Slippery Runs, Shifty Stops, Backward Steps, and Forward Hops: −2, −1, +1, +2, +5, and +6 Ribosomal Frameshifting

R.B. WEISS,* D.M. DUNN,* J.F. ATKINS,*† AND R.F. GESTELAND*

*Department of Human Genetics and Howard Hughes Medical Institute, University of Utah Medical Center, Salt Lake City, Utah 84132; †Department of Biochemistry, University College, Cork, Ireland

Frameshift mutations frequently express residual levels of gene activity; that is, they are often leaky. This leakiness can be used as a tool to define the functional components that affect the reading frame during gene expression (Atkins et al. 1972; Fox and Weiss-Brummer 1980; Weiss and Gallant 1983). Recent technological advances in the capability to efficiently build synthetic DNA sequences have facilitated the construction of small, defined "frameshift windows." These windows are regions where frameshift events can be detected and measured. The cloned synthetic window is fused onto the 5′ coding region of an active β-galactosidase gene that provides a sensitive monitor for the frameshift events. Fusions onto the *lacZ* gene have the advantages of simple colorimetric assays for β-galactosidase activity and little or no effect of the fused sequence on the specific activity or stability of the enzyme (Miller and Albertini 1983). A frameshift event also leaves a clue to its character in the protein sequence translated from the window's mRNA sequence. Recovery of the frameshift-containing β-galactosidase in sufficient yield and purity for determining its amino-terminal sequence provides hard evidence for the occurrence of a frameshift, and this sequence may be used to infer the kind of event generating the loss of reading frame.

Frameshift windows are defined at their 3′ border by a stop codon in the zero frame relative to the translation start and at their 5′ border by a stop codon in the monitored outgoing frame. Ribosomes can only enter the window by translating in the zero frame, and they can only exit by shifting to the −1 or +1 frame, only one of which results in the production of active β-galactosidase. A 2p, 3p, or 4p designation is given to windows that monitor the −1, 0, or +1 frame. Frameshift windows also have as a formal property the ability to determine the level (replication or transcription versus translation) at which the shift occurs. This property involves "reframing" the window's sequence so that the sequence remains the same but the frames in which the ribosome translates through it have both been shifted. Translational frameshifts should be sensitive to such reframing since ribosomes read frame, but the levels of transcriptional shifts or genetic reversion of the frameshift should not be altered by reframing, since the enzymatic machinery involved in these processes is not known to detect the reading frame.

This paper is confined to several key results from a construction project initiated several years ago, in which many *lacZ* frameshift windows have been built in and partially characterized. We focus on high-level translational shifts (> 0.1% of in-frame levels) here for two reasons: (1) They occur quite frequently and unexpectedly during construction of specific sequences and (2) interest in the mechanisms responsible for high-level shifts has recently been augmented by several cases of high-level ribosomal shifts that are necessary for proper gene expression and whose efficiency appears to be programmed in the mRNA (Dunn and Studier 1983; Craigen et al. 1985; Clare and Farabaugh 1985; Jacks and Varmus 1985; Mellor et al. 1985; Shimotohno et al. 1985; Jacks et al. 1987; Moore et al. 1987). The synthetic *lacZ* high-level shifts described here all occur within strings of repeated nucleotides (henceforth called strings) or between overlapping or nonoverlapping homologous codons. These strings and variably spaced homologous codons share the potential for correct decoding in the zero frame followed by a shift to a new frame that can be stabilized by good codon:anticodon pairing. This investigation has confirmed the location of several of these shifts and has begun to reveal a variety of events ranging from −2 to +6 nucleotides shifted relative to the zero frame. In all cases, good codon:anticodon pairing is available to the alleged shifting tRNA in both the incoming and outgoing frames, giving the paramount definition of the reading frame to the tRNA:mRNA interaction. However, mRNA sequence contexts can enhance the frequency of certain types of shifts, and the consequences of two mRNA contexts on these types of shifts are described.

EXPERIMENTAL PROCEDURES

Construction of *lacZ* Frameshifts

Oligonucleotides were synthesized by an Applied Biosystems 380A or B DNA synthesizer, and the crude material was cloned into a uniquely restricted pBR322 vector containing an engineered derivative of the *lacZ* gene. This places the synthetic sequence two codons downstream from the *Z* translation start, within in a region nonessential for β-galactosidase function. Inserts were verified by dideoxy sequencing of the plasmid DNA (Chen and Seeburg 1985). The primary *lacZ*

vector has *lac* operon sequence from codon 5 of *Z* through codon 66 of *Y*, replacing the *Eco*RI-*Ava*I section of pBR322. The *Hin*dIII/*Apa*I cloning sites and the ribosome-binding site (based on the sequence of *Escherichia coli* lipoprotein ribosome-binding site) were constructed synthetically. A 269-bp *Bam*HI fragment from plasmid containing a *Tac* promoter was inserted just upstream of the synthetic ribosome-binding site, and an 800-bp fragment containing a T7 promoter was inserted between the *Eco*RI site of the pBR backbone and the upstream *Bam*HI site bordering the *Tac* promoter. The bacterial strain used is *E. coli* K-12 SU1675 del- *prolac*, *recA56*/F′I^Q (rec$^-$ derivative of CSH26; Miller 1972).

Purification and Amino-terminal Sequencing of Frameshifted β-Galactosidase

Cell extracts of saturated cultures of *E. coli* SU1675 *prolac*, *recA56* grown in superbroth (Davis et al. 1980) were prepared by sonication or freeze-thawing in 50 mM KPO_4 (pH 7.4), 150 mM NaCl, 0.1% Tween-20, and 10 mM β-mercaptoethanol. β-Galactosidase was purified by either passage through an anti-β-galactosidase affinity column (Protosorb, Promega Biotech) or by precipitation with anti-β-galactosidase antibody (Cooper Biomedical) and protein-A-Sepharose CL-4B beads. When necessary, the sample was further purified by high-pH electroblotting from a 7.5% SDS-acrylamide gel onto an aminopropyl-glass fiber filter (Aebersold et al. 1986). The GF-C filters were prepared by heating in neat trifluoroacetic acid (TFA) (Pierce, Sequanal grade) on a heating block at 68–70°C for 45 minutes. They were air-dried in a hood on Whatman No. 1 sheets for 2 hours, and then treated with a freshly prepared solution of γ-aminopropyltriethoxysilane (Pierce, 5% in 95% aqueous acetonitrile) at room temperature for 5 minutes with agitation. The filters were then taken through three washes of acetonitrile (each filter being handled separately), blotted, and cured at 105–110°C for 1 hour; this procedure is based on that of Aebersold et al. (1986). The electroblotted filter was stained with 3,3-dipentyloxacarbocyanine iodide (Aebersold et al. 1986); the β-galactosidase band was excised from the filter and placed in the cartridge of an ABI 470a gas-phase protein sequencer equipped with an on-line ABI 120a high-performance liquid chromatography (HPLC) analyzer. The program 03RPTH was utilized, cartridge temperature set at 45°C, and 40% of the phenylthiohydantoin (PTH) amino acids from each cycle were chromatographed by the on-line HPLC.

β-Galactosidase Activity Measurements

Whole-cell assays were based on the procedure according to Miller (1972). Stationary phase cultures were diluted 1/40 into Luria broth (LB) + 2 mM isopropyl-β-D-thiogalactoside (IPTG), grown at 37°C with rapid shaking, and chilled to 0°C at an OD_{600} of 0.7–0.8. Whole-cell assays were done in a final volume of 1.0 ml Z-buffer, the assays were started with the addition of 0.2 ml of 4 mg/ml *O*-nitrophenyl-β-D-galactopyranoside and stopped with the addition of 0.5 ml of 1 M Na_2CO_3 (pH 11). All reactions were run at 28°C. The assay tubes were centrifuged before reading the OD_{420}.

RESULTS

Shifty Stops

Constructing frameshift mutations within the dispensable amino-terminal region of β-galactosidase takes advantage of the simple enzymatic assay for β-galactosidase activity (the range of in vivo detection spans at least five orders of magnitude); this permits easy measurement of the leakiness. Even low levels of frameshifted β-galactosidase produced from these constructs can be purified in one step by immunoaffinity chromatography, or in two steps using immunoprecipitation followed by SDS-acrylamide gel electrophoresis and electroblotting onto activated glass fiber filters. These techniques yield material suitable for amino-terminal sequencing, and the confinement of the frameshift window near the translation start puts this region within reach.

In an initial series of constructions, it was noted that placement of a string of repetitive nucleotides next to the 3′ stop codon border sometimes resulted in a sharp increase in the window's leakiness. This observation is shown in Table 1. When a GGG glycine codon is located 5′ to a stop codon (construct 4p101 CGG-GGG-UAA, Table 1) the reading frame shifts at the zero frame GGG by −2 nucleotides. This occurs at a level of approximately 2% relative to a similar in-frame construct (p900 series, Table 1). Comparison of the mRNA sequence of 4p101 with its protein sequence (Fig. 1a) leads to the inference that the glycine tRNA decoding the zero frame GGG shifts backwards by two nucleotides onto the 5′ overlapping GGG. This inference is strengthened by the decrease in the level of shifting observed when the first G in the 5′ overlapping GGG is altered (4p101 CGG-GGG-UAA:140 units versus 4p102 CAG-GGG-UAA:40 units and 4p103 CCG-GGG-UAA:12 units, Table 1). When the zero frame GGG is changed to a GGA glycine or GAA glutamate codon, −2 shifting is still observed at appreciable levels, and in the latter case, glutamate is incorporated at the shift site (data not shown), implying that a glutamate decoding tRNA reads the zero frame GAA glutamate codon and then shifts −2 onto the 5′ overlapping GGG glycine codon. Also, a sharp decrease in the level of shifting can be seen when the 3′ stop is changed to a sense codon in the 4p1100 series. Stop codons, especially UGA and UAA, elevate the level of this type of shift.

Enhancement of −1 frameshifting on 4 base G strings by a bordering 3′ stop codon is also seen (Table 1) by comparing 2p302 G-GGG-UAA:140 units with

Table 1. Frameshifting on Strings of Repetitive Nucleotides is Enhanced by a 3′ Stop Codon Border

Series		β-Galactosidase activity (whole-cell units)
	Met- Lys- Ser- Leu- Asp- Arg- Gly- [−2 shift] Gly- Lys- Gly-	
4p101	AUG-AAA-AGC-UUA-GAU-CGG-GGG-UAA-GGG-C	170
102	A	40
103	C	12
4p1103	AUG-AAA-AGC-UUA-GAC-CGG-GGG-UGA-CUG-UAA-G	230
1108	UAA	135
1122	UAG-GUG	40
1132	UGG-CUG-UAA	10
4p801	AUG-AAA-AGC-UUA-GAU-CGC-GCG-UGA-GGG-C	1
	Met- Lys- Ser- Phe- Asn- Leu- Gly- [−1 shift] Val-Lys-Gly-	
2p302	AUG-AAA-AGC-UUU-AAU-CUG-GGG-UAA-AGG-GC	140
2p2613	AUG-AAA-AGC-UUU-AAU-CUG-GGG-CUU-CAC-UAA-CGG-GC	8
2p403	AUG-AAA-AGC-UUU-AAU-CUU-GGG-UAA-AGG-GC	5
2p3901	AUG-AAA-AGC-UUU-AAC-UUA-AUC-UGG-GGC-UUC-ACU-ACC-UAA-	1
3p901	AUG-AAA-AGC-UUA-GAU-UGG-AAU-AAG-GGC-	10100
902	UC	10000
903	GG	4600

The ribosome-binding site, AUG start, and frameshift windows were derived from synthetic DNA inserts cloned into a plasmid-borne *lacZ* gene. The host is *E. coli* SU1675 F′I^{Q}. Cultures were grown in LB at 37°C with aeration from a 1:40 dilution of a fresh saturated culture until an OD_{600} of ~0.7; whole cell β-galactosidase assay conditions are described in Miller (1972), except assays were centrifuged to remove cell debris before measuring the OD_{420}. The amino acid sequences superimposed above certain constructs were determined by amino-terminal sequencing of the purified frameshift β-galactosidase from those constructs. The 5′ stop codon border is overlined and the 3′ stop codon border is underlined.

2p2613 G-GGG-CUU:8 units. In 2p301, the reading frame slips −1 at the zero frame GGG glycine codon (Fig. 1b). Construct 4p403 U-GGG-UAA:5 units suggests the importance of good base-pairing potential in the outgoing frame when compared with 2p302 G-GGG-UAA:140 units (Table 1). The influence of the zero frame 3′ stop codon in enhancing the rate of −2 and −1 shifting on G strings indicates that the shift is occurring during translation. However, the origin of a comparatively low level of activity of 2p2613 is unclear. To test whether the signal is produced mainly from translational shifting, the window is reframed in construct 2p3901. The same sequence from the 2p2613 window is contained in 2p3901 (plus extra nucleotides necessary for reframing) but the ribosome translates into the window in another frame, a frame in which G-GGG Gly codons are not encountered; amino-terminal sequence analysis indicates that the G-GGG sequence in 2p2613 is the shift site (data not shown). If the activity of 2p2613 results from the addition of a single G residue into the G-GGG sequence, then 2p2613 and 2p3901 should have equal levels of activity; however, if it results from a translational slip on the G-GGG sequence, then the activity of 2p3901 should be lower than 2p2613. The decreased activity of 2p3901 (1 unit, Table 1) confirms the translational component of the shift emanating from 2p2613 (8 units).

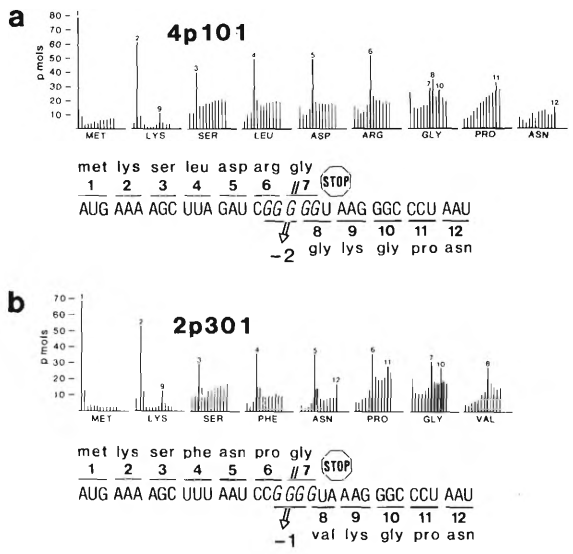

Figure 1. Amino-terminal sequence analysis of shifting on strings of repetitive nucleotides. Each panel shows the yield in pmols of PTH amino acids analyzed during each sequencing cycle; the data are not corrected for injection or base line. A number over a vertical line represents the major PTH amino acid recovered from that cycle. Underneath the panels are interpretations of the shifts seen by superimposing the protein sequence and the mRNA sequence.

If a GGG or GGA-decoding glycine tRNA can shift back −2 onto an overlapping GGG codon, and a GAA-decoding glutamate tRNA can shift back −2 onto an overlapping GGG, it might be expected that a GCG-decoding alanine tRNA would shift back −2 onto an overlapping GCG. 4p801 CGC-GCG-UGA (Table 1) demonstrates that this event is two orders of magnitude less frequent than the shift in 4p1103 CGG-GGG-UGA. One explanation is that the ability to backshift is tRNA-specific, whereas another view might require intermediate pairing, that is, a tRNA must step −1 with correct pairing before stepping −2, or an equivalent view would hold that purines may not pass purines on the opposite strand.

Inspection of sequences naturally occurring 5′ to ter-

mination codons of *E. coli* genes (from NIH GenBank and EMBL data libraries, January 1987) reveals that shifty stops of the type shown in Table 1 are rather prevalent, raising the prospect that in some cases they may be there for a purpose. Some potential uses for shifty stops might be required functional heterogeneity at the carboxyl terminus and translational coupling to downstream mRNA signals, for instance, translation start signals or mRNA degradation signals; such hypothetical uses of shifty stops remain to be tested. Why 3′ stop codon borders should enhance the rate of shifting on strings relative to sense codon borders is not known, but it may be a consequence of release factor action at stop codons, or the enhancement may occur prior to release factor binding.

tRNA Hopping

Can tRNAs move from their zero-frame codon to nearby similar codons, and do such tRNA "hops" occur? Construct 4p801 CGC-GCG-UGA suggests that tRNAs are not able to hop backward from their zero frame codon to a nearby similar codon at a high rate. However, in other constructs, several cases of forward hops appear to occur at high rates, in contrast to the attempted backward hop in 4p801. Table 2 lists the β-galactosidase activity, mRNA sequence, and protein sequence derived from three synthetic constructs in *lacZ* that are capable of testing sequences for tRNA hopping. One interpretation of the evidence is that in construct 2p4001 (GUG-UG) the valine tRNA decoding at GUG codon 7 hops onto the GUG overlapping codons 7 and 8, in construct 2p4101 (AAC-UCA-AU) the asparagine tRNA decoding at AAC codon 4 hops onto the AAU codon overlapping codons 5 and 6, and in construct 3p4201 (CUU-UAG-CUA) the leucine tRNA decoding the CUU at codon 7 hops over codon 8 and onto the CUA at codon 9, leading to shifts in the reading frame by +2, +5, and +6 nucleotides. The β-galactosidase levels produced from these constructs (ranging from 0.4 to 1.0% of average in-frame levels) implies that hopping occurs quite frequently. Another interpretation of the pattern of activities and protein sequences displayed by these constructs is that during expression, the mRNA is altered in some manner as to be translated into the observed protein sequence; reversion of the frameshift lesion cannot explain these hops, since genetic variation is not observed at the required level. The translational basis of the observed +2 and +5 hops is testable by reframing the windows. Construct 2p4001 is also a synthetic analog of the *trpE91* frameshift window (Atkins et al. 1983) first isolated in *Salmonella typhimurium* and for which many external −1 frameshift suppressors exist. Recently, one of these suppressors, *E. coli hopR*, has been shown to suppress via hopping at the GUG-UG (B. Falahee and J. Atkins, unpubl.).

The precise sequence requirements and generality of tRNA hopping remain uncharacterized; however, potential examples for utilizing a programmed hop during gene expression may be found among the plant RNA viruses. Tobacco mosaic virus (TMV) encodes a 183-kD read-through polypeptide synthesized by leaky termination at the TMV 126-kD polypeptide amber stop (Pelham 1978; Goelet et al. 1982), beet necrotic yellow vein virus RNA (Bouzoubaa et al. 1986), and turnip yellow mosaic virus RNA (Haenni et al. 1987) encode similar amber read-through polypeptides. In these three cases, the leaky amber terminator is flanked by CAA glutamine codons (CAA-UAG-CAA). It is possible these amber stops leak because of stop-hopping, similar to the leucine tRNA stop-hop seen in construct 4p4201 CUU-UAG-CUA.

Programmed Ribosomal Frameshifts

Two cases of programmed high-level frameshifting, *E. coli* peptide chain release factor 2 (RF2) and Rous sarcoma virus *gag-pol* polyprotein and related retroviral shifts, display similar motifs of shifting that appear related to the synthetic *lacZ* high-level shifts; however, the rates of 30% shifting and above indicate that they are doing something extra. The known retroviral shifts are −1 on pyrimidine or purine strings and their levels range from 3 to 30%; this high level requires mRNA sequences 3′ to the string, which may form stem-loop

Table 2. tRNA Hopping

Series		β-Galactosidase activity (whole-cell assays, % in-frame level)
	[+2 hop]	
A.	Met-Lys-Ser-Phe-Asp-Gly-Val . . . Arg-Leu	
2p4001-	AUG-AAA-AGC-UUU-GAU-GGA-GUG-UGA-GGU-UAA-	0.5
	[+5 hop]	
B.	Met-Glu-Ile-Asn- . . . Leu-Glu-Gly	0.4
2p4101-	AUG-GAG-AUU-AAC-UCA-AUC-UAG-AGG-GUA-	
	[+6 hop]	
C.	Met-Lys-Ser-Leu-Gly-Tyr-Leu- . . . Arg-Arg	1.0
3p4201-	AUG-AAA-AGC-UUA-GGG-UAU-CUU-UAG-CUA-CGA-CGG-	

The first AUG in each construct is the translation start codon, and the amino-terminal sequence determined for each construct is superimposed above the mRNA sequence. The β-galactosidase activity is given as a percentage of in-frame levels (in-frame level = average of 3p900 series, Table 1). The similar codons inferred to be involved in the hops are underlined.

structures (Jacks et al. 1987; Moore et al. 1987). The reading frame in RF2 mRNA shifts +1 on a CUU-U pyrimidine string (Craigen et al. 1985), and the rate of shifting is approximately 30% (Craigen and Caskey 1986). A mutational analysis of synthetic RF2 frameshift window (see below) suggests that both upstream and downstream elements flanking the CUU-U string enhance the rate of shifting. One possible recipe for programming high-level frameshifts derived from these examples is to begin with a codon:anticodon pair embedded within a string of repetitive mRNA nucleotides, and then perturb this inherently ambiguous pairing potential with outlying mRNA:ribosome interplay of various kinds.

The RF2 shift has been analyzed by synthetically reconstructing a 30-nucleotide mRNA sequence sufficient for high-level shifting near the 5′ end of the *lacZ* gene; Table 3 displays some of the variants of this sequence that define the components of the site. The 25% ratio of β-galactosidase produced from out-of-frame versus in-frame synthetic RF2-*lacZ* constructs (4p2101 versus 3p1201, Table 3) demonstrates that the high rate of shifting has been recreated within *lacZ*. The protein sequences (Fig. 2) of two synthetic derivatives, 4p2101, an analog of the RF2 window, and 4p2203, which changes the UGA stop to a UGG tryptophan codon, both suggest that a leucine tRNA decoding at the CUU-UGA/G sequence slips +1 at the string of pyrimidines.

The requirement for a string at the shift site is seen

Table 3. Variants of the Synthetic RF2-*lacZ* Frameshift Window

Series	β-Galactosidase activity (whole-cell units)
Val-Leu-Arg-Gly-Tyr-Leu-Asp-Tyr-Glu	
3p1201 -GUU-CUU-AGG-GGG-UAU-CUU-GAC-UAC-GAG-	24600
Val-Leu-Arg-Gly-Tyr-Leu -Asp-Tyr-Glu	
4p2101 -GUU-CUU-AGG-GGG-UAU-CUU-UGA-CUA-CGA-G	6600
4p1601 -GUU-CUU-AGG-AGG-UAU-CUU-UGA-CUA-CGA-G	1400
1602- U-G	660
1603- G-U	280
4p1701-GUU-CUU-AGG-GGG-CAU-CUU-UGA-CUA-CGA-G	2500
1706- GC-G	260
1716- CG-A	50
4p1801 G-UUC-UUA-GGG-GGU-UAU-CUU-UGA-CUA-CGA-G	390
1802- G	35
4p1901-GUU-CUU-AGG-GGG-UUA-CUU-UGA-CUA-CGA-G	9300
1905- CG	5400
Val-Leu-Arg-Gly-Tyr-Val -Asp-Tyr-Glu	
4p2001-GUU-CUU-AGG-GGG-UAU-GUU-UGA-CUA-CGA-G	1900
2003- UUA	60
2004- CUA	35
2005- GUA	25
2006- GUG	25
2007- AUA	15
4p2100-GUU-CUU-AGG-GGG-UAU-CUU-UGA-CUA-CGA-G	6600
2201- UAG	3800
2102- UAA	3600
2203- UGG	830
2103- UUA	560
2204- UUG	530
4p2301-GUU-CUU-AGA-GGG-UAU-CUU-UUG-CUA-CGA-G	45
Val-Leu-Arg-Gly-Tyr-Gly -Asp-Tyr-Glu	
4p2501-GUU-CUU-AGG-GGG-UAU-GGG-UGA-CUA-CGA-G	1500
2502- U	260
4p1302-UGU-CUU-AGG-GGG-UAU-CUU-UGA-CUA-CGA-G	5400
4p1403-GUG-GUU-AGG-GGG-UAU-CUU-UGA-CUA-CGA-G	4100
4p1501-GUU-CUA-AGG-GGG-UAU-CUU-UGA-CUA-CGA-G	3700
1504- GG	220
4p4301-UGG-UGG-CUU-AGG-GGG-UAU-CUU-UGA-CUA-C	5400
4302 GG-CGG	4200
4303 GG-GGG	170
4304 GG-AGG	140

The codon demarcating the 5′ border of these series is located 6 codons in from the AUG start codon. The host is *E. coli* SU1675 F′I^{Q}, cultures were grown in LB at 37°C with aeration from a 1:40 dilution of a fresh saturated culture until an OD_{600} of ~0.7; whole cell β-galactosidase assay conditions are described in Table 1. The amino acid sequences shown above certain constructs were determined by amino-terminal sequencing of the purified frameshift β-galactosidase. Critical residues in each series are underlined.

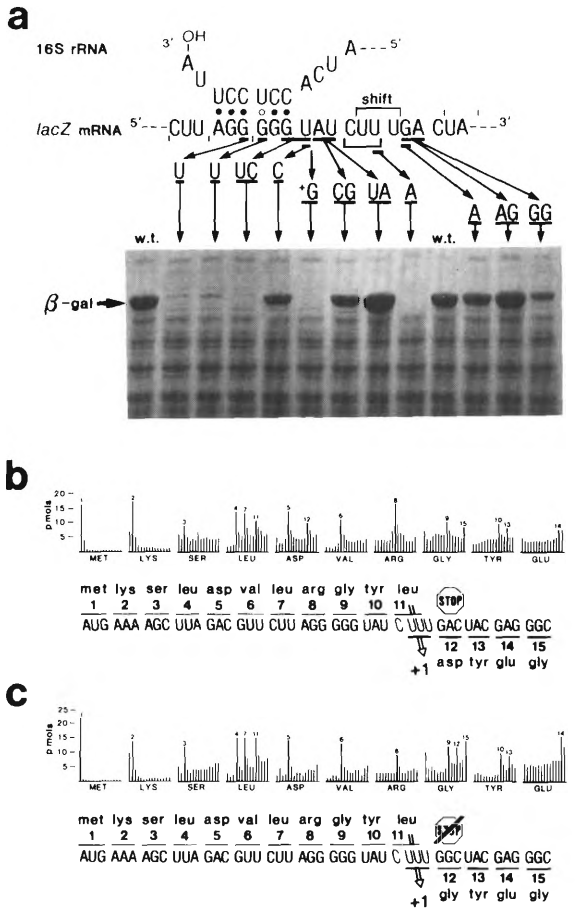

Figure 2. Variants of the synthetic RF2 frameshift window and amino-terminal sequencing of 4p2101 and 4p2203. The Coomassie-stained 7.5% SDS-PAGE gel (*a*) of whole cell lysates shows the amount of β-galactosidase produced from each construct in LB at 37°C; the w.t. designation refers to 4p2101. From left to right, the constructs shown are: 4p2101, 4p1602, 4p1603, 4p1707, 4p1701, 4p1802, 4p1905, 4p1901, 4p2004, 4p2101, 4p2102, 4p2201, and 4p2203 (see Table 3). Amino-terminal sequence analysis and interpretation (*b*) for 4p2101 and (*c*) for 4p2203.

with the constructs that disrupt the string (4p2003 UUA, 4p2004 CUA, 4p2005 GUA, 4p2006 GUG, and 4p2007 AUA; Table 1) and eliminate high-level shifting. It is also possible to substitute one string for another and still maintain high-level shifting (4p2001 GUU-UGA shifts a valine tRNA +1 at this position, and 4p2501 GGG-UGA shifts a glycine tRNA +1, Table 3). A decrease similar to that seen in Table 1 for stop to sense substitutions is seen in the 4p2100 and 4p2200 series (Table 3). This suggests that the CUU-UGA sequence is a shifty stop.

Decreases in the rate of shifting caused by changes upstream of the shifty stop are shown in series 4p1600-1900. Severe decreases are seen when single or multiple changes are introduced 3–7 nucleotides upstream of the CUU-UGA (4p1600 and 4p1700 series, Table 3). Changes within the UAU sequence located between this upstream element and the shifty stop do not grossly affect the level of shifting (4p1901 UUA and 4p1905 UCG, Table 3), but single nucleotide insertions within this region do have severe effects (4p1801 U-UAU and 4p1802 U-GAU, Table 3).

Substitutions further upstream of sequence only slightly alter the level of shifting (4p1300–1500 series, Table 3), except for the change in 4p1504, which decreases shifting by approximately 20-fold (Table 3). One explanation for this decrease is that this substitution creates an overlapping duplication of the region inferred to be critical in the 4p1600 and 1700 series. This region resembles the Shine-Dalgarno (SD) nucleotides of ribosome-binding sites, a region known to function via base pairing with nucleotides near the 3′ end of 16S rRNA (Shine and Dalgarno 1974; Steitz and Jakes 1975). In 4p1504, an SD-like sequence (GGAGG) overlaps the SD-like sequence inferred to be necessary for high-level shifting, implying perhaps that if the 16S rRNA pairs with this upstream GGAGG sequence it is then unavailable for pairing with the AGGGGG sequence necessary for shifting. Interference by upstream core SD-like sequences is tested explicitly in series 4p4300 (Table 3). When the sequences GGGGG or GGAGG are placed three nucleotides upstream of the AGGGGG, interference is evident (4p4303 GGGGG-170 units and 4p4304 GGAGG-140 units, Table 3), but when GGUGG or GGCGG is substituted here, only slight effects are observed (4p4301 GGUGG-5400 units and 4p4302 GGCGG-4200 units, Table 3). This pattern of SD-like effects on shifting implies that pairing between the mRNA and 16S rRNA may occur close to the decoding sites within elongating ribosomes and this pairing can enhance the rate of +1 shifting on properly spaced shifty stops. A comparison of 4p2204 UUG-530 units, a stop to a sense codon change, with 4p2301 AGAG-GG...UUG-45 units, a change of both the SD-like element and the stop codon, demonstrate that the enhanced level of shifting caused by the SD-like element is not confined to shifty stops but works on sense strings as well (see amino-terminal sequence of 4p2202, Fig. 2). Critical changes in the synthetic RF2 window are summarized in Figure 2, which displays the β-galactosidase levels produced by these mutants as seen by Coomassie-blue staining of an SDS-acrylamide gel of whole-cell lysates. The postulated pairing between the mRNA and 16S rRNA is also shown in Figure 2.

CONCLUSIONS

Application of technical improvements in the array of experimental approaches currently available has provided a fresh look at the basis of frameshift leakiness. The novel events observed include the stop codon effect upon shifting on strings, tRNAs that appear to hop forward on the mRNA from their zero-frame codon to a similar codon, and a postulated SD-like interaction within elongating ribosomes. Some of these events have evidently become grist for evolution's mill, having been turned into sophisticated control points for particular genes. The rationale for nature having con-

structed the RF2 shift site and bordering it with a UGA codon is to provide an efficient autoregulatory loop for RF2 expression (Craigen and Caskey 1986), since RF2 is involved in termination at UGA and UAA codons. In doing so, both the SD-like and stop codon effect on enhancing shifts on pyrimidine or purine strings have been used to advantage. Retroviruses have also taken advantage of enhanced shifting on strings by using 3′ sequence elements to elevate the level of shifting (Jacks et al. 1987, Moore et al. 1987). The rationale there appears designed for producing nested proteins in a defined ratio from a single translation start. How many more examples of these types of controls exist remains to be seen, but they contain the prospect of illuminating the design of the ribosome's translocation mechanism.

ACKNOWLEDGMENTS

Regina Zeikus and Bernadette Falahee provided excellent assistance in protein sequencing. We thank Al Dahlberg, Larry Gold, Tyler Jacks, Shalha Thompson, R.J. Goldberg, Bill Craigen, and Jon Gallant for helpful conversations. This research was supported by the Howard Hughes Medical Institute and National Science Foundation grant DMB-8408649. R. Weiss is a fellow of the Helen Hay Whitney Foundation.

REFERENCES

Aebersold, R.H., D.B. Teplow, L.E. Hood, and S.B.H. Kent. 1986. Electroblotting onto activated glass: High efficiency preparation of proteins from analytical SDS-PAGE for direct sequence analysis. *J. Biol. Chem.* **261:** 4229.

Atkins, J.F., D. Elseviers, and L. Gorini. 1972. Low activity of β-galactosidase in frameshift mutants of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **69:** 1192.

Atkins, J.F., B.P. Nichols, and S. Thompson. 1983. The nucleotide sequence of the first externally suppressible −1 frameshift mutant, and of some nearby leaky frameshift mutants. *EMBO J.* **2:** 1345.

Bouzoubaa, S., V. Ziegler, D. Beck, H. Guilley, K. Richards, and G. Jonard. 1986. Nucleotide sequence of beet necrotic yellow vein virus RNA-2. *J. Gen. Virol.* **67:** 1689.

Chen, E.Y. and P.H. Seeburg. 1985. Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. *DNA* **4:** 165.

Clare, J. and P. Farabaugh. 1985. Nucleotide sequence of a yeast Ty element: Evidence for an unusual mechanism of gene expression. *Proc. Natl. Acad. Sci.* **82:** 2829.

Craigen, W.J. and C.T. Caskey. 1986. Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature* **322:** 273.

Craigen, W.J., R.G. Cook, W.P. Tate, and C.T. Caskey. 1985. Bacterial peptide chain release factors: Conserved primary structure and possible frameshift regulation of release factor 2. *Proc. Natl. Acad. Sci.* **82:** 3616.

Davis, R.W., D. Botstein, and J.R. Roth. 1980. *Advanced bacterial genetics: A manual for genetic engineering*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Dunn, J.J. and F.W. Studier. 1983. Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J. Mol. Biol.* **166:** 477.

Fox, T.D. and B. Weiss-Brummer. 1980. Leaky +1 and −1 frameshift mutations at the same site in a yeast mitochondrial gene. *Nature* **288:** 60.

Goelet, P., G.P. Lomonosoff, P.J.G. Butler, M.E. Akam, M.J. Gait, and J. Karn. 1982. Nucleotide sequence of tobacco mosaic virus. *Proc. Natl. Acad. Sci.* **79:** 5818.

Haenii, A.L., M.D. Morch, G. Drugeon, R. Valle, R. Joshi, and T.M. Denial. 1987. Gene expression in turnip yellow mosaic virus. *UCLA Symp. Mol. Cell. Biol.* **54:** 149.

Jacks, T. and H.E. Varmus. 1985. Expression of Rous sarcoma virus *pol* gene by ribosomal frameshifting. *Science* **230:** 1237.

Jacks, T., K. Townsley, H.E. Varmus, and J. Majors. 1987. Two efficient ribosomal frameshifting events are required for synthesis of mouse mammary tumor virus *gag*-related polyproteins. *Proc. Natl. Acad. Sci.* **84:** 4298.

Mellor, J., S.M. Fulton, M.J. Dobson, W. Wilson, S.M. Kingsman, and A.J. Kingsman. 1985. A retrovirus-like strategy for expression of a fusion protein encoded by yeast transposon Ty1. *Nature* **313:** 243.

Miller, J. 1972. *Experiments in molecular genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Miller, J.H. and A.M. Albertini. 1983. Effects of surrounding sequence on the suppression of nonsense codons. *J. Mol. Biol.* **164:** 59.

Moore, R., M. Dixon, R. Smith, G. Peters, and C. Dickson. 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(2):** 480.

Pelham, H.R.B. 1978. Leaky UAG termination codon in tobacco mosaic virus RNA. *Nature* **272:** 469.

Shimotohno, K., Y. Takahashi, N. Shimizu, T. Gojobori, D.W. Golde, I.S.Y. Chen, M. Miwa, and T. Sugimura. 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.* **82:** 3101.

Shine, J. and L. Dalgarno. 1974. The 3′ terminal sequence of *Escherichia coli* 16S ribosomal rRNA: Complementarity to nonsense triplets and ribosome binding sites. *Proc. Natl. Acad. Sci.* **71:** 1342.

Steitz, J.A. and K. Jakes. 1975. How ribosomes select initiator regions in mRNA: Base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc. Natl. Acad. Sci.* **72:** 4734.

Weiss, R.B. and J.A. Gallant. 1983. Mechanism of ribosome frameshifting during translation of the genetic code. *Nature* **302:** 389.