Nucleic Acids Research, 1995, Vol. 23, No. 9 1487–1494

Prokaryotic ribosomes recode the HIV-1 *gag–pol* –1 frameshift sequence by an E/P site post-translocation simultaneous slippage mechanism

Julie A. Horsfield, Daniel N. Wilson, Sally A. Mannering, Frances M. Adamski and Warren P. Tate*

Department of Biochemistry and Centre for Gene Research, University of Otago, Dunedin, New Zealand

Received February 8, 1995; Accepted March 16, 1995

ABSTRACT

The mechanism favoured for –1 frameshifting at typical retroviral sites is a pre-translocation simultaneous slippage model. An alternative post-translocation mechanism would also generate the same protein sequence across the frameshift site and therefore in this study the strategic placement of a stop codon has been used to distinguish between the two mechanisms. A 26 base pair frameshift sequence from the HIV-1 *gag–pol* overlap has been modified to include a stop codon immediately 3′ to the heptanucleotide frameshift signal, where it often occurs naturally in retroviral recoding sites. Stop codons at the 3′-end of the heptanucleotide sequence decreased the frameshifting efficiency on prokaryote ribosomes and the recoding event was further depressed when the levels of the release factors *in vivo* were increased. In the presence of elevated levels of a defective release factor 2, frameshifting efficiency *in vivo* was increased in the constructs containing the stop codons recognized specifically by that release factor. These results are consistent with the last six nucleotides of the heptanucleotide slippery sequence occupying the ribosomal E and P sites, rather than the P and A sites, with the next codon occupying the A site and therefore with a post-translocation rather than a pre-translocation –1 slippage model.

INTRODUCTION

During translation ribosomes are able to maintain the reading frame with great accuracy, with frameshift errors occurring at a frequency of ~1 in 10 000 codons translated (1,2). It is now clear, however, that in certain specialized cases the ribosome can be deliberately induced to change frame in order to produce an alternative polypeptide. In these instances the event has been fixed as a form of translational control, rather than it representing an error in translation. The expression of a downstream gene is often dependent on a frameshifting event at a specific signal within the mRNA (3–5).

Frameshifting signals can programme the ribosome to slip either forwards or backwards on the mRNA. For example, frameshifting in the +1 direction occurs during translation of *Escherichia coli* release factor 2 (RF-2) (6), the yeast Ty element (7) and human antizyme (8), whereas –1 frameshifting occurs during translation of some bacterial insertion sequences (reviewed in 9), *E.coli* dnaX (10) and several plant and animal viruses (11,12). The frameshifting signals are *cis*-acting in the translated mRNA, for example, –1 frameshifting requires a 'slippery sequence' consisting of seven nucleotides in the format XXXYYYZ, where X is U, G or A, Y is U or A and Z is U, C or A (12,13), and in most cases there is an additional requirement for a downstream secondary structure, often a 'pseudoknot' (14–16).

Some frameshifting motifs are functional on both prokaryotic and eukaryotic ribosomes. A 52 nucleotide (nt) sequence containing the critical heptanucleotide U UUU UUA from the *gag–pol* overlap of HIV-1 was found to frameshift in the –1 direction on *E.coli* ribosomes and the sequence requirements have been analysed and discussed in detail (17). A 26 nt sequence was sufficient for frameshifting on yeast and mammalian ribosomes, with a secondary structural element downstream from the slippery sequence not essential for the recoding event (18), although this element was important in vertebrate cells *in vivo* to maintain relatively high efficiencies of frameshifting (19). More recently, an even shorter 17 nt sequence from the HIV-1 site was shown to mediate frameshifting on *E.coli* ribosomes (20).

In 1988, Jacks *et al.* (13), from studies of the Rous sarcoma virus *gag–pol* frameshift region, proposed 'simultaneous slippage' as a model for –1 ribosomal frameshifting on viral RNAs. This model requires tandem slipping of tRNAs in the P and A sites of the ribosome from the 0 frame codons (NNX) XXY YYZ to pair with the corresponding –1 frame codons (NN) XXX YYY Z (13). More recent evidence indicates that additional modes of frameshifting may be possible, at least under certain physiological conditions. Yelverton *et al.* (20) used a 15–17 nt sequence from the HIV-1 site and when leucine was made limiting, frameshifting on *E.coli* ribosomes was increased. This result is consistent with an alternative mode of slippage taking place at the P site before Leu-tRNA is bound, followed by overlapping reading by the next aminoacyl-tRNA. The latter implies slippage

* To whom correspondence should be addressed

at the P site alone and overlapping reading by the next aminoacyl-tRNA. The sequence data of Weiss *et al.* (17) supports this pre-translocation mechanism as a contributor to frameshifting at the HIV-1 site.

Stop codons have long been known to increase the shiftiness of runs of nucleotides on prokaryotic ribosomes (21) and they are also present naturally at many frameshift sites, including retroviral sites that frameshift in a –1 direction (22–24) and sites in prokaryotic and eukaryotic mRNAs that frameshift in a +1 direction (6,8). The simultaneous slippage model invokes a pre-translocational slippage through the A and P sites and in this model the stop codon would not reach a classical ribosomal site before the shift has occurred, hence a role for the stop codon is hard to envisage. One way in which a stop codon could influence the frameshift efficiency through a release factor-mediated mechanism would be if the slippage event took place after translocation has occurred, rather than before, i.e. from the E and P sites, when the A site would be filled with the stop codon. The sequence across the frameshift region in the protein would be the same for either the pre-translocation or the post-translocation events.

We have examined this idea using the HIV-1 frameshift site and by changing the naturally occurring glycine codon immediately 3′ to the heptanucleotide slippery sequence from a GGG sense codon to each of the three stop codons, as well as to other sense codons. Our results show that the stop codon can influence frameshifting efficiency through a RF-mediated mechanism and this provides strong evidence that the –1 shift occurs by a post-translocational mechanism.

MATERIALS AND METHODS

Reagents

Restriction enzymes and buffers, T4 ligase and polynucleotide kinase were purchased from New England Biolabs. Mouse monoclonal anti-TrpE antibody was obtained from Santa Cruz Biochemicals. Anti-sheep and anti-mouse alkaline phosphatase (AP)-conjugated antibodies were obtained from Sigma. Oligonucleotides were synthesized on site using a Applied Biosystems 380B DNA Synthesiser or purchased from Macromolecular Resources, Colorado State University. DNA sequencing was conducted using the Sequenase® kit from USB. Radioisotopes, Hybond transfer membranes and the enhanced chemiluminescence (ECL) detection kit were obtained from Amersham and nitrocellulose from Schleicher and Schuell. All other chemical reagents were obtained from Sigma.

Plasmids

The host plasmid pRFTH4 was constructed by C. Donly from pATH11 as described previously (25). In the plasmid RFTH4, the C-terminal *SalI–EcoRI* fragment of an RF-2 clone is fused directly behind the TrpE gene of pATH11. The TrpE and RF-2 portions of the plasmid are in-frame, such that a single fusion protein is made upon induction *in vivo*. The pHIV series was created from pRFTH4 by ligating annealed complementary oligonucleotides with *Eco*RI ends into an *Eco*RI site at the TrpE–RF-2 fusion border. Sequences of the oligonucleotide pairs were:

FSHIV,5′-AATTCGGACAGGCTAATTTTTTA$^{T}/_{G}{}^{A}/_{G}$GAA ATCTTG-3′ and 5′-AATTCAAGATCTTC$^{T}/_{C}{}^{A}/_{C}$TAAAAAATTAGCCT GT CCG-3′;

FSSTOP, 5′-AATTCGGACAGGCTAATTTTTTAT$^{A}/_{G}A^{A}/_{G}$ AGATCTTG-3′ and 5′-AATTCAAGATCT$^{C}/_{T}T^{C}/_{T}$ATAAAAAATTAGCCTG TCCG-3′;

FSNULL, 5′-AATTCGGACAGGCTAACGCCTTA$^{G}/_{T}$AGA AGATCTTG-3′ and 5′-AATTCAAGATCTTCT$^{A}/_{C}$TAAGGCGTTAGCCTGT CCG-3′.

The frameshift window of the HIV-1 *gag–pol* overlap is contained within the oligonucleotides and is the minimum sequence required for frameshifting as defined by Wilson *et al.* (18). A *Bgl*II restriction site was introduced 3′ to the frameshift window and was used to create a large *Bgl*II deletion within RF-2, so that only the 30 C-terminal amino acids remain downstream of the HIV frameshift window. An extra nucleotide was included in the oligonucleotides in order to place the partial RF-2 sequence in the +1 frame with respect to TrpE, such that a –1 frameshift is needed to express the RF-2 fragment. Redundancies introducing changes to the GGG glycine codon following the slippery sequence were introduced when synthesizing the oligonucleotides, so that stop codons would be incorporated at this position. Individual clones were identified by colony hybridization and plasmid DNA was isolated by standard miniprep methods and by caesium chloride gradient centrifugation (26). Sequence identity was confirmed by dideoxy sequencing across the frameshift window with a Sequenase® kit and protocol from Amersham. Plasmids were transformed into the appropriate *E.coli* strains by electroporation (2.5 kV, 5–6 ms) using the Electro Cell Manipulator® 600 from BTX. Plasmid pFJU336 is derived from pACYC184 (27) and overexpresses RF-2 from an IPTG-inducible *trc* promoter. This plasmid is selectable with chloramphenicol and contains the ColE1 origin of replication.

Bacterial strains and media

All plasmids were maintained and expressed in the *E.coli* strains TG-1 (26) or TB-1 (28) and additionally transformed for expression into the *E.coli* strain FJU112 (29). Strain FJU112 has wild-type ribosomes and no suppressor tRNAs which could compete with termination or frameshifting events. Cultures were maintained in LB medium plus 100 μg/ml ampicillin and grown for induction in M9 minimal medium supplemented with thiamine at 10 μg/ml and tryptophan at 20 μg/ml with 100 μg/ml ampicillin (26).

Induction of plasmid expression

Freshly transformed individual colonies were grown overnight in M9 supplemented with thiamine and tryptophan and 100 μg/ml ampicillin. For induction, 250 μl overnight culture was used to inoculate 5 ml fresh M9 plus thiamine and ampicillin, but without tryptophan. The cultures were incubated at 37°C with high aeration for 2 h and then indole acrylic acid (10 mg/ml in ethanol) was added to a final concentration of 0.5 mg/ml. The cultures were incubated for another 2 h, after which their optical density was measured at 600 nm. Equivalent numbers of cells were pelleted in a microfuge and resuspended in 100 μl lysis buffer (10 mM sodium phosphate, pH 7.2, 6 M urea, 1% β-mercaptoethanol, 1% SDS, 0.04% bromophenol blue). Dilutions of lysates (1:5–1:20 as appropriate) were used for subsequent gel analysis. *Escherichia coli* strain TG-1 was transformed and induced to express both pFJU336 (RF-2) and pHIV constructs simultaneously. Cultures were selected and maintained on 100 μg/ml ampicillin and 30 μg/ml chloramphenicol. pFJU336 was induced with 1 mM IPTG when cells had reached the log phase. After

further growth, equivalent numbers of induced cells were pelleted and resuspended in 100 µl cell lysis buffer. Except where indicated, results are derived from multiple experiments using three different isolates of bacterial colonies containing each construct.

Immunoanalysis of expression products

Appropriately diluted samples of bacterial lysates (5 µl) were separated by SDS–PAGE on 12% acrylamide gels (30) using a BioRad Mini-PROTEAN II Electrophoresis cell. Proteins were transferred onto nitrocellulose membranes for 1 h using a BioRad Mini Trans-Blot Electrophoretic Transfer cell in a buffer that contained 25 mM Tris, 192 mM glycine, pH 8.3. Membranes were blocked overnight or for 2 h in TBSM (20 mM Tris, pH 7.6, 75 mM NaCl, 0.05% Tween 20 with 1% skimmed milk powder). Proteins immobilized on membranes were incubated for 2 h with primary antibodies diluted in TBSM: anti-TrpE at 1/1000 dilution or sheep polyclonal anti-RF-2 at 1/2000 dilution. The membranes were washed (5 × 5 min with TBSM), after which secondary antibodies (anti-mouse-AP or anti-sheep-AP used at dilutions of 1/2000 and 1/4000 in TBSM respectively) were applied for a further 2 h. After washing three times for 5 min in TBS (20 mM Tris, pH 7.6, 75 mM NaCl, 0.05% Tween 20), the blots were developed by an AP catalysed reaction with a specific substrate, 83.3 µg/ml BCIP, 166.6 µg/ml NBT in 100 mM Tris, pH 9.5, 100 mM NaCl, 5 mM $CaCl_2$ or by ECL (Amersham). The ratio of expressed products was determined by laser densitometry (LKB and BioRad).

In vitro release factor assays

The release assay and components are as described by Tate and Caskey (31). Formation of a f[^{3}H]Met-tRNA–AUG–ribosome complex (Stage 1) involved incubating 50 pmol 70S ribosomes with 25 pmol f[^{3}H]Met-tRNA and 2.5 nmol AUG in buffer [20 mM Tris–HCl, pH 7.4, 150 mM NH_4Cl, 10 mM $Mg(OAc)_2$] for 20 min at 30°C. Stage 1 (5 µl of a 50 µl reaction) was then added to 4 nmol codon (UGA for RF-2 and UAG for RF-1) and the required amount of release factor in RF buffer [50 mM Tris–HCl, pH 7.2, 75 mM NH_4OAc, 30 mM $Mg(OAc)_2$]. The reaction was incubated at 20°C for 30 min and then terminated by the addition of 250 µl 0.1 M HCl. Released f[^{3}H]Met was then extracted with 1 ml ethyl acetate and radioactivity was measured in a triton–toluene scintillant. The isotope had a specific activity of 4000 c.p.m./pmol, the extraction efficiency was 70% and counting efficiency of the scintillation counter for [^{3}H] under these conditions was 30%.

RESULTS

Experimental system

Constructs were made by synthesizing complementary oligonucleotides spanning 26 bp of the HIV-1 frameshift site (18) flanked by *Eco*RI restriction sites, so that after annealing they could be cloned into a plasmid vector pRFTH4 between partial genes for two *E.coli* proteins, TrpE and a short C-terminal portion of RF-2. This fragment contains 30 C-terminal amino acids of RF-2 and has no apparent biological activity, but is recognized by a polyclonal antibody raised against the native molecule. In order to introduce stop codons into the HIV-1 frameshifting context, the

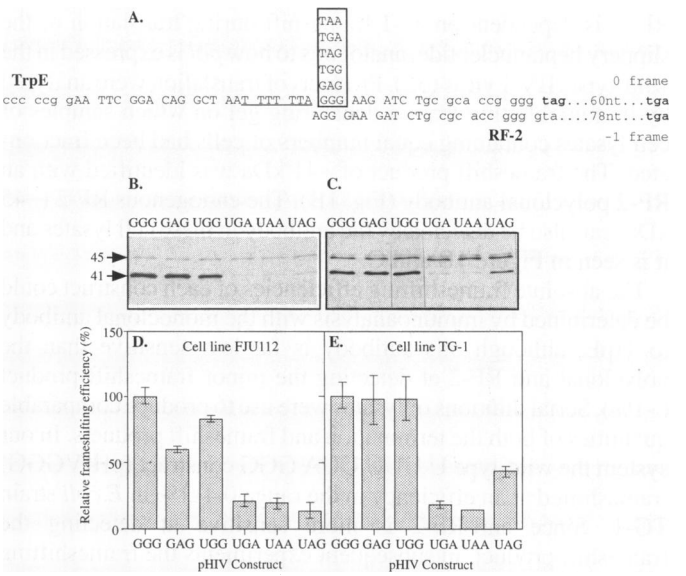

Figure 1. Frameshifting of pHIV constructs in the non-suppressing *E.coli* strain FJU112 and the supE44-containing strain TG-1. (**A**) DNA sequence of the pHIV series surrounding the HIV-1 *gag–pol* frameshift site: the cloned oligonucleotide sequence is given in capitals while the sequence of the upstream gene segment, *trpE*, and the C-terminal gene fragment of RF-2, *prfB*, appears in lower case. The upper triplet arrangement represents the 0 frame, the lower is the –1 frame arrangement. Sequence variations (boxed) are 3′ to the slippery heptanucleotide (underlined): the wild-type codon is GGG (glycine). Stop signals 3′ to the frameshift site in either frame appear in bold. (**B** and **C**) Anti-RF-2 immunoblots of *E.coli* lysates of strain FJU112 (B) and strain TG-1 (C) containing pHIV constructs expressed under the control of the *trp* promoter. The identity of the codon immediately 3′ to the slippery heptanucleotide is given above the panel. The three sense constructs appear on the left hand side of either panel, followed by the three 'stop' constructs. The frameshift product (41 kDa) and the endogenous *E.coli* RF-2 (~45 kDa) are arrowed. (**D** and **E**) Frameshifting efficiencies of the pHIV constructs expressed relative to the wild-type pHIVGGG construct in *E.coli* strains FJU112 (D) and TG-1 (E). Bands from anti-RF-2 immunoblots, as shown in (B) and (C), were scanned by laser densitometry and the analysis included a normalization for expression, as determined by anti-TrpE immunoblots (not shown). The bars depict an error of one standard deviation. The order of constructs is the same as for (B) and (C) and the codon following the heptanucleotide is listed below the panel.

oligonucleotides were synthesized with redundancies such that stop codons would appear immediately 3′ to the slippery heptanucleotide, replacing the GGG glycine codon that is present in the wild-type virus. The sequence variations are listed in Figure 1A. All three stop codons are represented and, in addition, sense codons TGG (Trp) and GAG (Val). The product predicted (~37 kDa) from these sense constructs is six amino acids longer than in the case of the stop constructs, as a result of termination occurring seven codons further downstream at a TAG (0 frame). When –1 frameshifting occurs the product is predicted to be 41 kDa, as a result of termination at a TGA a further 26 codons downstream (–1 frame) (Fig. 1A). The controls for these experiments were 'pHIVNull' constructs, where the slippery heptanucleotide was altered so that frameshifting was abolished at this site. All constructs were identified by sequencing selected clones.

Constructs were transformed into two *E.coli* strains: FJU112, which carries no suppressor tRNA, and TG-1, a typical laboratory strain which carries supE44, a suppressor of UAG stop codons. Expression was under the control of the *trp* promoter. Expression of the downstream reporter sequence, the last 30 amino acids of

RF-2, is dependent on a –1 frameshift during translation of the slippery heptanucleotide, analogous to how *pol* is expressed in the wild-type HIV-1 virus (32). Products of translation were analysed by immunoblotting from a denaturing gel on which samples of cell lysates containing equal numbers of cells had been fractionated. The frameshift product of ~41 kDa was identified with an RF-2 polyclonal antibody (Fig. 1B). The endogenous RF-2 (~45 kDa) can also be detected by the antibody in these cell lysates and it is seen in Figure 1B and C.

The absolute frameshifting efficiencies of each construct could be determined by immunoanalysis with the monoclonal antibody to TrpE, although this antibody is far less sensitive than the polyclonal anti-RF-2 at detecting the minor frameshift product (~1%). Serial dilutions of lysates were use to produce comparable quantities of both the termination and frameshift products. In our system the wild-type U UUU UUA GGG construct (pHIVGGG) frameshifted at an efficiency in the range 0.4–1% in *E.coli* strain TG-1. Since anti-RF-2 is more sensitive at detecting the frameshift product, in subsequent experiments the frameshifting efficiencies of the various constructs were expressed as a percentage of the wild-type pHIVGGG efficiency by comparative scanning of anti-RF-2 immunoblots. The monoclonal anti-TrpE was used to detect the major product (~37 kDa), for normalization of any variations in expression between constructs, since the gels were loaded with equivalent numbers of cells (data not shown).

Stop codons affect frameshifting efficiency

Stop codons inserted behind the HIV-1 slippery sequence influenced the frameshifting efficiency of this sequence on prokaryotic ribosomes. Figure 1B and C shows immunoblots of lysates of the cells FJU112 and TG-1 respectively, carrying the various constructs which have been probed with the polyclonal anti-RF-2. The frameshifting efficiencies relative to the wild-type pHIVGGG construct are shown in Figure 1D and E. The sense constructs pHIVUGG, pHIVGGG and pHIVGAG frameshifted at similar levels (41 kDa band). When UGA, UAA or UAG stop codons were substituted for the wild-type glycine codon (pHIVUGA, pHIVUAA and pHIVUAG respectively), frameshifting decreased to 10–20% of the wild-type pHIVGGG. An exception is pHIVUAG in TG-1 cells (Fig. 1C and E, last lane). In these cells, which carry the UAG suppressor SupE44, frameshifting is 4-fold higher than in the suppressor minus strain, although still lower than that obtained with the sense constructs.

Do stop codons affect frameshifting through a release factor-mediated mechanism or through a simple context effect?

Are these decreases in frameshifting efficiency due to a specific action of stop codons through a RF-mediated mechanism or merely to context effects of sequence changes around the frameshift site? We addressed this question by introducing a plasmid encoding the RF-2 gene into the bacteria containing the HIV constructs.

Modest overexpression of RF-2 from a low copy number plasmid further decreased the small amount of frameshift product from the constructs containing the stop codons (Fig. 2A). However, in all cells where the RF-2 plasmid was introduced expression from the HIV plasmids was ~2-fold lower, as

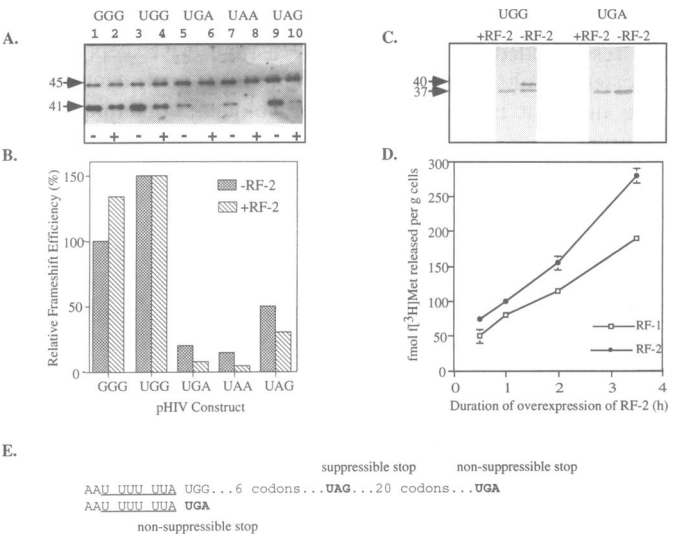

Figure 2. Effect of overexpression of RF-2 on frameshifting of pHIV constructs and on suppression of UAG codons in *E.coli* cell line TG-1. (**A**) Anti-RF-2 immunoblot of TG-1 cell lysates containing pHIV constructs expressed in the presence and absence of overexpressed RF-2. The identity of the pHIV constructs are given above the corresponding pairs of lanes (numbered 1–10) and the presence or absence of overexpressed RF-2 is indicated by plus and minus signs respectively. Frameshift (41 kDa) and RF-2 (45 kDa) proteins are arrowed. In each case, expression of the constructs containing RF-2 were consistently 2-fold lower than those without overexpressed RF-2, as determined from anti-TrpE immunoblots (not shown). (**B**) The effect of overexpression of RF-2 on the relative frameshift efficiencies of pHIV constructs. Frameshift efficiencies are expressed as a percentage of the wild-type pHIVGGG induced in the absence of overexpressed RF-2 (100%) and were obtained as described in the legend to Figure 1. (**C**) Anti-TrpE immunoblot of TG-1 lysates (equivalent numbers of cells) from constructs pHIVUGG and pHIVUGA with and without overexpressed RF-2. The suppression (~40 kDa) and termination (~37 kDa) products are arrowed. (**D**) Activity of both RF-1 and RF-2 increases on overexpression of RF-2. Release factor activity is expressed as fmol f[^{3}H]Met released/g cells determined for purified RF-1 (UAG-dependent release activity assayed in duplicate) and for purified RF-2 (UGA-dependent release activity assayed in triplicate) during the time course of RF-2 overexpression in *E.coli* cell line TB-1. (**E**) RNA sequence encompassing the HIV-1 frameshift site (underlined) and downstream 0 frame termination codons in the coding sequence which are either suppressible (UAG) or non-suppressible (UGA) by the TG-1 suppressor tRNA, supE44.

determined with the monoclonal anti-TrpE antibody. Overexpression of RF-2 decreased frameshifting efficiency 4- to 5-fold in the constructs containing UGA and UAA, after normalization for this variation in expression, but had little effect on frameshifting efficiency of the UGG or GGG sense constructs (Fig. 2B). This suggested that the effect of the stop codons was through an RF-mediated mechanism.

However, there was one confounding result; RF-2 overexpression also decreased frameshifting in the pHIVUAG construct (Fig. 2B) and yet the UAG codon is recognized by RF-1, and not RF-2. We have investigated this further and found that overexpression of RF-2 from an exogenous plasmid increases expression of endogenous RF-1. For example, suppression of UAG codons seen in TG-1 cells with the SupE44 UAG suppressor was also eliminated by overexpressing RF-2. In the absence of the plasmid expressing RF-2, two major products were detected from the sense constructs by the anti-TrpE antibody, illustrated in Figure 2C with UGG. The lower band is a product of ~37 kDa, as a result of termination at a UAG (see Fig. 2E, suppressible stop). However, this UAG termination codon is partially suppressed in TG-1 and so a further

band with an extra 20 amino acids is also observed; termination has now occurred at a non-suppressible UGA further downstream (see Fig. 2E). When RF-2 was overexpressed from the second plasmid, suppression at the UAG was abolished and so only the smaller band was seen (Fig. 2C, first lane). For comparison, the construct with the UGA stop codon is shown and this produced only one band in each case (slightly smaller than that obtained with the UGG construct, reflecting that it contains six fewer amino acids; see Fig. 2E). Termination occurs at the introduced non-suppressible UGA with this construct.

Furthermore, increased RF-1 functional activity is detected in extracts of cells overexpressing RF-2 from the introduced plasmid, as well as the expected elevated RF-2 activity. RF-1 and RF-2 were purified from samples of cells carrying the plasmid that were taken through an induction experiment. Their activities were measured with specific codons (RF-1, UAG; RF-2, UGA) in an *in vitro* termination assay (Fig. 2D). The activities of both factors increased over the duration of induction per gram of cells. An increased concentration of RF-1 protein, rather than a change in the specific activity of RF-1, was responsible for the increased activity. Control cultures did not show this change. We believe that expression of RF-2 influences regulation of RF-1 from the endogenous gene. Indeed, the levels of the two release factors expressed from endogenous genes are coordinated over a wide range of growth rates in *E.coli* by a mechanism yet to be elucidated (33). This coordinate regulation is now the subject of a separate investigation.

Our conclusion from these studies is that the RFs are influencing frameshifting efficiencies at the HIV site only when a stop codon is present immediately downstream from the heptanucleotide sequence (Fig. 1).

Up-regulation of the frameshifting by a defective release factor

We have further investigated the effect of release factors on frameshifting with the HIV constructs using the properties of the same series of constructs but where a near full-length gene encoding the *E.coli* RF-2 protein is present downstream of the HIV-1 frameshift sequence, rather than just a gene fragment encoding the RF-2 C-terminal 30 amino acids. This approach utilized a serendipitous finding that some plasmids in this series were able to express a defective release factor from a strong cryptic initiation site containing an AUG or an AUA codon introduced into the –1 frame in the frameshift window. Initiation from this codon produces only the RF-2 segment of the fusion protein, i.e. an RF-2 molecule with an intact C-terminus but missing the first 58 amino acids from the N-terminus, named RF-2Δ_{1-58}. The format of these constructs is illustrated in Figure 3A, with the new initiation site underlined. The products derived from such constructs are: a 38.5 kDa protein whose synthesis is terminated at an out-of-phase stop codon within the RF-2 sequence (0 frame); a 37 kDa protein where termination is determined by a nonsense codon at the frameshift site; a 72 kDa protein (the frameshift product) if there is a shift into the –1 frame; the novel protein of 33 kDa arising from the introduced cryptic initiation site in some of the constructs (see Fig. 3A).

The N-terminally truncated RF-2 (33 kDa) was seen as an abundant new product and was recognized by the RF-2 antibody (Fig. 3B). It was present only when the constructs had the first base of the glycine codon, G, replaced with T, so that ATG or ATA is created at the site. The antibody also detected the frameshift

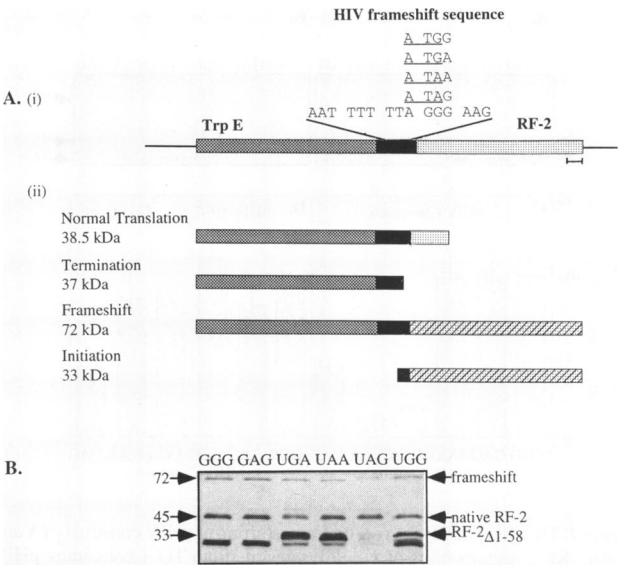

Figure 3. Proteins generated from a plasmid containing a near full-length RF-2 gene downstream of the HIV-1 frameshift site. (**A**) (i) General format of HIV constructs with a near full-length RF-2 gene 3′ of the frameshift sequence. The vector sequence is indicated by black lines at either end, the grey box is the upstream *trpE* gene, the black box is the cloned HIV frameshift oligonucleotide and the hatched box is the downstream RF-2 gene, *prfB*, which lacks 174 nt from the 5′ end. The bracket underneath the 3′ end of the RF-2 gene indicates the size and position of the fragment of RF-2 used as the reporter in the original series of experiments (Figs 1 and 2). The sequence spanning the HIV frameshift site is given above, with the –1 frame initiation codons underlined. (ii) Protein products derived from the constructs: 'normal translation' (38.5 kDa), where the product terminates out-of-frame in RF-2 (clear box); 'termination' (37 kDa), where the protein terminates at stop codons present at the frameshift site; 'frameshift' (72 kDa), where a shift in the –1 direction occurs at the slippery heptanucleotide and a full-length fusion protein is made; 'initiation' (33 kDa), where *de novo* initiation occurs within the frameshift site generating a novel RF-2 protein, RF-2Δ_{1-58}. (**B**) Anti-RF-2 immunoblot of *E.coli* lysates (strain FJU112) containing extended pHIV constructs. Two of the sense constructs, GGG and GAG, appear on the left, followed by the 'stop' constructs UGA, UAA and UAG and the sense construct UGG. The identity of the modified codon is given above the panel. The frameshift, native and initiation products are arrowed alongside.

product (72 kDa) and endogenous RF-2 (45 kDa). The 33 kDa protein was purified from cell extracts and its identity has been confirmed by N-terminal sequencing (Horsfield *et al.*, manuscript in preparation). The molecule was able to bind ribosomes *in vitro*, but had a specific activity for the release of a completed model peptide of only 0.1–1% of the native protein. We have deduced that *in vivo* it is a partially active RF, able to recognize stop codons, but unable to terminate protein synthesis effectively. In this way RF-2Δ_{1-58} would function as a competitive inhibitor of the endogenous RF-2 molecule. We have found previously that RF-2 with an N-terminal extension has the same properties, increasing frameshifting at the RF-2 site (25).

How does RF-2Δ_{1-58} influence frameshifting of the modified HIV sequences? The UGA and UAA constructs, when producing RF-2Δ_{1-58} (Fig. 4B and D), show a 5-fold increase in their frameshifting efficiencies relative to the constructs containing only the C-terminal fragment of RF-2 (Fig. 4A and C and also Fig. 1B). The frameshift bands seen are 72 kDa (Fig. 4B) and 41 kDa (Fig. 4A) for the two types of constructs (the endogenous RF-2 band is seen at 45 kDa in both sets). The UAG construct, pHIVUAG, which produces a low level of RF-2Δ_{1-58}, has a proportionally

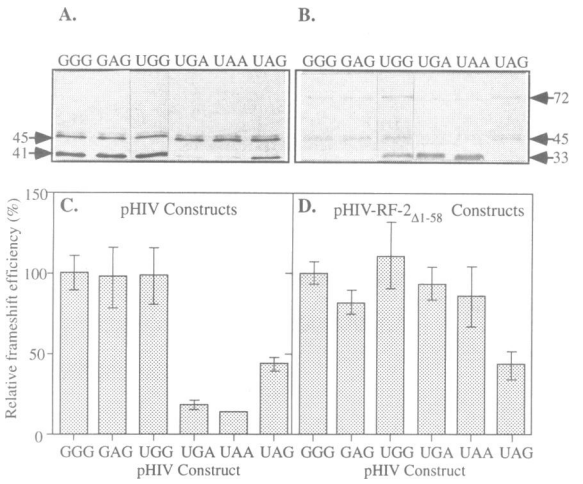

Figure 4. The effect of RF-2Δ_{1-58} on frameshifting of pHIV constructs. (**A** and **B**) Anti-RF-2 immunoblots of *E.coli* lysates of strain TG-1 containing pHIV constructs encoding (A; also Fig. 1C) the small C-terminal fragment of the RF-2 or (B) a near intact RF-2 (RF-2Δ_{1-58}) which carries a small deletion at its N-terminus, expressed under the control of the *trp* promoter as for Figure 1. The identity of the codon immediately 3′ to the slippery heptanucleotide is given above the panel. The three sense constructs appear on the left hand side of either panel, followed by the three 'stop' constructs. The sizes of the frameshift products (arrowed) are 41 kDa for the species containing only the small C-terminal fragment of RF-2 (A) and 72 kDa where the fusion protein contains TrpE and the near full-length RF-2Δ_{1-58} (B). The 45 kDa band is the endogenous *E.coli* RF-2 and the 33 kDa band appearing in (B) is the initiation product, RF-2Δ_{1-58}. (**C** and **D**) Frameshifting efficiencies of the pHIV constructs expressed relative to the wild-type pHIVGGG construct in the original construct series (C; also Fig. 1E) or in the presence of RF-2Δ_{1-58} (D). An analysis of the bands from anti-RF-2 immunoblots was carried out as described in the legend to Figure 1. The bars depict an error of one standard deviation. The order of constructs is the same as for (A) and (B).

similar level of frameshifting to that of the original construct containing only the 30 C-terminal amino acids. Since RF-2Δ_{1-58} is specific for UAA and UGA codons, we would not expect the UAG-containing construct to be influenced by this factor. The sense constructs pHIVGGG and pHIVGAG, which do not produce RF-2Δ_{1-58}, continue to frameshift at similar levels. The constructs gave comparable results in both the suppressing strain TG-1 and the non-suppressing strain, FJU112 (data not shown).

Together, the results suggest that a stop codon immediately downstream from the slippery sequence of the HIV-1 frameshift motif, in conjunction with a fully functional native or truncated defective RF, is able to exert direct influence on the frameshifting ribosome and modulate the frameshift event. Consistent with our observations of RF-mediated recognition was the effect of the suppressor tRNA in TG-1 on frameshifting efficiency of the UAG construct (shown in Fig. 1E compared with D). For the stop codon to be recognized by a release factor and influence frameshifting, it needs to be presented in the A site of the ribosome in the 0 frame. The implication of this is that a significant number of ribosomes which recode at the sequence must shift *after* the glycine codon of the wild-type HIV motif has appeared in the A site.

DISCUSSION

Stop signals at decoding sites

Stop signals occur naturally at many recoding sites and serve to enhance alternative events programmed by other *cis*-acting sequences (4,34). Stop codons following repetitive strings of nucleotides also stimulate –1, –2 and +1 frameshifts (21) and frameshifting efficiencies increase in proportion to the potential for aminoacyl-tRNAs to base pair in an alternative frame (22). Indeed, Weiss found that frameshifting efficiencies decreased 10-fold when a stop codon was replaced by a sense codon on a string of Gs (22).

How could stop signals affect frameshifting in the HIV-1 motif?

Studies of frameshift sites to date have shown that where stop codons exist as part of a frameshift sequence, they are either specifically involved in frameshifting (23,24) or have no apparent significant effect (13,35,36). The presence of stop codons immediately downstream of the heptanucleotide sequence in some –1 frameshift sites suggests a current or former active role for the stop codon in this position. Since stop codons are not present at a number of –1 frameshifting motifs, with a number of viral sites having a sense codon instead of a stop in that position, the influence of the stop codon may have been superseded by other *cis* elements in the motif, at least in these cases.

Where a stop codon exists at other recoding sites it has been assumed to prolong the pause for decoding that position and thereby increase the chance that frameshifting will occur. Unexpectedly, we found that stop signals immediately 3′ of the heptanucleotide sequence of the HIV-1 motif decreased (rather than increased or had no effect on) the frameshift efficiency in our studies. Among possible explanations is that frameshifting may occur at higher frequencies at positions downstream from the primary site as a result of the ribosome's difficulty in reading through the heptanucleotide motif. Placing a stop codon immediately after the motif would reduce this background by effectively shortening the frameshifting window. Indeed, a stop codon two codons down from the UUA shift site in the HIV-1 52 nucleotide motif does reduce the background signal (17).

Could it be then that the only role of the stop codon in natural frameshifting sites is to prevent minor secondary frameshifting on those ribosomes which had failed to frameshift at the primary site? Production *in vivo* of unwanted proteins in the cell would then be avoided. In that case, elevating the levels of RF in the cell should not eliminate frameshifting at the primary site. In our system an increased expression of plasmid-encoded RF-2 further decreased (by several fold) frameshifting from constructs carrying stop codons. In contrast, frameshifting in the constructs containing sense codons was not affected. These data imply that frameshifting at the primary site is also affected by stop codons and, since more efficient decoding of the stop codons has a profound influence, that it is not a simple context effect.

It may be that the HIV-1 sequence motif is a better context for decoding stop codons in the ribosomal A site by RFs than for decoding the GGG glycine codon by its tRNA. Indeed, this GGG codon is rarely used in *E.coli* (37) and may well be decoded more slowly than a stop codon. However, UGG and GAG also behave similarly in this position, although they are also not particularly commonly used. Perhaps, within the context of the HIV frameshift sequence, RFs decrease the 'decoding pause' in the A site relative to the aminoacyl-tRNAs as they decode their respective signals. An alternative explanation might be provided in the observation of non-randomness in the two codons immediately preceding *E.coli* stop signals and the suggestion that

RF-2 may interact with the last aminoacyl-tRNA during termination (38). While their results do not predict a favourable context for termination on the HIV sequence, it is possible that such an interaction might occur.

Can the effect of the stop signal be explained with current models?

How do our results impinge on current models to explain –1 frameshifting? The simultaneous slippage model proposed by Jacks *et al.* (13) (Fig. 5A) has proven to be invaluable in describing the event, as the model takes into account the protein sequence at the change in frame and the interactions between the codons and tRNAs when slippage occurs. In this model slippage occurs with two codons in tandem occupying the ribosomal A and P sites prior to the recoding event. In the HIV-1 motif, Leu is decoded and tandem slippage of Phe-tRNA and Leu-tRNA in the P and A sites is thought to occur before translocation. A modification of the simultaneous slippage model proposed by Weiss *et al.* (17) expands it to incorporate the hybrid three-site ribosome concept (39). The E site of the ribosome is involved in as much as the shift occurs post-transpeptidation, but pre-translocation, while the tRNAs are in E/P and P/A hybrid sites (17). In both cases the protein sequence at the point of shift is then Phe–Leu–Arg. The implication of this detailed description is that the codon immediately downstream from the heptanucleotide motif should not profoundly affect the slippage, apart from perhaps a general context effect. It is interesting that when this GGG was changed to a GCC, frameshifting was affected by about 60% (17).

An alternative model proposed by Yelverton *et al.* (20) (Fig. 5B) has been derived after sequencing the products of frameshifting when the bacteria had been starved for leucine and is invoked to account for part of the frameshifting under this condition. However, it is still thought to contribute to frameshifting in unstarved cells. In this new model, slippage occurs prior to Leu-tRNA binding to the UUA codon and therefore it is a P site slippage, with the Gly codon (replaced in our studies with a stop codon) downstream of the heptanucleotide motif remaining outside the classical site. After the event, two of the nucleotides of the Gly codon, GGG, enter the A site, now as part of the Arg codon, AGG. This protein sequence is different from the simultaneous slippage model (Phe–Phe–Arg compared with Phe–Leu–Arg).

In our studies it is this glycine codon, when changed to a stop signal, that is sufficient to decrease frameshifting dramatically in the presence of high levels of release factor. With our current understanding of ribosome function there appears no way of explaining our observations unless the stop codon is in the classical A site before the –1 shift occurs. This implies there must be peptide bond formation and a translocation event following binding of the Leu-tRNA to the UUA codon in the ribosomal A site, so that it becomes free to accommodate the next codon in the mRNA. The codon in the A site is then able to influence frameshifting at that site, whether it is a sense codon or a signal for termination. Indeed, a model involving the codon downstream of the heptanucleotide motif fits well with the observation that stop codons are frequently found in this position at recoding sites, in particular –1 frameshifting sites.

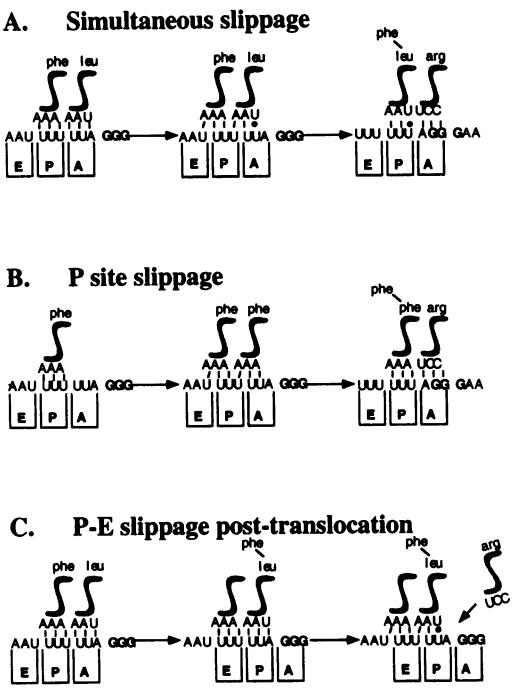

Figure 5. Hypothetical models describing the –1 frameshift event occurring at the HIV-1 slippery heptanucleotide. (**A**) The simultaneous slippage model proposed by Jacks *et al.* (13). The slippage occurs in tandem in the ribosomal A and P sites. Both tRNAs are involved, slipping backwards one nucleotide onto the corresponding codons in the –1 frame. Note that the GGG glycine codon does not appear in-frame in the A site at any time during the process. (**B**) The P site slippage model proposed by Yelverton *et al.* (20). The A site codon UUA remains unbound during the slip, while Phe-tRNA slips back one nucleotide in the P site to re-pair with the –1 frame codon. The next tRNA reads the overlapping –1 frame codon in the A site. The trans-frame product is distinguishable by protein sequencing from that produced by simultaneous slippage. This mechanism does not allow the GGG glycine codon to appear in the A site in-frame. (**C**) Post-translocational simultaneous slippage. In this model, frameshifting does not occur until after the glycine codon has appeared in the A site, when Phe-tRNA and Leu-tRNA are occupying the E and P sites respectively. Simultaneous slippage then occurs at the E and P sites, followed by decoding of the –1 frame codon by the next tRNA at the A site. Frameshifting by this model would generate a product indistinguishable by protein sequencing from that produced by simultaneous slippage.

A new model for ribosomal frameshifting on *E.coli* ribosomes: P–E site post translocation slippage

We propose a new model which would give the same protein sequence at the site as the original A–P site simultaneous slippage model (13). The model, shown in Figure 5C, proposes P–E site simultaneous slippage where the shift occurs after the translocation event. In this scenario the codons UUU UUA are decoded by Phe-tRNA and Leu-tRNA conventionally, after which translocation occurs. With the GGG Gly codon in the A site slippage occurs before the Gly-tRNA arrives at the A site. Within the HIV-1 recoding motif the context may be unfavourable for decoding GGG and a pause may result, which could act to promote frameshifting. The particular feature invoking slippage with an empty A site is also a feature of the Yelverton *et al.* model (20).

Simultaneous P–E site slippage would be feasible with respect to Moazad and Noller's E–P–A displacement model for the tRNA sites of the ribosome (39).The tRNAs appear to have contact regions with both 30S and 50S subunits at the A and P sites, which

presumably strengthen their interactions, whereas only the 50S subunit seems to make significant interaction with the tRNA in the E site. Hence, slippage from an E site position may be more favoured than from A or P. The fact that the stop codon/RF combination has the potential to abolish frameshifting at the HIV-1 site and therefore apparently compete effectively with the frameshifting event implies a significant proportion of the frameshifting previously attributed to A–P site simultaneous slippage is likely to result from this post-translocation mechanism we propose.

ACKNOWLEDGEMENTS

We thank Chris Brown for helpful discussions. JAH is supported by a postgraduate scholarship from the Health Research Council of New Zealand. WPT is an International Scholar of the Howard Hughes Medical Institute and is supported by grants from the Human Frontier Science Program, the NZ Health Research Council and the NZ Lotteries Board.

REFERENCES

1 Atkins,J.F., Elseviers,D. and Gorini,L. (1972) *Proc. Natl. Acad. Sci. USA*, **69**, 1192–1195.
2 Parker,J. (1989) *Microbiol. Rev.*, **53**, 273–298.
3 Gesteland,R.F., Weiss,R.B. and Atkins,J.F. (1992) *Science*, **257**, 1640–1641.
4 Tate,W.P. and Brown,C.M. (1992) *Biochemistry*, **31**, 2443–2450.
5 Farabaugh,P.J. (1993) *Cell*, **74**, 591–596.
6 Craigen,W.J., Cook,R.G., Tate,W.P. and Caskey,C.T. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 3616–3620.
7 Clare,J.J., Belcourt,M. and Farabaugh,P.J. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6816–6820.
8 Rom,E. and Kahana,C. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 3959–3963.
9 Chandler,M. and Fayet,O. (1993) *Mol. Microbiol.*, **7**, 497–503.
10 Tsuchihashi,Z. (1991) *Nucleic Acids Res.*, **19**, 2457–2462.
11 Atkins,J.F., Weiss,R.B. and Gesteland,R.F. (1990) *Cell*, **62**, 413–423.
12 Hatfield,D. and Oroszlan,S. (1990) *Trends Biochem. Sci.*, **15**, 186–190.
13 Jacks,T., Madhani,H.D., Masiarz,F.R. and Varmus,H.E. (1988) *Cell*, **55**, 447–458.
14 Brierley,I., Digard,P. and Inglis,S.C. (1989) *Cell*, **57**, 537–547.
15 Brierley,I. (1993) *Biochem. Soc. Trans.*, **21**, 822–826.
16 Ten Dam,E., Pleij,K. and Draper,D. (1992) *Biochemistry*, **31**, 11665–11676.
17 Weiss,R.B., Dunn,D.M., Shuh,M., Atkins,J.F. and Gesteland,R.F. (1989) *New Biologist*, **1**, 159–169.
18 Wilson,W., Braddock,M., Adams,S.E., Rathjen,P.D., Kingsman,S.M. and Kingsman,A.J. (1988) *Cell*, **55**, 1159–1169.
19 Parkin,N.T., Chamorro,M. and Varmus,H.E. (1992) *J. Virol.*, **66**, 5147–5151.
20 Yelverton,E., Lindsley,D., Yamauchi,P. and Gallant,J.A. (1994) *Mol. Microbiol.*, **11**, 303–313.
21 Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1987) *Cold Spring Harbor Symp. Quant. Biol.*, **52**, 687–693.
22 Weiss,R.B., Dunn,D.M., Atkins,J.F. and Gesteland,R.F. (1990) *Prog. Nucleic Acids Res. Mol. Biol.*, **39**, 159–183.
23 Brault,V. and Miller,W.A. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 2262–2266.
24 Gramstat,A., Prufer,D. and Rohde,W. (1994) *Nucleic Acids Res.*, **22**, 3911–3917.
25 Donly,B.C., Edgar,C.D., Adamski,F.M. and Tate,W.P. (1990) *Nucleic Acids Res.*, **18**, 6517–6522.
26 Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
27 Jorgensen,F., Adamski,F.M., Tate,W.P. and Kurland,C.G. (1993) *J. Mol. Biol.*, **230**, 41–50.
28 Stratagene (1994) *Stratagene Cloning Systems Catalogue, 1994*. Stratagene, CA, p. 282.
29 Jorgensen,F. and Kurland,C.G. (1990) *J. Mol. Biol.*, **215**, 511–521.
30 Laemmli,U.K. and Favre,M. (1973) *J. Mol. Biol.*, **80**, 575–599.
31 Tate,W.P. and Caskey,C.T. (1990) In Spedding,G. (ed.), *Ribosomes and Protein Synthesis: A Practical Approach*. Oxford University Press, Oxford, UK, pp. 81–100.
32 Jacks,T., Power,M.D., Masiarz,F.R., Luciw,P.A., Barr,P.J. and Varmus,H.E. (1988) *Nature*, **331**, 280–283.
33 Adamski,F.M., McCaughan,K.K., Jorgensen,F., Kurland,C.G. and Tate,W.P. (1994) *J. Mol. Biol.*, **238**, 302–308.
34 Tuite,M.F. and Stansfield,I. (1994) *Mol. Biol. Rep.*, **19**, 171–181.
35 Kim,K.H. and Lommel,S.A. (1994) *Virology*, **200**, 574–582.
36 Morikawa,S. and Bishop,D.H.L. (1992) *Virology*, **186**, 389–397.
37 Ikemura,T. (1985) *Mol. Biol. Evol.*, **2**, 13–34.
38 Arkov,A.L., Korolev,S.V. and Kisselev,L.L. (1993) *Nucleic Acids Res.*, **21**, 2891–2897.
39 Moazed,D. and Noller,H.F. (1989) *Nature*, **342**, 142–148.