Sea Urchin Hatching Enzyme (Envelysin): cDNA Cloning and Deprivation of Protein Substrate Specificity by Autolytic Degradation[†,‡,§]

Kohji Nomura,*[,||] Takeshi Shimizu,[⊥,▽] Hiroaki Kinoh,[#] Yutaka Sendai,[○] Mitsushi Inomata,[||] and Norio Suzuki[⊥]

Department of Biochemistry, Tokyo Metropolitan Institute of Gerontology, 35-2 Sakaecho, Itabashi-ku, Tokyo 173, Japan, Division of Biological Sciences, Graduate School of Science, Hokkaido University, Kita-ku, Sapporo 060, Japan, Department of Biochemistry, Toyama Medical Pharmaceutical University Faculty of Medicine, 2630 Sugitani, Toyama 930-01, Japan, and Research Institute for the Functional Peptides, 11-26 Minamisanban-cho, Yamagata-990, Japan

Received December 4, 1996; Revised Manuscript Received March 10, 1997[⊗]

ABSTRACT: The hatching enzyme (envelysin) of the sea urchin *Hemicentrotus pulcherrimus* was purified from the medium of hatched blastulae. By cDNA cloning its deduced amino acid sequence and molecular architecture were revealed. The 591-residue precursor with calculated M_r of 66 123 consists of an 18-residue signal sequence, a 151-residue propeptide, and a 422-residue mature enzyme with N-terminal catalytic and C-terminal hemopexin-like domains. As compared with that of *Paracentrotus lividus*, its amino acid sequence is 69% identical and 10% similar. They share typical structural features with the mammalian MMP gene family members: cysteine switch, zinc-binding signature, methionine-turn, Cys residues near both ends of hemopexin-like domain, etc. However, its propeptide has a 70-residue extra sequence with an Asp- and Glu-rich stretch, supposedly involved in the proenzyme activation by binding Ca^{2+} ions in seawater. The hinge region is also longer than those of most MMPs, with an extra sequence rich in Thr and Arg residues. Mature 50K enzyme is highly susceptible to autolytic cleavage at Gln-(503)-Leu(504), producing the 38K form retaining catalytic activity and substrate specificity against fertilization envelope. The 38K form and 15K fragment were coeluted from a gel-filtration column, suggesting that these two fragments are disulfide-bridged and that the tertiary structure is not much deviated. The 38K form further autolyzed to 32K form by cleaving Tyr(450)-Tyr(451) bond with the loss of protein-substrate specificity, retaining only nonspecific protease activity. Thus, the autolytic release of $^2/_3$ of the C-terminal domain reduced the highly specific enzyme to a common nonspecific protease, implying that the size and structure of almost the entire hemopexin-like domain is essential for the protein substrate specificity. Moreover, autolytic degradation of envelysins from the two species follow quite different pathways despite their high homology in structure. The 38K and 32K forms were inhibited by bovine TIMP-1 with different IC_{50} values, indicating that its inhibitory activity depends on the extent of the interaction with the C-terminal domain of the enzyme.

About 11 h after fertilization, the embryos of the sea urchin *Hemicentrotus pulcherrimus*, at mid-blastula stage, hatch out from the protective fertilization envelope (FE),[1] a huge complex of glycoproteins. During this hatching process FE is dissolved by the action of the protease historically called "hatching enzyme (HEz)" (EC 3.4.24.12) or "envelysin" by our designation (Nomura et al., 1991) with the aid of mechanical disruption by ciliary motion. The presence of HEz as a protease was first reported by J. Ishida (1936, 1985), and since then several groups have tried to purify and characterize HEz, but none could succeed in its purification in the strict sense of the word. The isolation procedures employing the conventional chromatographies of gel filtration, ion exchange, etc. were not successful for the isolation of intact HEz, i.e., the enzyme with the specificity and activity to rapidly dissolve the homologous FE in seawater or in buffers containing 0.5 M NaCl. The "isolated" enzymes in the past years could not dissolve FE in seawater but only in buffers with low ionic strength (Edwards et al., 1977; Takeuchi et al., 1979). The basic nature of HEz was reported by two or more groups to be chymotrypsin-like, because the enzyme, especially in a crude stage, was inhibited by the chymotrypsin-specific inhibitors, chymostatin, TPCK, etc.

[†]This work was supported in part by a Grant-in-Aid for Co-Operative Research (A) No. 07308051, by Grants-in-Aid for Scientific Research Nos. 02404006 and 06454925, from the Ministry of Education, Science and Culture of Japan, and by a grant from the Research Institute of Marine Invertebrates.

[‡] The cDNA and deduced amino acid sequences of *H. pulcherrimus* envelysin has been deposited with DDBJ/ EMBL/GeneBank under the accession number AB000719.

[§] This paper is dedicated to the late Dr. Juro Ishida who first reported the sea urchin hatching enzyme in 1936. He passed away on July 22, 1994, at the age of 86.

* To whom correspondence should be addressed. Tel: +81-3-3964-3241, ext 3068. FAX: +81-3-3579-4776. E-mail: nomura @ tmig.or.jp.

[||] Tokyo Metropolitan Institute of Gerontology.

[⊥] Hokkaido University.

[▽] Present address: Research Institute, International Medical Center of Japan, 1-21-1, Toyama, Shinjuku-ku, Tokyo 162, Japan.

[#] Toyama Medical Pharmaceutical University Faculty of Medicine.

[○] Research Institute for the Functional Peptides.

[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1997.

[1] Abbreviations: HEz, hatching enzyme; FE, fertilization envelope; MMP, matrix metalloproteinase; TIMP, tissue inhibitor of metalloproteinases; 3-D, three-dimensional; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; PCR, polymerase chain reaction; 2-ME, 2-mercaptoethanol; CHAPS, 3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate; HEPES, *N*-[2-hydroxyethyl]piperazine-*N*′-[2-ethanesulfonic acid]; PVDF, polyvinylidene difluoride.

S0006-2960(96)02979-0 CCC: $14.00

(Hoshi et al., 1979; Post et al., 1988). Moreover, the cleavage-site specificity was reported as Glu-specific, on the basis of results from experiments using the small peptide substrates such as Cbz-Glu-Phe and Cbz-Glu-Tyr (Barrett & Edwards, 1976). Another subject of controversy was whether the genetic information for HEz was of maternal origin or not (Barrett et al., 1978). This was basically settled by the experiments which by the use of hybrid andromerogons revealed that new mRNA synthesis from embryonic chromatin is required before hatching enzyme can be synthesized (Showman & Whiteley, 1980). Ten years later it was clearly solved by Northern blot analysis (Lepage & Gache, 1990).

The way out from the confusion and deadlock for more than half a century was at last found by employing procion-agarose chromatography for the purification of the HEz of Mediterranean sea urchin *Paracentrotus lividus* (Lepage & Gache, 1989). The purified 51K enzyme, in itself, retains the original activity and specificity to rapidly dissolve FE in seawater. A year later, by cDNA cloning, its deduced amino acid sequence of the preproenzyme was unveiled as a member of the matrix metalloproteinase (MMP) gene family (Lepage & Gache, 1990). The 587-residue precursor consists of an 18-residue signal sequence, a 148-residue propeptide, and a 421-residue mature enzyme made up of catalytic (N-terminal) and hemopexin-like (C-terminal) domains. As a member of MMP gene family it has a typical cysteine-switch consensus sequence PRCGVPD (Sanchez-Lopez, 1988; Springman et al., 1990) near the end of propeptide, and the catalytic site "zinc(-binding) signature," HEFGHSLGLYH (Jongeneel, et al., 1989). As compared with the mammalian collagenases, gelatinases, and stromelysins, HEz has an additional Asp- and Glu-rich stretch in the middle of propeptide, and another extra-sequence that is rich in Thr and Arg residues in the hinge region or linker peptide between the catalytic and hemopexin-like domains (Li et al., 1995). Later, they determined the total sequence of the 6332 bp genomic DNA comprising nine exons, with an altered cysteine-switch sequence PLCGVPDV, reportedly due to polymorphism (Ghiglione et al., 1994).

We previously reported about the purification of a 37K form HEz from *H. pulcherrimus* and the specificity at the cleavage site and the nearby residues using various peptides, including oxidized insulin B chain, substance P, and mastoparan (Nomura et al., 1991). The results indicated that the HEz preferentially cleaves the NH_2-side of hydrophobic residues, -Leu, -Ile, -Phe, -Tyr as well as -Ala and -Gln residues, similarly to stromelysin, but not so strictly as thermolysin. Thermolysin and HEz really revealed distinct cleavage patterns on the peptides, e.g., substance P, mastoparan, and LHRH. Later we found that HEz was reversibly inhibited by the synthetic peptide Ac-PR**C**GVPDV-NH_2 with the cysteine-switch consensus sequence but not by the peptide with D-Cys isomer, indicating the stereospecific nature of the binding of the peptides to the active site of HEz (Nomura & Suzuki, 1993).

Extensive research on mammalian MMPs in the past decades established some novel concepts on the structure and function of the metalloproteinases of this category. One of the striking features of MMP, except for matrilysin, is that their protein substrate specificity and the inhibition by specific inhibitors TIMPs are ascribed to the C-terminal hemopexin-like domain (Clark & Cawston, 1989; Murphy et al., 1992; Schnierer et al., 1993; Sanchez-Lopez et al., 1993; Baragi et al., 1994; O'Connell et al., 1994). Catalysis, another basic function of protease, is performed by the catalytic site or active center in the N-terminal catalytic domain, although the cleavage site specificity or the primary specificity is generally determined by the structure of specificity pocket close to the catalytic site. Thus, in this family of metalloproteinases, the two basic functions of a protease, catalysis and specificity, are ascribed respectively to the distinct domains, catalytic (N-terminal) and hemopexin-like (C-terminal) domains connected by the hinge region which is also involved in the determination of substrate specificity. This prominent type of molecular architecture for a distinct group of proteases is not found in ordinary proteases of digestive role or of bacterial origin. Supposedly, this structure was generated by the fusion of two genes in a later phase of protease evolution.

In order to widen our knowledge on the structure of HEz, the structural basis for substrate specificity, and the molecular mechanism of specific FE dissolution by HEz, we undertook the cDNA cloning of the *H. pulcherrimus* HEz, determined the sites of autolysis, and confirmed the differential inhibition of the two forms of HEz by TIMP-1.

EXPERIMENTAL PROCEDURES

Materials. The sea urchins, *H. pulcherrimus* were collected near Noto Marine Laboratory, Kanazawa University, on the Sea of Japan coast (Toyama Bay). The eggs and sperm were collected by artificial spawning induced by pouring 0.5 M KCl into opened body cavity. The concentrated medium of hatched blastulae, the starting material for HEz purification, and soft and hard FEs were prepared as reported previously (Nomura et al., 1990, 1991). Organic and inorganic chemicals were purchased from Wako Pure Chemical Ind. Ltd. (Osaka), Nacalai Tesque Inc. (Kyoto), or Sigma Chemical Co. (St. Louis, MO) except as otherwise noted. Chymostatin was purchased from Peptide Institute Inc. (Osaka). Red-agarose, i.e., reactive red 120 (type 3000-CL, cross-linked 4% beaded agarose), HEPES, and CHAPS were from Sigma, and the maleimide-activated keyhole limpet hemocyanin was from Boehringer-Mannheim (Tokyo). Recombinant bovine TIMP-1 was expressed in CHO cells and isolated as about 28K glycoprotein. It is identical to embryogenin-1 with embryogenesis-stimulating activity as previously reported (Sato et al., 1994). Superose 12-HR (10/30) and 12-PG (16/50) columns were from Pharmacia Biotech (Tokyo). The reverse phase C18 column for HPLC (Senshu-Pak) was purchased from Senshu Sci. Co. (Tokyo). The cDNA synthesis kit, Hybond-N membrane, [γ-^{32}P]ATP (111 TBq/mmol) and [α-^{32}P]dCTP (110 TBq/mmole) were products of Amersham International plc. (Amersham, U.K.). 7-DEAZA Sequencing Kit, version 2.0, was from Takara Biomedicals Co. (Kyoto). Sequenase, version 2.0, DNA sequencing kit was from United States Biochemical Co. (Cleveland, OH). Restriction enzyme, T4 DNA ligase, and other enzymes used for molecular cloning were purchased from Takara Biomedical Co. (Kyoto) or Toyobo Co. (Osaka). The random-primed DNA labeling kit was purchased from Boehringer-Mannheim (Indianapolis, IL). The plasmids, pBluescript II KS(+) and pBluescript II KS(−) were generously provided by Dr. Yoshitaka Nagahama at the National Institute for Basic Biology, Okazaki, Japan.

Fertilization and Embryo Culture. The collected eggs were washed with filtered seawater for a few times and

fertilized by adding diluted semen as previously described (Nomura et al., 1991), and the seawater medium was changed several times to remove sperm by settling and decantation/aspiration. Embryos were cultured in Millipore-filtered (0.45 μm) seawater at 20 °C, with gentle stirring at a population density of about 5000 eggs/mL. At the onset of hatching, stirring was stopped, the medium was decanted and aspired to 10% of the original volume, and the embryos were further incubated to almost complete hatching. After the embryos had settled, the medium was paper-filtered and processed for purification of HEz or was stored at −80 °C until purification.

Purification and Autolytic Degradation of Envelysin. Envelysin and its autolytic products were isolated from the medium of the hatched embryos as previously described (Nomura et al., 1991). In short, 0.5 M NaCl was added to 3 L of hatching supernatant prepared from about 30 females and concentrated to 150 mL by ultrafiltration on Amicon PM-10 membrane. It was then supplemented with 0.5% CHAPS and loaded on a red-agarose column (2.6 × 20 cm) equilibrated with buffer A (10 mM HEPES, 1 M NaCl, and 10 mM $CaCl_2$, 0.5% CHAPS, pH 8.2). It was eluted with buffer A and then by a gradient of ethylene glycol (0−60%) at a flow rate of 30 mL/h, and the fractions of 5 mL were collected. The fractions with hatching activity were combined and dialyzed against the starting buffer and concentrated to about 1 mL. It was further purified by gel filtration on a Superose 12PG column (1.6 × 50 cm) using the same buffer at the flow rate of 20 mL/h. Fractions of 1 mL were collected and stored at −30 °C. A portion of the purified enzyme was left for autolysis at 25 °C for up to 30 h. The chromatography fractions and the autolysis samples were analyzed by SDS−PAGE (Laemmli, 1970). The enzymes and autolysis samples were also applied to HPLC on a reverse-phase C18 column (Senshu-Pak, 4.6 × 25 cm) by the gradient elution with acetonitrile in 0.1% trifluoroacetic acid, monitoring the effluent at 225 nm.

The nonspecific protease fractions from red-agarose chromatography were combined, concentrated, and chromatographed on a Superose column, and the fractions with caseinolytic activity were collected, concentrated, and further chromatographed on a DEAE-Sephacel column (1 × 10 mm). The run-through fractions with high specific activity were combined and concentrated with a micro-ultrafiltration unit (Mol-Cut L, M_r 10 000 cutoff, Millipore Japan).

Measurement of Hatching and Caseinolytic Activities. The hatching enzyme activity was measured basically as previously described using ethanol-fixed four-cell stage embryos as substrate (Nomura et al., 1991). The fixed embryos were washed free of ethanol and equilibrated with assay buffer (10 mM HEPES, 0.5M NaCl, 10 mM $CaCl_2$, 10 mM $MgCl_2$, pH 8.2), and a 200 μL suspension of about 100 embryos in the assay buffer was placed in each well of a 96-well microplate. After addition of 10 μL enzyme sample, the mixture was incubated with occasional shaking for up to 30 h at room temperature, counting the number of completely denuded ("hatched") embryos at an adequate time interval on an inverted microscope. The number of the denuded embryos in each well was occasionally counted, and the time required to attain 50% and 100% hatching was recorded.

Nonspecific protease activity was measured in the following way. To 200 μL of 0.6% casein (Hammarsten, Merck) solution in 50 mM Tris-HCl (pH 7.5) containing 5 mM $CaCl_2$ and 0.1 M NaCl, was added 10 μL enzyme solution and incubated at 30 °C for 20 h. The reaction was stopped by adding 130 μL of 30% trichloroacetic acid, and the resulting solution was then vigorously mixed on a shaker and kept still for 10 min. The tubes were then microcentrifuged at 10 000 rpm for 10 min, and the supernatant was measured for A_{280} on a Shimadzu UV-160 spectrophotometer.

Amino Acid Sequence Analysis. Isolated HEz and its fragments derived by autolysis and nonspecific protease were analyzed directly or after HPLC for their N-terminal sequences on a protein sequencer, ABI model 476A, with the data processing software 610A.

Preparation of RNA. Total RNA was prepared from various developmental stages of the embryos by the LiCl method (Cathala et al., 1983). Poly(A)$^+$ RNA was then purified by two passages of the total RNA over a column of oligo(dT)-cellulose (Pharmacia Biotech, Uppsala, Sweden) (Davis et al., 1986).

cDNA Library Construction. A cDNA library (1.85 × 106 pfu) from poly(A)$^+$ RNA isolated from prehatching blastula stage embryos (cultured for 10−12 h at 20 °C) was constructed in λgt11 using the cDNA synthesis system and the cDNA cloning system λgt11 (Amersham, U.K.).

Polymerase Chain Reaction (PCR). Six degenerate oligonucleotide primers were synthesized on the basis of the partial amino acid sequences of the 38K form of HEz (NVAVTY for primer 2 and AIKNEL for primer 4), of the peptides generated by autolysis (FGNILY for primer 5 and IDRVEAV for primer 6), and of the known sequences highly conserved among the mammalian MMPs (CGVPDV for primer 1 and AAHEFGHS for primer 3). Namely, primer 1, 5′TG(C/T)GG(A/C/T/G)GT(A/C/G/T)CC(A/C/G/T)GA-(C/T)GT3′; primer 2, 5′AA(C/T)GT(A/C/G/T)GC(A/C/G/T)GT(A/C/G/T)AC(A/C/G/T/)TA3′; primer 3, 5′GC(A/C/G/T)CA(C/T)GA(A/G)(A/C/T)T(A/C/G/T)GG(A/C/G/T)CA3′; primer 4, 5′AT(C/T)TC(A/G)TT(C/T)TT(A/G/T)AT(A/C/G/T)GC3′; primer 5, 5′(A/G)TA(A/C/G/T)A(A/G)(A/G/T)AT(G/A)TT(A/C/G/T)CC(A/G)AA3′; and primer 6, 5′GC(C/T)TC(A/C/G/T)AC(A/C/G/T)C(G/T)(A/G)TC(A/G/T)AT3′.

These primers were used to amplify the total cDNA synthesized from *H. pulcherrimus* prehatching blastulae mRNA in a standard PCR using a RoboCycler 40 (Stratagene, U.S.A.). The PCR contained 100 ng of total cDNA, 50 mM KCl, 1.5 mM $MgCl_2$, 2.5 mM each of dATP, dCTP, dGTP, and dTTP, 5.0 units of Taq polymerase, 0.1 mM of each primer, and 10 mM Tris-HCl (pH 8.3). The reaction parameters were as follows: 1.5 min at 94 °C, followed by 35 cycles of 0.5 min at 94 °C, 0.5 min at 46 °C (first, second, and fifth cycles) or 37 °C (other cycles) and 1 min at 72 °C. An aliquot of the reaction mixture was reamplified under the same condition except for the annealing temperature (40 °C for the combination of primers 1−5, 2−5, and 1−4; 44 °C for primers 2−6; and 46 °C for primers 3−6), and the products were analyzed by agarose gel electrophoresis. Amplified DNA fragments were subcloned into the plasmid vector pBluescript, and each DNA insert was sequenced by the dideoxy method (Sanger et al., 1977) using 7-DEAZA sequencing kit, version 2.0, or Sequenase sequencing kit.

Molecular Cloning. A 1044 bp cDNA fragment generated by PCR was used as a probe to screen a *H. pulcherrimus* prehatching blastulae cDNA library. About 9 × 10^4 plaques were screened on replicate Hybond-N membranes with ^{32}P-labeled cDNA fragment. Positive clones were rescreened

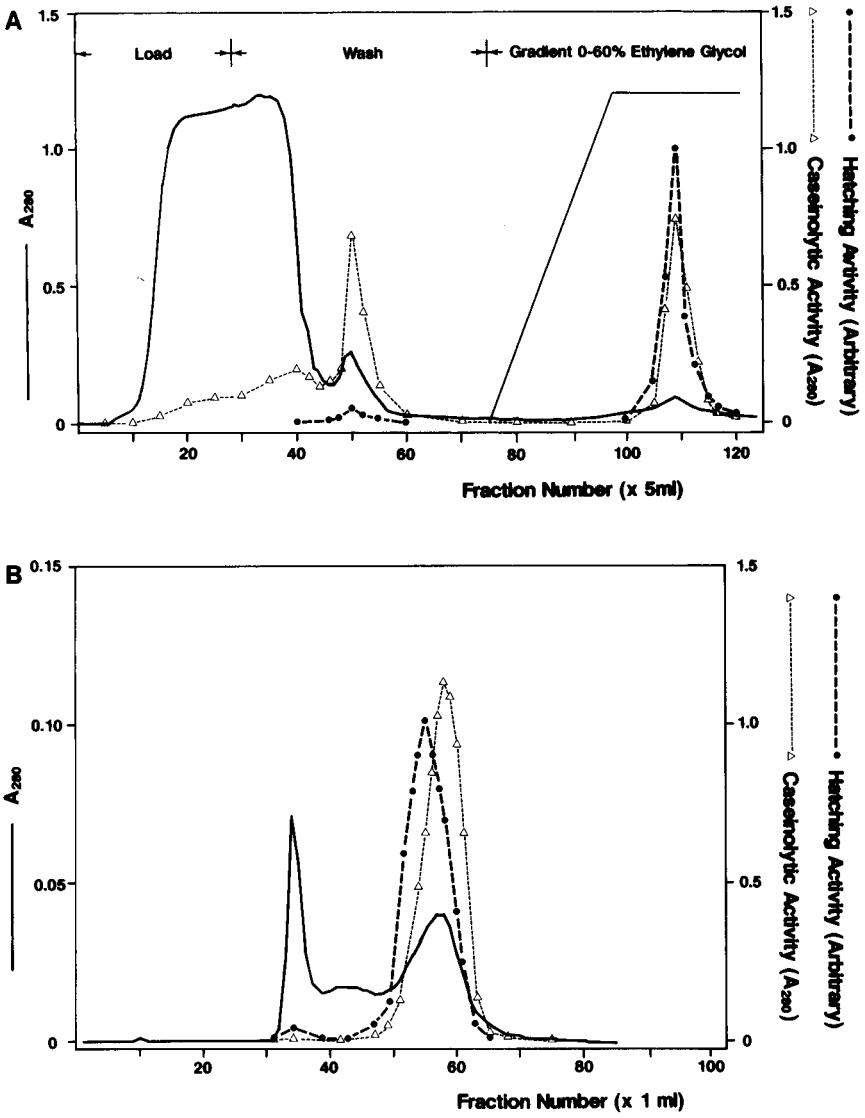

FIGURE 1: (A) Reactive-red-agarose chromatography of crude HEz. The incubation medium of the hatched embryos (3 L) was concentrated to 150 mL and loaded to the column. It was eluted by the starting buffer and by the gradient of ethylene glycol (0−60%). The nonspecific protease activity was eluted after the large run-through peak, and the HEz activity was eluted by the gradient of ethylene glycol (0−60%). (B) Superose 12 PG column chromatography. The fractions with hatching activity from the reactive red agarose chromatography (A) were combined and dialyzed against buffer A to remove ethylene glycol, concentrated to about 1 mL, and loaded to the column. It was eluted with buffer A, and the fractions with hatching and/or caseinolytic activities were collected and stored frozen at −30 °C.

until pure, eight positive clones were finally obtained, and the phage DNA was purified. The largest cDNA insert of the isolated clones, λHE4-1-1 with 2.5 kbp was subcloned into the plasmid vector pBluescript IIKS(+). Serial deletion mutants of the subclone were made according to the method of Yanisch-Perron et al. (1985). Nucleotide sequences were determined by the dideoxy chain termination method (Sanger et al., 1977) and analyzed on DNASIS software (Hitachi Software Engineering Co., Yokohama).

Northern Blot Analysis. Northern blot analysis was carried out using 1 μg of poly(A)$^+$ RNA isolated from prehatching blastulae. The size of RNA was estimated using the 0.24−9.5 kb RNA Ladder (GIBCO BRL, Gaithersberg, MD) as markers.

Hydropathy Plot. Hydropathy index for *H. pulcherrimus* envelysin was calculated by the method of Kyte and Doolittle (1982) at a window of 12 residues on DNASIS.

Antibody Preparation and Western Blot Analysis. A peptide with the N-terminal partial sequence of mature envelysin, YVTGGIAWPRN-C, was synthesized as an antigen for the preparation of an antibody specific to mature HEz. A Cys residue was attached to the C-terminus of the peptide to enable the conjugation onto maleimide-activated keyhole limpet hemocyanin. The antibody was prepared in a Japanese White Rabbit, and the antisera were stored at −80 °C. The enzyme samples were electrophoresed, electroblotted to a PVDF membrane at 5 °C, and incubated with antiserum and then with peroxidase-conjugated anti-rabbit IgG antibody. The antigens were visualized with diaminobenzidine as a substrate of the peroxidase.

RESULTS

Purification of Envelysin and Its Autolysis Products. The entire chromatogram of the Reactive-red-120-agarose chromatography for HEz purification is shown in Figure 1A for the first time, because it was not presented in the former reports (Lepage & Gache, 1989; Nomura et al., 1991). The large broad peak which passed through the column contains most of the FE component proteins and their breakdown products with little HEz activity. After the large peak, nonspecific protease fractions with only a slight hatching

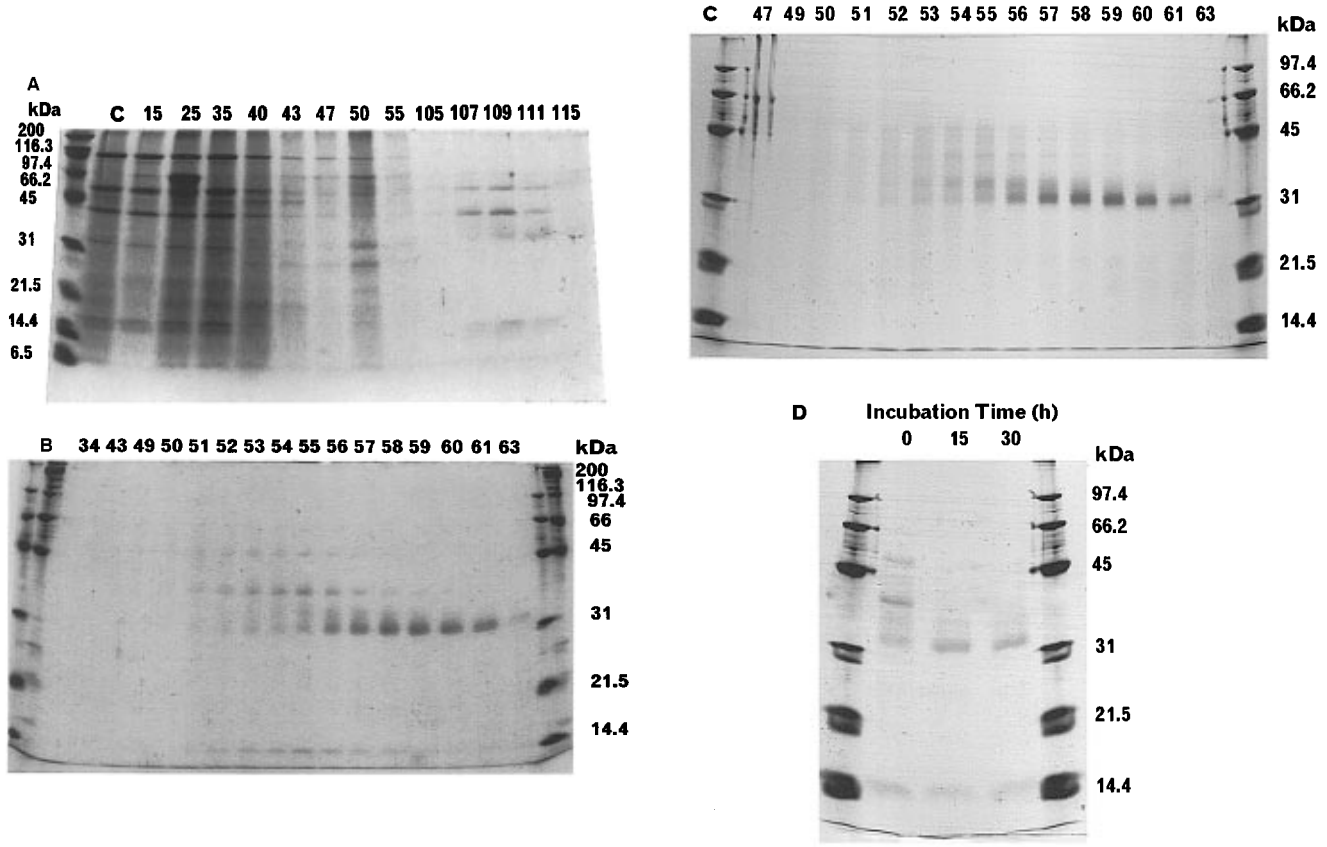

FIGURE 2: SDS−PAGE analysis of crude envelysin and the fractions eluted from red-agarose (A) and Superose 12PG [with 2-ME (B) and without 2-ME (C)], and of autolytic degradation products of purified HEz (D). In panel A samples were run on a 14% minislab gel, and in panels B, C, and D they were run on 13.5% gels of standard width with various height.

activity were eluted. A very small peak with both high hatching and caseinolytic activities emerged by the elution with 60% ethylene glycol after the gradient. On microplate assay, the fraction with the highest hatching activity (no. 109) dissolved the FEs of all the embryos in 30 min. The active fractions (nos. 105−111) revealed by SDS−PAGE the bands of 38K, 50K, 32K, and 15K in the order of band density (Figure 2A). The fractions with hatching activity were combined, dialyzed against the starting buffer, concentrated with Amicon PM10 membrane and Mol-Cut L to about 1.0 mL and applied to a Superose 12 PG column (1.6 × 50 cm). The resolution was better than Superose 12 HG (1.0 × 30 cm) mainly because of its larger column size. By this chromatography some contaminating substances with UV absorbance but without Coomassie stain were eliminated, and the purified enzyme was eluted in fractions 48−65 (Figure 1B). In the chromatogram, the peaks of hatching and caseinolytic activities did not coincide, but rather separated from each other by three fractions, the hatching activity being eluted faster. By SDS−PAGE (Figure 2B) the active fractions again revealed the bands of 50K, 38K, 32K, and 15K with a diffuse stain at ca. 34K. However, the relative density of 50K band became far lower than in Figure 2A, and the 32K band was the densest. This tendency was already observed for the concentrated enzyme solution before gel filtration. The 38K band was detected in the fractions 52−61, while the 32K band was in the fractions 54−63, corresponding to the hatching and caseinolytic activities, respectively. It was also revealed that the 38K and 15K bands were coeluted with parallel density for each fraction (51−59) and that the 32K protein was the last eluted (Figure 2B). These results suggest that mature HEz with the original M_r of 50K autolyzed mainly by a single cleavage to 38K and 15K forms linked by a disulfide bridge maintaining its hatching activity, fully or slightly less, and further via some intermediates or directly to apparently stable 32K form without specificity against FE during the hatching process and purification procedure. As described below, the mature HEz has a calculated M_r of 48 158 based on the deduced amino acid sequence without sugar. Thus, the purified HEz turned out to be a mixture of minor 50K and major 38K(+15K) forms retaining the hatching activity. The 38K(+15K) form corresponds to the 37K form in the previous study (Nomura et al., 1991). The faint diffuse band above 32K band suggests the presence of some other autolytic pathways. By SDS−PAGE without 2-ME (Figure 2C) no 15K band was detected in any fraction but a diffuse stain between 50K and 38K was detected in the fractions with HEz activity. The 32K band was detected in the same way as in Figure 2B. Although the 50K band was not detected as sharply as under reducing conditions, the result further supports the above conclusion for the 38K(+15K) form.

The nonspecific (caseinolytic) enzyme weakly bound to the red-agarose column (fractions around no. 50, Figure 1A) was further purified by Superose 12HR and DEAE-Sephacel chromatographies to homogeneity as revealed as a single band of M_r of 32K by SDS−PAGE (not shown). The enzyme ran through the DEAE-Sephacel column similarly to the *P. lividus* enzyme (Lepage & Gache, 1989).

Autolytic Degradation of Envelysin. Small amount of purified envelysin (fraction no. 54) with major 38K and minor 50K and 32K forms was incubated at 25 °C for up to 30 h and analyzed by SDS−PAGE. After 15 h, the 50K and 38K bands disappeared and only the 32K and 15K bands

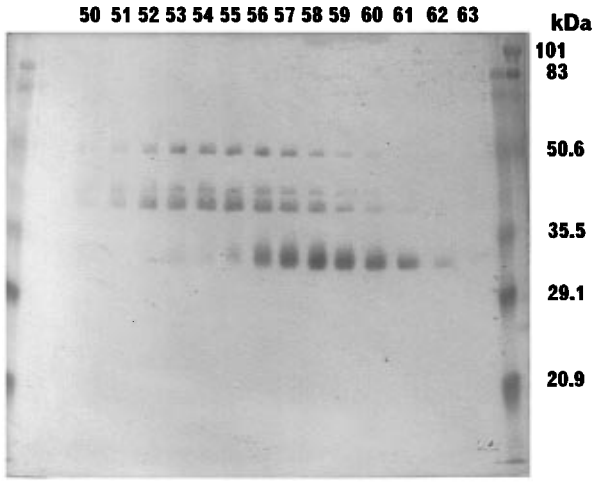

FIGURE 3: Immunoblot analysis of the fractions from Superose 12PG chromatography. The samples were electrophoresed as in Figure 2B, electrotransferred to a PVDF membrane, and treated with the antiserum against a peptide with the N-terminal 11-residue sequence of mature HEz.

and diffuse stain above 32K were detected (Figure 2D). The possibility of some other contaminating protease(s) is not completely ruled out, but it is unlikely. The 32K form did not digest itself further and did not produce any smaller band even after 30 h of incubation. The 32K product is considered to be identical with the nonspecific protease in Figure 1.

Immunoblot Analysis. The immunoblot of the Superose fractions using an anti-serum specific to the N-terminus of mature HEz revealed that the bands of 50K, 38K, 32K, and weak broad bands around 40K and 34K share a common antigenic determinant (Figure 3), confirming that these molecular species are derived from the mature 50K HEz by consecutive autolytic truncation of C-terminal fragments. In contrast, immunostain was not detected for the 15K band which was clearly Coomassie stained (Figure 2B). In this experiment, the density of the immunostain for each band was not parallel with that of Coomassie stain. Those with relatively high M_r, the 50K and ca. 40K bands, were stained more densely by antiserum than with Coomassie dye.

Partial Amino Acid Sequences of Envelysin. Portions of purified envelysin and nonspecific protease were directly applied to N-terminal sequence analysis. Because of inhomogeneity due to the autolysis products and some low M_r contaminants, the sequences were not determined unequivocally. The major barely readable sequences were (Y/N)(V/R)(T/P)(G/P)(G/S)(I/S)(A/R)- and (Y/N)(V/R)(T/G)(G/S)(G/S)(I/A)(A/G)-, respectively, suggesting the same N-terminal sequences for the two enzyme samples. Therefore, both enzyme samples were applied to HPLC and the resultant peaks were sequenced. Two peaks, major and minor, from the 32K nonspecific protease yielded the same N-terminal 51- and 39-residue sequences, YVTGGIAWPR-NVAVTYS-FGT-LSNDLNQNAI-KNEIRRAFQV-XDDXSGNXFR-N. We assumed that both peaks represent the whole body of 32K form, chromatograhically distinguished by slightly different post-translational modification or minor C-terminal truncation.

A portion of purified HEz, already containing autolysis fragments, was applied to HPLC. Only two of several isolated peaks yielded N-terminal sequence data. The peak that eluted at 8 min appeared to be a mixture of small fragments, but only the sequence YQTSTR- was determined.

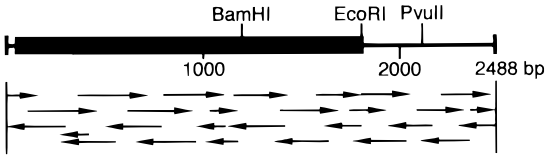

FIGURE 4: The restriction endonuclease map and sequencing strategy for a λHE4-1-1 cDNA insert. The map shows only the relevant restriction sites. The direction and length of the sequences determined are indicated by arrows. The deduced open reading frame is shown by the solid box.

The peak at 19 min was first sequenced as LQRRTRIRRY-FGNILYALID-REAV-, but some ambiguity was later found as LQRR(T/S)RIRRY-F(G/R)N(I/T/G)LYALID-(G/R)EAV- by re-examination of the sequencing data. These sequences, distinct from the N-terminus of mature HEz, are rationally assigned to the sites of autolytic cleavage. The peaks of 50K and 38K(+15K) forms were not found in the HPLC chromatogram, presumably because the gradient applied (0−60%) was not enough for the elution of active enzymes. The above four underlined partial sequences were used for the synthesis of oligonucleotide primers.

Isolation and Sequence Analysis of the cDNA Encoding H. pulcherrimus Envelysin. A cDNA library representing the mRNA from *H. pulcherrimus* prehatching embryos was constructed in λgt11 and screened with a 1044 bp PCR product which contained the predicted amino acid sequence 180−528:

NVAVTYSFGT-LSNDLSQTAI-KNELRRAFQV-WDDVSSLTFR-EVVDSSSVDI-
RIKFGSYEHG-DGISFDGQGG-VLAHAFLPRN-GDAHFDDSER-WTIGTNSGTN-
LFQVAAHEFG-HSLGLYHSDV-QSALMYPYYR-GYNPNFNLDR-DDIAGITSLY-
GRNTGSTTTT-TRRPTITRTT-TRRTTTRRTT-TQLATTQTTT-IRPPTYPTPP-
RQACTGSFDA-VIKDNSDRIY-ALAGRYYWRL-DQASPSWGWR-NRFGFDLPEN-
VDASFQNGIF-SYFFSGCYYY-YQTSTTRRRF-PRTPFNRRWV-GLPCDIDAVY-
KSGDSGTTYF-FKGRFVYKFS-SSNQLQRRSP-ISSYFRNTPY-ALRDGVEAV.

The analysis of 9×10^4 recombinants from an amplified cDNA library finally yielded 8 positively hybridizing clones. Among them, the clone λHE4-1-1 containing 2.5 kbp insert DNA was chosen for nucleotide sequence determination. The sequencing strategy for the clone is shown in Figure 4, and the complete nucleotide sequence and the deduced amino acid sequence are presented in Figure 5. The λHE4-1-1 cDNA insert was 2488 bp in length. The nucleotide sequences of six primers used for the probes to screen *H. pulcherrimus* HEz were all found in the sequence with some alteration: 526−542 for primer 1, 580−596 for primer 2, 633−650 for primer 3, 895−911 for primer 4, 1582−1600 for primer 5, and 1606−1626 for primer 6. We have assigned the initiation codon to the ATG at position 43−45, because there is an upstream in-frame stop codon, and furthermore this ATG is flanked by the sequence that fits Kozak's criteria for translation initiation codon, (A/C)-XXAUGG (Kozak, 1981). The initiation codon is followed by an open reading frame of 1770 bp. An in-frame stop codon occurs at the position 1817, and the 3′-untranslated region composed of 672 bp includes polyadenylation sites (AATAAA at positions 2102−2107, 2119−2124, and 2442−2447).

Deduced Amino Acid Sequence of Envelysin. The whole deduced amino acid sequence shown in Figure 5 indicates that the preproenzyme of 591 amino acid residues with the

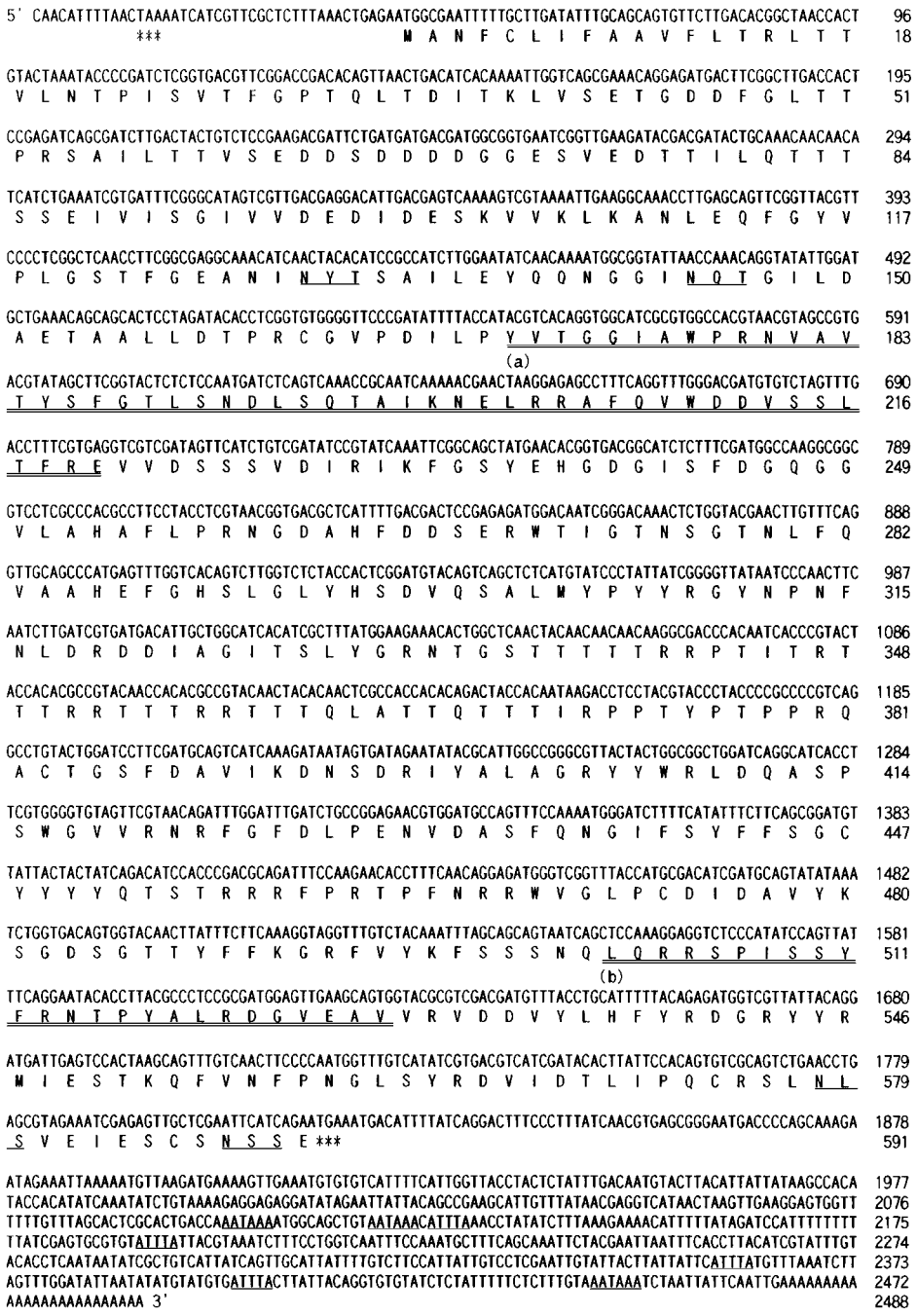

FIGURE 5: Complete nucleotide sequence and deduced amino acid sequence of the λHE4-1-1 cDNA insert. Nucleotide and amino acid residue numbers are indicated on the right side. The triple asterisks (***) denote the stop codons (TAA and TGA). Potentially N-linked carbohydrate binding sites [NX(S/T)] and polyadenylation signal sequence at the 3′-untranslated region (AATAAA) are underlined. The N-terminal amino acid sequences of the purified envelysin (YVTGG-) and of the peptide generated by autolysis (LQRRS-) are indicated by double underlines.

calculated M_r of 66 125, contains a hydrophobic region of about 20 amino acid residues in its N-terminal part, Asp- and Glu-rich region (EDDSDDDDGGESVED---VVDEDI-DES) in the propeptide, and four potentially N-linked glycosylation sites at residues 129−132 (NYT) and 144−146 (NQT) in the propeptide, and 578−580 (NLS) and 588−590 (NSS) near the C-terminus. The total number is the same, but their distribution is different from that of *P. lividus*: three in the propeptide and one near the C-terminus. The signal peptide is determined to be the residues 1−18, based on the method by von Heijne (1986). The N-terminal amino acid sequence of the purified enzyme (YVTGGIA-) suggests that the activation of proenzyme proceeded by the cleavage at Pro(169)-Tyr(170), yielding the 422-residue mature enzyme with the calculated M_r of 48 158 (without sugar), corresponding to the intact mature 50K band, but much larger than the major form of the purified envelysin with M_r 38K. This confirms that the smaller form is derived by truncation of its C-terminal portion of ca. 12 kDa size by autolytic degradation during hatching process and purification procedures. The cleavage occurred at the Gln-Leu (503−504) bond but the C-terminal fragment with the LQRRSPIS-SY- (504−591) sequence was not liberated from the major body of HEz, presumably due to the presence of intact disulfide bond as described below. The 15K band detected by SDS−PAGE may correspond to this fragment.

Near the end of propeptide, HEz has a typical cysteine-switch consensus sequence, PRCGVPD (Sanchez-Lopez,

1988) followed by the Ile residue, less popular than Val for this position as found in *P. lividus* HEz and most of the mammalian MMP family members. To our knowledge the PRCGVPDI sequence with Ile was first found in the stromelysin-3 of African clawed frog (Patterson et al., 1995). Some more structural features described for *P. lividus* (Lepage & Gache 1990) were also found in *H. pulcherrimus* HEz. The domain structure of HEz is summarized in Figure 7A. The mature enzyme has five Cys residues, all of which are in the C-terminal domain, much more than two in most mammalian MMPs. The two Cys residues Cys-383 and Cys-586 is highly probably assumed to be disulfide-bridged to accomplish the four-blade propeller structure of hemopexin-like domain, by analogy to the 3-D structure of full-length porcine synovial collagenase (Li et al., 1995).

By the DNASIS program, the isoelectric points of the 591-residue preproenzyme, the 422-residue (50K) intact mature enzyme, and the 334- (38K) and 281-residue (32K) autolyzed forms were calculated to be 5.09, 9.14, 9.47, and 8.10, respectively, showing drastic elevation of p*I* upon activation. The p*I* of the 32K form is lower than expected because it ran through DEAE-Sephacel column at its final purification step. As shown by the hydropathy plot in Figure 7B, the entire molecule of the preproenzyme contains very few hydrophobic block except in the N-terminal signal sequence and several residues in the propeptide, indicating that both propeptide and mature enzymes are rather hydrophilic. The interaction of HEz with the substrate FE, however, is very strong, and this was the reason for the mysterious difficulty in purification of this enzyme for more than half a century.

The prediction of the secondary structure was partially successful. The catalytic domains of mammalian MMPs in general have three α-helices: helix A near the N-terminus, helix B containing the active site His and Glu residues, and helix C at the end of catalytic domain (Dhanaraj et al., 1996). In HEz, on the assumption that it has helices at the same positions as MMPs, the helices A and B were successfully predicted, although they are slightly shorter than those in human fibroblast collagenase (Spurlino et al., 1994). The helix C, however, was not predicted at the corresponding position (Figure 7C), but instead a sheet was predicted. The C-terminal domain of mammalian MMPs has almost no helix, having instead a few one or $^3/_4$ turn of a helix-like stretch (Li et al., 1995; Libson et al., 1995). Therefore the prediction of α-helices in this domain of HEz is overestimated.

The deduced amino acid sequence of the prepro-form of *H. pulcherrimus* HEz is 69.1% identical and 10% similar to that of *P. lividus*, and if limited to the catalytic domain (residues 170−330, tentatively), the sequences are 79.5% identical. Other domains have different identities: 44% for signal sequence, 64% for propeptide, and 69% for hemopexin-like domain. As compared with human stromelysin-1 (Saus et al., 1988) it is 23.7% identical, and for each domain, the identities were 6% for signal peptide, 13% for propeptide, 50% for catalytic domain, and 15% for C-terminal domain. It is 23.2% identical and 10.9% similar to porcine synovial collagenase (Clarke et al., 1990) for the whole precursor sequence, and 16% identical for signal peptide, 13% for propeptide, 45% for catalytic domain, and 16% for hemopexin-like domain (Figure 6B).

Northern Blot Analysis. In order to determine the size of the mRNA for the λHE4-1-1 cDNA insert and to know when the mRNA for HEz is expressed in embryos, poly(A)$^+$ RNAs prepared from various developmental stages of *H. pulcherrimus* embryos were analyzed by Northern blot hybridization using a part of the λHE4-1-1 cDNA insert (nucleotide numbers 580−1626) as a probe. A single hybridization signal was detected at the position of 2.8 kb by poly(A)$^+$ RNA from the embryos cultured at 20 °C for 6−12 h, with a maximal signal for the 8 h-cultured embryos (Figure 8A). No signal was detected in unfertilized eggs or in the embryos either within 3 h postfertilization or after 14 h. The envelysin activity in the medium of normal embryos reached maximum at 12.5 h. In the medium of nude embryos, whose softened FEs were stripped soon after fertilization by forcing them through a nylon net (Nomura et al., 1990), the activity peak was at 11.5 h, an hour earlier than normal embryos (Figure 8B). Thus the period of 1 h is considered to be the time required to resolve the FE surrounding the embryo.

Effect of Inhibitors on Envelysin Activity and Hatching. Bovine TIMP-1 inhibited the hatching activity on ethanol-fixed embryos dose-dependently with the I_{50} value of about 110 nM. Using the human α_1-proteinase inhibitor as substrate as previously reported (Nomura & Suzuki, 1993), TIMP-1 in molar excess was shown to inhibit HEz activity completely for a practical period of 45 h. Caseinolysis by both 38K and 32K enzymes was also inhibited by TIMP-1 dose-dependently with apparent I_{50} values of 26 and 88 nM, respectively (Figure 9). Inhibition of 32K form to a lesser extent indicates that the binding of TIMP-1 to the C-terminal domain of HEz is necessary for effective inhibition (Bigg et al., 1994; Taylor et al., 1996). In contrast, chymostatin (1 mM) was not inhibitory at all on the hydrolysis of substance P, as analyzed by the method previously reported (Nomura et al., 1991). The hydrolyses of casein and substance P were also inhibited by $ZnCl_2$ dose-dependently as measured up to 5 mM. The normal hatching and later development of *H. pulcherrimus* embryos were also inhibited by 10 μM $ZnSO_4$. The same result was observed for $ZnCl_2$.

DISCUSSION

After the hatching enzyme was found by Ishida in 1936, more than half a century had passed before the isolation of an intact sea urchin HEz was successfully achieved by Lepage and Gache in 1989. A procedure suggestive of successful purification had been reported briefly and fully described later for the sea urchin *Strongylocentrotus purpuratus* (Roe & Lennarz, 1985, 1990) by a method with some steps different from those of other species (Lepage & Gache, 1989; Nomura et al., 1991). The failure to isolate intact HEz in earlier studies is due to excess steps of conventional chromatographies, during which the intact HEz autolyzed to a nonspecific protease without (or with defective) C-terminal hemopexin-like domain which is essential to the protein substrate specificity, well-known for mammalian MMPs except for matrilysin (Clark & Cawston, 1989; Alan et al., 1991; Murphy et al., 1992). The point was that they used the procion-agarose chromatography for the first step of purification to quickly separate the HEz from the bulk of FE degradation products. By this method the time for purification was drastically shortened, and thus the autolysis was practically suppressed during the isolation procedure. Procion-agarose or red-agarose, one of the dye-ligand affinity chromatography media, had been successfully employed for the purification of collagenases and stromelysins. A rabbit bone "proteoglycanase" was purified with "procion-red-

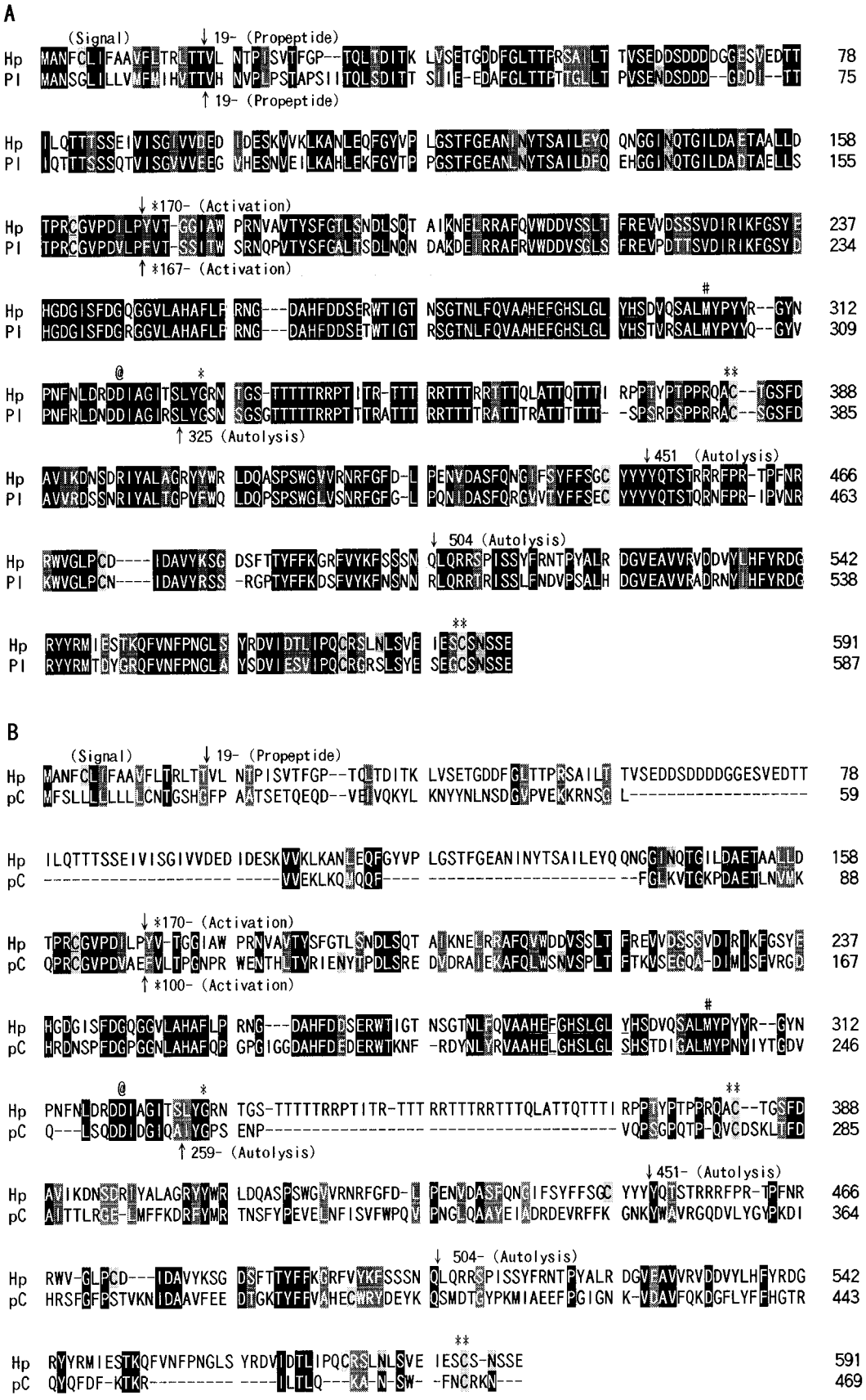

FIGURE 6: Sequence alignment of *H. pulcherrimus* envelysin with *P. lividus* envelysin (A), and with porcine synovial collagenase (B). The amino acid sequences of the two envelysins (Hp and Pl) were aligned first by the DNASIS software and further manually to obtain maximal identity by inserting blanks. Alignment with porcine collagenase (pC) in (B) was performed manually. The residues that are identical between the two are shown by white-on-black letters, and the residues of similar nature, belonging to any of the groups (G, A, S, T), (D, E), (N, Q), (K, R), (V, I, L, M), or (F, Y, W), are shown white-on-gray. The Cys residues and potentially N-glycosylated Asn residue are typed black-on-gray. The consensus sequences for cysteine switch and zinc signature are underlined. The symbols above the residues denote the following: *- - -*, tentative catalytic domain (Tyr 170−Gly 330 for Hp envelysin); #, Met residue in the "Met-turn" motif; @, Asp residue essential to Ca^{2+} -binding; **, the two Cys residues, supposedly disulfide-bridged in the hemopexin-like domain. The arrow heads indicate the cleavage sites for signal sequence, proenzyme activation and autolysis. The sequences and cleavage sites were adopted from Lepage and Gache (1990; *P. lividus* envelysin) and Richards et al. (1991) and Clark et al. (1995; porcine collagenase).

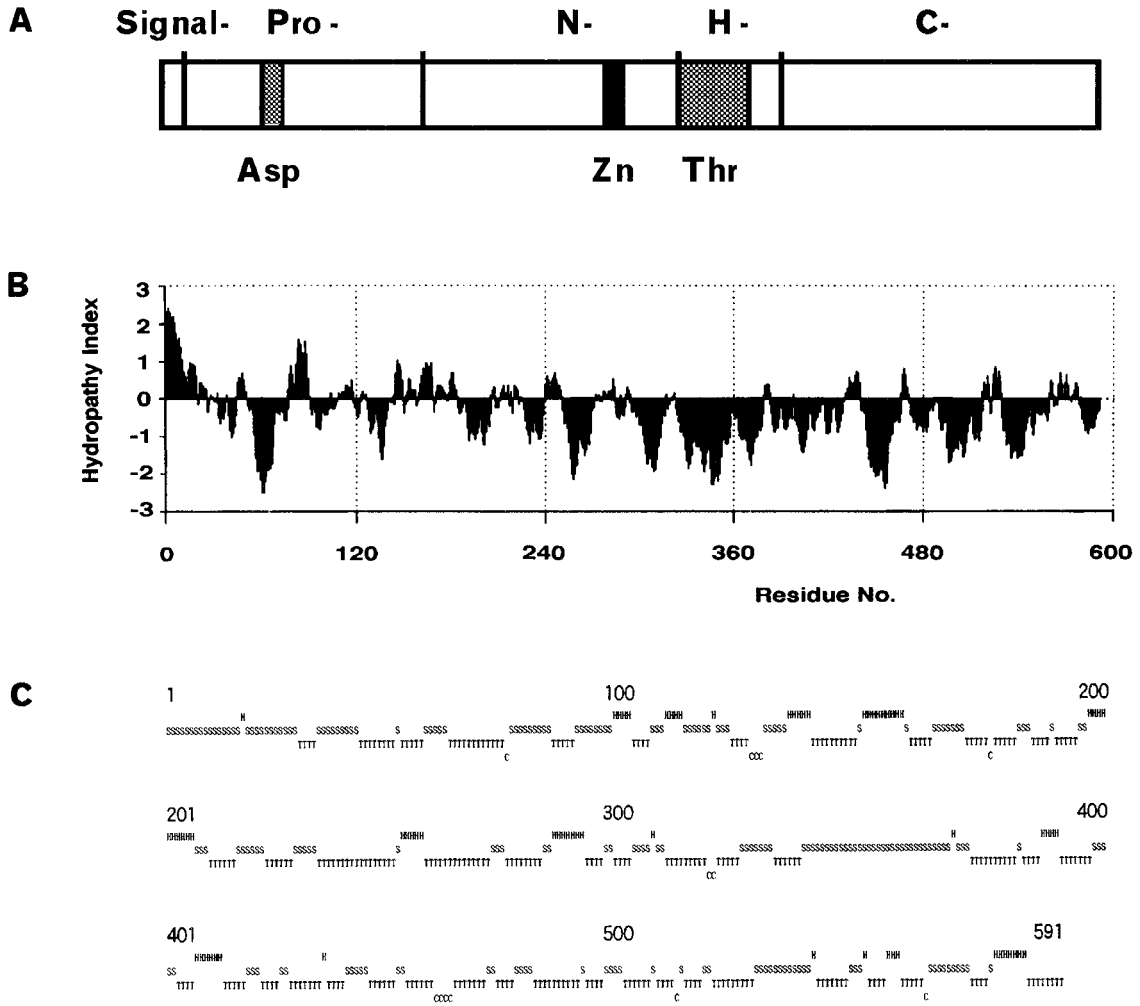

FIGURE 7: Domain structure (A), hydropathy plot (B), and predicted secondary structure (C) of *H. pulcherrimus* envelysin. (A) The bar indicating the total sequence is divided to signal peptide, pro-sequence, N-terminal domain, hinge region, and C-terminal domain. The characteristic Asp/Glu-rich stretch, zinc signature, and Thr-rich stretch are shown by shaded or black boxes. (B) The hydropathy index was calculated by the method of Kyte and Doolittle (1982) at a window of 12 residues. The horizontal axis with the same scale as in A separates the hydrophobic (positive) and hydrophilic (negative) regions. (C) The secondary structure was predicted by DNASIS software based on the method of Chou and Fasman (1978). The letters H, S, T, and C stand for Helix, Sheet, Turn, and Coil, respectively.

agarose" used at the final step (Galloway et al., 1983) and stromelysin was isolated from rabbit synovial fibroblasts with procion-red-Sepharose (Chin et al., 1985). Research on vertebrate collagenases, which led to the currently flourishing mammalian MMP studies, also have a long history with difficulty in purification after their findings in the involuting tadpole tail and rat uterus in the early 1960s (Woessner, J. F., 1962; Gross & Lapiere 1962).

Red-agarose chromatography provided an excellent method for purification of sea urchin HEz, as was proved by the successful isolation of *P. lividus* enzyme as a 51 kDa intact mature form without any autolyzed form (Lepage & Gache, 1989). In our study, however, it still takes many hours to remove 60% ethylene glycol by dialysis followed by concentration, during which the autolysis of intact enzyme proceeds to a measurable extent. The intact 50K mature enzyme was hardly isolated as a stable entity; rather, it rapidly autolyzed to 38K(+15K) and further to 32K forms during purification procedure, thus making the "purified" HEz as a mixture of 50K and 38K(+15K) or of 38K(+15K) and 32K molecular species as shown by SDS−PAGE of Superose fractions. In the previous study we isolated 37K enzyme as an intact HEz without knowing the significant role of the 15K fragment which was faintly detected in the SDS−PAGE plate (Nomura et al., 1990). To avoid autolysis more specific affinity chromatography with easier elution is preferable, e.g., active-site directed peptide hydroxamic acid columns without the use of ethylene glycol (O'Hare et al., 1995). Jaspisin, a HEz inhibitor produced by a sponge (Ikegami et al., 1994) may also be useful as a ligand for such column. Another tactic may be the use of some reversible inhibitors which suppress the autolysis during purification and storage. Zinc ion or some other recently developed inhibitors for MMPs in adequate concentration may be the candidates for such additives.

By the purification procedures of HEz so far reported, it is difficult to isolate proenzyme and thus to elucidate the mechanism of proenzyme activation, because the source of HEz was the supernatant of hatched blastulae medium, available only after the enzyme finished its role. Proenzyme and autolysis-resistant mutants of HEz in substantial amount may be obtained by the expression in adequate cells. Once obtained they will be of great value to study the mechanism of proenzyme activation and the crystal structure and will provide us much new information. Distribution of HEz protein in whole embryos was studied by immunolabeling (Lepage et al., 1992b).

The first report on the cDNA cloning of HEz appeared about 6 years ago (Lepage & Gache 1990); we show here the second cloning data. A part of the deduced amino acid

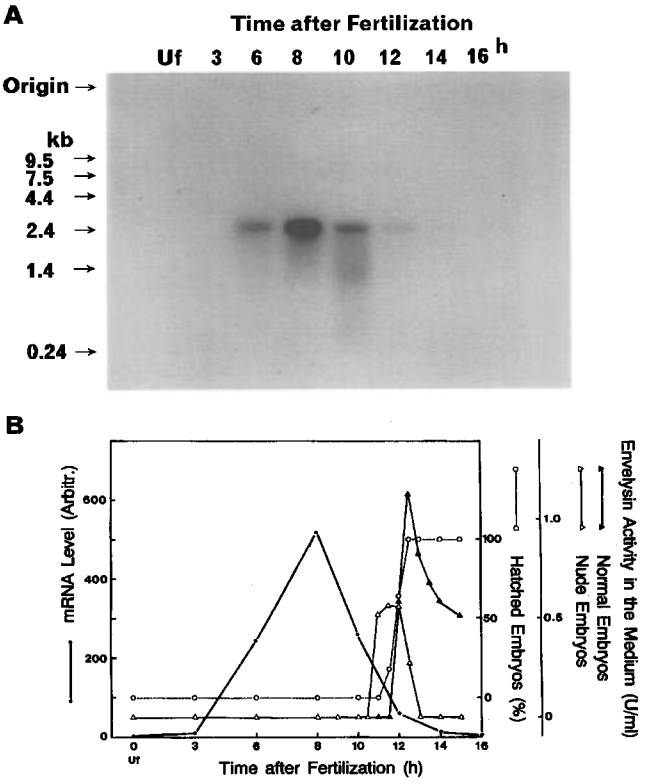

FIGURE 8: (A) Northern blot analysis. Poly(A)$^+$ RNA prepared from various developmental stages of *H. pulcherrimus* embryos cultured at 20 °C was hybridized with a portion (nucleotides 580−1626) of the λHE4-1-1 cDNA insert. Embryos were cultured at 20 °C up to 16 h after insemination. They developed through the stages of "early blastula (6 h)," "mid-blastula (8−10 h)," and the hatching period starting at 11.5 h. (B) The mRNA level in arbitrary units, the percentage of hatched embryos, and the envelysin activity in the media of normal and denuded embryos in arbitrary unit are plotted against the time after insemination.

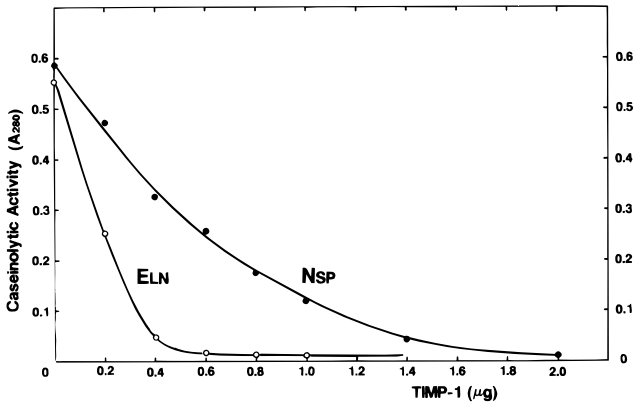

FIGURE 9: Inhibition of envelysin and nonspecific protease (autolyzed 32K form) by recombinant bovine TIMP-1. Activities of both types of enzyme are plotted against the weight of TIMP-1 in each tube. ELN, envelysin; NSP, nonspecific protease.

sequence (LQRRSPISSYFRNTPYALRDGVEA-) was not identical with that of the isolated autolysis fragment [LQRR-(T/S)RIRRYF(G/R)N(I/T/G)LYAL(I/G)D(G/R)VEAV-] possibly due to the limitation of peptide sequencing, particularly at the chromatography of PTH-amino acids, microheterogeneity in the enzyme protein from 30 individuals, polymorphism, or mutation during PCR amplification of DNA fragment. Similar inconsistency was observed between the cDNA and gene of *P. lividus* HEz (Ghiglione et al., 1994).

Two prepro-forms of HEz were revealed to have the sequence homology of 69% identity and 10% similarity in the total sequence. The value 69% is slightly lower than those for mammalian stromelysins. Namely, rat, mouse, and rabbit stromelysins-1 are 76%, 79%, and 85% identical, respectively, to the human enzyme, and stromelysins-1 of mouse and rat are 90% identical to each other (Henriet et al., 1992). In contrast, human fibroblast and neutrophil collagenases are only 58% identical (Devarajan et al., 1991; Goldberg et al., 1986).

On the basis of their high homology envelysins are rationally placed in the MMP gene family to which the subfamilies of mammalian MMPs and sea urchin envelysins belong (Sang & Douglas, 1996). They have common VAA sequence immediately before the zinc signature HEXGH-SLGLXHS. The cell wall lytic enzyme of *Chlamydomonas reinhardtii* (Kinoshita et al., 1992) may also be placed in this family as the third subfamily, although it has a VIM instead of a VAA sequence. Several metalloproteinases of M_r 55K, 50K, 42K, and 38K were found in a developing sea urchin embryo by gelatin zymography (Quigley et al., 1993), implying their roles in various phases of development other than hatching.

In addition to different susceptibility to autolytic degradation, the sites of autolysis are also quite different among the two enzymes. In *P. lividus* the reported autolysis site was Ser-Leu (325−326) in and near the end of catalytic domain (Lepage & Gache, 1989) similarly to porcine synovial collagenase (Clark et al., 1995), while in *H. pulcherrimus* the enzyme cleaved itself consecutively at Gln-Leu (503−504) and Tyr-Tyr (450−451), cleaving off about $^2/_3$ of the hemopexin-like domain to the apparently stable 32K form. The partial sequences on both sides of the autolytically attacked peptide bonds in *H. pulcherrimus* HEz are YYY-YQTST (448−455) and SSNQ-LQRR (500−507). The *P. lividus* HEz also has the same or similar sequences YYY-YQTST and SNNR-LQRR, but these sites were not cleaved at all. Instead, it was cleaved only at IAGIRS-LYGSNSGS (319−332), while the similar sequence IAGITS-LYGRNTGS in *H. pulcherrimus* was not attacked. The difference in the sites of autolytic attack may be due to a subtle difference in the folding of the peptide chains and/or in their total 3-D structures. Determination of exact 3-D structures by X-ray crystallography in future will be of great help to answer this question.

The SDS−PAGE pattern of the Superose fractions supports that the autolyzed 38K(+15K) form maintains the size of intact 50K form. However, by the second cleavage at Tyr-Tyr (450−451), the fragment Tyr-Gln (451−503) may be unable to maintain the secondary and tertiary structures even if the disulfide bond is intact. This may induce a drastic change in the total structure of hemopexin-like domain and linker peptide, and this may prevent the proper binding of the substrate protein to accommodate the scissile peptide bond to the active site in the catalytic domain. The difference in size of the two molecular species 38K(+15K) and 32K under non-denaturing condition is evident by the Superose fraction pattern (Figure 1B). It is also evident by SDS−PAGE that 32K enzyme does not accompany the second polypeptide chain linked by disulfide bond. The pattern of the band density for 38K and 15K in Figure 2B and the hatching activity profile in Figure 1B correlates well, suggesting that the 38K(+15K) enzyme is responsible for the hatching activity, in addition to the 50K intact enzyme in minor proportion.

In contrast to the 50K and 38K(+15K) forms, the 32K form was eluted from the red-agarose column with only a slight retardation after the bulk of FE degradation products (Figure 1A), indicating that the 32K form, lacking the $^2/_3$ of the hemopexin-like domain, has only a weak interaction with the reactive-red dye ligand. This means that the enzyme's interaction with the reactive-red dye is mainly via C-terminal hemopexin-like domain, but not the catalytic domain, suggesting that the dye is not catalytic site-directed in the chromatography, although it was really inhibitory to HEz (Roe & Lennarz, 1990). It is an interesting question whether or not the cleavage of small peptide substrates by the 32K form is inhibited by the dye.

TIMP-1 inhibited the 38K(+15K) form more effectively than the 32K form, suggesting that its interaction with the C-terminal domain of HEz is essential for inhibition. This result is in accordance with the report that the TIMP inhibition of 22K form (catalytic domain) of human fibroblast collagenase is less pronounced than whole molecule, confirming the role of C-terminal domain in the interaction with the inhibitor (Bigg et al., 1994; Taylor et al., 1996).

In the hemopexin-like domain there are four motifs highly conserved among the mammalian MMPs and two HEzs, at the starting point of each of the four β-propeller leaflets, mostly containing **DA**V triplet: SF**DA**V(386−390), (N/G)-V**DA**(S/A)(F/Y)(Q/E) (431−435), I**DA** (V/A)(Y/I) (474−477), and (V/I)(**E**/**D**)**A**V(V/F) (525−529). These VDA motifs are conserved in other MMPs and hemopexins, suggesting their role in initiating the β-structure in each leaflet, as shown in the 3D-structure of hemopexin and C-terminal domain of MMPs by Faber et al. (1995), Libson et al. (1995), and Gohlke et al. (1996).

The envelysins of the two sea urchin species are distinguished from mammalian MMPs in that HEzs have insertion of characteristic sequences: Asp- and Glu-rich region in the propeptide and Thr- and Arg-rich stretch in the hinge region. Such a Thr-rich stretch, even if very short, is also found in stromelysin and in the MMPs designated as 2D-1, 2D-19, and 2D-24 of regenerating newt limbs (Miyazaki et al., 1996).

The two autolytic cleavage sites identified in this work, -Tyr(451) and -Leu(503), supposedly correspond to the beginning of the third flat arrow in the second leaflet and to the beginning of the fourth arrow in the third leaflet, respectively, based on the reported structure (Faber et al., 1995; Libson et al., 1995). These positions may be susceptible to autolytic attack of HEz due to the location at the outer surface of the molecule near the exit end of the channel-like opening.

As to another function of hemopexin-like domain, it would be an interesting question whether or not intact or partially degraded envelysin and other MMPs of mammalian origin has the hyaluronidase activity, since hyaluronidase has been reported to be identical with hemopexin (Zhu et al., 1994).

The physiological significance for the high tendency to autolysis of HEz is difficult to elucidate, because degradation of FE in free seawater or more precisely in perivitelline space, does not seem to require subtle regulation of the enzyme action, in sharp contrast to that of extracellular matrix proteins in tissues by mammalian MMPs.

The specificity of a protease is on a hierarchy structure. As to HEz the primary specificity concerning the residue or short sequence on one or both side of scissile bond is already reported by us as cleaving the amino side of hydrophobic residues, -Leu, -Tyr, -Phe, as well as -Ala, -Asn, etc. (Nomura et al., 1991). However, the protein substrate specificity of HEz has not been well elucidated. Among the FE proteins, 70K ovoperoxidase and >100K proteins were reported to be rapidly degraded, while 100K, 50K, and 15K proteins are rather resistant (Nomura et al., 1991; Roe & Lennarz, 1990; Nomura & Suzuki, 1995). The FEs of the embryos from various species were dissolved at different rates depending on the phylogenetical distance, indicating the species specificity (Nomura et al., 1990). The cleavage site specificity, or S1′-specificity for most of metalloproteinases can be understood to be determined by the structure of "S1′ specificity pocket" (Stams et al., 1994) close to the catalytic center, but the higher order protein substrate specificity is rationally assumed to be determined by the interaction of the substrate protein(s) and the enzyme's C-terminal domain including the linker peptide for MMP family members. This implies that the scissile bond S1−S1′ in the substrate FE protein is somewhat distant from the putative "substrate specificity determinant" which interacts with the HEz's "substrate specificity site".

The region essential for the substrate specificity of neutrophil collagenase was studied using various C-terminally truncated collagenases and chimera enzymes between collagenase and stromelysin (Hirose et al., 1993). The truncation of C-terminal four residues including the Cys-464 reduced its collagenase activity to 62%. The region responsible to the specificity against native collagen was assigned to be minimally the 16 residues in the middle of hemopexin-like domain. The chimera enzyme NC(1−326)-ST(327−467) retains 16.1% of the native collagenase activity, while NC(1−301)-ST(302−467) has only 0.94% activity. In our experiment the 32K enzyme with the linker peptide and $^1/_3$ of the hemopexin-like domain does not have the specificity to FE. These data imply that the sequence to maintain the full size of the enzyme, except for a few residues at the C-terminus, contributes essentially to maintain the substrate specificity. It is thus rational that the autolyzed 38K(+15K) form of HEz, with a single nick at Gln-Leu (503−504) and supposedly maintaining its original tertiary structure, retains its substrate specificity against homologous FE as well as catalytic activity.

For the next step of the investigation on the specificity and structure, some autolysis-resistant mutant enzymes prepared by site-directed mutagenesis, replacing the cleavage site residues -Leu(503) and -Tyr(451) with non-hydrophobic or charged residues, e.g., Gly, Lys, Glu, would be of great value. By the use of these mutant enzymes, elimination of autolysis or alteration of autolysis sites will be confirmed. And furthermore, these stable mutants will be of great help for the study on the relationship of the protein−substrate specificity and the C-terminal hemopexin-like domain, and its detailed mechanism. Such a study has already been reported for human collagenase (O'Hare et al., 1995).

A recent paper on collagen−collagenase interaction (De Souza et al., 1996) is highly suggestive for the study on the specificity of HEz or on the HEz-FE interaction, in that the hinge region of HEz may mimic the conformation of its target protein(s) in FE. The mechanism of FE dissolution by HEz is the major theme of the next study, and for that purpose the knowledge of FE structure is essential and of great interest. Isolation and characterization of FE proteins are now in progress and will be reported in the near future.

The cDNA cloning of the two sea urchin HEzs revealed that they have a close relation to mammalian MMPs. Choriolysins (HCE and LCE) of the teleost *Oryzias latipes* (medaka), in contrast, belongs to the family of crayfish astacin (Yasumasu et al., 1992; Yamagami, 1996). It is of interest that a group of invertebrate enzymes belongs to a family of vertebrate enzymes and vice versa. It is widely accepted that the enzymes playing important roles in degradation of structural proteins are mostly metalloproteinases. The cDNA clones encoding another metalloproteinase BP10, a homologue of human BMP-1, were isolated from the sea urchin embryo cDNA libraries in two sea urchin species, *P. lividus* (Lepage et al., 1992a) and *S. purpuratus* (Hwang et al., 1994), with similar sizes of 64K and 72K, respectively. They are now known to play important roles in the early development, dorsal−ventral patterning. With the accumulation of more and more sequence data of HEzs and other metalloproteinases (e.g., Quigley et al., 1993), a new phylogenetic relation of animals in terms of metalloproteinases will be clarified.

ACKNOWLEDGMENT

We are indebted to Masahiro Matada at Noto Marine Laboratory, Kanazawa University, for collecting and maintaining the sea urchins.

REFERENCES

Allan, J. A., Hembry, R. M., Angal, S., Reynolds, J. J., & Murphy, G. (1991) *J. Cell Sci. 99*, 789−795.

Baragi, V. M., Friszar, C. J., Conroy, M. C., Ye, Q.-Z., Shipley, J. M., & Welgus, H. G. (1994) *J. Biol. Chem. 269*, 12692−12697.

Barrett, D., & Edwards, B. F. (1976) *Methods Enzymol. 45*, 354−373.

Barrett, D., Hursh, D. A., Landercasper, J., & McRorie, M. M. (1978) *ICN-UCLA Symp. Mol. Cell. Biol. 12*, 207−215.

Bigg, H. F., Clark, I. M., & Cawston, T. E. (1994) *Biochim. Biophys. Acta 1208*, 157−165.

Cathala, G., Savouret, J. F., Mondez, B., West, B. L., Karin, M., Martial, J. A., & Baxter, J. D. (1983) *DNA 2*, 329−335.

Chin, R. J., Murphy, G., & Werb, Z. (1985) *J. Biol. Chem. 260*, 12367−12376.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45−148.

Clark, I. M., & Cawston, T. E. (1989) *Biochem. J. 263*, 201−206.

Clark, I. M., Mitchell, R. E., Powell, L. K., Bigg, H. F., Cawston, T. E., & O'Hare, M. C. (1995) *Arch. Biochem. Biophys. 316*, 123−127.

Davis, L. G., Dibner, M. D., & Battery, J. F. (1986) *Basic Methods in Molecular Biology*, Elsevier, New York.

De Souza, S. J., Pereira, H. M., Jacchieri, S., & Brentani, R. R. (1996) *FASEB J. 10*, 927−930.

Devarajan, P., Mookhitar, K., Van Wart, H., & Berlinger, N. (1991) *Blood 77*, 2731−2738.

Dhanaraj, V., Ye, Q.-Z., Johnson, L. L., Hupe, D. J., Ortwine, D. F., Dunbar, J. B., Jr., Rubin, J. R., Pavlovsky, A., Humblet, C., Blundell, T. L. (1996) *Structure 4*, 375−386.

Edwards, B. F., Allen, W. R., & Barrett, D. (1977) *Arch. Biochem. Biophys. 182*, 696−704.

Faber, H. R., Groom, C. R., Baker, H. M., Morgan, W. T., Smith, A., & Banker E. N. (1995) *Structure 3*, 551−559.

Fini, M. E., Karmilowicz, M. J., Ruby, P. L., Beeman, A. M., Borges, K. A., & Brinckerhoff, C. E. (1987) *Arthritis Rheum. 30*, 1254−1264.

Galloway, W. A., Murphy, G., Sandy, J. D., Gavrilovic, J., Cawston, T. E., & Reynolds, J. J. (1983) *Biochem. J. 209*, 741−752.

Ghiglione, C., Lhomond, G., Lepage, T., & Gache, C. (1994) *Eur. J. Biochem. 219*, 845−854.

Gohlke, U., Gomis-Rueth, F.-X., Crabbe, T., Murphy, G., Docherty, A. J. P., & Bode, W. (1996) *FEBS Lett. 378*, 126−130.

Goldberg, G. I., Wilhelm, S. M., Kronberger, A., Bauer, E. A., Grant, A. G., & Eisen, A. Z. (1986) *J. Biol. Chem. 261*, 6600−6605.

Gross, J., & Lapiere, C. M. (1962) *Proc. Natl. Acad. Sci. U.S.A. 48*, 1014−1022.

Henriet, P., Rousseau, G. G., & Eckhout, Y. (1992) *FEBS Lett. 310*, 175−178.

Hirose, T., Patterson, C., Pourmotabbed, T., Mainardi, C. L., & Hasty, K. A. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2569−2573.

Hoshi, M., Moriya, T., Aoyagi, T., & Umezawa, H. (1979) *J. Exp. Zool. 209*, 129−134.

Hwang, S.-P. L., Partin, J. S., & Lennarz, W. L. (1994) *Development (Cambridge, U.K.) 120*, 559−568.

Ikegami, S., Kobayashi, H., Myotoishi, Y., Ohta, S., & Kato, K. H. (1994) *J. Biol. Chem. 269*, 23262−23267.

Ishida, J. (1936) *Annot. Zool. Jpn. 15* (4), 453−459.

Ishida, J. (1985) *Zool. Sci. 2*, 1−10.

Jongeneel, C. V., Bouvier, J., & Bairoch, A. (1989) *FEBS Lett. 242*, 211−214.

Kinoshita, T., Fukuzawa, H., Shimada, T., Saito, T., & Matsuda, Y. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4693−4697.

Kozak, M. (1981) *Nucleic Acids Res. 9*, 5233−5262.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105−132.

Laemmli, U. K. (1970) *Nature 257*, 680−685.

Lepage, T., & Gache, C. (1989) *J. Biol. Chem. 264*, 4787−4796.

Lepage, T., & Gache, C. (1990) *EMBO J. 9*, 3003−3012.

Lepage, T., Ghiglione, C., & Gache, C. (1992a) *Development (Cambridge, U.K.) 114*, 147−164.

Lepage, T., Sardet, C., & Gache, C. (1992b) *Dev. Biol. 150*, 23−32.

Li, J., Brick, P., O'Hare, M. C., Skarzynski, T., Lloyd, L. F., Curry, V. A., Clark, I. M., Bigg, H. F., Hazleman, B. L., Cawston, T. M., & Blow, D. M. (1995) *Structure 3*, 541−549.

Libson, A. M., Gittis, A. G., Collier, I. E., Marmer, B. L., Goldberg, G. I., & Lattman, E. E. (1995) *Nat. Struct. Biol. 2*, 938−942.

Miyazaki, K., Uchiyama, K., Imokawa, Y., & Yoshizato, K. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 6819−6824.

Murphy, G., Allan, J. A., Willenbrock, F., Cockett, M. I., O'Connell, J. P., & Docherty, J. P. (1992) *J. Biol. Chem. 267*, 9612−9618.

Nomura, K., & Suzuki, N. (1993) *FEBS Lett. 321*, 84−88.

Nomura, K., & Suzuki, N. (1995) *Arch. Biochem. Biophys. 319*, 525−534.

Nomura, K., Matsumoto, M., & Suzuki, N. (1990) *Biochemistry 29*, 4525−4534.

Nomura, K., Tanaka, H., Kikkawa, Y., Yamaguchi, M., & Suzuki, N. (1991) *Biochemistry 30*, 6115−6123.

O'Connell, J. P., Willenbrock, F., Docherty, A. J., Eaton, D., & Murphy, G. (1994) *J. Biol. Chem. 269*, 14967−14973.

O'Hare, M. C., Curry, V. A., Mitchell, R. E., & Cawston, T. E. (1995) *Biochem. Biophys. Res. Commun. 216*, 329−337 (1995).

Patterson, D., Hayes, W. P., & Shi Y.-B. (1995) *Dev. Biol. 167*, 252−262.

Post, L. L., Schuel, R., & Schuel, H. (1988) *Biochem. Cell Biol. 66*, 1200−1209.

Quigley, J. P., Braithwaite, R. S., & Armstrong, P. B. (1993) *Differentiation (Berlin) 54*, 19−23.

Richards, C. D., Rafferty, J. A., Reynolds, J. J., & Saklatvala, J. (1991) *Matrix 11*, 161−167.

Roe, J. L., & Lennarz, W. J. (1985) *J. Cell Biol. 101*, 470a.

Roe, J. L., & Lennarz, W. J. (1990) *J. Biol. Chem. 265*, 8704−8711.

Sanchez-Lopez, R., Nicholson, R., Gesnel, M.-C., Matrisian, L. M., & Breathnach, R. (1988) *J. Biol. Chem. 263*, 11892−11899.

Sanchez-Lopez, R., Alexander, C. M., Behrendtsen, O., Breathnach, R., & Werb, Z. (1993) *J. Biol.Chem. 268*, 7238−7247.

Sang, Q. A., & Douglas, D. A. (1996) *J. Protein Chem. 15*, 137−160.

Sanger, R., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.

Satoh, T., Kobayashi, K., Yamashita, S., Kikuchi, M., Sendai, Y., & Hoshi, H. (1994) *Biol. Reprod. 50*, 835−844 (1994).

Saus, J., Quinones, S., Otani, Y., Nagase, H., Harris Jr., E. D., & Kurkinen, M. (1988) *J. Biol. Chem. 263*, 6742−6745.
Schnierer, S., Kleine T., Gote, T., Hilleman, A., Knauper, V., & Tschesche, H. (1993) *Biochem. Biophys. Res. Commun. 191*, 319−326.
Showman, R. M., & Whiteley, A. H. (1980) *Dev., Growth Differ. 22*, 305−314.
Springman, E. B., Angleton, E. L., Birkedal-Hansen, H., & Van Wart, H. E. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 364−368.
Spurlino, J. C., Smallwood, A. M., Carlton, D. D., Banks, T. M., Vavra, K. J., Johnson, J. S., Cook, E. R., Falvo, J., Wahl, R. C., Pulvino, T. A., Wendoloski, J. J., & Smith, D. L. (1994) *Proteins: Struct., Funct., Genet. 19*, 8−109.
Stams, T., Spurlino, J. C., Smith, D. L., Wahl, R. C., Ho, T. F., Qoronfleh, M. W., Banks, T. M., & Rubin, B. (1994) *Nat. Struct. Biol. 1*, 119−123.
Takeuchi, K., Yokosawa, H., & Hoshi, M. (1979) *Eur. J. Biochem. 100*, 257−265.
Taylor, K. B., Windsor, L. J., Caterina, N. C. M., Bodden, M. K., & Engler, J. A. (1996) *J. Biol. Chem. 271*, 23938−23945.
von Heijne, G. (1986) *Nucleic Acids Res. 14*, 4683−4690.
Yamagami, K. (1996) *Zool. Sci. 13*, 331−340.
Yanish-Perron, C., Vieira, J., & Messing, J. (1985) *Gene 33*, 103−119.
Yasumasu, S., Yamada, K., Akasaka, K., Mitsunaga, K., Iuchi, I., Shimada, H., & Yamagami, K. (1992) *Dev. Biol. 153*, 250−248.
Whitham, S. E., Murphy, G., Angel, P., Rahmsdorf, H. J., Smith, B. J., Lyons, A., Harris, T. J. R., Reynolds, J. J., Harlot, P., & Docherty, A. J. P. (1986) *Biochem. J. 240*, 913−916.
Woessner, J. F., Jr. (1962) *Biochem. J. 83*, 304−314.
Zhu, L., Hope, T. J., Hall, J., Davies, A., Stern, W., Muller-Eberhard, U., Stern, R., & Parslow, T. G. (1994) *J. Biol. Chem. 264*, 32092−32097.

BI9629790