Characterization of Ribosomal Frameshift Events by Protein Sequence Analysis*

Timothy J. Dayhoff, John F. Atkins, and Raymond F. Gesteland

From the Howard Hughes Medical Institute, Departments of Biology and Human Genetics, University of Utah, Salt Lake City, Utah, and the Department of Biochemistry, University College Cork, Cork, Ireland

In cell-free protein synthesis studies with RNA from phage MS2 as template, normal Escherichia coli tRNAs, promote two base translocation at GCA alanine codons with a resultant shift of ribosomes to the minus one reading frame. Similarly, normal tRNAThr promotes two base translocation at CCG proline codons. These conclusions were reached by amino acid sequencing of tryptic peptides or cyanogen bromide fragments that contained the reading frame shift site. It is proposed that these frameshift events occur by a two-base pair interaction between the anticodons of these exceptional tRNAs and the noncognate codons.

It has recently become apparent that some proteins can be synthesized only if a specific ribosomal frameshift event occurs allowing decoding of adjacent sequences of their mRNAs in two different reading frames. The most convincing examples include the product of gene 10 of the DNA bacteriophage T-7 (1), the release factor 2 of Escherichia coli (2), a protein encoded by a yeast transposon (3, 4) and the gag-pol protein of Rous sarcoma virus (5). The translational complex must infrequently but reliably move in a nontriplet manner on the mRNA. One presumed manifestation of this is frameshift mutants which are sometimes leaky, i.e., a small proportion of the gene product is essentially normal (6-9). It appears that this leakiness is due neither to translation reinitiation (6) nor to unusual tRNAs (as occur in frameshift suppressor mutants) but rather to frameshifting by normal tRNAs that restore the correct reading frame.

The frequency of frameshifting can be manipulated by specifically altering the balance of certain tRNAs either by inhibition of aminoacylation of particular tRNAs (10-12) or by addition of purified tRNAs to an in vitro protein synthesizing system containing the normal complement of tRNAs (13). When bacteriophage MS2 RNA (14) is translated in vitro, ribosomal frameshifting just before the end of the MS2 coat protein gene causes the normal terminator to be bypassed. This results in synthesis of proteins longer than the coat protein which are readily detected by size fractionation (diagrammed in Fig. 1). A shift to the plus one frame gives rise to protein 5, a coat-lysis hybrid. Shifting to the minus one frame yields either protein 6 or 7. Addition of AGU/AGC decoding tRNAs (15, 16) enhances synthesis of both 6 and 7. This effect is ameliorated by tRNAAla, which decodes GCU/GCA/GCG (17), as if the serine tRNA is reading alanine codons to give a minus one shift. Similarly, ACU/ACC decoding tRNAThr (18, 19) appears to cause minus one frameshifting at proline codons in both the coat protein and synthetase genes of MS2 (13).

When a ribosome shifts reading frame, it will synthesize a hybrid protein until it encounters the first terminator in the new reading frame. The resultant protein will be shorter or longer than the normal product depending upon the location of the first downstream terminator in the new reading frame. For proteins 6 and 7, the “window” available for frameshifting is delimited by the zero frame terminator and by the first upstream terminator in the minus one frame.

One hypothesis is that minus one frameshifting occurs at all proline or alanine codons, and as shown in Fig. 2, there are many such codons in the window that could give rise to proteins 6 and 7. The hybrid proteins produced by each of these events would have an amino acid sequence different from the others at a few internal positions. Here we report amino acid sequence analyses of the reading frame shift sites of several hybrid proteins that pinpoint specific alanine and proline codons at which serine and threonine are inserted concomitantly with the shift in reading frame. This is diagrammed in Fig. 2. Further information on frameshifting at alanine and proline codons was inferred by less direct exper-

*This work was supported by the National Science Foundation Grant DMB-8408649 (to R. F. G. and J. F. A.) and by the Irish National Board for Science and Technology Grant 18/82 (to J. F. A.) and 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.

FIG. 1. MS2 frameshift proteins. A diagram of SDS-PAGE fractionation of proteins made in vitro. Proteins 6, 7, and 9 arise from shifts into the minus one frame near the 3’ end of the coat protein gene, and protein 5 is a coat-lysis hybrid resulting from a shift to the plus one frame. The 66K protein results from a shift to the minus one frame near the 3’ end of the synthetase gene.
frameshifting

FIG. 2. Frameshifting sites. A summary of the results of analyses of the frameshift products from the coat protein gene. The vertical lines mark the frameshift sites detected and are at alanine or proline codons. Frameshifting was not detected at the two proline codons marked with x. Numbers in parentheses indicate intervening codons. Stop codons are underlined and reading frames are indicated by the grouping of triplets.

MATERIALS AND METHODS

mRNAs—Preparation of phage RNAs was as described earlier (20). Phages GA and SP were provided by Dr. A. Hirashima, Yakult Honsha Co., Tokyo.

tRNAs—These were purchased from Subbriden (Seattle), the same source as used previously.

Cell-free Protein Synthesis—In vitro translations were performed as described earlier (13). The concentrations of added tRNAs were 160 \mu M for tRNA\text{Ser} or tRNA\text{Thr} (twice this for preparative reactions) and 200 \mu M of the various tRNAs used as competitors. Radiolabeled amino acids were the highest specific activities available from New England Nuclear and were used as described (13).

Protein and Peptide Isolation—Coat protein was extracted from phage MS2 using 67% acetic acid (21). RNA was removed by centrifugation, and coat protein was precipitated by neutralization with ammonium hydroxide at 0 °C. After 30 min the precipitate was collected by centrifugation, washed twice by resuspension in a one-half volume of water and centrifugation, and finally lyophilized.

Proteins were electroeluted from polyacrylamide gel slices as described (22), but unlabeled MS2 coat protein was added (1–2 mg/slice) and the concentration of sodium dodecyl sulfate was increased to 5% in order to keep the coat protein in solution.

Digestion with trypsin (13) was with a protein concentration of 5 mg/ml and an enzyme concentration of 50 \mu g/ml. For fractionation by HPLC, the lyophilized peptides were suspended in 0.3–0.4 ml of equilibration buffer (0.05% trifluoroacetic acid) and applied to an Ultrapak C-18 column (Altex). Elution was with a flow rate of 1 ml/min at room temperature and used a linear gradient of 0.05% trifluoroacetic acid in acetonitrile from 0–55% increasing at 1%/min except for a 5-min plateau at 20%. The optical density of the eluate was monitored at 210 nm, and 0.5 ml fractions were collected. The chromatographic profile for peptides from purified coat protein is shown in Fig. 3 where the numbered positions correspond to specific peptides identified in Table I. For samples containing radioactive peptides, 10% of each sample was assayed for radioactivity with a Beckman LS-8000. Fractions were pooled and lyophilized prior to further fractionation (Fig. 4). The same HPLC column and conditionswere used except that the gradient was linear, and the buffer contained either 10 mM ammonium acetate (organic phase, acetonitrile/water (4:1)) or 0.05% n-heptanfluorobutyric acid. Additional carrier coat protein tryptic peptides were added before fractionation.

Proteins were cleaved with cyanogen bromide (13), and the resultant peptides were lyophilized, dissolved in 200 \mu l of 6 M guanidine HCl, 0.1% trifluoroacetic acid, and fractionated by HPLC using a 300A pore Synchronpak C-18 column (Alltech) heated to 65 °C. The elution gradient was 1%/min of n-propyl alcohol (HPLC) grade.
Frameshifting

TABLE I

<table>
<thead>
<tr>
<th>Protein</th>
<th>Peptide number</th>
<th>Mobility</th>
<th>Amino acid sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coat</td>
<td>9T</td>
<td>0.21</td>
<td>E L I V(K)</td>
</tr>
<tr>
<td></td>
<td>6S</td>
<td>0.21</td>
<td>S P(R)</td>
</tr>
<tr>
<td></td>
<td>7T</td>
<td>0.12</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>7Sa</td>
<td>0.12</td>
<td>L(R)</td>
</tr>
<tr>
<td></td>
<td>7Sc</td>
<td>0.12</td>
<td>(H) L I D A G(H.S)K</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0.43</td>
<td>I T(H)V E D(R)K</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.29</td>
<td>E V Q L F(R)</td>
</tr>
</tbody>
</table>

Burick & Jackson Laboratories) in 0.1% trifluoroacetic acid with a flow rate of 1 ml/min. Radioactivity and absorbance were monitored simultaneously using a Radiometric flow detector, model IC, diverting 10% of the eluate to be mixed with scintillant (Flo-scint, IC, 0.5-ml fractions were collected.

Amino Acid Sequencing—For sequencing of short radiolabeled peptides, 0.4 mg of trypsin-digested MS2 coat protein was added as carrier; for the longer cyanogen bromide fragments, myoglobin was used. The radiolabeled peptides were sequenced on a Beckman model 890D sequencer, and the PTH derivatives were fractionated (23). The material from the peak corresponding to the PTH derivative of the radioactive amino acid(s) used was collected and the radioactivity determined by repeated 10-min counts using Aquasol II (New England Nuclear) as scintillant. In some experiments it was necessary to detect both serine and alanine, but the PTH derivatives of these (PTH-dehydroalanine and PTH-alanine derivatives, respectively) are difficult to separate. In these cases the two amino acids were differentially labeled (H and 13C). Phenylisothiocyanate-proline, an intermediate of Edman degradation of proline residues, is slow to hydrolyze. We could often anticipate proline positions based on mRNA sequence, and Edman degradation was augmented by extending the duration of heptafluorobutyric acid cleavage at the predicted positions.

RESULTS

To elucidate the frameshift events that led to the synthesis of the hybrid proteins 6 and 7, these proteins were synthesized in vitro in the presence of selected radioactive amino acids. They were purified by gel electrophoresis (Fig. 5), digested with trypsin, and the resultant peptides separated by HPLC. Nonradioactive coat protein from purified virions was added prior to tryptic digestion to provide internal standards for peptide fractionation. The coat protein peptides (Fig. 3) were identified with the known sequence by: (a) differential absorbance at 210 nm and 280 nm, indicating aromatic residues (data not presented), (b) cochromatography of coat protein tryptic peptides synthesized in vitro in the presence of selected L-[^3]H or L[^14]C amino acids, and (c) sequencing of nine of the 11 tryptic peptides of coat protein to confirm the assignments. The sequences of the coat protein peptides are given in Table I along with predicted sequences of the frameshift products shown in Fig. 2.

Comparative analyses of the profiles of radioactive tryptic peptides from frameshift proteins labeled with up to four amino acids (L[^3]H and L[^14]C or L[^35]S methionine) per experiment (employing a total of 11 amino acids) provided preliminary identification of the peptides. As expected, many peaks are coincident with those from coat protein marker peptides. Sequence analysis of peptides by Edman degradation and identification of the radioactive PTH amino acids showed that the amino termini of the proteins 6 and 7 were encoded by the zero frame (coat protein reading frame) and the carboxy...
**Frameshifting**

**Protein 6**

The synthesis of protein 6 is stimulated by addition of tRNAser, but not tRNAThr. Protein 6 was synthesized in several reaction mixtures with radioactive amino acids included. Since the frameshifting is predicted to occur by trRNAser reading an Ala codon, the key labels are serine and alanine. The tryptic peptides of protein 6, labeled with L-[^14]C]alanine, L-[^14]H]serine, and L-[^14]H]asparagine, are shown in Fig. 6A. The material from the peak designated 20 was further purified (Fig. 6, B and C) and sequenced (Fig. 7, peptide 20) giving Ser-X-Ala-Arg. This sequence can only arise from translation of the coat protein gene in the minus one frame where, by chance, Ala is the first codon in that window which allows synthesis of a product larger than coat protein. Translation in the minus one frame prior to this point would give a truncated fragment shorter than the coat protein. The last codon in the zero frame preceding the window (i.e., before the Ala codon) specifies Lys. Hence trypsin cleavage after this lysine will yield normal coat protein peptide 9 and, if the frameshift event occurs at the adjacent Ala codon, the next peptide (20, Table I) will be encoded entirely in the minus one frame. Peptide 9, eluting at 49 min in the chromatogram presented in Fig. 6, was identified by sequencing (data not shown). Additional sequencing of protein 6 tryptic peptides with various labeled amino acids identified two other peptides (21 and 23, Table...
Frameshifting

Protein 7

Protein 7 whose synthesis is stimulated by the addition of tRNAser is a mixture of at least three different hybrid sequences. Numerous peptide analyses were used to identify the junction peptides. For instance, l-[3H]arginine, l-[3H]serine, and l-[14C]alanine were used, and the tryptic peptides of purified protein 7 (as in Fig. 5B) were fractionated (Fig. 9). Three pools were made as indicated in Fig. 9A, and each was re fractionated as shown in Fig. 9b-g. The purified peptides, 14 from 9C, 15 from 9E, and 16 from 9F, were then sequenced, and the results are shown in Fig. 7. Peptide 14 has serine at positions 7 and 8 and arginine at position 10. The serine at position 7 is expected from zero frame translation (cf. coat peptide 11, Table I), but the serine at 8 replaces the succeeding alanine. This serine and the arginine at position 10 are predicted from translation in the minus one frame with the switch occurring at position 8. Peptide 15 has alanine at position 8 and serine at position 7, 10, and 11. The serine at 7 and the alanine at 8 are encoded in the zero frame (cf. coat peptide 11, Table I), while serine at 10 and 11 reflects a switch to the minus frame between 9 and 11. Sequencing of this peptide with the isoleucine residues labeled with tritium shows this amino acid at positions 8 and 9, both predicted from zero frame translation (data not presented). Hence the switch must occur at the alanine codon at position 10. Peptide 16 has alanine at position 8 and 10 and serine at 7 and 11. Here the alanine at position 10 defines the last residue from the zero frame, and the serine at position 11 is the first in the minus one frame. The results of this and other sequencing experiments (not shown) are as displayed in Table I and corroborate the three distinct reading frame junction peptides corresponding to frameshifting at each of the three Ala codons in the protein.

 Panel B, peptide 20, indicated by a bar, was pooled and further fractionated, panel B, using 10 mM ammonium acetate (as in Fig. 4C). The material from the resulting peak was again fractionated using heptfluorobutyric acid (Fig. 4D), and the peak material, panel C, was sequenced.

l-[3H]Serine-labeled protein 6 obtained from the purification in Fig. 5C was treated with cyanogen bromide, and the fragments were fractionated and sequenced. The results (Fig. 8) show serine at positions 11, 19, 23, and 31 as expected. The serine at position 11 is predicted from the zero frame while those at 12, 23, and 31 must come from the minus one frame. The spacing between the serine at 11 and those in the minus one frame means that the minus one frameshift must have occurred by a two-base translocation. Taken together with the data above on the tryptic peptides, this shows that the shift must occur at the first alanine codon (position 19 in this cyanogen bromide fragment where serine is inserted) in the frameshift window. The unexpected serine at position 13 is clearly in lower yield than that at 11 and 19. We think this is due either to contamination with another peptide or to mis-coding of the UGC codon at this position due presumably to triplet miscoding to explain any of our other data.

FIG. 6. Separation on HPLC of tryptic peptides from protein 6-Ser. Protein 6-Ser labeled with L-[14C]alanine, l-[3H]serine, and l-[3H]arginine was purified by SDS-PAGE as in Fig. 5 and was digested with trypsin. Panel A shows the resultant peptides fractionated as in Fig. 3. The absorbance profile is of tryptic coat peptides from the coat protein added as carrier. The fractions containing peptide 20, indicated by a bar, were pooled and further fractionated, panel B, using 10 mM ammonium acetate (as in Fig. 4C). The material from the resulting peak was again fractionated using heptfluorobutyric acid (Fig. 4D), and the peak material, panel C, was sequenced.

FIG. 7. Sequence of peptides from proteins 6-Ser and 7-Ser. The purified peptide 20 from protein 6-Ser in Fig. 6C and peptides 14, 15, and 16 from proteins 7-Ser a, b, and c, respectively, in Fig. 9, C, E, and G were subjected to Edman degradation.
FIG. 8. Sequence of a cyanogen bromide fragment of protein 6-Ser. A fragment of protein 6-Ser labeled with $[^3]$H]serine and containing the reading frame transition was purified by HPLC and sequenced.

![Sequence diagram]

FIG. 9. Separation on HPLC of tryptic peptides from protein 7-Ser. Protein 7-Ser labeled with $[^3]$H]alanine, $[^3]$H]serine, and $[^3]$H]arginine was purified by SDS-PAGE as in Fig. 5 and was digested with trypsin. Panel A shows the resultant peptides fractionated as in Fig. 3. The fractions containing peptides 14, 15, and 16, as indicated, were pooled individually and further fractionated as described in Fig. 6. The peak materials in panels C, E, and G were used for sequencing.

frameshift window (Fig. 2). The protein resulting from the first of these Ala codons is designated 7-Ser a and from the other two, 7-Ser b and 7-Ser c, in order.

Addition of tRNAThr also stimulates synthesis of protein 7 (designated 7-Thr) but not protein 6. Pro codon(s) were implicated by competition experiments as the site of tRNAThr-stimulated frameshifting (13). There are two Pro codons, CCG and CCC, within the last minus one frame coat protein gene window (the same window considered above). Protein 7-Thr was isolated from a reaction mixture containing $[^3]$H]proline, $[^3]$H]threonine, and $[^3]$H]asparagine as labels (purified as in Fig. 5D) and analyzed (Fig. 10) as described above. The rechromatographed material taken from the designated peak in Fig. 10C was sequenced (Fig. 11) identifying asparagine at positions 3 and 9 and threonine at position 4 which corresponds to peptide 13 in Table I. The asparagine at position 3 is encoded in the zero frame (cf. coat peptide 11, Table I) while the threonine at 4 and the asparagine at 9 are from the minus one frame, so the switch must occur at position 4 with threonine inserted at the CCG Pro codon. If frameshifting also occurs at the remaining Pro codon CCC (position 6 in peptide 11, Table I) then it is less than 5% of that occurring at the CCG, as the counts for the only peptide possibly having the expected sequence were barely above background (data not shown).

The synthesis of both proteins 6 and 7 requires a minus one frameshift event within the same window near the end of the coat protein gene. Both are expected to terminate at the same position and to have the same number of amino acids. The proteins only differ in the identity of a 15 amino-acid sequence early in the latter third of the protein. However, protein 6 migrates slower on SDS-PAGE. Although the mass of protein 6 is only about 200 daltons greater than protein 7, its electrophoretic mobility implies a difference of 1500 dal-

![Graphs and charts]

FIG. 10. Separation on HPLC of tryptic peptides from protein 7-Thr. Protein 7-Thr labeled with $[^3]$H]asparagine, $[^3]$H]threonine and $[^3]$H]proline was purified by SDS-PAGE as in Fig. 5 and digested with trypsin. Panel A shows the resultant peptides fractionated as in Fig. 3. The fractions containing peptide 13, as indicated, were pooled and further fractionated as described in Fig. 6. The peak material in panel C was used for sequencing.
it is necessary to locate and sequence the reading frame transition peptide.

Tryptic fragments of synthetase were unexpectedly recovered from the protein 9 profile discussed above. Among them a reading frame hybrid peptide was characterized (peptide α, Fig. 12, A–C) by finding [3H]thr at positions 6 and 10, [3H]asn at 11 (Fig. 11), and [3H]asp at position 5 (data not shown). This is consistent with a minus one shift occurring within a frameshift window early in the synthetase gene, producing a truncated product (9-Thra) which migrates on SDS-PAGE with protein 9-Thr from the coat protein gene window. The site for the shift, a CCG Pro codon, is shown at the far right of the diagram in Fig. 2.

FIG. 11. Sequence of peptides from proteins 7-Thr and 9-Thra. The purified peptide 13 from protein 7-Thr (Fig. 10) and a tryptic peptide from protein 9-Thra (a frameshift product from the synthetase gene) as shown in Fig. 12 were subjected to Edman degradation.

Peptide 13 Peptide α

FIG. 12. Separation on HPLC of tryptic peptides from protein 9-Thr. Protein 9-Thr labeled with L-[3H]asparagine, L-[3H]threonine, and L-[3H]proline was purified by SDS-PAGE as in Fig. 5 and digested with trypsin. Panel A shows the resultant peptides fractionated as in Fig. 3. The fractions containing the peptide indicated were pooled and further fractionated as described in Fig. 6. The peak material in panel C was used for sequencing, and the results are shown in Fig. 11.

Protein 9

Previous work led to the conclusion that tRNASer3-stimulated protein 9 resulted from frameshifting within the second to the last minus one frameshift window of the coat protein gene. tRNAThr3 also stimulates minus one frameshifts within this window. Two Pro codons are within the window, an upstream CCU and a downstream CCA. We have not located the reading frame transition peptides due to shifting at either. Peptide 8, however, which includes the proline encoded by CCU, is present, ruling out substantial frameshifting at this site. Peptide 9, which includes the other proline, is recovered in quite variable yield even from coat protein so its absence is not necessarily meaningful. The terminal peptide 11 is clearly deficient as expected. To show that tRNAThr3-stimulated protein 9 is not just a truncated product from the normal reading frame (which is a general problem when considering products shorter than their normal counterparts)
Frameshifting at Alanine and Proline Codons

Can frameshifting occur at all alanine and proline codons? The four alanine codons in the last minus one frameshift window of the MS2 coat protein gene are all GCA. In the case of proline, frameshifting within the same window was detected at a CCG codon but not a CCC codon. Frameshifting at other alanine and proline codons can be investigated with additional genes. It is convenient to look at other RNA phages since a representative of each of the four groups of E. coli RNA bacteriophages (24–26) has been sequenced. We have analyzed by gel electrophoresis the in vitro translation products of each synthesized in the presence of an excess of various tRNAs. This provides additional, although less direct, evidence for the frameshifting potential of other alanine and proline codons.

Alanine

GCC—The sequence of GA RNA (group II) (26) shows that the synthetase gene has a single GCC in its last minus one frame window, and its extension product should be 12 amino acids longer than synthetase. tRNASer, addition shows a very small but significant amount of a product of the predicted size (Fig. 13, lanes p–s).

GCG—The coat protein gene of Qβ (group III) has a single alanine codon GCC in its last minus one window (27), and frameshifting at this codon should give a protein 23 amino acids longer than the coat protein. We have not seen convincing frameshifting at this site (compare Fig. 13, lane h with lane i). The MS2 synthetase has both GCC and GCG alanine codons in its last minus one frame window and consistent with the previous cases tRNASer has a weak positive effect in stimulating apparent minus one frameshifting (see protein 66K in Fig. 3 in Ref. 13).

GCU—The GA coat protein gene has two alanine codons, a GCU and a GCG, in the last minus one window and tRNASer addition has a weak effect in stimulating synthesis of a product of the expected size, about 7000 daltons larger than the coat protein (Fig. 13, lane m). Based on the result seen with the GCG codon in the Qβ coat protein, it may be that it is the GCU of GA that is promoting this frameshifting.

GCA—Amino-acid sequencing data discussed above shows that tRNAser causes frameshifting at GCA.

In summary, these results show that tRNAser can promote frameshifting at GCA, GCU, and GCC alanine codons, but GCA is qualitatively more active as seen with MS2 proteins 6 and 7. This is apparent even when taking into account the three GCA codons at which frameshifting can occur to give protein 7.

Proline

CCG—We concluded above the tRNAThr causes minus one frameshifting at CCG proline codons. In at least one case (in the last minus one frameshift window in the MS2 coat protein gene) this effect is precluded by addition of tRNAPro (Fig. 13, lane f) which is thought to decode CCG, CCA, and CCU (28). (Note that tRNAPro on its own stimulates synthesis of 5 and 7 to a small extent, Ref. 13).

CCA—Whether tRNAThr causes frameshifting at CCA codons was uncertain. However, tRNAPro does stimulate synthesis of a predicted protein product from a minus one frameshift window in the coat protein gene of Qβ where CCA is the only proline codon (Fig. 13, lane j) and by tRNAPro, but to a lesser extent (Fig. 13, lane k).

CCU—We concluded above that CCU did not give frameshifting in the case of MS2 protein 9. The phage GA synthetase gene provided another opportunity to test this, and no frameshifting was detected (Fig. 13, lane x).

CCC—The synthetase gene of Qβ has a CCC in the last minus one frameshift window but little, if any, of the predicted product was seen (Fig. 13, lane z).

Phage GA has no proline codon in the last minus one frameshift window of its coat protein gene and reassuringly

![Frameshifting assay of group I, II, and III phage messenger RNAs.](https://example.com/fig13)

Fig. 13. Frameshift assay of group I, II, and III phage messenger RNAs. The products of in vitro protein syntheses incorporating L-[35S]methionine were separated by SDS-PAGE. The translation mixtures were programmed with the mRNAs indicated below the autoradiograms, and the tRNAs tested for frameshifting activity are identified below each lane. Lanes a–o, showing separations on 17.5% gels, utilized twice the exposure period of the 10% gels of lanes p–z. The coat protein of bacteriophage GA was identified by migration with protein from purified virions, but the prominent protein of slightly greater mobility is unidentified.
tRNAThr₃ does not cause minus one frameshifting in this 11 codon window (Fig. 13, lane o).

In summary, tRNAThr₃ stimulates frameshifting at CCG, probably at CCA, marginally if at all at CCC, but not at CCU.

**Isoaccepting tRNAs**

In the studies described above, tRNAThr₃ promoted frameshifting at proline codons. It is surprising that the isoacceptor tRNAThr, may be promoting some other frameshift event(s) since we see new proteins with MS2 (Fig. 13, lanes b and c) and with each of the other RNA phage messages (not shown). In addition tRNAPro precludes the effect of tRNAThr₃ but not of tRNAThr₁ (Fig. 13, lanes d and f), so there is a clear difference in activity of the two threonine tRNAs as revealed by this frameshift assay. The sequences of the two tRNAs differ by only eight nucleotides and only one of these is in the anticodon loop and stem (29). This is an A to G change at position 38, two removed from the anticodon. While position 37 is A in both cases, it is modified to t₆A in tRNAThr₁ and to mt₆A in tRNAThr₃. We think that the changes outside the anticodon loop and the identity of the modified base 37 are not likely reasons for the difference in behavior of tRNAThr₁ and tRNAThr₃. Rather the difference could be due to the change at position 38. tRNAThr₃ is virtually the only tRNA to have a G at position 38, and this position is likely to be important for decoding (31).

We have also detected a case of a tRNA causing presumed frameshifting at the codon of an isoacceptor: tRNAAla₁ (17) stimulates synthesis of an altered form of the Ga synthetase with slightly greater electrophoretic mobility (Fig. 13, lanes u and v). tRNAAla₂ (32) is a competitor (Fig. 13, lane v). We think that tRNAAla₁ with the anticodon 3'CGV(= uridine-5'-oxyacetic acid)⁵', which is thought to decode GCA, GCG, and GCU, is reading an tRNAAla₁ codon of GCU or GCC to give a minus one frameshift. We favor its reading GCC based on the windows defined by the GA sequence.

**Plus One Frameshifting by tRNASer₃ or tRNAThr₃**

The only case of tRNA enhanced frameshifting into the plus one frame that we have detected is tRNAGln and tRNAPro, stimulated synthesis of protein 5, the coat-lysis hybrid (see below). Neither tRNASer₃ nor tRNAThr₃ cause increased synthesis of protein 5 (13). None of the other frameshifting windows tested here gave evidence for any plus one frameshifting by these tRNAs.

**Discussion**

The frameshift window that we have examined in detail (the last minus one window in MS2 coat protein gene) has 24 codons, yet only two tRNAs out of a large number tested (36 of 50) convincingly promote frameshifting. Whatever the mechanism there must be something special about the codon-anticodon interaction of these tRNAs. We initially expected that this frameshifting would be due to out of frame triplet decoding but in each case examined in detail this has not been the explanation.

Amino-acid sequencing shows that the tRNAThr₃, with a 3' UGG 5' anticodon, must frameshift at CCG and CCA proline codons. And tRNASer₃, with a 3' UGG 5' anticodon, frameshifts at GCA, GCU, and GCC alanine codons. Both events involve a two-base translocation. This could be understood if the 5' two bases, 34 and 35, of the anticodon base pair with the 5' two bases of the codon (Fig. 14). Only in some cases (see Fig. 2) could there be pairing between the 3' base of the anticodon, 36, and the 3' base of the preceding codon. Therefore, triplet decoding with overlapping reading of one base is not required. A similar conclusion has been reached by experiments that directly alter the sequence of tRNAs.

Also, the possibility has been considered that the tRNA base 5' to the anticodon, U33, transiently pairs with the 3' base of the proline or alanine codon (33), and the data presented here are consistent with this model. The tRNA construction experiments show that this interaction is not necessary for this frameshift event. Thus, it seems that only bases 34 and 35 of the tRNA could be involved in conventional base pairing.

Presumably, the next available codon then begins with the third position base of the zero frame codon, resulting in a minus one shift in reading frame. In at least one case, however, the identity of the third base is important since CCA promotes shifting while CUC does not (note that the two codons are tested in different contexts and we have not made a direct comparison). We cannot simply explain the influence of the third position base in the zero frame on the decoding by the preceding tRNA. If a direct tRNA-tRNA interaction is involved, the hypermodification of base 37 might sterically interfere with the preceding tRNA. Codons beginning with A are decoded by tRNAs with t₆A or mt₆A at position 37 (34, 35). Indeed A is often the first nucleotide in the codon that must be recognized by the incoming tRNA. In two cases, however, it is G, and tRNAs which decode these latter codons are not hypermodified at position 37 (34, 35); therefore, this explanation is deemed unlikely. This problem of the role of downstream bases is not a new dilemma since effects of 3' mRNA sequence on nonsense and mis-sense suppression have been described and are not understood.

There are probably a number of different ways in which frameshifting can occur, and the approach used here would detect only a subset of these. For instance, we would not have detected frameshifting at, for example, AAA AAA by tRNALys reading out of frame; in this case the competitor would be the same tRNA and increasing its concentration would presumably have no effect.

We will now consider what role these frameshift events might play in the phage life cycle. Protein 5, a product of a plus one frameshift fusing the coat and lysis genes (13, 36, 37), is the only frameshifted protein found so far both in *vitro* and in *vivo*. We would be surprised if the minus one frameshift events detected *in vitro* did not also occur *in vivo*. The frameshifted proteins may not themselves have a specific...

---

**Fig. 14.** A model for frameshifting. This shows the proposed two-base pair interaction between the anticodon loops of tRNASer₃ and tRNAThr₃ and the noncognate codons that will give the observed minus one shift in reading frame.

---

² A. G. Bruce, J. F. Atkins, and R. F. Gesteland, manuscript in preparation.
function, rather it may be the progression of ribosomes in the new reading frame that serves some purpose. For example, we will consider the minus one frameshift event near the end of the synthetase gene of MS2 that results in synthesis of the 66K form of the protein rather than the normal 62K form. Perhaps the 66K protein has no specific function, but its synthesis permits a small proportion of ribosomes to progress nearer to the 3' end of the RNA where initiation of minus strand RNA occurs, and this presence influences or regulates RNA replication. Since only a small proportion of ribosomes enter the minus one frame and ultimately reach the 66K termination codon, this may permit coordination between synthetase translation and hence replicase formation and RNA replication. It is striking that the 66K terminator is coincident with a sequence which on the minus strand is CUCUUU. This sequence is identical to six contiguous bases of the boxA consensus sequence (38, 39). The NusA protein is known to bind to the boxA sequence in RNA (40) and thereby modulate the progression of RNA polymerase in transcription complexes (41-43). We suggest that in MS2 the NusA protein causes replicase pausing at the 66K terminator and that the approaching ribosomes interfere with the NusA-boxA interaction and thus promote progression of replication. The only support for a role of NusA in RNA phage replication is the presence of boxA sequences between codons 27 and 31 of the maturation gene of phages MS2, Qb, and GA (but not SP), approximately where evidence for a replication pause site has been found (review, 44). Interestingly, there is a boxA sequence just downstream from the synthetase initiator where the other two reading frames encode the overlapping lysis gene and the extended coat frameshift products described here. Robertson (44) has speculated that there is also a replication pause in this region, and again one could imagine that the synthesis of frameshift products is involved in the relief of pausing.

Acknowledgments—We thank W. Gray and R. Zeikus for expert help with protein sequencing, A. Hirashima for provision of GA and SP phages and their unpublished RNA sequences, N. Franklin, J. Murphy, and J. Gallant for discussions, B. Falahee for comments on the manuscript, and R. Weiss for discussion on models and mechanisms.

REFERENCES