HOKKAIDO UNIVERSITY

Title	Specificity of the medaka enteropeptidase serine protease and its usefulness as a biotechnological tool for fusion-protein cleavage
Author(s)	Ogiwara, Katsueki; Takahashi, Takayuki
Citation	Proceedings of the National Academy of Sciencesof the United States of America, 104(17), 7021-7026 https://doi.org/10.1073/pnas.0610447104
Issue Date	2007-04-24
Doc URL	https://hdl.handle.net/2115/30134
Type	journal article
File Information	PNAS104-17.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

Classification: Biological Sciences (Biochemistry)

Specificity of the medaka enteropeptidase serine protease and its usefulness as a biotechnological tool for fusion-protein cleavage

Katsueki Ogiwara and Takayuki Takahashi†

Laboratory of Molecular and Cellular Interactions, Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan

†*Correspondence: Takayuki Takahashi, Ph.D., Laboratory of Molecular and Cellular Interactions, Faculty of Advanced Life Science, Hokkaido University, Sapporo 060-0810, Japan*

Tel.: 81-11-706-2748; Fax: 81-11-706-4851; e-mail: ttakaha@sci.hokudai.ac.jp

Manuscript information: 24 text pages, 4 figures, and 1 table

Word and character counts: Abstract: 164 words; Paper: 39,025 characters

Abbreviations: EP, enteropeptidase; βNA, β-naphthyamide; MCA, 4-methylcoumaryl-7-amide; Boc, *t*-butyloxycarbonyl; OBzl, benzyloxycarbonyl(butanoil); Z, benzyloxycarbonyl.

Data deposition: The sequences reported in this paper have been deposited in the DDBJ database (accession nos. AB272104, AB272105, and AB272106)

We cloned two distinct cDNAs for enteropeptidase (EP) from the intestine of the medaka, *Oryzias latipes*, which is a small freshwater teleost. The mRNAs code for EP-1 (1036 residues) and EP-2 (1043 residues), both of which have a unique, conserved domain structure of the N-terminal heavy chain and C-terminal catalytic serine protease light chain. When compared with mammalian EP serine proteases, the medaka enzyme exhibited extremely low amidolytic activity for small synthetic peptide substrates. Twelve mutated forms of the medaka EP protease were produced by site-directed mutagenesis. Among them, one mutant protease, E173A, was found to have considerably reduced nonspecific hydrolytic activities both for synthetic and protein substrates without serious reduction of its Asp-Asp-Asp-Asp-Lys (D_4K)-cleavage activity. For the cleavage of fusion proteins containing a D_4K-cleavage site, the medaka EP proteases were shown to have advantages over their mammalian counterparts. Based on our present data, we propose that the E173A mutant is the most appropriate protease to specifically cleave proteins containing the D_4K cleavage sequence.

Keywords: enteropeptidase; cleavage specificity; biotechnology tool; medaka fish

Enteropeptidase (EP) (enterokinase, EC 3.4.21.9) is a two-chain, membrane-bound serine protease of the duodenal mucosa that converts trypsinogen to active trypsin (1). Trypsin thus produced then cleaves and activates other zymogens in pancreatic secretions, including chymotrypsinogen, proelastase, procarboxypeptidases, and some prolipases. Therefore, EP has been recognized to play a critical role in regulating protein digestion within the lumen of the gut. The biological importance of EP is demonstrated by the malabsorption and malnutrition of patients with primary EP deficiency, a genetic disorder resulting in no or little EP activity in the duodenum (2).

EP has been intensively studied in the last decade. To date, EP has been molecularly cloned from several sources, including cattle (3, 4), humans (5), pigs (6), rats (7), and mice (8). These studies provided much information on the structural details and organization of EP, and opened a path to further investigation of the molecular properties of this protease. For example, it was demonstrated that the N-terminal heavy chain is required for efficient activation of trypsinogen by the serine protease domain of the C-terminal light chain (9, 10). In addition, a recent study by Lu *et al.* (11) established the tertiary structure of the bovine EP catalytic domain. The study demonstrated the involvement of Lys99, which is situated in a unique exosite on the enzyme surface, in the specific cleavage of trypsinogen and similar peptidyl substrates. More recently, a mucin-like domain found in the heavy chain of EP has been shown to be a possible targeting signal for apical sorting of the protein (12).

EP is also of biotechnological interest because of the unique substrate specificity of the serine protease domain. The high degree of specificity exhibited by EP makes it a suitable reagent for cleaving bacterially produced proteins. Indeed, the catalytic domain

of bovine EP is now widely used as a valuable tool for this purpose (13).

It is generally believed that EP (or enteroptidase-like enzyme) is present in all vertebrates. This belief comes from the finding that in almost all vertebrate species a short peptide sequence of Asp-Asp-Asp-Asp-Lys (D_4K) is found in the presumed activation site of trypsinogens (14). However, no information on EP in vertebrates other than mammals has been made available to date. Here, we report on the isolation of cDNAs encoding EP of the medaka (*Oryzias latipes*), a freshwater teleost, and its expression in several tissues. The present study also describes some enzymatic properties of the catalytic serine protease domain. Surprisingly, the protease domain of medaka EP exhibited very limited amidolytic activity for any of the synthetic peptide substrates tested, indicating that the medaka protease itself is much more highly specific for the D_4K sequence than its mammalian counterparts. We further generated various mutant proteases of medaka EP by site-directed mutagenesis. Some of the mutated proteases exhibited cleavage specificity that was stricter than that of the wild-type enzyme, and may prove to be more effective tools for recombinant protein technology.

Results

cDNA Cloning and Expression of Medaka EP.

The present approach gave two distinct medaka *EP* cDNA clones, *EP-1* (3997-bp, AB272104) and *EP-2* (4036-bp, AB272105). The full-length *EP-1* cDNA clone contained an ORF that codes a protein of 1036 amino acids, while the *EP-2* clone codes a protein of 1043 amino acid residues (Fig. 5, supporting information.) The deduced amino acid sequence of the medaka EP was homologous to those of its mammalian

counterparts. As in mammalian EPs, unique domain structures were found in the N-terminal heavy chain of the fish protein (Fig. 1A). However, the extent of sequence identity between the medaka and mammalian EPs varies considerably from one domain to another: the identity is 21% in the mucin-like domain, 45% in LDLR domain 1, 41% in Cl r/s domain 1, 49% in the MAM domain, 57% in Cl r/s domain 2, 47% in LDLR domain 2, and 23% in the MSCR domain. The C-terminal serine protease domain of the fish EP exhibited 53% identity to mammalian EP serine proteases (Fig. 1B).

In RT-PCR analyses using primer sets specific for the two medaka EPs, we observed that the band intensities of amplified products were greater for EP-1 than for EP-2 at every PCR cycle (Fig. 5B, supporting information). We also performed RT-PCR using primers common to the two *EP* transcripts. Amplified products (1235 bp for *EP-1* and 1246 bp for *EP-2*) were gel-purified and cloned into pBluescript II KS(+), and the recombinant plasmids were transformed into *E. coli*, strain JM109. Forty-four clones were randomly picked for nucleotide sequencing; 26 clones were for EP-1 and 18 clones for EP-2. The results indicated that EP-1 is the dominant EP species expressed in the medaka intestine. The results of a Southern blot analysis support the presence of at least two distinct copies of the *EP* gene in the medaka (Fig. 5C, supporting information).

Northern blot analysis of EP using various fish tissues revealed that the intestine expresses an approximately 4-kb transcript, and this size is consistent with that of the full-length cDNA (Fig. 1C). Surprisingly, very strong signals at 1.3 kb and 1.5 kb were detected in the ovary and testis. Further analyses indicated that these were transcripts of 1090 bp (corresponding to 2908-3997 in AB272104) and 1241 bp (corresponding to

2757-3997 in AB272104). Both transcripts were found not to code for any functional protein. *In situ* hybridization analysis indicated that *EP* mRNA was localized in the cytoplasm of small, growing follicles in the ovary of mature female medaka (Fig. 6, supporting information). Neither Western blotting nor immunohistochemical analysis using antibodies specific for the medaka EP protease detected corresponding proteins. Therefore, no further study was conducted on the ovarian *EP* transcripts. We tentatively speculate that *EP* transcripts expressed in the fish ovary might play a role as non-coding RNAs in the oocytes of growing follicles (15).

In RT-PCR using primers common to the two species of medaka EP, transcripts were detected in the intestinal segments proximal to the stomach (Fig. 1D). *In situ* hybridization analysis localized EP expression to the intestinal epithelium (Fig. 1E). Western blot analysis (under reducing conditions) of the extract of medaka intestine, but not ovary and testis extract, using specific anti-EP antibodies against the catalytic domain detected a 36-kDa immunoreactive band (Fig. 1F, left). A polypeptide band of the same molecular mass was detected in both cytosolic and membrane fractions of the intestine (Fig. 1F, right). Western blotting of the intestine extract under nonreducing conditions gave no clear band (data not shown). Immunohistochemical analysis using the antibody demonstrated the epithelial localization of EP in the intestine (Fig. 1G). The extract of medaka intestine exhibited enzyme activity for the synthetic EP substrate Gly-Asp-Asp-Asp-Asp-Lys-β-naphthylamide (GD_4K-βNA). Using this activity as a marker, the apparent molecular mass of intact EP was estimated by gel filtration to be 440 kDa (Fig. 7A, supporting information). The fraction having GD_4K-βNA-hydrolyzing activity showed a 36-kDa polypeptide in Western blotting

under reducing conditions (Fig. 7B, left). Again, the same fraction did not show any clear band with the current antibody when analyzed under nonreducing conditions (Fig. 7B, right). Taken together, these results indicate that the fish intestine contains active, membrane-bound EP. Part of the molecule exists in the intestine in a soluble form that is probably detached from the epithelial cell membrane. The adult medaka intestinal epithelium contains most of the cell types (enterocytes, goblet cells, and enteroendocrine cells) observed in the small intestine of other vertebrates, but lacks crypts containing Paneth cells and intestinal stem cells (16). Our current data suggest that medaka EP is localized in the enterocytes in the proximal intestinal epithelium.

The Medaka EP Catalytic Serine Protease Domain Has Distinct Properties as Compared with its Mammalian Counterparts.

The active 32-kDa C-terminal serine protease domains of both EP-1 and EP-2 were prepared in order to characterize their enzymatic properties (Fig. 8A, supporting information). Both enzymes showed maximal activities for GD_4K-βNA at pH 8, but EP-1 was approximately three times more active than EP-2 (Fig. 8B). To examine the effects of EP-1 and EP-2 on the physiological substrate trypsinogen, a 866-bp medaka trypsinogen cDNA (AB272106), which codes for a protein of 242 amino acids (Fig. 9, supporting information), was obtained from the intestine. Using the sequence, a recombinant fusion protein of medaka trypsinogen was prepared. The trypsinogen was converted to active trypsin by EP-1 faster than by EP-2 (Fig. 8C). The behavior of the two proteases in response to various protease inhibitors was undistinguishable (Table 2, supporting information). From these results, together with the finding that EP-1 is the

dominantly expressed form in the intestine, we selected EP-1 for the following experiments.

The serine protease domain of medaka EP-1 cleaved GD_4K-βNA at a rate comparable to those of the porcine and bovine enzymes (Fig. 2A). Since, in addition to D_4K-cleavage activity, mammalian EP protease is known to have amidolytic activity for some MCA-containing peptide substrates (11), enzyme activities of medaka EP-1 protease were determined with these substrates. Surprisingly, the activities of the medaka protease for Boc-Glu(OBzl)-Ala-Arg-MCA, Z-Phe-Arg-MCA, and Pro-Phe-Arg-MCA were much lower than those of the EP proteases of mammalian origin (Fig. 2B). The kinetic parameters of the proteases for these substrates were determined (Table 1). Generally, the k_{cat}/K_m values of the medaka enzyme were 1-2 orders of magnitude smaller than those of the mammalian proteases for all MCA-containing synthetic substrates.

The proteolytic activity of the medaka protease was examined using gelatin (Fig. 2C), fibronectin (Fig. 2D), and laminin (Fig. 2E). For comparison, the mammalian proteases were also tested under the same conditions. Little or no hydrolysis of the proteins was observed with the fish enzyme, whereas these substrates were detectably hydrolyzed by the mammalian proteases. Finally, the fusion protein (48 kDa) containing an EP-cleavage site (available from Novagen) was tested with various EP proteases. Clearly, the medaka protease specifically cleaved the fusion protein to generate two polypeptides having expected molecular masses of 16 and 32 kDa (Fig. 2F). In contrast, the mammalian enzymes not only produced the two expected polypeptides but also further degraded the products, presumably due to extensive nonspecific proteolytic

activities. These results demonstrate that the medaka EP-1 protease intrinsically has much more strict cleavage specificity than its mammalian counterparts.

Active recombinant medaka EP-1 was stable at -20°C and 4°C; the initial enzyme activity was retained at both temperatures for at least six months with no detectable change in the electrophoretic pattern (data not shown). When medaka EP-1 alone was kept at 37°C at neutral pH, about 30% loss of enzyme activity was observed after 4 days of incubation (Fig. 10, supporting information). In a parallel experiment using bovine EP protease, a sharp decline in enzyme activity was seen after just a few hours of incubation at 37°C.

Some Medaka Mutant EP Proteases Exhibit Cleavage Specificity Stricter than That of the Wild-type Enzyme.

We further examined whether site-directed mutagenesis of the medaka EP-1 catalytic domain could produce a mutant protease with superior characteristics in terms of biotechnological interest. Because strict specificity of the medaka protease compared to its mammalian counterparts might be due to substitutions of residues conserved in the mammalian enzymes, we focused primarily on five amino acid positions where the residues differed from those of mammalian EP proteases. These five residues are all distant from the medaka EP-1 protease active site in a structural model (Fig. 3) based on a model of the bovine EP serine protease domain, which was crystallized in complex with an inhibitor analog of the activation peptide, Val-Asp-Asp-Asp-Asp-Lys-chloromethane (11). The residues were mutated to those conserved in the mammalian protease (K63R, T105E, F144S, E173K, and P193E). In

addition, the residues were replaced by those with a small side chain (K63A, T105A, F144A, E173A, and P193A) or those with a different charge (K63E and T105E) to examine the effects of side chains (Fig. 1B).

All 12 mutants could convert the recombinant medaka trypsinogen to its active enzyme (data not shown). K63R, T105E, F144S, E173K, and P193E showed reduced enzyme activity for GD_4K-βNA (Fig. 4A). The same held true for all the other mutants except F144A, which hydrolyzed both GD_4K-βNA and the three MCA-containing substrates at an elevated rate compared to the wild-type enzyme (Fig. 11, supporting information). As enzymes all exhibiting significantly reduced activity for the MCA substrates, but retaining GD_4K-βNA-hydrolyzing activity to a reasonable extent, four mutants (K63R, T105E, E173A, and P193E) were chosen for further characterization. K63R converted recombinant medaka trypsinogen to trypsin as fast as the wild-type enzyme, whereas the other mutants activated trypsinogen at a reduced rate (Fig. 12, supporting information). The mutant proteases were characterized by kinetic studies. Interestingly, E173A retained a k_{cat}/K_m value comparable to the wild-type enzyme for GD_4K-βNA. However, the k_{cat}/K_m values for the MCA-containing substrates were lower (Table 1).

Compared with mammalian EP proteases, the mutant proteases showed lower nonspecific proteolytic activity for human fibrinogen (Fig. 13A, supporting information). Nonspecific hydrolysis of human high-molecular-weight (HMW) kininogen occurred at a much slower rate with the mutant proteases than with the mammalian proteases, although some mutations slightly increased nonspecific cleavage (Fig. 13B, supporting information). Neither human fibronectin nor laminin was

hydrolyzed by the mutants (data not shown).

These results indicate that the substitution of glutamic acid by alanine at 173 caused a significant reduction in unwanted, nonspecific enzyme activities for both the synthetic and protein substrates without seriously deteriorating the mutant's cleavage specificity for the GD_4K sequence.

Medaka EP Protease Can be Used As an Efficient Biotechnological Tool for Cleavage of Recombinant Fusion Proteins.

To gain insight into the biotechnological usefulness of medaka EP serine protease, its effect on various fusion proteins containing a D_4K-cleavage site was examined. Medaka gelatinase A (17) was synthesized as a fusion protein containing a His tag and D_4K sequence at the N terminus in the *E. coli* expression system using the pET30 expression vector. Both the wild-type and mutant proteases converted the 60-kDa fusion protein to a 55-kDa protein (Fig. 4B). At the same substrate/protease ratio, mammalian EP serine proteases extensively digested the fusion protein.

Next, we synthesized a 35.5-kDa protein of human kallikrein 8 (hK8) in the same *E. coli* expression system. Digestion with all EP proteases commonly generated 31.5-kDa active hK8 by cleaving the D_4K sequence (Fig. 4C, upper). The EP protease-treated samples were directly assayed for hK8 activity using Pro-Phe-Arg-MCA, a good synthetic peptide substrate for hK8 (18). All the samples treated with the medaka or mammalian EP proteases exhibited Pro-Phe-Arg-MCA-hydrolyzing activity (Fig. 4C, lower). As expected, none of the medaka EP proteases (wild-type EP-1, K63R, T105E, E173A, or E193A) showed any significant enzyme activity. In contrast, considerable

enzyme activities were detected with porcine and bovine (Neb) EP proteases. Fusion protein that had been digested with the bovine (Nvg) protease had very low activity, presumably due to inactivation of the EP protease itself during incubation. These results demonstrate that the medaka EP protease used for cleaving the fusion protein has no serious effect on hK8 activity.

Finally, a human single-chain tissue-type plasminogen activator (tPA) fusion protein containing an 11-residue sequence of a His-tag/EP-susceptible site at the N terminus of mature tPA was generated by CHO cells and used as a substrate for medaka and mammalian EP proteases. The protein samples treated with the medaka wild-type or mutant EP proteases, but not with mammalian ones, showed two polypeptides (58 and 61 kDa) detectable with anti-human tPA antibodies (Fig. 4D, upper). The specific antibody for the His-tag sequence did not recognize the polypeptides (Fig. 4D, lower). These results indicate that the medaka proteases properly cleave the fusion protein at the EP-cleavage site to produce single-chain tPA, and also demonstrate that mammalian EP proteases extensively degrade the single-chain tPA fusion protein.

Taken together, our results suggest that the medaka proteases are more effective than their mammalian counterparts as fusion-protein cleavage enzymes for the preparation of desired recombinant proteins containing the D_4K cleavage sequence.

DISCUSSION

Almost all trypsinogen sequences known to date contain a highly conserved tetra-aspartate sequence preceding the lysine-isoleucine scissile peptide bond. Although EP is widely considered to play a role in trypsinogen activation in all vertebrate species,

there has been no report on EP from non-mammalian species. The present study is thus the first to report on the molecular and biochemical characterizations of EP from a lower vertebrate, the medaka.

Medaka EP mRNA exists in two distinct forms, *EP-1* and *EP-2*, in the intestine. A comparison of the entire amino acid sequences of EP-1 (1036 residues) and EP-2 (1043 residues) revealed a difference of only 22 amino acids, including an insertion of 7 residues in EP-2. Our data suggest that *EP-1* and *EP-2* mRNA are expressed at a ratio of approximately 3:2 in the intestine. It remains to be determined whether they are indeed translated at this ratio. Moreover, it is not known at present whether they have a discrete role *in vivo*.

To our surprise, medaka EP serine protease *per se* showed very low activity for three synthetic peptide substrates, Boc-Glu(OBzl)-Ala-Arg-MCA, Z-Phe-Arg-MCA, and Pro-Phe-Arg-MCA, that are rapidly hydrolyzed by the mammalian enzymes. A comparison of the kinetic parameters between medaka and mammalian EP proteases revealed that the remarkably low activities of the fish enzyme for these substrates were due to the reduced k_{cat} value. This finding led us to examine whether site-directed mutagenesis could produce a mutant enzyme with further reduced amidolytic activity for the synthetic substrates.

Twelve mutant medaka EP proteases were generated by mutagenesis at positions Lys63 (corresponding to Arg60f in bovine EP protease (11)), Thr105 (bovine Arg98), Phe144 (bovine Ser137), Glu173 (bovine Lys167), and Pro193 (bovine Glu185). These positions all occur on the surface of the medaka EP serine protease model (Fig. 3), and, with the exception of medaka EP protease position #105 (bovine #98), the residues are

located a considerable distance from the enzyme active site. A previous crystalline structural study (11) and other biochemical studies (10, 11) on mammalian EP protease indicate that the minimal recognition sequence for EP consists of Lys at P1 and Asp at P2 of the D_4K sequence. Therefore, the five residues we targeted are probably not essential for substrate recognition. As is clear from the medaka EP protease model, the structure around the S1 site highly resembles that of bovine EP protease, indicating that the catalytic mechanism and contacts with Lys at P1 are conserved with mammalian EP proteases. The interaction of Asp at P2 with the side chain of Lys106 (bovine Lys99) at S2 is also conserved, since GD_4K-βNA is hydrolyzed at a comparable rate by the medaka and bovine enzymes. However, the microenvironment around the S2 site is presumed to be different between medaka and mammalian EP proteases (Fig. 3). A slight difference in the structure around the S2 site might underlie the detectable difference in catalytic activities for MCA-containing peptide substrates. We indeed observed that the T105E mutant displayed significantly lower catalytic activity for GD_4K-βNA. A plausible explanation for this activity decrease is that replacement of Thr by Glu at position 105, which is next to Lys106 at the S2 site, could affect interaction of Asp at P2 with Lys106 (bovine Lys99) at S2.

In order to examine the effects of side chains of the target residues on catalytic activity, we replaced these by Ala. Mutagenesis had different effects on each of the enzyme activities. One of the mutants, E173A, was interesting in that it showed significantly lower activities than the wild-type enzyme toward all synthetic substrates tested. In addition, this mutant enzyme retained a low nonspecific proteolytic activity for protein substrates (HMW kininogen and fibrinogen), with no serious reduction in

D_4K-cleaving activity for fusion proteins (gelatinase A, hK8, and tPA). Since none of the targeted residues is essential for the EP protease catalytic activity, changes in the activity by mutagenesis are presumably due to conformational change brought about by substitution of residues at regions distant from the active site of the enzyme.

We note here that, as demonstrated with bovine EP (9), activity of the enzyme for D_4K-containing substrates is dependent on the heavy chain. The presence of the heavy chain had almost no effect on recognition of the small peptide substrate GD_4K-βNA, while this chain had profound effects on recognition of the macromolecular substrate trypsinogen (9). Furthermore, bovine EP is able to hydrolyze specifically several biologically active peptides *in vitro* (19), and substrate length greatly affects the catalytic efficiency of this hydrolysis. Activity dependency on the heavy chain and peptide substrate length remain to be determined for medaka EP.

Since EP serine proteases preferably hydrolyze the D_4K sequence, this motif has been intentionally introduced for the specific cleavage of fusion proteins. Bovine EP serine protease is now widely used for this purpose. The current system utilizing the bovine enzyme works reasonably well in many cases, but requires handling with great care. We often encounter the following difficulties. 1) Bovine EP protease primarily cleaves at the EP-cleavage site of recombinant fusion proteins; however, other peptide bonds are also hydrolyzed to a considerable degree by nonspecific proteolytic activity. This results in a low yield of the protein in question. 2) For preparing active recombinant proteases, the bovine EP protease employed for cleavage of the inactive fusion protein could be an obstacle. This is particularly serious when the proteases to be examined are ones with very low activity for synthetic and protein substrates.

Significant nonspecific activities of bovine EP protease often make it difficult to determine whether the target recombinant proteases have been successfully activated.

To address this latter difficulty, we examined the enzyme activities of medaka EP protease for synthetic MCA-containing substrates, which are known to be nonspecifically hydrolyzed by the mammalian counterparts, including bovine EP (11), although the synthetic substrates themselves are not physiologically relevant. As demonstrated in the present study, the serine protease domain of medaka EP has a stricter specificity for almost all of the substrates tested compared to the mammalian EP proteases. The disadvantages mentioned above associated with bovine EP could be overcome by using the medaka protease. Medaka wild-type EP protease would be adequate for the preparation of recombinant non-proteolytic enzymes. However, in view of the efficient cleavage at the D_4K site and the minimum nonspecific hydrolysis at peptide and amide bonds, we recommend use of the E173A mutant enzyme. Medaka wild-type EP protease and its mutant can be prepared in large quantity in the *E. coli* expression system. We believe that, using the medaka EP serine proteases as cleavage enzymes for fusion proteins containing the D_4K cleavage sequence, the desired recombinant proteins can be easily and effectively produced.

Materials and Methods

Preparation of the Recombinant EP Serine Protease Domain.

Two distinct medaka *EP* cDNA clones, *EP-1* (3997-bp, AB272104) and *EP-2* (4036-bp, AB272105), were obtained as described in Supporting Information.

A DNA fragment including the coding sequence for the medaka EP-1 or EP-2

catalytic domain was amplified by PCR using a pBluescript II plasmid containing cDNA of the catalytic domain as the template. The upper and lower primers were 5’-CGCGGATCCCAAGCTGGTGTGGTGGGTGG-3’ and 5’-CCCAAGCTTTCAGTCTAGATCTGAGAA-3’, respectively, which have *Bam*HI and *Hind*III sites at the respective 5’ termini. The product was ligated into the cloning site of a pET30a expression vector (Novagen, Madison, WI). Expression of the recombinant medaka EP catalytic domain in the *Escherichia coli* expression system was carried out as described previously (17). The medaka EP catalytic domain was produced as a fusion protein with an extra amino acid sequence of 50 residues at its N terminus; the vector-derived N-terminal stretch contained a His tag and an S-protein sequence. Harvested cells were lysed and the insoluble materials were dissolved in a solubilization buffer containing 6 M urea, 50 mM Tris·HCl (pH 7.6), and 0.5 M NaCl. Solubilized proteins were subjected to affinity chromatography on Ni^{2+}-Sepharose (GE Healthcare Biosciences, Piscataway, NJ), and eluted with the same buffer containing 50 mM histidine. Eluted recombinant proteins were renatured by dialysis against 50 mM Tris·HCl (pH 8.0). The fusion protein was then incubated in 50 mM Tris·HCl (pH 8.0) containing 0.5 M NaCl with trypsin immobilized on Sepharose 4B at room temperature for 1 h. The immobilized trypsin was then removed by filtration. The resulting sample, which contained not only active EP protease but also inactive enzyme protein, was fractionated on a column of Resource Q in AKTA Purifier (GE Healthcare Biosciences, Uppsala, Sweden) to remove inactive enzyme. A trace amount of trypsin often contained in the sample thus prepared was removed by passage through an aprotinin-Sepharose 4B column (Sigma, St. Louis, MO).

Active recombinant enzyme of the porcine EP serine protease domain (Ile800 to His1034) (6) was prepared basically according to the method described above. Bovine EP serine protease was obtained from Novagen and New England Biolabs (NEB, Schwalbach, Germany).

Site-directed Mutagenesis.

Site-directed mutagenesis of medaka EP-1 was carried out to produce various mutant proteases. For each mutant, two PCR products were first amplified with medaka EP-1 cDNA as a template using the following two primer combinations: the “upper” primer described above with the respective antisense primer (Table 3), and the “lower” primer described above with the sense primer (Table 3). Using a mixture of these amplified DNAs as the template, a second PCR was performed with the “upper” and “lower” primer. The PCR products were digested with *Bam*HI and *Hind*III, gel-purified, and ligated into the pET30a expression vector. All mutants were confirmed by DNA sequencing. The subsequent procedures for preparation of mutant proteases were the same as for the wild-type protein described above.

Hydrolysis of Proteins by the EP Catalytic Serine Protease Domain.

Human plasma fibronectin (Chemicon, Temecula, CA), human fibrinogen (Merk Biosciences, Tokyo, Japan), human high-molecular-weight (HMW) kininogen (Calbiochem, La Jolla, CA), mouse laminin (Biomedical Technologies Inc., Stoughton, MA), D_4K cleavage site-containing control protein (Novagen), medaka gelatinase A (17), trypsinogen (this study), human kallikrein 8 (hK8) (18), and human tissue-type

plasminogen activator (tPA) were incubated at 37°C in 20 mM Tris· HCl buffer (pH 7.4) containing 50 mM NaCl and 2 mM $CaCl_2$ with various EP serine proteases at ratios (w/w) ranging from 20:1 to 100:1. After incubation, samples were subjected to SDS-PAGE followed by Coomassie Brilliant Blue (CBB) staining. Gelatin zymography was conducted as described previously (17), except that gels were incubated in 20 mM Tris· HCl buffer (pH 7.4) containing 50 mM NaCl and 2 mM $CaCl_2$.

Computer Modeling of the Serine Protease Domain of Medaka EP.

Three-dimensional structure prediction was carried out by using Swiss Model. The serine protease domain of medaka EP was modeled in the 'optimize' mode after manual alignment of the sequence with bovine EP.

The methods used for the cloning of medaka EP-1 and EP-2 and trypsinogen; for preparations of recombinant proteins other than EP; and for RT-PCR, Northern blotting, Southern blotting, *in situ* hybridization, antibody production, Western blotting, immunohistochemistry, gel-filtration chromatography; as well as other enzymic methods are described in the Supporting Methods.

Acknowledgments

We thank Dr. M. Yao and I. Tanaka (Hokkaido University) for valuable comments on the structure of enteropeptidase. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (17370021).

References

1. Light, A. & Janska, H. (1989) *Trends Biochem. Sci.* **14**, 110-112.
2. Grishan, F. K., Lee, P. C., Lebenthal, E., Johnson, P., Bradley, C. A. & Greene, H. L. (1983) *Gastroenterology* **85**, 727-731.
3. LaVallie, E. R., Rehemtulla, A., Racie, L. A., DiBlasio, E. A., Ferenz, C., Grant, K. L., Light, A. & McCoy, J. M. (1993) *J. Biol. Chem.* **268**, 23311-13317.
4. Kitamoto, Y., Yuan, X., Wu, Q., McCourt, D. W. & Sadler, J. E. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 7588-7592.
5. Kitamoto, Y., Veile, R. A., Donis-Keller, H. & Sadler, J. E. (1995) *Biochemistry* **34**, 4562-4568.
6. Matsushima, M., Ichinose, M., Yahagi, N., Kakei, K., Tsukada, S., Miki, K., Kurokawa, K., Tashiro, K., Shiokawa, K., Shinomiya, K., *et al.* (1994) *J. Biol. Chem.* **269**, 19976-19982.
7. Yahagi, N., Ichinose, M., Matsushima, M., Matsubara, Y., Miki, K., Kurokawa, K., Fukamachi, H., Tashiro, K., Shiokawa, K., Kageyama, T., *et al.* (1996) *Biochem. Biophys. Res. Commun.* **219**, 806-812.
8. Yuan, X., Zheng, X., Lu, D., Rubin, D. C., Pung, C. Y. & Sadler, J. E. (1998) *Am. J. Physiol.* **274**, 342-349.
9. Lu, D., Yuan, X., Zheng. X. & Sadler, J. E. (1997) *J. Biol. Chem.* **272,** 31293-31300.
10. Mikhailova, A. G. & Rumsh, L. D. (1999) *FEBS Lett.* **442**, 226-230.
11. Lu, D., Fütterer, K., Korolev, S., Xinglong, Z., Tan, K., Waksman, G. & Sadler, J. E.

(1999) *J. Mol. Biol.* **292**, 361-373.

12. Zheng, X. & Sadler, J. E. (2002) *J. Biol. Chem.* **277**, 6858-6863.
13. Collins-Racie, L. A., McColgan, J. M., Grant, K. L., DiBlasio-Smith, E. A., McCoy, J. M. & LaVallie, E. R. (1995) *Biotechnology* **13**, 982-987.
14. Bricteux-Gregoire, S., Schyns, R. & Florkin, M. (1972) *Comp. Biochem. Physiol.* **42B**, 23-39.
15. Costa, F. F. (2005) *Gene* **357**, 83-94.
16. Rombout, J. H., Stroband, H. W. & Taverne-Thiele, J. J. (1984) *Cell Tissue Res.* **236**, 207-216.
17. Ogiwara, K., Takano, N., Shinohara, M., Murakami, M. & Takahashi, T. (2005) *Proc. Natl. Acad. Sci. USA* **102**, 8442-8447.
18. Rajapakse, S., Ogiwara, K., Takano, N., Moriyama, A. & Takahashi, T. (2005) *FEBS Lett.* **579**, 6879-6884.
19. Likhareva, V. V., Mikhailova, A. G., Vaskovsky, B. V., Garanin, S. K. & Rumsh, L. D. (2002) *Lett. Peptide Sci.* **9**, 71-76.

Figure legends

Fig. 1. Cloning, expression, and localization of medaka EP. *(A)* Schematic representation of the medaka EP domain structures. Medaka EP consists of a putative signal anchor (SA), a mucin-like domain, two low-density-lipoprotein receptor (LDLR) domains, two complement-component C1r or C1s (C1r/s) domains, a MAM domain (named for the motifs found in Meprin, *Xenopus laevis* A5 protein, and protein tyrosine phosphatase μ), a macrophage scavenger receptor (MSCR) domain, and a serine

protease domain with active-site residues of histidine (H), aspartate (D), and serine (S). The disulfide bond connecting the heavy and light chains is shown. *(B)* Amino acid sequence alignment of the EP serine protease domain. Amino acid residues are numbered based on the sequence of medaka EP (top numbers). For comparison, the bovine chymotrypsinogen (Chymo) residue numbers are included in parentheses at the bottom of each block. The arrow indicates a putative activation site between the heavy and light chains. The active-site residues (H, D, and S) are boxed. The positions of mutations are indicated by asterisks. *(C)* Expression of medaka EP mRNA in various tissues, analyzed by Northern blotting. The sizes of the detected mRNAs are shown at the left. The lower panel shows the results for medaka cytoplasmic actin mRNA as a control. *(D)* Expression of EP mRNA in the gastrointestinal tract, analyzed by RT-PCR. The medaka gastrointestinal tract was divided into eight pieces, from the stomach (lane 1) to the anus (lane 8), and the PCR products from each piece were electrophoresed. *(E)* *In situ* hybridization of EP mRNA in the medaka intestine. Neighboring sections of medaka intestine were hybridized with an EP antisense (left) or sense RNA probe (right). Scale bars: 100 μm. *(F)* Expression of the medaka EP protein, analyzed by Western blotting using the medaka anti-EP antibody. Extracts of the intestine, testis, and ovary (left), and of nuclear, membrane, and cytosolic fractions of the medaka intestine (right) were analyzed. The size of the EP protein detected is shown at the left. *(G)* Immunohistochemical analysis of EP in the medaka intestine, performed with the medaka anti-EP antibody (left). The control section was stained with the primary antibody previously treated with the antigen (right). Scale bars: 200 μm.

Fig. 2. Specificity of medaka EP-1 protease for peptide and protein substrates. *(A)* Active recombinant EP proteases were assayed using GD4K-βNA as a substrate, including bovine EP proteases obtained from Novagen (Nvg) and New England Biolabs (Neb). *(B)* Active recombinant EP proteases were assayed using various synthetic peptide substrates. *(C)* Active EP proteases were analyzed by gelatin zymography. *(D)* Fibronectin (4 μg) was incubated with active EPs (100 ng) at 37°C for 12 h. *(E)* HMW kininogen (10 μg) was incubated with active recombinant EPs (100 ng) at 37°C for 12 h. *(F)* Two μg of the control protein containing the D4K site were incubated with active recombinant EPs (100 ng) at 37°C for 1 h.

Fig. 3. Computer model of the putative three-dimentional structure of the serine protease domain of medaka EP. The model was created by using Swiss Model, starting from the crystallographically derived structure of bovine EP. The side chains of the active-site residues (Asp, His, and Ser) of the enzyme are shown in stick-model form in red; the substrate Asp-Asp-Asp-Lys, in blue; the substrate residue in the P2 position, in dark blue; and the side chains at the positions of point mutations, in green. The dashed boxes in black and purple indicate the putative S1 and S2 subsite regions, respectively.

Fig. 4. Specificity of mutant medaka EP proteases for peptide and protein substrates containing the D4K-cleavage sequence. *(A)* The specific activities of wild-type (EP-1) and mutant EP proteases were determined using GD4K-βNA. *(B)* A recombinant fusion protein of medaka gelatinase A (5 μg) was incubated separately with active EP proteases (100 ng) at 37°C for 1 h and analyzed by SDS-PAGE. *(C)* A recombinant fusion protein

of human kallikrein 8 (hK8) (5 μg) was incubated with active EP proteases (100 ng) at 37°C for 2 h. After incubation, the samples were analyzed by SDS-PAGE (upper panel), or assayed for activity with Pro-Phe-Arg-MCA (lower panel). *(D)* Culture media collected from two culture dishes (10 cm in diameter) of CHO cells transfected with the pCMV tag4 vector containing the human sctPA sequence were affinity-purified using Ni^{2+}-Sepharose, and the resulting eluate was treated separately with active EP proteases (100 ng) at 37°C for 1 h. The samples were then analyzed by SDS-PAGE/Western blotting using anti-human tPA antibody (upper panel) or anti-His probe antibody (lower panel). The positions of 58- and 61-kDa sctPA detected with the antibodies are shown.

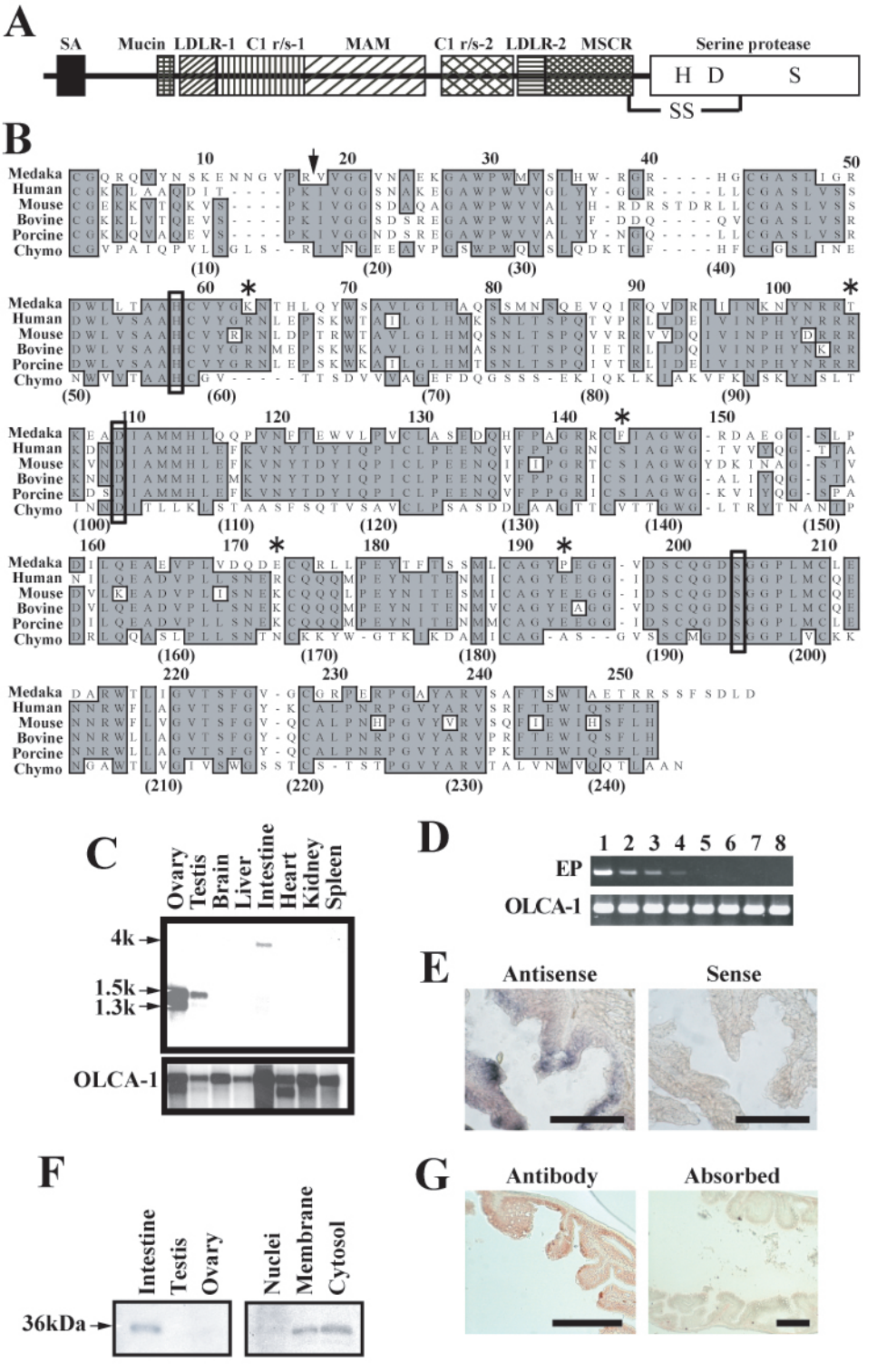
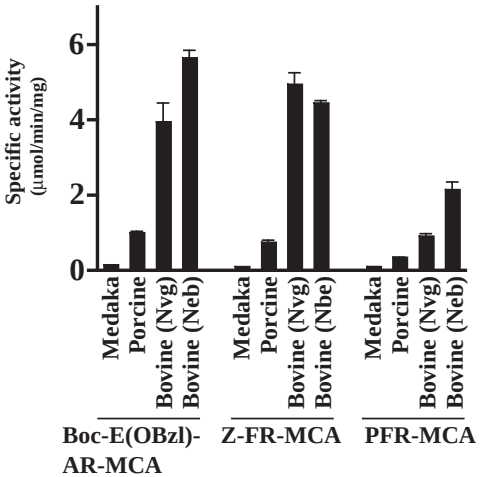
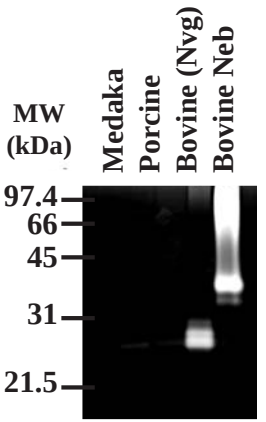


Fig. 1.
Ogiwara and Takahashi





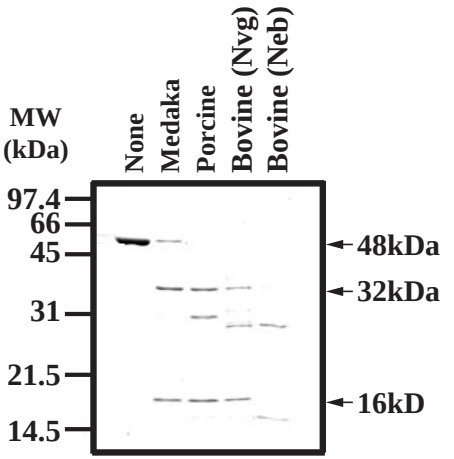


Fig. 2.
Ogiwara and Takahashi

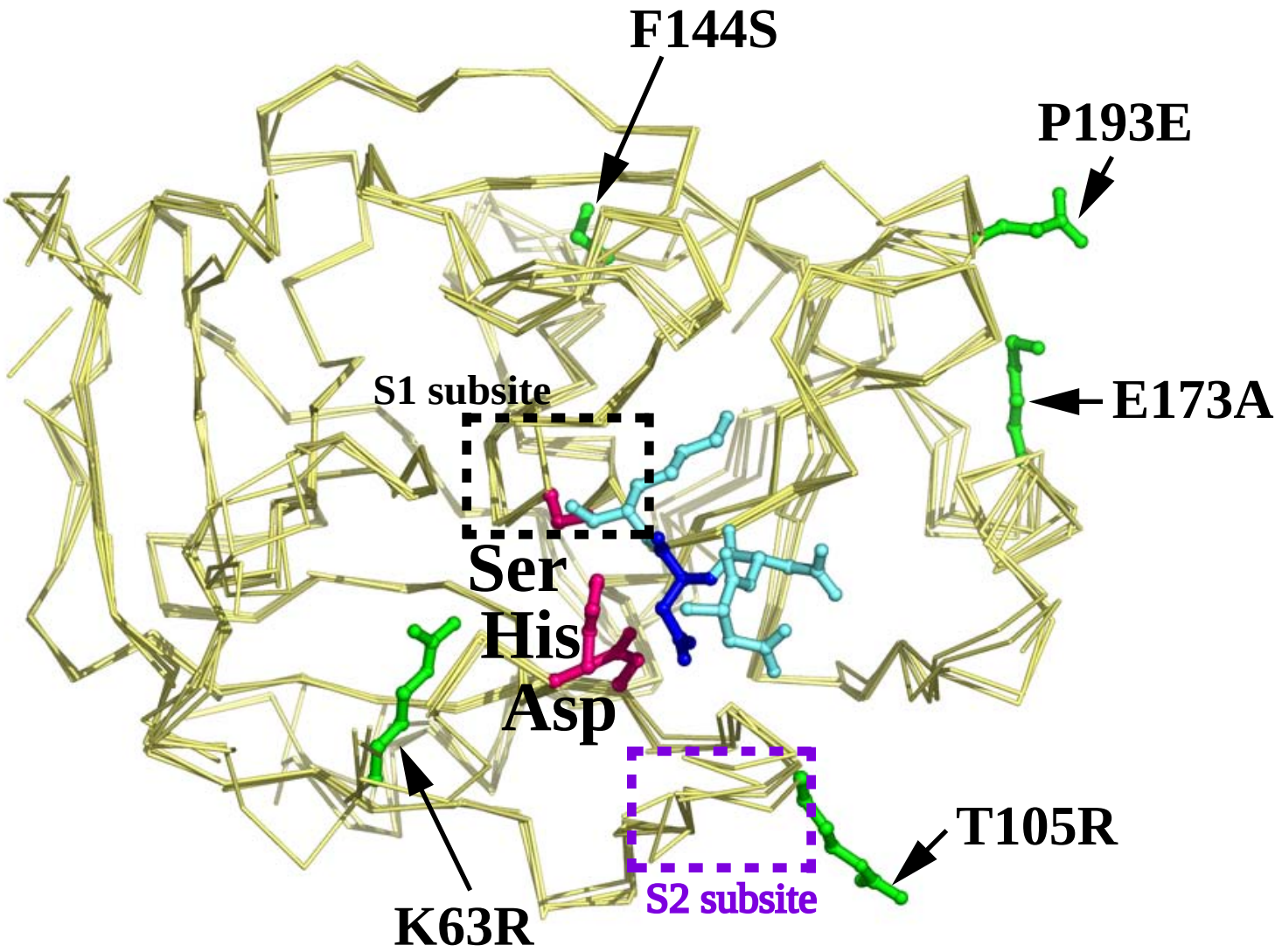


Fig. 3.
Ogiwara and Takahashi

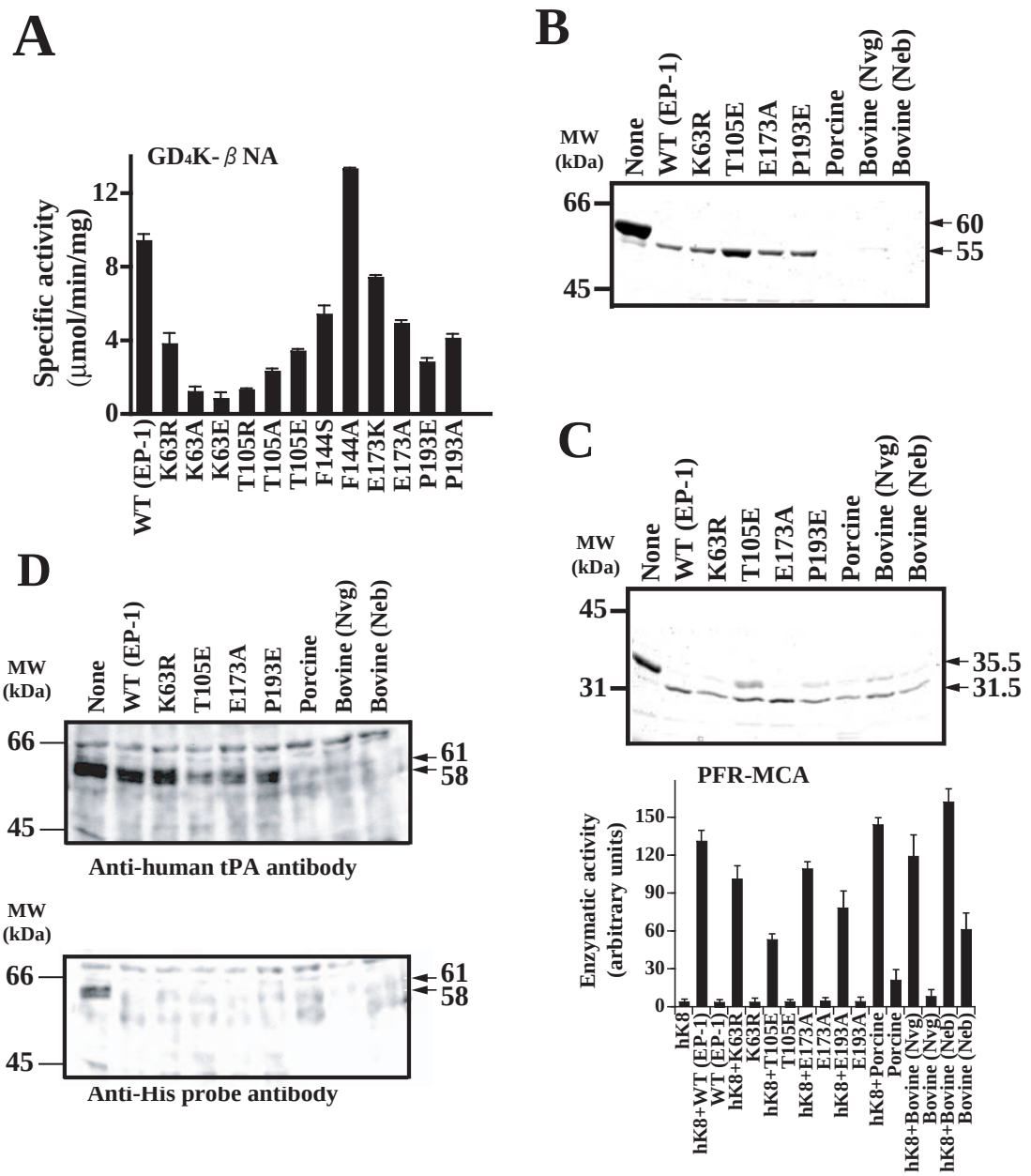


Fig. 4.
Ogiwara and Takahashi

Table 1.

Kinetic parameters of various EP proteases as measured on GD4K- β NA and MCA-containing substrates

	GD4K- β NA			Boc-E(OBzl)-AR-MCA			Z-FR-MCA			PFR-MCA		
	*K*m (mM)	*k*cat (min^{-1})	*k*cat/*K*m ($mM^{-1} \cdot min^{-1}$)	*K*m (mM)	*k*cat (min^{-1})	*k*cat/*K*m ($mM^{-1} \cdot min^{-1}$)	*K*m (mM)	*k*cat (min^{-1})	*k*cat/*K*m ($mM^{-1} \cdot min^{-1}$)	*K*m (mM)	*k*cat (min^{-1})	*k*cat/*K*m ($mM^{-1} \cdot min^{-1}$)
WT (EP-1)	0.7	940	1300	0.2	6.7	34	0.1	2.9	29	10	140	14
K63R	0.2	210	1100	1.2	12	10	0.1	1.4	14	1.1	11	10
T105E	0.4	260	650	1.3	11	9	0.1	2.0	20	1.3	16	12
E173A	0.3	320	1100	1.0	10	10	0.4	2.3	6	1.0	9.2	9
P193E	0.4	290	730	0.2	2.3	12	0.2	1.7	9	1.0	27	27
Porcine	0.4	530	1300	0.3	110	370	0.2	55	280	3.9	300	77
Bovine (Nvg)	0.8	770	960	0.5	740	1500	0.5	720	1400	4.0	790	200
Bovine (Neb)	0.5	1500	3000	0.4	570	1400	0.4	600	1500	2.9	1300	450

Supporting Information

Supporting Methods

Materials and Methods

cDNA Cloning of Medaka EP.

RNA was isolated from medaka intestine and ovary using Isogen (Nippon Gene, Tokyo, Japan). The first strand of cDNA was synthesized from total RNA from the medaka intestine using a SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). Two degenerate oligonucleotide PCR primers were synthesized based on the cDNA sequences of the conserved C-terminal catalytic protease domains in mammalian EPs (sense primer, 5'-TCIGC(C/T)GC(A/C)CACTG(C/T)GT(C/G)TA(C/T)(A/G)G(A/G)-3', which corresponds to the sequence around the active-site histidine, NH_2-Ser-Ala-Ala-His-Cys-Val-Tyr-Gly-COOH; and antisense primer, 5'-(G/T)A(A/G)TGG(C/T)CC(G/T)CC(A/T)GAATC(A/C)CCCTG-3', which corresponds to the sequence around the active-site serine, NH_2-Gln-Gly-Asp-Ser-Gly-Gly-Pro-Leu-COOH). cDNAs were amplified under the following conditions: 3 min at 94°C; 30 cycles of 0.5 min at 94°C, 0.5 min at 55°C, and 0.5 min at 72°C; and a 7-min final extension at 72°C. Fragments of about 0.5-kb in size were recovered from the PCR products by agarose gel purification and subcloned into pBluescript II KS(+) (Stratagene, La Jolla, CA). A 461-bp clone was obtained and was used as a probe for further screening of a medaka cDNA library.

A medaka intestine random cDNA library was constructed in λgt10 and was

packaged using Gigapack III packaging extract (Stratagene). Approximately 6×10^5 plaques from the library were transferred to nylon membranes (Schleicher & Schuell, Dassel, Germany) and hybridized at 65°C in a buffer containing 5×SSPE, 0.5% SDS, 5×Denhardt's solution (Wako, Osaka, Japan), and 100 μg/ml denatured salmon sperm DNA with the ^{32}P-labeled 461-bp PCR fragment described above. Filters were washed with increasing stringency, with a final wash of 0.1×SSC/0.1% SDS at 50°C. Phage DNA was subcloned into pBluescript (II) KS+ for sequencing. An EP clone containing a 2689-bp cDNA (nucleotides 611-3298) was obtained. Further screening was conducted with the same library using an EP 477-bp probe (nucleotides 630-1101), resulting in isolation of a 1364-bp cDNA containing the 5' portion of the EP sequence.

The 3' portion of medaka EP was obtained by 3' RACE (1) using the 3'-Full RACE Core Set (Takara, Tokyo, Japan). The sense primers used were 5'-GACATTCTACAGGAGGCTGAGGTT-3' (RACE1; nucleotides 2900 to 2923) and 5'-CGTCTCTTACCCGAGTACACCTTC-3' (RACE2; nucleotides 2951 to 2974). Two rounds of PCR reactions were performed under the conditions of 35 cycles of 0.5 min at 94°C, 0.5 min at 55°C, and 1 min at 72°C for the first round and 35 cycles of 0.5 min at 94°C, 0.5 min at 57°C, and 1 min at 72°C for the second round. The amplified products were then cloned into pBluescript II plasmid (Stratagene) and sequenced.

cDNA Cloning of Medaka Trypsinogen.

For medaka trypsinogen, two degenerate oligonucleotide PCR primers were synthesized based on the cDNA sequence of conserved regions in serine protease (sense primer, 5'-GT(G/T)(C/G)T(C/G/T)(A/T)C(A/T)GCTGC(C/T)CACTG-3', which corresponds to the amino acid sequence NH_2-Val-Leu-Thr-Ala-Ala-His-Cys-COOH;

antisense primer, 5’-(A/T)GGGCC(A/T)CC(A/T/G)GAGTC(A/T)CC-3’, which corresponds to the amino acid sequence NH_2-Gly-Asp-Ser-Gly-Gly-Pro-COOH). cDNAs were PCR-amplified under the conditions described for EP. A 435-bp fragment was cloned into pBluescript II KS(+) (Stratagene) and sequenced.

A 5’ portion of medaka trypsinogen was obtained by 5’ RACE (1) using the 5’-RACE system, Version 2.0 (Invitrogen). The antisense primers used were 5’-AGGAGGTGATGAACTG-3’ (GSP-1; nucleotides 273 to 288, AB272106), 5’-CTCGGTTCCGTCATTGTTCCGGGAT-3’ (GSP-2; nucleotides 249 to 272, AB272106), and 5’-CCAGACGCACCTCCACTCGGGACT-3’ (nested GSP; nucleotides 214 to 237, AB272106). Two rounds of PCR were performed under the conditions of 35 cycles of 0.5 min at 94°C, 0.5 min at 55°C, and 1 min at 72°C for the first PCR and 35 cycles of 0.5 min at 94°C, 0.5 min at 60°C, and 1 min at 72°C for the second PCR. The amplified products were then cloned into pBluescript II KS(+) plasmid (Stratagene) and sequenced.

A 3’ portion of medaka trypsinogen was obtained by 3’ RACE (1) using the 3’-Full RACE Core Set (Takara). The sense primers used were 5’-CATGATCACCAACTCCATGTTCTG-3’ (RACE1; nucleotides 545 to 568, AB272106) and 5’-TGGATACCTGGAGGGAGG-3’ (RACE2; nucleotides 572 to 589, AB272106). Two rounds of PCR were performed under the conditions of 35 cycles of 0.5 min at 94°C, 0.5 min at 55°C, and 1 min at 72°C for the first PCR and 35 cycles of 0.5 min at 94°C, 0.5 min at 57°C, and 1 min at 72°C for the second PCR. The amplified products were then cloned into pBluescript II KS(+) plasmid (Stratagene) and sequenced.

RT-PCR Analysis of EP Transcripts.

To identify the two distinct EP transcripts, *enteropeptidase-1* (*EP-1*) and *enteropeptidase-2* (*EP-2*), expressed in the medaka intestine, RT-PCR was conducted with KOD plus DNA polymerase (Toyobo, Osaka, Japan) using medaka intestine total RNA. The primers used were 5'-AGAACATCACAGGTGAACCGGTGA-3' (sense primer, nucleotides 1-24, AB272104) and 5'-TTCTGACATTCCTGAAGGGACAGC-3' (antisense primer, nucleotides 3930-3953, AB272104). PCR conditions were 2 min at 94°C, followed by 30 cycles of 30 sec at 94°C, 15 sec at 60°C, and 6 min at 68°C. The products were sequenced as described above. In some experiments, RT-PCR analyses were performed using specific primers: 5'-CAAGAACTACAACAGAAGAA-3' (sense) and 5'-GTGTATTGAGAAAAAGGTTGTTAA-3' (antisense) for *EP-1* (nucleotides 2719-3415, AB272104) and 5'-CAAGAACTACAACAGAAGAA-3' (sense) and 5'-CTGTACTAAGAAAAAATTTGTCAT-3' (antisense) for *EP-2* (nucleotides 2747-3443, AB272105). PCR conditions were 3 min at 94°C, followed by 20, 22, 24, 26, or 28 cycles of 30 sec at 94°C, 30 sec at 60°C, and 30 sec at 72°C.

For ovary 1.5- and 1.3-kb EP transcripts, RACE methods (1) were used. The sequence of the 5'-end was confirmed by 5' RACE using a 5'-RACE system (Invitrogen). The primers used were as follows: 5'-AGGTAACCAAGCAGAG-3' (nucleotides 3207-3222, AB272104) for the reverse transcriptase reaction, 5'-GAGAACGAGGAGCGCCTGGTCTCA-3' (nucleotides 3169-3192, AB272104) for the first PCR, and 5'-ATCCATGAAGTGAAAGCAGACACT-3' (nucleotides 3142-3165, AB272104) for the second PCR. PCR was performed for 35 cycles of 30 sec at 94°C, 30 sec at 55°C, and 2 min at 72°C. The 3' end of the transcripts was determined by 3' RACE (1) by means of a 3'-Full RACE Core Set (Takara), as

described above.

RT-PCR Detection of EP mRNA in the Gastrointestinal Tract.

The gastrointestinal tract was obtained from mature medaka (body sizes, 3-4 cm) and divided into eight pieces, each about 0.5 mm long. Specimens from five fish were combined for total RNA preparation. Aliquots of 2 μg of the total RNAs were used for reverse transcription. PCR was performed for 25 cycles using Ex *Taq* DNA polymerase (Takara) and the primers 5'-AGGACCAAACGGAACATTTC-3' (sense, nucleotides 802-821, AB272104) and 5'-GAGAGGGACGCAGGAGGA -3' (antisense, 1422-1439, AB272104).

Northern Blotting.

Two μg of poly(A) RNA from various tissues of the medaka were electrophoretically fractionated and transferred to a Nytran-plus membrane (Schleicher & Schuell). The blots were hybridized with ^{32}P-labelled cDNA fragments (EP, nucleotides 3359-3953, AB272104; trypsinogen, nucleotides 572-835, AB272106) in buffer containing 50% formamide, 5×0.15 M NaCl/8.65 mM NaH_2PO_4/1.25 mM EDTA (SSPE), 1% SDS, 5×Denhardt's solution, and 100 μg/ml denatured salmon sperm DNA. The membranes were washed twice in 2×SSC/0.05% SDS and then twice in 0.1×SSC/0.1% SDS at 50°C. As a control, medaka cytoplasmic actin (OLCA1) mRNA was detected with a ^{32}P-labeled 312-bp DNA fragment of the fish cDNA (2).

Southern Blotting

Medaka genomic DNA was extracted as described previously (3), with the

exception that whole-genomic DNA was purified from the entire medaka body. Twenty μg of genomic DNA were completely digested with various restriction enzymes. The digested DNA was fractionated on a 0.7% agarose gel and alkaline-transferred to a Nytran-plus membrane (Schleicher & Schuell). The blot was hybridized at 60°C for 16 h in 6×SSPE, 5×Denhardt's solution, 1% SDS, 10% dextran sulfate, and 100 μg/ml denatured herring sperm DNA with a ^{32}P-labeled 595-bp fragment of medaka EP cDNA (nucleotides 3359-3953, AB272104). The membrane was washed at 60°C in 0.1×SSC/0.1% SDS and exposed to Kodak Biomax Film.

***In situ* Hybridization.**

In situ hybridization was performed using frozen intestine and ovary sections (15 *μ*m) as described previously (4). RNA probes were prepared by *in vitro* transcription of reverse-transcriptase fragments of cDNAs with T3 or T7 RNA polymerase using a digoxigenin (DIG) RNA-labeling Kit (Roche Molecular Biochemicals, Mannheim, Germany). A 595-bp cDNA fragment (nucleotides 3359-3953, AB272104) was used as a specific probe. Hybridization was conducted at 50°C for 18 h in 50% formamide, 5×Denhardt's solution, 6×SSPE, and 0.5 mg/ml yeast transfer RNA. The sections were washed once at 50°C in 50% formamide/2×SSC for 30 min, once at 50°C in 2×SSC for 20 min, and twice at 50°C in 0.2×SSC for 20 min. The hybridization probes were detected using a Dig Nucleic Acid Detection Kit (Roche Molecular Biochemicals).

Preparation of Recombinant Proteins.

For preparation of medaka recombinant trypsinogen, a cDNA fragment (nucleotides

72-755, AB272106) containing the trypsinogen coding sequence, but without the putative signal sequence, was amplified by PCR using the following primers: 5’-CCGGAATTCCTTGACGATGACAAG-3’ and 5’-CCCAAGCTTTCAGTTGCTAGCCATGGT-3’. The PCR product was digested with *Eco*RI and *Hind*III, gel-purified, and ligated into the pET30a expression vector. The expression of recombinant medaka trypsinogen in the *Escherichia coli* expression system and its purification with an Ni^{2+}-Sepharose column were the same as for the wild-type EP protein described above. The purified recombinant protein was renatured by dialysis against 50 mM Tris·HCl (pH 8.0) and further purified with a column of Resource Q. These procedures yielded a fusion protein of medaka trypsinogen that had a vector-derived 52-residue peptide at its N terminus in addition to the 227-residue sequence of the fish trypsinogen. Thus, this recombinant fusion protein contained two EP-cleavage sites: one from the vector used and the other from trypsinogen itself.

For preparation of the insertional mutant of human tissue-type plasminogen activator (tPA), a cDNA for human tPA (5) was first obtained by RT-PCR from human ovary total RNA (Stratagene) using the primers 5’-CCCAAGCTTATGAAGAGAGGGCTCTGCTGT-3’ (sense-1) and 5’-CTTATCGTCATCATGATGATGATGATGGTGTCTGGCTCCTCTTCT-3’ (antisense-1) (BC007231). Using the cDNA as a template, two PCR products were amplified with following primer combinations: sense-1 and antisense-1; and 5’-CACCATCATCATCATCATGATGACGACGATAAGTCTTACCAAGTGATC-3’ (sense-2) and 5’-CCGCTCGAGTCACGGTCGCATGTTGTCACGAAT-3’ (antisense-2). Using a mixture of these amplified DNAs as templates, a second PCR was performed with the sense-1 and antisense-2 primers. The PCR products were digested with *Hind*III

and *Xho*I, then gel-purified and ligated into the pCMV tag4 mammalian expression vector (Stratagene). The resulting mutant was confirmed by DNA sequencing and transfected into CHO cells cultured in F-12 medium (Invitrogen) containing 10% fetal bovine serum (Biological Industries, Beit Haemek, Israel). Transfection was performed using Lipofectamin 2000 (GE Healthcare Biosciences, Uppsala, Sweden). The above procedure produced a human tPA fusion protein having 11 extra amino acid residues (His-His-His-His-His-His-Asp-Asp-Asp-Asp-Lys: a His-tag sequence followed by an EP-cleavage site) at the N terminus of the mature tPA. The fusion protein secreted from transfected CHO cells was collected from the culture medium using an Ni^{2+}-Sepharose column. Treatment of the fusion protein with EP proteases generated mature tPA without the 11-residue N-terminal peptide.

Recombinant human kallikrein 8 was prepared as described previously (6).

Recombinant medaka gelatinase A was prepared as described previously (4).

Enzyme Assays.

EP activity was routinely determined using the specific substrate Gly-Asp-Asp-Asp-Asp-Lys-β-naphthylamide (GD_4K-βNA) (Sigma, St. Louis, MO) according to the method of Mikhailova and Rumsh (7). Enzyme activity for various 4-methylcoumaryl-7-amide (MCA)-containing peptide substrates (Peptide Institute, Osaka, Japan) was determined by the method of Barrett (8). For kinetic studies, initial velocities, extrapolated from the plot of product versus time, were transformed into double-reciprocal plots (9). Maximum velocities (V_{max}) and K_m and k_{cat} values were obtained from the intercepts of these plots. For all experiments, the results of at least three separate determinations are shown. The active recombinant protein concentrations

were determined using the active-site titrant *p*-nitrophenyl-*p'*-guanidinobenzoate HCl (Sigma) according to the method described previously (10).

Production of Anti-medaka EP Protease Antisera.

The protein antigen was produced using a bacterial expression system with pET30a as described above. The recombinant protein eluted from an Ni^{2+}-Sepharose column was injected into rabbits. The specific antibody was affinity-purified using membranes onto which pure antigen was blotted (4).

Western Blotting and Immunohistochemistry.

Whole tissues of medaka intestine, ovary, and testis were separately homogenized in PBS containing 5 mM EDTA and protease inhibitor cocktail (Wako Chemicals, Osaka, Japan), and centrifuged at 18,000 × g for 10 min to obtain supernatant fractions. The supernatants were analyzed by Western blot analysis (4). For fractionation of medaka intestine, tissues were homogenized in 50 mM Tris·HCl (pH 7.4), 10 mM KCl, 10 mM $MgCl_2$, 1 μM dithiothreitol, 5 mM EDTA, and protease inhibitor cocktail, and centrifuged at 1,600 × g for 8 min. The pellet was collected as crude nuclei. The supernatant was further centrifuged at 100,000 × g for 30 min. The resulting supernatant and pellet were used as cytosolic and membrane fractions, respectively (11). The primary antibodies were affinity-purified EP protease antibodies as described above.

Intestine sections (15 μm) were cut on a cryostat and thaw-mounted onto slides coated with silan. Sections on slides that were fixed with 4% paraformaldehyde in PBS for at least 15 min were treated with 3% H_2O_2 in PBS. After blocking with BlockAce (Dainippon Seiyaku, Osaka, Japan) for 1 h at room temperature, each section was

incubated with a purified primary antibody for 1 h at room temperature, and was then washed with PBS. Bound antibodies were detected using DakoCytomatin EnVision^{+} System-labeled polymer-HRP anti-rabbit antibody (Dako, Carpinteria, CA) according to the manufacturer's instructions. Immunocomplexes were detected using an AEC Kit (Vector Laboratories, Burlingame, CA).

Gel-Filtration Chromatography.

Gel-filtration chromatography was performed using a HiLoad 16/60 Superdex 200-pg column (GE Healthcare Biosciences) equilibrated with 50 mM Tris·HCl (pH 8.0) and 0.2 M NaCl. Medaka intestine was homogenized in the same buffer containing 5 mM EDTA and protease inhibitor cocktail and centrifuged at 18,000 × g for 10 min to obtain the supernatant. The resulting supernatant was applied to the column at a flow rate of 24 ml/h. Fractions of 1 ml were collected and assayed for EP protease activity using GD_4K-βNA as a substrate. The active fractions were pooled and used for Western blotting. Calibration of the column was conducted using an HMW Gel-filtration Calibration Kit (GE Healthcare Biosciences).

Enzyme Stability.

One-hundred nanograms of medaka, porcine, and bovine enteropeptidase were separately incubated at 37°C in 20 mM Tris· HCl buffer (pH 7.4) containing 50 mM NaCl and 2 mM $CaCl_2$. The enzyme activity was measured at various time points (0 to 96 h) using GD_4K-βNA as a substrate.

Inhibitor Assay.

Active medaka enteropeptidase was preincubated with various inhibitors at 37℃ in 20 mM Tris· HCl buffer (pH 7.4) containing 50 mM NaCl and 2 mM $CaCl_2$. After incubation for 10 min, the enzyme activity was measured using GD_4K-βNA as a substrate.

References

1. Frohman, M. A., Dush, M. K. & Martin, G. R. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8998-9002.
2. Kusakabe, R., Kusakabe, T. & Suzuki, N. (1999) *Int. J. Dev. Biol.* **43**, 541-554.
3. Kimura, A., Yoshida, I., Takagi, N. & Takahashi, T. (1999) *J. Biol. Chem.* **274**, 24047-24053.
4. Ogiwara, K., Takano, N., Shinohara, M., Murakami, M. & Takahashi, T. (2005) *Proc. Natl. Acad. Sci. USA* **102**, 8442-8447.
5. Pennica, D., Holmes, W. E., Kohr, W.J., Harkins, R. N., Vehar, G. A., Ward, C. A., Bennett, W. F., Yelverton, E., Seeburg, P. H., Heyneker, H. L., *et. al.* (1983) *Nature* **301**, 214-221.
6. Rajapakse, S., Ogiwara, K., Takano, N., Moriyama, A. & Takahashi, T. (2005) *FEBS Lett.* **579**, 6879-6884.
7. Mikhailova, A. G. & Rumsh, L. D. (1999) *FEBS Lett.* **442**, 226-230.
8. Barrett, A. J. (1980) *Biochem. J.* **187**, 909-912.
9. Lineweaver, H. & Burk, D. (1934) *J. Am. Chem. Soc.* **56**, 658-663.
10. Chase, T. J. R. & Shaw, E. (1976) *Biochem. Biophys. Res. Commun.* **29**, 508-514.
11. Pie, D., Kang, T. & Qi, H. (2000) *J. Biol. Chem.* **275**, 33988-33997.

Fig. 5. Expression of two distinct *EP* transcripts in the medaka intestine. (*A*) Amino acid sequence alignment of EP-1 (upper) and EP-2 (lower) is shown. (*B*) RT-PCR analyses of the *EP-1* and *EP-2* transcripts were performed using specific primer pairs with total RNAs isolated from the medaka intestine. A transcript of medaka cytoplasmic actin-1 (OLCA-1) was amplified as a control. PCR cycle numbers are indicated at the top of the figure. (*C*) Southern blot analysis was performed using medaka genomic DNA (20 μg/lane) digested with various restriction enzymes as indicated.

Fig. 6. *In situ* detection of *EP* mRNA in the medaka ovary. Staining was performed with DIG-labeled antisense (*A and C*) and sense (*B and D*) probes. *(C)* The follicles indicated by the box in *(A)* are shown at higher magnification. (*D*) The follicles indicated by the box in (*B*) are shown at higher magnification. Scale bars: 500 μm in (*A*) and (*B*) and 200 μm in (*C*) and (*D*).

Fig. 7. Gel-filtration analysis of medaka intestine extracts. (*A*) The intestine extract was fractionated using a HiLoad 16/60 Superdex 200pg column. Fractions having GD_4K-βNA-hydrolyzing activity (indicated by a bar) were pooled. (*B*) The pooled active fraction was subjected to SDS-PAGE/Western blotting analysis under a reducing condition (left panel) or nonreducing condition (right panel) using anti-medaka EP protease antibody.

Fig. 8. Some enzymatic properties of recombinant medaka EP-1 and EP-2 protease. (*A*) The purity of purified recombinant medaka EP-1 and EP-2 protease was assessed by SDS-PAGE. Lane 1, medaka EP fusion protein; lane 2, medaka EP protease treated with

immobilized trypsin; lane 3, medaka EP protease purified using a resource Q column. (*B*) The enzyme activities of EP proteases were determined at various pHs using GD4K-βNA as a substrate. (*C*) Recombinant medaka trypsinogen was incubated with EPs for 15, 30, and 45 min at 37°C. After incubation, samples were analyzed by SDS-PAGE and visualized by CBB staining (upper panel). The relative amount of the active form of medaka trypsin at each time point (lower panel) was calculated based on the results shown in the upper panel. The results are presented as the means (±SD) of three separate experiments.

Fig. 9. Cloning and expression of medaka trypsinogen. (*A*) Amino acid sequence alignment of trypsinogen of the medaka, human (BAA08257), mouse (AAH61135), and salmon (CAA49676) is shown. A well-conserved D_4K-cleavage site for EP is indicated by the dashed box. Active-site residues (H, D, and S) are boxed. (*B*) The tissue distribution of medaka trypsinogen mRNA was analyzed by Northern blotting (upper panel); note that trypsinogen mRNA was detected in the testis, intestine, and spleen. The lower panel shows the detection of medaka cytoplasmic actin-1 (OLCA-1) mRNA as a control. The sizes of the detected mRNAs are shown at the right.

Fig. 10. Stability of EP protease. Medaka and mammalian EP proteases were incubated at 37°C in 20 mM Tris· HCl (pH 7.4), 0.2 M NaCl, and 2 mM $CaCl_2$. Aliquots of the reaction mixtures were taken at the indicated times for an activity assay using GD4K-βNA as a substrate. The enzyme activities relative to that at 0 time are shown.

Fig. 11. Specificity of mutant medaka EP proteases for synthetic peptide substrates. The

specific activities of wild-type (EP-1) and mutant EP proteases were determined using the synthetic peptide substrates.

Fig. 12. Activation of medaka trypsinogen by medaka wild-type (EP-1) and mutant EP proteases. Medaka recombinant trypsinogen was incubated separately with EP proteases at 37°C for 15, 30, and 45 min, and analyzed by SDS-PAGE followed by CBB staining (upper panel). The relative amount of the active form of medaka trypsin at each time point (lower panel) was calculated based on the results shown in the upper panel. The results are presented at the means (±SD) of three separate experiments.

Fig. 13. Specificity of mutant medaka EP proteases on protein substrates. *(A)* Fibrinogen (10 μg) was incubated with active EP proteases (100 ng) at 37°C for 12 h and analyzed by SDS-PAGE. *(B)* High-molecular-weight (HMW) kininogen (5 μg) was incubated with active EP proteases (100 ng) at 37°C for 2 h and analyzed by SDS-PAGE.

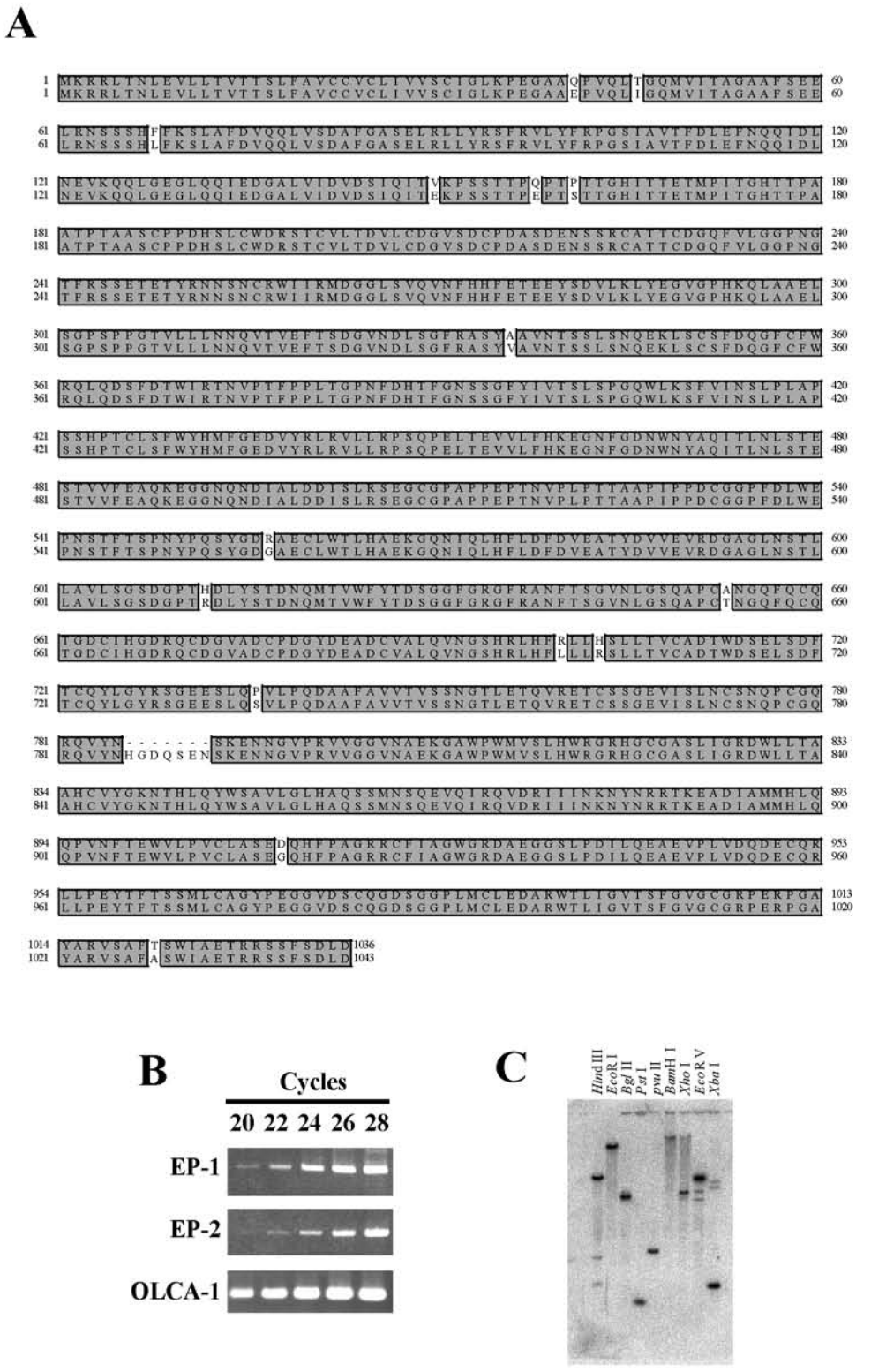


Fig. 5.
Ogiwara and Takahashi

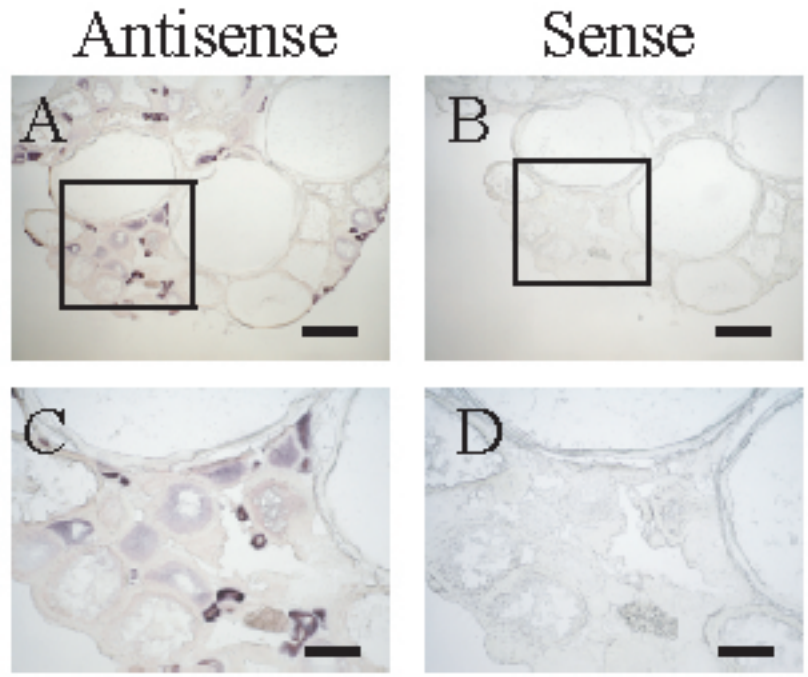


Fig. 6.
Ogiwara and Takahashi

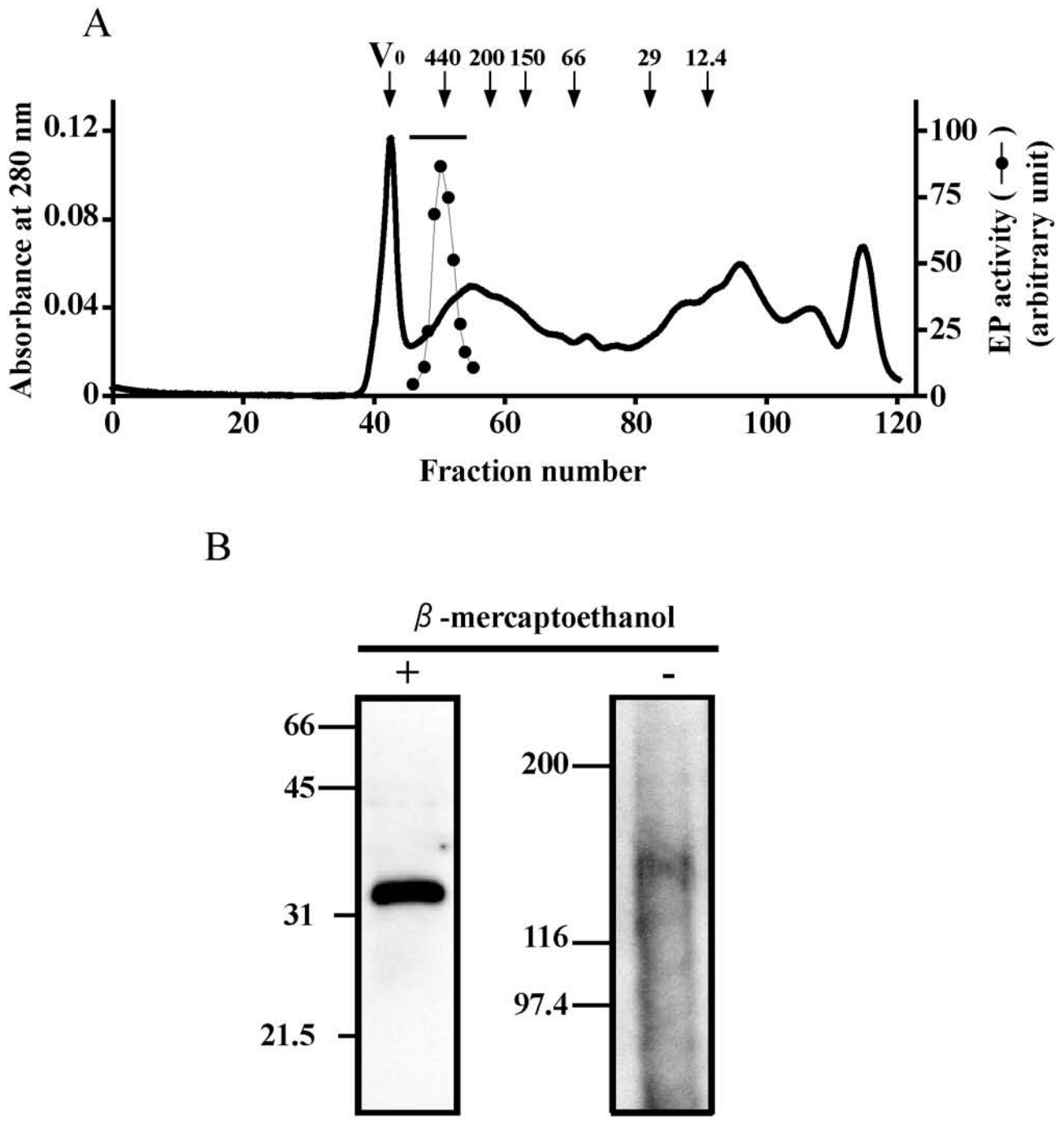


Fig. 7.
Ogiwara and Takahashi

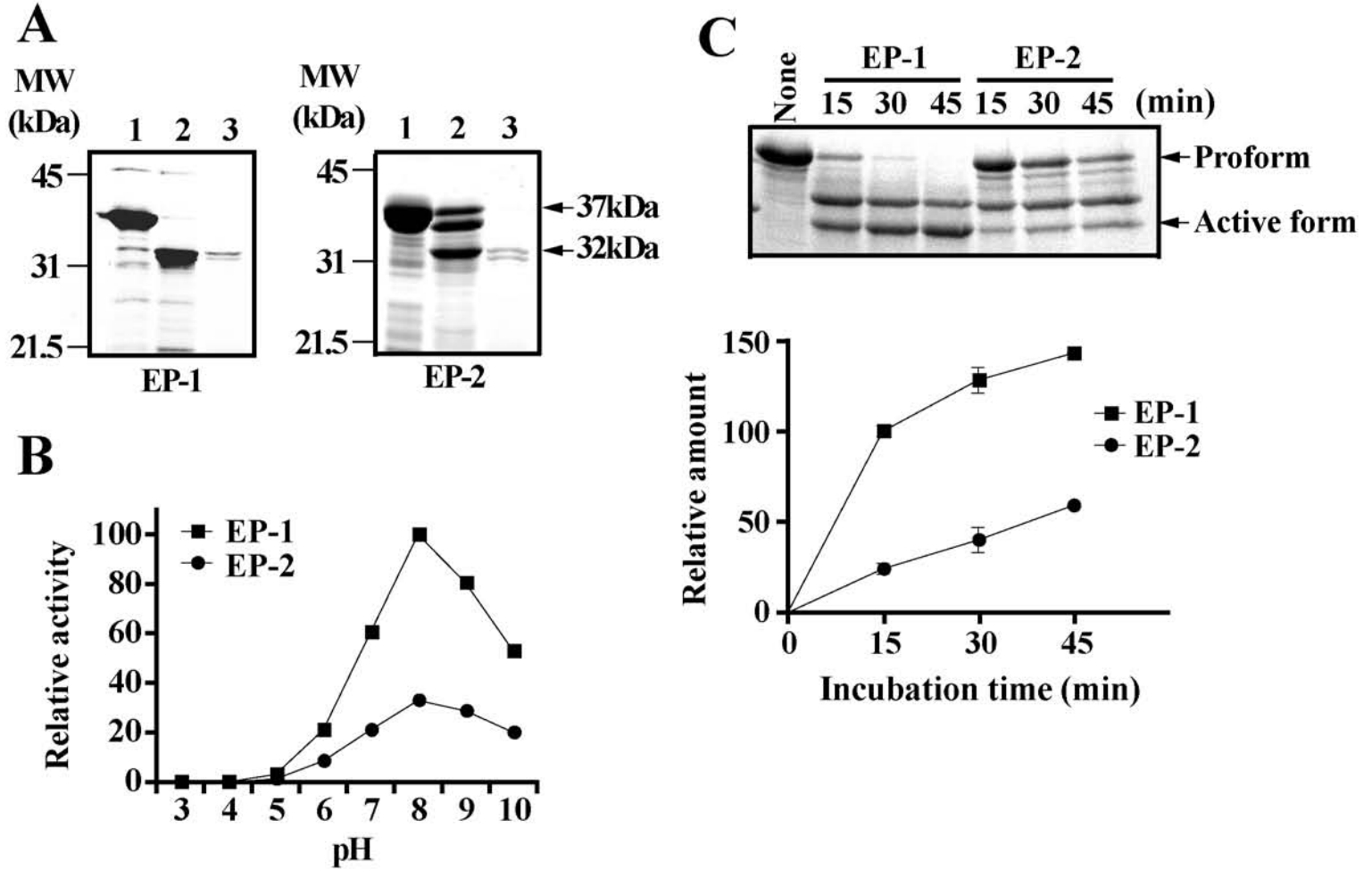


Fig. 8.
Ogiwara and Takahashi

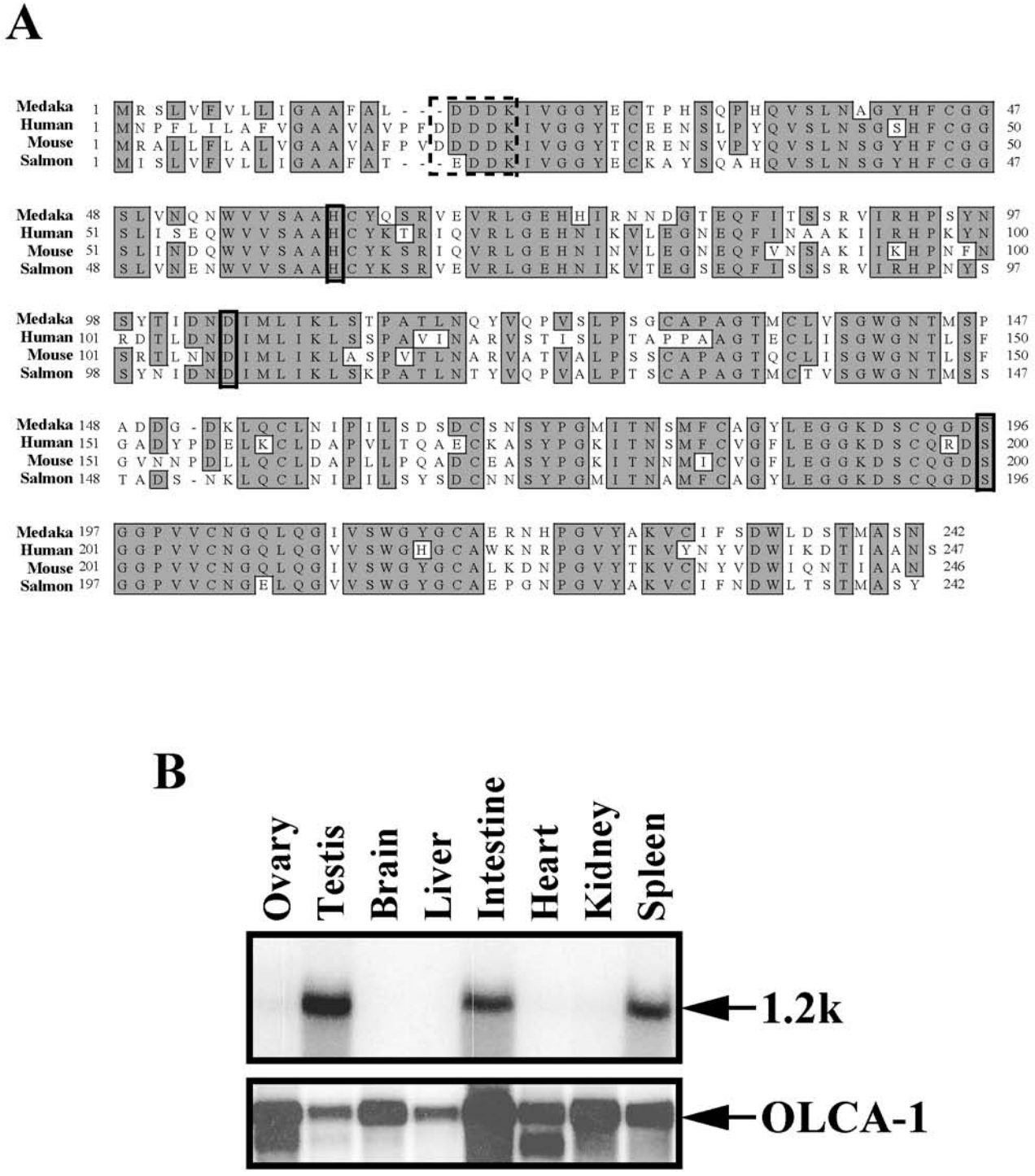


Fig. 9.
Ogiwara and Takahashi

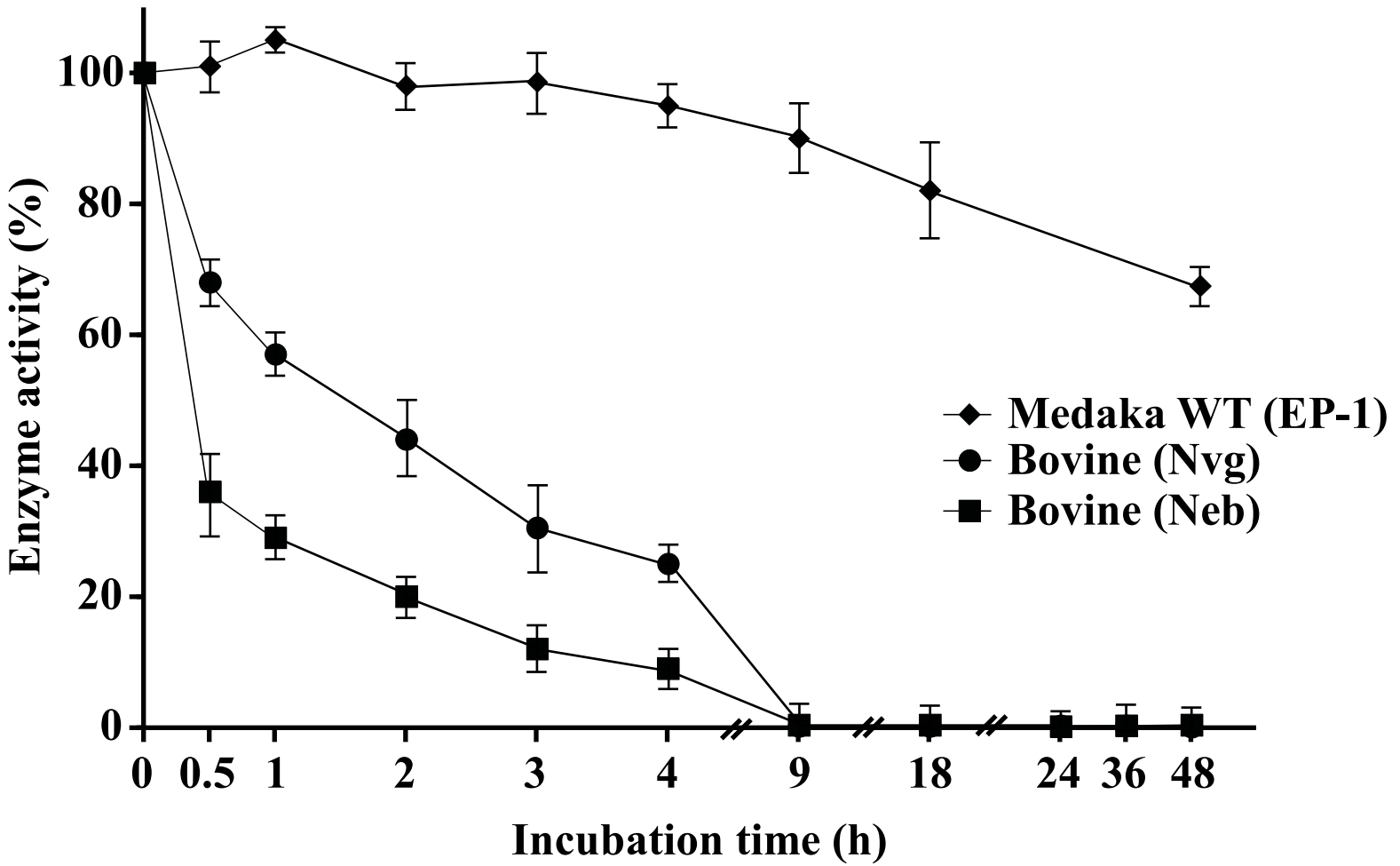


Fig. 10.
Ogiwara and Takahashi.

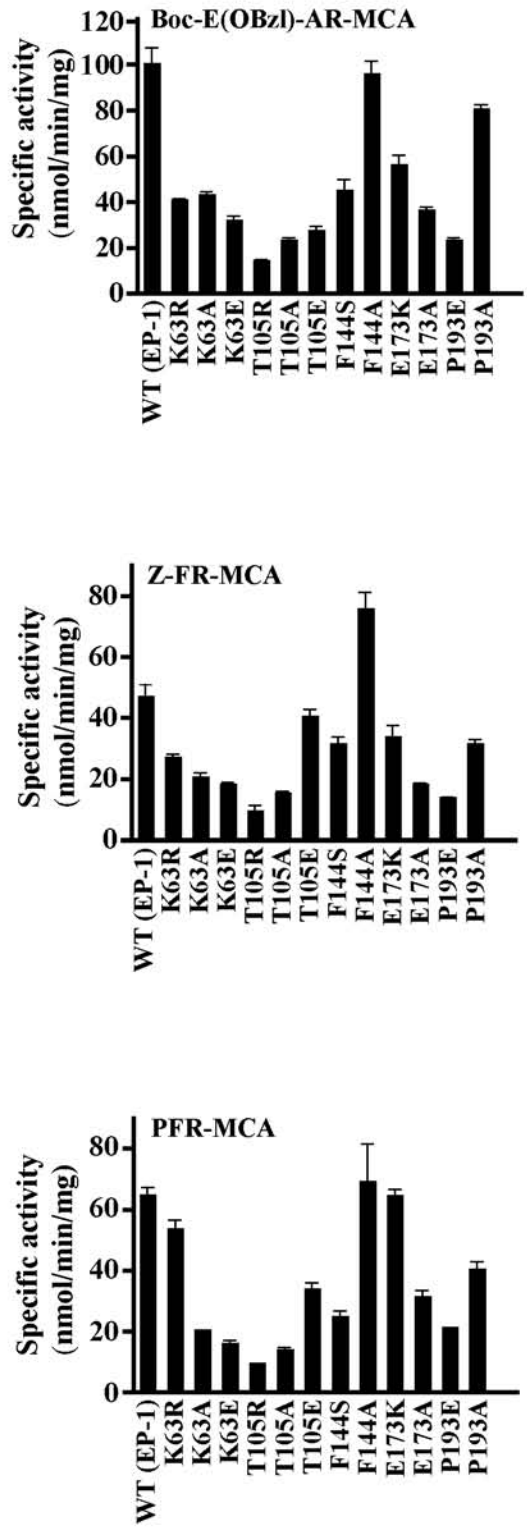


Fig. 11.
Ogiwara and Takahashi

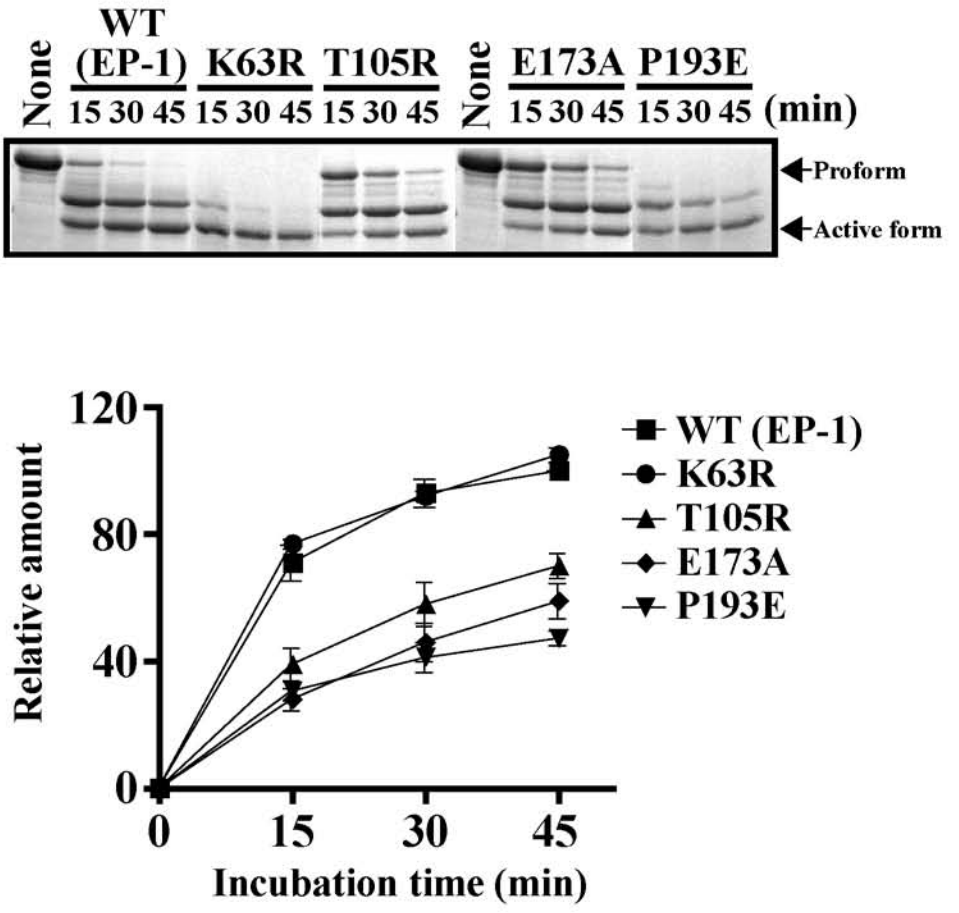


Fig. 12.
Ogiwara and Takahashi

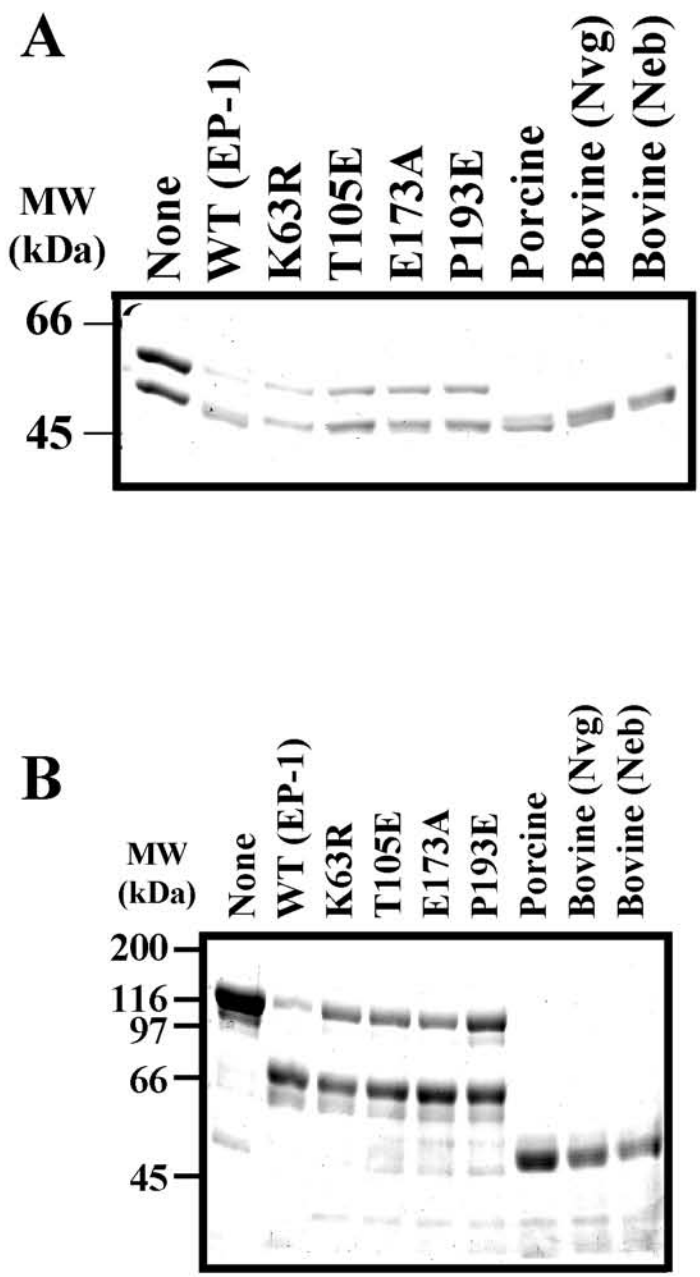


Fig. 13.
Ogiwara and Takahashi

Table 2. Effects of inhibitors on medaka EP-1 and EP-2 protease activity.

Inhibitor	Concentration	Inhibition (%)	
		EP-1	EP-2
EDTA	5.0 mM	5	10
DFP	0.2 mM	99	99
Benzamidine	1.0 mM	79	78
Antipain	0.1 mM	18	20
Leupeptin	0.1 mM	43	47
Chymostatin	0.1 mM	0	0
Aprotinin	0.01 mg/ml	0	5
SBTI	0.1 mg/ml	99	99
E-64	0.2 mM	0	0
Pepstatin	0.1 mM	0	3

Enzyme activities of medaka EP-1 and EP-2 protease were determined in the presence of various inhibitors using GD_4K-βNA as a substrate. Values are expressed as the percent inhibition of the respective control activity. Results are the averages of triplicate determinations.

Table 3. PCR primers used for mutant EP protease production by site-directed mutagenesis.

Mutant		Primer sequences
K63R	Sense	5'-GTCTATGGGAGGAACACACAC-3'
	Antisense	5'-GTGTGTGTTCCTCCCATAGAC-3
K63A	Sense	5'-GTCTATGGGGCGAACACACAC-3'
	Antisense	5'-GTGTGTGTTCGCCCCATAGAC-3'
K63E	Sense	5'-GTCTATGGGGAGAACACACAC-3'
	Antisense	5'-GTGTGTGTTCTCCCCATAGAC -3'
T105R	Sense	5'-AACAGAAGAAGGAAAGAGGCA-3'
	Antisense	5'-TGCCTCTTTCCTTCTTCTGTT -3'
T105A	Sense	5'-AACAGAAGAGCCAAAGAGGCA-3'
	Antisense	5'-TGCCTCTTTGGCTCTTCTGTT-3'
T105E	Sense	5'-AACAGAAGAGAAAAAGAGGCA-3'
	Antisense	5'-TGCCTCTTTTTCTCTTCTGTT-3'
F144S	Sense	5'-GGAAGAAGGTGTTCCATTGCAGGGTGG-3'
	Antisense	5'-CCACCCTGCAATGGAACACCTTCTTCC-3'
F144A	Sense	5'-GGAAGAAGGTGTGCCATTGCAGGGTGG-3'
	Antisense	5'-CCACCCTGCAATGGCACACCTTCTTCC-3'
E173K	Sense	5'-GTGGACCAGGATAAGTGCCAGCGTCTC-3'
	Antisense	5'-GAGACGCTGGCACTTATCCTGGTCCAC-3'
E173A	Sense	5'-GAGACGCTGGCACTTATCCTGGTCCAC-3'
	Antisense	5'-GAGACGCTGGCACGCATCCTGGTCCAC-3'
P193E	Sense	5'-TGTGCTGGATATGAAGAAGGCGGAGTT-3'
	Antisense	5'-AACTCCGCCTTCTTCATATCCAGCACA-3'
P193A	Sense	5'-TGTGCTGGATATGCTGAAGGCGGAGTT-3'
	Antisense	5'-AACTCCGCCTTCAGCATATCCAGCACA-3'