Proc. Natl. Acad. Sci. USA
Vol. 87, pp. 8301–8305, November 1990
Biochemistry

Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes

(mRNA structure/alternative initiator codons/scanning model/*in vitro* translation)

MARILYN KOZAK

Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Robert Wood Johnson Medical School, 675 Hoes Lane, Piscataway, NJ 08854

Communicated by Aaron J. Shatkin, August 21, 1990 (received for review July 30, 1990)

ABSTRACT **Recognition of an AUG initiator codon in a suboptimal context improves when a modest amount of secondary structure is introduced near the beginning of the protein-coding sequence. This facilitating effect depends on the position of the downstream stem–loop (hairpin) structure. The strongest facilitation is seen when the hairpin is separated from the preceding AUG codon by 14 nucleotides. Because 14 nucleotides corresponds to the approximate distance between the leading edge of the ribosome and its AUG-recognition center as measured by ribonuclease protection experiments, a likely explanation for the enhancing effect of a downstream hairpin is that secondary structure slows scanning, thereby providing more time for recognition of the AUG codon, and the facilitation is greatest when the 40S ribosome stalls with its AUG-recognition center directly over the AUG. The variable ability of mammalian ribosomes to initiate at non-AUG codons *in vitro* is also explicable by the presence or absence of a stem–loop structure just downstream from the alternative initiator codon. This may be relevant to recent reports of adventitious upstream initiation events at non-AUG codons in some vertebrate mRNAs that have structure-prone, G+C-rich leader sequences.**

Recognition of initiator codons by mammalian ribosomes is modulated by particular sequences flanking the AUG, or occasionally non-AUG, codon (1–3). Thus, when the first AUG codon occurs in an unfavorable primary sequence context, some 40S ribosomes bypass that site and initiate instead at an AUG triplet farther downstream (4). This "leakiness" can be suppressed to some extent by introducing a modest amount of secondary structure downstream from the first AUG codon (4).

The facilitating effect of downstream secondary structure can be fitted to the scanning model for initiation (5) by postulating that secondary structure slows the progression of the 40S ribosome-plus-factor complex, thereby providing more time for recognition of the preceding AUG codon. According to that hypothesis, the enhancement should depend critically on the position of the downstream hairpin structure. Here I show that the facilitating effect of a hairpin introduced near the beginning of the coding sequence indeed varies with the position of the hairpin and is maximal when 14 nucleotides intervene between the AUG codon and the base of the stem–loop structure.

An optimally positioned hairpin of moderate strength [−19 kcal/mol, calculated as the Gibbs free energy of formation (ΔG) (6)] also strikingly improves initiation at preceding non-AUG codons, such as GUG and UUG. This may bear on recent reports of adventitious upstream initiation events in the G+C-rich leader sequences of some mammalian mRNAs (7–15).

MATERIALS AND METHODS

Construction of Plasmids. These experiments were carried out with derivatives of the SP64-CAT (chloramphenicol acetyltransferase) plasmids described previously (4). A cassette mutagenesis approach was used as before (4) to introduce an ATG codon upstream from the normal CAT start site and to vary a small block of nucleotides that begins at the *Hin*dIII site, 10 nucleotides upstream from the first ATG codon (Fig. 1), and extends 2 to 32 nucleotides downstream from that ATG.

Plasmids J-atg-2(8334)CAT and J-atg-2(8336)CAT are identical except for the adaptor (lowercase letters in Fig. 1*A*, lines 1 and 2), which was introduced to shift the first ATG codon into the same reading frame as the second. Oligonucleotide 8336 can fold into a hairpin structure (ΔG = −19 kcal/mol; ref. 6) demarked by the horizontal arrows above the sequence (Fig. 1*A*, line 2); control transcripts containing oligonucleotide 8334 can form no stable secondary structure in that region. The other three constructs listed in Fig. 1*A* are identical to J-atg-2(8336)CAT except that the distance between the first ATG codon and the beginning of the hairpin has been increased to 8, 14, or 32 nucleotides; the plasmids are named accordingly.

The three plasmids listed in Fig. 1*B* are identical except for the sequence of the adaptor (lowercase letters, nucleotides +18 to +39), which was manipulated to change the stability of the downstream hairpin from −19 kcal/mol (oligonucleotide 8336) to −5.2 kcal/mol (oligonucleotide 8414, in which 3 nucleotides in the ascending limb of the stem have been changed) and back to −19 kcal/mol via compensatory changes in the descending limb (oligonucleotide 8416). This was achieved by recombination, at the *Bam*HI site, between J-atg-14(8336)CAT and the previously described plasmids SP64(8414)B34 and SP64(8416)B34 (4).

In the remainder of the text I will use U instead of T in describing the sequences of these constructs.

Other Methods. The procedures for purifying, sequencing, and manipulating plasmid DNA were described previously (4). *In vitro* transcription of *Ava* I-linearized plasmid DNA with SP6 polymerase (BRL) followed the general procedures of Melton *et al.* (16); specific reaction conditions are given in ref. 4. Capped transcripts were purified by phenol extraction and Sephadex G-50 chromatography. Each 30-μl translation reaction mixture contained 0.5 μg of mRNA, 50 μCi of [^{35}S]methionine (1000 Ci/mmol, DuPont/NEN Research Products; 1 Ci = 37 GBq), 10 μl of messenger-dependent reticulocyte lysate (BRL), and other components as described (4). Equal aliquots of each translation reaction were

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: CAT, chloramphenicol acetyltransferase.

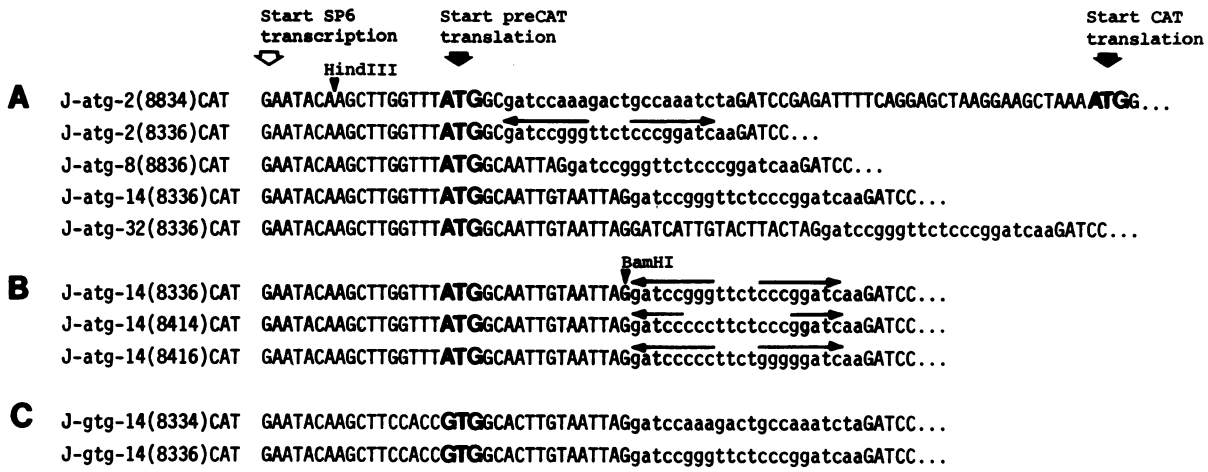

FIG. 1. Structures of inserts in SP64-based plasmids that direct translation of CAT and preCAT proteins. The origin of the parental SP64-CAT construct is detailed in ref. 4. All plasmids are identical upstream from the *Hin*dIII site and downstream from the ellipsis (. . .). Diverging arrows highlight the potential formation of hairpin structures downstream from the preCAT start site.

analyzed by polyacrylamide gel electrophoresis (4). Autoradiograms that had been exposed for 1 or 2 days were quantified by densitometry.

Reticulocyte lysates from some other commercial sources could be substituted for the BRL lysate with comparable results (17), provided that reaction conditions were adjusted to those worked out with the BRL lysate. Those reaction conditions were chosen because results thereby obtained *in vitro* were congruent with *in vivo* results (4).

RESULTS

Because my purpose was to test for features other than local context (1–3) that might enhance recognition of initiator codons *in vitro*, I began by placing the first AUG codon in a suboptimal context. Thus, in plasmids J-aug-2(8334)CAT and J-aug-2(8336)CAT (Fig. 1*A*, lines 1 and 2), the presence of U instead of the preferred A in position −3 demarks a rather weak initiation site; this enables ribosomes to produce two proteins—"preCAT," initiated from the first AUG codon, and CAT, initiated from the second AUG codon—by leaky scanning (Fig. 2, lanes 1 and 2). The question addressed in the next section was whether the ratio of preCAT to CAT protein would shift in response to sequence changes downstream from the preCAT start site. Specifically, I wanted to confirm and extend the notion (4) that downstream secondary structure might enhance recognition of the preceding AUG codon.

Varying the Position of a Downstream Hairpin. The only difference between plasmids J-aug-2(8334)CAT and J-aug-2(8336)CAT is the sequence of the 22-nucleotide adaptor (shown in lowercase letters in Fig. 1*A*) which follows the first AUG codon. Oligonucleotide 8336 can assume a stem–loop structure whereas oligonucleotide 8334 cannot. Comparison of protein yields from J-aug-2(8334)CAT and J-aug-2(8336)-CAT reveals no difference in the preCAT/CAT ratio (Fig. 2, lane 1 vs. lane 2), indicating that the potential formation of a hairpin 2 nucleotides beyond the first AUG codon neither helps nor hinders initiation from that AUG. That outcome was not unexpected inasmuch as a hairpin positioned so close to the first AUG codon would inevitably have to be melted before the advancing 40S ribosomal subunit could "see" the AUG triplet.

To assess the relationship between the position of a stem–loop structure and its ability to enhance recognition of the preceding AUG codon, four matched transcripts were tested: J-aug-2(8336)CAT, J-aug-8(8336)CAT, J-aug-14(8336)CAT, and J-aug-32(8336)CAT, where 2, 8, 14, and 32 refer to the number of nucleotides between the first AUG codon and the beginning of the hairpin structure. The results of translating these transcripts *in vitro* are shown in Fig. 2, lanes 2–5. As the distance of the hairpin from the first AUG codon was increased from 2 to 8 to 14 nucleotides, the ratio of preCAT to CAT protein gradually increased; indeed, ribosomes initiated almost exclusively at the preCAT start site in J-aug-14(8336)CAT, indicating that downstream secondary structure can compensate to a significant extent for absence of the preferred primary sequence around the AUG codon. The preCAT/CAT ratio in lane 5 was the same as in lanes 1 and 2, indicating that the facilitation is lost when the hairpin is moved too far downstream from the targeted AUG codon.

The simplest interpretation of these results is that a suitably positioned stem–loop structure slows the scanning 40S ribosomal subunit, thereby providing more time for recognition of the preceding AUG codon. The shift in the preCAT/CAT

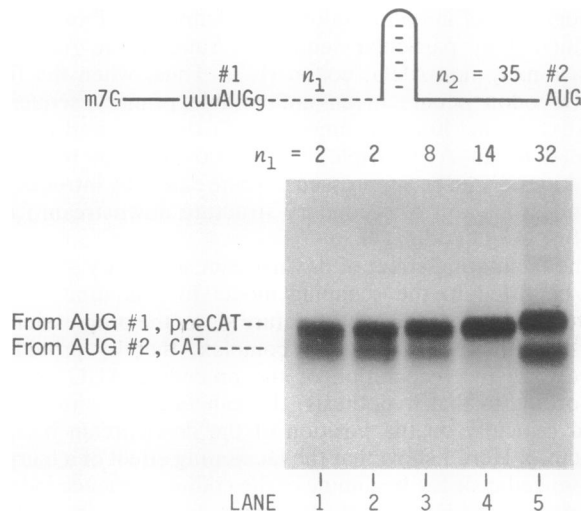

FIG. 2. Dependence of the enhancing effect of downstream secondary structure on its distance from the targeted AUG codon. Transcripts produced from the five plasmids diagrammed in Fig. 1*A* were translated *in vitro*; the [^{35}S]methionine-labeled polypeptides were analyzed by polyacrylamide gel electrophoresis and autoradiography. The control transcript used for lane 1 was J-aug-2(8334)CAT, which has no deliberate secondary structure between the preCAT and CAT start sites. The transcripts used for lanes 2–5 contained oligonucleotide 8336, which is predicted to form a stem–loop structure (ΔG = −19 kcal/mol); the distance (n_1) between the base of the hairpin and the preceding AUG codon is indicated above each lane of the autoradiogram.

ratio is evidently due to enhanced recognition of the AUG codon that precedes the hairpin, rather than to suppression of initiation from the second AUG, inasmuch as the distance of the hairpin from the second AUG codon is invariant in constructs that produce very different amounts of CAT protein. Moreover, the stem–loop structure formed by oligonucleotide 8336 is less stable than other hairpin structures that have been shown not to inhibit initiation (18, 19).

Disrupting and Restoring the Hairpin Structure. To determine whether it is indeed the hairpin in J-aug-14(8336)CAT, rather than the primary sequence downstream, that enhances initiation from the first AUG codon, I changed three nucleotides on the ascending side of the stem, thereby reducing its stability from −19 to −5.2 kcal/mol. The resulting construct, J-aug-14(8414)CAT, produced preCAT and CAT proteins in a ratio of 1.5:1 (Fig. 3, lane 2). Subsequent restoration of the −19 kcal/mol hairpin in J-aug-14(8416)CAT restored the preferential use of the first AUG codon (Fig. 3, lane 3). These results show that the presence of a downstream hairpin, but not the precise sequence of the hairpin, is required to promote initiation from the preceding AUG codon.

Initiation at Non-AUG Codons. Initiation at codons other than AUG is rare in higher eukaryotes (see *Discussion*). Even when an alternative initiator codon is in a favorable primary sequence context, which is a near-absolute requirement, initiation at non-AUG codons is unpredictable and usually inefficient. To test the possibility that the presence of downstream secondary structure might increase the efficiency of initiation at non-AUG codons, I inserted a GUG codon in place of the AUG codon at the preCAT start site. In J-gug-14(8336)CAT (Fig. 1*C*, line 2), a −19 kcal/mol hairpin occurs in the optimal position, 14 nucleotides downstream from the GUG codon; the control construct J-gug-14(8334)CAT has no deliberate secondary structure downstream. In the absence of secondary structure there was only weak recognition of the GUG codon (Fig. 4, lane 1, upper band marked preCAT) despite its favorable primary sequence context. The imposition of a downstream hairpin caused a 3-fold increase in the yield of preCAT protein initiated from the upstream GUG codon and a corresponding decrease in synthesis of CAT (Fig. 4, lane 2). Similar results were obtained with UUG,

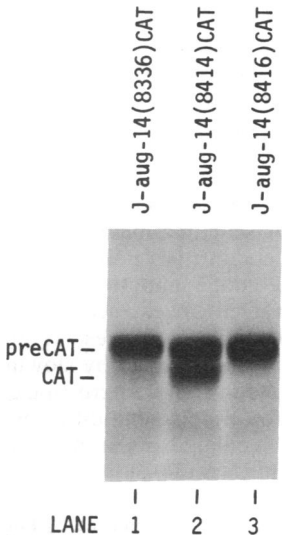

FIG. 3. Enhanced recognition of the preCAT start site correlates with the presence of a stem–loop structure. The autoradiogram shows, in lane 1, the translation of a transcript that contains the −19 kcal/mol hairpin; in lane 2, the stem was disrupted by changing 3 nucleotides; in lane 3, a −19 kcal/mol hairpin was restored by compensatory sequence changes. The sequences of the three plasmids are given in Fig. 1*B*.

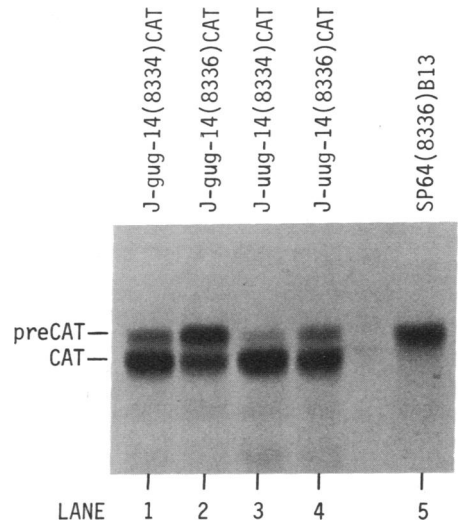

FIG. 4. Effect of downstream secondary structure on initiation at non-AUG codons. The sequences of transcripts that initiate preCAT translation from a GUG codon (lanes 1 and 2) are given in Fig. 1*C*; the only change for lanes 3 and 4 was the substitution of UUG for GUG. Comparison of lane 1 with lane 2, or lane 3 with lane 4, reveals a 3-fold increase in synthesis of preCAT protein when the structured oligonucleotide 8336 follows the GUG or UUG codon. Lane 5 shows the translation of a control construct that has an AUG codon in an optimal context at the preCAT start site.

which is an extremely weak alternative initiator codon in the absence of secondary structure (Fig. 4, lane 3); initiation at the upstream UUG codon increased 3-fold when the structure-prone oligonucleotide 8336 was introduced downstream (Fig. 4, lane 4). The control for this experiment was SP64(8336)B13 (see ref. 4) in which the 5′-proximal AUG codon occurs in a favorable primary sequence context; accordingly, preCAT was the only detectable protein product (Fig. 4, lane 5).

DISCUSSION

These experiments reveal that recognition of an AUG codon in a suboptimal context is higher when the adjacent downstream sequence is capable of assuming a stem–loop structure than when the downstream region is unstructured. Stated differently, when secondary structure is imposed downstream from the first AUG codon, recognition of the AUG codon by mammalian ribosomes becomes less dependent on the flanking primary sequence. Thus, with J-aug-14(8336)CAT, nearly all ribosomes initiated at the 5′-proximal AUG triplet, even though its flanking sequence (UGGUUU in positions −6 to −1) was far from optimal. Although context effects on initiation have been observed in many different laboratories (reviewed in ref. 5), the magnitude of the effects has varied; a possible explanation is that the contribution of downstream secondary structure differed from one construct to another. Some of the reported differences in efficiency of initiation at non-AUG codons (refs. 4 and 20; see below) might also be traced to different degrees of secondary structure near the start of the coding sequence.

The significance of downstream secondary structure in natural mRNAs appears to be twofold. (*i*) It probably explains why initiation is not a little leaky in the majority of mammalian mRNAs. Although 97% of vertebrate mRNAs have a purine in position −3 (ref. 3), few possess the full consensus sequence; thus, some feature in addition to primary sequence would seem to be required to explain the usual absence of leakiness. (*ii*) The contribution of downstream secondary structure may be most noticeable to the

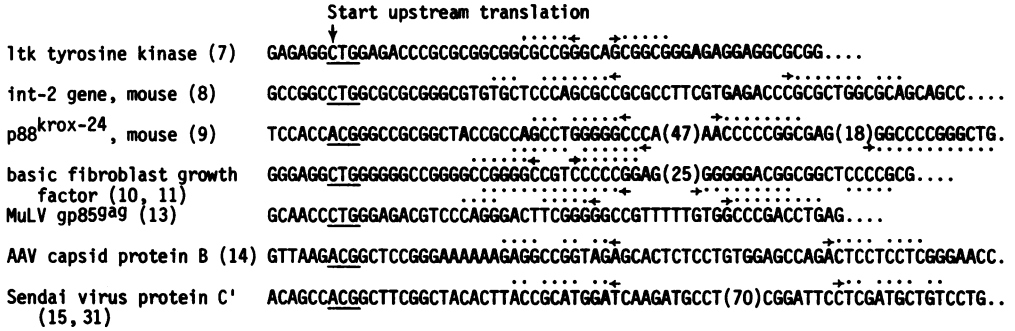

FIG. 5. Sequences of viral and cellular mRNAs that have been reported to initiate at an upstream CUG or ACG codon. References from which the sequences were taken are given in parentheses after the name of each protein. Rows of dots mark nucleotides that might form a stem–loop structure in a position that could facilitate initiation from the preceding CUG or ACG codon. For Krox-24 and fibroblast growth factor, two possible base pairings are noted, one by dots placed above and the other by dots placed below the sequence. Numbers in parentheses within some sequences indicate omitted nucleotides; thus, in some cases, the right arm of the hypothetical stem–loop structure lies some distance downstream. The predicted stem–loop structures are not unique, as discussed in the text. Note the favorable primary sequence (A or G in position −3 and G in position +4) around each alternative initiator codon. MuLV, murine leukemia virus; AAV, adeno-associated virus.

extent that it allows initiation from sites that would otherwise not be used, or not be used efficiently. Inspection of the handful of vertebrate mRNAs in which the AUG initiator codon occurs in the weakest primary sequence context (lacking the preferred nucleotide in positions −3 and +4) indeed reveals appropriately positioned downstream complementary sequences in most cases (21–24). If the positive effects of secondary structure, like the negative effects (25), can be modulated by changes in growth conditions, it may be no accident that growth factors and other crucial regulatory proteins are among the handful of vertebrate proteins initiated from "weak" AUG codons that depend (I predict) on downstream secondary structure.

Initiation at Codons Other Than AUG. Like the recognition of AUG codons that occur in an unfavorable primary sequence context, recognition of non-AUG codons—even when they occur in a favorable context—is strongly stimulated by downstream secondary structure (Fig. 4). This is of interest because of the growing number of natural mRNAs known to use alternative initiator codons (7–15), albeit usually inefficiently and usually as an adjunct to (not instead of) initiation at an AUG codon. With one exception (13) the aforementioned examples were documented *in vivo*, thereby precluding the very real possibility of artifacts induced by inappropriate reaction conditions *in vitro* (reviewed in ref. 4). In contrast with higher eukaryotes, where ribosomes can initiate (rarely and usually inefficiently) at non-AUG codons, the phenomenon is virtually undetectable in *Saccharomyces cerevisiae* (26, 27), perhaps because yeast show little in the way of context effects (26, 28, 29). In vertebrates, alternative initiator codons must occur in the optimal primary sequence context to be functional (4, 30).

A discrete downstream hairpin is necessary and (together with a favorable local context) sufficient to enhance recognition of the preceding GUG or UUG codon in synthetic constructs (Fig. 4); but, in the few natural vertebrate mRNAs that have been shown to initiate at non-AUG codons, the extraordinary G+C-richness of the downstream region probably allows formation of numerous alternative base-paired structures. Thus, although Fig. 5 points out one, or sometimes two, base-pairing possibilities in appropriate positions to influence initiation from the preceding ACG or CUG codon, the base-pairing schemes postulated in the figure are by no means unique. This underscores the difficulty of testing a rather simple hypothesis with natural mRNAs, and the corresponding utility of resorting to synthetic constructs. The general notion that mRNAs with G+C-rich leader sequences might be prone to initiating at adventitious upstream sites enables one to predict which other natural mRNAs might initiate at non-AUG codons, with interesting consequences. For example, it might be worthwhile to determine whether the long G+C-rich leader sequence on *syn* mRNA directs the synthesis of a 179-amino acid polypeptide that could initiate from a CUG codon near the 5′ end of the cDNA (32).

Finally, the fact that 80S ribosomes can penetrate secondary structures that are too stable to be melted by 40S initiation complexes (18) raises the interesting possibility that structure-prone G+C-rich leader sequences are not only required for, but actually necessitate, initiation from cryptic upstream sites.

Mechanistic Considerations. The simplest explanation for the facilitating effect of downstream secondary structure is that it slows the scanning 40S ribosomal subunit, thereby providing more time for recognition of the preceding AUG (or alternative) initiator codon. That hypothesis gains support from the finding that the facilitating effect of the hairpin depends on its distance from the targeted AUG codon and that 14 nucleotides is the optimal spacing. As measured by ribonuclease protection experiments (33), the distance between the ribosome's leading edge and the AUG-recognition center (i.e., the distance of the AUG codon from the 3′ edge of the ribosome-protected fragment) is 12 to 15 nucleotides. Thus, I imagine that a hairpin located 12 to 15 nucleotides downstream causes the scanning 40S ribosome to pause with its AUG-recognition center *right over the initiator codon* and, thus positioned, it is easier for the ribosome to initiate in the absence of the preferred context or even in the absence of the standard AUG initiator codon. Indeed, the fidelity of initiation might boil down to controlling the kinetics of scanning: anything that slows scanning might enhance recognition of the AUG initiator codon (to a point, beyond which further slowing might encourage initiation at spurious upstream sites); anything that enhances recognition of the initiator codon, such as flanking sequences (1, 2) or downstream secondary structure, might do so by slowing scanning.

The literature contains other interesting examples of modulation of translation by discontinuities in the rate of ribosome progression mediated apparently by conformational constraints in the mRNA (34–38).

Research in my laboratory is supported by Public Health Service Grant GM33915 from the National Institutes of Health.

1. Kozak, M. (1986) *Cell* **44**, 283–292.
2. Kozak, M. (1987) *J. Mol. Biol.* **196**, 947–950.
3. Kozak, M. (1987) *Nucleic Acids Res.* **15**, 8125–8148.
4. Kozak, M. (1989) *Mol. Cell. Biol.* **9**, 5073–5080.
5. Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241.
6. Tinoco, I., Borer, P. N., Dengler, B., Levine, M. D., Uhlen-

beck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246,** 40–41.
7. Bernards, A. & de la Monte, S. M. (1990) *EMBO J.* **9,** 2279–2287.
8. Acland, P., Dixon, M., Peters, G. & Dickson, C. (1990) *Nature (London)* **343,** 662–665.
9. Lemaire, P., Vesque, C., Schmitt, J., Stunnenberg, H., Frank, R. & Charnay, P. (1990) *Mol. Cell. Biol.* **10,** 3456–3467.
10. Florkiewicz, R. Z. & Sommer, A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 3978–3981.
11. Prats, H., Kaghad, M., Prats, A., Klagsbrun, M., Lelias, J., Liauzun, P., Chalon, P., Tauber, J. P., Amalric, F., Smith, J. A. & Caput, D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1836–1840.
12. Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. & Eisenman, R. N. (1988) *Cell* **52,** 185–195.
13. Prats, A.-C., DeBilly, G., Wang, P. & Darlix, J.-L. (1989) *J. Mol. Biol.* **205,** 363–372.
14. Becerra, S. P., Rose, J. A., Hardy, M., Baroudy, B. M. & Anderson, C. W. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 7919–7923.
15. Curran, J. & Kolakofsky, D. (1988) *EMBO J.* **7,** 245–251.
16. Melton, D. A., Krieg, P. A., Rebagliati, M. R., Maniatis, T., Zinn, K. & Green, M. R. (1984) *Nucleic Acids Res.* **12,** 7035–7056.
17. Kozak, M. (1990) *Nucleic Acids Res.* **18,** 2828.
18. Kozak, M. (1989) *Mol. Cell. Biol.* **9,** 5134–5142.
19. Kozak, M. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2850–2854.
20. Dasso, M. C. & Jackson, R. J. (1989) *Nucleic Acids Res.* **17,** 6485–6497.
21. Kawasaki, E. S., Ladner, M. B., Wang, A., Van Arsdell, J., Warren, M. K., Coyne, M. Y., Schweickart, V., Lee, M.-T., Wilson, K. J., Boosman, A., Stanley, E. R., Ralph, P. & Mark, D. F. (1985) *Science* **230,** 291–296.
22. Lopez, J. A., Chung, D. W., Fujikawa, K., Hagen, F. S., Papayannopoulou, T. & Roth, G. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 5615–5619.
23. Qian, S. W., Kondaiah, P., Roberts, A. B. & Sporn, M. B. (1990) *Nucleic Acids Res.* **18,** 3059.
24. Plowman, G. D., Green, J. M., McDonald, V. L., Neubauer, M. G., Disteche, C. M., Todaro, G. J. & Shoyab, M. (1990) *Mol. Cell. Biol.* **10,** 1969–1981.
25. Kozak, M. (1988) *Mol. Cell. Biol.* **8,** 2737–2744.
26. Donahue, T. F. & Cigan, A. M. (1988) *Mol. Cell. Biol.* **8,** 2955–2963.
27. Clements, J. M., Lax, T. M. & Sherman, F. (1988) *Mol. Cell. Biol.* **8,** 4533–4536.
28. Cigan, A. M., Pabich, E. K. & Donahue, T. F. (1988) *Mol. Cell. Biol.* **8,** 2964–2975.
29. Baim, S. B. & Sherman, F. (1988) *Mol. Cell. Biol.* **8,** 1591–1601.
30. Peabody, D. S. (1987) *J. Biol. Chem.* **262,** 11847–11851.
31. Giorgi, C., Blumberg, B. M. & Kolakofsky, D. (1983) *Cell* **35,** 829–836.
32. Semba, K., Nishizawa, M., Miyajima, N., Yoshida, M. C., Sukegawa, J., Yamanashi, Y., Sasaki, M., Yamamoto, T. & Toyoshima, K. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 5459–5463.
33. Kozak, M. (1977) *Nature (London)* **269,** 390–394.
34. Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. (1988) *Cell* **55,** 447–458.
35. Brierley, I., Digard, P. & Inglis, S. C. (1989) *Cell* **57,** 537–547.
36. Zinoni, F., Heider, J. & Bock, A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4660–4664.
37. Mita, K., Ichimura, S., Zama, M. & James, T. C. (1988) *J. Mol. Biol.* **203,** 917–925.
38. Shpaer, E. G. (1985) *Nucleic Acids Res.* **13,** 275–288.