

A Model for Nonstoichiometric, Cotranslational Protein Scission in Eukaryotic Ribosomes

Martin D. Ryan, Michelle Donnelly, Arwel Lewis, Amit P. Mehrotra, John Wilkie, and David Gani¹

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.

© 1999 Academic Press

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, encephalomyocarditis; FMOC, 9-fluorenylmethoxycarbonyl; FMDV, foot-and-mouth disease virus; GUS, β -glucuronidase; 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.



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 encephalomyocarditis (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

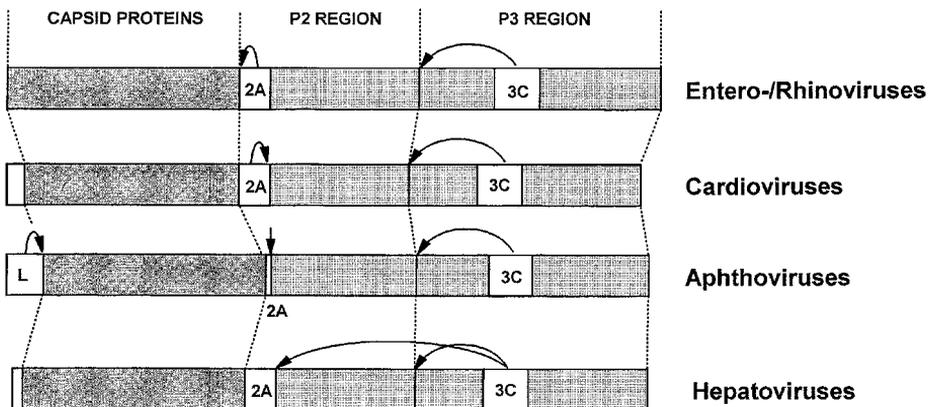


FIG. 1. Picornavirus primary polyprotein cleavages. Box regions indicate polyprotein domains.

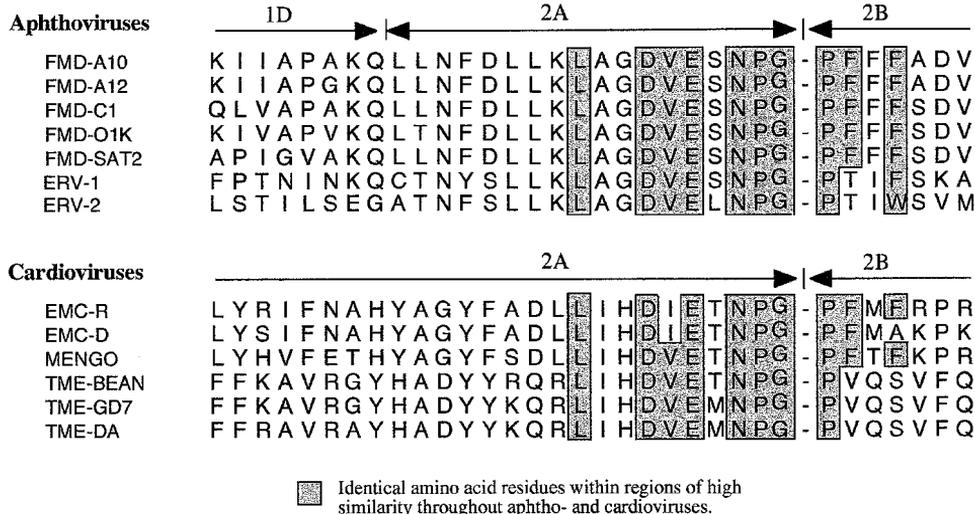


FIG. 2. Conserved amino acid residues in the aligned sequences of the aphthoviral and cardioviral 2A and 2B regions.

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 (^1H , 300 MHz; ^{13}C , 75.4 MHz), a Varian Gemini spectrometer (^1H , 200 MHz; ^{13}C , 50.3 MHz), a Varian Gemini spectrometer (^1CH , 300 MHz; ^{13}C , 75.4 MHz), and a Varian Unity Plus 500 spectrometer (^1H , 500 MHz; ^{13}C , 125.6 MHz). ^1H NMR spectra were referenced internally to $(\text{C}^2\text{H}_3)_2\text{SO}$ (δ 2.47), ^2HOH (δ 4.68) or C^2HCl_3 (δ 7.27). ^{13}C NMR were referenced to $(\text{C}^2\text{H}_3)_2\text{SO}$ (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 ^1H - 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 $\text{CH}^\alpha\text{-NH}$ 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 *Bg*III (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 *Af*III 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(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGACTCCAACCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTGAGGTTGGGGC)-5'.

pCAT2AGUS/04: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTCCAGTCCAACCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGGTCAGGTTGGGGC)-5'.

pCAT2AGUS/06.1: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGAGATTAACCCTGGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTCTCCTTGGGAC)-5'.

pCAT2AGUS/06.7: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGAGTTTAACCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTCAAATTGGGGC)-5'.

pCAT2AGUS/05.1: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGAGTCCCACCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGGTGGGCC)-5'.

pCAT2AGUS/05.3: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGAGTCCGAGCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGCTCGGAC)-5'.

pCAT2AGUS/05.7: 5'-d(CTAGAGGAGCATGCCAGCTGTTGAATTTTGACCTTCTTAAGCTTGC^{GGG}GAGACGTTCGAGTCCCAGCCCGGGCC)-3', 3'-d(TCCTCGTACGGTTCGACA^{ACTTAA}AACTGGAAGAATTCGAACGCCCTCTGCAGCTCAGGGTTCGGAC)-5'.

pMD2.2 series: 5'-d(CGAGTCCAACCCTG[C/T]GCCCTTTTTTTTTTACTAGTA)-3', 3'-d(TGCAGCTCAGGTTGGGAC[G/A]CGGGAAAAAAAAAATGATCATCTAGACCTAG)-5'.

pMD2.3 series: 5'-d(CGAGTCCAACCCTGGGNNNTTTTTTTTTTACTAGTA)-3', 3'-d(TGCAGCTCAGGTTGGGACCCNNNAAAAAAAAAAATGATCATCTAG)-5'.

pMD2.4 series: 5'-d(C[A/G/C][A[C/G]TCC[C/G]A[C/G]CCTGGGCCCTTTTTTTTACTAGTA)-3', 3'-d(TGCAG[T/C/G]T[C/G]AGG[C/G]T[C/G]GGACCCGGG-AAAAAAAATGATCATCTAG)-5'.

pMD2.6 series: 5'-d(CGAGTCCAACNNNGGCCCTTTTTTTTTTACTAGTA)-3', 3'-d(TGCAGCTCAGGTTGNNNCCCGGAAAAAAAATGATCATCTAG)-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(GCGCGCGACGTTCG[T/C]GGATAAGAAGGTCAAATTC)-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]AGGGTCAAATTC-AA)-3'.

Coupled Transcription/Translation in Vitro

Coupled TnT reactions were performed as per the manufacturer's instructions (Promega). Briefly, rabbit reticulocyte lysates (20 μ l) or wheat-germ extracts (20 μ l), each containing [³⁵S]-methionine (50 μ Ci; 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 (L¹LNFDLLKLAGDVESNPG¹⁸↓P¹⁹) in a single ORF ([CAT2AGUS]) (Fig. 3). Analyses of the translation products (TP) of this reporter polyprotein together against the control construct [CATGUS] using *in vitro* translation systems, RR lysates, and WG extracts were performed using PAGE, together with autoradiography and later using

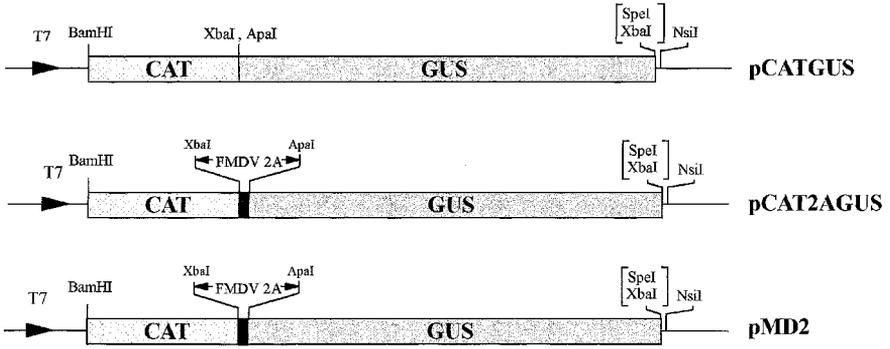


FIG. 3. Plasmid constructs. Boxed areas indicate the single open reading frames encoding the artificial polyproteins made up of the reporter genes CAT and GUS. Solid black boxes indicate 2A sequences.

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

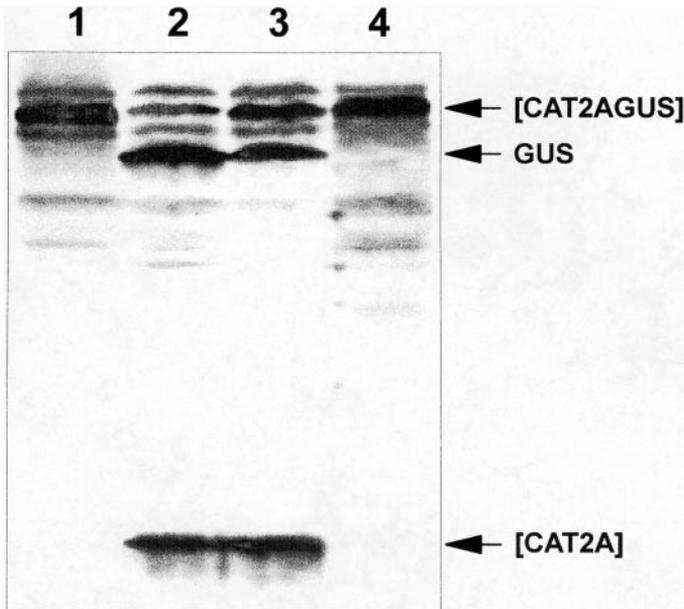


FIG. 4. Translation *in vitro*. A coupled transcription/translation rabbit reticulocyte system was programmed with pCATGUS (lane 1), pCAT2AGUS (lane 2), pCAT2AGUS/O5.1 (lane 3), and pMD2.6.11 (lane 4).

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).

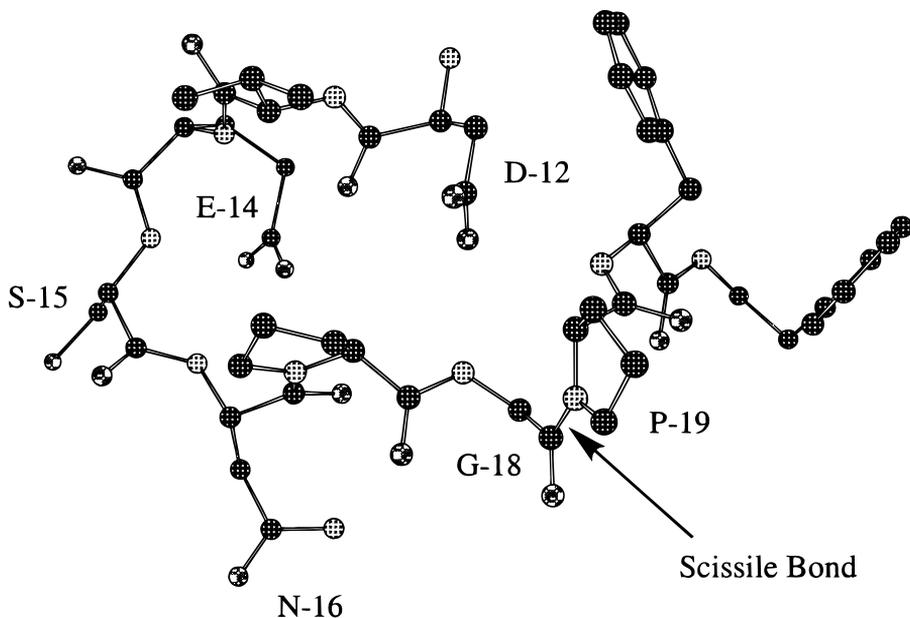
Molecular dynamics (10 ps equilibration, 500 ps data gathering, sampling every 1 ps) were performed on preminimized structures of the FMDV 2A region (acetyl-NFDLLKLAGDVESNPGPFFFA-NMe) using the AMBER molecular mechanics force field (10,11) and the Discover program (12) at a constant temperature of 300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

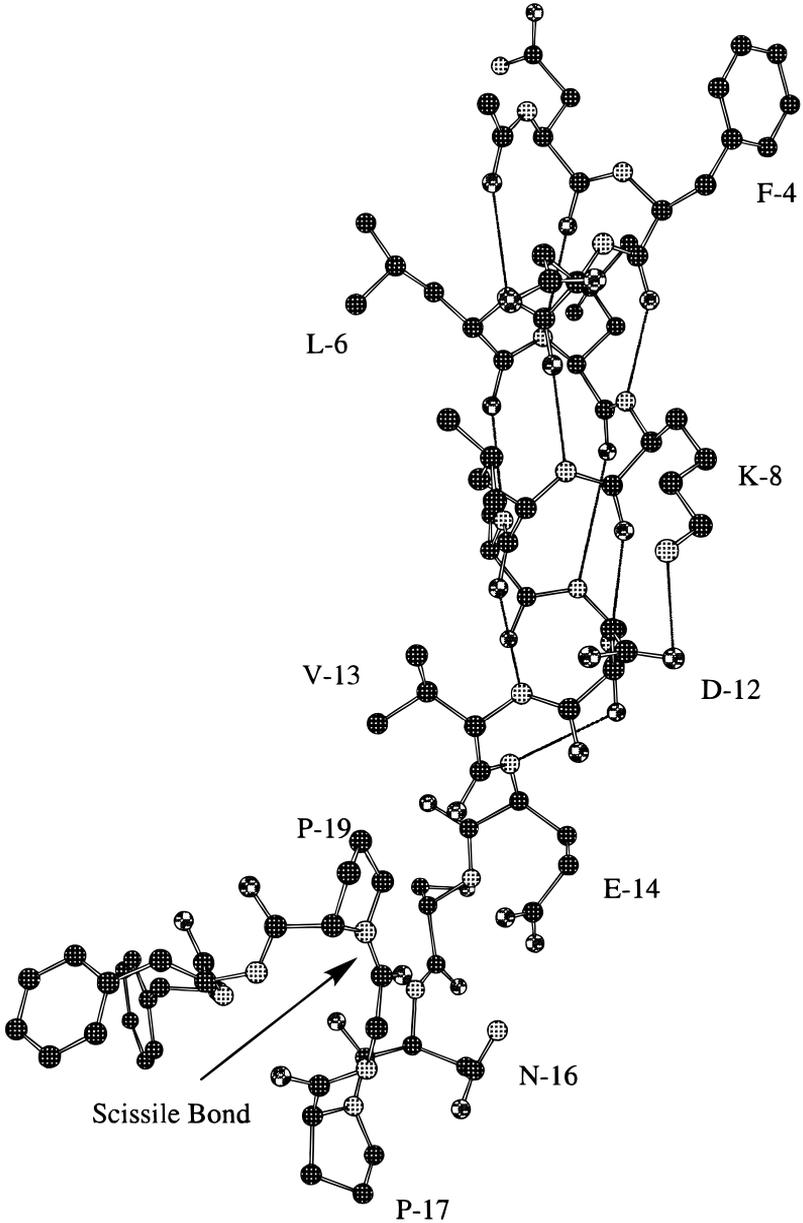
dynamics. In all cases, residues up to D-12 were placed in an α -helical conformation, while various starting conformations were used for the remaining C-terminal residues, including an all α -helical conformation and several random conformations. Also conformations with a *cis*-amide bond between G-18 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 Å 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.

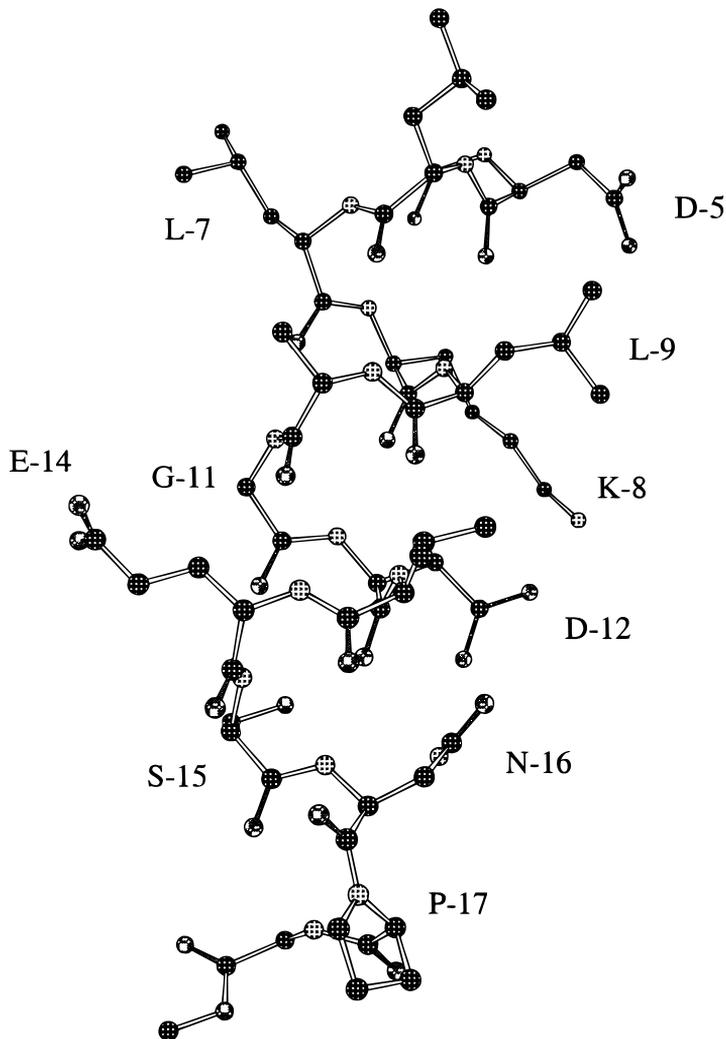


STRUCTURE 1. C-terminal portion of minimized stable structure.



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

		‘Cleavage’	
	FMDV	QLLNFDLLKLAGDVESNPGP	+
	pCAT2AGUS	QLLNFDLLKLAGDVESNPGP	+ A
	pMD2	QLLNFDLLKLAGDVESNPGP	+ B
1	pCAT2AGUS.09	-----LAGDVESNPGP	-
2	pCAT2AGUS.08	-----LKLAGDVESNPGP	+
3	pCAT2AGUS.02	-----DLLKLAGDVESNPGP	+
4	pCAT2AGUS.01	---NFDLLKLAGDVESNPGP	+
5	pMD3/11(a)	QLLNFDL P LKLAGDVESNPGP	-
6	pMD3/11(2)	QLLNFD P P L LKLAGDVESNPGP	-
7	pMD31/10(5);	QLLNFDLLKL A AGDVESNPGP	-
8	pMD31/10(3);	QLLNFDLLKL P AGDVESNPGP	-
9	pMD2.7.13	QLLNFDLLKLAG E VESNPGP	-
10	pMD2.7.15	QLLNFDLLKLAG Q VESNPGP	-
11	pCAT2AGUS/04	QLLNFDLLKLAGDV V SNPGP	+
12	pCAT2AGUS/03	QLLNFDLLKLAGDV D SNPGP	-
13	pCAT2AGUS/06.1	QLLNFDLLKLAGDV I NPGP	+
14	pCAT2AGUS/06.7	QLLNFDLLKLAGDV F NPGP	+
15	pCAT2AGUS/05.1	QLLNFDLLKLAGDV S H P GP	+
16	pCAT2AGUS/05.3	QLLNFDLLKLAGDV S E P GP	+
17	pCAT2AGUS/05.7	QLLNFDLLKLAGDV S Q P GP	-
18	pMD2.4.0	QLLNFDLLKLAGDV D S Q P GP	-
19	pMD2.4.14	QLLNFDLLKLAGDV D S D P GP	-
20	pMD2.4.20	QLLNFDLLKLAGDV Q S Q P GP	-
21	pMD2.4.24	QLLNFDLLKLAGDV Q S H P GP	-
22	pMD2.4.27	QLLNFDLLKLAGDV Q S E P GP	-
23	pMD2.4.19	QLLNFDLLKLAGDV N S H P GP	-
24	pMD2.4.28	QLLNFDLLKLAGDV N S Q P GP	+
25	pMD2.6.11	QLLNFDLLKLAGDVESN A GP	-
26	pMD2.6.13	QLLNFDLLKLAGDVESN T GP	-
27	pMD2.6.17	QLLNFDLLKLAGDVESN R GP	-
28	pMD2.2.5	QLLNFDLLKLAGDVESNP A P	-
29	pMD2.2.6	QLLNFDLLKLAGDVESNP V P	-
30	pMD2.3.1	QLLNFDLLKLAGDVESNP G A	-
31	pMD2.3.7	QLLNFDLLKLAGDVESNP G I	-
32	pMD2.3.8	QLLNFDLLKLAGDVESNP G S	-
33	pMD2.3.9	QLLNFDLLKLAGDVESNP G F	-
34	pMD2.3.12	QLLNFDLLKLAGDVESNP G I	-
35	42015	SRLLN F ADLL H L D IETNPGP	- C
36	42016	SRLLN F ADLL R L D IETNPGP	- C
37	pMD3/5	QLLNFDLL- I H -DVESNPGP	-
38	pMD3/6(c)	QLLNFDLL- H I -DVESNPGP	-

FIG. 5. Construction of a series of point, double, insertion, and deletion mutations, tested for activity in RR and WG translation systems.

and the translation products were analyzed by PAGE. The results, *vide infra*, were used to test and refine, reiteratively, various structural models. The shortest sequence able to support hydrolytic cleavage activity was the 13-mer, LKLAGDVESNPGP, entry 2 (Fig. 5).

An insertion mutation (entry 5, Fig. 5) in which a proline residue was introduced between L-6 and L-7 in FMDV 2A displayed no cleavage activity whatsoever. This result is consistent with the existence of helix-spanning residues D-5 to N-16 in FMDV 2A because a deletion mutation (entry 2) in which only L-7 to P-19 of FMDV 2A were retained, but in which the upstream sequence did not contain a nearby proline residue, was active (4). Presumably, the insertion mutation disrupts the ability of the sequence downstream of residue 6 to adopt a helical structure in the correct conformation relative to the upstream residues which in turn suggests that the helix is confined in some way. Interestingly, the interaction of D-5 and K-8 does not appear to be crucial.

The insertion of alanine or proline between residues L-9 and A-10 (entries 7 and 8) also gave completely inactive 2A regions and only the uncleaved CAT2AGUS mutant translation products were detected. The replacement of D-12 for glutamic acid (or glutamine) also gave completely inactive systems (e.g. entries 9 and 10) in accord with results obtained by Hahn and Palmenberg (15) for the EMC mutants H-12 and N-12 (Fig. 6). Certainly, it would seem that D-12 and N-16 in the wild-type sequences could only interact through the carboxylate group of D-12 and the N-H moiety of N-16. Given that K-8 also interacts with D-12 downstream, any change in the sidechain of residue 12 would be expected to destabilize two or more interactions. Thus, the lengthening of the sidechain to the carboxylate group in the E-12 mutant would be

	‘Cleavage’
DLL IHD IETNPGP	+
DLL IHH IETNPGP	-
DLL IHN IETNPGP	+ / -
DLL IHDF IETNPGP	-
DLL IHDV IETNPGP	+
DLL IHD I DTNPGP	+
DLL IHD I E ANPGP	+
DLL IHD I ET K PGP	+ / -
DLL IHD I ET N L GP	-
DLL IHD I ET N Q GP	+ / -
DLL IHD I ET N R GP	+ / -
DLL IHD I ET N P AP	-
DLL IHD I ET N P EP	-
DLL IHD I ET N P VP	-
DLL IHD I ET N P WP	-
DLL IHD I ET N P GL	+ / -
DLL IHD I ET N P GR	+ / -

FIG. 6. Cleavage of EMC 2A mutant sequences (data from Hahn and Palmenberg [15]).

expected to prevent E-12 from simultaneously forming strong interactions with K-8 and N-16. The Q-12 mutant would suffer the same fate but also could not act as a hydrogen bond acceptor to both K-8 and N-18 simultaneously and, indeed, was found to be inactive.

Residue E-14 is disposed on the opposite side of the α -helix (Structure 3) to the D-12 and N-16 residues and showed activity when replaced by a Gln residue and very slight activity when replaced by Asp in the FMDV sequence (entries 11 and 12). The D-14 mutant of EMC was also reported to show some activity (Fig. 6). Serine-15 can be substituted for either Thr or Met in wild-type sequences (Fig. 2), and the FMDV I-15 and F-15 mutants showed full activity (Fig. 5, entries 13 and 14) as did the A-15 mutant of EMC (Fig. 6,) (15) in accord with expectations.

According to the hypothetical model, N-16 serves as an H-bond donor to D-12 and is the last residue which possesses an N-H moiety that is part of the helix. The H-16 mutant displayed almost full hydrolytic activity as did the E-16 mutant (entries 15 and 16). Presumably the γ -carboxylate group in the latter mutant could interact with the D-12 residue through a water molecule as is common in the aspartic proteinases (16). The Q-16 mutant showed reduced 2A activity (entry 17) and double mutants in which E-14 was replaced by Asn or Gln and N-16 was replaced by His, Glu or Gln showed some but very low activity (entries 18–24).

It was argued above that P-17 could form a type VI reverse-turn structure in the active form of FMDV 2A and computer simulations had shown that this was a reasonable structure. The A-17, R-17, and T-17 mutants were all found to be totally inactive (Fig. 5, entries 25–27). Moreover, the L-17, R-17, and Q-17 mutants of EMC were found to be totally inactive (Fig. 6), (15). While the results do not prove that Structure 1 contains a type VI reverse turn, there does appear to be a requirement for proline, and only proline can populate the *cis*-rotameric form of the Asn-Pro amide bond to the extent that the polypeptide chain can fold back underneath the helix.

Glycine-18 is an absolutely conserved residue, and, as was expected, the A-18 and V-18 FMDV mutants were completely inactive (entries 28 and 29). Substitutions for Ala, Glu, Val, and Trp in the EMC sequence all gave protein products totally devoid of 2A activity (Fig. 6) (15). Finally, the last functional part of the 2A region, residue 19, is a conserved proline residue. In the FMDV sequence, all tested substitutions showed substantially reduced activities and only the S-19 and I-19 mutants showed any detectable activity whatsoever (entries 30–34). Similar results were obtained for L-19 and R-19 mutants of the EMC 2A region.

While it is almost impossible to prove a structure by SDM alone, all of the results are consistent with the proposed structural model. Moreover, several other lines of evidence support the notion that cleavage occurs cotranslationally, as is discussed below. However, two other constructs are worthy of further discussion in support of Structure 3. First, the EMC sequence contains Ile and His in positions 10 and 11, respectively, relative to the FMDV sequence (Fig. 2). When the positions of the residues were interconverted and Ile was swapped for Leu to give the mutant H-10, L-11, a totally inactive 2A region resulted. Similarly, when His was swapped for Arg to give the mutant R-10, L-11, no 2A cleavage activity was detected (entries 35 and 36) (17). Position 14 in both EMC and FMDV is occupied by a Glu residue and the introduction of a His (H-10) or Arg (R-10) residue in the *i* - 4 position would be

expected to stabilize a helical conformation through the formation of a salt bridge. Thus, either E-14 is required for some intermolecular interaction which is disrupted by its participation with H-10 or the disruption of the D-7 to H-11 interaction in wild-type EMC cannot be tolerated. Interestingly, the constructs in which the KLAG region of FMDV was replaced by IH, which gives a sequence very similar to that of EMC, were inactive, (Fig. 5, entry 37). The only difference in this sequence to the EMC sequence within the frame residue 6 onward is the presence of a Phe residue at position 6. It seems likely that F-6 is simply too big to bind in the conformation required for hydrolysis.

Thus, all of the mutants behaved in a manner consistent with proposed Structure, 3. However, there were several features of the translation assay that required further investigation.

Detailed Analyses of Translation Products

Careful analysis of the translation profiles of pCAT2AGUS (containing the wild-type 2A sequence) showed the presence of three products that migrated more slowly than the 70-kDa GUS cleavage product (TP2). By N-terminal truncation of the CAT gene we were able to demonstrate that these products were formed as a consequence of alternative initiation events within the CAT gene that had produced the “uncleaved” proteins. These products were all immunoprecipitated by anti-CAT antibodies and by phosphorimaging analyses it was possible to quantify the alternative initiations in both RR and WG. In both RR and WG more than 50% initiation took place in in-frame Met codons other than the initiation codon Met-1. The analyses of these N-terminally truncated forms of the polyprotein showed that they all cleaved to produce GUS and the corresponding deletion forms of [Δ CAT2A] with the same efficiency as the full-length [CAT2AGUS]. To evaluate the protein in the GUS gel band, due to internal initiation at this initiating codon, this codon was deleted and analyses showed no detectable contribution to the GUS product by internal initiation at this point within the mRNA (18).

Given that the internal initiation events caused the levels of radioactivity in the [CAT2A] (TP1) product gel band to be reduced compared to those for the GUS product (TP2), the molar ratios of the cleavage translation products, TP1 and TP2 (corrected for relative Met contents), were carefully measured. It was found that there were differences between batches of RR or WG and a large difference between RR and WG. In RR the TP2 was obtained in a very slight molar excess over TP1, whereas in WG TP1 was obtained in a 2:1 molar ratio compared to TP2. When the internal initiation effects were accounted for, the molar excess of TP1 over TP2 became apparent. In RR this ranged from a ratio of 2.4:1 to 5.3:1, while in WG the ratios ranged from 14:1 to 6:1. These observations were extended to the cardiovascular 2A constructs and, indeed, between 2.1- and 8-fold more TP1 than TP2 was produced in RR and between 17- and 23-fold more TP1 than TP2 was produced in WG.

Experiments designed to examine the differential rates of degradation of the [CAT2A] and GUS products showed that by translating for different times, and then arresting translations and incubating for different extended periods, cleavage occurred cotranslationally and not post-translationally and that the [CAT2A] and GUS products were stable to proteolysis by nonspecific proteinases in the translation systems for extended periods (4, 18). The major and particularly exciting finding of these studies was that the data were not consistent with a model whereby aphtho- or cardiovascular 2A proteins bring about a cotranslational proteolytic event to generate two fragments because the ratios of products TP1 and TP2 varied so widely.

To verify that the mRNA sequence was not responsible for the cleavage activity, the region encoding the 2A oligopeptide (boxed section, Fig. 7) was frame-shifted with respect to the reporter proteins, TP1 and TP2, by the insertion of two nucleotides (A and T) preceding the 2A region. The TP2 reading frame was restored by the insertion of a further nucleotide base (G) immediately following the 2A sequence. A point mutation needed to be introduced into the 2A region (G changed to T) in order to remove a stop codon (TGA) and thereby maintain the single, long open reading frame. Translation products derived from pAM2 showed a single, full-length translation product demonstrating that it is the peptide sequence that gives rise to 2A activity (19). Note that in this construct, pGFP2AGUS, green fluorescent protein (GFP) was used instead of CAT because it was found to give less internal initiation and therefore cleaner gels (19).

FMDV 2A Activity—An Alternative Hypothesis

Three mechanisms for cleavage had been considered during the course of the work: 2A-mediated cotranslational proteolysis, cotranslational proteolysis by an undescribed ribosome-associated host-cell proteolytic activity, or nonsynthesis of the peptide bond (4). Only the last of these mechanisms was consistent with the new data. Below we present an alternative hypothesis as to how FMDV 2A could bring about an apparent cotranslational cleavage in a mechanism whereby translation is terminated, to give TP1, and is then reinitiated, starting at the Pro-1 residue of TP2, to give TP2. Proline-1 is the N-terminal residue of the picornoviral protein 2B and, of course, was formerly regarded as the first residue beyond the cleavage site for a proteolytic reaction.

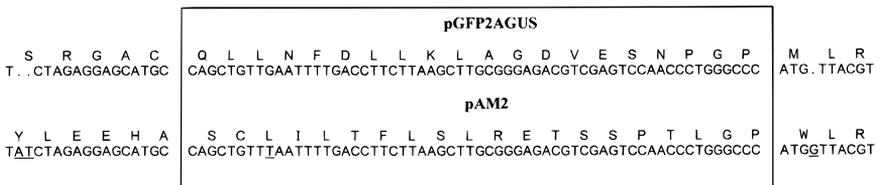
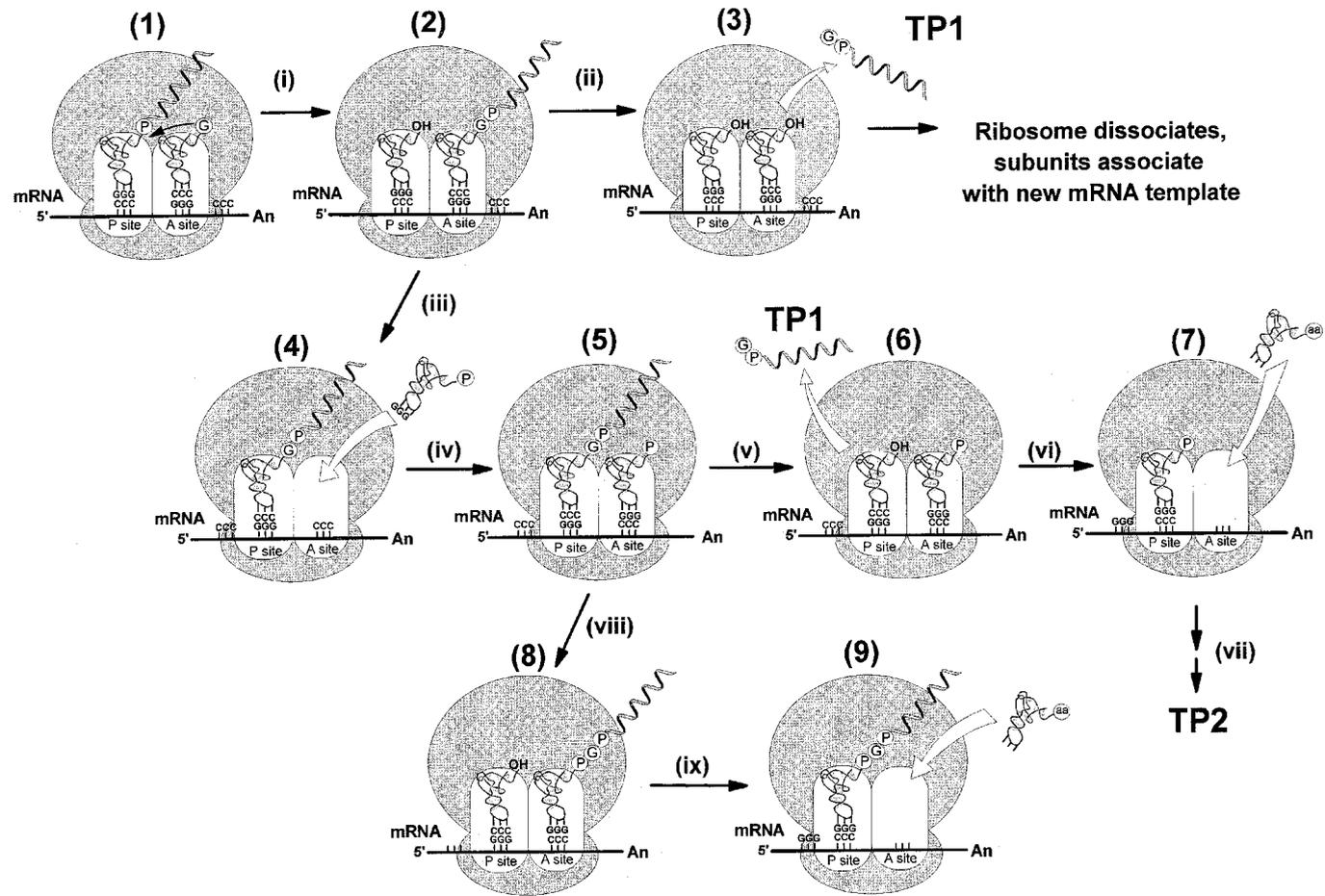


FIG. 7. Frame-shift of the 2A region. The sequence encoding 2A was frame-shifted (+2) with respect to CAT and GUS by the insertion of two nucleotides (A and T; underlined) upstream of the 2A sequence, and an additional base (G; underlined) immediately downstream of the 2A region to restore the GUS open reading frame. Note: A single open reading frame is maintained in construct pAM2.

To account for the observed unusual stoichiometry of the cleavage products whereby a molar excess of the protein N-terminal to 2A is produced (TP1) compared to the protein C-terminal to 2A (TP2), it was considered that the activity associated with aphtho- and cardiovirus 2A proteins might be that of an esterase rather than that of a proteinase. In this mechanistic model, the nascent 2A peptide mediates a hydrolytic attack on the 3'-O-adenosyl ester carbonyl group between the nascent polypeptide and the tRNA moiety of the peptidyl-tRNA complex *within the ribosome* in competition with nucleophilic attack by the amino group of Pro-19. Note that proline is by far the least nucleophilic of the proteinogenic amino acids and its substitution for all other amino acids tested resulted in the synthesis of uncleaved protein in both FMDV and EMCV constructs, *vide supra*.

This model must account for the three outcomes that were observed in the translation of the artificial polyproteins. First, that peptide bond formation proceeds throughout the length of the polyprotein to give uncleaved [CAT2AGUS]). Second, that ribosomal dissociation occurs at the C-terminal of the 2A sequence, TP1, within the template mRNA ORF, in order to account for the nonstoichiometric ratio of TP1 to TP2. Third, that ribosomal translocation from the ultimate (Gly) codon of TP1 to the adjacent (Pro) codon of the TP2 protein occurs, without peptide bond formation. This would account for the discrete GUS product arising in the absence of polyprotein proteolysis at this site. Furthermore, the model must explain the apparent discrepancy that when 2A is present in an artificial polyprotein context, an imbalance is observed in the ratio of the cleavage products, but when 2A is in its native polyprotein context, the cleavage products TP1 and TP2 are produced in equimolar quantities, *vide infra*.

The mechanistic model is summarized in Scheme 1. In complex (1), the ribosome contains a loaded Gly-tRNA molecule in the A site. In step i, normal peptidyl transfer of the acyl group of Pro-17 to the amino group of the Gly moiety occurs to give a deacylated tRNA^{Pro} molecule in the P site and the TP1 peptidyl tRNA molecule in the A site [complex (2)]. Since the mutation of the N-terminal proline residue of TP2 to primary amino acids results in the synthesis of uncleaved polyprotein, the rate of nucleophilic attack by 2A-activated water upon the peptidyl-tRNA^{Gly} ester linkage in the mutant forms of complex (2) must be slower than the combined rate of the ribosomal A to P site translocation step (catalyzed by elongation factor 2; eEF2), the binding of the cognate aminoacyl-tRNA into the A site (slow step), and the nucleophilic attack of the glycyl carbonyl group by the amino group of the correct aminoacyl-tRNA molecule (rapid step (20)). Thus, the mutation of Pro-1 of TP2 to a primary amino acid residue permits the aminoacyl-tRNA to outcompete the nucleophilic attack on the ester linkage by 2A-activated water in both the A site and the P site of the ribosome and reaction of complex (2) occurs to give the uncleaved product through steps iii, iv, viii, and ix in Scheme 1. The absolute requirement for cleavage activity of a proline residue in the first position of TP2 is explained by the poor nucleophilicity of proline. The secondary amino group is sterically hindered relative to the primary amino groups of other amino acids and is conformationally restrained due to its location in a five-membered pyrrolidine ring. Indeed, the poor nucleophilicity of proline compared to primary amino acids is well documented in peptide synthesis (21). Furthermore, 3'-O-prolyl adenosine has been found to be the worst 3'-O-aminoacyl



SCHEME 1. Proposed mechanism for polyprotein cleavage. See text for detailed discussion.

adenosine substrate for peptidyl transferase activity in *E. coli* ribosomes, along with 3'-O-glycyl adenosine (22).

Theoretical considerations predict that the nascent C-terminal regions of all polypeptides are present in the ribosomal exit tube as α -helical structures (23, 24). It is proposed that the orientation of the helix within the exit tube together with a type VI turn at the C-terminus of the 2A peptide (-NPG-) serves two purposes: first, to alter the position of the peptidyl-tRNA^{Gly} ester bond into a conformation which disfavors attack by the secondary amine of the incoming prolyl-tRNA; and second, to position the peptidyl-tRNA^{Gly} ester bond at the base of the 2A α -helix in the ribosomal exit tube into a suitable conformation for nucleophilic attack by a 2A-activated water molecule. In the model the structure of 2A assists in the generation of a nucleophilic species (probably a Mg²⁺ coordinated water molecule) that would attack the ester linkage to its own tRNAGly moiety. The kinetics of 2A-mediated cleavage of the peptidyl-tRNA^{Gly} ester bond are such that hydrolysis would occur in the P site of the ribosome. Thus, under normal circumstances (when 2A is present in its native polyprotein context) complex (2) would react as for normal protein synthesis and translocation would occur to move the peptidyl-tRNA^{Gly} molecule into the P site [complex (4), Scheme 1]. A prolyl-tRNA molecule would then bind into the A site, as for normal protein synthesis to give complex (5). Hydrolytic cleavage of the peptidyl-tRNA^{Gly} ester bond at this point would produce deacylated tRNA^{Gly} in the P site and free TP1, but would leave intact prolyl-tRNA in the A site [complex (6)]. Thus, the ribosome would exist in a state very similar to that during normal peptide bond synthesis [cf. complex (2)], except the peptidyl-tRNA in the A site would be replaced by prolyl-tRNA. Translocation of the prolyl-tRNA from the A site to the P site (catalyzed by eEF2), along with the deacylated tRNA from the P site to the exit (E) site (not shown) via complex (7), would proceed as normal and synthesis of a discrete downstream product could then ensue to give TP2. The crucial difference here would be that the nascent peptide TP1 would be released from the ribosome upon the completion of its synthesis as a discrete entity before translation of the downstream ORF.

The model presented above fits very well for 2A activity when it is present in its native polyprotein context. Under these circumstances the ratio of TP1:TP2 is 1:1. Clearly, no 2A-mediated hydrolysis can occur in the A site of complex (2) or in the P site of complex (4) prior to the binding of a prolyl-tRNA molecule else synthesis would abort, TP1 would be released, and the ribosomes would rebind at the start of the mRNA molecules and proceed with the translation afresh. Such a situation would lead to the unequal expression of TP1 and TP2 where TP1 would always be expressed at higher levels. While this does not happen in native viral sequences or in constructs possessing altered C-terminal domains downstream of the 2A region, artificial reporter systems in which the sequence immediately upstream of the N-terminal of 2A is altered show exactly such behavior. Note that as a control, phosphorimaging analysis of the translation products derived from the wild-type sequence pFMDP12ABC showed that no uncleaved material could be detected and that the cleavage products [P1-2A] and [2BC] were present in equal molar quantities. This was an important result in that translation factors, aminoacyl-tRNAs, metabolites, etc. were shown to be present in the *in vitro* translation systems (during the synthetic phase of the

translation reactions) at a level sufficient to synthesize an ORF longer than the artificial polyproteins described here, without premature termination of transcription/translation sufficient to give a spurious imbalance result.

To account for the observation that synthetic reporter systems show high TP1:TP2 ratios, it is, therefore, proposed that there is an intrinsic (slow) rate of 2A-mediated cleavage of the ester linkage in the A site of the ribosome. This effect is not normally observed because the reaction occurs in competition with the rapid translocation of the 2A peptidyl-tRNA^{Gly} from the A to P ribosome sites. We believe this effect does, however, come into play in two unusual situations, either when the 2A sequence is inserted into the artificial polyproteins or when eEF2 levels are very low such that slower A to P site translocation occurs (25). While it may be alluring to suggest that the structure of ribosomally bound 2A peptidyl-tRNA^{Gly} together with Mg²⁺ ions possesses all of the components required for hydrolytic cleavage, 2A activity, when present in its native polyprotein context, gives exactly 1:1 TP1:TP2. Therefore, prolyl-tRNA must be bound in the A site of complex (5) and it is quite reasonable to expect that components of the prolyl-tRNA molecule assist in conferring hydrolytic activity. For example, it is known that Mg²⁺ ions are required for binding aminoacyl oligonucleotides (26) and for peptidyltransferase activity (27) and it is highly likely that these serve to activate the ester bond as an electrophile by binding to either or both of the ester O-atoms. The prolyl-tRNA molecule may serve to contribute in stabilizing the Mg²⁺ complex(es) through direct interactions or through physically preventing the dissociation of the metal ions from the ester. Nevertheless, whatever the role of the prolyl-tRNA is in enhancing the activity of 2A, the attack by water is extremely efficient in wild-type sequences and almost completely suppresses normal peptide-bound formation.

While the FMDV 2A is able to mediate high levels of cleavage (95%) in the complete absence of other FMDV sequences (4), early work on 2A activity in recombinant FMDV polyproteins indicated that upstream sequences did play a role in maximizing the cleavage activity, in the sense that less uncleaved product was observed (7). This was confirmed by the observation that the inclusion of longer FMDV sequences upstream of 2A into the artificial polyprotein system increased cleavage activity (8). It is proposed that the effect of these upstream sequences is to alter the conformation of 2A and thereby the lability of scissile ester linkage. Thus, it is possible to envisage a situation where the rate of 2A-mediated hydrolysis in the A site (Scheme 1, step ii) is increased relative to its translocation to the P site (step iii). If cleavage occurred in the A site, then deacylated tRNA molecules would be present in both the P and A sites simultaneously. This situation is analogous to termination of translation where ribosome release from the template mRNA occurs. If, however, cleavage did not occur in the A site, 2A peptidyl-tRNA would be translocated from the A to the P site where either 2A could mediate cleavage or the incoming prolyl-tRNA would attack the ester linkage forming a peptide bond, as has been discussed above.

In summary, we propose that 2A in the native polyprotein context cleaves the peptidyl-tRNA bond in the P site [complex (5)], releasing the nascent peptide. In the artificial polyprotein systems or in a deletion form of an FMDV polyprotein in which the native upstream region is absent [pMR65; (7)], 2A-mediated cleavage of the ester bond may occur in the A site [complex (2)] which would lead to the release of a

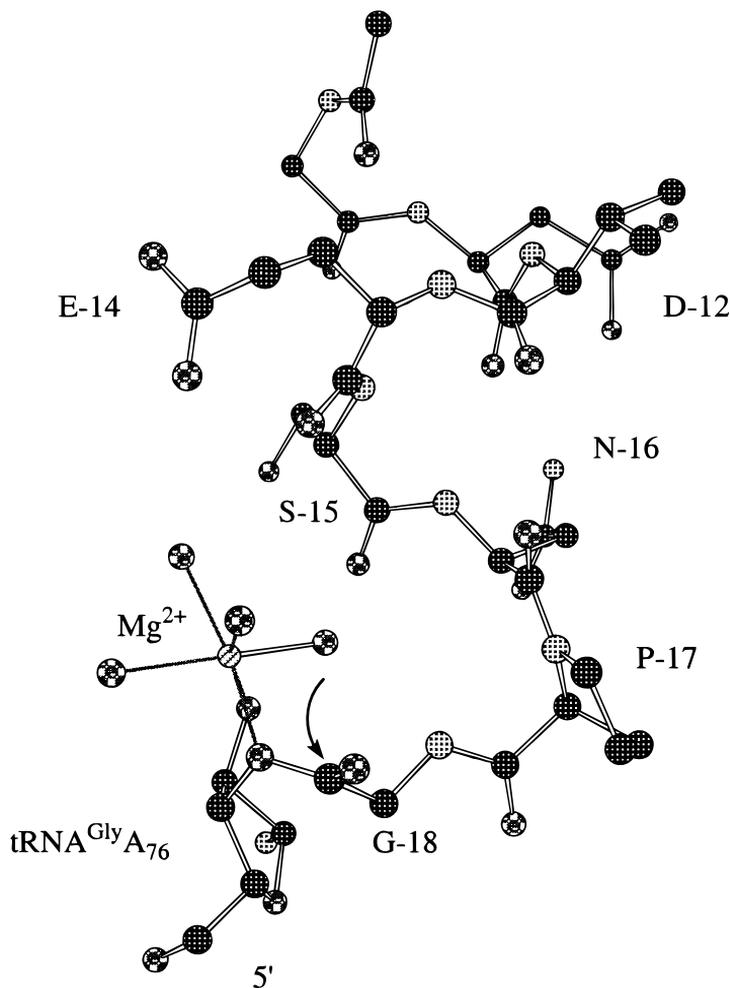
discrete product, TP1, with a Gly C-terminal residue and terminate translation. If cleavage in the A site does not occur, then following translocation to the P site (Scheme 1, step iii) the ester bond is attacked by prolyl-tRNA which leads to the synthesis of a full-length translation product. Alternatively, the ester bond could be attacked by 2A-activated water to give a C-terminal Gly as described above. Hence, the imbalance in the products arises from cleavage in the A site of the ribosome.

Under conditions where eEF2 activity is reduced, the rate of A to P site translocation (step iii) could be slowed such that 2A-activated water could attack its own peptidyl-tRNA ester linkage while in the A site. The consequences of such an attack are described above. Interestingly, translation studies on cardiovirus RNA using Krebs-2 cell-free extracts (containing low levels of eEF2 activity) showed a “translational barrier” in the central region of the genome (25). The translation products shown in the gels indicated that this barrier prevented translation of products downstream of cardiovirus 2A. The addition of eEF-2 greatly enhanced the synthesis of proteins C-terminal of the cardiovirus 2A region. The authors could not account for this effect of supplementation with eEF2. We can now interpret these data in light of the model. The low level of eEF2 in the Krebs cell extract leads to (cardiovirus) 2A attacking its own ester linkage to tRNA^{Gly} in the A site. The translational barrier observed, in our interpretation, is due to ribosome release in this region. Supplementation of the Krebs cell extract with eEF2 promotes translocation and the cardiovirus 2A cleaving the ester bond in the P site, leading to continued synthesis of downstream products. This effect may also explain the variation in the relative levels of the cleavage products we have observed between different batches of *in vitro* translation mixtures and between rabbit reticulocyte lysates and wheat germ extracts.

At a molecular level it seems extremely likely that the ester moiety is activated as an electrophile in exactly the same way as it is for peptidyl transfer, except the highly structured 2A peptide, which we suggest possesses a helix-turn structure, disturbs the position of the Mg²⁺ ion or ions such that peptidyl transfer is suppressed in favor of attack by a Mg²⁺-bound water molecule. There are numerous examples of metal ion-assisted phosphate ester hydrolyses in which a water molecule associated with the metal ion directly attacks the electrophile (28). Relevant to this discussion is the mechanism of inositol monophosphatase which employs two Mg²⁺ ions. One binds to peripheral phosphate O-atoms as a Lewis acid and the other binds to the bridging ester O-atom and positions an H-bonded hydroxide ion in the correct position for attack on phosphorus (29, 30). In this structure all of the ligands are O-atoms.

It is tempting to suggest that a Mg²⁺ ion can bind at the base of the 2A helix near the helix axis and coordinate to the acylated vicinal diol moiety of the adenosyl ribofuranosyl fragment at the end of the tRNA acceptor stem (31) (Structure 4). Such binding would position the Mg²⁺ ion in the negatively charged electric field of the helix dipole and position a coordinated water molecule perfectly for attack on the 3'-ester carbonyl group.

The presence of the Mg²⁺ ion is expected to significantly stabilize the structure of the helix through a charge-dipole interaction and two other effects might further stabilize the structure. First, as it seems certain that the G-18 residue is still attached to the tRNA^{Gly} molecule before hydrolysis occurs, rather than to a longer peptide that is subsequently processed through a proteinase activity, the G-18 residue would



STRUCTURE 4. Possible role of Mg²⁺ in hydrolytic cleavage by 2A sequence.

be firmly tethered at the base of the helix and not able to drift away as our unrestrained molecular dynamic simulations suggested. Moreover, the physical presence of the ribosomal polypeptide exit channel would restrain or at least retard polypeptide unfolding. Interestingly, if the acceptor stem (-CCA-) of the tRNA molecule retains its base-stacked helical conformation when bound to the ribosome, the penultimate cytidine base (C₇₅) could interact with the 2A peptide glutamic acid residue (E-14) through a *bis*-chelated H-bonding interaction between both γ-carboxy O-atoms and the N-3 and 4-amino moieties. In *E. coli* the base C₇₄ in tRNA is known to play a role in protein synthesis by forming a Watson–Crick base pair with G₂₂₅₂ of 23 S rRNA in the P site of the ribosome (32). The role of C₇₅ of tRNA is not known at this time (33), although it is known that the base is important for conferring protection

by P site-bound tRNA from kethoxal. The possible interaction between C₇₅ and E-14 is consistent with the finding that the Q-14, D-14, and N-14 mutants of FMDV 2A showed significantly lower levels of activity but clearly, without further structural information, a number of alternative interactions with rRNA could be possible. It should be noted that 2A activity is not expressed in prokaryotic ribosomes (8). Thus, the translational machinery of *E. coli* is unlikely to throw much light on the mode of action of 2A in the established eukaryotic systems of yeast, wheat, insects and mammals in which it has been demonstrated to operate highly effectively.

ACKNOWLEDGMENTS

We thank the Biomolecular Sciences Committee of the BBSRC for Grants B06902 and B05162 and the Wellcome Trust for a Wellcome Prize Studentship to A.L.

REFERENCES

1. Bazan, J. F., and Fletterick, R. J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 7872–7876.
2. Allaire, M., Cheraia, M. M., Malcolm, B. A., and James, M. N. G. (1994) *Nature* **369**, 72–76.
3. 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., and Worland, S. (1994) *Cell* **77**, 761–771.
4. Ryan, M. D. and Drew, J. (1994) *EMBO J.* **13**, 928–933.
5. Carroll, A. R., Rowlands, D. J., and Clarke, B. E. (1984) *Nucleic Acids Res.* **12**, 2461–2472.
6. Robertson, B. H., Grubman, M. J., Weddell, G. N., Moore, D. M., Welsh, J. D., Fischer, T., Dowbenko, D. J., Yansura, D., Small, B., and Kleid, D. G. (1985) *J. Virol.* **54**, 651–660.
7. Ryan, M. D., King, A. M. Q., and Thomas, G. P. (1991) *J. Gen. Virol.* **72**, 2727–2732.
8. Donnelly, M., Gani, D., Flint, M., Monaghan, S., and Ryan, M. D. (1997) *J. Gen. Virol.* **78**, 13–21.
9. Palmenberg, A. C. (1990) *Annu. Rev. Microbiol.* **44**, 603–623.
10. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) *J. Am. Chem. Soc.* **106**, 765–785.
11. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem.* **7**, 230–252.
12. InsightII, Discover, Biosym/MSI, San Diego, CA.
13. Piela, L., Nemethy, G., and Scheraga, H. A. (1987) *Biopolymers* **26**, 1587–1600.
14. Deber, C. M., Glibowicha, M., and Woolley, G. A. (1990) *Biopolymers* **29**, 149–157.
15. Hahn, H., and Palmenberg, A. C. (1996) *J. Virol.* **70**, 6870–6875.
16. Blundell, T., Jenkins, J., Pearl, L., Sewell, T., and Pederson, V. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.), de Gruyter, New York.
17. Thomas, P., and Ryan, M. D. (Unpublished results on recombinant influenza virus haemagglutinin).
18. Donnelly, M. (1997) Ph.D. dissertation, Univ. St. Andrews, Scotland.
19. Donnelly, M., Gani, D., and Ryan, M. D. (1998) Submitted for publication.
20. Spirin, A. S. (1985) *Prog. Nucleic Acids Res. Mol. Biol.* **32**, 75–114.
21. Lenman, M. M., Lewis, A., and Gani, D. (1997) *J. Chem. Soc. Perkin Trans. I*, 2297–2311.
22. Rychlik, I., Cerna, J., Chladek, S., Pulkrabek, P., and Zemlicka, J. (1970) *Eur. J. Biochem.* **16**, 136–142.
23. Spirin, A. S. (1985) *Prog. Nucleic Acids Res. Mol. Biol.* **32**, 75–114.
24. Lim, V. L., and Spirin, A.S. (1986) *J. Mol. Biol.* **188**, 565–577.
25. Svitkin, Y. V., and Agol, V. I. (1983) *Eur. J. Biochem.* **133**, 145–154.
26. Pestku, S., Hishizawa, T., and Lessard, J. L. (1970) *J. Biol. Chem.* **245**, 6208–6219.
27. Moldave, K. (1985) *Annu. Rev. Biochem.* **54**, 1109–1149.
28. Gani, D. and Wilkie, J. (1997) in *Structure and Bonding* (Hill, H. A., Sadler, P., and Thompson, A., Eds.) Vol. 89, pp. 133–175, Springer-Verlag, Heidelberg.
29. Wilkie, J., Cole, A. G., and Gani, D. (1995) *J. Chem. Soc. Perkin Trans. I*, 2709–2727.
30. Schulz, J., Wilkie, J., Beaton, M. W., Miller, D. J., and Gani, D. (1998) *Biochem. Soc. Trans.* **26**, 315–322.

31. Kim, S. H., Suddath, F. L., Quigley, G. J., McPherson, A., Sussman, J. L., Wang, A.H. J., Seeman, N.C., and Rich, A. (1974) *Science* **185**, 435–440.
32. Samaha, R., Green, R., and Noller, H. F. (1995) *Nature* **377**, 309–314.
33. Green, R. Samaha, R., and Noller, H. F. (1997) *J. Mol. Biol.* **266**, 40–50.
34. Evans, J. N. S. (1995) *in* Biomolecular NMR Spectroscopy, Oxford Univ. Press, New York.
35. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *in* Molecular Cloning: A Laboratory Manual (Nolan, C., Ed.), Cold Spring Harbor Laboratory Press, New York.