PATIL, ROHINI, M.S. Cloning and In Vitro Characterization of a Novel Ribosomal Frameshifting Site in the HIV-1 *Gag* Gene. (2011)
Directed by Dr. Ethan Will Taylor. 52 pp.

Most retroviruses and various other organisms have a capability of producing multiple protein products from a single RNA species using overlapping genes and taking advantage of ribosomal frameshifting (Jacks T. , 1990; Taylor, 1994). This phenomenon not only helps retroviruses overcome the coding density limitation imposed by their restricted genome size, but also to attain regulatory benefits from the proteins whose genes are present in the overlapping frames, which otherwise is unachievable by normal read through. The classical example is the retroviral *pol* gene, which in most retroviruses is overlapped by the 3' end of the *gag* gene, and can only be expressed as a result of -1 ribosomal frameshifting. The *pol* gene lacks a start codon, and thus can only be expressed via this mechanism. Retroviral -1 ribosomal frameshift sites (RFS) are characterized by two main structural features, an ideal slippery sequence of the pattern "X XXY YYZ", and a pseudoknot present downstream of the slippery site at a distance of up to 10-12 bases from the site. The triplets in the slippery site are the codons represented in the zero frame, and so when a frameshift occurs, the RNA translation machinery shifts by a single base in the 5' direction, where the triplets are then read in the -1 frame as "XXX YYY Z" (Jacks T. , 1988a). Based on this model, several novel frameshift sites have been identified by Taylor et al. in the

HIV genome. One such site is a theoretically predicted -1 RFS in the HIV-1 *gag* gene (EW, 1996). This site (*gag-fs*), which has not been experimentally validated or studied previously, has become the main focus of this project. A dual reporter assay was employed to study the frameshifting efficiency of the overlapping coding region present in the *gag* gene (*gag-fs*) and the constructs required for this assay were successfully made. Based on the results obtained from the assay an estimate of 24% frameshifting efficiency was observed in the wild type when compared to the mutated type (a 100% readthrough control).

CLONING AND IN VITRO CHARACTERIZATION OF A NOVEL RIBOSOMAL FRAMESHIFTING SITE IN THE HIV-1 *Gag* GENE

By

Rohini Patil

A Thesis Submitted to
the Faculty of The Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Greensboro
2011

Approved By

Dr.Ethan Will Taylor

Committee Chair

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of The Graduate School at The University of North Carolina at Greensboro.

Committee Chair Dr. Ethan.W.Taylor

Committee Members Dr. Greg Raner

Dr. Norman Chiu

Aug 1, 2011
Date of Acceptance by Committee

July 12, 2011
Date of Final Oral Examination

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

Retroviral Complexity

Retroviruses are well known for the mechanisms by which they have evolved from their genome, typically on the order of 10kb, leading into complex viral systems known to date. Studies of the molecular mechanisms of the viral life cycle has elucidated several mechanisms underlying the virus interactions with the host cellular components and how those host cell components interact within each other. When the process of retroviral replication came to light, it was realized that the 'central dogma of molecular biology' (that information can only go from DNA-> RNA-> Protein), was reversed in the case of retroviruses, which have the ability to copy RNA to DNA, via the enzyme reverse transcriptase. Based on the consequences of the infection in vitro and in vivo, retroviruses have been categorized into three subgroups: Oncovirus, lentivirus and Spumavirus. Morphogenically they need three main genes to replicate which are known as 1) the Gag gene, which produces viral structural proteins, 2) the Pol gene, which produces viral enzymatic proteins, and 3) the env gene, which produces envelope glycoproteins. In the proviral integrated DNA these genes are arranged linearly as

Gag-Pol-Env flanked by ‘long terminal repeats’ (LTR), which are required for an efficient transcription (CULLEN, 1991)

Under normal conditions, the essential property of translational regulation is to maintain the correct reading frame throughout the process, although some errors are known to occur at a rate of ~5× 10^{-5} per codon or less (Brierley, 1995)Because of the small size of their genomes, retroviruses are limited in their ability to encode multiple proteins, which could offer regulatory advantages. Given the way these viruses regulate their gene expression, the three possibilities by which they can overcome the size constraints are: RNA editing, RNA splicing, and programmed ribosomal frameshifting (Taylor, 1994) Frameshifting is one of the unique modes of gene expression exploited by retroviruses and other organisms, where a single RNA species can encode more than one protein. This is accomplished by using overlapping genes linked by a ribosomal frameshift site, thus providing a mechanism to produce variant proteins (Jacobs, 2007).

Introduction to Ribosomal Frameshifting

In this mechanism the ribosomal machinery progressing from the zero frame, is redirected into a new frame either in -1 nucleotide 5’ or +1 nucleotide 3’ direction resulting in the synthesis of a fusion protein encoded partially in the zero frame, and partially in the overlapping frame. Because of the inefficiency of frameshifting, this fusion protein is typically a minor product produced in addition

to the protein encoded entirely in the zero frame. Thus, the overall result of this process is somewhat similar to that of alternative RNA splicing (EW, 1996, Jacks, 1988). In some cases, the ribosomes can avoid a stop codon and continue translation, an example can be found in the junction of the gag and pol genes of murine leukemia virus and feline leukemia virus, where both genes are in the same reading frame separated by a stop codon, and where 5-10% of the ribsomes translate the stop codon at the end of the gag gene into a sense codon, which then enters the pol reading frame giving a fusion gag-pol polypeptide (ATKINS, 1991). This mechanism adds to the advantage of ever-evolving retroviruses to increase the protein coding density of their genomes and overcome the limitations of their restricted genomic size.

Structural Features

Structural features required for a -1 frameshift to occur are a ‘slippery site’, a heptamer which ideally is in the form of “X XXY YYZ” (or N NNW WWH as in the figure), where triplets represent codons in the zero reading frame, followed by a short spacer sequence, usually of less than 12 nucleotides, and an RNA pseudoknot structure downstream of the heptamer (Jacobs, 2007) (Figure 1).

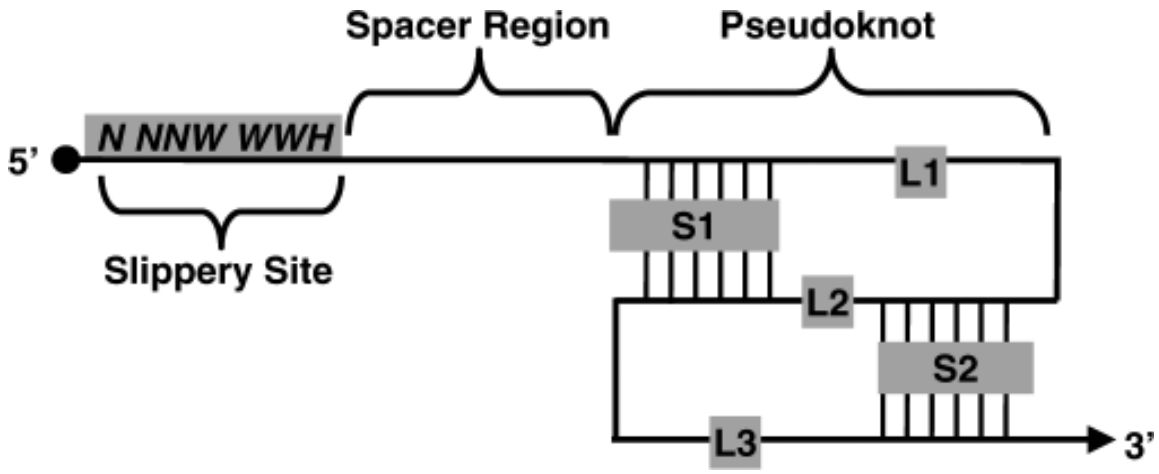


Figure illustrating the structural components of a typical RFS, "N NNW WWH" is an ideal slippery sequence with triplet codons shown in the zero frame, which is followed by a short nucleotide sequence and a pseudoknot structure with two stem structures labeled S1, S2; L1-L3 are single stranded loop regions. (Jacobs et al.).
Figure 1. Components of ribosomal frameshifting

Hypothesis and Aims

On the basis of this established model for -1 frameshifting, several potential novel frameshift sites and associated RNA structures were identified theoretically in HIV by Taylor et al (EW, 1996). One of these site is located in the Gag gene p17, overlapping the capsid protein coding region p24. The possibility that this is a functional -1 RFS site, and the potential function of the overlapping protein that it encodes, have never been explored experimentally. In theory, the predicted gag-fs sequence should be sufficient to produce -1 frameshifting at measurable efficiency, which in intact virus would permit the expression of a previously unidentified variant of the HIV gag protein.

The aim of this project is to clone the putative gag-fs FRS into a dual reporter vector along with two other variants of the same sequence as control constructs to assess the efficiency of the predicted frameshift site. The constructs will then be transfected into a mammalian cell line for their expression, from which their frameshifting efficiency will be measured by performing a dual – luciferase reporter gene assay.

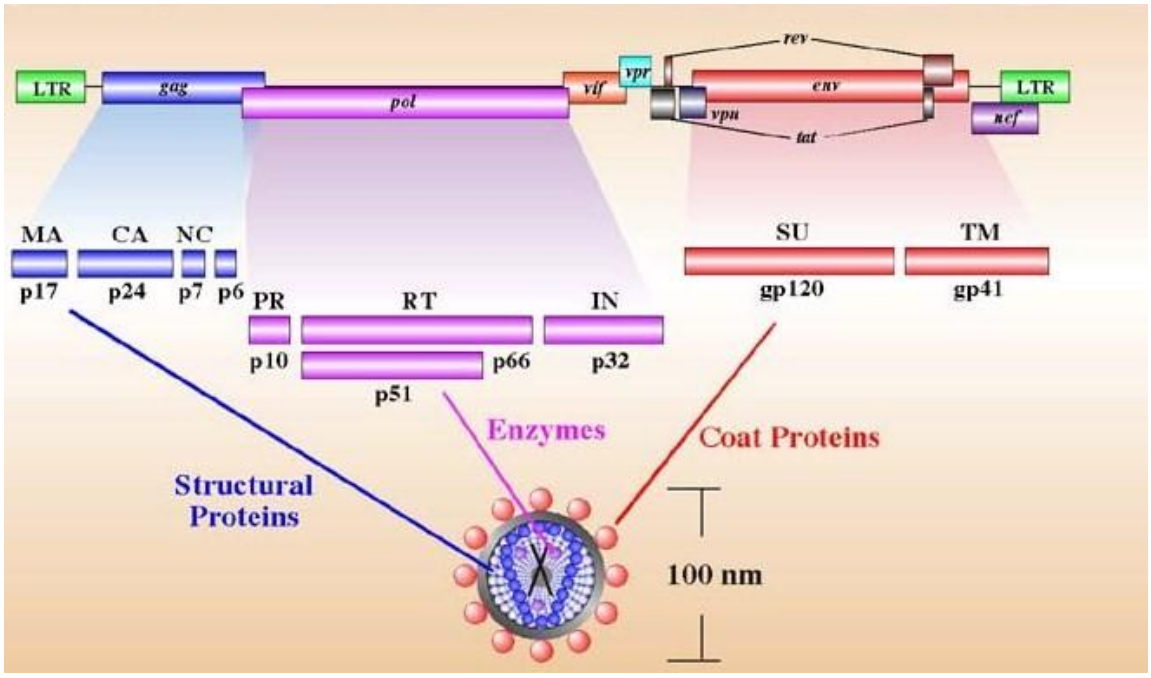
CHAPTER II

LITERATURE REVIEW

Ribosomal Frameshifting in Gag-Pol of HIV-1

In most of the retroviruses the Gag gene which codes for the structural proteins has an ORF at the 5' end where as the Pol gene which codes for the enzymatic proteins has an ORF that begins just before the 3'end of gag, but usually not in frame with the gag ORF. As a result, these retroviruses employ ribosomal frameshifting to translate the pol gene and generate the viral enzymes, required for the viral replication. Hence, the ribosomal frameshifting occurs where the genes are overlapped. The classical example for the ribosomal frameshift sites in most of the retroviruses fall at the junction of Gag-Pol, Gag-Pro, Gag-Pro-Pol genes (Jacks T. , 1990) . During the ribosomal frameshifting the RNA machinery is directed by the frameshift elements to shift the reading frame from 0 to -1 so that the hidden Pol gene with a different ORF could be expressed, which cannot be achieved by normal translation, because the pol gene lacks a functional 5' start codon. Gag is produced as a precursor protein commonly designated as p53, which yields three essential structural proteins: Matrix, Capsid, and Nucleocapsid. These proteins are produced upon the cleavage of the precursor p53 which is brought

about by the ‘viral proteases’, an enzymatic protein produced from pol gene. Other than protease, the Pol gene also yields reverse transcriptase, ribonuclease H and integrase which are essential enzymes required for the viral replication. The only way these enzymatic proteins can be expressed is through a ribosomal frameshift where the virus programs a shift in the read frame right before it enters into the pol region which is then expressed in fusion with gag, giving a combined polypeptide p160. This phenomenon not only determines the viral morphogenesis but also its propagation and infectivity (Ian Brierley, 2006).



Gag is processed as Matrix (MA), Capsid(CA) , Nucleocapsid (NC) and a small spacer protein p6, *Pol* produces Proteases (PR), Reverse transcriptase (RT), Integrase(IN), *Env* produces regulatory proteins gp120, gp41. As seen in the figure *gag* is overlaps the *pol* hence, the proteins enzymatic proteins formed are as a result of Ribosomal Frameshift. (Stewart, 2000).

Figure 2. Genomic organization in HIV

Occurrence of Frameshifting

It is important to emphasize that the ribosomal frameshifting process is inefficient at best; it generally occurs about 5-10% of the time in HIV-1, but the efficiency can vary widely between different viruses depending upon the structural details of the -1 RFS (Jacks T , 1988). It also depends on the ratio of relative amounts of structural or enzymatic components needed which are highly regulated and critical for the viral propagation (Dinman, 1992). Any alterations made to the efficiency of this frameshifting, which influences the ratio of the structural proteins to the enzymatic proteins needed, is also known to effect the viral assembly and its propagation; this was demonstrated using two endogenous retroviruses of the yeast *Saccharomyces cerevisiae*, where the ratio of gag to gag-pol protein ratio differed because of the induced errors in the frameshifting efficiency, resulting in reduced viral production which was caused by the incorrect ratio of the required proteins (Dinman, 1992). Ribosomal frameshifiting has also been encountered in eukaryotic positive single stranded RNA viruses, dsRNA yeast vruses, Plant RNA viruses and Bacteriophages (cordon et al, 1991; Levin et al, 1993).

Mutational studies done on structural elements

The three structural elements required as described earlier for frameshifting also determine its efficiency. The slippery site alone stimulates the process by 1%, whereas the downstream pseudoknot structure has 30- 50% enhancement on the ribosomal frameshifting (Giedroc, 2000). Various alternations made to the slippery site by mutagenesis have been shown to severely diminish the viral infectivity due to reduced levels of frameshifting (Biswas, 2004; Brierley I. , 1992). A mutational study was done on slippery site of Gag-Pol overlapping region of HIV type- 1, elucidating the requirement of the slippery site sequence 'UUUUUUA', explaining its unvarying sequence consistency in all other HIV-1 variants. Alterations made to the slippery site affected the amount of infectious viral particles produced (Biswas, 2004).

Although a slippery site is an essential element for the frameshifting, it does not serve the purpose completely on its own; for the process to occur efficiently, it needs a downstream RNA pseudoknot structure which is spaced at least 5 to 12 nucleotides from the slippery site. The requirement for the precise distance between the slippery sequence and the pseudoknot was well demonstrated by Brierley I. D., 1989 and Kollmus, 1994. The very first pseudoknot involved in frameshifting was observed in avian coroniavirus IBV (Brierley I. D., 1989). Several studies made on characterizing the RNA structure have developed many

models but mostly agree on a simple stem loop structure proposed by jacks et al, while others have proposed various interactions as an added extension to the stem-loop structure (Ian Brierley, 2006). It is the stability of the pesudoknot which determines the frameshifting process; the more stable the structure is, thenmore efficient the frameshifting will be. There have been various speculations about the mechanism of frameshifting, which led to the proposal of three models.

Models of Frameshifting (Ian Brierley, 2006)

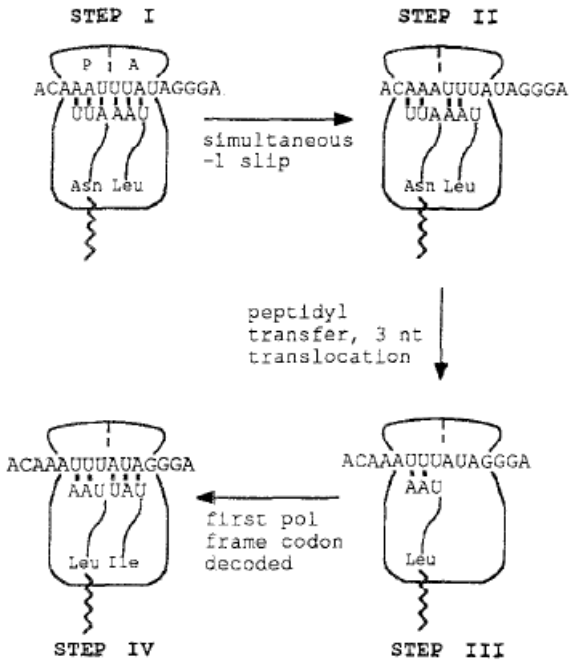
Binding Model

In this first model it was proposed that there may be some proteins which bind to the stimulatory RNA which in turn regulate the frameshifting process (Jacks T. , 1988a); however, there has been no evidence of such interactions so far**.**

Pausing Model

In this second model, it was proposed that when the ribosome encounters the pseudoknot, it is forced to pause at the slippery site; thus, this pausing of the ribosomes at the slippery site is directly correlated to the stability of the RNA pseudoknot structure, as the longer it pauses, the greater the probability of frameshifting would be, as this gives the ribsomes an ample time to conveniently adjust into the -1 frame (Jacks T. , 1988a).

This basis for the role of these structual elements in -1 frameshifting model was originally developed by Varmus and coworkers 3 (Jack, T; Madhani, H D; Masiarz, F R; Varmus, H E, 1988a). As outlined above, in their model (Slippage model of frameshifting in Gag-Pol region of Rous sarcome virus), the translating RNA machinery gets obstructed by the downstream pseudoknot just as the slippery sequence occupies the ribosome active site. During this pause, the mRNA shifts back a single base in 5' direction from the zero frame while the two tRNA molecules are bound to the ribosome, after which the mRNA can keep reading in the -1 frame (Figure2). A similar mechanism has been demonstrated in the gag-pol region of human immunodeficiency virus (Jacks et al 1988) and the gag-pro region of mouse mammary tumor virus (jacks 1987).



Step I of Figure 3, the two tRNA's, $tRNA_{Asn}$ and $tRNA_{Leu}$ are bound to codons AAU and UUA which are present in the Gag reading frame in the ribosomal P and A site respectively. After the -1 slippage (step II) of the mRNA, the tRNA's have been shifted back by one nucleotide relative 5 to the mRNA, translocating them into the Pol reading frame, where now, despite a mismatch in the 3rd or "wobble" position of the codons, the tRNA anticodons pair with AAA and UUU respectively (step III). Then, following peptidyl transfer, the next Pol gene codon AUA is brought into frame (step IV) where it is decoded, and reading proceeds in the -1 (Pol gene) frame (figure from Jacks et al)

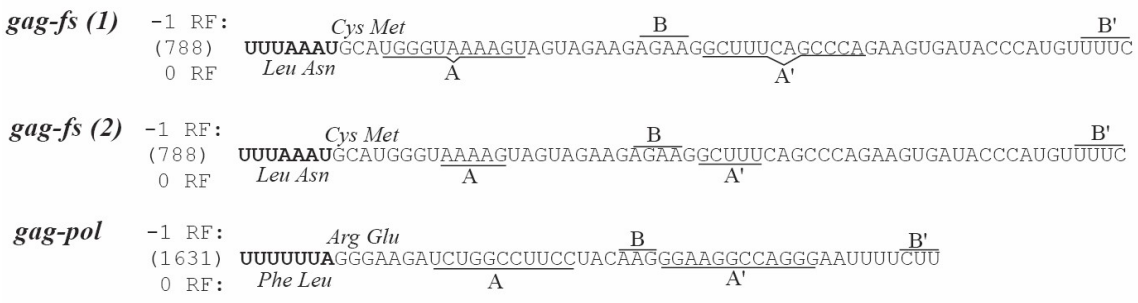
Figure 3. Slippage model of frameshifting

Unwinding Model

In this third model, which is really an extension of the second model, it was speculated that there were some stimulatory structural features to the pseudoknot which stabilizes the secondary structure, and it takes longer than usual for the ribosomal associated RNA helicases to unwind as it comes across those structural features (Yusupova, 2001; Takyar, 2005), for example a triplex structure formed by the interactions between stem and loops structures, such as, a kink found between stem 1 and 2 found in the pseudoknot structure of MMT virus (Shen, 1995) or a frameshifting determinant inside or adjacent to the stem 2 in the pseudoknot structure of SARS Co-Virus (Ian Brierley, 2006), where both serve as part of an important frameshifting signal. The resistance of these signaling structures towards the unwinding by the helicases, allows the ribosomes to pause while it gets ready to unwind the pseudoknot.

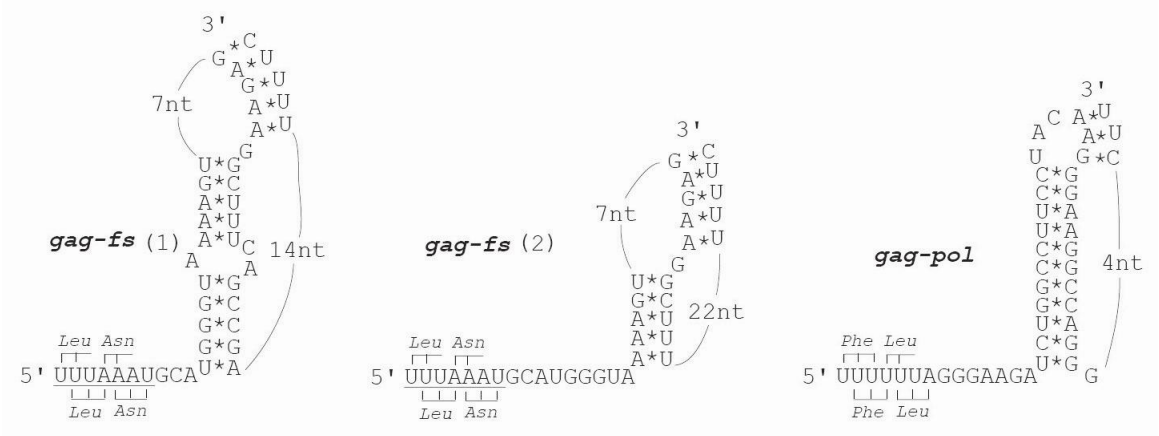
Frameshifting Site identified previously

On the basis of this established model for -1 frameshifting, several potential novel frameshift sites and associated RNA structures were identified theoretically in HIV by Taylor et al. The frameshift site of interest to our laboratory falls within the N-terminal side of the capsid protein. This RFS site was first predicted by Taylor et al. A number of primate retroviruses have an ideal or near-ideal -1 shift sequence followed by an RNA structure in the gag coding region, about 800 bases upstream from the known gag-pol RFS site, which is also shown in the following figures. In HIV-1, the heptameric shift sequence for this novel gag-fs is ideal (U UUA AAU). A potential PK with an asymmetric bulge in its 5' stem begins 3 nucleotides after the shift site. The predicted amino acid sequence at the shift site is LN/CM, and the hypothetical protein encoded in the -1 frame potentially extends another 51 amino acids. The -1 shift sequence and PK are highly conserved in HIV-1 subtype B, and fairly well conserved in other HIV-1 subtypes except subtype A, where only about half of the sequences have these features intact. The PK structure has two alternatives, one (shown as gag-fs 2) is predicted by a semi-manual method, which contains a bulge in the first helical stem; the other (shown as gag-fs 1) is as predicted by a computer program. The latter is simply a substructure variant of the previously predicted PK.



The known gag-pol frameshift site is shown for comparison. A:A' and B:B' correspond respectively to the 5' and 3' stems of a predicted pseudoknot. (See fig. 5 for a 2-D representation). The predicted protein sequence spanning the frameshift site is shown at left (LeuAsnCysMet for the gag-fs site). Two possible variants for the gag-fs pseudoknot are shown.

Figure 4. Schematic representation of a novel -1 RFS



RNA secondary structure representation of two possible variants of the predicted gag-fs site, compared to the known gag-pol RFS structure. The gag-fs site features an ideal slippery sequence (UUUAAAU) and a pseudoknot structure. The amino acids below the sequence represent the P and A-site tRNAs for the slippery site codons, as they are read in the zero frame; they are also shown above the sequence in their positions after the one-base shift to the left, now in the -1 frame. The third figure is the well known gag-pol RFS site of HIV.

Figure 5. RNA secondary structures predicted in *Gag*-fs site

The role of the HIV *Gag* gene

Gag, as described previously, is both one of the raw materials and also an architect of the assembly of the structural components of the virus particles and their release in later infectious stages (as reviewed in Göttlinger, 2001) . This process can be described in three steps, starting with the binding step when the *Gag* precursor(Pr55) is processed, the matrix targets the plasma membrane with the help of a myristic acid moiety present on the N-terminal side. This region, which is hydrophobic and highly conserved, is believed to bind the gag precursor to the plasma membrane, and this is achieved by the association of N-terminal proximal basic residues with the anionic phospholipids of the plasma membrane. Once the *Gag*-precursor is anchored to the plasma membrane it gradually starts budding out; during this process the maturation occurs, the polyprotein precursor Pr55 is proteolytically cleaved into matrix, capsid, nucleocapsid and P6 accordingly, along with two spacer proteins SP1, SP26. Once the protein particles are formed the RNA genome is assembled in order by the Capsid(CA) protein which contains two main domains N-terminal domain which helps in forming a mature core and the C-terminal domain that helps in particle assembly and also in the formation of the core. Capsid protein is believed to have major binding sites for its interaction with the cellular components of the plasma membrane, the N-terminal CA has an extended loop where it can bind to the cyclophilin A which

gets incorporated and helps in the viral assembly. The C-terminal side of the capsid, which contains four helices, helps in dimerization of the capsid proteins. In addition to the viral assembly and core formation, the C-terminal capsid extends to P6 protein which separates capsid from the nucleocapsid region. It is believed that the extended capsid C-terminal end and P6 both form a helical structure and this junction helps in budding off of the virus particle. The RNA is packaged along with nucleocapsid at the center which is surrounded by conical capsid protein, while matrix forms the outermost layer remains in contact with the membrane.

Expected results and significance

Two other novel frameshift sites in HIV-1 that were identified by similar methods were found to be functional in vitro, despite having non-ideal slippery sites, and to encode functional protein variants (Lijun Zhao and Taylor, 2000). Given that the gag-fs site has the ideal pattern established for retroviral frameshifting, we expect that the gag-fs site will be active, with an efficiency likely to be at least a few percent. The probability of this outcome is also supported by the presence in HIV infected cells of a 22Kd band labeled by gag antibodies in a few previous studies, which did not correspond to known gag cleavage products, and corresponds exactly to the expected mass of the fusion protein produced by translation of gag followed by a -1 frameshift at the gag-fs site (Buchacher et al)

CHAPTER III

METHODS AND MATERIALS

Construction of Plasmids

PNL4-3 was used in the cloning the desired gene of interest which is present in the capsid region of Gag gene of HIV. Three different versions of the plasmid were made which are described as in the following

Sequence of HIV gene of Interest

"GGTCAGCCAAAATTACCCTATAGTGCAGAACATCCAGGGGCAAATGGTACA
TCAGGCCATATCACCTAGAACTTTAAATGCATGGGTAAAAGTAGTAGAAGAG
AAGGCTTTCAGCCCAGAAGTGATACCCATGTTTTCAGCATTATCAGAAGGAG
CCACCCCACAAGATTTAAACACCATGCTAAACACAGTGGGGGGACATCAAGC
AGCCATGCAAAT"

Gag-fs-WT (Wild Type)

The wild type was amplified using primers Gag-716Fwd (5'-AGATCTGG TCAGCCAAAAT TAC C -3') and Gag-934R (5'-CTCGAGATTTGCATGGCT GCTTG-3') adding Bgl II and Xho1 restriction sites to the primers respectively; This construct was made such that the luciferase gene is in -1 frame with the betagalactosidase, this way there's always a signal obtained from luciferase activity whenever frameshifting occurs in the wild type.

Gag-fs-MUT(+) the mutant construct

This construct served as a positive control where an extra base 'T' was inserted right after the slippery site 'TTTAAA' to set the luciferase gene into zero frame with the betagalactosidase gene giving a 100% read through of both the reporter genes.The insertion was achieved by a 2-step PCR procedure . In the first step, a forward primer Gag-736Fwd(5'-AGAT CTGGTCAGCCAAAAT TACC-3'*)* and a reverse primer Gag-*fs*-R810 (5'-CTACTTTTACCCATGCAATTTA AAG-3'*)* were used adding Bgl II restriction site to the forward primer, to obtain a 100Bp product and another set of primers were used subsequently in the same step; a forward primer Gag-Fs Fwd (5'-TAGAACTTT AAATTGCATG GGTAAAA GTA-3')and a reverse primer Gag-934 R (5'-CTCGAGATTTGCATGGCTGCTTG-3') , in obtaining a 150 Bp product where Xho1 restriction site was added to the reverse primer . The PCR products were then gel purified and used as templates in the second step of the PCR using primers Gag-736 Fwd and Gag-934R to get a successful insertion.

Gag-fs-UGA

This construct, which served as a negative control was amplified using primers Gag-736Fwd (5'-AGATCTGGTCAGCCAAAATTACC-3') and Gag-R UGA (5'-CTCGAGGGCTCATTCTGATAATGCTGA-3') where a base was mutated; the purpose of this mutation was to introduce an in frame stop codon at the 3' end of

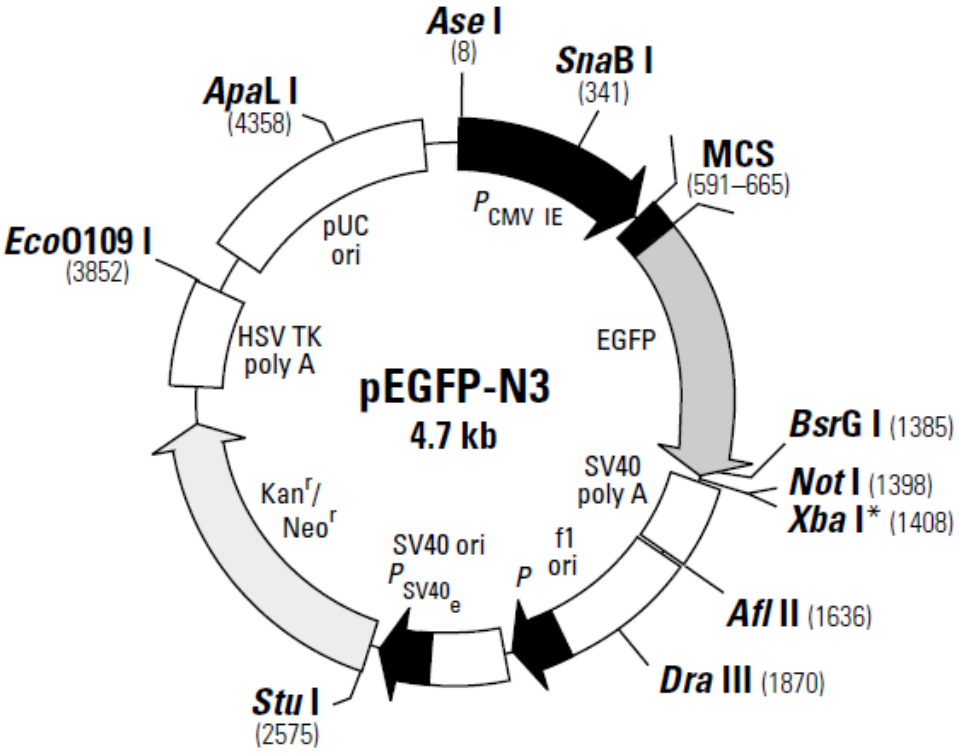
the insert so that this construct only reads beta-gal. The results obtained from this control plasmid in comparison to the wild type and mutant eliminates back ground light signals thus normalizing the values for the wild type and the mutated type.

Methods

Standard protocols were followed for all the molecular biology techniques, the resultant PCR products from the three PCR reaction (Figs 8&9) were purified (using Sigma PCR cleanup kit), then analyzed on 2% agarose gel. An overnight ligation reaction of the Purified DNA inserts was set up with pGEM-T vector (from Promega) which provided a better efficiency for ligation of the PCR products. The vector also has a Beta-lactamase gene (produces blue colonies) , provides a convenient way for better selection of the right recombinant plasmids (white colonies). The ligated constructs where then transformed into JM109 cells using a standard heat-shock method (promega), transformed cells were then plated on agar plates with ampiclin/Iptg/X-gal. Colonies were then harvested from blu/white screening followed by an incubation period of six to several hours in LB media with ampicilin. Successfully grown plasmid colonies were then purified (promega plasmid purification kit) followed by a double restriction digestion with BglII and xho1(Figs.10,11,12). The resultant inserts of the appropriate sizes which were obtained on the gel were then gel purified and ligated into another expression vector LacZ-LuC following the above transformation procedures which were then

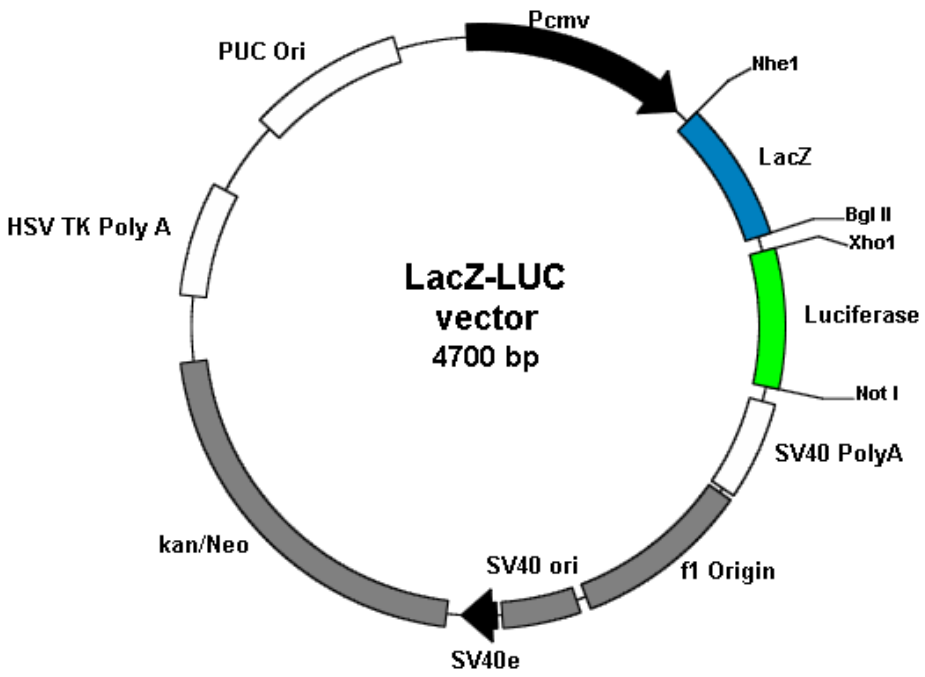
digested with the restriction enzymes for the analysis of successful ligation (Figs 13,14,15).

LacZ-Luc is a dual expression vector with two reporter genes Beta-Galactosidase and Luciferase, with neomycyin/kanamycin resistance. These reporter genes were built on pEGFP-N3 vector; the Lac Z gene for the Beta-Galactosidase was cloned from the pHIV-LACZ and inserted upstream of the multi cloning site replacing EGFP coding region between Nhe1 and Bgl II restriction sites; luciferase which was cloned from pGL-1 Vector (promega) was inserted downstream of the multi-cloning site between Xho1 and Not I sites.



The back bone of pEGFP on which two reporter genes were engineered Upstream and downstream of the multicloning site (Figure from Clontech technical sheet)

Figure 6. Schematic representation of pEGFP vector



The general built of the pEGFP vector modified into LacZ-Luc vector when the LacZ and Luciferase reporter genes were added

Figure 7. Schematic representation of Dual reporter vector

Cell Transfection with lipofectamine

HEK/T17 cells were used for our transfection procedures which were grown in Dubelco minimal essential media with 10% FBS, 4mM L-Glutamine, 1mM Sodium Pyruvate. Approximately 35,000 cells were seeded per well in a 96 well plate with a 0.3cm^2 surface area, in six replicates and were incubated till the cells reached 90% confluency. Lipofectamine 2000 (invitrogen) was used in transfecting the cells with ~0.2µg of appropriate plasmid DNA. Six replicates of Gag-fs WT, Gag-fs MUT(+), Gag-fs-UGA were transiently transfected with HEK

cells along with mock transfection. The set up was incubated for a period of 48 hrs at 37 degree C with 5% CO_2.

Dual Reporter Assay

NovaBright™ β-galactosidase and firefly luciferase dual enzyme reporter gene chemiluminescent detection system (Invitrogen), was used for all the assays. This kit contains substrates luciferin and galacton for the detection of luciferase and beta-galactosidase activities respectively in a single reaction well. After 48 hrs of incubation, the cells were lysed using a lysis buffer (invitrogen) with an incubation period of 10mins and the cell lysate were assayed for the respective reporter activity. In all the assays the reporter activities were quantitated in a Synergy HT Multi-Mode microplate reader. This assay involves the addition of an assay buffer A, which contains all the components to enhance the luciferase reaction, then buffer B (which contains luciferin and galacton plus substrates) was added to initiate the the luciferin/luciferase reaction which produces a luminescent signal which decays in a minute so therefore the luminescence for the luciferase is read soon after the injection of buffer B. After 30-60min of incubation, the luciferase reaction is quenched by the addition of a light emission accelerator sapphire II (which increases the pH and enhances the light intensity) and the luminescent signal from Galacton plus/ Betagal reaction is measured which decays with a half life of 180mins. During this phase the luciferase light signal are very

low due to the rapid decay of luciferase signal doesnot interfere with the betagal signals (Martin., 1996).

The recoding efficiency or the Frameshifting efficiency was calculated as follows

$$\text{Frameshifting Efficiency } \% = \left\{ \frac{(\text{Luciferase}_{\text{test}} / \beta\text{-Galactosidase}_{\text{test}})}{(\text{Luciferase}_{\text{Control}} / \beta\text{-Galactosidase}_{\text{Control}})} \right\} \text{ X } 100$$

The above formula was employed in converting the relative luminescent units of the reporter genes into % frameshifting efficiency.

The ratios of luciferase and galactosidase of the test and mutated constructs were normalized by subtracting the intensities (RLU) of luciferase/Betgal of the negative control (*gag*-fs-UGA) from the test and the mutated (positive control) constructs respectively.

Assay with pure Enzymes

The assay was performed in a similar manner as mentioned above to generate a standard curve for luciferase and betagal activities which also corrects for the non-linearity encountered in normal cellular assay. A stock solution (1mg/ml) of pure enzymes were prepared in 0.1M sodium phosphate pH 7.0, 01% BSA and a series of dilution of the stock were prepared in the cell culture medium

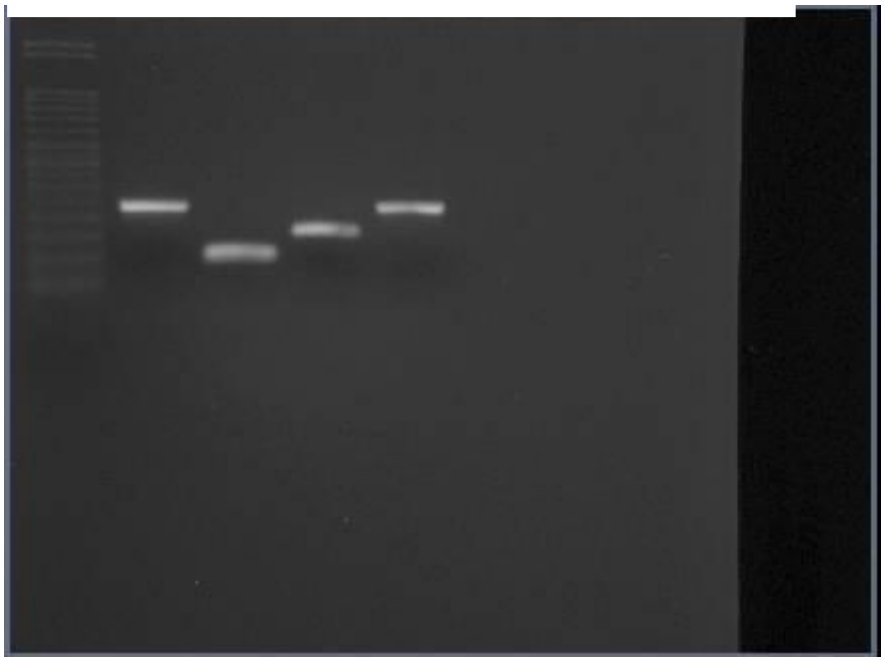
(DMEM containing 10%FBS, 4mM L-Glutamine, 1mM Sodium Pyruvate) ranging from 0.01ng to 15ng. The experiment was carried out in triplicates using the same protocol as mentioned above

CHAPTER IV

RESULTS AND DISCUSSION

(50bp- 3000bp ladder was used in all the gels)

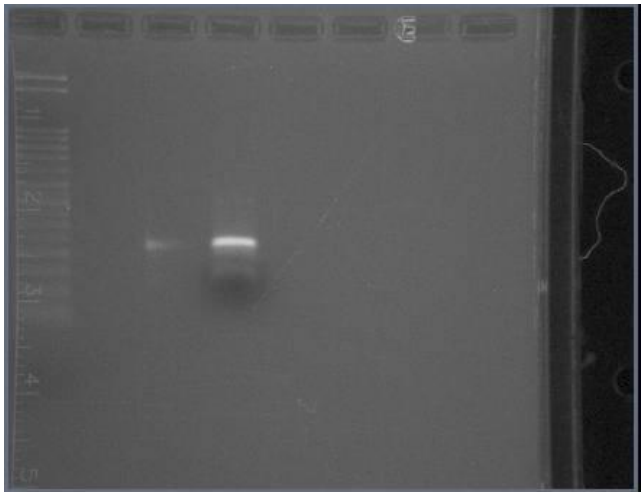
1 2 3 4 5 <--- Lanes



Lane 1 showing the ladder (50bp- 3kbp) used, **lane 2** represents the pcr band for the Wild type construct *Gag*-fs-WT at about 250bps; **Lane 3 and 4** are the pcr bands of Mutated construct *Gag*-fs-M(+) from the 1st step of the PCR; **lane 3** is about100bps, **lane 4** is about 200bps; **Lane 5** is the pcr band representing the negative construct *Gag*-UGA which is about 250bps.

Figure 8. Gel results of pcr products

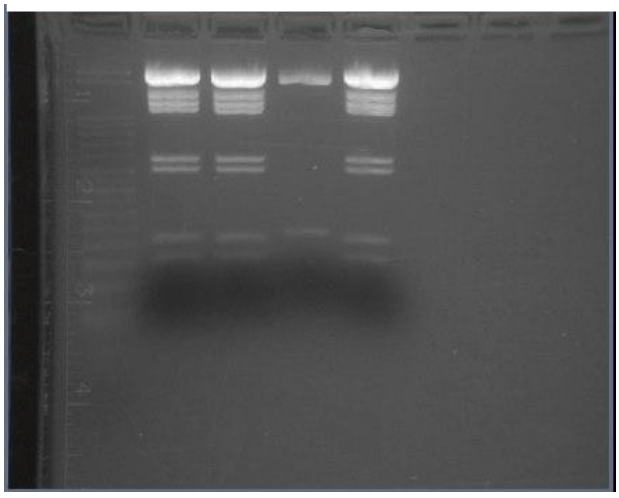
1 2 3 4 <--- Lanes



Lane 4 representing the pcr band for *Gag*-fs M(+) the mutated construct, after the 2nd step of the PCR at about 250bps; **Lane 1** is the ladder.

Figure 9. Gel results of the mutated construct

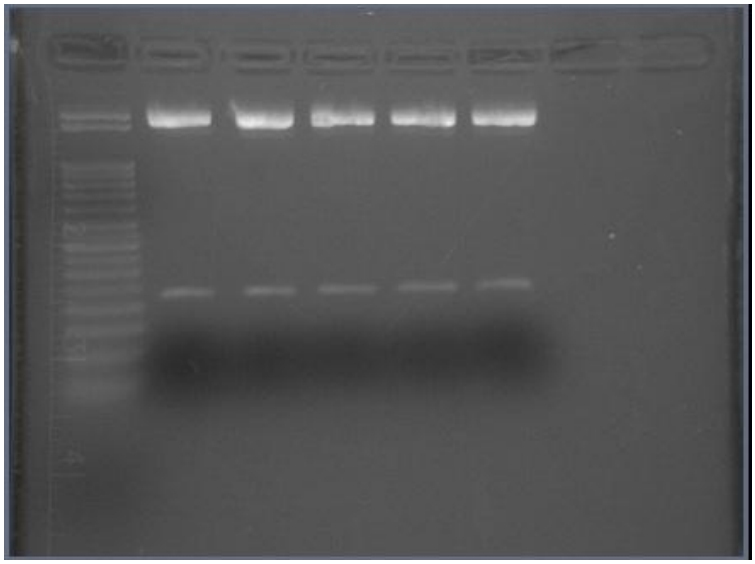
1 2 3 4 5 <--- Lanes



Gel results obtained after double digestion of *Gag*-fs-WT in pGEM-T with BglII and Xho1. **Lane 1** showing the ladder; **lane 3** representing the wild tpe at about 250bps indicating a successful ligation of the insert with pGEM-T; **Lane 2,3 & 5** indicating unsuccessful ligation.

Figure 10. Retriction digest analysis of the Wild type

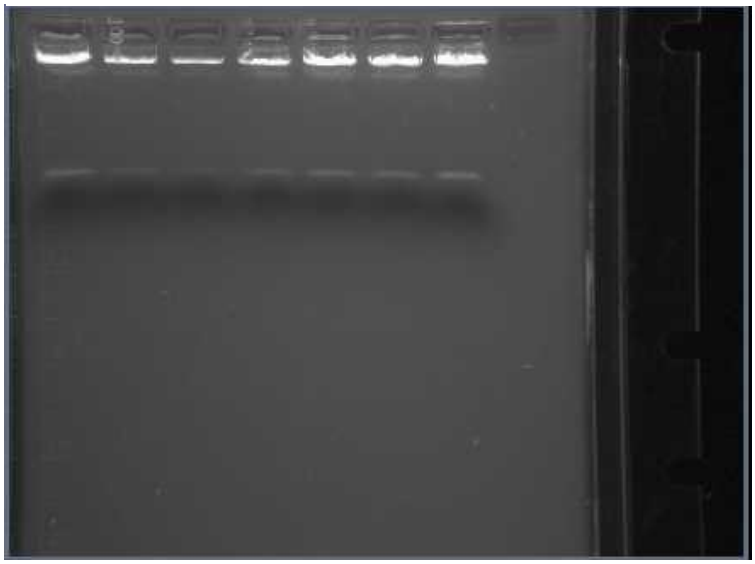
1 2 3 4 5 6 <--- Lanes



Gel results obtained after double digestion of Gag-fs-Mut(+) in pGEM-T with BglII and Xho1; **Lane 1** showing the ladder; **All the lanes (2-6)** contain a band at 250bps representing the mutated type, hence indicating a successful ligation of the insert with pGEM-T in all the isolated plasmids; Bands at about 3kbps represent the digested pGEM-T vector.

Figure 11. Restriction digest analysis of mutated type

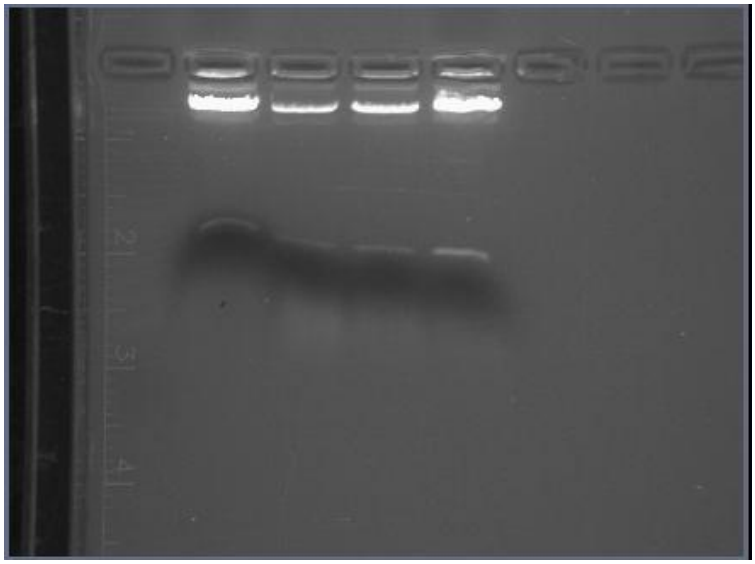
1 2 3 4 5 6 7 <--- Lanes



Gel results obtained after double digestion of Gag-fs-UGA in pGEM with BglII and Xho1; All the lanes contain a band at 250bps representing the Gag-fs UGA insert, indicating a successful ligation with pGEM-T in all the isolated plasmids; Bands at about 3kbps represent the digested pGEM-T vector.

Figure 12. Resctriction digest analysis of *Gag*-fs UGA

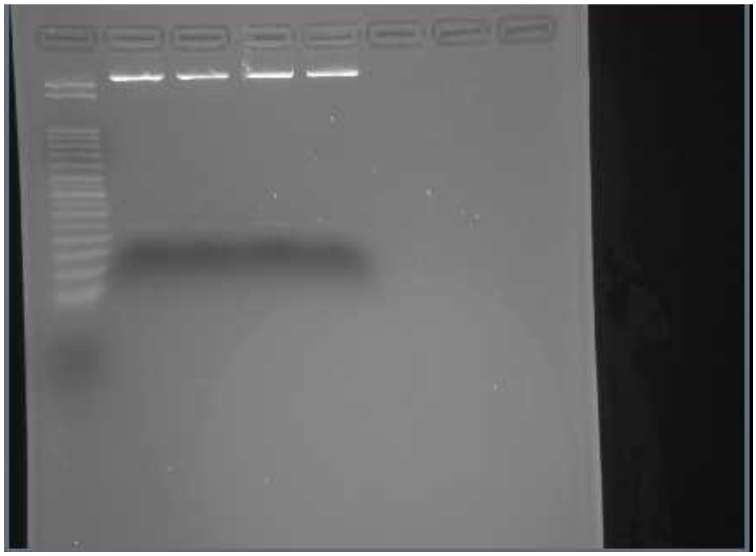
1 2 3 4 5 <--- Lanes



Gel results obtained after double digestion of *Gag*-fs-WT in LacZ-Luc with BglII and Xho1; Lane 1 showing the ladder (faintly visible); All the lanes contain a band at 250bps representing the *Gag*-fs WT insert, indicating a successful ligation with the dual reporter vector in all the isolated plasmids; Bands above 3kbps represent the digested LacZ-Luc vector

Figure 13. Restriction digest analysis of Wild type in LacZ-Luc

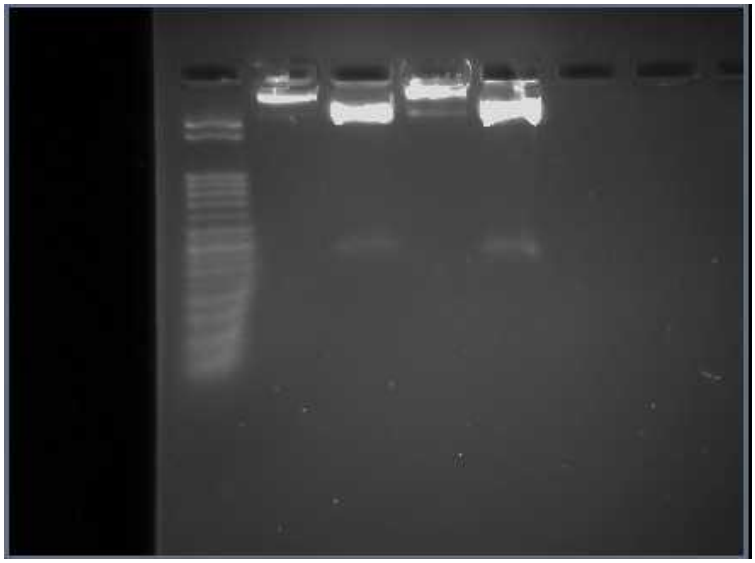
1 2 3 4 5 <--- Lanes



Gel results obtained after double digestion of *Gag*-fs-Mut(+) in LacZ-Luc with BglII and Xho1; Lane 1 showing the ladder; All the lanes contain a band at 250bps representing the *Gag*-fs MUT(+) insert, indicating a successful ligation with the dual reporter vector in all the isolated plasmids; Bands above 3kbps represent the digested LacZ-Luc vector.

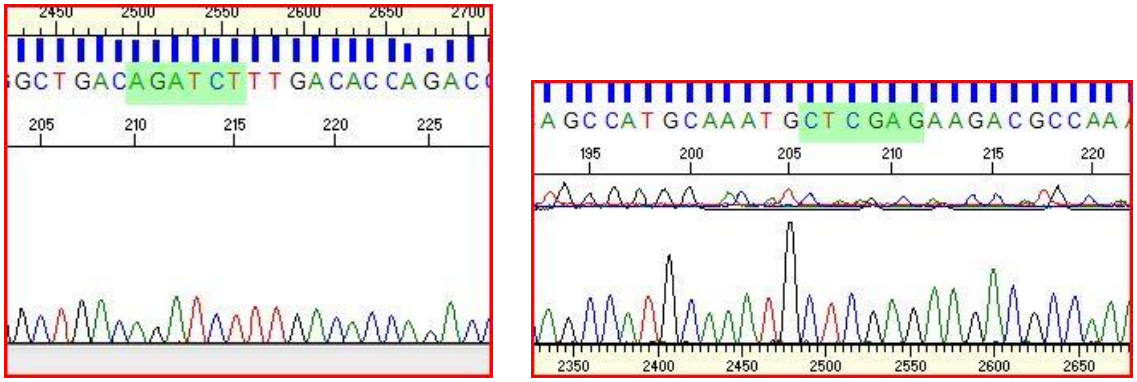
Figure 14. Restrcition digest analysis of Mutant in LacZ-Luc

1 2 3 4 5 <--- Lanes

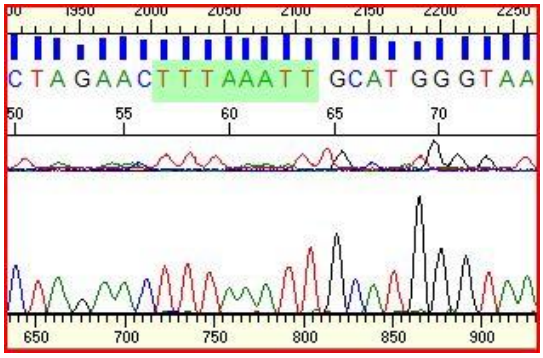


Gel results obtained after double digestion of Gag-fs-UGA in LacZ-Luc with BglII and Xho1; Lane 1 is the ladder (50bp-3Kbp); Lane 2 & 4 above 3kbps represent the undigested vector, Lane 3& 5 represent contain bands at 250bps showing insert *Gag*-fs-UGA and the cut LacZ-Luc at above 3kbps.

Figure 15. Restriction digest analysis of *Gag*-fs-UGA- LacZ-Luc

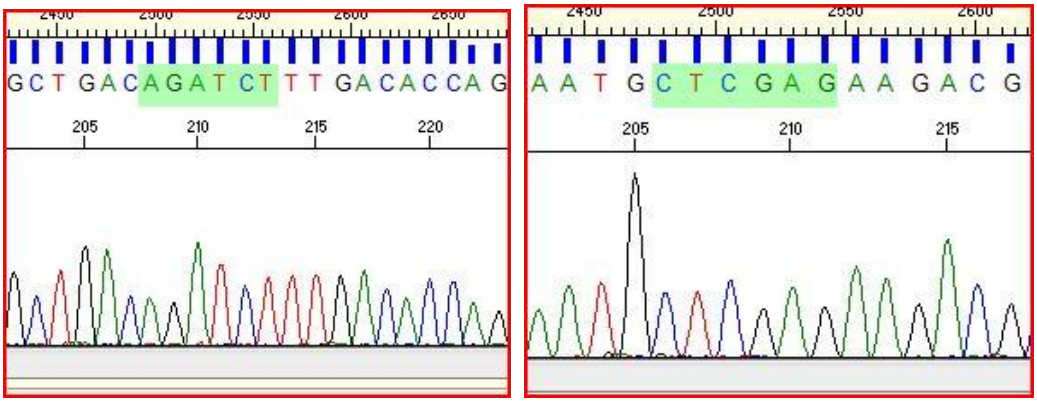


The highlighted sequences representing BglII site ‘AGATCT’, and the Xho1 site ‘CTCGAG’, confirming a successful ligation of the insert *Gag*-fs-MUT in the LacZ_LUC vector.

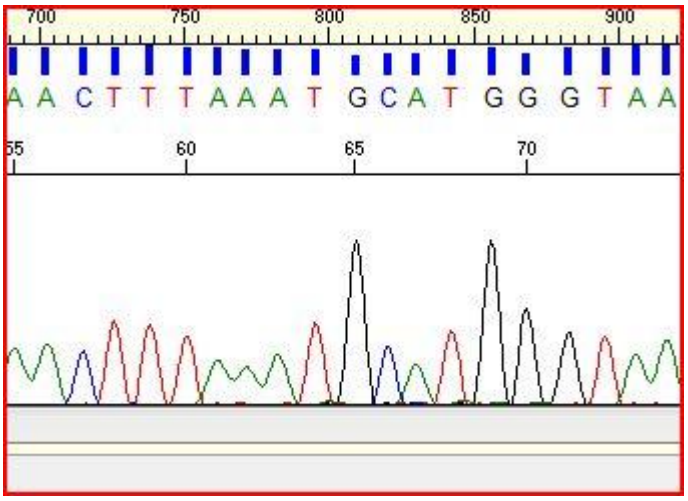


The highlighted sequence represent the slippery site “TTTAAAT”, with a successful insertion of an extra base ‘T’ right after the slippery site in the final construct Gag-fs-Mut(+) in LacZ-LUC.

Figure 16. Seq Verification of Mutated Construct



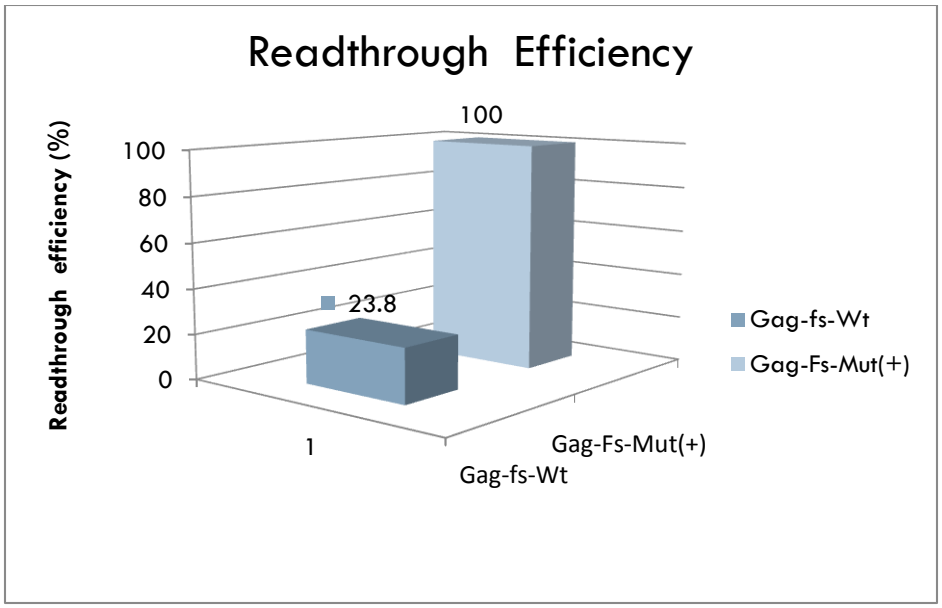
The highlighted sequences representing BglII site ‘AGATCT’, and the Xho1 site ‘CTCGAG’, onfirming a successful ligation of the insert Gag-fs-WT in the LacZ_LUC vector.



The sequence ‘TTTAAT’ in the above figure represent the slippery site of the wild type construct in LacZ-Luc.

Figure 17. Sequence verification of *Gag*-fs-WT

Sequence results obtained from wake forest institute.



The test construct *Gag*-fs-WT showed a 24% frameshifting efficiency when compared to 100% activity seen in the mutatant construct *Gag*-fs-Mut(+)

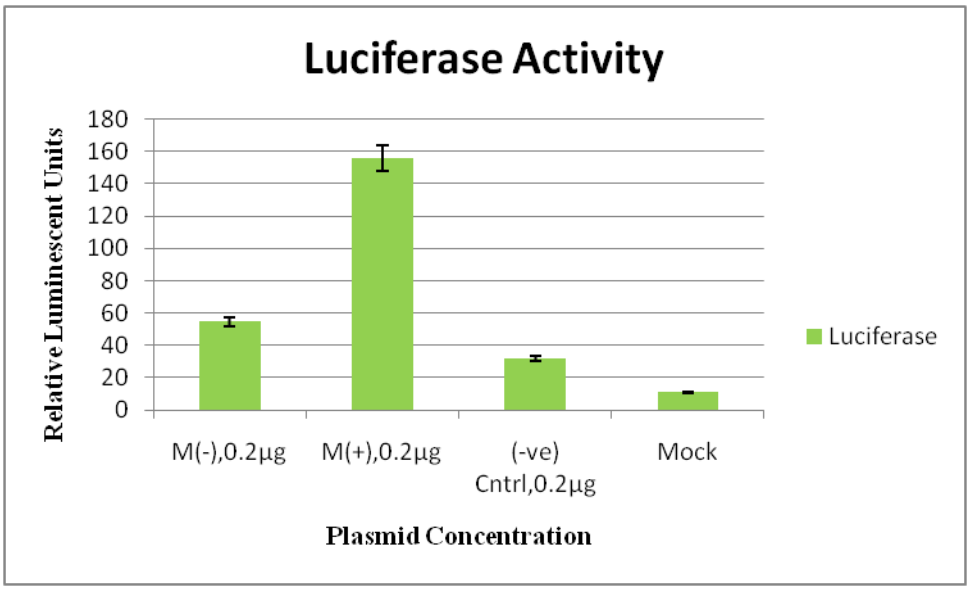
Figure 18. Readthrough Efficiency from dual reporter assay

.

Constructs	**Frameshifting/ Readthrough Efficiency (%)**
***Gag*-fs-WT**	23.8+/- 4.6
***Gag*-fs-Mut(+)**	100 +/- 18.6

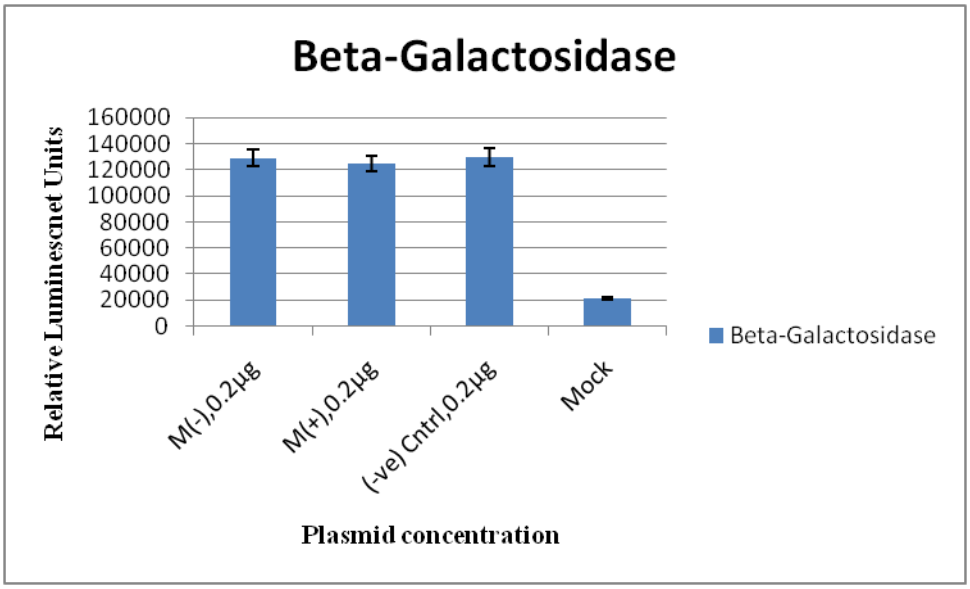
Average percent of frameshifting (+/- SEM) measured using the constructs expressing either wild type (Gag-fs-Wt) or the mutated type (Gag-fs-Mut(+), 100% readthrough control). Results are normalized to 100% for the Gag-fs-Mut(+) construct (n=6, p=0.002) .

Figure 19. Percent frameshifting efficiency



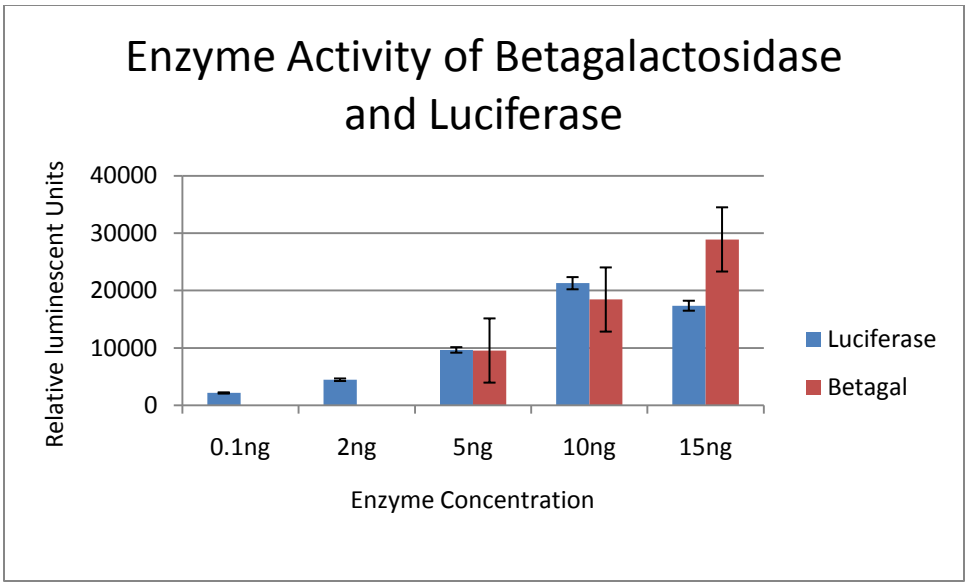
The avegare luciferase activity (RLU, +/- SD) measured from 0.2µg of wild type M(-), Mutant constructs M(+) and Negative control(-ve) from the dual reporter assay; M(-) showed some activity at ~50 (RLU); M(+) activity at ~150 (RLU); Negative control(-ve) showing activity at ~30 (RLU); baseline luminescenece provided my mock transfected cells.

Figure 20. Luminescence results of luciferase activity



The average BetaGalactosidase activity (RLU, +/- SD) measured from 0.2µg of wild type M(-), mutatant M(+) and negative (-ve) constructs from the dual reporter assay; the results show that there is an equal amount of activity of Beta-galactosidase taking place in all the three constructs; mock transfected cells provide a baseline luminescence.

Figure 21. Luminescence results of Beta-Galactosidase



The relative light signal (RLU +/- SD) produced from the assay with pure enzymes luciferase (blue bars) and betagalactosidase(maroon bars). The amount of signal produced from both enzymes at 5ng and 10ng concentrations are nearly equivalent but it varies a little at 15ng concentration; luciferase activity at 0.1ng and 2ng produce a light intensity between a 1000 to 5000 (RLU) .

Figure 22. Pure enzymatic activities from the assay

Discussion

Based on the results obtained from the dual reporter assay (Fig. 20) in the wild type construct *Gag* –fs WT there seems to be a fair amount of frameshifting (24%) into the -1 reading frame when the wild type is compared to the mutated construct *Gag*-fs-Mut(+), which is intended to read in -1 frame all the time, defining 100% translation.

One interesting observation made during these studies was that although approximately equivalent amounts of the purified enzymes (beta-galactoidase and luciferase) produced an equivalent light signal (measured in RLU) in the standard

curve experiment (Fig.23), in our actual assays with both wild type and mutant constructs (Figs. 21,22), the light intensity from the luciferase was consistently over 1000 fold *less* than that from beta-galactosidase. This is difficult to explain, because both the wild type and mutant 100% readthrough constructs express fusion proteins in which there is one of each of the two domains. Thus signal intensities for the 2 enzymes should be equivalent. The best explanation for this observation is that there may be a protein-protein interaction in the fusion protein, enhanced by the tethering together of the two domains, and that the geometry of the interaction partially occludes the luciferase active site, or causes an allosteric effect leading loss or reduction in luciferase activity. However, this effect would be similar in the fusion proteins produced by either construct, so the system still has utility for demonstrating that frameshifting is occurring.

However, this strongly suggests that the fundamental design of the frameshift assay construct using these two reporter genes may be flawed, and this is not an ideal system for our objective of the quantitation of frameshift efficiency.

There are also limitations on the use of non-mammalian reporter genes in such a system, as other investigators have suggested that beta-galactosidase is not ideal for this purpose (Grentzmann., 1998). In previous work from the Taylor lab, using ^{35}S-Met labeling of framesifted vs. non-frameshifted proteins, with beta-galactosidase as a downstream reported gene, bands corresponding to truncated

proteins were observed (Fig. 5 in Ref. Taylor, 2000). Because these truncated proteins would likely be enzymatically inactive, this phenomenon would interfere with the quantitation that we are attempting to achieve in the current study.

Thus, there are two major conclusions to this study:

1. An active ribosomal frameshift site in the capsid coding region of HIV-1 has been demonstrated, and an estimate (24%) of frameshift efficiency has been obtained.

2. Another conclusion is that further work needs to be done in the design of assays and vectors for the quantitation of frameshiftin

REFERENCES

1. ATKINS, J. F. (1991). Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine. *Proc. Natl. Acad.* , Vol. 88, pp. 6991-6995.
2. Biswas, P. (2004). The Human Immunodeficiency Virus Type 1 Ribosomal Frameshifting Site Is an Invariant Sequence Determinant and an Important Target for Antiviral Therapy. *JOURNAL OF VIROLOGY* , p. 2082–2087.
3. Brierley, I. D. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. *Cell* , 57, 537–547.
4. Brierley, I. (1992). Mutational Analysis of the "Slippery-sequence" Component of a Coronavirus Ribosomal Frameshifting Signal. *J. Mol. Biol.* , *227*, 463-479.
5. Brierley, I. (1995). Ribosomal frameshifting on viral RNAs. *Journal of General Virology* , 1885 1892.
6. Chris S. Martin., P. A. (1996). Dual Luminescence-Based Reporter Gene Assay for Luciferase andb-Galactosidase. *BioTechniques* , 21:520-524.
7. Cullen, B. R. (1991, March). Human Immunodeficiency Virus as a Prototypic Complex Retrovirus. *JOURNAL OF VIROLOGY* , 1053-1056.
8. Dinman, J. D. (1992). Ribosomal Frameshifting Efficiency and gag/ gag-pol Ratio Are Critical for M1 Double Stranded RNA Virus Propagationl. *Journal of Virology* , p. 3669- 3676.
9. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine. (1991). *Proc. Natl. Acad.* , Vol. 88, pp. 6991-6995.

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

11. Giedroc, D. P. (2000). Structure, Stability and Function of RNA Pseudoknots Involved in Stimulating Ribosomal Frameshifting. *J. Mol. Biol.* , 298, 167-185.

12. Göttlinger, H. G. (2001). HIV-1 Gag: a Molecular Machine Driving Viral Particle Assembly and Release. *HIV Sequence Compendium 2001* , pp. 2-28.

13. Grentzmann., G. (1998). A dual-luciferase reporter system for studying recoding signals. *RNA* , 4: 479-486.

14. Ian Brierley, F. J. (2006). Programmed ribosomal frameshifting in HIV-1 and the SARS–CoV. *Virus Research* , (119) 29–42.

15. Jacks, T. (1988). Signals for Ribosomal Frameshifting in the. *Cell* , Vol. 55, 447-458.

16. Jacks, T. (1988a). Signals for ribosomal frameshifting in the Rous sarcoma virus Gag-Pol region. *Cell* , 55, 447–458.

17. Jacks, T. (1990). Translational suppression in gene expression in retroviruses and retrotransposons. *Curr. Top. Microbiol. Immunol.* , 157, 93-124.

18. Jacobs, J. L. (2007). Identification of functional, endogenous programmed 1 ribosomal frameshift signals in the genome of Saccharomyces cerevisia. *Nucleic Acids Research* , Vol. 35, No. 1 165–174.

19. KOLLMUS, H. (1994). The Sequences of and Distance between Two cis-Acting Signals Determine the Efficiency of Ribosomal Frameshifting in Human Immunodeficiency Virus Type 1 and Human T-Cell Leukemia Virus Type II In Vivo. *JOURNAL OF VIROLOGY* , p. 6087-6091.

20. Lijun Zhao, A. G. (2000). Molecular modeling and in vitro activity of an HIV-1-encoded glutathione peroxidase. *Proc Natl Acad* , 97(12): 6356–6361.

21. Shen, L. T. (1995). The structure of an RNA pseudoknot that causes efficient frameshifting in mouse mammary tumor virus. *J. Mol. Biol.* , 247, 963–978.

22. Stewart, S. A. (2000). Human Immunodeficiency Virus Type 1 Vpr Induces Apoptosis through Caspase Activation. *Journal of Virology* , Vol. 74, No. 7: 3105-3111.

23. Takyar, S. H. (2005). mRNA helicase activity of the Ribosome. *Cell* , 120, 49–58.

24. Taylor, E. W. (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* , pp 2637–2654.

25. Taylor, E. W. (2000). Nutrition, HIV, and Drug Abuse: The Molecular Basis of a Unique Role for Selenium. *Journal of Acquired Immune Deficiency Syndromes* , 25:S53–S61.

26. Yusupova, G. Y. (2001). The path of messenger RNA through the ribosome. *Cell* , 106, 233–241.