

## *Escherichia coli* translational initiation factor IF3: A unique case of translational regulation

LARRY GOLD, GARY STORMO, AND ROGER SAUNDERS

Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO 80309

Communicated by William B. Wood, August 1, 1984

**ABSTRACT** The *Escherichia coli* translational initiation factor IF3 is encoded by an mRNA that has an unusual ribosome binding site. We have explored a mechanism that may account for the translation of IF3 and that provides regulation of the quantity of IF3 relative to ribosomes.

*Escherichia coli* has many small regulatory loops that tightly control the levels of various important proteins. Such loops often involve regulation of gene expression at the post-transcriptional level (1, 2). In this paper, we propose a regulatory circuit that maintains appropriate levels of the *E. coli* translational initiation factor IF3 (3). The proposal simultaneously predicts a precise orientation of four different domains of 16S RNA within the 30S ribosomal particle.

### RESULTS AND DISCUSSION

**The Problem.** IF3 is responsible (at least) for dissociating 70S ribosomes that are not making protein, so as to allow the 30S subunit to reengage an mRNA and initiate translation (3). Purified IF3 is required for 70S ribosome dissociation *in vitro* and for high-level translation of natural mRNAs (3). When a 30S particle, bound to IF3, initiates translation, the normal mechanism is as shown in Fig. 1. Although the nucleotides surrounding initiating codons in *E. coli* are not random, we do not understand the contribution of that information to the 1000-fold differences in rates of translation between different mRNAs (1, 5, 6). Specifically, we do not know whether any 16S RNA domains (besides the 3' end) are used to achieve mRNA selection. Nevertheless, we have always believed that 16S RNA contains other domains that potentiate mRNA binding, at least for some mRNAs.

The sequence of the mRNA encoding the initiation factor IF3 was recently published (7); the translational initiation domain breaks many "rules" followed by most other *E. coli* mRNAs (1, 5, 6). The sequence of the IF3 mRNA around the initiation codon is (7)

\* \* \* \* \* † † † † † § ‡ ‡ ‡ ‡ ‡  
G-G-A-G-G-A-A-U-A-A-G-G-U-AUU-A-A-A-G-G-C-G-G

The Shine-Dalgarno region (\*) is the only segment of this mRNA that is conventional (8); the initiation codon is AUU (§), which is unique among *E. coli* mRNAs. A second potential Shine-Dalgarno sequence (†), with improper spacing for the initiation codon, sits between the probable Shine-Dalgarno sequence and the AUU; extra potential Shine-Dalgarno regions within ribosome binding sites are very rare (1, 5, 6). Lastly, a "structurogenic" sequence (‡) sits just 3' to the AUU; such guanosine-rich regions are very rare in *E. coli* ribosome binding sites (exclusive of the Shine-Dalgarno region). Thus, this mRNA breaks several rules that were deduced through a comparison of more than 150 other *E. coli* ribosome binding sites. The IF3 message is translated at a

high rate *in vivo* (9). We have attempted to account for IF3 expression by seeking to understand its translation.

IF3 expression is known to be regulated. IF3 levels increase, along with the other components of the translational machinery, when growth rates increase (9). That is, *E. coli* must respond to transiently decreased levels of IF3 by increasing the relative rate of IF3 expression, which is itself needed for the translation of all other mRNAs in the cell. Similarly, excess IF3 should selectively diminish the relative rate of IF3 production. However, IF3 is not feedback regulated by the same mechanism as are ribosomal proteins (2). Cells carrying plasmids encoding IF3 yield more IF3 protein in response to increased plasmid copy number (10).

A constant IF3/ribosome ratio could be maintained if IF3 were translated in a mode independent of IF3 function; cells could adjust the relative amount of IF3 made as a result of competition between the IF3 mRNA and all others in the cell. High IF3 levels would stimulate all other mRNAs, leaving few ribosomes for IF3 translation; low IF3 levels would make available many ribosomes for IF3-independent translation of IF3 itself. We propose that the odd sequences around the IF3 initiation codon (and the odd initiation codon itself) provide molecular components of the regulatory mechanism; the specific model involves interactions with certain sequences of 16S RNA.

**Juxtaposition of Certain 16S RNA Domains.** The secondary structure of 16S RNA (Fig. 2) is now well established (11). The figure highlights the regions that, we will argue, lie close to the groove that captures mRNA during translational initiation (as in Fig. 1).

We first note a crosslinking experiment between acetylvalyl-tRNA and 16S RNA. The acetylvalyl-tRNA was bound into the P site, after which photoactivation gave a specific crosslink to cytidine-1400 (12). The location of cytidine-1400 (Fig. 3), relative to the P site, has rotational ambiguity, dependent on the precise details of the photoactivation process [the tRNA has a modified base at the 5' position of the anticodon, and the crosslink is between that modified base and cytidine-1400 (13)]. We have selected among the possible orientations of cytidine-1400, relative to the tRNA in the P site, according to our idea about IF3 translation. In Figs. 4 and 5, orientations and polarities are based on the same idea.

Hearst and his colleagues have studied psoralen intramolecular crosslinking of 16S RNA. Most psoralen crosslinks confirm the 16S RNA secondary structure predicted by the Woese-Noller phylogenetic arguments (11). In addition, a crosslink between uridines-1506 and -956 has been found (14). Thus the cytidine-980 domain of 16S RNA may be brought into the neighborhood of the mRNA groove (Fig. 4). The orientation of the cytidine-980 domain has rotational ambiguity; the orientation shown is consistent with the proposed secondary structure of 16S RNA.

Another crosslink involves the 3' uridine of a tetranucleotide messenger (U-U-U-U) and both guanosine-462 and -474 of 16S RNA (16). The tetranucleotide mRNA was positioned in the decoding site on the 30S particle (Fig. 5). Guanosine-

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

Abbreviation: IF3, initiation factor 3.

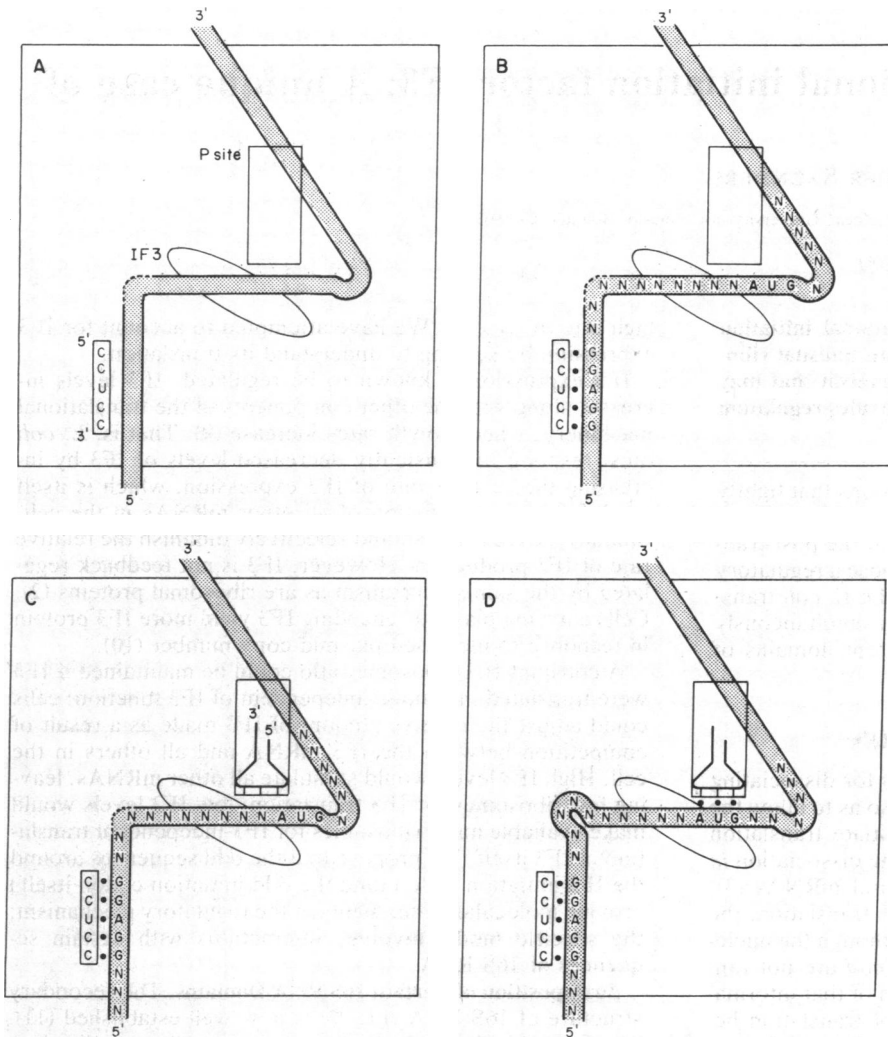


FIG. 1. Normal translational initiation. A conventional representation of a 30S particle (with the 3' end of 16S RNA and the ribosomal P site) is shown in A. A shaded track is shown to accommodate mRNA. The IF3 protein is depicted as a "football" in the middle of the 30S particle. Initiation can begin by mRNA binding into the groove, using the Shine-Dalgarno interaction (B). As shown by Jay *et al.* (4), bound mRNA can facilitate initiator tRNA binding (C), which can lead to codon-anticodon base pairing (D). The ribosomal A site is not shown in this figure but is used in Fig. 5. The A site could be shown like the P site, situated just to the right of the P site.

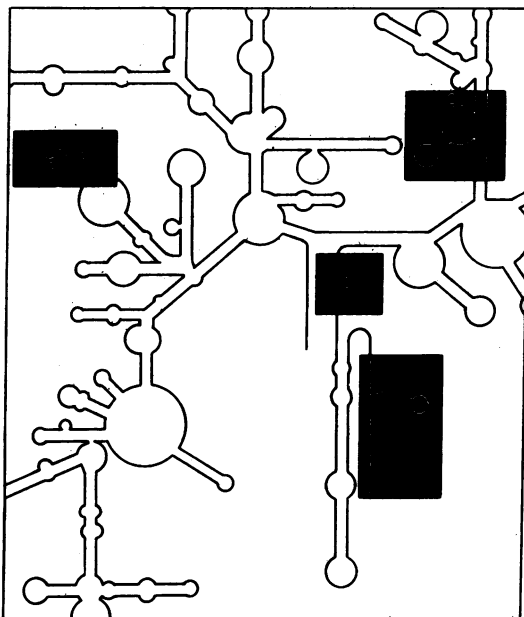


FIG. 2. Secondary structure of *E. coli* 16S RNA. The proposed secondary structure of 16S RNA is given, according to Woese *et al.* (11). Black boxes are placed around regions of the molecule that lie close to the mRNA binding domain and that, we argue, are involved in IF3 expression.

474 is not more than 15 Å from guanosine-462, and those guanosines are nearly on the same side of the helix.

In summary, four 16S RNA domains (the 3' end, cytosine-1400, cytosine-980, and guanosine-462/guanosine-474) are near each other on the surface of a 30S particle (Fig. 6). This surface is known to participate in mRNA binding, 50S particle binding (18, 19), and binding of the initiation factor IF3 (3).

**The Model.** We now examine the proposed IF3-independent initiation of the IF3 mRNA (Fig. 7). We first note that the 16S RNA domains, as we have positioned them (Fig. 6), have an interesting relationship to the mRNA binding groove of the 30S particle (compare Fig. 7A with Fig. 1A). The binding of IF3 mRNA (Fig. 7B) occurs with and depends on the unfolding of the guanosine-462/guanosine-474 helix; the bound message is base paired to 16S RNA at 20 of 24 nucleotides. By virtue of the Shine-Dalgarno interaction, as in the normal initiation transitional event (Fig. 1B to C), initiator tRNA joins the complex (Fig. 7C). Base pairing with the initiator tRNA (as in Fig. 1D) is facilitated by the cytosine-980 region and the renaturation of the guanosine-462/guanosine-474 helix (Fig. 7D). The proposed interactions rationalize every odd feature of IF3 mRNA, including the AUU initiation codon. In fact, the AUU of the mRNA participates in a perfect 10-base duplex. We imagine that IF3 binding to the 30S particle (as in Fig. 1) interferes with the proposed IF3-independent mode; Fig. 1 is drawn so as to show IF3 occluding some of the 16S RNA nucleotides used in the IF3-independent mode.

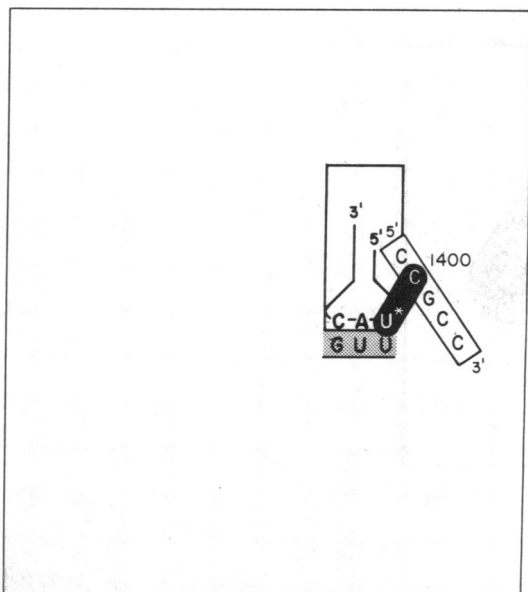


FIG. 3. The cytosine-1400 crosslink. Acetylvalyl-tRNA [anticodon = cmo<sup>3</sup>U(34)-A-C] is shown in the ribosomal P site, base paired with an mRNA (GUU); the shading indicates the mRNA track (as in Fig. 1). Photoactivation of the 70S ribosome-tRNA-mRNA complex leads to crosslinking (shown in black) between the modified base of the anticodon and cytosine-1400 of 16S RNA (12, 13).

**Buffering of IF3 Translation.** How do the proposed interactions meet the objective of buffered IF3 expression? Translational termination (of all messages) is followed by the release of free 30S and 50S particles (22). The free 30S particle (Fig. 8) can react with either the IF3 protein (A), a 50S particle (B), or the IF3 mRNA (C). When the level of IF3 protein is low relative to the number of 30S particles, IF3-independent translation of the IF3 mRNA brings about an

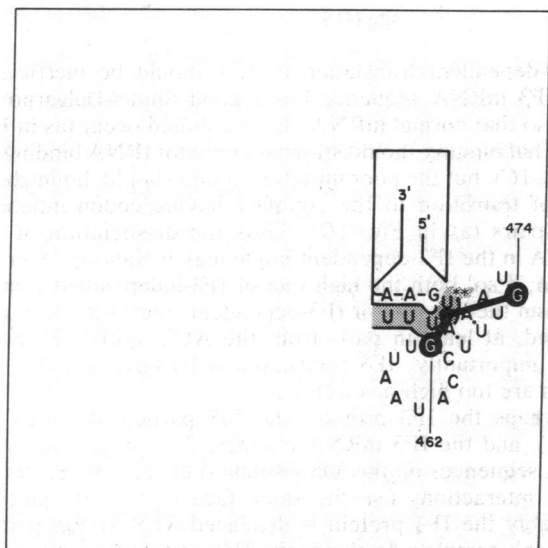


FIG. 5. The guanosine-462/guanosine-474 crosslink. A tetranucleotide mRNA (within the mRNA track) analog (U-U-U-U\*\*) base paired with tRNA<sup>Phe</sup> in the ribosomal A site is shown. The modified message contains a reactive (guanine-specific) adduct on the 2' position of the 3'-most ribose. After incubation with 30S particles (containing <sup>32</sup>P-labeled 16S RNA) the mRNA analogue was found to be crosslinked to guanosines-462 and -474 (Fig. 2). The crosslinking to both guanosine-462 and guanosine-474 (shown in black) is consistent with the secondary structure of 16S RNA (11). The data for this experiment are from Wagner *et al.* (16) and Wagner and Gassen (17).

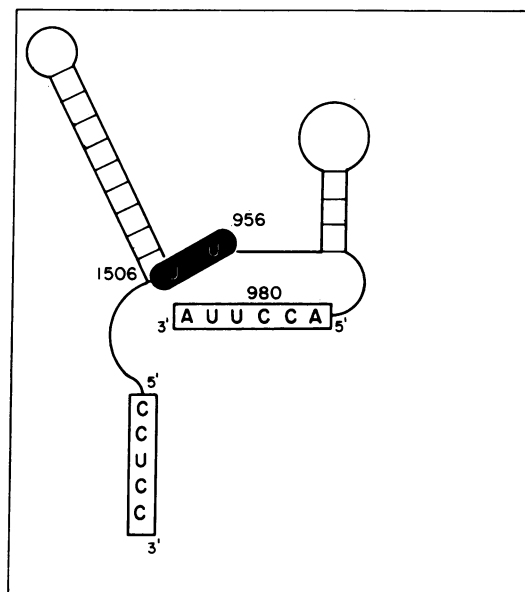


FIG. 4. The uridine-956/uridine-1506 crosslink. A psoralen crosslink identified by Thompson and Hearst (14) is shown in black. Deproteinized 16S RNA was crosslinked. The same crosslink has been identified by electron microscopy (15). Thompson and Hearst proposed a precise interaction (involving base pairing between the two domains) that accounts for the capacity of psoralen to crosslink the two uridines (see their Fig. 7). Our representation of the two domains maintains the secondary structure of 16S RNA proposed by Woese *et al.* (11), since we are focused on the 16S RNA nucleotides 978-983. Our representation is consistent with the Thompson and Hearst structure.

increase in IF3 protein level; this is the means by which IF3 protein levels respond to increased growth rate (9). When the growth rate is decreased, IF3 expression should decrease and excess IF3 would be diluted. While IF3 is in excess, few 30S particles will be able to initiate translation in the IF3-independent mode. Alternatively, IF3 mRNA might directly bind to 70S particles in the absence of IF3 protein. The interactions of Fig. 7 would still be used. Some data (25, 26) support this path to IF3-independent initiation of the IF3 mRNA.

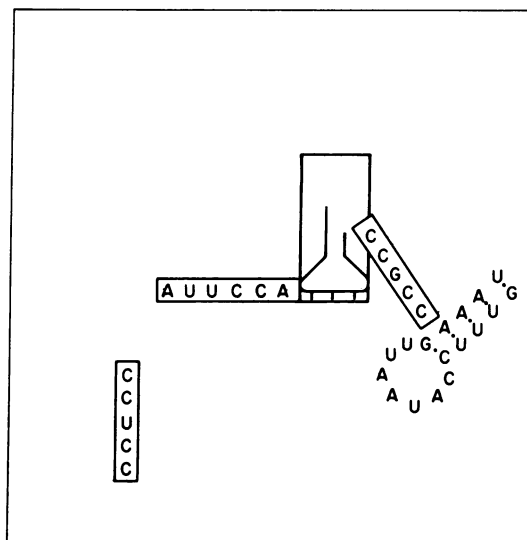


FIG. 6. Composite of the 16S RNA domains and the P site. The regions identified by crosslinking are superposed onto one diagram of a 30S particle.

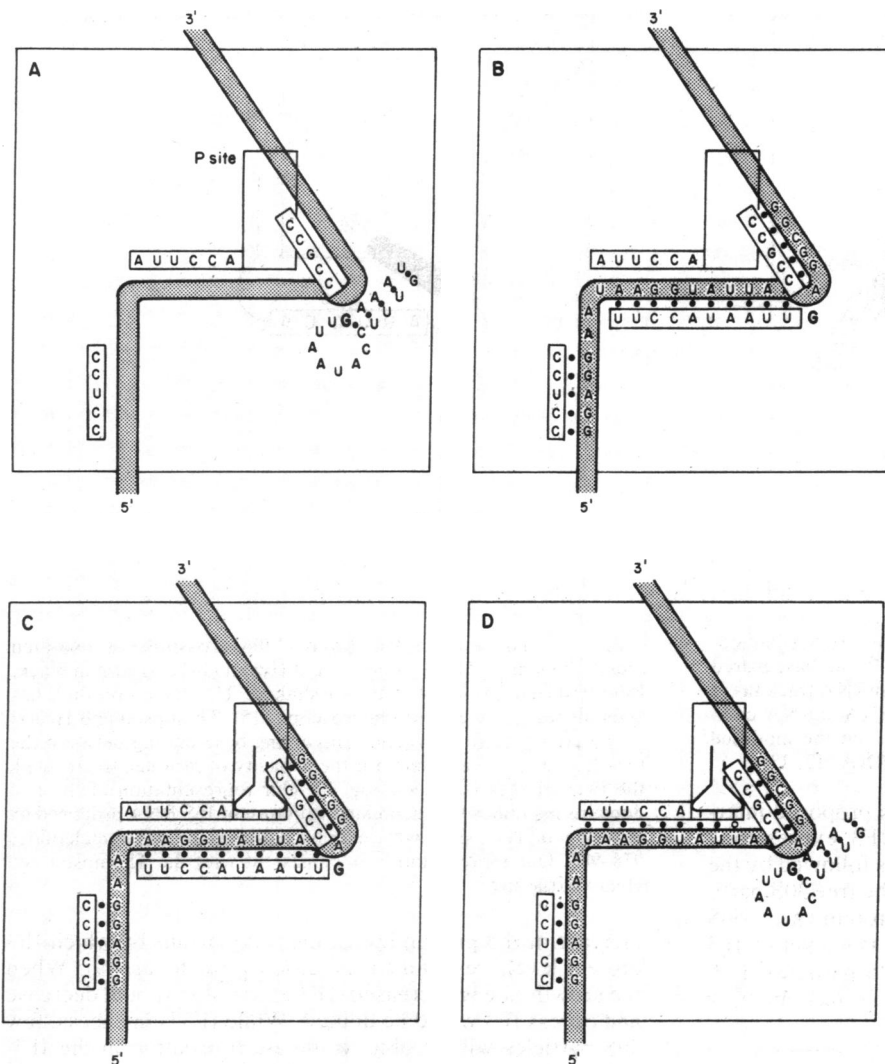


FIG. 7. IF3-independent initiation on IF3 mRNA. Note the relationship between the relevant 16S RNA domains and the potential mRNA track, which is shaded (A). Bound IF3 mRNA is depicted in B. Initiator tRNA is next bound to the complex (C). Codon-anticodon base pairs are then formed, along with mRNA-cytosine-980 base pairs, as the guanine-462/guanine-474 helix reforms (D). [The initiator tRNA of *E. coli* can recognize AUG, GUG, AUA, UUG, and AUU (1); one experiment suggests that even UUU can be recognized (20, 21). In an extreme case, *in vitro*, an mRNA with a good Shine-Dalgarno domain can lead to initiator tRNA binding even when no obvious initiator lies in the P site (4).]

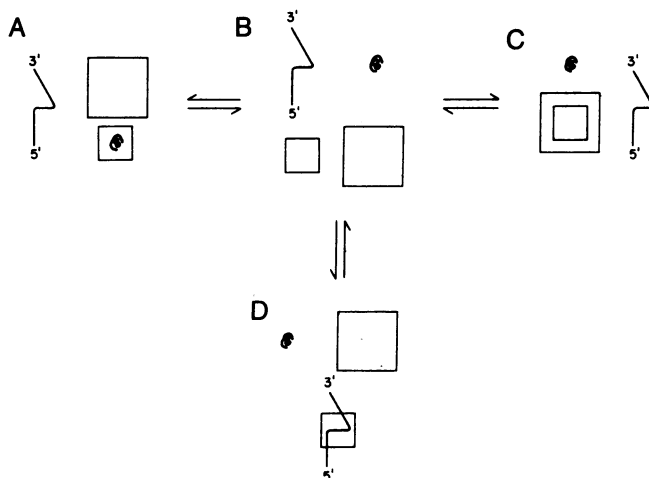


FIG. 8. Linked equilibria between IF3 mRNA, IF3 protein, 30S particles, and 50S particles. (B) The free components (IF3 protein, IF3 mRNA, the 30S particle, and the 50S particle). (C) The 30S and 50S particles can interact to form a 70S particle. (A) The 30S particle and IF3 protein can interact to prepare the 30S particle for normal translational initiation. (D) The 30S particle can interact directly with IF3 mRNA to form a pre-initiation complex. This complex corresponds to Fig. 7B. The 30S particle with IF3 protein (A) participates in translational initiation on most *E. coli* messages, through the Shine-Dalgarno interaction. In this mode, IF3 mRNA will be discriminated against because of the initiation codon AUU.

IF3-dependent translation of IF3 should be inefficient. The IF3 mRNA sequence has a good Shine-Dalgarno region, so that normal mRNA binding should occur (as in Fig. 1B). That binding should stimulate initiator tRNA binding (as in Fig. 1C), but the poor initiation codon should diminish the rate of transition to the complex having codon-anticodon base pairs (as in Fig. 1D). Thus the dissociation of IF3 mRNA in the IF3-dependent mode may be favored over initiation; if so, both the high rate of IF3-independent translation and the low rate of IF3-dependent translation would be derived, at least in part, from the AUU initiation codon. Most importantly, IF3 translation is buffered whether IF3 levels are too high or too low.

Perhaps the IF3 protein, the 50S particle (through 23S RNA), and the IF3 mRNA compete for some of the same RNA sequences on the 30S particle (but see refs. 8, 19). All three interactions use the same face of the 30S particle. Probably the IF3 protein is displaced when a 50S particle joins the complex. In this view, IF3 protein bound to a 30S particle could provide mRNA selectivity by preventing RNA-RNA interactions that are usually not productive but can lead (quickly and preferentially) to IF3 expression in the absence of IF3.

**Evolution of the Regulatory Loop.** The present IF3 mRNA sequence is only two transversions away from a sequence whose conventional translation would yield an IF3 protein missing the amino-terminal hexapeptide (Fig. 9). In fact, the IF3 protein in the cell is found as a mixture of long and short forms (23); the short form is fully active and missing exactly

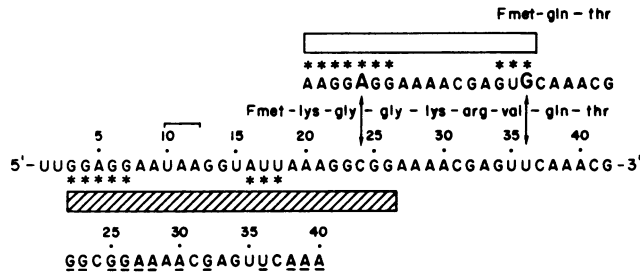


FIG. 9. Evolution of the IF3 mRNA. The IF3 mRNA sequence is shown (numbered to nucleotide 42). The termination codon for the upstream *thrS* gene is at position 10–12. The present Shine–Dalgarno domain is at 3–7, and the initiating AUU is at 16–18. Our discussion thus far has centered on nucleotides 3–26. The upper line of RNA sequence shows the sequence from position 20 to position 42 as it might have once looked. With two transversions, we would have a good Shine–Dalgarno domain (position 20–26) and a good initiation codon (position 34–36), with proper spacing. The hypothetical IF3' protein sequence is also shown beginning fMet-Gln-Thr. For comparison, nucleotides 22–40 of the present sequence are placed under nucleotides 3–21. The sequences share 12 of 19 bases (underlined, lower RNA sequence), which suggests that a small duplication and subsequent mutations led to the present IF3 regulatory sequences. Simultaneously, the IF3' protein was elongated by six amino acids.

six amino acids from the amino terminus (24). The short form begins with Val-Gln-Thr and thus results from proteolysis (10) rather than translational initiation (through the remnant of the old initiation domain). Since a possible primordial ribosome binding site is still visible, and the present amino terminus is not essential, we think the extra codons evolved to provide a buffered regulatory circuit. The nucleotides probably are for mRNA function rather than protein function. Furthermore, the relationship between nucleotides 3–21 and 22–40 suggests that creation of a small duplication and subsequent point mutations may have led to the evolution of the present IF3 mRNA (Fig. 9).

**Conclusion.** We have presented a detailed model for IF3-independent translation of the IF3 mRNA in *E. coli*. The model is based on chemical crosslinking data and the odd nucleotide sequence around the IF3 ribosome binding site. It is so precise that tests will be simple to construct, and it has value both as a predictor of a new kind of regulatory loop and as a predictor of 30S particle topography.

We thank the Alfred Benzon Foundation for arranging a meeting in June 1983 at which these ideas began to take serious form. We appreciate the useful conversations we had with Ed Brody, Gary Craven, Rick Gourse, Marianne Grunberg-Manago, John Hershey, Larry Kahan, Jim Ofengand, Masayasu Nomura, and Mathias Springer. Sid Shinedling helped us state our ideas clearly, and Kathy Piekarski was crucial to the preparation of the manuscript. Jan Lo-

gan developed the figures, which made the presentation possible. This work was supported by National Institutes of Health Grant GM28685.

1. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) *Annu. Rev. Microbiol.* **35**, 365–403.
2. Campbell, K. M., Stormo, G. D. & Gold, L. (1983) in *Gene Function in Prokaryotes*, eds. Beckwith, J., Davies, J. & Galant, J. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), pp. 185–210.
3. Hershey, J. W. B. (1980) *Cell. Biol.* **4**, 1–68.
4. Jay, E., Seth, A. K. & Jay, G. (1980) *J. Biol. Chem.* **255**, 3809–3812.
5. Stormo, G. D., Schneider, T. & Gold, L. (1982) *Nucleic Acids Res.* **10**, 2971–2996.
6. Stormo, G. D., Schneider, T., Gold, L. & Ehrenfeucht, A. (1982) *Nucleic Acids Res.* **10**, 2997–3011.
7. Sacerdot, C., Fayat, G., Dessen, P., Springer, M., Plumbridge, J. A., Grunberg-Manago, M. & Blanquet, S. (1982) *EMBO J.* **1**, 311–315.
8. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
9. Howe, J. G. & Hershey, J. W. B. (1983) *J. Biol. Chem.* **258**, 1954–1959.
10. Lestienne, P., Dondon, J., Plumbridge, J. A., Howe, J. G., Mayaux, J.-F., Springer, M., Blanquet, S., Hershey, J. W. B. & Grunberg-Manago, M. (1982) *Eur. J. Biochem.* **123**, 483–488.
11. Woese, C. R., Gutell, R., Gupta, R. & Noller, H. F. (1983) *Microbiol. Rev.* **47**, 621–669.
12. Prince, J. B., Taylor, B. H., Thurlow, D. L., Ofengand, J. & Zimmermann, R. A. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 5450–5454.
13. Ehresmann, C., Ehresmann, B., Millon, R., Ebel, J.-P., Nurse, K. & Ofengand, J. (1984) *Biochemistry* **23**, 429–437.
14. Thompson, J. F. & Hearst, J. E. (1983) *Cell* **32**, 1355–1365.
15. Wollenzein, P. L., Hearst, J. E., Thammana, P. & Cantor, C. R. (1979) *J. Mol. Biol.* **135**, 255–269.
16. Wagner, R., Gassen, H. G., Ehresmann, C. H., Stiegler, P. & Ebel, J. P. (1976) *FEBS Lett.* **67**, 312–315.
17. Wagner, R. & Gassen, H. G. (1975) *Biochem. Biophys. Res. Commun.* **65**, 519–529.
18. Herr, W., Chapman, N. M. & Noller, H. F. (1979) *J. Mol. Biol.* **130**, 433–449.
19. Vassilenko, S. K., Carbon, P., Ebel, J. P. & Ehresmann, C. (1981) *J. Mol. Biol.* **152**, 699–721.
20. van der Laken, K., Bakker-Steeneveld, H., Berkhout, B. & van Knippenberg, P. H. (1980) *Eur. J. Biochem.* **104**, 19–23.
21. van der Laken, K., Bakker-Steeneveld, H. & van Knippenberg, P. (1979) *FEBS Lett.* **100**, 230–234.
22. Martin, J. & Webster, R. E. (1975) *J. Biol. Chem.* **250**, 8132–8139.
23. Suryanarayana, T. & Subramanian, A. R. (1977) *FEBS Lett.* **79**, 264–268.
24. Brauer, D. & Wittmann-Liebold, B. (1977) *FEBS Lett.* **79**, 269–275.
25. Goss, D. J., Parkhurst, L. J. & Wahba, A. J. (1982) *J. Biol. Chem.* **257**, 10119–10127.
26. Chaires, J. B., Hawley, D. A. & Wahba, A. J. (1982) *Nucleic Acids Res.* **10**, 5681–5693.