Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Cell, Vol. 62, 117-126, July 13, 1990, Copyright © 1990 by Cell Press

A Nascent Peptide Is Required for Ribosomal Bypass of the Coding Gap in Bacteriophage T4 Gene *60*

Robert B. Weiss,* Wai Mun Huang,† and Diane M. Dunn*
* Howard Hughes Medical Institute
and Department of Human Genetics
† Department of Cellular, Viral and Molecular Biology
The University of Utah
Salt Lake City, Utah 84132

Summary

Bacteriophage T4 DNA topoisomerase gene *60* contains a 50 nucleotide untranslated region within the coding sequence of its mRNA. Translational bypass of this sequence by elongating ribosomes has been postulated for the mode of synthesis of an 18 kd polypeptide specified by the split coding segments. Ribosome bypass of the untranslated region also occurs when a segment of gene *60* is fused to *lacZ* and expressed in E. coli. The efficiency of bypass in these gene *60–lacZ* fusions approaches 100%. Here, mutations that delete, insert, or substitute nucleotides from gene *60–lacZ* fusions are examined. Essential features necessary for high level gap bypass emerging from this analysis are a *cis*-acting nascent peptide sequence, a short duplication bordering the gap, and a stop codon contained in a stem-loop structure at the 5′ junction of the gap.

Introduction

Bacteriophage T4 DNA topoisomerase is a multiple subunit enzyme consisting of polypeptides encoded by genes *39*, *52*, and *60*. The enzyme complex is an ATP-dependent type II DNA topoisomerase that is capable of relaxing superhelical DNA but not supertwisting it (Stetler et al., 1979; Liu et al., 1979). Gene *60* encodes the 18 kd subunit required for tight complex formation of the three subunits and allows reconstitution of the ATP-dependent catalytic topoisomerase activity. T4 amber mutants defective for gene *60* display a cold-sensitive phenotype as well as a delay in the onset of phage DNA replication (Mufti and Bernstein, 1974).

Within the coding region of gene *60*, an internal 50 nucleotide untranslated coding gap persists in the mRNA found in T4-infected Escherichia coli and in mRNA from cloned versions of gene *60* expressed in uninfected cells. These unusual 50 nucleotides separate the first 46 codons of gene *60* from the last 114 codons. The extent of this coding gap within the mRNA was delineated by sequence analysis of the mRNA and the 18 kd protein produced from a cloned version of gene *60* (Huang et al., 1988). Alignment of the polypeptide and mRNA sequences revealed the 50 nucleotide untranslated gap separating codons 46 and 47.

Initial experiments (Huang et al., 1988) showed that gap bypass could be observed when gene *60* was transcribed and translated in vitro and when a 258 nucleotide segment of gene *60* was fused to a *lacZ* reporter gene and expressed in E. coli. The gene *60–lacZ* fusion allowed quantitation of the amount of bypass: this level was gauged to be at least 70%. In none of these conditions could a spliced mRNA product be detected. The absence of spliced mRNA, especially where bypass was known to occur efficiently, supports the postulate that passage over the gap occurs during translation.

Efficient ribosomal bypass of large segments of mRNA sequence within a coding region has little precedent among known translational phenomena. As an initial approach to understanding the atypical structure and expression of gene *60*, a mutational analysis was undertaken to ascertain the mRNA sequence elements required for efficient gap bypass. The results of this analysis are described below and indicate that the gap sequence alone is not sufficient for bypass and that an extensive portion of the 5′ flanking sequence, including the nascent polypeptide chain, is required for high level bypass.

Results

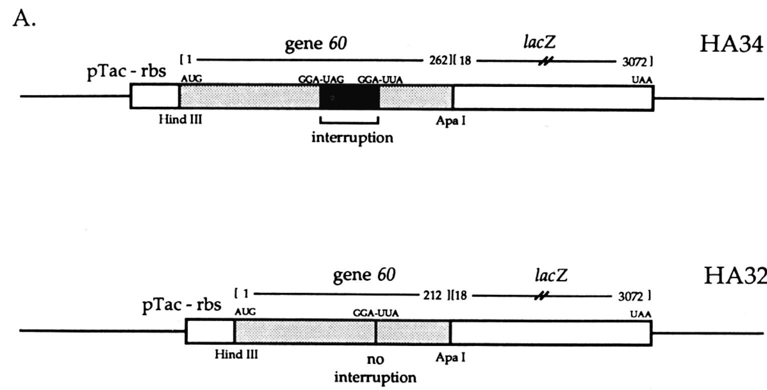
A gene *60–lacZ* reporter system was used to define the mRNA sequence features that allow efficient bypass of the 50 nucleotide internal coding gap. Two initial fusions demonstrated that efficient gap bypass can be observed in E. coli. Gene *60* sequences in these fusions, and all subsequent fusions described, were assembled and cloned from synthetic oligonucleotides.

HA34 contains 261 nucleotides of the gene *60* coding sequence, from codon 1 to codon 70 (the gap spans nucleotides 139–189; Figure 1). HA32 is identical except for a precise deletion of the 50 nucleotide gap. The extent of bypass in HA34 is estimated from the level of β-galactosidase activity. The gene *60–lacZ* fusion, with the gap intact, produces β-galactosidase at a relative level of 94% compared with that of a deletion variant that lacks the gap precisely (HA34, gap^+ = 14,150 units versus HA32, gap^- = 15,050 units; Figure 1). The congruence in physical amount and size of β-galactosidase produced from HA32 and HA34 in vivo is shown in Figure 1B, where polypeptides from whole-cell extracts are separated in an SDS–polyacrylamide gel. Amino-terminal sequence analysis of β-galactosidase produced from HA34 and HA32 (data not shown) indicates that both constructs begin translation at the expected start codon.

Minimum Gene *60* Segment Necessary for Efficient Gap Bypass

The extent of gene *60* sequence required for efficient bypass is examined by deleting segments flanking the gap. A modified gene *60–lacZ* fusion, termed BA3, is the parent construct for the ensuing variants. BA3 contains gene *60* sequence from the ATG start codon to 5 nucleotides past the coding gap (codon 1 to codon 48), and the sequence is fused to the sixth codon of *lacZ*. Three minor alterations

A.



B.

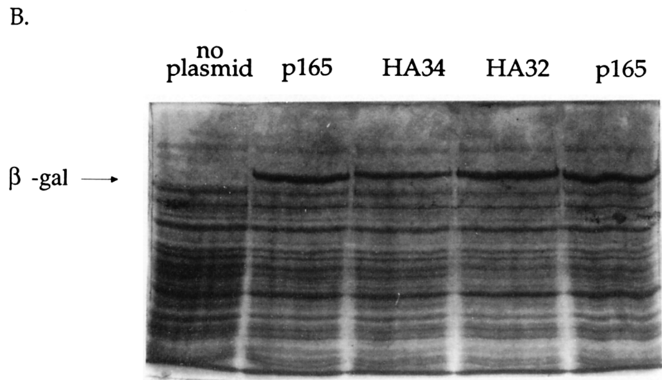


Figure 1. Structure and Expression of Gene *60–lacZ* Fusions in E. coli

(A) Diagram of the gene *60–lacZ* fusions. HA34 and HA32 show the approximate locations of the expression signals and coding regions of these constructs. The gene *60* sequence was cloned from synthetic oligonucleotides into a pBR322-based modified *lacZ* gene described previously (Weiss et al., 1987).

(B) SDS–polyacrylamide gel of total cell protein from strains with and without gene *60–lacZ* fusions. Cultures (0.05 ml) in mid-log phase (OD_{600} = 0.5) were lysed by boiling in 10 mM Tris–HCl (pH 6.9), 0.2% SDS, and 10 mM dithiothreitol, and the proteins were separated on a 7.5% SDS–polyacrylamide gel and stained with Coomassie blue. The position of β-galactosidase produced from the plasmid p165, which encodes only six in-frame residues between the HindIII and ApaI sites on the parent vector of HA34, is indicated with an arrow.

are incorporated to facilitate cloning: a 12 nucleotide 5′ extension (ATG-AAA-AGC-TTA) creating a flanking HindIII site and a T to A transversion at nucleotide 87 plus an A to T transversion at nucleotide 89 creating an internal BglII site. The amount of β-galactosidase expression from BA3, which retains only two amino acids from gene *60* beyond the gap, is 98% compared with that of an exact gap deletion variant. The result indicates that neither the deletion of gene *60* sequence beginning 5 nucleotides 3′ of the gap nor the cloning alterations affect the level of bypass (Figure 2).

The precise location of the 3′ junction of the coding gap is defined by constructs shown in Figure 2. When nucleotides 195 to 197 (AUU) are replaced by UAA, efficient bypass is still observed (BA7-2, 62%; Figure 2 [all values shown are relative amounts of β-galactosidase activity compared with that of the precise gap deletion variant of BA3]). This UAA substitution places a stop codon in the same frame as the downstream *lacZ* coding sequence, indicating that bypass proceeds from nucleotide 151 to at least nucleotide 198. When a stop codon replaces nucleotides 201 to 203, bypass is substantially reduced (BA7-3,

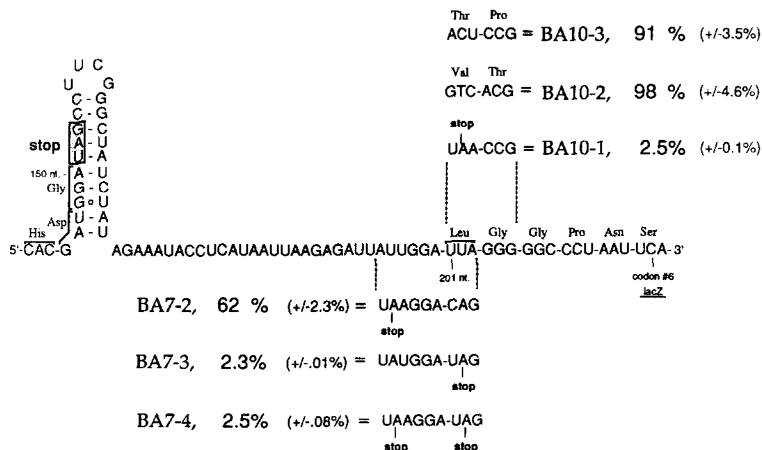


Figure 2. Localization of the 3′ End of the Coding Gap

Shown are the gap nucleotides and local flanking sequence of the gene *60–lacZ* construct BA3 and derivatives. The percentages indicate relative levels of β-galactosidase activity compared with that of the exact gap deletion variant of BA3 (absolute level = 19,000 units) plus the value of one standard deviation unit from assays performed in triplicate. Nucleotides are numbered from the A of the AUG start codon of these constructs.

Figure 3. 5′ Deletions of Gene *60–lacZ* Fusions

Nucleotides 1 to 150 of the modified 5′ segment of gene *60* from construct BA3 and the corresponding peptide sequence are shown. The 50 nucleotide coding gap sequence is unadulterated gene *60* sequence, which continues for 5 nucleotides 3′ of the gap junction: UUA-GGG-GGC-CCU-AAU-UCA (gene *60* sequence is underlined and the UCA is codon 6 of *lacZ*). The relative levels of β-galactosidase activity compared with that of the exact gap deletion variant of BA3 (absolute level = 19,000 units) plus one standard deviation unit from assays performed in triplicate are shown above each deletion construct. The deletion constructs were assembled from 3′ overlapping oligonucleotides, extended, and cloned into HindIII- and BglII-restricted (HB series) or HindIII- and ApaI-restricted (HA1) BA3 vector.

2.3%; BA7-4, 2.5%; BA10-1, 2.5%; Figure 2), indicating that translation resumes after nucleotide 197 but before nucleotide 201. The efficient bypass observed with the sense codon replacement (BA10-2, 98%; BA10-3, 91%; Figure 2) indicates that sequences 3′ of the gap can be substantially altered. The role of the GGA located at nucleotides 198 to 200 is addressed below.

Unlike the 3′ flanking sequence, deletion of gene *60* sequence located 5′ of the gap substantially reduces expression of β-galactosidase from these fusions (Figures 3 and

Figure 4. Internal Deletions and Insertions in the 5′ Segment

(A) Deletions were created in BA3 from the BglII site with BAL-31 nuclease, followed by trimming with T4 DNA polymerase and recircularization via blunt-end ligation and transformation. The relative levels of β-galactosidase activity shown are in comparison with that of the exact gap deletion variant of BA3.

(B) Insertion variants BG1 and BG3a were generated by cloning overlapping oligonucleotides containing BglII sticky ends into BglII-cut BA3 vector. An XhoI site located in the BG3a construct was used to generate BG3b via XhoI and BglII restriction and fill-in of BG3a.

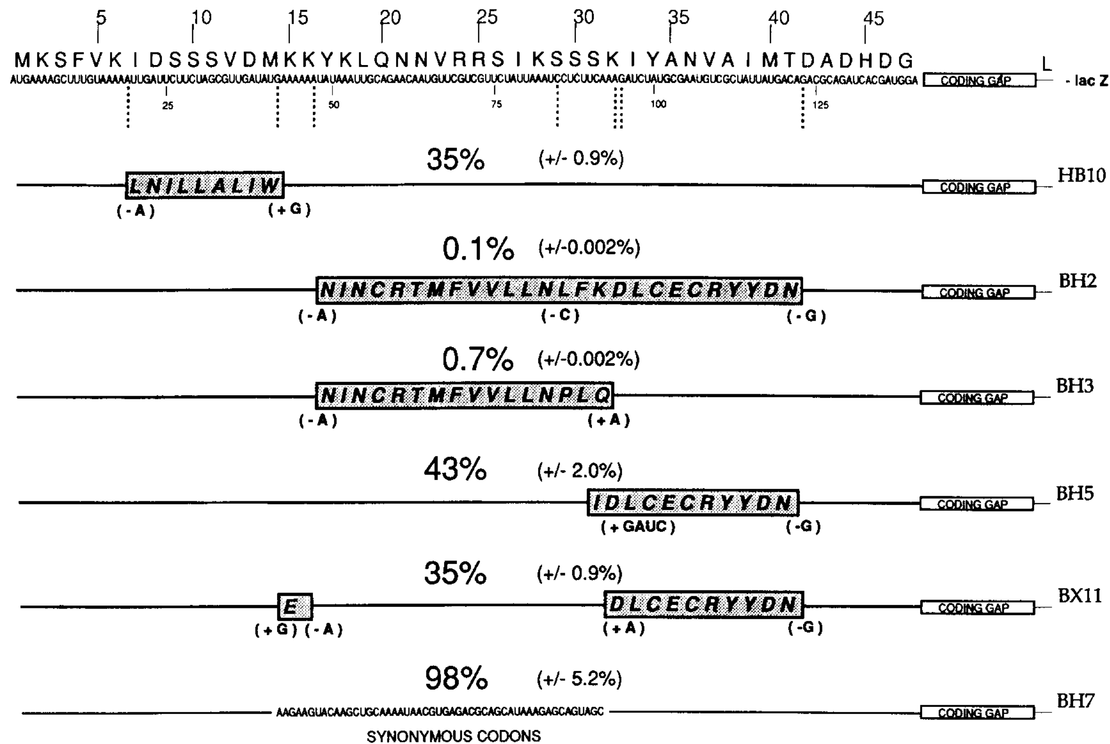


Figure 5. Combinations of Small Insertion and Deletion Mutations That Alter Sections of the Reading Frame of the 5′ Segment

The nucleotide sequence and zero frame peptide sequence of BA3 are included for reference. Each "phase" mutant shows the combinations of insertions and deletions used to generate the shifts in reading frame, as well as the amino acid sequence translated as a result of the frameshifts. The relative levels of β-galactosidase activity shown are in comparison with that of the exact gap deletion variant of BA3. Variants HB10, BH3, and BH7 were generated in a single step, analogous to the assembly of the HB series (Figure 3). Variants BH2, BH5, and BX11 were generated in a single step from variant BH1 (not shown), which contains the $-G_{123}$ mutation.

4). Deletions that progessively remove nucleotides from the 5′ end of the gene *60* sequence cluster into four incremental levels of β-galactosidase expression. Deletions of 24 or 30 nucleotides decrease expression by less than 2-fold. More substantial decreases are seen with deletions of 39, 42, or 45 nucleotides (10-fold), 48 or 51 nucleotides (100- to 300-fold), and 132 nucleotides (5000-fold).

Deletions originating internally in the 5′ gene *60* coding segment also show incremental levels of decreased expression (Figure 4). Deletion of 27, 42, and 51 nucleotides results in 5- to 12-fold decreases, while deletion of 54, 57, 63, and 78 nucleotides results in 70-to 300-fold decrements. BB17 (12%; Figure 4A) lacks the 51 nucleotides preceding the GGA codon bordering the 5′ junction of the coding gap, indicating that these nucleotides are not essential for moderate levels of gap bypass. The expression of BB17 (−51 internal nucleotides, 12% activity) as compared with that of HB8 and HB2, which lack 48 and 51 5′ nucleotides and show <1% activity, suggests that the extent of bypass is not determined simply by the length of the the 5′ coding segment. Internal insertions into the the 5′ coding segment of 12 and 24 nucleotides are tolerated well (BG1 and BG3b; Figure 4B), while a longer insertion of 48 nucleotides results in a 50-fold reduction in expression (BG3a, 1.9%).

The 5′ Segment Encodes a *Cis*-Acting Peptide

The β-galactosidase activities displayed by the 5′ and internal deletions suggest that a sequence element in the central region of the 5′ segment is necessary for efficient bypass. The possibility that a peptide sequence translated from this region is involved in gap bypass was tested with a series of compensating insertions and deletions placed throughout the 5′ segment. These combinations of insertions and deletions route the ribosome in and out of different reading frames through the 5′ region. The levels of gap bypass in these "phase" mutations are shown in Figure 5.

In a 5′ segment encoding 47 amino acids, alteration of amino acids 7 to 15 (HB10, 35%) or 34 to 43 (BH5, 43%) has only a slight effect on gap bypass. However, altering amino acids 17 to 42 (BH2, 0.1%) or 17 to 32 (BH3, 0.7%) causes a large reduction in the amount of bypass. That this reduction in bypass results from the reading frame change and not the single base insertion and deletion generating the phase change is shown by BX11 (Figure 5). In BX11, the two mutations from BH3 are retained, but are now combined with two flanking mutations such that residues 17 to 32 are translated from their original frame. The level of bypass observed with BX11 (35%) is similar to that observed with BH5 (43%), which frameshifts amino

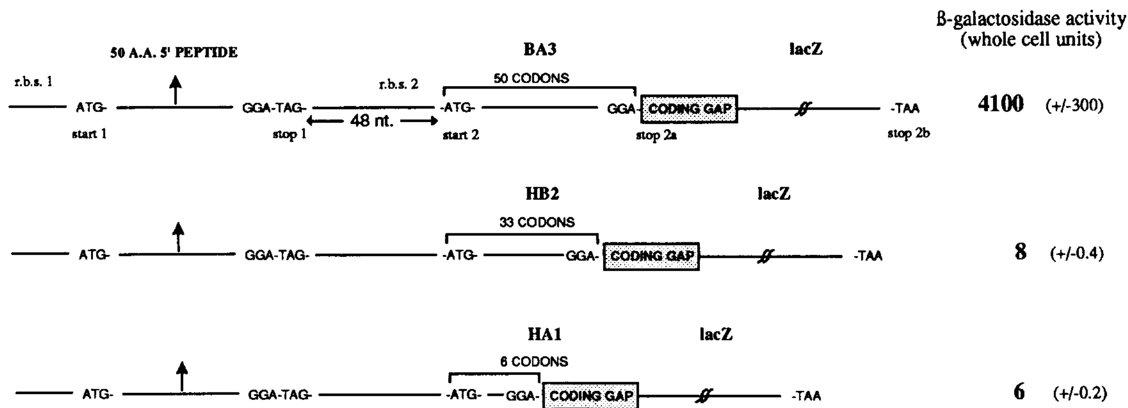


Figure 6. Lack of Complementation of 5′ Segment Deletions by Supplying the 5′ Segment Peptide in *Trans*

The plasmids carrying dicistronic variants were assembled by ligating a 1.9 kb PstI–XhoI fragment containing the 5′ end of the amp^R gene through the tac promoter and 5′ gene *60* segment from BX5a (Figure 9) to 5.7 kb BanHI–PstI fragments of BA3, HB2, or HA1 containing the gene *60–lacZ* fusions from their ribosome binding sites through the 3′ end of the amp^R gene. The XhoI–BamHI junction was created by fill-in and blunt-end ligation of the two sites.

acids 34 to 43. A likely explanation for the reduction in bypass observed with BH2 and BH3 is that amino acids 17 to 32 are not translated from their original reading frame.

The necessity for original frame translation of codons 17 to 32 may be a consequence of the translated peptide or simply the act of translating those particular codons. For instance, it may be the rate of translation across this sequence that activates gap bypass. Such a requirement should be sensitive to codon choice. Codons for amino acids 17 to 32 were replaced by synonomous codons (BH7 = 98%, 24 of 52 nucleotides substituted; Figure 5), with no effect on bypass. Thus, the peptide translated from this sequence can be implicated in efficient gap bypass.

The peptide encoded by the 5′ segment could be affecting gap bypass either in *cis* or *trans*. If the peptide can function efficiently in *trans*, provision of the 5′ segment peptide from an autonomous coding region should restore activity of the 5′ deletions. Figure 6 shows the activities of dicistronic constructs with the 5′ segment translated upstream of active and inactive gene *60-lacZ* fusions. The lack of substantial *lacZ* activities in these variants indicates that the primary effect of the 5′ segment peptide occurs in *cis*.

The loss of gap bypass seen with the phase variants BH2 and BH3 implicates amino acid residues 17 to 32 as activators of gap bypass, but does not indicate which features of the KYKLQNNVRRSIKSSS sequence are important. A preliminary assessment of the information content within this sequence is made by seven random replacements of amino acids 17 to 32; the result is shown in Table 1. The replacements exhibit activities ranging from 0.1% to 5.0%, each relative to their exact gap deletion variant. Although it is not possible to define exact features of the peptide by these variants, the slight reactivation of HB12-GG and HB12V suggests that it may consist of a relatively simple sequence combination.

Gap Insertions, Deletions, and Substitutions

The initial formulation of the ribosomal bypass model anticipated that certain features of the gap sequence would be important for bypass (Huang et al., 1988). Prominent features include a short duplication of the nucleotides U-GGA-U at the 5′ and 3′ junctions of the gap, a potentially stable stem-loop structure at the 5′ end of the gap, a stop codon at the 5′ junction, and a possible secondary structure that would bring the 5′ and 3′ ends of the gap into close proximity. The contributions of each of these features is evaluated with gap variants.

Figure 7 displays the gap sequence in two possible forms: a 10 bp stem structure at the 5′ junction or an interdigitated double stem-loop that brings the 5′ and 3′ ends of the gap into close proximity. The locations and activities of three deletion variants within the 50 nucleotide gap are shown. Deletions of nucleotides 174 to 177 or nucleotides 169 to 180 (XA1 = 31%; XA3 = 27%) allow gap bypass at moderately high levels, and a deletion of nucleotides 172 to 197 (XA7 = 2.8%) still retains a small, but signifi-

Table 1. Random Replacements of Amino Acids 17 to 32 of the 5′ Segment Peptide

Construct	Amino Acids 17 to 32	Relative β-Galactosidase Activity (%)[a]
BA3	KYKLQNNVRRSIKSSS	98 ± 2.1
HB12GG	SPQQLLRTTRHECSGC	5.0 ± 0.1
HB12V	ISLLPSIGYPMISDAK	4.1 ± 0.3
HB12Z	SNLEDYSGEPLRFLNS	1.5 ± 0.03
HB12BB	RSRPACLLLAFDLLFA	0.8 ± 0.03
HB12P	PLHARDATPTQQPTPH	0.7 ± 0.02
BH3	NINCRTMFVVLLNPLQ	0.7 ± 0.002
HB12R	GERQPLEAQAKSAPTQ	0.1 ± 0.003
HB12HH	SQETRDAAGLYPKLGD	0.1 ± 0.001

[a] The HB12 series was generated by extension of 3′ overlapping oligonucleotides followed by cloning into HindIII- and BglII-cut BA3 vector DNA. One oligonucleotide was synthesized with equal amounts of A, G, C, and T phosphoramidites added for the nucleotides encoding residues 17 to 32. The β-galactosidase activities from the HB12 series were measured relative to exact gap deletion variants of each construct.

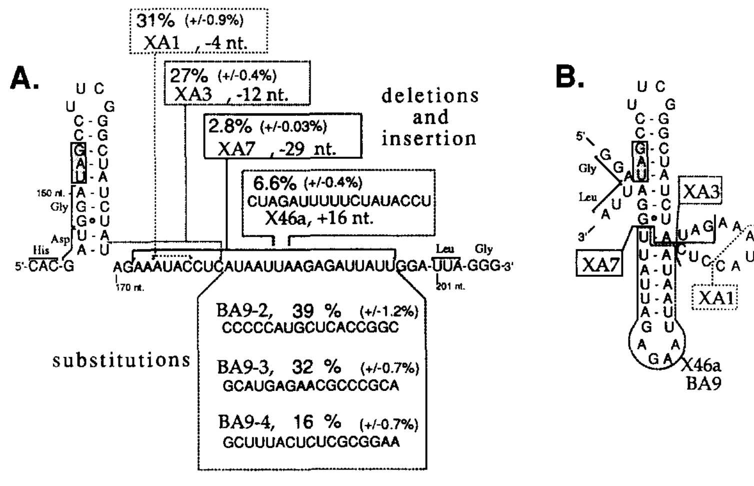


Figure 7. Insertion, Deletion, and Substitution Variants of the Gap Sequence

The gap sequence is shown with a potential stem-loop formed at the 5′ junction and the stop codon at the 5′ junction highlighted. The deleted, substituted, or inserted nucleotides are indicated by the arrows and boxes. The deletion variants were built by inserting overlapping oligonucleotides with XhoI and ApaI sticky ends into XhoI- and ApaI-cut BX5a vector DNA. The deletion variants contain the BX5a loop substitution (see Figure 9), while the BA9 series and X46a do not. The relative levels of β-galactosidase activity shown are in comparison with that of the exact gap deletion variant of BA3. Also shown are the locations of the deletions and substitutions on a hypothetical secondary structure of the gap sequence that brings the 5′ and 3′ junctions into close proximity.

cant amount of gap bypass. Substitutions of nucleotides 181 to 197 have moderate effects (BA9-2, 39%; BA9-3, 32%; BA9-4, 16%), and a 16 nucleotide insertion (X46a = 6.6%) shows a reduced amount of bypass. The degree of bypass retained by these variants suggests that the 3′ half of the gap is malleable.

Figure 8 displays the activities associated with various loop changes, stem changes, and stop codon substitutions at the 5′ end of the gap. Replacement of the stop codon by either of two sense codons (plus the compensating changes to maintain stem base pairs) results in reduced bypass (BX15, UGG = 11%; BX1q, UAC = 1.7%). Disruption of potential base pairs at the top of the stem results in 10- to 30-fold decreases in activity (BX1d, BX1z, and BX1j; Figure 8). This effect can be reversed by restoring potential base pairs with compensating changes (BX1x = 70%). Changes in the length of the stem and the size and sequence of the loop also decrease activity. Altering the loop sequence (BX5a = 57%) or doubling its size (BX5b = 35%) causes moderate decreases, while length-

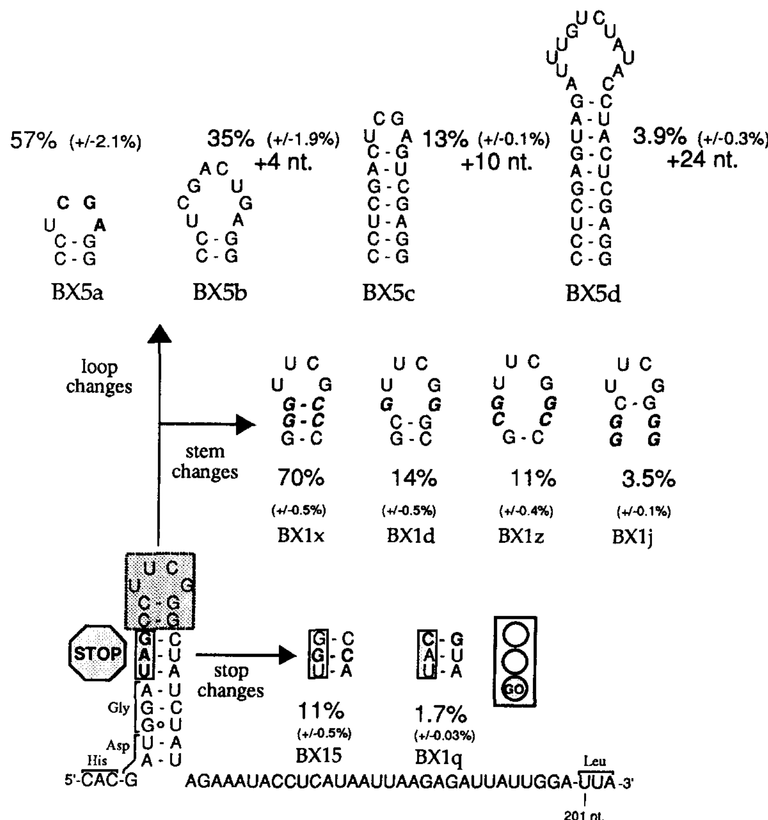


Figure 8. Potential Stem-Loop Changes and Stop to Sense Changes at the 5′ End of the Gap Sequence

The relative levels of β-galactosidase activity shown are in comparison with that of the exact gap deletion variant of BA3.

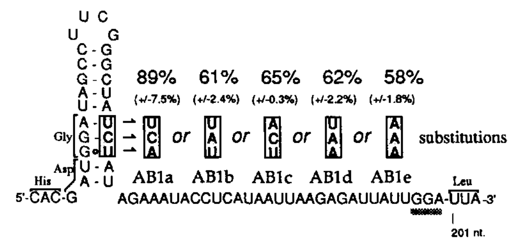
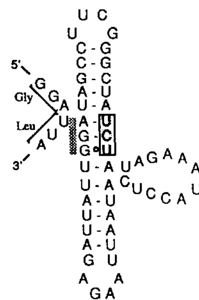



Figure 9. Ribosome Jumping Model versus mRNA Bridging Gap Model

Substitutions of the UCU sequence at positions 15 to 17 of the gap are displayed along with the relative levels of β-galactosidase activity in comparison with that of the exact gap deletion variant of BA3. The UCU sequence is shown base paired to either the 5′ GGA (jumping model) or the 3′ GGA (bridging model).

ening the potential stem results in more substantial decreases (BX5c, +10 nucleotides = 13%; BX5d, +24 nucleotides = 3.9%). The increase in gap size caused by the longer stems leads to similar decreases in bypass efficiency relative to an insertion elsewhere in the gap (X46a, +16 nucleotides = 6.6% versus BX5c and BX5d; Figure 7).

Jumping versus Bridging the Gap

The depiction of the gap sequence in two possible forms (Figure 7) suggests two distinct models of gap bypass: ribosome jumping and mRNA bridging (Huang et al., 1988). A ribosome jumping model would involve dissociation of the peptidyl-tRNA from codon 46 (GGA) and binding of the peptidyl-tRNA to the GGA located at the 3′ junction. An mRNA bridging model would invoke formation of an mRNA structure that brings the 5′ and 3′ ends of the gap into close physical proximity. An essential distinction between these two models is how each locates (or defines) the 3′ end of the gap. The jumping model, in its simplest form, locates the 3′ end via codon–anticodon dissociation of the peptidyl-tRNA from the 5′ GGA followed by codon–anticodon binding to the 3′ GGA after passage over the gap sequence. The bridging model pins the 3′ end of the gap close to the 5′ end via base pairing between the UCU gap sequence (nucleotides 165 to 167) and the 3′ GGA (nucleotides 198 to 200).

Alteration of gap nucleotides 165 to 167 (UCU) would help differentiate these two views of gap bypass. These nucleotides can potentially base pair with the GGA codon at the 5′ junction or the GGA at the 3′ junction. Five substitutions of the UCU sequence that would partially or fully disrupt base pairing with GGA are shown in Figure 9. Their activities range from 58% to 89%, indicating that this UCU is not involved in a base pairing interaction essential for bypass. A ribosome jumping model would not necessarily require occlusion of the 5′ GGA via base pairing with the UCU gap sequence. Other factors, such as the nascent peptide segment or the stop codon and top of the 5′ stem, may be essential components for initiating effective dissociation of the peptidyl-tRNA from the 5′ end of the junction. However, the lack of an effect from the UCU substitutions strongly disfavors an mRNA bridging model like the one shown in Figure 9.

A clear prediction of a ribosome jumping model is a requirement for matched codons located at the junctions of the gap. Figure 10 shows the activities of constructs with

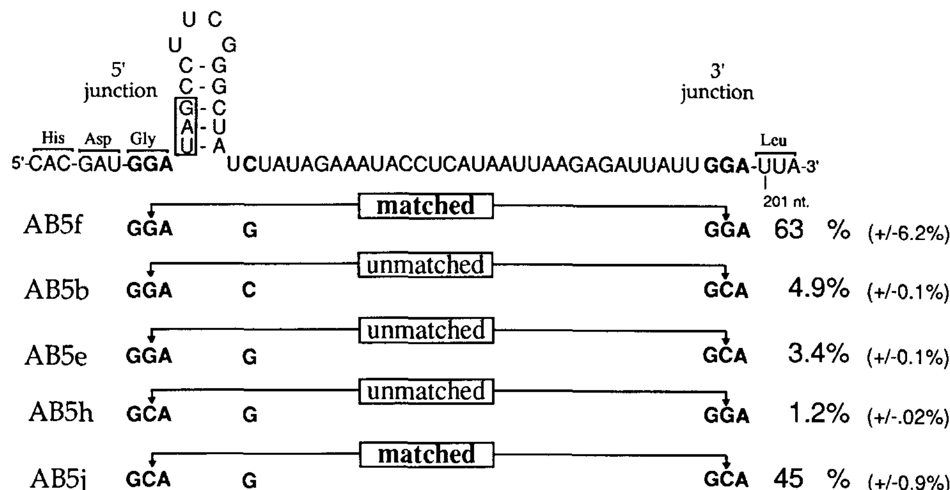


Figure 10. Matched versus Unmatched Codons at the 5′ and 3′ Junctions of the Gap

The relative levels of β-galactosidase activity shown are in comparison with that of the exact gap deletion variant of BA3. The AB5 series contains GGA or GCA codons at the 5′ and 3′ junctions, as well as a C or G at the UCU gap nucleotides 15 to 17. Whether or not the junction codons are matched or unmatched is indicated.

matched or unmatched GGA (Gly) and GCA (Ala) codons at either end of the gap. High levels of gap bypass are observed when 5′ and 3′ junction codons are matched (GCA-GCA, AB5j = 45.3%) but not when unmatched (GGA-GCA, AB5b = 4.9% or GCA-GGA, AB5h = 1.2%).

Discussion

The results presented suggest that E. coli ribosomes are capable of bypassing the coding gap of gene *60* by responding to certain sequence features within and preceding the gap. Mutational analysis implicates the nascent peptide, a stop codon within a short stem at the beginning of the gap, and the distance separating the duplicated flanking codons as key features. The possibility that the juxtaposition of such elements can elicit a ribosomal jump over 50 mRNA nucleotides, at near unitary efficiency, merits further consideration. The critical gap alterations for disrupting bypass occur at the junctions. The 5′ junction displays requirements for a stop codon contained in a short stem-loop structure, plus a penultimate codon matched to the 3′ junction. Disruption of stem base pairs progressively diminishes bypass (14% to 3.5% bypass), which can then be restored by compensating base changes (70% bypass). The sequence of the loop and closing base pair of the stem (CUUCGG) have been observed to confer unusual thermal stability to stem structures (Tuerk et al., 1988). Although changes in the loop sequence have only minor effects (35% to 57% bypass), the occurrence of an extraordinarily stable stem at the 5′ junction may be necessary for maximal bypass.

The 3′ junction appears to have only two easily discernible components: the matched GGA codon and an optimal distance from the 5′ junction. Disruption of the matched set of GGA codons at either the 5′ or 3′ junction diminishes bypass significantly (1.2% to 4.9% bypass; Figure 10). High level bypass can be restored by compensating base changes that create matched GCA alanine codons at the 5′ and 3′ junctions (45% bypass). In both cases, there are no GGA or GCA codons in the coding gap. Inclusion of out-of-frame GGA codons within the gap also reduces the amount of bypass (R. B. W., unpublished data), and the one extensive substitution of gap nucleotides 31 to 47 with anomalously low activity (BA9-4, 16% bypass) contains an internal GGA.

An alternative view of the gene *60* gap bypass mechanism postulated a physical association between the 5′ and 3′ junctions (Huang et al., 1988). The involvement in gap bypass of a possible gap secondary structure that could bring the junctions into close proximity is not supported by the mutational analysis described here (Figures 7 and 9). Of the 50 gap nucleotides, 36 can be deleted or substituted with only minor effects on the level of gap bypass: nucleotides 6 to 9 (57% bypass), 15 to 17 (58% to 89% bypass), and 31 to 47 (16% to 39% bypass) can be substituted, and nucleotides 19 to 30 (27% bypass) can be deleted. This pliable character of the gap suggests that it is not designed to align rigidly the 5′ and 3′ junctions within the ribosome's decoding sites. However, the existence of a mechanism that would physically associate the synonymous codons flanking the gap cannot be completely excluded.

The requirement for a matched set of codons at the 5′ and 3′ junctions is compatible with a relatively simple view of gap bypass, in which the peptidyl-tRNA efficiently dissociates from this codon, the ribosome traverses the gap, and the peptidyl-tRNA binds to the 3′ GGA codon, where normal elongation resumes. The dissociation of a growing peptidyl-tRNA is not completely without precedent. At low levels on certain sequences, tRNAs appear to dissociate from the mRNA and bind to a synonymous codon elsewhere on the mRNA. Such tRNA "hops" were first observed in E. coli, where the residual levels of β-galactosidase activity of certain *lacZ* frameshift and nonsense mutants resulted from a ribosome step size of +5 or +6 nucleotides (Weiss et al., 1987). In these cases, matched sets of codons occurred at the junctions of the mRNA sequence that had been passed over. Direct implication of tRNA in these bypass events has been obtained with mutant valine tRNAs, the *hopR* alleles, which enhance the rate of hopping at certain valine codons (Falahee et al., 1988; O'Connor et al., 1989). One feature of tRNA hops is their directionality: all observed hops have been forward (i.e., toward the 3′ end of the mRNA), leading to deletions of amino acids from the translated polypeptide. The other pertinent observation from tRNA hopping is the inference of a "localized" scanning mechanism (O'Connor et al., 1989), which suggests that the dissociated peptidyl-tRNA can, at high efficiency, resume translation at the take-off site if no suitable landing site is found. If the analogy of gap bypass to low-level tRNA hopping is correct, the key question becomes what other factors can account for the enormous increase in the distance and efficiency over which such a mechanism can operate.

Stop codons have been previously observed to enhance frameshifting and tRNA hopping at the penultimate codon (Weiss et al., 1987; Curran and Yarus, 1989; O'Connor et al., 1989). The stop codon may be enhancing the dissociation of the peptidyl-tRNA from the GGA codon at the 5′ junction. The contribution of the 5′ stem has less precedent. From models of leader function in bacterial attenuators (Landick and Yanofsky, 1988) and certain inducible antibiotic operons (Gryczan et al., 1980; Horinouchi and Weisblum, 1980), as well as certain −1 ribosomal frameshift sites (Jacks et al., 1988; Brierley et al., 1989), it is thought that ribosomes melt mRNA secondary structure before it reaches the ribosomal A (aminoacyl or acceptor) site. The formation of the 5′ junction stem requires circumvention of this activity and may lead to occlusion of the stop codon in the A site. Occlusion of the mRNA at the 5′ junction may also facilitate dissociation of the P (peptidyl or donor) tRNA and/or passage over this sequence by a dissociated tRNA.

Another gap feature is the apparent distance requirement of the 3′ junction from the 5′ junction. Either shortening the gap to 21 nucleotides or lengthening it to 66 nucleotides reduces bypass (XA7 = 2.9%; X46a = 6.6%; Figure 7). If the jumping model is correct, it implies that there may exist a mechanism for ribosome movement along mRNA independent of tRNA movement, as well as

tracking mechanisms other than codon–anticodon pairing. An analogy between ribosomes jumping the coding gap and scanning 40S subunits postulated for eukaryotic initiation can be made readily. During scanning it is thought that the 40S ribosomal subunit tracks along an mRNA until the first codon–anticodon match is found (Cigan et al., 1988). One potential impetus for gap bypass in a scanning model may be the movement of ribosomes through the 5′ coding segment. A ribosome situated at the 5′ junction could be pushed through the gap as the trailing ribosomes continued to elongate. The distance such a push would propel a ribosome may be roughly equal to the region of mRNA protected by a ribosome from nuclease digestion, which is approximately 35 bases (Kang and Cantor, 1985). The distance separating the junctions may also reflect the need to swap the gap sequence into an mRNA binding track, perhaps placing the 3′ GGA codon close to the P site in analogy to prokaryotic ribosome binding sites. A requirement for trailing ribosomes and the implication that the ribosome can maintain threading of the mRNA without tRNA–mRNA pairing should be testable.

An alternative to this type of scanning or tracking mechanism is a local search model. In low level tRNA hopping and in suppressor tRNA hopping, it has been postulated that the hopping tRNA can return to the take-off codon if no suitable landing codon is encountered. An experimental distinction between scanning and a local search should be possible by further characterization of landing site choice.

Significance of the 5′ Coding Segment and Nascent Peptide

The mutational analysis of the 5′ coding segment identified a region of the nascent peptide that is critical for gap bypass.

The ribosome is capable of protecting between 30 and 40 residues of the nascent polypeptide chain from proteolytic digestion (Smith et al., 1978), and it has been surmised that the chain exits the ribosome through a defined channel (Bernabeu and Lake, 1982; Yonath et al., 1988). Residues 17 to 32 of the nascent peptide may thus be within the ribosome as the decoding sites reach the 5′ junction. Action of the nascent chain within the ribosome may distinguish the gene *60* nascent peptide from other examples of functional effects of nascent chains, such as signal recognition particle arrest of elongation (Wolin and Walter, 1988) and the autoregulated instability of β-tubulin mRNA (Yen et al., 1988). How might a nascent peptide within the ribosome, in terms of a jumping model, facilitate take-off or landing?

One possibility is that the nascent chain is capable of binding the gap sequence, perhaps "bootstrapping" the ribosome over the gap. However, the malleability of the gap sequence revealed by the mutational analysis makes it difficult to imagine where this binding site would be located. Another possibility is that the nascent chain transiently plugs the exit channel, perhaps destabilizing the peptidyl-tRNA–mRNA complex as the chain piles up within the ribosome. The mild effects of large internal insertions and deletions make this model less plausible. A third possibility is that the nascent peptide binds the ribosome with functional consequences, in analogy to certain antibiotics that inhibit protein synthesis.

One relevant antibiotic example is the basic oligopeptide edeine, which blocks P site binding of tRNA (Szer and Kurylo-Borowska, 1970). Edeine also inhibits recognition of AUG start codons by scanning 40S ribosomal subunits during initiation on eukaryotic mRNA (Kozak and Shatkin, 1978) and mimics the chemical footprint observed by P site tRNA binding to 30S particles (Moazed and Noller, 1987). Destabilizing P site tRNA binding may be a key event in triggering a ribosomal jump. If the nascent peptide sequence can activate bypass by destabilizing P site tRNA binding, then such a mechanism may have interesting use as a sequence-specific local modulator of translational accuracy.

Ribosomes do not usually jump 50 nucleotides while translating. The patterns found by this mutational analysis suggest that the required mRNA components to elicit this event are not very exotic. Thus, other genes may also be taking advantage of the possibilities of high level ribosome jumping. The puzzle of why the gene *60* coding sequence contains a high level ribosome jump was not addressed in this study. There is no genetic evidence to implicate a regulatory function to the jump. The mystery of its existence remains unsolved.

Experimental Procedures

Assembly and Cloning of Gene *60* Sequence into a Reporter *lacZ* Gene

Cloning of synthetic oligonucleotides into the HindIII and ApaI sites in a *lacZ* reporter gene and verification by plasmid DNA sequencing were as described (Weiss et al., 1987). The *lacZ* gene (from codon 6 through the AvaI site at codon 66 of *lacY*, 3317 nucleotides) is located on pBR322, replacing the EcoRI–AvaI segment, and transcribed clockwise from a tac promoter (Weiss et al., 1987). Oligonucleotides were synthesized on an ABI 380B DNA synthesizer and gel purified prior to cloning. The initial HA34, HA32, and BA3 gene *60–lacZ* fusions were constructed by PCR amplification of a set of four overlapping oligonucleotides, followed by cleavage of recessed HindIII and ApaI sites and cloning. All subsequent constructions used the internal BglII site created in BA3 as a cloning site for oligonucleotides with HindIII and BglII or BglII and ApaI sticky ends. Lengthy inserts were generated by a fill-in reaction with DNA polymerase I (Klenow fragment) of oligonucleotides with 3′ overlaps of 12 to 15 nucleotides, followed by cleavage of recessed cloning sites. The set of internal deletions was created by limited BAL-31 digestion of BglII-cut BA3 plasmid DNA.

Frameshifted β-galactosidase was purified from saturated broth cultures grown at 37°C with a 0.5 ml bed volume anti–β-galactosidase column (Protosorb, Promega), eluted with 3 vol of 0.1 M $NaCO_2$ and spun-dialyzed in Centricon 30 ultrafiltration cartridges (Amicon). The amino acid sequences were determined on an ABI 470a/900/120 liquid-pulse protein sequencer equipped with on-line HPLC analyzer and data analysis modules.

β-Galactosidase Activity Measurements

Whole-cell β-galactosidase assays were based on the procedure according to Miller (1972). All plasmid constructs were maintained in E. coli SU1675 $F'I^Q$ (Weiss et al., 1987). Overnight cultures were diluted 1:20 into M9 minimal media plus 0.2% glucose, 0.4% casamino acids, 0.01% thiamine, and 2 mM IPTG, grown at 37°C with rapid shaking, and chilled to 0°C at an OD_{600} of 0.6–0.8. Whole-cell assays were done at 28°C, in a final volume of 1.0 ml of buffer Z; the assays were started with the addition of 0.2 ml of 4 mg/ml O-nitrophenyl-β-galactoside and stopped with the addition of 0.5 ml of 1 M Na_2CO_3 (pH 11.0). The assay tubes were centrifuged before reading the A_{420}. The whole-cell

unit definition = (1000 × A_{420})/(OD_{600} × time [min] × vol [ml]). All assays were measured in triplicate and the standard deviation was calculated; the activities are presented as a percentage of in-frame β-galactosidase activity measured from a related construct.

Acknowledgments

We thank Ray Gesteland for encouraging this effort and Marilyn Kozak for discussion about the effect of edeine on scanning 40S ribosomal subunits. W. M. H. is supported by a grant from the National Institutes of Health. This work was supported by the Howard Hughes Medical Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 16, 1990; revised April 25, 1990.

References

Bernabeu, C., and Lake, J. A. (1982). Nascent polypeptide chains emerge from the exit domain of the large ribosomal subunit: immune mapping of the nascent chain. Proc. Natl. Acad. Sci. USA *79*, 3111–3115.

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

Cigan, A. M., Feng, L., and Donahue, T. F. (1988). $tRNA_{imet}$ functions in directing the scanning ribosome to the start site of translation. Science *242*, 93–97.

Curran, J. F., and Yarus, M. (1989). Rates of aminoacyl-tRNA selection at 29 sense codons *in vivo*. J. Mol. Biol. *209*, 65–78.

Falahee, M. B., Weiss, R. B., O'Connor, M., Doonan, S., Gesteland, R. F., and Atkins, J. F. (1988). Mutants of translational components that alter the reading frame by two steps foward or one step back. J. Biol. Chem. *263*, 18099–18103.

Gryczan, T. J., Grandi, G., Hahn, J., and Dubnau, D. (1980). Conformational alteration of mRNA structure and the posttransciptional regulation of erythromycin-induced drug resistance. Nucl. Acids Res. *8*, 6081–6097.

Horinouchi, S., and Weisblum, B. (1980). Posttranscriptional modification of mRNA conformation: mechanism that regulates erythromycin-induced resistance. Proc. Natl. Acad. Sci. USA *77*, 7079–7083.

Huang, W. M., Ao, S., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D., and Fang, M. (1988). A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene *60*. Science *239*, 1005–1012.

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

Kang, C., and Cantor, C. R. (1985). Structure of ribosome-bound messenger RNA as revealed by enzymic accessibility. J. Mol. Biol. *181*, 241–251.

Kozak, M., and Shatkin, A. J. (1978). Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. J. Biol. Chem. *253*, 6568–6577.

Landick, R., and Yanofsky, C. (1988). Transcription attenuation. In *Escherichia coli* and *Salmonella typhimurium*: Cellular and Molecular Biology, F. C. Neidhardt, ed. (Washington, D.C.: American Society of Microbiology), p. 1276.

Liu, L., Liu, C., and Alberts, B. M. (1979). T4 DNA topoisomerase: a new ATP-dependent enzyme essential for initiation of T4 bacteriophage DNA replication. Nature *281*, 456–461.

Miller, J. H. (1972). Experiments in Molecular Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Moazed, D., and Noller, H. F. (1987). Interactions of antibiotics with functional sites in 16S rRNA. Nature *327*, 389–394.

Mufti, S., and Bernstein, H. (1974). The DNA-delay mutants of bacteriophage T4. J. Virol. *14*, 860–871.

O'Connor, M., Gesteland, R. F., and Atkins, J. F. (1989). tRNA hopping: enhancement by an expanded anticodon. EMBO J. *13*, 4315–4323.

Smith, W. P., Tai, P.-C., and Davis, B. D. (1978). Interaction of secreted nascent chains with surrounding membrane in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA *75*, 5922–5925.

Stetler, G., King, G., and Huang, W. M. (1979). T4 DNA-delay proteins, required for specific DNA replication, form a complex that has ATP-dependent DNA topoisomerase activity. Proc. Natl. Acad. Sci. USA *76*, 3737–3741.

Szer, W., and Kurylo-Borowska, Z. (1970). Effect of edeine on aminoacyl-tRNA binding to ribosomes and its relationship to ribosomal binding sites. Biochem. Biophys. Acta *224*, 477–486.

Tuerk, C., Gauss, P., Thermes, C., Groebe, D. R., Gayle, M., Guild, N., Stormo, G., D'Aubenton-Carafa, Y., Uhlenbeck, O. C., Tinoco, I., Brody, E. N., and Gold, L. (1988). CUUCGG hairpins: extraordinarily stable RNA secondary structures associated with various biochemical processes. Proc. Natl. Acad. Sci. USA *85*, 1364–1368.

Weiss, R. B., Dunn, D. M., Atkins, J. F., and Gesteland, R. F. (1987). Slippery runs, shifty stops, backward steps and foward hops: −2, −1, +1, +5 and +6 ribosomal frameshifting. Cold Spring Harbor Symp. Quant. Biol. *52*, 687–693.

Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., and Gesteland, R. F. (1988). Reading frame switch caused by base-pair formation between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis. EMBO J. *7*, 1503–1507.

Wolin, S. L., and Walter, P. (1988). Ribosome pausing and stacking during translation of eukaryotic mRNA. EMBO J. *7*, 3559–3569.

Yen, T. J., Machlin, P. S., and Cleveland, D. W. (1988). Autoregulated instability of β tubulin mRNAs by recognition of the nascent amino terminus of β tubulin. Nature *334*, 580–584.

Yonath, A., and Wittman, H. G. (1988). Crystallographic and imagery construction studies on ribosomal particles from bacterial sources. Meth. Enzymol. *169*, 95–116.