JOURNAL OF VIROLOGY, Dec. 1988, p. 4586–4593
0022-538X/88/124586-08$02.00/0

Vol. 62, No. 12

# Poliovirus Proteinase 3C: Large-Scale Expression, Purification, and Specific Cleavage Activity on Natural and Synthetic Substrates In Vitro

MARTIN J. H. NICKLIN,[1]† KEVIN S. HARRIS,[1] PETER V. PALLAI,[2] AND ECKARD WIMMER[1]*

*Department of Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794-8621,[1] and Boehringer Ingelheim Pharmaceuticals Inc., Ridgefield, Connecticut 06877[2]*

Received 30 June 1988/Accepted 29 August 1988

**Proteinase 3C of poliovirus type 2 (Sabin) was expressed at 4% total protein in *Escherichia coli*. The protein was soluble and could be purified by a simple scheme. It was weakly active on the capsid precursor P1 (expressed in vitro), which contains two cleavage sites. The products of processing of P1 were 1ABC and 1D (VP1). The activity was insensitive to Triton X-100. Crude extracts of cells infected with poliovirus type 1 (Mahoney) gave strong processing and yielded 1AB (VP0), 1C (VP3), and 1D in the same assay system but were sensitive to detergent. 3C from cell extracts that was separated from its precursors resembled the recombinant proteinase in its activity. Recombinant 3C cleaved the peptide dansyl-Glu-Glu-Glu-Ala-Met-Glu-Gln-Gly-Ile-Thr-Asn-Lys-$NH_2$ at the Gln-Gly bond. We conclude that 3C is merely the core of the Gln-Gly-cleaving activity which processes P1 in vivo and that there is probably a hydrophobic contact between a larger 3C precursor and its P1 substrate which allows the second processing reaction: 1ABC, 1D → 1AB, 1C, 1D.**

Poliovirus is a picornavirus and is the best-studied representative of the genus *Enterovirus*, which includes a variety of serious pathogens. Enteroviruses are also broadly related to the rhinoviruses, which are etiological agents of the common cold. These genera have essentially the same genomic structure, and within this group results for a single virus are mostly of general significance. The genome of poliovirus is a single ~7.5-kilobase RNA molecule that includes an open reading frame large enough to encode a 247-kilodalton polyprotein (14). The polyprotein is cotranslationally processed by at least two viral proteinases, mapping to the 3C and 2A regions of the viral genome (9, 29). Most of the processing sites are between Gln and Gly residues (14), and it has been shown (8) that the polypeptide sequence of the viral product 3C is essential for this process and furthermore that 3C releases itself from longer precursors (9) by cleavage at the two Gln-Gly bonds which flank the polypeptide. Two other polypeptides that contain all of the 3C sequence are present at high concentration in infected cell extracts. The most abundant is 3CD, which also contains the whole 3D polymerase sequence and appears to be a relatively stable precursor of both 3C and 3D. The other polypeptide, 3C′, is derived from 3CD by cleavage at a 2A-specific site (Tyr-Gly) and contains a small segment of the 3D protein. It has not been shown previously that 3C itself is catalytically active, although it releases itself from small engineered precursors (9, 12).

On the basis of a comparison of the encoded amino acid sequences of related viruses, the likely catalytic residues of 3C were suggested (1). These results were in agreement with the identification of 3C from encephalomyocarditis virus (EMCV) as a cysteine proteinase by biochemical methods (22). Site-directed mutagenesis has been used to confirm the essential function of Cys-147 and His-161 (12).

It appears that any evolutionary relationship between 3C and nonviral proteinases (notably, cysteine proteinases in particular) is remote, although it has been suggested that 3C is homologous to the trypsinlike family of serine proteases (7; J. F. Bazan and R. J. Fletterick, Proc. Natl. Acad. Sci., in press). The apparently strong specificity of 3C and its homolog toward the *P1* residue (glutamine), of itself, is a remarkable feature in a cysteine proteinase. Together, these observations indicate that the active site of 3C is a suitable target for the design of antiviral drugs that would have a minimal effect on healthy proteolysis by host enzymes.

Other workers have described expression systems for 3C, but it has been necessary to develop a *trans* assay for 3C activity to show that any purification scheme would yield active material for biochemical characterization and crystallography. Furthermore, it has been reported that 3C expressed in *Escherichia coli* is insoluble. We (18) and others (32) have reported methods for assaying Gln-Gly cleavage activity (QG-ase) in infected cell extracts by *trans*-cleavage of the capsomer precursor P1 (Fig. 3A), which we express in vitro. When infected cell extracts are used as the source of QG-ase, the products of processing of P1 are as seen in vivo, namely, P1 → 1AB, 1C, 1D. We used this assay for protein expressed in *Escherichia coli*. Ultimately, it will be more convenient to use a peptide substrate for 3C, and we demonstrate here that a peptide can be processed specifically by the recombinant enzyme.

## MATERIALS AND METHODS

**Bacterial culture medium.** M medium was M9 medium (17) containing 2 g of $NH_4Cl$ per liter and supplemented with 2 g of Casamino acids per liter, 10 g of glucose per liter, 20 μM ferric chloride per liter, and 0.1 g of sodium ampicillin per liter.

**Buffers.** The pHs of all buffers were determined at 22°C. Buffer A was 100 mM NaCl–40 mM Tris hydrochloride (pH 7.9). Buffer B was 100 mM KCl–20 mM Tris hydrochloride–10 mM $MgCl_2$–5 mM dithiothreitol (DTT)–1 mM EDTA (pH 7.9). Buffer C was 40 mM Tris hydrochloride–1 mM DTT (pH 7.9). Buffer D was 100 mM NaCl–20 mM *N*-2-hydroxy-

* Corresponding author.
† Present address: Institute of Molecular Pathology, A-1030 Vienna, Austria.

ethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES)–NaOH–1 mM EDTA–1 mM DTT (pH 7.4). Buffer E was 10 mM Tris hydrochloride–10 mM NaCl–1 mM $MgCl_2$ (pH 7.4). Buffer F was 10% (vol/vol) glycerol–100 mM KCl–10 mM HEPES–KOH–1 mM EDTA–1 mM DTT (pH 8.0). Buffer G was 140 mM KCl–20 mM HEPES–KOH–2 mM DTT–5 mM EDTA (pH 7.2). Buffer S (1×) was 12% (vol/vol) glycerol–76 mM Tris hydrochloride–4 mM Tris–25 mM DTT–1% sodium dodecyl sulfate.

**Gel electrophoresis of protein samples.** Samples were prepared in buffer S. All electrophoresis was performed in 0.1% sodium dodecyl sulfate. Except in the analysis of P1 processing, all gels were 10 to 20% polyacrylamide gradient gels (40 parts acrylamide to 1 part methylenebisacrylamide [bis]). Otherwise, gels were homogeneous 15% acrylamide gels (175 parts acrylamide to 1 part bis). The buffer system used has been described previously (18).

**Preparation of substrate peptide dansyl-EEEAMEQGIT NK-$NH_2$.** The substrate peptide dansyl-EEEAMEQGITNK-$NH_2$, of which the first 11 amino acids correspond to the cleavage site between poliovirus polypeptides 2A and 2B, was synthesized with an automatic synthesizer (SAM2; Biosearch) on methylbenzhydrylamine resin by standard peptide chemistry. The dansyl group was coupled to the peptide on the resin before HF cleavage by treating the free amino terminus with excess dansyl chloride. The peptide was purified to homogeneity by reverse-phase high-pressure liquid chromatography.

**Construction of 3C expression plasmid pMN35.** The T7 expression vector pAR2113 (24) and the expression strain *E. coli* BL21(DE3) (26) were kind gifts of J. Dunn and F. W. Studier (Brookhaven National Laboratory, Upton, N.Y.). Restriction endonucleases were obtained from New England BioLabs, Inc., and T4 DNA ligase was obtained from Bethesda Research Laboratories, Inc. Standard procedures were used in the construction of plasmids. Oligonucleotides were synthesized on a Microsyn 1450A apparatus (Systec Inc.). Poliovirus 3C cDNA was derived from a partial cDNA clone [pVS(2)2501] of the Sabin strain of poliovirus type 2 (28).

**Preparation of bacterial paste.** *E. coli* BL21(DE3) was transformed with plasmid pMN35. After 12 h of incubation at 37°C, a colony was picked to inoculate 50 ml of prewarmed Lennox broth and grown with shaking at 300 rpm at 37°C. After about 3.5 h, the culture (0.5 $A_{600}$) was added to 0.5 liter of M medium in a 2-liter flask and incubated as before for about 4 h, after which the $A_{600}$ reached 0.5 again. The culture was injected into 9.5 liters of M medium and grown with stirring (900 rpm) and aeration (16 liters/min) in a 14-liter fermentor (Microferm II; New Brunswick Scientific Co., Inc.). The pH was maintained between 7.3 and 7.8 by addition of 10 M NaOH. When the $A_{600}$ of a 1/5 dilution of culture (into 0.1 M NaCl) reached 0.7, expression was induced by adding 0.4 mM isopropyl-β-D-thiogalactopyranoside. After 2.5 h, the culture was harvested. Crushed ice (5 kg) was added to the culture with stirring. The bacteria were collected by centrifugation at 3,500 × *g* for 15 min. The pellets were suspended in 1 liter of buffer A and recentrifuged. The final pellet was stored at −80°C.

**Purification of recombinant 3C.** All procedures for purification of recombinant 3C were performed in a room refrigerated to 4°C, unless stated otherwise. Bacterial cell paste (15 g) was thawed and suspended to 40 ml with buffer A. Bacteria were lysed by two passages through a French pressure cell at 70 MPa. The lysate was mixed with a further 40 ml of buffer B (fraction 1) and centrifuged for 2 h at 360,000 × *g* in a Beckmann 60 Ti rotor (at 59,000 rpm) to yield a clear supernatant (70 ml), which was decanted and diluted to 200 ml with water. A 1.2-ml volume of 1 M Tris base was added dropwise with stirring (fraction 2). The final pH of a sample at room temperature was 8.3. This solution was layered on a column (180 ml) of DEAE-cellulose (Whatman DE-52) that had been preequilibrated with buffer C. The sample was eluted with buffer C. The first 130 ml of effluent was discarded, and the next 320 ml was collected. A 2.5-ml volume of 0.5 M EDTA-$Na_3H$ and 5 ml of 0.5 M morpholinoethanesulfonic acid were added dropwise with stirring (fraction 3). Ammonium sulfate was added, with stirring, at 0.55 g/ml to fraction 3 over 15 min, and precipitation was allowed to proceed for 12 h. Precipitated protein was collected by centrifugation at 6,000 × *g* for 60 min at 8°C. The supernatant was discarded, and the pellet was carefully drained and suspended in buffer D to less than 4 ml (fraction 4B). Insoluble material was removed by centrifugation at 4,000 × *g*. To concentrate the protein, 2 g of solid ammonium sulfate was added with gentle agitation until dissolved. The mixture was allowed to stand at 0°C for 30 min, and then the precipitate was collected by centrifugation at 10,000 × *g* at 4°C for 30 min. The pellet was drained, and the precipitate was redissolved in 1 ml of buffer D (fraction 5). In a separate purification, fraction 5 was run on a column (55 by 1.6 cm) of superfine Sephadex G-75 (Pharmacia) in buffer D at 6 ml/h. Peak fractions (10.5 ml total) were identified by electrophoresis and pooled. Protein was collected by ammonium sulfate precipitation (0.5 g added per ml), followed, after 2 h, by centrifugation at 10,000 × *g* for 30 min. The final pellet (fraction 6) was redissolved to about 1 ml in buffer D. When purified, the concentration of 3C was estimated by measuring the $A_{280}$. The protein contains no tryptophan and seven tyrosine residues per molecule of $M_r$ 20,000; thus, the extinction coefficient at 280 nm is about 8,400 $M^{-1}$ $cm^{-1}$ and the $A_{280}$ 0.1% value is about 0.42 $M^{-1}$ $cm^{-1}$ (6).

**Preparation and fractionation of infected HeLa cell lysate.** HeLa R19 spinner cells ($1.2 \times 10^9$) were infected with poliovirus type 1 (Mahoney) at 100 PFU per cell, as described previously (8). At 4 h later, the cells were swollen in 10 ml of buffer E at 0°C. Cells were broken by 15 strokes of a tight-fitting Dounce homogenizer. The homogenizer was drained and rinsed with 2 ml of the same buffer which was pooled with the lysate. The lysate (15 ml) was centrifuged at 10,000 × *g* for 30 min. The supernatant (12 ml) was collected (S-10). A 10-ml volume of S-10 was centrifuged at 300,000 × *g* for 60 min. The remaining 2 ml of S-10 was mixed with 0.5 ml of 50% (vol/vol) glycerol and stored at −80°C. The second supernatant (S-300; 9.5 ml) was mixed with 2.5 ml of 50% glycerol. A 10-ml volume of this fraction was mixed with 4 g of ammonium sulfate. The precipitate was collected by centrifugation at 10,000 × *g* for 30 min and suspended to 1.5 ml in buffer F. Insoluble material was removed by centrifugation at 10,000 × *g* for 5 min. A portion (0.2 ml) of the supernatant was loaded on a Superose 12 HR 10/30 column (Pharmacia) and eluted at 0.2 ml/min in buffer F. Fractions (0.5 ml each) were collected up to 24 ml. Each fraction between elution volumes 11 and 17 ml was concentrated to approximately 75 μl by centrifugal ultrafiltration for 2 h at 5,000 × *g* (Centricon; Amicon Corp.).

**Correlated assays on P1 and peptide cleavage.** Polypeptide P1 was supplied in reticulocyte lysate translation mixtures, as described previously (18), in which 40 g of pMN22 RNA per ml had been translated at 30°C for 60 min in the presence

of 1 mCi of [$^{35}$S]methionine per ml, followed by addition of nonradioactive methionine to 50 μM.

To test recombinant 3C with thiol-reactive inhibitors (E-64 and 1,3-dibromoacetone), the 3C (0.1 ml of a 0.5 mM solution in buffer D) was treated with fresh reducing agent (10 mM DTT), incubated at 30°C for 10 min, and then separated from free reducing agent by gel filtration into buffer G without DTT at 4°C. The effluent which did not contain low-molecular-weight thiol was pooled and diluted to 0.5 ml. To 90 μl of the enzyme solution (about 100 μM) was added 10 μl of buffer G without DTT (control) or a fresh solution of E-64 (to give 0.2 mM), dibromoacetone (to give 0.15 mM), or Triton X-100 (to give 1%). The mixtures were incubated at 30°C for 15 min. From each reaction mixture, 10 μl was withdrawn and added to 20 μl of buffer G–10 mM DTT, except for the final mixture, which was added to buffer G–10 mM DTT–1% Triton X-100. In parallel experiments, 3- and 10-μl volumes of infected S-10 were diluted to 30 μl with buffer G–10 mM DTT and a further 10 μl of infected S-10 was diluted with buffer G–10 mM DTT–1% Triton X-100. P1 translation mixture (4 μl) was then added with mixing, and the samples were incubated at 30°C for 2 h. Protein was precipitated by thorough admixture of 0.3 ml of acetone and collected by centrifugation; it was then drained and dried under vacuum. Pellets were redissolved in buffer S and analyzed by electrophoresis in 15% acrylamide gels. Gels were fluorographed with $En_3Hance$ as described by the manufacturer (New England Nuclear Corp.) and used to expose preflashed X-ray film ($A_{350}$ = 0.2) for 18 h at −80°C. Densitometry was performed on an LKB gel scanner.

Meanwhile, 10 μl of DTT solution (to give 10 mM) was added to the remaining 90 μl of the enzyme-inhibitor mixtures, followed by 10 μl of the peptide substrate solution (0.83 mM). These mixtures were also incubated for 2 h at 30°C. At the end of the incubation, the mixture was diluted 10-fold with water and analyzed on Mono-Q (see Fig. 5B, legend).

**Quantification of protein products.** The integrated absorbance of the P1, 1ABC, 1D, and, in some cases, 1C peaks was recorded. To correct for errors in loading and preparation of the reactions and deviations in lane width, the values of P1, 1ABC, and 1D were summed and then the fraction of the sum which each band represented was calculated. To convert to relative molar quantities ($x_i$) for each polypeptide ($i$), these values were divided by the number of methionine residues in each protein. In this construction, 1D contains 5, 1ABC contains 20, and P1 contains 25 Met residues.

## RESULTS

**Expression of proteinase 3C in BL21(DE3).** The cDNA segment encoding the 3C region of poliovirus type 2 (Sabin strain) (28) was placed in the vector pAR2113 (24) under control of a T7 promoter. The initiation codon of T7 gene 10 was immediately followed by the coding sequence of 3C, followed by a termination codon. The completed plasmid, pMN35, was transferred into the expression strain, BL21 (DE3). The expression system is described in detail elsewhere (24, 26). Briefly, the plasmid contains a T7 promoter inserted into the *Bam*HI site of pBR322, which directs transcription by T7 RNA polymerase, the gene 1 product, toward the promoter of the *bla* gene. Gene 1 itself is inserted into the host chromosome, where its expression is controlled by the *lac* UV5 promoter, and is thus inducible by isopropyl-β-D-thiogalactopyranoside. Induction of gene 1 leads to rapid and overwhelming transcription of plasmid-specific RNA.

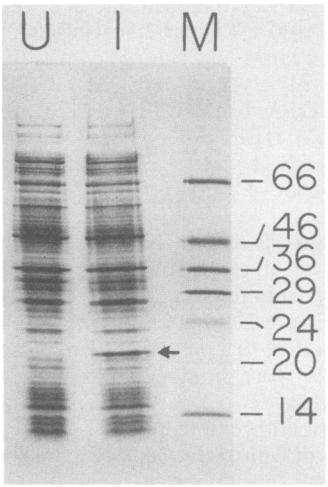


FIG. 1. Analysis on sodium dodecyl sulfate-polyacrylamide gel electrophoresis of biosynthesis of 3C in batch fermentation. Samples were withdrawn from the fermentor immediately before induction (lane U) and 2.5 h after induction (lane I) of expression with isopropyl-β-D-thiogalactopyranoside. Lanes U and I were each loaded with the protein derived from a 10-liter culture. The arrow in lane I indicates the expressed 3C. Protein was visualized by staining with Brilliant Blue R. Lane M contained molecular size markers. The numbers on the right indicate apparent molecular size (in kilodaltons).

To prepare enough cell mass to purify recombinant 3C preparatively, a benchtop fermentor was used as described in Materials and Methods. Figure 1 illustrates the induction of expression.

**Purification of recombinant 3C.** Since the yield of 3C was not very great, a simple purification scheme (detailed in Materials and Methods; the results are shown in Table 1 and Fig. 2) was required that could be readily repeated to yield sufficient material for X-ray crystallographic analysis. Only 5% of all the protein extracted from bacteria was not adsorbed by DEAE-cellulose (Table 1; Fig.2, compare lanes 2 and 3). Roughly 50% of this was 3C. The major contaminant was not precipitated by ammonium sulfate (Fig. 2,

TABLE 1. Purification of biosynthetic 3C from 15 g of *E. coli* BL21 (DE3) (pMN35) cell paste

| Fraction no. | Description | Vol (ml) | Concn (mg/ml) | Total protein (mg) | % yield | % purity |
|---|---|---|---|---|---|---|
| 1 | Crude lysate | 80 | 8.7[a] | 690[a] | 100 | 4 |
| 2 | S-360 | 200 | 1.2[a] | 240[a] | 65 | 8 |
| 3 | DEAE-cellulose effluent | 320 | 0.03[a] | 10[a] | 50 | 50 |
| 5 | Protein in second $(NH_4)_2SO_4$ precipitation | 1.3 | 3.4[a]; 6.8[c] | 4.3[a]; 8.9[c] | 29[b] | 90 |

[a] Determined by Bio-Rad protein assay (dye binding).
[b] Calculated from 3C yield in fraction 5, total protein in fraction 1, and densitometric scan of Fig. 2, lane 1.
[c] Determined by $A_{280}$ from the estimated extinction coefficient (see Materials and Methods).

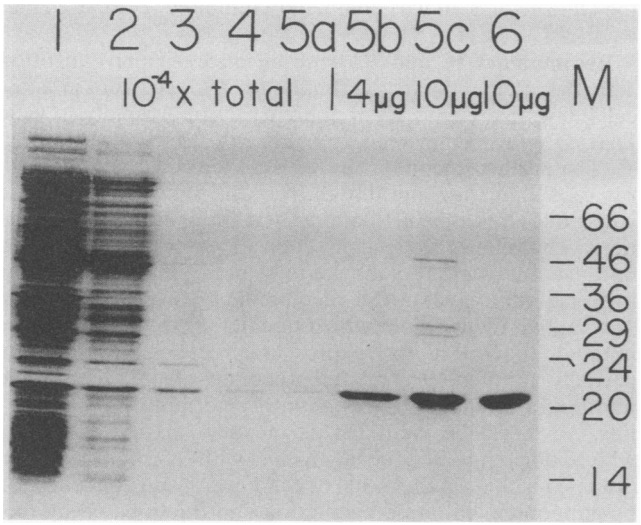


FIG. 2. Purification of recombinant 3C. For lanes 1 to 5a; a fixed proportion (1/10,000) of certain fractions (as indicated) was loaded in buffer S. Lanes 5b and 5c were loaded with 4 and 10 μg (by $A_{280}$), respectively, of fraction 5. Lane 6 contained 10 μg of fraction 6. Protein was visualized by staining with Brilliant Blue R. The numbers on the right (M) indicate apparent molecular sizes (in kilodaltons).

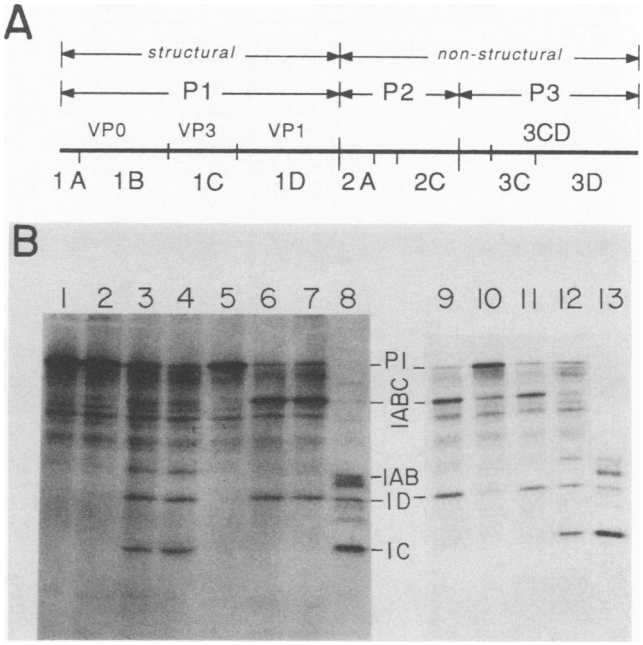


FIG. 3. Proteolytic processing of poliovirus proteins. (A) Schematic representation of the poliovirus polyprotein (heavy line). Individual polypeptides (not drawn to scale) are designated according to the nomenclature of Rueckert and Wimmer (25). The polyprotein is divided conceptually into the following three domains: P1, the capsid region which is eventually processed into the four capsid proteins 1A, 1B, 1C, and 1D (also called VP4, VP2, VP3, and VP1, respectively); P2, which yields 2A, 2B (not indicated), and 2C; and P3, which yields 3A, 3B (VPg; not indicated), 3C, 3D (the polymerase), and a variety of alternative cleavage products. Proteinase 2A (29) (and its precursors) cleaves at specific Tyr-Gly pairs (the P1-P2 site and a site in 3D). With the exception of the 1A-1B cleavage site, all other cleavages are made by 3C or its precursors (9). For a complete processing map, see Pallansch et al. (19). Note that 1AB is cleaved to 1A and 1B when the viral RNA is encapsidated and probably involves neither 2A nor 3C (2). The reaction has not been reproduced in vitro. (B, left panel) Comparison of processing of P1, expressed in vitro, by crude extracts of poliovirus-infected cells and by purified recombinant 3C. See Materials and Methods for conditions. Lanes: 1, no QG-ase; 2, no QG-ase and 1% Triton X-100; 3, 3 μl of infected S-10; 4, 10 μl of infected S-10; 5, 10 μl of S-10 and 1% Triton X-100; 6, about 25 μM 3C; 7, about 25 μM 3C and 1% Triton X-100; 8, extract of infected cells labeled for 3 h postinfection. 1AB, 1C, and 1D are indicated with arrows. (B, right panel) Effect of preincubation with thiol-reactive inhibitors on about 0.1 mM recombinant 3C for 15 min. Lanes: 9, 0.2 mM E-64; 10, 0.15 mM 1,3-dibromoacetone; 11, no inhibitor; 12, infected S-10 as in lane 4; 13, markers (as in lane 8). The proteinase was diluted to about 25 μM in the reaction mixture.

compare lanes 3 and 4). After reprecipitation from a small volume, 3C was 90% pure by densitometry (Fig. 2, lanes 5a, b, and c), and material of this quality was used in most of our experiments. Ninety-eight percent pure material was obtained by adding a single step of molecular size chromatography (Fig. 2, lane 6) with around 90% recovery. The overall recovery of protein was 30%, and one can isolate 8 mg of 3C from 15 g of wet *E. coli* cell paste.

The N-terminal sequence of the protein product was determined up to residue 10 by automated gas phase Edman degradation. The size of signal was commensurate with the sample loaded, and the data were unambiguous and did not deteriorate appreciably over the 10 cycles. The sequence determined was Met-Gly-Pro-Gly-Phe-Asp-Tyr-Ala-Val-Ala, i.e., the predicted amino terminus of 3C but with an additional methionine. Some of the purified 3C was used to raise antibodies in rabbits. The antisera strongly recognized 3C from cells infected with type 1 poliovirus (see Fig. 4A; compare lanes 31 and 32 with the recombinant 3C standards on the right of the blot).

**Activity of recombinant 3C on capsid precursor P1 expressed in vitro.** In preliminary experiments, crude extracts were prepared from induced BL21(DE3)(pMN35) and BL21(DE3)(pMN42), of which the latter plasmid leads to expression of the 3C protein from EMCV. No specific cleavage of poliovirus P1, expressed in vitro, was obtained with the latter expression system, but it was highly active against the EMCV 3C substrate, LVP0 (data not shown). The precursor polypeptide LVP0 of EMCV was also expressed in vitro by us, as previously described (20). Processing of poliovirus P1 by the bacterial lysate containing poliovirus 3C (the former) was qualitatively the same as that observed later with the purified protein (data not shown).

We used in vitro translation products of synthetic mRNA encoding the capsomer precursor, P1, derived from poliovirus type 1 (Mahoney) cDNA (18) to monitor QG-ase activity. After 2 h of incubation in buffer, no apparent change in the products of the translation was observed. Figure 3B, lane 1, shows the products of the control incubation. In lanes 3 and 4, infected S-10, 3 and 10 μl, respectively, was used as the source of QG-ase. As reported previously (18), processing yielded the three capsomer proteins 1AB, 1C, and 1D. It is clear that the comparative rates of processing at the two cleavage sites are similar, since all three capsomer proteins were observed even while processing of P1 was partial (lane 3). In lane 6, about 25 μM purified biosynthetic 3C was included in the incubation, giving almost complete cleavage of P1 but yielding only 1ABC and 1D, with no detectable cleavage of 1ABC to 1AB and 1C. We were able to detect a trace of 1C (identified by its mobility on gel electrophoresis)

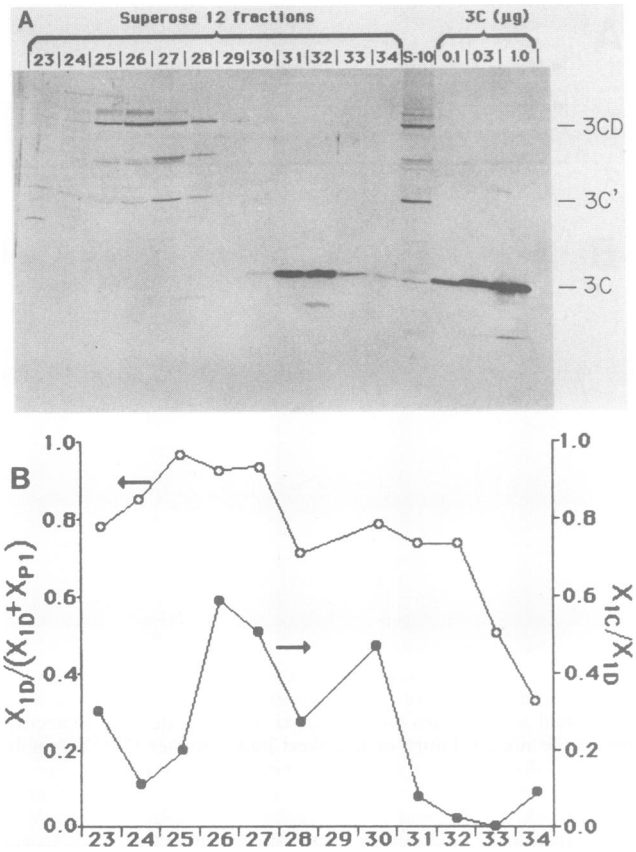


FIG. 4. Fractionation of soluble QG-ase from poliovirus-infected cells on fast-protein liquid chromatography with Superose 12 HR10/30 (bed volume, 23.6 ml). The sample size was 0.2 ml, and fractions were 0.5 ml each but were concentrated to 75 μl (see Materials and Methods). Effluent from 11 to 17 ml (fractions 23 to 34) was analyzed in detail. (A) Immunoblotting of 10 μl of each fraction (after sodium dodecyl sulfate-polyacrylamide gel electrophoresis) was performed as described by Kräusslich et al. (16), except that the transfer was for 2 h only at 0.2 A. Rabbit anti-3C serum raised against the recombinant protein was used at a 1/300 dilution. Detection was with goat anti-rabbit immunoglobulin G coupled to alkaline phosphatase, with indolyl phosphate-Nitroblue Tetrazolium chloride as the indicator substrate system. The positions of the 3C-related proteins in the infected S-10 are labeled at the right margin. (B) A 10-μl volume of each fraction was incubated with 2 μl of P1 translation mixture for 3 h at 30°C. The gel was prepared, fluorographed, and scanned as described in the text, and the extent of cleavage (0 to 1) to yield 1D (open circles) and the molar ratio of 1C/1D (filled circles) were calculated. The arrows indicate the ordinates corresponding to open or filled circles.

at still higher 3C concentrations (data not shown). We estimate that the concentration of 3C-related proteins in the reaction mixture analyzed in lane 3 was about 1/100 of that present in the mixture analyzed in lane 6 (compare the S-10 lane with the 3C standards in Fig. 4A). Although recombinant 3C was active against P1, its activity was apparently low and restricted largely to cleavage of the 1ABC-1D bond. In separate experiments, we were able to determine the rate constant, $k_{cat}/K_m$, to be 80 $M^{-1}$ $S^{-1}$ for the interaction of 3C with P1 (data not shown).

When the same experiments were performed in the presence of 1% Triton X-100, no qualitative difference could be observed (Fig. 3B, compare lanes 6 and 7) when recombinant 3C was used as the QG-ase. Under the same conditions, the QG-ase of S-10 was completely inhibited (compare lanes 4 and 5).

**Recombinant 3C and 3C from infected cells have identical activities.** Recombinant 3C (unlike the QG-ase from infected cells) does not cleave P1 of poliovirus type 1 rapidly to yield 1AB, 1C, and 1D; instead, it yields only 1ABC and 1D (Fig. 3). The purified recombinant 3C is derived from the poliovirus type 2 genome and differs from the 3C in cells infected with poliovirus type 1 (Mahoney) at three amino acid positions. This should have no effect on P1 processing, however, because intertypic recombinants of poliovirus readily grow in tissue culture (23). The recombinant 3C also retains the N-terminal methionine which results from the synthetic initiation codon in the expression vector. We therefore partially purified 3C and its precursors from the soluble fraction of infected cell extracts to determine their properties. 3C has been identified as the only polioviral protein which does not sediment mainly with membranes (27). Soluble proteins in the S-300 of cell lysate were concentrated by ammonium sulfate precipitation and fractionated by molecular size. In a preliminary experiment, the size range over which the bulk of QG-ase was eluted was determined. A finer analysis was made with three assays (Fig. 4) over the range of QG-ase as follows. (i) We performed an immunoblotting analysis (Fig. 4A) with anti-3C serum. (ii) We determined the proportion of P1 that was cleaved to release 1D (Fig. 4B, open circles) in an assay. (iii) We calculated the molar ratio of 1C to 1D ($x_{1C}/x_{1D}$; Fig. 4B, filled circles). Since the measurements involved in calculating the third set of data were complicated, large proportional experimental errors were expected; thus, the difference in $x_{1C}/x_{1D}$ between fractions 28 and 30 is probably not significant. The fractions which contained the peak of QG-ase activity (26 to 28) in the P1 assay did not contain 3C but did contain 3C', 3CD, and possibly other 3C-related precursors. 3CD was also detected in fraction 30 in an immunoblot treated with 10 times more concentrated anti-3C serum. It is also clear that fractions 31 and 32 contain abundant 3C and undetectable levels of 3C precursors and generate very little 1C in a P1 cleavage assay, despite extensive processing to yield 1ABC and 1D. The activity of the 3C (which is approximately 1 μM) in these fractions therefore accords well both quantitatively and qualitatively with that of recombinant 3C.

**A peptide substrate specifically processed by 3C.** Two procedures were used to detect the products of cleavage of the fluorescent peptide dansyl-EEEAMEQGITNK-$NH_2$ by recombinant 3C. In one approach (Fig. 5A and B), the peptide was subjected to extended digestion (50 μM peptide–12.5 μM 3C in buffer D for 14 h at 30°C), and all of the new UV-absorbing products which did not arise from the buffer alone were isolated by high-pressure liquid chromatography and analyzed by fast-atom bombardment mass spectroscopy (data not shown). Total conversion of the parent material (Fig. 5A, peak 1) to peaks 2 and 3 (Fig. 5B) was observed. The molecular masses of peaks 2 and 3 are consistent with the expected masses of dansyl-EEEAMEQ and GITNK-$NH_2$, respectively, the two products of cleavage at the Gln-Gly bond.

In separate experiments, the peptide (83 μM peptide–50 μM 3C in buffer D for 2 h at 30°C) was approximately 50% converted by recombinant 3C to a new fluorescent material with affinity for Mono-Q greater than that of the parent peptide (Fig. 5C and D and legend). After 2 h of incubation, 0.5 ml of acetone at room temperature was added to an identical reaction mixture to denature the protein and stop the reaction. The reaction mixture was dried in vacuo, and

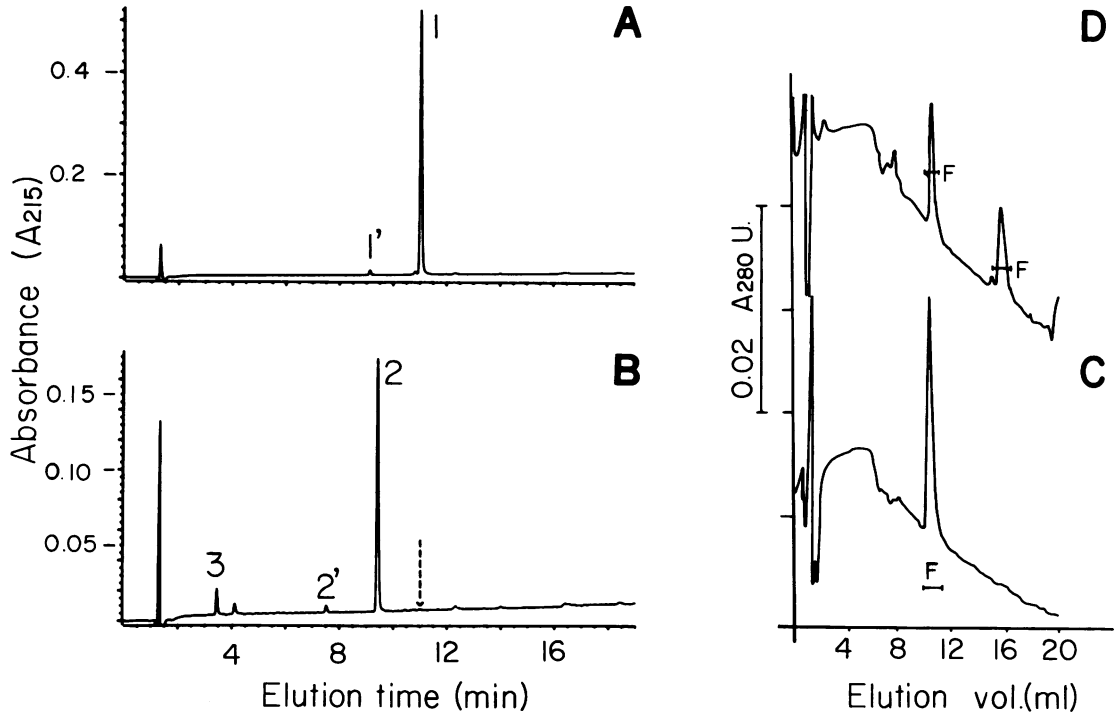


FIG. 5. Processing of the peptide dansyl-EEEAMEQGITNK-$NH_2$ by biosynthetic 3C. (A) Reverse-phase analysis of the substrate peptide. (B) Analysis of cleavage reaction. The peptide (50 μM) was incubated with 12.5 μM 3C for 14 h at 30°C in a final volume of 0.1 ml of buffer C. Samples were run on a Vydac C-18 column in a linear 60-ml gradient at 2 ml/min, containing 5 to 100% solution B mixed in solution A. Solution A was 0.1% trifluoroacetic acid in water, and solution B was 0.1% trifluoroacetic acid in 40% water-60% acetonitrile. The peak fractions were pooled and analyzed by fast-atom bombardment mass spectroscopy. Peak 1 yielded a molecular ion of $M_r$ 1,610.6, as expected of the uncleaved peptide. Peak 2 yielded a molecular ion of $M_r$ 1,098, as expected of the peptide dansyl-EEEAMEQ. Peptide 3 yielded a molecular ion of $M_r$ 531, as expected of GITNK-$NH_2$. Peaks 1′ and 2′ yielded molecular ions of $M_r$s 1,626.7 and 1,114, respectively, as expected from the methionine sulfoxide derivatives of peaks 1 and 2. (C and D) Ion-exchange analysis. The peptide (83 μM) was incubated in buffer G with either no enzyme (C) or 50 μM 3C in 0.1 ml (D) for 2 h. The mixtures were diluted 10-fold and resolved on fast-protein liquid chromatography with Mono-Q HR5/5. After sample application, the column was washed with 3 ml of 20 mM bis–Tris hydrochloride (pH 6.7), followed by a 17-ml gradient from 0 to 50% 20 mM bis–Tris hydrochloride–1 M NaCl (pH 6.7) mixed with 20 mM bis–Tris hydrochloride (pH 6.7). The fluorescent fractions were detected by standing the tubes over a 330-nm transilluminator and are indicated by horizontal bars labeled F. The reaction mixture was also analyzed as described in Results.

the dry pellet was extracted with dimethylformamide and centrifuged, and the solution was transferred to a new tube for drying in vacuo. A portion of this material was analyzed by gas phase automated Edman degradation for six cycles. The sequence Gly-Ile-Thr-Asn-Lys was obtained (data not shown), with no amino acid derivative detected in cycle 6. Small contaminating peaks (running with the derivatives of Asp and Ala in cycle 1, with Gln in cycle 2, and with Ile in cycle 3) did not represent any other sequence within the substrate peptide. Together, these data show that the enzyme preparation cleaved the synthetic peptide at the Gln-Gly cleavage site.

**Correlation of activity of inhibitors in P1 and peptide assays.** In separate experiments, recombinant 3C (at approximately 90 μM) was pretreated with the two cysteine proteinase inhibitors L-*trans*-expoxysuccinylleucylamido(4-guanidinobutane) (E-64; 0.2 mM [3]) and 1,3-dibromoacetone (0.15 mM [11]). A control reaction contained only enzyme and buffer (see Materials and Methods). After addition of excess thiol, the mixtures were tested for activity in both the P1 and peptide assays. In the control assays, P1 was almost completely converted to 1ABC and 1D (Fig. 3B, lane 11) and the fluorescent peptide was detected under UV illumination as running almost completely in the position of the product peak, as in Fig. 5D (data not shown). The pattern of neither assay was changed (Fig. 3B, lane 9) by preincubation of 3C with E-64; hence, E-64 does not appear to inhibit the enzyme under these conditions. However, after treatment with dibromoacetone (Fig. 3B, lane 10), P1 was only weakly cleaved, whereas in fast-protein liquid chromatography roughly half of the fluorescent peptide migrated in the position of the unmodified peptide and half migrated in the position of the processed peptide (data not shown). Thus, dibromoacetone appeared to inhibit the activity of 3C in both reactions.

## DISCUSSION

**Expression of 3C in *E. coli*.** Ivanoff et al. (12) have previously reported efficient expression (15% of total protein) of type 1 (Mahoney) poliovirus 3C from a *trp* promoter in *E. coli*, but the product was described as largely insoluble. Large amounts of internal initiation products are observed when poliovirus type 1 cDNA is used for expression of 3C (9, 12; M.J.H.N. and E.W., unpublished data). We expressed 3C derived from poliovirus type 2 (Sabin) cDNA, which does not give internal initiation in the 3C region, probably because of a single-base difference upstream of Met-27. The level of expression of the protein product of our system was only 4% by densitometry of a stained gel (Fig. 1,

compare lanes U and I), but it was largely extracted into a high-speed supernatant during purification (Fig. 2, lanes 1 and 2) and was easily purified. Other workers have also recently expressed type 2 3C, derived from the same cDNA clone, from an *E. coli* promoter (A. Nomoto, personal communication). However, like Ivanoff et al., they have found that their product is insoluble but active after renaturation.

The N-terminal methionine of our recombinant protein was unexpectedly retained (4, 30). However, as described below, the recombinant and the natural 3C seemed comparable in activity.

**The processing reaction.** Jackson (13) has shown in synchronized in vitro translations of EMCV RNA that the precursors of 3C are active in cleaving the capsid precursor in this system. EMCV 3C, however, is fully active upon the complex EMCV capsid precursor (20). There has previously been no clear report that 3C of poliovirus alone is a QG-ase. Here we show that recombinant 3C is proteolytically active. However, pure 3C cleaves only one of the two peptide bonds in P1 which are cleaved in vitro by the QG-ase from infected cell extracts (Fig. 3B) and are readily processed in vivo. Ypma-Wong and Semler (32) have reported previously that on translation in vitro (where the concentration of enzyme and substrate would be in the range of $10^{-8}$ M), the entire 3CD region must be translated for processing of P1 (synthesized from the same mRNA) to be observed, whereas the P2 and P3 regions apparently require only the 3C region itself. Our kinetic data (M.J.H.N. and E.W., unpublished data) show that 3C would need to be 2 orders of magnitude more concentrated to produce detectable processing of P1 in this system. More recent work, in which P1, P2, and P3 substrates were translated independently of the proteinase activity, has supported these findings (13a, 31a; H.-G. Kräusslich, C. Hellen, M. J. H. Nicklin, and E. Wimmer, unpublished data). It is likely, therefore, that the $K_m$ of 3C for the P2 and P3 substrates is much lower than that for P1. We have shown that soluble precursors of 3C, when partially purified from infected cells, are highly active on P1, although 3C itself is apparently absent from these fractions. These precursors cleave P1 fully to 1AB, 1C, and 1D. The 3C from infected cells shows weak activity similar to that of recombinant 3C, yielding 1ABC and 1D. The implication of these findings is that 3C lacks a domain required for effective interaction with P1. Our observation that the QG-ase from infected cells is rendered inactive (or merely as inactive as 3C) upon the P1 substrate by inclusion of detergent in the reaction mixture, whereas high concentrations of 3C are unaffected, suggests that the interaction of the functional QG-ase and P1 is mediated by a hydrophobic interaction. In this respect it would be relatively simple to test for the involvement of the N-terminal myristoyl moiety of P1 (5, 21).

It has been reported that 3CD is not active as a polymerase (31); hence, the processing (by the QG-ase) of 3CD to 3C and 3D represents a potential regulatory step in which the P1 QG-ase is inactivated and the polymerase is activated.

**Poliovirus 3C as a peptidase.** Clearly, the most desirable form of assay for a proteinase is a peptidolytic assay. It would also open up a convenient route for quantitative studies of the primary specificity of the enzyme and the investigation of inhibitors. An esterolytic assay of poliovirus proteinase 2A has been reported (15), but none has been reported for 3C. The peptide sequence EEEAMEQGITN was selected to represent the QG-ase-cleavable site between polypeptides 2A and 2B, which is efficiently cleaved in vivo. The peptide synthesized was dansyl-EEEAMEQGITNK-$NH_2$. The reaction catalyzed by 3C yielded two peptides with molecular masses corresponding to those of the products of cleavage between Gln and Gly, and amino-terminal sequencing of all dimethylformamide-soluble peptides present in the mixture yielded a single N-terminal sequence corresponding to the C-terminal product, GITNK, as expected from the Gln-Gly specificity of the enzyme on its protein substrates. Synthetic peptides not containing a proper QG site were entirely resistant to even large amounts of 3C (P.V.P. and E.W., unpublished data). E-64, a highly specific, irreversible inhibitor of the papain superfamily (including many lysosomal cysteine proteinases) (3), apparently did not react with recombinant 3C to inhibit its activity in either the peptidolytic or the P1 cleavage assay. Dibromoacetone, on the other hand, inhibited the enzyme correspondingly. The latter inhibitor is not likely to exhibit much specificity, although it may cross-link the active-site Cys and His residues (11).

## ACKNOWLEDGMENTS

We thank P. Kissel for synthesis of the oligonucleotides, M. G. Murray for uncoupling of the peptide, F. Burkhardt for high-pressure liquid chromatographic analysis, G. Hansen for mass spectroscopy, and M. Schmidt for preparation of rabbit anti-3C. We are grateful to R. Jackson. B. Enger-Valk, and B. L. Semler for communicating their data before publication. We also thank H. Toyoda, F. W. Studier, R. Cabelli, H. G. Kräusslich, and C. Hellen for stimulating discussions; A. Kameda and J. Bradley for preparing figures; L. Zywacki for help in preparing the manuscript; and C. Helmke for photography.

Gas-phase polypeptide sequencing was performed by the Center for Analysis and Synthesis of Macromolecules at Stony Brook, supported by National Institutes of Health grant RR02427 and the Center for Biotechnology, State University of New York. This work was supported in part by National Institutes of Health grants A115122 and CA28146. M.N. is fellow DRG-848 of the Damon Runyon-Walter Winchell Cancer Fund.

## LITERATURE CITED

1. **Argos, P., G. Kamer, M. J. H. Nicklin, and E. Wimmer.** 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. Nucleic Acids Res. **12:** 7251–7276.
2. **Arnold, E., M. Luo, G. Vriend, M. G. Rossmann, A. C. Palmenberg, G. 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.
3. **Barrett, A. J., A. A. Khembhavi, M. A. Brown, H. Kirschke, C. G. Knight, C. G. Tamai, and K. Hanada.** 1982. L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. Biochem. J. **201:**189–198.
4. **Ben-Basset, A., K. Bauer, S. Chang, K. Myambo, A. Boosman, and S. Chang.** 1987. Processing of the initiation methionine from proteins: properties of the *Escherichia coli* methionine aminopeptidase and its gene structure. J. Bacteriol. **169:**751–757.
5. **Chow, M., J. F. E. Newman, D. Filman, J. M. Hogle, D. J. Rowlands, and F. Brown.** 1987. Myristylation of picornavirus capsid protein UP4 and its structural significance. Nature (London) **327:**482–486.
6. **Fasman, G. D. (ed.).** 1977. CRC handbook of biochemistry and molecular biology, 3rd ed., vol. 1, 187–189. CRC Press, Inc., Boca Raton, Fla.
7. **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.
8. **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.

9. **Hanecak, R., B. L. Semler, H. Ariga, C. W. Anderson, and E. Wimmer.** 1984. Expression of a cloned gene segment of poliovirus in *E. coli*: evidence for autocatalytic production of the viral proteinase. Cell **37**:1063–1073.
10. **Hartley, B. S.** 1970. Strategy and tactics in protein chemistry. The first BDH lecture. Biochem. J. **119**:805–822.
11. **Husain, S. S., and G. Lowe.** 1968. Evidence for histidine in the active site of papain. Biochem. J. **108**:855–859.
12. **Ivanoff, L. A., T. Towatari, J. Ray, B. D. Korant, and S. R. Petteway, Jr.** 1986. Expression and site-specific mutagenesis of the proteinase 3C protease in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **83**:5392–5396.
13. **Jackson, R.** 1986. A detailed kinetic analysis of the *in vitro* synthesis and processing of encephalomyocarditis virus products. Virology **149**:114–127.

13a. **Jore, J., B. de Geus, R. J. Jackson, P. H. Powels, 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. **Kitamura, N., B. L. Semler, P. G. Rothberg, G. R. Larsen, C. J. Adler, A. J. Dorner, E. A. Emini, R. Hanecak, J. J. Lee, S. van der Werf, C. W. Anderson, and E. Wimmer.** 1981. Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature (London) **291**:547–553.
15. **König, H., and B. Rosenwirth.** 1988. Purification and partial characterization of poliovirus protease 2A by means of a functional assay. J. Virol. **62**:1243–1250.
16. **Kräusslich, H. G., M. J. H. Nicklin, H. Toyoda, D. Etchison, and E. Wimmer.** 1987. Poliovirus proteinase 2A induces cleavage of eucaryote initiation factor 4F polypeptide p220. J. Virol. **61**: 2711–2718.
17. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Nicklin, M. J. H., H. G. Kräusslich, H. Toyoda, J. J. Dunn, and E. Wimmer.** 1987. Poliovirus polypeptide precursors. Expression *in vitro* and processing by 3C and 2A proteinases. Proc. Natl. Acad. Sci. USA **84**:4002–4006.
19. **Pallansch, M. A., O. Kew, B. L. Semler, D. R. Omilianowski, C. W. Anderson, E. Wimmer, and R. R. Rueckert.** 1984. Protein processing map of poliovirus. J. Virol. **49**:873–880.
20. **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.
21. **Paul, A. V., A. Schultz, S. E. Pincus, S. Oroszlan, and E. Wimmer.** 1987. Capsid protein VP4 of poliovirus is N-myristylated. Proc. Natl. Acad. Sci. USA **84**:7827–7831.
22. **Pelham, H. R. B.** 1978. Translation of encephalomyocarditis virus RNA *in vitro* yields an active proteolytic processing enzyme. Eur. J. Biochem. **85**:457–461.
23. **Romanova, L. I., E. A. Tolskaya, M. S. Kolesnikova, and V. I. Agol.** 1980. Biochemical evidence for intertypic genetic recombination of polioviruses. FEBS. Lett. **118**:109–110.
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. **Rueckert, R. R., and E. Wimmer.** 1984. Systematic nomenclature of picornavirus proteins. J. Virol. **50**:957–959.
26. **Studier, F. W., and B. A. Moffat.** 1986. Use of T7 RNA polymerase to direct selective, high-level expression of cloned genes. J. Mol. Biol. **189**:113–130.
27. **Tershak, D. R.** 1984. Association of poliovirus proteins with the endoplasmic reticulum. J. Virol. **52**:777–783.
28. **Toyoda, H., M. Kohara, Y. Kataoka, T. Suganama, T. Omata, N. Imura, and A. Nomoto.** 1984. Complete nucleotide sequences of all three poliovirus serotype genomes: implication for genetic relationship, gene function and antigenic determinants. J. Mol. Biol. **174**:561–585.
29. **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.
30. **Tsunawa, S., J. W. Stewart, and F. Sherman.** 1985. Amino-terminal processing of mutant forms of yeast *iso*-1-cytochrome *c*. (The specificities of methionine aminopeptidase and acetyltransferase). J. Biol. Chem. **260**:5382–5391.
31. **Van Dyke, T. A., and J. B. Flanegan.** 1980. Identification of poliovirus polypeptide p63 as a soluble RNA-dependent RNA polymerase. J. Virol. **35**:732–740.

31a. **Ypma-Wong, M. F., P. G. Dewalt, V. H. Johnson, J. G. Lamb, and B. L. Semler.** 1988. Protein 3CD is the major proteinase responsible for cleavage of the P1 capsid precursor. Virology **166**:265–270.

32. **Ypma-Wong, M. F., and B. L. Semler.** 1987. *In vitro* molecular genetics as a tool for determining the differential cleavage specificities of the poliovirus 3C proteinase. Nucleic Acids Res. **15**:2069–2088.