

# The cleavage activities of aphthovirus and cardiovirus 2A proteins

Michelle L. L. Donnelly,<sup>1</sup> David Gani,<sup>1</sup> Mike Flint,<sup>2</sup> Susan Monaghan<sup>2</sup> and Martin D. Ryan<sup>3</sup>

<sup>1</sup> School of Chemistry, University of St Andrews, Purdie Building, North Haugh, St Andrews KY16 9ST, UK

<sup>2</sup> Department of Biochemistry, School of Biology and Medical Sciences, University of St Andrews, Irvine Building, North Street, St Andrews KY16 9AL, UK

<sup>3</sup> BBSRC Institute for Animal Health, Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF, UK

The primary 2A/2B polyprotein cleavage of aphtho- and cardioviruses is mediated by their 2A proteins cleaving C-terminally. Whilst the aphthovirus 2A region is only 16 aa (possibly 18 aa) long, the cardiovirus 2A protein is some 150 aa. We have previously shown that foot-and-mouth disease virus (FMDV) 2A is able to mediate cleavage in an artificial (chloramphenicol acetyltransferase/FMDV 2A/ $\beta$ -glucuronidase [CAT-2A-GUS]) polyprotein system devoid of any other FMDV sequences with high (~85%), although not complete, cleavage. In this paper we show that insertion of upstream FMDV capsid protein 1D sequences increases the activity. In addition, we have demonstrated that the cardiovirus Theiler's murine encephalomyelitis virus (TME) 2A protein, when linked to GUS in a single ORF, is able to cleave at its own C terminus with high

efficiency – if not completely. The C-terminal 19 aa of TME 2A, together with the N-terminal proline residue of protein 2B, were inserted into the CAT/GUS artificial polyprotein system (in a single ORF). This recombinant [CAT- $\Delta$ TME2A-GUS] polyprotein was able to mediate cleavage with high (~85%) efficiency – directly comparable to the activity observed when FMDV 2A was inserted. A similar insertion into [CAT-GUS] of the C-terminal 19 aa of the cardiovirus encephalomyocarditis virus (EMC) 2A, together with the N-terminal proline residue of protein 2B, produced a [CAT- $\Delta$ EMC2A-GUS] polyprotein which also mediated cleavage at ~85%. Analysis of the products of expression of these artificial polyproteins in a prokaryotic translation system did not, apparently, reveal any GUS cleavage product.

## Introduction

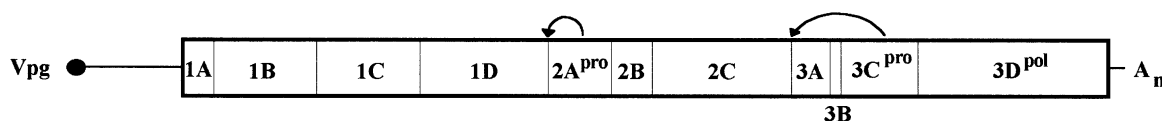
Picornavirus genomes contain a single, long, open reading frame encoding a polyprotein of some 225 kDa, although full-length translation products are not normally observed due to extremely rapid 'primary' intramolecular cleavages (in *cis*) mediated by virus encoded proteinases. Similarities between cellular serine proteinases and both the 2A (entero-, rhinoviruses) and 3C proteinases (all picornaviruses) can be demonstrated by sequence alignments (Gorbalenya *et al.*, 1986; Bazan & Fletterick, 1988), or by structural analyses (Allaire *et al.*, 1994; Matthews *et al.*, 1994). 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; Fig. 1). This cleavage is mediated by

a virus encoded proteinase (2A<sup>pro</sup>), of some 17 kDa, cleaving at its own N terminus (Toyoda *et al.*, 1976; Sommergruber *et al.*, 1989) and recently demonstrated to be a zinc-containing enzyme (Sommergruber *et al.*, 1994). In contrast, the aphtho-, or foot-and-mouth disease viruses (FMDV) and cardiovirus polyproteins undergo a primary polyprotein cleavage at the C terminus of the 2A region between the capsid protein precursor [(P1–2A) – aphthoviruses; (L-P1–2A) – cardioviruses] and 2BC and P3 (Fig. 1). Precursors spanning the 2A/2B cleavage site are not detected during native polyprotein processing. Whilst the cardiovirus 2A protein (~15 kDa) is comparable in size to the 2A proteinases of the entero- and rhinovirus groups, no sequence similarity is observed. Although 2A proteins are highly conserved amongst Theiler's murine encephalomyelitis (TME) viruses and amongst encephalomyocarditis (EMC) viruses, only the C-terminal region is highly conserved across the cardiovirus group. The C-terminal region of cardiovirus 2A is, however, highly similar to the much shorter 2A region of

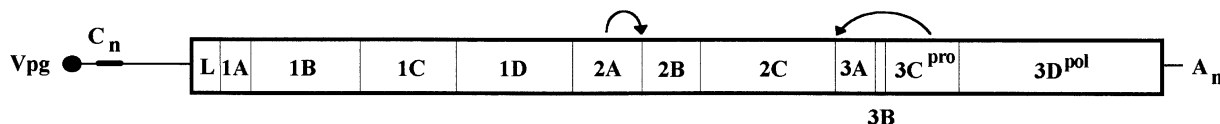
**Author for correspondence:** Martin D. Ryan.

Fax +44 1334 463400. e-mail martin.ryan@st-and.ac.uk

## Entero-, Rhinoviruses



## Cardioviruses



## Aphthoviruses

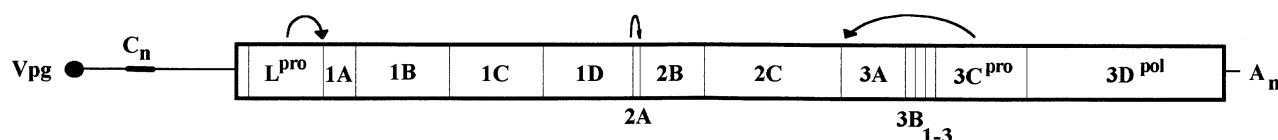


Fig. 1. Picornavirus primary polyprotein cleavages. The 5' non-coding region is capped by an oligopeptide VPg (or 3B). The single, long, open reading frame and polyprotein organization are shown (boxed areas) for the entero-, rhino-, cardio- and aphthovirus groups. Arrows indicate primary polyprotein cleavage sites and the virus proteins responsible.

FMDV (see Fig. 5a). The FMDV 2A region is totally conserved amongst all aphthovirus genomic RNAs sequenced to date (N. Knowles, personal communication) – at variance with published cDNA sequences (Carroll *et al.*, 1984; Robertson *et al.*, 1985). Interestingly, the three C-terminal residues of aphtho- and cardiovirus 2A proteins are completely conserved (-NPG-) whilst the N-terminal residue (proline) of the 2B proteins of both groups is, again, completely conserved.

The FMDV 2A oligopeptide is cleaved from P1-2A by 3C<sup>pro</sup>, although more efficiently by 3CD<sup>pro</sup> (Ryan *et al.*, 1989). In the literature, the C terminus of capsid protein 1D was determined by C-terminal degradation of radiolabelled 1D protein (Bachrach *et al.*, 1973; Kurz *et al.*, 1981) and combined with N-terminal amino acid sequencing of the 2B protein (Robertson *et al.*, 1985; Palmenberg *et al.*, 1992) the 2A region of FMDV was generally accepted as 16 aa long. In the opinion of this author (M.R.) this is possibly incorrect. Analysis of the occurrence of known FMDV 3C<sup>pro</sup> scissile amino acid (Q/X) pairs shows that in all FMDV sequences available such a pair is present 2 aa N-terminal of the site determined (L/N) by C-terminal analysis of 1D – decreasing the length of 1D by 2 aa and increasing the length of the 2A region to 18 aa. Interestingly, the recently sequenced equine rhinoviruses (ERV) 1 and 2 both possess an 'FMDV-like' short 2A region with high sequence identities together with an amino acid pair forming a potential 3C<sup>pro</sup> cleavage site in the same position as in FMDV (ERV-1 – Q/C; ERV-2 – E/G; see Fig. 5).

Previous analyses of recombinant FMDV polyproteins showed the FMDV 2A sequence (together with the N-terminal

proline of protein 2B) to retain activity when either the upstream or downstream contexts were replaced, but that the upstream context did, apparently, influence the activity (Ryan *et al.*, 1991). The contribution of any other FMDV polyprotein domains to this activity was eliminated by the finding that when the FMDV 2A region, together with the N-terminal proline of protein 2B, was inserted in between the reporter proteins CAT (chloramphenicol acetyltransferase) and GUS ( $\beta$ -glucuronidase) in a single, long, open reading frame a cotranslational cleavage was observed (Ryan & Drew, 1994). In this study the cleavage between CAT2A and GUS was estimated at ~ 85%. Similarly, analysis of recombinant FMDV polyproteins had indicated that replacement of the upstream sequences depressed the level of cleavage from 100% to ~ 85% (Ryan *et al.*, 1991). Other studies have shown; (i) deletion of the N-terminal 66% of EMC 2A did not abrogate cleavage at the 2A/2B site (Palmenberg *et al.*, 1992), (ii) mutations within the conserved -NPGP- sequence at the extreme C terminus of EMC 2A abolished cleavage activity (Palmenberg *et al.*, 1992), (iii) microsequence analyses showed the 2A/2B cleavage site as a glycyl-prolyl pair (-NPG/P-; Robertson *et al.*, 1985; Palmenberg *et al.*, 1992), (iv) proteolysis at the 2A/2B site of TME virus was highly efficient when only 2A and 2B sequences were present (Batson & Rundell, 1991) and (v) that insertion of 4 aa towards the N terminus of TME 2A did not apparently affect polyprotein processing (Roos *et al.*, 1989).

FMDV 2A shows a high degree of amino acid similarity with the C-terminal region of cardiovirus 2A. Since the FMDV

2A oligopeptide sequence is able to mediate cleavage with high efficiency independently of all other FMDV sequences, one hypothesis we wished to test was that the equivalent domain of cardiovirus 2A should possess a similar activity. Indeed, if this proved to be the case, why do cardioviruses possess a 2A region of 150 aa whilst those of aphthoviruses are a mere 18 aa?

In this paper we describe experiments designed to determine the influence of the upstream sequences on aphtho- and cardiovirus 2A activity by insertion of sequences derived from FMDV protein 1D into sequences encoding the CAT2AGUS polypeptide in a position immediately upstream of 2A – restoring the native upstream context of 2A. These data were compared with those derived from constructs encoding the entire TME 2A polypeptide linked to GUS, or the C-terminal 18 aa of TME 2A inserted between CAT and GUS.

## Methods

■ **Materials.** All restriction enzymes and coupled transcription/translation (TNT) rabbit reticulocyte lysates were obtained from Promega whilst oligonucleotides were obtained from commercial suppliers (R & D Systems, GIBCO-BRL). Plasmid pTMEGDVII was a kind gift of T. D. K. Brown.

■ **Plasmid constructs.** A list of oligonucleotide primers used in this study is given in Table 1.

(a) **FMDV 2A constructs.** pMD2. Plasmid pGUS 17/22 (Ryan & Drew, 1994) was doubly restricted with *Eco*RI and *Bam*HI. The double-stranded oligonucleotide adaptors OR88 and OR89 were inserted giving an additional 7 aa N-terminal of the natural initiating methionine of GUS – creating plasmid pGUS 17/22/88/89. This plasmid was doubly restricted with *Apa*I and *Xmn*I and the restriction fragment

encoding the modified GUS gene isolated. Plasmid pCAT2AGUS was similarly restricted and the fragment containing the CAT2A encoding sequences isolated. Upon ligation of these two fragments plasmid pMD2 was created.

pTG393. Sequences encoding the entire FMDV capsid protein 1D, other than the N-terminal 32 aa, together with the 2A region were amplified from plasmid pMR15 (Ryan *et al.*, 1989) using oligonucleotides OR393 and OR48. The PCR product was doubly restricted with *Xba*I and *Apa*I and ligated into pMD2 similarly restricted.

pTG394. Sequences encoding the C-terminal 39 aa of protein 1D together with the 2A region were amplified from plasmid pMR15 (Ryan *et al.*, 1989) using oligonucleotides OR394 and OR48. The PCR product was doubly restricted with *Xba*I and *Apa*I and ligated into pMD2 similarly restricted.

pTG395. Plasmid pTG394 was doubly restricted with *Age*I and *Xba*I and the large fragment gel purified. Oligonucleotides TG5 and TG6 were annealed (100 fmol each) and ligated to the purified restriction fragment.

pUCCATGUS. Plasmid pCATGUS was restricted with *Bam*HI and *Nde*I and the fragment encoding CATGUS purified and ligated into pUC8 similarly restricted.

pUCCAT2AGUS. Plasmid pCAT2AGUS was restricted with *Bam*HI and *Nde*I and the fragment encoding CAT2AGUS purified and ligated into pUC8 similarly restricted.

(b) **Cardiovirus constructs.** pTME2AGUS. The 2A region of TME strain GDVII was amplified from plasmid pT7GDVII by PCR using primers OR98 (introducing a *Bam*HI restriction site adjacent to the ATG start codon) and OR99. When the PCR product was restricted with *Bam*HI, however, a *Bam*HI site was present within the TME 2A sequence – created by a T → C mutation at nt 3937 (Pevear *et al.*, 1988; Law & Brown, 1990). The PCR was repeated using oligonucleotides OR98a (introducing a *Bgl*II restriction site adjacent to the ATG start codon)

**Table 1.** Oligonucleotides used in this study

The oligonucleotides were used either for PCR amplification of DNA, or as double-stranded adaptors.

| Oligonucleotide | Nucleotide Sequence (5' → 3')                                        |      |
|-----------------|----------------------------------------------------------------------|------|
| OR48            | 5' - GATATCGGCGTCGGCCAGCATGATGGC                                     | - 3' |
| OR98            | 5' - TTTTGGATCCATGCCCCGATCTCTCTACCGC                                 | - 3' |
| OR98a           | 5' - TTTTGTAGATCTATGCCCCGATCTCTCTACCGC                               | - 3' |
| OR99            | 5' - TTTTGGATCCGGAGGGCCCGGGGTTTCAATTTCTACATC                         | - 3' |
| ORM1            | 5' - GCGCGCTCTAGATACCATGCTGACTACTAC                                  | - 3' |
| ORM2            | 5' - GCGCGCGGGCCCGGGGTTT [A/G] TTTC                                  | - 3' |
| ORM3            | 5' - GCGCGCGGATCCATGGAGAAAAA                                         | - 3' |
| ORM4            | 5' - GCGCGCGGATCCTCACCCGGGGTTGGACTC                                  | - 3' |
| OR393           | 5' - TTTTCTCTAGATCGTTTCATCATGGACAGA                                  | - 3' |
| OR394           | 5' - TTTTCTCTAGAGTCACCGAGTTGCTTTAC                                   | - 3' |
| TG5             | 5' - CTAGAGCATGCGCA                                                  | - 3' |
| TG6             | 3' - TCGTACGCGTGGCC                                                  | - 5' |
| EMC-F           | 5' -CTAGACACTACGCTGGTTACTTTGCGGACCTACTGATTCATGACATTGAGACAAATCCAGGGCC | - 3' |
| EMC-R           | 3' - TGTGATGCGACCAATGAAACGCCTGGATGACTAAGTACTGTAACCTGTCTTAGGTC        | - 5' |

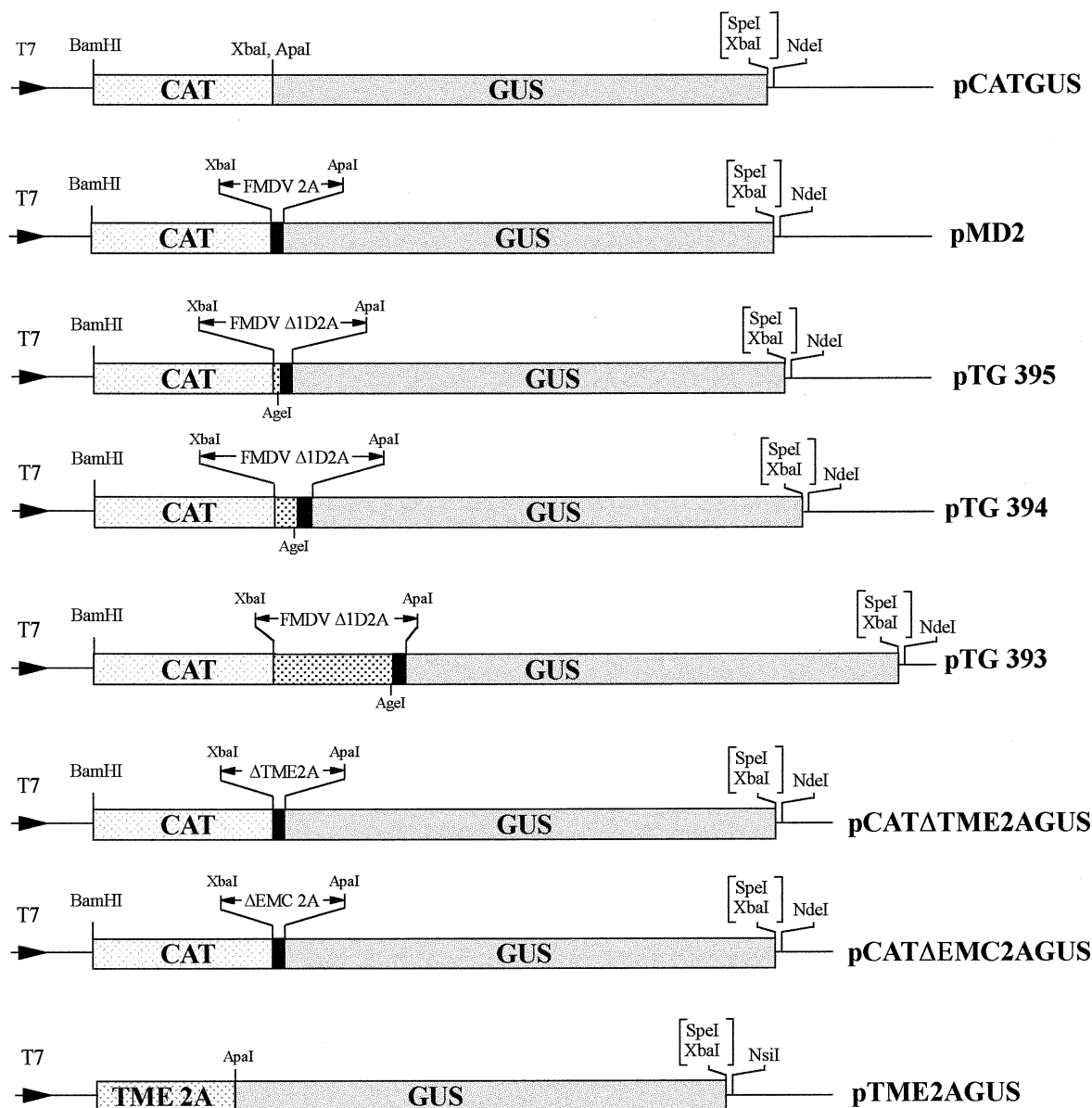


Fig. 2. Plasmid constructs. Boxed areas indicate the single open reading frames encoding the artificial polyproteins; solid black boxes, 2A sequences; dark stippling, FMDV 1D sequences; lighter stippling, entire TME 2A.

and OR99. The PCR product was doubly restricted with *Bgl*II and *Apa*I and ligated into pCATGUS doubly restricted with *Bam*HI and *Apa*I.

**pCATΔTME2AGUS.** The 2A region of TME was amplified by PCR from pTME2AGUS using oligonucleotide primers ORM1 and ORM2. The PCR product was doubly restricted with *Xba*I and *Apa*I and ligated into pMD2 similarly restricted. Nucleotide sequencing differentiated between pCATΔTME2AGUS constructs with a TME strain BeAn 2A sequence (see Fig. 5) – designated pCATΔTME2AGUS (T), or those with strains DA and GDVII 2A sequence – designated pCATΔTME2AGUS (M).

**pUCTME2AGUS.** DNA encoding TME 2A was amplified by PCR using oligonucleotides OR98a and OR99. The PCR product was doubly restricted with *Bgl*II and *Apa*I and ligated into plasmid pUCCAT2AGUS

restricted with *Bam*HI and *Apa*I (removes CAT-2A and replaces with TME 2A).

**pCATΔEMCV2AGUS.** Plasmid pMD2 was doubly restricted with *Xba*I and *Apa*I and the large restriction fragment gel purified. Oligonucleotides EMC-F and EMC-R were annealed (100 fmol each) and ligated together with the purified restriction fragment.

**(c) Control constructs. pCAT2A.** The CAT2A region of pMD2 was amplified by PCR using oligonucleotide primers ORM3 and ORM4. The PCR product was restricted with *Bam*HI and ligated into pGEM-11zf(+) similarly restricted.

**pUCGUS.** pGUS 17/22 (Ryan & Drew, 1994) was doubly restricted with *Bam*HI and *Nde*I and the fragment encoding GUS purified and ligated into pUC8 similarly restricted.

■ **Translation *in vitro*.** Coupled transcription/translation systems, either rabbit reticulocyte lysates or wheatgerm extracts, were programmed with unrestricted plasmid DNA (1 µg) and incubated at 30 °C for 45 min. Translation reactions were arrested by the addition of SDS sample loading buffer.

■ **Prokaryotic expression.** Plasmid DNA was used to transform *E. coli* strain JM109. Cells were grown at 37 °C with agitation for 2 h when IPTG was added to a final concentration of 1 mM. Cells were incubated for a further 2 h, pelleted, and resuspended in ice-cold PBS. Cells were disrupted by sonication and the cell debris was removed by centrifugation. The supernatant was analysed by SDS-PAGE (10%) and proteins transferred to a PVDF membrane by Western blotting. The distribution of GUS protein was determined by probing the blot with anti-GUS antibodies (Clontech) and bound antibodies were visualized by enhanced chemiluminescence (ECL) as per the manufacturer's instructions (Amersham).

■ **Densitometric analyses.** Translation reactions were analysed by SDS-PAGE (10%) and the distribution of radiolabel was determined either by autoradiography or by phosphorimaging using a Fujix BAS 1000. Incorporation of radioactivity into specific products was quantified directly by the latter method.

## Results

Plasmid constructs used in this study are shown in Fig. 2 whilst the translation profiles derived from these constructs are shown in Fig. 3. All of the artificial polyproteins referred to

below are encoded by a single open reading frame. The translation products indicated were identified previously by immunoprecipitation (Ryan & Drew, 1994). Analyses of N-terminally truncated constructs showed that other products were derived from internal initiation within the CAT gene (Met<sup>77</sup> and Met<sup>165</sup>; data not shown).

The translation profiles derived from (i) pTG393 (encoding the CAT gene, almost the entire FMDV capsid protein 1D, FMDV 2A, the N-terminal proline of FMDV protein 2B and GUS) and (ii) pTG394 (encoding CAT, the C-terminal 39 aa of 1D, 2A, the N-terminal proline of FMDV protein 2B and GUS) showed the level of 2A/GUS cleavage to be > 99% – complete within the sensitivity of the densitometric analysis (Fig. 3). The [CATΔ1D2A] cleavage products derived from pTG393 and pTG394 are of the expected molecular mass. We determined that the pTG394 product migrating more quickly than [CAT2A] is derived from internal initiation at Met<sup>77</sup> of CAT (data not shown). This compares with the easily detected uncleaved form of the artificial polyprotein of pMD2 (encoding CAT, 19 aa corresponding to FMDV 2A, the N-terminal proline of 2B and GUS [CAT-2A-GUS]). In the translation profile derived from pTG395, encoding CAT, only the C-terminal 5 aa of capsid protein 1D (-APVKQ-), 2A, the N-terminal proline of 2B and GUS showed higher cleavage activity than

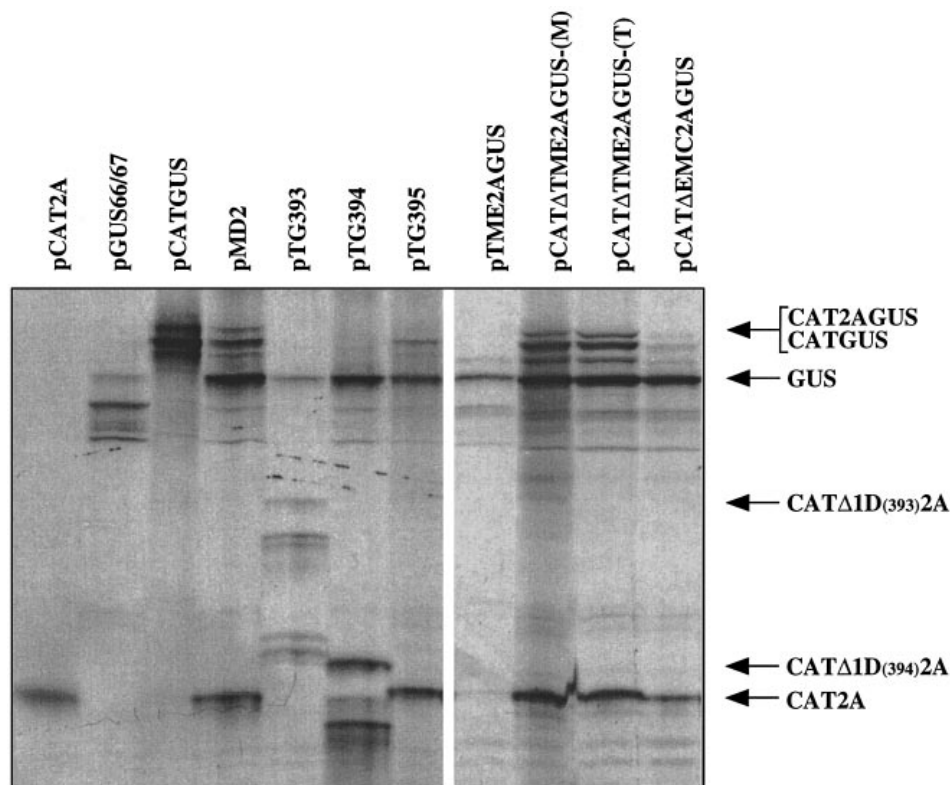


Fig. 3. Translation *in vitro*. Coupled transcription/translation rabbit reticulocyte lysates were programmed with the plasmid constructs indicated. Translation products were analysed by SDS-PAGE (10%). Control, uncleaved and cleaved translation products are indicated, the [CATΔ1D2A] cleavage products derived from pTG393 and pTG394 being indicated appropriately.

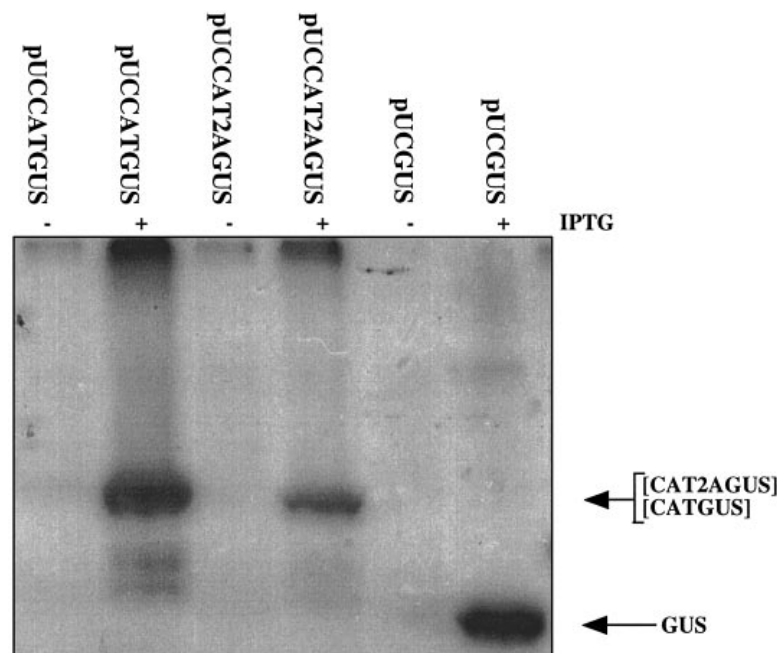


Fig. 4. Prokaryotic expression. Artificial polyproteins were expressed in *E. coli* and the cleavage activities assessed by probing a Western blot of total cellular proteins with anti-GUS antibodies.

pMD2, although some uncleaved form could be detected – cleavage activity of pTG395 was estimated at  $\sim 98\%$ . The [CAT $\Delta$ 1D2A] product in this case migrated only somewhat more slowly than [CAT2A].

The translation profile derived from pTME2AGUS, encoding the entire TME 2A protein, the N-terminal proline of TME protein 2B and GUS, showed the level of cleavage to be  $> 99\%$ , again ostensibly complete cleavage. The translation profiles derived from pCAT $\Delta$ TME2AGUS (M) and (T), encoding CAT, the C-terminal 18 aa of the TME 2A protein (either strain BeAn or DA), the N-terminal proline of TME protein 2B and GUS, showed a very similar pattern to that of pMD2;  $\sim 85\%$  cleavage. Similarly, the translation profile derived from pCAT $\Delta$ EMC2AGUS, encoding the C-terminal 18 aa of the EMC 2A protein, the N-terminal proline of EMC protein 2B and GUS, showed  $\sim 85\%$  cleavage.

Prokaryotic expression of plasmids encoding the artificial polyprotein [CAT-2A-GUS] and the control constructs pUCCGUS and pUCCATGUS was achieved by simply cloning the inserts into pUC8 restricted with *Bam*HI and *Nde*I – using the *lac* promoter and ribosome binding site upstream of the  $\alpha$ -peptide of  $\beta$ -galactosidase sequences within the pUC plasmid. The proteins expressed contained, therefore, the N-terminal 10 aa of the  $\alpha$ -peptide of  $\beta$ -galactosidase. Expression products were visualized by Western blotting and probing with anti-GUS antibodies. An attempt was made to probe with anti-CAT antibodies but cross-reactivities with other *E. coli* proteins masked the signal. The results shown in Fig. 4 indicate that the GUS protein was readily detected together with the [CATGUS] control construct, but only in the uncleaved form of [CAT-2A-GUS]. Neither FMDV 2A (this paper), the complete TME 2A, nor the truncated form of TME 2A appears

to cleave in *E. coli* (data not shown). These data are consistent with prior observations seen in other types of artificial polyprotein systems; glutathione S-transferase/2A/GUS [GST-2A-GUS], kanamycin/2A/GUS [KAN-2A-GUS] and the  $\alpha$ -peptide of  $\beta$ -galactosidase/2A/GUS [ $\beta$ -GAL-2A-GUS] (data not shown; B. McKeown, 1996).

## Discussion

Examination of an alignment of cardiovirus 2A proteins shows a high degree of identity between TME viruses, between the EMC viruses but not between EMC and TME viruses – with the exception of the C-terminal region which, in addition, has identity with the FMDV and ERV 2A regions (Fig. 5a). We have shown that the FMDV 2A region together with the N-terminal proline of protein 2B is able to mediate cleavage in a completely heterologous context (Ryan & Drew, 1994). In this paper we extend this observation demonstrating that the corresponding C-terminal regions of both the TME and EMC cardioviruses possess the same cleavage activity.

The original observation that the C-terminal region of the EMC 2A protein could produce a co-translational cleavage was made in a recombinant influenza virus haemagglutinin construct. A sequence encoding the C-terminal region of the EMC 2A protein (14 aa; -FADLLIHDITNPG-) together with the N-terminal proline of protein 2B was inserted between the HA<sub>1</sub> and HA<sub>2</sub> domains of haemagglutinin, maintaining the single open reading frame. Analysis of translation products derived from this construct showed that  $\sim 17\%$  of the translation products were cleaved to [HA<sub>1</sub>-2A] and [HA<sub>2</sub>] (G. P. Thomas & M. Ryan, unpublished observations).

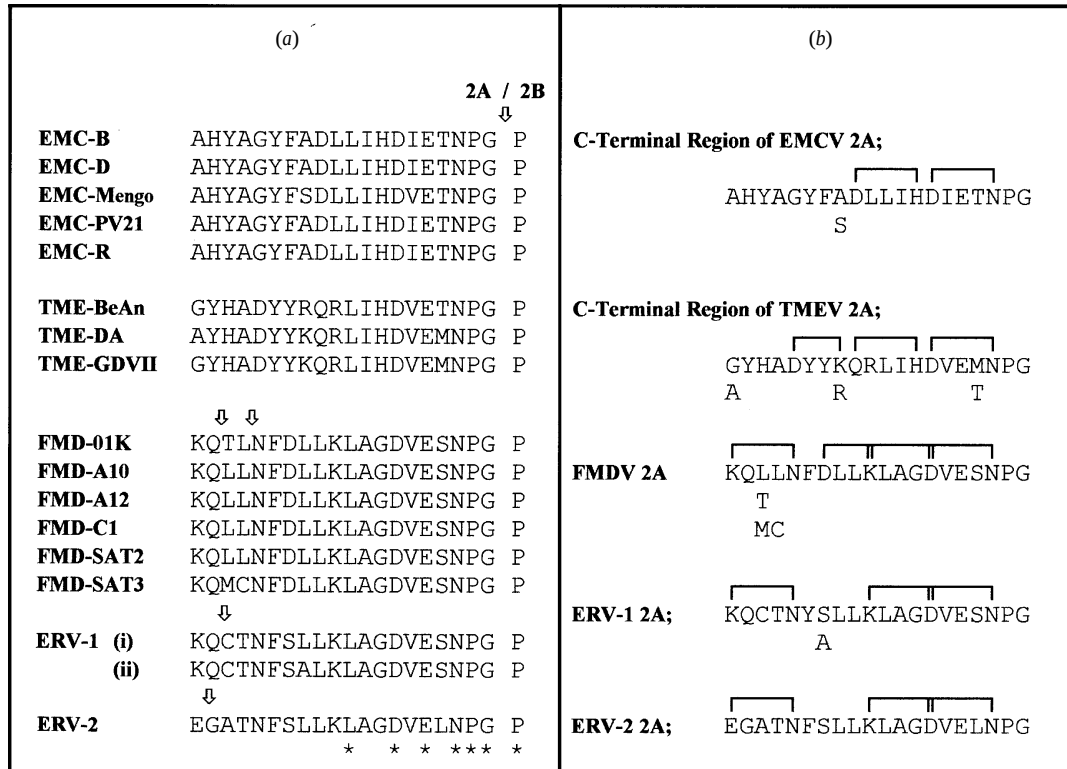


Fig. 5. Sequence alignments of 2A proteins. The C-terminal region of the 2A protein of cardioviruses (TME, EMC) and the aphthovirus 2A region are shown together with the 2A/2B cleavage site. The previously reported FMDV 1D/2A cleavage site is arrowed (-L/N-; Bachrach *et al.*, 1973; Kurz *et al.*, 1981) as is the cleavage site suggested by sequence alignment (-Q/Xaa-), and by comparison with other amino acid pairs known to be cleaved by FMDV 3C<sup>pro</sup>. ERV-1 sequences are from (i) Li *et al.* (1996) and (ii) G. Wutz and others (unpublished; accession no. X96870). The ERV-2 sequence is from G. Wutz and others (unpublished; accession no. X96871; panel a). The FMDV 2A amino acid side-chain/side-chain interactions identified by molecular modelling are indicated, and, by inference, similar interactions in the TME, EMC, ERV-1 and ERV-2 sequences (panel b).

The role of the upstream sequences was highlighted previously (Ryan *et al.*, 1991) and here we show this effect to be localized to within the C-terminal region of capsid protein 1D – consistent with analysis of the EMC 2A activity (Palmenberg *et al.*, 1992). The high level of cleavage activity observed in the [CAT-ΔTME2A-GUS] and [CAT-ΔEMC2A-GUS] polyproteins raises the question what is the function of the remaining ~ 85% of the cardiovirus 2A protein? It seems improbable that its function is to enhance the C-terminal region's cleavage activity to 100% when FMDV 2A can achieve this with a total length of between 24–55 aa. It seems more probable that another function, other than a role in 2A/2B cleavage, could be invoked. Unlike the entero- and rhinoviruses the 2A protein of both aphtho- and cardioviruses forms part of the primary polyprotein cleavage product containing the capsid proteins ([L-P1-2A] – cardioviruses; [P1-2A] – aphthoviruses) and is cleaved from capsid protein 1D during this precursor processing. The nature of the different primary processing patterns between the entero-/rhinoviruses and the aphtho-/cardioviruses implies that any function of the cardiovirus 2A other than the primary 2A/2B cleavage could

be either *cis*-acting in capsid protein precursor processing/ encapsidation, or, perhaps more improbably, *trans*-acting in replication since it is cleaved from the replicative polyprotein domains (P2, P3) in a primary cleavage reaction.

The reason for the inability of cardio- and aphthovirus 2A proteins to mediate cleavage in a prokaryotic system is not clear. The rapidity of 2A mediated cleavage, the wide range of eukaryotic expression systems in which 2A mediated cleavage is observed (Ryan & Drew, 1994) and the stability of the small proportion of uncleaved material in eukaryotic systems argue for a co-translational cleavage – the nature of which appears to be specific for either 80S ribosomes, or a eukaryotic cellular milieu. A number of artificial polyproteins have been examined for cleavage in prokaryotic systems (including polyproteins composed only of soluble *E. coli* proteins and 2A) and we did not observe any detectable cleavage products.

Using secondary structural prediction algorithms the 2A region of FMDV (and the corresponding C-terminal region of the TME and EMC 2A proteins) was predicted to form helices capped at their C termini with a tight turn (-PGP-). This has been more stringently tested by dynamic molecular modelling

studies on FMDV 2A (J. Wilkie, personal communication) which reveal a predicted structure wherein an amphipathic  $\alpha$ -helix is stabilized by a number of side-chain/side-chain interactions (Fig. 5*b*) and hydrophobic interdigitations. Interestingly, one such interaction occurs across the (proposed) 1D/2A 3C<sup>pro</sup> cleavage site (-Q/Xaa-) between the capsid protein 1D C-terminal lysine and an asparagine residue. A constraint on the N-terminal region of FMDV 2A (and the C terminus of 1D) is to maintain the structural requirements of a 3C<sup>pro</sup> cleavage site to permit this highly conserved oligopeptide to be removed from the C terminus of capsid protein 1D – a region that will ultimately form a surface (immunogenic) part of the capsid.

Our experiments with synthetic peptides have failed to show a proteolytic activity although cleavage of a synthetic peptide (NH<sub>2</sub>-Asn-Pro-Gly-Pro-COOH) into Asn-Pro and Gly-Pro dipeptides has been reported (Palmenberg, 1990). The position of this cleavage does not, however, correspond to that observed in biological systems (-Asn-Pro-Gly↓Pro-) but may be giving clues as to the structure of this tight turn motif 'capping' the (putative) helix.

In this paper the activities of the C-terminal region of the cardiovirus type 2A and the aphthovirus type 2A regions were analysed and shown to be indistinguishable. The mechanism by which these oligopeptide sequences mediate cleavage remains unclear. Site-directed mutagenesis studies in our laboratory have identified mutable residues and in most cases these correspond to residues which vary amongst wild-type sequences or between the two virus groups. The data presented here advance our understanding of the mechanism of 2A-mediated cleavage inasmuch that we show that related sequences possess an identical activity, and that the sequences that can bring about complete cleavage (as far as we can detect) are contained within a region somewhat longer than the FMDV, EMC and TME 2A regions described above but certainly shorter than the entire cardiovirus 2A region.

The authors wish to thank staff at the I.A.H. and Institute of Biochemistry, University of Vienna together with W. Sommergruber, E. Wimmer and A. Palmenberg for useful discussions and interest. This work was supported entirely by the Biotechnology and Biological Sciences Research Council.

## References

- Allaire, M., Chernaia, M. M., Malcolm, B. A. & James, M. N. G. (1994). Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. *Nature* **369**, 72–76.
- Bachrach, H. L., Swaney, J. B. & vande Woude, G. F. (1973). Isolation of the structural polypeptides of foot-and-mouth disease virus and analysis of their C-terminal sequences. *Virology* **52**, 520–528.
- Batson, S. & Rundell, K. (1991). Proteolysis at the 2A/2B junction in Theiler's murine encephalomyelitis virus. *Virology* **181**, 764–767.
- Bazan, J. F. & Fletterick, R. J. (1988). Viral cysteine proteases are homologous to the trypsin-like serine proteases: structural and functional implications. *Proceedings of the National Academy of Sciences, USA* **85**, 7872–7876.
- Carroll, A. R., Rowlands, D. J. & Clarke, B. E. (1984). The complete nucleotide sequence of the RNA coding for the primary translation product of foot and mouth disease virus. *Nucleic Acids Research* **12**, 2461–2472.
- Gorbalenya, A. E., Blinov, V. M. & Donchenko, A. M. (1986). Poliovirus-encoded proteinase 3C: a possible evolutionary link between cellular serine and cysteine proteinase families. *FEBS Letters* **194**, 253–257.
- Kurz, C., Forss, S., Kupper, H., Strohmaier, K. & Schaller, H. (1981). Nucleotide sequence and corresponding amino acid sequence of the gene for the major antigen of foot-and-mouth disease virus. *Nucleic Acids Research* **9**, 1919–1931.
- Law, K. M. & Brown, T. D. K. (1990). The complete nucleotide sequence of the GDVII strain of Theiler's murine encephalomyelitis virus (TMEV). *Nucleic Acids Research* **18**, 6707–6708.
- Li, F., Browning, G. F., Studdert, M. J. & Crabb, B. S. (1996). Equine rhinovirus 1 is more closely related to foot-and-mouth disease virus than to other picornaviruses. *Proceedings of the National Academy of Sciences, USA* **93**, 990–995.
- McKeown, B. (1996). *A study of the 2A and 3C mediated cleavage events in the processing of the foot-and-mouth disease virus polyprotein*. PhD thesis, Queen's University, Belfast, UK.
- Matthews, D. A., Smith, W. W., Ferre, R. A., Condon, B., Budahazi, G., Sisson, W., Villafranca, J. E., Janson, C. A., McElroy, H. E., Gribskov, C. L. & Worland, S. (1994). Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding Site, and means for cleaving precursor polyprotein. *Cell* **77**, 761–771.
- Palmenberg, A. (1990). Proteolytic processing of picornaviral polyprotein. *Annual Review of Microbiology* **44**, 603–623.
- Palmenberg, A. C., Parks, G. D., Hall, D. J., Ingraham, R. H., Seng, T. W. & Pallai, P. V. (1992). Proteolytic processing of the cardioviral P2 region: primary 2A/2B cleavage in clone-derived precursors. *Virology* **190**, 754–762.
- Pevear, D. C., Borkowski, J., Calenoff, M., Oh, C. K., Ostrowski, B. & Lipton, H. (1988). Insights into Theiler's virus neurovirulence based upon a genomic comparison of the neurovirulent GDVII and less virulent BeAn strains. *Virology* **165**, 1–12.
- Robertson, B. H., Grubman, M. J., Weddell, G. N., Moore, D. M., Welsh, J. D., Fischer, T., Dowbenko, D. J., Yansura, D. G., Small, B. & Kleid, D. G. (1985). Nucleotide and amino acid sequence coding for polypeptides of foot-and-mouth disease virus type A12. *Journal of Virology* **54**, 651–660.
- Roos, R. P., Kong, W.-P. & Semler, B. L. (1989). Polyprotein processing of Theiler's murine encephalomyelitis virus. *Journal of Virology* **63**, 5334–5353.
- Ryan, M. D. & Drew, J. (1994). Foot-and-mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO Journal* **13**, 928–933.
- Ryan, M. D., Belsham, G. J. & King, A. M. Q. (1989). Specificity of substrate-enzyme interactions in foot-and-mouth disease virus polyprotein processing. *Virology* **173**, 35–45.
- Ryan, M. D., King, A. M. Q. & Thomas, G. P. (1991). Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence. *Journal of General Virology* **72**, 2727–2732.
- Sommergruber, W., Zorn, M., Blaas, D., Fessl, F., Volkmann, P., Maurer-Fogy, I., Pallai, P., Merluzzi, V., Matteo, M., Skern, T. &



**Keuchler, E. (1989).** Polypeptide 2A of human rhinovirus type 2: identification as a protease and characterization by mutational analysis. *Virology* **169**, 68–77.

**Sommergruber, W., Casari, G., Fessl, F., Seipelt, J. & Skern, T. (1994).** The 2A proteinase of human rhinovirus is a zinc containing enzyme. *Virology* **204**, 815–818.

**Toyoda, H., Nicklin, M. J. H., Murray, M. G., Anderson, C. W., Dunn, J. J., Studier, F. W. & Wimmer, E. (1986).** A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* **45**, 761–770.

---

Received 20 June 1996; Accepted 29 August 1996