Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8947–8951, October 1993
Biochemistry

Translation of reovirus RNA species m1 can initiate at either of the first two in-frame initiation codons

(mRNA/protein synthesis/initiation-codon usage/recombinant vaccinia virus expression system)

MICHAEL R. RONER, LISA A. RONER, AND WOLFGANG K. JOKLIK

Department of Microbiology, Duke University Medical Center, Durham, NC 27710

Contributed by Wolfgang K. Joklik, June 21, 1993

ABSTRACT **The m1 species of reovirus RNA, which encodes the minor protein component μ2, possesses two initiation codons, one "strong" according to Kozak rules and preceded by 13 residues (IC1), the other "weak" and located 49 codons downstream of the first (IC2). In reovirus-infected cells only IC2 is used, but initiation from IC1 can be activated, and efficiency of initiation from either initiation codon modulated over a wide range, by coupling unrelated sequences to either or both ends of m1 RNA. For example, when the M1 genome segment is cloned into the thymidine kinase gene of vaccinia virus in such a way that various "irrelevant" stretches of nucleotides comprising restriction endonuclease cleavage sites or promoter remnants are coupled to the 5′ end of m1 RNA, translation of the resultant transcripts is also initiated at IC2, with frequencies controlled by the nature of the attached sequences. However, in rabbit reticulocyte lysates these same transcripts are translated from IC1 as well as from IC2, and transcripts in which m1 RNA is preceded by long sequences of encephalomyocarditis virus RNA (from the T7 polymerase-controlled pTM1 vector) are translated exclusively from IC1. By contrast, m1 RNA itself is translated only from IC2. It appears that the most important factor that controls the extent to which translation is initiated from IC1 and IC2 is their "availability," which is likely to be a function of the extent to which the regions on either side of them interact with each other (and also, to a lesser extent, with the 3′ untranslated region) either directly or via interaction with host cell proteins. The effects described here are of considerable potential significance when genetic material is rearranged as a result of translocations, insertions, deletions, and amplifications—that is, when sequences that are normally separated are brought into apposition.**

The present investigation arose from efforts to obtain sizable amounts of reovirus protein μ2, a minor core component of unknown function (probably <12 molecules per core), which is encoded by the M1 genome segment. The expression system chosen was vaccinia virus strain WR into the thymidine kinase (TK) gene of which we cloned the M1 genome segment preceded by the bacteriophage SP6 promoter (inoperative in this system), under the control of the cowpox virus ATI protein gene promoter (1, 2). This is a strong late promoter that has been used to elicit the formation of large amounts of several reovirus proteins including σ1 (3) and λ2 (4), as well as λ1, μNS, σ2, σNS, and σ3 (unpublished results). However, the recombinant vaccinia virus that contained M1 elicited the formation of only very small amounts of μ2.

We therefore set out to modify the SP6 component of the 5′ untranslated region of the SP6–M1 construct in order to generate versions of it that would elicit increased levels of μ2 expression. This investigation gained additional interest and relevance in view of an unusual feature of m1 RNA translation. The m1 RNA species possesses a "strong" [according to Kozak rules (5)] initiation codon starting at residue 14 [initiation codon 1 (IC1)]; the next in-phase AUG is "weak" and is located 49 codons downstream (IC2). In reovirus-infected cells m1 RNA is translated very inefficiently, with an efficiency not exceeding 2% that of s4 RNA, the most efficiently translated reovirus RNA (6). However, in rabbit reticulocyte lysates m1 RNA is translated 4–6 times more efficiently than s1 RNA, which *in vivo* is translated about 10 times more efficiently than m1 RNA. Further, the 5′-terminal untranslated sequence of m1 is capable of increasing the translation of the s1 RNA open reading frame (ORF) as efficiently as that of s4 RNA (7).

The modified clones described above were (*a*) inserted into the TK gene of vaccinia virus under the control of the cowpox ATI protein gene promoter and (*b*) inserted into pGEM-4Z and transcribed *in vitro* with SP6 polymerase. The ability of the resultant m1 RNAs to be translated was then determined. The results were surprising: depending on the nature of the sequences that preceded its own intact 5′ untranslated region, the translation of m1 RNA was initiated either from the "strong" first AUG or from the "weak" second AUG, or from both.

MATERIAL AND METHODS

Virus. Reovirus serotype 3 strain Dearing was used. The virus was grown in mouse L929 fibroblasts and purified as described by Smith *et al.* (8). The M1 genome segment of this virus was cloned into pBR322 by Cashdollar *et al.* (9, 10) and has been sequenced (11).

Constructs. The RNA species whose translation efficiencies were to be determined *in vitro* and *in vivo* were constructed as described (3, 4, 7).

Digestion with BAL-31 Nuclease. pUC18-SP6-M1 (7) was linearized with *Eco*RI restriction endonuclease and incubated at 65°C for 10 min in BAL-31 digestion buffer. BAL-31 nuclease was then added and digestion was allowed to proceed for various periods of time (up to 10 min). Samples were taken at various times, deproteinized with phenol/chloroform, and precipated with ethanol, and their 5′ termini were sequenced (12). Three molecules of varying length (#29, #44, and #42) were selected and their 5′-terminal portions, up to the *Sal* I site at position 881 in the M1 ORF, were inserted into pGEM-4Z (Promega) at the *Sma* I site. The M1 gene was then reconstructed with the appropriate segment derived from pUC18-SP6-M1.

Construction of Recombinant Vaccinia Viruses. Recombinant vaccinia viruses with TK genes containing various M1 RNA constructs under the control of the ATI protein gene promoter were constructed as described (2–4, 13). In brief,

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.

Abbreviations: IC1 and IC2, initiation codons 1 and 2; TK, thymidine kinase; ORF, open reading frame.

DNA excised from pUC18-SP6-M1 and #29, #44, and #42 DNAs excised from pGEM-4Z were cloned into p2100 EH (2, 4). The resultant plasmids were lipofected into TK$^-$143 cells infected with vaccinia virus strain WR at a multiplicity of 0.01 plaque-forming unit per cell [25 μg of plasmid together with 25 μl of Lipofectin (GIBCO/BRL) in 2 ml of Eagle's minimal essential medium (MEM-E) (GIBCO/BRL) per 25-cm^2 flask]. Recombinant vaccinia viruses were selected and identified by standard procedures (see above).

***In Vitro* Transcription.** mRNAs were transcribed as described by Promega with SP6 RNA polymerase. The mRNAs were capped by reducing the GTP concentration and adding the cap analog m^{7}G(5')ppp(5')G (Pharmacia) to the reaction mixtures as described by Promega.

***In Vitro* Translation.** Translation was carried out in a nuclease-treated rabbit reticulocyte lysate system (Promega) as described by the supplier, employing limiting amounts of mRNAs in 7.5-μl reaction mixtures containing [^{35}S]methionine.

Labeling of Cells Infected with Reovirus. L929 cells (1.6×10^6) seeded into 60-mm dishes were incubated at 37°C overnight. The cells were then infected with reovirus at a multiplicity of 10 in 0.5 ml of MEM-E. After 8 hr the medium was removed and replaced with 2 ml of MEM-E containing EXPRE^{35}S^{35}S protein labeling mix (NEN) at 250 μCi/ml (1 μCi = 37 GBq). The cells were harvested 24 hr later, pelleted at 4000 $\times$ *g*, and suspended in 500 μl of phosphate-buffered saline (140 mM NaCl/2.7 mM KCl/8 mM Na_2HPO_4/1.5 mM KH_2PO_4, pH 7.2).

Immunoprecipitation of Protein μ2. The Nonidet P-40 cell lysis method of Gonzalez and Burrone (14) was used to prepare samples for immunoprecipitation. In brief, cells were lysed by the addition of an equal volume of 200 mM Tris·HCl, pH 8/500 mM NaCl/1% (vol/vol) Nonidet P-40. After 10 min at 4°C the lysed cells were centrifuged at 12,000 $\times$ *g* for 2 min. Antibody was added to the supernatants first, followed 60 min later by activated *Staphylococcus aureus* cells which were pelleted after 30 min and washed four times with 100 mM Tris·HCl, pH 8/500 mM LiCl/1% (vol/vol) Nonidet P-40/0.1% (wt/vol) bovine serum albumin. The samples were then incubated at 100°C for 3 min in Laemmli sample buffer (15) and analyzed by electrophoresis in 7.5% polyacrylamide gels containing SDS, followed by autoradiography.

RESULTS

The origin of this work was the attempt to create a recombinant vaccinia virus containing the reovirus serotype 3 strain Dearing M1 genome segment that would express sufficient amounts of protein μ2 to permit isolation and purification. This has been achieved for reovirus proteins such as σ1 (3), λ2 (4), λ3 (16), and λ1, μNS, σ2, σNS, and σ3 (unpublished results). Ideally, the protein in question should be expressed in amounts sufficient to yield a band stainable with Coomassie blue when extracts of cells infected with the recombinant vaccinia virus are electrophoresed in SDS/polyacrylamide gels (indicating that it corresponds to about 1% of total cellular protein).

Translation *in Vivo*. When the M1 genome segment, in the form of SP6M1 excised from pUC18-SP6-M1 (7), was cloned into the TK gene of vaccinia virus strain WR under the control of the ATI protein gene promoter as described by Mao and Joklik (4) for genome segment L2, using techniques described by Mackett *et al.* (13) and Patel *et al.* (2), the resulting recombinant vaccinia virus elicited the synthesis of only small amounts of protein μ2. Appropriate experiments showed that the problem was not the amount of RNA transcribed, but rather the efficiency with which it was translated.

In order to generate a recombinant vaccinia virus that would elicit the formation of larger amounts of μ2, it was decided to remove some of the SP6 promoter sequence upstream of the M1 ORF (which was of course inoperative in the recombinant vaccinia virus), in case it adversely affected translation efficiency. Plasmid pUC18-SP6-M1 was therefore linearized with *Eco*RI and digested with BAL-31 nuclease for various lengths of time; the resulting digestion products were then cloned into the *Sma* I site of pGEM-4Z and thence into p2100 EH, from which they were recombined into the vaccinia virus genome.

The ability of the resulting recombinant vaccinia viruses to express protein μ2 was then tested. These included viruses containing constructs from which various portions or all of the SP6 promoter had been eliminated, as well as a construct from which, in addition, all the 5' untranslated region of m1 RNA as well as the AU of the first AUG initiation codon had been removed. The sequences of the transcripts generated by these viruses are shown in Fig. 1.

The μ2 proteins synthesized in cells infected with these four recombinant vaccinia viruses are shown in Fig. 2. There were two noteworthy features. (*i*) The size of protein μ2 produced by the first three constructs, all of which contained an intact ORF, was exactly the same as that produced by #42, from which translation cannot be initiated at the first AUG and presumably starts at the second AUG, 49 codons downstream. Thus even in the presence of an intact first AUG, translation of the M1 ORF cloned into the vaccinia virus TK gene is initiated from the second AUG. (*ii*) The virus containing #29 from which most of the SP6 promoter sequence had been eliminated elicited the formation of almost 8 times more μ2 than the virus containing the SP6–M1 sequence, and the recombinant vaccinia virus containing #44 expressed half as much (Table 1). The amount of μ2 formed in cells infected with #42 was 25% higher than that formed in cells infected with the #44 construct.

Translation *in Vitro*. To examine more closely the factors that control translation of m1 RNA, the translation of the constructs described above was examined *in vitro* in rabbit reticulocyte lysates. In addition, the translation patterns of two other RNAs were examined: RNA transcribed from a construct engineered so that authentic m1 RNA was formed, and RNA transcribed from the M1 genome segment inserted

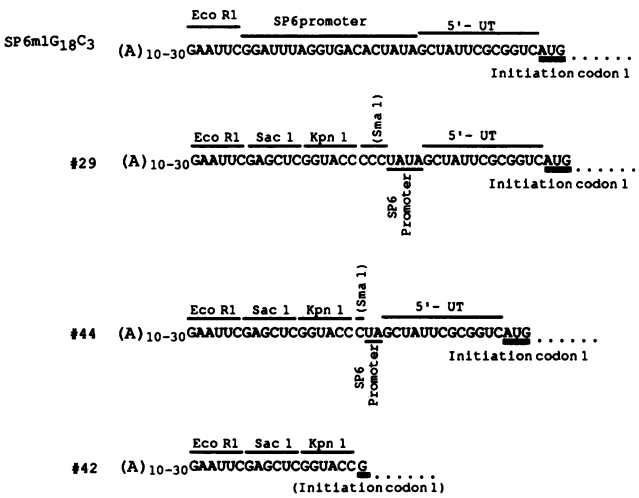

FIG. 1. Structure of the m1-containing transcripts formed in cells infected with recombinant vaccinia viruses with M1 inserts in their TK genes. Note that the *Sma* I remnant is variable. At their 3' ends these transcripts possessed the m1 untranslated region followed by an 18-residue oligo(G) tract and about 160 residues of the cowpox virus ATI protein gene terminating in a poly(A) tail as described by Antczak *et al.* (17). 5'-UT, 5' untranslated region.

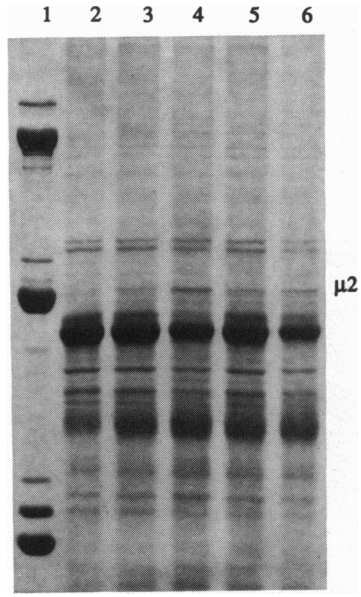

FIG. 2. The μ2 proteins synthesized in TK⁻143 cells infected with recombinant vaccinia viruses into which the SP6M1, #29, #44, and #42 DNAs (Fig. 1) had been cloned under the control of the ATI protein gene promoter. TK⁻143 cells (1.6×10^6) seeded into 60-mm dishes were incubated at 37°C overnight and then infected with the various recombinant vaccinia viruses at a multiplicity of 10. The cells were harvested 24 hr later, pelleted at $2000 \times g$, and suspended in 500 μl of phosphate-buffered saline. Lane 1, reovirus; lane 2, uninfected TK⁻143 cells; lanes 3–6, TK⁻143 cells infected with SP6m1 (lane 3), #29 (lane 4), #44 (lane 5), and #42 (lane 6) RNA. The gel was stained with Coomassie blue. The position of protein μ2 was established by Western blotting using antiserum against reovirus serotype 3.

into pTM1 in which it is under the control of the phage T7 promoter (18). The sequences of the RNAs examined are shown in Fig. 3.

When these RNAs were translated in rabbit reticulocyte lysates the following results were obtained. Translation of authentic m1 RNA and of T7m1 each yielded one protein, but they differed in size, the size difference being consistent with translation initiating at either the second or the first initiation codon. Translation of $m1G_{18}C_3$ RNA yielded both proteins, the smaller predominating with a ratio of about 3:1 (Fig. 4*a* and Table 1). Translation of #29 and #44 yielded much larger amounts of the same two proteins, with ratios of about 1:1

Table 1. Relative amounts of protein μ2 translated from the first and second initiation codons of various forms of m1 RNA

Construct	Relative translational efficiency: First initiation codon	Second initiation codon
	Translation *in vitro*	
$m1G_{18}C_3$	0.3	0.75
#29	5.0	4.8
#44	6.25	11.0
#42	0	1.7
m1	0	1.0
T7m1	1.1	0
	Translation *in vivo*	
SP6m1	0	1.0
#29	0	7.7
#44	0	3.8
#42	0	4.8

Relative translational efficiencies under normalized conditions were determined by quantitative scanning of autoradiograms. See *Materials and Methods* and refs. 6 and 7.

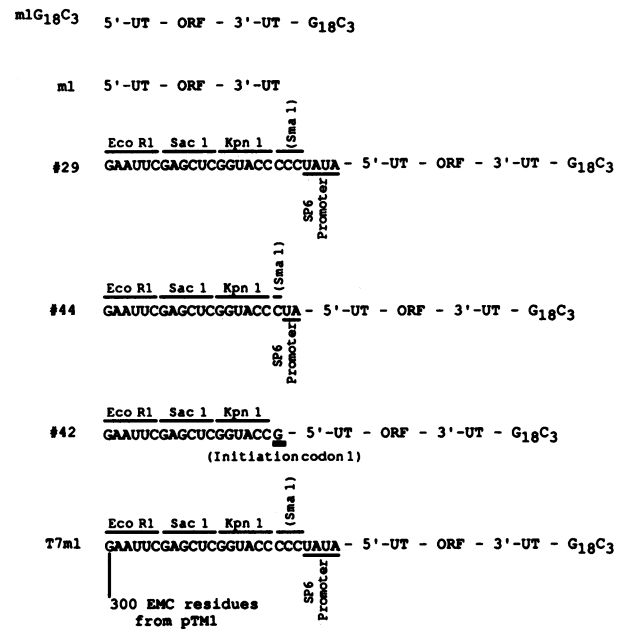

FIG. 3. Structures of the m1-containing transcripts tested *in vitro*. All except two were transcribed under the control of the SP6 promoter and were terminated by 18-residue oligo(G) tails and a remnant of a *Sma* I restriction site. The exceptions were authentic m1 RNA transcribed by reovirus cores and the RNA transcribed by T7 polymerase when the M1 genome segment was cloned into pTM1 (18). EMC, encephalomyocarditis virus; UT, untranslated region.

and 2:1 in favor of the smaller, respectively; and translation of #42 yielded only one protein, which was the same size as the smaller yielded by #29, #44, and $m1G_{18}C_3$ (Fig. 4*b* and Table 1).

Authentic m1 RNA is thus translated, in rabbit reticulocyte lysates, from the second initiation codon; when preceded by about 300 nucleotides of encephalomyocarditis virus RNA, it is translated from the first initiation codon. Further, the fact that translation of authentic m1 RNA is initiated at the second initiation codon, whereas that of $m1G_{18}C_3$ is initiated at both, is highly significant because it indicates that the nature of the 3′ untranslated region (as well as that of the 5′ untranslated region) can specify initiation codon usage.

The relative efficiencies of translation of these RNAs fell into two groups (Table 1). RNA species m1, $m1G_{18}C_3$, T7m1, and #42 were all translated with roughly equal efficiencies; RNA species #29 and #44 were translated 10–17 times more efficiently.

Nature of the μ2 Protein Formed in Cells Infected with Reovirus. The crucial question arising out of these observations was the nature of the μ2 protein formed in reovirus-infected cells. Since the amount of μ2 produced in infected cells formed only a minor band in the area where the major proteins μNS, μNSC, and μ1C also formed bands, extracts of cells infected with reovirus and labeled with [^{35}S]methionine were first precipitated with monoclonal antibody 5 against μNS/μNSC and monoclonal antibody 12 against protein μ1/μ1C (19), and the supernatant was then immunoprecipitated with antiserum against reovirus serotype 3 (Fig. 5). The only μ-class reovirus protein that was precipitated was the small form of protein μ2, which is therefore the form of protein μ2 that is formed in infected cells. We recommend that it be designated protein μ2 (78,024 Da, 687 amino acids) (11) and that the form of protein μ2 whose translation is initiated at the first initiation codon be designated μ2L (83,244 Da, 736 amino acids).

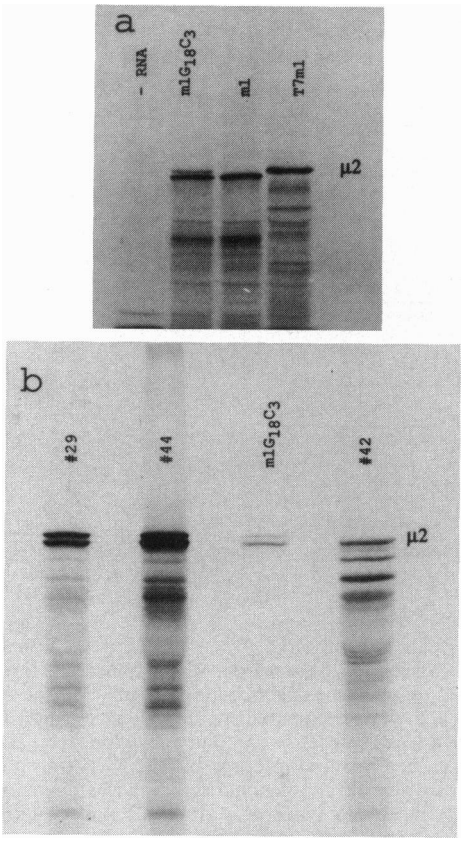

FIG. 4. Translation *in vitro* of the RNAs shown in Fig. 3. For details see *Materials and Methods* and refs. 6 and 7. The label was [^{35}S]methionine. The results of two experiments are shown (*a* and *b*). Both of the two largest proteins (in the region marked μ2) reacted with antiserum against reovirus serotype 3 in Western blotting. So did some of the smaller proteins, which are products of translation initiating at internal AUG codons, which is well known to occur in rabbit reticulocyte lysates. Note that the pattern of these smaller proteins is identical for m1$G_{18}C_3$ and m1, which have identical 5′ untranslated regions, and quite different from that for T7m1. This indicates that the proteins smaller than μ2 that react with reovirus serotype 3-specific antiserum are not breakdown products.

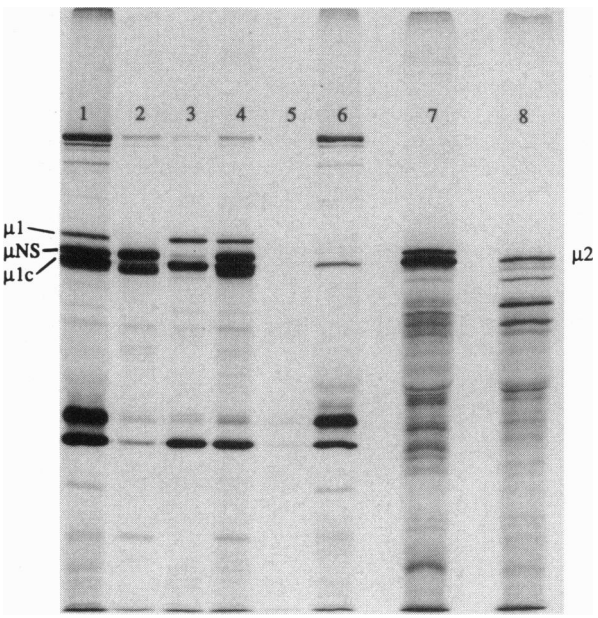

FIG. 5. Translation of the form of protein μ2 that is produced in reovirus-infected L929 cells is initiated at the second initiation codon. For details of infection, labeling, and immunoprecipitation, see text. Lane 1, L929 cells infected with reovirus; lanes 2–6, proteins precipitated from extracts of infected cells with the following antibodies/antisera: lane 2, monoclonal antibody (mAb) 5 against proteins μNS/μNSC; lane 3, mAb 12 against μ1/μ1C; lane 4, mAb 5 plus mAb 12; lane 5, mAb 5 plus mAb 12 again, demonstrating completeness of removal of μ1, μ1C, μNS, and μNSC; lane 6, antiserum against reovirus serotype 3 after two successive precipitations with mAb 5 plus mAb 12 [which, in addition to μ2, also precipitates the λ proteins, σNS, and σ3 associated with σNS (17)]. Lanes 7 and 8 provide size markers: #29 and #42, respectively, translated *in vitro* as illustrated and described in Fig. 4.

DISCUSSION

The results presented here demonstrate that the translation of the m1 species of reovirus RNA can be initiated either at the first (IC1) or at the second (IC2) in-frame initiation codon, or at both. In reovirus-infected cells only IC2 is used, and the same is true in cells infected with recombinant vaccinia viruses expressing the M1 genome segment. *In vitro* the situation is different. Here translation of authentic m1 RNA is also initiated at IC2; but this can be modified in at least three ways. (*i*) When a $G_{18}C_3$ tail is attached to the 3′ end of m1 RNA, about one-third of the translation initiations occur at IC1. Thus, the 3′ end, as well as the 5′ end, can influence translation initiation. (*ii*) When, in addition to the $G_{18}C_3$ tail, sequences totally unrelated to any reovirus sequences precede the 5′ untranslated region of m1 RNA, the translation frequency can be greatly increased, both from IC1 and from IC2. Thus, unrelated sequences at the 5′ end can profoundly influence frequency/efficiency of initiation both from the first and from the second in-frame initiation codon. (*iii*) Unrelated sequences can modulate initiation frequency over a wide range: for example, when 300 nucleotides of encephalomyocarditis virus RNA are coupled to the 5′ terminus of m1 RNA, initiation from IC2 is completely shut down and initiation occurs exclusively at IC1.

How could such foreign sequences, located upstream of authentic 5′ untranslated sequence on the one hand and downstream of authentic 3′ untranslated sequence on the other (the $G_{18}C_3$ tails), influence where translation is initiated? This is a problem that is quite different from that dealing with the control of translation frequency/efficiency by sequences immediately adjacent to and surrounding initiation codons as investigated by Kozak (20) and many others. It is of considerable potential importance in relation to rearrangements of genetic material such as translocations, insertions, amplifications, and deletions, whenever sequences that are normally separated are brought into apposition.

Such newly "acquired" sequences located upstream of normal 5′ untranslated sequence, most likely exert their effect by interacting with sequences that surround initiation codons and also with 3′ untranslated sequence (since m1 and m1$G_{18}C_3$ RNAs exhibit different translation initiation behavior). We have carried out extensive computer modeling analyses of the putative secondary structures at the 5′ ends of the various RNA species tested here and found that all were capable of yielding highly base-paired structures involving sequences at least as far as 200 nucleotides downstream of IC2. The problem was that all these structures were of roughly equal stability and that there was no correlation between ΔG and frequency of initiation at either initiation codon. It is likely, therefore, that host proteins are involved in the mechanisms by which these sequences exert their effects. These proteins are unlikely to be species-specific because infection of mouse (L929 and OST7-1), rabbit (RK13), hamster (BHK), monkey (Vero), and human (143B) cells with the recombinant vaccinia viruses described above always yield the same protein with initiation at IC2. Thus the

reason why translation of the same transcripts initiates at IC2 in mouse L fibroblasts but at IC1 as well as at IC2 in rabbit reticulocyte lysates is not the species difference of any putative interacting cellular proteins, but rather the fact that the ability of cellular proteins to interact with such RNA structures is likely to be different *in vitro* and *in vivo*.

Finally, how might the protein-stabilized interaction of sequences surrounding initiation codons with upstream sequences regulate where translation is initiated? The most likely mechanism is by controlling how tightly the regions surrounding initiation codons are involved in base pairing—that is, by controlling initiation-codon accessibility. The nature of the conformational features of such base-paired RNA structures (which would be unique to each mRNA species) and of the host proteins that might be involved in these interactions remains to be elucidated.

This work was supported by Research Grant R01 AI08909 from the National Institutes of Health and Research Grant 3233 from the Council for Tobacco Research.

1. Patel, D. D., Pickup, D. J. & Joklik, W. K. (1986) *Virology* **149**, 174–189.
2. Patel, D. D., Ray, C. A., Drucker, R. P. & Pickup, D. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9431–9436.
3. Banerjea, A. C., Brechling, K. A., Ray, C. A., Erikson, H., Pickup, D. J. & Joklik, W. K. (1988) *Virology* **167**, 601–612.
4. Mao, X. & Joklik, W. K. (1991) *Virology* **185**, 377–386.
5. Kozak, M. (1981) *Nucleic Acids Res.* **9**, 5233–5252.
6. Gaillard, R. K., Jr., & Joklik, W. K. (1985) *Virology* **147**, 336–348.
7. Roner, M. R., Gaillard, R. K., Jr., & Joklik, W. K. (1989) *Virology* **168**, 292–301.
8. Smith, R. E., Zweerink, H. J. & Joklik, W. K. (1969) *Virology* **39**, 791–810.
9. Cashdollar, L. W., Esparza, J., Hudson, G. R., Chmelo, R., Lee, P. W. K. & Joklik, W. K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 7644–7648.
10. Cashdollar, L. W., Chmelo, R., Esparza, J., Hudson, G. R. & Joklik, W. K. (1984) *Virology* **133**, 191–196.
11. Wiener, J. R., Bartlett, J. A. & Joklik, W. K. (1989) *Virology* **169**, 293–304.
12. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
13. Mackett, M., Smith, G. L. & Moss, B. (1984) *J. Virol.* **49**, 857–864.
14. Gonzalez, S. A. & Burrone, O. R. (1991) *Virology* **182**, 8–16.
15. Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685.
16. Starnes, M. C. & Joklik, W. K. (1993) *Virology* **193**, 356–366.
17. Antczak, J. B., Patel, D. D., Ray, C. A., Ink, B. S. & Pickup, D. J. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 12033–12037.
18. Moss, B. (1991) *Science* **252**, 1662–1667.
19. Lee, P. W. K., Hayes, E. C. & Joklik, W. K. (1981) *Virology* **108**, 134–146.
20. Kozak, M. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 385–402.