Olivera, B. M., McIntosh, J. M., Cruz, L. J., Luque, F. A., & Gray, W. R. (1984) *Biochemistry 23*, 5087-5090.
Olivera, B. M., Gray, W. R., Zeikus, R., McIntosh, J. M., Varga, J., Rivier, J., de Santos, V., & Cruz, L. J. (1985) *Science (Washington, D.C.) 230*, 1338-1343.
Olivera, B. M., Cruz, L. J., de Santos, V., LeCheminant, G. W., Griffin, D., Zeikus, R., McIntosh, J. M., Galyean, R., Varga, J., Gray, W. R., & Rivier, J. (1987) *Biochemistry 26*, 2086-2090.
Pappone, P. A. (1980) *J. Physiol. (London) 306*, 377-410.
Poser, J. W., & Price, P. A. (1979) *J. Biol. Chem. 254*, 431-436.
Price, P. A. (1983) *Methods Enzymol. 91*, 13-17.
Sato, S., Nakamura, H., Ohizumi, Y., Kobayashi, J., & Hirata, Y. (1983) *FEBS Lett. 155*, 277-280.
Stenflo, J., & Suttie, J. W. (1977) *Annu. Rev. Biochem. 46*, 157-172.
Ui, N. (1979) *Anal. Biochem. 97*, 65-71.
Waterman, M. S., Smith, T. F., & Beyer, W. A. (1976) *Adv. Math. 20*, 367-387.
Weiss, R. E., & Horn, R. (1986) *Science (Washington, D.C.) 233*, 361-364.
Yanagawa, Y., Abe, T., & Satake, M. (1986) *Neurosci. Lett. 64*, 7-12.
Yanagawa, Y., Abe, T., & Satake, M. (1987) *J. Neurosci. 7*, 1498-1502.

Human Rhinovirus 3C Protease: Cloning and Expression of an Active Form in *Escherichia coli*

Randell T. Libby,*,‡ David Cosman,‡ Marion K. Cooney,§ Janet E. Merriam,‖ Carl J. March,‖ and Thomas P. Hopp‖

Departments of Molecular Biology and Protein Chemistry, The Immunex Corporation, 51 University Street, Seattle, Washington 98101, and Department of Pathobiology SC-38, University of Washington, Seattle, Washington 98195

Received February 9, 1988; Revised Manuscript Received April 6, 1988

ABSTRACT: A cDNA encoding the viral protease from the 3C region of human rhinovirus type 14 was expressed in *Escherichia coli* through the use of a periplasmic secretion vector. The recombinant protease contained an eight amino acid N-terminal extension that enabled its detection by a specific antibody. It was expressed at a level of approximately 1 mg/L of *E. coli* culture. Biological activity of the protease was assessed in vitro by using a chemically synthesized peptide consisting of a consensus picornavirus protease cleavage site, Arg-Ala-Glu-Leu-Gln-Gly-Pro-Tyr-Asp-Glu. The peptide was cleaved by the recombinant protease at the Gln-Gly bond, generating the product peptides Arg-Ala-Glu-Leu-Gln and Gly-Pro-Tyr-Asp-Glu, which could be separated from the substrate peptide by reversed-phase HPLC. An in vitro assay for the rhinovirus 3C protease was developed by observing the rate of disappearance of the substrate peak from chromatograms of the supernatants of digestion mixtures.

Human rhinoviruses (HRVs) belong to one genus of the family picornaviridae and are the leading cause of the common cold in man (Stott et al., 1972; Gwaltney, 1975). Other members of the picornavirus family include the entero (polio, echo, coxsachie, hepatitis A), aphtho- [foot and mouth disease (FMDV)], and cardioviruses [mengo, encephalomyocarditis (EMCV)] (Cooper et al., 1978). These latter groups have much in common with the rhinoviruses including their occurrence as icosehedral virions containing a single-stranded RNA genome. The antigenic features on the surfaces of picornaviruses are complex, and among HRVs, 100 serotypes have been identified (Hamparian et al., 1987). The high incidence of the common cold may be explained, at least in part, by this antigenic variability and the cocirculation within the community of several serotypes simultaneously (Gwaltney, 1975). Hence, conventional vaccination programs are virtually useless against the rhinoviruses, making a search for alternative methods for inactivation of these viruses essential.

Controlled proteolysis appears to be crucial to the life cycle of the rhinoviruses, as is evident from the fact that at least two, and probably three, proteolytic enzymes are encoded in the viral genome (Arnold et al., 1987). Like other picornaviruses, the rhinovirus contains a single, positive stranded RNA genome of about 7500 nucleotides (Rossman et al., 1985; Medappa et al., 1971; Dimmock, 1966; Brown et al., 1970; Stanway et al., 1984; Callahan et al., 1985) that encodes a single long open reading frame that is translated into a large precursor protein of more than 200 000 daltons. This polyprotein is processed into the four virion polypeptides VP1, VP2, VP3, and VP4 as well as a number of noncapsid proteins by cleavages made by the viral proteases.

This report deals with the cloning and expression of the cDNA that encodes the rhinovirus 3C protease. A novel expression vector was developed that places an eight amino acid immunoreactive "peptide" extension N-terminal to the 3C protease gene. The peptide extension was found to be useful for partial purification of the recombinant protease. Moreover, we have developed a synthetic substrate, based on the predicted HRV-14 polyprotein cleavage sites (Callahan et al., 1985), for analysis of the recombinant protease activity.

* Author to whom correspondence should be addressed at his present address: Department of Genetics, SK-50, University of Washington, Seattle, WA 98195.
‡ Department of Molecular Biology, The Immunex Corp.
§ University of Washington.
‖ Department of Protein Chemistry, The Immunex Corp.

0006-2960/88/0427-6262$01.50/0 © 1988 American Chemical Society

The work described herein, therefore, will serve as a basis in the future for analyzing compounds that may inhibit the viral protease in vivo.

Materials and Methods

General Methods. Restriction endonucleases, T4 DNA polymerase, T4 polynucleotide kinase, T4 DNA ligase, and the Klenow fragment of DNA polymerase I were obtained from either Bethesda Research Laboratories (Gaithersburg, MD) or New England Biolabs (Beverly, MA), and used according to the manufacturer's recommendations. Oligonucleotides were synthesized on an Applied Biosystems Model 380A DNA synthesizer (Foster City, CA) by the phosphoramidite method (Beaucage & Caruthers, 1981). Oligonucleotides were purified by elution from preparative 18% polyacrylamide gels. Purified oligonucleotides were phosphorylated with T4 polynucleotide kinase. Peptides were synthesized on an Applied Biosystems Model 430A peptide synthesizer by the standard Merrifield solid-phase method (Barany & Merrifield, 1979) with *N*-α-*tert*-butyloxycarbonyl-protected amino acids and standard side-chain protection. Peptides were cleaved from the resin by treatment with hydrofluoric acid and purified by high-performance liquid chromatography (HPLC).

DNA restriction enzyme fragments were resolved by electrophoresis on 0.8% low-gelling agarose (FMC, Rockland, ME). Following electrophoresis, bands containing the appropriate DNA fragments were excised and purified through Elutip columns (Schleicher & Schuell, Needham, NH). Calcium chloride competent *E. coli* strain JM107 or RR1 was prepared and transformed according to established procedures (Maniatis et al., 1982).

Preparation of RNA. Purified HRV-14 was used to infect HeLa cells at a multiplicity of infection of 1.0. After 14–18 h RNA was harvested from the infected cells by standard techniques (Maniatis et al., 1982). Polyadenylated RNA was then selected by oligo(dT)–cellulose chromatography.

Strains and Expression Vectors. E. coli strain JM107 (*Δlac, pro, thi, strA, endA, sbcB15, hspR4/F', traD36, lacI*^{q}*-ZΔM15*) was used as the host for all expression plasmids. The expression vector pIN-III-OmpA$_3$ (Ghrayeb et al., 1984) was kindly provided by Dr. Masayori Inouye (State University of New York, Stony Brook, NY). The pIN-III-OmpA$_3$ vector (Figure 2) is transcriptionally driven by the tandem lipoprotein promoter (*lpp*^{p}) and lactose promoter–operator (*lac*^{po}) regions of *E. coli*. Regulation of the tandem *lpp*^{p}/*lac*^{po} is controlled by the lactose repressor, encoded by the *lacI* gene, cloned elsewhere on the plasmid. Expression of cloned genes is induced by addition of isopropyl β-D-thiogalactoside (IPTG). The signal peptide encoding sequence of the *OmpA* gene is located distal to the *lpp*^{p}/*lac*^{po} transcriptional regulatory unit and serves to direct the cloned gene product out of the cytoplasm. The unique restriction sites *Eco*RI, *Hin*dIII, and *Bam*HI, located downstream of the *OmpA* gene signal sequence, serve as sites for insertion of foreign genes. The vector pRL 6-6-85 is a derivative of pIN-III-OmpA$_3$ in which the *Xba*I/*Hin*dIII fragment encompassing the *OmpA* sequence and multiple restriction sites was replaced with a synthetic oligonucleotide. The synthetic oligonucleotide (80-mer) reconstructed the *OmpA* signal regions (amino acids 1–20) and placed a *Hin*dIII restriction site at the 3′ end of the *OmpA* signal (Figure 2). The use of pRL 6-6-85 as an expression vector fuses the foreign gene directly to the *OmpA* signal peptide. Thus, pRL 6-6-85 results in direct expression of the HRV-14 protease without supernumerary amino acids at the N-terminus, once the signal peptide is removed.

Construction of Double-Stranded cDNA. Double-stranded cDNA was synthesized from rhinovirus-infected cell mRNA by standard procedures (Barany & Merrifield, 1979; March et al., 1985) with the following modifications. First, strand cDNA synthesis was primed with an oligonucleotide complementary to the 3′ end of the 3C protease gene. This oligonucleotide (RV-71, Table I) was used at a concentration of 5 μg/mL. Following first and second strand cDNA synthesis (Gubler & Hoffman, 1983), oligo(dC), tails were added by terminal transferase and the cDNA was annealed to oligo(dG)-tailed *Pst*I-cut pBR322. The annealed DNA was transformed into *E. coli* strain RR1.

Screening of Positive Clones. Several thousand tetracycline-resistant colonies were screened by colony hybridization with an oligonucleotide probe specific for the 5′ end of the protease gene (RV-21) (Table I). Plasmid DNAs from positive colonies were isolated, restriction endonuclease mapped, and sequenced by standard techniques (Barany & Merrifield, 1979; Sanger et al., 1977).

Assembly of the Complete HRV-14 Protease Gene. First, a 251 base pair (bp) *Ava*II/*Eco*RV DNA fragment encompassing the 5′ end of the protease gene was purified from one of the positive clones above. This fragment was inserted into pBR322 restricted at the *Cla*I and *Eco*RV sites by using synthetic complementary oligonucleotides (RV-5′/24/25, Table I) generating the vector pRV-5′. A vector containing the complete 3′ end of the protease gene was constructed by cloning a 211 bp *Eco*RV/*Bgl*II (partial) fragment into pBR322 restricted at the *Eco*RV and *Hin*dIII sites through the aid of complementary oligonucleotides (RV-3′/493/497, Table I). The resulting vector, pRV-3′, thus provides the 3′ end of the protease gene and a translational termination codon. Final assembly of the complete gene was accomplished by subcloning the 277 bp *Cla*I/*Eco*RV fragment from pRV-5′ and the 300 bp *Eco*RV/*Hin*dIII fragment from pRV-3′ into pGB34 restricted at the *Cla*I and *Hin*dIII site. The fully assembled vector, pGB34-RV, thus served as the source of the complete protease gene.

Construction of the HRV-14 Protease Expression Vector. The plasmid pGB34-RV, containing the complete HRV-14 3C protease gene, was restricted with *Ava*II and *Hin*dIII. The resulting 560 bp fragment was inserted through the aid of synthetic oligonucleotides (RV-17, Table I) into the pIN-III-OmpA$_3$ expression vector at the *Eco*RI and *Hin*dIII sites. The resulting construct, pJL 7-27-81 (Figure 2), was transformed into *E. coli* strain JM107 (containing *lacI*^{q} on the F′) and expression of the protease gene induced following addition of 2 mM IPTG. Samples were prepared according to the procedure of Laemmli (1977) for electrophoresis on polyacrylamide gels. The final expression vector designed to simplify the detection and purification of the protease was constructed as follows: Plasmid pJL-27-81 was restricted at the *Ava*II and *Sal*I sites, and the 1.4 kbp fragment containing the protease gene was purified as previously described. Next, plasmid pRL 6-6-85 (Figure 2) was restricted with *Hin*dIII and *Sal*I and the vector fragment ligated to the *Ava*II/*Sal*I fragment (above) through the aid of an oligonucleotide duplex (RV-29, Table I) which specifies the octapeptide sequence Asp-Tyr-Lys-(Asp)$_4$-Lys fused to the N-terminus of the protease gene. The resulting construct, pKL 6-29-86 (Figure 2), specifies a periplasmically secreted form of the rhinovirus 3C protease containing an eight amino acid N-terminal peptide extension.

Preparation of Anti-Peptide Monoclonal Antisera. The anti-peptide hybridoma was produced according to established

procedures (Gillis & Henney, 1981) using, as immunogen, human interleukin 2 bearing an N-terminal Asp-Tyr-Lys-$(Asp)_4$-Lys extension. This antibody has been shown to react exclusively with the N-terminus of proteins bearing the peptide sequence and does not react with extracts derived from uninduced cultures of *E. coli* containing the expression vector pKL 6-29-86 (details to be published elsewhere).

Partial Purification of Recombinant HRV-14 3C Protease. Pelleted *E. coli* cells containing periplasmic 3C protease were lysed by suspending the cell pellet from 500 mL of culture in 5 mL of phosphate-buffered saline (PBS) containing lysozyme (0.5 mg/mL) and 5 mM EDTA, freezing and thawing three times, and Dounce homogenization. After centrifugation at 25000*g* for 45 min, the pellets were homogenized in an additional 20 mL of PBS and centrifugation was repeated. Finally, the proteases, which remained with the particulate fraction, was resuspended for assay by homogenization in 50 mL of buffer containing 20 mM Tris, 3 mM magnesium acetate, 150 mM potassium acetate, and 2 mM dithiothreitol and adjusted to pH 7.6 with glacial acetic acid. Proteolytic activity was stable in this buffer when frozen at −70 °C for up to 1 year.

Protease Digestions. Proteolysis of the synthetic substrate peptide was initiated by mixing 0.5 mL of the protease suspension with 0.5 mL of a solution containing substrate peptide at 0.4 mg/mL and bestatin, an amino peptidase inhibitor, at 2 mg/mL in the same buffer used to resuspend the protease. After a 0-min aliquot was removed, the mixture was incubated at 37 °C and aliquots were removed at regular intervals thereafter. Each 100-μL aliquot was mixed with 100 μL of 10% (w/v) trichloroacetic acid in order to stop the digestion and to precipitate the protease and other *E. coli* solids. Following centrifugation (3 min, Eppendorf) samples of the supernatant were sufficiently clarified to be injected onto the HPLC.

Other Procedures. Polyacrylamide gel electrophoresis was carried out on 10–20% polyacrylamide gradient gels (Laemmli, 1977). Western transfer were performed as previously described (Towbin et al., 1979), using a 1:100 dilution of purified anti-peptide monoclonal antisera (unpublished experiments) as the primary antibody, and horseradish peroxidase conjugated goat anti-mouse IgG (Bio-Rad Laboratories) as the secondary antibody. Peptide sequence analysis was performed on an Applied Biosystems Model 470A gas-phase sequencer (March et al., 1985) using a Hewlett-Packard 1090 HPLC and a diisopropylethylamine-based chromatogram for PTH-amino acid detection. HPLC analysis of digests was performed on a Brownlee Aquapore RP300 reversed-phase column (2.1 × 220 mm) eluted at 0.8 mL/min with 0.1 (v/v) trifluoroacetic acid and a 2–35% (v/v) gradient of acetonitrile in 35 min.

RESULTS

Cloning of the HRV-14 Protease Gene. Messenger RNA was isolated from HRV-14 infected cells and used to construct a cDNA library in pBR322. First strand cDNA synthesis was primed with an HRV-14-specific oligonucleotide (RV-71) from the 3′ end of the 3C protease gene. However, of five transformants that hybridized to the RV-21 oligonucleotide (Table I), which is specific for the 5′ end of the protease gene, none contained an exact copy of the RV-71-specific sequence. DNA sequence analysis of these five cDNA clones showed that the body of the 3C protease gene matched exactly with that previously published, but that there were several deletions and point mutations at the 3′ end of the gene (data not shown). Accordingly, we decided to assemble the complete gene from one of the positive clones above with the use of synthetic oligonucleotides to reconstruct the 3′ end of the gene as outlined in Figure 1. The resulting fully assembled vector, pGB34-RV (Figure 1), thus contained the complete HRV-14 protease gene and was used as the source of DNA for construction of an expression vehicle.

Table I: Oligonucleotides Used in Constructions and as Probes

Oligonucleotide	Sequence
RV-71 (3′-end primer)	5′ CCTTGTTTCTCTACAAAATATTGTTTTTTAAGTTG AGCTGAAAATCCTTGTCTTCCATTACCGCCAACATG
RV-21 (5′end probe)	5′ GGACCAAACACAGAATTTGCA
RV-5′/24/25	5′ CGATAACACAGGAACAGATCTATG TATTGTGTCCTTGTCTAGATACCTG
RV-3′/493/497	5′ GATCTTTGGT.....AAACAATA AAACCA.....TTTGTTATTCGA
RV-17	5′ AATTCCAGTTGTGCAAG GGTCAACACGTTCCTG
RV-29	5′ AGCTGACTACAAAGACGATGACGATAAAG CTGATGTTTCTGCTACTGCTATTTCCTG

[a] RV-71 is complementary to nucleotides 5717–5787 of HRV-14 (Stanway et al., 1984), and RV-21 represents nucleotides 5240–5260 of HRV-14. RV-5′/24/25 provide a translational start codon (overlined) for the protease gene. RV-3′/493/497 encode the 3′ end of the protease gene from positions 5714 to 5785, followed by a translational termination codon (overlined). RV-17 represents nucleotides 5224–5243 of HRV-14. RV-29 specifies the octapeptide sequence [i.e., Asp-Tyr-Lys-$(Asp)_4$-Lys] containing a 3′ overhang compatable with the *Ava*II site at position 5243 of HRV-14.

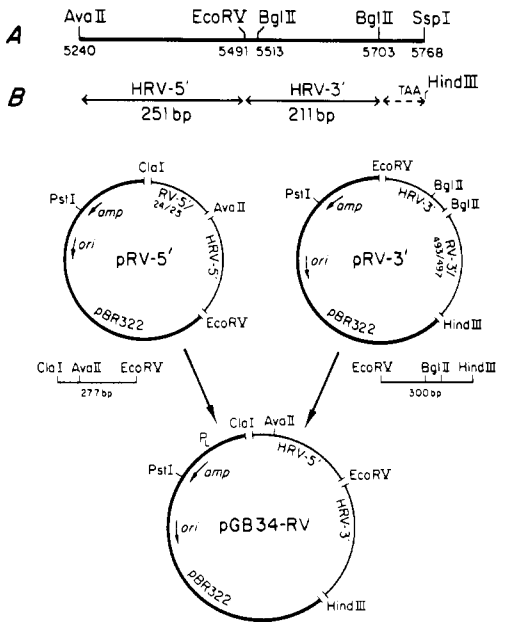

FIGURE 1: Schematic diagram detailing the cloning of the HRV-14 3C (protease) gene. The topmost line (A) represents a restriction map of the cDNA encoding the HRV-14 protease gene. Numbers below the line represent map positions in the 3C gene as previously described. (Stanway et al., 1984). The bottom line (B) represents the reconstructed gene using the 251 bp *Ava*II/*Eco*RV fragment representing the 5′ end of the protease gene (HRV-5′) or the 211 bp *Eco*RV/*Bgl*II fragment representing most of the 3′ end of the protease gene (HRV-3′). The broken line represents the extreme 3′ portion of the protease gene specified by the synthetic oligonucleotides RV-3′-493/497 containing a translational stop codon (TAA). The complete protease gene was constructed from pRV-5′, pRV-3′, and RV-3-493/497 and assembled into pGB34 as described under Materials and Methods.

Expression of the 3C Protease Gene. As expression of many toxic enzymes such as nucleases and proteases cannot be carried out in the cytoplasm of *E. coli*, we sought to guide the

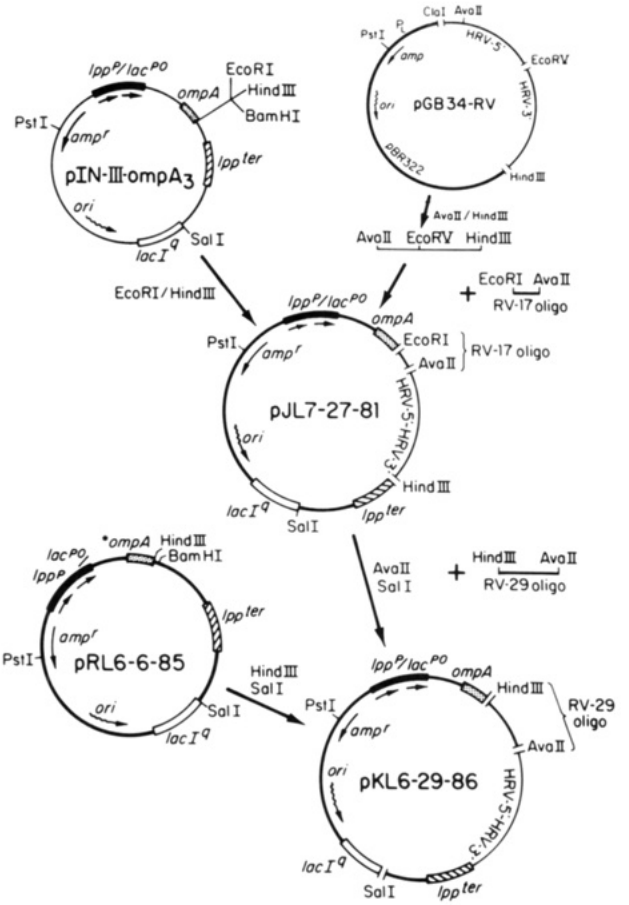

FIGURE 2: Schematic diagram detailing the construction of the 3C protease expression vectors as described under Materials and Methods. Abbreviations are *lpp*^{p}, lipoprotein promoter; *lac*^{po}, lactose promoter–operator; OmpA, outer-membrane protein A signal sequence; *OmpA, synthetic outer-membrane protein A signal sequence containing a *Hin*dIII site at the signal processing site; *lpp*^{ter}, lipoprotein terminator; *lac*^{Iq}, lactose repressor; and *ori*, origin of replication.

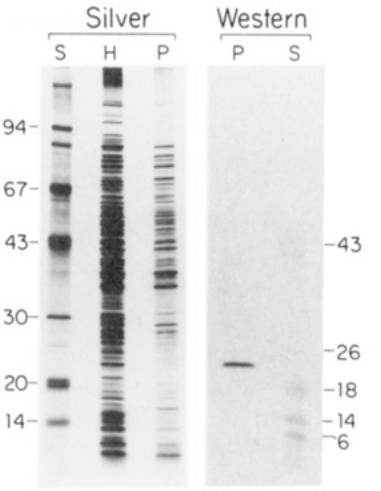

FIGURE 3: Polyacrylamide gel analysis of *E. coli* product. Left panel: Silver-stained gel of whole *E. coli* homogenate (H) and washed pellet (P). Right panel: Western transfer showing the immunoreactive peptide band present in the washed pellet (P). Extracts from lanes H and P, left panel, were prepared from the same number of cells. Molecular weights of standard proteins (S) are shown in the margins.

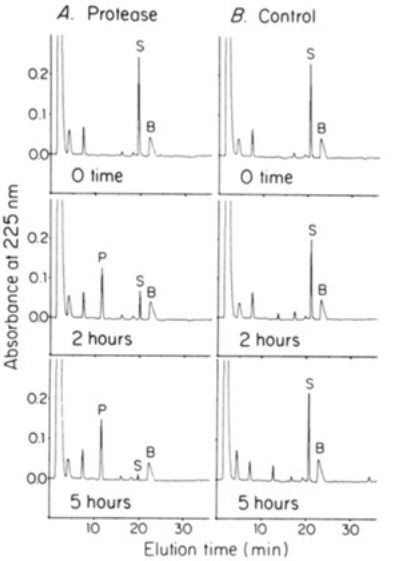

FIGURE 4: HPLC analysis of protease activity. (A) Incubation of substrate peptide (S) in the presence of the marker peptide–3C protease containing pellet. (B) Incubation in the presence of a pellet from the GM-CSF expression culture. Bestatin (B) serves as an internal standard. The products of hydrolysis were found in the region marked by the letter P. Absorbance was monitored at 225 nm.

protease out of the cytoplasm. Thus, expression was accomplished through the use of an outer-membrane protein OmpA signal peptide based *E. coli* secretion vector, pIN-III-OmpA$_3$ (Ghrayeb et al., 1984). The initial HRV-14 3C protease expression vector, pJL 7-27-81, was constructed as outlined in Figure 2. Following transformation of pJL 7-27-81 into *E. coli* strain JM107, the protease gene was induced following addition of IPTG. However, the level of recombinant protease specified by pJL 7-27-81 was relatively low despite attempts to express the recombinant protease in alternate strains of *E. coli*. Thus, detection of the recombinant protein from the background of host proteins was difficult. We therefore utilized a marker peptide expression system that was designed to enhance detection and subsequently simplify purification of the protease. Plasmid pJL 7-27-81 was reengineered as described under Materials and Methods, to yield pKL 6-29-86, which specified periplasmic production of the rhinovirus protease containing the eight amino acid N-terminal "peptide" extension. The protease specified by *E. coli* strain JM107 harboring plasmid pKL 6-29-86 was induced as described above, and peptide-reactive proteins were deteted following Western transfer to nitrocellulose and treatment with an anti-peptide monoclonal antibody as described under Materials and Methods. As demonstrated in Figure 3, a major immunoreactive band of approximately 21 000 Da representing the recombinant marker peptide–protease was clearly observed following Western blotting of the induced culture (Figure 3, right panel), while the same band was not observed following immunoblotting of an uninduced extract of *E. coli* (data not shown). Silver staining of extracts from induced cultures, on the other hand, did not produce a readily apparent band (Figure 3, left panel). However, by comparison to standard marker peptide–proteins on other Western transfers, we were able to estimate the amount of peptide–3C protease produced at approximately 1 mg/L of *E. coli* culture.

Partial Purification of the Flag–Protease. Because the 3C protease was localized in the membrane fraction of lysed cells of *E. coli*, it was difficult to obtain fully purified preparations, due to loss of material that remained associated with centrifuged membrane pellets. Extractions with several detergents were not successful in solubilizing protease activity. However, we found that the extraction procedure described under Materials and Methods was capable of removing all contaminating *E. coli* endopeptidase activities while leaving the active 3C protease in the membrane pellet. Resuspension of washed

Table II: Derivation of a Consensus 3C Protease Substrate from Proposed Cleavage Sites Found in Picornaviral Polyproteins[c]

HRV-14 Proteins[a]	Cleavage Site									
1A/1B	-Ala-	Pro-	Ala-	Leu-	Asn/	Ser-	Pro-	Asn-	Val-	Glu-
1B/1C	-Ser-	Ile-	Val-	Pro-	Gln/	Gly-	Leu-	Pro-	Thr-	Thr-
1C/1D	-Val-	Ala-	Leu-	Thr-	Glu/	Gly-	Leu-	Gly-	Asp-	Glu-
2A/2B	-Ile-	Ala-	Glu-	Glu-	Gln/	Gly-	Leu-	Ser-	Asp-	Tyr-
2B/2C	-Tyr-	Ile-	Glu-	Arg-	Gln/	Ala-	Asn-	Asp-	Gly-	Trp-
2C/3A	-Glu-	Thr-	Leu-	Phe-	Gln/	Gly-	Pro-	Val-	Tyr-	Lys-
3A/3B	-Phe-	Ala-	Gln-	Thr-	Gln/	Gly-	Pro-	Tyr-	Ser-	Gly-
3B/3C	-Pro-	Val-	Val-	Val-	Gln/	Gly-	Pro-	Asn-	Thr-	Glu-
3C/3D	-Phe-	Val-	Glu-	Lys-	Gln/	Gly-	Glu-	Val-	Ile-	Ala-
Consensus	-Phe-	Ala-	Glu-	Leu-	Gln/	Gly-	Pro-	Tyr-	Asp-	Glu-
Occurrences[b]	6	11	5	5	30	25	17	4	4	7
Substrate	Arg-	Ala-	Glu-	Leu-	Gln-	Gly-	Pro-	Tyr-	Asp-	Glu
Position	P5	P4	P3	P2	P1	P1'	P2'	P3'	P4'	P5'

[a] The numbering scheme (e.g., 1A/1B) indicates that the viral protein 1A is separated from viral protein 1B by cleavage of the bond indicated by the slash (e.g., Asn/Ser) in the sequence at right. [b] The number of times that the indicated amino acid occurred at that position in the polyproteins of HRV-14, HRV-2, encephalomyocarditis virus, and poliovirus is indicated. [c] Amino acid sequences represent a complete list of HRV-14 polyprotein sites. Cleavage sites are those proposed by Callahan [see Callahan et al. (1985)].

pellets by Dounce homogenization allowed an adequate rate of substrate hydrolysis for subsequent assay development. As can be seen in Figure 4, no endopeptidase activity was expressed in comparable washed pellets from *E. coli* bearing an irrelevant plasmid that was induced identically to the pKL 6-29-86 transformed *E. coli* strain. Traces of aminopeptidase activity capable of degrading the substrate peptide were eliminated by the inclusion of bestatin in the reaction mixture.

Synthesis of a Consensus 3C Protease Substrate. The partially purified recombinant marker peptide–3C protease was used to identify an appropriate substrate for an in vitro protease assay. It has been recognized that the rhinovirus protease cleaves peptide bonds between glutamine and glycine residues (Callahan et al., 1985). Accordingly, we designed a synthetic substrate that incorporated these residues. To assure that the chemically synthesized substrate would be optimal for use in an assay, it was desirable to include residues on either side of the *Gln-Gly* pair that would be conducive to a high rate of cleavage. As it seems likely that such conducive residues would occur more frequently than nonconducive residues in the predicted cleavage regions of the polyproteins of a number of picornaviruses, we decided to use a consensus of the surrounding amino acids as illustrated in Table II. In addition to the HRV-14 cleavage sites shown, we also considered cleavage sites from rhinovirus type 2 (Skern et al., 1985), encephalomyocarditis virus (Palmenberg et al., 1984), and poliovirus type 1 (Kitamura et al., 1981), giving a total of 33 cleavage sites. As expected, *Gln* and *Gly* occurred with great frequency in the P1 and P1′ positions on either side of the cleavage site, and *Pro* occurred frequently at position P2′. All other positions contained an essentially random set of amino acids, with the possible exception of position P4 where *Ala* occurred 11 times. The consensus sequence shown in Table II was used in the chemical substrate synthesis, with one alteration. The *Phe* at position P5 was replaced with *Arg* in order to increase the likelihood that the peptide would be soluble in aqueous media, including the TCA supernatants obtained in the assay. In order to determine if the synthetic substrate is specifically cleaved, we partially purified the marker peptide–3C protease (as described below) and tested it in our proteolytic activity assay.

Proteolytic Activity Assay. Figure 4 demonstrates the results obtained when the synthetic substrate was utilized in a typical 3C protease activity assay. The substrate peptide peak detected by HPLC decreases in size over time during incubation with the partially purified marker peptide–protease (as observed from Figure 3, right panel, lane P), while a peak appears over time that represents the two product peptides which coelute with trifluoroacetic acid and a gradient of acetonitrile as described under Materials and Methods. We have confirmed the authenticity of the site of cleavage by isolating and sequencing the components of the product peak and found the two expected sequences *Arg-Ala-Glu-Leu-Gln* and *Gly-Pro-Tyr-Asp-Glu*, confirming that only the *Gln-Gly* bond was cleaved. Integration of the peak area of the substrate and/or product peptides allows for estimation of protease activity. However, the presence of an active 3C protease is usually apparent visually from the dramatic change in peak size. Several small peaks appear in the control digests overtime, but these do not interfere with either the substrate peak or the product peak. In control digest, 500 μL of homogenized pellet from cultures expressing the human granulocyte–macrophage colony stimulating factor (GM-CSF) (Libby et al., 1987) was similarly incubated with the substrate. As observed in Figure 4 (panel B), no peak appears at the product elution time in the control digest. Thus the digestion products observed in the presence of the recombinant protease are not due to host- or vector-related proteases produced following induction with IPTG. Moreover, bestatin is unaltered during the digestion and is soluble in the TCA supernatants. This provides a convenient internal standard because the bestatin peak remains constant as the digestion progresses and therefore serves as a control for sample handling and injection variability.

Discussion

In order to study the rhinovirus 3C protease effectively, it is necessary to obtain quantities of the molecule that permit a variety of experiments to be carried out. Furthermore, it is necessary to have a supply of the protease that is reproducible and more abundant than the quantities that have been obtained in the past from virally infected cells or by in vitro synthesis in reticulocyte lysates. We initially attempted to express the 3C protease in vectors lacking bacterial signal sequences and found that the protease was rapidly degraded in the cytoplasm of *E. coli* (data not shown). We have overcome this instability through use of the OmpA signal sequence from the *E. coli* secretion vector, pIN-III-OmpA$_3$ (Ghrayeb et al., 1984). Thus, fusing the OmpA signal sequence to the N-termini of the 3C protease cDNA served to redirect the viral protease from the cytoplasm to the outer membrane, a region known to be low in host proteases. Hence, expression of the protease from plasmid pJL 7-27-81 (Figure 2) was stable and was found to specifically cleave the synthetic substrate at the predicted Gln/Gly pair (data not shown). However, as a result of the low level of expression from plasmid pJL 7-27-81, detection of the 3C protease from the background of host proteins was difficult (see Figure 3, left panel). We, therefore, constructed plasmid pKL 6-29-86 (Figure 2) which fuses an eight amino acid immunoreactive "marker" peptide onto the N-terminal end of the protease gene (just downstream of the OmpA signal sequence). As such, the marker peptide–3C protease is stably expressed and was found to cleave our synthetic substrate at the predicted *Gln–Gly* pair, as did the non-peptide-bearing version of the recombinant 3C protease specified by plasmid pJL 7-27-81. Thus, fusion of an eight amino acid peptide to the N-termini of the 3C protease

did not interfere with its ability to specifically cleave the substrate and was found to be useful in its purification. The marker peptide–protease, thus, will serve as source material for future studies aimed at obtaining homogenous, active 3C protease in quantity. It is interesting to note that we have only been able to demonstrate activity in the protease when it is associated with the membrane fraction of *E. coli* extracts. Attempts to obtain soluble protease activity by detergent extraction and other means have not yet been successful. It would appear either that the activity of the 3C protease is very sensitive to inactivating conditions or that activation of the protease may be dependent upon an association with membranes which may be a natural property of the protease, as has been previously suggested for this protease (Nicklin et al., 1986) and for poliovirus 2A protease (Toyoda et al., 1986). This also implies that the picornavirus life cycle may require that proteolytic processing of the viral polyprotein be accomplished at a specific location inside the infected cells, a location that is determined by a membrane-bounded compartment for which the 3C protease has an affinity.

Another important aspect of this work is that this is the first reported instance of a picornavirus protease acting on a small substrate rather than its usual large polyprotein target. The availability of a small chemically synthesized substrate now makes it much more convenient to do repetitive assays in order to study the activity of the 3C protease. Large quantities of this substrate are readily prepared by the Merrifield procedure, eliminating the need for substrates prepared from virally infected cells or reticulocyte lysates, and eliminating the need for handling radioactive products that are typically obtained from these systems. A further advantage of a chemical synthesis approach is that new analogues can be produced at a rate of approximately one per day, which is a great improvement over the rate of new analogue production that could be expected when using in vitro mutagenesis approaches on the polyprotein. This should lead to a rapid elucidation of the critical portions of the substrate for interaction with the protease. A further advantage is that chemical synthesis is not limited to the 20 usual amino acids but can also incorporate D-amino acids, α-carbon-substituted amino acids, and other nonphysiological moieties in order to test the specificity determinants of the 3C protease. This ability will also aid in a preliminary search for inhibitor substance as well. Such a capability, combined with the ability to make large batches of active protease and to store the product as stable aliquots at −70 °C, should accelerate our progress toward understanding the biological role of the 3C protease.

The amino-terminal marker peptide sequence has played an important part in this work because it has several unique properties. It is extremely hydrophilic (Hopp, 1986) and therefore does not engage in unfavorable folding interactions with the rest of the molecule that might lead to denaturation as has been seen for many other fusion protein products (Schoner et al., 1985). Because an antibody is available that reacts specifically with the marker peptide sequence, it is relatively simple to detect and quantify expression products, even when they are not expressed in high levels, as in this case. Finally, the marker peptide sequence is compatible with cleavage by signal peptidase to remove the OmpA signal peptide, a necessary step on the secretion pathway. In addition, it is likely that the antibody to the marker peptide will be of use in further efforts to purify the protease to homogeneity once a means has been found to extract it in an active form.

ACKNOWLEDGMENTS

We thank Masayori Inouye (State University of New York at Stony Brook) for use of the pIN-III-OmpA_3 secretion vector, Paul Conlon (Immunex) for supplying the anti-peptide monoclonal antibody, and Gary Braedt for use of plasmid pGB-34. Teresa Chiaverotti (University of Washington) and Janis Wignall (Immunex) provided excellent technical assistance. This work was completed in part by R.T.L. at the University of Washington. We are indebted to Peggy Green and Linda Troup for preparation of the manuscript.

Registry No. Protease 3C, 115203-96-8.

REFERENCES

Arnold, E. Luo, M., Vriend, G., Rossman, M. G., Palmenberg, A. C., Parks, G. D., Nicklin, M. J. H., & Winmer, E. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 21–25.

Barany, G., & Merrifield, R. B. (1979) in *Peptides: Analysis, Synthesis, Biology* (Gross, E., & Meienhofer, J., Eds.) pp 3–254, Academic, New York.

Beaucage, S. L., & Caruthers, M. H. (1981) *Tetrahedron Lett. 22*, 1859–1863.

Brown, F., Newman, J. F. E., & Stott, E. J. (1970) *J. Gen. Virol. 8*, 145–148.

Callahan, P. L. Mizutani, S., & Colonno, R. J. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 732–736.

Cooper, P. D., Agol, V. I., Bachrach, H. L., Brown, F., Gheudon, Y., Gibbs, A. J., Gillespie, J. H., Lonberg-Holm, K., Melnick, J. L., Mohanty, S. B., Povey, R. C., Rueckert, R. R., Schaffer, F. L., & Tyrrell, D. A. J. (1978) *Intervirology 10*, 165–180.

Dimmock, N. J. (1966) *Nature (London) 209*, 792–794.

Ghrayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y., & Inouye, M. (1984) *EMBO J. 3*, 2437–2442.

Gillis, S., Henney, C. S. (1981) *J. Immunol. 126*, 1978–1984.

Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.

Gwaltney, J. M. (1975) *Yale J. Biol. Med. 48*, 17–45.

Hamparian, V. V., Colonno, R. J., Cooney, M. K., Dick, E. C., Gwaltney, J. M., Jr., Hughes, J. H., Jordan, W. S., Kapikian, A. Z., Mogabgab, W. J., Monto, A., Phillips, C. A., Rueckert, R. R., Schieble, J. H., Stott, E. J., & Tyrrell, D. A. J. *Virology 159*, 191–192.

Hopp, T. P. (1986) *J. Immunol. Methods 88*, 1–18.

Kitamura, N., Semler, B. L., Rothberg, P. G., Larsen, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., Hanecak, R., Lee, J. L., vanderWerf, S., Anderson, C. W., & Wimmer, E. (1981) *Nature (London) 291*, 547–553.

Laemmli, U. K. (1977) *Nature (London) 227*, 680–685.

Libby, R. T., Braedt, G., Kronheim, S. R., March, C. J., Urdal, D. L., Chiaverotti, T. A., Tushinski, R. J., Mochizuki, D. Y., Hopp, T. P., & Cosman, D. (1987) *DNA 6*, 221–229.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning, Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

March, C. J., Mosley, B., Larsen, A., Cerretti, D. P., Braedt, G., Price, V., Gillis, S., Henney, C. S., Kronheim, S. R., Grabstein, K., Conlon, P. J., Hopp, T. P., & Cosman, D. (1985) *Nature (London) 315*, 641–646.

Medappa, K. C., McLean, C., & Rueckert, R. R. (1971) *Virology 44*, 259–270.

Nicklin, M. J. H., Toyoda, H., Murray, M. G., & Wimmer, E. (1986) *Bio/Technology 4*, 33–42.

Palmenberg, A. C., Kirby, E. M., Janda, M. R., Drake, N. L., Duke, G. M., Potratz, K. F., & Collett, M. S. (1984) *Nucleic Acids Res. 12*, 2969–2985.

Rossman, M. G., Arnold, E., Erickson, J. W., Frankenberger, E. A., Griffith, J. P., Hecht, H.-J., Johnson, J. E., Kamer, G., Luo, M., Mosser, A. G., Rueckert, R. R., Sherry, B.,

& Vriend, G. (1985) *Nature (London) 317*, 145-153.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Schoner, R. G., Ellis, L. F., & Schoner, B. E. (1985) *Bio/Technology 3*, 151-154.
Skern, T., Sommergruber, W., Blaas, D. Gruendler, P., Fraundorfer, F., Pieler, C., Fogg, I., & Kuechler, E. (1985) *Nucleic Acids Res. 13*, 2111-2126.
Stanway, G., Hughes, P. J., Mountford, R. C., Minor, P. D., & Almond, J. W. (1984) *Nucleic Acids Res. 12*, 7859-7875.
Stott, E. J., & Killington, R. A. (1972) *Annu. Rev. Microbiol. 26*, 503-524.
Tessier, L.-H., Sondermeyer, P., Faure, T., Dreyer, D., Benavente, A., Villeval, D., Courtney, M., & Lecocq, J.-P. (1984) *Nucleic Acids Res. 12*, 7663-7675.
Towbin, H., Staehelin, T., & Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4350-4354.
Toyoda, H., Nicklin, M. J. H., Murray, M. G., Anderson, C. W., Dunn, J. J., Studier, F. W., & Wimmer, E. (1986) *Cell (Cambridge, Mass.) 45*, 761-770.

Placental Anticoagulant Proteins: Isolation and Comparative Characterization of Four Members of the Lipocortin Family†

Jonathan F. Tait,‡,§ Mitsuhiko Sakata,‡ Brad A. McMullen,‡ Carol H. Miao,‡ Takayuki Funakoshi,‡ Lee E. Hendrickson,‡ and Kazuo Fujikawa*,‡

Departments of Biochemistry and Laboratory Medicine, University of Washington, Seattle, Washington 98195

Received February 29, 1988; Revised Manuscript Received April 18, 1988

ABSTRACT: Previously we isolated and characterized a placental anticoagulant protein (PAP or PAP-I), which is a Ca^{2+}-dependent phospholipid binding protein [Funakoshi et al. (1987) *Biochemistry 26*, 5572] and a member of the lipocortin family [Funakoshi et al. (1987) *Biochemistry 26*, 8087]. In this study, three additional anticoagulant proteins (PAP-II, PAP-III, and PAP-IV) were simultaneously isolated from human placental homogenates prepared in the presence of 5 mM ethylenediaminetetraacetic acid. The isoelectric points of PAP-I, PAP-II, PAP-III, and PAP-IV were 4.8, 6.1, 5.9, and 8.1, respectively, and their apparent molecular weights were 32 000, 33 000, 34 000, and 34 500, respectively. Amino acid sequences of cyanogen bromide fragments of these proteins showed that PAP-III was a previously unrecognized member of the lipocortin family, while PAP-II was probably the human homologue of porcine protein II and PAP-IV was a derivative of lipocortin II truncated near the amino terminus. Comparative studies showed that all four proteins inhibited blood clotting and phospholipase A_2 activity with potencies consistent with their measured relative affinities for anionic phospholipid vesicles. However, PAP-IV bound to phospholipid vesicles approximately 160-fold more weakly than PAP-I, while PAP-II and PAP-III bound only 2-fold and 3-fold more weakly. These results increase to six the number of lipocortin-like proteins known to exist in human placenta. The observed differences in phospholipid binding may indicate functional differences among the members of the lipocortin family despite their considerable structural similarities.

Phospholipid surfaces are essential components in several key reactions of the coagulation cascade (Mann, 1984; Zwaal et al., 1986). As such, they represent a potential regulatory point for physiological or pharmacological control of blood coagulation. Recently, we described the identification and purification of a candidate anticoagulant protein from human placenta termed placental anticoagulant protein (PAP or PAP-I)[1] (Funakoshi et al., 1987a). This protein inhibits the extrinsic and intrinsic pathways of blood coagulation and binds specifically to anionic phospholipid surfaces in the presence of Ca^{2+} (Funakoshi et al., 1987a; Kondo et al., 1987). Iwasaki et al. (1987) have also purified and characterized the same anticoagulant protein (termed inhibitor of blood coagulation) from human placenta, and Reutelingsperger et al. (1985) have isolated a similar protein (vascular anticoagulant) from human umbilical cord arteries. This protein has also been purified from human placenta during studies of potential substrates of the epidermal growth factor receptor/kinase, and it has been called endonexin II (Haigler et al., 1987; Schlaepfer et al., 1987).

Protein and cDNA sequence data (Funakoshi et al., 1987a,b; Iwasaki et al., 1987; Schlaepfer et al., 1987) show that PAP-I is a member of a recently described family of Ca^{2+}-dependent phospholipid binding proteins variously termed lipocortins (Wallner et al., 1986; Huang et al., 1986), calpactins (Saris et al., 1986; Kristensen et al., 1986; Glenney, 1986a), proteins I, II, and III (Gerke & Weber, 1984; Shadle et al., 1985), calelectrins (Walker et al., 1983; Sudhof et al., 1984, 1988), annexins (Geisow, 1986), p35 and p36 (Fava & Cohen, 1984; De et al., 1986; Gerke & Weber, 1984; Glenney & Tack, 1985), chromobindins (Creutz et al., 1987), or calcimedins (Moore & Dedman, 1982; Smith & Dedman, 1986). The physiological roles of these proteins are presently unknown. They have been proposed to participate in membrane fusion and exocytosis (Creutz, 1981; Geisow & Burgoyne, 1982; Sudhof et al., 1982), cytoskeleton–membrane linkage (Walker,

† Supported by NIH Grant HL 16919.
‡ Department of Biochemistry.
§ Department of Laboratory Medicine.

[1] Abbreviations: C_6-NBD-PC, 1-palmitoyl-2-[[*N*-(7-nitro-2,1,3-benzoxadiazol-4-yl)amino]caproyl]phosphatidylcholine; EDTA, ethylenediaminetetraacetic acid; FITC, fluorescein-5-isothiocyanate; FPLC, fast protein liquid chromatography; HEPES, *N*-(2-hydroxyethyl)-piperazine-*N'*-2-ethanesulfonic acid; HPLC, high-performance liquid chromatography; IC_{50}, concentration of protein causing 50% inhibition of binding; PAP, placental anticoagulant protein; PC, phosphatidylcholine; PS, phosphatidylserine; SDS, sodium dodecyl sulfate.

0006-2960/88/0427-6268$01.50/0 © 1988 American Chemical Society