Site-Specific Mutations at a Picornavirus VP3/VP1 Cleavage Site Disrupt In Vitro Processing and Assembly of Capsid Precursors

GRIFFITH D. PARKS* AND ANN C. PALMENBERG

Institute for Molecular Virology and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53706

Received 19 June 1987/Accepted 13 August 1987

Most proteolytic cleavages within the picornavirus polyproteins are carried out by viral protease 3C. For encephalomyocarditis virus, the protease 3C-catalyzed processing occurs between Gln-Gly or Gln-Ser amino acid pairs which are flanked by proline residues, but the sequence-specific constraints on recognition and cleavage by the enzyme are not completely understood. To examine alternative cleavage site sequences, we constructed a cDNA plasmid which expresses the viral L-P1-2A capsid precursor in vitro and introduced site-specific mutations into the Gln-Gly pair at the VP3/VP1 junction. The altered protein substrates were tested for cleavage activity in assays with protease 3C. The encephalomyocarditis virus 3C processed Gln-Ala as efficiently as its natural sites but did not cleave Gln-Val, Gln-Glu, Lys-Gly, Lys-Ala, Lys-Val, Lys-Glu, or Pro-Gly combinations. Displacement of the flanking proline residue by an engineered insertion slowed but did not prevent cleavage at this site. Also, a mutant defective in processing at the VP3/VP1 junction was unable to form 14S pentameric assembly intermediates in vitro.

Encephalomyocarditis (EMC) virus is the prototype of the cardiovirus subgroup of picornaviruses. Its genome consists of a single strand of messenger-sense RNA (16). Viral proteins are expressed from a large open reading frame encoding a giant precursor protein (approximately 255,000 daltons [Da]) which is processed through a series of proteolytic cleavages to produce both capsid and nonstructural proteins (4, 10). Most aspects of the viral life cycle, from replication to assembly, are controlled by the cleavage cascade. The processing events can be divided into three phases: a cotranslational primary cleavage which releases the capsid precursor from the growing polypeptide chain, a series of sequential polyprotein cleavages catalyzed by viral protease 3C, and a final capsid maturation cleavage occurring as the last step in virion assembly after association of the viral RNA with the protein shell (5, 10, 13, 23).

Most of these processing events are carried out by the 3C enzyme (3, 10). 3C is a cysteine-type protease (1) which is capable of monomolecular self-cleavage reactions within the P3 region precursors (3ABCD to 3AB, 3C, and 3D) (14, 19) and also responsible for bimolecular processing steps within the P1 and P2 regions of the polyprotein (3, 13, 24). For the EMC virus 3C protease, cleavage of the polyprotein normally occurs between glutamine-glycine or glutamine-serine amino acid pairs which are usually flanked by proline residues (12), but the mechanism which limits 3C protease processing to specific sites is not known.

We have developed a cell-free system for the analysis of cardiovirus processing events based on cDNA clones which encode defined portions of the EMC virus genome, including an active 3C protease (15). As a convenient source of capsid region substrate, we have also constructed a plasmid which efficiently expresses the L-P1-2A precursor protein in vitro. To examine the sequence-specific constraints on the 3C enzyme, site-specific mutations were introduced into the glutamine-glycine pair at the 1C/1D cleavage site and tested for their effects in cell-free protease assays.

MATERIALS AND METHODS

Nomenclature. EMC virus proteins are designated by the standard L-4-3-4 format (18) except for viral capsid polypeptides 1AB (VP0), 1C (VP3), and 1D (VP1), for which the traditional nomenclature is used interchangeably. Capsid precursor L-P1-2A was previously referred to as A1.

EMC virus 3C protease purification fractions F2 and F4 have been described elsewhere (15). Briefly, EMC virus-infected HeLa cell lysate was clarified by high-speed centrifugation (F2), and the protein in the resulting supernatant was fractionated by ion-exchange chromatography. The 3C protease-containing protein eluted by 0.35 M NaCl was further fractionated by gel filtration chromatography (F4). Fluorograms of the F4 fraction revealed a single radiolabeled 3C protease band.

Proteolytic cleavage site nomenclature follows that outlined by Kraut (6). A site is generalized as $\dots P_2-P_1-P_1'-P_2' \dots$, where scission occurs at the peptide bond between the P_1 and P_1' residues.

Construction of plasmid DNA. Restriction enzymes, T4 DNA ligase, and the Klenow fragment of DNA polymerase I were purchased from New England BioLabs, Inc. DNA manipulations were done by standard procedures (9). All bacterial transformations were to ampicillin resistance, with *Escherichia coli* HB101 or C600.

Plasmid pE13 encodes the complete L fragment sequences of EMC virus, and its construction will be described in detail elsewhere (G. Duke and A. Palmenberg, manuscript in preparation). Briefly, pE13 contains EMC virus-specific sequences ranging from base 320 near the viral poly(C) tract to the 3' poly(A) tail. The EMC virus fragment was derived by primer extension with reverse transcriptase on virion RNA and is contained within a pSPT18 parental vector (Pharmacia Fine Chemicals). Plasmids pE3T11 and pEM3 have been described elsewhere (12, 15).

To construct clone pE51D2A, DNA from pEM3 (12) was digested with *Pvu*II, and the 929-base-pair (bp) fragment (bases 3297 to 4226, which encode part of VP1, all of 2A, and

* Corresponding author.

FIG. 1. Structure of the EMC virus genome and clone pE5A1. Relevant restriction sites and portions of the EMC virus genome contained within pE5A1 are illustrated. Abbreviations: T7 and PL, T7 RNA polymerase promoter and polylinker cloning sites from the parental vector, respectively; VPg, viral protein covalently attached to the 5' end of the genomic RNA; 5' NC, EMC virus 5' noncoding sequences. Symbol: [stippled box], proteins (L-P1-2A and p8) produced from pE5A1, as described in the text.

part of 2B) was purified by gel electrophoresis. *Bam*HI linkers (catalog number 1003; New England BioLabs) were added to the ends of the DNA by incubation with T4 DNA ligase. After gel filtration, the DNA fragment was treated with *Ava*I and *Bam*HI, and the resulting 720-bp segment was isolated by gel electrophoresis. In a separate reaction, 6 μ g of pEM3 was digested with *Xmn*I and *Ava*I, and the 851-bp fragment (bases 2655 to 3506, which encode VP1) was gel purified. The *Xmn*I-*Ava*I and *Ava*I-*Bam*HI fragments were ligated (T4 DNA ligase, 10 U, 12 h, 12°C) together with 0.5 μ g of plasmid pE3T11, which had been previously digested with *Bal*I and *Bam*HI. Colonies were screened for the correct orientation of the inserted 1D-2A(2B) fragment by restriction enzyme analysis.

To construct pE5A1, 2.0 μ g of pE51D2A DNA was digested to completion with *Mst*II and *Bam*HI. The resulting 1,532-bp fragment (bases 2698 to 4230, which encode VP1 to the mid 2B region) was purified by gel electrophoresis before ligation (18 U, 16°C, 5 h) into 0.5 μ g of pE13 (15), which had been previously cut with *Mst*II, *Nru*I, and *Bam*HI. After transformation, bacterial colonies were screened for the desired construction by digestion with restriction enzymes.

For construction of pE5A1M, 6 μ g of pE5A1 was digested with *Mst*II. After precipitation with ethanol, the DNA was dissolved in 18 μ l of buffer (30 mM sodium acetate [pH 5.0], 0.1 M NaCl, 2 mM ZnCl₂). The mixture was heated to 70°C for 5 min and then held at 4°C for 1 min before the addition of 1 μ l (15 U) of mung bean nuclease (New England BioLabs). After incubation at 22°C for 10 min, NaCl was added to a concentration of 0.2 M, and the sample was extracted with phenol-chloroform (1:1), precipitated with ethanol, and then treated with ligase (50 U, 12 h, 22°C) before use in bacterial transformations.

Likewise, clone pE5A1K was constructed from pE5A1 (4 μ g). However, after digestion with *Mst*II, the four deoxynucleotides were added to a final concentration of 0.4 mM, and the sample was incubated at 22°C for 20 min with 10 U of DNA polymerase Klenow fragment and then precipitated with ethanol. The resulting pellet was dissolved in buffer and treated with ligase, as described above, before it was used in transforming bacteria. Plasmid DNA contained within ampicillin-resistant colonies was screened for loss of the *Mst*II site, and a representative construction was confirmed by dideoxynucleotide sequencing (12).

Site-specific mutagenesis. Oligonucleotide-directed muta-

genesis was done essentially as described elsewhere (26). The 557-bp *Xba*I fragment, which encodes the EMC virus 1C/1D cleavage site, was excised from pE5A1 by digestion with *Xba*I and purified by gel electrophoresis before insertion into the *Xba*I site of the replicative form of M13mp19. The orientation of the insert was such that the single-stranded bacteriophage DNA contained EMC virus-coding sequences. After plaque purification, the phage DNA was enriched in uracil (7) by infection of Ung⁻ Dut⁻ bacteria (*E. coli* RZ1032; gift from D. Lichtenstein). After uracil enrichment, phage DNA (0.5 μ g) was used as a template for in vitro mutagenesis by primer extension with 5 U of DNA polymerase Klenow fragment, 6 U of T4 DNA ligase (22°C, 2 h), and a synthetic oligonucleotide (3 to 30 pmol, 3'-CCTCGG GANTTCMTCATCTTTTGC-5', where N is G, C, or T and M is C, A, T, or G). The underlined T within the primer sequence was incorporated to destroy the *Mst*II restriction site and was used to distinguish wild-type from mutant DNA in the subsequent cloning steps. After transformation of *E. coli* JM101 with the double-stranded DNA, resulting phage isolates were screened by dideoxynucleotide sequencing and the desired M13 recombinants were plaque purified. The frequencies of primer-generated single and double mutants were approximately 80 and 10%, respectively. When necessary, a second round of mutagenesis was carried out as before by using the same primer mix with uracil-enriched mutant DNA.

The mutant EMC virus fragments were excised by *Xba*I digestion of the replicative form of M13 and substituted for the corresponding wild-type sequences in pE5sub (a subclone containing the EMC virus VP3-2A region within a pSPT18 vector). After confirming the orientation of the subcloned *Xba*I insert, the 1,100-bp fragment which encoded the mutant cleavage site was removed from pE5sub by digestion with *Spe*I and *Acc*I, purified by gel electrophoresis, and inserted in place of the corresponding sequences of pE5A1 for expression. The site-specific mutations within the resulting pE5A1-derived clones were reconfirmed by sequence analysis before use in protease assays.

In vitro transcription and cell-free translation. Purified plasmid DNA was linearized with *Bam*HI before transcription. Plasmid DNA (1 μ g) was transcribed in vitro (25- μ l reaction) by using 40 to 60 U of T7 RNA polymerase (Bethesda Research Laboratories, Inc.) exactly as described previously (15). After 1 h at 37°C, samples were extracted

FIG. 2. Expression and processing of L-P1-2A produced from pE5A1 and pE5A1M. RNA transcripts from pE5A1 and pE5A1M were used to program translation reactions (W.T. and -Gln, lanes 1 and 4, respectively). Identical portions (2 μ l) were held with 3 μ l of buffer (phosphate-buffered saline; 20) and with 3 μ l of fraction F4 protease at 30°C for 1 h (lanes 2 and 5) or 12 h (lane 3 and 6). Lane 7, Products of a 12-h incubation of pE5A1M-derived proteins (2 μ l) with fraction F2 protease (3 μ l). Lane M, EMC virion RNA translation products. The arrow indicates the position of the uncleaved capsid precursor 1CD (VP3/VP1).

with phenol-chloroform (1:1) and precipitated with ethanol. The resulting pellets were dried under vacuum and dissolved in 10 μ l of water.

A portion (2 to 3 μ l) of the transcription products was used to program cell-free translation (30- μ l reaction) in reticulocyte extracts (22). [35 S]methionine (1,100 Ci/mmol) was included at a concentration of about 1 μ Ci/ μ l. EMC virion RNA was isolated as described previously (17). All reactions were terminated at the desired time by adding cycloheximide and pancreatic RNase (0.33 mg/ml each).

Sucrose gradient analysis of capsid assembly intermediates. Cell-free protein synthesis directed by RNA transcripts from plasmids pE5A1 and pE5A1M was carried out as described above. Translation extracts (25 μ l, approximately 7×10^6 acid-insoluble cpm) were mixed with 10 μ l of nonradiolabeled 3C protease produced by in vitro translation of RNA transcripts from clone pE5P3 (15). After incubation (22 h, 30°C), samples were diluted with buffer, layered onto 5 to

FIG. 3. Construction of pE5A1K and pE5A1M. A portion of the nucleotide and amino acid sequences (one-letter codes) surrounding the VP3/VP1 (1C/1D) junction are shown. The procedures used to engineer an insertion (pE5A1K) and a deletion (pE5A1M) near the VP3/VP1 cleavage site are outlined. The *Mst*II and *Sma*I restriction sequences are underlined. The histidine (H) inserted into the protein sequence of pE5A1K is shown in boldface type.

20% (wt/wt) sucrose gradients, and centrifuged exactly as described previously (11). Fractions (350 μ l) were collected from the bottom of the tubes, and portions (75 μ l) from each fraction were assayed for trichloroacetic acid-insoluble radioactivity (11).

Gel electrophoresis. Analytical polyacrylamide gel electrophoresis was carried out as described previously (13) except that the *N,N'*-methylenebisacrylamide concentration was increased from 0.2 to 0.26% (wt/vol) to aid in the separation of viral proteins 1D and 3ABC. Radiolabeled proteins were visualized by autoradiography of dried slab gels (13).

RESULTS

Synthesis of clone-derived L-P1-2A substrate. We constructed a cDNA plasmid encoding the EMC virus L-P1-2A capsid region precursor. The key elements of pE5A1 (EMC virus, 5' noncoding, plus A1 protein) are outlined in Fig. 1. This plasmid contains EMC virus-specific sequences which extend from the 3' side of the viral poly(C) tract (base 320) to the middle of the 2B protein-coding region (base 4220). Plasmid DNA was transcribed in vitro, and the resulting products were used to program cell-free translation in reticulocyte extracts.

RNA derived from pE5A1 directed the synthesis of a

FIG. 4. Proteolytic processing of L-P1-2A produced from pE5A1 and pE5A1K. Plasmids pE5A1 (W.T.) and pE5A1K (+His) were used to program cell-free translations, and samples (1 μ l) were saved from each reaction (lanes 0). Portions of the translation extracts (20 μ l, approximately 1.5×10^7 acid-insoluble cpm) were mixed with 5 μ l of phosphate-buffered saline and 10 μ l of nonradiolabeled fraction F2 protease (15) and held at 30°C. Samples (4 μ l) were removed at the indicated times (in minutes, lanes 5 to 75) and analyzed by gel electrophoresis (13).

protein (Fig. 2, lane 1) with an electrophoretic mobility matching that of authentic L-P1-2A (lane M). The identity of the band migrating with the mobility of L-P1-2A was further supported by immunoprecipitation with polyclonal antiserum specific for ME virions, a serologically identical cardiocivirus (data not shown).

The EMC virus-specific DNA sequences contained within

pE5A1 predict that cleavage at the 2A/2B junction should produce a small 2B-derived polypeptide (molecular weight, 8,000), in addition to L-P1-2A. We have consistently observed a small peptide (p8) on our gels (see Fig. 2) which maps by deletion experiments to the 3' end of the cloned open reading frame of pE5A1 and is not the result of a second translational initiation event (data not shown, manuscript in preparation). Our data suggest that accurate and efficient 2A/2B cleavage occurs in pE5A1 clone-derived viral polypeptides and support the contention that primary scission of the EMC virus polyprotein does not require any proteins from the 3' half of the genome (including protease 3C) (4, 13, 22).

Protein precursors produced from pE5A1 were tested for the ability to be processed by 3C protease in assays with purified fraction F4 protease (15). The resulting capsid region profile, including peptides P1-2A, P1, 1ABC, 1AB, 1D, 1C, 2A, and L, was nearly identical to that produced by cell-free translation of EMC virion RNA (Fig. 2, lanes 2, 3, and M). Thus, L-P1-2A, produced from pE5A1, is accurately processed by exogenous 3C enzyme. Interestingly, incubation with protease fractions sometimes also resulted in the disappearance of protein p8 from the gels (e.g., Fig. 2, lane 7 and Fig. 6, lanes 2 and 4). We do not know the agent responsible for this loss (possibly 3C protease), but we suspect that the 2B-derived protein fragment may simply be unstable during prolonged reactions.

Effect of insertion or deletion of amino acids at a 3C cleavage site. The nucleotide sequence of EMC virus cDNA contains a unique *Mst*II restriction site in the region encoding the 1C/1D cleavage point (Fig. 3) (12), a junction known to be processed by the 3C protease (13). The amino acids at this cleavage site (Pro-Gln/Gly-Val) have been identified for mengovirus, a closely related cardiocivirus (25). By taking advantage of the *Mst*II site within pE5A1, we produced modified plasmids which express L-P1-2A protein precursors containing mutations at the 1C/1D site.

The steps in the construction of pE5A1M and pE5A1K are outlined in Fig. 3. Briefly, DNA from pE5A1 was digested with *Mst*II, and the resulting 3-base overhang was either

FIG. 5. Site-specific mutagenesis of the VP3/VP1 cleavage site. A segment of the EMC virus sequences encoding the 1C/1D cleavage site is shown. The upper strand represents the 24-base synthetic primer mix used during mutagenesis of this region, with the nucleotide and amino acid substitutions shown as vertical columns. All primers contained a nucleotide substitution from T to G at position 3 of the Gln codon (*). The *Mst*II restriction sequence is underlined.

FIG. 6. Proteolytic processing of L-P1-2A precursors containing site-specific mutations at the 1C/1D junction. Oligonucleotide-directed mutants were constructed as described in the text. DNA from plasmids containing Gln-Ala (QA) or Gln-Glu (QE) mutations at the 1C/1D cleavage sites were transcribed *in vitro*, and the resulting products were translated in reticulocyte extracts. Portions (2 μ l) of the translation mixtures were either saved (lanes 1 and 3) or mixed with 3 μ l of phosphate-buffered saline and 3 μ l of nonradio-labeled F2 protease (lanes 2 and 4). After 1 h of incubation at 30°C, samples were analyzed on sodium dodecyl sulfate slab gels. The arrow indicates the position of the 1CD precursor protein.

removed by treatment with a single-stranded specific nuclease to generate pE5A1M (pE5A1, mung bean nuclease) or repaired by reaction with DNA polymerase to yield pE5A1K (pE5A1, Klenow fragment). The 1C/1D cleavage sites encoded within pE5A1M and pE5A1K are altered by the deletion of a glutamine (Pro-Gly-Val) or the insertion of a histidine (Pro-His-Gln-Gly-Val), respectively.

To test whether the mutated 1C/1D sites could still be processed, RNA transcripts produced from pE5A1M or pE5A1K were used to program protein synthesis in reticulocyte extracts. Each clone produced an L-P1-2A precursor protein as well as p8 (Fig. 2 and 4). Nonradiolabeled 3C protease was incubated with the clone-derived substrates, and the resulting cleavage products were analyzed by gel electrophoresis and autoradiography.

Processing of protein pE5A1 (wild type [WT]) yielded a detectable 1D band after 1 h of incubation with protease (Fig. 2, lane 2). Under similar conditions, precursors from pE5A1M (minus Gln) showed no sign of cleavage at the

1C/1D site (lane 5). Longer incubation (12 h) with purified 3C protease (lane 6) or with highly active fraction F2 protease (lane 7) also failed to produce 1D protein with substrates from pE5A1M. Rather, the mutant processing profiles included an abnormal amount of precursor 1CD (Fig. 2, arrow), a protein usually seen in very low levels when WT capsid precursors are processed (lanes 3 and M). The appearance of 1AB, L, and 2A in the mutant-directed samples indicates that only 1C/1D processing was inhibited by the engineered alteration. These experiments support the contention that the glutamine residue in the P₁ position of the EMC virus 1C/1D site plays an essential role in processing by the 3C protease (3).

Clone pE5A1K was used to determine whether 3C protease could process a 1C/1D site containing an insertion which displaced the flanking (P₂) proline residue. A typical time course of 3C protease-catalyzed reactions with precursor protein encoded by pE5A1 (WT) or pE5A1K (plus His) is shown in Fig. 4. As early as 5 min after the addition of 3C protease, pE5A1 (WT) samples contained detectable amounts of 1D, while reactions with the histidine insertion mutant (plus His) produced few precursors which had been cleaved past the P₁ stage (compare 5-min lanes). After further incubation, however, the WT and plus His cleavage profiles were virtually identical (45- and 75-min lanes). Repetition of this experiment always gave similar results. We conclude that 3C protease-catalyzed processing of the 1C/1D junction was slowed, but not prevented, by the engineered histidine insertion.

Oligonucleotide-directed mutagenesis of the 1C/1D cleavage site. To further test the effect of cleavage site alterations on 3C protease activity, we employed oligonucleotide-directed mutagenesis to introduce specific substitutions into the P₁ and P₁' positions of the 1C/1D site. Mutants were generated *in vitro* by using M13 phage and synthetic primers. The

FIG. 7. Summary of capsid region cleavage site sequences tested for EMC virus 3C protease activity. Site-specific mutations were constructed at the 1C/1D proteolytic cleavage site (Gln/Gly). The altered capsid precursors containing P₁ (Gln, Lys, or Pro) and P₁' (Gly, Ala, Val, or Glu) combinations at this site were assayed for the ability to be processed (+ or -) in reactions with 3C protease (see the legend to Fig. 6). The Gln-Ser^a pair occurs naturally at the EMC virus 1AB/1C junction and is included for completeness. The Pro-Gly^b combination is the result of an engineered deletion and may not be directly analogous to the substitutions. ND, Not determined.

FIG. 8. Sucrose gradient analysis of pE5A1- and pE5A1M-derived viral proteins. Cell-free translation products directed by plasmids pE5A1 and pE5A1M were incubated with 3C protease and analyzed on 5 to 20% sucrose gradients as described in Materials and Methods. (A) Gradient profile of pE5A1-derived proteins. $\rightarrow$, Position of the rabbit hemoglobin marker (4.2S; 11). (B) Gradient profile of pE5A1M-derived proteins. (C) A sample (200 μ l) from fraction 22 of the gradient shown in panel A was concentrated and analyzed by gel electrophoresis (13).

primer mix (Fig. 5) was selected to maximize the variation in amino acid pairs which could be easily constructed. After isolation and characterization of the mutant clones by sequence analysis, protein precursors containing 1C/1D substitutions were expressed *in vitro* and then tested for cleavage with 3C protease.

Results with two such mutants are shown in Fig. 6. Cell-free expression of the Gln-Ala (QA) and Gln-Glu (QE) isolates each produced proteins L-P1-2A and p8 (lanes 1 and 3). However, subsequent reactions with 3C protease indicated a marked difference in the processing of these substrates. Cleavage of the 1C/1D site (appearance of 1C and 1D) was evident with the Gln-Ala sample (lane 2) but not with precursors containing a Gln-Glu (lane 4). A Gln-Ala but not a Gln-Glu therefore can function *in vitro* in place of the normal Gln-Gly pair.

Five other mutant dipeptide combinations were also isolated, characterized, and tested as substrates for 3C protease. These included Gln-Val, Lys-Gly, Lys-Ala, Lys-Val, and Lys-Glu pairs. The results of assays with precursors containing these mutations are summarized in Fig. 7. None of the additional 1C/1D mutant combinations were cleaved at this site, giving processing profiles identical to that obtained with the Gln-Glu isolate (data not shown). For unknown reasons, we had difficulty isolating specific mutant combinations with glutamate in the P₁ position. However, experiments designed to test the effects of glutamate substitutions as well as other dipeptide sequences are under way.

Cleavage at the 1C/1D site is a prerequisite to assembly of capsid pentamers. One of the first steps in picornavirus capsid assembly is the formation of 14S pentameric structures by the association of five viral capsid protomers (1AB, 1C, and 1D) (2, 16). Cell-free translation of EMC virion RNA produces 14S pentamers during prolonged incubation (11). To determine whether capsid precursors defective in processing at the 1C/1D junction could also form 14S pentamers, proteins expressed from pE5A1M were assayed for assembly activity *in vitro*. Radiolabeled L-P1-2A capsid precursors were synthesized from RNA transcripts derived from pE5A1M or from pE5A1 and then incubated with nonradiolabeled 3C protease. The resulting cleavage profiles were analogous to those shown in Fig. 2 (lanes 3 and 7). Samples

were then fractionated on sucrose gradients (Fig. 8) to identify assembly intermediates.

In agreement with previous work (11), proteins derived from pE5A1 gave a large promoter peak (5S), migrating slightly faster than the globin marker (4.2S) and a smaller peak of approximately 14S (Fig. 8A). This 14S complex was composed exclusively of VP0, VP1, and VP3, as expected for authentic capsid pentamers (Fig. 8C) (11, 16). In contrast, the sample derived from pE5A1M yielded only a protomer peak, with no detectable 14S material (Fig. 8B).

DISCUSSION

The cleavage reactions in the capsid region of EMC virus are catalyzed by the 3C protease (4, 13, 15). As a convenient source of substrate for the analysis of these reactions, we constructed a cDNA clone which expresses the viral L-P1-2A precursor *in vitro*. Previously, cardiovirus capsid region substrates were obtained by stopping cell-free translations of virion RNA after short incubation periods, a method requiring precise timing and the addition of inhibitors to prevent synthesis of 3C protease (13, 22). We have seen no evidence for the spontaneous processing of clone-derived L-P1-2A into P1-2A or P1, as reported with precursors obtained from limited translations (21). Since cleavage of the leader protein (L) from L-P1-2A is now known to be catalyzed by 3C protease (15), we suggest that small amounts of P3 region proteins could have contaminated these previous preparations.

Plasmid pE5A1 contains EMC virus sequences spanning the 2A/2B primary cleavage site. Cell-free expression of these sequences produced the proteins expected from accurate primary cleavage, providing direct experimental evidence that this cardiovirus cotranslational scission does not require any proteins encoded downstream of 2B. The primary EMC virus cleavage event occurs by a mechanism different from that of poliovirus (4, 10, 23), and although we have not yet identified the agent responsible for the 2A/2B scission, our clones now provide an excellent cell-free system for further study of this event.

Our analysis of EMC virus 3C protease cleavages in the capsid region has focused on the 1C/1D junction for several

reasons. First, for cardioviruses, the location of this particular 3C protease cleavage site has been confirmed by direct amino acid sequencing (25). Second, within the normal EMC virus P1 region processing cascade, two 3C protease-directed cleavages (L/P1 and P1/2A) invariably occur before the 1C/1D scission (4, 15, 21), providing positive internal controls which assure that engineered alterations have not grossly distorted the L-P1-2A substrates. Third and most important, the position of a unique *Mst*II restriction site within the DNA encoding the 1C/1D junction allowed the easy construction of two useful plasmid isolates containing an insertion and a deletion.

With one of these mutant plasmids (pE5A1K), we found that 3C protease could effectively process a 1C/1D site where the flanking (P_2) proline was displaced by an insertion, albeit at a somewhat slower rate than in reactions with WT sequences. Flanking proline residues are found at nearly all cardiovirus 3C protease-catalyzed cleavage sites (8, 12). Our experiments with the histidine insertion mutant do not clarify whether a neighboring proline contributes to sequence or structural recognition by the enzyme, but they do show that exact proximity to the 1C/1D site is not essential. We are currently constructing site-specific substitutions at the P_2 position of the 1C/1D site to directly test the influence of flanking residues on 3C protease processing.

Most picornavirus 3C protease-catalyzed cleavage sites have glutamine (or glutamate) in their P_1 positions (3, 12). Deletion of the glutamine codon from the 1C/1D site of EMC virus L-P1-2A prevented processing at this junction. Likewise, site-specific mutations which substituted lysine for glutamine inactivated this site. Although we tested only a few P_1 mutations (Lys and Pro), our data are consistent with sequence-specific constraints (Gln or Glu?) on residues at this position.

In contrast, results obtained with mutants encoding substitutions at the P_1' position indicate that the specificity at this location may be somewhat more flexible. When the normal glycine residue was replaced by alanine, efficient cell-free 1C/1D processing still occurred, even though this amino acid does not appear naturally at any EMC or mengovirus 3C protease cleavage site (8, 12). Unrestrained substitution is not permissible, as P_1' mutations encoding valine or glutamate inactivated 1C/1D processing. We suggest that P_1' sequence constraints at functional EMC virus 3C protease sites are probably limited to small, uncharged amino acids (Gly, Ala, or Ser). The P_1' residue alone is not sufficient to define a cleavage site, however, since Lys-Ala and Lys-Gly mutants were inactive in our assays. Therefore, activity is clearly dictated by acceptable amino acids at both the P_1 and P_1' positions (Fig. 7). It is important to emphasize, however, that our results reflect the analysis of the EMC virus 1C/1D junction and may not necessarily hold true for other 3C protease cleavage sites.

The 3C protease-catalyzed proteolytic cascade within EMC virus capsid precursor proteins normally occurs in a defined stepwise manner, with sequential release of L, 2A, 1D, and then 1AB plus 1C (4, 21, 22). While testing our defective substrates for activity, it was somewhat surprising to detect efficient 1AB/1C cleavage, a reaction usually preceded by 1C/1D scission (15). The mechanism by which the ordered steps are selected is still unknown, but it is now apparent that the usual *in vitro* 3C protease processing pattern for this region is preferential but not obligatory.

Proteolytic processing has been suggested to play an important role in the formation of picornavirus capsid structures (2, 16). Although no EMC virus capsid assembly

intermediates composed of uncleaved P1 proteins have ever been isolated, direct evidence for a cleavage prerequisite to assembly was never completely established. We have tested this hypothesis by using mutant pE5A1M (minus Gln), which produces capsid precursors defective in 1C/1D processing. The P1 and 1AB plus 1CD cleavage products, produced by 3C protease processing of pE5A1M-derived proteins, were unable to form stable pentameric assembly structures, one of the first steps in capsid formation (16). The results are in complete agreement with crystallographic data from picornavirus virions, indicating that interaction of the processed ends may be critical for pentamer stability (2, 8).

ACKNOWLEDGMENTS

We thank Roland Rueckert for helpful discussions and also Gregory Duke for the gift of pE13 plasmid DNA.

This work was supported by Public Health Service grant AI-17331 from the National Institutes of Health to A.C.P. G.D.P. was supported by American Cancer Society grant MV-33 to R. R. Rueckert.

LITERATURE CITED

- Argos, P., G. Kamer, M. J. H. Nicklin, and E. Wimmer. 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Res.* 12:7251-7267.
- Arnold, E., M. Luo, G. Vriend, M. G. Rossmann, A. C. Palmenberg, G. D. Parks, M. J. H. Nicklin, and E. Wimmer. 1987. Implications of the picornavirus capsid structure for polyprotein processing. *Proc. Natl. Acad. Sci. USA* 84:21-25.
- Hanecak, R., B. L. Semler, C. W. Anderson, and E. Wimmer. 1982. Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. *Proc. Natl. Acad. Sci. USA* 79:3973-3977.
- Jackson, R. J. 1986. A detailed kinetic analysis of the *in vitro* synthesis and processing of encephalomyocarditis virus products. *Virology* 149:114-127.
- Jacobson, M. F., J. Asso, and D. Baltimore. 1970. Further evidence on the formation of poliovirus proteins. *J. Mol. Biol.* 49:657-669.
- Kraut, J. 1977. Serine proteases: structure and mechanism of catalysis. *Annu. Rev. Biochem.* 46:331-358.
- Kunkel, T. A. 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. *Proc. Natl. Acad. Sci. USA* 82:488-492.
- Luo, M., G. Vriend, G. Kamer, I. Minor, E. Arnold, M. G. Rossmann, U. Boege, D. G. Scraba, G. M. Duke, and A. C. Palmenberg. 1987. The atomic structure of Mengo virus at 3.0 Å resolution. *Science* 235:182-191.
- Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Nicklin, M. J. H., H. Toyoda, M. G. Murray, and E. Wimmer. 1986. Proteolytic processing in the replication of polio and related viruses. *Biotechnology* 4:33-42.
- Palmenberg, A. C. 1982. *In vitro* synthesis and assembly of picornaviral capsid intermediate structures. *J. Virol.* 44:900-906.
- Palmenberg, A. C., E. M. Kirby, M. R. Janda, N. L. Drake, G. M. Duke, K. F. Potratz, and M. S. Collett. 1984. The nucleotide and deduced amino acid sequences of the encephalomyocarditis viral polyprotein coding region. *Nucleic Acids Res.* 12:2969-2985.
- Palmenberg, A. C., M. A. Pallansch, and R. R. Rueckert. 1979. Protease required for processing picornaviral coat protein residues in the viral replicase gene. *J. Virol.* 32:770-778.
- Palmenberg, A. C., and R. R. Rueckert. 1982. Evidence for intramolecular self-cleavage of picornaviral replicase precursors. *J. Virol.* 41:244-249.

15. Parks, G. D., G. M. Duke, and A. C. Palmenberg. 1986. Encephalomyocarditis virus 3C protease: efficient cell-free expression from clones which link viral 5' noncoding sequences to the P3 region. *J. Virol.* **60**:376-384.
16. Rueckert, R. R. 1985. Picornaviruses and their replication, p. 705-736. In B. N. Fields (ed.), *Virology*. Raven Press, New York.
17. Rueckert, R. R., and M. A. Pallansch. 1981. Preparation and characterization of encephalomyocarditis virus. *Methods Enzymol.* **78**:315-325.
18. Rueckert, R. R., and E. Wimmer. 1984. Systematic nomenclature of picornavirus proteins. *J. Virol.* **50**:957-959.
19. Semler, B. L., V. H. Johnson, P. G. Dewalt, and M. F. Ypma-Wong. 1987. Site-specific mutagenesis of cDNA clones expressing a poliovirus proteinase. *J. Cell. Biochem.* **33**:39-51.
20. Sherry, B., and R. Rueckert. 1985. Evidence for at least two dominant neutralization antigens on human rhinovirus 14. *J. Virol.* **53**:137-143.
21. Shih, C.-Y. T., and D. S. Shih. 1981. Cleavage of the capsid protein precursors of encephalomyocarditis virus in rabbit reticulocyte lysates. *J. Virol.* **40**:942-945.
22. Shih, D. S., C. T. Shih, D. Zimmern, R. R. Rueckert, and P. Kaesberg. 1979. Translation of encephalomyocarditis virus RNA in reticulocyte lysates: kinetic analysis of the formation of virion proteins and a protein required for processing. *J. Virol.* **30**:472-480.
23. Toyoda, H., M. J. H. Nicklin, M. G. Murray, C. W. Anderson, J. J. Dunn, F. W. Studier, and E. Wimmer. 1986. A second virus-encoded proteinase involved in proteolytic processing of poliovirus polyprotein. *Cell* **45**:761-770.
24. Ypma-Wong, M. F., and B. L. Semler. 1987. In vitro molecular genetics as a tool for determining the differential cleavage specificities of the poliovirus 3C proteinase. *Nucleic Acids Res.* **15**:2069-2088.
25. Ziola, B. R., and D. G. Scraba. 1976. Structure of the Mengo virion: amino- and carboxyl-terminal analysis of the major capsid polypeptides. *Virology* **71**:111-121.
26. Zoller, M. J., and M. Smith. 1983. Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors. *Methods Enzymol.* **100**:468-500.