The EMBO Journal vol.7 no.1 pp.245-251, 1988

Ribosomal initiation from an ACG codon in the Sendai virus P/C mRNA

Joseph Curran and Daniel Kolakofsky

Département de Microbiologie, Faculté de Médecine, Université de Genève, CMU, Genève, Switzerland

Communicated by D.Kolakofsky

The Sendai virus P/C mRNA expresses the P and C proteins from alternate reading frames. The C reading frame of this mRNA, however, is responsible for three proteins, C', C and Y, none of which appear to be precursors to each other *in vivo*. Using site-directed and deletion mutagenesis of the P/C gene cloned in SP6 and *in vitro* translation of the mRNAs, we show that the 5' most proximal initiation codon of the mRNA is an ACG at position 81, responsible for C' synthesis. The succeeding initiation codons, all ATGs, are responsible for the P protein (position 104), the C protein (position 114) and the Y protein(s) (either positions 183 or 201). Examination of the relative molar amounts of the C', P and C proteins found *in vivo* suggests that an ACG in an otherwise favorable context is almost as efficient for ribosome initiation as an ATG in a less favored context, but only 10-20% as efficient as an ATG in a more favored context. The judicious choice of increasingly more favorable initiation codons in the P/C gene allows multiple proteins to be made from a single mRNA.

Key words: Sendai virus/site-directed mutagenesis/ribosomes

Introduction

Sendai virus, a model parainfluenza virus, encodes its entire genetic information within a non-segmented 15-kb negative strand (-) RNA (Kingsbury, 1974; Kolakofsky *et al.*, 1974). Upon infection, six positive-sense mRNAs are transcribed from the genome by the virion-associated polymerase. These transcripts are largely monocistronic, encoding the NP, M, F, HN and L proteins. The second mRNA in the transcriptional map, the P mRNA, however, has been demonstrated to be polycistronic. In addition to encoding the P protein, it also codes for a pair of non-structural proteins, C and C' (Dethlefsen and Kolakofsky, 1983; Giorgi *et al.*, 1983) which share tryptic peptides (Lamb and Choppin, 1978; Etkind *et al.*, 1980). The P protein is thought to be initiated at the 5' proximal ATG of the mRNA at position 104, and this open reading frame (ORF) continues for 568 codons (see Figure 1). Antisera raised against a synthetic peptide representing the last 16 amino acids (aa) of this ORF specifically recognizes the P protein. A second ATG is found at position 114 and this +1 ORF continues for 204 codons. Antisera raised against the last 12 aa of this ORF specifically recognizes the C and C' proteins (Curran *et al.*, 1986). Neither of the ATGs which start the P or C ORFs is in the context considered most favorable for initiation (A/GNNATGG), but the ATG of the C ORF (position 114) is considered more favorable in that it contains a purine at position -3, whereas the ATG of the P ORF does not (Kozak, 1986).

Although it is clear that these overlapping ORFs give rise to the P and C proteins, the manner in which the two C proteins are generated has remained problematical. Pulse-chase and canavanine poisoning experiments *in vivo* (Lamb *et al.*, 1976; Curran and Kolakofsky, 1987) have failed to demonstrate a product-precursor relationship between C and C', suggesting that they may be the result of independent ribosomal initiation. C and C' are also made *in vitro* from viral mRNA, and although their ratio can vary widely here from one translation reaction to another, their kinetics of accumulation again suggests that each is independently initiated. One obvious possibility is therefore that the large C' protein (as estimated by PAGE) starts on the second ATG of the mRNA at position 114 and that the smaller C protein begins on either the third or fourth ATGs at positions 183 and 201, which are also in the C ORF. However, since the ATGs at 183 and 201 contain pyrimidines at both positions -3 and +4, whereas the C protein on the other hand is about five times as abundant as either P or C' *in vivo*, such a situation would require that the most abundant protein of the P/C mRNA starts on either the third or fourth ATG, which are also considered the least favorable for initiation. Another possibility is that C and C' are inverted on SDS-PAGE relative to their chain lengths (Curran *et al.*, 1986). In order to unambiguously determine the manner in which C and C' are generated, we have systematically mutated a cloned copy of the P/C gene in the SP6 expression vector, and examined these mRNAs by *in vitro* translation.

Results

Using the Harris strain of Sendai virus to infect BHK cells, we have previously found that the [^{35}S]methionine labeled C protein band is about four times as intense as the P protein band, whereas the C' band was slightly less intense than P, under conditions where these proteins are equally stable (Curran *et al.*, 1986). Figure 2 shows a similar analysis except that a collection of five Sendai strains has been used. Although the C and C' proteins have slightly different electrophoretic mobilities among the five strains, it is clear that in all strains the C protein band is more intense than P, whereas C' is similar in intensity to P in strains Fus and E 9072 (lanes 5 and 6) but is somewhat reduced in the three other strains. Virus infected cells also contain a band marked Y, whose electrophoretic mobility mirrors that of the C and C' proteins among the five strains. The Y protein is yet another product of the P/C gene, since it is also made *in vitro* from mRNA derived from an SP6 clone (Curran and Kolakofsky, 1987).

In order to determine the ATGs which initiate C and C', the first three ATGs (positions 104, 114 and 183) of the P/C

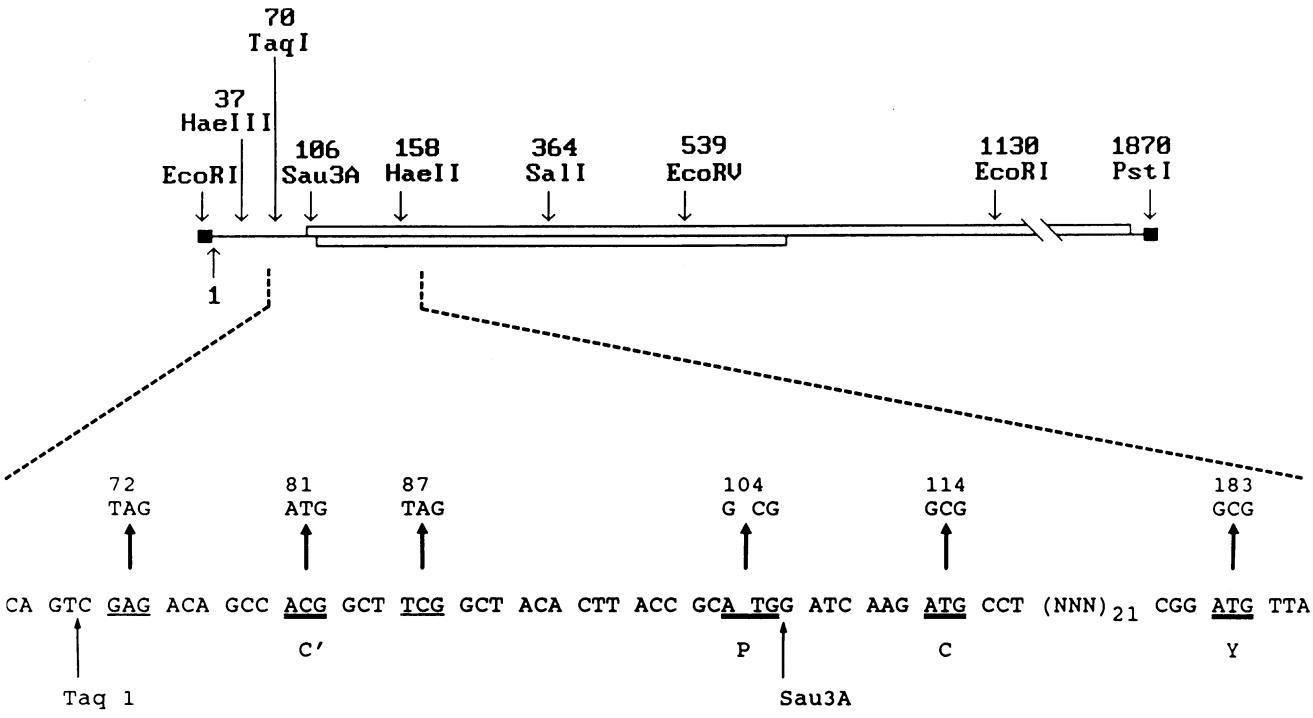

Fig. 1. Organization of the Sendai virus P/C gene in SP65. **Upper panel:** the P/C gene and the restriction sites used either during the construction of the subclones or to linearize the plasmids prior to transcription are shown. The enclosed region above the line indicates the P reading frame and below the line the C reading frame. The closed squares represent additional sequences present at the beginning and ends of the gene which are not part of the viral P/C sequences. **Lower panel:** the sequence surrounding the initiation sites of the P, C, C′ and Y proteins determined from this work are shown. Superimposed on this sequence are the positions of restriction sites used to produce the deletion subclone series, and the sites of mutagenesis. All numbers refer to their position relative to the natural start of the P/C mRNA, and to the first base of each codon indicated.

gene cloned in SP6 were each changed in turn to GCG, such that an internal methionine would be changed to leucine. (Since all manipulations of the gene were carried out as DNA and for clarity, T is always used instead of U regardless of whether it refers to DNA or RNA.) mRNA from each mutant gene was then translated *in vitro*. As shown in Figure 3, when the ATG at 104 is changed, C and C′ are made in the same relative proportions as from wild-type mRNA, but the P protein is selectively lost. When the second ATG is changed (ATG114/GCG), P and C′ continue to be made, but the C protein is selectively lost. A curious consequence of the loss of C is the increase in the amount of Y, which is also apparent when wheatgerm rather than reticulocyte extracts are used for translation, except that here the Y protein appears as a doublet band (right side, Figure 3). When the third ATG is changed (ATG183/GCG), P, C and C′ all continue to be made at wild-type levels. The Y protein(s) is also a very minor product of this mRNA in wheatgerm translations, but overexposure of the gel indicates that the top band of Y is the only protein which is selectively lost in this mutant (not shown). These results indicate that the first and second ATGs start the P and C proteins respectively, but that the third ATG starts the top band of the Y doublet seen upon wheatgerm translation. The bottom band of the doublet is then presumably initiated on the fourth ATG of the mRNA which follows six codons downstream (position 201). Translation of the C′ protein, on the other hand, is unaffected by changing the first three ATGs of the P/C mRNA to GCG.

To determine whether the different electrophoretic mobilities of C and C′ were due to sequences at the N or C termini of these proteins, two constructs were also linearized with *Eco*RV, which cuts at position 539 or 132 codons after the start of the C ORF. These constructs were chosen because they have lost the ATG at 104 and therefore do not translate the P protein, since the truncated P protein might interfere with the interpretation of the results. The construct *Sau*3A (described below) makes only the C protein whereas ATG104/GCG makes both C and C′. The mRNA from this latter *Eco*RV cut construct was found to make two proteins whose mobility is consistent with their being the truncated forms of C and C′, whereas the *Eco*RV cut *Sau*3A construct only made the lower of these two bands. This experiment suggests that the sequences responsible for the different mobilities of C and C′ lie in the N termini of these proteins, and that C′ is most probably the larger of the two.

Since these results appear to conflict with our previous report that antisera to the N-terminus of the C ORF recognized C but not C′ (Curran *et al.*, 1986), these experiments were repeated, but with the SP6 mRNAs rather than infected cell RNA. The antisera were also purified on a protein A sepharose column so that more and cleaner antibodies could be used. Besides wt mRNA which translates both C and C′, the *Sau*3A and ATG114/GCG constructs which make only C and C′ respectively were also used. The results (Figure 4) clearly show that antisera to the N-terminal as well as the C-terminal peptides of the C ORF immunoprecipitate both C and C′, indicating that both proteins contain the N- and C-terminal sequences of this ORF. In addition, antisera to the C-terminus of the ORF also recognizes the Y proteins, but only when they are overexpressed (ATG114/ACG), whereas antisera to the N-terminus do not, consis-

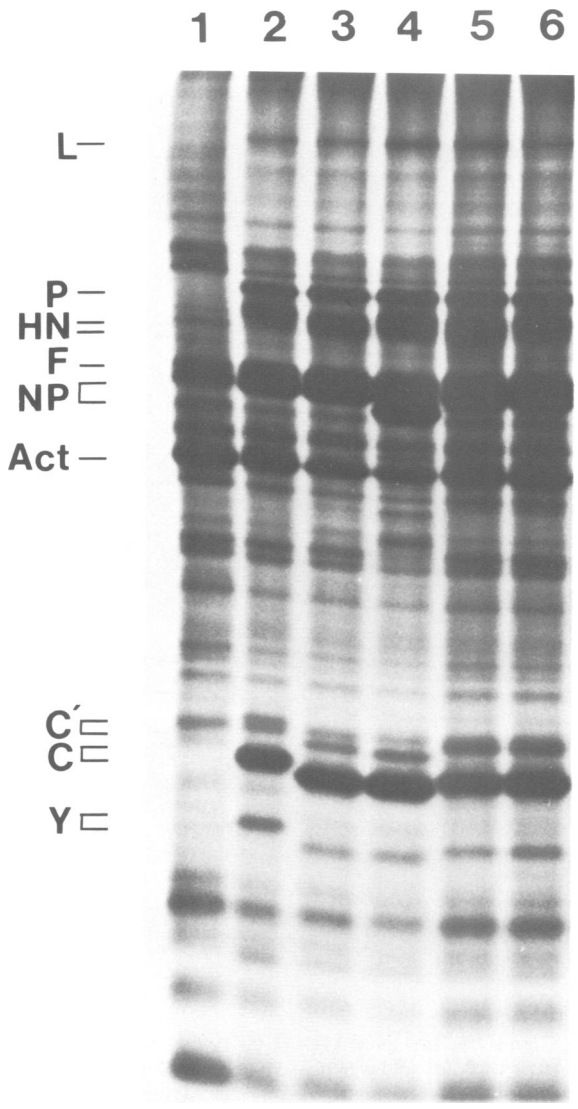

Fig. 2. Analysis of proteins in cells infected with five different Sendai virus strains. BHK cells were infected with five different Sendai virus strains and labeled *in vivo* with [^{35}S]methionine from 16 to 18 h p.i. Total cell extracts were analyzed directly on a 15% acrylamide−SDS gel and visualized by fluorography. **Lane 1**, mock infected cell control; **lane 2**, Harris strain; **lane 3**, MN strain; **lane 4**, Z strain; **lane 5**, Fushima strain; **lane 6**, E9072 strain. The various virus-specific proteins are indicated by letters; Act refers to cellular actin.

tent with the Y proteins starting on the third and fourth ATGs (183 and 201). This result should be considered tentative, however, since the C-terminal peptide antiserum reacts more weakly with Y than it does with either C or C′.

Since the C′ protein and its N-terminal fragment are both larger than the C protein and its N-terminal fragment (as estimated by SDS−PAGE) and since both C and C′ reacted with antisera specific to both ends of the C ORF, this suggests that the C′ protein is initiated upstream of the C ATG at 114. To examine this possibility more directly, three deletion mutants were prepared in which the first 37 (*Hae*III), 70 (*Taq*I) and 106 (*Sau*3A) nt of the P/C mRNA were removed and the remainder joined to SP6 sequences within the polylinker. When the mRNAs from these constructs were translated (left side, Figure 5), deletion of nt 1−37 could not be distinguished from wild-type mRNA (not shown) whereas deletion of nt 1−70 (*Taq*I) severely reduced C′ expression relative to wild type but did not eliminate it. In other experiments, *Taq*I mRNA translated C′ at 10−50% the rate as wild-type mRNA (not shown). Deletion of the first 106 nt (*Sau*3A) eliminated P expression as expected, but it also totally eliminated C′. The overexpression of Y as C′ is progressively lost is also clear in this deletion series. Since C′ is not initiated at the ATG at 104, this suggests that C′ is initiated between nt 71 and 106, and most likely closer to position 71 since the *Taq*I mRNA expresses less C′ than wild type.

Examination of the sequences upstream of the C ORF shows that the 5′ extension of this frame does not contain any stop codons and could therefore be translated if ribosomes were to initiate at a non-ATG codon. In particular, just downstream of the *Taq*I site lies the Thr codon ACG (position 81) in the same reading frame as C. This codon has been found to act as a ribosomal initiation site in a somewhat analogous situation in adeno-associated virus (AAV) (Bacerra *et al.*, 1985). To examine whether the ACG codon was likely to initiate the C′ protein, we have changed the second codon following the ACG (position 87) from TCG to TAG, as well as the third codon upstream of the ACG (position 72) from GAG to TAG, since these positions required only a single base change to introduce the amber codons. When these two mRNAs were translated (right side, Figure 5), the introduction of an amber codon upstream of the ACG (GAG72/TAG) was found to have no effect on the translation of P, C and C′, whereas when this codon was placed downstream of the ACG (TCG87/TAG), the P and C protein were made as expected, but C′ expression could not be detected. This result then places the initiation site of C′ between positions 72 and 87, and is consistent with the notion that C′ is initiated at the ACG codon at position 81. To examine this possibility more directly, we next changed the ACG at 81 to an ATG. This artificial initiation site is now in an excellent context for ribosomal initiation, with a purine at −3, a G at +4, as well as Cs in positions −1 and −2. When this mRNA was translated, P synthesis could no longer be detected and C synthesis decreased 5- to 10-fold, whereas C′ synthesis increased dramatically. The co-migration of C′ made from ACG81/ATG mRNA with that made from wt mRNA further suggests that C′ is initiated at the ACG at 81.

Discussion

A combination of site-directed and deletion mutants of the Sendai virus P/C gene has confirmed our previous expectation that the P protein is initiated on the 5′ proximal ATG at position 104. The second ATG at position 114 is clearly responsible for initiation of the C protein, whereas the third and fourth ATGs at 183 and 201 are most likely responsible for initiation of the Y protein(s). The initiation site of the C′ protein, however, maps between positions 72 and 87 as defined by in-frame amber codons. Although there is no ATG in this region, there is an ACG codon at position 81 in frame with the C ORF, which is in a context which is otherwise highly favorable for ribosomal initiation. When this ACG is changed to ATG, C′ is now virtually the only product of this mRNA. The SV ACG codon is also similar to the ACG codon responsible for the initiation of the AAV B protein. There is in fact a remarkable coincidence between

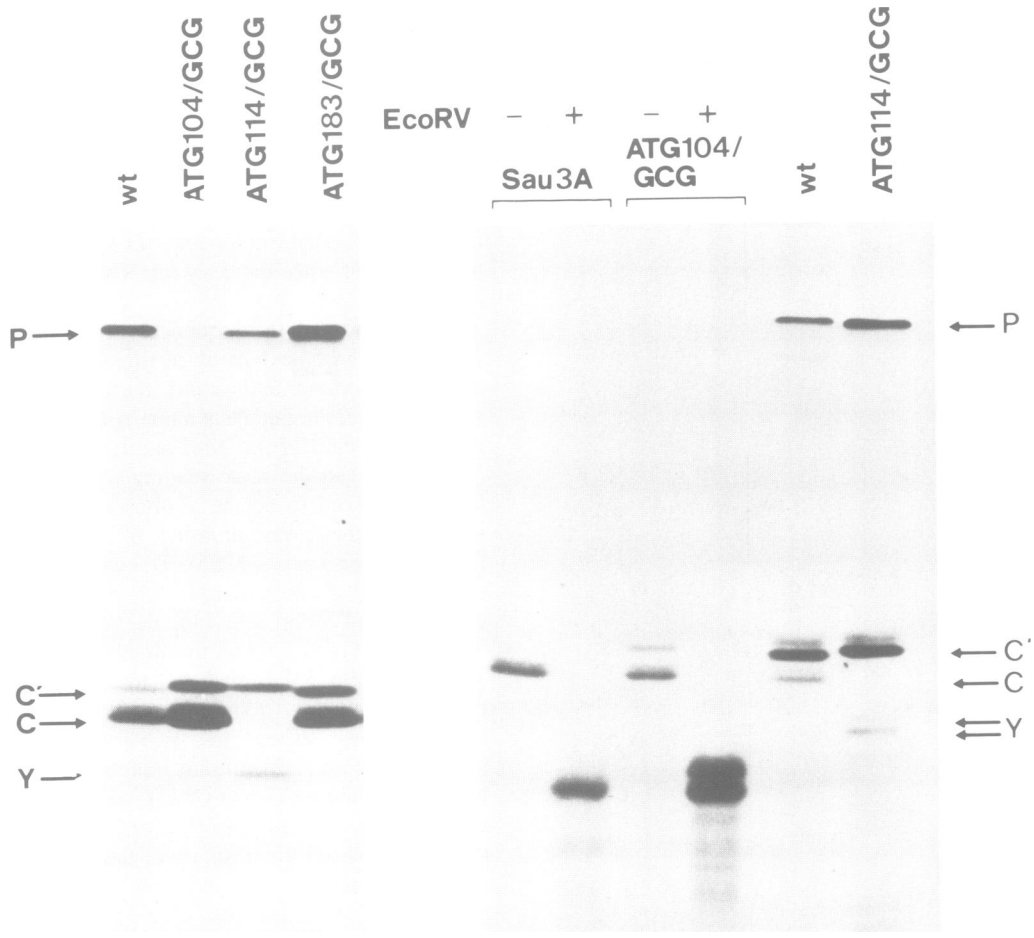

Fig. 3. Left-hand panel: site-directed mutagenesis of the first three ATGs of the P/C gene. The first three ATGs of the P/C wild-type sequence (positions 104, 114 and 183) were individually changed to a GCG leucine codon. The mutated and wild-type plasmids were linearized with *PstI*, transcribed with SP6 polymerase and the mRNAs translated in a rabbit reticulocyte lysate in the presence of [^{35}S]methionine. Protein products were resolved on 15% polyacrylamide−SDS gels and visualized by fluorography. **Right-hand panel:** the N-terminal sequences of C and C′ are responsible for their different mobilities. Plasmid subclones *Sau*3A and ATG104/GCG were linearized with *Eco*RV as well as *PstI* (see Figure 1), transcribed with SP6 polymerase and the mRNAs translated in wheat germ extracts to examine the C-terminally truncated forms of the C/C′ proteins. Proteins were resolved as in the left-hand panel. Translations of the wild-type and ATG114/GCG mRNAs are also shown for reference.

the Sendai C′ and C proteins and the AAV B and C proteins. Both pairs of proteins are C-terminal nested sets in which the smaller protein starts on an ATG whereas the larger protein starts on a ACG upstream. Further both ACGs are not only in the most favored context for ribosome initiation with a purine at position −3 and a G at +4, but positions +5 to +10 are also identical except that position +7 is a T in Sendai virus and a C for AAV. Thus, although we have not formally shown that the Sendai C′ protein starts on the ACG codon (which will require N-terminal sequencing of the protein), we can see no other alternative which is reasonable.

Besides AAV and SV, two other instances where an ACG codon serves to initiate proteins in eucaryotic systems have been reported. A bacteriophage T7 mRNA in which the natural ATG has been changed to ACG was found to be translated in both rabbit reticulocyte lysates and wheat germ extracts at between 12 and 18% of the wild-type mRNA (Anderson and Buzash-Pollert, 1985). In another case, the natural ATG initiation codon of the mouse DHFR gene was changed to ACG, and upon transfection into COS cells was found to make this protein at about 5% of the control level. The context of the ACG was also shown to be important in this system, since changing the purine at −3 and the G at +4 severely reduced the ACG directed DHFR expression (Peabody, 1987). There is thus little doubt that ACG can serve as an initiation codon in eucaryotic systems, albeit at reduced levels relative to ATG.

What is peculiar to the SV C and C′ story is that, like AAV where initiation at an ACG codon is used to produce an N-terminally elongated form of the protein initiated with ATG, this situation is also superimposed on an mRNA which contains an overlapping gene for the P protein. Since the P protein initiates on an ATG which lies in between the two initiation codons for C′ and C, ribosomes scanning from the 5′ end of the mRNA would encounter the initiation sites for C′, P and C in this order. Previous estimates as well as densitometric analysis of the gel shown in Figure 2 (and normalizing for Met content) shows that *in vivo* C′ is made at half to one times the amount of P, whereas C is four to five times as abundant as P. This would indicate that the C′ ACG at position 81 is initiated at 50−100% of the rate as the P

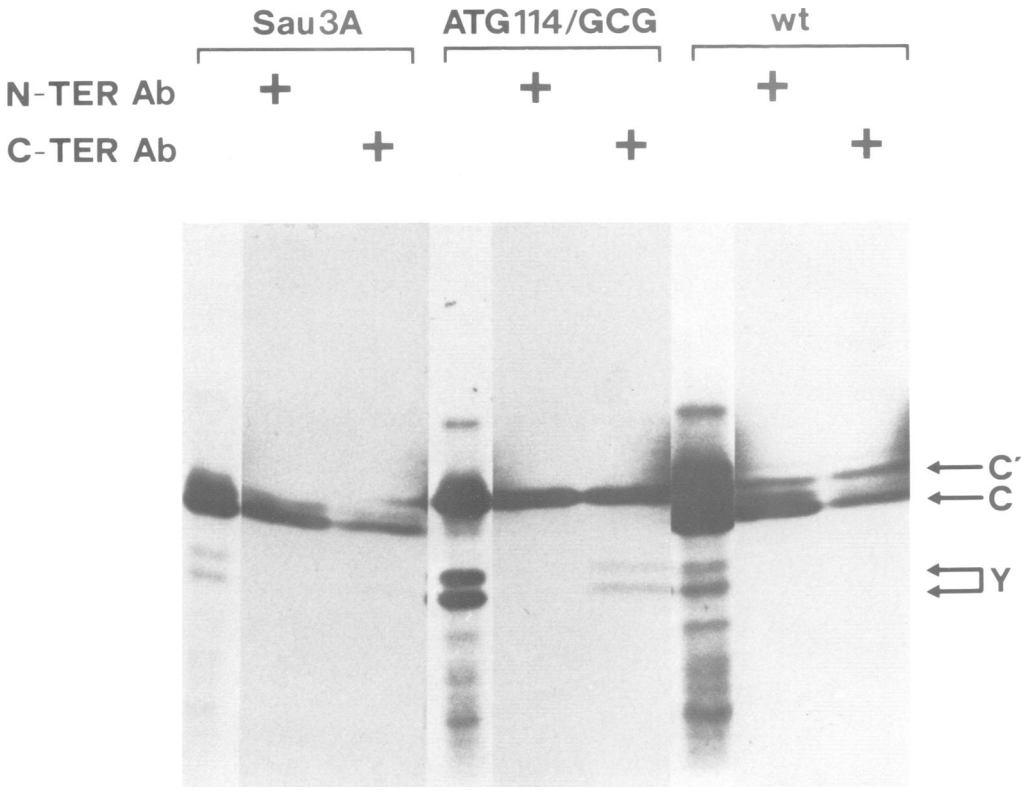

Fig. 4. Immunoprecipitation of the C and C′ proteins with peptide antiserum. Peptide antiserum raised to the N-terminal and C-terminal 12 aa of the C ORF was used to immunoprecipitate the proteins expressed by subclones *Sau*3A, ATG114/GCG, and wild-type in WGE in the presence of [^{35}S]methionine. Antigen−antibody complexes were recovered, resolved on 15% polyacrylamide gels and visualized by fluorography. Only the lower portion of the gel is shown.

ATG at 104, but only 10−20% of the rate of the C ATG at 114. Since the ACG-initiated AAV B protein is made at ~10% of the rate of the ATG-initiated AAV C protein (Becerra *et al.*, 1985), our results indicate that the SV ACG codon functions at a similar level relative to the ATG at 114 as do the ACG and ATG in the AAV system, but almost as well as the P ATG at 104. An ACG in an otherwise favorable context is thus almost as efficient for ribosomal initiation as a 5′ proximal ATG which is in a less favored context.

Although it is unexpected that the major product of the P/C mRNA starts from the third initiation codon relative to the 5′ end of the mRNA (the C ATG at 114), these results are not inconsistent with the scanning model for ribosomal initiation. According to the work on the model preproinsulin gene (Kozak, 1986 and references therein), the major context determinants for efficient ribosomal initiation are a purine at position −3, followed by a G at +4. The initiation sites for C′, P and C on the SV P/C mRNA are thus found in increasingly more favorable contexts for ribosomal initiation. Further, when the C′ ACG at 81 which naturally lies in the most favored context for initiation is changed to an ATG, initiation from the two downstream ATGs virtually ceases, presumably because there are now few ribosomes which have scanned past position 81 without having initiated polypeptide chains. The strongest initiation codon of the mRNA, the C ATG at position 114, curiously does not lie in the most favorable context, in that although it contains an A at −3, it has a C rather than a G at +4. This might then explain how some ribosomes can reach the ATGs at positions 183 and 201, which we believe are responsible for initiation of the Y protein(s). Consistent with this hypothesis, we note that whenever C′ or C expression is reduced *in vitro* either due to deletions or to site-directed mutations, Y protein increases as a consequence. The choice of these particular initiation codons for the SV C′, P and C proteins allows these proteins as well as Y to be made from the same mRNA, and suggests that, similar to other viral systems, e.g. AAV, the relative expression of these proteins must be tightly regulated.

Materials and methods

Construction of an SP6 subclone expressing the P/C mRNA

Using the cDNA library prepared by Dowling *et al.* (1983), a full-length P/C clone was initially constructed by combining an *Rsa*I−*Sal*I fragment from pSL2 representing all of the start of the P/C gene and the upstream 25 nt of the NP gene (which does not contain an ATG codon) to the *Sal*I−*Pst*I fragment of pSN5 (Giorgi *et al.*, 1983) representing the remainder of the gene except for the last 26 nt of the 3′ non-translated region, but including the poly(C) tail of ~20 nt. This combined fragment was then inserted into the *Sma*I−*Pst*I site of pUC19 and then subcloned into the expression vector SP65 (Melton *et al.*, 1984) to yield SP65 P/C. Plasmid DNA was linearized with *Pst*I and transcribed to produce a capped mRNA using the SP6 polymerase. The conditions for this reaction were essentially as described by Konarska *et al.* (1984). Between 1 and 2 μg of DNA was incubated in 40 mM Tris pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM DTT, 250 μM UTP/CTP/ATP, 25 μM GTP, 10 units RNasin, 500 μM $m^7G(5')ppp(5')G$ and 1−2 units SP6 polymerase (Boehringer), at 40°C for 1 h in a total volume of 20 μl. The RNA was then precipitated directly by the addition of 3 volumes of ethanol.

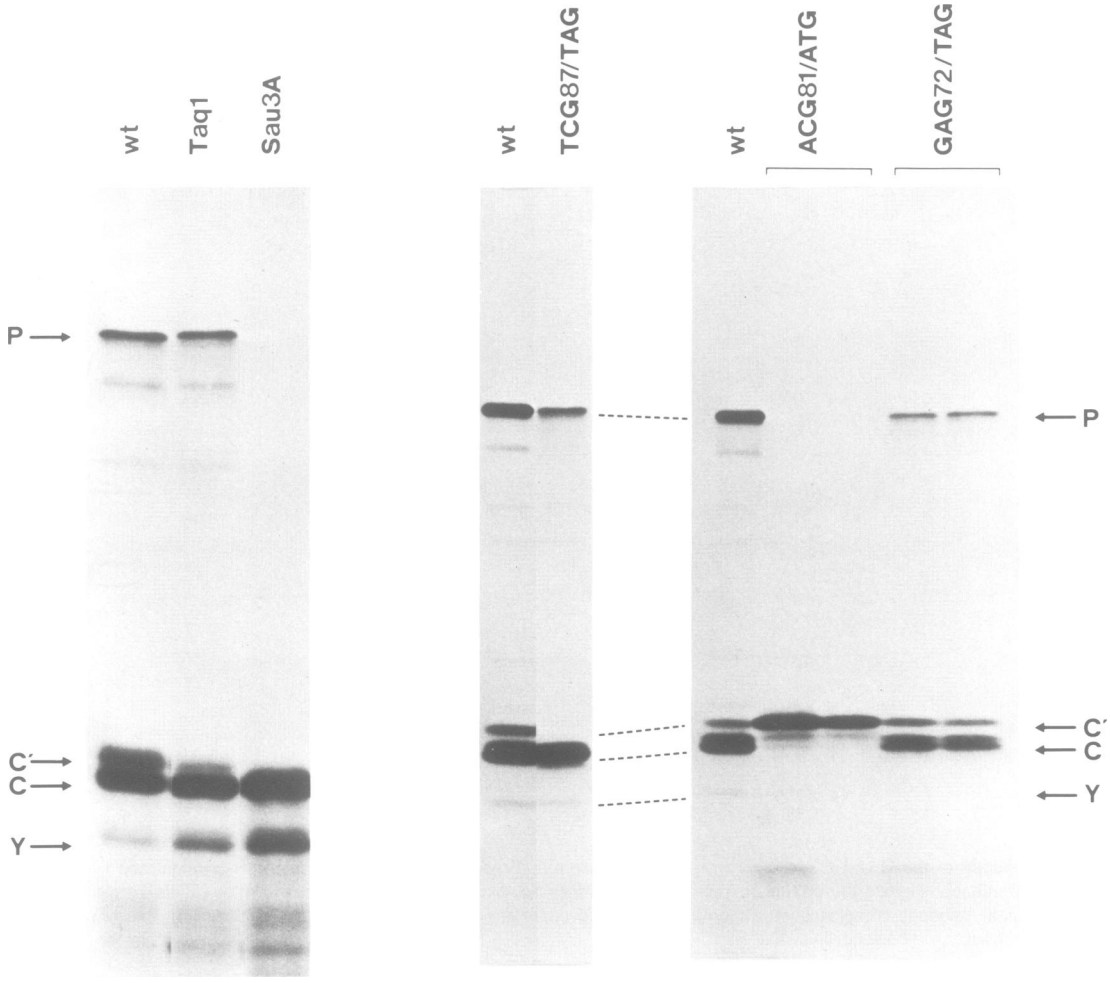

Fig. 5. Location of the C′ initiation codon. **Left-hand panel:** expression of the deletion series. Regions of the 5′ sequence were deleted from the wild-type plasmid by use of convenient restriction sites. The first 70 nt were deleted using a *Taq*I site and the first 106 nt (which included the P initiation codon) were deleted using a *Sau*3A site (see Figure 1). These constructs were than linearized with *Pst*I, transcribed with SP6 polymerase and translated in WGE in the presence of [^{35}S]methionine. Protein products were compared to wild type on 15% polyacrylamide−SDS gels and visualized by fluorography. **Right-hand panel:** site-directed mutagenesis upstream of the C ORF. A translational stop codon (TAG) was inserted in the C reading frame either upstream (position 72) or downstream (position 87) of the ACG81 codon. In addition, the ACG81 codon was changed to an ATG codon. All these subclones were linearized with *Pst*I, transcribed and their mRNAs expressed in RRL in the presence of [^{35}S]methionine. Proteins were resolved on 15% polyacrylamide−SDS gels. The double lanes for ACG81/ATG and GAG72/TAG represent protein profiles from two independent clones.

In vitro translation and immunoprecipitation

One fiftieth of the RNA produced in a 20 μl SP6 reaction was translated in either a wheat germ extract (Amersham) or a rabbit reticulocyte lysate (Promega Biotec) under the conditions recommended by the supplier. The *in vitro* products were analyzed on SDS−polyacrylamide gels and viewed by fluorography.

Immunoprecipitations were performed essentially as described by Curran *et al.* (1986). Briefly, 5−10 μl of the *in vitro* translation mix was resuspended in 150 μl of RIPA buffer (150 mM NaCl, 1% deoxycholate, 1% Triton X100, 0.1% SDS, 10 mM Tris pH 7.8) and precipitated with a maximum of 20 μl of peptide antiserum (the IgG fraction of which had been previously purified by binding to a protein A sepharose 4B column and eluted in citrate buffer pH 2.5). After dialysis against PBS, the antiserum was stored at −20°C. Antibody−antigen complexes were recovered on protein A sepharose and analyzed on SDS-polyacrylamide gels.

Site-directed mutagenesis

The procedure employed to introduce base changes into the SP65 P/C sequence was supplied by Dr Toni Gautier (Biogen SA, Geneva). This method involved the formation of a 'gapped' plasmid similar to the method described previously in the phage M13 system (Kramer *et al.*, 1984). Starting with the SP65 P/C wild-type plasmid, cuts were made around the site for mutagenesis such that a fragment no greater than 10% of the total plasmid (i.e. 400 bp; in all constructs described this required a *Sac*I/*Sal*I digestion) was excised, and the remainder of the linearized vector purified by agarose gel electrophoresis. A second sample of SP65 P/C plasmid DNA was linearized with *Bgl*II which cuts opposite the site for mutagenesis. The two linear plasmids were then mixed in equimolar amounts and denatured in 20 μl of 0.2 M NaOH at 22°C for 10 min before adding 200 μl H_2O, 25 μl 1 M Tris pH 8 and 50 μl 0.1 M HCl. This mixture was incubated at 65°C for 2−3 h. The re-annealing was checked on a 1% agarose/TBE gel (90 mM Tris pH 8.3, 90 mM borate, 2.5 mM EDTA) where the gapped plasmid migrated slightly faster than nicked SP65 P/C and slower than linear SP65 P/C plasmid DNA, both of which were used as markers. The gapped form was then electroeluted.

Approximately 20 pmol of phosphorylated oligodeoxynucleotide [17−24 mers were employed with the required change(s) from wild type located in the exact centre] was mixed with 0.1 pmol of gapped plasmid in 30 μl of 50 mM NaCl, 10 mM Tris pH 7.4, 1 mM EDTA. Heteroduplex formation was allowed to take place at 10°C below the T_m of the oligomer (35°C) for 1−2 h. The plasmid was then 'filled-in' and the ends ligated simultaneously using 5 units of the Klenow fragment and 1 unit of T4 DNA ligase (Boehringer) in 50 mM Tris pH 7.4, 10 mM $MgCl_2$, 10 mM DTT, 1 mM spermidine, 1 mM ATP and 1 mM dNTPs, at 15°C for 12 h. The plasmids were then used to transform *E.coli* (HB101) cells. Subclones were screened by the method of Grunstein and Hogness (1975), using the end-labeled oligonucleotide used to prepare the mutagenized subclones as probes. Nitrocellulose filters were washed to a final stringency of 1 × SSC

at 50°C for 15–30 min. Minilysates were then prepared from positive colonies using the boiling method of Holmes and Quigley (1981) and used to retransform HB101, and colonies positive to the probe were selected a second time. New minilysates were then prepared and one third of the DNA was linearized with *Pst*I, recovered by precipitation with ethanol, transcribed using SP6 polymerase, and translated under the conditions outlined above. Finally, to confirm the constructions, the DNA was sequenced around the site of the intended mutation(s) using ddNTPs and an oligonucleotide primer located downstream from the mutation. Plasmid sequencing was performed essentially as described by Haltiner *et al.* (1985) using one third of the plasmid DNA isolated from a 1.5 ml overnight culture.

Construction of the 5′ deletion series

Clone P/C *Hae*III was constructed to delete the first 37 nt from the wild-type P/C mRNA. Starting from the wt plasmid, an *Eco*RI/*Sal*I fragment (the start of the insert to a unique *Sal*I site at position 364) was isolated which contained a single *Hae*III site located at position 37. This fragment was digested with *Hae*III and the 327 bp *Hae*III/*Sal*I fragment was then ligated back into the wild-type SP65 P/C plasmid in which the first 364 bp of P/C sequence had been removed by a *Sac*I/*Sal*I digestion. This final ligation took place in two steps. The *Sal*I ends were first ligated together, then the remaining overhangs were blunted with Klenow before the second ligation.

Clone P/C *Taq*I contained a 70 nt deletion from the 5′ end of the P/C mRNA. A 158 bp *Eco*RI/*Hae*II (unique site at position 158) fragment isolated from the wild-type plasmid was digested with *Taq*I and the 88 bp *Tap*I/*Hae*II portion was ligated back to the remainder of the P/C gene contained on a *Hae*II/*Pst*I fragment. This DNA was then inserted into the *Sma*I/*Pst*I sites of SP65 in a two-step ligation. The *Pst*I ends were first ligated together and then the *Hae*II end was blunted with Klenow so that the second ligation could occur.

Clone P/C *Sau*3A contained a 106 nt deletion from the 5′ end of the P/C mRNA. This was constructed by first isolating the *Eco*RI/*Sal*I fragment mentioned earlier. The fragment was then digested with *Sau*3A (position 106) and the 258 bp *Sau*3A/*Sal*I portion was ligated back to the remainder of the P/C gene contained on a *Sal*I/*Pst*I fragment. The DNA was then inserted into the *Bam*HI/*Pst*I sites on SP65.

Note: All deletions resulted in a loss of additional sequences present on the wild-type plasmid originating from the neighbouring NP gene. In addition, all map references refer to positions on the P/C gene and do not include the additional sequences mentioned above.

Acknowledgements

We are grateful to Bob Lamb for providing us with the collection of Sendai virus strains. This work was supported by a grant from the Swiss National Science Fund.

References

Anderson,C.W. and Buzash-Pollert,E. (1985) *Mol. Cell. Biol.*, **5**, 3621–3624.
Becerra,S.P., Rose,J.A., Hardy,M., Baroudy,B.M. and Anderson,C.W. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 7919–7923.
Curran,J.A., Richardson,C. and Kolakofsky,D. (1986) *J. Virol.*, **57**, 684–687.
Curran,J.A. and Kolakofsky,D. (1987) *J. Gen. Virol.*, **68**, 2515–2519.
Dethlefsen,L. and Kolakofsky,D. (1983) *J. Virol.*, **46**, 321–324.
Dowling,P.C., Giorgi,C., Roux,L., Dethlefsen,L.A., Galantowicz,M., Blumberg,B.M. and Kolakofsky,D. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 5213–5216.
Etkind,R.R., Cross,R.K., Lamb,R.A., Merz,D.C. and Choppin,P.W. (1980) *Virology*, **100**, 22–33.
Girogi,C., Blumberg,B.M. and Kolakofsky,D. (1983) *Cell*, **35**, 829–836.
Grunstein,M. and Hogness,D. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961–3965.
Haltiner,M., Kempe,T. and Tjian,R. (1985) *Nucleic Acids Res.*, **13**, 1015–1025.
Holmes,D.S. and Quigley,M. (1981) *Anal. Biochem.*, **114**, 193–197.
Kingsbury,D.W. (1974) *Med. Microbiol. Immunol.*, **160**, 73–83.
Kolakofsky,D., Boy de la Tour,E. and Delius,J. (1974) *J. Virol.*, **13**, 261–268.
Konarska,M.M., Padgett,R.A. and Sharp,P.A. (1984) *Cell*, **38**, 731–736.
Kozak,M. (1986) *Cell*, **44**, 283–292.
Kramer,W., Drutsa,V., Jansen,H.W., Kramer,B., Pflugfelder,M. and Fritz,H.J. (1984) *Nucleic Acids Res.*, **12**, 9441–9456.
Lamb,R.A., Mahy,B.W.J. and Choppin,P.W. (1976) *Virology*, **69**, 116–131.
Lamb,R.A. and Choppin,P.W. (1978) *Virology*, **84**, 469–478.
Melton,D.A., Krieg,P.A., Rebagliati,M.R., Maniatis,T., Zinn,K. and Green,M.R. (1984) *Nucleic Acids Res.*, **12**, 7035–7056.
Peabody,D.S. (1987) *J. Biol. Chem.*, **262**, 11847–11851.

Received on November 2, 1987

Note added in proof

N-terminal sequencing of [^{35}S]methionine labeled C′ protein made *in vitro* from wild-type mRNA for 13 cycles showed a single methionine at position 11. This result is consistent with initiation of the C′ protein at the ACG/81 codon followed by removal of the initiating methionine, leaving the methionine at position 11 from the AUG/114 codon.