Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 7668–7672, October 1990
Biochemistry

Secondary structure of the ribosome binding site determines translational efficiency: A quantitative analysis

(translational initiation/RNA-helix stability)

MAARTEN H. DE SMIT AND J. VAN DUIN

Department of Biochemistry, Gorlaeus Laboratories, Leiden University, P.O. Box 9502, 2300 RA Leiden, The Netherlands

Communicated by Peter H. von Hippel, July 6, 1990

ABSTRACT **We have quantitatively analyzed the relationship between translational efficiency and the mRNA secondary structure in the initiation region. The stability of a defined hairpin structure containing a ribosome binding site was varied over 12 kcal/mol (1 cal = 4.184 J) by site-directed mutagenesis and the effects on protein yields were analyzed *in vivo*. The results reveal a strict correlation between translational efficiency and the stability of the helix. An increase in its ΔG^0 of −1.4 kcal/mol (i.e., less than the difference between an A·U and a G·C pair) corresponds to the reduction by a factor of 10 in initiation rate. Accordingly, a single nucleotide substitution led to the decrease by a factor of 500 in expression because it turned a mismatch in the helix into a match. We find no evidence that exposure of only the Shine–Dalgarno region or the start codon preferentially favors recognition. Translational efficiency is strictly correlated with the fraction of mRNA molecules in which the ribosome binding site is unfolded, indicating that initiation is completely dependent on spontaneous unfolding of the entire initiation region. Ribosomes appear not to recognize nucleotides outside the Shine–Dalgarno region and the initiation codon.**

There is good evidence that mRNA secondary structure is a key factor in determining the efficiency of translational initiation in prokaryotes (1–3). The expression of several genes in *Escherichia coli* appeared inversely related to the stability of the secondary structure of their ribosome binding sites (4–9).

Here, we present a quantitative analysis of the relationship between the stability of a local secondary structure and the level of gene expression *in vivo*. As a model system, we have chosen the coat protein gene of RNA bacteriophage MS2. (*i*) Analysis with structure-sensitive enzymes and chemicals as well as phylogenetic sequence comparison have provided evidence that its ribosome binding site adopts a defined hairpin structure (Fig. 1) (10). This enabled us to alter the stability of this structure in a predictable way through site-directed mutagenesis and monitor the effects on expression. (*ii*) The high translational efficiency of this gene allowed detection of the protein output over a range of four orders of magnitude.

MATERIALS AND METHODS

Bacterial Strains. *E. coli* K-12 strain BW313 (12) was used in the mutagenesis procedure. M13 phages were grown on JM101 (13). Expression was measured in strain M5219 (14), encoding the thermosensitive λ repressor (*c*I857) and the transcriptional antitermination factor N. All strains were grown in LC broth.

Mutagenesis. Oligonucleotide-directed mutagenesis was performed on M13mp11 (13) carrying the *Xba* I(1303)–*Bam*HI(2057) fragment of MS2 cDNA (15). Procedures were essentially as described (12, 16). Mutations were detected by dideoxynucleotide sequencing (13).

Plasmids. Expression of the coat gene (MS2 coordinates 1335–1725) was determined using plasmids carrying the MS2 cDNA fragment *Eco*RI(103)–*Bam*HI(2057) under control of the P_L promoter from phage λ, in vector pPLa236 (14).

Western Blot Analysis. Cultures of the appropriate clones were grown at 28°C to an OD_{650} of 0.2, followed by induction at 42°C for 20 min. Samples of 1 ml were centrifuged and the pellets were boiled in SDS buffer (17). Serial dilutions were fractionated by SDS/PAGE (17) on a 15% gel and blotted onto nitrocellulose. Translation products were immunodetected by using antibodies raised against SDS-denatured MS2 coat protein and alkaline phosphatase-linked goat anti-rabbit antibodies (Sigma).

RESULTS

The secondary structure of the initiation region of the MS2 coat protein gene is shown in Fig. 2 with the mutations introduced. All mutations leave the Shine–Dalgarno (SD) region (GGAG) and the amino acid sequence of the coat protein intact. In Fig. 3 a representative Western blot of five mutants is shown, illustrating how relative expression levels were determined.

We have verified that the observed differences in expression arise at the translational level, by cloning a reporter gene immediately downstream of the MS2 information. For this purpose we used a cDNA copy of the coat protein gene of tobacco rattle virus (18). Expression of this protein was largely independent of the yield of MS2 coat protein and under full control of the P_L promoter as judged by Western blot analysis (not shown).

Only Helices Stronger than the Wild Type Reduce Expression. In the first set of mutants base-pair III was varied (Fig. 2). The 12 mutants obtained (mutants 2–13, Table 1) show that weakening of the helix has no detectable effect on coat protein production. In contrast, replacement of the A·U pair by C·G or G·C reduces expression to 20% and 4%, respectively (mutants 5 and 11). This reduction cannot be due to the nucleotide substitutions *per se*, since the same mutations are present individually in the destabilized mutants.

Multiple Stabilizing Mutations Close the Ribosome Binding Site Completely. To obtain the maximal stabilization allowed by the amino acid sequence, we simultaneously changed base-pair III to G·C, base-pair VI to U·A, and base-pair XII to G·C. In this mutant (mutant 14), coat protein production had dropped below the limits of detection, i.e., to less than 0.003% of the wild type. Clearly, the highly efficient ribo-

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: SD, Shine–Dalgarno.

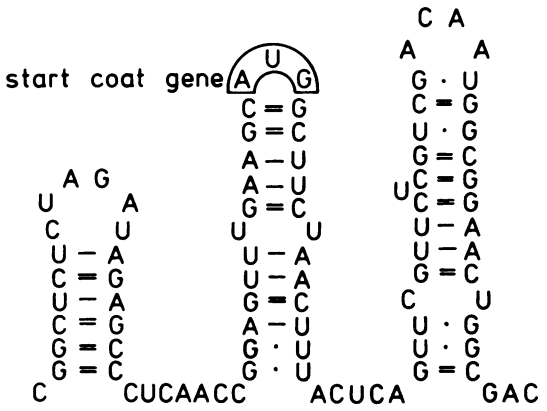

FIG. 1. Secondary structure of the translational initiation region of the coat gene of bacteriophage MS2 (10). The ribosome binding site (i.e., the region initiating ribosomes protect against RNase) extends from position −20 to position +18 relative to the adenosine of the AUG codon (11). –, A·U pair; =, G·C pair; ·, G·U pair.

some binding site of the coat gene can be shut off completely by a secondary structure of sufficient stability.

Compensatory Mutations Restore Expression. To confirm that translation in mutant 14 is prevented by the proposed secondary structure, we introduced a number of second-site destabilizing mutations. First, base-pair III was opened up by changing G·C into A·C (mutant 17). This mutation indeed raised expression from below detection to 3% of the wild type. Replacement of the U·A base pair at position VI by an A·A or C·A mismatch resulted in an expression level of 0.05–0.1% (mutants 15 and 16). When both base pairs were disrupted simultaneously, wild-type expression was restored (mutants 18 and 19), again demonstrating that expression is related to the secondary structure, rather than to the nucleotide sequence.

Changes in Expression by Single Substitutions at Base-Pair VI or XII. When the terminal G·U base pair (XII) was altered to G·C, expression dropped to 6% (mutant 20), confirming the participation of this base pair in the helical structure. The importance of the U·U mismatch in the middle of the helix was examined by replacing it by C·U, A·U, and U·A. As expected, the change to C·U had no effect at all (mutant 21). In the A·U and U·A mutants, on the other hand, expression was reduced to a mere 0.2–0.3% (mutants 22 and 23). Mismatches drastically affect helix stability and the dramatic changes in expression reflect this fact. Evidently, the mismatch plays an essential role in allowing efficient translation.

Expression in Destabilized Mutants Is Not Limited by a Cellular Component. Our results indicate that the strength of the secondary structure of the ribosome binding site determines its translational efficiency. To approach this apparent relationship in a more quantitative manner, we have calculated the stability (ΔG_f^0 or free energy of formation) of the coat initiator hairpin in each of our mutants by using the parameters of Freier *et al.* (19). As shown in Table 1, a good correlation between the ΔG_f^0 of the hairpin and coat protein production is indeed observed for the stabilized mutants.

Interestingly, destabilization does not raise expression over the wild-type level. In view of the high efficiency of promoter and ribosome binding site, we considered that some

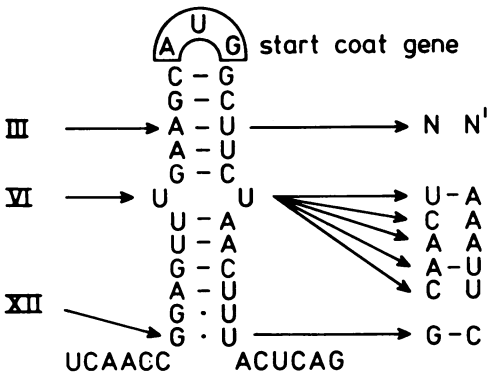

FIG. 2. Mutations introduced in the coat initiator hairpin. Mutated base pairs are denoted by roman numerals (see Table 1). –, Watson–Crick base pair; ·, G·U pair.

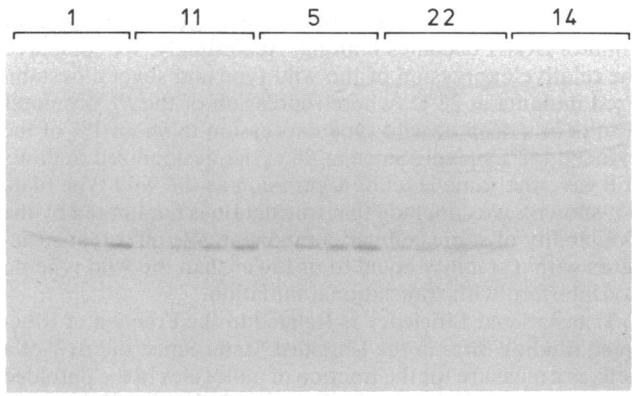

FIG. 3. Representative Western blot showing the coat protein in serial dilutions of extracts of five mutants. Ratios between the amounts of cell extract loaded were 1:8:32:1024:1024 for mutants 1, 11, 5, 22, and 14, respectively, as indicated. Within each mutant series, every lane contains twice as much extract as the previous one. Mutant numbers correspond to Table 1. More accurate estimates of relative protein production were obtained by comparing mutants with similar expression levels.

Table 1. Mutations introduced in the MS2 coat initiator hairpin

Mutant	Base pair III	Base pair VI	Base pair XII	ΔG_f^0, kcal/mol	Relative expression
wt	(A-U)	(U U)	(G·U)	−5.8	100
2	G G	—	—	−2.4	100
3	G A	—	—	−2.4	100
4	G·U	—	—	−5.2	100
5	G-C	—	—	−7.8	4
6	A A	—	—	−2.4	100
7	A C	—	—	−2.4	100
8	U·G	—	—	−5.4	100
9	U U	—	—	−2.4	100
10	U C	—	—	−2.4	100
11	C-G	—	—	−7.3	20
12	C U	—	—	−2.4	100
13	C C	—	—	−2.4	100
14	G-C	U-A	G-C	−14.1	<0.003
15	G-C	A A	G-C	−10.6	0.1
16	G-C	C A	G-C	−10.6	0.05
17	A C	U-A	G-C	−8.7	3
18	A C	A A	G-C	−5.2	80
19	A C	C A	G-C	−5.2	100
20	—	—	G-C	−8.6	6
21	—	C U	—	−5.8	100
22	—	A-U	—	−9.4	0.2
23	—	U-A	—	−9.3	0.3

Base-pair numbering is as in Fig. 2. Base pairs in the wild-type (wt) structure are shown in parentheses. Dashes indicate the presence of the wild-type base pair in the mutants. ΔG_f^0 of the hairpin was calculated using the energy parameters of Freier *et al.* (19) (1 cal = 4.184 J). Relative expression (% of wild type) was determined by comparing serial dilutions of induced cultures on Western blots, as shown in Fig. 3. Symbols between bases are as follows: –, Watson–Crick base pair; ·, G·U pair; space, mismatch.

cellular component such as ribosomes, initiation factors, or initiator tRNA becomes limiting. Accordingly, we compared the relative expression of the wild type and several destabilized mutants at 28°C, where repression of the P_L promoter results in a drop of wild-type expression to about 1% of the induced (42°C) level. Since at 28°C the destabilized mutants still gave the same level of expression as the wild type (data not shown), we conclude that translation is not limited by the availability of some cellular component. We infer that structures with a stability equal to or lower than the wild type do not interfere with translational initiation.

Translational Efficiency Is Related to the Fraction of Ribosome Binding Sites in the Unfolded State. Since the ΔG_f^0 of a helix is a measure for the fraction of molecules in the unfolded state, our results suggest that translational efficiency is directly determined by the availability of unfolded ribosome binding sites. This proposal can be verified, using the thermodynamic relationships

$$\Delta G_f^0 = -RT \ln K_f \quad [1]$$

and

$$f_u = 1/(K_f + 1), \quad [2]$$

where R is the gas constant, T is the temperature (315 K), K_f is the equilibrium constant of helix formation, and f_u is the fraction of unfolded RNA. Combining these two equations yields the following relationship between f_u and ΔG_f^0

$$\ln f_u = -\ln(e^{-\Delta G_f^0/RT} + 1). \quad [3]$$

In Fig. 4A, this relationship is shown as a plot of $\ln f_u$ against ΔG_f^0. The resulting curve has two asymptotes, one with a slope of $1/RT$ (for $\Delta G_f^0 \rightarrow -\infty$) and the other forming a plateau where all molecules are unfolded ($f_u \rightarrow 1$ for $\Delta G_f^0 \rightarrow +\infty$).

If the efficiency of initiation is related to the fraction of unfolded ribosome binding sites, then plotting the natural logarithm of the expression against ΔG_f^0 should yield a similar curve. Fig. 4B shows that this is indeed found. The maximal expression level discussed in the previous section corresponds to the plateau and linear regression based on the lower eight points (solid line) produced a slope that is virtually equal to $1/RT$ (dashed line).

Binding of Ribosomes Shifts the Equilibrium to the Unfolded Side. Although the experimental curve (Fig. 4B) has the same shape as the theoretical curve (Fig. 4A), it is displaced along the horizontal axis. One cause for this displacement is that we have not yet taken into account that the binding of 30S subunits to the unfolded RNA shifts the equilibrium in the direction of unfolding. Indeed, the displacement of the experimental curve to the left means that the level of expression is higher than expected from the calculated f_u.

In Fig. 5, 30S subunit binding is added to our model as a second equilibrium. This is justified because the subsequent step in initiation appears relatively slow (20, 21). Since the protein yield is proportional to the concentration of 30S subunit–mRNA complex (21, 22), the relative expression (E) equals the fraction of mRNA present in this complex. By using the definitions of the equilibrium constants K_f and K_{30S}, it can be shown that

$$E = (K_{30S}[30S])/(1 + K_f + K_{30S}[30S]). \quad [4]$$

In analogy to Fig. 4, we have plotted E against ΔG_f^0 for various 30S subunits–mRNA affinities (ΔG_{30S}^0) (Fig. 6). The degree to which the 30S subunits displace the equilibrium evidently depends on their concentration and on their affinity for the unfolded ribosome binding site. As expected, high affinities shift the expression curve further to the left. The concentration of free 30S subunits was set at 8.5 μM (23, 24) for all curves.

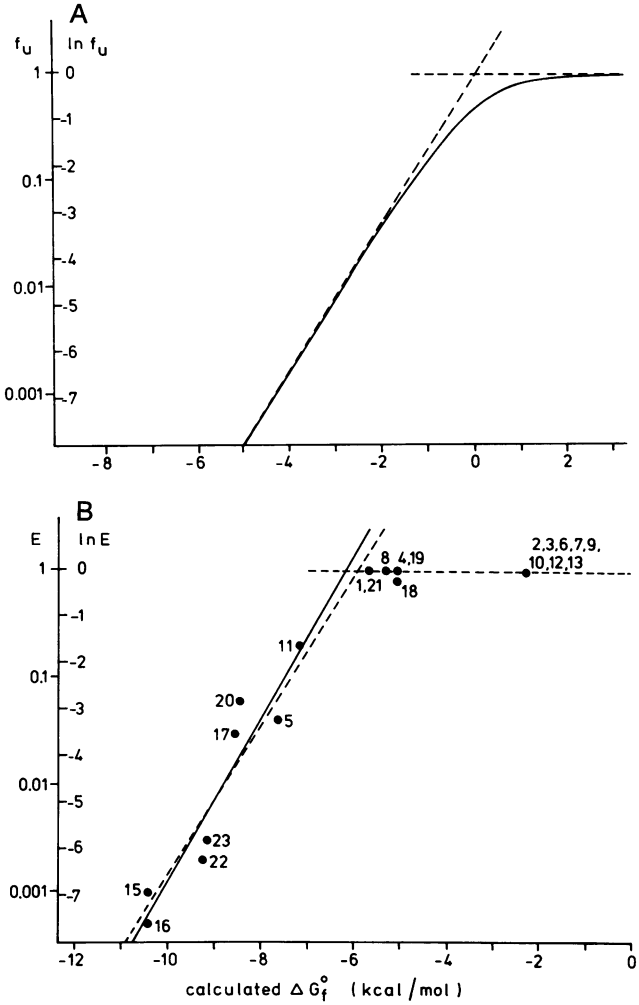

FIG. 4. (A) Relationship between the stability of a helix and the fraction of molecules in the unfolded form. Dashed lines indicate the asymptotes. (B) Relationship between the stability of the coat initiator hairpin and coat protein production. The solid line is based on linear regression of the lower eight points. Dashed lines indicate the theoretical slope of $1/RT$ and the plateau where $E = 1$ (see text).

Entering our data points in Fig. 6 suggests that the ribosome binding site of the coat gene has a ΔG_{30S}^0 of about −13.5 kcal/mol. This compares reasonably well with the average −10.5 kcal/mol determined *in vitro* for other efficient ribosome binding sites (21, 22, 25). The difference may in part arise by the inability of the available thermodynamic parameters to predict the absolute stability of the coat gene initiator hairpin *in vivo* in its natural RNA context. Alternatively, the affinity of ribosomes for this site may just be very high.

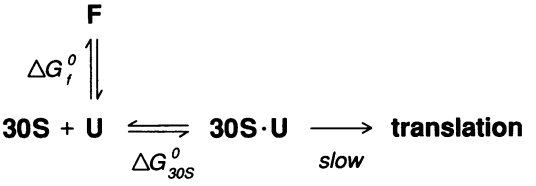

FIG. 5. Two equilibria that determine the fraction of mRNA bound to 30S subunits. ΔG_f^0 is the free energy of helix formation; ΔG_{30S}^0 is the free energy of binding of a 30S subunit to the unfolded mRNA.

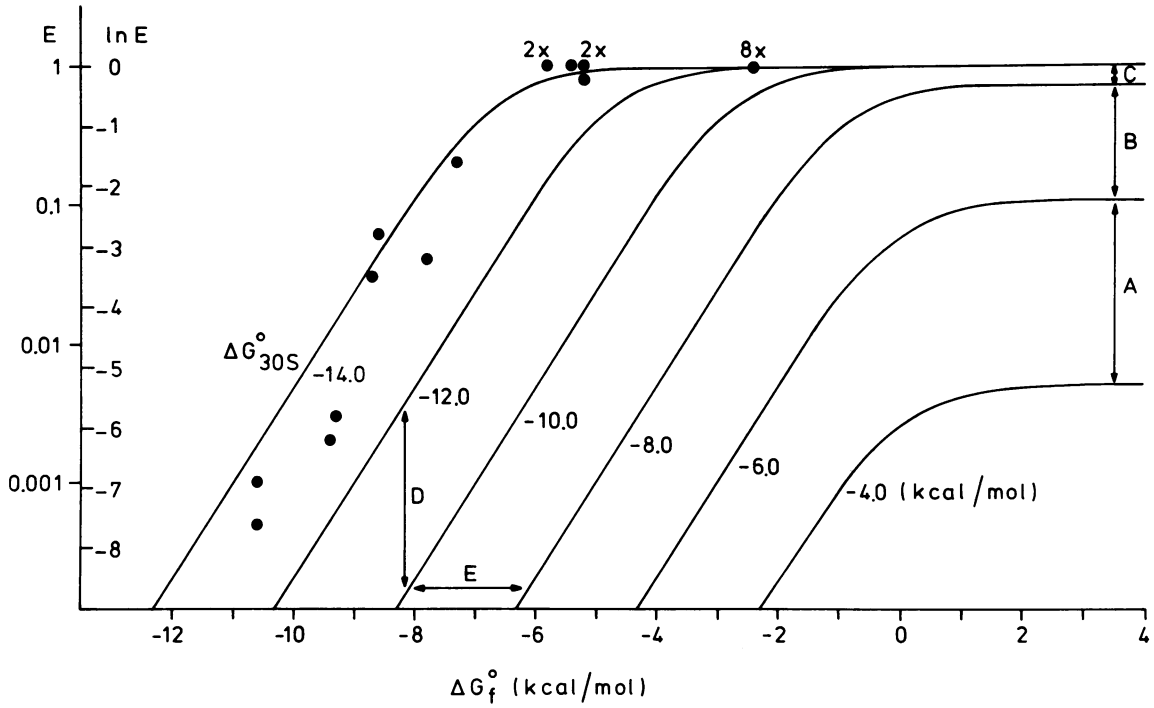

FIG. 6. Relationship between the stability of the secondary structure in a ribosome binding site and the relative expression predicted by the scheme of Fig. 5, for various affinities ($\Delta G^0_{30S} = -RT \ln K_{30S}$; [30S] = 8.5 μM). Experimental data are shown as in Fig. 4*B*. Arrows A–E are explained in the text.

DISCUSSION

All of our nucleotide substitutions affect expression only by changes in helix stability, irrespective of their position in the hairpin. Initiation apparently requires spontaneous unfolding of the whole ribosome binding site. These findings show that ribosomes do not recognize mRNA secondary structures and they are also contradictory to earlier ideas that the location of an AUG codon in the loop of a hairpin or exposure of just the SD region could facilitate initiation (26–28). Our results, therefore, support the view that ribosomes only bind to single-stranded RNA (1–3, 29).

Figs. 4 and 6 show that a decrease in ΔG^0_f by 1.4 kcal/mol can reduce translation by a factor of 10. This strong effect is most conspicuous in mutant 22 where elimination of the mismatch stabilizes the initiator helix by −3.6 kcal/mol. As a result expression drops by a factor of 500.

Contribution of Ribosome–mRNA Affinity to Efficiency of Translation. Our concept of mRNA binding implies that the efficiency of translational initiation is not necessarily proportional to the association constant of ribosomes with their binding site on the messenger (K_{30S}) (1). Fig. 6 shows that this is only true for the sloped part of the curves, where expression is limited by mRNA structure (arrow D) or when the mRNA has a low affinity for ribosomes (arrows A and B). mRNAs whose expression is not limited by secondary structure (the plateaus in Fig. 6) are already saturated with 30S subunits at a moderate affinity (e.g., $\Delta G^0_{30S} = -10$ kcal/mol) and an increase in affinity does not raise expression any further (arrow C). This explains how a 10-fold difference in K_{30S} can result in only a 2-fold difference in expression (25).

The affinity contributes to the translation of structured mRNAs in that each increase of 1 kcal/mol in ΔG^0_{30S} helps to "overcome" secondary structure of 1 kcal/mol in the ribosome binding site (arrow E). This may explain why increasing the strength of the SD interaction only raises the expression if the ribosome binding site is involved in base pairing (30). Similarly, the requirement for ribosomal protein S1 in translation of structured mRNAs is probably directly related to its positive effect on the affinity of 30S subunits for mRNA (for review, see refs. 31 and 32). Indeed, ribosomal protein S1 seems able to complement a weak SD interaction (11, 33).

These considerations also explain why secondary structures, stable enough to confer protection against nucleases, do not necessarily impair the translational efficiency (7, 34). For example, the wild-type hairpin of the MS2 coat start is cut by the double-strand-specific cobra venom endonuclease (10), in accordance with the calculation that only a minor fraction of the helices are unfolded. Nevertheless, because of the high affinity virtually all molecules are bound by ribosomes (Fig. 4*B*). It is noteworthy that the strength of the hairpin is just on the verge of becoming inhibitive (Fig. 4*B*). Its evolution has apparently resulted in the highest stability compatible with maximal ribosome loading.

Ribosomes Do Not Recognize Nucleotides Outside the Start Codon and the SD Region. Not all start codons preceded by an SD sequence function as translational initiation sites. Statistical analyses have revealed a nucleotide bias in true ribosome binding sites that is absent from such "false starts" and it has often been suggested that ribosomes recognize this biased sequence (36–39). The difference in efficiency between true and false starts was shown to be maintained after transfer to another RNA context. This was interpreted as evidence that the main determinants for discrimination lie in the nucleotide sequence (40). Our results seem at odds with these proposals. Including data from others (41), we find that mutations at positions −6, −3, +6, +9, +12, +13, and +15 relative to the AUG codon only affect expression if they change the stability of the secondary structure. Apparently, outside the SD region and the initiation codon, nucleotide changes *per se* have little or no effect on expression. The paradox would be resolved if the sequence bias arose from the need to expose true ribosome binding sites and to shield false ones. The high preference for adenosine residues in true sites has indeed been ascribed to their low potential to form stable structures (29). Moreover, there is statistical evidence that false sites have a potential for base pairing that is higher than random (42).

Turner Rules Reliably Predict Differential Helix Stability *in Vivo*. Our data points follow the predicted slope of $1/RT$ (Fig. 4), which indicates that the published free energy increments associated with base-pair stacking and mismatches (19) are close to their true values *in vivo*. The recently proposed parameters for mismatches (35) do not yield a qualitatively different picture.

We have examined two sets of expression data (4, 5) and these also fit the theoretical curves of Fig. 6 very well (data not shown). This supports the general validity of our findings.

We thank Dr. W. Fiers and coworkers for the MS2 cDNA and expression vectors, Dr. T. Kunkel for strain BW313, and Gerco Angenent for TRV cDNA. Molly Hughes is acknowledged for her assistance and support in the initial stages of this work.

1. Stormo, G. D. (1986) in *Maximizing Gene Expression*, eds. Reznikoff, W. & Gold, L. (Butterworth, Boston), pp. 195–254.
2. Gold, L. (1988) *Annu. Rev. Biochem.* **57,** 199–233.
3. de Smit, M. H. & van Duin, J. (1990) *Prog. Nucleic Acid Res. Mol. Biol.* **38,** 1–35.
4. Hall, M. N., Gabay, J., Débarbouillé, M. & Schwartz, M. (1982) *Nature (London)* **295,** 616–618.
5. Buell, G., Schultz, M.-F., Selzer, G., Chollet, A., Movva, N. R., Semon, D., Escanez, S. & Kawashima, E. (1985) *Nucleic Acids Res.* **13,** 1923–1938.
6. Tessier, L.-H., Sondermeyer, P., Faure, T., Dreyer, D., Benavente, A., Villeval, D., Courtney, M. & Lecocq, J.-P. (1984) *Nucleic Acids Res.* **12,** 7663–7675.
7. Schmidt, B. F., Berkhout, B., Overbeek, G. P., van Strien, A. & van Duin, J. (1987) *J. Mol. Biol.* **195,** 505–516.
8. Spanjaard, R. A., van Dijk, M. C. M., Turion, A. J. & van Duin, J. (1989) *Gene* **80,** 345–351.
9. Tomich, C.-S. C., Olson, E. R., Olsen, M. K., Kaytes, P. S., Rockenbach, S. K. & Hatzenbuhler, N. T. (1989) *Nucleic Acids Res.* **17,** 3179–3197.
10. Skripkin, E. A., Adhin, M. R., de Smit, M. H. & van Duin, J. (1990) *J. Mol. Biol.* **211,** 447–463.
11. Steitz, J. A. (1979) in *Biological Regulation and Development*, ed. Goldberger, R. F. (Plenum, New York), Vol. 1, pp. 349–399.
12. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 488–492.
13. Messing, J. (1983) *Methods Enzymol.* **101,** 20–78.
14. Remaut, E., Stanssens, P. & Fiers, W. (1981) *Gene* **15,** 81–93.
15. Devos, R., van Emmelo, J., Contreras, R. & Fiers, W. (1979) *J. Mol. Biol.* **128,** 595–619.
16. Berkhout, B., Schmidt, B. F., van Strien, A., van Boom, J., van Westrenen, J. & van Duin, J. (1987) *J. Mol. Biol.* **195,** 517–524.
17. Laemmli, U. K. (1970) *Nature (London)* **227,** 680–685.
18. Angenent, G. C., Posthumus, E. & Bol, J. F. (1989) *Virology* **173,** 68–76.
19. Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T. & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9373–9377.
20. Gualerzi, C., Risuleo, G. & Pon, C. L. (1977) *Biochemistry* **16,** 1684–1689.
21. Ellis, S. & Conway, T. W. (1984) *J. Biol. Chem.* **259,** 7607–7614.
22. Draper, D. E. (1987) in *Translational Regulation of Gene Expression*, ed. Ilan, J. (Plenum, New York), pp. 1–26.
23. Forchhammer, J. & Lindahl, L. (1971) *J. Mol. Biol.* **55,** 563–568.
24. Gouy, M. & Grantham, R. (1980) *FEBS Lett.* **115,** 151–155.
25. Calogero, R. A., Pon, C. L., Canonaco, M. A. & Gualerzi, C. O. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 6427–6431.
26. Iserentant, D. & Fiers, W. (1980) *Gene* **9,** 1–12.
27. Selker, E. & Yanofsky, C. (1979) *J. Mol. Biol.* **130,** 135–143.
28. Queen, C. & Rosenberg, M. (1981) *Cell* **25,** 241–249.
29. Looman, A. C., Bodlaender, J., Comstock, L. J., Eaton, D., Jhurani, P., de Boer, H. A. & van Knippenberg, P. H. (1987) *EMBO J.* **6,** 2489–2492.
30. Munson, L. M., Stormo, G. D., Niece, R. L. & Reznikoff, W. S. (1984) *J. Mol. Biol.* **177,** 663–683.
31. Thomas, J. O. & Szer, W. (1982) *Prog. Nucleic Acid Res. Mol. Biol.* **27,** 157–187.
32. Subramanian, A.-R. (1983) *Prog. Nucleic Acid Res. Mol. Biol.* **28,** 101–142.
33. Roberts, M. W. & Rabinowitz, J. C. (1989) *J. Biol. Chem.* **264,** 2228–2235.
34. Rosa, M. D. (1981) *J. Mol. Biol.* **147,** 55–71.
35. Jaeger, J. A., Turner, D. H. & Zuker, M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7706–7710.
36. Scherer, G. F. E., Walkinshaw, M. D., Arnott, S. & Morré, D. J. (1980) *Nucleic Acids Res.* **8,** 3895–3907.
37. Stormo, G. D., Schneider, T. D. & Gold, L. M. (1982) *Nucleic Acids Res.* **10,** 2971–2996.
38. Stormo, G. D., Schneider, T. D., Gold, L. & Ehrenfeucht, A. (1982) *Nucleic Acids Res.* **10,** 2997–3011.
39. Schneider, T. D., Stormo, G. D., Gold, L. & Ehrenfeucht, A. (1986) *J. Mol. Biol.* **188,** 415–431.
40. Dreyfus, M. (1988) *J. Mol. Biol.* **204,** 79–94.
41. Precup, J. & Parker, J. (1987) *J. Biol. Chem.* **262,** 11351–11355.
42. Ganoza, M. C., Kofoid, E. C., Marlière, P. & Louis, B. G. (1987) *Nucleic Acids Res.* **15,** 345–360.