JOURNAL OF VIROLOGY, Mar. 1995, p. 1727–1733 Vol. 69, No. 3
0022-538X/95/$04.00+0


# Cleavage Specificity of Purified Recombinant Hepatitis A Virus 3C Proteinase on Natural Substrates

TINA SCHULTHEISS,[1]† WOLFGANG SOMMERGRUBER,[2] YURI KUSOV,[1] AND VERENA GAUSS-MÜLLER[1]*

*Institute of Medical Microbiology, Medical University of Lübeck, 23538 Lübeck, Germany,*[1] *and Ernst-Boehringer Institut für Arzneimittelforschung, 1120 Vienna, Austria*[2]

Received 22 July 1994/Accepted 2 December 1994

**Hepatitis A virus (HAV) 3C proteinase expressed in *Escherichia coli* was purified to homogeneity, and its cleavage specificity towards various parts of the viral polyprotein was analyzed. Intermolecular cleavage of the P2-P3 domain of the HAV polyprotein gave rise to proteins 2A, 2B, 2C, 3ABC, and 3D, suggesting that in addition to the primary cleavage site, all secondary sites within P2 as well as the 3C/3D junction are cleaved by 3C. 3C-mediated processing of the P1-P2 precursor liberated 2A and 2BC, in addition to the structural proteins VP0, VP3, and VP1-2A and the respective intermediate products. A clear dependence on proteinase concentration was found for most cleavage sites, possibly reflecting the cleavage site preference of 3C. The most efficient cleavage occurred at the 2A/2B and 2C/3A junctions. The electrophoretic mobility of processing product 2B, as well as cleavage of the synthetic peptide KGLFSQ*AKISLFYT, suggests that the 2A/2B junction is located at amino acid position 836/837 of the HAV polyprotein. Furthermore, using suitable substrates we obtained evidence that sites VP3/VP1 and VP1/2A are alternatively processed by 3C, leading to either VP1-2A or to P1 and 2A. The results with regard to intermolecular cleavage by purified 3C were confirmed by the product pattern derived from cell-free expression and intramolecular processing of the entire polyprotein. We therefore propose that polyprotein processing of HAV relies on 3C as the single proteinase, possibly assisted by as-yet-undetermined viral or host cell factors and presumably controlled in a concentration-dependent fashion.**

Polyprotein processing is required for the temporal liberation of the structural and nonstructural proteins of picornaviruses prior to genome replication and capsid assembly. This mechanism of regulation of gene expression usually is mediated by proteinases (more than one) which are encoded as integral parts of the viral polyprotein. Except for hepatitis A virus (HAV), all picornaviruses studied so far have evolved at least two proteolytically active centers which are parts of domains P2 and P3 of the viral polyprotein and are encoded by the genomic regions 2A and 3C (for reviews, see references 8, 12, and 17). The enterovirus and rhinovirus protein 2A is the proteinase responsible for the so-called primary cleavage which cotranslationally separates the precursor of the structural proteins P1 from that of the nonstructural proteins. Protein 2A of the cardioviruses and aphthoviruses catalyzes cleavage of the P1-2A/2B site by an enzyme-independent mechanism (12, 25). Except for the maturation cleavage which splits VP0 into VP2 and VP4, all subsequent cleavages within the polyprotein are mediated by 3C proteinase or its precursors. In enteroviruses, the secondary cleavage cascade is characterized by competition of proteinases $3C^{pro}$ and $3CD^{pro}$, the latter being modified in its substrate specificity by the large polymerase portion. Because of its slow growth in cell culture and its mostly noncytopathic course of infection, HAV is clearly set apart from other picornaviruses and has therefore been assigned to the newly created genus *Hepatovirus* (9). Sequence alignments as well as experimental evidence argue for a different regulation of proteolytic processing in HAV, which might account for some of the exceptional features of the virus (7, 14, 16). One important aspect of these differences is the fact that protein 2A apparently lacks any proteolytic activity whereas the autocatalytic activity of 3C has been shown in various recombinant systems (10, 11, 14, 28). In addition, intermolecular cleavage activity of 3C was shown by incubation of the recombinant enzyme with substrates generated by translation in vitro (11, 13, 19, 21, 26).

Recombinant expression of viral proteinases in *Escherichia coli* has been largely facilitated by the pET expression system first described by Rosenberg et al. (24); see also references 20, 22, and 26. With this system for the expression of different parts of P2 and P3 of the HAV polyprotein, the role of 3C in the context of its flanking regions has been investigated (19, 26). These data demonstrated the ability of 3C proteinase to cleave the HAV polyprotein at all examined sites within domains P1, P2, and P3, supporting the notion that 3C is the only virus-encoded proteinase which is able to regulate the cascade of proteolytic processing of the entire HAV polyprotein by itself or in concert with as-yet-unidentified viral or host cell factors.

In order to study the cleavage pattern and regulatory pathways involved in the proteolytic cleavage cascade of HAV, we have focused on investigations that apply a highly purified recombinant HAV 3C proteinase. With a nested set of polyprotein substrates spanning all parts of the HAV polyprotein, it was possible to study 3C cleavage specificity and to define as-yet-unidentified products. Furthermore, the system allowed the evaluation and ranking of cleavage efficiencies at different sites in the HAV polyprotein. Fidelity of the processing products formed by intermolecular cleavage was warranted by comparison with products liberated from the HAV polyprotein by intramolecular cleavage. Among the processing products, proteins VP1-2A, 2A, and 2B were identified by comigration with the respective recombinant proteins expressed from

* Corresponding author. Mailing address: Institut für Medizinische Mikrobiologie, Medizinische Universität zu Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany. Phone: 49 451 500 4085. Fax: 49 451 500 3637.

† Present address: Laboratory of Infectious Diseases, NIAID, Bethesda, MD 20892-0740.

the PCR-amplified genomic region, and thus their localization on the viral genome was determined. Precursor protein VP1-2A (PX), also found in HAV-infected cells, was a stable cleavage product generated by the intermolecular action of 3C (2, 6, 18). Our data suggest that the cleavage site preference of 3C proteinase might be responsible for the temporal regulation of HAV polyprotein processing.

## MATERIALS AND METHODS

**Construction of plasmids.** Proteinase 3C was derived from construct pET-3CD* as described elsewhere (19, 26). The genomic region spanning all of 3C and the 70 N-terminal amino acids of 3D (termed 3D*) was inserted into the *Bam*HI site of the expression vector pET-3a. To construct pP1-2ABC*, the cDNA plasmid pHAV/7, containing the complete nucleotide sequence of HAV strain HM175, was restricted at nucleotide (nt) 4962 (*Eco*RI), blunt ended, and reannealed. pT7-HAS15 contains the complete cDNA of HAV strain HAS-15, which was excised from plasmid pHAV53 and replaced in pGEM-7Zf(+) (26). pP1-2AB* was generated by restriction of pT7-HAS15 at nt 3954 with *Xba*I and blunt ending and reannealing of the cDNA. pET-ΔVP1-P2-P3* and the respective construct encoding the proteolytically inactive mutant (pET-ΔVP1-P2-P3*μ) were described before (26). pT7-ΔVP1-2ABC+ contains the HAV sequences from nt 626 to 736 and from nt 2801 to 4760 of HAV strain MBB (16). Δ indicates an N-terminally truncated protein, and * and + indicate C-terminally truncated proteins.

Other 3C substrates and reference proteins corresponding to intermediate or end products of polyprotein processing were obtained by PCR subcloning. pET-VP1-2A and pET-VP1-2AB were constructed by PCR amplification of nt 2202 to 3236 and nt 2202 to 3989, respectively, of HAV strain HM175. The primers used encoded the respective terminal nucleotide sequences, start or stop codons, and restriction sites for insertion of the PCR fragments into the expression vector pET-3 (24). cDNA clones encoding HAV proteins 2A, 2B, and 2C were obtained in the same manner by amplification of the genomic regions between nt 2964 and 3236 (2A), nt 3237 and 3989 (2B), and nt 3990 and 4994 (2C).

**Combined transcription and translation of cDNA in vitro.** Circular cDNA (0.5 to 1 μg) was used in a 15-μl reaction mixture containing TNT reticulocyte lysate supplemented according to the instructions of the manufacturer (Promega). Proteins were labeled for 90 min at 30°C with [$^{35}$S]methionine (>37 MBq/mmol; New England Nuclear), with 1.4 MBq/15-μl reaction mixture. Products of translation and processing were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) followed by fluorography.

**Bacterial expression and purification of recombinant 3C.** Bacterial expression of 3C with plasmid pET-3CD* has been described before (19, 26). Induced bacteria were harvested by centrifugation, and 3C was purified by ion-exchange chromatography according to the protocol of Malcolm et al. (21). Purified 3C was dissolved in a buffer containing 50 mM Tris-HCl (pH 8.5), 2.5 mM EDTA, and 2 mM dithiothreitol, and its final concentration was adjusted to 0.85 mg/ml. Long-term storage was at −70°C. For two-dimensional PAGE, isoelectric focusing was performed with nonequilibrated pH gel electrophoresis (NEPHGE) according to the method of Bravo and was followed by SDS-PAGE (4). Gels were either silver stained or analyzed by immunoblot with anti-3C (26).

**Assay of proteinase activity on polyproteins and synthetic peptides.** Purified proteinase 3C was incubated with radiolabeled translation products for 4 h at 37°C or for the time periods indicated. Processing reactions were stopped by the addition of gel sample buffer, and the products were analyzed by SDS-PAGE and fluorography. Immunoprecipitation of the processing products was performed as described before (16, 19, 26).

The synthesis of peptide 2A-2B (KGLFSQ*AKISLFYT, where * designates the putative cleavage junction between 2A and 2B) and the competition assay for the determination of the $(V_{max}/K_m)_{rel}$ value were performed as described elsewhere (27). As a reference peptide, Ac-ELRTQ*SFS-NH2 was used (kindly provided by B. A. Malcolm; see reference 21). The assay was performed in 175 μl containing 10 μM highly purified HAV 3C and 400 μM each peptide substrate in a buffer of 50 mM Tris-HCl (pH 8.0), 2.5 mM EDTA, and 2 mM dithiothreitol at 37°C. Aliquots were removed at 4, 8, 16, 32, 60, 180, and 360 min, and the reaction was stopped by transfer of 25-μl aliquots into UTRAFREE-MC tubes (nominal molecular weight limit, 10,000; Millipore) containing 200 μl of 0.5 M $HClO_4$. After incubation on ice for 15 min and centrifugation (30 min, 4°C, 14,000 rpm), the filtrate was directly used for reverse-phase high-pressure liquid chromatography analysis. The extent of substrate conversion was determined by analysis of the peak areas, and the data were processed as described elsewhere (27).

## RESULTS

**Large-scale expression and purification of HAV recombinant 3C proteinase.** To study the substrate specificity of the HAV 3C proteinase, the required amounts of enzyme were produced by recombinant expression in the pET system (24). Construct pET-3CD* codes for a fusion protein of 34 kDa which is autocatalytically cleaved to yield soluble proteinase of an apparent molecular mass of 26 kDa (19, 26). A 3-mg sample of purified 3C was obtained from one liter of induced bacteria (optical density at 600 nm, 1.2). On the basis of the estimated isoelectric point of 3C (pI = 8.5), a purification scheme with DEAE-Sephadex A-25 and carboxymethyl (CM)-Sepharose CL-6B was applied (21). Highly purified 3C was eluted from CM-Sepharose at approximately 500 mM NaCl and subsequently analyzed by two-dimensional electrophoresis. Because of the high pI value of 3C, separation by NEPHGE was chosen for the first dimension, followed by SDS-PAGE. As shown in Fig. 1, 3C was purified to near homogeneity, as judged from the absence of contaminating proteins, and could be separated into equal amounts of two isoforms migrating with slightly different pIs. Both isoforms were identified by immunoblot with an anti-3C serum (Fig. 1A). The presence of isoforms was observed before (1, 5).

**Proteolytic processing of P1-P2, P2, and P2-P3 precursor polypeptides mediated by purified HAV 3C proteinase.** To examine proteolytic processing at various sites within the HAV polyprotein, different parts of the polyprotein encoded by overlapping cDNA segments were used as substrates of purified 3C proteinase (Fig. 2). The respective polypeptides were generated in a coupled in vitro transcription/translation system. Parts of the HAV polyprotein were preferred over the complete precursor molecule (P1-P2-P3) because it was shown that in vitro translation of HAV RNA resulted in the formation of aberrant initiation products which obscured the pattern of cleavage products (15, 16). Translation products representing P1-2ABC*, P1-2AB*, ΔVP1-P2-P3*μ, and ΔVP1-2ABC+ migrated with apparent molecular masses of 156, 120, 124, and 71 kDa, respectively (indicated by dots [pre] in Fig. 3, lanes 1a, 2a, 3a, and 4a, respectively). Polypeptides smaller than the expected size originated from internal initiation of translation (15, 16). All translation products were incubated with purified 3C for 4 h at a final enzyme concentration of 200 μg/ml. Polypeptides derived by processing from precursors P1-2ABC* and P1-2AB* were almost identical and included VP3, VP0, two forms of VP1-2A, VP0-VP3, and P1-2A (indicated by open asterisks in Fig. 3, lanes 1b and 2b). In addition, a polyprotein of 63 kDa which represents 2BC* was observed among the processing products of P1-2ABC* (Fig. 3, lane 1b). In some experiments, 2C* and 2B were detectable among the products of P1-2ABC*, suggesting that cleavage within this polyprotein at the 2B/2C junction is less efficient than cleavage within the framework of the ΔVP1-P2-P3*μ polypeptide (Fig. 3, lane 3b). Small amounts of 2A produced by cleavage of P1-2ABC* and P1-2AB* could be detected in gels of higher acrylamide concentration (see Fig. 5A). Identification of products was achieved by immunoprecipitation (not shown) as well as comigration with polypeptides generated by in vitro expression of the respective PCR-amplified genomic regions (Fig. 3, VP1-2A [lane 5], 2B [lane 6] 2C [lane 7]; data for 2A not shown). This supports our earlier suggestion that VP1-2A and 2B are represented by amino acids 492 to 836 and 837 to 1087, respectively, of the HAV polyprotein, and their mappings therefore differ from those originally proposed (7). The origin of the two forms of VP1-2A is unclear. The patterns produced by cleavage of P1-2ABC* and P1-2AB* suggest that HAV 3C proteinase catalyzes cleavage of the HAV polyprotein at sites VP0/VP3, VP3/VP1, and 2A/2B and to a lesser extent also at the VP1/2A and 2B/2C junctions as observed before (26).

Cleavage of the proteolytically inactive substrate ΔVP1-P2-P3*μ was assessed in the same manner. As recently shown, processing products were identical to those generated by au-

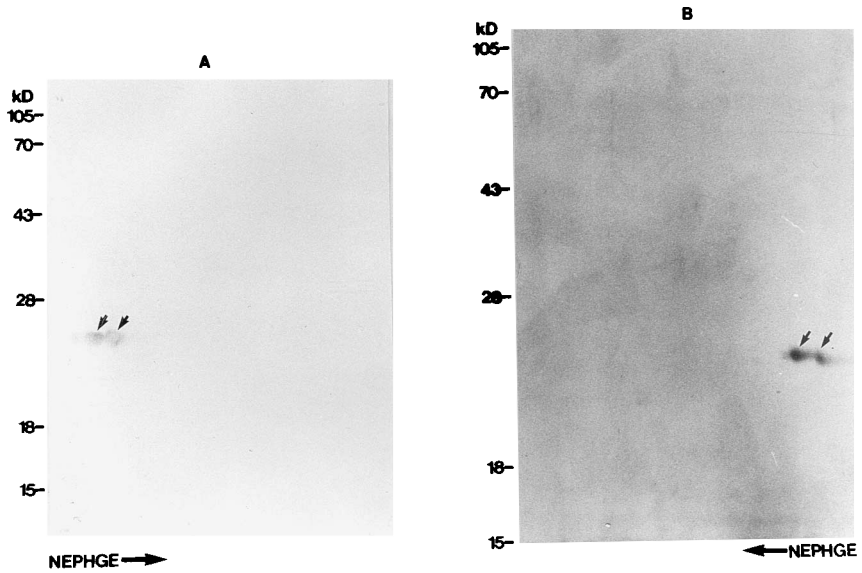


FIG. 1. Two-dimensional separation of purified HAV proteinase 3C analyzed by immunoblot (A) and silver staining (B). 3C purified by ion-exchange chromatography was separated by NEPHGE (first dimension) and SDS-PAGE (second dimension). Molecular mass standards are marked on the sides.

toproteolysis of the respective unmutated and proteolytically active polypeptide precursor ΔVP1-P2-P3* (2). Proteins 2BC, P3*μ, 2C, 3ABCμ, 2B, and ΔVP1-2A were the most prominent products (indicated by filled asterisks in Fig. 3, lane 3b). 2A, as a protein of approximately 10 kDa, could be resolved on a 15% gel (see Fig. 4). P3*μ and 3ABCμ were usually found as stable intermediates of intermolecular processing, and further cleavage of 3ABCμ was observed in only a few experiments (see Fig. 4B). This suggests that cleavage at the 3C/3D junction occurred in a part of the P3*μ polypeptides whereas sites within 3ABCμ were not efficiently processed in this system (13, 26). 2A, ΔVP1-2A, 2BC+, 2C+, and 2B were also produced by 3C-mediated cleavage of ΔVP1-2ABC+ (indicated by filled asterisks in Fig. 3, lane 4b; 2C+ and 2B migrate at the same position). The results with regard to intermolecular cleavage of the P2 and P2-P3 proteins (including the C terminus of VP1) confirm our earlier conclusions that sites 2A/2B and 2C/3A are

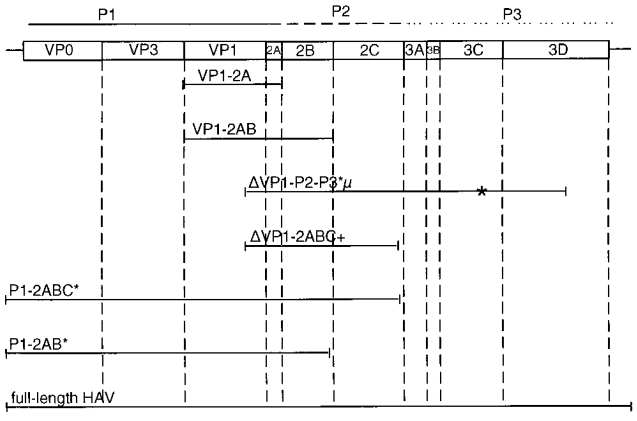


FIG. 2. Schematic representation of overlapping coding regions representing parts of the HAV genome. △ indicates an N-terminally truncated protein, and * and + indicate C-terminally truncated proteins; ✻ indicates a Cys residue of the active site mutated to Ala. None of the constructs encodes an active proteinase, and proteins derived from them can therefore serve as substrates for intermolecular cleavage mediated by proteinase 3C.

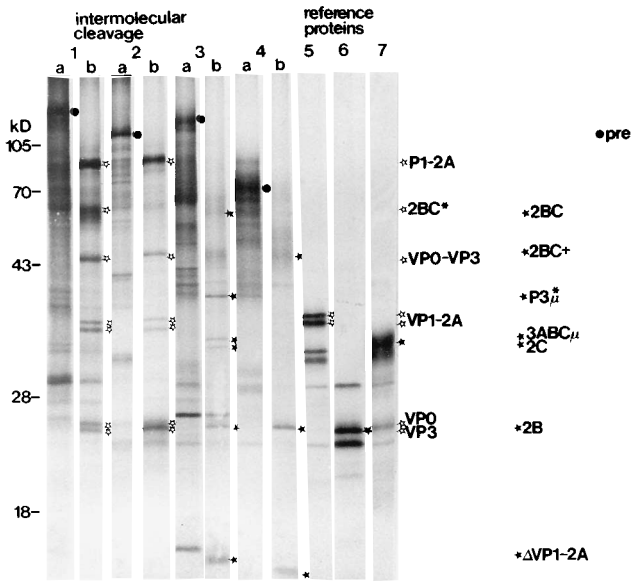


FIG. 3. Cell-free translation and proteolytic processing by purified recombinant 3C of P1-2ABC* (lanes 1), P1-2AB* (lanes 2), ΔVP1-P2-P3*μ (lanes 3), and ΔVP1-2ABC+ (lanes 4). cDNAs of HAV plasmids depicted in Fig. 2 were transcribed and translated in a cell-free system for 90 min to yield radiolabeled substrates (lanes a; pre [dot] marks uncleaved precursor polypeptides). For subsequent proteolytic cleavage, purified 3C was added at a final concentration of 200 μg/ml and incubation was continued for 4 h at 37°C (lanes b). Cleavage products of P1-2ABC* (lane 1b) and P1-2AB* (lane 2b) are marked by open asterisks, and those of ΔVP1-P2-P3*μ (lane 3b) and ΔVP1-2ABC+ (lane 4b) are marked by filled asterisks. In vitro translation of VP1-2A (lane 5), 2B (lane 6), and 2C (lane 7) is also shown. Molecular mass standards are marked on the left. Polypeptides 2C and 3ABC migrated in a reversed position compared with results shown in our earlier report (25). ΔVP1-2A, derived from ΔVP1-P2-P3*μ and ΔVP1-2ABC+, migrated with different mobility because of heterogeneous N termini.

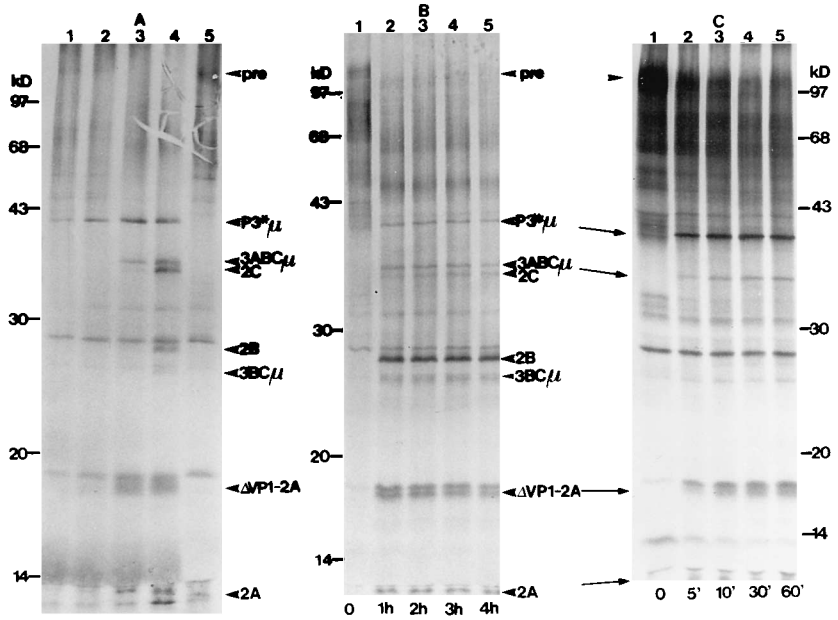


FIG. 4. Proteolytic processing of ΔVP1-P2-P3*μ with application of various concentrations of purified 3C (A) and with various incubation periods (B and C). (A) Radiolabeled ΔVP1-P2-P3*μ was incubated for 4 h at 37°C with amounts of purified proteinase 3C as follows: 0.2 μg/ml (lane 1), 2 μg/ml (lane 2), 20 μg/ml (lane 3), and 200 μg/ml (lane 4); lane 5, uncleaved substrate. (B and C) Proteinase 3C at 200 μg/ml (B) and 20 μg/ml (C) was used to cleave the same substrate for the time intervals indicated. Molecular mass standards and HAV proteins are marked.

completely cleaved by 3C proteinase whereas junction 2B/2C is processed only partially (26). Compared with the bacterial expression system, sites within 3ABC were processed very inefficiently in the cell-free system (10, 16, 26). Our data indicate that processing of all nine potential cleavage sites within the HAV polyprotein is catalyzed by 3C proteinase alone. However, the efficiency of cleavage varies for each site and depends on the expression systems used.

**Substrate specificity and cleavage efficiency of 3C.** Initial experimental evidence for the substrate specificity of 3C was provided by cleavage of synthetic peptides encompassing the potential cleavage sites in the HAV polyprotein (13). From this study, it was concluded that sites within P1 and P3 are cleaved by 3C less efficiently than those within P2, sites 2B/2C and 2C/3A being the best. To further elucidate the efficiency of 3C in catalyzing cleavage at the nine sites within the frame of the polyprotein, the influence of time and 3C concentration was studied. Preferential cleavage among the various sites was observed when increasing amounts of purified 3C were tested on the substrates described above. Figure 4A shows the cleavage pattern derived from substrate ΔVP1-P2-P3*μ following incubation with increasing concentrations of 3C. P3*μ was cleaved off the substrate at an enzyme concentration as low as 0.2 μg/ml (Fig. 4, lane 1). Ten times more enzyme was needed for the liberation of ΔVP1-2A and 2A (Fig. 4, lane 2). 3ABCμ appeared when the concentration of 3C was raised to 20 μg/ml (Fig. 4, lane 3). Products 2C and 2B were detectable when 3C was applied at a concentration of 200 μg/ml (Fig. 4, lane 4). These data suggest that sites 2C/3A and 2A/2B (at amino acid position 836/837 of the HAV polyprotein) are very efficiently cleaved. Both junctions may be considered as primary targets of 3C within the HAV polyprotein when the enzyme is added in *trans*. The order of subsequent cleavages was 3C/3D and 2B/2C, as implied by the amounts of enzyme required for their efficient cleavage.

To examine the kinetics of processing, ΔVP1-P2P3*μ was subjected to intermolecular cleavage by purified 3C (200 μg/ml, incubation periods ranging from 1 to 4 h; Fig. 4B). No hierarchic order of processing could be established by this approach since all polypeptides produced at this enzyme concentration appeared at the shortest time interval tested. Even 1/10 of the enzyme (20 μg/ml) and shorter incubation periods (5 to 60 min) applied to the same substrate indicated that all cleavages performed at a given enzyme concentration occurred very rapidly (within 5 min). Uncleaved intermediate products, however, were not processed any further even after prolonged incubation periods (Fig. 4C). Supplementation of *trans*-cleavage assays with S10 extracts of HAV- or mock-infected cells had no influence on the extent of cleavage within the P2 and P3 domains of the HAV polyprotein (23). Preincubation of the purified enzyme with reticulocyte lysate excluded the possibility that instability of the enzyme in the cell-free system is responsible for the lack of 3C activity after prolonged incubation periods. Furthermore, subsequent addition of alternate substrates to a cleavage assay mixture indicated that the enzyme is not inhibited by accumulated products (data not shown).

Processing of the P1-P2 domain of the HAV polyprotein was assessed in the same manner with P1-2ABC* as the substrate (Fig. 5A and B). The site most readily cleaved by 3C was the 2A/2B junction, liberating P1-2A and 2BC* (Fig. 5A, lane 3). Further cleavage within P1-2A to yield VP3-VP1-2A, p55 (VP0-VP3 and VP3-VP1), VP0, VP3, 2A, and two forms of VP1-2A was achieved by incubation with the 10-fold higher concentration of 3C (Fig. 5A, lane 4). These cleavages occurred only in a small portion of the precursor molecules, since the amount of P1-2A did not decrease significantly with increasing concentrations of 3C. By immunoprecipitation, further processing intermediates derived from cleavage of P1-2ABC* were identified. As shown in Fig. 5B, anti-VP1 reacted with P1-2A, P1, VP3-VP1-2A, VP3-VP1 (p55), and VP1-2A. The high-molecular-weight cleavage intermediates were also

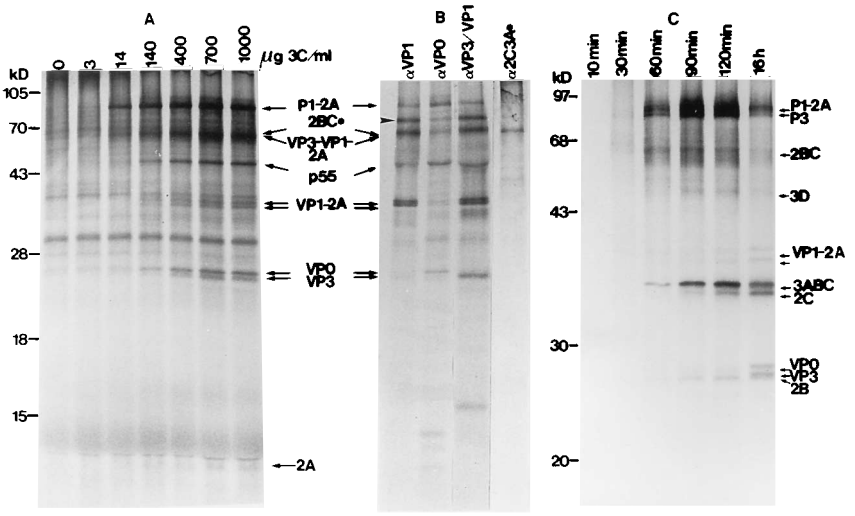


FIG. 5. (A) Polypeptides produced by intermolecular cleavage of P1-2ABC* mediated by various concentrations of purified 3C. (B) Identification of cleavage products by immunoprecipitation, with anti-VP1, anti-VP0, anti-VP3-VP1, and anti-2C3A*. The arrowhead marks P1. (C) Kinetics of cell-free translation and proteolytic processing of the complete HAV polyprotein derived from cDNA clone pT7-HAS15. Samples were analyzed by SDS-PAGE on a 15% gel and fluorography. Molecular mass standards are shown on the left. HAV proteins are indicated.

precipitated by anti-VP3-VP1 and anti-VP0 in addition to the respective end products. Anti-2C3A* recognized 2BC*, a polypeptide of 63 kDa which comigrates with VP3-VP1-2A. Essentially similar processing patterns were produced when increasing concentrations of 3C were applied on P1-2AB* and ΔVP1-2ABC* as substrates. However, no order of processing could be deduced from kinetic studies with P1-P2 precursors (data not shown).

To compare intra- and intermolecular processing of the HAV polyprotein, full-length HAV cDNA was transcribed and translated in the cell-free system with pT7-HAS15 as a template (Fig. 5C). After 60 min of translation, P1-2A, P3, 2BC, 3D, and 3ABC appeared as processing products followed by 2C and 2B (90 min). After 120 min, VP1-2A was detected, and after 16 h, VP0 and VP3 were detected. VP1 was not found. These data allow us to establish a hierarchy of processing which supports our above-mentioned findings on intermolecular processing, indicating that within the complete HAV polyprotein sites 2A/2B, 2C/3A, and 3C/3D are most readily cleaved. Processing of junctions VP0/VP3 and VP3/VP1 was catalyzed by 3C after long incubation periods and only to a small extent, whereas VP1-2A and 3ABC remained uncleaved under these conditions.

**Processing at the VP1/2A and 2A/2B cleavage sites.** Among the genera of the picornavirus family, the site of primary cleavage within the polyprotein differs. Primary cleavage in the entero- and rhinovirus polyprotein occurs at the C terminus of P1, whereas in the cardio- and aphthovirus group the primary cleavage product is P1-2A, resulting from cleavage at the 2A/2B junction (8, 12, 17). Not only sequence alignment but also our experimental data suggest that HAV is unique in employing only one proteinase for the cleavage of all sites in its primary translation product and that no proteolytic activity seems to be encoded in the P2 domain of the HAV polyprotein (7, 14, 16, 26).

Our results described above show that cleavage of the VP1/2A junction in polyproteins ΔVP1-P2-P3*µ and ΔVP1-2ABC+ was efficiently catalyzed by 3C. Cleavage at this site was less pronounced in substrates P1-2ABC* and P1-2AB*, implying that structural features of the precursor molecules might limit the accessibility of this junction. To test the conformational prerequisites required for cleavage at site VP1/2A, shorter constructs encoding only VP1-2A and VP1-2AB were expressed by cell-free translation and subjected to proteolysis. In both cases, the largest protein corresponds in size to the expected precursor (VP1-2A and VP1-2AB, respectively), and the polypeptides of smaller sizes most likely were synthesized from internal translation initiation sites which were mapped at the appropriate distances from the initiation codon (indicated by the suffix i in Fig. 6). Since the proportion of the higher- to the lower-molecular-weight proteins remained unchanged during time course experiments, it can be concluded that no proteolytic activity resides either in VP1-2A or in VP1-2AB (data not shown).

Processing at the putative primary cleavage sites (VP1/2A and 2A/2B) mediated by 3C was assessed by subjecting substrates VP1-2A and VP1-2AB to increasing amounts of purified 3C (Fig. 6). Additional protein bands were not observed when VP1-2A was incubated with 3C (Fig. 6, lanes 1 and 2). However, processing of VP1-2AB gave rise to 2B and polypeptides that comigrated with the two most prominent translation products of VP1-2A (Fig. 6, lanes 3 to 9). For complete cleavage of the 2A/2B junction, 140 µg of 3C per ml was necessary (Fig. 6, lane 7).

Cleavage at the 2A/2B junction was further assessed by subjecting the synthetic peptide KGLFSQ*AKISLFYT to purified 3C. The prediction of this putative 2A/2B cleavage signal is based on an alignment of all known 3C cleavage sites. Correct processing of the proposed junction between Gln and Ala in the 2A-2B peptide was verified by amino acid sequencing of the cleavage products (underlined sequence). Furthermore, in a competition assay with peptide Ac-ELRTQ*SFS-$NH_2$ (2B/2C junction) as a reference, the 2A-2B peptide was efficiently processed [$(V_{max}/K_m)_{rel}$ = 0.85 ± 0.14 (mean ± standard deviation)]. The results indicate that the proposed 2A/2B site serves as a substrate of 3C and its cleavage within the polyprotein liberates a 2B polypeptide of approximately 26 kDa. Our findings that processing of P1-2ABC* or the complete polyprotein did not yield mature VP1, but rather P1 and 2A or VP1-2A, are consistent with the lack of VP1-2A cleav-

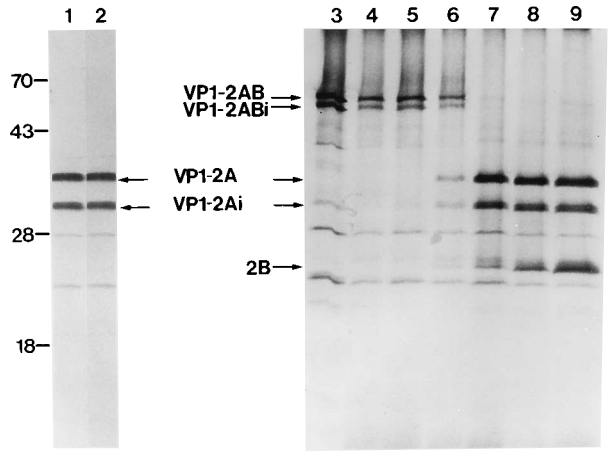


FIG. 6. Cell-free translation and intermolecular cleavage of VP1-2A (lanes 1 and 2) and VP1-2AB (lanes 3 to 9). Purified 3C was applied at a concentration of 400 μg/ml (lane 2) or in various concentrations to cleave VP1-2AB (uncleaved, lanes 1 and 3): 1.4 μg/ml (lane 4), 3 μg/ml (lane 5), 14 μg/ml (lane 6), 140 μg/ml (lane 7), 400 μg/ml (lane 8), and 700 μg/ml (lane 9). Samples were analyzed by SDS-PAGE on a 15% gel and fluorography. Molecular mass standards (in kilodaltons) are shown on the left. HAV proteins are indicated.

age. In combination, these results provide circumstantial evidence that P1-P2 processing at the VP3/VP1 site, and thus the liberation of the N terminus of VP1, renders the VP1/2A site uncleavable, resulting in the stable intermediate VP1-2A, which has been observed as a precursor of VP1 in HAV-infected cell culture (2, 6, 18).

## DISCUSSION

Recently, we and others provided evidence that the activity profile of HAV 3C proteinase differs from that of other picornaviruses in catalyzing cleavage of the primary as well as the secondary sites in the HAV polyprotein and that it is the only virus-encoded proteinase of HAV (15, 16, 26). In these reports, the specific activity of 3C proteinase could not be assessed in detail since either crude bacterial extracts or the self-cleaving system containing precursors of 3C was used. In other studies, purified recombinant proteinase was tested for its activity towards synthetic peptides and polypeptides of the P1 domain of the HAV polyprotein (13, 21). Utilizing a purified enzyme for intermolecular cleavage assays, we now show that proteinase 3C alone is indeed sufficient to cleave all junctions in the HAV polyprotein. The extent of intermolecular cleavage at different sites within the HAV polyprotein is dependent on the concentration of 3C applied. Junctions P1-2A/2B and P2/P3 were most efficiently cleaved. Cleavages mediated by 3C within the P1 and P3 domains were performed to a lesser extent. Differences in affinities to various cleavage sites might be responsible for the regulated liberation of the viral structural and nonstructural proteins by 3C. As shown earlier, the range of products obtained by intermolecular cleavage of the proteolytically inactive precursor polypeptide ΔVP1-P2-P3*μ was identical to the proteins found after intramolecular processing of the respective ΔVP1-P2-P3* polypeptide and comparable with processing products of the complete HAV polyprotein (26). In addition, processing products VP0, VP3, and VP1-2A correspond to the respective polypeptides detected in HAV-infected cells (data not shown here; see references 2, 6, 18, and 29).

The specificity of 3C was assessed by identification of the processing products employing immunoprecipitation and comigration with recombinant reference proteins. In all instances, cleavage at the nine putative sites could be shown, although to various extents. Processing product 2B was particularly interesting, since it migrated at a mobility different from what has been proposed earlier (7). Comigration with a recombinant protein expressed from the PCR-amplified genomic region between nt 3237 and 3989 suggested that the N terminus of 2B comprises amino acids AKIS and a following chain and is located further upstream in the HAV polyprotein. This observation is in line with the cleavage of a synthetic oligopeptide representing the respective junction. The relative cleavage efficiency of the 2A-2B peptide used here was higher than that determined for a peptide site proposed earlier (7, 13). Comigration with the PCR-derived translation product strongly suggested that VP1-2A comprises amino acids 492 to 836 of the HAV polyprotein. Since a protein of similar electrophoretic mobility (38 kDa) was observed in HAV-infected cells, it can be assumed that processing of in vitro-synthetized substrates by purified 3C truly reflects the processing pattern in vivo (2, 6, 18). The relative stability of VP1-2A in vivo and in vitro may argue for a function of VP1-2A different from that of VP1.

Among the products derived by inter- and intramolecular processing of the in vitro-produced P3 domain, 3ABC was a stable intermediate, implying that sites 3A/3B and 3B/3C are not efficiently cleaved when produced in the cell-free system. It is interesting to note that 3BC and 3C are generated when this region is expressed in vivo (10, 11, 26, 28). Most probably, either the reticulocyte lysate system is devoid of factors necessary to cleave 3ABC or proteins produced in the cell-free system are subject to misfolding (3).

The relative cleavage efficiency of 3C at the various sites was evaluated by comparison of the product patterns formed when the various domains of the HAV polyprotein were exposed to increasing concentrations of the purified enzyme. The amounts of intermediate and end products varied significantly depending on the enzyme concentration used and allowed ranking of the cleavage sites. This was most clearly observed in the liberation of P3*μ, which was detected at the lowest 3C concentration tested (12 nM), from the ΔVP1-P2-P3*μ precursor whereas, for instance, 2B appeared at an enzyme concentration as high as 12 μM. The apparently very efficient intermolecular cleavage at site P2/P3 is in accordance with data obtained in translation kinetics of the respective autoproteolytically active protein, in which P3 was the first processing intermediate to be detected (11, 14, 16, 26). Additionally, our finding is consistent with cleavage experiments with a synthetic peptide which represents the 2C/3A junction (13).

In addition to the respective intermediate products, cleavage of the P1-P2 domain and parts of it resulted in the expected end products VP0, VP3, VP1-2A, 2A, 2B, and 2C, indicating that all sites of this region of the polyprotein are recognized as substrates of 3C. In contrast to another report, mature VP1, a protein of approximately 33 kDa, was not observed as processing product derived from either P1-P2, P1-2AB, VP1-2A, or VP1-2AB (21). The fact that we had obtained clear evidence for cleavage of the VP1/2A site by using substrate ΔVP1-P2-P3*μ suggests that the N-terminal region of VP1 plays an important role in the regulation of processing at the VP1/2A site (26). It can be concluded that processing at both termini of VP1 is mutually exclusive in the in vitro system, resulting in either P1 and 2A or VP1-2A. Obviously, P1 is not further processed under the experimental conditions used. Since both VP1 and VP1-2A are produced under in vivo conditions, we assume that host cell factors are involved in the liberation of mature VP1 (2, 3, 18).

When full-length RNA was translated and cotranslationally

cleaved by 3C, evidence that processing occurred in an ordered manner was obtained. As P1-2A, P3, and 2BC were the first products visible, it can be concluded that junctions 2A/2B and 2C/3A are the preferred sites for cleavage. Longer periods were required to separate 2B and 2C and ultimately to process P1-2A. These data imply that junctions 2A/2B and 2C/3A are the sites of primary cleavage which are mediated by the only virus-encoded proteinase 3C that subsequently catalyzes all secondary cleavages also. Whether these cleavages occur by a mono- or bimolecular reaction mechanism and whether precursors of 3C are involved in these reactions cannot be proven by translation and processing of the complete polyprotein.

Our attempts to rank the susceptibilities to intermolecular cleavage of the various sites within the HAV polyprotein by time course studies failed, since even short time intervals could not resolve different processing patterns. Although no time dependence could be established, a pronounced effect of enzyme concentration was observed for all substrates tested. These data support the order found for autocatalytic processing of the polyprotein in vitro. Partition of the primary translation product between domains P2 and P3 was equally as efficient as cleavage between P1-2A and 2BC. This is manifested by the liberation of P1-2A from P1-2ABC* and P1-2AB* at an enzyme concentration similar to that necessary to cleave site 2C/3A within polyprotein ΔVP1-P2-P3* as well as the respective synthetic peptide. Our evidence that intermolecular processing of the HAV polyprotein depends on the proteinase concentration rather than on the incubation period favors the idea that the course of events in the infectious cycle may be controlled by the amount of 3C liberated from the polyprotein. While the initial production of nonstructural proteins required for RNA replication might be effected by only small quantities of proteinase, later processing events liberating the structural proteins possibly depend on larger concentrations of 3C.

## ACKNOWLEDGMENTS

Plasmid pHAV53, containing the full-length cDNA of HAV strain HAS-15, was a generous gift of E. Snezhkov, Institute of Bioorganic Chemistry, Moscow, Russia.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG SFB 367, Project B7). T.S. was a recipient of a fellowship granted by the state of Schleswig-Holstein.

## REFERENCES

1. **Allaire, M., M. M. Chernaia, B. A. Malcolm, and M. N. G. James.** 1994. Picornaviral 3C cysteine proteinases have a fold similar to chymotrypsin-like serine proteinases. Nature (London) **369:**72–75.
2. **Anderson, D. A., and B. C. Ross.** 1990. Morphogenesis of hepatitis A virus: isolation and characterization of subviral particles. J. Virol. **64:**5284–5289.
3. **Blair, W. S., X. Li, and B. L. Semler.** 1993. A cellular factor facilitates efficient 3CD cleavage of the poliovirus P1 precursor. J. Virol. **67:**2336–2343.
4. **Bravo, R.** 1984. Two-dimensional gel electrophoresis: a guide for the beginner, p. 4–34. *In* J. E. Celis and R. Bravo (ed.), Two-dimensional gel electrophoresis of proteins, 1st ed. Academic Press, New York.
5. **Chernaia, M. M., B. A. Malcolm, M. Allaire, and M. N. G. James.** 1993. Hepatitis A virus 3C proteinase: some properties, crystallization and preliminary crystallographic characterization. J. Mol. Biol. **234:**890–893.
6. **Cho, M. W., and E. Ehrenfeld.** 1991. Rapid completion of the replication cycle of hepatitis A virus subsequent to reversal of guanidine inhibition. Virology **190:**770–780.
7. **Cohen, J. I., J. R. Ticehurst, R. H. Purcell, A. Buckler-White, and R. M. Baroudy.** 1987. Complete nucleotide sequences of hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses. J. Virol. **61:**50–59.
8. **Dougherty, W. G., and B. L. Semler.** 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. Microbiol. Rev. **57:**781–822.
9. **Francki, R. I. B., C. M. Fauquet, D. L. Knudson, and F. Brown.** 1991. Classification and nomenclature of viruses. Arch. Virol. **2**(Suppl.)**:**320–326.
10. **Gauss-Müller, V., D. Jürgensen, and R. Deutzmann.** 1991. Autoproteolytic cleavage of recombinant 3C proteinase of hepatitis A virus. Virology **182:** 861–864.
11. **Harmon, S. A., W. Updike, X.-Y. Jia, D. F. Summers, and E. Ehrenfeld.** 1992. Polyprotein processing in *cis* and in *trans* by hepatitis A virus 3C protease cloned and expressed in *Escherichia coli*. J. Virol. **66:**5242–5247.
12. **Jackson, R. J.** 1989. Comparison of encephalomyocarditis virus and poliovirus with respect to translation initiation and processing in vitro, p. 51–71. *In* B. L. Semler and E. Ehrenfeld (ed.), Molecular aspects of picornavirus infection and detection. American Society for Microbiology, Washington, D.C.
13. **Jewell, D. A., W. Swietnicki, B. M. Dunn, and B. A. Malcolm.** 1992. Hepatitis A virus proteinase 3C substrate specificity. Biochemistry **31:**7862–7869.
14. **Jia, X.-Y., E. Ehrenfeld, and D. F. Summers.** 1991. Proteolytic activity of hepatitis A virus 3C protein. J. Virol. **65:**2595–2600.
15. **Jia, X.-Y., G. Scheper, D. Brown, W. Updike, S. Harmon, O. Richards, D. F. Summers, and E. Ehrenfeld.** 1991. Translation of hepatitis A virus RNA *in vitro* aberrant internal initiations influenced by the 5′ noncoding region. Virology **182:**712–722.
16. **Jürgensen, D., Y. Y. Kusov, M. Fäcke, H.-G. Kräußlich, and V. Gauss-Müller.** 1993. Cell-free translation and proteolytic processing of the hepatitis A virus polyprotein. J. Gen. Virol. **74:**677–683.
17. **Kräusslich, H.-G., and E. Wimmer.** 1989. Viral proteinases. Annu. Rev. Biochem. **57:**701–754.
18. **Kusov, Y. Y., Y. A. Kazachkov, G. K. Dzagurov, G. A. Khozinkskaya, M. S. Balayan, and V. Gauss-Müller.** 1992. Identification of precursors of structural proteins VP1 and VP2 of hepatitis A virus. J. Med. Virol. **37:**220–237.
19. **Kusov, Y. Y., W. Sommergruber, M. Schreiber, and V. Gauss-Müller.** 1992. Intermolecular cleavage of hepatitis A virus (HAV) precursor protein P1-P2 by recombinant HAV proteinase 3C. J. Virol. **66:**6794–6796.
20. **Liebig, H.-D., E. Ziegler, R. Yan, K. Hartmuth, H. Klump, H. Kowalski, D. Blaas, W. Sommergruber, L. Frasel, B. Lamphear, R. Rhoads, E. Kuechler, and T. Skern.** 1993. Purification of two picornaviral 2A proteinases: interaction with eIF-4γ and influence on *in vitro* translation. Biochemistry **32:** 7581–7588.
21. **Malcolm, B. A., M. C. Sherman, D. A. Jewell, J. R. Stratton-Thomas, K. B. Thudium, R. Ralston, and S. Rosenberg.** 1992. Expression and characterization of recombinant hepatitis A virus 3C proteinase. Biochemistry **31:**3358–3363.
22. **Nicklin, M. J. H., K. S. Harris, P. V. Pallai, and E. Wimmer.** 1988. Poliovirus proteinase 3C: large-scale expression, purification, and specific cleavage activity on natural and synthetic substrates in vitro. J. Virol. **62:**4586–4593.
23. **Probst, C., and V. Gauss-Müller.** Unpublished data.
24. **Rosenberg, A. H., B. N. Lade, D. Chui, S. Lin, J. J. Dunn, and F. W. Studier.** 1987. Vectors for selective expression of cloned DNAs by T7 RNA polymerase. Gene **56:**125–135.
25. **Ryan, M. D., and J. Drew.** 1994. Foot-and mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. EMBO J. **13:**928–933.
26. **Schultheiss, T., Y. Y. Kusov, and V. Gauss-Müller.** 1994. Proteinase 3C of hepatitis A virus (HAV) cleaves the HAV polyprotein P2-P3 at all sites including VP1/2A and 2A/2B. Virology **198:**275–281.
27. **Sommergruber, W., H. Ahorn, A. Zoephel, I. Maurer-Fogy, F. Fessl, G. Schnorrenberg, H.-D. Liebig, D. Blaas, E. Kuechler, and T. Skern.** 1992. Cleavage specificity on synthetic peptide substrates of human rhinovirus 2 proteinase 2A. J. Biol. Chem. **267:**22639–22644.
28. **Tesar, M., I. Pak, X.-Y. Jia, O. C. Richards, D. F. Summers, and E. Ehrenfeld.** 1994. Expression of hepatitis A virus precursor protein P3 *in vivo* and *in vitro*: polyprotein processing of the 3CD cleavage site. Virology **198:**524–533.
29. **Updike, W. S., M. Tesar, and E. Ehrenfeld.** 1991. Detection of hepatitis A virus proteins in infected BS-C-1 cells. Virology **185:**411–418.