The EMBO Journal vol.9 no.11 pp.3753-3759, 1990

Initiation of encephalomyocarditis virus RNA translation: the authentic initiation site is not selected by a scanning mechanism

Ann Kaminski, Michael T.Howell and Richard J.Jackson

Department of Biochemistry, University of Cambridge, UK

Communicated by T.Hunt

The authentic initiation site on encephalomyocarditis virus (strain R) RNA is the 11th AUG codon (at nt. 834) from the 5′ end of the viral RNA, the downstream of the two AUGs in the sequence . . . .ACGAUGAUAAUAUG-GCC. . . In order to assess the role of ribosome scanning in the selection of the correct initiation site, transcripts of a construct comprising the viral 5′ noncoding sequence fused to a reporter gene were translated *in vitro* and the relative frequency of initiation at these two AUG codons, AUG-10 and AUG-11, was assayed. When deletions from the 5′ end were made that retained less than 100 nucleotides of the viral 5′-noncoding sequence, initiation was highly cap-dependent and AUG-10 was utilized in preference to AUG-11, consistent with the scanning ribosome model. On the other hand, when the constructs included the entire 574 nt. segment of the 5′ noncoding region situated between the poly(C) tract of the virion RNA and nt. 834, initiation was cap-independent and occurred exclusively at AUG-11, with negligible use of AUG-10. These results suggest that structural features of the 574 nucleotide segment cause initiating ribosomes to bind directly to AUG-11, without scanning the immediate upstream sequences that include AUG-10.

Key words: encephalomyocarditis virus/internal entry/scanning ribosome/translation initiation

Introduction

It has recently been shown that when part of the 5′ untranslated region of picornaviral RNAs, either from encephalomyocarditis virus (EMCV) or poliovirus, is placed in the intercistronic region of a dicistronic mRNA it confers independent translation on the downstream cistron in that the product of the downstream cistron is made in higher yield and may appear earlier than that of the upstream cistron (Pelletier and Sonenberg, 1988; Jang *et al.*, 1988, 1989). This independent translation of the downstream cistron required a substantial segment of the picornaviral 5′ untranslated region: current evidence indicates that the necessary sequence of EMCV RNA extends from about nt. 260 to nt. 834 (Jang *et al.*, 1988, 1989), whilst that of poliovirus RNA is from about nt. 140 to nt. 630 (Pelletier and Sonenberg, 1988). These observations have been interpreted as showing initiation on the picornaviral RNA segment via an internal entry mechanism that does not involve linear scanning from the physical 5′ end of the RNA (Jackson, 1988; Jang *et al.*, 1988; Pelletier and Sonenberg, 1988). However, the details of how this internal entry mechanism might operate and the reasons why it should require several hundred nucleotides of picornaviral RNA sequence are obscure. In the case of poliovirus, it was originally proposed that ribosomes bind first to a particular sequence centrally located within the picornavirus RNA segment, possibly a sequence devoid of secondary structure, and are then translocated, probably by a scanning mechanism, from this binding site to the initiation codon (Pelletier and Sonenberg, 1988). This model has the attraction of showing some features in common with the conventional scanning ribosome model (Kozak, 1989). The problem, however, is that picornaviral 5′ noncoding regions have many apparently silent AUG codons and a high degree of (predicted) secondary structure, two features which would be expected to inhibit the scanning ribosome from reaching the correct initiation site.

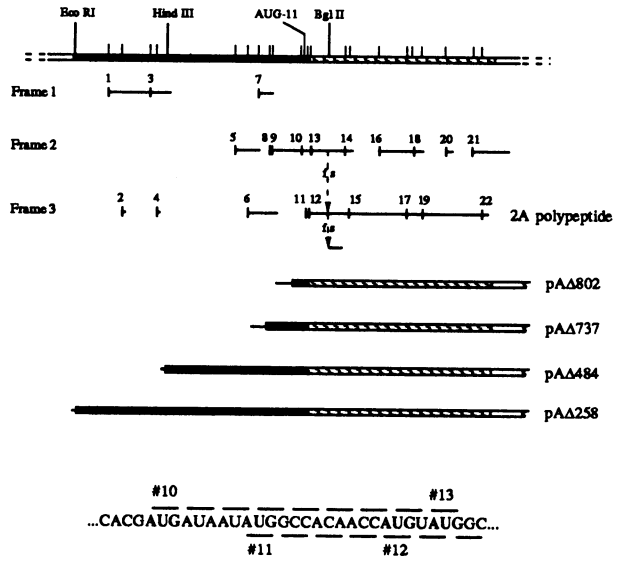
In the case of EMCV (strain R) RNA, the authentic initiation site at nt. 834 is the eleventh AUG codon downstream from the poly(C) tract. The disposition of the upstream AUG codons, particularly AUG-10, makes this a particularly good system to test whether ribosomes select the correct initiation site by scanning from an upstream entry site. AUG-10 is situated a short distance upstream of the correct site (AUG-11) but in a different reading frame (Figure 1). As the local sequence context of AUG-10 is quite favourable for initiation (. . .ACGAUGA. .), one might expect a scanning ribosome to initiate at AUG-10 in preference to AUG-11. In this study we assay the use of both of these two AUG codons on mRNAs with various lengths of EMCV 5′ untranslated sequences. We show that when the leader is short, and initiation operates by the conventional cap-dependent scanning mechanism, AUG-10 is used in preference to AUG-11, as expected. However, when the 5′-noncoding segment consists of the 574 nucleotide segment of EMCV RNA previously shown to allow internal initiation when inserted in the intercistronic region of a dicistronic mRNA (Jang *et al.*, 1988, 1989), initiation is cap-independent and is exclusively at AUG-11 rather than AUG-10. This observation eliminates any model which involves selection of AUG-11 by strictly linear scanning of the RNA from an upstream entry site.

Results

Strategy

The starting construct used in this work was pS32A3 of Jang *et al.* (1988), which was renamed pAΔ258 for this work (to conform with the standard nomenclature of our plasmids according to the length of EMCV sequences deleted from the 5′ end of the virion RNA). This is a monocistronic derivative consisting of nt. 259–848 of EMCV RNA followed by the coding region of poliovirus polypeptide 2A and then two tandem stop codons (Figure 1), in a vector carrying the bacteriophage T7 promoter. The authentic EMCV initiation codon (at nt. 834 of the virion RNA) is the 11th

AUG from the 5′ end of the RNA transcribed from this clone. Initiation at this site should produce a polypeptide consisting of poliovirus 2A with an N-terminal extension of five amino acids corresponding to the N-terminus of the EMCV polyprotein. Initiation at AUG-12 would give a polypeptide lacking four amino acids of this extension, but translation from AUG-10 or AUG-13 would give only short products (Figure 1).



Length of translation product (amino acid residues)

Initiation site	- frame-shift	+ frame-shift
AUG-10	39	159
AUG-11	155	30
AUG-12	151	26
AUG-13	31	151

Fig. 1. Diagram of the plasmid constructs. The upper line represents the relevant segment of the parent construct, pAΔ258, with EMCV sequences depicted in black, poliovirus 2A coding sequences in cross-hatching, and with vertical bars denoting the AUG codons. The three lines below this show the ORFs present in each of the three frames, with AUG codons denoted by vertical bars and numbered consecutively from the 5′ end. The broken lines designated 'fs' show the switch in reading frames caused by in-fill of the *Bgl*II site. The lower part of the figure depicts the deletion mutants studied (pAΔ484, pAΔ737, pAΔ802), with the EMCV and poliovirus 2A sequences depicted as above, and transcribed polylinker sequences symbolized by single lines. The nomenclature of these deletion derivatives is based on the number of residues of the EMCV RNA sequence deleted (from the 5′ end of the virion RNA sequences): the authentic initiation site, AUG-11, is at nt. 834 of the virion RNA sequence. The nucleotide sequence between AUG-10 and AUG-13 is shown at the bottom, and the sizes of the translation products from the four AUG codons with or without the frame-shift are tabulated.

We wished to assess the utilization of AUG-10 (and AUG-13) without perturbing the sequence or structure of the RNA in this region with four closely spaced AUG codons. This was achieved by taking advantage of a unique *Bgl*II site situated some 60 residues downstream of AUG-11 (Figure 1). Digestion of the plasmids with *Bgl*II followed by in-filling and religation engineered a frame-shift in the 2A coding region with the outcome that initiation at AUG-10 or AUG-13 would now give full-length translation products (Figure 1). Constructs bearing the frame-shift are denoted by the suffix 'fs', e.g. pAΔ258fs.

Besides the parent plasmid, we made three constructs with different lengths deleted from the 5′ end of the EMCV RNA sequence (Figure 1), each being prepared with or without the frame-shift-inducing *Bgl*II in-fill. Each plasmid was transcribed to generate either capped or uncapped RNAs, thus generating sixteen different RNA preparations: four different lengths of 5′ untranslated EMCV RNA leader, each with and without the frame-shift in-fill, and each in capped and uncapped form. These sixteen preparations were translated over a wide range of mRNA concentrations, assayed for total [^{35}S]methionine incorporation and the translation products analysed by gel electrophoresis and autoradiography using a 5−30% polyacrylamide gradient gel which resolved translation products arising from initiation at AUG-10, AUG-11 and AUG-13 (Figure 2). The yield of the different products was assayed by scanning densitometry of different exposures of the film obtained from a similar gel except that the samples showing the highest incorporation were diluted prior to electrophoresis so that all the bands were more comparable in intensity. Products initiated at AUG-11 and AUG-12 (155 and 151 amino acid residues respectively) could not be resolved sufficiently to allow reliable determination of their individual yields, and they are therefore considered as one product initiated at AUG-11.

As our approach depends on a comparison of the translation efficiency and the products synthesized from different RNA preparations, we not only checked the integrity of all RNA samples, but also studied at least two, and in some cases up to five, independent preparations of each RNA

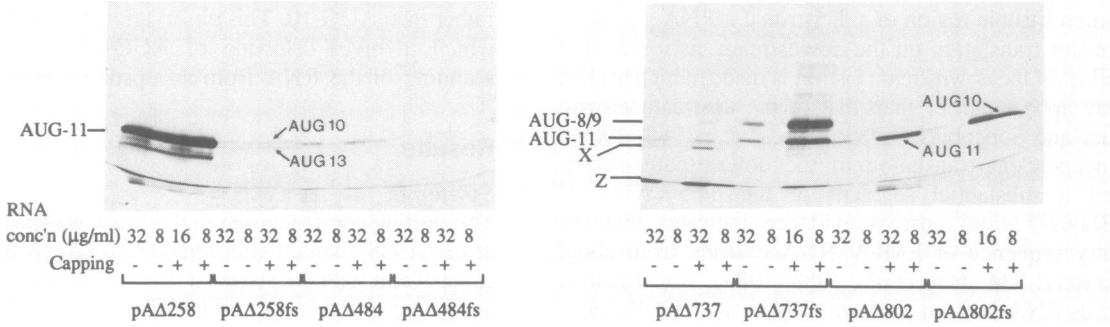


Fig. 2. Products of translation of deletion and frame-shift derivatives. Capped and uncapped mRNA derivatives were translated for 90 min at 4, 8, 16 and 32 μg/ml. Samples corresponding to 1 μl translation assay mix were separated by electrophoresis on 5−30% polyacrylamide gradient gels, which were dried and autoradiographed. (The use of a gradient gel is the cause of the pronounced curvature of the figure.) The figure shows the results obtained at two different RNA concentrations, 8 μg/ml and 32 μg/ml except in those cases where high RNA levels were strongly inhibitory, when the results of the assay at 16 μg/ml are shown. Products are designated according to the initiating AUG codon. The nature of bands X and Z is discussed in the text.

species. A summary of the results of a typical analysis of the sixteen RNA preparations is shown in Table I, which presents the ^{35}S incorporation observed at 8 μg/ml RNA and at 32 μg/ml, except in those cases where high concentrations of RNA inhibited significantly, in which case the data for 16 μg/ml are shown. The yield of products from the different AUG codons is tabulated relative to the use of AUG-11 in the full-length parent uncapped pAΔ258 construct assayed at 8 μg/ml.

Relative efficiency of initiation at AUG-10 and AUG-11

The results show that RNAs with the longest length of EMCV 5′ untranslated RNA (pAΔ258) were translated efficiently with virtually all initiation at AUG-11. The frame-shifted version of this construct (pAΔ258fs), gave very low yields of two products which were assigned on the basis of size to initiation at AUG-10 and AUG-13. Sub-optimal levels of Mg^{2+} inhibited initiation at AUG-10, without causing an absolute increase in the frequency of initiation at AUG-13, while Mg^{2+} concentrations above the optimum had the converse effect (data not shown). The conclusion that low Mg^{2+} concentrations cause a slight shift in the preference of initiation sites towards downstream sites is supported by analyses of the products of translation of pAΔ258 using a 24−30% gradient gel, which partially resolves the products initiated at AUG-11 and AUG-12. At Mg^{2+} concentrations above the optimum, only the product initiated at AUG-11 was seen, but at sub-optimal Mg^{2+} levels there appeared to be two bands, in roughly equal yield, corresponding to products from AUG-11 and AUG-12 (data not shown).

Although capping slightly increased the yield of specific products from pAΔ258 mRNA (Table I), the maximum stimulation (17%) was far less than the >7-fold stimulation resulting from capping of some of the shorter RNAs discussed below. We consider it unlikely that the small stimulation seen with the pAΔ258 RNAs reflects a true cap-dependence of the actual initiation process, and it could well arise from effects such as an influence of capping on the stability of the mRNA during the assay. We will therefore classify any translation where capping the mRNA causes <50% stimulation as 'cap-independent'.

Constructs in which the length of EMCV untranslated region (UTR) was reduced to 349 nucleotide residues (pAΔ484) were translated very inefficiently indeed in agreement with the results of Jang *et al.* (1988). Such translation as occurred was cap-independent and was assignable to initiation at AUG-11 on pAΔ484, and AUG-10, but not AUG-13, on pAΔ484fs (Figure 2 and Table I).

The next stage of truncation reduced the EMCV UTR sequences to a length of 96 residues (pAΔ737), which retains two upstream AUG codons (AUG-8 and AUG-9) in addition to the four (nos. 10−13) discussed previously (Figure 1). Translation of these mRNAs was highly cap-dependent, particularly with pAΔ737fs. Initiation on this mRNA was assigned to AUG-8/9 and AUG-10 on the basis of product size. (AUG-8 and AUG-9 are in the same reading frame only nine nucleotides apart, and so products of initiation at these two sites are unlikely to be resolved.) According to the published nucleotide sequence (Palmenberg *et al.*, 1984), products initiated at AUG-8/9 should terminate just before AUG-10, but on sequencing our clones we find a run of seven consecutive A residues at a point 60 nt. upstream of AUG-11 where the published sequence has only six, which puts AUG-8/9 in the same unbroken reading frame as AUG-10, in agreement with the results of our translation assays (Figure 2). (We do not know whether this is a discrepancy in the original sequence, or whether the extra A residue has arisen by replicative stuttering during propagation of the plasmids in *Escherichia coli*, but we have found the extra residue in all our constructs.) Clearly AUG-8/9 were the preferred initiation sites in this RNA, but some initiation also occurred at AUG-10, perhaps because the poor contexts of AUG-8 (. . .CACAUGC. . .) and AUG-9 (. . . .UACAUGU. . .) caused some ribosomes to scan past these two codons and reach AUG-10 (Kozak, 1986, 1989).

Table I. Efficiency of overall translation and of utilization of specific initiation sites

Clone used for mRNA synthesis	Capping	RNA concentration (μg/ml)	Incorporation (c.p.m./μl)		Relative yield of labelled products initiated at[a]:		
			− frame shift	+ frame shift	AUG-10	AUG-11[b]	Others
pAΔ258	−	8	64 046	2283	0.9	100.0	AUG-13: 0.7
	−	32	80 201	2468	1.3	81.7	AUG-13: 1.0
	+	8	69 985	977	1.3	117.1	AUG-13: 1.0
	+	32	80 226[c]	984	1.4	90.1[c]	AUG-13: 1.2
pAΔ484	−	8	5465	30	0.1	1.0	
	−	32	5190	1373	0.5	2.3	
	+	8	1314	2389	0.2	1.7	
	+	32	3043	2767	0.5	2.5	
pAΔ737	−	8	2050	2441	0.7	0.5	AUG-8/9: 0.8 X: 0.2
	−	32	18 248	20 880	7.9	2.3	AUG-8/9: 9.8 X: 1.2
	+	8	12 184	105 210	42.6	1.1	AUG-8/9: 98.1 X: 1.2
	+	32	91 480	140 790[c]	56.3[c]	4.3	AUG-8/9: 128.4[c] X: 5.5
pAΔ802	−	8	1369	1176	0.3	0.5	
	−	32	2576	6490	2.2	5.5	
	+	8	23 415	24 460	28.1	28.8	
	+	32	39 796	56 330[c]	61.4[c]	27.0	

[a]Determined by scanning densitometry of autoradiograms similar to Figure 2 but with different loadings and different exposures, and expressed relative to the yield of product synthesized from AUG-11 of pAΔ258 RNA at 8 μg/ml. The yields of products initiated at AUG-8/9, AUG-10 and AUG-13 were determined using the constructs with the frame-shift, the yields of all other products from assays using constructs lacking the frame-shift.

[b]May include some initiation at AUG-12 (see text).

[c]Determined at 16 μg/ml RNA because incorporation was strongly inhibited at 32 μg/ml.

The corresponding mRNA without the frame-shift in-fill (i.e. pAΔ737) was translated much less efficiently. Two products of about the size of 2A were produced in low yield. The larger was assigned to initiation at AUG-11, on the basis of its size. The identity of the smaller (designated X in Figure 2) remains uncertain as it is definitely too small to be the product from AUG-12, but seems too large to have originated from AUG-15 (Figure 1). However, most of the [^{35}S]methionine incorporation using the mRNA lacking the frame-shift is into a small polypeptide running at the front (designated Z in Figure 2). It seems likely that this product originates from initiation at AUG-8 (and/or AUG-9) and AUG-10, as was seen with pAΔ737fs, but which in this case, in the absence of the frame-shift, terminates within the first part of the 2A coding sequence (Figure 1).

The ultimate truncations of the EMCV 5′ UTR removed both AUG-8 and AUG-9. The translation of these RNAs (pAΔ802) was highly stimulated by capping, and pAΔ802fs was translated slightly more efficiently than pAΔ802. Thus, with this shortened 5′ leader, initiation occurs at both AUG-10 and AUG-11, with the use of the former predominating. This rather poor discrimination between two neighbouring AUG codons is reminiscent of the influenza B virus NA/NB mRNA where initiation takes place at two closely spaced AUG codons situated seven nucleotides apart in different reading frames (Williams and Lamb, 1989). The interesting feature of our constructs is that there is clearly much less discrimination between AUG-10 and AUG-11 when the 5′ UTR is the shorter length of pAΔ802 than the longer length in pAΔ737.

The results described above indicate two different modes of translation of the RNA depending on the length of the 5′ UTR originating from EMCV RNA. With the two shortest leaders tested, translation is highly cap-dependent, which implies the operation of the conventional scanning mechanism, and occurs at both AUG-10 and AUG-11, with the former predominating. This result shows that there are no peculiarities about AUG-10 that prevent its recognition by the scanning ribosome in preference to AUG-11. However, when the full EMCV 5′ UTR segment necessary for internal initiation is present, we see efficient cap-independent initiation almost exclusively at AUG-11, with absolutely minimal use of AUG-10 or AUG-13. We conclude that in this case, ribosomes select AUG-11 without scanning the immediate upstream segment which includes AUG-10.

Our results also re-emphasize the critical point that the efficient and exclusive use of AUG-11 in pAΔ258 cannot arise from translation of fragmented degradation products, since none of the shorter uncapped mRNAs we have tested, which serve as mimics of the partial degradation products of this mRNA, showed a product pattern and yield resembling that obtained from pAΔ258.

Scanning from the 5′ end of the EMCV noncoding region is very inefficient

Although these results show that scanning is not involved in the initiation at AUG-11 on mRNA transcribed from pAΔ258, the question arises as to the fate of those ribosomes which do scan from the 5′ end of this mRNA. To examine this we constructed pNS-H3, in which the influenza virus NS cDNA (lacking the whole 5′ UTR and the first 21 nt. of the coding segment) was inserted at the *Hin*dIII site of pAΔ258, so that it would be in-frame with AUG-1 and AUG-3 to act as a reporter for initiation at these two sites (Figure 1). As a control, the identical fragment of the NS cDNA was inserted at the same site but in reverse orientation (pSN-H3). Another control was pNS-R1, in which the entire NS cDNA (complete with its 5′ UTR) is inserted in the correct orientation at the *Eco*RI site of pAΔ258 (Figure 1), placing it upstream of the whole EMCV segment. Capped and uncapped transcripts of these constructs were translated over a wide range of RNA concentrations. Representative results are shown in Figure 3.

When the NS gene was placed upstream of the EMCV segment (pNS-R1), it was translated in a cap-dependent manner (Figure 3, lanes 2 and 3), and at an efficiency comparable with the translation of monocistronic NS mRNA (data not shown). In contrast, the efficiency of translation of the downstream cistron coding for polypeptide 2A was not influenced by capping, nor was it affected by the presence of the NS ORF at the extreme 5′ end (Figure 3, lanes 1−3). It should be noted that the 2A polypeptide (six methionine residues) from the downstream cistron was synthesized in higher yield than the NS protein (nine methionine residues) from the upstream cistron (Figure 3, lanes 2 and 3).

Insertion of the NS coding segment in the middle of the EMCV 5′ UTR (pNS-H3) drastically reduced overall translation (Figure 3, lanes 4 and 5), but on prolonged exposure of the autoradiogram a number of products were detected (Figure 3, lanes 9 and 10). The two largest products were assigned on the basis of size to initiation at AUG-1 and AUG-3, which are both in-frame with the NS coding sequences. Initiation at these two sites is likely to be via the scanning mechanism, as the synthesis of these products was stimulated by capping, particularly in the case of initiation at AUG-1. The use of both AUG codons is likely to be related to the fact that AUG-1 has a very poor context (UAUAUGU), so that a fair proportion of scanning

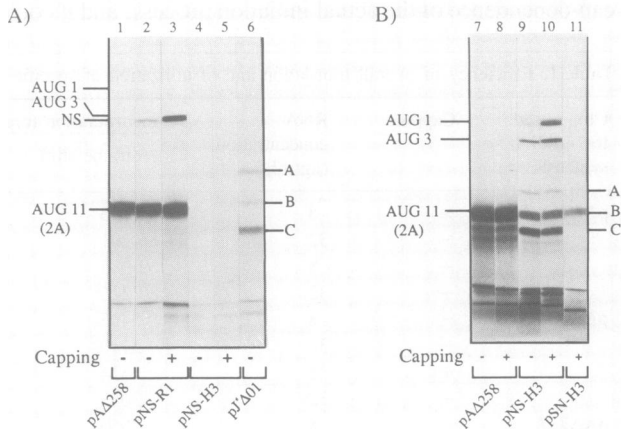


Fig. 3. Products of translation of influenza virus NS cDNA/pAΔ258 fusion constructs. Capped or uncapped transcripts were translated for 90 min at 32 μg/ml (except lane 1, 16 μg/ml). Samples corresponding to 1 μl of translation assay were analysed on a 20% polyacrylamide gel, which was dried and autoradiographed. The products of initiation at AUG-1, AUG-3 and AUG-11 (polypeptide 2A) of pAΔ258 (Figure 1) are shown, as is the full-length influenza virus NS protein (NS). The nature of the smaller products A, B and C translated from internal initiation sites in the NS cDNA sequence is discussed in the text. The exposure of the autoradiogram shown in panel B was 5-fold longer than for panel A, in order to reveal products synthesized in low yield from pNS-H3 and pSN-H3.

ribosomes might be expected to by-pass this codon and continue scanning to AUG-2 (GCAAUGU) or AUG-3 (GGAAUGC).

The combined yield of the products initiated at AUG-1 and AUG-3 in translations of pNS-H3 mRNA was much lower than the yield of NS polypeptide from pNS-R1 (Figure 3, lanes 2–5). This suggests that scanning does not penetrate effectively as far as AUG-3, but an alternative explanation is that AUG-3 is not recognized efficiently by the scanning ribosomes, either because of the mediocre context of this AUG codon or because translation of the short ORF following AUG-2 (Figure 1) interferes with initiation at AUG-3. If a high proportion of ribosomes scan past AUG-3 without initiating at this site, we should expect to see translation products initiated at sites within the NS ORF. Figure 3 (lane 6) shows a control designed to reveal such products by translating the (uncapped) transcript of an NS cDNA derivative from which the authentic initiation site has been deleted (pJ′Δ01). In this case, and also with constructs in which the initiation site is retained but a nearby frame-shift causes termination of translation 22 codons downstream, three products are routinely observed (A–C), of which C is always the most abundant and B the least. In translations of pNS-H3 mRNA, normal exposure of the autoradiogram fails to reveal significant synthesis of A–C (Figure 3, lanes 4 and 5), but they are apparent on prolonged exposure (Figure 3, lanes 9 and 10). Again, C is more abundant than A, whilst the yield of B is hard to estimate because it migrates close to the 2A polypeptide product. Two additional experiments confirmed that these products were translated from the NS ORF rather than the downstream 2A ORF: (i) a polyclonal antiserum against the NS protein precipitated these three products but not the product designated as 2A, and (ii) linearization of the pNS-H3 plasmid between the NS and 2A ORFs prior to transcription abolished the synthesis of 2A but had no effect on the production of A, B and C (data not shown).

We conclude that although some initiation does occur at internal sites within the NS ORF of pNS-H3 , the yield of such products is not in great excess over the yield of products initiated at AUG-1 and AUG-3 (Figure 3, lane 10), and is much less than the yield of NS protein translated from pNS-R1 (Figure 3, lanes 2–4). This implies that the main feature of scanning through the 5′ proximal part of the EMCV sequence of pAΔ258 is its inefficiency in penetrating as far as AUG-3, rather than any unusual 'leakiness' causing most ribosomes to by-pass AUG-3. The reason for the poor penetration of the scanning process on pAΔ258 is likely to be the high degree of secondary structure of this region of the EMCV RNA as predicted by biochemical probing experiments and phylogenetic comparisons (Pilipenko *et al.*, 1989). It should be remembered that pAΔ258 is lacking not only the poly(C) tract of EMCV virion RNA, but also the extreme 5′ proximal sequences, a segment which also has a high degree of predicted secondary structure (Vartapetian *et al.*, 1983). Scanning from the 5′ end of the uncapped virion RNA is therefore likely to be even less effective than with pAΔ258 mRNA.

Insertion of the NS cDNA in the middle of the EMCV 5′ UTR (in pNS-H3) resulted in a severe reduction in the yield of product translated from the downstream cistron coding for polypeptide 2A (Figure 3). As the 2A product migrates close to polypeptide B, initiated at an internal site in the NS ORF, the reduction was difficult to quantify precisely, but was estimated to be some twenty-fold. The related construct with the NS cDNA in reverse orientation (pSN-H3) gave only one product, 2A, synthesized at 2–5% of the yield obtained with pAΔ258 lacking the insert (Figure 3, lane 11). This reduction in the yield of 2A caused by insertion of the NS cDNA at the *Hin*dIII site in either orientation is almost certainly due to the disruption of secondary and tertiary structure elements in the EMCV 5′ UTR which are essential for internal initiation at AUG-11. In the case of poliovirus, it has been shown that very much shorter insertions (as little as four bases) in the 5′ noncoding region can drastically reduce infectivity, for reasons that seem to be due entirely to defective translation (Kuge and Nomoto, 1987; Trono *et al.*, 1988).

Discussion

The presence in EMCV (strain R) of an out-of-frame AUG codon (AUG-10) with a favourable local sequence context and situated just upstream of the authentic initiation site (Palmenberg *et al.*, 1984), provides a good test of whether ribosomes select the correct site by scanning from an upstream entry site. It should be noted that an AUG codon at or near the position of AUG-10 is not a conserved feature of cardioviruses, and is found in only two out of nine strains of the closely related species, foot and mouth disease virus (Sangar *et al.*, 1987).

That there is nothing peculiar about the local sequence context of AUG-10 which could explain why it should be overlooked by a scanning ribosome, is shown by our assays of the two most extreme truncated versions of the RNA, where initiation occurs with higher frequency at AUG-10

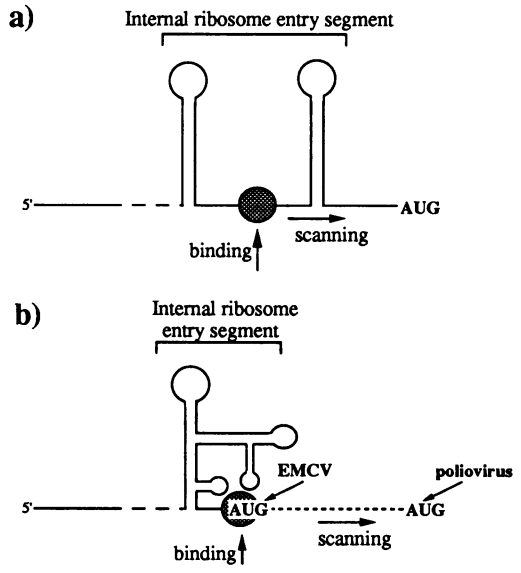


Fig. 4. Models for internal initiation on picornavirus RNAs. In model (**a**), based on that of Pelletier and Sonenberg (1988), ribosomes first bind to a potentially unstructured segment centrally located in the segment of the 5′ untranslated region necessary for internal initiation, and then scan from this entry site to the authentic initiation site. In model (**b**), based on the results described here, the ribosome entry site is located at the 3′ extremity of the segment necessary for internal initiation. Scanning from this binding site occurs on poliovirus RNA, but not on EMCV RNA where the authentic initiation site is located at the actual ribosome entry point.

rather than AUG-11, and in a cap-dependent fashion indicative that the normal scanning process is operative. If AUG-10 is utilizable when ribosomes are scanning from the 5′ end of the mRNA, but is silent when the extended EMCV 5′ UTR is present and the internal entry mechanism is operative, it follows that in the latter case the selection of AUG-11 cannot be made by a process in which the ribosome scans the mRNA in a strictly linear manner from a site upstream of AUG-10. This leaves open the possibility that ribosomes bind first to an upstream entry site and are then translocated by some non-scanning transfer process from the initial binding site to AUG-11. However, this type of transfer process is quite unprecedented and is hard to imagine unless the folding of the RNA brings the upstream binding site into close proximity to AUG-11, in which case one might envisage a mechanism with no transfer step whatsoever, but a direct binding of the ribosome to a composite site consisting of far-upstream sequences and AUG-11 located in close proximity as a result of the RNA folding (Figure 4b; Howell *et al.*, 1990).

Although the translation of poliovirus RNA is also considered to take place by a similar mechanism of internal ribosome entry (Pelletier and Sonenberg, 1988), there are some indications that scanning, or a process akin to scanning, does play a role in the selection of the correct initiation site, which is at nt. 743 in poliovirus type 1 (Kitamura *et al.*, 1981). In contrast to the very precise selection of AUG-11 and rejection of the nearby AUG-10 and AUG-13 for initiation of translation of EMCV RNA, initiation on poliovirus RNA is much more flexible in that it appears to select the first AUG codon downstream of about nt. 600, regardless of the actual position of this AUG or the preceding sequences (other than the immediate local context). Thus, a deletion from nt. 600−727 had no influence on viral infectivity (Kuge and Nomoto, 1987), and a deletion from nt. 631−733 did not affect the ability of the poliovirus 5′ UTR to promote independent initiation of the downstream cistron of a dicistronic mRNA (Pelletier and Sonenberg, 1988). Insertions in this region of the poliovirus 5′ UTR also have no influence on viral infectivity or *in vitro* translation, except in two circumstances. An insertion with a stable hairpin secondary structure motif inhibited initiation at nt. 743 *in vitro* (Pelletier and Sonenberg, 1988). In addition, an insert at nt. 702 reduced virus growth, giving a small plaque phenotype, if the insert had an AUG codon with an unfavourable local sequence context. All the large plaque pseudo-wild-type revertants carried point mutations that removed the inserted AUG codon (Kuge *et al.*, 1989). These observations all point to a mechanism of initiation site selection in which ribosomes select the first AUG codon downstream of about nt. 600, by scanning from an entry point situated upstream of nt. 600 but certainly downstream of nt. 390 (Howell *et al.*, 1990).

The models for initiation on EMCV and poliovirus RNAs could be unified by postulating that in both cases a length of several hundred nucleotides of the 5′ UTR, probably representing a particular RNA superstructure, directs the ribosomes to bind and start scanning at a particular point located at the 3′ end of this segment. In the case of EMCV RNA, this point of entry and commencement of scanning is at AUG-11, or between AUG-10 and AUG-11, so that the latter is selected with absolutely minimal scanning over only (at most) a few residues, whilst AUG-10 is silent. In the case of poliovirus RNA translation, the ribosome is directed to bind and start scanning from a point situated between nt. 390 and 600. Such a model can also accommodate the case of foot and mouth disease virus, in which two initiation sites are used (Sangar *et al.*, 1987). Given the strong sequence conservation between these viruses and EMCV (Pilipenko *et al.*, 1989), we imagine that ribosome scanning starts at or just before the upstream of the two sites, which is equivalent to AUG-11 in EMCV (strain R). As this has a mediocre local context in all foot and mouth disease virus strains except the A_{10} serotype (Sangar *et al.*, 1987), we propose that many ribosomes scan past that site and initiate at the next AUG codon situated 84 nucleotide residues downstream, which has a good local sequence context in all strains.

Our results, and the implications discussed above, put a new perspective on how internal initiation on picornaviral RNAs might operate. An implicit feature of previous ideas was a relatively small internal ribosome binding site centrally located within the long picornaviral 5′ noncoding region, from which the ribosomes were translocated, probably by a scanning process, to the authentic initiation site (Figure 4a; Pelletier and Sonenberg, 1988). Our work suggests that a very large RNA structure, or superstructure, directs a rather precise binding of ribosomes to a point situated at the 3′ extremity of this structure and presumably determined by its proximity to the critical structural motifs (Figure 4b). In the case of EMCV and related viruses, this ribosome entry or binding site is the AUG initiation codon itself, or is not more than ~5 residues upstream, so that initiation site selection requires very little, if any, scanning.

Materials and methods

Construction of plasmids

The starting plasmid was pS32A3 of Jang *et al.* (1988) renamed pAΔ258 for this work. pAΔ258 DNA was linearized with *Bgl*II, in-filled using the Klenow fragment of DNA polymerase I, and religated to produce pAΔ258 fs.

pAΔ484 and pAΔ484fs were made from pAΔ258 and pAΔ258fs, respectively, by removal of the *Eco*RI−*Hin*dIII fragment, in-filling and ligating.

pAΔ737 was made by digesting pAΔ258 with *Apa*LI, in-filling all fragments, and gel purifying the 898 bp fragment, which was ligated with *Sma*I cut pGEM-1 (Promega). A clone with the correct orientation with respect to the SP6 polymerase promoter was selected. pAΔ737fs was created from pAΔ484fs in a similar way.

The insert of pAΔ737 was excised by digestion with *Eco*RI and *Hin*dIII, gel purified and ligated with *Eco*RI/*Hin*dIII cut pGEM2, to create a clone with the correct orientation with respect to the T7 polymerase promoter. This was used to make pAΔ802, by first digesting with *Hin*dIII to cut between the promoter and the insert, followed by digestion with *Bal*31 for various times, and finally *Eco*RI digestion to release the truncated insert. Inserts judged to have been resected to a suitable degree were ligated with *Sma*I/*Eco*RI cut pGEM-2. A number of the resulting clones were sequenced, and from these pAΔ802 was selected. pAΔ802fs was constructed from pAΔ802 in exactly the same way as pAΔ258fs was made from pAΔ258.

The influenza virus NS cDNA−pAΔ258 fusion plasmids were made using pJ10 and pJ12: pJ12, described previously (Dasso and Jackson, 1989b), has the entire NS cDNA of strain A/PR/8/34 inserted in the *Eco*RI site of pGEM-2, and pJ10 is the reverse orientation of the same construction. pNS-R1 was made by digesting pJ12 with *Eco*RI, gel purifying the 890 bp representing the entire NS cDNA, ligating this into pAΔ258 cut at the *Eco*RI site (Figure 1), and selecting a clone with the correct orientation with respect to the T7 promoter. To construct pNS-H3 and pSN-H3, pJ10 was digested with *Hin*dIII, and the ends in-filled using DNA polymerase I Klenow fragment, before gel purification of the 884 bp fragment representing the whole vector polylinker, the entire cDNA 3′ noncoding region, and the coding region minus the 5′ proximal 21 nt. This was ligated with pAΔ258 which had been cut with *Hin*dIII (Figure 1) and in-filled with DNA

polymerase 1 Klenow fragment. A clone (pNS-H3) with the NS ORF in the correct orientation, in-frame with AUG-1 and AUG-3, and a clone with the insert in reverse orientation (pSN-H3) were selected. Plasmid pJ'Δ01, in which the 5' noncoding segment and the first 21 nt. of the coding region of the NS cDNA have been deleted, was constructed by digesting pJ12 with *Hin*dIII, gel purifying the large fragment, and religating.

Plasmids were propagated by standard methods (Maniatis *et al.*, 1982), without chloramphenicol amplification, using *E.coli* TG1. Plasmid DNA for use in transcription reactions was purified by digestion with RNase A, followed by proteinase K digestion, extraction with phenol/chloroform and ethanol precipitation in the presence of 2 M ammonium acetate.

The sequence of the region immediately downstream of the T7 promoter, or SP6 promoter where appropriate, of each construct (corresponding to the 5' proximal ~150 nucleotides of the mRNA) was determined essentially as described by Hattori and Sakaki (1986), using a primer complementary to the T7 or SP6 promoter (Promega).

Transcription reactions

Plasmid DNAs for use in transcription reactions were completely linearized using either *Dra*I (pAΔ258, pAΔ258fs, pAΔ484, pAΔ484fs, pNS-R1, pNS-H3, pSN-H3), or *Eco*RI (pAΔ737, pAΔ802 and pAΔ802fs), or *Hin*dIII (pAΔ737fs), followed by extraction with phenol then with chloroform, and ethanol precipitation.

Both capped and uncapped transcripts were produced from each plasmid. Transcription reactions (usually 0.05 ml) for the production of uncapped transcripts were as follows: 40 mM Tris−HCl, pH 8.0; 15 mM $MgCl_2$; 1 mM ATP, CTP, GTP, UTP; 40 μCi/ml [α-^{32}P]UTP; 5 mM dithiothreitol; 2000 U/ml RNasin; 80 μg/ml linearized template DNA; 50 μl/ml T7 RNA polymerase (activity >40 U/μl). The reaction was incubated at 37°C for 90 min. Transcription reactions using SP6 RNA polymerase (for pAΔ737fs) were carried out under identical conditions except for 10 mM dithiothreitol, 6 mM $MgCl_2$, 2 mM spermidine and 40 μl/ml SP6 RNA polymerase (activity 10 U/μl).

For the production of capped transcripts, the reactions included 0.5 mM m^7 GpppG and the GTP concentration was only 0.1 mM for the first 30 min incubation, after which a further 0.3 mM GTP was added and incubation continued for a further 60 min. Approximately two-thirds of the transcripts produced under these conditions are capped (Dasso and Jackson, 1989a).

The extent of RNA synthesis was calculated from the percentage of radioactivity incorporated from [α-^{32}P]UTP, and RNA was isolated from the bulk of the reaction mix as described previously (Dasso and Jackson, 1989a).

Translation assays and analysis of translation products

Rabbit reticulocyte lysates were prepared and treated with micrococcal nuclease as described previously (Jackson and Hunt, 1983). The translation assays were identical to those previously described by Jackson and Hunt (1983), except for the presence of 4 mM 2-aminopurine, which was added to block the inhibitory effect of any double-stranded RNA contaminants in the *in vitro* transcript preparations. Total incorporation of labelled methionine into acid-precipitable protein was determined as described previously (Jackson and Hunt, 1983). To analyse the products by gel electrophoresis, one volume of translation assay was mixed with an equal volume of 100 μg/ml ribonuclease A, 10 mM EDTA (pH 7.2), incubated for 10 min at room temperature and then diluted with 8 vol. of the sample buffer described previously (Jackson, 1986). Aliquots (usually 10 μl) of this diluted material were loaded onto polyacrylamide slab gels, either 20% acrylamide gels as described previously (Dasso and Jackson, 1989b), or gradient gels of composition varying linearly from 5% (w/v) acrylamide, 0.26% (w/v) methylenebisacrylamide to 30% (w/v) acrylamide, 0.045% (w/v) methylenebisacrylamide, the gradient being stabilized with glycerol. Gel and running buffers were as described previously (Jackson, 1986). Gels were subsequently stained with Coomassie blue R250, dried and exposed to Hyperfilm β-max (Amersham International). The autoradiographs were scanned using a Transidyne 2955 Scanning Densitometer; the gel loadings and the exposure times were adjusted to be within the linear response range of the film.

Acknowledgements

We thank Tim Hunt for helpful comments; Sung-Key Jang and Eckard Wimmer for the gift of clones and for advice in a collaboration supported by a grant from NATO Scientific Affairs Division; Franco Carlotti for the construction of pJ'Δ01; and Steve Inglis for polyclonal antiserum against the influenza virus NS protein. This work was supported by a project grant from the Medical Research Council, and M.T.H. acknowledges a research studentship from the same organization.

References

Dasso,M.C. and Jackson,R.J. (1989a) *Nucleic Acids Res.*, **17**, 3129−3144.
Dasso,M.C. and Jackson,R.J. (1989b) *Nucleic Acids Res.*, **17**, 6485−6497.
Hattori,M. and Sakaki,Y. (1986) *Anal. Biochem.*, **152**, 232−238.
Howell,M.T., Kaminski,A. and Jackson,R.J. (1990) In Brinton,M.A. and Heinz,F.X. (eds), *New Aspects of Positive-strand RNA Viruses*. ASM Publications, Washington, DC, pp. 144−151.
Jackson,R.J. (1986) *Virology*, **149**, 114−117.
Jackson,R.J. (1988) *Nature*, **334**, 292−293.
Jackson,R.J. and Hunt,T. (1983) *Methods Enzymol.*, **96**, 50−74.
Jang,S.K., Krausslich,H.-G., Nicklin,M.J.H., Duke,G.M., Palmenberg,A.C. and Wimmer,E. (1988) *J. Virol.*, **62**, 2636−2643.
Jang,S.K., Davies,M.V., Kaufman,R.J. and Wimmer,E. (1989) *J. Virol.*, **63**, 1651−1660.
Kitamura,N., Semler,B.L., Rothberg,P.G., Larsen,G.R., Adler,C.J., Dorner,A.J., Emini,E.A., Hanecak,R., Lee,J.J., van der Werf,S., Anderson,C.W. and Wimmer,E. (1981) *Nature*, **291**, 547−553.
Kozak,M. (1986) *Cell*, **44**, 283−292.
Kozak,M. (1989) *J. Cell Biol.*, **108**, 229−241.
Kuge,S. and Nomoto,A. (1987) *J. Virol.*, **61**, 1478−1487.
Kuge,S., Kawamura,N. and Nomoto,A. (1989) *J. Virol.*, **63**, 1069−1075.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
Palmenberg,A.C., Kirby,E.M., Janda,M.R., Drake,N.L., Duke,G.M., Potratz,K.F. and Collett,M.S. (1984) *Nucleic Acids Res.*, **12**, 2969−2985.
Pelletier,J. and Sonenberg,N. (1988) *Nature*, **334**, 320−325.
Pilipenko,E.V., Blinov,V.M., Chernov,B.K., Dmitrieva,T.M. and Agol,V.I. (1989) *Nucleic Acids Res.*, **17**, 5701−5711.
Sangar,D.V., Newton,S.E., Rowlands,D.J. and Clarke,B.E. (1987) *Nucleic Acids Res.*,**15**, 3305−3315.
Trono,D., Andino,R. and Baltimore,D. (1988) *J. Virol.*, **62**, 2291−2299.
Vartapetian,A.B., Mankin,A.S., Skripkin,E.A., Chumakov,K.M., Smirnov,V.D. and Bogdanov,A.A. (1983) *Gene*, **26**, 189−195.
Williams,M.A. and Lamb,R.A. (1989) *J. Virol.*, **63**, 28−35.

Received on June 22, 1990