Bioorganic Chemistry **27,** 55–79 (1999)
Article ID bioo.1998.1119, available online at http://www.idealibrary.com on IDEAL®

A Model for Nonstoichiometric, Cotranslational Protein Scission in Eukaryotic Ribosomes

Martin D. Ryan, Michelle Donnelly, Arwel Lewis, Amit P. Mehrotra, John Wilkie, and David Gani[1]

School of Chemistry and Centre for Biomolecular Sciences, Purdie Building, University of St. Andrews, Fife, KY16 9ST, United Kingdom

Received August 17, 1998

The aphthovirus 2A region apparently responsible for the hydrolytic cleavage of a single large polyprotein at a Gly-Pro linkage is only 18 amino acid residues long and is evidently not a proteinase. Here we describe the construction of reporter recombinant polyproteins and provide the results of further mutagenesis experiments designed to test the functions of specific amino acid residues within the foot-and-mouth disease virus (FMDV) 2A region. These results show that a Gly-Pro amide bond is not actually synthesized. The result can be rationalized into a kinetic and structural model for cotranslational aphtho- and cardiovirus polyprotein cleavage in which hydrolysis is mediated by a ribosomally bound 2A polypeptidyl-tRNA molecule at its own 3′-O acyl adenosyl ester linkage. The possible role of the 3-D structure of the 2A polypeptide in preventing peptide bond formation but in allowing the synthesis of the downstream polypeptide sequence is discussed within the context of the new findings.

INTRODUCTION

Picornavirus genomes contain a single, long open reading frame (ORF)[2] encoding a polyprotein of some 225 kDa. Full-length translation products are not normally observed due to rapid “primary” intramolecular cleavages mediated by virus-encoded proteinases. The primary P1/P2 polyprotein cleavage in entero- and rhinoviruses is

[1] Author to whom correspondence should be addressed at School of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom).

[2] Abbreviations used: CAT, chloramphenicol acetyl transferase; COSY, 2-D homonuclear chemical-shift correlation spectroscopy; DCM, dichloromethane; DMSO, dimethyl sulfoxide; eEF2, elongation factor 2; EMC, encephalomycarditis; FMOC, 9-fluorenylmethoxycarbonyl; FMDV, foot-and-mouth disease virus; GUS, β-glucoronidase; HMQC, heteronuclear multiple quantum coherence; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser enhancement spectroscopy; ORF, open reading frame; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; PyBOP, benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate; RR, rabbit reticulocytes; SDM, site-directed mutagenesis; TFA, trifluoroacetic acid; TLC, thin-layer chromatography; TME, Theiler’s murine encephalomyelitis; TP, translation product; WG, wheat germ.

0045-2068/99 $30.00

mediated by the 2A proteinase cleaving at its own N-terminus. Similarities between cellular serine proteinases and the 2A proteinase can be observed by sequence alignments (*1*) or by structural analyses (*2*, *3*). The primary 2A/2B polyprotein cleavages of aphtho- and cardioviruses are, similarly, mediated by their 2A proteins and cleave at the C-terminal (Fig. 1).

The cardiovirus and aphthovirus 2A regions are some 150 and only 18 amino acid residues long, respectively. It is now evident that neither are proteinase enzymes, vide infra. While the cardiovirus 2A protein (ca 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 among Theiler's murine encephalomyelitis (TME) viruses and among encephalomycarditis (EMC) viruses, only the C-terminal region is highly conserved across the cardiovirus group (Fig. 2). The C-terminal region of cardiovirus 2A is, however, highly similar to the much shorter 2A region of foot-and-mouth disease virus (FMDV). The FMDV 2A region is totally conserved among all aphthovirus genomic RNAs sequenced to date (*4*), at variance with published cDNA sequences (*5*, *6*). Interestingly, the last three amino acids at the carboxy termini of aphtho- and cardiovirus 2A proteins are completely conserved (-NPG-), while the N-terminal proline residue of the 2B proteins of both groups is, again, completely conserved (Fig. 2).

Analysis of recombinant FMDV polyproteins synthesized using the eukaryotic translation systems of rabbit reticulocytes (RR) or wheat germ (WG) has shown that the replacement of sequences downstream of the Gly-Pro scissile bond does not impair 2A-mediated cleavage activity. However, the replacement of sequences upstream of 2A reduces activity slightly, to ca. 90% (*7*).

Moreover, certain site-specific substitutions at positions quite remote from the Gly-Pro scissile bond completely turn off activity. The aphtho- and cardiovirus 2A-mediated cleavage activity is, therefore, quite distinct from that of the entero- and rhinoviruses and appears to be quite distinct from any other known proteolytic activities, viral or cellular. The recent finding that aphtho- and cardiovirus 2A-mediated

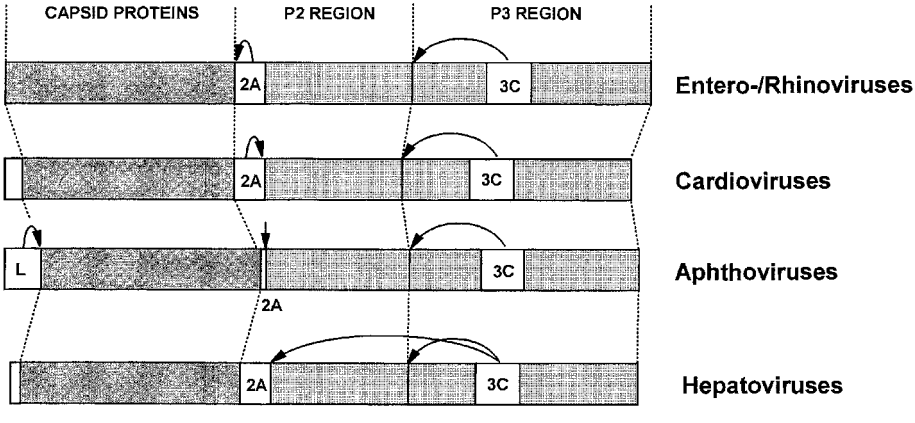


FIG. 1. Picornavirus primary polyprotein cleavages. Box regions indicate polyprotein domains.

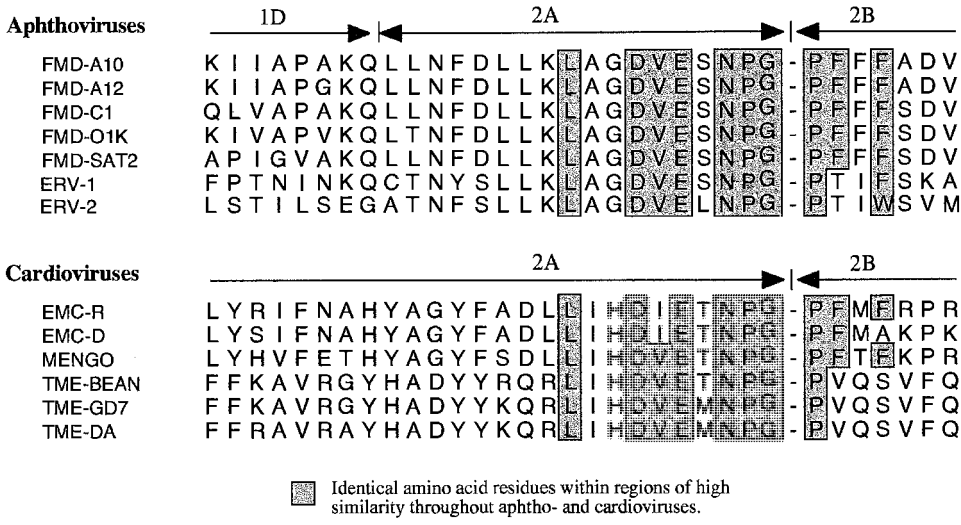


FIG. 2. Conserved amino acid residues in the aligned sequences of the aphthoviral and cardioviral 2A and 2B regions.

cleavage activity is not expressed in prokaryotic translation systems, coupled to the fact that recombinant truncated TME 2A possessing only the last 39 amino acid residues of the native sequence (ca. 130 residues) is fully active in RR and WG translation systems, indicated that the cleavage event occurs cotranslationally in the eukaryotic ribosome (*8*). Here we describe the construction of reporter recombinant polyproteins and provide the results of further mutagenesis experiments designed to test the functions of specific amino acid residues within the FMDV 2A region. These can be rationalized into a kinetic and structural model for cotranslational aphtho- and cardiovirus polyprotein cleavage in which hydrolysis is mediated by a ribosomally bound 2A polypeptidyl-tRNA molecule at its own 3′-O acyl adenosyl ester linkage.

EXPERIMENTAL

NMR spectra were recorded on a Bruker AM-300 spectrometer (^{1}H, 300 MHz; ^{13}C, 75.4 MHz), a Varian Gemini spectrometer (^{1}H, 200 MHz; ^{13}C, 50.3 MHz), a Varian Gemini spectrometer (^{1}CH, 300 MHz; ^{13}C, 75.4 MHz), and a Varian Unity Plus 500 spectrometer (^{1}H, 500 MHz; ^{13}C, 125.6 MHz). ^{1}H NMR spectra were referenced internally to $(C^2H_3)_2SO$ (δ 2.47), ^{2}HOH (δ 4.68) or C^2HCl_3 (δ 7.27). ^{13}C NMR were referenced to $(C^2H_3)_2SO$ (δ 39.70) or C^2HCl_3 (δ 77.5). Mass spectra and accurate mass measurements were recorded on a VG 70-250 SE or on a VG Platform. Protected amino acid precursors and resins were purchased from Calbiochem–Novabiochem Ltd. (Beeston, Nottingham, UK). Solid-phase peptide synthesis on Wang resin was performed using FMOC-protected amino acids and PyBOP as the coupling reagent. All solvents were of the highest purity available or were redistilled before use.

General Procedure for Peptide Preparation

Polypeptides (50–100 mg) were prepared using solid phase chemistry and ''double couplings'' were employed for the reactions to form acyl prolines. The products were cleaved from the Wang resin using DCM/TFA/water/triethylsilane (50:42:5:3) and were examined by analytical HPLC on a C-18 column and by single-dimensional ^{1}H- and ^{13}C-NMR spectroscopy and by ES–mass spectrometry. In each case the required compound was obtained in at least 80% purity as a mixture of oligopeptides in which incomplete coupling reactions at each position accounted for 1–2% or less of each of several contaminants. These samples were, therefore, sufficiently homogeneous to perform more detailed NMR structural studies and test the oligopeptides for self-cleavage activity.

Analysis by NMR Spectroscopy

The sequence NFDLLKLAGDVESNPGPFFF which corresponds to the natural FMDV 2A–2B junction and also the variant aminoisobutyryl-DLLKLAGDVESNPG-PFTFAF were prepared and subjected to a structural examination by NMR using HMQC, COSY, and NOESY techniques in a range of solvents including chloroform, DMSO, TFE, and aqueous methanol. In each case the peptides appeared to exist as random coils and in no case was there evidence for the formation of helical structures, as assessed by searching for small CH^{α}-NH coupling constants of $J = 4.2$ Hz or less (*34*) or any significant NOE cross-peaks for residues that were not directly connected. Several truncated sequences also showed random coil conformations.

Assessment of Self-Cleavage in Synthetic Peptides

The synthetic polypeptides (1 mM) were incubated in buffered aqueous or aqueous ethanolic solutions at pH 4.0 to 9.5 in twelve 0.5 pH unit steps, each in the absence or presence of magnesium chloride (10 mM), imidazole (50 mM), urea (2 M), sodium chloride (20 mM), potassium chloride (50 mM), sodium hexanoate (5 mM), or Triton X-100 (2%). The incubations were stored at 37°C for several days and were assayed by the periodic removal of aliquots of the solution for TLC analysis on cellulose plates eluting with 19:1 isopropanol/aqueous ammonia (0.1 M) and developing with ninhydrin spray. No new bands were detected corresponding to any cleavage products.

Plasmid Constructs

Mutations in the region encoding 2A were introduced either into pCAT2AGUS (*4*) or a minor modification of this construct, pMD2 (*8*). Mutations in pCAT2AGUS were introduced by restriction of the plasmid with *Xba*I and *Apa*I, agarose gel purification of the large DNA restriction fragment, and ligation with a double-stranded oligonucleotide "adapter" molecule, as indicated below. Using the same strategy, mutations were introduced into pMD2 restricted with *Aat*II and *Bgl*II (pMD2.2, −2.3, −2.4, and −2.6 series mutations). The pMD2.7 series of mutations were produced by restricting pMD2 with *Aat*II, and mung bean nuclease treatment to remove the overhang and then a second restriction with *Afl*II before ligation with the oligonucleotide adapter molecule. Plasmids pMD3/5, pMD3/6(c), pMD31/10, and pMD31/11 were constructed using a different strategy. Sequences encoding the CAT gene together with

some of the 2A sequences were amplified by PCR using the forward primer ORM31 and a reverse primer, either ORM10 or ORM11. The PCR products were restricted with *Bam*HI and *Hind*III. Following gel purification the doubly restricted PCR products were ligated into pMD2 similarly restricted. All molecular biological manipulations were performed using standard procedures (*35*). Nucleotides introducing mutations are underlined. Degeneracies in the oligonucleotides are indicated by brackets. Sequence of all constructs were confirmed by automated DNA sequencing using a Perkin–Elmer 3100.

pCAT2AGUS/03: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGACTCCAACCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTGAGGTTGGGGC)-5′.

pCAT2AGUS/04: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCCAGTCCAACCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGGTCAGGTTGGGGC)-5′.

pCAT2AGUS/06.1: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGAGATTAACCCTGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTCTCCTTGGGAC)-5′.

pCAT2AGUS/06.7: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGAGTTTAACCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTCAAATTGGGGC)-5′.

pCAT2AGUS/05.1: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGAGTCCCACCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGGTGGGCC)-5′.

pCAT2AGUS/05.3: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGAGTCCGAGCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGCTCGGAC)-5′.

pCAT2AGUS/05.7: 5′-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGCGGGAGACGTCGAGTCCCAGCCCGGGCC)-3′, 3′-d(TCCTCGTACGGTCGACAACTTAAAACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGGTCGGAC)-5′.

pMD2.2 series: 5′-d(CGAGTCCAACCCTG[C/T]GCCCTTTTTTTTTACTAGTA)-3′, 3′-d(TGCAGCTCAGGTTGGGAC[G/A]CGGGAAAAAAAAATGATCATCTAGACCTAG)-5′.

pMD2.3 series: 5′-d(CGAGTCCAACCCTGGGNNNTTTTTTTTTTACTAGTA)--3′, 3′-d(TGCAGCTCAGGTTGGGACCCNNNAAAAAAAAAATGATCATCTAG)-5′.

pMD2.4 series: 5′-d(C[A/G/C[A[C/G]TCC[C/G]A[C/G]CCTGGGCCCTTTTTT-TTTACTAGTA)-3′, 3′-d(TGCAG[T/C/G]T[C/G]AGG[C/G]T[C/G]GGACCCGGG-AAAAAAAAATGATCATCTAG)-5′.

pMD2.6 series: 5′-d(CGAGTCCAACNNNGGGCCCTTTTTTTTTACTAGTA)-3′, 3′-d(TGCAGCTCAGGTTGNNNCCCGGGAAAAAAAAATGATCATCTAG)-5′.

pMD2.7 series: 5′-d(TTAAGCTTGCGGGGA[C/G]AGGT)-3′, 3′-d(CGAACGC-CCCT[C/G]TCCA)-5′.

pMD3/5: ORM31; 5′-d(GCGCGCGGATCCGATGGAGAAAAAA)-3′, ORM5; 5′-d(GCGCGCGACGTCG[T/C]GGATAAGAAGGTCAAAATTCAA)-3′.

pMD3/6(i): ORM6; 5′-d(GCGCGCGACCTGGATG[T/C]GAAGAAGCTGAAA-ATT)-3′.

pMD31/10: ORM10; 5′-d(GCGCGCGACGTCTCCCGCGG[G/C]AAGCTTAAG-AAGCTG)-3′.

pMD31/11a: ORM11; 5′-d(GCGCGCCTTAAGGG[G/C]AGGGTCAAAATTC-AA)-3′.

Coupled Transcription/Translation in Vitro

Coupled TnT reactions were performed as per the manufacturer's instructions (Promega). Briefly, rabbit reticulocyte lysates (20μl) or wheat-germ extracts (20μl), each containing [^{35}S]-methionine (50 μCi; Amersham), were programmed with unrestricted plasmid DNA (1 mg) and incubated at 30°C for 45 min.

Cleavage 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.

RESULTS AND DISCUSSION

At the outset of our research in this area we believed we were investigating a novel proteolytic cleavage event mediated by a short oligopeptide sequence in an enzyme-independent manner, a mechanism which it appeared might prove to be unique.

Construction of Artificial Self-Processing Polyproteins

To demonstrate the ability of FMDV to function independently of all other FMDV sequences and also to facilitate the study of its activity by site-directed mutagenesis (SDM), an artificial ''reporter gene'' polyprotein was constructed with chloramphenicol acetyltransferase (CAT) and β-glucuronidase (GUS) genes flanking 2A (L^1LNFDLLKLAGDVESNPG$^{18}\Downarrow$P^{19}) in a single ORF ([CAT2AGUS]) (Fig. 3). Analyses of the translation products (TP) of this reporter polyprotein together against the control construct [CATGUS] uing *in vitro* translation systems, RR lysates, and WG extracts were performed using PAGE, together with autoradiography and later using

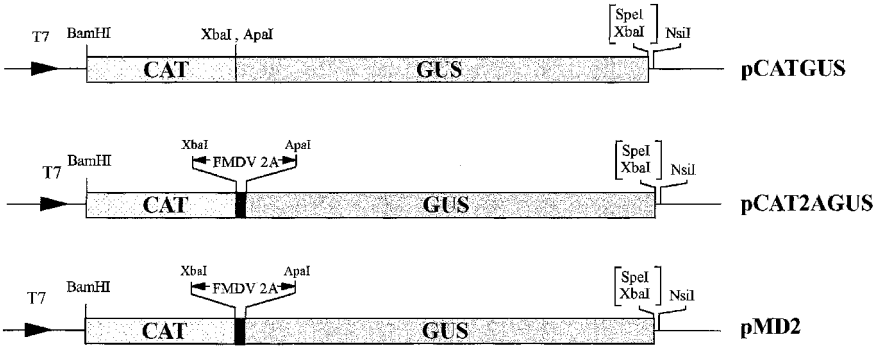


FIG. 3. Plasmid constructs. Boxed areas indicate the single open reading frames encoding the artificial polyproteins made up of the reporter genes CAT and GUS. Solid black boxes indicate 2A sequences.

phosphorimaging densitometry. The autoradiograms showed that while the [CATGUS] construct produced the expected single major product, the construct encoding [CAT2A-GUS] produced three major products. These were (i) the uncleaved forms [CAT2A-GUS] (comprising some 5–10% of the products) and the cleavage products; (ii) translation product 1 (TP1), [CAT2A]; and (iii) translation product 2 (TP2), GUS (Fig. 4). Note that the choice of genes was governed by the hope that such a reporter

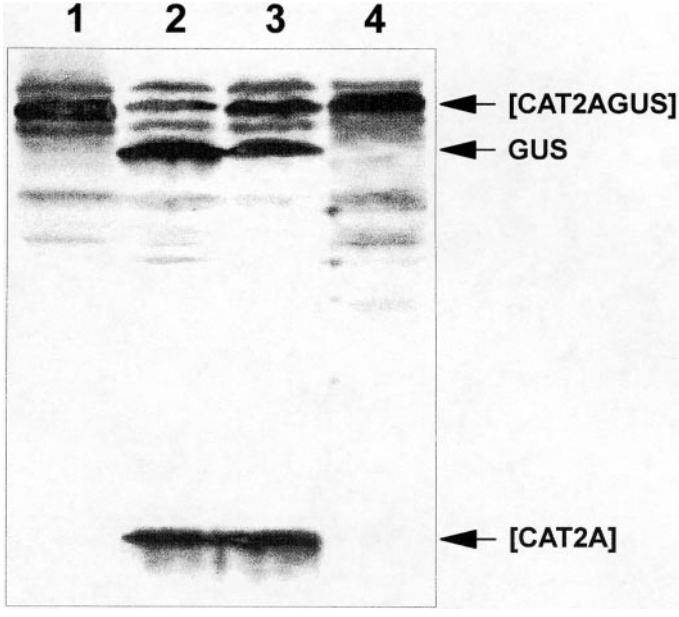


FIG. 4. Translation *in vitro.* A coupled transcription/translation rabbit reticulocyte system was programed with pCATGUS (lane 1), pCAT2AGUS (lane 2), pCAT2AGUS/O5.1 (lane 3), and pMD2.6.11 (lane 4).

polyprotein system would enable us, by inspection of *Escherichia coli* colony phenotypes (blue/white on X-gluc media; chloramphenicol$^{S/R}$), to perform a semi-random saturating SDM analysis of the 2A region. However, more recently we demonstrated that the 2A-mediated cleavage was entirely specific for eukaryotic (80S) and not prokaryotic (70S) ribosomes (*8*) so that the analysis of 2A activity would be restricted to the more cumbersome *in vitro* transcription/translation analysis using RR or WG.

Additionally we investigated the cleavage activity of the C-terminal region of cardiovirus 2A protein (highly similar to the FMDV 2A region) from EMCV and TMEV for comparison with FMDV 2A. Indeed, all of these sequences mediated cleavage with high efficiency (~95%). The full-length TME 2A protein linked to GUS cleaved to even higher levels (~100%). All of these constructs were tested for activity in *E. coli* and in no case were any of these found to be able to cleave (*8*).

The observation that sequences upstream of FMDV 2A were not critical for, but could influence the level of, cleavage was investigated by the insertion of FMDV capsid protein 1D sequences (present immediately upstream of 2A in the native polyprotein) into the [CAT2AGUS] artificial polyprotein. This showed that the insertion of the 1-D C-terminal 39 amino acids increased cleavage from ca. 95% to ca. 100% (*8*).

Site-Directed Mutagenetic, Chemical, and Modeling Studies

Inspection of the 19-amino-acid sequence suggested a helical structure possibly ending in a reverse turn. Several long peptides spanning the 2A region were synthesized and fully characterized. However, none of these showed any interesting structural properties in solution as determined by NMR spectroscopy. Moreover, the sequence NFDLLKLAGDVESNPGPFFF which corresponds to the natural FMDV 2A–2B junction (and also the variant aminoisobutyryl-DLLKLAGDVESNPGPFTFAF) was prepared by chemical synthesis and was assayed for self-cleavage under 400 different conditions of pH, ionic strength, various mixtures of different monovalent and divalent cations, imidazole, surfactants, and denaturants. In each case, no cleavage occurred as determined by either TLC or HPLC analysis. The tetrapeptide NPGP, in our hands, also failed to give cleavage products upon prolonged incubation in buffered solvents, in contrast to earlier reports (*9*).

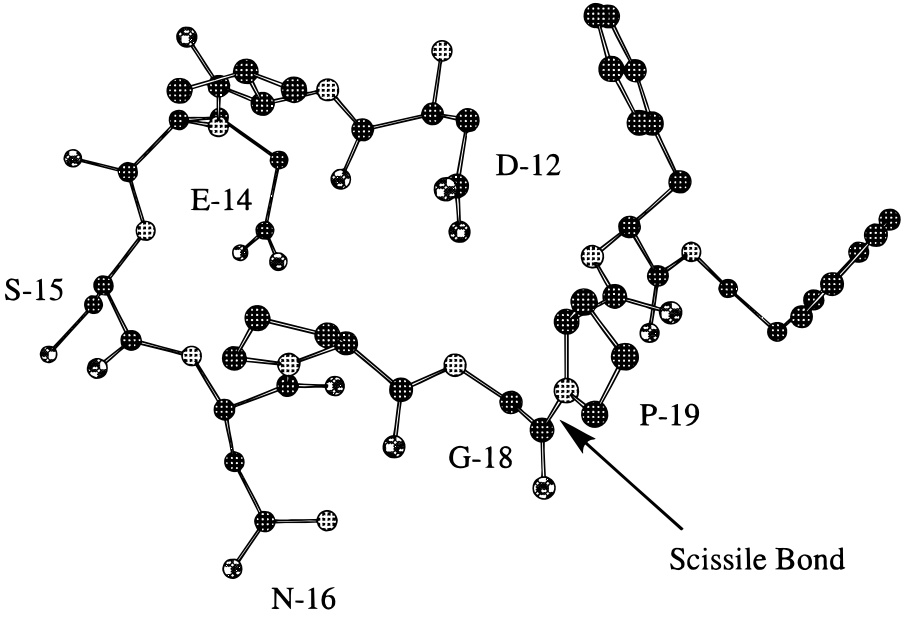
Molecular dynamics (10 ps equilibration, 500 ps data gathering, sampling every 1 ps) were performed on preminimized structures of the FMDV 2A region (acetyl-NFDLLKLAGDVESNPGPFFFA-NMe) using the AMBER molecular mechanics force field (*10,11*) and the Discover program (*12*) at a constant temperature of 300 K. No explicit solvent molecules were included and a low, distance-dependent, dielectric constant was used for all calculations ($\varepsilon = 4r$). Visualization and analysis of resulting structures were carried out using the analysis module of Insight (*12*).

First the idea that the 2A sequence started with an α-helix was considered. The largely hydrophobic residues except for D-5, K-8, and D-12 which could form salt bridges, vide infra, were certainly consistent with this notion. Residues downstream of residue D-12 seemed unlikely to exist in a stable helical conformation as the side chains of the polar residues E-14, S-15, and N-16 could disrupt the intrabackbone hydrogen bonds while P-17 and P-19 possess no amide hydrogen atoms. Therefore, various starting conformations were generated, minimized, and then subjected to

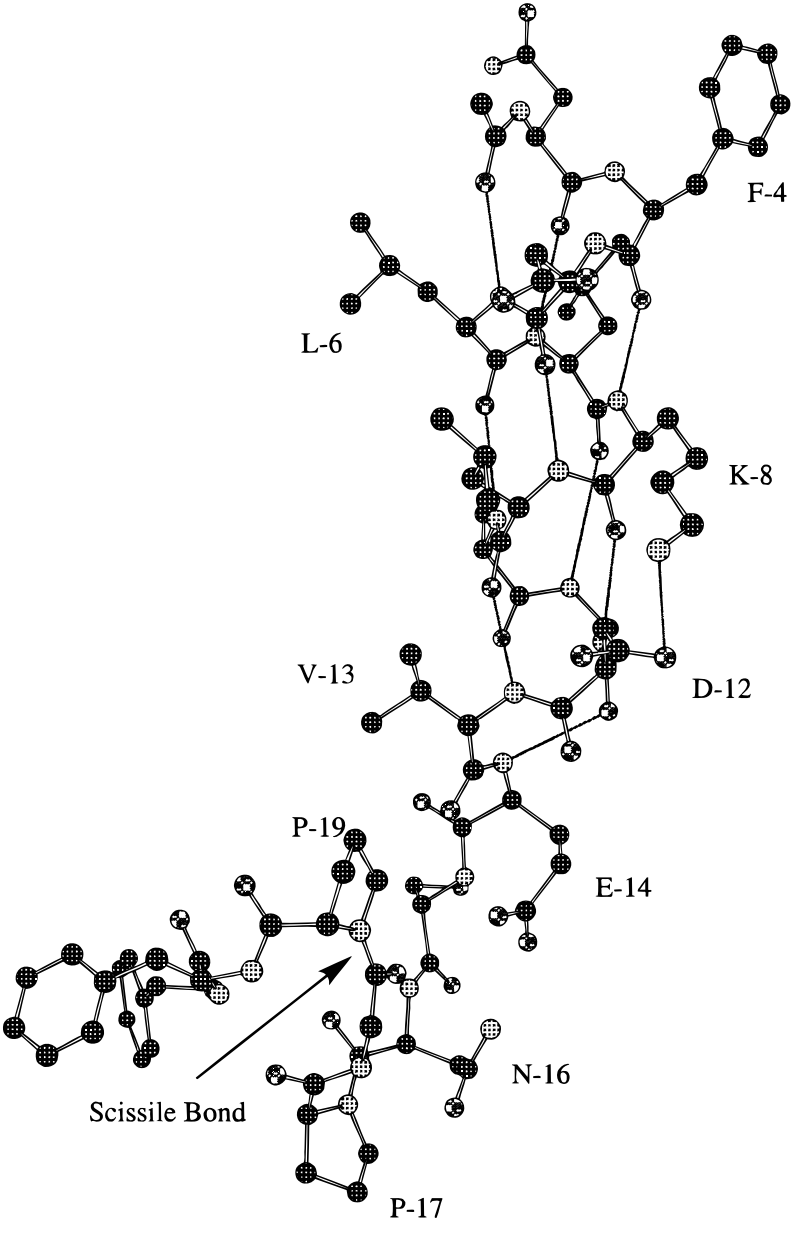
dynamics. In all cases, residues up to D-12 were placed in an α-helical conformation, while various starting conformations were used for the remaining C-terminal residues, including an all α-helical conformation and several random conformations. Also conformations with a *cis*-amide bond between G-18 and P-19 were considered. In the all α-helical case, the structure of the polypeptide downstream of residue D-12 was disrupted during optimization at E-14, S-15, and N-16 and the helical structure disintegrated further during dynamics. The N-terminal portion of the helix, however, remained present throughout all of the dynamic simulations.

The dynamic simulations produced a number of different conformers, but showed some consistent features in addition to the N-terminal helix (N-3 to D-12). Once formed, a charge triad between D-12, K-8, and E-14 proved very stable and a tight turn, formed from E-14 to N-16, served to bring the scissile amide bond (G-18, P-19) close to the side chain of D-12 (Structure **1**). However, it was far from clear how the structure might support the hydrolytic cleavage of the G-P amide bond. No consistency in the four residues following P-19 was observed. Although the fine details of the structure varied from one simulation to another, the conformers produced tended to be very stable.

In the case of the simulation from the optimized all-helical conformation (Structure **2**), the RMS deviation in atomic coordinates was less than 1.5 A between any two optimized conformations over the last 380 ps of the simulation. Conformations taken from this stable period of the simulation were optimized and used to guide and rationalize the structural role of variations in the wild-type sequences and in the choice of site-specific mutants to test the importance of interactions shown in the structure.

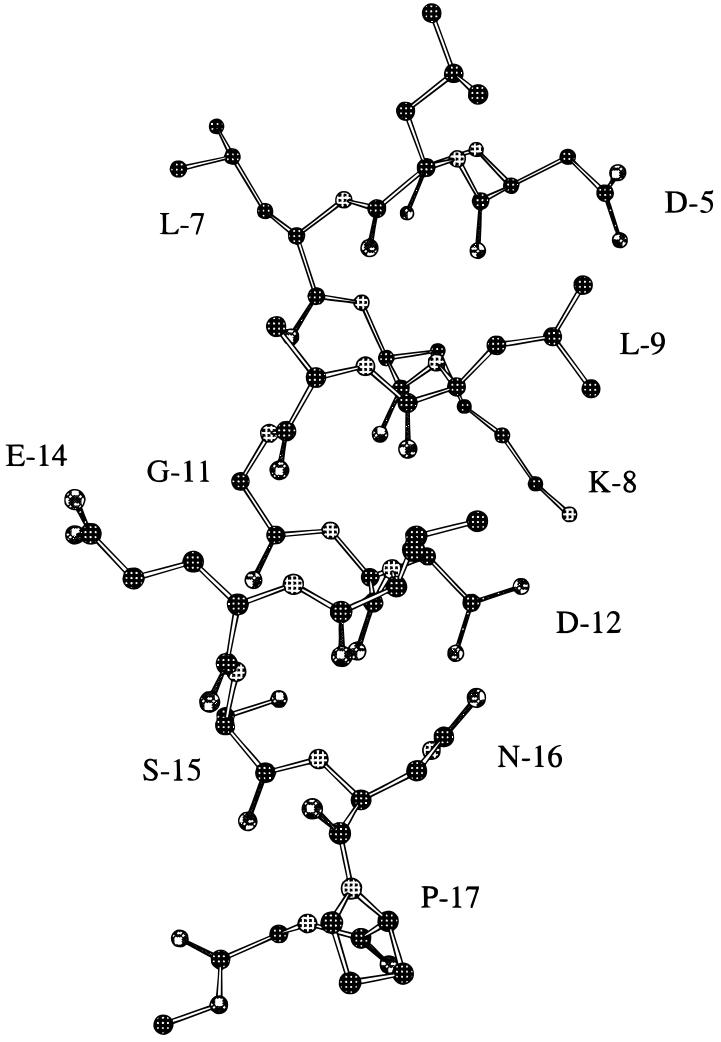


STRUCTURE 1. C-terminal portion of minimized stable structure.



STRUCTURE 2. Optimised all helical conformation.

Other structural models were constructed on the basis of the primary structure of the 2A region. Of these an α-helix type VI reverse-turn structure (Structure **3,** and primary structural variants thereof) seemed to fit with the conserved amino acid residues in the aligned sequences of the aphtho- and cardioviral 2A regions (Fig 2.)



STRUCTURE 3. α-Helix type VI reverse-turn conformation.

In aphthoviral sequences many features were apparently consistent with the putative Structure **3** including the positions of residues D-5, K-8, D-12, and N-16 that would be aligned along one side of the α-helical segment and which were present and stable in dynamic simulations of the FMDV 2A peptide, Structure **2,** vide supra. These form two salt bridges [(D-5 and K-8 in an i + 3 arrangement) and (K-8 and D-12 in an i + 4 arrangement] and an i + 4 H-bonding interaction between D-12 and N-16, respectively. Note that the long side-chain of the K-8 residue allows the ε-amino group to reside between the D-5 and D-12 residues.

In the cardioviral sequences, G-11 in the aphthoviral sequences is replaced by a

histidine residue and in each sequence there is an $i - 4$ residue containing an H-bond acceptor, D-7 in EMC and MENGO and Q-7 in TME, that could potentially interact with the H-11 residue. Interestingly the TME sequences appear to have the potential to form two sets of $i + 4$ side-chain interactions, between Q-7 and H-11 and between R-8 and D-12, which might stabilize a helical structure under certain conditions. However, the shear diversity of the sequences in the N-terminal region of 2A, residues 1–11, strongly suggested that the presence of a helical structure was the important structural feature and that interactions of the side chains of these residues with other molecules was not important.

Residues 9–19 in the sequences of both aphtho- and cardioviruses are highly conserved, other than residue 11. Aspartic acid-12 and N-16 would interact in an α-helix, but the amino acid following N-16 in the sequence is a completely conserved proline residue (P-17). Clearly, P-17 does not possess an N-H moiety and is, therefore, unable to act as a hydrogen bond donor (*13*). Indeed, it is known that proline residues disrupt α-helices in solution and cause significant bends in transmembrane helices spanning hydrophobic bilayers (*14*). In simulations the α-helical conformation of Structure **2** could not be propagated beyond the D-12 residue because the H-bonding partner for the carbonyl group of V-13 is absent in P-17. However, if the side chain of the N-16 residue does form an H-bond with D-12 in the $(i - 4)$ position, then a 180° rotation about the C^{α}-CO (the ψ angle) of N-14 would allow a type VI reverse turn to exist in which the P-17 residue, in its *cis*-rotomeric form, would line up the N-H moiety of the G-18 residue such that it could H-bond with the carbonyl O-atom of S-15 in the helix, Structure **3**.

This structural arrangement is absolutely unique because only a proline residue could both disrupt an α-helix and exist in a stable *cis*-rotomeric form such that the extending peptide chain is forced to fold underneath the helix. In this structure the G-18 carbonyl group is held at the bottom of the axis of the helix and the V-13 and P-17 carbonyl O-atoms are close enough to H-bond to a single water molecule or chelate a metal ion, vide infra, which would further stabilize the structure. Interestingly, the model predicts that the side-chain hydroxy group of S-15 and the α-carboxamide and γ-carboxy groups of E-14 do not form intramolecular interactions which could stabilize Structure **3**. Note that glutamic acid is conserved in the natural sequences, but that S-15 is not (Fig. 2). Nevertheless, in Structure **3**, aside from the G-18 carbonyl O-atom and the side chain of E-14, every other main-chain and side-chain functional group up to residue P-19 forms an intramolecular H-bond in what appears to be a stable structure. Molecular modeling and dynamic simulations confirmed that Structure **3** was reasonable but that further stabilization would be required for it to exist for significant periods in free solution. Structure **3** in which the P-19 residue is replaced by a methyl ester was optimized and subjected to molecular dynamics at 300 K for 400 ps. The C-terminal of the peptide twisted away from the bottom of the helix dipole but the helix was stable for prolonged periods. This is expected in the absence of H-bond donors for the carbonyl O-atoms.

Since at this stage we did not know whether cleavage was mediated by an exogenous proteolytic activity or some unimolecular process, a series of point, double, insertion, and deletion mutations were constructed (see Experimental). These are summarized in Fig. 5. The mutants were tested for activity in RR and WG translation systems