# Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein

# Martin D.Ryan<sup>1</sup> and Jeff Drew

AFRC Institute for Animal Health, Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF, UK

<sup>1</sup>Corresponding author

Communicated by R.J.Jackson

We describe the construction of a plasmid (pCAT2AGUS) encoding a polyprotein in which a 19 amino acid sequence spanning the 2A region of the foot-and-mouth disease virus (FMDV) polyprotein was inserted between the reporter genes chloramphenicol acetyl transferase (CAT) and  $\beta$ -glucuronidase (GUS) maintaining a single, long open reading frame. Analysis of translation reactions programmed by this construct showed that the inserted FMDV sequence functioned in a manner similar to that observed in FMDV polyprotein processing: the CAT2A-GUS polyprotein underwent a cotranslational, apparently autoproteolytic, cleavage yielding CAT-2A and GUS. Analysis of translation products derived from a series of constructs in which sequences were progressively deleted from the N-terminal region of the FMDV 2A insertion showed that cleavage required a minimum of 13 residues. The FMDV 2A sequence therefore provides the opportunity to engineer either whole proteins or domains such that they are cleaved apart cotranslationally with high efficiency.

Key words: autoproteolytic/cotranslational cleavage/foot-and-mouth disease virus/polyprotein

#### Introduction

Linking proteins in the form of self-processing polyproteins is a strategy adopted in the replication of many viruses, notably the positive strand RNA viruses. In the picornaviridae a single long open reading frame (ORF) encodes a polyprotein of some 225 kDa, although full-length translation products are not normally observed due to extremely rapid 'primary' intramolecular (cis) cleavages mediated by virus encoded proteinases. A number of such proteinases have been partially characterized and are thought to be related, both structurally and in catalytic mechanism, to cellular proteinases (Bazan and Fletterick, 1988). In the case of the entero- and rhinoviruses, a primary cleavage occurs between the P1 capsid protein precursor and the replicative domains of the polyprotein (P2, P3; Figure 1A). This cleavage is mediated by a virus encoded proteinase (2A<sup>pro</sup>), of some 17 kDa, cleaving at its own N-terminus (Toyada et al., 1976; Sommergruber et al., 1989).

The aphtho-, or foot-and-mouth disease (FMDV), viruses form a distinct group within the picornaviridae. FMDV polyprotein undergoes a primary polyprotein cleavage at the C-terminus of the 2A region between the capsid protein precursor (P1-2A) and replicative domains of the polyprotein

2BC and P3 (Figure 1B). Precursors spanning the 2A/2B cleavage site are not detected during native polyprotein processing. The 2A region of the FMDV polyprotein was demonstrated, however, to be only 16 amino acids long (Robertson *et al.*, 1985), shown in Figure 1C. The predicted amino acid sequence of this region, in contradiction with previous reports (Ryan *et al.*, 1991), is totally conserved amongst all aphthovirus genomic RNAs sequenced to date (N.Knowles, personal communication).

In an initial characterization of the FMDV 2A region, we reported the processing properties of a series of recombinant FMDV polyproteins, and showed that the FMDV 2A/2B cleavage activity was mediated by residues located within a 19 amino acid sequence spanning the FMDV 2A region (Ryan et al., 1991). These data did not, however, eliminate the involvement of additional sequence elements elsewhere in the FMDV polyprotein (although this seemed improbable), nor was it clear to what extent the 2A region could function independently of the physical environment provided by the FMDV polyprotein. To address this question we have studied proteolysis associated with FMDV 2A in a completely foreign context, that of a synthetic polyprotein designed such that two reporter genes flank sequences from the FMDV 2A region of the polyprotein. In this paper we describe the design and construction of plasmids pCATGUS and pCAT2AGUS and the analysis of the proteolytic processing properties of the recombinant polyproteins when expressed in rabbit reticulocyte lysates and human HTK-143 cells.

## Results

The reporter genes, chloramphenicol acetyltransferase (CAT) and β-glucuronidase (GUS) were amplified by PCR using oligonucleotide primers such that restriction sites were created at both termini. Individual genes were cloned (pCAT20/21, pGUS12/23; Figure 2) and also assembled together (pCATGUS; Figure 2) to create a single ORF encoding the artificial polyprotein [CATGUS]. Coding sequences from the FMDV 2A region were assembled in the plasmid vector pGEM 7zf+ in such a way that a series of unique restriction sites was created throughout the sequence (pMR90; Figure 3). The fragment of DNA containing the FMDV sequences was excised by digestion with restriction enzymes XbaI and ApaI and inserted between the CAT and GUS genes of pCATGUS to retain a single ORF and form construct pCAT2AGUS (Figures 2 and 3).

Translation studies were performed in two systems: (i) a coupled transcription/translation (TnT) rabbit reticulocyte system and (ii) a human cell line (HTK-143) infected with the recombinant vaccinia virus vTF7-3 expressing T7 RNA polymerase (Fuerst *et al.*, 1986). Plasmid DNAs of constructs described above were used to programme the TnT rabbit reticulocyte lysate system or transfected into HTK-143 cells which had previously been infected with vTF7-3.

Translation directed by plasmid pCAT14/21 showed a

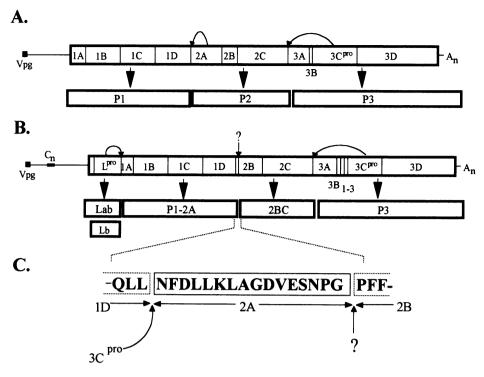


Fig. 1. Picornavirus primary polyprotein cleavages. The 5' non-coding region is capped by a small protein VPg (or 3B). The single long ORF and polyprotein organization are shown (boxed areas) for both entero- and rhinovirus groups (A) and aphthoviruses (B). Arrows indicate sites of primary cleavage and the virus-encoded proteinases responsible, where known. Primary cleavage products are shown below. The amino acid sequence spanning the aphthovirus 2A region of the polyprotein is shown, the 2A oligopeptide being cleaved from the capsid protein 1D by 3C<sup>pro</sup> in an intermolecular reaction occurring at a later stage of polyprotein processing (C).

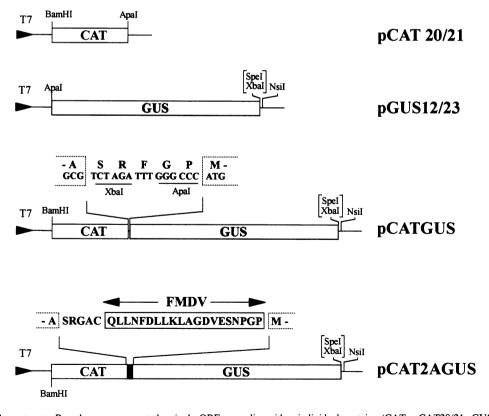


Fig. 2. CAT/GUS constructs. Boxed areas represent the single ORFs encoding either individual proteins (CAT, pCAT20/21; GUS, pGUS12/23) or the artificial polyproteins [CATGUS] and [CAT2AGUS]. All plasmids were based on pGEM transcription vectors.

major polypeptide of the expected molecular weight for CAT (25.7 kDa) in both rabbit reticulocyte lysates (Figure 4A) and cells (Figure 5A). This protein was immunoprecipitated

with anti-CAT antibodies but not anti-GUS or anti-2A antibodies (Figures 4B and 5B). Translation programmed with pGUS12/23 produced a protein which migrated at the

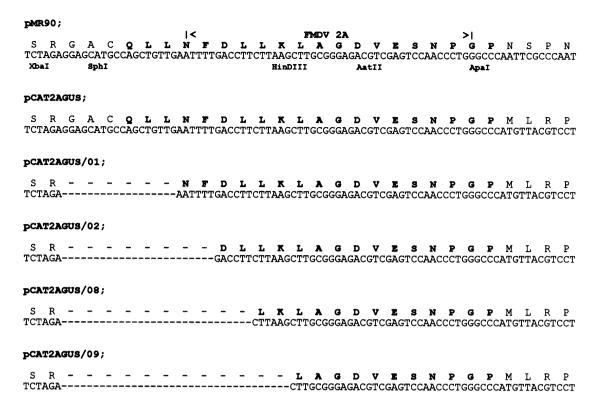


Fig. 3. Sequence of FMDV 2A region insertions. FMDV amino acid sequences are shown in bold. FMDV 2A sequences are indicated. Plasmid pMR90: residues -QLL- correspond to the three C-terminal residues of capsid protein 1D, residues -NFDLLKLAGDVESNPG- to the 2A region, and the proline residue to the N-terminal residue of protein 2B. Plasmid pMR90 was constructed such that a number of unique restriction sites (shown below the nucleotide sequence) were created in the 2A coding sequence, facilitating the deletion of 2A sequences in constructs pCAT2AGUS/01, /02, /08 and /09.

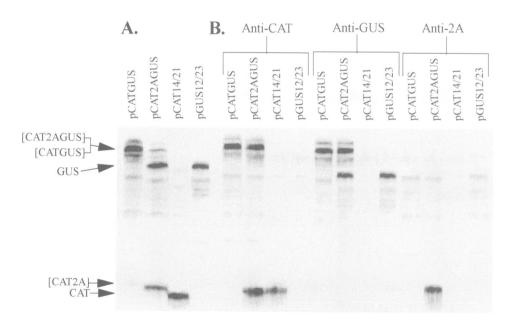


Fig. 4. Translation products of CAT/GUS constructs. TnT rabbit reticulocyte lysate systems were programmed with plasmid DNA from constructs shown in Figure 3 (A). The identity of rabbit reticulocyte lysate translation products was confirmed by immunoprecipitation with either anti-CAT, -GUS or -2A antibodies (B).

expected molecular weight for GUS (70.4 kDa) in both translation systems (Figures 4A and 5A). This protein was immunoprecipitated by antibodies directed against GUS but not with the other antibodies (Figures 4B and 5B).

Translation programmed with pCATGUS showed a major polypeptide of the size expected for the [CATGUS] polyprotein (96.3 kDa) in both translation systems (Figures 4A

and 5A). The [CATGUS] polyprotein was immunoprecipitated by both anti-CAT and anti-GUS antibodies but not anti-2A antibodies (Figures 4B and 5B). In contrast, translation directed by pCAT2AGUS showed three major translation products in both translation systems (Figures 4A and 5A). The uppermost band comigrated with [CATGUS] and corresponded to the expected [CAT2AGUS] translation

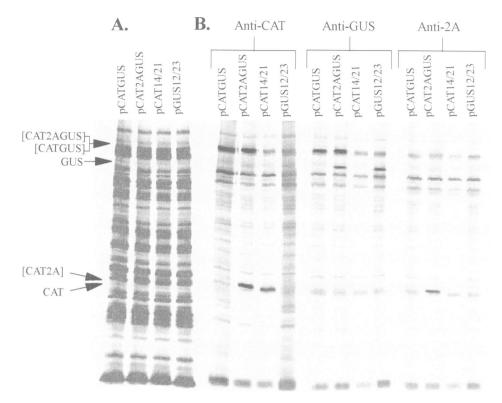


Fig. 5. Translation products of CAT/GUS constructs. Human HTK-143 cells were infected with vaccinia virus vTF7-3 and subsequently transfected with plasmid DNA from constructs shown in Figure 3 (A). The identity of translation products was confirmed by immunoprecipitation with either anti-CAT, -GUS or -2A antibodies (B).

product. The [CAT2AGUS] polypeptide product was immunoprecipitated by anti-CAT and anti-GUS but not anti-2A antibodies (Figures 4B and 5B). The second band comigrated with GUS and was immunoprecipitated only by antibodies directed against GUS (Figures 4B and 5B), corresponding to a GUS cleavage product. The lower band migrated somewhat more slowly than CAT and was immunoprecipitated by anti-CAT and anti-2A antibodies but not anti-GUS antibodies (Figures 4B and 5B), corresponding to the [CAT2A] cleavage product.

Densitometric analysis of the distribution of radiolabel in the rabbit reticulocyte lysate translation reaction programmed with pCAT2AGUS (Figures 4A and 6A) showed that ~80% of the [CAT2AGUS] translation product was cleaved to [CAT2A] and GUS. This analysis was not performed on the cellular translation products due to comigration with other vaccinia virus or host-cell proteins (Figure 5A). Densitometric analysis would not be valid when performed on immunoprecipitations as antibodies may not react as efficiently with proteins in the context of a polyprotein as individual cleavage products.

In a further experiment a larger rabbit reticulocyte lysate translation reaction was incubated for different periods, aliquots removed and translation arrested by the addition of RNase A and cycloheximide. These aliquots were further incubated for different periods. The translation profiles obtained are shown in Figure 6B. Densitometric analysis of the distribution of radiolabel showed that the proportions of all translation products are constant throughout the period of synthesis and subsequent incubations.

The unique restriction sites present within the FMDV 2A sequence were used to create a series of N-terminally truncated versions of FMDV 2A to define more closely those

sequences necessary to mediate cleavage (pCAT2AGUS/01, /02, /08 and /09; Figure 3). Plasmid DNAs from these constructs were used to programme the TnT rabbit reticulocyte lysate system. Analysis of the translation products (Figure 6A) showed that cleavage of CAT2AGUS occurred in reactions programmed with constructs pCAT2AGUS/01 (~75%), pCAT2AGUS/02 (~70%) and pCAT2AGUS/08 (~65%) but could not be detected in reactions programmed with construct pCAT2AGUS/09.

CAT and GUS activity assays were performed on both rabbit reticulocyte and cell lysate translation mixtures. Translation reactions programmed with pCATGUS or pCAT2AGUS CAT contained approximately equal levels of CAT activity (data not shown). In the case of reactions programmed with pCATGUS all of the CAT is present as the polyprotein [CATGUS], whereas in the case of pCAT2AGUS, 20% is present as uncleaved [CAT2AGUS] and 80% of the CAT protein is in the form of CAT with a C-terminal extension of FMDV 2A ([CAT2A]). This constitutes indirect evidence that [CAT2A] possessed CAT activity. Similarly, GUS activities were equal in cell lysates or translation mixtures programmed with pCATGUS or pCAT2AGUS, indicating that the GUS cleavage product (with an additional N-terminal proline residue) is also active (data not shown). Correlation of enzyme activities with levels of protein was not possible as the commercially available CAT ELISA kit used did not recognize CAT in the form of a [CATGUS] polyprotein.

#### **Discussion**

Rabbit reticulocyte lysate translation mixtures and vTF7-3 infected HTK-143 cells programmed with pCATGUS

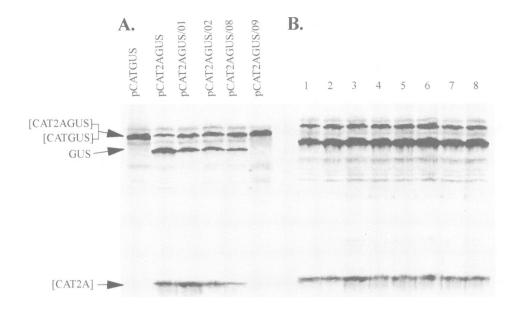


Fig. 6. Cleavage of N-terminally truncated 2A insertions and stability of translation products. TnT rabbit reticulocyte lysate systems were programmed with plasmid DNA from constructs shown in Figure 3 (A). TnT rabbit reticulocyte lysate systems were programmed with pCAT2AGUS. Translation reactions were performed for either 30 min (lanes 1-4) or 60 min (lanes 4-8) and then translation arrested. Mixtures were then either frozen (lanes 1 and 5) or further incubated for 30 min (lanes 2 and 6), 60 min (lanes 3 and 7), 90 min (lanes 4 and 8) and then frozen (B).

produced the single major polyprotein product [CATGUS], whereas pCAT2AGUS produced three major products. The [CAT2AGUS] product was identified by its reaction with anti-CAT and anti-GUS antibodies. This polyprotein did not, however, react with anti-FMDV 2A antibodies—presumably the FMDV 2A sequence epitope(s) is masked in this context. The other products were identified by immunoprecipitation as GUS and [CAT2A]. The simplest explanation for these data is that the FMDV 2A oligopeptide sequence present within the [CAT2AGUS] synthetic polyprotein is able to mediate cleavage at its C-terminus in a fashion similar to its role in native FMDV polyprotein processing. It is clear that FMD 2A does not require other domains within the FMDV polyprotein to function and is an autonomous element capable of mediating cleavage, even in a completely foreign context. This cleavage occurs in a range of heterologous expression systems: rabbit reticulocytes, human HTK-143 cells (this paper), wheat-germ extracts (data not shown) and insect cells (Roosien et al.,

FMDV 2A is clearly able to mediate cleavage and its small size may lead one to the conclusion that this sequence is merely a substrate for a host-cell proteinase. The FMD 2A region does, however, show high similarity to the C-terminal region of the ~ 10-fold larger 2A protein of another genus of the picornaviridae, the cardioviruses. The N-terminal residue of aphtho- and cardiovirus 2B proteins, the P1' residue, is invariantly proline. In both virus groups a primary cleavage occurs between the capsid protein precursor (P1-2A) and the replicative domains of the polyprotein (2BC and P3) at the C-terminus of the 2A region of the polyprotein (Figure 1). In neither case are precursors spanning the 2A/2B cleavage site detected during native polyprotein processing. Studies on cardiovirus polyprotein processing have shown the 2A/2B cleavage to possess characteristics of a cotranslational cleavage in cis-insensitivity to dilution and resistance to a range of proteinase inhibitors (Palmenberg

and Rueckert, 1982; Jackson, 1986; Batson and Rundell, 1991). The oligopeptide comprising the C-terminal 14 residues of the EMC 2A protein together with the N-terminal proline residue of EMC protein 2B also mediates cleavage in a foreign context, although to a lower level than FMD 2A (G.P.Thomas and M.D.Ryan, in preparation). Taken together, these data are highly suggestive that this cleavage is not mediated by a cellular proteinase, or, that such an enzyme would be tightly coupled to translation to account for the rapidity of the cleavage. We believe FMDV 2A represents a polyprotein region whose sole function is a single cleavage at its own C-terminus, functioning only *in cis*, after which it is cleaved from the capsid protein 1D by another virus encoded proteinase, 3C<sup>pro</sup> or 3CD<sup>pro</sup> (Ryan *et al.*, 1989).

No precursors spanning the FMDV 2A/2B cleavage site can be detected during FMDV polyprotein processing (Ryan et al., 1989), although deletions upstream of the 2A region reduce cleavage at the 2A/2B site to ~80% (Ryan et al., 1991). Analysis of the N-terminally truncated forms of FMDV 2A show that the 13 residue oligopeptide (-LKL-AGDVESNPGP-) is able to mediate cleavage whereas the 11 residue oligopeptide (-LAGDVESNPGP-) is not. Residues comprising the N-terminal portion of the FMDV 2A region whilst being influential in, but not critical for, cleavage activity have an additional role in forming the 1D/2A cleavage site utilized by 3C<sup>pro</sup> to cleave the 2A oligopeptide from the capsid protein 1D at a later stage of FMDV polyprotein processing (Figure 1C).

Interestingly, the uncleaved [CAT2AGUS] molecule appears not to undergo subsequent cleavage and is stable in reticulocyte lysates. This does not, however, resolve the question as to whether 2A is an 'active' element or a 'passive' substrate, since it may be argued that the stability of the uncleaved form (not precursor, *sensu stricto*) may reflect either competence to cleave itself or be cleaved by a cellular enzyme. Our data indicate that a certain proportion of

molecules are competent to cleave, or be cleaved, this reaction occurring in a rapid cotranslational manner and not at a later, post-translational, stage.

We have demonstrated that FMDV 2A does not act as a substrate for a proteinase located elsewhere within the FMDV polyprotein (L<sup>pro</sup>, 3C<sup>pro</sup>) or absolutely require other FMDV polyprotein domains for activity. Three explanations may account for a cotranslational cleavage associated with such a short sequence: (i) FMDV 2A functions as a substrate for a cellular proteinase, which, to account for the observed cleavage kinetics, would need to be closely coupled to translation, (ii) the FMDV 2A sequence in some manner disrupts the normal peptide bond formation during translation, or (iii) the FMD 2A sequence possesses an entirely novel type of proteolytic activity.

An element as small as FMDV 2A may possess a 'single turnover' activity since it would not require those domains present in enzymes necessary for the repeated regeneration of nucleophilic species within active sites, the modulation or activation of (proteolytic) activity, or substrate recognition in trans. In which case, FMDV 2A together with the N-terminal proline residue of protein 2B may represent an entirely novel, enzyme-independent, type of protein cleavage in vivo.

Although the mechanism of FMDV 2A mediated cleavage is not yet understood, its utility is apparent. The strategy of self-processing polyproteins exemplified in the positive stranded RNA viruses may be adapted in many applications to enable the co-ordinated and stoichiometric expression of multiple proteins from a single ORF.

#### Materials and methods

The following reagents were obtained from commercial suppliers: TAQ polymerase (Boehringer Mannheim); T7 polymerase, T4 DNA ligase, restriction enzymes, pGEM vectors, TnT rabbit reticulocyte lysate (Promega); plasmid pRAJ255 (Clontech); [35S]methionine, 35S-Express (NEN DuPont); anti-CAT, anti-GUS antibodies (5'-3' Inc.). Oligonucleotides were synthesized on an Applied Biosystems 381A using phosphoramidite chemistry. Anti-2A antiserum raised against the synthetic peptide NH<sub>2</sub>-LLNFDLLKLAGDVESNPGP-COOH was a kind gift of Drs T.Doel and T.Collen.

#### Reporter gene constructs

pMR90. Oligonucleotides OR5 (5'-CCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGT-3') and OR6 (5'-CTCCCGCAAGCTTAAGAAGTCAAAATTCAACAGCTGGCATG-3') when annealed form an Sphl—AaiII linker encoding the three C-terminal amino acids of FMDV capsid protein 1D (-Q-L-L-) together with the first 10 amino acids of FMDV 2A. Plasmid pMR77 (Ryan et al., 1991) was restricted with SphI and AaiII and the OR5/OR6 linker inserted to form plasmid pMR84. The SphI—ApaI restriction fragment encoding FMDV 2A was ligated into pGEM7zf+ similarly restricted, forming pMR90. The pMR90 construct encoding FMDV 2A and flanking sequences are shown in Figure 3A.

pCAT20/21. The CAT gene of pBR328 was amplified by the PCR using oligonucleotides CAT20 (5'-GCGCGGATCCATGGAGAAAAAATGACTGGA-3') which introduced a BamHI restriction site immediately prior to the CAT coding sequence, and CAT21 (5'-ATATAGGGCCCAAATCTA-GACGCCCCGCCCTGCAA-3') which removed the stop codon and introduced XbaI and ApaI restriction sites immediately following the CAT coding region, shown in Figure 3. PCR products were restricted with BamHI and ApaI and ligated into pGEM7zf+ similarly restricted.

pGUS12/23. The GUS of pRAJ255 was amplified by the PCR using oligonucleotides GUS12 (5'-CGCGCGGGGCCCATGTTACGTCCT-GTAG-3') which introduced an ApaI restriction site immediately prior to the GUS coding region, and GUS23 (5'-ATATAGGGCCCAAATCTA-GATTGTTTGCCTCCCTG-3') which introduced ApaI and XbaI restriction

sites as described above (Figure 3). PCR products were restricted by ApaI and XbaI and ligated into pGEM5zf+ vector restricted with ApaI and SpeI.

pCATGUS. The BamHI-ApaI restriction fragment of pCAT20/21 containing the CAT gene was ligated together with the ApaI-NsiI restriction fragment of pGUS12/23 containing the GUS gene into pGEM11zf+ vector restricted with BamHI and NsiI.

pCAT2AGUS. pCATGUS was restricted with ApaI and XbaI and the ApaI – XbaI restriction fragment of pMR90, encoding FMDV 2A, was inserted.

pCAT2AGUS/01, 02, 08, 09. Double stranded oligonucleotide 'adaptor' molecules were formed by annealing two oligonucleotides (50 pM each) and ligated into pCAT2AGUS restricted with XbaI and HindIII. pCAT2A-GUS/01, oligonucleotides OR64 (5'-CTAGAAATTTTGACCTTCTTA-3') and OR65 (5'-AGCTTAAGAAG-GTCAAAATTTT-3'). pCAT2AGUS/02, oligonucleotides OR66 (5'-CTAGAGACCTTCTTA-3') and OR67 (5'-AGCTTAAGAAGGTCT-3'). pCAT2AGUS/08, oligonucleotides OR82 (5'-CTAGACTT-AAGCTTGCGGGAGACGT-3') and OR83 (5'-CTC-CCGCAAGCTTAAGT-3'). pCAT2AGUS/09, oligonucleotides OR84 (5'-CTAGACTTGCGGGAGACGT-3') and OR85 (5'-CTC-CCGCAAGCTTAGGT-3').

#### Transcription and translation

Plasmid constructs were used to programme a TnT system. Translation reactions (50  $\mu$ l) were performed as per the manufacturer's instructions. Briefly, plasmid DNA (1  $\mu$ g) was used to programme rabbit reticulocyte lysates containing [ $^{35}$ S]methionine (50  $\mu$ Ci) and incubated at 30°C for 1 h. Translation reactions were arrested by the addition of RNase A (1  $\mu$ g) and cyclohexamide (50  $\mu$ g). Protein products were analysed by 11.5% SDS-PAGE (Laemmli, 1970).

#### Expression in human HTK-143 cells

Transient expression of plasmid constructs was as described (Fuerst *et al.*, 1986). Briefly, HTK-143 cells were infected with a recombinant vaccinia virus (vTF7-3) expressing T7 RNA polymerase. Plasmids were transfected with plasmid constructs 1 h later and cells incubated for 20 h. Cells were then incubated in methionine/cysteine free media for 1 h before replacing the media with that containing <sup>35</sup>S-Express. Cells were harvested after 2 h and proteins analysed by 10% SDS-PAGE and autoradiography.

## Immunoprecipitation

Radiolabelled proteins were immunoprecipitated using the method of Barrett et al. (1989).

## **Acknowledgements**

The authors would like to thank Drs A.King and C.Bostock for critical reading of the manuscript and Dr A.Palmenberg for stimulating discussions. This work was supported entirely by the Agriculture and Food Research Council.

#### References

Barrett, T., Belsham, G.J., Subbarao, S.M. and Evans, S.A. (1989) *Virology*, **170**, 11–18.

Batson, S. and Rundell, K. (1991) Virology, 181, 764-767.

Bazan, J.F. and Fletterick, R.J. (1988) Proc. Natl Acad. Sci. USA, 85, 7872-7876.

Fuerst, T.R., Niles, E.G., Studier, F.W. and Moss, B. (1986) *Proc. Natl Acad. Sci. USA*, 83, 8122-8126.

Jackson, R.J. (1986) Virology, 149, 114-127.

Laemmli, U.K. (1970) Nature, 227, 680-685.

Palmenberg, A.C. and Rueckert, R.R. (1982) J. Virol., 41, 244-249.

Robertson, B.H. et al. (1985) J. Virol., 54, 651-660.

Roosien, J., Belsham, G.J., Ryan, M.D., King, A.M.Q. and Vlak, J.M. (1990) J. Gen. Virol., 71, 1703-1711.

Ryan, M.D., Belsham, G.J. and King, A.M.Q. (1989) *Virology*, **173**, 35-45. Ryan, M.D., King, A.M.Q. and Thomas, G.P. (1991) *J. Gen. Virol.*, **72**, 2727-2732.

Sommergruber, W., Zorn, M., Blass, D., Fessl, F., Volkmann, P.T. and Kuechler, E. (1989) *Virology*, **58**, 893-899.

Toyada, H., Nicklin, M.J.H., Murray, M.G., Anderson, C.W., Dunn, J.J., Studier, F.W. and Wimmer, E. (1986) Cell, 45, 761-770.

Received on October 15, 1993