HIV pol EXPRESSION VIA A RIBOSOMAL FRAMESHIFT

Alan J. Kingsman, Wilma Wilson and Susan M. Kingsman

Department of Biochemistry
South Parks Road
Oxford
OX1 3QU
U.K.

INTRODUCTION

The genetic relationships of the gag and pol genes of all retroviruses are approximately the same and the strategy for expression of the protein products of these genes is also strongly conserved (Weiss et al., 1982) (e.g. Figure 1). The gag and pol genes are adjacent and in many cases the 3' end of gag and the 5' end of pol overlap by up to a few hundred nucleotides. Where there is an overlap pol is generally in the -1 translational phase with respect to gag. Both genes are expressed from the full length genomic RNA to produce two primary translation products, a GAG precursor protein and a GAG:POL fusion precursor protein. The production of the fusion protein is achieved by the gag and pol reading frames being brought into translational phase. For several years it was assumed that this translational shift was mediated by a splice and the absence of any evidence for this was explained by proposing that the splice was small and therefore hard to detect (Weiss et al., 1982). However, in 1985 two pieces of data suggested that the splicing hypothesis was wrong. First, in the retrovirus-like yeast transposon Ty it was shown that frameshifting between the TYA gene, a gag analogue, and the TYB gene, a pol analogue, was not due to splicing (Mellor et al., 1985; Clare and Farabaugh, 1985). Secondly, Jacks and Varmus (1985) showed that RSV frameshifting could be achieved when an RNA synthesised in vitro was used in an in vitro translation system. This clearly excluded the possibility of a splice and suggested that the frameshift was due to some event at the ribosome. The phenomenon is now refered to as ribosomal frameshifting.

Some retroviruses fuse the products of gag and pol by a different mechanism. For example in MLV the gag and pol genes are adjacent and in phase but separated by a UAG termination codon.

NATO ASI Series, Vol. H 49
Post-Transcriptional Control of Gene Expression
Edited by J. E. G. McCarthy and M. F. Tuite

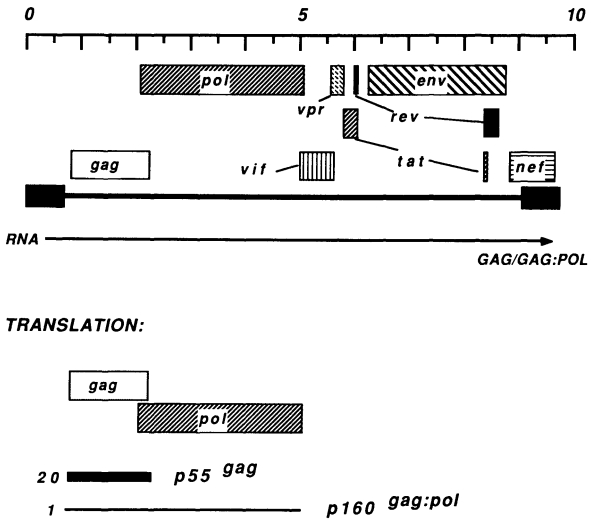

Figure 1 The genetic organisation of HIV-1. The upper figure shows the HIV map and the lower figure shows the simple translation product of the full length RNA, p55 and the frameshift product p160.

Production of a GAG:POL fusion protein is achieved by suppression of termination by a glutamine-tRNA (Yoshinaka et al., 1985).

Clearly in retroviruses with gag and pol out of phase frameshifting is essential for the expression of the enzyme activities of the virus, protease, reverse transcriptase, RNaseH and integrase. Also the frequency of shifting, generally about 5%, determines the relative levels of GAG and POL proteins in the cell. In addition, the attachment of the POL proteins to the GAG proteins either via the shift, or by termination suppression, not only achieves the genetic economy commonly seen in viruses but also ensures that the enzyme activities are packaged into the virus, as it is the GAG proteins that assemble into the viral core.

In HIV-1, gag and pol overlap by 241 nucleotides with pol in the -1 phase with respect to gag (Figures 1 and 2) (Ratner et

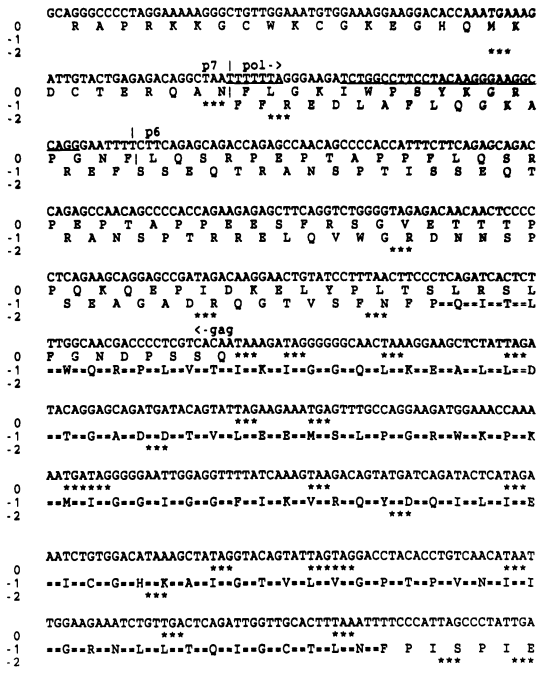

Figure 2. The nucleotide sequence of the gag:pol overlap region of HIV. The sequence is derived from the IIIB isolate reported by Ratner et al (1985). The frameshift site and the potential stem- loop region are underlines. 0, -1 and -2 denote translational reading phases where gag is arbitrarily given 0. The limits of the gag and pol open reading frames are marked as are cleavage sites that give rise to the p7 and p6 mature GAG proteins. The amino acid sequence of the viral protease is marked by '='. Stop codons are maked by ***.

al., 1985; Wain-Hobson et al., 1985; Sanchez-Pescador et al., 1985). The GAG precursor protein, $Pr55^{gag}$, is the primary product of simple translation of full length genomic viral RNA. The GAG:POL fusion protein, $Pr160^{gag:pol}$, also a precursor, is the product of frameshifted translation of the same full length RNA. The frequency of the shift is about 5% and therefore the relative abundance of the two precursors is about 20:1 respectively, although this has only been determined in vitro (Jacks et al., 1988a; Wilson et al., 1988).

The HIV overlap begins almost exactly at the 3' end of the p7 coding region. Protein p6 is encoded entirely by the overlap

in the GAG phase and the protease coding sequence in the POL phase overlaps p6 by 12 codons (Figure 2).

SEQUENCE REQUIREMENTS FOR THE GAG:POL RIBOSOMAL FRAMESHIFT

The standard assay for frameshifting uses an *in vitro* translation reaction to translate an SP6 generated RNA (Jacks and Varmus, 1985). The RNA is constructed with a preshift sequence,

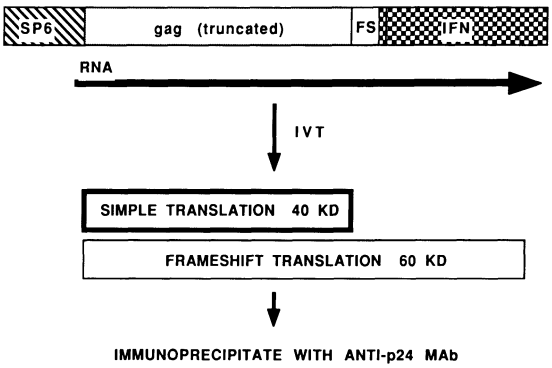

Figure 3. An SP6 frameshift assay system. IVT = *in vitro* translation. In this case the truncated gag is a pre-shift sequence, FS marks the position of a candidate shifting sequence and IFN (interferon) is a post-shift sequence that adds about 20 kD onto the 40kd GAG protein. This system is taken from Wilson *et al*. (1988).

a candidate shift site and then a post-shift sequence (Figure 3). Shifting is detected either by the presence of a second higher molecular weight band in an immunoprecipitation with an antibody directed against the protein encoded by the preshift sequence or by expression of a reporter gene that is dependent on the shift.

Frameshifting must occur in the *gag*:*pol* overlap region so that the shift into the -1, *pol* phase is achieved before the

ribosome reaches the stop codon of the gag open reading frame. Jacks et al. (1988a) showed that shifting occured at codon 3 of the pol open reading frame with the sequence TTT.TTA.GGG being read as PHE.LEU.GLY in GAG and as PHE.LEU.ARG in the GAG:POL fusion. However, the data reveal that there is also substantial (at least 30%) phenylalanine in the second position of the shifted product giving an alternative shift site sequence of PHE.PHE.ARG (see later). These data were in agreement with the observation of Wilson et al. (1988) who showed that the sequences required for HIV frameshifting were located within the first 16 nucleotides of the overlap region.

The observation that such a short stretch of the overlap region was required for shifting was surprising in the context of what was known about other retroviral requirements for frameshifting (Jacks et al., 1987; Jacks et al., 1988b). In almost every case where a virus makes use of frameshifting as a gene expression strategy the shift site or putative shift site is followed within a very short distance, usually less that 10 nucleotides, by a region of secondary structure (Table 1 and Figure 4)). The secondary structure may be a simple stem-loop or a pseudoknot in which the loop sequence is capable of base-pairing with another region downstream to form a complex structure (Jacks et al., 1988b; Brierley et al., 1989). Stem-loop structures that may form downstream of a selection of retroviral shift sites, including those of the HIV/SIV family are shown in Figure 4. It is possible that these secondary structures are recognised by soluble "shifting factors' or by ribosomal components that mediate the shift. A simpler suggestion is that these structures slow the ribosome making it more likely that it will 'slip' back at the shift site. In RSV there is good evidence (Jacks et al., 1988b) that the downstream secondary structure is required from mutational studies that show that destabilisation of the stem-loop substantially decreases frameshifting efficiency and unpublished data (cited in Jacks et al. 1988b) indicates that ribosomal pausing may occur at the stem loop. In HIV, however, although a stem-loop structure exists (Figure 4) it does not lie within the first 16 nucleotides of the overlap region. It would seem therefore that HIV frameshifting does not require ribosomal stalling and is therefore not mediated by the same mechanism as is used in RSV (Wilson et al., 1988). We will return to this later.

Table 1. Frameshift classes.

RETROVIRUS/ RETROELEMENT	OVERLAP	FRAMESHIFT SEQUENCE	DISTANCE TO SS.
Class I		X XXY YY	
		A AAU UU	
RSV	gag/pol	ACA AAU UUA UAG	7
SRV-1	pro/pol	GGA AAU UUU UAA	8
MPMV	pro/pol	GGA AAU UUU UAA	8
		G GGU UU	
Mouse IAP	gag/pol	CUG GGU UUU CCU	6
		G GGA AA	
SRV-1	gag/pro	CAG GGA AAC GAC	8
MPMV	gag/pro	CAG GGA AAC GGG	8
Visna	gag/pol	CAG GGA AAC AAC	7
		U UUA AA	
BLV	pro/pol	CCU UUA AAC UAG	7
HTLV-1	pro/pol	CCU UUA AAC CAG	7
HTLV-2	pro/pol	CCU UUA AAC CUG	7
Class II		X XXX XX	
		U UUU UU	
HIV-1	gag/pol	AAU UUU UUA GGG	8
HIV-2	gag/pol	GGU UUU UUA GGA	5
SIV	gag/pol	GGU UUU UUA GGC	4
gypsy	gag/pol	AAU UUU UUA GGG	8
		A AAA AA	
MMTV	gag/pro	UCA AAA AAC UUG	8
BLV	gag/pro	UCA AAA AAC UAA	8
HTLV-1	gag/pro	CCA AAA AAC UCC	7
HTLV-2	gag/pro	GGA AAA AAC UCC	8
EIAV	gag/pol	CCA AAA AAC GGG	10
EXCEPTION			
MMTV	pro/pol	CAG GAU UUA UGA	5
Ty1-15	TYA/TYB	CAU CUU AGG CCA GAA	

Table 1. Two classes of retroviral and retroelement frameshifting Examples of class I and II demonstrated or suspected shift sites are listed. The sequences are grouped to show gag or gag-equivalent codons. The last column shows the distance from the actual or purtative frameshift sequence to a region of downstream secondary structure. References for nucleotide sequences: RSV, Rous sarcoma virus (Schwartz et al., 1983); SRV-1, simian retrovirus type 1 (Power et al., 1986); MPMV, Mason-Pfizer monkey virus (Sonigo et al., 1986); 17.6 (Saigo et al., 1984); Visna virus (Sonigo et al., 1985); Mouse IAP, mouse intracisternal A particle (Meitz et al., 1987); BLV, bovine leukemia virus (Sagata et al., 1985; Rice et al., 1985); HTLV-1, human T cell leukemia virus type 1 (Hiramatsu et al., 1987); HTLV-2, human T cell leukemia virus type 2 (Shimotohno et al., 1985; Mador, Panet and Honigman, 1989); HIV-1 (Ratner et al., 1985); HIV-2 (Guyader et al., 1987); SIV, simian immunodeficiency virus (Chakrabarti et al., 1987); gypsy (Marlor et al., 1986); MMTV, mouse mammary tumor virus (Jacks et al., 1987; Moore et al., 1987); EIAV, equine infectious anemia virus (Stephens et al., 1986); Ty1-15 (Mellor et al., 1985; Wilson et al., 1986; Clare, Belcourt and Farabaugh, 1988).

Shift sites are generally regarded as being heptanucleotide sequences. Examination of these sequences reveals that they fall into two broad groups (Table 1). In what we will call Class I shift sites the first six nucleotides are of the general sequence X.XXY.YY where the stops represent GAG codons. In Class II shift

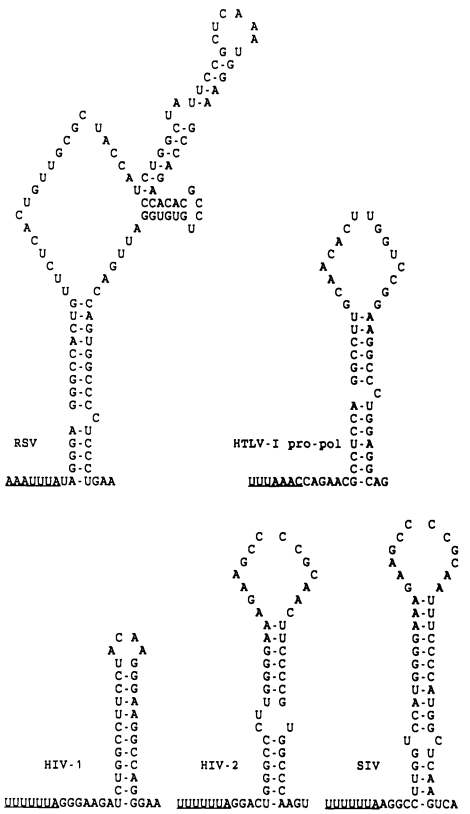

Figure 4. A selection of potential secondary structures found associated with retroviral frameshift sites. In each case the heptanucleotide shift site is underlined. Sequence information is the same as for Table 1.

sites the general sequence is X.XXX.XX (Table 1). HIV has a heptanucleotide sequence of U.UUU.UUA. Jacks et al (1988a) have shown that mutating the terminal UUUA to CUUA or UCUA dramatically reduces frameshifting efficiency. On the basis of this and the fact that UUUA appears in many shifty sequences it has been suggested that this tetranucleotide is a key signal in determining shifting. However, Wilson et al (1988) have shown that deletion of three of the six T/Us in the heptanucleotide also disrupts shifting even though reading frame and the UUUA sequence are preserved. More subtle changes in the 6T/U sequence also disrupt shifting. For example mutation of the third T/U or the sixth T/U to any of the other three nucleotides substantially

reduces the shift. However, mutation of the A at position 7 has no effect except in the case of a change of A to T/U when frameshifting increases. This, and other studies (Wilson et al., unpublished data), suggest that the signal for shifting in HIV is the 6T/U sequence rather than UUUA.

THE MECHANISM OF RIBOSOMAL FRAMESHIFTING IN RETROVIRUSES

Most studies of viral frameshifting have been carried out on viruses with sequences of the type that we have called Class I, such as RSV (Jacks et al., 1988b) and IBV (Brierley et al., 1989). In both of these, downstream secondary structure, in the form of a stem-loop or a pseudoknot, is required for efficient shifting. The only virus with a Class II sequence that has been analysed is HIV and this does not require a region of secondary structure just downstream of the shift site, even though one exists (Wilson et al., 1988; Madhani et al., 1988). It is tempting, and perhaps useful, to look for an explanation for this in the potential shiftiness of the shift sites alone. In Class I, such as the RSV gag:pol frameshift site, there are two short, adjacent, homopolymeric runs of three nucleotides of the general structure X XXY YY (Table 1). The phase relationship of these two triplets to the gag and pol reading frames is always the same irrespective of sequence composition. In RSV the heptanucleotide sequence AAAUUUA is the shift site and is thought to mediate shifting through a -1 slip of codon:anticodon interactions at both the A and P sites (Figure 5, Class I) (Jacks et al., 1988b). Following the slip the $tRNA^{Leu}$ (UUA) and the $tRNA^{Asn}$ (AAU) would be held on the RNA by 2 out of 3 base pairs each. Normal translocation would then take place and translation would proceed, in phase with pol, to the end of pol. Such a mechanism can be brought about by any adjacent, homopolymeric triplets as long as the distribution of the triplets with respect to the gag and pol reading phases is the same as in RSV (Table 1). This mechanism does not allow a shift to the +1/-2 phase. In Class II, such as the HIV gag:pol frameshift site, there is a single long homopolymeric run of six nucleotides of the general structure X XXX XX (Table 1). Like Class I the relationship of these six nucleotides to the gag and pol reading phases is conserved and is the same as the relationship of the two triplets in Class I (Table 1). In HIV the heptanucleotide shift sequence is UUUUUUA. Shifting to -1 can be achieved by a mechanism almost identical

to that proposed for Class I (Figure 5, Class IIa). In this case the $tRNA^{Phe}$ (UUU) and $tRNA^{Leu}$ (UUA) would both slip back one nucleotide in the A and P sites. The only difference between Class I and Class II would be that after the slip one of the tRNAs, $tRNA^{Phe}$, would be held on the RNA by three out of three base pairs. $tRNA^{Leu}$ would be held, as in Class I, by a 'two-out-

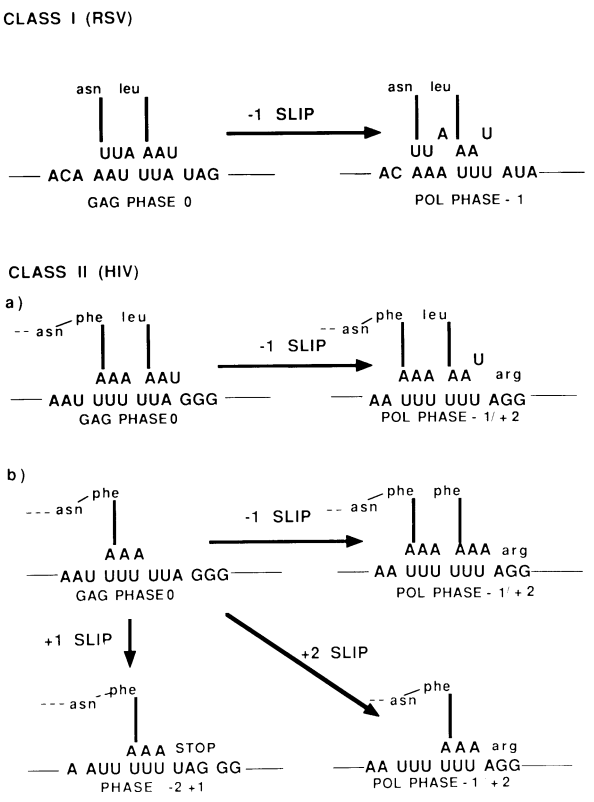

Figure 5. Class I and class II shifting. See text for explanation.

of-three' association. In Class II, therefore, the slip is maintained by five base pairs as opposed to four in Class I. Consequently, the slip may be more stable. The amino acid sequence over the frameshift region that would be predicted by this mechanism is ASN-PHE-LEU-ARG (Figure 5), the sequence determined for this region by Jacks et al. (1988a).

Slippage at the A and P sites is not the only mechanism open to a Class II shift sequence. It is possible that in HIV, for example, the $tRNA^{Phe}$ (UUU) slips during translocation (Figure 5, Class IIb) prior to the $tRNA^{Leu}$ (UUA) entering the A site. In this case the $tRNA^{Phe}$ slips to -1, maintaining three out of three base pairing, and exposing a free UUU codon in the A site. Rather than $tRNA^{Leu}$ (UUA) entering the A site, a second $tRNA^{Phe}$ (UUU) enters and the -1 shift is completed. The slip is maintained by six out of six base pairs and is likely, therefore, to be quite stable. This mechanism predicts that the amino acid sequence over the HIV frameshift region would be ASN-PHE-PHE-ARG. Both A and P site shifting and translocation shifting seem equally plausible for HIV. If both occured then a mixture of two gag:pol products would be produced differing at one amino acid position at the frameshift site. Close examination of the sequence data of Jacks et al (1988a) shows exactly that. At least 30% of the HIV-1 shifts produced the sequence ASN-PHE-PHE-ARG. We would suggest, therefore, that the frameshift in HIV might produce microheterogeneity in the gag:pol fusion protein.

Translocation slippage of the $tRNA^{Phe}$ could also produce a +2 slip (Figure 5, Class IIb) resulting in the amino acid sequence ASN-PHE-ARG. This apparently is not seen (Jacks et al., 1988a). Furthermore the HIV frameshift sequence could mediate shifting into the +1/-2 phase as well as into the -1/+2 phase. The translocation slip of $tRNA^{Phe}$ (UUU) could accommodate this via a +1 slip (Figure 5, Class IIb). As there is a UAG termination codon in the +1/-2 phase immediately after the shift site in HIV this observation predicts that there should be a truncated gag precursor protein of about 48 kD present at about 5% of the level of the authentic p55 GAG precursor protein. Clearly this is testable.

Our scheme suggests that Class II sequences are much shiftier than Class I and this may be reflected in differences in their respective requirements for downstream secondary structures. The requirements for HIV frameshifting appear to be simple. The virus makes use of a generally shifty sequence, T/U_6, to express its pol gene and this is all that is necessary. In contrast, RSV not only requires a shifty sequence, A_3T/U_3, but also a region of RNA secondary structure immediately downstream of the shift site (Jacks et al., 1988b). We would suggest that this is not a fundamental difference. Jacks et al (1988b) have

shown that in RSV the sequence of the region that forms the secondary structure is not critical to shifting efficiency as long as the potential for secondary structure is maintained. This argues against sequence specific cellular factors being involved in a complex mechanism for RSV frameshifting. Class I frameshift sites are likely to be less efficient than Class II sites. There are fewer opportunities for shifting and the slip may be less stable in Class I sites. Viruses with class I, low efficiency, shifty sites may have therefore evolved to use a downstream secondary and possibly tertiary RNA structure (Jacks et al., 1988b; Brierley et al., 1989) to act as a ribosomal stalling site. During the stall there is an increased probability of a shift at the low efficiency site. In contrast, viruses with a class II, high efficiency site do not require a ribosomal stalling site to achieve shifting at the frequency required to deliver the correct ratio of GAG and GAG:POL fusion.

The notion that there are two classes of retroviral and retroelement frameshift sites is easily tested but it also needs to be qualified. There are shift sites that do not conform to either Class I or Class II. The MMTV -1 pro:pol shift site is one and may represent a member of a third class. The retrotransposon, Ty, also does not fit into the scheme. Its TYA and TYB gene products are fused by frameshifting occuring in the sequence, CATCTTAGGCCAGAA (in Ty1-15), (Wilson et al., 1986; Clare, Belcourt and Farabaugh, 1988). This mediates a +1 shift and may, therefore, resemble the +1 shifts observed in some prokaryotic systems, such as the E.coli RF2 gene (Craigen and Caskey, 1986; Curran and Yarus, 1988).

The number of shift sites that have been analysed is remarkably small, Ty, RSV, MMTV, IBV and HIV. A broader study will reveal whether the division of shifting into these two Classes is appropriate or whether the story is more complex.

ACKNOWLEDGEMENTS

We would like to thank our colleagues in the Virus Molecular Biology Group for comments on the manuscript. The MRC supports our work on HIV frameshifting.

REFERENCES

Brierley,I., Digard,P. and Inglis,S.C. (1989). Characterization of an efficient coronavirus ribosomal frameshifting signal: requirement for an RNA pseudoknot. Cell 57, 537-547.

Chakrabarti,L., Guyader,M., Alizon,M., Daniel,M.D., Desrosiers,R.C., Tiollais,P. and Sonigo,P. (1987). Sequence of simian immunodeficiency virus from macaque and its relationship to other human and simian retroviruses. Nature 328, 543-547.

Clare, J. and Farabaugh, P. (1985) Nucleotide sequence of a yeast Ty element: evidence for an unusual mechanism of gene expression. Proc. Natl. Acad. Sci. USA 82, 2829-2833.

Clare,J.J., Belcourt,M. and Farabaugh,P.J. (1988). Efficient translational frameshifting occurs within a conserved sequence of the overlap between the two genes of a yeast Ty1 transposon. Proc. Natl. Acad. Sci. USA 85, 6816-6820.

Craigen,W.J. and Caskey,C.T. (1986). Expression of peptide chain release factor 2 requires high-efficiency frameshift. Nature 332, 273-275.

Curran,J.F. and Yarus,M. (1988). Use of tRNA suppressors to probe regulation of Escherichia coli release factor 2. J. Mol. Biol. 203, 75-83.

Guyader,,M., Emerman,M., Sonigo,P., Claver,F., Montagnier,L. and Alizon,M. (1987). Genome organization and transcription of the human immunodeficiency virus type 2. Nature 326, 662-669.

Hiramatsu,K., Nishida,J., Naito,A. and Koshikura,H. (1987). Molecular cloning of the closed circular provirus of human T cell leukemia virus type 1: a new open reading frame in the gag-pol region. J.Gen. Virol. 68, 213-218.

Jacks,T. and Varmus,H.E. (1985). Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science 230, 1237-1242.

Jacks,T., Townsley,K., Varmus,H.E. and Majors,J. (1987). Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus gag-related polyproteins. Proc. Natl. Acad. Sci. USA 84,4398-4302.

Jacks,T., Power,M.D., Masiarz,F.R., Luciw,P.A., Barr,P.J. and Varmus,H.E. (1988a). Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331, 280-283.

Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988b). Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell 55, 447-458.

Madhani,H.D., Jacks,T. and Varmus,H.E. (1988). Signals for the expression of the HIV pol gene by ribosomal frameshifting. In The Control of HIV pol Gene Expression, R.Franza, B.Cullen and F.Wong-Staal,eds. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 119-125.

Mador,N., Panet,A and Honigman,A. (1989). Translation of gag, pro, and pol gene products of human T-cell leukemia virus type 2. J.Virol. 63, 2400-2404.

Marlor,R.L. Parkhurst,S.M. and Corces,V.G. (1986). The Drosophila melanogaster gypsy transposable element encodes putative gene products homologous to retroviral proteins. Mol. Cell. Biol. 6, 1129-1134.

Meitz,J.A., Grossman,Z., Leuders,K.K. and Kuff,E.L. (1987). Nucleotide sequence of a complete mouse intracisternal A-

particle genome: no relationship to known aspects of particle assembly and function. J. Virol. 61, 3020-3029.
Mellor,J., Fulton,A.M., Dobson,M.J., Wilson,W., Kingsman,S.M. and Kingsman,A.J. (1985). A retrovirus-like strategy for expression of a fusion protein encoded by the yeast transposon, Ty1. Nature 313, 243-246.
Moore,R., Dixon,M. Smith,R., Peters,G. and Dickson,C. (1987). Complete nucleotide sequence of a milk-transmitted mouse mammary tumor virus: two frameshift suppression events are required for translation of gag and pol. J. Virol. 61, 480-490.
Power,M.D. Marx,P.A. Bryant,M.L., Gardner,M.D., Barr,P.J. and Luciw,P.A. (1986). Nucleotide sequence of SRV-1, a type D simian acquired immune deficiency syndrome retrovirus. Science 231, 1567-1572.
Ratner,L., Haseltine,W., Patarca,R., Livak,K.J., Starcich,B., Josephs,S.F., Doran,E.R., Rafalski,J.A., Whitehorn,E.A., Baumeister,K., Ivanoff,L., Petteway,Jr.,S.R. Pearson,M.L., Lautenberger,J.A., Papas,T.S., Ghrayeb,J., Chang,N.T., Gallo,R.C. and Wong-Staal,F. (1985). Complete nucleotide sequence of the AIDS virus, HTLV-III. Nature 313, 277-284.
Rice,C.M. and Strauss,J.H. (1981). Nucleotide sequence of the 26S mRNA of sindbis virus and deduced sequence of the encoded virus structural proteins Proc. Natl. Acad. Sci. USA 78, 1062-1066.
Sagata,N., Yasunaga,T., Tsuzuku-Kawamura,J., Ohishi,K., Ogawa,Y. and Ikawa,Y. (1985). Complete nucleotide sequence of the genome of bovine leukemia virus: its evolutionary relationship to other retroviruses. Proc. Natl. Acad. Sci. USA 82, 677-681.
Saigo,K., Kugimiya,W., Matsuo,Y., Inouye,S., Koshioka,K. and Yuki,S. (1984). Identification of the coding sequence for a reverse transcriptase-like enzyme in a transposable genetic element in Drosophila melanogaster. Nature 312, 659-661.
Sanchez-Pescador,R., Power,M.D., Barr,P.J., Steimer, K.S., Stempien,M.M., Brown-Shimmer,S.L., Gee,W.W., Renard,A., Randolph,A., Levy,J.A., Dina,D. and Luciw,P.A. (1985). Nucleotide sequence and expression of an AIDS-associated retrovirus (ARV-2). Science 227,484-492.
Schwartz,D.E., Tizard,R. and Gilbert,W. (1983). Nucleotide sequence sequence of Rous sarcoma virus. Cell 32, 853- 869.
Shimotohno,K., Takahashi,Y., Shimizu,N., Gojobori,T., Golde,D.W., Chen,I.S.Y., Miwa,M. and Sugimura,T. (1985). Complete nucleotide sequence of an infectious clone of human T-cell leukemia virus type II: an open reading frame for the protease gene. Proc. Natl. Acad. Sci. USA 82, 3101-3105.
Sonigo,P., Alizon,M., Staskus,K., Klatzmann,D., Cole,S., Danos,O., Retzel,E., Tiollais,P., Haase,A. and Wain-Hobson,S. (1985). Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus. Cell 42, 369-382.
Sonigo,P., Barker,C., Hunter,E. and Wain-Hobson,S. (1986). Nucleotide sequence of Mason Pfizer monkey virus: an immunosuppressive D-type retrovirus. Cell, 45, 375-385.
Stephens,R,M., Casey,J.W. and Rice,N.R. (1986). Equine infectious anemia virus gag and pol genes: relatedness to visna and AIDS virus. Science 231, 589-594.
Wain-Hobson,S., Sonigo,P., Danos,O., Cole,S. and Alizon,M.

(1985). Nucleotide sequence of the AIDS virus, LAV. Cell 40, 9-17.

Weiss,R., Teich,N., Varmus,H. and Coffin,J. (1982). Molecular Biology of Tumor Viruses. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Wilson,W., Malim,M.H., Mellor,J., Kingsman,A.J. and Kingsman,S.M. (1986). Expression strategies of the yeast transposon Ty: a short sequence directs ribosomal frameshifting. Nucl. Acids Res. 14, 7001-7015.

Wilson,W., Braddock,M., Adams,S.E., Rathjen,P.D., Kingsman,S.M. and Kingsman,A.J. (1988). HIV expression strategies: ribosomal frameshifting is directed by a short sequence in both mammalian and yeast systems. Cell 55, 1159-1169.

Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) Murine leukaemia virus protease is encoded by the gag-pol gene and is synthesised through suppression of an amber termination codon. Proc. Natl. Acad. Sci. USA 82, 1618-1622.