Cleavage of foot-and-mouth disease virus polyprotein is mediated by residues located within a 19 amino acid sequence

Martin D. Ryan,* Andrew M. Q. King and G. Paul Thomas

AFRC Institute for Animal Health, Pirbright Laboratory, Ash Road, Pirbright, Woking, Surrey GU24 0NF, U.K.

The 2A region of the foot-and-mouth disease virus (FMDV) polyprotein is only 16 amino acids in length. During synthesis of the FMDV polyprotein a primary proteolytic processing event occurs between the 2A and 2B regions of the polyprotein. The activity responsible for this cleavage is not known but it is thought that either an unidentified virus-encoded proteinase may be responsible, or that 2A acts as a substrate for a host cell proteinase. A series of recombinant FMDV polyproteins has been constructed in which sequences to the N- or C-terminal side of the 2A region have been deleted. Analysis of the processing of these polyproteins shows that a 19 amino acid sequence spanning 2A is sufficient to mediate polyprotein cleavage at a site immediately C-terminal to 2A, whereas deletions extending into the 2A region prevent cleavage.

Introduction

The picornavirus family contains four main genera, the entero-, rhino-, cardio- and aaphtho- or foot-and-mouth disease (FMDV) viruses. The FMDV genome is an ssRNA molecule of mRNA sense of approximately 8500 nucleotides. A single long open reading frame of 2332 codons would encode a polyprotein of 259K, but full-length translation products are not observed because the nascent polyprotein undergoes primary (cotranslational) cleavages which produce four products, L, P1-2A, 2BC and P3 (Fig. 1). All picornavirus primary polyprotein cleavages examined so far are mediated by virus-encoded proteinases and are characteristically extremely rapid intramolecular (cis) events (Fig. 1). However, secondary cleavages are thought to be a combination of a series of intra- and intermolecular (trans) cleavages.

The cardio- and aphthoviruses contain a non-structural protein (L) in the N-terminal region of their polyproteins. In aphthoviruses this protein has been identified as a proteinase which cleaves at its C terminus (Strebel & Beck, 1986) and exists in two forms (Lab and Lb) arising from alternative initiation sites (Clarke et al., 1985). The 1A/1B cleavage occurs at a late stage in virus morphogenesis and is concomitant with encapsidation of virion RNA (vRNA). The 2C/3A primary cleavage in cis (Fig. 1) and subsequent cleavages (secondary processing) are mediated by the 3C proteinase (3C\textsuperscript{pro}), and produce processing intermediates and mature proteins.

The 2A proteins of entero- and rhinoviruses have been identified as proteinases (Toyada et al., 1986; Sommergruber et al., 1989) of approximately 17K and are thought to be structurally similar to the small subclass of trypsin-like serine proteinases (Bazan & Fletterick, 1988). The 2A proteinases (2A\textsuperscript{pro}) of the entero- and rhinoviruses perform a primary polyprotein cleavage at their N terminus, between the P1 (capsid protein precursor) and P2 regions (Fig. 1). The 2A protein of the cardiovirus encephalomyocarditis virus (EMCV) is of similar size, and evidence suggests that this protein is a proteinase mediating a primary cleavage, although in this case the cleavage is at its C terminus (Jackson, 1986). This EMCV 2A/2B cleavage has been shown to be both extremely rapid and resistant to a range of proteinase inhibitors (Jackson, 1986). The EMCV 2A protein shows no significant sequence similarity with 2A proteinases of either the entero- or rhinoviruses. In a manner similar to EMCV, a primary cleavage which is mediated by an unknown agent occurs at the C terminus of the FMDV 2A region. In FMDV, however, the 2A region of the polyprotein is only 16 amino acids in length (Robertson et al., 1985), although it does show similarity with the C-terminal region of the 2A protein of EMCV (Fig. 4).

We have shown previously that processing at the 2A/2B cleavage site does not depend upon either of the known FMDV proteinases, L and 3C (Belsham et al., 1990), and have speculated that the 2A peptide may itself have proteolytic activity (Ryan et al., 1989). The purpose of this investigation was to identify more precisely those regions of the FMDV polyprotein involved in the 2A/2B cleavage. Plasmids based on the transcription vector pSP64 (Melton et al., 1984) were constructed in which specific portions of the FMDV polyprotein had been deleted. Cell-free translation reactions were programmed with transcripts from these plasmids and were...
analysed for evidence of cleavage at the 2A/2B site of recombinant polyproteins.

**Methods**

All deletions and insertions used in FMDV cDNA plasmid constructions maintain the reading frame (Fig. 2). Restriction enzyme sites referred to below use the numbering systems described by the original authors (Forss et al., 1984; Carroll et al., 1984). pMR15 encodes the entire FMDV O1K polyprotein coding region in the transcription vector pSP64 (Ryan et al., 1989). pMR53 contains the HindIII to NruI (nucleotides 742 to 1161) fragment of FMDV O1K cDNA and NruI to EcoRI (nucleotides 371 to 4359) fragment of FMDV A10 cDNA inserted into HindIII and EcoRI double-digested pSP64. pMR77: the EcoRI (nucleotides 5149 to poly linker, 3' terminus of FMDV sequences) fragment was deleted from pMR15. pMR64: the ApaI (nucleotides 3666 to 6498) fragment was deleted from pMR15. pMR65: the Smal to PvuII (nucleotides 1640 to 2816) fragment was deleted from pMR53. pMR66: the AartII (nucleotides 3073 to 3649) fragment was deleted from pMR77. pMR67: the ApaI (nucleotides 1009 to 2875) fragment of pMR53 was inserted at the ApaI site (nucleotide 3666) of pMR77 in the orientation indicated in Fig. 1. Before transcription, pMR15 and pMR64 were linearized using HpaI to restrict sequences 3' to those of FMDV; all other FMDV constructs were linearized using EcoRI. Transcription and translation reactions were as described previously (Ryan et al., 1989). Briefly, translation reactions were terminated after 1 h by the addition of RNase and incubated for a further hour in the presence or absence of a processing extract (PE) prepared from FMDV-infected cells.
The 2A/2B cleavage of FMDV polyprotein

Results

Rabbit reticulocyte lysates were programmed with transcripts derived from FMDV cDNA constructs (Fig. 2); the resultant protein profiles are shown in Fig. 3. Constructs were made using cDNA from two serotypes of FMDV, O1K and A10, owing to the distribution of convenient restriction sites. The polyprotein processing characteristics of the two viruses are, however, identical. Recombinant proteins formed by deletion or truncation are indicated by A throughout the text.

Translation of pMR15 transcripts, encoding the entire FMDV serotype O1K coding region, resulted in the normal polyprotein processing pattern. The primary cleavage products P1-2A, 2BC and P3 were formed (Fig. 3), together with processing intermediates (1ABC, 1CD and 3BCD complex) and mature products (1AB, 1C, 1D, 2C, 3CD and 3D).

Plasmid pMR64 was constructed so that the majority of sequences C-terminal to 2A were removed by replacing all of the 2B region, other than the N-terminal proline residue, with 3D sequences (Fig. 2 and 4). Translation of pMR64 transcripts resulted in three major products, L, P1-2A and a cleavage product corresponding to A3D, the truncated form of protein 3D (Fig. 3). Apparently, complete cleavage occurred even though all P2 sequences other than the N-terminal proline of protein 2B had been deleted. Addition of PE resulted in the subsequent processing of P1-2A to produce intermediates 1ABC and 1CD, as well as the mature products 1AB, 1C and 1D (Fig. 3).

Plasmid pMR65 was constructed so that sequences immediately N-terminal to 2A were removed by deletion of most of 1C and all but the two C-terminal residues of protein 1D (Fig. 2 and 4). Transcripts of pMR65 produced the L protein, a doublet corresponding to the
2BC–Δ3A protein (cf. profile obtained with constructs pMR53 and pMR77) and a protein that migrated more slowly than protein 1AB, which corresponded to the truncated form of the P1-2A precursor, 1ABAC–2A (Fig. 3). Therefore, cleavage still occurred, even when 1D sequences upstream of 2A (as described above) were deleted. Unlike pMR15, pMR64 and the control constructs pMR77 and pMR53, in which this cleavage went to completion, translation products of construct pMR65 also contained some of the 1ABAC–2ABC–Δ3A precursor form, a protein in which cleavage at the 2A/2B boundary had not occurred and which migrated slightly more slowly than P1-2A. The reduced efficiency of 2A/2B cleavage after expression of pMR65 indicated that sequences N-terminal of 2A influence cleavage activity. Densitometric analysis showed that approximately 90% of pMR65 translation products were cleaved at the 2A/2B site. Addition of PE to the products of pMR53 resulted in limited processing of the 2BC–Δ3A protein to 2BC. Similar trimming of the 1ABΔC–2ABC–Δ3A precursor was evident at the 2C/3A site upon addition of PE (Fig. 3). In the case of pMR53, the full-length P1-2A precursor was processed to the intermediates 1ABC and 1CD, and the mature products 1AB, 1C and 1D (which in FMDV serotype A10 comigrated in this gel system), whereas the 1ABΔC–2A deletion form of P1-2A from pMR65 was not processed.

To test whether deletions extending into the 2A region (but retaining the 2A/2B cleavage site) would affect cleavage, plasmid pMR66 was constructed in which sequences encoding most of 1D and the first nine residues of 2A were deleted from the control plasmid pMR77 (Fig. 2 and 4). Translation of pMR66 transcripts resulted in the accumulation of a high Mr product corresponding in size to 1ABCΔAΔ2A–2BC–Δ3A, the uncleaved precursor molecule (Fig. 3). Therefore, removal of the first nine amino acids of 2A prevented cleavage of the polyprotein at the 2A/2B site. The addition of PE resulted in no further processing, suggesting that 2A/2B cleavage did not occur in trans. Changes introduced into such a small region may, however, affect the ability of the novel protein to act as a substrate in trans. The data also confirm previous observations that processing by 3Cpro in trans is extremely sensitive to changes in substrate conformation (Ryan et al., 1989).

Construct pMR67 contains sequences encoding the L, P1-2A and 2BCA–3A proteins of FMDV O1K (cf. pMR77), but also encodes a substantial proportion of the P1-2A protein of FMDV A10 (Δ1BCD–2A) inserted immediately C-terminal to the 2A/2B boundary of the FMDV O1K polyprotein (Fig. 2). This created a duplication of most of the capsid region, albeit from a different serotype. The FMDV A10 sequences are
The translation products included the L, P1-2A, and therefore flanked by two 2A sequences, one derived from FMDV O1K the other from FMDV A10 (Fig. 2 and 4). The translation products included the L, P1-2A, and 2BC–A3A proteins derived from the FMDV O1K coding regions (cf. pMR53 and pMR77) and an additional product corresponding to A1B1CD-2A derived from the FMDV O1K coding sequences. The fact that both FMDV O1K P1-2A and FMDV A10 A1B1CD-2A were resistant to PE. As mentioned above, processing by A1B1CD-2A derived from the FMDV A10 sequences proceeded to completion (Fig. 3). Addition of PE resulted in cleavage of the FMDV O1K P1-2A precursor to mature capsid proteins. Precursor polypeptide A1B1CD-2A derived from the FMDV A10 sequences was resistant to PE. As mentioned above, processing by 3C Pro in trans is highly dependent on substrate conformation, and this N-terminal truncation rendered this precursor an unsuitable substrate for exogenous 3C Pro.

Discussion

In many viruses segregation of encapsidation and replicative functions is achieved by temporal control of transcription and translation. In other viruses this is achieved by encoding such functions on separate genome segments or by synthesis of subgenomic transcripts. In the case of viruses which have adopted the replication strategy of encoding their proteins in the form of a polypeptide, proteins which are biochemically or temporally associated characteristically are clustered in discrete domains of the polypeptide. These polypeptide domains are cleaved early in their synthesis by virus-encoded proteinases which act to separate the distinct functions. All picornaviruses have a cleavage activity associated with 2A which appears to have the function of separating the capsid protein precursor from the replicative functions of the polypeptide. Whereas 2A Pro of enteroviruses has other activities in trans, the sole function of FMDV 2A may be this primary polypeptide cleavage.

We show here that some or all of the first nine residues of 2A are required for cleavage (pMR66). Although replacement of sequences N-terminal to 2A reduced the level of cleavage (pMR65), replacement of sequences downstream of the N-terminal proline residue of 2B did not affect cleavage [pMR67]. Therefore, the sequence mediating the 2A/2B cleavage is located within a 19 amino acid region, but is larger than the C-terminal seven residues of 2A.

The C-terminal region of 1D is variable in sequence (although some residues are conserved) and forms a surface antigenic feature on the virus capsid. The conserved 2A peptide is adjacent to this region and during processing of the P1-2A precursor, 2A is cleaved from 1D by 3C Pro, or more efficiently by 3CD Pro (Ryan et al., 1989). This provides an additional constraint on the 2A region, i.e. the need to form a suitable substrate for 3C Pro. Although the amino acid pairs cleaved by FMDV 3C Pro are more varied than those cleaved by other picornavirus 3C Pro's, the pair forming the 1D/2A cleavage site (L/N) is unusual.

The mechanism by which 2A/2B cleavage occurs in the aphtho- and cardioviruses remains unclear. Two alternative explanations are that 2A represents a self-cleaving sequence, or that it acts as a recognition sequence for a host cell enzyme. Such a host cell activity would need to be (i) both highly efficient and tightly coupled with translation to achieve the rapidity and yield of the 2A/2B cleavage reaction described above, (ii) conserved amongst a range of cell types, (iii) as active in rabbit reticulocyte (Ryan et al., 1989; Grubman & Baxt, 1982) and wheatgerm translation systems (data not shown) as in cells, (iv) show unaltered levels of activity in rabbit reticulocytes diluted up to 320-fold with buffer (Palmengren & Ruekert, 1982), (v) be able to recognize and cleave the 2A/2B site in a variety of sequence contexts (although not pMR67; Fig. 3), (vi) resistant to a range of proteinase inhibitors (Jackson, 1986) and (vii) act on only a proportion (about 90%) of the recombinant polypeptide encoded by pMR65, a level which remains constant throughout translation (data not shown). Riboso...
somess have been shown to possess a trypsin-like endoproteinase activity (Korant, 1977) which would at least represent the probable cellular location, if not the specific activity, of a proteinase required for such a rapid and efficient cleavage mechanism.

The self-cleavage hypothesis would invoke cleavage as an intrinsic property of 2A. The 16 amino acids of FMDV 2A might represent an active site/substrate couple cleaving at its own C terminus. The 3C\textsuperscript{pro}s of picornaviruses, which possess activity both in trans and in cis are able to cleave at their own C termini, a 'substrate' only 36 amino acids from the active site nucleophile. FMDV 2A might represent an enzyme which has dispensed with those additional sequences necessary for both (i) regulation or modification of enzyme activity and (ii) recognition and binding a substrate in trans. The 2A\textsuperscript{pro} of enteroviruses mediates host cell shutoff by cleavage of the p220 component of the cap-binding protein complex (Etchison et al., 1982). In FMDV this activity is found not in the 2A region, but in a proteinase unique to FMDV, the L protein (Lloyd et al., 1988).

An alternative self-cleavage mechanism is that the 2A region forms a structure which is unstable under physiological conditions. Secondary structure predictions and modeling studies indicate that the 2A sequence may form an \( \alpha \)-helix of seven to 11 residues followed by strong turn-forming residues around the scissile bond. Interestingly, P. V. Pallai has observed cleavage of the synthetic tetrapeptide N-P-G-P (corresponding to the C-terminal region of EMCV and FMDV 2A) to N-P and G-P when incubated in alkaline buffer alone (cited in Palmenberg, 1990).

Whichever, if any, of the alternative mechanisms discussed above proves to be correct, the 2A/2B cleavage of the aptho- and cardioviruses is unique. Synthetic peptide and site-directed mutagenesis studies are in progress to determine those residues within 2A which are critical for cleavage and to elucidate the mechanism of this novel cleavage.

We are grateful to Drs C. Bostock and J. McCauley for their constructive comments, to Drs P. Freemont and D. Stewart for help with structural predictions and modelling studies, and to Dr A. Palmenberg for useful discussions. This work was supported in part by the Ministry of Agriculture, Food and Fisheries.

References


(Received 8 May 1991; Accepted 25 July 1991)