Seattle, WA, for supplying the mammalian expression vector.

References

Bar, R. S., Booth, B. A., Boes, M., & Dake, B. L. (1989) *Endocrinology 125*, 1910–1920.

Blum, W. F., Jenne, E. W., Reppin, F., Kietzmann, K., Ranke, M. B., & Bierich, J. R. (1989) *Endocrinology 125*, 766–772.

Czech, M. P. (1989) *Cell 59*, 235–238.

De Meyts, P., Gu, J.-L., Shymko, R. M., Kaplan, B. E., Bell, G. I., & Whittaker, J. (1990) *Mol. Endocrinol. 4*, 409–416.

Ellis, L., Sissom, J., & Levitan, A. (1988) *J. Mol. Recognit. 1*, 25–31.

Graham, F. L., & van der Eb, A. J. (1973) *Virology 52*, 456–467.

Hollenberg, M. D. (1985) in *Insulin—Its Receptor and Diabetes* (Hollenberg, M. D., Ed.) pp 57–83, Marcel Dekker, New York and Basel, Switzerland.

Kobilka, B. K., Kobilka, T. S., Daniel, K., Regan, J. W., Caron, M. G., & Lefkowitz, R. J. (1988) *Science 240*, 1310–1316.

Lax, I., Bellot, F., Howk, R., Ullrich, A., Givol, D., & Schlessinger, J. (1989) *EMBO J. 8*, 421–427.

Marshall, S., Heidenreich, K. A., & Horikoshi, H. (1985) *J. Biol. Chem. 260*, 4128–4135.

Martin, J. L., & Baxter, R. C. (1986) *J. Biol. Chem. 261*, 8754–8760.

Okayama, H., & Berg, P. (1983) *Mol. Cell. Biol. 3*, 280–289.

Scahill, S. J., Devos, R., van der Heyden, J., & Fiers, W. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 4654–4658.

Scatchard, G. (1949) *Ann. N.Y. Acad. Sci. 51*, 660–672.

Seino, S., Seino, M., Nishi, S., & Bell, G. I. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 114–118.

Shibasaki, Y., Sakura, H., Odawara, M., Shibuya, M., Kanazawa, Y., Akanuma, Y., Takaku, F., & Kasuga, M. (1988) *Biochem. J. 249*, 715–719.

Toyoshige, M., Yanaihara, C., Hoshino, M., Kaneko, T., & Yanaihara, N. (1989) *Biomed. Res. 10*, 139–147.

Ullrich, A., Bell, J. R., Chen, E. Y., Herrera, R., Petruzzelli, L. M., Dull, T. J., Gray, A., Coussens, L., Liao, Y.-C., Tsubokawa, M., Mason, A., Seeburg, P. H., Grunfeld, C., Rosen, O. M., & Ramachandran, J. (1985) *Nature (London) 313*, 756–761.

Ullrich, A., Gray, A., Tam, A. W., Yang-Feng, T., Tsubokawa, M., Collins, C., Henzel, W., Le Bon, T., Kathuria, S., Chen, E., Jacobs, S., Francke, U., Ramachandran, J., & Fujita-Yamaguchi, Y. (1986) *EMBO J. 5*, 2503–2512.

Waugh, S. M., DiBella, E. E., & Pilch, P. F. (1989) *Biochemistry 28*, 3448–3455.

Wedekind, F., Baer-Pontzen, K., Bala-Mohan, S., Choli, D., Zahn, H., & Brandenburg, D. (1989) *Biol. Chem. Hoppe-Seyler 370*, 251–258.

Whittaker, J., & Okamoto, A. (1988) *J. Biol. Chem. 263*, 3063–3066.

Yarden, Y., & Ullrich, A. (1988) *Biochemistry 27*, 3113–3119.

Yip, C. C., Hsu, H., Patel, R. G., Hawley, D. M., Maddux, B. A., & Goldfine, I. D. (1988) *Biochem. Biophys. Res. Commun. 157*, 321–329.

Articles

Primary Structure of the Major Pepsin Inhibitor from the Intestinal Parasitic Nematode *Ascaris suum*†

Mark R. Martzen,*,‡ Brad A. McMullen,‡ Nancy E. Smith,§ Kazuo Fujikawa,‡ and Robert J. Peanasky§

Department of Biochemistry, University of Washington, Seattle, Washington 98195, and Department of Biochemistry and Molecular Biology, University of South Dakota School of Medicine, Vermillion, South Dakota 57069

Received December 27, 1989; Revised Manuscript Received May 1, 1990

ABSTRACT: The major pepsin inhibitor from *Ascaris suum* was isolated by affinity chromatography and chromatofocusing. Its amino acid sequence was determined by automated Edman degradation of peptide fragments. Peptides were produced by chemical and enzymatic cleavage of pyridylethylated protein and were purified by reverse-phase high-performance liquid chromatography. The inhibitor consists of 149 residues with the following sequence: QFLFSMSTGP^{10}FICTVKDNQV^{20}FVANLPWTML^{30}EGDDIQVGKE^{40}-FAARVEDCTN^{50}VKHDMAPTCT^{60}KPPPFCGPQD^{70}MKMFNFVGCS^{80}VLGNKLFIDQ^{90}KYVR-DLTAKD^{100}HAEVQTFREK^{110}IAAFEEQQEN^{120}QPPSSGMPHG^{130}AVPAGGLSPP^{140}PPPSFCTVQ^{149}. It has a molecular weight of 16 396. All cysteines are engaged as disulfide bonds: Cys(13)–Cys(59), Cys(48)–Cys(66), and Cys(79)–Cys(146). The protein is probably composed of two domains connected by a short hydrophobic region. This is the first aspartyl protease inhibitor of animal origin that has been sequenced. The sequence has no significant homology with any other known protein.

One-fourth of the world's population and virtually all pigs are infected by the intestinal endoparasitic nematode *Ascaris* (Muller, 1979). This constitutes an important medical, agricultural, and economic problem (Levine, 1980; Mahmoud, 1989). Current control is by chemical antihelminthics directed at adult worms who are capable of propagating themselves by passing hundreds of thousands of eggs each day. As long as infected hosts live in an environment prone to fecal–oral contamination, reinfestation is imminent and chemical de-

† This work was supported by Research Grants AI10992 and HL16919 from the National Institutes of Health.

* Author to whom correspondence should be addressed.

‡ University of Washington.

§ University of South Dakota School of Medicine.

0006-2960/90/0429-7366$02.50/0

worming is of limited use.

It has been suggested that protease inhibitors in the parasite are critical to their survival, perhaps even to their species specificity, and are logical sites to attack to biologically control parasite migration through infested hosts (Martzen et al., 1985, 1986; Peanasky et al., 1987). This approach will interrupt the life cycle of the parasite and prevent its development to the adult stage. To achieve this goal, each of the parasite's inhibitors should be isolated and characterized.

Intestinal parasitic nematodes exist in a hostile proteolytic environment. At the turn of the century, extracts of *Ascaris suum* were found to inhibit the proteolytic activities of crude pancreatic preparations (Weinland, 1903). The presence of "anti-enzymes" (protease inhibitors) in these extracts was proposed to explain the resistance of the parasite to the digestive enzymes of its host. This assumption, however, has never been proven, and alternative functions for the inhibitors have been suggested (Peanasky et al., 1974, 1987; Martzen et al., 1985). Families of isoinhibitor proteins have been isolated from *A. suum*. They inactivate the pancreatic proteases chymotrypsin, elastase, trypsin, and carboxypeptidases A and B and have been recently reviewed (Peanasky et al., 1987). *A. suum* inhibitors of the gastric enzyme pepsin have also been reported (Abu-Erreish & Peanasky, 1974a,b). Of all these inhibitors the pepsin inhibitors are present in the lowest amounts, approximately 0.7% of total inhibitor protein, and have been the most difficult and labor intensive to purify. Consequently, they are the last to undergo primary sequence analysis.

Proteolytic enzymes are classified on the basis of their catalytic mechanisms. All proteases belong to one of four mechanistic classes and are either serine, cysteine, metallo or aspartyl proteases (Barrett, 1980). Aspartyl proteases have been isolated from a variety of sources (Barrett & McDonald, 1980; Bedi et al., 1983; Kay, 1985; Tang & Wong, 1987) and are implicated in the onset of pathological conditions including gastric ulcers, hypertension, and neoplasia. In addition, mutational analysis has recently suggested that the HIV-1 protease is also functionally homologous with aspartyl proteases (Loeb et al., 1989).

All of the aspartyl proteases examined to date are similar in their three-dimensional structures (Subramanian et al., 1977; Hsu et al., 1977; Andreeva et al., 1978) and share many physicochemical properties as recently reviewed (Valler et al., 1985). Despite these similarities, however, each enzyme must also have evolved subtle distinctions in structure to enable it to operate in its own particular microenvironment. In order to investigate these differences and to characterize the catalytic mechanism of action of these proteases, naturally occurring aspartyl protease inhibitors can be a valuable tool. In this paper we report the purification of the major pepsin inhibitor from *A. suum*, designated PI-3,[1] by a new affinity chromatography/chromatofocusing protocol and its complete amino acid sequence.

Experimental Procedures

Materials. Cyanogen bromide was obtained from Pierce. Tributylphosphine and 4-vinylpyridine were obtained from Aldrich. Pyroglutamate aminopeptidase and endoproteinase Asp-N were purchased from Boehringer Mannheim. Endopeptidase Lys-C and ammonium 7-fluorobenzo-2-oxa-1,3-diazole-4-sulfonate were obtained from Wako Pure Chemicals, Osaka, Japan. Commercial pepsin was purified prior to use (Humphreys & Fruton, 1968), and substrate hemoglobin was prepared from outdated red blood cells (Drabkin, 1949). Carboxypeptidase Y was a gift from Dr. M. Ottsen, Carlsberg Laboratories, Copenhagen.

Collection of Parasites. Live *A. suum* were obtained from the abattoir and transported in the salt solution of Baldwin and Moyle (1947). This solution is isosmotic with the parasite's pseudocoelomic fluid. At the lab worms were washed in water and purged for 30 min in 0.9% saline. The body walls of the adult worms were then prepared as described by Peanasky and Laskowski (1960) and stored at −20 °C.

Inhibitor Assay. Pepsin (6 μg) in 15 mM acetate, pH 5, and inhibitor (1–3 μg) were adjusted to pH 2.0 in a final volume of 1 mL and incubated at 37 °C for 5 min. Residual pepsin was determined by hemoglobin digestion as previously described (Abu-Erreish & Peanasky, 1974a). One unit of inhibitor activity was defined as that amount which inactivates 1.0 μg of pepsin under the conditions of the assay. Specific activity was defined as units of inhibitor per 1.0 unit of absorbance of the protein as measured through a 1-cm light path at 280 nm. The protein concentration of crude solutions was determined by the biuret method (Gornall et al., 1949) or by the Lowry et al. (1951) technique. Albumin and pepsin were used to standardize the methods.

Preparation of Affinity Gel. Aminohexyl-Sepharose (AH-Sepharose) was prepared as described by March et al. (1974). CNBr-activated Sepharose 4B (250 mL of packed gel) was reacted with 6.3 g of 1,6-hexanediamine at 4 °C. At 20 h, 15 g of 2-aminoethanol was added to block unreacted sites. Gels were routinely examined for evidence of aminohexyl group coupling by the technique of Inman and Dintzis (1969) as modified by Cuatrecasas (1970). AH-Sepharose was then mixed with 2.5 g of pepsin in 208 mL of water at pH 4.5 and stirred for 5 min at 24 °C with a cone drive stirrer. At this time 125 mg of 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide (EDC) was dissolved in 42 mL of H_2O at pH 4.5. The EDC was added dropwise to the gel slurry with continuous stirring, and the reaction was allowed to continue for 1 h. The reaction was stopped with sequential washes of (1) 6 L of H_2O (pH 4.5), (2) 6 L of 50 mM NaCl (pH 4.5), (3) 6 L of H_2O (pH 4.5), and (4) 1 L of 0.15 M sodium acetate (pH 5.0). The amount of pepsin coupled to the AH-Sepharose gel was then estimated by a hemoglobin digestion assay (Valentova et al., 1975) or by direct titration with known quantities of *A. suum* pepsin inhibitor. One gram of suction-filtered affinity gel bound approximately 1500 units of inhibitor.

Purification of Inhibitors. Cases of severe allergic reaction and anaphylactic shock as a result of exposure to *A. suum* have been reported (Stites et al., 1982) and must be considered in the design of purification protocols. The procedure of Abu-Erreish and Peanasky (1974a) for the initial preparation of pepsin inhibitors was followed with modifications. All steps were at 4 °C unless noted. Frozen *Ascaris* body walls were thawed and homogenized (4 mL of ultrapure H_2O/g of tissue) for 3 min in a Waring Blendor. Three 100-g packages of worms were processed in a batch. The homogenate was stirred for 20 min and then centrifuged at 28000*g* for 15 min. The supernatant was centrifuged in a type 19 rotor 54000*g* in a Beckman L8-80 ultracentrifuge.

[1] Abbreviations: PI-3, the major pepsin inhibitor from *Ascaris suum*; PE-PI-3, pyridylethylated PI-3; kDa, kilodalton(s); EDC, 1-ethyl-3-[3-(dimethylamino)propyl]carbodiimide; CF, chromatofocusing; PBE, polybuffer exchanger gel; PAGE, polyacrylamide gel electrophoresis; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); HPLC, high-performance liquid chromatography; PTH, phenylthiohydantoin; CpY, carboxypeptidase Y; TFA, trifluoroacetic acid; DTE, dithioerythritol; SBD-F, ammonium 7-fluorobenzo-2-oxa-1,3-diazole-4-sulfonate; CD, circular dichroism.

The ultracentrifuge supernatant was filtered through Whatman No. 1 paper, and the last 15% of supernatant in each tube repelleted at 280000*g* for 150 min in the L8-80 ultracentrifuge. The ultracentrifuge supernatants were then pooled, adjusted to pH 1.9 with 6 N HCl, and incubated at 37 °C for 75 min. The incubated solution was cooled to 25 °C on ice and adjusted to pH 5.5 with 10 N NaOH. The resulting precipitate was removed by centrifugation at 28000*g* for 10 min.

The clarified supernatant was adjusted to 0.65 saturation with ammonium sulfate at 0 °C (Noda & Kuby, 1957), the pH after salt addition was adjusted to pH 5.3, and the solution was held overnight at 4 °C. Precipitated protein was pelleted at 28000*g* for 20 min. The pellet was then dissolved in a minimal volume of H_2O at pH 2.1, Tris base added to a final concentration of 100 mM, and the pH readjusted to 2.1. The role of Tris is unknown, but it stabilizes the inhibitor significantly.

Any precipitate that remained after redissolving was removed by centrifuging at 48000*g* for 20 min. The final protein solution was adjusted to 1/15 the original volume of the combined ultracentrifuge supernatants and was centrifuged at 280000*g* for 150 min and stored at −17 °C. A sufficient amount of pepsin–AH-Sepharose affinity gel, determined by its binding capacity, was then washed for 25 min with a 10-fold (w/v) excess of 0.1 M Tris/0.1 M sodium acetate adjusted to pH 1.95. Washing was repeated twice. Crude *A. suum* pepsin inhibitor was adjusted to a minimal concentration of 100 units/mL at pH 1.95. Inhibitor and gel were stirred at pH 1.95 for 5 min. The gel slurry was then adjusted to pH 5.0 by dropwise addition of 1 N NaOH and adsorption continued for 30 min at 4 °C. It is not clear why this procedure is necessary, but any other protocol resulted in a poor release of inhibitor from pepsin. Affinity gel was separated from the suspension by filtration on a medium sintered glass funnel. Nonspecifically adsorbed protein was removed by three washes with a 10-fold (w/v) excess of 0.1 M Tris/0.1 M sodium acetate/0.2 M NaCl (pH 3.5), followed by a H_2O wash.

Inhibitor was desorbed by three 1-h elutions with 0.1 M Tris/1 M NaCl/0.5 M ethanolamine (pH 10) at 37 °C. Each eluate was adjusted to pH 1.95 with 2.5 N HCl, centrifuged at 5000*g* for 10 min to remove any precipitate, and readjusted to pH 6.0 with 2.5 N NaOH.

The pooled inhibitor solution was concentrated to a minimal volume under nitrogen at 25 psi in an Amicon cell fitted with a YM-5 membrane. The protein solution was diluted with H_2O and concentrated three more times. *A. suum* pepsin inhibitor proteins in H_2O were then made 0.025 M in imidazole hydrochloride at pH 7.5 for chromatofocusing. Polybuffer exchanger gel, PBE 94, was equilibrated in 0.025 M imidazole hydrochloride (pH 7.5). The chromatofocusing column (CF column) was packed and equilibrated at 4 °C. The inhibitor sample was then gravity charged onto the column. Polybuffer 74 (25 mL) was adjusted to pH 4.0 with 0.5 N HCl, brought to a volume of 200 mL, degassed, and applied to the CF column. When Polybuffer was exhausted, residual protein was released with 1 M NaCl and the column was reequilibrated. Focused inhibitor peaks were concentrated by ultrafiltration.

Removal of Polybuffer Ampholines from PI-3. The concentrated fractions were adjusted to 0.85 saturation with solid ammonium sulfate, placed on ice for 30 min, and centrifuged at 48000*g* for 30 min (Yoshimoto & Laskowski, 1982). Precipitates were washed twice with 0.85 saturated ammonium sulfate solution and then dissolved in H_2O, desalted, and concentrated by ultrafiltration. This removed in excess of 96% of the ampholines present. In addition, pepsin inhibitor 3 (PI-3) used for primary sequence analysis was further chromatographed on AG-501-X8(D) bifunctional ion-exchange resin (Baumann & Chrambach, 1975). A resin column (1 × 10 cm) was equilibrated with 0.001 N HCl at 4 °C and 10500 units of inhibitor applied. Fractions of 1.75 mL were collected at 38 mL/h and assayed for inhibitor. Some residual ampholines from the previous step were still detected at 230 nm, well separated from inhibitor.

Polyacrylamide Gel Electrophoresis. Purified proteins were analyzed by nondissociating, discontinuous polyacrylamide gel electrophoresis (PAGE) as described by Ornstein (1964) and Davis (1964). The high-pH buffer system used stacked at pH 8.3 and separated at pH 9.5 on the 12.5% gels. Inhibitors were visualized by staining with 1% Naphthol Blue Black in 7% acetic acid. Destaining was in 7% acetic acid. Visualized protein bands were positively identified as pepsin inhibitor by adaptation of the cytochemical staining technique of Furihata et al. (1975). CF column fractions which contained more than one inhibitor species when analyzed by PAGE were further resolved by DEAE-Sephadex chromatography as previously described (Abu-Erreish & Peanasky, 1974a).

Sequence Determination. Pepsin inhibitor 3 (PI-3) was analyzed for the presence of free sulfhydryl groups with 5,5′-dithiobis(2-nitrobenzoic acid) in the presence of SDS, urea, or both (Ellman, 1959). Reduction and pyridylethylation of PI-3 with 4-vinylpyridine was performed by the procedure of Friedman et al. (1970). PE-PI-3 (1 mg) was digested overnight at room temperature in 0.5 mL of 70% formic acid containing 2% CNBr. Digestion of PE-PI-3 (1 mg) with endopeptidase Lys-C, or endoproteinase Asp-N, was performed by incubating the protein in 0.5 mL of 0.1 M NH_4HCO_3 overnight at 37 °C with 1:1000 (w/w) or 1:100 (w/w) enzyme, respectively. Pyroglutamate aminopeptidase digestion of intact PI-3 (0.5 mg) was performed in 0.5 mL of 100 mM potassium phosphate/10 mM EDTA/5 mM DTE/5% glycerol (pH 8.0). A ratio of 1 unit of enzyme/mg of protein was used and the mixture incubated at 37 °C overnight. The resulting peptides from the above digests were separated on a Waters HPLC system using an Altex Ultrapore RPSC column (0.46 × 7.5 cm) or a Waters μBondapak C_{18} column (0.49 × 30 cm). The carboxyl-terminal sequence of PI-3 was verified with carboxypeptidase Y (CpY) as described by Hayashi (1977). The amino acid compositions of the peptide hydrolysates (6 N HCl, 110 °C for 24 h) and CpY digests were determined by a Waters Picotag system according to the method of Bidlingmeyer et al. (1984). Sequence analyses were performed with an Applied Biosystems 477A protein sequencer connected to an on-line PTH analyzer (Model 120A). Repetitive yields of 90–95% were routinely obtained.

Determination of Disulfide Bridges. Cystine-containing peptides were detected by the procedure of Sueyoshi et al. (1985). Fluorescence intensities were measured in a Perkin-Elmer LS-5 fluorescence spectrophotometer with excitation at 385 nm and emission at 515 nm. Sequence analyses of cystine-containing peptides were performed as described by Marti et al. (1987).

Circular Dichroism Spectra. CD spectra of PI-3 were taken with a Jasco J-500 A spectropolarimeter with Jasco DP-501N data processor and TRIO Model CO-1530 A oscilloscope. Protein (1.5 mg/mL) was dissolved in 50 mM Tris/150 mM NaCl (pH 7.4). Dichroically neutral quartz cylindrical cuvettes were used, and CD spectra were the average of eight accumulations smoothed by the data processor. Far-UV

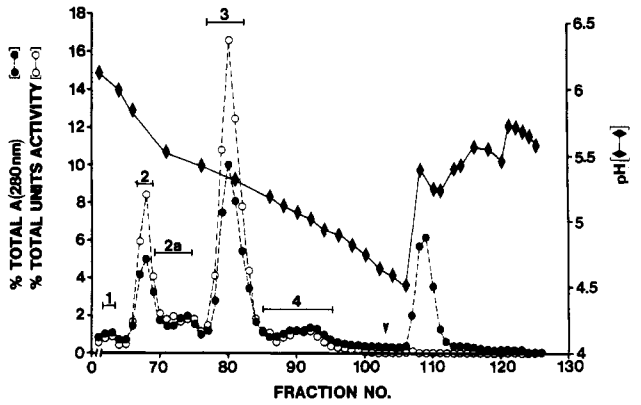

FIGURE 1: Affinity-adsorbed *A. suum* pepsin inhibitors released by 0.5 M ethanolamine/0.1 M Tris/1 M NaCl (pH 10.0) buffer were separated by chromatofocusing. Released pepsin inhibitor (63 840 units) was desalted, concentrated to 27 mL, made 0.025 M in imidazole hydrochloride (pH 7.5), and then applied at 4 °C to a Polybuffer exchanger chromatofocusing column (1 × 40 cm) equilibrated with the same buffer. After the protein (total A_{280nm} = 20.7) was applied to the column with the imidazole buffer, elution was begun with Polybuffer 74 (25 mL) adjusted to pH 4 and diluted to 200 mL. Flow rate was 21.5 mL/h, and each fraction collected was 2.6 mL. At the arrow 1 M NaCl was applied. Five pepsin inhibitor peaks were released. The protein peak released by 1 M NaCl did not inhibit pepsin.

(190–240 nm) was scanned for secondary structure and near-UV (250–300 nm) for the CD of aromatic amino acids. All measurements were at ambient temperature.

Results and Discussion

The overall recovery of pepsin inhibitors by the present affinity chromatography/chromatofocusing protocol (Table I) was 70% compared to 44% by the technique of Abu-Erreish and Peanasky (1974a). This increased yield was obtained by modifying the ultracentrifugation step and by replacing the sequential column chromatography steps on Bio-Gel P-30, Cellex-SE, and DEAE-cellulose with pepsin–AH-Sepharose affinity chromatography. In addition, introduction of the affinity column reduced the purification time of the inhibitors from months to weeks and greatly minimized exposure of laboratory personnel to crude *A. suum* extracts.

Inhibitor released from the affinity gel had an average specific activity of 3070 units/mg of protein. This is a 750-fold increase in purification over the previous step and is comparable to the material of Abu-Erreish and Peanasky (1974a) after sequential column chromatography on Cellex-SE and DEAE-Sephadex. Affinity chromatography, unlike DEAE-Sephadex chromatography, did not separate the inhibitors into homogeneous fractions. This was undertaken by chromatofocusing (CF) chosen as a separation technique for its speed and reported high resolution and recoveries. Ten separate CF columns were run with recoverable activity consistently in excess of 90%. Pepsin inhibitor, isolated by desorption from an affinity gel, had a reproducible chromatofocusing profile (Figure 1). Assay of the CF columns indicated the following: two major inhibitor peaks (2 and 3), two minor inhibitor peaks (1 and 4), a shoulder of activity on peak 2 (2a), and a protein peak released with 1 M NaCl at the end of the pH gradient with no inhibitor activity.

Table I: Recovery of Pepsin Inhibitors from 20 kg of *A. suum*

step	total units	% recovery	specific activity
22500*g* supernatant	568 000	(100)	0.6
ultracentrifuge supernatants	562 000	99	0.8
pH 5.5 supernatant	561 500	99	1.8
0–65% $(NH_4)_2SO_4$	539 500	95	4.1
affinity chromatography	437 000	77	3070
chromatofocusing column (total)	397 500	70	
peak 1	5 500	1	2900
peak 2	79 500	14	4200
peak 3	261 000	46	4240
peak 4	51 000	9	2800

Individual protein peaks from 10 columns were combined and analyzed on 12.5% PAGE gels at pH 9.5 (not shown). Only peak 3, identified as the major pepsin inhibitor (PI-3), was homogeneous. All of the other peaks were cross contaminated to various degrees. Polybuffer ampholines were removed from each peak, and the nonhomogenous peaks were refocused through various pH ranges. This selectivity enriched individual inhibitors but still failed to completely resolve them. This indicated that the chromatofocusing step, while rapid and successful for the major pepsin inhibitor (PI-3), was not a suitable substitute for DEAE-Sephadex resolution. Subseqeunt resolution of the inhibitors on DEAE-Sephadex after affinity chromatography (not shown) resulted in homogeneous inhibitor proteins in the following distribution: PI-1 (1%), PI-2 (20%), PI-3 (66%), and PI-4 (13%).

The primary structure of PI-3 was then studied by using homogeneous inhibitor isolated by chromatofocusing. Automated Edman degradation of intact PE-PI-3 failed to detect sequence, indicating a blocked NH_2 terminal. Amino acid analysis of the intact protein (Table II) revealed 6 Met residues/protein, and digestion by CNBr was chosen. Five fragments were isolated by HPLC (Figure 2A). Four of these fragments were sequenced and found to be homogeneous. The

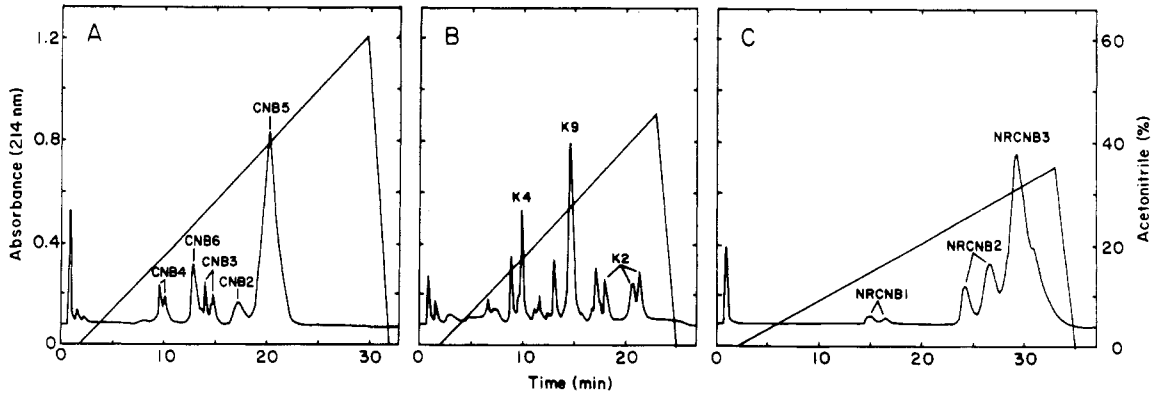

FIGURE 2: Separation of peptide fragments obtained in three digests of *A. suum* pepsin inhibitor 3. Peptides were separated with an Altex Ultrapore C3 reverse-phase separation column (0.46 × 7.5 cm) connected to a Waters HPLC system. The gradient was composed of 0.1% trifluoroacetic acid (solvent A) and 0.08% trifluoroacetic acid in 80% acetonitrile (solvent B). Flow rate was 1.5 mL/min. (A) CNBr digest of PE-PI-3 (CNB); (B) endopeptidase Lys-C digest of PE-PI-3 (K); (C) CNBr digest of intact (nonreduced) PI-3 (NRCNB).

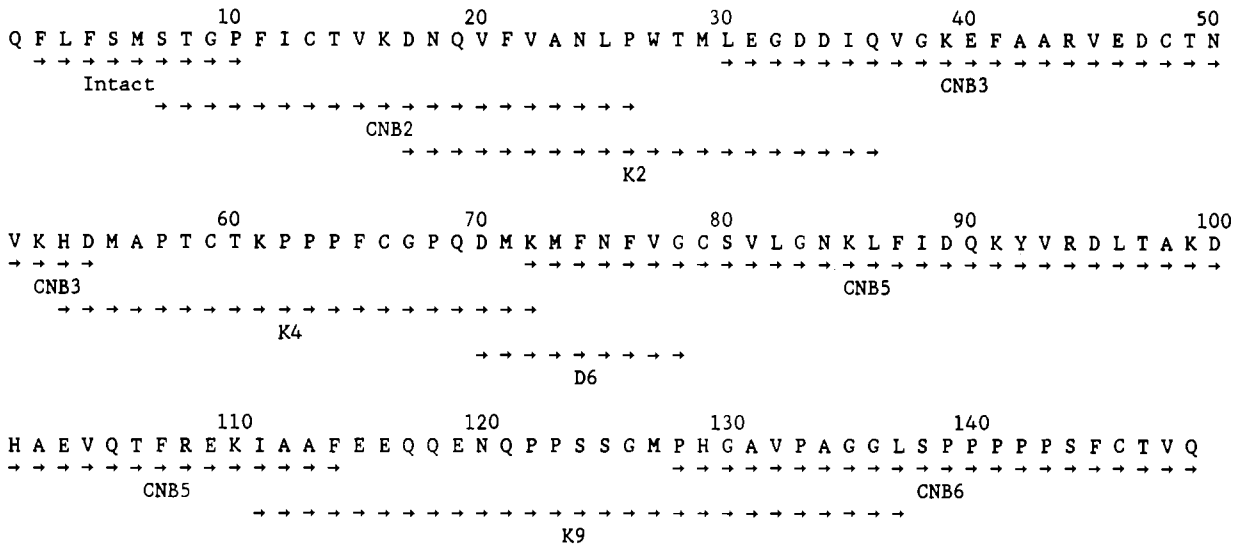

FIGURE 3: Summary of proof of sequence of *A. suum* PI-3. Results marked with (→) were identified by Edman degradation. Intact = pyroglutamate aminopeptidase treated PI-3; CNB = cyanogen bromide generated peptides; K = endopeptidase Lys-C generated peptides; D = endoproteinase Asp-N generated peptide.

Table II: Amino Acid Composition of Pepsin Inhibitor 3 from *A. suum*

amino acid	mol of residues/mol of protein: from acid hydrolysates[a]	mol of residues/mol of protein: from amino acid sequence
aspartic acid (D + N)	15.4	15
threonine	8.7	9
serine	6.9	7
glutamic acid (E + Q)	18.2	18
proline	17.2	16
glycine	9.5	10
alanine	10.1	10
half-cystine	4.8[b]	6
valine	11.9	12
methionine	5.4	6
isoleucine	4.2	4
leucine	7.8	7
tyrosine	1.2	1
phenylalanine	12.5	12
histidine	2.9	3
lysine	9.5	9
arginine	3.1	3
tryptophan	ND[c]	1
total amino acids		149
M_r		16396

[a] Average of duplicate determinations. Calculated by assuming a molecular weight of 16500. [b] Determined directly. [c] Not determined.

results are shown in Figure 3. The fifth fragment (CNB5) was also subjected to sequence analysis and a ragged NH_2 terminus found. The major sequence of this fragment began at position 74 and the minor sequence began at position 72. Interpreting the chromatograms of each cycle was not difficult, since the ratio of the two sequences was 7:3 and the minor sequence disappeared after 20 cycles.

Amino acid analysis identified CNB6 as the COOH-terminal peptide as no homoserine was detected in the sample. A carboxypeptidase Y digestion of CNB6 rapidly released Gln and Val together, followed by the release of a group of three amino acids: Phe, *S*-β-(4-pyridylethyl)cysteine, and Thr. No additional amino acids were detected after 30 or 120 min (not shown). These results agreed with those obtained by sequence analysis of CNB6 and indicated that CNB6 was completely sequenced and that the carboxy-terminal amino acid is Gln.

The peptide from the NH_2-terminal (CNB1) was not isolated but was found in a CNBr digest of intact PI-3 (Figure 2C). Amino acid analysis of NRCNB1 (nonreduced CNB1) had the following composition: Z_1, S_1, L_1, F_2, and homo-S_1. The presence of Z in the amino acid composition provided a simple explanation for the blocking of NRCNB1. If Z was a Gln located at the amino terminus, it could have cyclized to form pyroglutamic acid and thus would not provide a free α-amino group for sequence analysis. To verify this assumption, intact PE-PI-3 was digested with pyroglutamate aminopeptidase, the mixture was desalted by HPLC, and the sequence was analyzed. Phe was found at the NH_2 terminal, and an overlap was obtained with CNB2. As a confirmation, an additional fraction of *Ascaris* pepsin inhibitors from a separate purification, which did not use the chromatofocusing step, was directly subjected to sequencing. This sample contained a Gln at its amino terminus, indicating that the material that we used was blocked during its isolation. The remaining overlaps were constructed by isolating peptides from an endopeptidase Lys-C digest of PE-PI-3. Figure 2B shows the HPLC chromatogram of this digest, and three peptides (K2, K4, and K9) were used for the sequence proof. Unfortunately, K4 had only a single residue overlap with CNB5. A digest with endoproteinase Asp-N was subsequently analyzed. After the digestion products were separated by HPLC (not shown), the peptide of interest was sequenced. Peptide D6 began at Asp (70) and continued to Gly (78), strengthening the overlap between K4 and CNB5 (Figure 3).

Attempts to react PI-3 with Ellman's reagent (DTNB) indicated no free sulfhydryls. This suggested that all six Cys residues in PI-3 were involved in forming disulfide bonds. To determine the bond pairing, intact PI-3 was cleaved with CNBr and the CNBr fragments were separated by HPLC (Figure 2C). In addition to the blocked NH_2-terminal peptide (NRCNB1), fragments NRCNB2 and NRCNB3 were isolated. Both fragments contained cystine/2 residues as determined by fluorometric assay with SBD-F. Automated Edman degradation of NRCNB2 indicated that Cys (13) was paired to Cys (59) and Cys (48) was paired to Cys (66) (Table III). The third disulfide bond was located in NRCNB3, and automated Edman degradation confirmed that Cys (79) was paired with Cys (146).

Initial attempts at secondary structure estimates of PI-3 were then made. Circular dichroism spectra of PI-3 in both the far-UV and near-UV regions were taken (not shown). CD spectra between 190 and 240 nm suggested the presence of α-helix in the inhibitor's secondary structure with high negative

Table III: Identification of Disulfides in PI-3

peptide	structural data[a]	disulfide bond, Cys–Cys
NRCNB2	7 S T G P F I C T V K D N Q V F V A N L P 26 56 A P T X T K P P P F X G P Q D 70 30 L E G D D I Q V G K E F A A R V E D C T N V 51	13–59 and 48–66
NRCNB3	74 F N F V G X S V L G N K L F I D Q K Y V R D 95 128 P H G A V P A G G L S P P P P P S F C T V Q 149	79–146

[a] Standard abbreviations for the amino acids were used except that X denotes a Cys residue expected from the known sequence but not observed and C denotes a di-Cys residue.

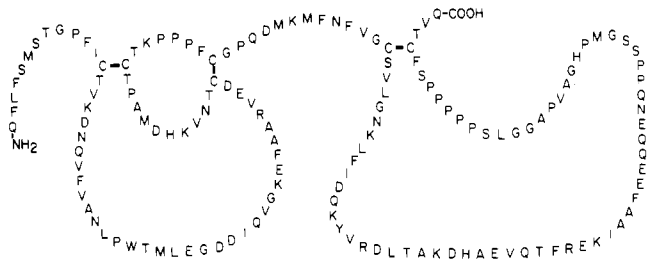

FIGURE 4: Amino acid sequence for *A. suum* PI-3 with disulfide bridge pairing.

ellipticity at 208 nm and high positive ellipticity at 193 nm. This was a surprising observation as all of the other classes of *Ascaris* protease inhibitors analyzed to date appear to be β-sheets and random coils with β-turns by the Chou and Fasman (1978) algorithms.

CD spectra between 250 and 320 nm were also analyzed. Ellipticities in the 250–320-nm region arise from circular dichroism associated with aromatic amino acids and disulfide bonds (Kahn, 1979). CD spectra of PI-3 in this region confirmed the presence of both the aromatic amino acids, Phe, Tyr, and Trp, and disulfide bonds in the molecule.

Hydropathy (Kyte & Doolittle, 1982) and net charge distribution (Hopp & Woods, 1981) of the PI-3 molecule were also analyzed. These analyses indicated two broadly hydrophilic regions of the molecule connected by a short hydrophobic stretch of amino acids with two regions of high negative charge distribution, each corresponding to one of the hydrophilic domains.

A preliminary computer representation of PI-3 after energy minimization (Brooks et al. 1983) using the Polygen/CHARMm program and the QUANTA system for graphical modeling was also made. An elongated molecule of two domains was predicted, agreeing with the hydropathy and net charge analysis and the disulfide bridge assignments. A two-dimensional representation of the amino acid sequence of PI-3 with its disulfide bridge assignments is shown in Figure 4.

As more detailed information about the molecule is obtained, by techniques such as epitope mapping, a more refined estimation of the molecule will be possible. At present, the complete molecule is too large for 2D NMR; however, X-ray crystallographic analysis of the molecule is underway.

In summary, the *A. suum* PI-3 inhibitor was isolated by a new affinity chromatography/chromatofocusing protocol. Analysis of this molecule, the only aspartyl protease inhibitor of animal origin sequenced to date, suggests that it is composed of two domains and that these domains are separated in space on the basis of charge. Experiments to resolve the structure of the molecule and to identify its inhibitory reactive site are currently in progress.

Added in Proof

It has just come to our attention that a protein from potatoes has been reported to inhibit both trypsin and cathepsin D (Mareš et al., 1989). This potato protein does not show significant sequence homology with the inhibitor from *Ascaris*. The *Ascaris* inhibitor does not inhibit trypsin (Abu-Erreish & Peanasky, 1974a) or cathepsin D (Valler et al., 1985).

Acknowledgments

Specimens of *Ascaris suum* for this study were collected with the kind permission of John Morrell & Co. (Sioux Falls, SD) at their abattoir. We thank Drs. John Erickson and Gene Homandberg for collecting the CD data and Dr. Don Babin for some complete amino acid composition analyses. We express our gratitude to Dr. Earl Davie for encouragement and assistance and to Drs. Ralph Bottenus, Steve Leytus, and Joost Meijers for valuable discussions and advice. We also thank Lois Swenson and Dianna Olson for their assistance in the preparation of the manuscript.

Supplementary Material Available

A table showing the PTH-amino acid yield at each Edman cycle of each peptide used in determining the sequence (1 page). Ordering information is given on any current masthead page.

References

Abu-Erreish, G. M., & Peanasky, R. J. (1974a) *J. Biol. Chem. 249*, 1558–1565.

Abu-Erreish, G. M., & Peanasky, R. J. (1974b) *J. Biol. Chem. 249*, 1566–1571.

Andreeva, N. S., Federov, A. A., Gutschina, A. E., Riskulov, R. R., Schutzkever, N. E., & Safro, M. G. (1978) *Mol. Biol. U.S.S.R. (Engl. Transl.) 12*, 922–935.

Baldwin, E., & Moyle, V. (1947) *J. Exp. Biol. 23*, 227–291.

Barrett, A. J. (1980) in *Protein Breakdown in Health and Disease* (Evered, D., & Whelan, J., Eds.) Ciba Foundation Symposium, Vol. 75, pp 1–13, Excerpta Medica, Amsterdam.

Barrett, A. J., & McDonald, J. K. (1980) in *Mammalian Proteases: A Glossary and Bibliography*, Vol. 1, pp

303–356, Academic Press, New York.
Baumann, G., & Chrambach, A. (1975) *Anal. Biochem. 69*, 649–651.
Bedi, G. S., Balwierczok, J., & Back, N. (1983) *Biochem. Pharmacol. 32*, 2071–2077.
Bidlingmeyer, B. A., Cohen, S. A., & Tarvin, T. C. (1984) *J. Chromatogr. 336*, 93.
Brooks, B. R., Brucceleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.
Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.
Cuatrecasas, P. (1970) *J. Biol. Chem. 245*, 3059–3065.
Davis, B. J. (1964) *Ann. N.Y. Acad. Sci. 121*, 404–427.
Drabkin, D. L. (1949) *Arch. Biochem. Biophys. 21*, 224–232.
Ellman, G. L. (1959) *Arch. Biochem. Biophys. 82*, 70–77.
Friedman, M., Krull, L. H., & Cavins, J. F. (1970) *J. Biol. Chem. 245*, 3868–3871.
Furihata, C., Sasajima, K., Kazama, S., Kogure, K., Kawachi, T., Sugimura, T., Tatematsa, M., & Takahashi, M. (1975) *J. Natl. Cancer Inst. 55*, 925–929.
Gornall, A. G., Bardawill, C. J., & David, M. M. (1949) *J. Biol. Chem. 177*, 751–766.
Hayashi, R. (1977) in *Methods in Enzymology* (Hirs, C. H. W., & Timasheff, S. N., Eds.) Vol. 47, pp 84–93, Academic Press, New York.
Hopp, T. P., & Woods, K. R. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 3824–3828.
Hsu, I. N., Delbaere, L. T. J., & James, M. N. G. (1977) *Nature (London) 266*, 140–145.
Humphreys, R. E., & Fruton, J. S. (1968) *Proc. Natl. Acad. Sci. U.S.A. 59*, 519–525.
Inman, J. K., & Dintzis, H. M. (1969) *Biochemistry 8*, 4074–4082.
Kahn, P. C. (1979) in *Methods in Enzymology* (Hirs, C. H. W., & Timasheff, S. N., Eds.) Vol. 61, pp 339–378, Academic Press, New York.
Kay, J. (1985) in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., Ed.) pp 1–18, de Gruyter, Berlin.
Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.
Levine, N. D. (1980) *Nematode Parasites of Domestic Animals and of Man*, pp 256–295, Burgess Publishing Co., Minneapolis, MN.
Loeb, D. D., Hutchison, C. A., III, Edgell, M. H., Farmerie, W. G., & Swanstrom, R. (1989) *J. Virol. 63*, 111–121.
Lowry, O. H., Rosebrough, N. J., Farr, L. A., & Randall, R. J. (1951) *J. Biol. Chem. 193*, 265–275.
Mahmoud, A. A. F. (1989) *Science 246*, 1015–1022.
March, S. C., Parikh, I., & Cuatrecasas, P. (1974) *Anal. Biochem. 60*, 149–152.
Mareš, M., Meloun, B., Pavlík, M., Kostka, V., & Baudyš, M. (1989) *FEBS Lett. 251*, 94–98.
Marti, T., Rosselet, S. J., Titani, K., & Walsh, K. A. (1987) *Biochemistry 26*, 8099–8109.
Martzen, M. R., Geise, G. L., Hogan, B. J., & Peanasky, R. J. (1985) *Exp. Parasitol. 60*, 139–149.
Martzen, M. R., Geise, G. L., & Peanasky, R. J. (1986) *Exp. Parasitol. 61*, 138–145.
Muller, R. (1979) in *Parasites and Western Man* (Donaldson, R. J., Ed.) p 90, University Park Press, Baltimore.
Noda, L., & Kuby, S. A. (1957) *J. Biol. Chem. 226*, 541–549.
Ornstein, L. (1964) *Ann. N.Y. Acad. Sci. 121*, 321–349.
Peanasky, R. J., & Laskowski, M. (1960) *Biochim. Biophys. Acta 37*, 167–169.
Peanasky, R. J., Abu-Erreish, G. M., Gaush, C. R., Homandberg, G. A., O'Heeron, D., Linkenheil, R. K., Kucich, U., & Babin, D. R. (1974) in *Bayer Symposium V: Proteinase Inhibitors* (Fritz, H., Tschesche, H., Greene, L. J., & Truscheit, E., Eds.) pp 649–666, Springer-Verlag, Berlin.
Peanasky, R. J., Martzen, M. R., Homandberg, G. A., Cash, J. M., Babin, D. R., & Litweiler, B. (1987) in *Molecular Paradigms for Eradicating Helminthic Parasites*, pp 349–366, Alan R. Liss, New York.
Stites, D. P., Stobo, J. D., Fudenberg, H. H., & Wells, J. V. (1982) *Basic and Clinical Immunology*, p 685, Lange Medical Publications, Los Altos, CA.
Subramanian, E., Swan, I. D. A., Liu, M., Davies, D. R., Jenkins, J. A., Tickle, I. J., & Blundell, T. L. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 556–559.
Sueyoshi, T., Miyata, T., Iwanaga, S., Toyo'oka, T., & Imai, K. (1985) *J. Biochem. (Tokyo) 97*, 1811–1813.
Tang, J., & Wong, R. N. S. (1987) *J. Cell. Biochem. 33*, 53–63.
Valentova, O., Turkova, J., Lapka, R., Zima, J., & Coupek, J. (1975) *Biochim. Biophys. Acta 403*, 192–196.
Valler, M. J., Kay, J., Aoyagi, T., & Dunn, B. M. (1985) *J. Enzyme Inhib. 1*, 77–82.
Weinland, E. (1903) *Z. Biol. 44*, 1–15.
Yoshimoto, M., & Laskowski, M., Sr. (1982) *Prep. Biochem. 12*, 235–254.