Separation of native and truncated forms of poliovirus protease 3C produced in *Escherichia coli*

László POLGÁR,*¶ Ferenc ERDÉLYI,†‖ Éva HAJNAL,† Miklós LÖW‡ László GRÁF,† and Bruce D. KORANT§

* Institute of Enzymology, Biological Research Center, Hungarian Academy of Sciences, P.O. Box 7, H-1518 Budapest, Hungary;† Department of Biochemistry, L. Eötvös University, Puskin u. 3, H-1088 Budapest, Hungary;‡ G. Richter Chemical Works Ltd., P.O. Box 27, H-1475 Budapest 10, Hungary; and § Du Pont–Merck Pharmaceutical Co., Du Pont Experimental Station, Wilmington, DE 19880-0328, U.S.A.

Poliovirus protease 3C is a cysteine enzyme that is essential for the processing of the viral precursor polyprotein containing structural proteins and enzymes, including the protease itself. We have constructed the plasmid pSD/PV3C which produced protease 3C as inclusion bodies when expressed in *Escherichia coli*. In addition to the full-length protease, a truncated form was also generated, starting from an internal initiation site (Met-27). The enzyme was renatured by dilution of a 6 M guanidinium chloride solution of the inclusion bodies, and the proteins were precipitated from the diluted solution with ammonium sulphate. By extracting the precipitate with a buffer solution, the full-length enzyme could be completely separated from its N-terminally truncated form. Size-exclusion chromatography of the extracted protease 3C resulted in an active enzyme which appeared homogeneous by SDS/PAGE. For measuring the activity of the protease, a spectrofluorimetric method was devised to monitor the hydrolysis continuously, which is simpler and more precise than the h.p.l.c. technique used previously.

INTRODUCTION

The processing of viral proteins by virally encoded proteases is of key importance in the maturation of many viruses (Kraeusslich and Wimmer, 1988; Palmenberg, 1990; Lawson and Semler, 1990). Hence these enzymes are possible targets for antiviral agents (Korant, 1990). Poliovirus is the most extensively studied member of the clinically and economically important *Picornaviridae*. Replication of poliovirus is controlled by a series of proteolytic cleavages of a large precursor polyprotein. The proteolytic processing is accomplished by two viral gene products, proteases 2A and 3C (Hanecak et al., 1982; Toyoda et al., 1986). Most of the cleavages are catalysed by protease 3C, which contains 183 amino-acid residues and exhibits specificity for the Gln-Gly bond. The enzyme is a cysteine protease, but its relationship to the well-known papain family is unclear, and no physical studies have been reported to date. Its activity can be determined with synthetic oligopeptides and target virus proteins by using h.p.l.c. techniques and SDS/PAGE respectively.

It has previously been shown that some *Escherichia coli* systems produced a truncated polypeptide chain, in addition to the normal gene product protease 3C or protease 3C elongated by 25 amino-acid residues at the C-terminus (Hanecak et al., 1984; Ivanoff et al., 1986; Korant and Towatari, 1986; Sablina and Antonov, 1991). The smaller molecule was an internal initiation product beginning at the methionine residue at position 27. The present paper shows that the absence of the N-terminal portion of protease 3C substantially affects the physical properties of the molecule, and this renders it possible to separate readily the mature and truncated forms. Furthermore, a simple spectrofluorimetric method has also been devised for characterization of the kinetic properties of the renatured protease 3C.

MATERIALS AND METHODS

Restriction endonucleases and T4 ligase were obtained from Amersham International. A Superose 12 column was purchased from Pharmacia.

General methods

Plasmids were isolated by the modified boiling method (Del Sal et al., 1988) and the bacteria were transformed as described (Hanahan, 1983). DNA fragments were separated by agarose gel electrophoresis and purified with a Geneclean kit as recommended by the manufacturer (Bio 101 Inc.). N-terminal amino-acid sequencing was carried out with a gas-phase automatic sequencer (type ABI 450).

SDS/PAGE was performed on 15% (w/v) slab gels (1 mm × 150 mm × 100 mm) using the discontinuous buffer system (Laemmli, 1970). Gels were stained for 2 h in 0.05% (w/v) Coomassie Brilliant Blue R250 in methanol/acetic acid/water (5:1:4, by vol.), and destaining was carried out in methanol/acetic acid/water (1:1:8, by vol.).

Western blotting was performed according to Towlin et al. (1979). The antiserum was raised against truncated protease 3C isolated from an SDS/polyacrylamide gel.

Bacterial strain and plasmids

E. coli JM109 was used to express poliovirus protease 3C. For constructing the expression vector the following two plasmids were employed. (1) pExc derived from pBR322 by insertion of a tryptophan promoter and a Shine–Dalgarno sequence (Ivanoff et al., 1986). This plasmid contains a modified gene of poliovirus protease 3C (*PV3C*) that starts with Met instead of Gly-1, and

Abbreviations used: Nle, norleucine; Phe(NO_2), 4-nitrophenylalanine; Boc, t-butyloxycarbonyl; Fmoc, fluoren-9-ylmethyloxycarbonyl; OSu, *N*-succinimidyloxy; NBz, 2-aminobenzoyl; IPTG, isopropylthiogalactoside; DTE, dithioerythritol.

‖ Present address: Agricultural Biotechnology Center, Gödöllö, P.O. Box 170, H-2101 Hungary.
¶ To whom correspondence should be addressed.

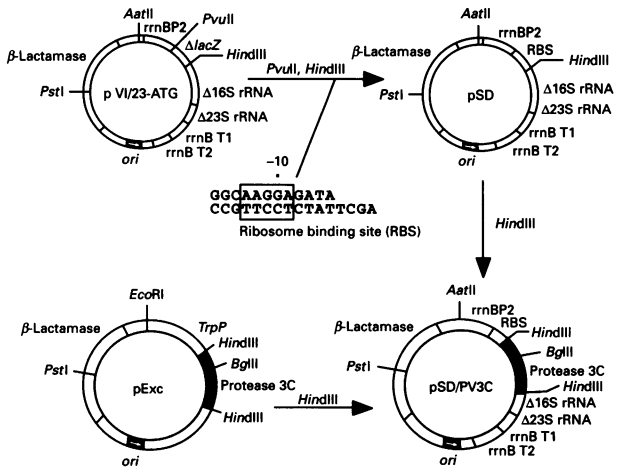

Figure 1 Schematic representation of the strategy used to generate plasmid pSD/PV3C

has a Gly-184 to stop mutation (Ivanoff et al., 1986). (2) pI/23-ATG, a derivative of pBR322 containing a ribosomal rrnBP2 promoter combined with a *lac* operator sequence, portions of 16S rRNA (1603–1787) and 23S rRNA (6383–6617) coding sequences as translational stabilizers, and the rrnB T1 and T2 terminator sequences (Boros et al., 1986).

Construction of pSD/PVC3

The protease gene from the less effective pExc construction was inserted into the plasmid pVI/23-ATG which has a stronger, easily inducible promoter. To this end the plasmid pVI/23-ATG was digested with *Pvu*II and *Hin*dIII, and the large fragment of the vector was isolated. A synthetic linker containing an effective Shine–Dalgarno sequence (see Figure 1) was ligated into the cleaved vector named pSD. Plasmid pExc was digested with *Hin*dIII and the isolated *PV3C* fragment was inserted into the *Hin*dIII site of plasmid pSD. The orientation of the *PV3C* insert was screened by agarose gel electrophoresis of the *Bg*lII–*Pst*I digest.

Isolation of inclusion bodies

E. coli cells harbouring the pSD/PV3C plasmid were grown in Luria–Bertoni medium containing 100 μg/ml ampicillin. For expression of protease 3C, the culture, in the logarithmic phase (attenuence values of 0.3–0.4), was treated with 10 μM isopropylthiogalactoside (IPTG) for 6 h or with 250 μM IPTG for 2 h at 37 °C. The cells were centrifuged and the pellet was suspended in 50 mM phosphate buffer (one-twentieth of the original volume), pH 7.5, containing 5 mM DTE, 10 mM EDTA. The cell suspension was sonicated with an MSE sonicator over a salted ice bath, until the attenuance at 550 nm decreased to about 20% of the initial value. It was then centrifuged for 15 min at 6000 *g*, the pellet washed with the sonication buffer, and stored at −18 °C.

Renaturation

Inclusion bodies (500 mg) were dissolved in 10 ml of 50 mM phosphate buffer, pH 7.5, containing 10 mM DTE, 10 mM EDTA and 6 M guanidinium chloride. Because guanidinium chloride acidifies the medium, the pH of the buffer was adjusted to a pH value of 7.5 with about 200 μl of 1 M NaOH. After a 15 min incubation at room temperature, the opalescent solution was centrifuged at 12500 *g* for 5 min. The clear supernatant (total absorbance value of 120 when measured at 280 nm in 0.1 M NaOH) was diluted 10-fold at room temperature with 50 mM phosphate buffer, pH 8.0, containing 5 mM Na_2SO_3 and 5 mM EDTA. The insoluble material generated on dilution was removed by centrifugation at 15000 *g* for 20 min at 4 °C.

Precipitation and extraction

The renatured enzyme obtained in the preceding step (total absorbance value of 48 at 280 nm) was precipitated by the addition of 0.55 g of ammonium sulphate to 1.0 ml of solution. After standing at 4 °C for 60 min, the precipitate was collected by centrifugation (12500 *g*, 5 min) and extracted with 8 ml of 50 mM phosphate buffer, pH 7.0, containing 5 mM DTE and 5 mM EDTA. The enzyme was concentrated by ultrafiltration (total absorbance of 7.5 at 280 nm) and stored in the presence of 10% (v/v) glycerol at −80 °C or purified further with gel chromatography.

F.p.l.c. analysis

A sample (200 μl) of the solution described above (total absorbance value of 3) was applied to a Superose 12 column equilibrated with 50 mM phosphate buffer, pH 6.8, containing 10 mM cysteine, 5 mM EDTA and 10% (v/v) glycerol. The flow rate was 0.4 ml/min and the active protein peak eluted at 35.5 min. It was concentrated on an Amicon PM-10 membrane and stored at −80 °C.

Synthesis of substrate

The fluorogenic substrate NBz-Nle-Nle-Nle-Glu-Ala-Leu-Phe-Gln*Gly-Pro-Phe(NO_2)-OH (* represents a scissile bond) was prepared by solid-phase synthesis on aminomethylated polystyrene derivatized with 4-(hydroxymethyl)phenoxyacetic acid. The coupling steps were performed by prepared 1-hydroxybenzotriazole esters of fluoren-9-ylmethyloxycarbonyl (Fmoc) amino-acid residues prepared *in situ*. Boc-NBz-OSu was used to introduce the NBz group. The crude product obtained by trifluoroacetic acid treatment of peptide–resin was dissolved in hot acetic acid. The NBz-peptide precipitated on cooling was shown to be homogenous by h.p.l.c. and gave the correct amino-acid analysis.

Determination of activity

The activity of protease 3C was measured spectrofluorimetrically by using our oligopeptide substrate. A 5-fold dilution of the substrate (2 mg/ml in dimethylformamide) with 50 mM phosphate buffer, pH 6.8, was prepared. A sample (10 μl) of this solution was added to the reaction mixture of 540 μl final volume, which also contained 50 mM phosphate buffer, pH 6.8, 1 mM EDTA, 1 mM DTE and enzyme (0.2 μM). Initial rates were monitored at 25 °C using a Jasco FP 777 spectrofluorimeter using excitation and emission wavelengths of 330 nm and 420 nm respectively. One unit of activity corresponds to the amount of enzyme which catalyses the hydrolysis of 1 μmol of substrate/ml per min.

The enzyme concentration was calculated from the absorbance at 280 nm by using an M_r value of 20000 and A_{280} (0.1%) of 0.42 (Nicklin et al., 1988).

RESULTS AND DISCUSSION

Analysis of the pSD/PV3C products

The plasmid pSD/PV3C contains an rrnBP2 promoter combined with a *lac* operator. Induction with 250 μM IPTG for 2 h or 10 μM IPTG for at least 6 h led to the largest yield of inclusion bodies. The *PV3C* gene products were about 40% of the total cellular proteins.

SDS/PAGE showed that the new construction produced two polypeptide chains; this is in agreement with previous observations (Hanecak et al., 1984; Ivanoff et al., 1986; Korant and Towatari, 1986; Sablina and Antonov, 1991). These two protein bands were the only ones detected when tested by Western-blot analysis.

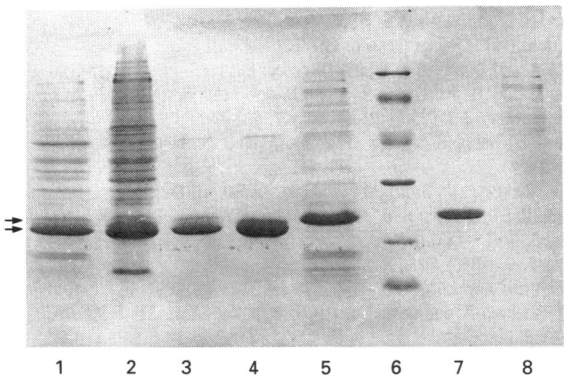

Figure 2 SDS/PAGE of protease 3C

Lane 1, inclusion bodies; lane 2, precipitate formed on dilution of the inclusion bodies dissolved in guanidinium chloride; lane 3, ammonium sulphate precipitate of renatured enzyme; lane 4, insoluble fraction of the ammonium sulphate precipitate; lane 5, soluble fraction of the ammonium sulphate precipitate; lane 6, M_r standard (Pharmacia) consisting of (from bottom upward): lactalbumin, M_r 14400; soybean trypsin inhibitor, M_r 20000; carbonic anhydrase, M_r 30000; ovalbumin, M_r 43000; BSA, M_r 67000; and phosphorylase *b*, M_r 94000; lane 7, peak B from f.p.l.c. analysis; lane 8, peak A from f.p.l.c. analysis. Arrows indicate bands giving positive results on Western blotting.

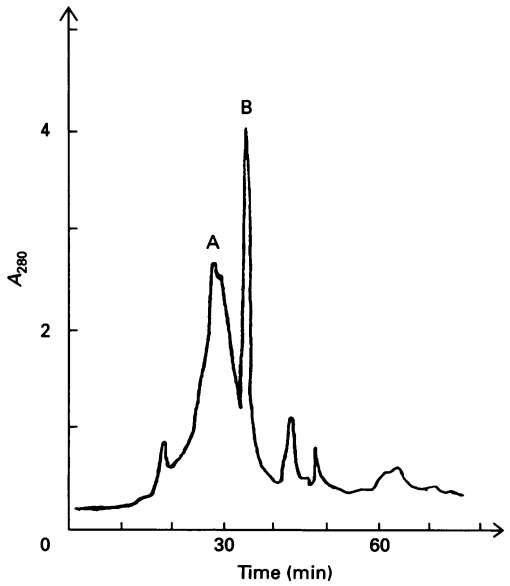

Figure 3 F.p.l.c. of protease 3C on a Superose 12 column

The sharp peak B contains the active enzyme.

Before the N-terminal-sequence analysis of the two polypeptide chains, the inclusion bodies were partially purified by washing with 1 M urea and 70% (v/v) ethanol. N-terminal sequencing of this product showed that the two proteins represented the full-length protease and its internal initiation product starting with Met-27 respectively. Based on the sequencing data, the molar ratio of the complete protease 3C and the truncated form was about 1:10. Amounts of amino-acid residues originating from contaminating proteins in the inclusion bodies were not significant.

Apparently, the translational efficiency of the internal ribosome binding site in the mRNA (AAGGGAGAGUUCACUAUG) is much higher than the one (AAGGAGAUAAGCUUAUG) synthesized to regulate the production of the full-length protease (see Figure 1), despite the fact that the latter was designed as an optimal *E. coli* ribosomal binding site (Min et al., 1988). The ratio found for the two gene products, in addition to the relative affinity of the two ribosome binding sites to the 3'-end of *E. coli* 16 rRNA, may be strongly dependent on the expression system itself. While proteins expressed in a soluble form probably undergo rapid and differential proteolytic degradation in the cytoplasm (cf. Ivanoff et al., 1986), the formation of inclusion bodies prevents such degradation (Mitraki and King, 1989).

Isolation of protease 3C

Pure protease 3C has been isolated in only a few cases. According to one approach the folded enzyme was isolated from the cell lysate with a rather low yield (Nicklin et al., 1988). The other approach used inclusion bodies dissolved in 8 M urea and dialysed against a buffer lacking urea (Baum et al., 1991). We have found that during dialysis some of the truncated polypeptide chains precipitated, resulting in an increased proportion of the full-length enzyme in solution. The partial precipitation might be due to the less efficient folding of the truncated polypeptide compared with the full-length enzyme. When the inclusion bodies were dissolved in 8 M urea or 6 M guanidinium chloride, and diluted rather than dialysed, only the full-length protease could be extracted from the ammonium sulphate precipitate (Figure 2). This method has the additional advantage of being simpler and faster than dialysis. It should be noted that despite the unfavourable ratio of mature to truncated enzymes produced by the present construction a considerable quantity of protease 3C was obtained (2.5 mg from 500 mg of inclusion bodies or 9–10 mg from 1 litre of growth medium), because of the high level of production of the recombinant proteins.

The final purification step was performed by chromatography on a Superose 12 column. Most of the impurities (peak A of Figure 3) eluted before protease 3C. Peak A contained a variety of proteins (Figure 2), those of small relative molecular masses were presumably in aggregates. The broad peak A was immediately followed by the sharp peak B, indicating the elution of a homogeneous protein of about M_r 25000. The homogeneity was confirmed by SDS/PAGE (Figure 2).

Activity of protease 3C

The catalytic activity of the viral protease is usually measured with natural protein substrates by detecting the cleavage products by SDS/PAGE. Quantitative results can be obtained by using synthetic peptide substrates and by determining the reaction products of h.p.l.c. For such measurements peptides containing 12–16 residues are generally used. These are composed of amino-acid residues flanking poliovirus cleavage junctions (Nicklin et al., 1988; Pallai et al., 1989; Baum et al., 1991). Clearly, the most desirable assay for precise kinetic measurements monitors the

hydrolysis continuously by spectrophotometric or spectrofluorimetric methods. Therefore, we synthesized a peptide derivative of the 2C/3A junction of the poliovirus polyprotein, one of the most susceptible sites for protease 3C (Pallai et al., 1989). The oligopeptide contained an aminobenzoyl (NBz) (anthranilyl) and a *p*-nitrophenylalanine [Phe(NO_2)] label attached to the N- and C-terminus respectively. A similar intramolecular fluorescence-energy transfer proved to be successful with an oligopeptide substrate for the human immunodeficiency virus protease (Cheng et al., 1990). When the Gln-Gly bond of the present substrate is hydrolysed, the intramolecular quenching is abolished and the increase in fluorescence can be monitored. The substrate concentration in the assay mixture was higher than K_m so that initial rates were independent of an increase or decrease in the substrate concentration by a factor of 2. The K_m value could not be determined because it was too low to measure under the assay conditions although the substrate concentration (5.3 μM) was lower by an order of magnitude than that used with other substrates having the same quenching system (50 μM; Cheng et al., 1990). At 25 °C, 1 mg of protease 3C hydrolysed 240 nmol of substrate/ml per min, corresponding to 0.24 i.u. This reaction was completely inhibited when protease 3C was treated with iodoacetamide, an irreversible inhibitor of cysteine enzymes. This indicates that the hydrolysis was indeed performed by the recombinant cysteine protease. The cleavage of the Gln-Gly bond by the enzyme was confirmed by h.p.l.c. which gave a retention time for one of the products identical with that of NH_2-Gly-Pro-Phe(NO_2)-OH.

We thank Dr. A. Patthy (ABC, Gödöllö) for N-terminal-sequence analysis and Dr. A. Pethö for helpful discussions. Thanks are also due to Ms J. Fejes and Ms. I. Szamosi for skilful technical assistance. The financial support by grant OTKA I/3 1360 is gratefully acknowledged.

Received 29 May 1992/1 October 1992; accepted 13 October 1992

REFERENCES

Baum, E. Z., Bebernitz, G. A., Palant, O., Mueller, T. and Plotch, S. J. (1991) Virology **185**, 140–150
Boros, I., Lukacsovich, T., Baliko, G. and Venetianer, P. (1986) Gene **42**, 97–100
Cheng, Y.-S. E., Yin, F. H., Foundling, S., Blomstrom, D. and Kettner, C. A. (1990) Proc. Natl. Acad. Sci. U.S.A. **87**, 9660–9664
Del Sal, G., Naufioletti, G. and Schneider, C. (1988) Nucleic Acid Res. **16**, 9878
Hanahan, D. (1983) J. Mol. Biol. **166**, 557–580
Hanecak, R., Semler, B. L., Anderson, C. W. and Wimmer, E. (1982) Proc. Natl. Acad. Sci. U.S.A. **79**, 3973–3977
Hanecak, R., Semler, B. L., Ariga, H., Anderson, C. W. and Wimmer, E. (1984) Cell **37**, 1063–1073
Ivanoff, L. A., Towatari, T., Jasodhara, R., Korant, B. D. and Petteway, S. R., Jr. (1986) Proc. Natl. Acad. Sci. U.S.A. **83**, 5392–5396
Korant, B. D. (1990) Ann. N.Y. Acad. Sci. **616**, 252–257
Korant, B. D. and Towatari, T. (1986) Biomed. Biochim. Acta **45**, 1529–1535
Kraeusslich, H.-G. and Wimmer, E. (1988) Annu. Rev. Biochem. **57**, 701–754
Laemmli, U. K. (1970) Nature (London) **227**, 680–685
Lawson, M. A. and Semler, B. L. (1990) Curr. Top. Microbiol. Immunol. **161**, 49–87
Min, K. T., Kim, M. H. and Lee, D.-S. (1988) Nucleic Acids Res. **16**, 5075–5088
Mitraki, A. and King, J. (1989) Bio/Technology **7**, 690–697
Nicklin, M. J. H., Harris, K. S., Pallai, P. V. and Wimmer, E. (1988) J. Virol. **62**, 4586–4593
Pallai, P. V., Burkhardt, F., Skoog, M., Schreiner, K., Bax, P., Cohen, K. A., Hansen, G., Palladino, D. E. H., Harris, K. S., Nicklin, M. J. and Wimmer, E. (1989) J. Biol. Chem. **264**, 9738–9741
Palmenberg, A. C. (1990) Annu. Rev. Microbiol. **44**, 603–623
Sablina, E. P. and Antonov, V. K. (1991) FEBS Lett. **283**, 291–294
Towlin, H., Staehelin, T. and Gordon, T. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 4350–4353
Toyoda, H., Nicklin, M. J., Murray, M. G., Anderson, C. W., Dunn, J. J., Studier, F. W. and Wimmer, E. (1986) Cell **45**, 761–770