Ribosomes can slide over and beyond “hungry” codons, resuming protein chain elongation many nucleotides downstream

Jonathan A. Gallant* and Dale Lindsley

Department of Genetics, University of Washington, Box 357360, Seattle, WA 98195-7360

Communicated by Charles Yanofsky, Stanford University, Stanford, CA, September 9, 1998 (received for review July 1, 1998)

ABSTRACT In cells subjected to moderate aminoaacyl-tRNA limitation, the peptidyl-tRNA–ribosome complex stalled at the “hungry” codon can slide well beyond it on the messenger RNA and resume translation further downstream. This behavior is proved by unequivocal amino acid sequence data, showing a protein that lacks the bypassed sequence encoded between the hungry codon and specific landing sites. The landing sites are codons cognate to the anticodon of the peptidyl-tRNA. The efficiency of this behavior can be as high as 10–20% but declines with the length of the slide. Interposition of “trap” sites (nonproductive landing sites) in the bypassed region reduces the frequency of successful slides, confirming that the ribosome–peptidyl-tRNA complex passes through the untranslated region of the message. This behavior appears to be quite general: it can occur at the two kinds of hungry codons tested, AUA and AAG; the sliding peptidyl-tRNA can be any of three species tested, phenylalanine, tyrosine, or leucine tRNA; the peptidyl component can be either of two very different peptide sequences; and translation can resume at any of the three codons tested.

Protein synthesis is conventionally pictured as the orderly, sequential addition of amino acids to a growing peptide chain dictated by the sequence of triplets in the messenger RNA. However, alternatives to this linear progression are possible and can occur at surprisingly high frequencies in certain sequence contexts and/or under certain physiological conditions. Limitation for an aminoaacyl-tRNA, in particular, raises the probabilities of alternative paths at “hungry” codons cognate to the limiting species. The alternative paths described so far include binding of noncognate aminoaacyl-tRNA species (reviewed in refs. 1–3) and frameshifting either to the left (−1) or the right (+1) (reviewed in refs. 4, 5). In addition, peptidyl-tRNA release, normally a low-frequency event, may occur at hungry codons (6–8).

In this communication, we describe another alternative path the ribosome can follow at a hungry codon: sliding over it and moving several codons downstream, resuming translation further along the message. We encountered this phenomenon in the course of studying ribosome behavior at a sequence conducive to leftward frameshifting.

MATERIALS AND METHODS

Bacteria were cultured at 37°C under forced aeration in M63-glucose medium (9). To retain comparability with ref. 9, the medium was supplemented with leucine, threonine, arginine, and histidine, even though the strain used in the present experiments is prototrophic. The host was strain C92 (reA+, spoT−) carrying the O6 deletion of the early part of lacZ (10). The deletion, provided to us as an F+ plasmid by Jon Beckwith (Harvard Medical School, Boston), was introduced into the C92 chromosome by homogenotization. For unknown reasons, identical plasmids yield twice as much β-galactosidase in this strain as in the CP78/79 lineage we have used previously (9). This proved useful in purifying material for protein sequencing.

All methods of DNA construction and analysis were as described (9). Engineered plasmids (Fig. 1) were transformed into the host strain with selection for the plasmid’s bla gene (1 mg/ml carbenicillin). To measure enzyme synthesis, cultures were grown overnight from single colonies with carbenicillin, diluted, and grown into logarithmic phase. When an OD700 (Beckman spectrophotometer) of 0.2 was reached, the cultures were induced [2 mM isopropyl β-D-thiogalactopyranoside (IPTG), 2.5 mM cAMP], and aliquots were grown in parallel in the absence and presence of 72 μg/ml isoleucine hydroxamate (ILHX) to inhibit isoleucine tRNA charging. Immediately before induction and after about one doubling, samples were withdrawn for measurement of β-galactosidase and protein, as described (9). The plasmid vector containing the engineered forms of the lacZ gene also carries a lacP′ regulator gene.

RESULTS

Fig. 2A shows the coding-strand sequence of the relevant region of a lacZ reporter-gene construct we call L-16 (TTT), in which the sequence at positions 9–15 is very prone to leftward frameshifting during limitation for isoleucyl tRNA (9). However, leftward frameshifting at this position, or anywhere else, cannot give rise to β-galactosidase in this construct because the β-galactosidase protein is encoded in the other alternative, or rightward (+1), reading frame, the translation product of which is shown. Moreover, this reading frame cannot be accessed by a rightward shift at any isoleucine codon because it is blocked by two terminator triplets located between the isoleucine codons of the 0 reading frame and position 26, as shown. Nonetheless, partial isoleucine-tRNA starvation elicited a more than 40-fold increase in β-galactosidase synthesis with this construct (Table 1), an increase comparable to that observed with LefShf, a reporter of simple frameshifting at the same sequence (Table 1; ref. 9).

How do ribosomes starved for isoleucyl tRNA negotiate this messenger sequence so as to synthesize β-galactosidase? To find out, we purified the β-galactosidase and determined its amino acid sequence in the region of interest (Fig. 2B). The sequence following the enterokinase cut site is as expected (Val-Ser-Tyr-Phe) up to the position of the hungry AUA isoleucine codon at positions 13–15; it then continues (Glu-Asp-Asn-Leu-Glu-Ile-Pro-Gly- and four further amino acids (Fig. 2A and B). Evidently, ribosomes stalled at positions 13–15 simply hopped over 16 nucleotides, landing so far downstream that the sequence of triplets in the messenger RNA and resume translation furth
the UUU codon was changed to UUC, which is recognized by the same phenylalanine tRNA, ILHIX-induced hopping was equally efficient. When the UUU was changed to GUU, complementary to the phenylalanine tRNA anticodon at only the second and third nucleotides, it was reduced by a factor of 6. Protein sequencing of the β-galactosidase (data not shown) confirmed that it still reflected a hop to the triplet following the GUU codon. When the UUU was changed to GAA or GCG, different at two or three positions, hopping was drastically reduced, by a factor of 20. Finally, when we changed the takeoff site to TAC and likewise changed the landing site to TAC, in construct TAC-16 (TAC), we observed hopping virtually as efficient as the cases in which both sites encode phenylalanyl-tRNA. Thus, hopping induced by a hungry codon depends on the ability of the P-site codons of the takeoff and landing sites to bind the same tRNA or peptidyl-tRNA.

In another set of experiments, we varied the distance between the takeoff and landing sites’ phenylalanine codons. The results (Table 1) show that the frequency of successful hops decreases with distance. When the takeoff and landing sites were separated by 7, 15, or 16 nucleotides, enzyme synthesis was about 20% that of the wild-type control; when separated by 22 nucleotides (bypassing two slightly different sequences), it was 10% of the wild-type control; and when separated by 39 or 40 nucleotides, it was only about 1% of the wild-type control. This effect of the spacing of the takeoff and landing site codons is also consistent with the hypothesis that these codons determine the successful resumption of translation. The results also suggest that this is insensitive to reading frame. Hops that landed in the original reading frame (15 and 39) were not notably more efficient than those of similar distance (16 and 40, respectively) that landed in the rightward reading frame. The similar results with the two reporters of 22-nucleotide hops suggest, in fact, that the frequency of success is not much affected by the sequence the ribosome bypasses without translating.

On the other hand, the ribosome–peptidyl-tRNA complex presumably does interact with the region it bypasses so as to detect potential landing sites—that is, codons (evidently in any frame) cognate to the peptidyl-tRNA. If this view is correct, then the number of ribosomes arriving at a productive landing site should be reduced by the interposition of nonproductive landing sites between the takeoff site and the productive landing site. Tests of this proposition are shown in Table 1.

The construct designated E-Trap-16 is identical to L-16 (GT) except for a 1-base change that introduces a UUU landing trap (followed two codons downstream by a terminator) seven nucleotides beyond the takeoff site. This change reduced β-galactosidase synthesis 54% relative to that of L-16. L-Trap-16 differs from L-16 at just two nucleotides, such that a UUU trap (in the wrong reading frame) is introduced at a later position, 11 nucleotides beyond the takeoff site. This change reduced enzyme synthesis by 58% relative to L-16. In E-Trap-22 and L-Trap-22, we interposed UUU traps (in the wrong reading frame) at early and late positions between the takeoff and landing sites of L-22. These changes reduced β-galactosidase synthesis by 71% and 67%, respectively.

Thus, UUU traps in the bypassed region can eliminate about two-thirds of the ribosomes that would otherwise arrive at a productive landing site and yield β-galactosidase. We conclude that the ribosome–peptidyl-tRNA complex does not in fact hop over the bypassed region, but rather slides through it, scanning for triplets complementary to the anticodon of the peptidyl-tRNA as it proceeds.

The β-galactosidase produced by constructs that demanded slides of 22 and 39 nucleotides respectively was purified and sequenced as in Fig. 2B. In both cases (Fig. 3) the amino acid sequence confirmed that β-galactosidase was made by the expected movements to the UUU landing site, which is followed by the same downstream sequence in each case.
In all of the preceding cases, the takeoff site was the 28th codon of a specific and rather peculiar message. Our reporter oligonucleotides all are inserted into a vector that carries an amino acid sequence of the protein made in isoleucine-tRNA limited products in all three reading frames are shown, aligned under their corresponding triplets, with termination at nonsense triplets indicated by an *.

Predicted translation products in all three reading frames are shown, aligned under their corresponding triplets, with termination at nonsense triplets indicated by an *.

The hungry isoleucine codon is italicized. The shifty sequence T TTC ATA is at positions 9–15. Amino acids in boldface are those found in the protein made in isoleucine-tRNA limited products in all three reading frames are shown, aligned under their corresponding triplets, with termination at nonsense triplets indicated by an *.

The enterokinase cut site, shown by arrows. The 0-frame sequence on the N-terminal side of the cut site is indicated schematically. Predicted translation products in all three reading frames are shown, aligned under their corresponding triplets, with termination at nonsense triplets indicated by an *.

The observed effect is due to an increased rate of expression of the cognate isoleucine-tRNA.

**Table 1.** Differential rates of codon-jumping induced by ILHX

<table>
<thead>
<tr>
<th>Construct name (Landing site sequence)</th>
<th>Translation behavior</th>
<th>Differential rates EU/mg protein × 10³</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Control</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>352 ± 16 (3)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>216 ± 2 (2)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → GAT</td>
<td>217 ± 45 (3)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TAC</td>
<td>147 ± 1.5 (2)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TAC</td>
<td>402 ± 11 (4)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>567 ± 63 (4)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>495 ± 34 (2)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>322 ± 26 (5)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>308 ± 39 (3)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>200 ± 17 (4)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>56 ± 5.8 (2)</td>
</tr>
<tr>
<td>L-16 (TTT)</td>
<td>TTC → 16 → TTT</td>
<td>60 ± 2.5 (2)</td>
</tr>
</tbody>
</table>

Under "Translation behavior" we indicate the takeoff-site codon for peptidyl-tRNA, followed by an arrow symbolizing the hop, its length, and the landing P-site. The data are reported as average differential rates of enzyme synthesis, in Enzyme Units per mg protein × 10⁴, ± SEM, with the number of replicate experiments in parentheses. In the case of construct L-16 (TTT), simple isoleucine starvation by means of valine inhibition (11) elicited an increase in the differential rate from 444 to 12,900, closely similar to the effect of ILHX.

*Zero-frame control.
†Simple left shifter.
‡Slide reporter.
specific to this unusual sequence, we inserted the oligonucleotide of L-16 (TTT) into another vector with a completely different upstream sequence, to construct pXA-L16. Fig. 4 shows the two L-16 constructs with different upstream sequences and compares their activities. In pXA-L16, enzyme synthesis was 1/3–1/4 that of L-16, but ILHX stimulated enzyme synthesis by virtually the same factor, 45-fold. Thus, the induction of sliding by a hungry AUA codon is largely independent of upstream sequence, and hence of the nature of the peptide part of the sliding peptidyl-tRNA.

In all of the cases described so far, we have induced sliding at a hungry AUA codon by limiting the acylation of isoleucine tRNA. Is there anything special about the hungry AUA codon, or can sliding be initiated at a different hungry codon? The experiments shown in Fig. 5 address this question. Here, we have altered the sequence of L-16 to replace AUA by AAG, coding for lysine tRNA. In addition, we changed the P-site codon to UUA to test still another peptidyl-tRNA for sliding behavior, providing a UUA landing site 16 nucleotides downstream. Finally, we changed the triplet following the landing site to test the influence of that sequence element. Fig. 5A shows the sequence of the coding strand in this construct, which we call L-9-16-L9 and indicates its β-galactosidase activity. Lysyl tRNA limitation by lysine hydroxamate promoted a 30-fold increase in β-galactosidase synthesis encoded by this construct. Fig. 5B shows the amino acid sequence of the enzyme made under this condition, confirming the slide from the takeoff sequence UUAAAG to the triplet following the landing site UUA. Note that the first amino acid following the slide in this construct is histidine (CAC), in contrast to the glutamic acid (GAA) in the previous set. We have also confirmed sliding by amino acid sequence analysis in one other case (data not shown), where the first amino acid after the slide is lysine (AAA). Thus, to a first approximation, sliding is independent of the amino acid with which peptide-chain elongation resumes.

We conclude that ribosome sliding can be promoted when the ribosome is stalled at either AUA or AAG, and perhaps at any hungry codon; that the sliding ribosome can carry peptidyl-tRNA of the phenylalanyl, tyrosyl, or leucyl varieties, and very likely any peptidyl-tRNA, to a cognate landing site downstream; and that translation then resumes with any aminacyl-tRNA encoded in the triplet following the landing site.

**DISCUSSION**

The constructs under discussion in this paper—including reporters both of frameshifting and of sliding—support about 63.10^24 as much enzyme synthesis as the lacZ control in growing, uninhibited cells (Table 1). The reporters and the control share the same sequence in the promoter, the ribosome loading region, the first 78 nucleotides of the lacZ coding sequence, and, indeed, nearly all of the lacZ coding sequence. As a consequence, transcription and translation should occur identically in reporters and their controls up to the 0 reading frame terminator triplets of the former (placed by design just downstream of the hungry codon) and beyond that if frame-
shifting or sliding occurs. Therefore, the $6 \times 10^{-4}$ ratio of enzyme synthesis almost certainly means that about 1 ribosome in 1,600 performs the acrobatics necessary to generate active $\beta$-galactosidase. Similarly, normalizing the rate of enzyme synthesis to that in the 0 reading frame control under conditions of isoleucyl-tRNA limitation should yield the frequency of ribosomes that performs the feat under this condition.

Our isoleucyl-tRNA limitation regime (72 $\mu$g/ml ILHX) reduces growth rate $\sim$80% imposed on relA mutant cells in which errors of translation of all kinds at hungry codons are exacerbated (1–3). Under these conditions, the average differential rate of enzyme synthesis in the frameshift or shorter slide-reporters increases about 40-fold (Table 1). Moreover, this finding is to be contrasted with an 85% decrease observed in the lacZ control because of effects on both transcription of the lacZ gene (14–16) and translation of the other isoleucine codons in the message (2, 3, 17, 18). Normalizing to the lacZ control indicates that the frequency of sliding in our L-7, L-15, and L-16 reporters increases from about $6 \times 10^{-4}$ to between 0.18 and 0.23.

These values are one to two orders of magnitude higher than the frequency of even shorter hops by normal tRNAs reported by Weiss et al. (12) and O’Connor et al. (13), and even somewhat higher than the frequencies observed in certain anticondon loop tRNA mutants that displayed increased hopping (13). Most, but not all of the hops described by these workers were over a nonsense triplet, which is partly consistent with the view that a (presumed) translational pause promotes the hop. Kane et al. (19) reported suggestive evidence for a high-frequency (20%) hop of +6 nucleotides in a mammalian gene overexpressed in Escherichia coli from a takeoff point preceding a very rare AGG codon, so that this event too could be rationalized in terms of ribosome pausing. Our results provide more compelling evidence that ribosome pausing underlies these unusual ribosome movements: limitation for aminocyl-tRNA converts an ordinary sense codon into a rare AGG codon in the message (2, 3, 17, 18). Normalizing to the lacZ control because of effects on both transcription of the lacZ gene (14–16) and translation of the other isoleucine codons in the message (2, 3, 17, 18). Normalizing to the lacZ control indicates that the frequency of sliding in our L-7, L-15, and L-16 reporters increases from about $6 \times 10^{-4}$ to between 0.18 and 0.23.

FIG. 5. Leaping in construct L'-16-L' induced by stalling the ribosome at a lysine codon with peptidyl-leucyl-tRNA in the P site. (A) Coding strand sequence of the relevant region of L'-16-L' as in Fig. 3A, with the predicted amino acid sequence shown up to the takeoff point and after the predicted landing site. Differential rates of $\beta$-galactosidase synthesis are reported for one doubling of growth in the control and in 50 $\mu$g/ml lysine hydroxamate (each in the presence of 100 $\mu$g/ml every amino acid except lysine as in ref. 9), which inhibits growth slightly less than 72 $\mu$g/ml ILHX. (B) Cycle bar plot of Edman N-terminal sequence analysis, as in Fig. 2B. Two liters of cells were induced and subjected to lysine hydroxamate (50 $\mu$g/ml) inhibition. After about one doubling, the cells were harvested, and $\beta$-galactosidase was purified and sequenced as in Figs. 2B and 3B, at the Stanford Protein and Nucleic Acid Facility.

promoted sliding, but about 1/7 as effectively, although the reduction in growth rate was comparable (data not shown). The mechanism by which the relA ppGpp system influences ribosome behavior and accuracy at hungry codons is unclear (2, 3).

There may be sequence contexts around or between the takeoff and landing sites that affect the frequency of sliding, although these preliminary results suggest that these effects are not great. Three different triplets immediately after the landing P site, and three different pairs of takeoff/landing P sites all permitted sliding. No striking secondary-structure sequences are evident between the takeoff and landing sites. Most of the efficient sliders happen to have a UGAA U-turn motif overlapping the landing site, but three (those with TTC, TAC, and TTA takeoff and landing P sites) do not. There is obviously a rich field for much more detailed investigation of sequence constraints affecting this unusual ribosome behavior.

Our results are in some ways analogous to the giant hop of 50 nucleotides discovered by Weiss and colleagues (20, 21) in translation of the gene 60 topoisomerase of phage T4. In this case, the ribosome movement depends on both a terminator codon at the takeoff point and a stem/loop secondary structure just after it (21). Our results demonstrate that starvation-induced pausing at a sense codon can promote analogous ribosome movement very efficiently as well, independent of secondary structure in the untranslated region. The T4 topoisomerase movement also was mysteriously sensitive to the amino acid sequence of the nascent peptide slightly upstream of the takeoff point (21). The two upstream sequences we have tested have virtually nothing in common, as far as we can tell, with each other or with the T4 topoisomerase region in question. Thus, sliding in this study is seemingly not associated with any unique upstream elements or with any obvious consensus-sequence elements.

In each of the cases described above (12, 13, 20, 21) there are matched codons at the take-off and landing sites, and their necessity was demonstrated by mutagenesis in the case of the 50-nucleotide hop in T4 topoisomerase (21). Our results confirm the importance of such matching sites and further demonstrate that they pertain to the movement of peptidyl-tRNA between the P sites at takeoff and landing positions, elicited in the present system by withholding aminocyl-tRNA cognate to the codon in the takeoff position’s aminocyl site.
Our results and those just reviewed all differ in regard to the matching takeoff and landing sites from a controversial series of reports by Engelberg-Kulka and colleagues on another translational anomaly in *E. coli* (22–25). These workers identified an abnormal foreshortened translation product of the trpR gene (22), and later of a trpR-lacZ fusion (23), which they ascribed to a 55-nucleotide hop (23–25). Surprisingly, their analysis shows no matching codons at the putative takeoff and landing sites. Recently, Willis *et al.* (26) carried out a restudy of the phenomenon and were unable to detect any hopping at all; they concluded instead that the aberrant protein is in fact produced in one case by termination and downstream reinitiation, and in another by (+1) frameshifting without hopping.

Our results, like those of Weiss and colleagues (12, 13, 20, 21), emphasize the importance of matching P sites and therefore seem incompatible with the analysis of Engelberg-Kulka and her colleagues. On the other hand, we did detect some low, residual β-galactosidase activity—about 1% of the lacZ + control—in constructs without matching sites (Table 1). We have been unable to determine, from amino acid analysis, whether these low values reflect leaping to nonhomologous sites. The most likely nonhomologous site is the TTA triplet in the β-galactosidase reading frame 28 nucleotides downstream of the takeoff site; this would be complementary to the peptidyl-Phe-tRNA’s anticodon at the first and second positions. Amino acid sequence data will be necessary to analyze these and other putative nonhomologous slides. In any case, the hungry codon phenomenon greatly expands the repertoire of slides calling for analysis.

The sequence changes we have examined affect β-galactosidase synthesis in control (uninhibited) cells and in ILHX-inhibited cells in a correlated fashion (*r* = 0.9 for Table 1 data). This suggests that most of the background enzyme synthesis is caused by the slide events documented with ILHX inhibition. If there is this much sliding in a small region encompassing less than 1% of the lacZ message, then several percent of normal β-galactosidase may include slides somewhere in translation, and an even larger fraction of the ribosomes will have wandered out of reading frame somewhere along the message. These considerations are consistent with earlier studies showing that more than 20% of the ribosomes that initiate translation on the lacZ message fail to stay the course, suggesting a high frequency of processivity errors on a message this long (27–31).

This tendency of ribosomes to wander is greatly stimulated at hungry codons, particularly in relA− cells. High-efficiency translation of heterologous genes in host cells can easily generate hungry codons, with the attendant missense errors, processivity errors, frameshifts (32, 33), and, as the present results show, slides. We think the biotechnology industry would do well to be cognizant of these phenomena and the possible risks they pose.

Dr. Jenny Atkinson carried out some preliminary experiments on a different subject which led to those described in this communication. John Miller performed the experiments on leaping during valine inhibition summarized in the legend of Table 1. Bill Hellberg helped with preparation of the manuscript. This work was supported by Grant GM13626-31 from the National Institutes of Health.