M^{-1} s^{-1}, and the stopped-flow results imply that interaction with the wild-type thioredoxin is considerably more rapid (greater than 10^7 M^{-1} s^{-1}). Since the rate constant for electron transfer (k_3) is approximately half that of the reaction with wild-type thioredoxin, it appears that the mutation has only a small effect on this process.

Unlike the reaction with wild-type thioredoxin, reduction of the mutant is enhanced by ionic strength up to 200 mM NaCl. Increased salt would be expected to shield the negative charge on the glutamic acid and facilitate binding to the reductase. As seen from Table II, K_m for K36E is lower at 200 mM NaCl, and the k_{cat} value, i.e., the electron-transfer rate constant, is the same as with wild-type thioredoxin. This is consistent with a decrease in the small inhibitory electrostatic influence of the negative carboxyl ion on this process.

The data clearly show that rapid reaction kinetics can be used to study electron transfer between thioredoxin and reductase. Combined with site-directed mutagenesis of residues which are believed to be in the protein interaction site, additional information can thus be obtained on the forces which optimize this interaction.

Registry No. Thioredoxin reductase, 9074-14-0; lysine, 56-87-1; glutamic acid, 56-86-0.

References

Ghisla, S., & Massey, V. (1989) *Eur. J. Biochem. 181*, 1–17.

Gleason, F. K., & Holmgren, A. (1988) *FEMS Microbiol. Rev. 54*, 271–298.

Gleason, F. K., Lim, C.-J., Gerami-Nejad, M., & Fuchs, J. A. (1990) *Biochemistry 29*, 3701–3709.

Hazzard, J. T., McLendon, G., Cusanovich, M. A., & Tollin, G. (1988) *Biochem. Biophys. Res. Commun. 151*, 429–434.

Huber, P. W., & Brandt, K. G. (1980) *Biochemistry 19*, 4568–4575.

Kallis, G.-B., & Holmgren, A. (1980) *J. Biol. Chem. 255*, 10261–10265.

Katti, S. K., LeMaster, D. M., & Eklund, H. (1990) *J. Mol. Biol. 212*, 167–184.

Lim, C.-J., Gleason, F. K., & Fuchs, J. A. (1986) *J. Bacteriol. 168*, 1258–1264.

Luthman, M., & Holmgren, A. (1982) *Biochemistry 21*, 6628–6633.

Massey, V., Gibson, Q. H., & Veeger, C. (1960) *Biochem. J. 77*, 341–351.

Moore, E. C., Reichard, P., & Thelander, L. (1964) *J. Biol. Chem. 239*, 3445–3452.

Myers, R. M., Terman, L. S., & Maniatis, T. (1985) *Science 229*, 242–247.

O'Donnell, M. E., & Williams, C. H. (1983) *J. Biol. Chem. 258*, 13795–13805.

O'Donnell, M. E., & Williams, C. H. (1985) *Biochemistry 24*, 7617–7621.

Pigiet, V. P., & Conley, R. R. (1977) *J. Biol. Chem. 252*, 6367–6372.

Porque, P. G., Baldesten, A., & Reichard, P. (1970) *J. Biol. Chem. 245*, 2363–2370.

Russel, M., & Model, P. (1988) *J. Biol. Chem. 263*, 9015–9019.

Simondsen, R. P., Weber, P. C., Salemme, F. R., & Tollin, G. (1982) *Biochemistry 21*, 6366–6375.

Thelander, L. (1967) *J. Biol. Chem. 242*, 852–859.

Williams, C. H., Jr. (1976) *Enzymes (3rd Ed.) 13*, 89–173.

Wong, K. K., Vanoni, M. A., & Blanchard, J. S. (1988) *Biochemistry 27*, 7091–7096.

cDNA Cloning of Porcine Brain Prolyl Endopeptidase and Identification of the Active-Site Seryl Residue†

Denise Rennex,‡ Brian A. Hemmings, Jan Hofsteenge,* and Stuart R. Stone

Friedrich Miescher-Institut, Postfach 2543, CH-4002, Basel, Switzerland

Received August 8, 1990; Revised Manuscript Received October 29, 1990

ABSTRACT: Prolyl endopeptidase is a cytoplasmic serine protease. The enzyme was purified from porcine kidney, and oligonucleotides based on peptide sequences from this protein were used to isolate a cDNA clone from a porcine brain library. This clone contained the complete coding sequence of prolyl endopeptidase and encoded a polypeptide with a molecular mass of 80 751 Da. The deduced amino acid sequence of prolyl endopeptidase showed no sequence homology with other known serine proteases. [^{3}H]Diisopropyl fluorophosphate was used to identify the active-site serine of prolyl endopeptidase. One labeled peptide was isolated and sequenced. The sequence surrounding the active-site serine was Asn-Gly-Gly-Ser-Asn-Gly-Gly. This sequence is different from the active-site sequences of other known serine proteases. This difference and the lack of overall homology with the known families of serine proteases suggest that prolyl endopeptidase represents a new type of serine protease.

Prolyl endopeptidase (PE)[1] is a cytoplasmic protease (Dresdner et al., 1982; Green & Shaw, 1983) that is ubiquitously distributed among mammalian tissues (Yoshimoto et al., 1979; Kato et al., 1980). It is one of several proteases that are specific for proline [reviewed in Walter et al. (1980)]. However, in contrast to other proline-specific proteases that act as aminopeptidases, carboxypeptidases, or dipeptidases, PE cleaves peptide bonds on the C-terminal side of prolyl residues within peptides that are up to approximately 30 amino acids long (Walter & Yoshimoto, 1978; Taylor & Dixon, 1980;

† The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05311.

* Author to whom correspondence should be addressed.

‡ Present address: Division of Hematology–Oncology, New England Medical Center, 750 Washington St., Box 832, Boston, MA 02111.

[1] Abbreviations: Ac, acetyl; CH_2Cl, chloromethane; DFP, diisopropyl fluorophosphate; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PE, prolyl endopeptidase; pNA, *p*-nitroanilide; PTH, phenylthiohydantoin; TLCK, tosyllysine chloromethyl ketone; TPCK, tosylphenylalanine chloromethyl ketone; Z, benzyloxycarbonyl.

0006-2960/91/0430-2195$02.50/0

Koida & Walter, 1976). It does not cleave native or denatured proteins larger than this size (Moriyama et al., 1988; Taylor & Dixon, 1980).

PE is inactivated by inhibitors of cysteine proteases, such as *N*-ethylmaleimide, *p*-(chloromercuri)benzoate (Yoshimoto et al., 1983; Zolfaghari et al., 1986), and an active-site-directed peptidyl diazomethane (Green & Shaw, 1983). However, it is classified as a serine protease, on the basis of the inactivation of the enzyme by diisopropyl fluorophosphate (DFP) (Yoshimoto et al., 1977), a compound that reacts with the reactive serine in the active site of enzymes such as serine proteases and serine esterases (Cohen et al., 1967). PE that is inactivated by [^{3}H]DFP incorporates 1 mol of diisopropyl fluorophosphate/mol of protein (Yoshimoto et al., 1977, 1983), and the amino acid that is labeled when the enzyme from lamb brain is inactivated by [^{32}P]DFP has been identified as a seryl residue (Yoshimoto et al., 1981).

Although PE is classified as a serine protease, it differs from the established families of these enzymes in its substrate specificity and subcellular location. With a few exceptions, the mammalian serine proteases whose structures have been determined belong to the trypsin family and are located mainly extracellularly or in secretory granules and vesicles (Neurath, 1989). In this paper, we report the deduced amino acid sequence of porcine PE and the identity of the active-site serine. The sequence of the enzyme showed no homology to other proteases, and the amino acid sequence that surrounded the active-site serine was different from those of other serine proteases. Therefore, PE appears to represent a new type of serine protease.

Experimental Procedures

Materials. DEAE-Sepharose CL-6B, PBE94, Polybuffer 74, and FPLC columns were obtained from Pharmacia (Uppsala, Sweden). Porcine liver and porcine brain cDNA libraries were obtained from Clonetech Laboratories, Inc. (Palo Alto, CA). TPCK-treated trypsin and TLCK-treated chymotrypsin were from Worthington (Malvern, PA). [^{3}H]DFP was from Du Pont–New England Nuclear (Wilmington, DE). Rainbow-colored molecular weight markers were from Amersham Corp. (Amersham, U.K.). All other chemicals were of the highest purity commercially available.

Purification Procedure. All purification steps except FPLC were carried out at 4 °C. One kilogram of porcine kidney was homogenized in a Waring blender with 4 L of 50 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, 1 mM EDTA, 1 mM DTT, 20 mg/L soybean trypsin inhibitor, 1 mM benzamidine, 2 mg/L leupeptin, and 20 000 units/L kallikrein inhibitor. The homogenate was centrifuged for 30 min at 7000*g*. The pH of the supernatant was brought to 5.0 by the addition of 0.5 M acetic acid, stirred for 1 h, and centrifuged at 7000*g* for 30 min. The pH of the supernatant was brought to 7.5 by the addition of 0.5 M NaOH, and solid ammonium sulfate was added to 50% saturation. The solution was stirred for 45 min and centrifuged for 30 min at 7000*g*. The supernatant was brought to 80% saturation with ammonium sulfate and stirred and centrifuged as described above. The pellet was dissolved in 50 mM Tris-HCl, pH 7.5, containing 1 mM EDTA and 1 mM DTT (buffer A) and dialyzed against the same buffer. The dialysate was batch-adsorbed to about 300 mL of DEAE-Sepharose CL-6B for 1 h. The resin was washed with buffer A, packed into a column (5 cm × 15 cm), and eluted with a gradient of 0–0.4 M NaCl in buffer A. The active fractions were pooled and concentrated in an Amicon ultrafiltration apparatus using a PM10 membrane. The concentrated fractions were loaded onto a Sephadex G-100 gel filtration column (5 cm × 90 cm) that had been equilibrated with 25 mM piperazine buffer, pH 6.3, containing 50 mM NaCl, 1 mM EDTA, and 1 mM DTT. Fractions of 10 mL were collected. The active fractions were pooled and loaded onto a PBE94 chromatofocusing column (1.25 cm × 23 cm) that had been equilibrated with 25 mM piperazine buffer, pH 6.3, containing 1 mM EDTA and 1 mM DTT. The column was eluted with 400 mL of 10-fold-diluted Polybuffer 74 solution, pH 4.5, containing 1 mM EDTA and 1 mM DTT. The chromatofocusing step was repeated at room temperature using an FPLC Mono P column and the same equilibration and elution buffers. In order to remove Polybuffer ions, ammonium sulfate was added to the pooled active fractions to give a final concentration of 1.7 M, and these were loaded onto an FPLC phenyl-Superose column that had been equilibrated with buffer A containing 1.7 M ammonium sulfate at room temperature. The enzyme was eluted with a gradient of 1.7–0 M ammonium sulfate in buffer A.

Preparative SDS–Polyacrylamide Gel Electrophoresis. SDS–polyacrylamide gel electrophoresis was performed by the method of Laemmli (1970). During preparative electrophoresis, the procedure was altered to protect the protein from chemical modification (Hunkapiller et al., 1983). The protein was heated with an equal volume of loading buffer at 60 °C for 15 min and loaded onto a 10% SDS–polyacrylamide gel. The cathode running buffer contained 2 mM sodium thioglycolate. The protein band corresponding to PE was visualized by soaking the gel in 1 M KCl. The band containing the enzyme was excised and the protein eluted as described by Hunkapiller et al. (1983). SDS was removed from the protein by using the method of Konigsberg and Henderson (1983).

Protein Sequence Analysis. The protein was reduced and carboxymethylated by the method of Crestfield et al. (1963) and digested with trypsin. The digest was fractionated on a C_8 reverse-phase column (Brownlee, Santa Clara, CA) using a linear gradient of acetonitrile (0–70%) (v/v) in 0.1% (v/v) trifluoroacetic acid over 240 min at a flow rate of 1 mL/min.

Amino acid sequence determination was performed by using an Applied Biosystems gas-phase sequencer Model 470A equipped with an on-line PTH analyzer. The program 03RPTH supplied by the manufacturer was used. Amino acid analysis was performed by the method of Knecht and Chang (1987), after hydrolysis of the polypeptide in the vapor of 6 N HCl at 110 °C for 18 h.

Measurement of Enzyme Activity. The activity of PE was measured by using the fluorogenic substrate Z-Gly-Pro-7-amido-4-methylcoumarin. The 2.0-mL reaction mixture contained 50 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, 0.02% (w/v) Tween 20, 2.5% (v/v) dimethylformamide, and 50 μM substrate. The assay was performed at 37 °C, and the release of the fluorophore, 7-amino-4-methylcoumarin, was measured by using an excitation wavelength of 383 nm and an emission wavelength of 455 nm.

One unit of PE activity is defined as the amount of enzyme that released 1 μmol of 7-amino-4-methylcoumarin/min.

Protein Assay. Protein was measured by the method of Bradford (1976) using bovine serum albumin as a standard.

Inactivation of PE by Diisopropyl Fluorophosphate. The inactivation of PE by DFP was studied in the presence of the substrate Z-Gly-Pro-pNA. The assays were carried out in 50 mM HEPES buffer, pH 7.4, containing 0.1 M NaCl, 1.0 mM EDTA, 5.0 mM DTT, 0.02% (w/v) Tween 20, and 2% (v/v) methanol as well as variable concentrations of substrate and

DFP. The assays were initiated by addition of the enzyme. Experiments were conducted with six different substrate concentrations ranging from 22 to 131 μM. The substrate concentration was determined spectrophotometrically from its extinction at 342 nm (Lottenberg & Jackson, 1983). At each substrate concentration, progress curve data were collected for six different concentrations of DFP from 0 to 40 μM. Stock solutions of DFP were made up in 2-propanol; the amount of DFP was measured by weight. The desired concentration of DFP was obtained by a 20–50-fold dilution into the cuvette immediately before the start of the assay.

The inactivation of PE by DFP can be represented by the following scheme:

$$\text{PE} + \text{DFP} \xrightarrow{k_i} \text{PE–DIP}$$

where PE–DIP and k_i represent the inactive PE–diisopropyl phosphate complex and the second-order inactivation rate constant, respectively. In the presence of substrate, the progress curve data for the inactivation process can be described by eq 1 [see Gray and Duggleby (1989), Tian and Tsou (1982), Walker and Elmore (1984), and Stone and Hofsteenge (1985)] where P is the concentration of p-nitroaniline formed

$$P = (v_0/k'I)[1 - \exp(-k'It)] \quad (1)$$

by the enzyme at time t, v_0 is the rate of p-nitroaniline formation in the absence of DFP, k' is the apparent second-order rate constant for the inactivation, and I is the concentration of DFP. Progress curve data for a series of DFP concentrations at a particular substrate concentration were fitted to eq 1 by nonlinear regression to yield estimates for k'. In all cases, only data obtained at less than 10% substrate utilization were used in the analysis.

If the DFP competes with the substrate, the value of k' will depend on the substrate concentration as given by eq 2 where

$$k' = \frac{k_i}{1 + S/K_m} \quad (2)$$

S is the substrate concentration and K_m is the Michaelis constant for the substrate. Values of k' obtained at different substrate concentrations were weighted according to the squared reciprocal of their standard error and fitted to the reciprocal form of eq 2 by weighted linear regression in order to demonstrate that DFP was competitive with the substrate and to obtain an estimate of k_i.

An estimate for K_m was also independently determined in initial velocity studies using methods previously described (Hofsteenge et al., 1986).

Active-Site Labeling with [^3H]DFP. PE was purified as described above, except that the acid precipitation step was omitted. The enzyme obtained from the phenyl-Superose column was transferred into 50 mM Tris-HCl buffer, pH 7.5, containing 150 mM NaCl, 1 mM EDTA, and 1 mM DTT. A 10-fold molar excess (20 μM) of [^{3}H]DFP (in propylene glycol, final concentration 6% v/v) was added, and the enzyme was incubated at 37 °C for 2 h. During this time, samples were removed and assayed for loss of activity and incorporation of the inhibitor into the enzyme. A control reaction was performed in which the enzyme was incubated with propylene glycol alone. Incorporation of the labeled inhibitor was followed by SDS–polyacrylamide gel electrophoresis. For each time point, the reaction was stopped by the addition of loading buffer that had been heated to 95 °C. The samples were run on a 10% gel and visualized by fluorography using 1 M sodium salicylate (Chamberlin, 1979).

The labeled protein was dialyzed at 4 °C against 50 mM ammonium bicarbonate, then denatured by heating at 80 °C for 10 min, and digested with 5% (w/w) chymotrypsin for 2 h at 37 °C. The digest was acidified by the addition of trifluoroacetic acid and chromatographed on a reverse-phase C_{18} column (Vydac, Hesperia, CA) using a linear gradient of 7–35% (v/v) acetonitrile in 0.1% (v/v) trifluoroacetic acid over 80 min at a flow rate of 1 mL/min. Fractions of 1 mL were collected, and the radioactivity of each fraction was measured. The major radioactively labeled peptide was repurified by using a linear gradient of 24.5–35% (v/v) acetonitrile over 60 min at a flow rate of 1 mL/min.

cDNA Cloning. The oligonucleotide TC(T/C)TG(T/C)-TGGGG(A/G)TAGGC(A/G)TT(A/G)TA(A/G)AACATG was used to screen about 10^6 plaques each of an amplified porcine liver and an LLC-PK_1 porcine kidney cell cDNA library (Hemmings et al., 1986). For this purpose, the oligonucleotide was end-labeled with [γ-^{32}P]ATP (>5000 Ci/mmol) and hybridized with the filters (1×10^6 cpm/mL) in 6 × SSC/2 × Denhardt's/0.05% sodium pyrophosphate/0.1% SDS/100 μg/mL denatured herring sperm DNA at 37 °C for 16 h. The filters were washed at 45 °C in 3 × SSC/0.1% SDS for 45 min, followed by a final wash at 50 °C for 15 min under the same conditions.

A 200 bp *Eco*RI/*Pvu*II fragment of one of the porcine liver cDNA clones (pl1) was isolated and used to screen a porcine brain cDNA library. The fragment was labeled by the random priming method of Feinberg and Vogelstein (1983) and hybridized with the filters at 60 °C in 6 × SSC/2 × Denhardt's/0.05% sodium pyrophosphate/0.1% SDS/100 μg/mL denatured herring sperm DNA for 16 h. The filters were washed at 60 °C in 1 × SSC/0.1% SDS for 1.5 h.

DNA Sequence Analysis. cDNA fragments were subcloned into M13mp18 and sequenced by the dideoxynucleotide chain termination method (Sanger et al., 1977) using Sequenase (United States Biochemical Corp., Cleveland, OH) and [α-^{35}S]thio-dATP. The pig brain clone (pb1) was sequenced on both strands using specific oligonucleotide primers. Regions containing a high content of G and C residues were sequenced by using dITP in place of dGTP (Barnes et al., 1983).

Northern Analysis. RNA was purified by using the acid/guanidinium thiocyanate/phenol/chloroform extraction method of Chomczynski and Sacchi (1987) followed by precipitation with 2.5 M LiCl. Twenty micrograms of total RNA from porcine brain, heart, kidney, liver, skeletal muscle, and ovaries and from the LLC-PK_1 cell line was fractionated in a formaldehyde/1.0% agarose gel and transferred electrophoretically onto a Zetaprobe membrane. The membrane was baked at 80 °C under vacuum for 2 h and then prehybridized at 42 °C for 2 h with 5 × SSC/5 × Denhardt's/50% formamide/2% SDS/0.05% sodium pyrophosphate, containing 0.25 mg/mL denatured herring sperm DNA. Hybridization was performed overnight at 42 °C in the same solution containing the pl1 insert that had been labeled by the method of Feinberg and Vogelstein (1983). The filter was washed at 60 °C for 1.5 h in 1 × SSC/0.1% SDS and autoradiographed at −70 °C.

Analysis of Nucleotide and Amino Acid Sequences. The nucleotide sequence and the deduced amino acid sequence of the pb1 clone were compared with the EMBL (release 20), Genbank (release 61.0), and NBRF (nucleic acid release 35.0 and protein release 22.0) databases using the algorithm of Lipman and Pearson (1985).

Results

Purification of PE from Porcine Kidney. PE was purified from porcine kidney 4700-fold with a final yield of 7.5% (Table I). Enzyme activity was assayed by using the specific substrate Z-Gly-Pro-7-amido-4-methylcoumarin (Yoshimoto et

Table I: Purification of Prolyl Endopeptidase[a]

fraction	protein (mg)	act. (units)[b]	% yield	sp act. (units/mg)	purification (*x*-fold)
crude	51800	133	100	2.6×10^{-3}	1
acid pptn	26800	97	73	3.6×10^{-3}	1.4
ammonium sulfate pptn	11100	78	58	7.0×10^{-3}	2.7
DEAE–Sepharose	659	50	38	7.6×10^{-2}	30
G100	229	45	34	0.20	77
PBE94	20.3	30	23	1.50	580
Mono P	1.68	18	14	10.8	4200
phenyl-Superose	0.82	10	7.5	12.2	4700

[a] PE was purified from porcine kidney as described under Experimental Procedures. Details of the methods used for enzyme and protein assays are also given under Experimental Procedures. [b] One unit was defined as the amount of enzyme that cleaved 1 μmol of Z-Gly-Pro-7-amido-4-methylcoumarin/min under the assay conditions used.

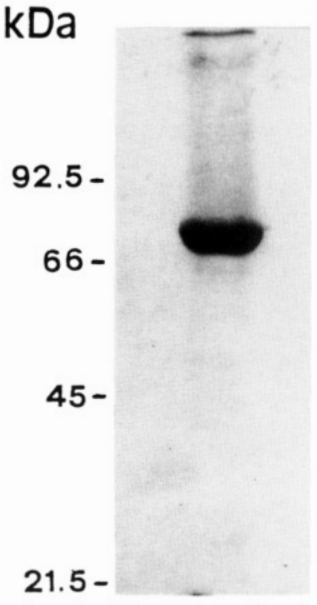

FIGURE 1: SDS–polyacrylamide gel electrophoresis of porcine kidney PE. The purified protein was electrophoresed on a 10% polyacrylamide gel as described under Experimental Procedures and visualized by staining with Coomassie blue. The positions of the molecular weight markers are indicated.

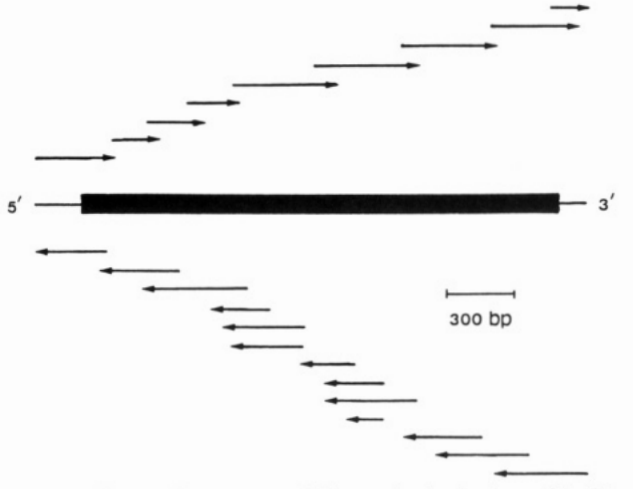

FIGURE 2: Sequencing strategy of the porcine brain clone pb1. The insert of the pb1 clone was subcloned into M13mp18 and sequenced by using the dideoxynucleotide chain termination method (Sanger et al., 1977), as described under Experimental Procedures. The filled-in box represents the coding sequence of the clone, and the arrows represent the direction and extent of sequencing.

al., 1979). SDS–polyacrylamide gel electrophoresis showed one major band that migrated with an apparent molecular mass of 74.5 kDa. Both the native protein and the protein that had been reduced and carboxymethylated were resistant to Edman degradation, which suggested that the N-terminal residue was blocked. To obtain amino acid sequence data, the enzyme was first purified to homogeneity by using preparative gel electrophoresis (Figure 1). The electroeluted protein was reduced and carboxymethylated and digested with trypsin. Peptides were purified by reverse-phase HPLC and sequenced.

cDNA Cloning of PE. A set of oligonucleotides was synthesized, based on the sequence of the peptide G-M-F-Y-N-A-Y-P-Q-Q-D-. This set contained a mixture of 64 oligonucleotides that were partly based on codon preference (Lathe, 1985), and it was used to screen both porcine liver and kidney cell cDNA libraries. The choice of the libraries that were screened was based on the observations that PE activity is present in all tissues and that the enzymes from different tissues are biochemically identical and immunologically indistinguishable (Hersh, 1981; Andrews et al., 1982; Yoshimoto et al., 1988). Several clones were isolated from the porcine liver library, and the clone (pl1) that contained the largest insert was sequenced (data not shown). This clone contained one open reading frame encoding a polypeptide of 598 amino acids, that had a molecular mass of about 67 kDa. Sequences identical with those of 13 of the 17 tryptic peptides could be found in the deduced amino acid sequence of the clone, which confirmed that it encoded PE. However, the clone did not contain the complete coding sequence of the enzyme as the molecular mass of the polypeptide encoded by it was lower than that of the purified enzyme and it lacked an initiating methionine codon.

A 200 bp *Eco*RI/*Pvu*II fragment from the 5′ end of pl1 was used to isolate several clones from a porcine brain cDNA library, and the largest of these (pb1) was sequenced (Figure 2). The 2.5-kb insert of this clone contained one open reading frame that began at the most 5′ ATG codon (Figure 3) and encoded a polypeptide of 710 amino acids that had a molecular mass of 80 751 Da. The nucleotide sequence surrounding this ATG codon conformed to the consensus sequence for an initiating methionine codon (Kozak, 1984). The correctness of the deduced amino acid sequence of the pb1 clone was supported by the amino acid composition and partial sequence analysis of the purified enzyme. The amino acid composition calculated from the deduced amino acid sequence agreed well with the one determined experimentally for the enzyme from porcine kidney (Table II). Additionally, sequences identical with all tryptic peptides could be located in the deduced amino acid sequence. A sequence heterogeneity was found at position 29 of the deduced amino acid sequence. The sequence of one peptide contained a histidyl residue instead of the predicted alanyl residue. However, in a subsequent tryptic digest of PE from porcine muscle, a peptide with an alanyl residue in this position was isolated and sequenced. The amino acid residues that could not be identified during protein sequencing of the tryptic peptides were found to be cysteinyl residues in all cases. This indicated that the carboxymethylation of the protein was

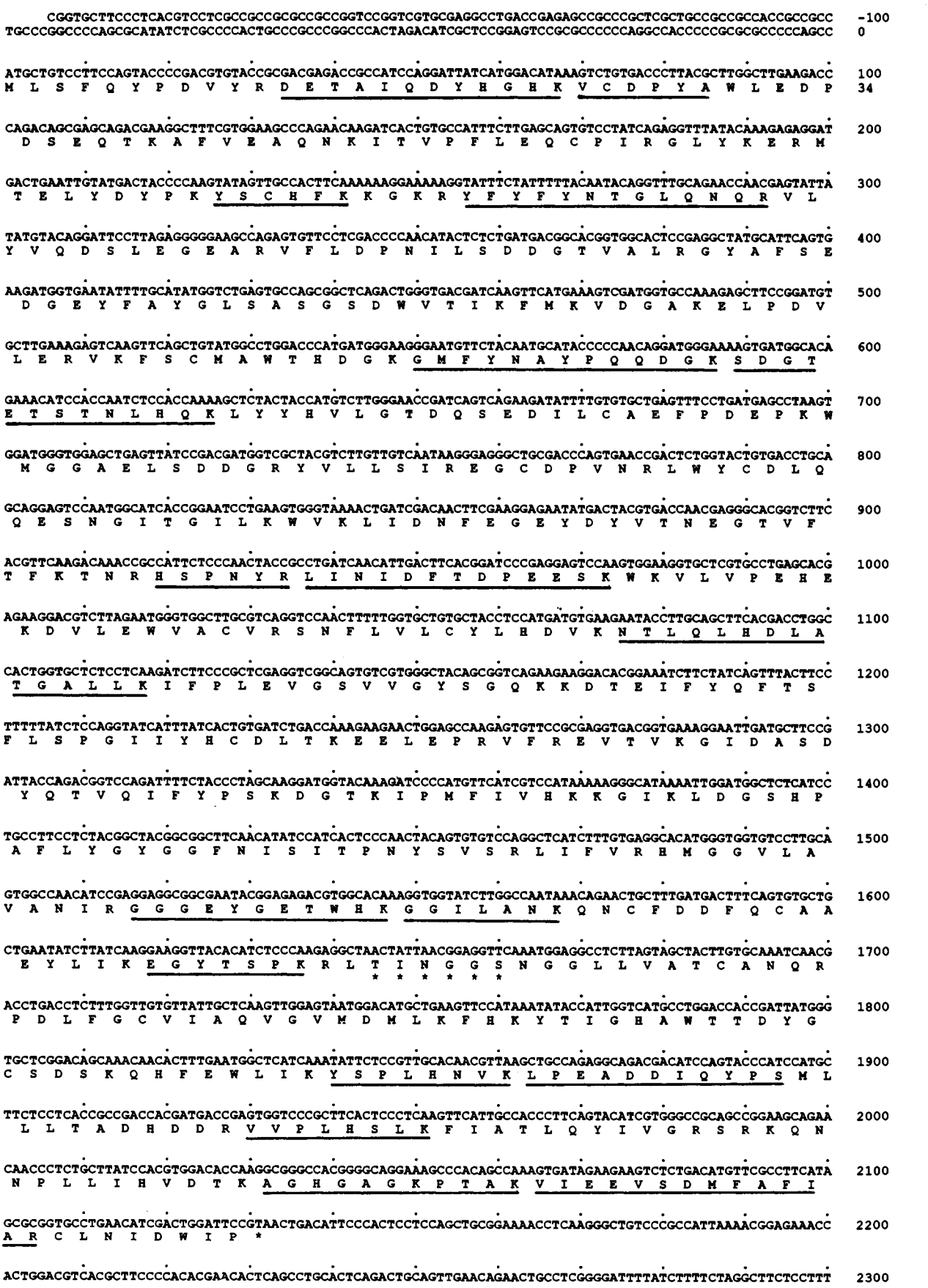

FIGURE 3: Nucleotide sequence and deduced amino acid sequence of the porcine brain clone pb1. The complete nucleotide sequence of the pb1 clone is shown. Number +1 of the nucleotide sequence was assigned to the first nucleotide of the proposed initiator codon. The deduced amino acid sequence in the single-letter code is shown below the nucleotide sequence. The amino acid sequences that are underlined correspond to the sequenced tryptic peptides. The amino acids marked with asterisks correspond to the partial sequence of the [^{3}H]DFP-labeled peptide.

Table II: Amino Acid Composition of the Entire Protein and [^{3}H]DFP-Labeled Peptide[a]

amino acid	protein[b]	[^{3}H]DFP-labeled peptide
Asx	82.9 (79)	1.6 (2)
Glx	72.5 (69)	0.36 (0)
Cys	12.6 (16)	
Ser	38.8 (38)	1.2 (1)
Gly	57.5 (57)	3.9 (4)
His	24.1 (23)	
Arg	27.8 (25)	
Thr	34.8 (38)	1.0 (1)
Ala	40.3 (36)	0.41 (0)
Pro	33.4 (31)	0.29 (0)
Tyr	31.3 (41)	
Val	42.0 (48)	0.22 (0)
Met	11.9 (13)	
Ile	33.3 (39)	1.2 (1)
Leu	60.0 (63)	2.2 (2)
Phe	35.3 (36)	
Lys	44.4 (47)	
Trp	ND (12)	

[a] Amino acid analysis was performed by the method of Knecht and Chang (1987). The values expected from the deduced amino acid sequence are given in parentheses. [b] Average of four determinations.

only partially successful, as PTH-(carboxymethyl)cysteine can be identified in analyses of the PTH-amino acids by the procedure that was used. The amino terminus of PE was blocked, and the sequence of the first tryptic peptide began only at position 12 of the deduced amino acid sequence. However, a CNBr fragment of the protein that began at the leucine at position 2 has been sequenced (data not shown) which suggests that the N-terminal residue of the mature protein was methionine.

In addition to the clones pb1 and pl1, a 2.2-kb clone (pk1) was isolated from a library prepared from the porcine kidney cell line LLC-PK$_1$. The nucleotide sequence of the 2144 bp insert of this clone began at position 437 of the pb1 clone and contained a putative polyadenylation site 11 bp from its 3′ end (data not shown; the sequence will be provided upon request). The length of the overlapping clones was 2779 bp (Figure 3).

Northern Analysis. The insert of the clone pl1 was used as a probe in Northern analysis of total RNA from porcine tissues and from the porcine kidney cell line LLC-PK$_1$. One species of RNA of about 3.0 kb was detected in all tissues examined and in the cell line (Figure 4). The size of the mRNA for PE that was detected by Northern analysis was approximately equal to the total length of the overlapping clones (2779 bp) plus an estimated 200 bases, that comprise the average length of a poly(A) tail (Darnell, 1982). Therefore, it is likely that the clones pl1, pb1, and pk1 covered the entire coding and noncoding regions of the mRNA encoding PE. The level of this mRNA varied little among the tissues and cell type examined.

Inactivation of PE by DFP. The inactivation of PE by DFP was followed in the presence of the substrate Z-Gly-Pro-pNA. Figure 5A shows a typical set of data. Analysis of these and other data as described under Experimental Procedures yielded estimates of the apparent second-order inactivation rate constant (k') at six different substrate concentrations. If DFP inactivates PE by modifying an active-site residue, it should be competitive with the substrate. In this case, the value of k' should vary with the substrate concentration as described by eq 2 under Experimental Procedures, and a plot of $1/k'$ against the substrate concentration should yield a straight line. Figure 5B shows that this was indeed the case. Analysis of these data yielded estimates of 825 ± 56 M^{-1} s^{-1} and 53 ± 6 μM for k_i and K_m, respectively. If DFP and the substrate compete with one another, the value determined for the K_m from the data of Figure 5B should correspond with that determined from initial velocity studies. An estimate for K_m of 48 ± 2 μM was obtained from such studies (data not shown), and the good agreement between these two estimates supports the proposal that DFP competes with the substrate. The

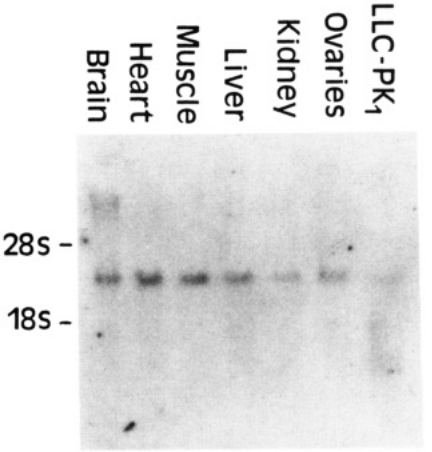

FIGURE 4: Northern analysis of RNA from porcine tissues and from the LLC-PK$_1$ cell line. Twenty micrograms of total RNA from each tissue and from the cell line were fractionated in a formaldehyde/1.0% agarose gel and transferred onto a Zetaprobe membrane. The membrane was hybridized with the insert of the pig liver clone pl1, as described under Experimental Procedures, washed, and autoradiographed at −70 °C. The positions of 18S and 28S rRNA are indicated.

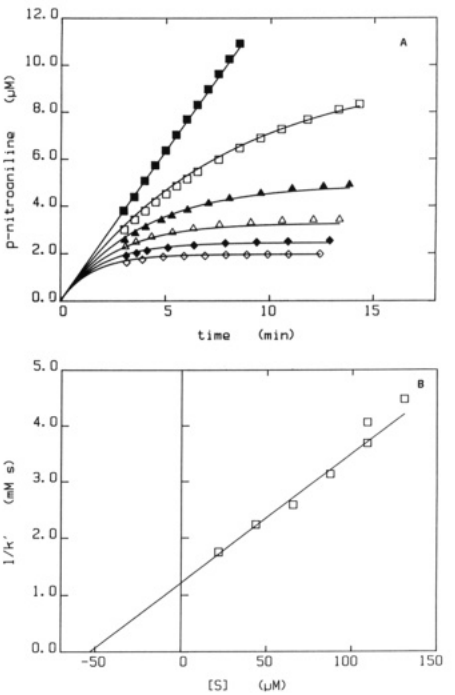

FIGURE 5: Kinetic analysis of the inactivation of PE by DFP. (A) Assays were performed as described under Experimental Procedures with concentrations of 109 μM Z-Gly-Pro-pNA and 0.39 nM enzyme. DFP was present at the following concentrations: 0 (■), 7.9 (□), 15.8 (▲), 23.7 (△), 31.6 (♦), and 39.5 μM (◇). The data were analyzed by nonlinear regression according to eq 1, and the lines drawn show the fit of the data to this equation. Data points prior to 2.8 min have been omitted for the sake of clarity. (B) Dependence of the apparent second-order inactivation rate constant (k') on the substrate concentration. Values of k' obtained from analysis of data such as those shown in Figure 5A were obtained at six different substrate concentrations; the reciprocals of these values are plotted. The data were analyzed by weighted linear regression according to the reciprocal form of eq 2, and the line drawn represents the fit of the data to this equation.

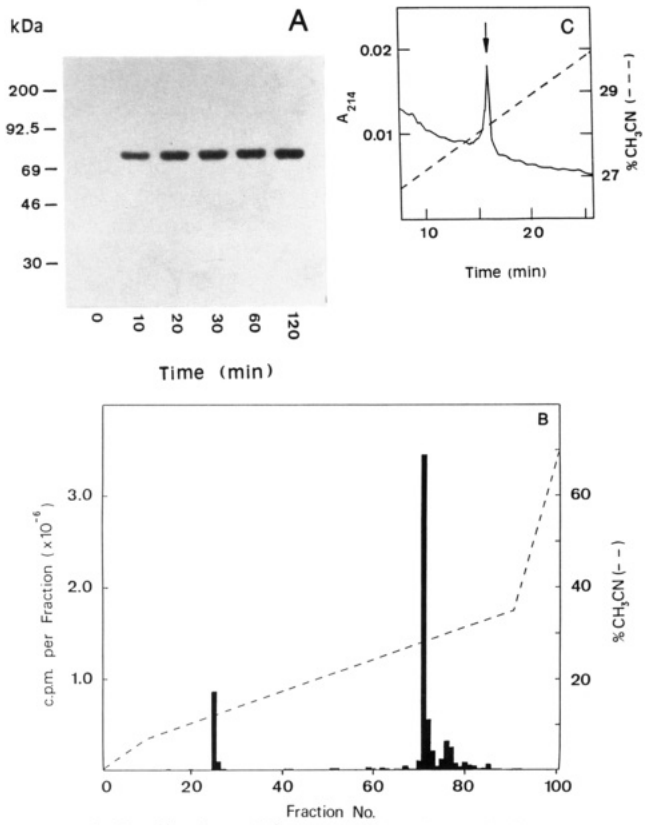

FIGURE 6: Purification of [^{3}H]DFP-labeled peptide from PE. (A) SDS–polyacrylamide gel electrophoresis of [^{3}H]DFP-labeled PE. Samples taken at the indicated times were electrophoresed on a 10% polyacrylamide gel, and the radioactive bands were visualized by fluorography as described under Experimental Procedures. The markers indicate the positions of the standards that were used to estimate the apparent molecular mass of the labeled protein. (B) A chymotryptic digest of [^{3}H]DFP–PE was chromatographed on a C_{18} reverse-phase HPLC column using the indicated linear gradient of acetonitrile in 0.1% (v/v) trifluoroacetic acid (---). The amount of radioactivity contained in each fraction is shown. The radioactive peak at fractions 25–26 eluted at the same position as [^{3}H]DFP. (C) The material from fraction 71 of Figure 6B was rechromatographed on a C_{18} reverse-phase HPLC column using the indicated linear gradient of acetonitrile in 0.1% (v/v) trifluoroacetic acid (---). One radioactive peptide (indicated by the arrow) was detected and analyzed by determination of its amino acid sequence and composition.

estimate of 825 M^{-1} s^{-1} for k_i compares well with that previously determined for the inactivation of PE by DFP at 30 °C (Andrews et al., 1980).

Active-Site Labeling by [^{3}H]DFP. The enzyme was completely inactivated by incubation with a 10-fold molar excess (20 μM) of [^{3}H]DFP for 2 h at 37 °C (data not shown). Analysis by SDS–polyacrylamide gel electrophoresis of samples of the reaction mixture showed one labeled protein that migrated with an apparent molecular mass equivalent to that of PE (Figure 6A). The protein was cleaved with chymotrypsin and fractionated by reverse-phase HPLC, and samples of the fractions were measured for radioactivity (Figure 6B). The recovery of label after chromatography was 84%, and it was distributed over three peaks, one of which contained 65% of the amount of radioactivity that was recovered from the column. One of the minor radioactive peaks, that contained 14.5% of the radioactive label, eluted at the same position as [^{3}H]DFP itself. The major peak was rechromatographed using a shallower gradient of acetonitrile (Figure 6C), and the radioactive peptide material was analyzed by amino acid sequencing and amino acid analysis. The partial amino acid sequence of the peptide, with the yield in picomoles shown in parentheses, was threonine (26), isoleucine (33), asparagine (26), glycine (16), glycine (7.5), and serine (ND). This corresponded to residues 549–554 of the deduced amino acid sequence. The amino acid composition of the peptide (Table II) confirmed the amino acid sequencing data and showed the peptide to comprise residues 549–559 of the deduced amino acid sequence. The incorporation of [^{3}H]DFP into the labeled peptide was 1.48 mol of DFP/mol of peptide. Ser-554 is the only residue that could be reasonably expected to react with [^{3}H]DFP. Moreover, Yoshimoto et al. (1981) have previously shown that the residue in PE that is labeled by [^{32}P]DFP is a serine. Thus, all the data are consistent with Ser-554 being the active-site serine of PE.

Discussion

PE was purified from porcine kidney and oligonucleotides based on the sequences of the tryptic peptides from this protein were used to isolate a cDNA clone that contained the complete coding sequence of PE. The value of the molecular mass calculated from the deduced amino acid sequence (80.7 kDa) is higher than the experimentally determined value of 74.5 kDa (Figure 1). This latter value is similar to those previously obtained for the enzyme isolated from porcine tissues (67–74 kDa; Odya et al., 1987; Moriyama & Sasaki, 1983; Soeda et al., 1984; Hauzer et al., 1984). The reason for this discrepancy is not clear, but it should be pointed out that partial protein sequence data were obtained spanning residues 2–702 of the deduced amino acid sequence which would yield a minimum molecular mass of 79.6 kDa. The identity of the clone was confirmed by the observations that the deduced amino acid sequence contained the sequences of all the tryptic peptides that were sequenced and that the amino acid composition deduced from the cDNA clone was in agreement with that determined experimentally from the purified porcine kidney enzyme. However, the amino acid compositions of both the protein and that deduced from the cDNA clone differ considerably in many residues from those reported by Hauzer et al. (1984) and Moriyama and Sasaki (1983) for porcine PE. Considerable differences also exist between the amino acid compositions determined by these two groups of investigators. The initial steps of the purification procedure employed in this paper were similar to the ones used by Moriyama and Sasaki (1983) and Hauzer et al. (1984). The purification schemes differ in the final steps where we applied chromatofocusing and hydrophobic interaction chromatography (Table I) to obtain a pure protein. The overall degree of purification (4700-fold) was similar to that obtained by Moriyama and Sasaki (1983) for the enzyme from porcine liver (5200-fold). However, the specific activity of our preparation after the final step (12.2 units/mg) is comparable to the highest values obtained (Yoshimoto et al., 1983, 1988), whereas that reported by Moriyama and Sasaki (1983) is approximately 10-fold lower. Hauzer et al. (1984) did not report a specific activity for their purified enzyme. Given these uncertainties, it is possible that the enzyme purified here is different from the ones obtained by Moriyama and Sasaki (1983) and Hauzer et al. (1984).

The enzyme contains 16 half-cystinyl residues. At present, it is not clear whether these occur as free thiols, or whether disulfide bridges are present. However, PE is found in the cytoplasm, and disulfide bridges are rare in cytoplasmic proteins (Schulz & Schirmer, 1979). The formation of disulfide bridges occurs on the luminal side of the endoplasmic reticulum (Freedman, 1984, 1989), and these bonds are generally found in secreted proteins (e.g., trypsin, thrombin, and chymotrypsin).

Partial cDNA clones that encode PE were also isolated from porcine liver and LLC-PK_1 porcine kidney cell libraries. The

nucelotide sequence of the clones from the brain, liver, and kidney cell libraries differed only at three positions. One of these was in the 3′ noncoding region, and the other two were in the third codon position within the coding sequence and did not cause any changes in the deduced amino acid sequence. This sequence data confirm the results of biochemical and immunological studies that indicate that the enzymes from different tissues are indistinguishable (Hersh, 1981; Andrews et al., 1982; Yoshimoto et al., 1988). Northern analysis of RNA from porcine tissues showed that the mRNA encoding PE was present in approximately equal amounts in all tissues examined. The data from Northern analysis are consistent with measurements of the enzyme activity in human, rat, and rabbit tissue extracts which showed that activity was present in all tissues examined and that there was only a slight variation in specific activity among the different tissues (Orlowski et al., 1979; Hersh, 1981; Yoshimoto et al., 1979).

PE is classified as a serine protease on the basis of inactivation by DFP. Kinetic analysis of this inactivation showed that DFP was competitive with the substrate and, thus, established that DFP was modifying an active-site residue. The known serine proteases can be classified into three different families that are named after the representative enzymes: trypsin, subtilisin, and carboxypeptidase Y (Neurath, 1989; Breddam, 1986). Within each family of serine proteases, the primary structures of the enzymes are homologous, and the tertiary structures that have been determined are similar (Wright et al., 1969; Betzel et al., 1988; Remington et al., 1988; Read & James, 1988). Enzymes from different families are not homologous (Wright et al., 1969; Breddam, 1986). The amino acids that surround the active-site seryl residue are conserved in each family, and are Gly-Asp-*Ser*-Gly-Gly-Pro for the trypsin family, Gly-Thr-*Ser*-Met-Ala for the subtilisin family, and Gly-Glu-*Ser*-Tyr-Ala for the carboxypeptidase Y family. These sequences have been used to identify new proteases as serine proteases and to assign them to one of the existing families. The amino acid sequence around the active-site serine of PE was Gly-Gly-*Ser*-Asn-Gly-Gly, which is different from those of the other known families. This difference, together with the lack of overall sequence homology with members of the other three families of enzymes, suggests that PE represents a new type of serine protease. It is interesting to note that although the sequence around the active-site serine in PE is different from those of the known families, it conforms to the consensus sequence Gly-X-*Ser*-X-Gly that was proposed for the active site of mammalian serine proteases and serine esterases by Brenner (1988). The active-site sequence of enzymes of the trypsin family also conforms to this sequence. However, among proteases of the subtilisin and carboxypeptidase Y families, the second glycyl residue is replaced by an alanine.

Although the tertiary structures of enzymes from the trypsin family are different from those of the subtilisin family, the active sites of both families of enzymes contain a catalytic triad of Ser, His, and Asp residues (Kraut, 1977). During catalysis, the hydroxyl group of the serine acts as a nucleophile in the attack on the peptide bond. The histidine, which was first identified by labeling with chloromethane inhibitors (Shaw et al., 1965), acts as a general acid–base catalyst to activate the nucleophilic group and to aid the expulsion of leaving groups. The aspartic acid residue helps to stabilize the charged tetrahedral intermediates formed in the reaction (Fink, 1987). Despite the lack of any other three-dimensional homology, the positions of the amino acids that form the catalytic triad in both the trypsin and sutilisin families are superimposable (Higaki et al., 1987; Betzel et al., 1988). It is unknown whether PE contains the same catalytic triad. It is inactivated by the peptidyl chloromethane inhibitor Z-Gly-Pro-CH_2Cl (Yoshimoto et al., 1977), which suggests that a histidyl residue may also be located in the active site. However, these reagents also react with the active-site cysteine of cysteine proteases and with thiol groups in general (Shaw, 1989). Work is currently underway to identify the site of reaction of Ac-Ala-Ala-ProCH_2Cl with PE, in order to determine if the active site of this enzyme also contains a histidyl residue that could form part of a catalytic triad.

Acknowledgments

We thank Renate Matthies for her expert technical assistance. We also thank Werner Zürcher and Dr. Josef Jiricny for the synthesis of the oligonucleotides and Dr. Yee-Sim Khew-Goodall for help with the Northern analysis.

References

Andrews, P. C., Hines, C. M., & Dixon, J. E. (1980) *Biochemistry 19*, 5494–5500.
Andrews, P. C., Minth, C. D., & Dixon, J. E. (1982) *J. Biol. Chem. 257*, 5861–5865.
Barnes, W. M., Bevan, M., & Son, P. H. (1983) *Methods Enzymol. 101*, 98–122.
Betzel, C., Pal, G. P., & Saenger, W. (1988) *Eur. J. Biochem. 178*, 155–171.
Bradford, M. (1976) *Anal. Biochem. 72*, 248–254.
Breddam, K. (1986) *Carlsberg Res. Commun. 51*, 83–128.
Brenner, S. (1988) *Nature 334*, 528–530.
Chamberlain, J. P. (1979) *Anal. Biochem. 98*, 132–135.
Chomczynski, P., & Sacchi, N. (1987) *Anal. Biochem. 162*, 156–159.
Cohen, J. A., Oosterbaan, R. A., & Berends, F. (1967) *Methods Enzymol. 11*, 686–702.
Crestfield, A. M., Moore, S., & Stein, W. H. (1963) *J. Biol. Chem. 238*, 622–627.
Darnell, J. E. (1982) *Nature 297*, 365–371.
Dresdner, K., Barker, L. A., Orlowski, M., & Wilk, S. (1982) *J. Neurochem. 38*, 1151–1154.
Feinberg, A. P., & Vogelstein, B. (1983) *Anal. Biochem. 132*, 6–13.
Fink, A. L. (1987) in *Enzyme Mechanisms* (Page, M. I., & Williams, A., Eds.) pp 159–177, The Royal Society of Chemistry, London.
Freedman, R. B. (1984) *Trends Biochem. Sci. 9*, 438–441.
Freedman, R. B. (1989) *Biochem. Soc. Trans. 17*, 331–335.
Gray, P. J., & Duggleby, R. G. (1989) *Biochem. J. 257*, 419–424.
Green, G. D. J., & Shaw, E. (1983) *Arch. Biochem. Biophys. 225*, 331–337.
Hauzer, K., Barth, T., Servitova, L., & Jost, K. (1984) *Coll. Czech. Chem. Commun. 49*, 1847–1853.
Hemmings, B. A., Schwarz, M., Rao Advani, S., & Jans, D. A. (1986) *FEBS Lett. 209*, 219–222.
Hersh, L. B. (1981) *J. Neurochem. 37*, 172–178.
Higaki, J. N., Gibson, B. W., & Craik, C. S. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 615–621.
Hofsteenge, J., Taguchi, H., & Stone, S. R. (1986) *Biochem. J. 237*, 243–251.
Hunkapiller, M. W., Lujan, E., Ostrander, F., & Hood, L. E. (1983) *Methods Enzymol. 91*, 227–236.
Kato, T., Okada, M., & Nagatsu, T. (1980) *Mol. Cell. Biochem. 32*, 117–121.
Knecht, R., & Chang, J.-Y. (1987) *Anal. Chem. 58*, 2375–2379.

Koida, M., & Walter, R. (1976) *J. Biol. Chem. 251*, 7593–7599.
Konigsberg, W. H., & Henderson, L. (1983) *Methods Enzymol. 91*, 254–259.
Kozak, M. (1984) *Nucleic Acids Res. 12*, 857–872.
Kraut, J. (1977) *Annu. Rev. Biochem. 46*, 331–358.
Laemmli, U. K. (1970) *Nature 227*, 680–684.
Lathe, R. (1985) *J. Mol. Biol. 183*, 1–12.
Lipman, D. J., & Pearson, W. R. (1985) *Science 227*, 1435–1441.
Lottenberg, R., & Jackson, C. M. (1983) *Biochim. Biophys. Acta 742*, 558–564.
Moriyama, A., & Sasaki, M. (1983) *J. Biochem. (Tokyo) 94*, 1387–1397.
Moriyama, A., Nakanishi, M., & Sasaki, M. (1988) *J. Biochem. (Tokyo) 104*, 112–117.
Neurath, H. (1989) in *Proteolytic Enzymes: a practical approach* (Beynon, R. J., & Bond, J. S., Eds.) pp 1–13, IRL Press, Oxford.
Odya, C. E., Dally, R. D., & Georgiadis, K. E. (1987) *Biochem. Pharmacol. 36*, 39–49.
Orlowski, M., Wilk, E., Pearce, S., & Wilk, S. (1979) *J. Neurochem. 33*, 461–469.
Read, R. J., & James, M. N. G. (1988) *J. Mol. Biol. 200*, 523–551.
Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., & Neurath, H. (1988) *Biochemistry 27*, 8097–8105.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Schulz, G. E., & Schirmer, R. H. (1979) in *Principles of Protein Structure*, Springer-Verlag, Heidelberg.
Shaw, E. (1989) *Adv. Enzymol. Relat. Areas Mol. Biol. 63*, 271–347.
Shaw, E., Mares-Guia, M., & Cohen, W. (1965) *Biochemistry 4*, 2219–2224.
Soeda, S., Ohyama, M., & Nagamatsu, A. (1984) *Chem. Pharm. Bull. 32*, 1510–1516.
Stone, S. R., & Hofsteenge, J. (1985) *Biochem. J. 230*, 497–502.
Taylor, W. L., & Dixon, J. E. (1980) *Biochem. Biophys. Res. Commun. 94*, 9–15.
Tian, W.-X., & Tsou, C.-L. (1982) *Biochemistry 21*, 1028–1032.
Walker, B., & Elmore, D. T. (1984) *Biochem. J. 221*, 277–280.
Walter, R., & Yoshimoto, T. (1978) *Biochemistry 17*, 4139–4144.
Walter, R., Simmons, W. H., & Yoshimoto, T. (1980) *Mol. Cell. Biochem. 30*, 111–127.
Wright, C. S., Alden, R. A., & Kraut, J. (1969) *Nature 221*, 235–242.
Yoshimoto, T., Orlowski, R. C., & Walter, R. (1977) *Biochemistry 16*, 2942–2948.
Yoshimoto, T., Ogita, K., Walter, R., Koida, M., & Tsuru, D. (1979) *Biochim. Biophys. Acta 569*, 184–192.
Yoshimoto, T., Simmons, W. H., Kita, T., & Tsuru, D. (1981) *J. Biochem. (Tokyo) 90*, 325–334.
Yoshimoto, T., Nishimura, T., Kita, T., & Tsuru, D. (1983) *J. Biochem. (Tokyo) 94*, 1179–1190.
Yoshimoto, T., Oyama, H., Koriyama, N., & Tsuru, D. (1988) *Chem. Pharm. Bull. 36*, 1456–1462.
Zolfaghari, R., Baker, C. R. F., Canizaro, P. C., Feola, M., Amirgholami, A., & Behal, F. J. (1986) *Enzyme 36*, 165–178.