0273-2289/00/88/0159/$14.00

Enteropeptidase

Structure, Function, and Application in Biotechnology

ANNA G. MIKHAILOVA* AND LEV D. RUMSH

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, GSP-7, Moscow, 117871 Russia, E-mail: anna@enzyme.siobc.ras.su

Abstract

A preparative method for purification of enteropeptidase (enterokinase) (EC 3. 4. 21. 9) is developed. A highly purified form of this enzyme is stabilized by calcium ions and does not contain any other proteolytic enzyme contaminations. These enteropeptidase preparations were successfully used for cleavage of a variety of fusion proteins containing the tetraaspartyl-lysyl sequence in an arbitrary position on the polypeptide chain. A series of substrates was methodically studied, which resulted in the suggestion that the peptide and fusion protein substrates (K_m = 200 μ*M* and 125 μ*M*, respectively) were bound to the enzyme through the linker $(Asp)_4Lys$ at the binding site on the light chain of enteropeptidase. Much more efficient hydrolysis of the natural substrate trypsinogen (K_m = 2.4 μ*M*) testifies to a significant contribution of other sites of the substrate and the enzyme in productive binding. A variation in the enzyme's unique specificity was shown to be a result of the autolysis caused by the loss of calcium ions; the cleavage sites were determined. The truncated enzyme containing the C-terminal fragment 466–800 of its heavy chain and the intact light chain does not distinguish the natural substrate trypsinogen, fusion protein, or peptide substrates. These results suggest that the N-terminal fragment 118–465 of the enteropeptidase heavy chain contains a secondary substrate-binding site that interacts directly with trypsinogen.

Index Entries: Enteropeptidase (enterokinase); fusion (recombinant) proteins; autolysis; specificity; purification; affinity chromatography; inhibitors.

Abbreviations: BPI, basic bovine protease inhibitor; STI, soybean trypsin inhibitor; MPG-PA-2000, porous glass chemically coated with poly (4-nitrophenyl acrylate); OG, *n*-octyl-β-D-glucopyranoside; BAPNA, α-*N*-bensoyl-DL-arginine 4-nitroanilide; GD_4K-NA, Gly-Asp-Asp-Asp-Asp-Lys-β-naphthylamide; Gdn-Bz-ONp, 4'-guanidinobenzoate 4-nitrophenyl ester; Gdn-Bz-OMum, 4'-guanidinobenzoate 4-methylumbelliferil ester.

*Author to whom all correspondence and reprint requests should be addressed.

Introduction

Enteropeptidase (enterokinase) (EC 3.4.21.9), a highly specific processing protease, occupies the top of a cascade of reactions activating the digestion enzymes *(1)*. Its first reaction is the conversion of trypsinogen into trypsin. Enteropeptidase, catalyzing trypsinogen activation, exhibits its unique properties for high efficiency hydrolysis of the polypeptide chain after the N-terminal tetraaspartyl-lysil sequence (scheme).

$$\overset{\text{Enteropeptidase}}{\overline{P_5\ \ P_4\ \ P_3\ \ P_2\ \ P_1 \downarrow P_1'}}$$

H-Val-Asp-Asp-Asp-Asp-Lys — Ile-Val-Gly-Gly-Tyr-…(trypsinogen) →

→ H-Ile-Val-Gly-Gly-Tyr-... (trypsin)

Liepnieks and Light *(2)* discovered that BPI was the only inhibitor suitable as an affinity sorbent ligand for the bovine enzyme; the method for enteropeptidase purification was developed and its basic properties were studied *(2–4)*.

The bovine enteropeptidase molecule contains approx 35% carbohydrate residues and consists of two disulfide-linked chains: heavy (120 kDa) and light (47 kDa) *(2,3,5)* (Fig. 1). This enzyme is synthesized as a single-chain precursor of 1035 amino acid residues; active enteropeptidase was cleaved after Arg-800 to produce a disulfide-linked heterodimer *(5)*. The total amino acid sequences of proenteropeptidases were also determined from porcine *(6)*, human *(7)*, and rat *(8)* cDNA sequences.

The light chain of the bovine enzyme is a trypsin-like serine protease containing a full set of amino acid residues necessary for the formation of the active site (Ser-987, His-841, and Asp-892), the trypsin specificity pocket that binds P_1 residue (Asp-981, Gly-1008, and Gly-1018), and the binding site for the linker $(Asp)_4$ sequence of the substrate (Lys-886—Arg-887—Arg-888—Lys-889) *(5)* (Fig. 1). The isolated light chain was obtained after the partial reduction and alkylation of the disulfide bonds of bovine enteropeptidase *(3)*. Recently, cDNA encoding one catalytic (light) chain of bovine enteropeptidase was expressed in various cell cultures (*Esherichia coli*, yeast, and mammalian cells) *(9–12)*; moreover, the recombinant bovine proenteropeptidase variants were expressed, purified, and activated by trypsin (*12*). In all cases with retaining the specificity of the intact enteropeptidase, the light chain was less active toward trypsinogen significantly in comparison with the native enzyme, and was inhibited not only by BPI but also by STI *(3,12)*.

The high degree of specificity exhibited by enteropeptidase makes it an ideal reagent for cleaving the specially constructed fusion proteins, containing the $(Asp)_4$Lys sequence between the carrier protein domain and the target product. For naturally occurring proteins, this sequence is found only in N-terminal activation peptides of various trypsinogens *(13)*; then,

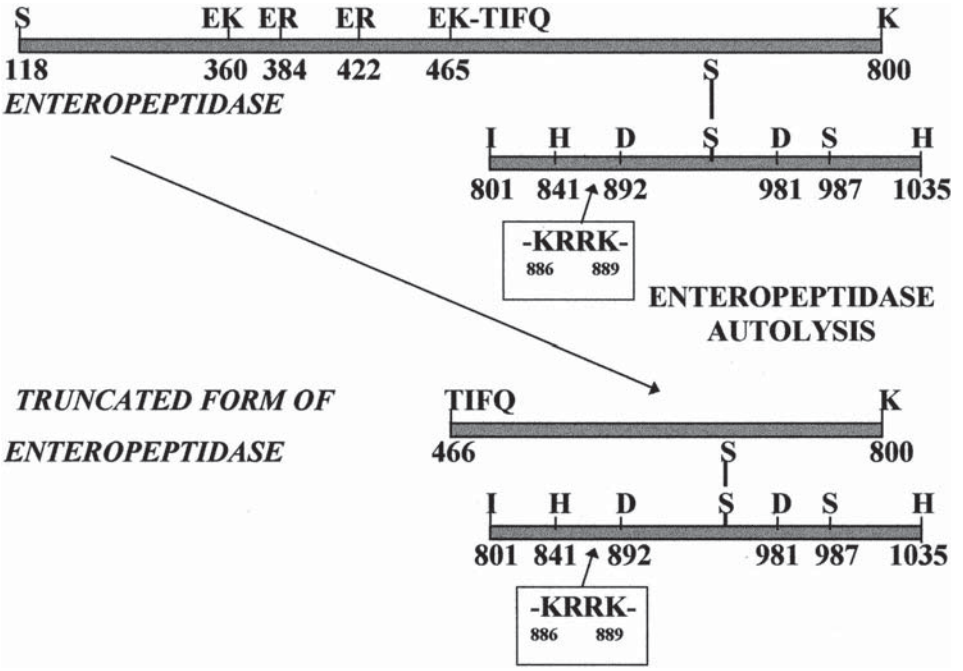

Fig. 1. Schematic structure of bovine enteropeptidase and its truncated form.

the enteropeptidase recognition site lies completely on the N-terminal side of the scissile bond, allowing any downstream fusion partner to retain its native amino terminus. Besides, the enzyme was found to be permissive to all amino acid residues in the P_1' position with the exception of Pro *(10,14)*.

For this purpose, enteropeptidase must be highly purified. Protease contaminations present in commercial preparations often degrade nonspecifically the fusion proteins *(15)*. Minor amounts of unwanted cleavage in some cases *(16,17)* indicated that traces of contaminating proteolytic activities are still present in the enteropeptidase samples highly purified by the Liepnieks and Light protocol *(2)*.

This article deals with a convenient preparative method for enteropeptidase purification; the preparations contained no other proteolytic enzyme contaminations. A variety of fusion proteins containing the $(Asp)_4Lys$ sequence were hydrolyzed by these enteropeptidase preparations. The K_m and k_{cat} values for the series of substrates were determined; the kinetic data indicated that the peptide and artificial protein substrates were bound to the enzyme through the linker $(Asp)_4Lys$ sequence on the light (catalytic) chain; the secondary binding site may participate in more efficient hydrolysis of the physiological substrate trypsinogen. The properties of the truncated enzyme, the autolysis product, suggest that this secondary binding site lies within the region of N-terminal fragment 118–465 of the enteropeptidase heavy chain.

Materials and Methods

Preparation of Affinity Sorbents

Trypsin-Sepharose 4B (sorbent 1), STI-Sepharose 4B (sorbent 2), and BPI-Sepharose 4B (sorbent 3) were prepared by the cyanogen bromide activation of the sorbents *(18)*. MPG-PA-2000-trypsin (sorbent 1A) was prepared according to the method in ref. *19*.

Purification of Bovine Enteropeptidase

Fresh bovine duodena were obtained from a local slaughterhouse and sliced down the middle, and the mucosa lining the intestinal wall was removed by gentle scraping. Mucosal cells were stored at –70°C.

Extract 700–800 g of thawed duodenal mucosa with 2.1 L of 0.1 *M* acetic acid (pH 4.0) containing 1% Triton X-100 and 20 m*M* $CaCl_2$ for 4 h with gentle stirring. All operations were performed at 4°C. The extract was centrifuged at 17,500*g* for 30 min. The supernatant was diluted 10-fold with 10 m*M* sodium acetate (pH 4.0) containing 50 m*M* $CaCl_2$ (buffer A) and concentrated by ultrafiltration to 1.5–2 L ("Pelicon," Amicon). The protein solution was mixed with 300 g of settled CM-32, equilibrated with buffer A, for 15–17 h by gentle shaking. The cellulose was recovered by centrifugation at 3000*g*. The supernatant was concentrated to 0.8–1 L by ultrafiltration, adjusted to pH 8.0 with 1 *M* Tris, and made 0.5 *M* NaCl and 0.1% OG. Three columns were connected in consecutive order, containing affinity sorbents 1 or 1A (20 mL), 2 (50 mL), and 3 (100 mL),respectively, equilibrated with 10 m*M* Tris-HCl, pH 8.0, containing 0.5 *M* NaCl , 50 m*M* $CaCl_2$, and 0.1% OG (buffer B). The protein solution was applied to this system of affinity columns at a flow rate 5 mL/h. The columns were washed with buffer B and precolumns 1 and 2 were disconnected. Column 3 was eluted with 0.1 *M* sodium formate (pH 3) containing 0.1 *M* NaCl, 50 m*M* $CaCl_2$, and 0.1% OG (buffer C). The fractions containing enteropeptidase were pooled, concentrated by ultrafiltration (Centricon-30, Amicon), or diluted 100-fold with H_2O and concentrated in the same manner. Enteropeptidase preparations were stored in 50% glycerol at –20°C without freezing.

The inhibitor of enteropeptidase and trypsin from duodenum (DI) was eluted from affinity sorbents 1 or 1A *(20)*, and the remaining duodenase from sorbent 2 *(21)*. The bulk of duodenase may be obtained from CM-32 cellulose by a linear gradient of 0.0–0.5 *M* NaCl in buffer A *(21)*.

Protein Determination

The protein concentration was estimated by the Bradford method *(22)* with Bio-Rad Protein Assay Kit, Standard 1 (bovine IgG) or spectrophotometrically by A_{235}–A_{280} difference *(23)*.

Activity Assay

Enteropeptidase activity was determined from the activation of trypsinogen according to Liepnieks and Light *(2)* with some modifications. One

enteropeptidase unit will produce 1.0 nmole of trypsin from trypsinogen in the activation mixture per minute. The mixture contained enteropeptidase (<0.03 U/mL), 0.1 *M* sodium acetate, pH 5.0, 50 m*M* $CaCl_2$ (buffer D), and 50 μL of trypsinogen solution (1.0 mg/mL of 1 m*M* HCl) in a total volume of 0.5 mL and was incubated for 30 min at 37°C. The reaction was quenched with 50 μL of 1 *M* HCl. The tryptic activity of 75-μL aliquots from the incubation mixture was measured spectrophotometrically at 405 nm and 25°C with BAPNA *(24)* (0.5 mg/mL) in 0.1 *M* Tris-HCl, pH 8.0, containing 50 m*M* $CaCl_2$ (buffer E) and 10% dimethylsulfoxide in the total volume of 1.5 mL. The degree of trypsinogen autoactivation (not exceeding 5–10% from enteropeptidase-derived trypsin amounts) was taken into account in the calculations. The calibration curve was plotted using known trypsin concentrations. The molarity of trypsin solutions was estimated by active-site titration with Gdn-Bz-ONp (*25*).

Active-Site Titration of Enteropeptidase

The concentration of enteropeptidase solutions was determined by active-site titration with Gdn-Bz-OMum according to *(2)* using trypsin titrated with Gdn-Bz-ONp as a standard. The activity of 100% pure enteropeptidase in our experiments is 1840 U/mg.

Fusion Protein Cleavage

Enteropeptidase stock solution (40–100 U/mL) was incubated with nine (or more) volumes of substrate solution in 10 m*M* Tris-HCl, pH 8.0, at 37°C for 10–20 h. The best results were obtained using recombinant protein concentrations >1 mg/mL. The hydrolysis was monitored by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) according to Laemmly *(26)* under reducing conditions (3% SDS and 2% β-mercaptoethanol) or by HPLC *(27)*.

Hydrolysis of Synthetic Peptide Substrate

Kinetic parameters for the GD_4K-NA cleavage were determined as described previously for the human enteropeptidase (*28*). A Hitachi MPF-4 spectrofluorimeter was used to follow the β-naphthylamin release, with excitation at 330 nm and emission at 410 nm. A 1-cm cuvet contained 1 mL of 0.1–1 m*M* GD_4K-NA in 50 m*M* Veronal buffer, pH 8.4, and 20 m*M* $CaCl_2$ at 37°C. The reaction was initiated by adding 10 μL of enteropeptidase solution (final concentration 1.4 n*M*).

Hydrolysis of PrA-P26

Kinetic parameters for the cleavage of fusion protein PrA-P26 were determined as described previously *(27)*. Assays (100–250 μL) contained 26–130 μ*M* PrA-P26 and 14 n*M* enteropeptidase. At different times, 15–20-μL samples were removed and the hydrolysis was monitored by SE chromatography on an Ultropac TSK G-2000 SW column (0.75 × 60 cm; LKB) in

Beckman HPLC system at a flow rate of 0.5 mL/min, eluent 25 m*M* sodium phosphate (pH 6.6) containing 0.2 *M* NaCl (buffer F). The initial velocity of substrate hydrolysis was estimated as a decrease in the area of PrA-P26 peak in the course of the incubation with the enzyme.

Kinetic Studies of Trypsinogen Activation

Kinetic parameters for the trypsinogen activation were determined at pH 5.0 and 8.0, 37°C. Trypsinogen autoactivation was controlled for all substrate concentrations and was insignificant for ≤15-min incubations. The rates of autoactivation were taken into account in the calculations.

Assays (0.5 mL) contained trypsinogen (0.8–8 μ*M*) in buffer D or buffer E at 37°C. The reaction was initiated by adding enteropeptidase (0.07 n*M*). At different time intervals, not exceeding 15 min, 50-μL samples were removed and the tryptic activity was measured as described in "Activity Assay."

Kinetic Parameters Determination

Michaelis–Menten constants (K_m and k_{cat}) were determined from linear plots constructed according to Eisenthal and Cornish-Bowden *(29)*. For all kinetic measures, duplicate determinations were made at six different concentrations of each substrate. Under the assay conditions, the consumption of substrates was <20%.

Autolysis of Enteropeptidase in EDTA Presence

The effects of EDTA on the enteropeptidase activity and structure were studied using the enzyme preparations purified in the absence of $CaCl_2$ in the crude extract and buffers A–C, or with Boehringer lyophilized enteropeptidase (Boehringer, Mannheim).

SDS-PAGE of enteropeptidase samples according to Laemmly *(26)* was carried out in 7% gel under reducing (3% SDS and 2% β-mercaptoethanol) or nonreducing (3% SDS) conditions; the samples were incubated in the loading buffer at 55–60°C for 5–6 h for denaturation. Protein bands were detected by Coomassie brilliant blue G-250 staining or transferred to an Immobilon membrane (Millipore) by electroblotting and N-terminal amino acid sequencing of autolysis products was carried out by automatic Edman degradation with an Applied Biosystem 470 A gas-phase sequencer.

Results

Purification

First, we tried the Light method for purification of enteropeptidase, and faced a lot of difficulties. One of the disadvantages of our preparations was trypsin- and chymotrypsin-like contaminations. In the search for optimal purification of enteropeptidase, a new serine protease displaying dual proteolytic activity (trypsin-like and chymotrypsin-like) was found in

bovine duodenal mucosa *(21)*. The protease (named duodenase) had a common inhibitor, BPI, with enteropeptidase, and, therefore, affinity chromatography on BPI-sepharose (sorbent 3) could not separate these enzymes. Duodenase appeared to be a dominant endopeptidase of bovine duodenal mucosa, and its content in mucosa homogenate was 0.3–0.4% of all proteins *(21)*. The content of enteropeptidase is only 0.015–0.03%. In our opinion, the traces of duodenase are the main cause of unwanted fusion protein cleavage. Duodenase, fortunately, is not so strictly specific as enteropeptidase. STI was found to be a good inhibitor for this enzyme in contrast to enteropeptidase. Duodenase purification on STI-sepharose was developed *(21)*, therefore we used the affinity sorbent (sorbent 2 in our purification scheme) for removing the admixture of duodenase; enteropeptidase did not bind to this sorbent. This approach with two consecutive columns containing two different immobilized inhibitors proved to be a useful tool for separating the high-specific enzyme from the related enzyme with the broad specificity. We also used sorbents 2 and 3 for successful purification of human tissue kallikrein *(20)*.

In the attempt to increase the amounts of active enteropeptidase after the affinity chromatography on sorbent 3, we used the detergent OG, preventing the membrane proteins and especially glycoproteins denaturation and inactivation. Indeed the yield of the affinity-purified enteropeptidase was a twofold increase.

The method developed by Light *(2)* contained numerous steps of ammonium sulfate fractionation, pH changes, and acidic precipitation. In our attempts to design a simpler and more rapid method of enteropeptidase purification, we found that at the first purification step an acid (pH 4.0) extraction of duodenal mucosa instead of pH 8.0 extracts (in the Light method) significantly decreased the amount of ballast proteins in the homogenate (5- to 10-fold) without the loss of the enteropeptidase activity. This made the ammonium sulfate and acidic precipitations unnecessary. The protein solution could be subjected to ion-exchange chromatography immediately after ultrafiltration.

We chose CM-cellulose instead of DE-cellulose in the Light method. Although the anion-exchanger bound the enteropeptidase *(2)*, the cation-exchanger was shown to bind the main part of the contamination proteins (80–90%) at pH 4.0, including duodenase, whereas the bulk of the enteropeptidase activity remained in solution. The simultaneous purification of duodenase and enteropeptidase was possible. Duodenase could be eluted from the cellulose by 0.0–0.5 *M* NaCl gradient *(21)*; extra amounts of enteropeptidase sorbed by CM-32 (10–20%) were separately eluted by this gradient. The 10-fold purification of enteropeptidase (Table 1) was achieved by this approach.

The next problem was associated with some peculiarities of affinity chromatography on BPI-sepharose. For a long time, in the literature *(2)* and in our experiments, the incomplete binding of enteropeptidase on the affinity sorbent has been observed. Despite the significant excess of immobi-

lized inhibitor, the fraction of the enzyme activity (15–44%, but in some mucosa lots >90%) did not bind to the affinity sorbent, and remained unbound on repeated chromatography on BPI-sepharose. Light *(2)* could not explain this phenomenon. This problem was solved when we suggested that enteropeptidase forms a complex with some unknown inhibitor in the duodenal mucosa. The presence of such an inhibitor could not be detected under the conditions of the activity assay of the enzyme: the high concentration of trypsinogen and low concentration of enteropeptidase, pH 5.0. Enteropeptidase became distributed among the soluble and immobilized inhibitors in the presence of BPI-sepharose; the ratio of the bound and unbound enzyme depended on the inhibitor amount in the mucosa lot. We isolated this inhibitor (DI) by affinity chromatography on immobilized trypsin, that in turn confirmed this hypothesis *(20)*. The N-terminal amino acid sequence 1–19 of DI was highly homologous to this region of BPI *(20)*. The Kunitz-type inhibitor DI is, in our opinion, a physiological inhibitor of enteropeptidase. The DI content in duodenum is usually 40-fold (mol) higher than enteropeptidase content *(20)*. For complete binding of enteropeptidase to sorbent 3 (BPI-sepharose) we added precolumn 1 with immobilized trypsin to our system of affinity columns. The advantage of this approach is that the side-product, DI, obtained after elution from immobilized trypsin, could serve as a ligand for affinity sorbent 3 to substitute for commercial BPI preparations. Removal of DI increased the yield of enteropeptidase in the affinity chromatography step up to >50%; 6000–7000-fold purification was achieved (Table 1), yielding 3–4 mg of highly purified enteropeptidase from 1 kg of mucosa.

The last modification of our purification procedure was made when we found that the second truncated form of enteropeptidase appeared in addition to its native form if the purified enzyme preparations had been stored for several months. The truncated enteropeptidase resulted from the autolysis caused by the loss of calcium ions by the enzyme *(30)*. It consists of the portion of the heavy chain bound by the disulfide bridge with the catalytic (light) chain. This form of enteropeptidase exhibited only several percent of the initial activity toward trypsinogen, but completely retained its activity toward the fusion proteins and GD_4K-NA and the number of active sites by titration with Gdn-Bz-OMum *(30)*. The known procedures of enteropeptidase purification including the Light method *(2)* did not use Ca^{2+} at all. We decided to deal only with the natural enzyme, despite the fusion proteins cleavage being independent of this autolysis: it is convenient to determine the enteropeptidase activity by trypsinogen activation. Besides, the truncated form was shown to be more susceptible to aggregation and denaturation with the entire loss of activity. By isolation and purification of enteropeptidase in the presence of $CaCl_2$ at the very first step, we obtained a single high-molecular form of the enzyme that cannot be converted into the truncated form even upon a long incubation with EDTA.

Table 1
Purification of Bovine Enteropeptidase from 700 g of Duodenal Mucosa

Purification step	Protein (mg)	Activity[c] (U)	Specific activity (U/mg)	Yield (%)	Purification (-fold)
Crude homogenate supernatant	33,000[a]	9370	0.28	(100)	(1.0)
Ultrafiltration	15,600[a]	9130	0.59	97.4	2.11
CM-32 chromatography supernatant	2940[a]	8300	2.82	88.5	10.07
Affinity chromatography dialysed and concentrated peak from BPI-Sepharose	2.35[b]	4330	1843	46.2	6580

[a]Protein was determined according to Bradford method *(22)*.
[b]Protein was determined spectrophotometrically *(23)*.
[c]The activity assay was carried out as described in "Materials and Methods."

Thus, we developed a preparative procedure to obtain a highly purified bovine enteropeptidase (Table 1). The activity of our preparations was an order of magnitude higher than the activity of those described, including the preparations obtained by the Light method *(2)*. The preparations of enteropeptidase in 50% glycerol were stable for several years at –20°C.

Purity Control of Enteropeptidase Preparations

The homogeneity of the enzyme preparations was confirmed by SDS-PAGE on 7% gel according to Laemmli *(26)* under nonreducing conditions (Fig. 2, second lane; single band) and reducing conditions (Fig. 2, third lane; two bands corresponding to heavy and light chains) and by SE-HPLC on a Toyo Soda TSK G 4000 SW column (0.75 × 60 cm) in buffer F, flow rate 0.5 mL/min (single peak, data not shown).

No visible degradation was observed on an SDS-PAGE gel when bovine serum albumin or cytochrome C (1 mg/mL) was incubated with enteropeptidase (100 U/mL) under the conditions for fusion protein cleavage, confirming the enzyme preparations did not contain any protease contamination.

The absence of duodenase traces was shown by the inability of purified enteropeptidase preparations to hydrolyze the best duodenase substrates, α-*N*-succinilalanilalanilprolilphenilalanin 4-nitroanulide and α-*N*-tosylglycilprolylarginine 4-nitroanilide *(31)*.

The N-terminal sequencing of 15 amino acid residues of the heavy chain of this enzyme was performed for the first time; it was shown that Ser-118 of the precursor is the starting residue of the heavy chain of the active bovine enteropeptidase *(30)* (Fig. 1). Previously a removal of 1–117 sequence from the proenteropeptidase molecule in the course of processing was found only

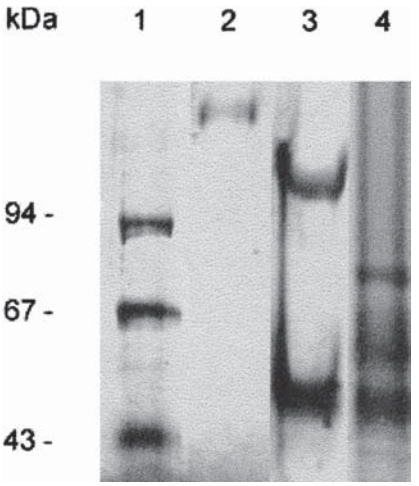

Fig. 2. SDS-PAGE of enteropeptidase. Samples (15–20 μg) were loaded onto 7% gel and electrophoresed according to Laemmly *(26)* under reducing (1,3,4) and nonreducing (2) conditions: (1), marker proteins; (2,3), native Ca^{2+}-containing enteropeptidase; (4), autolyzed sample of noncontaining Ca^{2+} enteropeptidase after incubation with EDTA.

Table 2
Construction of Fusion Proteins-Enteropeptidase Substrates[a]

Fusion protein	Carrier protein (kDa)	Linker	Target polypeptide (kDa)
GAL-ENK	β-Galactosidase (116)	-$(Asp)_4$Lys-	[Leu]^{5}Enkephalin (0.556)
GAL-VIP	β-Galactosidase (116)	-$(Asp)_4$Lys-	Vasoactive intestinal peptide (VIP 17-28) (1.35)
TNF-P	Tumor necrosis factor (16.8)	-$(Asp)_4$Lys-	Substance P analogs (1.4)
INF-HAV	γ-Interferon fragment (11)	-$(Asp)_4$Lys-	Protein of VP_1 region of Hepatitus A virus (44)
INF-HIV	γ-Interferon fragment (11)	-$(Asp)_4$Lys-	Arg-Proteinase HIV I (13)
PrA-P26	Protein A fragment (12)	-$(Asp)_4$Lys-	Ala-Recoverin (P26) (26)
PrA-GKN	Protein A fragment (12)	-$(Asp)_4$Lys-	Ala-Guanilatkinase (22)
PrA-FaD	Protein A fragment (12)	-$(Asp)_4$Lys-	Differentiation factor (8.2)
PrA-OSK1	Protein A 27 domain (14)	-$(Asp)_4$Lys-	Toxin from *Orthrochirus scrobiculosus* (4)
PrA-ST	Protein A 27 domain (14)	-$(Asp)_4$Lys-	Toxin from scorpion (4)
TRX-P2	Thioredoxin-His_6Tag (18)	-$(Asp)_4$Lys-	Modified destabilase *N*-terminus (2)
TRX-DS	Thioredoxin-His_6Tag (18)	-$(Asp)_4$Lys-	Destabilase (14)

[a]Expression and purification of fusion proteins were described in *(27,32–36)*.

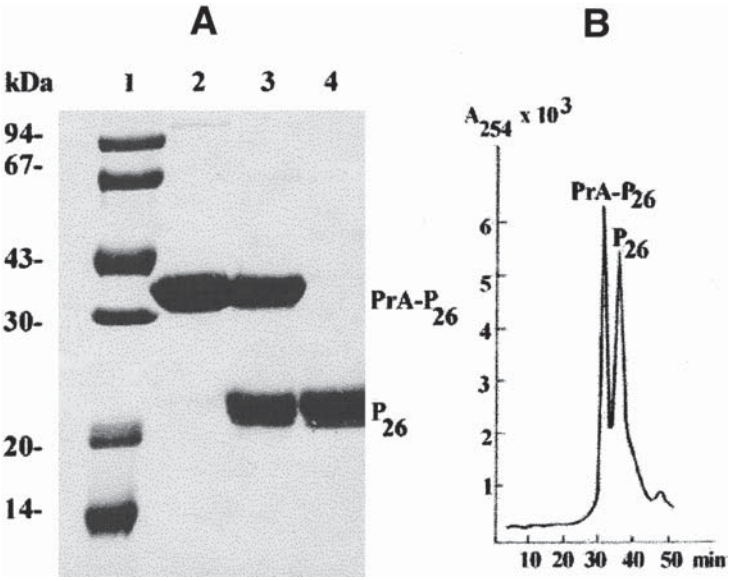

Fig. 3. Hydrolysis of PrA-P26 (2×10^{-4}*M*) by enteropeptidase (6.5 n*M*); **(A)** SDS-PAGE, 15% gel, reducing conditions: (1), marker proteins; (2), control, incubation in the absence of enteropeptidase for 17 h; (3,4), hydrolysis products, 2.5 h and 17 h, respectively; **(B)** SE-HPLC on an Ultropac TSK G-2000 SW column (0.75 cm x 60 cm), flow rate 0.5 mL/min, 25 m*M* sodium phosphate (pH 6.6) containing 0.2 *M* NaCl; incubation time 2.5 h.

for the porcine enzyme *(6)*. In all other cases the intrinsic N-terminal sequence of the heavy chain of the mature active enzyme remains unknown.

Fusion Protein Cleavage

For the first time, we successfully used our enteropeptidase preparations in 1987 for cleavage of fusion protein GAL-ENK *(32)*. The variety of recombinant proteins (Table 2) containing the tetraaspartyl-lysyl sequence in arbitrary positions of the polypeptide chain were then studied. The expression and purification of these fusion proteins were described in refs. *27* and *32–36*. In Figs. 3 and 4, some examples of such substrates digestion with enteropeptidase are shown. Soluble fusion proteins, like GAL-ENK, and especially those secreted, like PrA-P26, are found to be preferable in comparison with the proteins produced as insoluble inclusion bodies. For an effective enteropeptidase cleavage of such substrates its concentration must be high (1–10 mg/mL) *(27)*. The concentration of inclusion body proteins after refolding was only 0.1–0.2 mg/mL and such proteins as TNF-P, INF-HAV, and INF-HIV appeared to be poor substrates for enteropeptidase *(27)*. We arrived at the conclusion that the best N-terminal fusion partner for this expression system was the modified protein A, providing the secretion of soluble and properly folded target proteins (Table 2). Another expression system utilizing the enteropeptidase linker $(Asp)_4Lys$

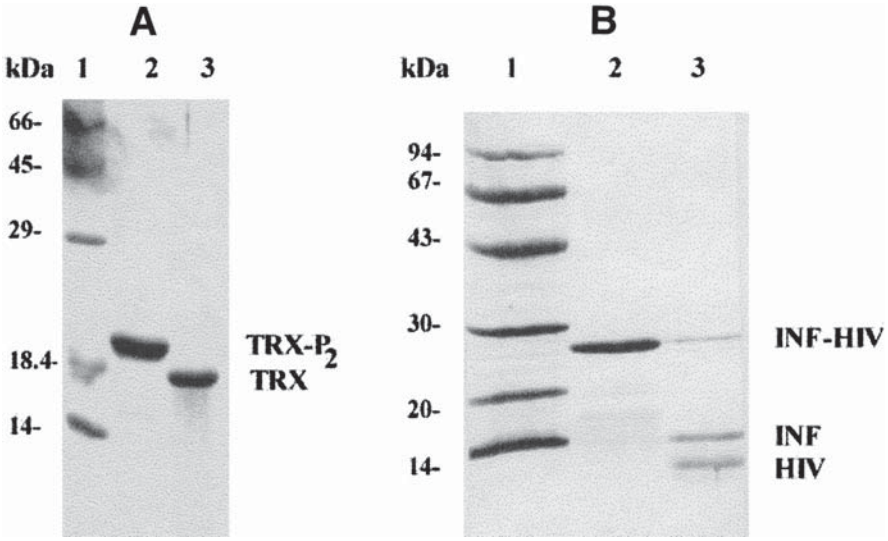

Fig. 4. Hydrolysis of fusion proteins by enteropeptidase (30 n*M*); SDS-PAGE under reducing conditions: **(A)** 15% gel, TrX-P2 (1×10^{-4} *M*); 1h; **(B)** gradient gel 10-20%, INF-HIV (1×10^{-5} *M*); 5 h; (1), marker proteins; (2), control, incubation without enteropeptidase; (3), hydrolysis products.

was invented by LaVallie *(10,11,14)*, where the *E. coli* cytoplasmic protein thioredoxin (TRX) and its secreted homolog DsbA as a N-terminal fusion partner were shown to direct proper folding and especially the disulfide bond formation of the C-terminal fusion heterologous proteins. The authors also stated that only secretion of the fusion protein was an essential element, which led to the generation of active target products *(11)*. The constructions with TRX as a carrier were also successfully cleaved by our enteropeptidase preparations (Table 2).

Kinetics of Substrate Cleavage by Enteropeptidase

One of the best fusion proteins-enteropeptidase substrates appeared to be PrA-P26. It was convenient to monitor its hydrolysis by means of SE-HPLC (Fig. 3). This allowed us to use it for studying the enzymatic activity of enteropeptidase, for example, inhibitory analysis *(20)*, and so on. It should be mentioned that the kinetics of enteropeptidase action was previously difficult to study, owing to the absence of a sufficient number of various substrates; besides, for lowering the risk of the trypsinogen autoactivation at pH 8.0, the enteropeptidase activity from the rate of trypsinogen activation was usually determined at pH 5.0–6.0 rather than at pH-optimum 7.0–8.0 *(2)*. By using recombinant proteins with the $(Asp)_4Lys$ sequence, we significantly increased the number of substrates suitable for experimental studies of this highly specific enzyme and for the first time determined the catalytic parameters for hydrolysis of the fusion protein with enteropeptidase *(27)*. A series of substrates was methodically studied

Table 3
Kinetic Constants for Hydrolysis of Substrates by Enteropeptidase[a]

Substrate	pH	$K_m \times 10^4$ (M)	k_{cat} (min^{-1})	$k_{cat}/K_m \times 10^{-6}$ (M^{-1} min^{-1})
GD_4K-NA	8.4	2.0	1000	5.0
PrA-P26	8.0	1.25	157	1.26
Trypsinogen	8.0	0.024	700	290
	5.0	0.018	435	240
	5.6 *(2)*	0.17	380	22

[a]The reactions were carried out as described in "Materials and Methods." Data for trypsinogen activation at pH 5.6 are from ref. *2*.

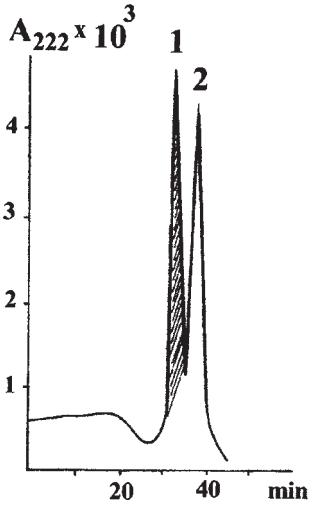

Fig. 5. SE-HPLC of the calcium-free enteropeptidase preparation after long-term storage at 20°C; (1), native enzyme; (2) truncated enzyme. A Toyo Soda TSK G 4000 SW column (0.75 cm × 60 cm) in 25 m*M* sodium phosphate (pH 6.6) containing 0.2 *M* NaCl, flow rate 0.5 mL/min.

(27,30): the natural substrate trypsinogen, the recombinant protein containing the linker sequence $(Asp)_4Lys$, and a low-molecular substrate, GD_4K-NA. Moreover, we determined for the first time the kinetic parameters of the trypsinogen activation not only at pH 5.0, but also at the optimal pH for the purpose of a more correct comparison of the hydrolysis parameters of various substrates (Table 3).

Autolysis of Enteropeptidase Heavy Chain

The second form of enteropeptidase appearing during extended storage of the purified enteropeptidase in the absence of $CaCl_2$ enzyme preparations (several months) or their incubation with EDTA (several hours) had a lower mobility by SE-HPLC (Fig. 5) as compared to the native protein. A measurement of the proteolytic activity showed that this modified

enteropeptidase possessed only several percent of the initial activity toward trypsinogen, but completely retained its activity toward fusion protein substrates, GD_4K-NA, and an active site titrant Gdn-Bz-OMum *(27,30)*. An increase in the retention time under gel chromatography conditions indicates that an apparent decrease (approx twofold) in the molecular mass accompanied the enzyme modification. Sufficient acceleration of the process in the presence of complexon EDTA (but not of phenanthroline) suggests that the loss of the calcium ions by the protein was the reason for this phenomenon. That was confirmed by isolation and purification of enteropeptidase in the presence of $CaCl_2$.

After SDS-PAGE in reducing conditions (Fig. 2, fourth lane) and electroblotting of fragments of the heavy chain on Immobilon, the N-terminal amino acid sequences of the autolysis products were determined. The high-molecular protein band (70–80 kDa) was shown to contain the C-terminal half of the heavy chain beginning with Thr-466 together with the N-terminal half, beginning with Ser-118 (it is the first amino acid of active enteropeptidase heavy chain *[30]*). So the cleavage site of Ca^{2+}-dependent autolysis of the heavy chain is the peptide bond formed by the carboxyl of the Lys-465 residue. In the proenteropeptidase amino acid sequence *(5)* Glu residue is located in position 464. It should be noted that the heavy chain does not contain the enteropeptidase linker sequence $(Asp)_4Lys$ for the enzyme attack *(5)*. But the middle part of the heavy chain (domain 359–465) contains not only Lys-465, but also Lys-360, Arg-384, and Arg-422 preceded by Glu residues *(5)* (Fig. 1). Sequencing of the broad protein band 56–66 kDa (Fig. 2, fourth lane) showed that these autolysis products began with Ser-118 and probably were shorter N-terminal fragments of the heavy chain derived after its cleavage in these sites.

Discussion

The K_m and k_{cat} values were established for various enteropeptidase substrates: the natural substrate trypsinogen, recombinant protein containing the linker sequence $(Asp)_4Lys$ and low-molecular substrate GD_4K-NA *(27,30)*. Recently the kinetic constants for cleavage of GD_4K-NA and trypsinogen activation were determined for native and recombinant variants of bovine enteropeptidase *(12)*. The reported values are in close agreement with our data.

We believe that the discrepancies in the K_m and k_{cat} values for the trypsinogen activation at acidic pH between our and literature data (Table 3) arise as a result of the admixture of the second, low-active to trypsinogen form in the enteropeptidase preparation obtained earlier in the absence of Ca^{2+} *(2)*.

The precise kinetic data of fusion protein hydrolysis by native enteropeptidase and recombinant light chain are absent in the literature. The rates of the cleavage of various TRX-containing fusions *(10,11)* differed significantly one from another not only in the case of the two-chain enzyme, but also in the case of its isolated light chain. These results suggest that the matter of

the $(Asp)_4Lys$ availability to the enzyme active site may cause this substrate difference. It should be mentioned that the rate of PrA-P26 hydrolysis under these conditions is the highest one, indicating the linker sequence of this substrate to be fully accessible to the enteropeptidase active site.

The light chain of enteropeptidase is a trypsin-like serine protease. The role of the heavy chain of enteropeptidase was discussed in literature. There is an opinion that it serves only to anchor the catalytic subunit to the intestinal brush border and orients it toward the lumen *(2,11)*. However, a significant difference in the enzymatic properties of the native two-chain molecule and the isolated light chain, detected not only with respect to trypsinogen activation (the 100-fold reduced activity of the latter) but also to the protein protease inhibitors reactivity, led Lu et al. *(12)* to the conclusion on the heavy chain participation in the recognition of macromolecular substrates and inhibitors.

We believe the role of the enteropeptidase heavy chain may be partially clarified by the results of our experiments. The comparison of the kinetic parameter values for the enteropeptidase substrates series (Table 3), on the one hand, and the light chain structure and properties *(3,5,12)*, on the other hand, clearly indicate that the binding of $(Asp)_4Lys$ sequence of all the substrates occurs in the binding site in the light chain, whereas physiological substrate trypsinogen is additionally fixed in some secondary binding site located in the heavy chain.

We have found the second form of enteropeptidase, the product of autolysis, caused by the loss of calcium ions by the protein. This truncated form of the enzyme contains the C-terminal part 466-800 of the heavy chain, linked by Cys-788-Cys-912 disulfide bond with light chain (Fig. 1). The truncated enzyme properties toward trypsinogen and low-molecular substrates *(30)* seem similar in many aspects to the isolated light chain *(3,5,12)*. These results may help define more exactly the enteropeptidase heavy chain participation in trypsinogen recognition: the secondary substrate-binding site that interacts directly with trypsinogen *(12)* may be located in the *N*-terminal fragment 118-465 of the heavy chain.

Acknowledgments

The authors thank L. A. Chupova and T. I. Muravieva for their assistance in protein purification; and N. A. Potapenko and Yu. F. Leonova for amino acid sequencing. This work was supported by the Russian Foundation for Basic Research (no. 99-04-48362) and by the State Scientific Program "Novel Methods for Bioengineering: Protein Engineering" (no. 03.0003N-335).

References

1. Kunitz, M. (1939), *J. Gen. Physiol.* **2,** 429–446.
2. Liepnieks, J. J. and Light, A. (1979), *J. Biol. Chem.* **254,** 1677–1683.
3. Light, A., and Fonseca, P. (1984), *J. Biol. Chem.* **259**, 13195–13198.

4. Light, A., Savithri, H. S., and Liepnieks, J.J. (1980), *Anal. Biochem*. **106,** 199–206.
5. Kitamoto, Y., Yuan, X., Wu, Q., McCourt, D. W., and Sadler, J. E. (1994), *Proc. Natl. Acad. Sci. USA* **91,** 7588–7592.
6. Matsushima, M., Ichinose, M., Yahagi, N., Kakei, N., Tsukada, S., Miki, K., et al. (1994), *J. Biol. Chem*. **269,** 19,976–19,982
7. Kitamoto, Y., Veile, R. A., Donis-Keller, H., and Sadler, J. E. (1995), *Biochemistry* **34,** 562–568.
8. Yahagi, N., Ichinose, M., Matsushima, M., Matsubara, Y., Miki, K., Kurokawa, K., et al. (1996), *Biochem. Biophys. Res. Commun*. **219,** 806–812.
9. LaVallie, E. R., Rehemtulla, A., Racie, L. A., DiBlasio, E. A., Ferenz, C., Grant, K. L., Light, A., and McCoy, J. M. (1993), *J. Biol. Chem.* **268,** 23,311–23,316.
10. Collins-Racie, L. A., McColgan, J. M., Grant, K. L., DiBlasio-Smith, E. A., McCoy, J. M., and LaVallie, E. R. (1995), *Bio/Technology* **13,** 982–987.
11. Vozza, L. A., Wittwer, L., Higgins, D. R., Purcell, T. J., Bergseid, M., Collins-Racie, L. A., LaVallie, E. R., and Hoeffler, J. P. (1996), *Bio/Technology* **14,** 77–81.
12. Lu, D., Yuan, X., Zheng, X., and Sadler, J. E. (1997), *J. Biol. Chem*. **272,** 31,293–31,300.
13. Guy, O., Bartelt, D. C., Amic, J., Colomb, E., and Figarella, C. (1976), *FEBS Lett*. **62,** 150–153.
14. LaVallie, E. R., DiBlasio, E. A., Kovacic, S., Grant, K. L., Schendel, P. F., and McCoy, J. M. (1993) *Bio/Technology* **11,** 187–193.
15. Dykes, C. W., Bookless, A. B., Coomber, B. A., Noble, S. A., Humber, D. C., and Hobden, A. N. (1988), *Eur. J. Biochem* **174,** 411–416.
16. Hopp, T. P., Prickett, K. S., Libby, R. T., March, C. J., Cerretti, D. P., Urdal, D. L., and Conlon, P.G. (1988), *Bio/Technology* **6,** 1204–1210.
17. Uegaki, K., Nemoto, N., Shimizu, M., Wada, T., Kyogoku, Y., and Kobayashi, Y. (1996), *FEBS Lett.* **379,** 47–50.
18. March, S. C., Parikh, I., and Cuatrecasas, P. (1974), *Anal. Biochem*. **60,** 149–152.
19. Kudryavtseva, N. E., Zhigis, L. S., Zubov, V. G., Vulfson, A. N., Maltsev, K. V., and Rumsh, L. D. (1995), *Khim. Pharm. Zh.* **1,** 61–-63 (in Russian).
20. Mikhailova, A. G., Evtyukova, N. G., Chupova, L. A., and Rumsh, L. D. (1998), *Vopr. Med. Khim.* **4,** in press (in Russian).
21. Zamolodchikova, T. S., Vorotyntseva, T. I., and Antonov, V. K. (1995), *Eur. J. Biochem.* **227,** 866–872.
22. Bradford, M. (1976), *Anal. Biochem*. **72,** 248–254.
23. Whitaker, J. R. and Granum, P. E. (1980), *Anal. Biochem*. **109,** 156–159.
24. Kassel, B. (1970), *Methods Enzymol*. **19,** 844–853.
25. Chase, T., Jr. and Shaw, E. (1970), *Methods Enzymol*. **19,** 20–27.
26. Laemmli, U. K. (1970), *Nature* **227,** 680–685.
27. Mikhailova, A. G., Shibanova, E. D., Rumsh, L. D., and Antonov, V. K. (1994), *Bioorg. Khim.* **20,** 883–893 (in Russian).
28. Grant, D. A. W. and Hermon-Taylor, J. (1979), *Biochim. Biophys. Acta* **567,** 207–215.
29. Eisenthal, R. and Cornish-Bowden, A. (1974), *Biochem. J.* **139,** 715–720.
30. Mikhailova, A. G. and Rumsh, L. D., (1998), *Bioorg. Khim.* **24,** 282–287 (in Russian).
31. Zamolodchikova, T. S., Sokolova, E. A., Alexandrov, S. L., Mikhaleva, I. I., Prudchenko, I. A., Morozov, I. A., et al. (1997), *Eur. J. Biochem.* **249,** 612–621.
32. Dobrynin, V. N., Boldyreva, E. F., Filippov, S. A., Chuvpilo, S. A., Korobko, V. G., Vorotyntseva, T. I., et al. (1987), *Bioorg. Khim.* **13,** 119–121 (in Russian).
33. Kutusov, M. A., Schmuckler, B. E., Suslov, O. N., Zargarov, A. A., and Abdulaev, N. G. (1992), *Bioorg. Khim.* **18,** 119–121 (in Russian).
34. Gaidarov, I. O., Suslov, O. N., Ovchinnikova T. V., and Abdulaev,N. G. (1994), *Bioorg. Khim.* **20,** 367–381 (in Russian).
35. Dergousova, N. I., Amerik, A. Yu., Volynskaya, A. M., and Rumsh, L. D. (1996), *Appl. Biochem. Biotechnol*. **61,** 97–106.
36. Kostanyan, I. A., Astapova, M. V., Starovoytova, S. M., Dranitsina, S. M., and Lipkin, V. M. (1994) *FEBS Lett.* **356,** 327–329.