Nucleic Acids Research, 1993, Vol. 21, No. 8 1837–1843

Analysis of effects of tRNA:message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site

James F.Curran
Department of Biology, Wake Forest University, PO Box 7325, Winston-Salem, NC 27109, USA

Received January 8, 1993; Revised and Accepted March 17, 1993

ABSTRACT

The codon that is in-frame prior to +1 frameshifting at the *E.coli prfB* (RF2 gene) frameshift site is randomized to create thirty-two variants. These alleles vary 1000-fold in frameshift-dependent expression in fusions to *lacZ*. Frameshifting is more frequent at sites where the in-frame codon ends in uridine, as if third position wobble pairs to message uridine facilitate slippage into the +1 frame. Consistent with other studies of programmed frameshift sites, efficient frameshifting depends on stable message:tRNA base pairs after rephasing. For complexes with mispairs, frameshift frequency depends on the nature, number, and position of mispairs. Central purine:purine mispairs are especially inhibitory. Relative stabilities of +1 rephased complexes are estimated from published data on the stabilities of tRNA:tRNA complexes. Stability correlates with frameshifting over its entire range, which suggests that stability is an important determinant of the probability of translation of the rephased complex.

INTRODUCTION

Programmed ribosomal frameshifting dramatically shows that message sites differ greatly in the capacity to transmit ribosomes into specific frames [for reviews, see (1,2)]. At these sites, and with frequencies of tens of percent, tRNAs and message realign to redirect translation specifically into either the +1 or −1 frame. Most programmed frameshifts are of two types. First are slippages of peptidyl-tRNA by one-nucleotide 3′ when the ribosomal A site does not contain tRNA. This +1 slippage places a new triplet in the ribosomal A site, allowing for translation in the +1 frame. Examples occur in *E.coli prfB* (3) and yeast Ty (4,5). Next are 'simultaneous-slippages' of tRNAs within both the A and P sites by one nucleotide 5′, which realigns the message in the −1 frame. The second type are found in eukaryotic viruses (2,6 and references therein; 7,8), and in *E.coli dnaX* (9−11), phages (12,13), and insertion sequences (14).

Specific features that contribute to high-frequency have been identified for many programmed frameshift sites. For most simultaneous-slippage sites, frameshifting is associated with stable secondary or tertiary message structures that are precisely located downstream of the slippage sites (6,15,16). Pausing and frameshifting both depend on the estimated stabilities of those structures (17,18), which supports the proposal that the structures halt ribosomes allowing time for tRNA:message rephasing (6). In *E.coli prfB* and yeast Ty, the 'pause' is the time required for ribosomal selection of RF2 or rare aminoacyl-tRNA, respectively, at an empty A site. In these systems, increasing rate of translation of the A site codon decreases frameshifting (19−23). That property of the *prfB* frameshift site is exploited to estimate relative rates of tRNA selection at various codons substituted at that site (21,24). Also for *prfB*, high frequency frameshifting requires that a specific run of purines upstream of the frameshift site interact with the Anti-Shine-Dalgarno sequence of 16 S rRNA (25), which may catalyze rephasing (20).

Programmed high-frequency frameshift sites all allow for stable base pairing between message and tRNA following rephasing. Mutations that create a requirement for unstable pairs can decrease frameshifting (6,26). The range of acceptable pairs, however, is greater than required during selection of aminoacyl-tRNA. For example, wobble pairs between G and U occur in both the first position during +1 frameshifting by the peptidyl-tRNA in *prfB* (3) and yeast Ty (23). For natural simultaneous-slippage sites, Watson:Crick pairs usually occur at the first two positions in both the A and P sites, though examples of G:U (27,28) and G:A (29) pairs for the P site central position are known. The suitability of pairs at the first position of the P site recently determined for variants of a coronavirus frameshift site (30) partially correlates with the *in vitro* stabilities of those pairs (31,32). These results suggest that frameshift probability is related to the stability of the rephased complex.

There is evidence that frameshift frequency may also depend on the stability of the tRNA:message complex before tRNA:message slippage occurs. For example, efficient simultaneous-slippage frameshifting is associated with relatively weak base pairs in the third position of the A site (17,30,33,34), or with hypomodification of anticodon loop bases (35). In addition, sequences that give maximal efficiencies differ between *E.coli* and eukaryotic expression systems (34), perhaps reflecting differences in tRNA structures between these genetic systems that affect slippage into the −1 frame (17).

Here, the role of message:anticodon pairing stability in frameshifting is explored for *prfB*. *prfB* is ideal for a study of the effects of tRNA:message stability because that frameshift site is simpler than those for simultaneous-slippage, because it

requires that a single tRNA rephase and establish a new frame while in a single ribosomal coding site. Furthermore, the sequences and decoding properties of *E.coli* tRNAs are relatively well characterized [see, eg., (36)], which facilitates assignment of frameshifting properties to specific anticodon bases. Frameshift frequencies are determined for thirty-two variants. tRNA:message pairing is shown to greatly affect frameshift frequency. The data suggest that wobble pairing to message U facilitates slippage into the +1 frame, and that stability in the +1 frame facilitates translation in that phase.

MATERIALS AND METHODS

The parental *prfB/lacZ* reporter plasmid used for site-directed mutagenesis of the *prfB* frameshift site is plac/RF-UAG and has been described (20). Site-directed mutagenesis was performed as described (37). The oligonucleotide used for mutagenesis is: GATCCCGTAGCTANNNATACCCCCTAAGGA, where NNN are ambiguous nucleotides corresponding to the randomized codon at the frameshift site. This oligonucleotide is complementary to the frameshift site depicted in Figure 1. β-galactosidase activities were determined using modifications of the Miller assay as described (38).

RESULTS AND DISCUSSION

A set of thirty-two *prfB/lacZ* fusion variants

A set of 32 variants of a *prfB/lacZ* fusion were created, each of which requires a different set of tRNA:message base pairs in the +1 frame. To achieve this variety, the codon XYZ in Figure 1 was randomized by site-directed mutagenesis. For +1 frameshifting, peptidyl-tRNAs must rephase from their cognate in-phase codon, XYZ, onto YZU at XYZU (Figure 1). Frameshifting occurs during the pause prior to translation of the UAG codon 3′ to XYZ. Because all alleles have this same UAG codon 3′ to XYZ, I assume that the pause is similar for all alleles.

For each allele, the message:anticodon base pairs that are expected to form in the +1 frame are listed in Table 1. At the first position of the +1 rephased complex, 14 of the 16 possible base pairs occur; missing are G:G and A:U message:anticodon pairs. At the second position, 15 of the 16 possible combinations occur; missing is G:C. At the third position, the message nucleotide is always the 'U' from the UAG codon 3′ to the XYZ codon. For various alleles, this U is paired with U, G, C, and the modified bases Q, S, D, E and F, which are identified in the legend to Table 1. Adenine is absent because this nucleotide is not present in the wobble position of mature tRNA. Inosine is also not represented in our data set. Only one *E.coli* tRNA, $tRNA_1^{Arg}$, contains inosine, and none of its respective codons (CGU, CGC and CGA) were obtained from the mutagenesis.

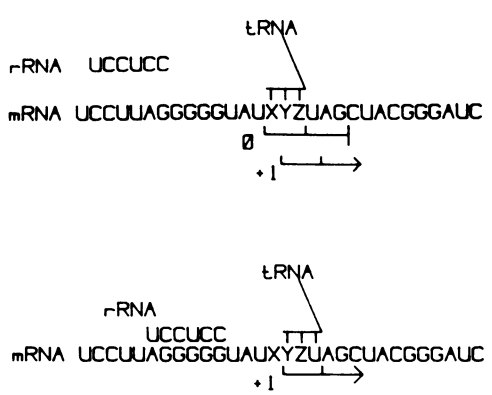

Figure 1. tRNA:message rephasing occurs by slippage of the peptidyl-tRNA paired at XYZ rightward one nucleotide to pair at YZU. **A:** Prior to the shift, peptidyl-tRNA is paired to the XYZ in-phase codon. The message triplet in the A site is UAG. The message purine run upstream of the frameshift site is not paired to the anti-Shine–Dalgarno of 16 S rRNA. **B:** During frameshifting, the anti-Shine–Dalgarno sequence of 16 S rRNA pairs to the run of purines upstream of the shift site and the peptidyl-tRNA rephases onto YZU. Translation of the next triplet, AGC, allows for the synthesis of β-galactosidase in the +1 phase.

Frameshift frequencies of the thirty-two alleles

The frameshift-dependent β-galactosidase activities for these 32 alleles are ranked in order of activity in Table 1. β-galactosidase activities range over 1000-fold, from a high of 8047 β-galactosidase units down to 7 units. The activities for six alleles (CUU, GUU, UUA, GUA, GUG, and AUA) are similar to those from similar *lacZ/prfB* fusions studied (26). Also consistent with that and other previous studies, the frameshift frequencies in Table 1 appear to depend on number of standard base pairs in the +1 rephased complex. For example, all but one of the alleles that frameshifts more frequently than 5% (≥ 833 units) have three standard pairs. Comparisons among alleles in this large set may facilitate understanding how specific base pairs in the rephased complex affect frameshift probability. Alleles that have U in the third position of the in-phase triplet generally frameshift more often than other alleles. Below, this shift-prone tendency for U-ending alleles is attributed to weaker wobble pairing at the in-phase codon that enhances slippage into the +1 frame. Because this difference may not be related to stability of the rephased complex, I restrict specific comparisons to alleles within either the U-ending or C/A/G-ending sets.

U-ending alleles with activities ≥ 833 units: The acceptability of first position U:G is demonstrated by CUU, which is the codon at the natural, *prfB* site and gives the highest activity of the entire set. The acceptability of U:G in the first position has been previously noted for *prfB* (19,26) and for Ty (23). CCU has 1/10 the activity of CUU, which suggests that in the center position U:G pairs are less acceptable than Watson:Crick pairs.

U-enders with activities of ≤ 260 units: UGU, CAU and AUU all have similar activities. It seems odd that AUU, whose only mispair is a first position U:U, is as affected as UGU and CAU, which have first position pairs between A and G, and second position py:py pairs. First position pairs between G and A may not be detrimental to frameshifting. A first position G:A pair occurs in the −1 realigned complex of the natural equine arteritis viral frameshift (8). The same pair only slightly reduces frameshift efficiencies for variants of a coronavirus frameshift site (30). In addition, first:third pairs between A and G are marginally stable in anticodon:anticodon complexes (32). The central py:py pairs may not seriously destabilize +1 complexes for UGU and CAU. In support of this idea are observations that central U:U pairs do not destabilize anticodon:anticodon complexes (32). The first position U:U for AUU may be as inhibitory as the combinations of two unusual, but not terribly detrimental, base pairs for UGU and CAU.

The rephased complexes GAU (57 units) and CAU (250 units) have identical pairs, except at the first position. The five-fold higher activity for CAU suggests that A:G is more acceptable than A:C in the first position. UCU differs from GCU also only

at the first position. The several-fold greater activity for UCU might, therefore, be attributed to the greater acceptability of C:A over C:C pairs in the first position.

C/A/G-enders: Only CCC, which forms three standard pairs in the +1 complex, has relatively high activity. The next highest activities are observed for UUC, GCC, and ACC, which each have two normal pairs and a one mispair. For UUC, the central C:A pair apparently renders this complex relatively unstable. GCC and ACC have first position C:C and C:U pairs, respectively. The detrimental effect of C:C was noted for above for GCU. The relatively low activity for ACC seems out of place because its only apparent defect is a first position C:U; above a first position U:C allows efficient frameshifting for GUU. Perhaps the pairs between U and C are sensitive to context or base pair orientation. Others report irregular behaviors for pairs between U and C during frameshifting by coronavirus variants (30). There, first position C:U decreases frameshifting either 4 or 14-fold for two different constructs, while first position U:C causes a 17-fold reduction in frameshift efficiency.

C/A/G-enders with activities ≤ 57 units: Here, all alleles have at least two mispairs and/or a central purine:purine pair. Central purine:purine pairs may be incompatible with frameshifting in *prfB*. Not one of nine alleles that has a purine:purine mispair in the center has an activity greater than 37 units. Included in this group is CCA, which has a very low activity (20 units) despite having the potential for normal pairs at the first and third positions. Perhaps the bulky central A:G pair hinders normal pairing at one or both neighboring positions. A measure of the relative acceptability of central pairs can be made from comparisons of CCC, UUC and CCA, which all have good pairing at the two outer positions but differ in the central position: C:G > C:A > A:G.

To summarize these data: With respect to acceptability in the first position, Watson:Crick = U:G > G:A = A:G > C:A > C:C. In addition, U:U is detrimental, but cannot be ranked because the data do not permit the necessary direct comparisons. This order of acceptability is in rough agreement with that observed at the first position of the P site in a coronavirus (30). In the second position, Watson:Crick pairs are clearly best; U:G, U:U, and U:C are all acceptable; C:A is significantly inhibitory and central purine:purine pairs are worse.

The very poor activities for alleles with central purine:purine pairs contrasts with observations for acceptable P site pairing at simultaneous-slippage frameshift sites. A central G:A pair is required in the P site during programmed −1 frameshifting in bacteriophage T7 gene 10 (29). In addition, RSV variants that require P site A:A and A:G central pairs have only about 5-fold reduced frameshifting (6). These sites frameshift orders of magnitude more frequently than the *prfB* alleles that have central purine:purine pairs. The dramatic difference in activities may

Table 1. Frameshifting into the +1 frame for *prfB* alleles

Shift site	Message:anticodon pairs for the +1 realignment	β-gal	F(%)	Estimated stability A	B
UAU CUU UAG	U:G, U:A, U:G	8047	48	−4	−4
CCC	C:G, C:G, U:G	1874	11	−5	−7
UUU	U:A, U:A, U:G	1811	11	−5	−5
GUU	U:C, U:A, U:V*	1746	10	−3	−3
CCU	C:G, U:G, U:V*	833	5	−3	−4
UGU	G:A, U:C, U:G	260	1.5	−1	−1
CAU	A:G, U:U, U:Q	250	1.5	−2	−2
AUU	U:U, U:A, U:G	240	1.4	−3	−3
UUC	U:A, C:A, U:G	103	0.6	−2	−2
GCC	C:C, C:G, U:V	84	0.5	−2	−3
UCU	C:A, U:G, U:V	80	0.5	−1	−1
ACC	C:U, C:G, U:G	69	0.4	−3	−4
GAU	A:C, U:U, U:Q	57	0.3	−2	−2
CAG	A:G, G:U, U:C	53	0.3	0	0
CAA	A:G, A:U, U:S	50	0.3	−3	−3
GAA	A:C, A:U, U:S	41	0.2	−3	−3
GAG	A:C, G:U, U:S	40	0.2	−1	−1
GUC	U:C, C:A, U:G	39	0.2	0	0
GAC	A:C, C:U, U:Q	38	0.2	−1	−1
AGC	G:U, C:C, U:G	38	0.2	−1	−1
UUA	U:A, A:A, U:D	37	0.2	0	0
GCU	C:C, U:G, U:V	36	0.2	0	0
UAC	A:A, C:U, U:Q	35	0.2	−1	−1
GGA	G:C, A:C, U:E	30	0.2	−1	−2
GUG	U:C, G:A, U:V	26	0.15	1	1
ACG	C:U, G:G, U:V	23	0.14	1	1
CCA	C:G, A:G, U:V	20	0.12	−1	−2
ACA	C:U, A:G, U:V	14	0.08	1	1
CUG	U:G, G:A, U:C	12	0.06	1	1
UCA	C:A, A:G, U:V	10	0.06	1	1
AUA	U:U, A:A, U:F	9	0.05	2	2
GUA	U:C, A:A, U:V	7	0.04	1	1

A, G, C and U are the standard bases; Q = queuosine; S = 5-methylaminomethyl-2-thiouridine; and V = uridine-5-oxyacetic acid; D is an unidentified derivative of A that pairs with G and A; E is an unidentified derivative of U that pairs with A; and F is a modified pyrimidine that pairs with A. *GUU and CCU may be also be read by relatively rare tRNAs that have G in the wobble position; U:G and U:V pairs have similar stabilities (32). β-gal are β-galactosidase units. F is frameshift frequency, relative to a pseudowildtype *lacZ* allele that produces 16,800 β-galactosidase units (20). Estimated Stabilities are described in Table 2.

Table 2. Relative stabilities of base pairs used to estimate tRNA:message stability

Message:anticodon base pair	Relative stability Smaller values mean more stable
G:C*, C:G*, A:U, U:A	−2
U:Q, U:S, U:V; first and third position U:G, G:U; and central U:U	−1
U:C, C:U, U:D, U:E, U:F; first and third position U:U; first position A:C, C:A, pu:pu; central U:G, G:U	0
C:C pairs in any position; C:A, A:C pairs in the middle position	+1
Pu:Pu pairs in the middle position	+2

Where A, G, C and U are the standard bases; Q = queuosine; S = 5-methylaminomethyl-2-thiouridine; and V = uridine-5-oxyacetic acid; D is an unidentified derivative of A that pairs with G and A; E is an unidentified derivative of U that pairs with A; and F is a modified pyrimidine that pairs with A. Most scores for individual pairs are assigned from the relative stabilities of tRNA:tRNA complexes (32). A few pairs were not studied by Grosjean; those assignments are explained below.

The most stable pairs are assigned a relative stability of −2, where like free energy differences, lower numbers mean more stable. G:C and A:U pairs have similar strengths in anticodon:anticodon complexes, so we assigned the most stable score of −2 to Watson:Crick pairs involving these standard nucleotides. In the first and third positions, U:G and G:U wobble pairs are scored as −1. In the third position, wobble pairs U:V and U:Q are also scored as −1. Also receiving this moderately stable score are U:U pairs in the middle position, where they are unexpectedly stable. U:S pairs are relatively stable, therefore those pairs are scored as −1.

Pairs that are marginally stable are scored as '0'. These pairs include first position purine:purine pairs and pairs between C and A. In the center position, G:U pairs are only marginally stable. Grosjean does not provide information on the stability of complexes with U:C, or for U:U in the first and third positions, or for U pairing with the modified bases D, E, and F in the wobble position. It is assumed that they are marginally stable, or at least do not destabilize complexes.

Central A:C, C:A and C:C pairs in any position are assigned a score of +1 because they destabilize tRNA:tRNA complexes.

Central purine:purine mispairs are assigned a score of +2 because those pairs are more severe than other central mispairs.

*For stability estimates assuming that pairs with G and C are more stable than those between A and U, G:C and C:G are assigned a score of −3.

reflect different roles for P site tRNAs in the two frameshift mechanisms. In general, for RSV mutations that disrupt stable pairing in the P site are less detrimental than mutations in the A site (6), as if a requirement for P site stability is relaxed because the responsibility for establishment of the new frame is shared with the A site tRNA. For *prfB*, in contrast, a central purine:purine pair may severely inhibit that single tRNA from establishing the +1 frame.

Within Table 1 consistent patterns emerge: Codons that have third position uridine are more shift-prone than other codons. High frameshift frequencies require stable +1 pairing. Decreases in frameshifting correspond with the number, type and positions of mispairs. These consistencies suggest that this large data set may be useful for exploring relationships between tRNA:message stability and frameshift frequency.

Relationships between frameshifting and +1 complex stability

Frameshifting requires that the triplet following the +1 complex be translated (AGC in Figure 1), which allows for continued protein synthesis in the +1 phase. A plausible role for tRNA:message stability is to hold the frame long enough for translation of that rephased A site triplet. If so, then frameshift frequency should be systematically related to stability. The exact relationship cannot be defined because frameshift frequency is also determined by the probability of the peptidyl-tRNA slipping from its codon in the '0' frame into the +1 frame. Precise values for +1 stability are also not known. However, an exact model and precise stability values are not necessary to observe general relationships. Because Table 1 contains a large data set in which both parameters vary over wide ranges, strong general relationships will be apparent in plots between frameshifting and reasonably approximate indices of +1 complex stability.

Relative stabilities of P site tRNA:message complexes are estimated from the data of Grosjean and collaborators (32), who estimate relative strengths of base pairs in anticodon:anticodon complexes. For two reasons, the Grosjean data provide the best available models of tRNA:message interaction. First, the Grosjean data are obtained for complexes with anticodons, whose pairing properties are strongly affected by structural features of anticodon arms. For example, anticodon function is modulated affected by base modifications (39), by inclusion within a hairpin loop (40), and by the identities of neighbor nucleotides in the anticodon arm (41). All of these structural features are neatly accounted for in Grosjean's study of paired tRNAs. Second, the Grosjean data provide estimates of the relative stabilities of many different cognate and noncognate base pairs at both the outer and central anticodon positions. Again, such detailed information about anticodon pairing is not available elsewhere. An unexpected finding from the Grosjean study is that pairs between G and C are not detectably more stable than pairs between A and U in tRNA:tRNA complexes. Because this surprising finding has not been confirmed for tRNA:message pairs, I show two analyses.

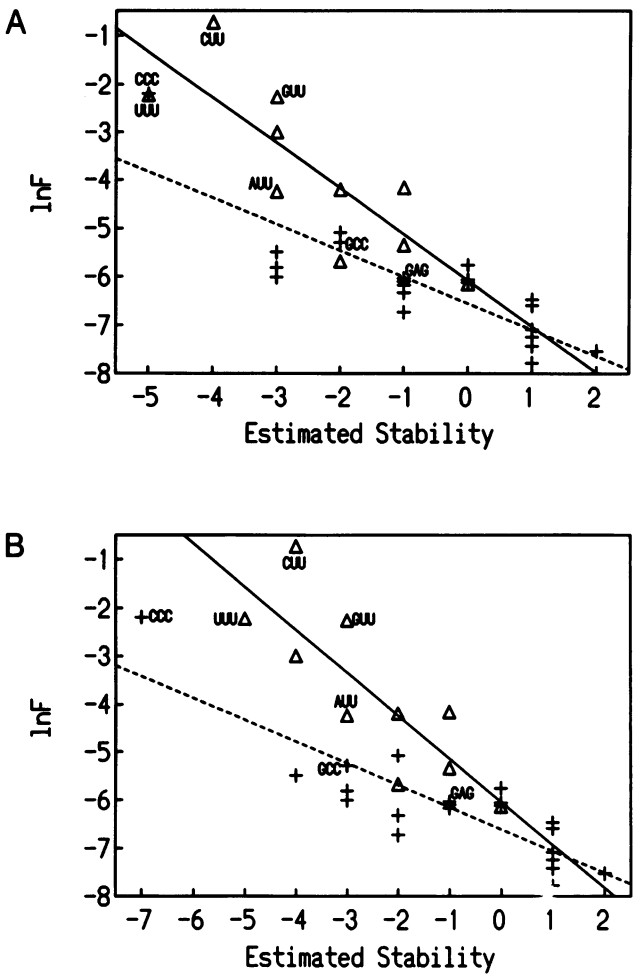

Figure 2. The natural logs of frameshift frequencies (F) from Table 1 are plotted *versus* estimated stabilities of the +1 rephased complexes (Table 1). Alleles in which the in-phase codon has a third position uridine (triangles, solid lines), other alleles (crosses, dashed lines). **A**: stabilities estimated assuming all Watoson:Crick pairs are equally stable. **B**: stabilities estimated assuming that G:C and C:G are more stable than pairs between A and U.

One assumes equivalence of Watson:Crick pairs, the other assumes that G:C and C:G pairs are more stable than A:U and U:A pairs. Both analyses support the same conclusions.

One limitation of the Grosjean assay is that stability differences between very unstable mispairs cannot be distinguished. For example, most central mispairs prevent tRNA:tRNA complex formation and, therefore, differences in stabilities among those complexes could not be measured. Thus the Grosjean data cannot account for the observation from Table 1 that alleles required to form central purine:purine pairs in the +1 frame have very low frameshift frequencies. To compensate for this deficiency in the Grosjean data, I assign an especially unstable score to central purine:purine pairs. All base pair assignments are listed in Table 2. Relative stabilities of entire +1 complexes are taken as sums of values of component pairs. Estimated complex stabilities for each allele are listed in the righthand columns of Table 1.

To detect relationships between stability and frameshifting, in Figures 2 the natural logs of the frameshift frequencies from Table 1 are plotted *versus* estimated +1 stability. In Figure 2A, stabilities are estimated assuming that all Watson:Crick pairs have equal strengths. In Figure 2B, stabilities are estimated assuming that pairs between G and C are more stable than A:U or U:A pairs. Despite these differences in +1 stability estimates, two general features are quite clear in both figures. First, alleles in which the in-phase codon ends in U (triangles) plot consistently higher than other alleles (crosses). The second feature is that lnF is related to estimated +1 stability. These two general features are also apparent in many other plots in which specific base pair strengths are systematically either increased or decreased by one unit (not shown). Their persistence strongly suggests that these two trends are not artifacts of the base pair strength assignments; instead, these two features are likely related to frameshifting. The precise relationships between stability and lnF are not known; however, because it facilitates description of the plots, linear regression 'best fit' lines are plotted. Lines are calculated separately for U-enders and C/A/G-enders to facilitate distinction between these groups. All lines have high linear correlation coefficients ($r \geq 0.82$), which shows that best fit lines are suitable aids for discussion.

Variation about the best fit lines may have more than one cause including, no doubt, imprecision of the stability estimates. Another plausible cause for variation is idiosyncratic properties of specific tRNAs or codons that may affect frameshift frequency independently of +1 stability. I first discuss the two general features evident in these plots. Then in the final section I discuss one idiosyncratic allele, CUU, which has an especially high frameshift frequency and is the codon at the natural *prfB* frameshift site.

Rephased stability correlates with frameshift frequency

Both U-ending and C/A/G-ending codon groups show apparent relationships between frameshift frequency and stability that span the entire range of frameshift activity. Those relationships strongly argue that +1 stability directly contributes to frameshift frequency. Variation about the best fit lines is similar for both U-enders and C/A/G-enders, which suggests that lnF for both groups is similarly dependent on +1 complex stability. The simplest interpretation is that stability of the rephased P site tRNA:message complex is an important determinant of the probability that translation will continue in the +1 phase.

tRNAs paired to codons that have third position uridine may be more likely to slip into the +1 frame

As a group, the U-ending codons are clearly more shift-prone than C/A/G-ending codons. For example, all U-enders plot above all C/A/G-enders at estimated stabilities of −3 and −1 (Figures 2). Relatively high frameshifting for U-enders suggests that an effect associated with the third position U is superimposed on the dependence on estimated stability. Two explanations for this effect are (1) systematic underestimation of relative +1 stabilities for U-enders, and (2) increased slippage into the +1 frame from codons with third position uridine.

The first possibility is less likely. If rephased complexes for U-ending alleles are more stable than expected from the Grosjean data, then stabilities of base pairs that are specific to those complexes are underestimated. The only pairs unique to +1 complexes for U-enders are central pairs with message U, which are already assigned relatively stable scores (Table 2). All four central U:N pairs occur among the U-enders, and the high

activities of U-enders is not associated with any specific central U:N pair. Consider central U:A pairs for example. CUU, UUU, GUU and AUU will all form a central U:A pair following slippage into the +1 phase, but those four alleles bracket the best fit lines for U-enders in Figures 2. Therefore, the stability of central U:A is not misjudged relative to the other central U:N pairs. In fact, no particular central U:N pair is associated with a distinctive frameshift tendency. Thus, to attribute the high activities of U-enders to underestimated +1 stabilities, one must assume that *all* central U:N pairs are considerably more stable than expected from the Grosjean data (Table 2). This would require for example that central U:U is at least as stable, and that U:A is substantially more stable, than central C:G, G:C and A:U, which seem unlikely.

A simpler explanation for the high activities of U-enders is that wobble pairing to third position uridine facilitates tRNA:message slippage into the +1 frame. In the alleles examined here, third position uridines are decoded by wobble pairing to either G, Q, or V. All of those wobble pairs are considerably weaker than Watson:Crick base pairs in anticodon:anticodon complexes (32). In contrast, most of the C/A/G-ending codons in Table 1 have relatively stable Watson:Crick pairing in the third position of their respective preshift complexes. Others argue that unstable pairing in the A site facilitates frameshifting during simultaneous-slippage (7,17,30,42). Evidence supporting that argument is that the highest frameshift frequencies occur at sites where the third position base pair is likely to be relatively weak (17,30,34), and/or read by unmodified tRNAs (35). Additionally, replacement of A- or U-rich A site codons with G- or C-rich triplets reduces frameshift frequency, as if the stronger C:G and G:C pairs inhibit slippage (7,30). In the current study, that elevated frameshift frequencies are associated with weaker U-wobble pairing strongly supports the idea that unstable preshift complexes can facilitate frameshifting and extends that proposition to the *prfB* frameshift mechanism. Perhaps weaker U-wobble pairing eases dissociation of the anticodon from the in-phase triplet so that it may sample pairing in the +1 frame.

It is worth note that while it may generally facilitate frameshifting, third position U is not essential for high frequency frameshifting. For example, CCC has a frameshift frequency of 11% (Table 1). Additionally, in a similar study GGG frameshifts with an efficiency of about 6% (26). In these cases, the lack of a third position U may be compensated by particularly strong rephased complexes (see CCC in Figures 2). I also point out that third position wobble pairs to message G and C are not detectably associated with increased frameshifting. GAG and GCC, which have G:S and C:V wobble pairs in their respective preshift complexes, do not plot distinctly from other C/A/G-ending alleles in Figures 2.

RF2 may have selected a uniquely shift-prone site for programmed frameshifting

The allele with the natural *prfB* shifty codon, CUU, frameshifts greater than four times more often than every other allele in the set (Table 1). One weaker frameshifter is UUU, is a U-ender that is predicted to be more stable than CUU in the +1 frame. Other less frequent shifters include various other pyrimidine runs such as CCC, CCU, UCU and UUC. Clearly, at least at the *prfB* programmed shift site, CUU is frameshift-prone beyond that predicted from having a third position U and a relatively high estimated rephased stability. One feature that may contribute to especially high frameshifting is greater than expected stability of the rephased complex. We previously compared alleles in which CUU is followed by codons starting with either U or C, and which, therefore, require either U:G or C:G third position pairs in their respective rephased complexes (21). Frameshift frequency is not lower when a U:G pair is required, which suggests that frameshifting at CUU is not limited by rephased complex stability. Molecular features that may contribute to unusual rephased stability are not clear, but it may be relevant that the tRNA cognate to CUU, $tRNA_2^{Leu}$, does not contain a bulky modification at the base 3′ to the anticodon (m_1G37). Because bulky adducts at base 37 may restrict first position wobble pairing (43,44,39), it is plausible that the first position U:G pair for $tRNA^{Leu}$ is more stable than predicted from the Grosjean data, in which U:G strengths were estimated using tRNAs that have threonine in carbamoyl linkage to the 6-amino group of A37 (32). Whatever its origin, the apparently unique tendency for high frequency frameshifting may have led to the use of CUU for the *prfB* autoregulatory mechanism.

ACKNOWLEDGEMENTS

I thank Beverly Gross and Karen Scharret for technical assistance. I am also grateful Renée Schwartz and Mike Yarus for helpful discussions. This work was supported by NSF grant DMB-8904708.

REFERENCES

1. Atkins, J.F., Weiss, R.B., and Gesteland, R.F. (1990) *Cell* **62**, 413–423.
2. Hatfield, D., and Oroszlan, S. (1990) *TIBS* **15**, 186–190.
3. Craigen, W.J., Cook, R.G., Tate, W.P., and Caskey, C.T. (1985) *Proc. Natl. Acad. Sci., U.S.A.* **82**, 3616–3620.
4. Mellor, J., Fulton, S.M., Dobson, W.W., Kingsman, S.M., and Kingsman, A.J. (1985) *Nature* **313**: 243–246.
5. Clare, J., and Farabaugh, P. (1985) *Proc. Natl. Acad. Sci., U.S.A.* **82**: 2829–2833.
6. Jacks, T., Madhani, H.D., Masiarz, F.R., and Varmus, H.E. (1988) *Cell* **57**, 447–458.
7. Dinman, J.D., Icho, T., and Wickner, R.B. (1991) *Proc. Natl. Acad. Sci., U.S.A.* **88**: 174–178.
8. den Boon, J.A., Snidjer, E.J., Chirnside, E.D., DeVries, A.A.F., Horizinek, M.C.,and Spaan, W.J.M. (1991) *J. Virol.* **65**, 2910–2920.
9. Blinkowa, A.L., and Walker, J.R (1990) *Nucleic Acids Res.* **18**, 1725–1729.
10. Flower, A.M., and McHenry, C.S. (1990) *Proc. Natl. Acad. Sci., U.S.A.* **87**, 3713–3717.
11. Tsuchihashi, Z., and Kornberg, A. (1990) *Proc. Natl. Acad. Sci., U.S.A.* **87**, 2516–2520.
12. Condron, B.G., Atkins, J.F., and Gesteland, R.F. (1991) *J. Bacteriol.* **173**: 6998–7003.
13. Sipley, J., Stassi, D., Dunn, J., and Goldman, E. (1991) *Gene Expression* **1**: 127–136.
14. Vögele, K., Schwartz, E., Welz, C., Schlitz, C., and Rak, B. (1991) *Nucleic Acids Res.* **19**, 4377–4385.
15. Ten Dam, E.B., Pleij, C.W.A., and Bosch, L. (1990) *Virus Genes* **4**: 121–136.
16. Brierley, I., Rolley, N.J., Jenner, A.J., and Inglis, S.C. (1991) *J. Mol. Biol.* **220**: 889–902.
17. Tsuchihashi, Z. (1991) *Nucleic Acids Res.* **19**, 2457–2462.
18. Tu, C., Tzeng, T-H., and Bruenn, J.A. (1992) *Proc. Natl. Acad. Sci., U.S.A.* **89**, 8636–8640.
19. Craigen, W.J., and Caskey, C.T. (1986) *Nature* **322** 273–275.
20. Curran, J.F., and Yarus, M. (1988) *J. Mol. Biol.* **203**, 75–83.
21. Curran, J.F., and Yarus, M. (1989) *J. Mol. Biol.* **209**, 65–77.
22. Sipley, J., and Goldman, E. (1992) *Proc. Natl. Acad. Sci., U.S.A. in press.*
23. Belcourt, M.F., and Farabaugh, P.J. (1990) *Cell* **62**: 339–352.
24. Pedersen, W.T., and Curran, J.F. (1991) *J. Mol. Biol.* **219**: 231–241.

25. Weiss, R.B., Dunn, D.M., Dahlberg, A.E., Atkins, J.F., and Gesteland, R.F. (1988) *EMBO J* **7**, 1503–1507.
26. Weiss, R.B., Dunn, D.M., Atkins, J.F., and Gesteland, R.F. (1987) *Cold Spring Harbor Symp. Quant. Biol.* **52**, 687–693.
27. Jacks, T. Townesly, K., Varmus, H.E., and Majors, J. (1987) *Proc. Natl. Acad. Sci., U.S.A.* **84**: 4298–4302.
28. Moore, R., Dixon, M., Smith, R., Peters, G., and Dickson, C. (1987) *J. Virol.* **61**: 480–490.
29. Condron, B.G., Gesteland, R.F., and Atkins, J.F. (1991) *Nucleic Acids Res.* **19**: 5607–5612.
30. Brierley, I., Jenner, A.J., and Inglis, S.C. (1992) *J. Mol. Biol.* **227**, 463–479.
31. Freier, S.M. Kierzek, R. Jaeger, J.A., Sugimoto, N., Caruthers, M.H., Neilson, T., and Turner, D.H. (1986) *Proc. Natl. Acad. Sci., U.S.A.* **83**: 9373–9377.
32. Grosjean, H.J., DeHenau, S., and Crothers, D.M. (1978) *Proc. Natl. Acad. Sci., U.S.A.* **75**, 610–614.
33. Chamorro, M., Parkin, N., and Varmus, H.E. (1992) *Proc. Natl. Acad. Sci., U.S.A.* **89**: 713–717.
34. Weiss, R.B., Dunn, D.M., Shuh, M., Atkins, J.F., and Gesteland, R.F. (1989) *The New Biologist* **1**: 159–169.
35. Hatfield, D., Feng, Y.-X., Lee, B.J., Rein, A., Levin, J.G., and Oroszlan, S. (1989) *Virology* **173**: 736–742.
36. Björk, G.R. (1987) In Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M., and Umbarger, H.E., (eds.), *Escherichia coli* and *Salmonella typhimurium*: Cellular and molecular biolgy, American Society for Microbiology Publications, Washinton D.C., pp. 719–731.
37. Curran, J.F., and Yarus, M. (1986) *Proc. Natl. Acad. Sci., U.S.A.* **83**, 6538–6542.
38. Curran, J.F., and Yarus, M. (1987) *Science* **238**, 1545–1550.
39. Björk, G.R. (1992) In: Hatfield, D.L., Lee, B.J. and Pirtle, R.M. (eds.), Transfer RNA in Protein Synthesis, CRC Press, Boca Raton, pp. 23–85.
40. Grosjean, H., Söll, D.G., and Crothers, D.M. (1976) *J. Mol. Biol.* **103**: 499–519.
41. Yarus, M.,Cline, S.W., Wier, P., Breeden, L., and Thompson, R.C. (1986) *J. Mol. Biol.* **192**: 235–255.
42. Weiss, R.B., Dunn, D.M., Atkins, J.F., and Gesteland, R.F. (1990) *Prog. Nucl. Acids Res. Mol. Biol.* **39**, 159–183.
43. Jukes, T.H. (1973) *Nature* **246**: 22–26.
44. Wilson, R.K., and Roe, B.A. (1989) *Proc. Natl. Acad. Sci., U.S.A.* **86**: 409–413.