uridine in the anticodon loop of tRNA-lys. Some suggestions indicate that a thiolated

uridine will lead to higher propensity for frameshifting in bacterial and mammalian translation systems (Agris et al. 1997). Other studies suggest a decrease in frameshifting due to a more rigid anticodon loop that interacts more readily with its corresponding codon (Sundaram et al. 2000). To give some insight to this issue the role selenium and sulfur play in frameshifting was studied. The frameshift assay developed is useful in determining a role that selenium or sulfur may play in frameshifting. To determine such a role, the frameshift assay was carried out as previously described under incubation with varying amounts of a selenium donor compound and a sulfur donor compound. Upon literature review, it was evident that the sulfur donor for incorporation of sulfur in several biological systems is L-cysteine. A candidate commonly used to introduce selenium in selenobiology experiments is sodium selenite. Hence the effects of L-cysteine and sodium selenite on the frameshift assay were quantitated to give some understanding of the role sulfur and selenium play in frameshifting and in HIV's mechanisms of survival as a whole.

Materials and Methods

Frameshift Assay Constructs

The frameshift assay vectors were constructed as described in earlier sections. All test frameshift sequences were cloned and placed upstream of a cloned firefly luciferase gene (Promega) and downstream of a green fluorescent protein gene (Clontech). These were arranged accordingly in a PCI vector (Promega) and expression was assessed.

Cell Culture and Transfection of Cell Lines

293-T cells were grown in dubelco minimal essential medium and deviated in a few components depending on the element to be tested. The cells were grown in 24 well plates. For L-cysteine studies 293-T cells were grown in DMEM lacking L-cysteine, L-methionine, L-glutamine, sodium bicarbonate and sodium pyruvate. The media in L-cysteine studied sample wells was supplemented with 10% Defined fetal bovine serum (Hyclone), which contains approximately 1.09μM L-cysteine, 2mM L-glutamine, 1.5g/L sodium bicarbonate and 1.0 mM sodium pyruvate and 1mM methionine. The methionine was decreased to limit possible interference of sulfur donation by L-methionine, though this had to be balanced with regular methionine translational requirements. The sample wells were supplemented accordingly with L-cysteine ranging in concentration from 0 to 3mM. Cells were starved of cysteine for 24 hours and subsequently incubated for 48 hours at 37º C with 5% CO_2 with media changes and L-cysteine supplementation every 24 hours.

For sodium selenite studies 293-T cells were grown in DMEM lacking a selenium source. The media was supplemented with Defined fetal bovine serum (Hyclone), which supplements the media with 0.05uM sodium selenite. The sample wells were supplemented accordingly with sodium selenite ranging in concentration from 0 to 3uM. Cells are starved for 24 hours and subsequently incubated for 48 hours at 37º C with 5% CO_2 with media changes and sodium selenite supplementation every 24 hours.

Cell lines were grown to 95% confluency. Lipofectamine (invitrogen) was used to transfect the cell lines in each well with 0.5μg of PGenvL (-) DNA in Opti-MEM (Invitrogen) reduced serum media. Following a 24-hour transfection period, the media

of transfected cells is replaced with DMEM and supplemented with L-cysteine or sodium selenite and the cells are incubated at 37º C with 5% CO_2 for the 48-hour post transfection period.

Reporter Gene Assay

The transfected cells were lysed by incubating the cells with 200μl of glo-lysis buffer (Promega) per well for 5 minutes. Crude extracts are centrifuged and supernatant is recovered. The supernatant is assayed for reporter gene activity.

Luciferase activity and Green fluorescent protein fluorescence is quantitated in a Fluostar Optima Galaxy plate reader. 10μl of the cell lysate is added to 40μl of Promega luciferase assay buffer, which contains luciferin as the active luminescent agent. Cell lysates are pipetted into a white luminescent reading plate. The luciferase assay buffer is dispensed by the pump into the 96 well plate containing cell lysates, and luminescence is recorded over a 12 second interval. Background luminescence is normally 3 to 4 fold lower than test samples. Fluorescence of green fluorescent protein is measured using a Fluostar Optima Galaxy, or Polarstar Optima plate reader. Cell lysate samples are excited at 488nm wavelength and relative intensity of emission at approximately 507nm is recorded.

Comparing fluorescence of green fluorescent protein and luminescence of luciferase quantitates Frameshifting. Luciferase translation measures frameshifting while green fluorescent protein translation standardizes the assay. Therefore frameshifting is represented by luciferase luminescence as a factor of green fluorescent protein fluorescence.

Results and Discussion

Effects of Sulfur on Frameshifting

To determine the effects of sulfur on frameshifting, the media of PGenvL (-) transfected 293-T cell sample wells was supplemented with cysteine as previously mentioned. 293-T cells were chosen based on the high transfection efficiency of the cell line and literature recommendation that the kidney has a significant number of sulfurtransferases, and a suitable number of available sulfur transfer sites available in tRNA relative to the number of sulfurtransferases (Harris et al. 1978). The recommended concentration of cysteine in cell culture media is approximately 0.2mM L-cysteine. Increasing the concentration of L-cysteine caused the ratio of frameshifting to decrease as shown in Figure 3-21. This decrease cannot be attributed to effects of cysteine on translation, because several cases exist where the rate of luciferase activity and GFP fluorescence was actually highest for the samples supplemented with the highest amounts of cysteine, but their ratios of frameshifting still proved to be lower than for samples supplemented with lower concentrations of cysteine.

Effects of Selenium on Frameshifting

To determine the effects of selenium on frameshifting, the media of PGenvL (-) transfected 293-T cells were supplemented with sodium selenite as previously mentioned. Sodium selenite is toxic to most cell lines in high doses, but is necessary for efficient cell culture. The recommended concentration of sodium selenite in cell culture media is approximately 0.05 μM. The selenium component of the DMEM media is from the fetal bovine serum added to the media, which makes up 10% of the total media volume. By increasing the concentration of sodium selenite, the ratio of

frameshifting was found to decrease as shown in Figure 3-22. This decrease could not be attributed to effects of sodium selenite on translation, because several cases exist where the rate of luciferase activity and GFP fluorescence were actually highest for the samples supplemented with the highest amounts of sodium selenite, but their ratios of frameshifting still proved to be lower than for samples supplemented with lower concentrations of cysteine.

Hungry Arginine Codon and Frameshifting

The hungry codon phenomenon is another means by which organisms may regulate translation. Hungry codons describe a situation where certain codons present a pause to the ribosome during translation. These codons generally have acceptor tRNAs that are rare in normal physiology. This process is aided by a deficient concentration of the associated amino acid.

A hypothetical hungry arginine codon was discovered in the envelope gene of HIV-1 through bioinformatic search studies conducted on the HIV genome (Taylor et al. 1996). As mentioned earlier, the envelope frameshift sequence was found to be involved in the expression of a hidden glutathione peroxidase gene in the HIV genome (Zhao et al. 2001). The hungry arginine codon was found in the slippery sequence of the envelope frameshift sequence and is represented by the nucleotide sequence base notation AAAAAGA. The adenine quadruplet was determined to be involved in slippery frameshift sequence hungry codon processes (Gallant et al. 1993), and in the case of the envelope frameshift sequence, is followed by an AGA arginine codon.

The severe acute respiratory syndrome coronavirus (SARS) genome was recently analyzed using the same techniques, which were used to analyze the HIV

genome. An embedded gene was also discovered in the SARS genome with similar structural elements. This gene was hidden in the genome of SARS in a –1 frame and would require a frameshift event to be encoded. The slippery sequence of the gene frameshift sequence found in the SARS gene was identical to the sequence previously discovered preceding the hidden GPx gene in the envelope region of the HIV genome.

Physiological arginine concentrations are important to different biological functions and mechanisms. Arginine deficiency has been shown to cause a decrease in nitric oxide, which in turn causes an increase in hydroxy radical oxidative stress (Aoyagi et al. 2003). Nitric oxide is a known mediator of the immune system and, in addition to its precursor, arginine, are found to be associated with immunodeficiency. An early study of viral replication showed that arginine deficient media is capable of causing a marked increase in viral replication (Henle et al. 1968). Therefore, the role of arginine in virology is one that should not be overlooked.

Materials and Methods

Frameshift Assay Constructs

The frameshift assay vectors were constructed as described in previous sections. All test frameshift sequences were cloned and placed upstream of a cloned firefly luciferase gene (Promega) and downstream of a green fluorescent protein gene (Clontech). These were arranged accordingly in a PCI vector (Promega).

Cell Culture and Transfection of Cell Lines

293-T cells were grown in dubelco minimal essential medium and deviated in a few components depending on the compound to be studied. The cells were grown in 24 well tissue culture plates. For arginine deficiency studies 293-T cells were grown in

DMEM lacking L-arginine, and L-lysine, and supplemented appropriately as the experiments dictated. The media in L-arginine sample wells was supplemented with 10% volume FBS and 0.8mM L-lysine. The sample wells were supplemented accordingly with either 0.1mM or 1mM of L-arginine. Cells were incubated for 48 hours at 37º C with 5% CO_2 with media changes and L-arginine supplementation every 24 hours. Cell lines were grown to 95% confluency. Lipofectamine (invitrogen) was used to transfect the cell lines in each well with 0.5μg of PGenvL (-) in Opti-MEM (Invitrogen) reduced serum media. Following a 24-hour transfection period, the media of transfected cells was supplemented with L-arginine and the cells are incubated at 37º C with 5% CO_2 for the 48-hour post transfection.

Reporter Gene Assay

Transfected cells were lysed by incubating with 200μl of glo-lysis buffer (Promega) per well for 5 minutes. Crude extracts are centrifuged and supernatant is recovered. The supernatant is assayed for reporter gene activity.

Luciferase activity and Green fluorescent protein fluorescence is quantitated in a Fluostar Optima Galaxy plate reader. 10μl of the cell lysate is added to 40μl of Promega luciferase assay buffer, which contains luciferin as the active luminescent agent. Cell lysates are pipetted into a white luminescent reading plate. The luciferase assay buffer is dispensed by the pump into the 96 well plate containing cell lysates and luminescence is recorded over a 12 second interval. Background luminescence is normally 3 to 4 fold lower than test samples. Fluorescence of green fluorescent protein is measured using a Fluostar Optima Galaxy, or Polarstar Optima plate reader. Cell

lysate samples are excited at 488nm wavelength and relative intensity of emission at approximately 510nm is recorded.

Comparing fluorescence of green fluorescent protein and luminescence of luciferase quantitates Frameshifting. Luciferase translation measures frameshifting while green fluorescent protein translation standardizes the assay. Therefore frameshifting is represented by luciferase luminescence as a factor of green fluorescent protein fluorescence.

Results and Discussion

The effect of L-arginine deficient media on frameshifting facilitated by the envelope frameshift sequence was assessed. Wells supplemented with the normal 1mM media concentration of L-arginine are compared with wells supplemented with 0.1mM L-arginine. It was found that with the PGenvL (-) construct there was a significant decrease in frameshifting with an increase in L-arginine concentration supplementation ($p<0.005$) as shown in Figure 3-23. In comparison the two concentrations of L-arginine had no significant effects on frameshifting on the PGgpL (-) frameshift sequence as shown in Figure 25.

References

Aoyagi K et al. (2003) “Inhibition of arginine synthesis by urea: A mechanism for arginine deficiency in renal failure which leads to increased hydroxyl radical generation” Molecular and Cellular Biochemistry 244 (1): 11-15

Agris, P. F., R. Guenther, et al. (1997). "Unconventional structure of tRNA(Lys)SUU anticodon explains tRNA's role in bacterial and mammalian ribosomal frameshifting and primer selection by HIV-1." Rna 3(4): 420-8.

Benas, P., G. Bec, et al. (2000). "The crystal structure of HIV reverse-transcription primer tRNA(Lys,3) shows a canonical anticodon loop." Rna 6(10): 1347-55.

Berger, E. A., P. M. Murphy, et al. (1999). "Chemokine receptors as HIV-1 coreceptors: Roles in viral entry, tropism, and disease." Annual Review of Immunology 17: 657-700.

Ching, W., (1984) “Occurrence of selenium-containing tRNAs in mouse leukemia cells” Proc. Natl. Acad. Sci. 81:3010-3013

Durant P.C., Davis D.R., et al. (1999).“Stabilization of the anticodon stem-loop of tRNA(LyS,3) by an A(+)-C base-pair and by pseudouridine” Journal of Molecular Biology 285 (1): 115-131

Gallant J., Lindsley D., et al. (1993) “Ribosome frameshifting at hungry codons-sequence rules, directional specificity, and possible relationship to mobile element behavior” Biochemical Society Transactions 21 (4): 817-821

Harris, C.,(1978). “Mammalian tRNA sulfurtransferases: properties of the enzyme in rat liver” Nucl. Acids Research 5(2)599-613

Henle, W., Henle., C (1968) Effect of arginine deficient media on herpes type virus associated with cultured burkitt tumor cells” Journal of Virology 2 (3): 182

Isel, C., R. Marquet, et al. (1993). "Modified nucleotides of tRNA(3Lys) modulate primer/template loop-loop interaction in the initiation complex of HIV-1 reverse transcription." J Biol Chem 268(34): 25269-72.

Isel, C., C. Ehresmann, et al. (1995). "Initiation of reverse transcription of HIV-1: secondary structure of the HIV-1 RNA/tRNA(3Lys) (template/primer)." J Mol Biol 247(2): 236-50.

Landau, N. R., M. Warton, et al. (1988). "The envelope glycoprotein of the human immunodeficiency virus binds to the immunoglobulin-like domain of CD4." Nature 334(6178): 159-62.

Marquet, R., C. Isel, et al. (1995). "tRNAs as primer of reverse transcriptases." Biochimie 77(1-2): 113-24.

Ratner, L., W. Haseltine, et al. (1985). "Complete nucleotide sequence of the AIDS virus, HTLV-III." Nature 313(6000): 277-84.

Sprinzi, M., Horn, C. et al. (1998). “Compilation of tRNA sequences and sequences of tRNA genes.” Nucleic Acids Research, 26(1): 148-153

Sundaram M., Durant P.C., et al. (2000) “Hypermodified nucleosides in the anticodon of tRNA(Lys) stabilize a canonical U-turn structure” Biochemistry 39 (50): 15652-15652

Taylor, E. W., C. S. Ramanathan, et al. (1994). "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(17): 2637-54.

Taylor, E.W., Ramanathan, C.S., et al. (1996) “A general method for predicting potential new genes in nucleic acid sequences: application to the human immunodeficiency virus.“ In: Computational Medicine, Public Health and Biotechnology, M. Witten, Ed. World Scientific, Singapore. Ser. Math. Biol. Med. 5: 285-309.

Taylor, E. W., A. Bhat, et al. (1997). "HIV-1 encodes a sequence overlapping env gp41 with highly significant similarity to selenium-dependent glutathione peroxidases." Journal of Acquired Immune Deficiency Syndromes and Human Retrovirology 15(5): 393-394.

Taylor E.W., Cox A.G., Zhao L.J., et al. (2000) “Nutrition, HIV, and drug abuse: The molecular basis of a unique role for selenium” Journal of Acquired Immune Deficiency Syndromes 25: S53-S61 Suppl. 1

Watts, M. T. and I. Tinoco, Jr. (1978). "Role of hypermodified bases in transfer RNA. Solution properties of dinucleoside monophosphates." Biochemistry 17(12): 2455-63.

Zhang W., Ramanathan C.S., et al. (1999). “Selenium-dependent glutathione peroxidase modules encoded by RNA viruses”Biological Trace Element Research (2): 97-116

Turner, B. G. and M. F. Summers (1999). "Structural biology of HIV." J Mol Biol 285(1): 1-32.

Zhao LJ, Cox AG, Ruzicka JA, et al.(2000) ”Molecular modeling and in vitro activity of an HIV-1-encoded glutathione peroxidase” Proc of the Natl Acad of Sci of the United states of America 97 (12): 6356-6361

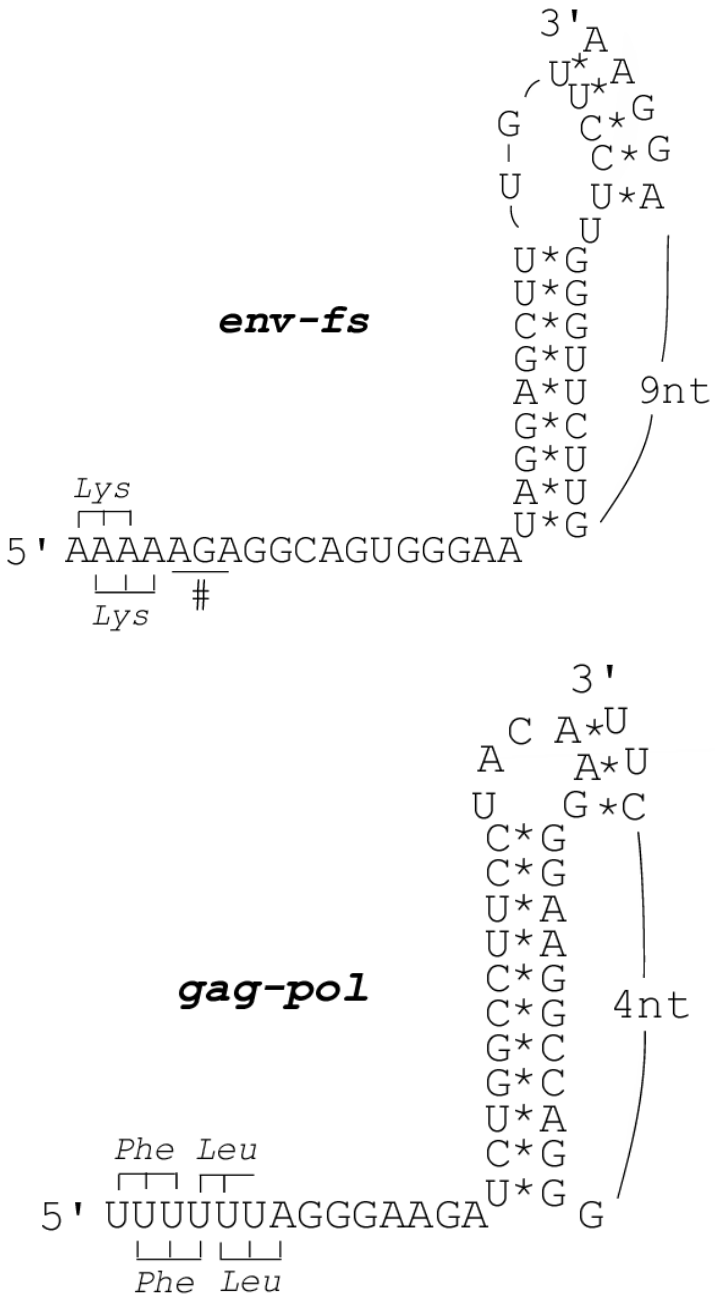


Figure 3-1

Envelope and gag-pol frameshift sequences. 2D schematic of the slippery sequence and pseudoknot frameshift sequences in the envelope (proposed) and gag-pol (Jacks et al. 1988) of HIV-1. Proposed env-fs slippery sequence contains the adenine quadruplet followed by the arginine codon.

Table 3-1

Primers designed for construction of frameshift assay vectors. These primers were designed for PCR cloning of the luciferase gene as well as frameshift sequence. They were also used to incorporate restriction sites, and stop codons at the end of frameshift sequences.

Oligos designed and synthesized for PCR

All primers are presented as synthesized in the 5' → 3' direction

Pos-Luc-Sal I
TTG GCC GTC GAC ATG GAA GAC GCC AAA AAC ATA AAG
Sal I

Neg-Luc-Sal I
TTG GCC GTC GAC AAA TGG AAG ACG CCA AAA ACA TAA AG
Sal I

Luc-Not I
AAC CGG GCG GCC GCT TTA CAA TTT GGA CTT TCC GCC CT
Not I

Env-fs-Fw
TTG GCC TCT CGA GAG AGA GAA AAA AGA GCA GTG GGA
Xho I

Env-fs-Rv
TTC CGG GTC GAC TTA GAT AGT GCT TCC TGC TGC
Sal I

GP-fs-Fw
TTG GCC TCT CGA GAG GCT AAT TTT TTA GGG AAG ATC
Xho I

GP-fs-Fw
TTC CGG GTC GAC TTA GCT CTG AAG AAA ATT CCC AGG
Sal I

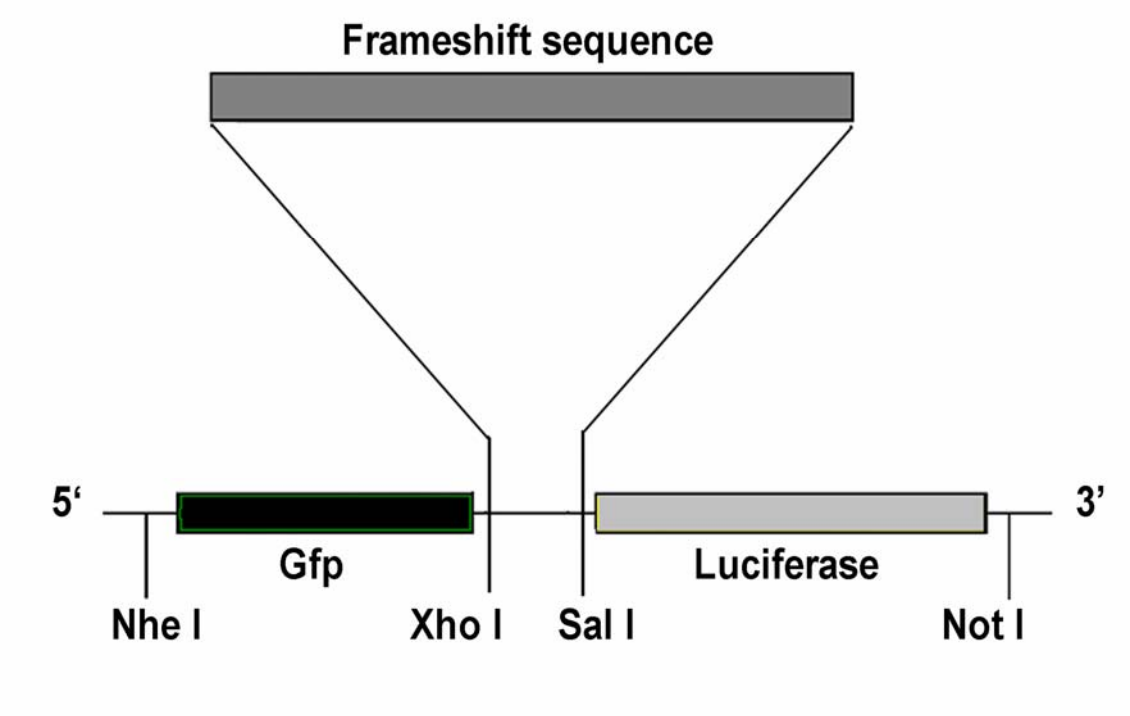


Figure 3-2

A map of the frameshift assay construct used in the frameshift assay. The map describes the organization of gene components for all (+) and (-) assaying vectors**.**

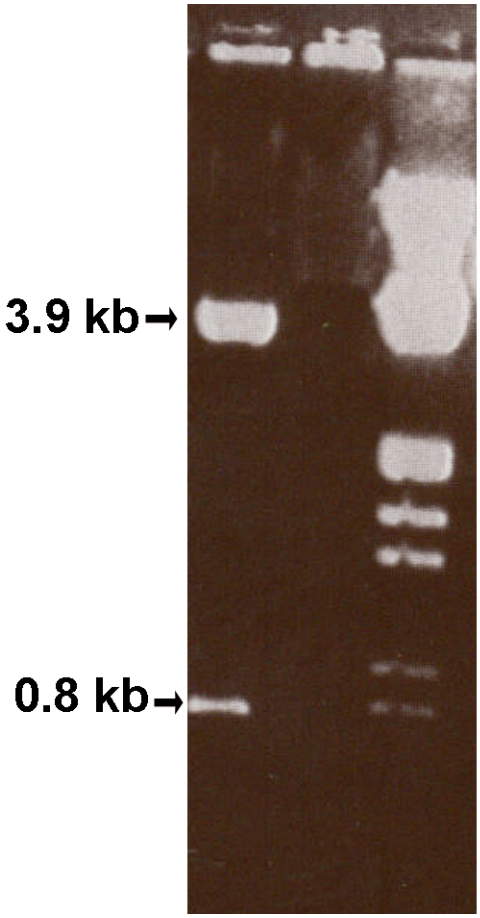


Figure 3-3

Restriction digest of PEGFP-C1. Agarose gel showing PEGFP-C1 digested with NheI and Bgl II in 1st lane and DNA ladder in 3rd lane. Bands were detected at 3.8 KB and 0.8 KB by ethidium bromide and UV detection and correspond to the GFP gene and remainder of the PEGFP-C1 vector respectively.

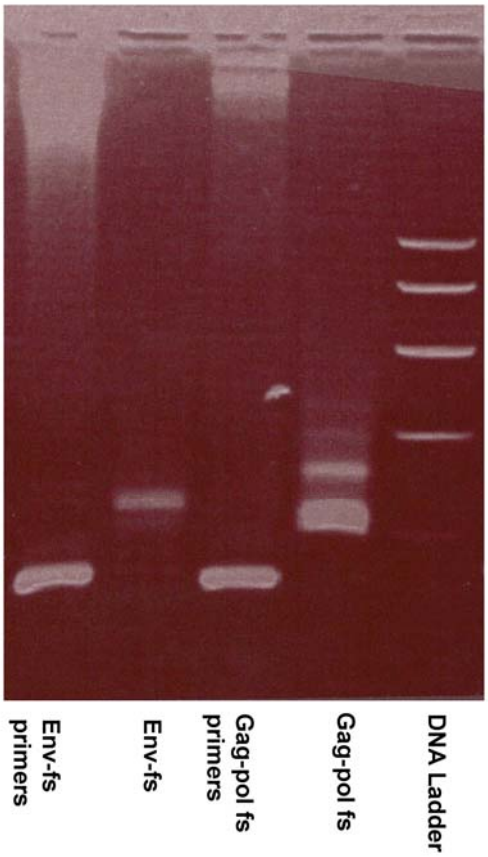


Figure 3-4

PCR amplification of gag-pol and envelope frameshift sequences. Agarose gel showing results of PCR amplification of gag-pol-fs and envelope-fs. Lanes 1 and 3 show results of PCR experiments using 5’ and 3’ primers with no template.

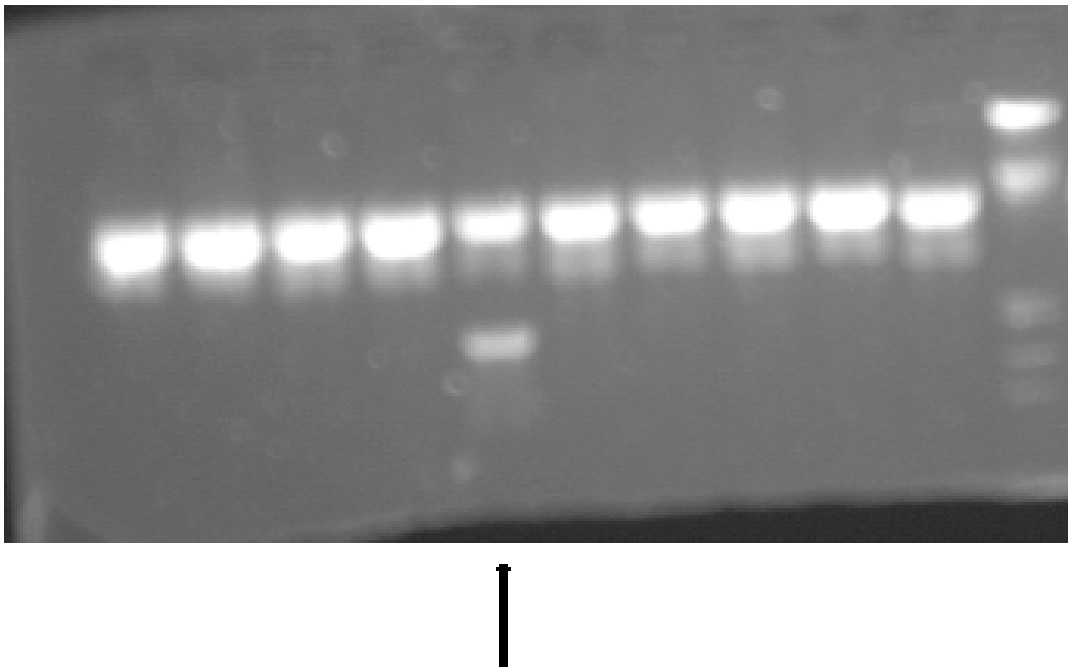


Figure 3-5

PCR identification of positive clones. Agarose gel showing PCR identification of postive clones with primers Env-fs-Fw, GP-fs-Fw and Luc-Not I. PGenvL (-) clones are in lanes 1 - 5 and PGgpL (-) clones lanes 6 – 10. A positive PGenvL (-) clone was detected in Lane 5.

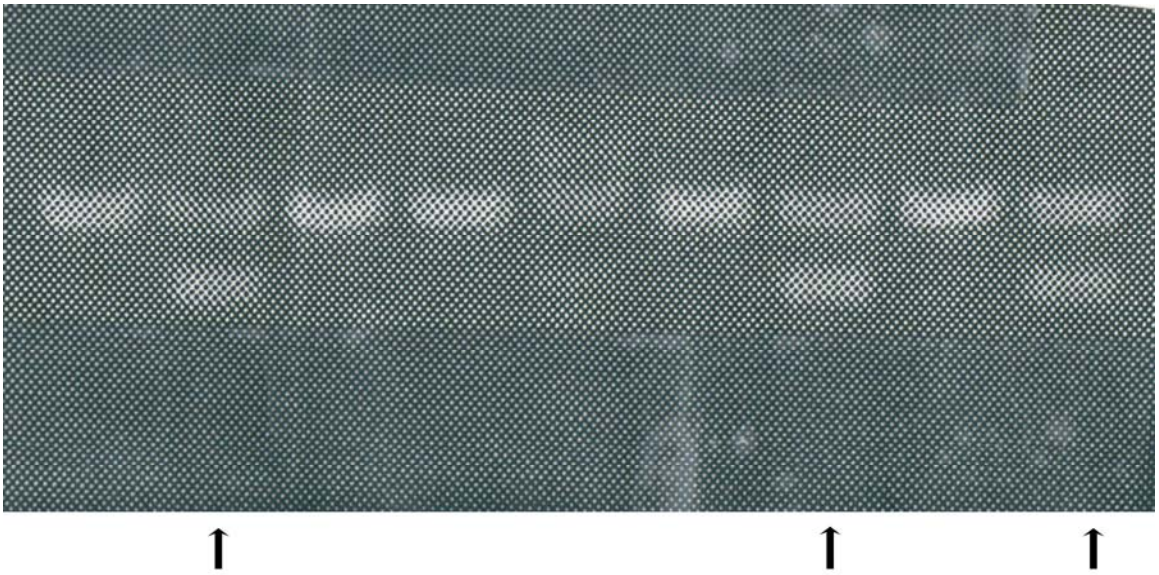


Figure 3-6

PCR identification of positive clones. Agarose gel showing PCR identification of positive clones with primers GP-fs-Fw, GP-fs-Rv and Luc-Not I. PGgpL (-) clones are in lanes 1 – 4 and 6 - 8. Lane 5 electrophoresis of DNA marker and lane 9 is positive control with T7-Luc vector (Promega) with Pos-Luc-Sal I and Neg-Luc-Sal I primers. 2 positive PGgpL (-) clones found in (lanes 2,7).

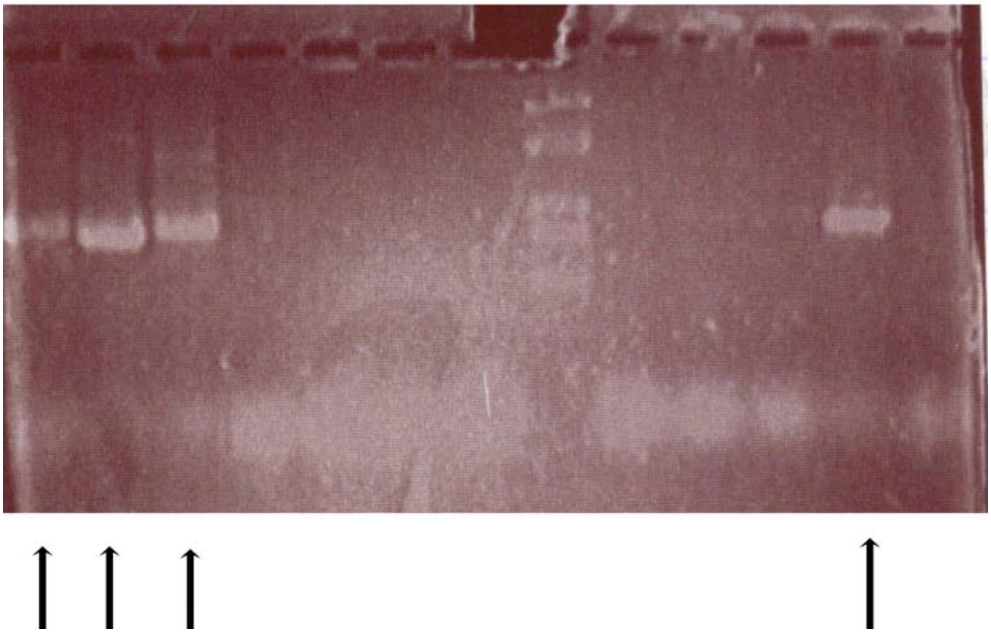


Figure 3-7

PCR identification of positive clones. Agarose gel showing PCR identification of positive clones with primers GP-fs-Fw, and Luc-Not I. PGgpL (+) clones are in lanes 1 – 7 and 10 - 11. Lane 12 is PCR of T7-Luc with Pos-Luc-Sal I and Neg-Luc-Sal I as positive control. DNA marker in lane 8. 3 positive PGgpL (+) clones found in lanes 1,2 and 3.

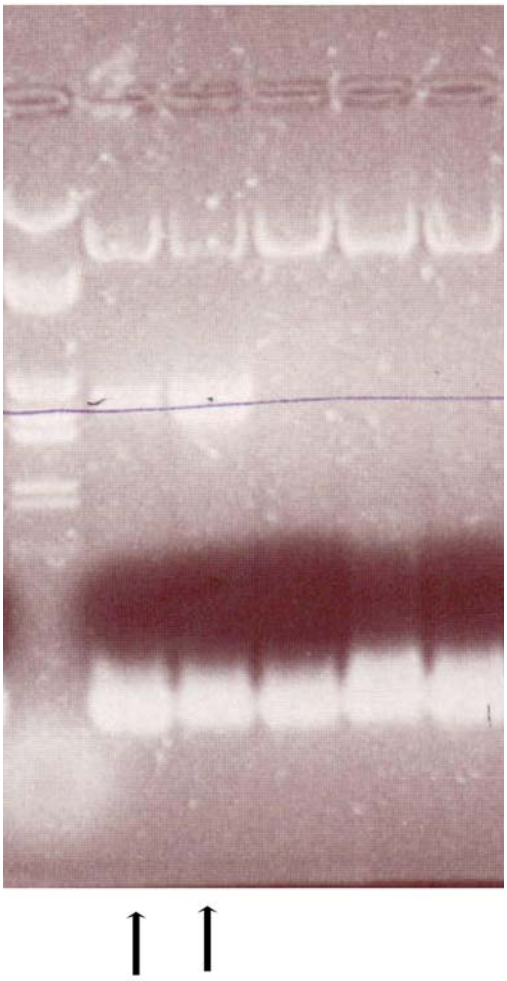


Figure 3-8

PCR identification of positive clones. Agarose gel showing PCR identification of positive clones with primers Env-fs-Fw, and Luc-Not I. PGenvL (+) clones are in lanes 2 – 6. Lanes 2 and 3 show positive clones representing amplified luciferase gene + frameshift sequence.

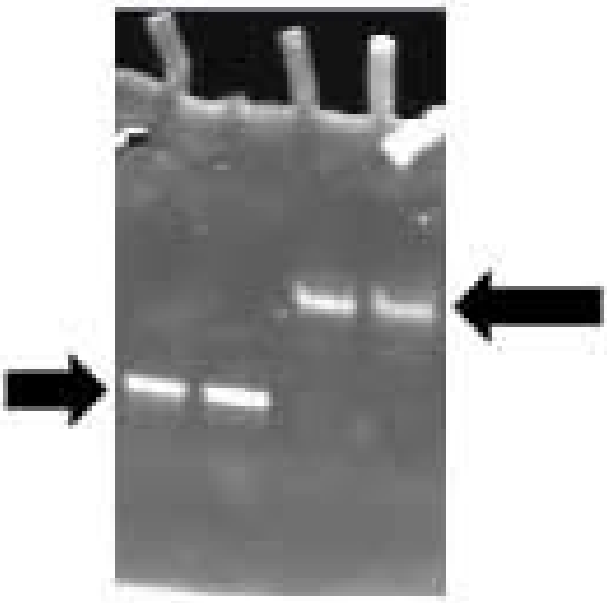
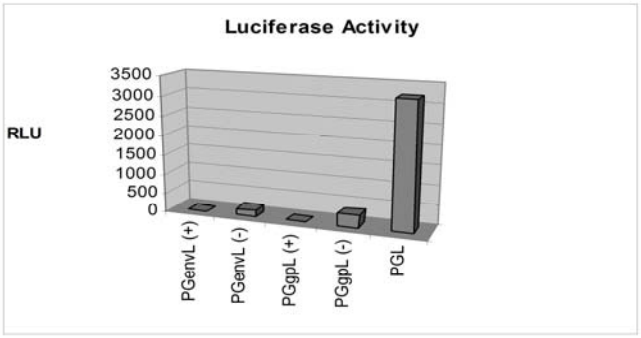
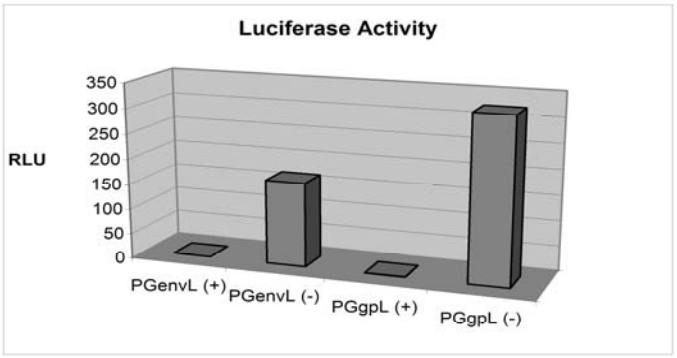


Figure 3-9

Fluorescent imaging of assay vector cell lysates. Native acrylamide gel of lysates in lanes 1 and 2 show GFP alone, lane 3 and 4 show a GFP-luciferase fusion protein. Image was captured with a 520nm filter.





Construct	Luciferase activity in Relative Light Units Average +/- Standard Deviation
PGenvL (+)	0.011 +/- 0.004
PGenvL (-)	166 +/- 20.4
PGgpL (+)	0.60 +/- 0.16
PGgpL (-)	324 +/- 65
PGL	3230 +/- 262

Figure 3-10

Luminescence from cell lysates expressing frameshift assay constructs.

Luminescence of samples was measured on a TD 20/20 luminometer. Cell lysates of PGenvL (+) and PGgpL (+) expressing cells show virtually no luminescence. ($P<0.001$ for PGenvL (+) and PGenvL (-), $P<0.0001$ for PGenvL (+) and PGL, $P<0.001$ for PGgpL (+) and PGgpL (-), $P<0.0001$ for PGgpL (+) and PGL).

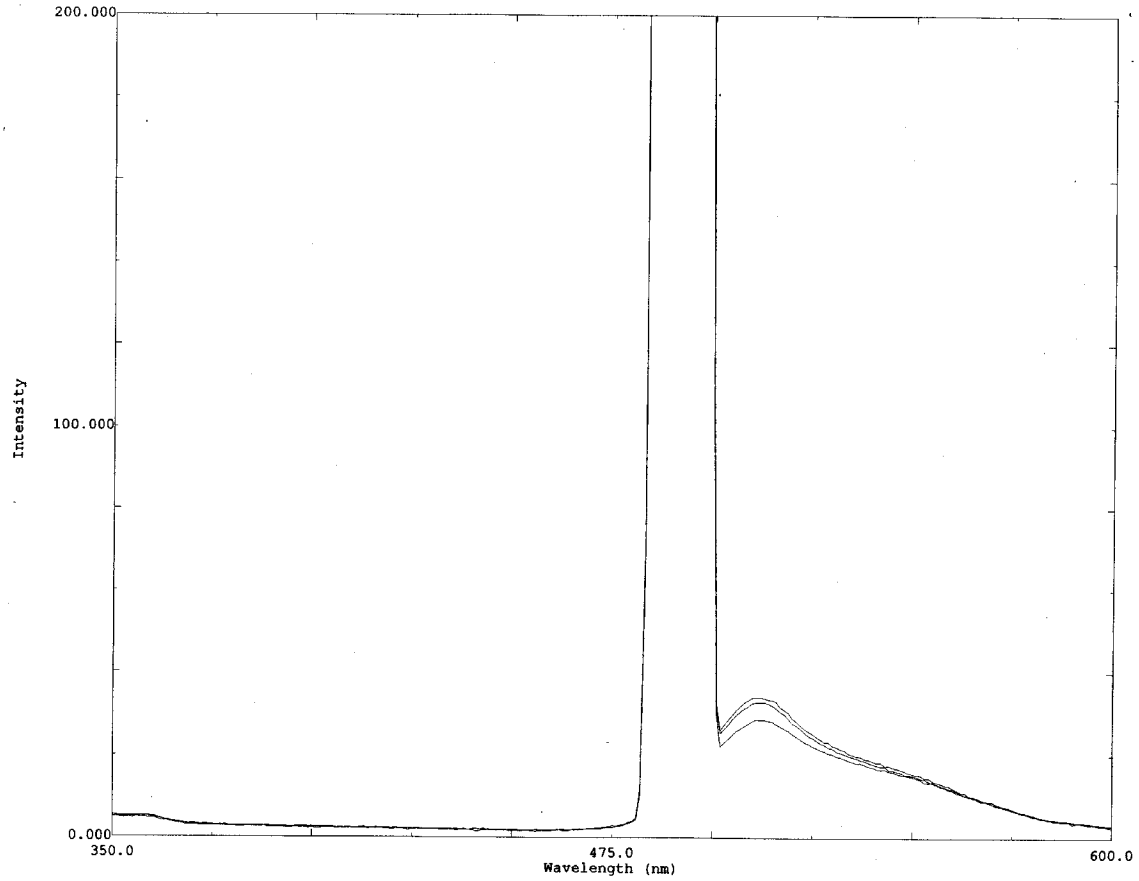


Figure 3-11

GFP fluorescence emission spectra of PGenvL (+) lysates. Fluorescence spectra from three PGenvL (+) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

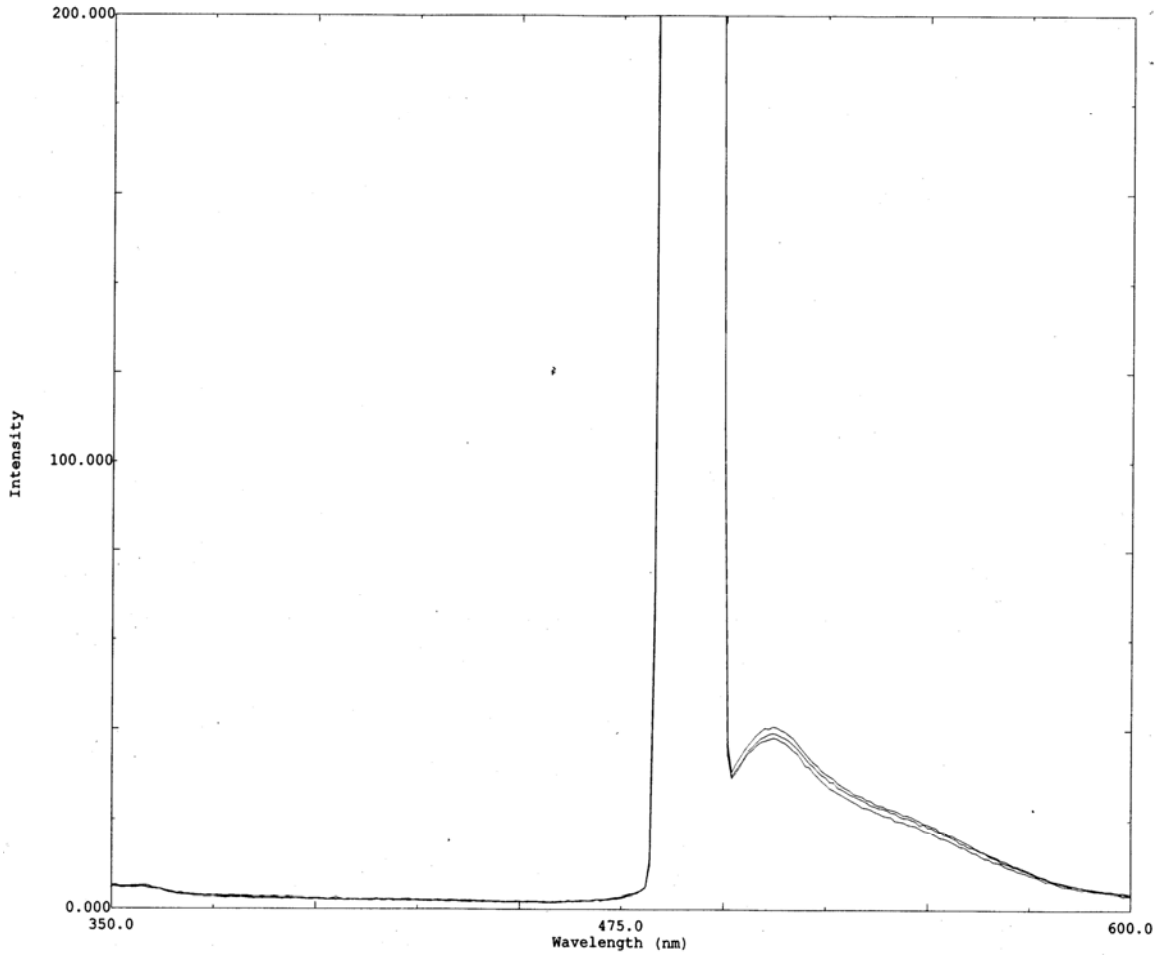


Figure 3-12

GFP fluorescence emission spectra of PGenvL (-) lysates. Fluorescence spectra from three PGenvL (-) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

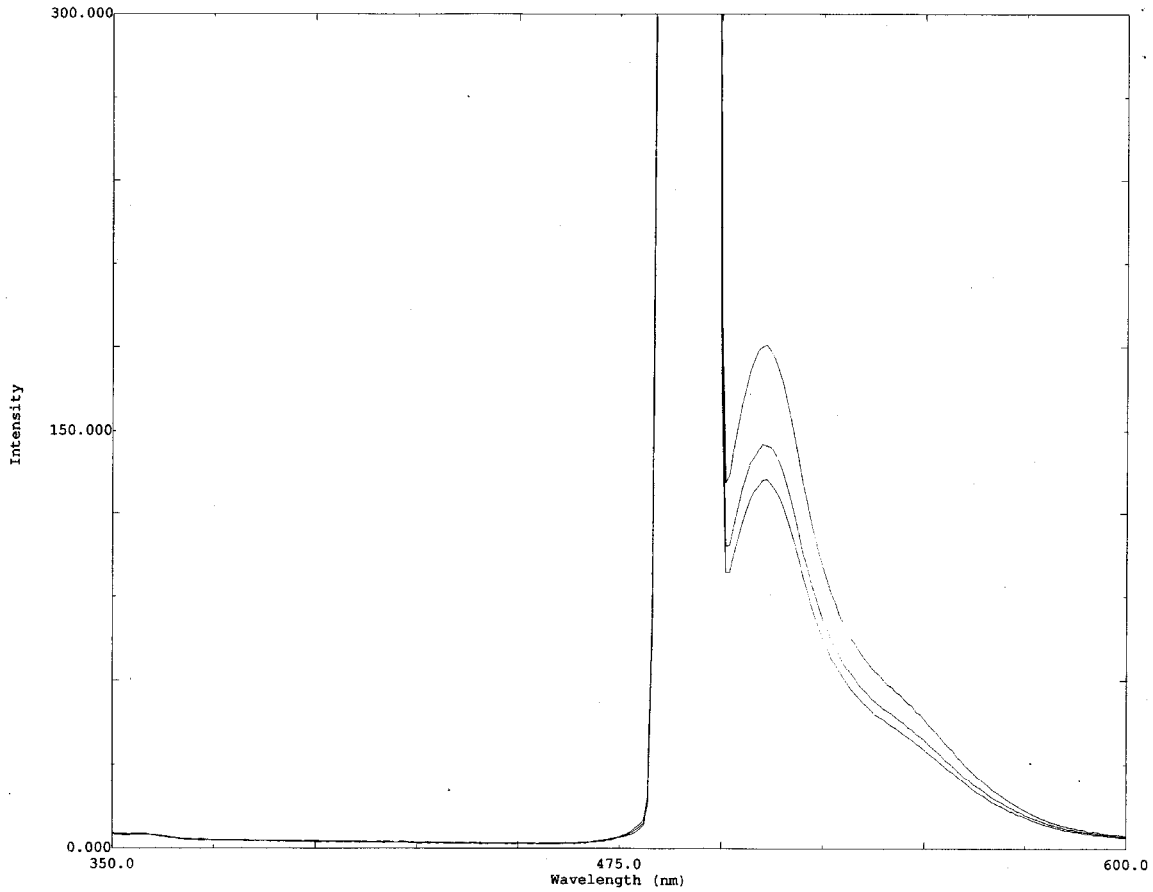


Figure 3-13

GFP fluorescence emission spectra of PGgpL (+) lysates. Fluorescence spectra from three PGgpL (+) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

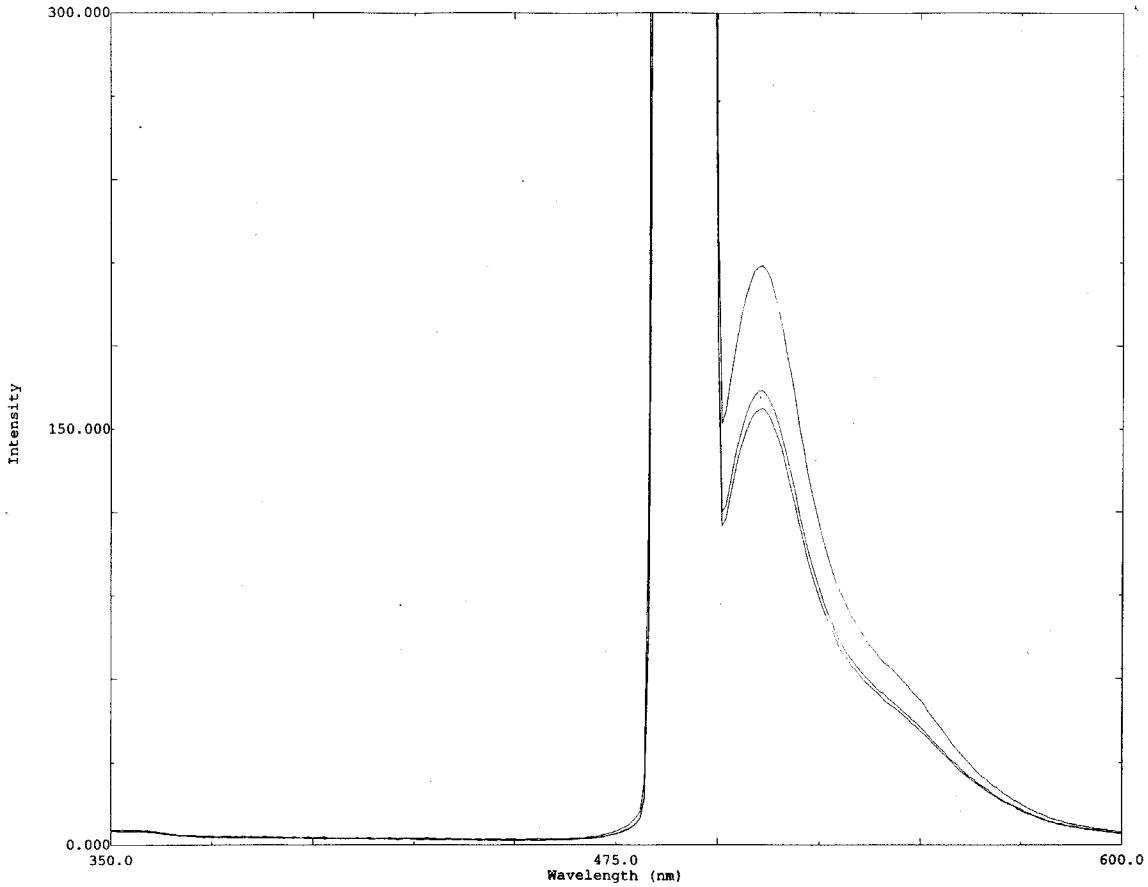


Figure 3-14

GFP fluorescence emission spectra of PGgpL (-) lysates. Fluorescence spectra from three PGgpL (-) expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.

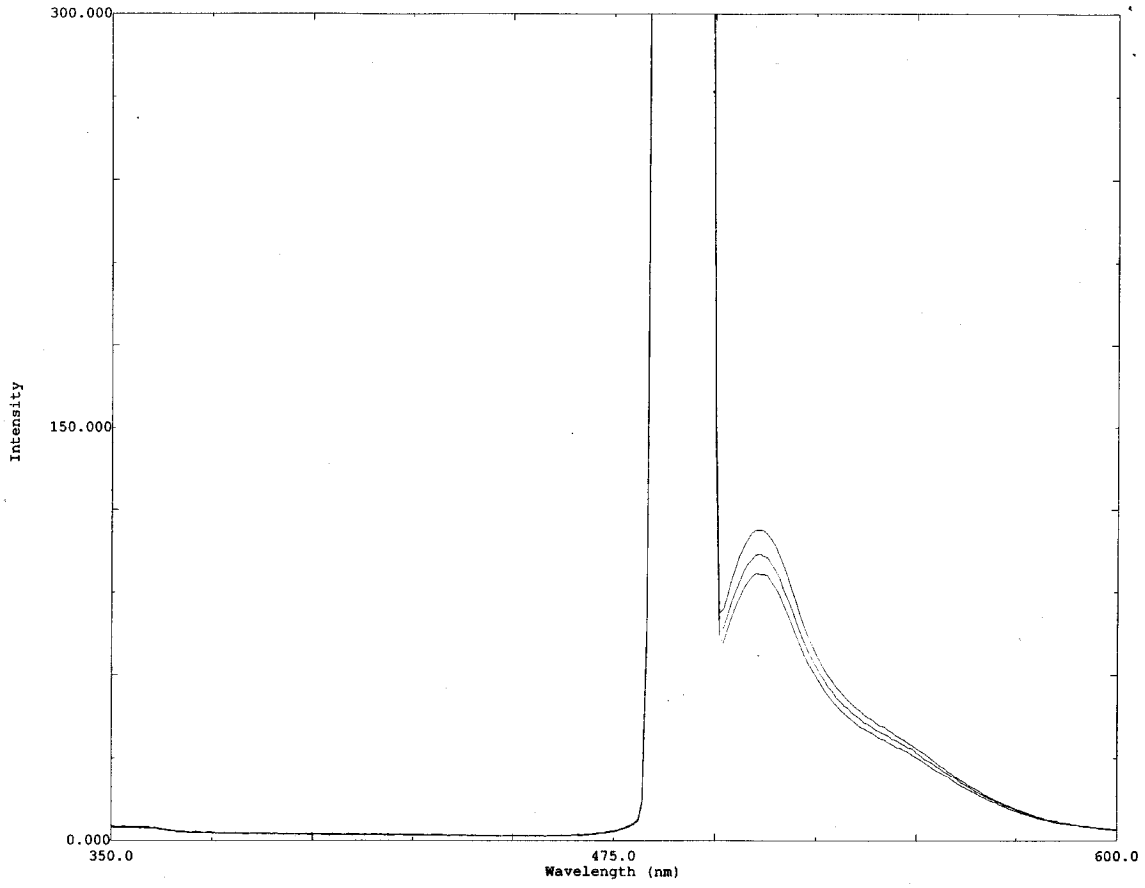
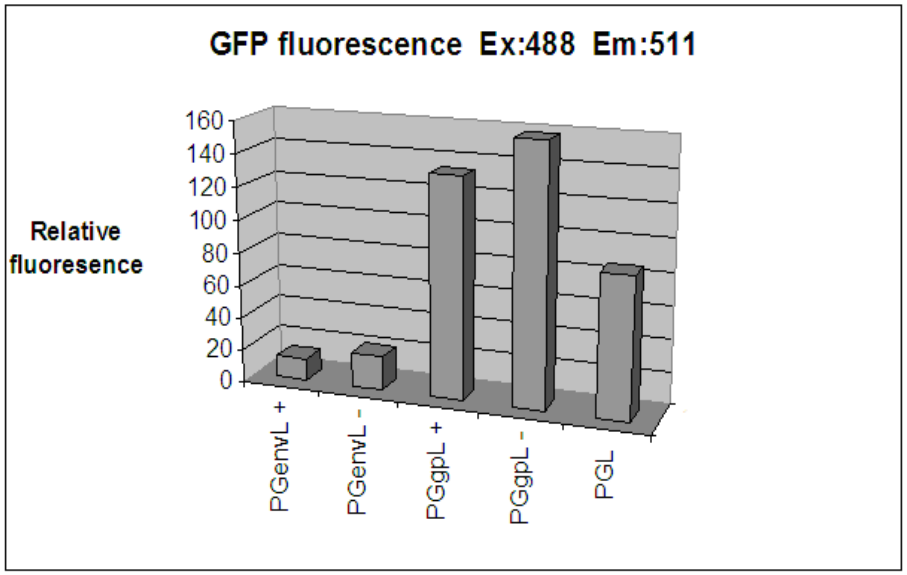


Figure 3-15

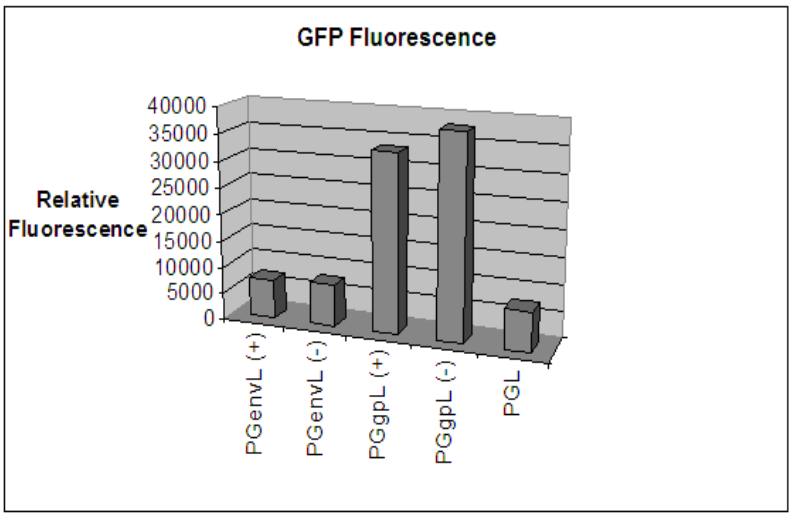
GFP fluorescence emission spectra of PGL lysates. Fluorescence spectra from three PGL expressed, cell sample wells, excited at 488nm and measured on the Shimadzu RF-5301 fluorometer. Emission peak was at 511nm.



Construct	Green Fluorescent Protein Relative Fluorescence Excitation: 488 Emission: 511 Average +/- Standard Deviation
PGenvL (+)	14 +/- 2.8
PGenvL (-)	21+/- 1.3
PGgpL (+)	134 +/- 25
PGgpL (-)	159 +/- 28
PGL	86 +/- 7.9

Figure 3-16

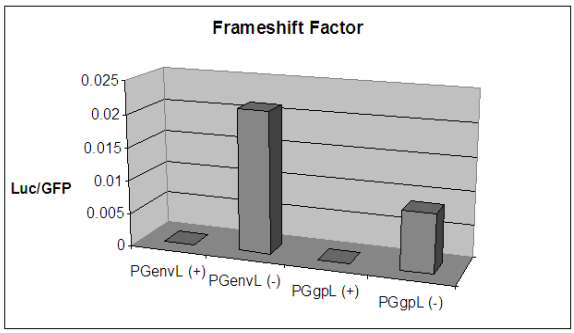
Relative fluorescence of frameshift construct cell lysates. GFP fluorescence was similar for constructs with similar frameshift sequences.

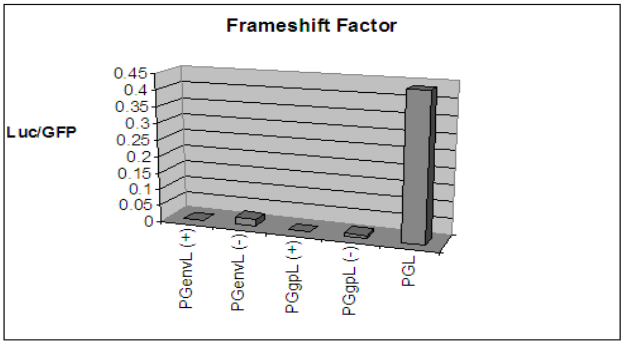


Vector	GFP Fluorescence Average +/- Standard Deviation
PGenvL (+)	7370 +/- 449
PGenvL (-)	7800 +/- 980
PGgpL (+)	34000 +/- 5000
PGgpL (-)	38000 +/- 12000
PGL	7400 +/- 720

Figure 3-17

GFP fluorescence. GFP fluorescence was measured on a Fluostar Optima plate reader for enhanced quantitation of GFP fluorescence and calculation of frameshift efficiency.





Construct	Frameshift Factor	Frameshift Efficiency
PGenvL (+)	1.5×10^{-6}	0.0003%
PGenvL (-)	0.02	5%
PGgpL (+)	1.8×10^{-5}	0.004%
PGgpL (-)	0.009	2%
PGL	0.44	100%

Figure 3-18

Frameshift Factor and Efficiency of frameshift constructs. The frameshift factor is calculated by determining the ratio of luciferase activity to GFP fluorescence. Frameshifting efficiency is a constructs' frameshift factor as a percentage of PGL frameshift factor.

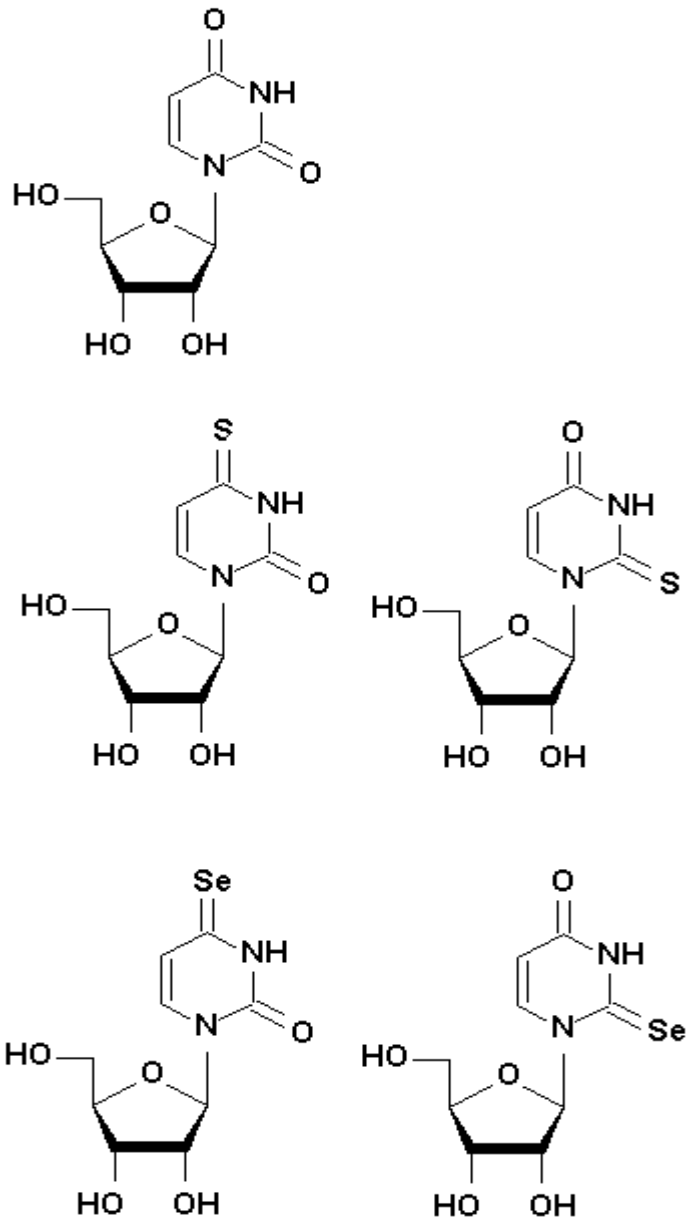


Figure 3-19

Selenium and sulfur substituted uridine. Hypermodified uridine is commonly found in nature to contain sulfur or selenium in place of oxygen in the 2’ or 5’ positions of the base.

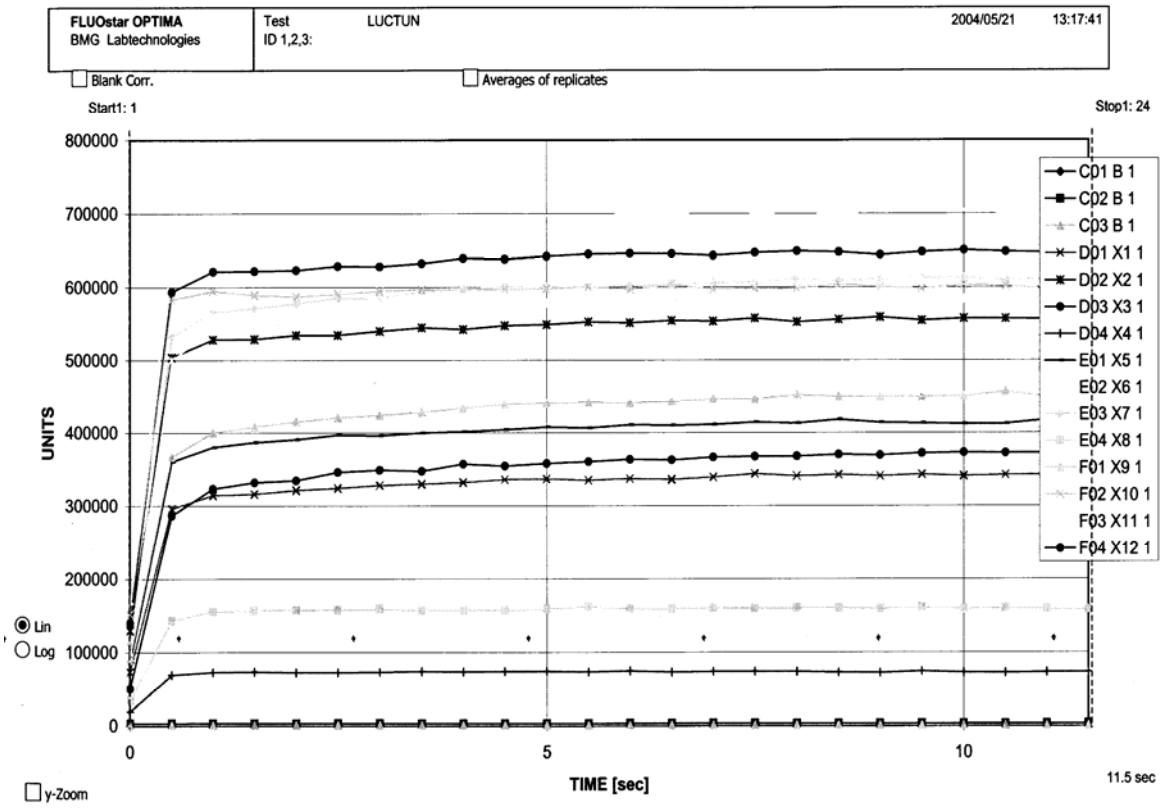


Figure 3-20

Luminescence of constructs expressing cell lysates. Luminescence of cell lysates was quantitated in a FLUOstar Optima plate reader read over a 12 second period.

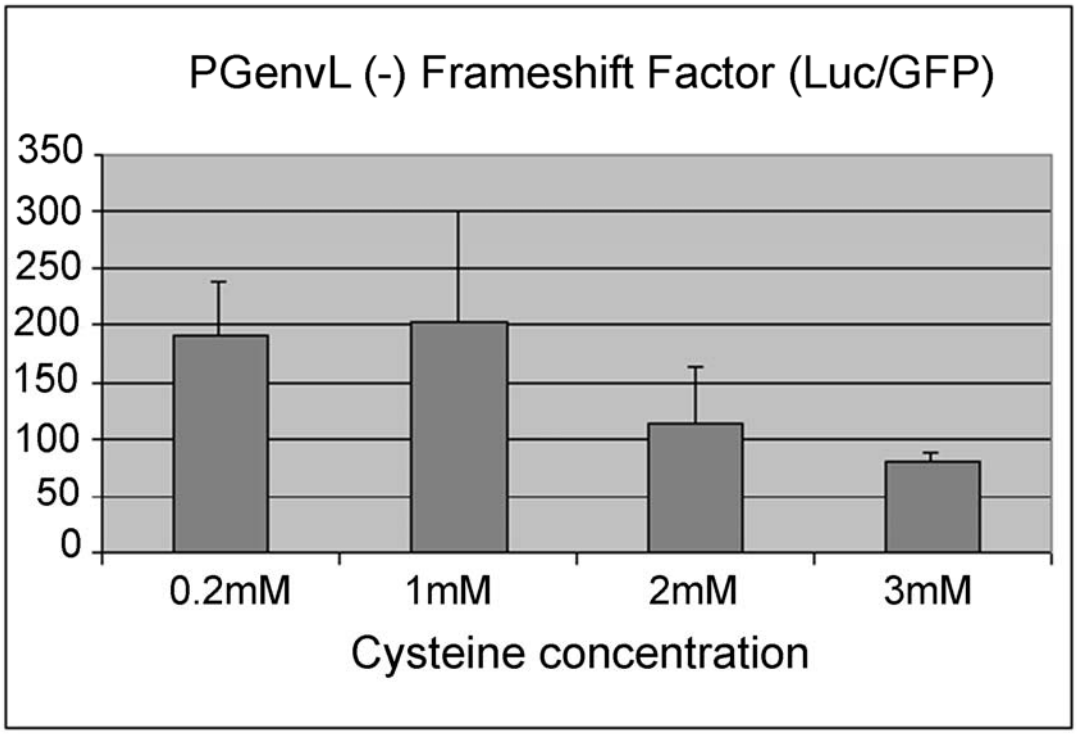


Figure 3-21

The Frameshift Factor of PGenvL (-) with cysteine supplementation. Sample wells were supplemented with cysteine concentrations ranging from 0.2mM to 3mM. The Frameshift Factor of PGenvL (-) was found to decrease with an increase in cysteine concentrations.

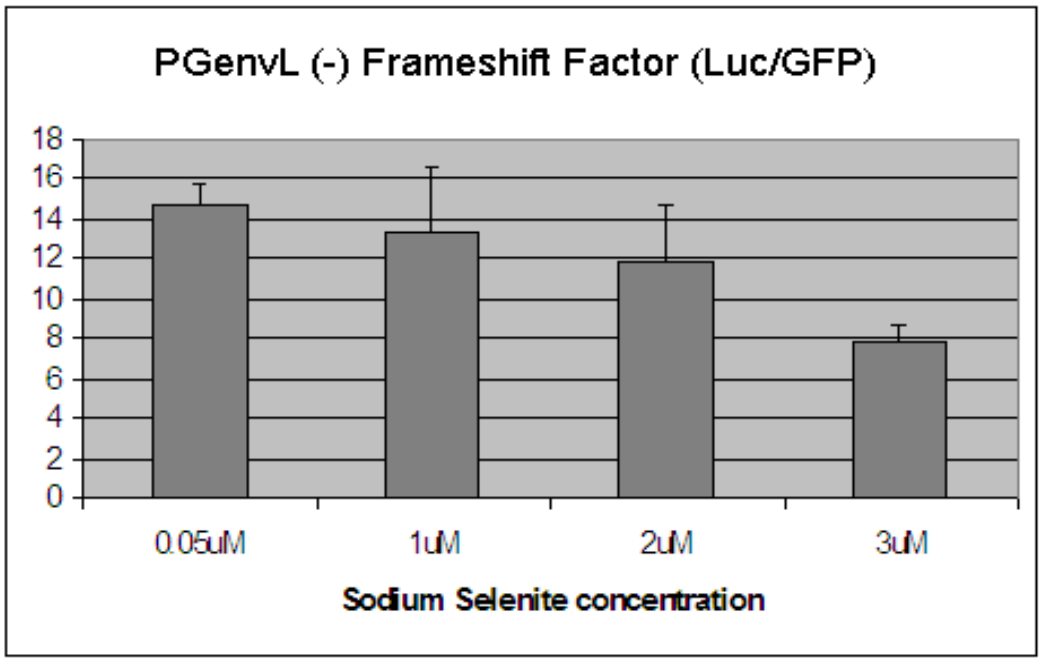
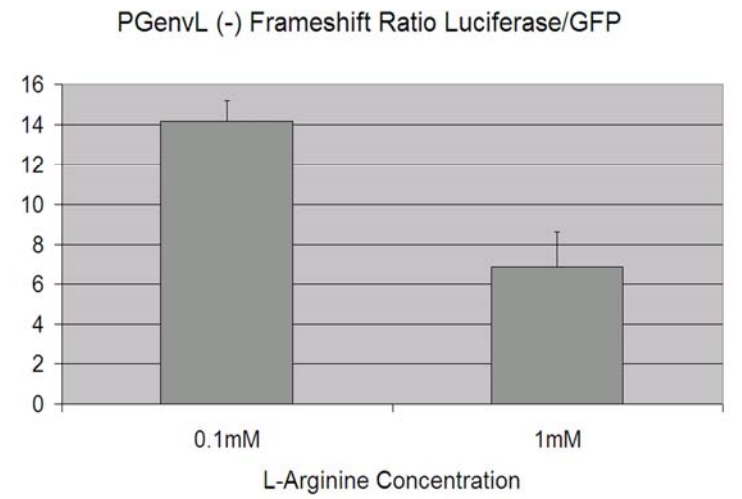


Figure 3-22

The Frameshift Factor of PGenvL (-) with sodium selenite supplementation.

Sample wells were supplemented with cysteine concentrations ranging from 0.05uM to 3uM. The Frameshift Factor of PGenvL (-) was found to decrease with an increase in sodium selenite concentration.



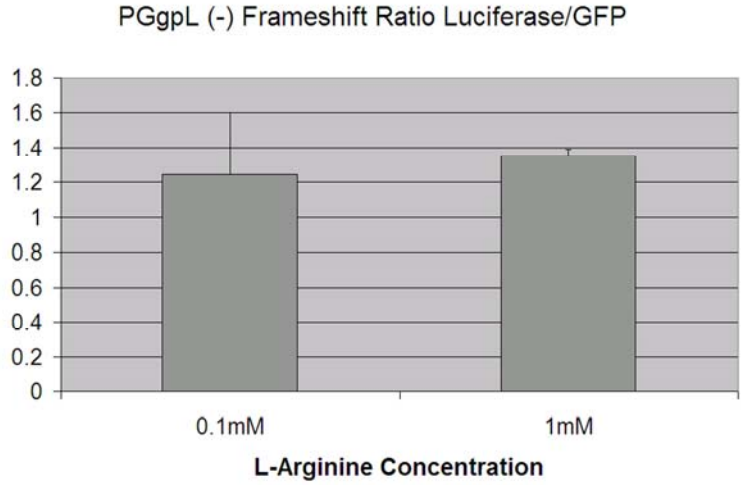


Figure 3-23

Effects of L-arginine on frameshifting. Frameshift factor of PGenvL (-) shows an increase in frameshift factor with L-arginine deficient media (while the comparable construct for the known HIV-1 *gal-pol* frameshift site, PGgpL (-), shows no significant change in frameshift factor in response to differences in L-arginine concentration ($P<0.005$ for PGenvL (-), $P>0.1$ for PGgpL (-))

CHAPTER 4

RIBOSOMAL FRAMESHIFTING IN THE NEF GENE OF HIV-1 AND CHEMOTAXIS OF NOVEL NEF ISOFORMS[1]

[1] Olubajo, B and Taylor, E.W. to be submitted to Journal of AIDS Research and Human Retrovirology

Overview

This chapter contains a review of ribosomal frameshifting, the properties and functions of nef, and construction of constructs for selenocysteine incorporation studies. Furthermore, the development of a GFP- β-galactosidase, dual reporter frameshift assay system is discussed and frameshifting of the proposed nef frameshift sequence is assessed. Several constructs were made to study proteins that would result from the proposed frameshift genes in nef of HIV and incorporation of selenocysteine. The main objectives of the experiments conducted in this study was to: 1. Validate the frameshift sequence in the nef region of HIV-1. 2. Examine chemotactic ability of nef and frameshift resultant nef isoforms. 3. Construct vectors for the study of selenocysteine incorporation into HIV nef.

Introduction and Review

The nef gene of HIV encodes one of the earliest expressed HIV proteins made from the genome. Nef is the largest of HIV's accessory proteins and is found to be expressed in significant quantities early in HIV infection (Kim et al. 1989). Nef, which stands for negative factor, was named according to early reports indicating that nef caused a decrease in HIV transcription. Currently, nef is believed to be important for several aspects of viral replication.

The molecular weight of HIV-nef is reported in literature as ranging between 25 and 35 KD. It is difficult to determine the exact basis for the disparity of nef's molecular weight in different studies, but this underscores the complexity of characterizing nef and understanding its multitude of functions.

Cells infected with the HIV virus have been shown to contain nef molecules localized to the nucleus, cytoplasm and in extracellular space (Fujii et al. 1996, Murti et al. 1993). Nef has also been reported as a monomer, and in multimeric forms (Fujii et al. 1996). It is found in virion particles and is cleaved by HIV protease in the virion (Pandori et al. 1996).

Nef associates with several cellular kinases and can be serine or threonine phosphorylated (Arnold et al. 1997, Sawai et al. 1994). Nef has been shown to induce T-cell activation and also to block IL-2 expression (Sawai et al. 1994, Schraeger et al. 1995). Studies have demonstrated that HIV-1 Nef acts within infected cells to increase the infectivity of progeny viruses (Spina et al. 1994, Miller et al. 1994).

Nef has been found to down regulate the expression of a chemokine transcription factor NFκB (Luria et al. 1991). Nef also down-modulates MHC class 1 expression on the surface of T-cells (Schwartz et al. 1996), and causes a decrease in the number of CD4 molecules on immune cells (Garcia et al. 1992).

The Nef protein of HIV-1 has been ascribed numerous functions and characteristics. The aforementioned roles have been assigned to nef as well as other biological functions not stated in this chapter and attributes currently being studied. As previously indicated, Nef has also been presented in literature as a protein of varying molecular weight. Though the concept of one protein playing so many roles in a biological system cannot be dismissed as absolutely impossible, it is still one that requires cautious analysis. The multitude of functions ascribed to the nef protein can raise questions about the possibility of different isoforms or variants of the nef protein, which could in turn account for specific functions. Different isoforms of the nef protein

could easily be overlooked during nef function or characterization studies if researchers are unaware of transcriptional or translational modifications that viruses exploit in encoding viral genes. The possibility of such a trend may be evident in the discrepancy of nef's molecular weight that is reported in the literature.

Translational modifications to RNA sequences are common occurrences in encoding viral genes, and can be carried out by different mechanisms. One such method employed by viruses in efficient utilization of DNA sequences is ribosomal frameshifting.

Ribosomal frameshifting is a method used by several organisms to maximize protein output from DNA sequences. It is a process that is very common in viruses, and is important for maximum utilization of viral genomes. Frameshifting requires a suitable heptanucleotide "slippery" sequence and an RNA pseudoknot.

Slippery sequences and pseudoknots have been shown to be involved in frameshifting in several cases either as a pair, or individually. Slippery sequences are heptanucleotide RNA sequences, which, in many cases, have a less efficient interaction with the anticodon loop of the corresponding tRNAs in the ribosome during protein translation. Slippery sequences generally conform to the sequence pattern X-XXY-YYZ, where X, Y or Z represents a single purine or pyrimidine base.

Pseudoknots are tertiary RNA structures that are formed by base pairing between a secondary loop structure and complement bases outside the loop. This results in an RNA structure composed of two loops and two stems. The two stems coaxially stack to form a quasi-continuous RNA helix while the two loops span the major and minor grooves of the helix.

Though frameshifting is not completely understood, current opinions suggest that a frameshift begins when a ribosome encounters a pseudoknot in the mRNA to be encoded. In the process of the ribosome unraveling the pseudoknot, a pause in translation occurs which in turn facilitates slippage of the ribosome at a suitable slippery sequence. The ribosome can slip either forward (+change) or backward (- change) and a different sequence of codons is interpreted from there on.

Frameshifiting can also serve as a means of regulating protein translation. The pol protein of HIV-1 is encoded as a fusion to the gag protein through –1 frameshifting (Jacks et al. 1988). Frameshifting in the gag-pol gene of HIV involves the heptanucleotide slippery sequence UUUUUUA and a pseudoknot, which can adopt two conformations as shown in Figure 4-1. Frameshifting in the gag-pol gene of HIV-1 has been shown to keep the ratio of gag to pol translation at 20:1 (Jacks et al. 1988). This ratio was later shown to be important for HIV infectivity when compared to HIV particles induced to express gag and pol proteins at ratios other than the physiological ratio (Shehu-Xhilaga et al. 2001).

Computational analysis of the HIV-1 genome indicates the potential for a -1 frameshift in the nef coding region (Taylor et al. 1996). This is supported by detection of an ideal heptameric UUUAAAG slippery sequence as seen in Figure 4-2, which would be capable of promoting ribosome slippage unaided, and a potential RNA pseudoknot structure. The slippery sequence and pseudoknot were highly conserved in 95% of HIV variants in the database, and the mutant variant 5% had deviations from the ideal slippery sequence, but would still be capable of frameshifting (Taylor et al. 1997).

The frameshift sequence in the nef gene of HIV-1 was one of several frameshift sequences found in the HIV genome. These frameshift sequences serve as a major component of a proposed viral selenoprotein theory (Taylor et al. 1994,1996). The pseudoknot involved in frameshifting of the nef-fs gene was found upstream of novel reading frames in the nef gene. The slippery sequence involved in nef-fs frameshifting was found 7 bases upstream of the proposed pseudoknot and would involve leucine and lysine codons in the A and P site during translation of both the 0 and –1 frames (Taylor et al. 1996,1997), a feature that could facilitate ribosome slippage.

Evaluation of the nef-fs pseudoknot sequence using the RNA structure analysis program mFOLD, confirmed secondary structure in the pseudoknot, which supported the novel frameshift sequence in the HIV nef gene (Taylor et al. 1996,1997). The program was used to search for stable RNA hairpin structures in the HIV genome. Regions of 20 base pairs upstream or downstream from the base of the hairpin structure were searched for complementary bases in the loop of the RNA hairpin structure (Taylor et al. 1996,1997).

The frameshift sequence was tested with a cell based frameshift assay. The frameshift assay was developed to test the validity of hypothesized frameshift sequences. The nef frameshift sequence was studied with a β-galactosidase reporter system. β-galactosidase was cloned and placed in a vector downstream of the green fluorescent protein gene, which was included to add bulk to a frameshift fusion protein. The nef frameshift sequence was cloned into the vector between the GFP gene and the β-galactosidase gene in a –1 frame, and the assaying vector was expressed in MDCK

cells. β-galactosidase activity was quantitated in determination of the frameshifting capacity of the putative nef frameshift sequence.

Materials and Methods

Frameshift Assay Constructs

The β-galactosidase gene was digested from PSV-BGal (Promega) with Kpn I and BamH I restriction endonucleases. The digested β-galactosidase gene was ligated into a Kpn I and BamH I digested PEGFPC-1 vector (Clontech). This PGB vector was the negative control vector in the assay.

The nef frameshift sequence was cloned from PBH10 (HIV genome from NIH AIDS Reagent Program) by polymerase chain reaction adding Hind III and Kpn I sites to the 5' and 3' ends of the cloned nef frameshift sequence. The frameshift sequence was digested with Hind III and Kpn I and subsequently ligated to the, similarly digested, PGB construct. Hence, the frameshift sequence was placed between the GFP and β-galactosidase genes. A map of these vectors is illustrated in Figure 4-5, and results of vector construction are shown in Figures 4-3 and 4-4. All vectors were confirmed with sequencing.

Transfection of cell lines

MDCK cells were grown in Dubelco Minimal Essential Medium with 2mM L-glutamine, 1.5g/L sodium bicarbonate and 1.0 mM sodium pyruvate. The final media volume is adjusted to contain 10% fetal bovine serum. Cell lines were grown in 35mm diameter, 6 well plates to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 2μg of appropriate vector for 6 hours. Cells were

transiently transfected with the PGB and PGnefB constructs. The cells were incubated at 37º C with 5% CO_2 for 48 hours.

Cell based frameshift Assay

Transfected cells are lysed by incubating with 400μl of a detergent based Reporter Lysis Buffer (Promega) per well. The cells are scraped and exposed to 3 freeze thaw cycles. Crude extracts are centrifuged and supernatant is recovered. The supernatant is assayed for reporter gene activity.

β-galactosidase activity is quantitated by incubating 150μl of the cell lysates with an equal volume of a β-galactosidase assaying buffer, which contained the indicator enzyme substrate o-nitrophenol-β-d-galactopyranoside (ONPG). The enzyme reaction was incubated at 37 degrees Celsius for 1-3 hours. Following incubation, a yellow color was visible with cell lysates that contain –1 frameshift expressed β-galactosidase. The reaction is stopped with sodium carbonate and absorbance is quantitated at 420nm. The Lowry assay is done to quantitate total protein content as a means of standardizing the assay.

Results and Discussion

The nef frameshift construct PGnefB showed significant β-galactosidase activity. The control PGB construct also exhibited β-galactosidase activity, which was unexpected, since the β-galactosidase gene was in a –1 frame in relation to the GFP gene. The PGnefB construct did show increased β-galactosidase activity in comparison to the PGB construct. The nef-fs sequence resulted in an increased expression of β-galactosidase as indicated by the β-galactosidase activity profile exhibited by MDCK

cells transfected with PGnefB. The results were not definitively conclusive, because the control PGB vector also displayed β-galactosidase activity as illustrated in Figure 4-6.

Potential Selenoproteins encoded from the HIV nef gene

Eukaryotic selenoprotein synthesis

Selenocysteine, which is often referred to as the "21st amino acid," is encoded by the codon UGA. The UGA codon is most commonly recognized and efficient as, the termination codon for protein translation. The dual nature of the UGA codon creates one of the main setbacks in identification of selenoproteins.

Selenocysteine incorporation into proteins differs from the standard mechanisms of amino acid incorporation. Selenocysteine incorporation requires cis and trans acting components that are exclusive to selenoprotein synthesis.

Selenoprotein synthesis requires a secondary RNA structure known as a selenocysteine insertion sequence element (SECIS) (Low et al. 1996). SECIS elements are long RNA hairpin structures that consist of internal loops and bulges as shown in Figure 4-7. Their are sequence specific RNA motifs of the SECIS element that reside mostly at the base of the hairpin loop, and depending on the type of SECIS element, a triple adenine sequence resides either in the main loop of the hairpin structure, or as a component of the internal loop (Gu et al. 1997, Low et al. 1996).

Selenoprotein synthesis requires trans acting elements that play different roles in selenocysteine incorporation. Selenophosphate is thought to be the donor compound of selenium to the process. TRNA-Sec, a selenocysteine specific tRNA, is charged with serine, and selenophosphate donates selenium to the mechanism to convert the serine to selenocysteine (Low et al. 1996). In the bacterial mechanism, three genes sel A, sel

B and sel D are involved in encoding proteins involved in the selenocysteine incorporation mechanism. Sel A encodes a selenocysteine synthase involved in converting the serine charged tRNA-Sec to a selenocystyl-tRNA-Sec. The selD gene encodes a selenophosphate synthetase involved in generation of selenophosphate, while selB encodes a selenocysteine specific elongation factor (Low et al. 1996).

Other cis and trans acting elements may exist and are yet to be discovered. The process of selenocysteine incorporation in prokaryotes is better defined, but eukaryotic homologs of some defined components of the prokaryotic mechanism have not been ascertained or still remain obscure.

Selenobiology, in many ways, is still in its early stages of discovery. Further investigation is required for a more comprehensive understanding of selenoproteins and to facilitate unlocking probable therapeutic potential.

Chemotaxis and Nef

Viruses encode for proteins that are very similar to physiologic chemokines, and use these proteins to fool the immune system into favoring the conditions suitable for the virus' survival. The Epstein-Barr virus, for example, encodes a chemokine homologous to IL-10, which the virus uses to inhibit macrophage activation. Some poxvirus species encode molecules that are similar to cytokines IL-1 (Alcami et al. 1992, Spriggs et al. 1992), IFN-gamma (Davies et al. 1993) and TNF (Alcami et al. 1995), all of which are involved in chemokine processes. These are all believed to serve as competitive inhibitors of the host responses to these cytokines. The character of these and other viruses, along with the documented association between HIV and chemokine

receptors, further support the hypothesis that HIV may have chemokine homologs, which should not be overlooked.

Computational analysis of the HIV-1 genome identified several potential selenoproteins in the HIV-1 genome (Taylor et al. 1994). These genes were found to contain highly conserved selenocysteine codons (Taylor et al. 1994). These conserved UGA codons discovered in the HIV-1 genome could not be attributed to translation termination of known HIV-1 proteins.

Conserved UGA codons exist in a –1 frameshift resultant reading frame of the HIV nef gene, that cannot be explained by previously determined requirements of the nef coding sequence (Taylor et al. 1994, Taylor et al. 1997). Preceding the –1 reading frame of the nef gene is a conserved initiation codon that is followed by a short sequence homologous to a signal peptide translocation sequence (Nadimpalli et al. 1997). Two conserved selenocysteine codons appear in the nef –1 frame. Furthermore, a suitable +1 frameshift sequence is found after these UGA codons (Nadimpalli et al. 1997). Hence, with a –1 and +1 frameshift, a functional chemotactic cassette could be dropped into the nef gene.

Analysis of the frameshift product shows the nef-fs gene could encode a gamma chemokine homolog. The sequence shows similarity through sequence alignment to 20 CXC, CC and gamma or C-type chemokines. It shows the closest similarity to the CXC chemokine IL 8 (Nadimpalli et 1997). The 2 conserved cysteines of the chemokines align with 2 UGA codons of the nef frameshift sequence.

The alignment score for nef-fs vs. the entire alignment was greater than 4.5 SD higher than the average score of randomized sequences of similar size and

composition. Protein structure fingerprint database probing was also used to model protein structure. The top hit was the CXC chemokine interleukin 8 at 4.8 SD significance, using a gap penalty of 2 and a gap extension penalty of 0.2. Homology modeling through threading techniques based on interleukin-8 crystal structure was also found to be energetically favorable. (Nadimpalli et al. 1997)

To examine the feasibility of the proposed chemokines, constructs were developed with mutated versions of the nef gene. The mutations were made to simulate the product of a –1 frameshift, a +1 frameshift, and a –1, +1 frameshift. The constructs were expressed in a eukaryotic cell line, and chemotaxis ability was assessed.

Materials and Methods

Vector Construction

The nef gene of HIV was cloned from the PBH10 vector (HIV genome from NIH AIDS Reagent Program). The nef gene contained a premature termination codon, which had to be mutated from UGA to UGG. Polymerase Chain Reaction site directed mutagenesis was used to convert the stop codon to a tryptophan codon by a PCR mutation procedure shown in Figure 4-8. The process involved the use of two overlapping primers that were designed to include complement bases of the mutation. Primers corresponding to the beginning and end of the nef gene were also designed and synthesized with the 5' primer incorporating a Nhe I restriction site, and the 3' primer a Sal I restriction site.

A segment of the nef gene preceding the mutation site and beginning with the nef gene initiation region was amplified by PCR with Nhe I-7701 and the 3' → 5' mutation overlapping primer StopI-8528. Stop2-8516, The 5' → 3' mutation overlapping

primer, and a Sal I- 8923, a primer corresponding to the end of nef gene, were used to amplify the region of the nef gene following the mutation. Both segments were then subjected to a PCR reaction in which no separate template is added. In this step, the two double stranded gene segments are melted to their single strand components, and are then allowed to re-anneal during the PCR process. The overlapping complementary portions of the annealed single strands served as primers and the remainder of the strand serves as a template. In the final step, the primers corresponding to the start (Nhe I-7701) and end of the nef gene (StopI-8528) are used to amplify the new gene, which included the mutation.

A similar process was used to add mutations to nef constructs involved in simulating a frameshifted gene product, except for the double mutant construct, which included a repeat of the mutant segment synthesis step. The three mutant constructs Nef-pos, Nef-neg and the double mutant construct Nef-pos-neg, involved the previously described process using Nef-wt as a first step template. These mutants involved the nef gene primers Nhe1-8001 and KpnI-8773, which correspond to the beginning and end of the respective mutant nef gene. The primers Nef-8433 and Nef-8419 were complementary overlapping primers for the –1 frameshift mutants while Nef-8526 and Nef 8508 are overlapping primers to induce a +1 frameshift gene product conformation.

Inserting the deiodinase SECIS element in the PEGFP-CI vector between, Kpn I and Hind III restriction sites, led to a vector labeled PG-SECIS. Each mutated nef gene was incorporated into the eukaryotic expression vector PG-SECIS, by digesting the vector and the mutant nef gene PCR products with Nhe I and Sal I and subsequent

ligation of the mutant genes with digested PG-SECIS. These experiments are illustrated in Figures 4-8 – 4-15, and illustrated in tables 4-2 and 4-3.

Eukaryotic Expression of Constructs

293-T cells were grown in Dubelco minimal essential medium with 2mM L-glutamine, 1.5g/L sodium bicarbonate, 1.5μM sodium selenite and 1.0 mM sodium pyruvate. The final media volume is adjusted to contain 10% fetal bovine serum. Cell lines were grown in 35mm diameter, 6 well plates to 95% confluency. Lipofectamine (Invitrogen) was used to transfect the cell lines in each well with 2μg of appropriate vector for 6 hours. The cells were incubated at 37º C with 5% CO_2 for 48 hours.

Chemotaxis Assay

The cell samples were lysed with 400μl of Phosphate Buffered Saline by sonication and freeze thaw cycles. The samples were centrifuged and the supernatant was assayed for chemoattractant properties. 96-well chemotaxis-plates (Neuroprobe) were used to assess chemoattractant properties of nef mutant proteins in cell lysates as shown in Figure 4-16. 30ul of cell lysate is placed in the well or chemoattractant chamber of the chemotaxis plates. The 5uM-pore size filter was gently placed over the meniscus of the test chemoattractant cell lysate.

The next step involved placing a suspension of Jurkat cells on the filter above the chemoattractant. A 70μl drop of Jurkat cell solution, at a concentration of 10^6 cells/ml was placed on the filter above each well containing a chemoattractant as shown in Figure 4-16. The assay was incubated at 37º C for 5 hours, and following the incubation period, the cells were counted with a hemacytometer under a phase-contrast microscope.

Results and Discussion

Some lysates of the recombinant nef-expressing cells appeared capable of inducing Jurkat cell chemotaxis. Cells lysates with transfected Nef-wt vector exhibited significant chemotaxis over control lysates in accordance with previous findings in the literature (Shutt et al. 1999, Koedel et al. 1999). Cells transfected with the Nef-pos vector also exhibited the highest chemotaxis of Jurkat cells in the assays, with the cell count almost double the number of migrated cells found in the wells containing unmutated nef cell lysates as shown in Figure 4-17. The remaining nef mutant lysates had no significant chemotaxis in comparison with lysates from control transfected and untransfected cells.

Future Directions

Analysis of the nef gene indicates a highly conserved selenocysteine codon, known commonly as the termination codon for nef, is preceded by a cysteine codon and was found to be highly conserved in several HIV strains. A UAA termination codon is evident, 100 bases following the UGA codon and both codons fall in the same reading frame. The presence of cysteine and selenocysteine in close proximity at the c-terminal of a protein is a hallmark of selenium containing oxidoreductases, enzymes involved in biological electron transfer, reducing or oxidizing reactions. For this reason, our laboratory was interested in selenocysteine incorporation and testing oxidoreductase activity in nef of HIV. To test this hypothesis 2 plasmids were constructed, to assess potential selenocysteine incorporation in the nef gene of HIV.

Materials and Methods

The first vector was constructed by restriction digest of the Nef-wt construct and the PG-SECIS vector with Nhe I and Sal I restriction endonucleases and subsequent ligation of the fragments. This vector therefore had the nef gene in a eukaryotic expression vector, with the deiodinase SECIS element and was named Nef-wt-SECIS.

The second vector was made to test the incorporation of selenocysteine into the nef protein with a proposed HIV SECIS element (Taylor et al. 1994). The Nef-HIV-SECIS vector contains the predicted HIV-SECIS, directly following the nef gene, extending into the 3' LTR, and would contain the known conserved RNA motif AUGA that is required for selenocysteine incorporation. These vectors will be useful in understanding any hidden functionality involved in the highly conserved UGA codon, as well as the potential SECIS element. Map of vector construction and results are shown in Figures 4-18 and 4-19.

References

Alcami, A. and G. L. Smith (1992). "A Soluble Receptor for Interleukin-1-Beta Encoded by Vaccinia Virus - a Novel Mechanism of Virus Modulation of the Host Response to Infection." Cell 71(1): 153-167.

Alcami, A. and G. L. Smith (1995). "Cytokine Receptors Encoded by Poxviruses a Lesson in Cytokine Biology." Immunology Today 16(10): 474-478.

Arnold, S., P. Franken, et al. (1997). "The crystal structure of HIV-1 Nef protein bound to the Fyn kinase SH3 domain suggests a role for this complex in altered T cell receptor signaling." Structure 5(10): 1361-72.

Davies, M. V., H. W. Chang, et al. (1993). "The E3l and K3l Vaccinia Virus Gene-Products Stimulate Translation through Inhibition of the Double-Stranded Rna-Dependent Protein-Kinase by Different Mechanisms." Journal of Virology 67(3): 1688-1692.

Fujii, Y., K. Otake, et al. (1996). "Clustered localization of oligomeric Nef protein of human immunodeficiency virus type 1 on the cell surface." FEBS Lett 395(2-3): 257-61.

Garcia, J. V. and A. D. Miller (1992). "Downregulation of cell surface CD4 by nef." Res Virol 143(1): 52-5.

Gu, Q. P., M. A. Beilstein, et al. (1997). "Conserved features of selenocysteine insertion sequence (SECIS) elements in selenoprotein W cDNAs from five species." Gene 193(2): 187-96.

Jacks, T., M. D. Power, et al. (1988). "Characterization of ribosomal frameshifting in HIV-1 gag-pol expression." Nature 331(6153): 280-3.

Kim, S. Y., Byrn. R et al. (1989). “Temporal aspects of DNA and RNA synthesis during human immunodeficiency virus infection: Evidence for differential gene expression.” Journal of Virology 63: 3708-3713

Koedel, U., B. Kohleisen, et al. (1999). "HIV type 1 Nef protein is a viral factor for leukocyte recruitment into the central nervous system." Journal of Immunology 163(3): 1237-1245.

Low, S. C. and M. J. Berry (1996). "Knowing when not to stop: selenocysteine incorporation in eukaryotes." Trends Biochem Sci 21(6): 203-8.

Luria, S., I. Chambers, et al. (1991). "Expression of the type 1 human immunodeficiency virus Nef protein in T cells prevents antigen receptor-mediated induction of interleukin 2 mRNA." Proc Natl Acad Sci U S A 88(12): 5326-30.

Miller, M. D., M. T. Warmerdam, et al. (1994). "The human immunodeficiency virus-1 nef gene product: a positive factor for viral infection and replication in primary lymphocytes and macrophages." J Exp Med 179(1): 101-13.

Murti, K. G., P. S. Brown, et al. (1993). "Highly localized tracks of human immunodeficiency virus type 1 Nef in the nucleus of cells of a human CD4+ T-cell line." Proc Natl Acad Sci U S A 90(24): 11895-9.

Nadimpalli, R.G., Zhang, W., et al. (1997) “The HIV-1 nef transframe protein has significant sequence and structural similarity to chemokines, as assessed by threading (inverse folding), sequence analysis, and molecular modeling.” Antiviral Research, 34(2): A57

Pandori, M. W., N. J. Fitch, et al. (1996). "Producer-cell modification of human immunodeficiency virus type 1: Nef is a virion protein." J Virol 70(7): 4283-90.

Sawai, E.T. , Baur, A. et al. (1994) “Human immunodeficiency virus type 1 Nef associates with a cellular serine kinase in T lymphocytes.” Proc. Natl Acad. Sci. 91: 1539-1543

Schrager, J. and Marsh, J. (1999) “HIV-1 Nef increases T cell activation in a stimulus-dependent manner “ Proc. Natl. Acad. Sci. 96: 8167-8172.

Schwartz, O., V. Marechal, et al. (1996). "Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein." Nat Med 2(3): 338-42.

Shehu-Xhilaga, M, Crowe, S. et al. (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.

Shutt, D. C. and D. R. Soll (1999). "HIV-induced T-cell syncytia release a two component T-helper cell chemoattractant composed of Nef and Tat." Journal of Cell Science 112(22): 3931-3941.

Spina, C., Kwoh, T. (1994) “The importance of nef in the induction of human immunodeficiency virus type 1 replication from primary quiescent CD4 lymphocytes” Journal of Exp. Med., 179 (1), 115-123

Spriggs, M. K., D. E. Hruby, et al. (1992). "Vaccinia and Cowpox Viruses Encode a Novel Secreted Interleukin-1-Binding Protein." Cell 71(1): 145-152.

Taylor, E. W., C. S. Ramanathan, et al. (1994). "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(17): 2637-54.

Taylor, E.W., Ramanathan, C.S., et al. (1996) “A general method for predicting potential new genes in nucleic acid sequences: application to the human immunodeficiency virus.“ In: Computational Medicine, Public Health and Biotechnology, M. Witten, Ed. World Scientific, Singapore. Ser. Math. Biol. Med. 5: 285-309.

Taylor, E. W., R. G. Nadimpalli, et al. (1997). "Genomic structures of viral agents in relation to the biosynthesis of selenoproteins." Biological Trace Element Research 56(1): 63-91.

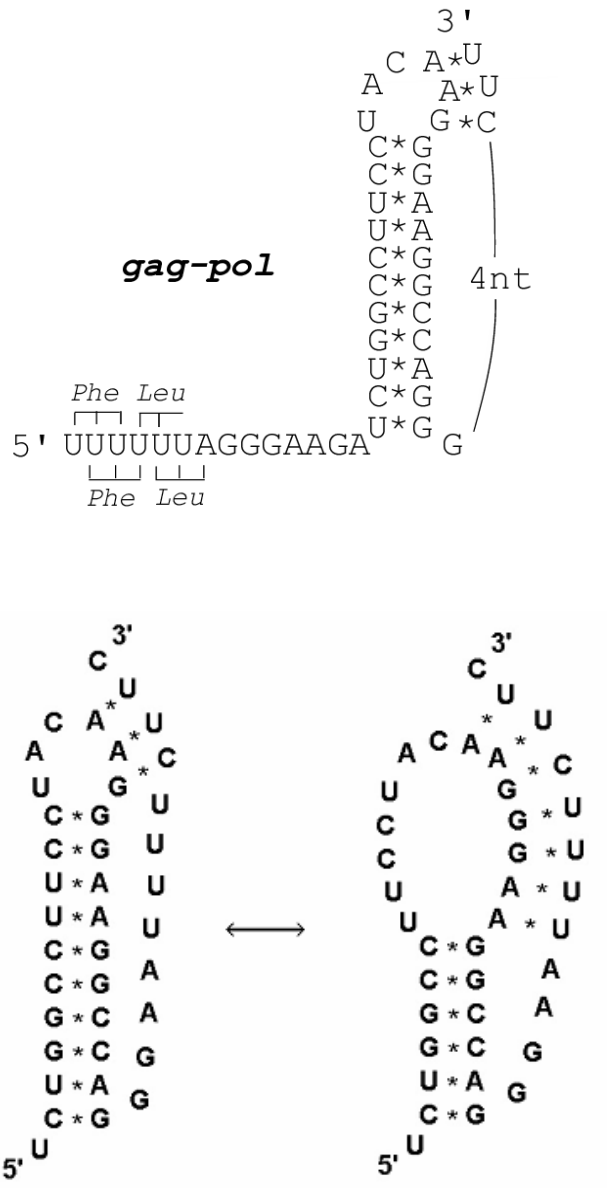


Figure 4-1

Diagram of known gag-pol frameshift sequence. The diagram also shows two possible conformations that the pseudoknot could adopt (Taylor et al. 1994).

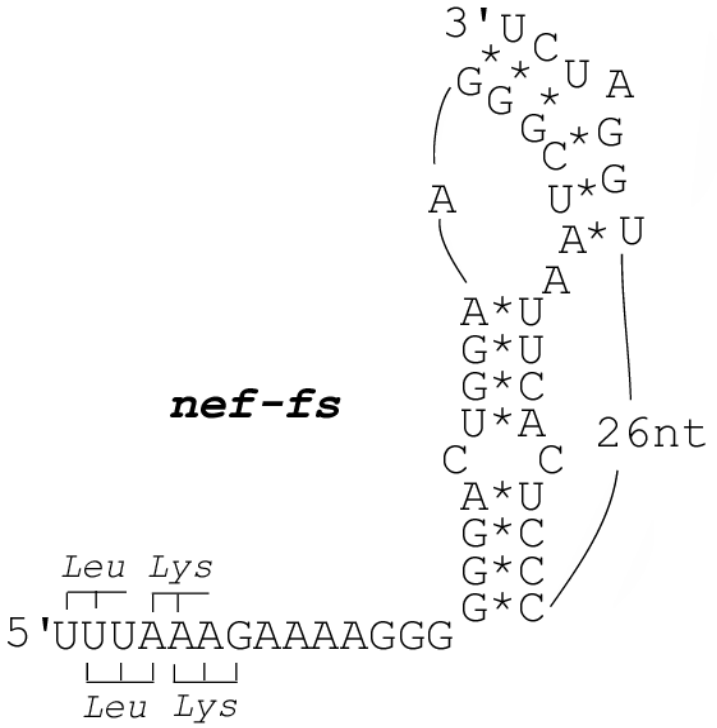


Figure 4-2

Proposed nef frameshift sequence. The putative frameshift sequence in the nef gene of HIV-1 has a suitable frameshift sequence that conforms to the XXX YYY Z pattern of slippery sequences. The predicted pseudoknot would also be capable of facilitating the pause in translation.

Table 4-1

Primers for construction of PGnefB. These primers were used to amplify proposed nef frameshift sequence from PBH10 vector.

Oligos designed for PCR of nef frameshift sequence

Nef-fs-Kpn I

TTG CGG GGT ACC AGA TCC ACA GAT AAA GGA TA
Kpn I

Nef-fs-Hind III

TTG GCC CAA GCT TCT GTA GAT CTT AGC CAC
Hind III

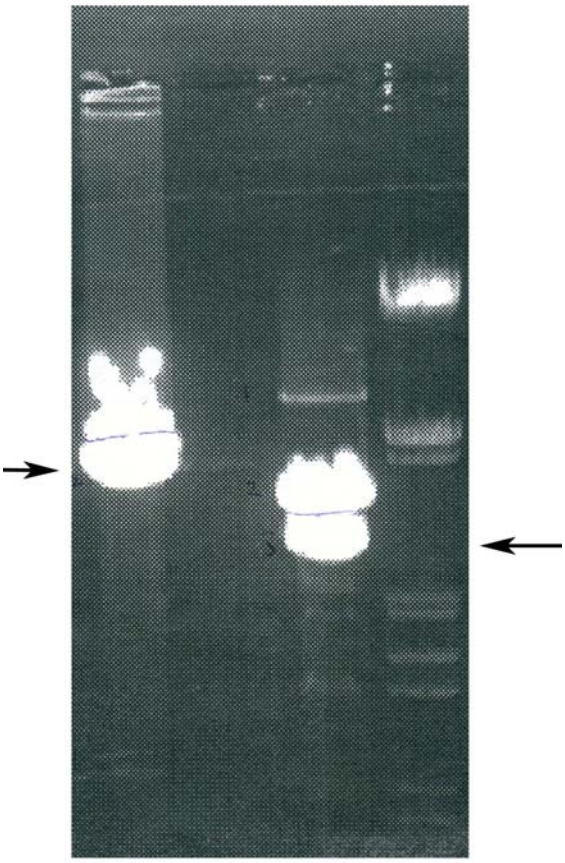


Figure 4-3

Restriction digest of PSV-Bgal and PEGFP-C1. PSV-Bgal (Promega) is in lane 1 and PEGFP-CI (Clontech) in lane 3. Lane 4 is DNA marker, and arrows indicate fragment that was later gel-purified.

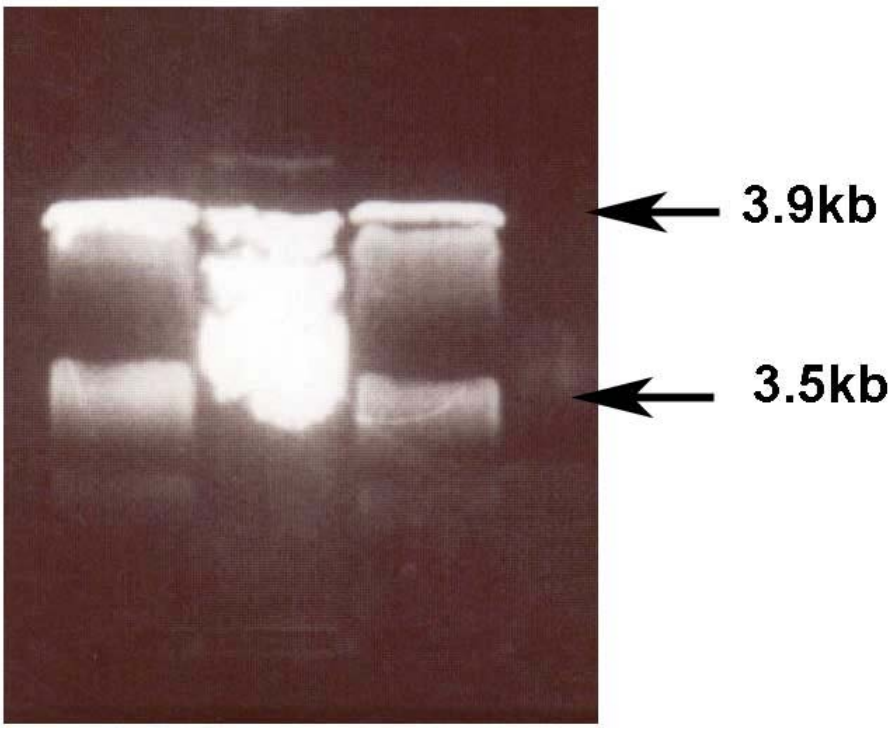


Figure 4-4

Analytical digest of PGnefB and PGB. PGnefB and PGB are in lanes 1 and 3 respectively. 3.9 KB fragment corresponds to PEGFP-CI fragment, while 3.5kb fragments corresponds to β-galactosidase gene.

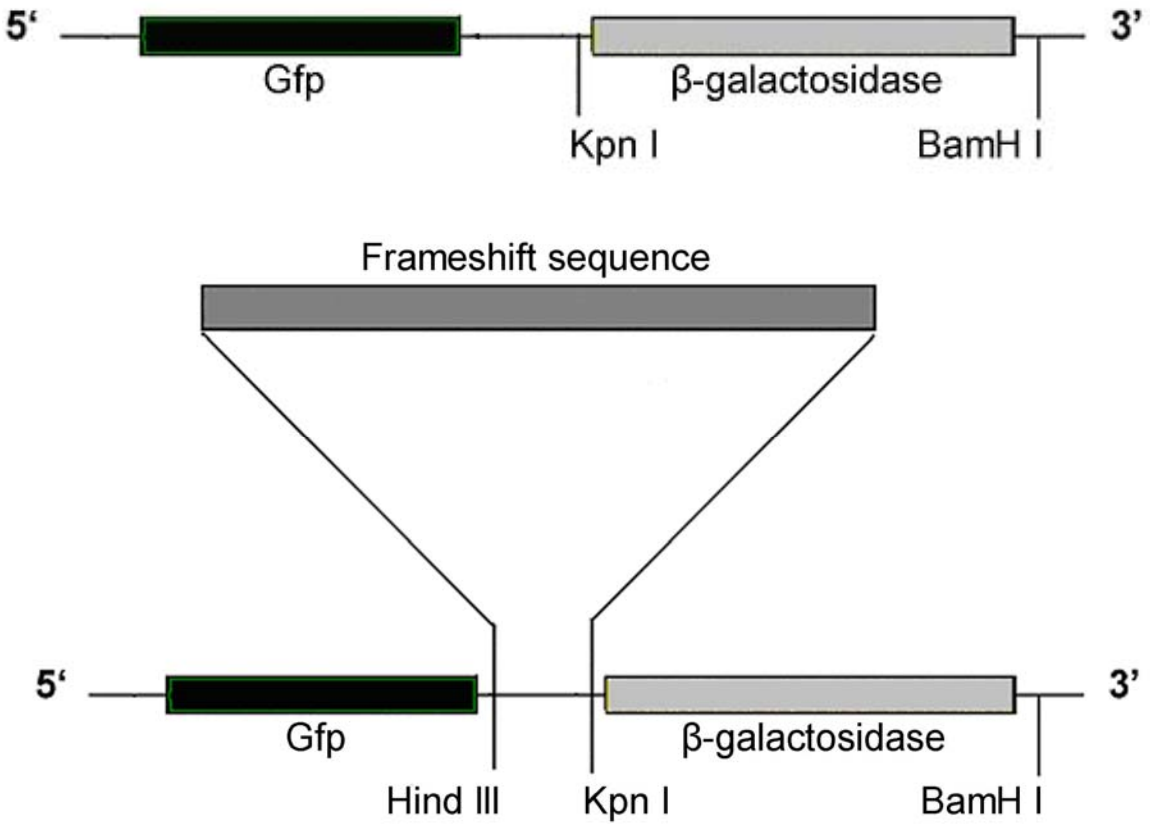
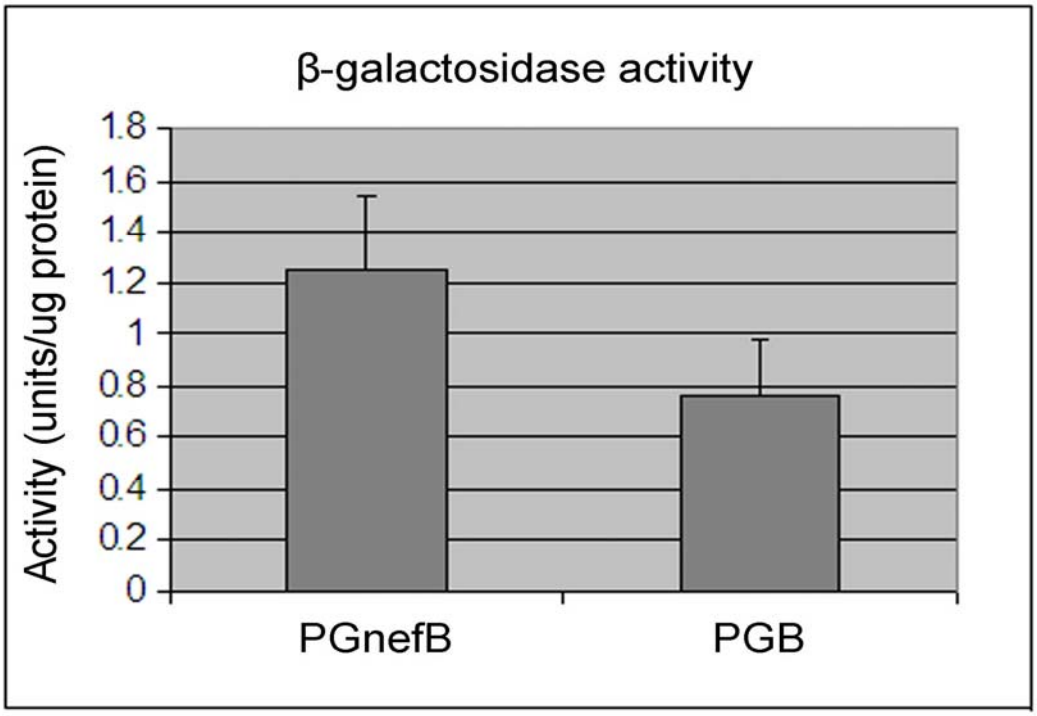


Figure 4-5

Frameshift construct map of PGB and PGnefB constructs. β-galactosidase is embedded in a –1 frame of upstream GFP.



Construct	Activity (units/µg protein)
PGnefB	1.25 +/- 0.292
PGB	0.77 +/- 0.208

Figure 4-6

β-galactosidase activity of frameshift constructs. Frameshift constructs were expressed in MDCK cells. PGnefB exhibited minimal β-galactosidase activity in comparison with PGB.

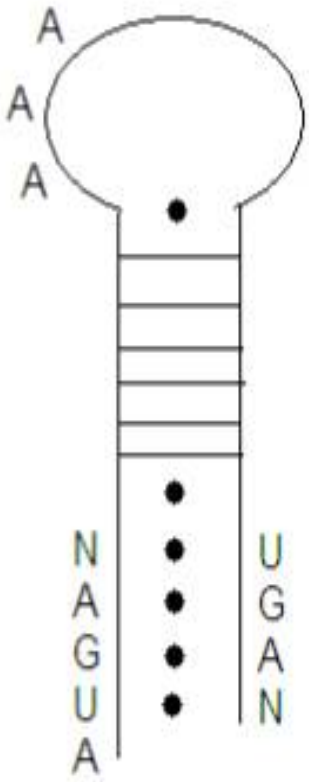


Figure 4-7

Eukaryotic SECIS element. The non Watson-Crick base-pairing shown at the stem, and the AAA RNA motif in the loop are required for SECIS function in selenocysteine incorporation.

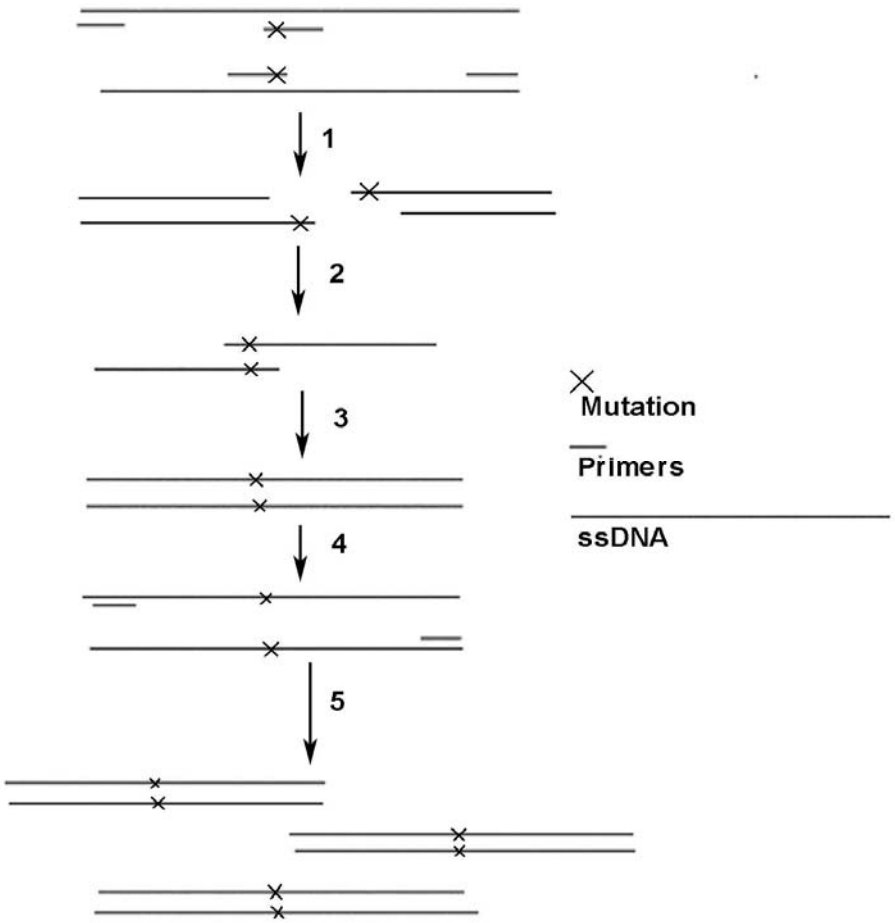


Figure 4-8

Schematic of mutation inducing PCR. 1. PCR amplification of gene of interest with 5', 3' and overlapping primers. 2. PCR amplification of new DNA segments with dNTP alone during which 3. The strands serve as templates of one another and form the new gene fragment that includes the mutation of interest. 4. The 5' and 3' primers are added to the PCR reaction, and 5. The mutated gene is amplified by PCR.

Table 4-2

Primers used for PCR mutation. The premature termination in BH10 was mutated from UGA to UGG with these primers.

Oligos designed for mutation of premature stop in nef gene of BH10

Nhe I-7701 (nef 5' primer)
TTG GCC GCT AGC GTT AGG CAG GGA TAT TCA CCA TTA
Nhe I

Stop1-8528 (overlap primer)
TTC TGC **C**AA TCA GGG AAG TAG CCT TGT GTG
Mutated base

Stop2-8516 (overlap primer)
CTG ATT **G**GC AGA ACT ACA CAC CAG GGC
Mutated base

Sal I-8923 (nef 3' primer)
AAC CGG GTC GAC GGC TCA GAT CTG GTC TAA CCA GA
Sal I

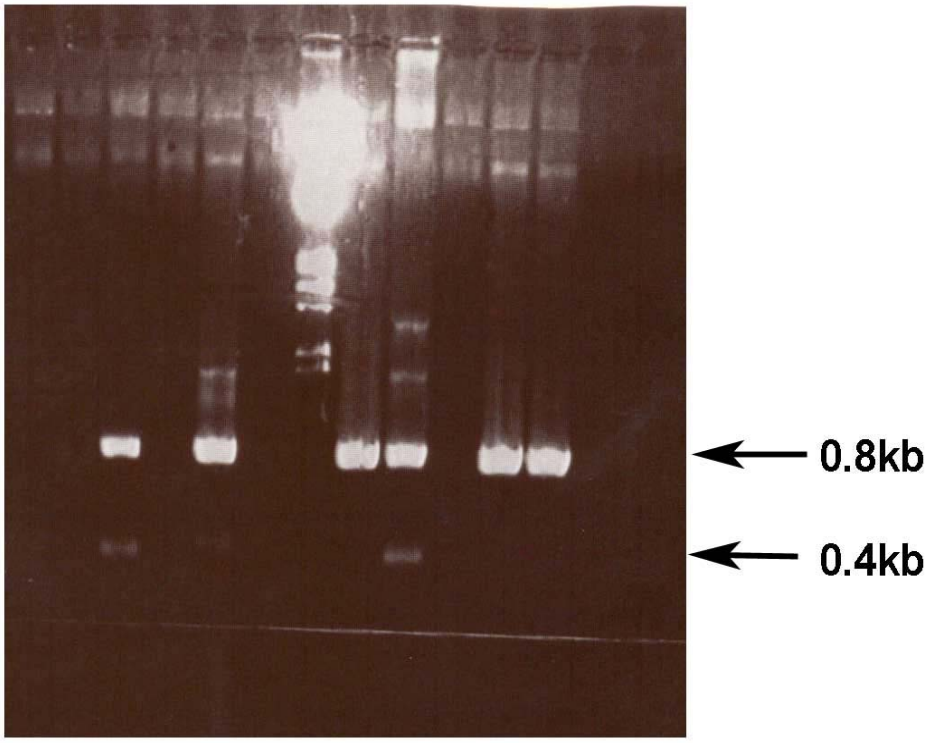


Figure 4-9

Electrophoresis of PCR fragments from mutation inducing PCR. Agarose gel shows electrophoresis of 800bp and 400bp PCR fragments from several PCR reactions at varying temperatures, purified for Nef-wt vector.

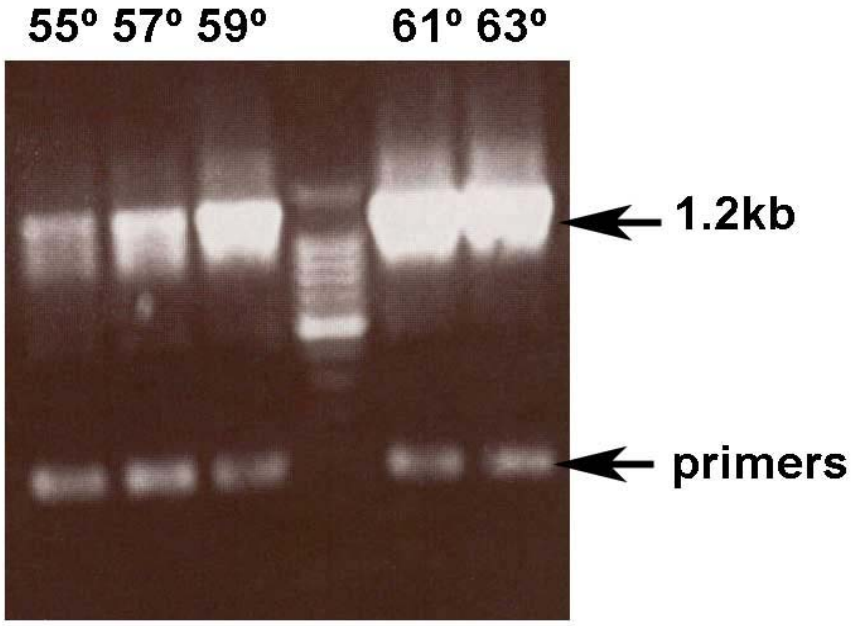


Figure 4-10

Electrophoresis of the Nef gene. The agarose gel shows PCR products of the nef gene with premature stop codon mutated. The gel shows an increase in PCR product with increasing PCR annealing temperature.

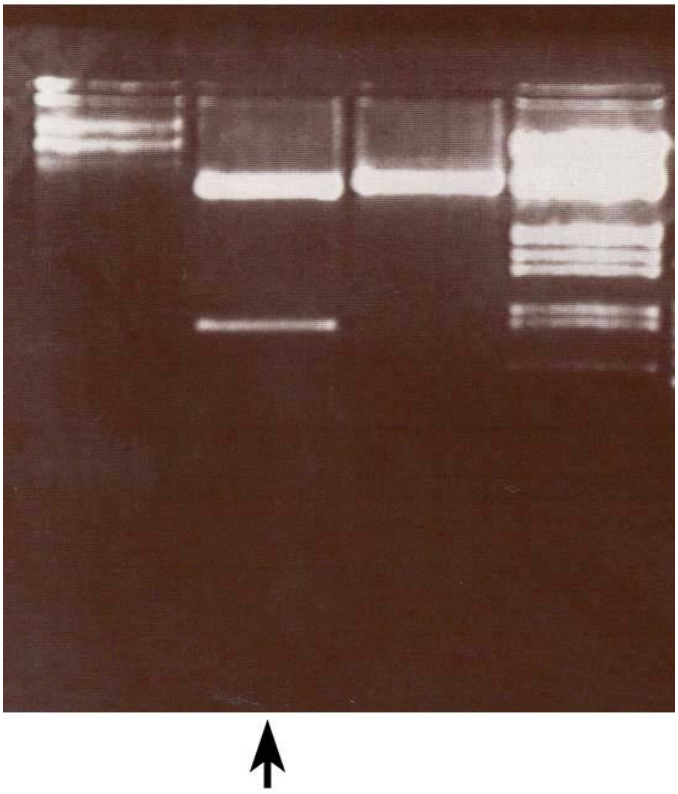


Figure 4-11

Analytical restriction digest of Nef-wt vector. The positive clone was identified by Nhe I, BamH I restriction digest. The 3 clones tested are in lanes 1 – 3, the positive clone is in lane 2, and lane 4 is DNA marker. The nef gene is the lower molecular weight band in lane 2.

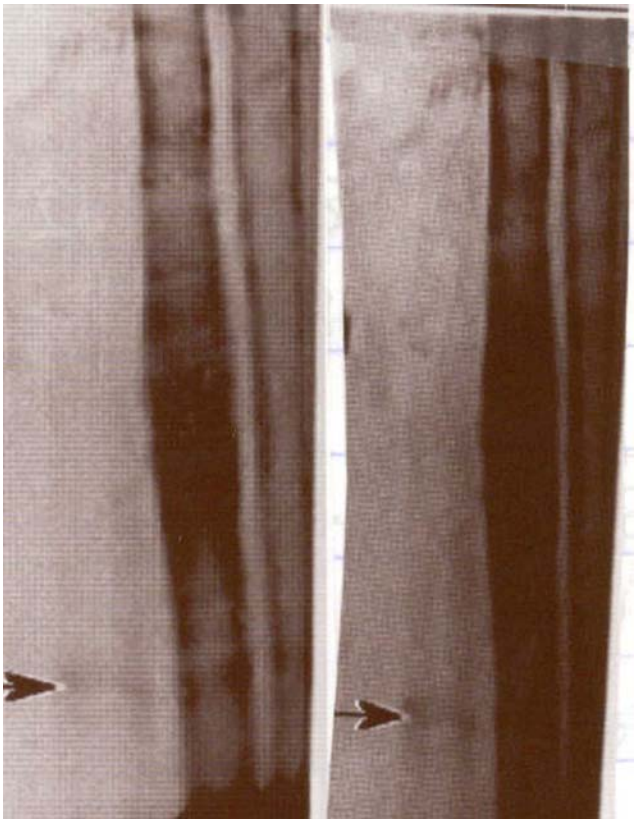


Figure 4-12

Western blot of Nef protein. The Figure shows a western blot of lysates from cells expressing Nef-wt vector. Lane 2 is smear of cell lysate subjected to nef-antibodies on the same membrane.

Table 4-3

Primers designed to construct nef gene mutants by PCR mutation. These primers were involved in inducing a frameshifted mutant conformation in the nef gene for expression of proposed chemokines.

Oligos designed for chemokine nef isoforms.

Nhel 8001 (Nef gene 5' primer)
TTG GCC GCT AGC AAG AAT AGT GCT GTT AGC TTG
NheI

Nef-8433 (Overlap primer set 1)
CTT TTC TTT T**T**A AAA AGT GGC TAA GAT CTA CAG
Added base

Nef-8419 (Overlap primer set 1)
TTT TTA AAA **A**GA AAA GGG GGG ACT GGA AGG GCT
Added base

Nef-8526 (Overlap primer set 2)
CTG CTA ATC **AG**G AAG TAG CCT TGT GTG TGG TAG
Deleted base (guanidine)

Nef-8508 (Overlap primer set 2)
CTA CTT C**CT** GAT TAG CAG AAC TAC ACA CCA
Deleted base (Cytosine)

KpnI-8773 (Nef gene 3' primer)
AAC CGG GGT ACC TCA GCA GTT CTT GAA GTA CT
KpnI

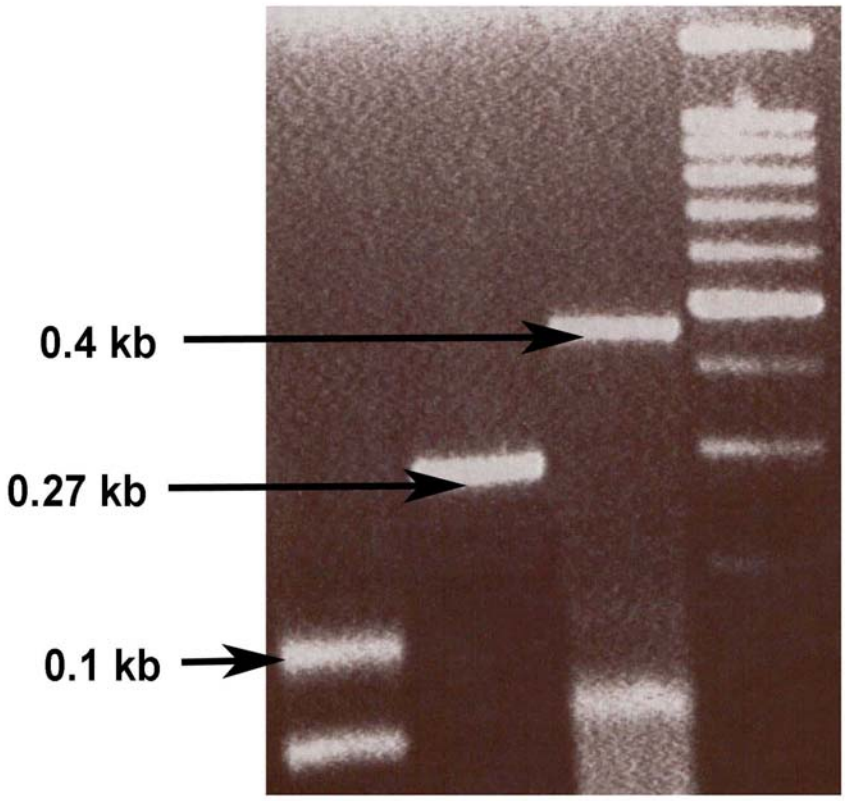


Figure 4-13

Electrophoresis of PCR fragments for nef mutant constructs. The agarose gel shows fragments that were later combined through PCR to form Nef-pos, Nef-neg and Nef-pos-neg vectors.

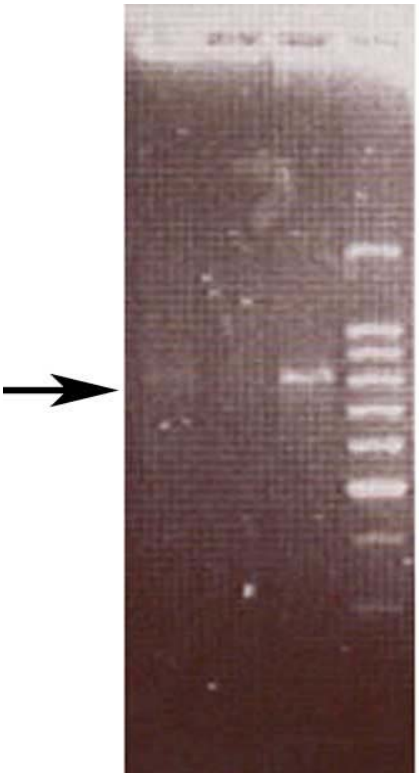


Figure 4-14

PCR fragments of full mutant nef constructs. The arrow indicates a 0.8kb DNA fragment corresponding to the nef gene. These fragments were later incorporated into PG-SECIS vector that includes the deiodinase SECIS element and were confirmed by sequencing.

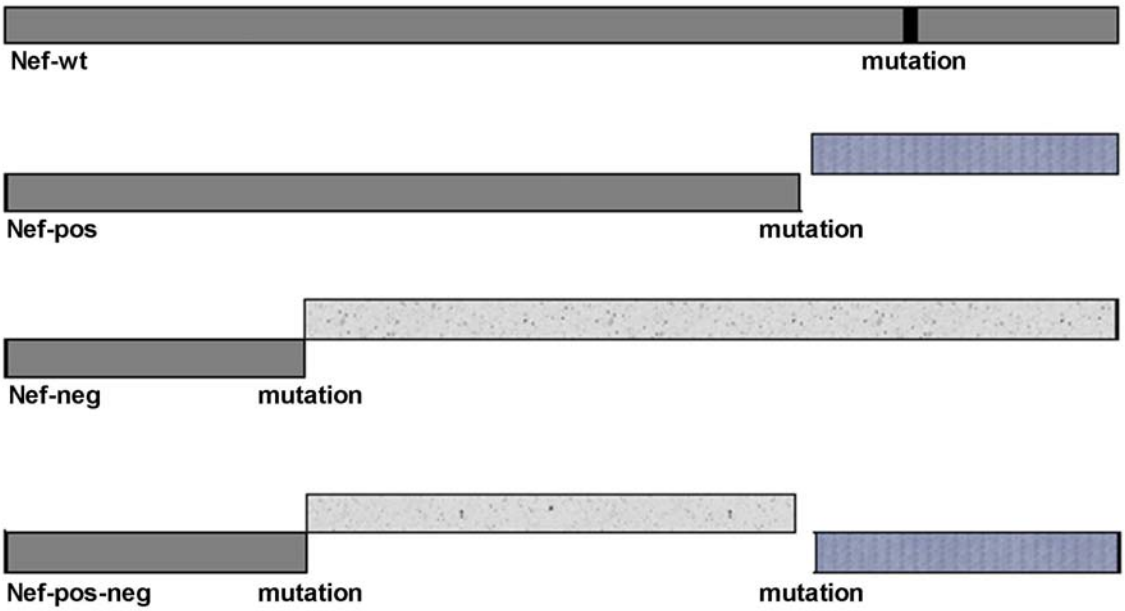
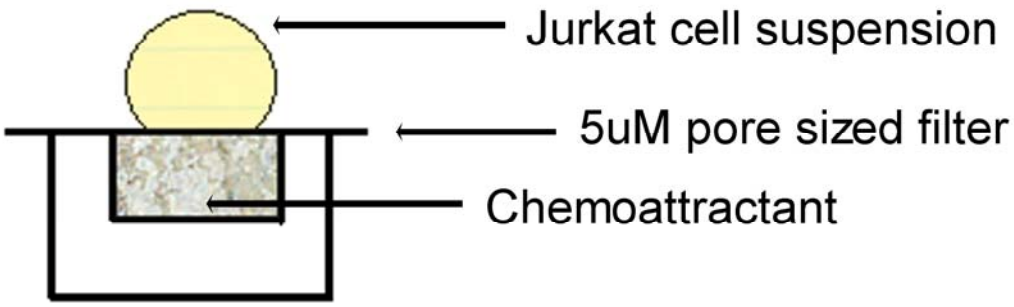


Figure 4-15

Map of recombinant nef gene constructs. The Nef-wt construct has a mutated premature stop codon, Nef-pos is +1 frameshift induced construct, Nef-neg is a –1 frameshift induced construct, and Nef-pos-neg has the homologous chemokine cassette that would be the result of a –1 and +1 frameshift.



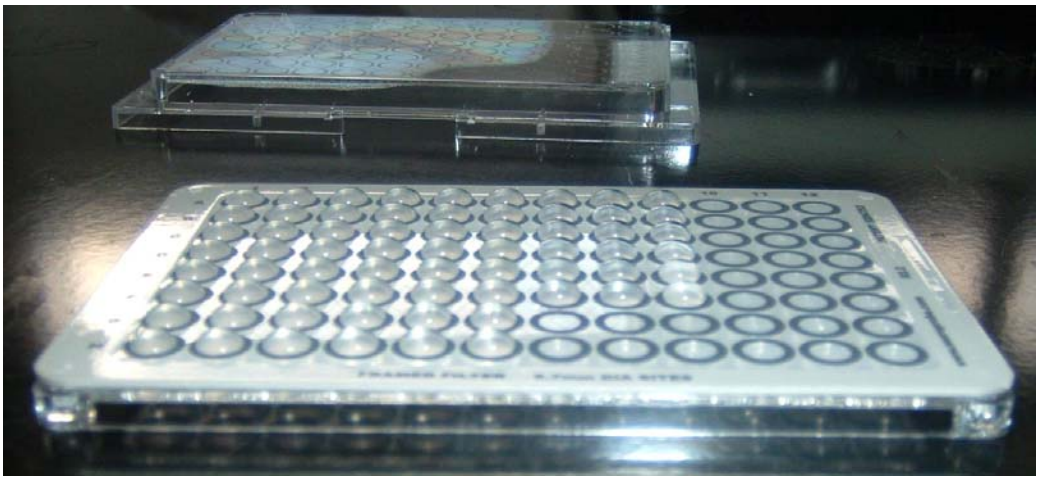


Figure 4-16

Picture of 96-well plate (Neuroprobe) chemotaxis assay. The lower chamber was filled with the chemoattractant (cell lysates with nef mutants) and the Jurkat cells drop is placed on the filter above each well.

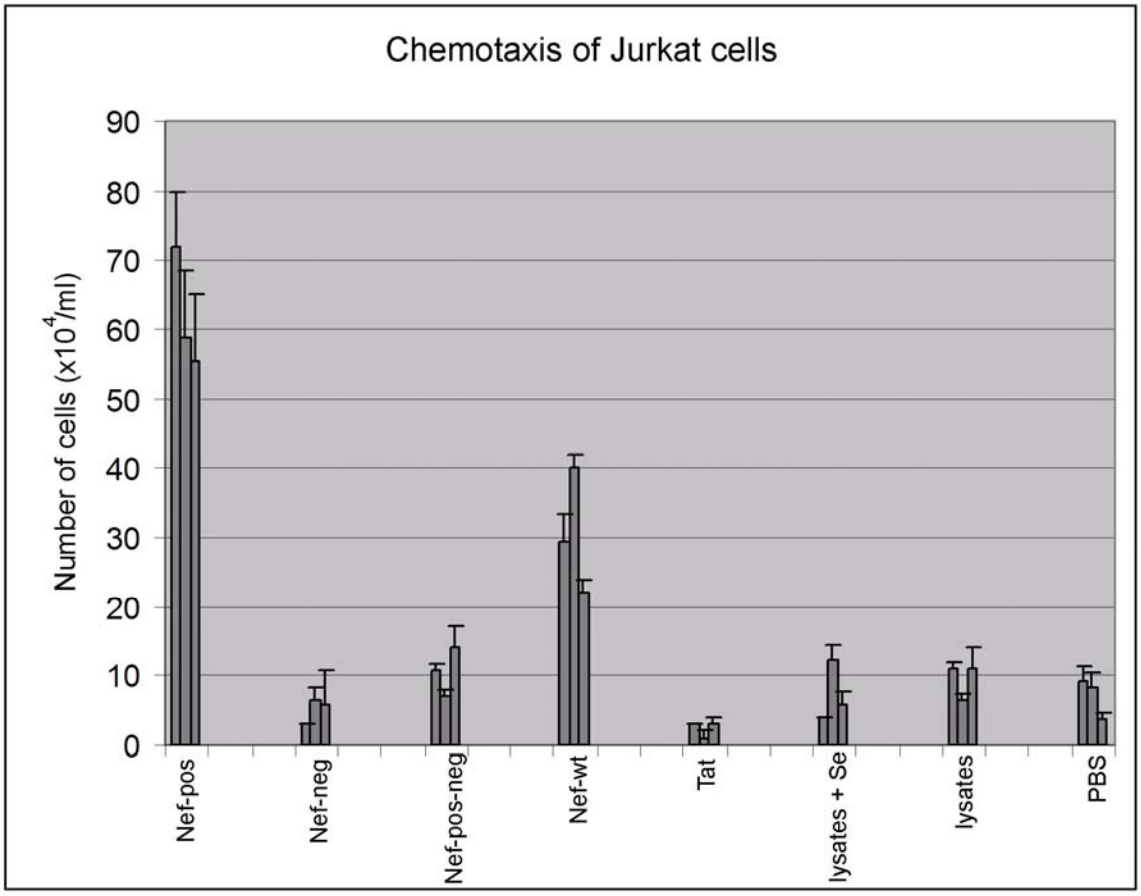


Figure 4-17

Migrated cells in chemotaxis assay. The graph illustrates the number of Jurkat cells that migrated across the filter towards the chemoattractant. Each bar represents the average of three quadrants counted on a hemacytometer from a single well, and each lysate is represented by a triplet of bars. The lysates from Nef-pos construct expressing cells exhibited almost 100% increase in migration in comparison to the regular recombinant nef expressing cells (Nef-wt). The other lysates, showed no significant migration above baseline.

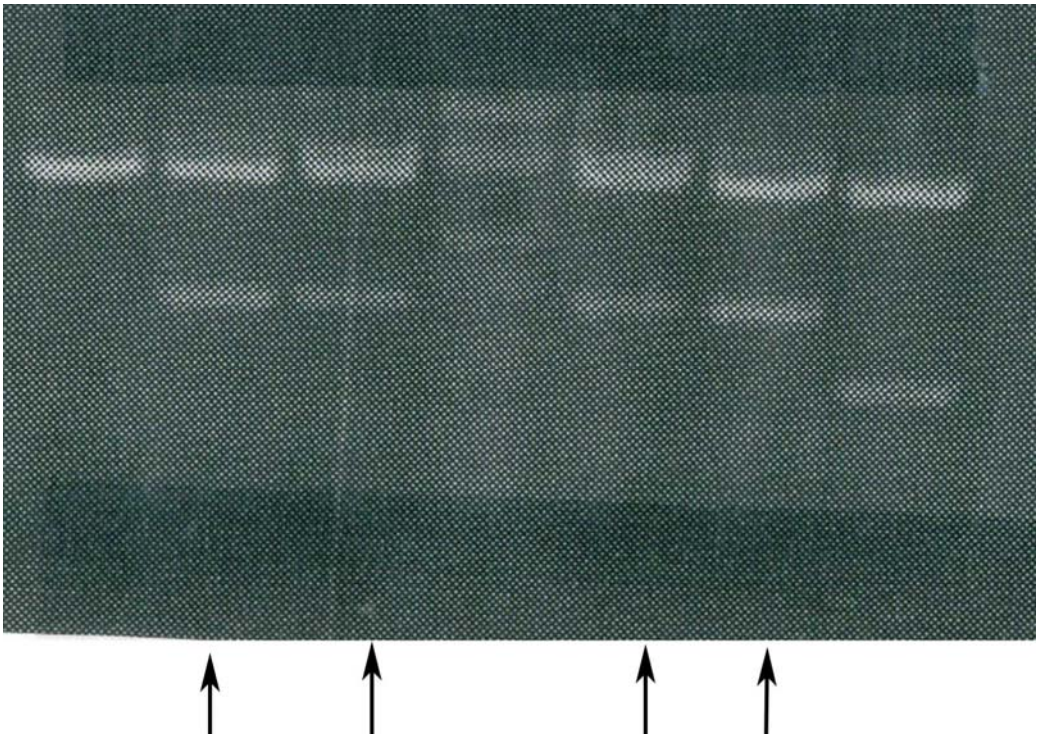


Figure 4-18

Analytical restriction digest of Nef-wt-SECIS. Clones were digested with Nhe I and BamH I for positive identification. Lane 1 is a negative clone, while lanes with arrows were positive clones with 6 being a negative control (PEGFP-CI digested with the same enzymes).

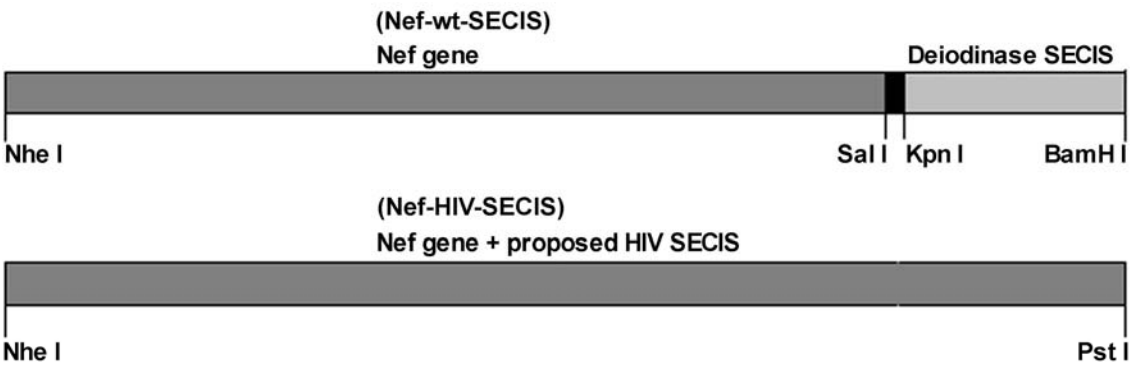


Figure 4-19

Map of constructs for selenocysteine incorporation. These constructs were made to study selenocysteine incorporation at highly conserved UGA codons in at HIV-nef C-terminus, as well as potential thioredoxin activity.

CHAPTER 5

CONCLUSIONS AND DISCUSSION

The mechanisms of HIV survival pose a complicated conundrum to biomedical research. The inefficient HIV polymerase leads to a virus that mutates rapidly, and this evolution of the virus breeds more resistant strains. The complexities of frameshifting, HIV-specific post-translational processing, and other means of HIV protein production that may still remain undiscovered, support the potential for previously undetected HIV proteins that may be expressed at low levels. Hence, an exhaustive analysis of the HIV genome, supported by concentrated approaches in deciphering proteins that may be under strict regulatory control, or exist only in minimal quantities, must be utilized in uncovering hidden mechanisms of HIV survival. The efforts of experiments documented in this thesis, attempted to take such an approach, and though some aspects may not be absolutely indisputable, in the least scenario, new questions have been raised that may assist in complementing the way that HIV research is approached.

The frameshift sequence detected in the protease gene of HIV-1 showed a capability to support ribosome slippage during translation. The –1 frame embedded luciferase gene was expressed in constructs that included the protease frameshift sequence. Furthermore, an increase in cysteine and sodium selenite caused a decrease in frameshifting with the protease frameshift sequence, but the possibility of these compounds playing a direct role in hypermodified base synthesis for tRNA incorporation is still inconclusive and warrants future studies.

The earliest reports of gag-pol frameshifting report efficiencies at approximately 5% *in vivo*, but literature reports of HIV gag-pol frameshifting have been reported as low as 0.2%. The difference in frameshifting efficiency found in different publications, underscores the complexity of the process. The frameshift sequences studied in frameshifting experiments tend to be cloned at different lengths, and in the process, regions required for efficient frameshifting may have been deleted.

Similar results were observed for the envelope frameshift sequence, which constantly displayed frameshifting that was double the known gag-pol frameshift sequence. The envelope frameshift sequence also exhibited a decrease in frameshifting in the presence of sodium selenite and cysteine.

Furthermore, it is of considerable functional interest that the expression of the putative HIV-1 GPx protein encoded by *env-fs* is under the transcriptional regulation of an amino acid, arginine, whose codon was previously predicted to serve as a hungry codon in the *env-fs* frameshift sequence (Taylor et al. 1996). In this case, an 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 (Zhao et al. 2000, Cohen et al. 2004). 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 (Aoyagi et al. 2003). 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 (Cohen et al. 2004). The significance of these observations for HIV pathogenesis in relation to dietary selenium status remains to be elucidated.

The frameshift sequence of HIV nef, which has previously shown capability in inducing frameshifting (Nadimpalli et al. 1997), showed minimal frameshifting in comparison to controls as is indicated by expression of the, –1 frame embedded, β-galactosidase gene. Unfortunately the numerous attempts made at incorporating the nef frameshift sequence into the gfp-luciferase assay vector were unsuccessful.

The constructs that expressed the nef mutants displayed interesting results. The constructs Nef-neg and Nef-pos-neg which we initially expected to express chemoattractant versions of nef, showed no notable increase in cell migration in comparison with untransfected-cell lysates or PBS. The possibility of these constructs expressing a chemokine cannot be totally dismissed, because the purified protein alone may have a different profile. Furthermore, a protein translate represented by the internal methionine in nef that follows the proposed signal peptide needs to be studied to fully dispute the possibility of an embedded nef chemokine gene.

Expression of the nef gene alone (Nef-wt) was capable of inducing chemotaxis. Results that supported the chemoattractant nature of the Nef-pos mutant were also interesting. This nature of the Nef-pos expressed cell lysates, may be further support of reports conferring chemotactic activity to nef, and may indicate that the chemotactic portion of nef must reside in the first half of the nef gene. Furthermore the role of the

high sodium selenite concentration in expression in these cells may play a role that requires future analysis.

References

Aoyagi K et al. (2003) “Inhibition of arginine synthesis by urea: A mechanism for arginine deficiency in renal failure which leads to increased hydroxyl radical generation” Molecular and Cellular Biochemistry 244 (1): 11-15

Cohen I., Boya P., (2004) “Anti-apoptotic of the gluthathione peroxidase homologue encoded by HIV-1” Apoptosis 9:181-182

Taylor, E.W., Ramanathan, C.S., and Nadimpalli, R.G. (1996) “A general method for predicting potential new genes in nucleic acid sequences: application to the human immunodeficiency virus.“ In: Computational Medicine, Public Health and Biotechnology, M. Witten, Ed. World Scientific, Singapore. Ser. Math. Biol. Med. 5: 285-309.

Zhao LJ, Cox AG, Ruzicka JA, et al..(2000)”Molecular modeling and in vitro activity of an HIV-1-encoded glutathione peroxidase” Proc of the Natl Acad of Sci of the United states of America 97 (12): 6356-6361