JOURNAL OF VIROLOGY, Dec. 1989, p. 5037–5045
0022-538X/89/125037-09$02.00/0

Vol. 63, No. 12

Cleavage of Small Peptides In Vitro by Human Rhinovirus 14 3C Protease Expressed in *Escherichia coli*

MICHAEL G. CORDINGLEY,* R. BRUCE REGISTER, PIA L. CALLAHAN, VICTOR M. GARSKY, AND RICHARD J. COLONNO

Department of Virus and Cell Biology, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486

Received 27 June 1989/Accepted 14 August 1989

The 3C region of human rhinovirus 14 was expressed in *Escherichia coli*. The microbially synthesized protease was functional, since the expressed precursor underwent autoproteolytic processing to generate mature molecules of the expected molecular weight and antigenicity. Mutation of the putative active-site Cys-146 residue to an alanine resulted in the synthesis of unprocessed precursor molecules. Large quantities of the 20-kilodalton protease were purified by a simple purification protocol, and the resulting molecule was shown to be biologically active in vitro against synthetic peptides corresponding to the 2C-3A cleavage site. This site was cleaved with high efficiency and fidelity and was used to generate kinetic data on the 3C protease. The protease exhibited sensitivity to Zn^{2+}, was capable of cleaving five of seven rhinovirus cleavage site peptides tested with variable efficiency, and could distinguish authentic substrate peptides from control peptides containing the dipeptide cleavage sequence pair Gln-Gly.

Human rhinoviruses (HRVs) are members of the picornavirus family and are the major causative agents of the common cold in humans (9). Other members of the picornavirus family include the enteroviruses (polioviruses, coxsackieviruses, echoviruses, and hepatitis A virus), aphthoviruses (foot-and-mouth disease virus), and cardioviruses (mengo virus and encaphalomyocarditis virus). In common with all picornaviruses, HRVs translate their RNA genome directly into a large viral polyprotein precursor which must undergo a series of controlled proteolytic cleavages to generate the functional viral gene products (22). Indeed, it is now believed that all the proteolytic enzymes involved in the replication of picornaviruses are encoded in the viral genome. The existence of such virus-specific proteases with no known cellular homologs suggests that these enzymes may be attractive targets for antiviral chemotherapeutic agents, since the identification of 100 distinct HRV serotypes (10) precludes the production of an effective vaccine capable of neutralizing all serotypes.

The genome of HRVs is a positive-stranded RNA molecule of approximately 7,200 nucleotides that encodes a single long open reading frame encoding a viral polyprotein of 243 kilodaltons (kDa) (4, 28). Based on sequence alignments, the organization of the viral proteins within the polyprotein appears to be very similar to that of other picornaviruses (4, 28), and proteolytic cleavage is predicted to occur at sites analogous to those previously demonstrated for poliovirus (23). Studies with poliovirus have demonstrated that the initial proteolytic cleavage of the polyprotein is performed by the viral protein 2A, which cleaves the structural gene precursor, P1, from the nonstructural gene precursor P2-P3 region (3, 31). A second proteolytic activity has been associated with the 3C region of the poliovirus genome and is responsible for all but one of the remaining cleavages of the viral precursor (11). These cleavages were found to occur predominantly between the amino acids glutamine and glycine (21). The requirement of this particular dipeptide sequence is not as stringent for rhinoviruses, since variant cleavage sites have been predicted (4, 28).

The final cleavage event in the processing of the viral polyprotein precursor is the maturation cleavage, which occurs during encapsulation of the viral RNA and results in the cleavage of VP0 into the structural proteins VP4 and VP2. This cleavage is unique in that no extrinsic protease molecule appears to be involved. Instead, a model has been previously proposed that cleavage is catalyzed by a serine protease-like mechanism in which the genomic RNA in close proximity with the maturation cleavage site acts as a base in concert with a nucleophyllic serine residue in the amino terminal portion of VP2 (1).

The product of the 3C region of poliovirus was first conclusively identified as a viral protease by Hanecak et al. (11), who showed that expression of a segment of poliovirus cDNA containing 3C resulted in release of the 3C species by a protease whose active site resided within the 3C region. Sequence comparison of the 3C regions from related viruses indicated that a highly conserved domain containing a cysteine residue may be the active center of the enzyme, and the sensitivity of viral polyprotein processing to inhibitors such as iodoacetamide and *N*-ethylmaleimide suggests that this is a member of the class of cysteine proteases (8, 29). More recent analysis has suggested that although picornavirus proteases are functionally similar to thio proteases, they are structurally related to the trypsinlike family of serine proteases (2).

To gain more information about the 3C protease of HRVs, we have expressed a cDNA containing the 3C region of HRV 14 (HRV-14) in *Escherichia coli*. The expressed protease was purified and shown to specifically cleave small peptides containing authentic cleavage sequences. The substrate specificity in vitro and its relevance to viral polyprotein cleavage in vivo are discussed.

MATERIALS AND METHODS

Bacterial strains and plasmid constructions. The polymerase chain reaction (PCR) was used to generate selected

* Corresponding author.

segments of the HRV-14 cDNA for cloning into the *E. coli* expression vector pKK233-2 (Pharmacia, Inc.). Template DNA for all PCRs was plasmid SP64 (RV-14) (18), a complete cDNA clone of HRV-14 which was cut with *Pst*I to release the insert cDNA. The oligonucleotide primers used for PCR amplification of cDNA were designed to incorporate an in-frame ATG and *Nco*I restriction site at the 5′ end of the cDNA segment and to incorporate a termination codon (TAA) and *Hin*dIII site at the 3′ end of the amplified sequence. These sequence changes were introduced by incorporation of minimal mismatches into primers of 20 to 25 bases (see below). The amplified HRV cDNA sequences corresponded to nucleotides 5240 to 5786 (pR202), 5090 to 5935 (pR207), and 5090 to 7170 (pR209). PCRs were carried out by using the GeneAmp DNA amplification kit on a Perkin Elmer Cetus DNA Thermal Cycler. A cycle included a 1-min 94°C denaturation step, a 2-min 37°C annealing step, and a 5-min 72°C extension. Thirty cycles were used. Resulting PCR fragments were digested with *Nco*I and *Hin*dIII, electrophoresed on a 1% agarose gel, and purified with Geneclean (BIO 101). Purified fragments were then ligated into the *E. coli* expression vector pKK322-2 (Pharmacia) that had been similarly digested and purified. Ligated DNA was transformed into competent *E. coli* JM109 (Stratagene, Inc.). All restriction enzymes and ligase were purchased from New England BioLabs, Inc.

Oligonucleotide primers. Deoxyoligonucleotides used in the constructions described above include the 5′ and 3′ oligonucleotides CCA CCA TGG TTA TTA ATA CTA TTC and CAA ATT AAG CTT CTA TAT TGT TAA CC, respectively, for pR209; the 5′ and 3′ oligonucleotides GTT GCC ATG GGA CCA AAC ACA GAA TTT G and CTA AAG CTT AGC CTT GTT TCT CTA C, respectively, for pR202; and the 5′ and 3′ oligonucleotides CCA CCA TGG TTA TTA ATA CTA TTC and CAA TTA AGC TTA CAG TCT GGG ATC, respectively, for pR207.

Analysis of bacterial proteins. Bacteria were routinely grown by inoculation of small cultures from colonies maintained on minimal medium agar plates containing 100 μg of ampicillin per ml to ensure the maintenance of the F′ episome of the host *E. coli* JM109. Cultures were induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG) in mid-log phase and grown for 3 to 10 h before being harvested for analysis of bacterial proteins. Samples were prepared in Laemmli sample buffer (15) and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) with 15% polyacrylamide gels. The gels were subjected to Western immunoblotting (30) with a rabbit 3C antiserum generated against a synthetic peptide corresponding to amino acid residues 53 to 69 of the HRV-14 3C protease. The gels were stained with Coomassie brilliant blue or with a silver staining kit available from Separation Sciences.

Pulse-chase analysis of bacterial proteins. Bacterial cultures (20 ml) were grown in M9 salts containing ampicillin until the optical density at 600 nm reached 0.1 and then induced by the addition of 1 mM IPTG. Following incubation for 1 h, [^{35}S]methionine (20 μCi/ml) was added for 1 min, and then 1/2 culture volume of warm chase medium (M9 salts containing 0.4 mg of methionine per liter) was added and the mixture was incubated for an additional 90 min. At the designated times, samples (4 ml) were diluted twofold with stop buffer (40 mM sodium phosphate [pH 7.0], 2 mg of methionine per ml, 2% formaldehyde) on ice. Bacteria were collected by centrifugation and washed twice with 1 ml of cold phosphate-buffered saline prior to preparation of labeled proteins. The pellet was suspended in 50 μl of Laemmli sample buffer, heated for 5 min at 100°C, and centrifuged for 5 min in a microcentrifuge. The supernatant was diluted 10-fold with RIPA buffer (10 mM sodium phosphate [pH 7.6], 100 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.5% sodium deoxycholate, 0.1% SDS) prior to immune precipitation with anti-3C rabbit serum and analysis by SDS-PAGE.

Site-directed mutagenesis of HRV 3C. Oligonucleotide-directed mutagenesis was carried out essentially as described by Gibbs et al. (7), except that the template for mutagenesis was single-stranded phagemid DNA. An *Eco*RI-*Hin*dIII fragment of pR207, containing the *trp-lac* promoter and HRV sequences, was ligated into the pGEM3Zf– vector (Promega Biotec) to generate pR226, which was transformed into *E. coli* DH5αF′ (Bethesda Research Laboratories, Inc.). Single-stranded mutagenesis template was rescued from this strain by superinfection with helper M13 strain R408 (Stratagene) and prepared by standard methods. Expression of protein from wild-type pR226 (226–Cys-146) and the mutant pR226 (226–Ala-146) was assayed directly in *E. coli* DH5αF′ as described above.

Purification of HRV-14 3C. A 0.2-liter culture of strain 207 was grown in M9 salts plus ampicillin supplemented with glucose (0.2%) and thiamine (1 μg/ml) to an optical density at 600 nm of 0.45. This culture was used to inoculate six 0.5-liter cultures of Luria broth containing 100 μg of ampicillin per ml and 1 mM IPTG; the cultures were shaken vigorously until they reached an optical density at 600 nm of 1.5. The resulting 20 g of bacterial cell paste was stored at −80°C. The purification protocol for HRV 3C protease was essentially as described for that of poliovirus 3C (19). Briefly, the frozen cell pellet was suspended in 60 ml of 40 mM Tris hydrochloride (pH 7.9)–0.1 M NaCl and the suspension was passed three times through a Stanstead Press. The 75-ml lysate (containing 1.7 g of protein) was diluted with an equal volume of 20 mM Tris hydrochloride (pH 7.9)–10 mM $MgCl_2$–0.1 M KCl–5 mM dithiothreitol, 1 mM EDTA and centrifuged for 2.5 h at 55,000 rpm in a Beckman 60 Ti rotor, the supernatant (800 mg of protein in 140 ml) was diluted to 0.4 liter with distilled water, and the pH was adjusted to 8.3 with 1 M Trizma base. This material was mixed with a 0.2-liter slurry of DEAE-Sephacel preequilibrated with 40 mM Tris hydrochloride (pH 7.9)–1 mM dithiothreitol and stirred for 15 min. Unbound material was subsequently recovered in the supernatant by filtering the slurry through a sintered glass funnel. The ion-exchange gel was suspended once in 0.2 liter of preequilibration buffer, and this second supernatant was pooled with the first. Then 4.7 ml of 0.5 M EDTA and 9.3 ml of 0.5 M morpholinoethanesulfonic acid (MES) were added to the unbound material (0.6 liter) and ammonium sulfate precipitated (0.55 g/ml). Precipitation proceeded for 14 h at 4°C, and the insoluble material was collected by centrifugation at 8,000 × *g* for 30 min. The precipitate was suspended in 30 ml of 20 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES) (pH 7.9)-NaOH–100 mM NaCl–1 mM EDTA–1 mM dithiothreitol and dialyzed against three changes of 50 to 100 volumes of the same buffer, centrifuged to remove insoluble material, and frozen in small aliquots at −80°C. This material, which contained 30 mg of protein and was nominally 90% pure 3C protease as determined by SDS-PAGE and silver staining, was used in all protease assays. A parallel purification was carried out to prepare an equivalent protein fraction from a control bacterial strain.

Peptide synthesis. The peptides were prepared by solid-phase synthesis (17), using a double-coupling protocol for the introduction of all amino acids on the model 430A ABI

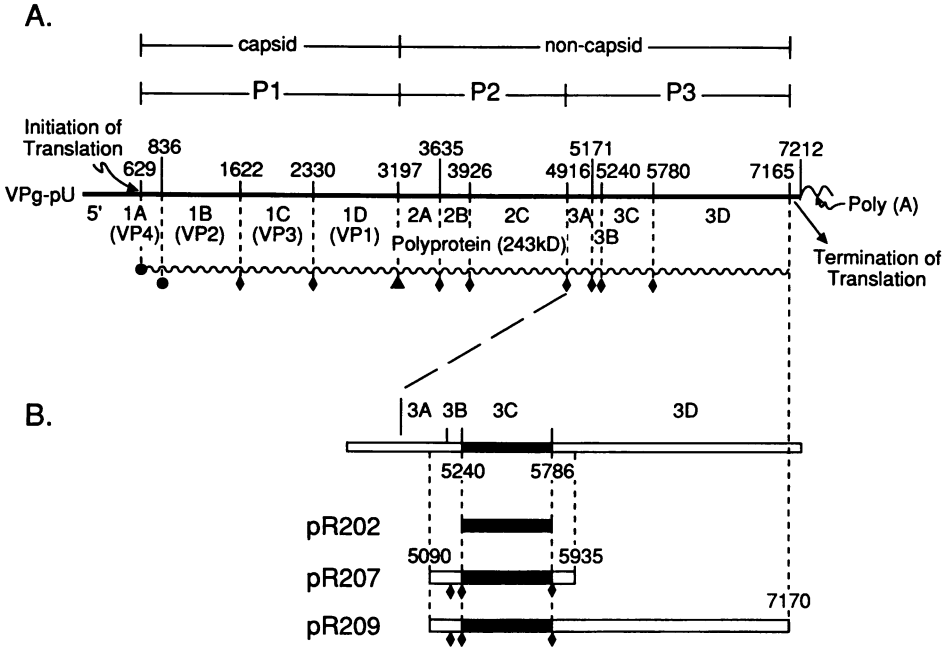

FIG. 1. Cloning of partial HRV cDNAs in *E. coli*. (A) Organization of the HRV-14 genome. The positive single-stranded RNA genome of HRV-14 is shown (—), with the locations of viral products indicated by nucleotide position; the colinear polyprotein is also shown (~~~). Cleavage sites for protease 2A (▲) and for protease 3C (◆) are indicated. (B) Structure of HRV-14 cDNA in the *E. coli* expression vector pKK233-2. Numbers represent nucleotide positions on the HRV-14 genome. Each DNA fragment has an additional methionine codon inserted immediately upstream of the 5′ nucleotide and a TAA termination codon immediately downstream of the inserted segment. Bacteria containing the plasmids pR202, pR207, and pR209 are designated strains 202, 207, and 209, respectively. The strain containing pKK233-2 is designated the control strain.

automated peptide synthesizer. Deprotection and removal of the peptide from the resin support were effected by treatment with liquid HF (27). The peptide was purified by preparative high-pressure liquid chromatography (HPLC) on a reverse-phase C_{18} silica column (25 cm by 6.4 cm) with a 0 to 50% gradient of acetonitrile containing 0.1% trifluoroacetic acid (26). Homogeneity was demonstrated by analytical HPLC, and identity was confirmed by amino acid analysis.

Peptide cleavage assays. Except when noted, all cleavage assay mixtures contained (in 0.1 ml) 40 mM Tris (pH 7.2), 1 mM dithiothreitol, 0.1% Nonidet P-40 at 30°C, 100 μM substrate peptide, and 1 μg of protease 3C per ml. Reactions were terminated by adding an equal volume of 20% acetic acid and centrifuging the mixture for 5 min in a microcentrifuge to remove insoluble material prior to analysis by reverse-phase HPLC on a 10-cm Vydac C_{18} column. Cleavage products were eluted with a 30-min linear gradient of 0 to 95% acetonitrile in 0.1% trifluoroacetic acid. A_{210} was measured, and data were recorded and analyzed by using a Waters Photodiode Array Detector 990 and Waters data analysis software. Rates of peptide hydrolysis were determined under conditions of constant substrate concentration, and the initial accumulation of cleavage products was measured before 10% reduction in substrate concentration was observed. Determinations were carried out in duplicate and often at two time points (10 and 20 min) to ensure linear accumulation of product peptides over time. The absorbance of product peptides was determined empirically for the 2C-3A peptide by measurement of substrate depletion and product accumulation in a cleavage reaction in which greater than 50% cleavage had taken place. Generally, it was found that the A_{210} of substrate was approximately equal to the sum of A_{210} of the products for each peptide used in these experiments.

RESULTS

Expression of HRV-14 3C protein. The predicted map locations and derivation of each of the HRV-14 proteins were previously determined by sequence homology to poliovirus (4) and are shown in Fig. 1A. A series of DNA fragments generated by PCR from a full-length cDNA clone of HRV-14 were expressed in *E. coli* by creation of in-frame initiation and termination codons at positions flanking the inserted region of the viral polyprotein. In these vectors, expression is driven from the IPTG-inducible *trp-lac* promoter and the expected primary translation product is defined by the region of the viral precursor inserted plus an additional N-terminal methionine residue. Figure 1B summarizes the structures of the three vectors constructed: pR202 contains the 3C region alone (nucleotides 5240 to 5786) and is expected to encode a 20-kDa protein, whereas pR207 and pR209 encode larger portions of the polyprotein precursor spanning part of regions 3A-3B-3C-3D and have reading frames encoding expected products of 38 and 86 kDa, respectively. Bacteria containing plasmids pR202, pR207, and pR209 are designated strains 202, 207, and 209, respectively.

Induced cultures of bacterial strains carrying each construct were harvested, and total cell lysates were examined for expression of the HRV proteins by SDS-PAGE and Western blotting analysis with an antiserum generated against a 3C-specific peptide. As expected, a prominent 20-kDa species was evident in extracts of strain 202 (Fig. 2A, lane 2) but not in extracts from the control strain, which

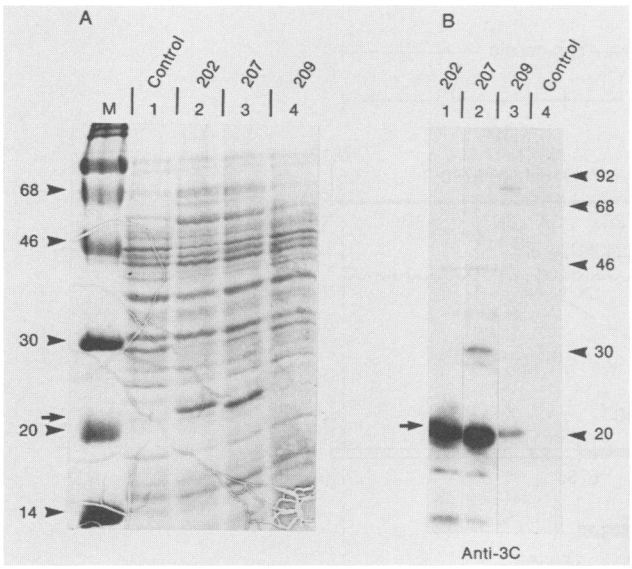

FIG. 2. SDS-PAGE analysis of expression of HRV-14 3C protease in *E. coli*. Bacterial strains which express cDNAs containing the HRV-14 3C region and a control strain containing the parental expression vector pKK233-2 were induced with IPTG and incubated for 10 h, after which total cell extracts were prepared and analyzed on an SDS–15% polyacrylamide gel. (A) Coomassie brilliant blue staining of bacterial proteins; (B) Western blot analysis of a parallel gel containing serum generated against a peptide from the HRV-14 protease 3C region. Lanes include the control bacterial strain (lanes Control); strain 202 (lanes 202); strain 207 (lanes 207); strain 209 (lanes 209); and protein markers (lane M).

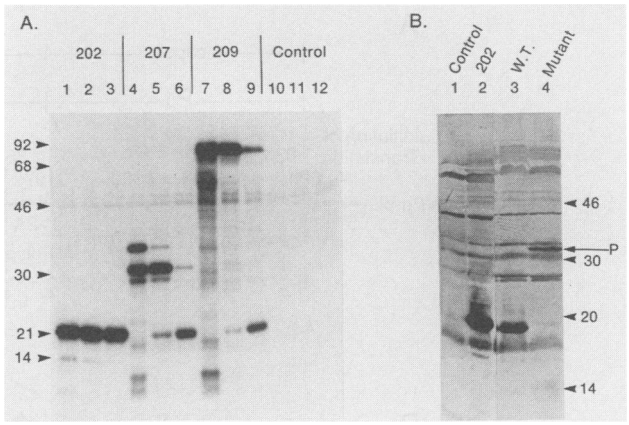

FIG. 3. Analysis of 3C protease processing in *E. coli*. (A) Pulse-chase analysis. Bacteria were induced with 1 mM IPTG for 30 min, at which time they were pulsed for 30 s with [^{35}S]methionine. Samples were removed immediately (lanes 1, 4, 7, and 10) or after a chase period of 10 min (lanes 2, 5, 8, and 11) or 90 min (lanes 3, 6, 9, and 12). Labeled bacterial products were immune precipitated with anti-3C antiserum and analyzed on an SDS–15% polyacrylamide gel. Lanes: 1 to 3, strain 202; 4 to 6, strain 207; 7 to 9, strain 209; 10 to 12, control strain. (B) Western blot analysis of total protein from the following strains: control (lane 1), 202 (lane 2), 226–Cys-146 (lane 3), and 226–Ala-146 (lane 4). Strain 226–Cys-146 contains the wild-type 3C region and is designated W.T.; strain 226–Ala-146 contains a mutant 3C region and is designated Mutant; P indicates the position of the unprocessed 226–Ala-146 precursor.

contained the parental expression vector (Fig. 2A, lane 1). This species was immune reactive with 3C antiserum (Fig. 2B, lane 1) and had the expected size for a protein derived from the 3C region. A species of slightly higher mobility was evident in extracts of strain 207 (Fig. 2A, lane 3) and barely detectable in strain 209 (Fig. 2A, lane 4) by Coomassie staining. The Western blot confirmed that these species were also immune reactive with 3C serum (Fig. 2B). The stained gel showed no evidence of the larger 3C precursor expected from strains 207 and 209; however, Western blot analysis revealed low levels of potential precursors (Fig. 2B, lanes 2 and 3). We speculated that in these strains the primary translation product does not accumulate to high levels and the 20-kDa 3C species must be generated by proteolytic processing of the larger precursor molecule. The slightly faster migration of 3C species generated in strains 207 and 209 compared with strain 202 is consistent with this hypothesis, since 3C synthesized in strain 202 is expected to have an additional N-terminal methionine residue compared with the 3C generated by autoproteolytic cleavage.

To confirm this hypothesis, we performed a pulse-chase analysis of induced cultures, in which proteins radiolabeled during a 1-min pulse were subjected to immune precipitation with the 3C antiserum at various times during the chase period. Strain 202 gave rise to a 20-kDa labeled protein that was stable throughout the 90-min case period (Fig. 3A, lanes 1 to 3). The first labeled species of strain 207 to appear corresponded to the expected primary translation product of 38 kDa. Smaller species of approximately 30 and 27 kDa, which also accumulate during the initial labeling period, are clearly derived by processing from this initial product, since their accumulation after a 10-min chase period (Fig. 3A, lanes 5 and 6) coincides with a decline in levels of the larger species. These species appear to be intermediates in proteolytic processing, leading to the accumulation of the mature 20-kDa species. The longer stability of the larger, 30-kDa intermediate compared with the 27-kDa species indicates that these two species represent alternative intermediates in the processing pathway and do not have a precursor-product relationship, since each one is apparently processed directly into mature 3C protease. On the basis of the available 3C protease cleavage sites in the 207 precursor (Fig. 1A and B), the larger intermediate may correspond to the 3B-3C intermediate having a predicted molecular mass of 30 kDa. A similar situation is also evident with strain 209 in that a large precursor corresponding in size to the expected primary translation product of 86 kDa is evident at the earliest time point (Fig. 3A, lane 7). This species is short lived and is rapidly processed to the 20-kDa species (Fig. 3A, lanes 8 and 9). These results are consistent with the conclusion that 3C protease accumulation is derived by processing from the larger precursor species.

These data clearly show that 3C protease is generated in strains 207 and 209 by proteolytic processing from the precursor encoded by the HRV cDNAs. To confirm that this process, which proceeds by a series of ordered cleavages, is autocatalytic rather than a result of the activity of endogenous bacterial proteases, a cDNA containing a mutant 3C protease was expressed in *E. coli*. Site-directed mutagenesis was used to change the putative active-site cysteine residue (Cys-146) to an alanine. The mutation was engineered into the 3C region of the cDNA expressed in strain 207 to generate strain 226–Ala-146. Western blot analysis of proteins synthesized by *E. coli* 226–Ala-146, containing the mutant 3C region (Fig. 3B, lane 4), shows that the 38-kDa primary translation product is not processed. In contrast, expression of cDNA containing the wild-type (Cys-146) 3C region (strain 226–Cys-146) results in accumulation of only the mature 20-kDa protease molecule (Fig. 3B, lane 3).

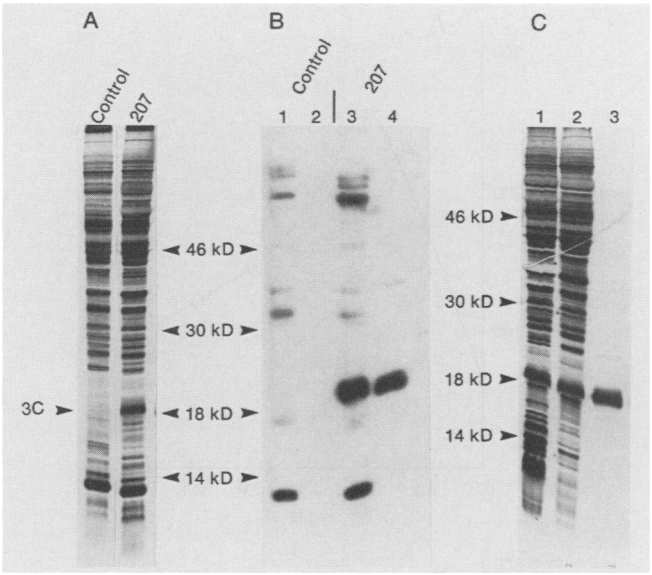

FIG. 4. Purification of recombinant 3C protease from *E. coli*. (A) PAGE and silver staining of total soluble protein from induced control strain and strain 207, which expresses protease 3C. (B) Western blot analysis with 3C antiserum of total soluble protein from control strain (lane 1) and strain 207 (lane 3) and derived purified fractions from control strain (lane 2) and from strain 207 (lane 4) containing 3C protease. (C) PAGE and silver staining of total strain 207 protein (lane 1); total soluble protein (lane 2) and protease 3C purified by DEAE-column chromatography and ammonium sulfate precipitation (lane 3). Equivalent proportions of total material were loaded in each lane.

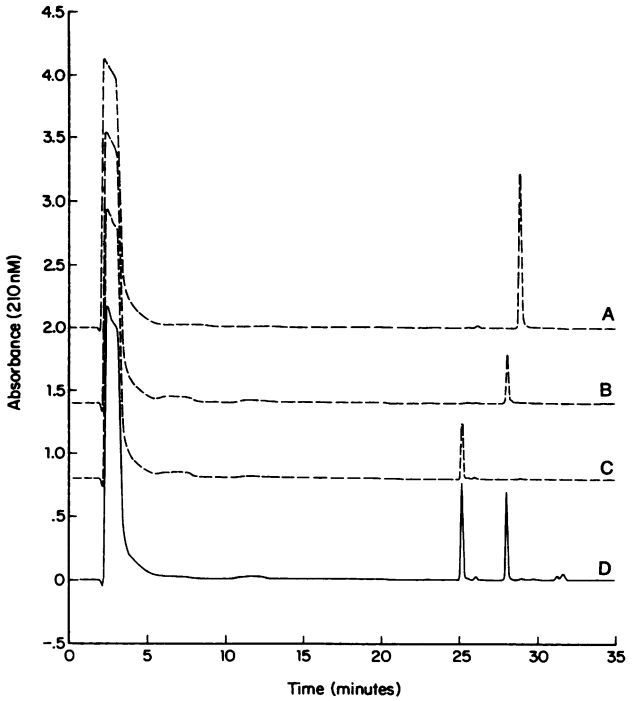

FIG. 5. Specific cleavage of synthetic peptide substrate by purified recombinant HRV protease. Reverse-phase HPLC analysis of peptide cleavage products on a C_{18} column. Peptides were resolved in a linear gradient from 0 to 95% acetonitrile in 0.1% trifluoroacetic acid. The substrate peptide corresponding to 16 amino acids surrounding the 2C-3A cleavage site of HRV-14 (NH_2-DSLETLFQ-GPVYKDLE) was incubated with purified 3C protease (trace D) or with parallel fractions purified from control bacteria (trace A). Marker peptides corresponding to putative cleavage products (NH_2-DSLETLFQ [trace B] and NH_2-GPVYKDLE [trace C]) were separated under identical conditions.

Processing of the primary translation product is therefore autocatalytic, being dependent on the synthesis of an active 3C protease. An analogous mutation of the poliovirus 3C gene product also results in failure of that molecule to undergo proteolytic processing in *E. coli* (12).

Purification of 3C protease. Strain 207 was chosen as a source for purification of recombinant 3C protease, since this strain expresses high levels of functional protease. The protease species was evident as a 20-kDa species representing 3 to 5% of total protein in induced cells containing pR207 but absent in cells containing the parental expression vector pKK233-2 (Fig. 4A). The protease partitioned in the soluble fraction of the bacterial cell lysate, and a simple purification protocol based upon the abnormally high predicted isoelectric point was used. This protocol consisted of anion-exchange chromatography at pH 8.3 so that the vast majority of bacterial proteins are retained on the column and only such highly anionic proteins as 3C flow through. The unbound material from this column was concentrated by ammonium sulfate precipitation. Figure 4B shows a Western blot of the initial soluble fraction from induced cultures of the control bacterial strain and strain 207 which expresses the 3C product. An equivalent proportion of the fractionated material from each strain was also analyzed; the results confirmed the recovery of the 20-kDa immune-reactive species in the purified fraction from strain 207 extract.

Silver staining was used to visualize the polypeptide composition of crude total lysate, soluble proteins prior to fractionation, and purified 3C protease fractions from strain 207 (Fig. 4C, lanes 1 to 3, respectively). From 1.7 g of total protein in a cell lysate, approximately 30 mg of protease was recovered at a level of purity of nominally 90%. This represented an approximate recovery of >50% of the 20-kDa species as judged from stained gels and protein concentrations in peak fractions. An assessment of the recovery of activity was not made since the enzymatic activity of the crude material was not measured.

Enzymatic activity of biosynthetic 3C protease. Previous studies by Libby et al. demonstrated that HRV-14 3C protease activity present in a particulate membrane fraction of expressing *E. coli* cells could cleave a small peptide (16). Therefore, synthetic substrates designed to mimic authentic viral polyprotein cleavage sites were used to determine whether purified 3C protease also was capable of specifically cleaving peptides in vitro. A 16-amino-acid peptide with a sequence corresponding to the 2C-3A cleavage site of the HRV-14 polyprotein was chemically synthesized. Purified fractions of 3C protease or equivalently prepared material from a control bacterial strain which does not contain the 3C protease coding region were incubated with the substrate peptide, and products of the reaction were separated by reverse-phase HPLC. Figure 5, trace A, shows the separation of reaction products following incubation of the peptide with control extract. The HPLC profile of this material shows a single peak characteristic of the 2C-3A peptide substrate and therefore demonstrates that the control fraction contains no activity capable of hydrolyzing this substrate. In contrast, incubation of the peptide with the purified 3C protease resulted in total loss of the substrate peak and

detection of two distinct separable peaks in apparently equal molar concentrations (Fig. 5, trace D). This enzymatic activity was totally inhibited by the addition of $ZnCl_2$ to 0.2 mM in the reaction mixture (data not shown). The sensitivity of cleavage activity to Zn^{2+} is a characteristic feature of picornavirus proteases (14) and provides additional support for our conclusion that HRV 3C protease is responsible for the observed activity in vitro. The product peaks from the 2C-3A cleavage were isolated and subjected to amino acid analysis. The composition of each peak corresponded to that expected for products generated by cleavage between the Gln-Gly sequence of the substrate peptide. Additional confirmation of the fidelity of cleavage was that synthetic peptides, NH_2-DSLETLFQ (Fig. 5, trace B) and NH_2-GPVYKDLE (Fig. 5, trace C), which correspond to the expected products of cleavage, coeluted with the protease-generated products. We conclude from these results that the microbially expressed 3C protease retains the cleavage specificity of the viral enzyme observed in vivo, since it hydrolyzed the amide bond between the Gln and Gly residues.

Characteristics of biosynthetic 3C protease activity. Preliminary characterization of the in vitro enzymatic activity of the HRV-14 3C protease was carried out with the 2C-3A peptide as substrate and used HPLC separation of substrate peptide and cleavage products to quantitate the rate of peptide hydrolysis. Peptide cleavage occurred optimally at pH 7.0, and the activity is relatively sensitive to both acidic and alkaline conditions (Fig. 6A). The kinetic constants of the enzyme with this substrate were derived from Lineweaver-Burk (Fig. 6B) and Eadie-Hofstee plots and found to have a V_{max} of 1 μM/min per mg and a K_m of 250 μM.

Substrate specificity of recombinant protease. To determine whether the purified enzyme can recognize and cleave other HRV-14 cleavage sites, we generated a series of synthetic peptides that were representative of other predicted polyprotein cleavage sites of HRV-14 (Table 1). In addition, control peptides with sequences that do not constitute authentic cleavage sites but do contain a Gln-Gly amino acid pair were synthesized. Control peptide 1 represents an authentic HRV-14 sequence from within the 3C protease coding region, which contains a Gln-Gly pair but is not recognized as a cleavage site in vivo. Control peptide 2 was arbitrarily designed to generate a heterologous 1B-1D site containing the N-terminal end of 1B joined to the C-terminal end of 1D. Cleavage of the various substrate peptides was determined by HPLC analysis, and rates of hydrolysis were measured at the same substrate concentration. The ability of the protease to cleave each peptide and the relative rates of hydrolysis are taken as a rough estimate of how well the substrate is recognized and cleaved by the expressed 3C protease. The data represented in Table 1 reveal that the purified 3C protease had measurable cleavage rates against five of the seven peptides tested. However, the rate of cleavage of these peptides varied within a ca. 300-fold range (Table 1).

No cleavage was detected when control peptides were used as substrates. This result indicated that the mere presence of a Gln-Gly amino acid pair does not fulfill the requirements for substrate recognition and cleavage. The two predicted cleavage sites found to be resistant to cleavage, 1C-1D and 2B-2C, both represent variant sites in that the predicted cleavage site is composed of a Glu-Gly (1C-1D) and a Gln-Ala (2B-2C), instead of the highly conserved Gln-Gly. Consequently, it cannot be ruled out that in vitro cleavage of short peptides requires the presence of a Gln-Gly amino acid pair in the substrate. The peptides that show the highest rates of cleavage overtly show little similarity, with the exception that they all contain a proline residue at the +2 position that flanks the Gln-Gly cleavage site (Table 1).

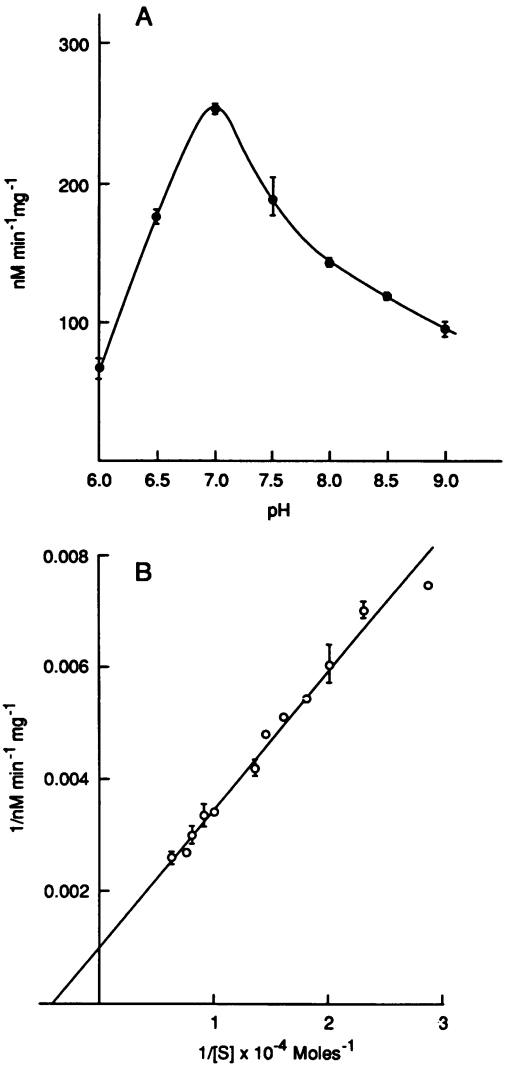

FIG. 6. Enzymatic activity of *E. coli* synthesized 3C protease in vitro. The activity of the purified enzyme was determined for the 2C-3A substrate peptide by HPLC analysis of peptide cleavage products (see Materials and Methods). Error bars represent 2 standard deviations from the mean. (A) Dependence of protease activity on pH. (B) Lineweaver-Burke plot for derivation of kinetic constants. K_m = 250 μM; V_{max} = 1 μM/min per mg.

DISCUSSION

The molecular and biochemical characterization of picornavirus proteases has been limited for many years by the fact that enzyme activity had to be derived from infected cell extracts. Subsequent studies, initially with encephalomyocarditis virus, demonstrated that these enzymes could be purified and that their enzymatic activities could be maintained after storage at low temperature (25). In addition, more recent studies have demonstrated that high-level expression of picornavirus proteases in *E. coli* can be obtained and that the purified enzymes are capable of cleaving small synthetic peptides (16, 19, 20). Studies of the

TABLE 1. HRV polyprotein cleavage sites

Cleavage site	Amino acids	Peptide sequence	Hydrolysis (nM/min per mg)	Relative rate[a] of hydrolysis
1C-1D	560–575	SQTVALTE–GLGDELEE	0	<0.0005[b]
2A-2B	995–1010	LECIAEEQ–GLSDYITG	8	0.024
2B-2C	1093–1108	VPYIERQ–ANDGWFRK	0	<0.0005[b]
2C-3A	1422–1437	DSLETLFQ–GPVYKDLE	335	1.00
3A-3B	1507–1522	YKLFAQTQ–GPYSGNPP	147	0.44
3B-3C	1530–1545	TLRPVVVQ–GPNTEFAL	37	0.11
3C-3D	1712–1727	KQYFVEKQ–GQVIARHK	1	0.003
Control 1	(3C)	HVGGNGRQ–GFSAQLKK	0	<0.0005[b]
Control 2	(1B-1D)	RSKSIVPQ–GLGDELEE	0	<0.0005[b]

[a] Rates of peptide hydrolysis are expressed relative to that of 2C-3A and are the mean of two or more independent determinations (see Materials and Methods). Rates were measured under conditions of constant substrate concentration (less than 10% substrate depletion) and at two time points to confirm the linearity of product accumulation.

[b] No cleavage detectable.

proteases encoded by HRVs have been quite limited. Recently, a 3C precursor of HRV-14 was synthesized in *E. coli* maxicells and shown to autocatalytically process itself to a mature 3C protease (5). Concurrently, a study by Libby et al. (16) demonstrated that the HRV-14 3C protease could be expressed in *E. coli* and that membrane-associated enzyme could cleave a consensus cleavage site peptide of 10 amino acids. However, the 3C protease could not be purified from membranes in an active form.

The studies described here demonstrate that HRV-14 3C polypeptides expressed in *E. coli* contained proteolytic activity capable of autoproteolytically processing and releasing mature 20-kDa protease molecules. Pulse-chase analysis demonstrated that synthesis of the stable 3C species resulted from processing of larger precursor polypeptides via a series of discrete intermediates. Mutation of the Cys-146 residue in the putative active site of the protease to Ala resulted in synthesis of precursor molecules unable to undergo autoprocessing (Fig. 3B), confirming that the proteolytic processing observed was a result of autocatalysis and was not due to processing by an *E. coli* enzyme. Similar observations have been made with the protease encoded by the 3C region of poliovirus (12).

The HRV-14 3C protease synthesized in *E. coli* was evidently functional and was found to be present in the soluble fraction of lysates of *E. coli* cells. A simple purification protocol was used to purify the protease to >90% from a bacterial strain which expressed the protease at between 3 and 5% of total cell protein. This material was enzymatically active and was capable of specifically hydrolyzing synthetic peptide substrates containing the predicted sequence of viral polyprotein cleavage sites. In a manner similar to that found with the viral enzyme in vivo, the proteolytic activity was found to cleave the Gln-Gly bond of substrate peptides and was sensitive to the divalent cation Zn^{2+}. Utilization of a series of synthetic peptides representing the cleavage sites deduced from the sequence of the HRV-14 genome allowed us to examine the substrate specificity of the purified enzyme in vitro. The recombinant enzyme cleaved five of seven predicted protease 3C cleavage site peptides tested at rates that varied over a 300-fold range into the rank order 2C-3A > 3A-3B > 3B-3C > 2A-2B > 3C-3D. There was a notable inability to detect cleavage of either of two control peptides that did not represent authentic cleavage sites but contained a Gln-Gly amino acid pair. These results confirm the notion that amino acid residues flanking the Gln-Gly pair are crucial for recognition and cleavage of substrate. Furthermore, it confirms that this element of substrate recognition is also of critical importance, even when the cleavage site is presented to the enzyme as a short peptide. It is conceivable that cleavage site availability in the polyprotein precursor may also be dramatically affected by the conformation and tertiary structure of the surrounding protein sequences. Although detailed analysis of these requirements awaits further investigation, it is clear that the complex substrate requirement of the enzyme is maintained with short peptide substrates.

Comparison of the cleavage rates obtained with several substrate peptides provided initial insight regarding the important features of the HRV-14 3C protease substrates. The most prominent feature of substrates that are hydrolyzed efficiently in vitro is the presence of a proline at position +2. This is consistent with the observation that a proline or a helix-destabilizing residue is found in close proximity to many observed 3C cleavage sites in picornaviruses (22). However, a proline residue at position −2 is clearly not sufficient to confer substrate activity, since control peptide no. 2 (Table 1) is not a substrate. Another conserved residue was found at position −4, where susceptible peptides invariably contain a Thr, Ala, or Val residue. These amino acids represent relatively conservative substitutions in a position which has been proposed to be preferably Ala by comparison of a broad range of picornavirus cleavage sites (22). Such statistical analysis of cleavage sites yields only limited data and does not take into account the different efficiencies with which some cleavage sites are recognized and cleaved.

The observation that the 2C-3A peptide was the most efficiently cleaved substrate peptide is not surprising, since this cleavage site is between the P2 and P3 region of the viral precursor and is generally believed to be one of the first observed cleavages by 3C protease in poliovirus-infected cells (22). The most stable, but cleavable, peptide substrate represented the 3C-3D cleavage site. We have noted that in HRV-14-infected cells a 3C-D (protease-polymerase) species remains prominent late in infection (unpublished data). In addition, kinetic analysis of the accumulation of mature 3C protease in *E. coli* 209 is consistent with the conclusion that the 3C-3D cleavage site is a poor substrate, since a product corresponding to 3C-3D exhibited a relatively slow turnover in the pulse-chase analysis.

No measurable cleavage was detected for peptides representing the 1C-1D and 2B-2C sites. The failure to cleave the 1C-1D site is not unexpected, since recent studies have suggested that the 3C-3D precursor protein is the molecule required to cleave within the structural P1 region (13, 32).

This was further confirmed by the finding that chimeric 3C-3D proteins, in which the poliovirus 3C was substituted with the 3C protease of either HRV-14 or coxsackievirus B3, were incapable of cleaving the P1 region of poliovirus (6). However, the failure to cleave the 2B-2C peptide remains unexplained. Recent studies with chimeric picornavirus genomes have shown that the HRV-14 3C protease was capable of cleaving the 2B-2C Gln-Gly site of poliovirus (6). Since only authentic peptides containing Gln-Gly cleavage sites had detectable cleavage rates in vitro, it is possible that other sequences, such as the HRV-14 Gln-Ala, are cleaved too poorly to be detected. Non-Gln-Gly sites may exhibit a greater dependency on other secondary and tertiary features, as was recently described for poliovirus (33). It is interesting that the 3C protease of encephalomyocarditis virus would process in vitro-synthesized P3 sequences containing Gln-Ser or Gln-Cys, but not those containing Gln-Thr, Gln-Ile, Gln-Tyr, Arg-Gly, or Leu-Gly, when these residues were substituted into normal cleavage contexts (24).

The availability of a system to test the substrate specificity of the HRV-14 protease systematically with short peptide substrates in vitro should allow rapid illumination of the features of substrates which satisfy the obviously complex substrate requirements of this protease.

ACKNOWLEDGMENTS

We thank Peter Kniskern and Carl Bennett for synthesizing oligonucleotides and Ken Stauffer for providing peptides.

LITERATURE CITED

1. **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.
2. **Bazan, J. F., and R. J. Fletterick.** 1988. Viral cysteine proteases are homologous to the trypsin-like family of serine proteases: structural and functional applications. Proc. Natl. Acad. Sci. USA **85**:7872–7876.
3. **Bernstein, H. D., P. Sarnow, and D. Baltimore.** 1986. Genetic complementation among poliovirus mutants derived from an infectious cDNA clone. J. Virol. **60**:1040–1049.
4. **Callahan, P. L., S. Mizutani, and R. J. Colonno.** 1985. Molecular cloning and complete sequence determination of the RNA genome of human rhinovirus type 14. Proc. Natl. Acad. Sci. USA **82**:732–736.
5. **Cheah, K.-C., S. Sankar, and A. G. Porter.** 1988. Expression and processing of human rhinovirus type 14 polypeptide precursors in Escherichia coli maxicells. Gene **69**:265–274.
6. **Dewalt, P. G., M. A. Lawson, R. J. Colonno, and B. L. Semler.** 1989. Chimeric picornavirus polyproteins demonstrate a common 3C proteinase substrate specificity. J. Virol. **63**:3444–3452.
7. **Gibbs, J. B., I. S. Sigal, M. Poe, and E. M. Scolnick.** 1984. Intrinsic GTPase activity distinguishes normal oncogenic *ras* p21 molecules. Proc. Natl. Acad. Sci. USA **81**:5704–5708.
8. **Gorbalenya, A. E., V. M. Blinov, and A. P. Donchenko.** 1986. Poliovirus-encoded proteinase 3C: a possible evolutionary link between cellular serine and cysteine proteinase families. FEBS Lett. **194**:253–257.
9. **Gwaltney, J. M., Jr.** 1982. Rhinoviruses, p. 491–517. *In* A. S. Evans (ed.), Viral infection of man: epidemiology and control, 2nd ed. Plenum Publishing Corp., New York.
10. **Hamparian, V. V., R. J. Colonno, M. K. Cooney, E. C. Dick, J. M. Gwaltney, Jr., J. H. Hughes, W. S. Jordan, Jr., A. Z. Kapikian, W. J. Mogabgab, A. Monto, C. A. Phillips, R. R. Rueckert, J. H. Schieble, E. J. Scott, and D. A. J. Tyrrell.** 1987. A collaborative report: rhinoviruses—extension of the numbering system from 89 to 100. Virology **159**:191–192.
11. **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.
12. **Ivanoff, L. A., T. Towatari, J. Ray, B. D. Korant, and S. R. Petteway, Jr.** 1986. Expression and site-specific mutagenesis of the poliovirus 3C protease in Escherichia coli. Proc. Natl. Acad. Sci. USA **83**:5392–5396.
13. **Jore, J., B. De Geus, R. J. Jackson, P. H. Pouwels, and B. E. Enger-Valk.** 1988. Poliovirus protein 3CD is the active protease for processing of the precursor protein P1 in vitro. J. Gen. Virol. **69**:1627–1636.
14. **Korant, B. D., and B. E. Butterworth.** 1976. Inhibition by zinc of rhinovirus protein cleavage: interaction of zinc with capsid polypeptides. J. Virol. **18**:298–306.
15. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
16. **Libby, R. T., D. Cosman, M. K. Cooney, J. E. Merriam, C. J. March, and T. P. Hopp.** 1988. Human rhinovirus 3C protease: cloning and expression of an active form in Escherichia coli. Biochemistry **27**:6262–6268.
17. **Merrifield, R. B.** 1963. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. J. Am. Chem. Soc. **85**:2149–2154.
18. **Mizutani, S., and R. J. Colonno.** 1985. In vitro synthesis of an infectious RNA from cDNA clones of human rhinovirus type 14. J. Virol. **56**:628–632.
19. **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.
20. **Pallai, P. V., F. Burkhardt, M. Skoog, K. Schreiner, P. Bax, K. A. Cohen, G. Hansen, D. E. H. Palladino, K. S. Harris, M. J. Nicklin, and E. Wimmer.** 1989. Cleavage of synthetic peptides by purified poliovirus 3C proteinase. J. Biol. Chem. **264**:9738–9741.
21. **Pallansch, M. A., O. M. Kew, B. L. Semler, D. R. Omilianowski, C. W. Anderson, E. Wimmer, and R. R. Rueckert.** 1984. The protein processing map of poliovirus. J. Virol. **49**:873–880.
22. **Palmenberg, A. C.** 1987. Picornaviral processing: some new ideas. J. Cell. Biochem. **33**:191–198.
23. **Palmenberg, A. C., and R. R. Rueckert.** 1982. Evidence for intramolecular self-cleavage of picornaviral replicase precursors. J. Virol. **41**:244–249.
24. **Parks, G. D., J. C. Baker, and A. C. Palmenberg.** 1989. Proteolytic cleavage of encephalomyocarditis virus capsid region substrates by precursors to the 3C enzyme. J. Virol. **63**:1054–1058.
25. **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.
26. **Rivier, J., R. McClintock, and H. Anderson.** 1984. Reversed-phase high-performance liquid chromatography: preparative purification of synthetic peptides. J. Chromatogr. **288**:303–328.
27. **Sakakibara, S., Y. Shimoniski, Y. Kishida, M. Okada, and H. Sugihara.** 1967. Use of anhydrous hydrogen fluoride in peptide synthesis. I. Behavior of various protective groups in anhydrous hydrogen fluoride. Bull. Chem. Soc. Jpn. **40**:2164–2167.
28. **Stanway, G., P. J. Hughes, R. C. Mountford, P. D. Minor, and J. W. Almond.** 1984. The complete nucleotide sequence of a common cold virus: human rhinovirus 14. Nucleic Acids Res. **12**:7859–7875.
29. **Summers, D. F., E. N. Shaw, M. L. Stewart, and J. V. Maizel.** 1972. Inhibition of cleavage of large poliovirus-specific precursor proteins in infected HeLa cells by inhibitors of proteolytic enzymes. J. Virol. **10**:880–884.
30. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheet: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.

31. **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.
32. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major poliovirus proteinase responsible for cleavage of the P1 capsid precursor. Virology **166**:265–270.
33. **Ypma-Wong, M. F., D. J. Filman, J. M. Hogle, and B. L. Semler.** 1988. Structural domains of the poliovirus polyprotein are major determinants for proteolytic cleavage at Gln-Gly pairs. J. Biol. Chem. **263**:17846–17856.