J. Biochem. 121, 1010-1017 (1997)

Pepsinogens and Pepsins from House Musk Shrew, *Suncus murinus*: Purification, Characterization, Determination of the Amino-Acid Sequences of the Activation Segments, and Analysis of Proteolytic Specificities[1]

Yuichi Narita,* Sen-ichi Oda,* Akihiko Moriyama,† Osamu Takenaka,‡ and Takashi Kageyama‡,2

*Laboratory of Animal Management, School of Agricultural Sciences, Nagoya University, Nagoya 464-01; †Division of Biomolecular Science, Institute of Natural Sciences, Nagoya City University, Nagoya 467; and ‡Department of Cellular and Molecular Biology, Primate Research Institute, Kyoto University, Inuyama, Aichi 484

Received for publication, October 24, 1996

Three pepsinogens, namely, pepsinogens A, C-1, and C-2, were purified from gastric mucosa of adult house musk shrew (*Suncus murinus*) by conventional chromatographic and gel filtration procedures. The molecular masses were 40, 39, and 41 kDa for pepsinogens A, C-1, and C-2, respectively. Pepsinogen C-2 contains an Asn-linked carbohydrate chain(s) of about 2 kDa. Each pepsinogen was converted to pepsin through an intermediate form under acidic conditions. By NH_2-terminal sequence analysis of these protein species, the amino acid sequences of activation segments (proparts) of pepsinogens A and C-1 were determined to be LYKVPLVKKKSLRQNLIENGLLKDFLAKHNVNPASKYFPTE and KVTKVTLKKF-KSIRENLREQGLLEDFLKTNHYDPAQKYHFGDF, respectively. The similarity of these two sequences is nearly 50%. Each pepsin cleaved preferentially peptide bonds between hydrophobic and aromatic amino acids, or bonds on either side of these amino acids. Although each activation segment had several sites susceptible to pepsin action, activation proceeded by limited cleavages of the segment, presumably due to the steric inflexibility of the segment in native pepsinogen. The activity of pepsin A was inhibited completely in the presence of a more than equimolar amount of pepstatin, while a hundred-molar excess amount of pepstatin was needed for the complete inhibition of the activity of pepsins C-1 and C-2.

Key words: activation segment, house musk shrew, pepsin, pepsinogen, proteolytic specificity.

Pepsinogens are zymogens of pepsins, the gastric aspartic proteinases in vertebrates. They have been purified from the gastric mucosa of various animals and the presence of multiple forms has been demonstrated (*1, 2*). Currently, pepsinogens are classified into four groups, namely, pepsinogens A and pepsinogens C (progastricsins) in adult animals, and prochymosins and pepsinogens F in neonatal animals (*1, 3*). Pepsinogens A and C have been shown to have diverged from a common ancestor during the early period of vertebrate evolution (*4*). Their relative levels in gastric mucosa are different between vertebrates especially between mammals. Both type-A and type-C pepsinogens are synthesized in primates such as man (*5*) and Asian macaque monkeys (*6*), and artiodactyls such as cow (*7, 8*) and pig (*9*). However, type-A pepsinogens are nearly the only components of pepsinogens of rabbit (order Lagomorpha) (*10*) and Asiatic black bear (order Carnivora) (*11*), and type-C pepsinogens are the exclusive components of pepsinogens of mouse (*12*), rat (*13*), and guinea pig (*14*) (order Rodentia). Since these variations in the expression of different types of pepsinogens are thought to be correlated to the phylogeny of mammals, it is of interest to investigate pepsinogens from more kinds of mammals.

Various types of pepsinogens have been shown to be activated to the respective pepsins autocatalytically at acidic pH. The release of the activation segment (the propart of pepsinogen whose length ranges from 35 to 47 residues among pepsinogens) from the N-terminal part of pepsinogen is the major process of activation. Very limited cleavages of the activation segment are common in different types of pepsinogens (*15-20*). Although the cleavage sites of pepsinogen on activation have been thought to be related to the hydrolytic specificity of the pepsin moiety of each pepsinogen, this relationship has not been clarified in detail.

We were interested in investigating pepsinogens of house musk shrew and in comparing them with those from other mammals, since house musk shrew belongs to the order Insectivora, which is generally accepted as a group of mammals retaining many primitive characters (*21*) and is thought to be a group central to the question of higher

[1] This work was supported in part by a Grant-in-Aid (No. 07640900) from the Ministry of Education, Science, Sports and Culture of Japan.
[2] To whom correspondence should be addressed. Tel: +81-568-63-0578, Fax: +81-568-63-0085, E-mail: kageyama@smtp.pri.kyoto-u.ac.jp
Abbreviation: FPLC, fast protein liquid chromatography.

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

relationships among placental mammals (*22*). Here we report the occurrence of both type-A and C pepsinogens in house musk shrew. Since shrew pepsinogen A released its activation segment by cleavages at different positions from those of other mammalian pepsinogens, we analyzed these cleavage sites in detail, focusing on their relation to the proteolytic specificity of pepsin. Amino acid compositions, NH_2-terminal amino acid sequences, and some enzymatic properties are also described.

MATERIALS AND METHODS

Materials—House musk shrews (*Suncus murinus*) of the laboratory KAT line were obtained from the Research Institute of Environmental Medicine, Nagoya University. This line was derived from a wild population in Katmandu, Nepal (*23*). Stomachs were removed from five adult individuals immediately after death by exsanguination under deep anesthesia and were stored at $-20^\circ C$ until use. Bovine hemoglobin substrate powder was purchased from Worthington Diagnostic Systems (Freehold, NJ); oxidized B chain of insulin from Sigma (St. Louis, MO); pepstatin from Peptide Institute (Minoh); DEAE-Sephacel, Sephadex G-100, and a Mono Q column HR 5/5 were from Pharmacia LKB Biotech. (Uppsala, Sweden); polyvinylidene fluoride membrane (Immobilon-P^{SQ}) from Millipore (Bedford, MA); *N*-glycanase from Genzyme (Cambridge, MA), and a silver staining kit from Wako Pure Chem. (Osaka). Fragments of the activation segment of type-A pepsinogen were synthesized by Bio-Synthesis (Lewisville, TX). Other chemicals were of reagent grade.

Assay of Proteolytic Activity—Potential pepsin activity of pepsinogen, as well as pepsin activity, was determined at pH 2.0 and 37°C with a solution of about 2% hemoglobin as the substrate, by the method of Anson and Mirsky (*24*). One unit of enzymatic activity was defined as the amount of enzyme that catalyzed an increase of 1.0 in the absorbance at 280 nm per minute under the assay conditions. The pH-dependence of activity was determined in a similar manner in the range of pH 1.0 to 5.0 (adjusted with 0.2 N HCl). The final chloride concentration in each assay mixture was adjusted to 0.1 M by addition of NaCl. The influence of pepstatin on the digestion of hemoglobin at pH 2.0 was examined by similar procedures to those described previously (*25*).

Purification of Pepsinogens—All procedures were performed at 0-4°C except for FPLC, which was performed at room temperature. Chromatography on DEAE-Sephacel and gel filtration were carried out in 0.01 M sodium phosphate buffer, pH 7.0. FPLC was carried out in 0.01 M Tris buffer, pH 7.0.

Step 1. Preparation of the crude homogenate supernatant: Five frozen stomachs (total weight, 2.9 g) of adult shrews were thawed in running water and washed with 0.9% NaCl. The mucosa (total weight, 1.4 g) was stripped from the muscle and homogenized with 19 ml of 0.01 M sodium phosphate buffer, pH 7.0. The homogenate was centrifuged at $15{,}000\times g$ for 50 min. The supernatant was used for further purification.

Step 2. Chromatography on DEAE-Sephacel: The supernatant was applied to a column (1.5 cm × 30 cm) of DEAE-Sephacel. Potential pepsin activity was not found in the flow-through fractions. The adsorbed proteins were eluted with a 1-liter linear gradient of 0-0.5 M NaCl. Proteolytic activities against hemoglobin were detected as two peaks and a small shoulder following the second peak. The first peak was a mixture of a type-A pepsinogen and a type-C pepsinogen (pepsinogen C-2). The second peak mainly consisted of a type-C pepsinogen (pepsinogen C-1). A proteinase(s) in the shoulder was not purified because of its small amount, although some procedures for purification were examined.

Step 3. Gel filtration: Fractions of each peak were combined separately and dialyzed against saturated ammonium sulfate solution adjusted to pH 7.0. The precipitated protein was dissolved in a small volume of 0.01 M sodium phosphate buffer, pH 7.0, and then gel-filtered on a column (1.5 cm × 100 cm) of Sephadex G-100. Fractions with proteolytic activity in each gel filtration were combined separately.

Step 4. FPLC: The protein solutions were subjected to FPLC on a Mono Q column HR 5/5 for final purification. Proteins were eluted with a linear gradient of NaCl from 0 to 0.5 M over the course of 35 min at a flow rate of 1.0 ml/min. The mixture of pepsinogens A and C-2 was separated completely by FPLC.

Electrophoresis of Proteins—Aliquots of the purified pepsinogens were subjected to non-denaturing PAGE essentially according to the methods of Ornstein (*26*) and Davis (*27*), and SDS-PAGE according to Laemmli (*28*). Proteins were stained with Coomassie Brilliant Blue and/or silver staining. Molecular masses of proteins on SDS-PAGE were determined with the following proteins as standards: rabbit muscle phosphorylase *b* (97.4 kDa), bovine serum albumin (66.2 kDa), chicken ovalbumin (42.7 kDa), bovine carbonic anhydrase (31.0 kDa), soybean trypsin inhibitor (21.5 kDa), and chicken lysozyme (14.4 kDa).

Deglycosylation of Pepsinogens—Each pepsinogen (*ca.* 0.2 μg) was incubated with 0.3 unit of *N*-glycanase for 18 h at 37°C in 30 μl of 0.01 M sodium phosphate buffer, pH 7.0, containing 0.5% SDS, 50 mM 2-mercaptoethanol, and 1.3% Triton X-100. The decrease in molecular mass of pepsinogen was analyzed by SDS-PAGE.

Activation of Pepsinogens and Isolation of Intermediate Forms and Pepsins—The procedure for activation was the same as that used by Kageyama and Takahashi for Japanese monkey pepsinogens (*15*). To determine the N-terminal sequences of pepsin and the intermediate form of pepsinogen and pepsin, the activation reaction was stopped at an appropriate time by the addition of one-fourth volume of 0.25 M Tris-HCl buffer, pH 8.0, containing 10% SDS, 2.8 M 2-mercaptoethanol, and 40% glycerol. The mixture was subjected to SDS-PAGE, and proteins were transferred to a polyvinylidene fluoride membrane. The bands of intermediate forms and pepsins were cut out of the membrane and subjected to NH_2-terminal amino acid sequence determination.

Hydrolysis of Peptides—Hydrolysis of oxidized insulin B chain and synthetic activation peptides by pepsins was carried out as follows. The reaction mixture contained 0.2 M NaCl-HCl buffer, pH 2.0, 50 μM peptide, and 24 μg pepsin. The total volume was 200 μl. The reaction mixture was incubated at 37°C for 1 h, and the reaction was stopped by the addition of 600 μl of 3% perchloric acid. Products of hydrolysis were separated by high-pressure liquid chro-

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

matography on ODS-120T (Tosoh, Tokyo) and subjected to amino acid analysis to determine the sites of cleavage.

Determination of Protein Concentration—The concentrations of protein in the solutions of the enzymes at each step in the purification were determined by measuring the absorbance at 280 nm with bovine serum albumin as a standard. The amount of each purified pepsinogen and pepsin was determined by using the molar extinction coefficient calculated from the amino acid composition and the molecular mass.

Amino Acid Analysis—Samples for amino acid analysis (10 μg of protein) were hydrolyzed under HCl vapor at 150°C for 1 h in a Waters PICO-TAG™ Work Station (Millipore). The samples were analyzed with an amino acid analyzer (model 835; Hitachi, Tokyo).

Amino Acid Sequence Determination—NH_2-terminal amino acid sequence was determined by using an automatic protein sequencer (model 477A; Applied Biosystems, Foster City, CA, USA).

TABLE I. **Purification of house musk shrew pepsinogens.**

Step	Protein (mg)	Activity (units)	Specific activity (units/mg protein)	Yield (%)
1. Supernatant of crude homogenate	81	208	2.6	100
2. DEAE-Sephacel				
Pepsinogens A+C-2	6.4	87	14	42
Pepsinogen C-1	3.7	56	15	27
3. Sephadex G-100				
Pepsinogens A+C-2	3.6	71	20	34
Pepsinogen C-1	1.7	31	18	15
4. FPLC				
Pepsinogen A	1.5	31	21	15
Pepsinogen C-1	0.61	15	25	7.2
Pepsinogen C-2	0.05	1.4	28	0.67

RESULTS

Purification of Pepsinogens—The result of purification is summarized in Table I. The DEAE-Sephacel chromatography separated the proteolytic activity into 2 major peaks and a small shoulder following the second peak (Fig. 1). The first peak contained pepsinogen A and a minor component of pepsinogen C (pepsinogen C-2). The second peak contained the major component of pepsinogen C (pepsinogen C-1). A proteinase(s) contained in the shoulder was not purified, but it seemed to be procathepsin E, as deduced from the apparent molecular mass of 80 kDa on Sephadex G-100 gel filtration and the affinity to Con-A Sepharose. Final purification of pepsinogens A and C was achieved by FPLC on a Mono Q column after gel-filtration on Sephadex G-100 (Fig. 2). Pepsinogens A and C-2 were separated completely by FPLC. Thus, 3 pepsinogens, namely pepsinogens A, C-1, and C-2 were obtained. The relative levels of pepsinogens A, C-1, and C-2 based on weight in the gastric mucosa were calculated to be 63, 35, and 2%, respectively, from the distribution of their activities in DEAE-Sephacel chro-

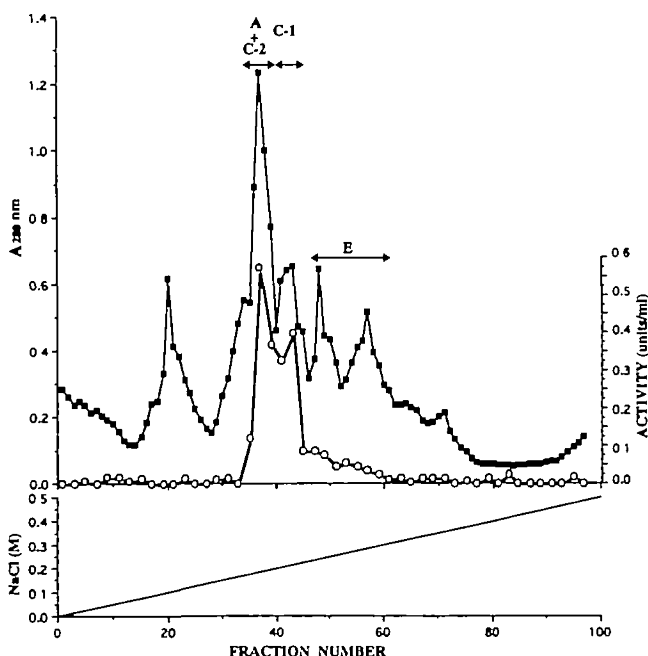

Fig. 1. **Chromatography of crude homogenate supernatant of house musk shrew gastric mucosa on a column (1.5 cm×30 cm) of DEAE-Sephacel.** The column was equilibrated with 0.01 M sodium phosphate buffer, pH 7.0, and proteins were eluted with a linear gradient of NaCl in the same buffer. Fractions of 10 ml were collected. The fractions under the bars were pooled separately. ○, proteolytic activity; ■, absorbance at 280 nm. A, C-1, C-2, and E stand for pepsinogens A, C-1, C-2, and cathepsin E-like proteinase, respectively.

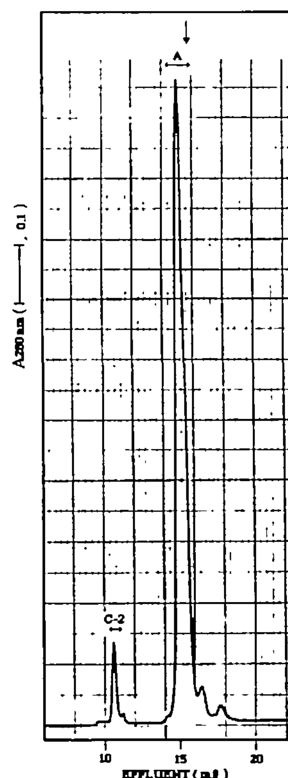

Fig. 2. **FPLC on a Mono Q column HR 5/5 for the final purification of the mixture of pepsinogens A and C-2.** The column was equilibrated and eluted with 0.01 M Tris-HCl buffer, pH 8.0, and proteins were eluted with a linear gradient of NaCl in the same buffer. The flow rate was 1.0 ml/min. Abbreviations are the same as those in Fig. 1. The downward-pointing arrow indicates the elution position of pepsinogen C-1 under the same conditions.

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

matography and FPLC and their specific activities in Table I. Each purified pepsinogen was homogeneous on non-denaturing (Fig. 3) and denaturing PAGE (Fig. 4).

Molecular Masses—The molecular masses of pepsinogens A, C-1, and C-2 were determined to be 40, 39, and 41 kDa, respectively, by SDS-PAGE. These molecular masses were the same under reducing and non-reducing conditions. The molecular mass of pepsinogen C-2 was reduced to 39 kDa after treatment with *N*-glycanase, while those of pepsinogens A and C-1 did not change after similar treatment, showing that pepsinogen C-2 has an Asn-linked carbohydrate chain(s) of about 2 kDa (Fig. 4).

Amino Acid Compositions—The amino acid compositions of the three pepsinogens are given in Table II. Relatively high levels of Asx and Glx in all of them were consistent with the fact that they are acidic proteins. The composition of pepsinogen A was different from those of pepsinogens C-1 and C-2, which resembled each other. The most remarkable differences were found in the ratios of Glx/Asx and Leu/Ile. Their values were higher in pepsinogens C-1 and C-2 than in pepsinogen A. These differences are characteristic of pepsinogens A and C from other animal sources, such as Japanese monkey (shown in Table II for comparison). The difference in the levels of basic residues such as Lys between pepsinogens C-1 and C-2 was appreciable.

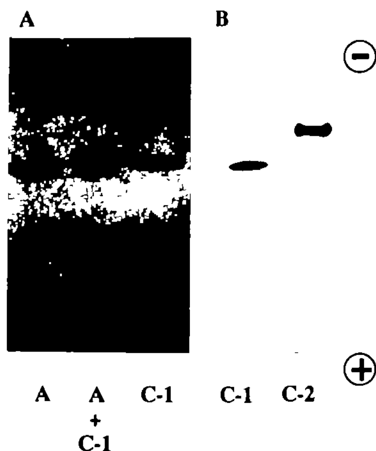

Fig. 3. **Non-denaturing PAGE of purified pepsinogens.** The concentration of the gel was 10%. Tris-glycine buffer, pH 8.3, was used for electrophoresis. The bottom is the position of bromphenol blue. An amount corresponding to 2 μg (A) or 0.2 μg (B) of protein was loaded in each lane. Protein was stained with Coomassie Brilliant Blue R-250 (A) and by the method of silver staining (B). Abbreviations are the same as those in Fig. 1.

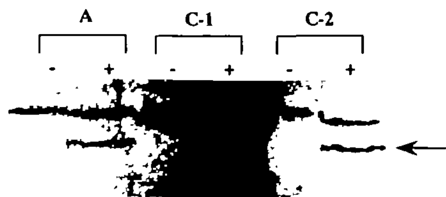

Fig. 4. ***N*-Glycanase treatment of purified pepsinogens.** Each pepsinogen (0.1 μg) was digested with *N*-glycanase and the products were subjected to SDS-PAGE. The gel was silver-stained. The bands shown by the arrow are due to *N*-glycanase. − and + show samples non-digested and digested with glycanase, respectively. Abbreviations are the same as those in Fig. 1.

Activation of Pepsinogens and Structure of Activation Peptides—The activation process of each pepsinogen was analyzed by SDS-PAGE (Fig. 5). Each pepsinogen was converted to an intermediate form in a few minutes, and then gradually converted to pepsin. The rate of conversion

TABLE II. **Amino acid compositions of purified house musk shrew pepsinogens.** Amino acid compositions of Japanese monkey pepsinogens A and C derived from the amino acid sequences (*44, 45*) are shown for comparison.

Amino acid	Number of residues per molecule[a]				
	House musk shrew			Japanese monkey	
	A	C-1	C-2	A	C
Asp+Asn	45.6	33.0	34.2	24+16	13+17
Thr	24.9	32.0	27.5	26	27
Ser	44.3	29.9	31.7	46	39
Glu+Gln	38.6	48.6	48.7	14+19	18+26
Pro	16.2	18.5	15.9	21	19
Gly	38.7	40.9	43.5	36	37
Ala	23.3	20.9	26.3	20	22
Cys	*6	*6	*6	6	6
Val	17.9	15.8	21.7	27	28
Met	5.4	7.5	10.4	4	7
Ile	24.6	14.2	12.9	29	16
Leu	29.6	33.9	29.6	30	33
Tyr	16.4	21.1	17.1	18	23
Phe	18.2	27.5	20.9	17	21
Lys	8.8	7.4	10.5	7	8
His	1.7	3.4	3.2	3	3
Arg	4.8	4.4	4.9	5	5
Trp	*5	*5	*5	5	6
Total	*370	*370	*370	373	374
Glx/Asx	0.85	1.47	1.42	0.83	1.47
Leu/Ile	1.20	2.39	2.30	1.03	2.06

[a]The values were calculated on the assumption that the total number of residues was 370 and the numbers of Cys and Trp were 6 and 5, respectively. Each assumed value is marked with an asterisk. The Thr and Ser values were corrected for losses of 13 and 17%, respectively.

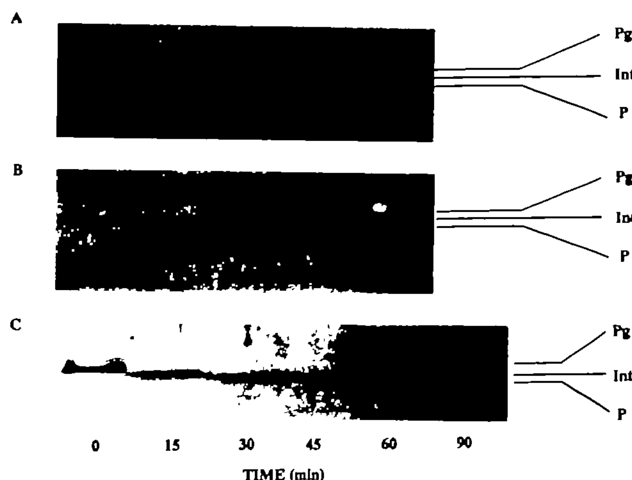

Fig. 5. **Time course of pepsinogen activation analyzed by SDS-PAGE.** Pepsinogens (0.1 mg/ml in the case of pepsinogens A and C-1, and 0.01 mg/ml in the case of pepsinogen C-2) were activated at 14°C, and pH 2.0. (A) pepsinogen A, (B) pepsinogen C-1, and (C) pepsinogen C-2. Pg, pepsinogen; P, pepsin; Int, intermediate form. When pepsinogen C-1 was activated at the initial concentration of 0.01 mg/ml, the result was essentially the same as that in (C).

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

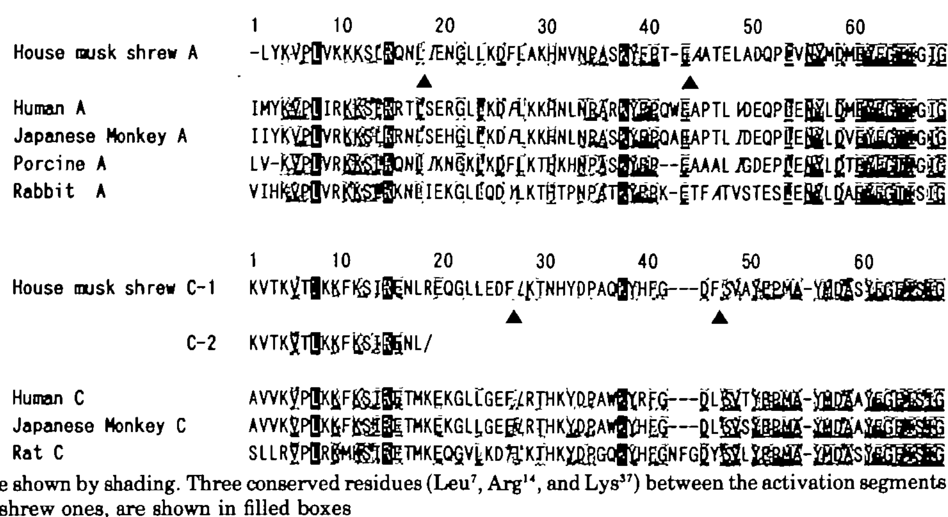

Fig. 6. **NH_2-terminal amino acid sequences of house musk shrew pepsinogens A and C.** Amino acid sequences for human pepsinogens A (*30*) and C (*32*), Japanese monkey pepsinogens A (*44*) and C (*45*), porcine pepsinogen A (*46*), rabbit pepsinogen A (*10*), and rat pepsinogen C (*47*) are shown for comparison. The numberings of pepsinogens A and C are based on the sequences of human pepsinogen A and rat pepsinogen C, respectively. – stands for a deleted residue. The italic letters show the NH_2-terminal residues of intermediate forms and pepsins. The arrowheads indicate the positions of cleavage sites for activation of house musk shrew pepsinogens. Common residues among these pepsinogens in each group are shown by shading. Three conserved residues (Leu[7], Arg[14], and Lys[37]) between the activation segments of known animal pepsinogens, including shrew ones, are shown in filled boxes

increased when the initial concentration was increased. This showed the involvement of the intermolecular reaction at a high concentration of pepsinogen. The NH_2-terminal amino acid sequences of pepsinogen A, the intermediate form, and pepsin A were determined to be LYKVPLVKK-KSLRQNLIEN-, IENGLLKDFLAKHNVNPASKYFPTE-AA-, and AATELADQPLVNYMDMEYFGTIGIG-, respectively. The sequences of pepsinogen C-1, the intermediate form, and pepsin C-1 were KVTKVTLKKFKSIRENLRE-QGLLEDFLK-, LKTNHYDPAQKYHFGDFSVAYE-, and SVAYEPMAYMDASYFGEISIG-, respectively. The sequence of pepsinogen C-2 was KVTKVTLKKFKSIRE-NL-. From these results, the complete sequences of the activation segments in pepsinogens A and C-1 were elucidated (Fig. 6).

Enzymatic Properties of Pepsins—Hemoglobin digestion: House musk shrew pepsins A and C were active at low pH values, like those from other animal sources. Optimal pHs were about 2.3 for pepsin A and about 3.0 for pepsins C-1 and C-2 (Fig. 7A). Specific activity of pepsin A was about half of those of type-C pepsins in the range of pH 1-4.

Hydrolytic specificity for insulin B chain and synthetic activation peptides: Oxidized insulin B chain was hydrolyzed by pepsins A and C-1 at various sites (Fig. 8). The sites of cleavage by both pepsins were almost the same, although the rates of cleavage at some of these sites differed considerably. The P1 and P1′ positions of the cleavage sites were occupied mainly by hydrophobic and aromatic amino acids. Two synthetic activation peptides of pepsinogen A, namely, peptides 11-30 and 33-51, respectively, were prepared and subjected to hydrolysis by pepsins A and C-1. They were cleaved at various sites (Fig. 8). Hydrophobic and aromatic amino acids were located at the P1 and P1′ positions or either side of these positions. The cleavage of activation peptide 11-30 by pepsin A was less specific than by pepsin C-1, and several bonds were cleaved, including the Leu[16]-Ile[17] bond, which was the cleavage site on activation of native pepsinogen A. It is noteworthy that pepsin C-1 did not cleave the Leu[16]-Ile[17] bond. The cleavage of activation peptide 33-51 by pepsins A and C-1 occurred mainly at two sites. Although the cleavage of the Glu[45]-Leu[46] bond was common to both pepsins, the cleavage of the Glu[41]-Ala[42] bond, which was the cleavage site on activation of native pepsinogen A, was specific to pepsin A. These results showed that activation segments had various potential sites of cleavage, although the cleavage sites on activation of native pepsinogen are highly restricted. The results also showed that the cleavage sites on activation of native pepsinogen A were susceptible to the action of pepsin A, but not pepsin C.

Inhibition by pepstatin: House musk shrew pepsins were inhibited by pepstatin (Fig. 7B). The sensitivities to pepstatin were very different between pepsin A and type-C pepsins. Pepsin A was inhibited almost completely in the presence of over an equimolar amount of pepstatin, while type-C pepsins needed about 100-fold molar excess of pepstatin for complete inhibition.

DISCUSSION

One type-A pepsinogen and two type-C pepsinogens were purified from the gastric mucosa of house musk shrew. The expressional pattern in house musk shrew was similar to those in primates (*5, 6*) and artiodactyls (*8, 9*). Since house musk shrew is thought to retain primitive characters of mammals, its expressional pattern of pepsinogen might reflect the expressional pattern of the ancestral mammal. Expressional patterns of pepsinogens in rabbit (*3, 10*) and rodents (*12-14*) are thought to have shifted to the extremes after the divergence of these mammals from the common ancestor. Although rabbit and rodents are used widely in experimental gastric physiology, house musk shrew might be a more suitable animal for such studies, especially as a model of man, considering the similarity of expressional patterns of pepsinogens between house musk shrew and man. Since house musk shrew has already been established as an experimental animal and its breeding is easy (*23*), its utilization in place of rabbit and rodents can be recommended.

Multiple isozymogens have been found in both type-A and type-C pepsinogens (*1, 2*). The multiplicity is quite extreme in type-A pepsinogen. For example, the occurrence of five, four, and six pepsinogens A has been reported in man (*29*), monkeys (*6*), and rabbit (*10*), respectively. Although some of these isozymogens are generated by post-translational modifications such as phosphorylation,

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

Fig. 7. **(A) Effect of pH on the activity of purified pepsins against hemoglobin in the range of pH 1 to 5. (B) The inhibitory effect of pepstatin on purified pepsins.** ●, pepsin A; ▲, pepsin C-1; □, pepsin C-2.

Insulin B chain	FVNEHLC.GSHL VEAL Y LVC.GERGF F YTPKA
Pepsin A	
Pepsin C-1	
PgA Peptide 11-30	SLRQNL I EN GLLKD F L AKHN
Pepsin A	
Pepsin C-1	
PgA Peptide 33-51	PASKYFPTE A ATE LADQPL
Pepsin A	
Pepsin C-1	

Fig. 8. **Cleavage sites of oxidized insulin B chain and synthetic activation peptides by pepsins A and C-1.** These peptides were cleaved at pH 2.0. The conditions of the cleavage are described in detail in "MATERIALS AND METHODS." The cleavages that generate a yield of more than 20% are shown as major cleavage sites (↑). Others are shown as secondary cleavage sites (↑).

most of them are known to be products of different genes. Multiple genes for type-A pepsinogens have been shown to be generated by gene duplications (*30, 31*). In contrast to these multiplicities of type-A pepsinogens, house musk shrew expresses only one type-A pepsinogen, although the possibility cannot be excluded completely that a different type-A pepsinogen(s) is synthesized in a small amount and could not be detected in the present study. This simplest expressional pattern might suggest that gene duplication events did not occur during the evolution of house musk shrew. Since type-A pepsinogen is the major type of pepsinogen in mammals except for rodents, house musk shrew should be very useful for studying the expressional regulation of the gene for type-A pepsinogen.

The multiplicities of type-C pepsinogens in mammals are less frequent than those of type-A pepsinogens. Two pepsinogens C have been reported in man (*20*), although only one gene for pepsinogen C is known (*32, 33*). In house musk shrew, two type-C pepsinogens, namely, pepsinogens C-1 and C-2, were also found, although pepsinogen C-1 was expressed highly predominantly. They had slightly different amino acid compositions, and pepsinogen C-2 showed more basic character than pepsinogen C-1, presumably due to its relatively high content of Lys. This difference might suggest that the two type-C pepsinogens in house musk shrew are products of different genes. Although pepsinogen C-2 was a minor component, it is unusual in being a carbohydrate-containing pepsinogen. To date, glycosylation has been found in Japanese monkey pepsinogen A-4 (*34*), chicken pepsinogen A (*35*) and prochymosin (*36*), turtle pepsinogen A (*37*), and frog pepsinogen C (*38*). In monkey and chicken pepsinogens A, carbohydrate chains are attached to Asn^{73} and Asn^{113}, respectively. House musk shrew pepsinogen C-2 is the second example of a carbohydrate-containing type-C pepsinogen.

We should discuss the structural features of the activation segment and the cleavage sites on activation. The segment has a universal and essential role in stabilizing the pepsin moiety at neutral pH, and the involvement of basic residues has been anticipated (*39, 40*). The common residues of the activation segments of known animal pepsinogens except for shrew pepsinogens are restricted to 4 residues, namely, Pro^6, Leu^7, Arg^{14}, and Lys^{37} (the residue numbering is based on the sequence of human pepsinogen A, see Fig. 6) (*2, 3*). Lys^{37} is especially important for providing electrostatic stabilization through hydrogen bonding for the negative charges on active-site aspartic acids (*39, 40*), and this residue is conserved in shrew pepsinogens. A remarkable substitution of Pro→Thr at position 6 was found in shrew pepsinogens C (Fig. 6). Since the Pro^6-Leu^7 dipeptide participates in the β-structure of the NH_2-terminal sequence of pepsinogen to interact with the pepsin moiety (*41*), this substitution might cause the β-structure of shrew pepsinogens C to be more rigid than those of other animal pepsinogens although the precise effect of this substitution is unclear at present. Thus, completely conserved residues among the activation segments of pepsinogens, including shrew ones, are reduced to 3 residues, namely, Leu^7, Arg^{14}, and Lys^{37}, and they are

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

thought to be essential core residues of the activation segments. Pepsin and the intermediate form between pepsinogen and pepsin were shown to be generated by the specific cleavage of the activation segment of pepsinogen. The cleavage of the Leu^{16}-Ile^{17} bond of pepsinogen A is the same as those of porcine (*17*) and bovine (*42*) pepsinogens A, but different from those of pepsinogens A from other sources such as monkey (*15*) and man (*18, 19*). The site of cleavage to generate pepsin A is the Glu^{41}-Ala^{42} bond. This cleavage causes the NH_2-terminal sequence to be a few residue longer in shrew pepsin A than in other mammalian pepsins A. The cleavage sites that generate pepsin C and the intermediate form between pepsinogen C and pepsin C are Phe^{43}-Ser^{44} and Phe^{26}-Leu^{27}, respectively, and these sites are similar to those in human (*20*) and monkey (*4*) pepsinogens C. Since the cleavage sites on activation are highly restricted and, moreover, are different between shrew and other animal pepsinogens A, we examined how the cleavage sites are determined in pepsinogens. From the results on the hydrolysis of synthetic activation peptides of shrew pepsinogen A, the cleavage sites of pepsinogen A are susceptible to the action of pepsin A, but not pepsin C. This shows that the specificity of pepsin is an important factor to determine the cleavage sites. Synthetic activation peptides, however, were cleaved at various other sites which were scarcely cleaved on activation of native pepsinogen A. This suggests that steric inflexibility of the activation segment of native pepsinogen might restrict the access of its own active site to the segment, and, thus, result in the limited cleavages of the segment. Therefore, the tertiary structure of native pepsinogen might be another important factor to determine the cleavage sites on activation.

Comparison of the NH_2-terminal amino acid sequences of pepsinogens between house musk shrew and other mammals is informative concerning the phylogeny of house musk shrew, an insectivore. The numbers of amino acid substitutions of the activation segments of pepsinogens A between house musk shrew and other mammals ranged from 12 to 18, and those of pepsinogens C from 14 to 19 (Fig. 6). These values are similar to those reported previously between mammalian orders except for Insectivora (*4*). The present results support the recent proposal that Insectivora should be recognized as an animal group which should be dealt with using the same principle as other mammalian orders, and not as a particular group containing the ancestors of some other orders (*21*). Molecular evolutional analyses based on the nucleotide sequences of type-A and C pepsinogens should give clearer results on the phylogeny of house musk shrew, and are in progress in our laboratory.

Finally, we should mention the characteristics of enzymatic properties of house musk shrew pepsinogens and pepsins compared with those of other mammals. The specific activity of pepsin A for hemoglobin digestion was similar to those of monkey enzymes (*6*). The two-fold higher specific activities of pepsins C-1 and C-2 than that of pepsin A are consistent with the results reported in man (*18*) and monkey (*25*). This may show that type-C pepsins have high hydrolytic activities toward hemoglobin. Alternatively, since the hemoglobin digestion method is essentially based on the measurement of the content of Tyr and Trp in trifluoroacetic acid-soluble peptides of the digest (*24*), type-C pepsins might have high hydrolytic specificities to release peptides containing these amino acids. High susceptibility of pepsin A to pepstatin and the 100-fold lower susceptibility of pepsin C are also common to pepsins A and C of other mammals (*25*). Different susceptibilities to pepstatin are convenient to distinguish mammalian pepsins A and C, as was the case in the present study. These different susceptibilities as well as different hydrolytic specificities between type-A and C pepsins can presumably be rationalized in terms of the structural differences between the active sites of these pepsins (*43*).

REFERENCES

1. Foltmann, B. (1981) Gastric proteinases—structure, function, evolution and mechanism of action. *Essays Biochem.* **17**, 52-84
2. Takahashi, K., Tanji, M., Yakabe, E., Hirasawa, A., Athauda, S. B.P., and Kageyama, T. (1995) Non-mammalian vertebrate pepsinogens and pepsins: isolation and characterization in *Aspartic Proteinases: Structure, Function, Biology, and Biomedical Implications* (Takahashi, K., ed.) pp. 53-65, Plenum Press, New York
3. Kageyama, T., Tanabe, K., and Koiwai, O. (1990) Structure and development of rabbit pepsinogens. Stage-specific zymogens, nucleotide sequences of cDNAs, molecular evolution, and gene expression during development. *J. Biol. Chem.* **265**, 17031-17038
4. Kageyama, T. and Takahashi, K. (1985) Monkey pepsinogens and pepsins. VII. Analysis of the activation process and determination of the NH_2-terminal 60-residue sequence of Japanese monkey progastricsin, and molecular evolution of pepsinogens. *J. Biochem.* **97**, 1235-1246
5. Tang, J. (1970) Gastricsin and pepsin. *Methods Enzymol.* **19**, 406-421
6. Kageyama, T. (1994) Asian macaque pepsinogens and pepsins. *J. Med. Primatol.* **23**, 375-381
7. Chow, R.B. and Kassell, B. (1968) Bovine pepsinogen and pepsin. Isolation, purification, and some properties of the pepsinogen. *J. Biol. Chem.* **243**, 1718-1724
8. Martin, P., Trieu-Cuot, P., Collin, J-C., and Ribadeau Dumas, B. (1982) Purification and characterization of bovine gastricsin. *Eur. J. Biochem.* **122**, 31-39
9. Ryle, A.P. (1970) The porcine pepsins and pepsinogens. *Methods Enzymol.* **19**, 316-336
10. Kageyama, T. and Takahashi, K. (1984) Rabbit pepsinogens. Purification, characterization, analysis of the conversion process to pepsin and determination of the NH_2-terminal amino-acid sequences. *Eur. J. Biochem.* **141**, 261-269
11. Kageyama, T., Moriyama, A., and Takahashi, K. (1983) Purification and characterization of pepsinogens and pepsins from Asiatic black bear, and amino acid sequence determination of the NH_2-terminal 60 residues of the major pepsinogen. *J. Biochem.* **94**, 1557-1567
12. Esumi, H., Sato, S., Sugimura, T., and Furihata, C. (1978) Purification of mouse pepsinogens by pepstatin-affinity chromatography. *FEBS Lett.* **86**, 33-36
13. Furihata, C., Saito, D., Fujiki, H., Kanai, Y., Matsushima, T., and Sugimura, T. (1980) Purification and characterization of pepsinogens and a unique pepsin from rat stomach. *Eur. J. Biochem.* **105**, 43-50
14. Kageyama, T., Ichinose, M., Tsukada, S., Miki, K., Kurokawa, K., Koiwai, O., Tanji, M., Yakabe, E., Athauda, S.B.P., and Takahashi, K. (1992) Gastric procathepsin E and progastricsin from guinea pig. Purification, molecular cloning of cDNAs, and characterization of enzymatic properties, with special reference to procathepsin E. *J. Biol. Chem.* **267**, 16450-16459
15. Kageyama, T. and Takahashi, K. (1982) Monkey pepsinogens and pepsins. VI. One-step activation of Japanese monkey pepsinogen to pepsin. *J. Biochem.* **92**, 1179-1188
16. Takahashi, K. and Kageyama, T. (1985) Multiplicity and intermediates of the activation mechanism of zymogens of gastric

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012

aspartic proteinases in *Aspartic Proteinases and Their Inhibitors* (Kostka, V., ed.) pp. 265-282, Walter de Gruyter & Co., Berlin
17. Kageyama, T. and Takahashi, K. (1983) Occurrence of two different pathways in the activation of porcine pepsinogen to pepsin. *J. Biochem.* **93**, 743-754
18. Kageyama, T. and Takahashi, K. (1980) Isolation of an activation intermediate and determination of the amino acid sequence of the activation segment of human pepsinogen A. *J. Biochem.* **88**, 571-582
19. Kageyama, T., Ichinose, M., Miki, K., Athauda, S.B., Tanji, M., and Takahashi, K. (1989) Difference of activation processes and structure of activation peptides in human pepsinogens A and progastricsin. *J. Biochem.* **105**, 15-22
20. Foltmann, B. and Jensen, A.L. (1982) Human progastricsin. Analysis of intermediates during activation into gastricsin and determination of the amino acid sequence of the propart. *Eur. J. Biochem.* **128**, 63-70
21. Shigehara, N. (1985) A historical review of the classification of mammalian order Insectivora (in Japanese with English summary) in *Suncus murinus, Biology of the Laboratory Shrew* (Oda, S., Kitoh, J., Ota, K., and Isomura, G., eds.) pp. 11-19, Japan Sci. Soc. Press, Tokyo
22. Novacek, M.J. (1992) Mammalian phylogeny: shaking the tree. *Nature* **356**, 121-125
23. Oda, S., Koyasu, K., and Shrestha, K.C. (1992) Breeding of the house musk shrew, *Suncus murinus* originating from a wild population in Katmandu, Nepal (in Japanese). *Ann. Res. Inst. Environ. Med.* **43**, 239-240
24. Anson, M.L. and Mirsky, A.E. (1932) The estimation of pepsin with hemoglobin. *J. Gen. Physiol.* **16**, 59-63
25. Kageyama, T. and Takahashi, K. (1976) Pepsinogen C and pepsin C from gastric mucosa of Japanese monkey. Purification and characterization. *J. Biochem.* **80**, 983-992
26. Ornstein, L. (1964) Disc electrophoresis. I. Background and theory. *Ann. N.Y. Acad. Sci.* **121**, 321-349
27. Davis, B.J. (1964) Disc electrophoresis. II. Method and application to human serum proteins. *Ann. N.Y. Acad. Sci.* **121**, 404-427
28. Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685
29. Athauda, S.B.P., Tanji, M., Kageyama, T., and Takahashi, K. (1989) A comparative study on the NH_2-terminal amino acid sequences and some other properties of six isozymic forms of human pepsinogens and pepsins. *J. Biochem.* **106**, 920-927
30. Sogawa, K., Fujii-Kuriyama, Y., Mizukami, Y., Ichihara, Y., and Takahashi, K. (1983) Primary structure of human pepsinogen gene. *J. Biol. Chem.* **258**, 5306-5311
31. Zelle, B., Evers, M.P.J., Groot, P.C., Bebelman, J.P., Mager, W. H., Planta, R.J., Pronk, J.C. Meuwissen, S.G.M., Hofker, M.H., Eriksson, A.W., and Frants, R.R. (1988) Genomic structure and evolution of the human pepsinogen A multigene family. *Hum. Genet.* **78**, 79-82
32. Hayano, T., Sogawa, K., Ichihara, Y., Fujii-Kuriyama, Y., and Takahashi, K. (1988) Primary structure of human pepsinogen C gene. *J. Biol. Chem.* **263**, 1382-1385
33. Pals, G., Azuma, T., Mohandas, T.K., Bell, G.I., Bacon, J., Samloff, I.M., Walz, D.A., Barr, P.J., and Taggart, R.T. (1989) Human pepsinogen C (progastricsin) polymorphism: evidence for a single locus located at 6p21.1-pter. *Genomics* **4**, 137-145
34. Kageyama, T. and Takahashi, K. (1978) Monkey pepsinogens and pepsins. III. Carbohydrate moiety of Japanese monkey pepsinogens and the amino acid sequence around the site of its attachment to protein. *J. Biochem.* **84**, 771-778
35. Baudys, M. and Kostka, V. (1983) Covalent structure of chicken pepsinogen. *Eur. J. Biochem.* **136**, 89-99
36. Hayashi, K., Agata, K., Mochii, M., Yasugi, S., Eguchi, G., and Mizuno, T. (1988) Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prochymosin. *J. Biochem.* **103**, 290-296
37. Hirasawa, A., Athauda, S.B.P., and Takahashi, K. (1996) Purification and characterization of turtle pepsinogen and pepsin. *J. Biochem.* **120**, 407-414
38. Yakabe, E., Tanji, M., Ichinose, M., Goto, S., Miki, K., Kurokawa, K., Ito, H., Kageyama, T., and Takahashi, K. (1991) Purification, characterization, and amino acid sequences of pepsinogens and pepsins from the esophageal mucosa of bullfrog (*Rana catesbeiana*). *J. Biol. Chem.* **266**, 22436-22443
39. James, M.N.G. and Sielecki, A.R. (1986) Molecular structure of an aspartic proteinase zymogen, porcine pepsinogen, at 1.8 Å resolution. *Nature* **319**, 33-38
40. Sielecki, A.R., Fedorov, A.A., Boodhoo, A., Andreeva, N.S., and James, M.N.G. (1990) Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution. *J. Mol. Biol.* **214**, 143-170
41. Hartsuck, J.A., Koelsch, G., and Remington, S.J. (1992) The high-resolution crystal structure of porcine pepsinogen. *Proteins: Struct. Funct. Genet.* **13**, 1-25
42. Harboe, M., Andersen, P.M., Foltmann, B., Kay, J., and Kassell, B. (1974) The activation of bovine pepsinogen. Sequence of the peptides released, identification of a pepsin inhibitor. *J. Biol. Chem.* **249**, 4487-4494
43. Fujinaga, M., Chernaia, M.M., Tarasova, N.I., Mosimann, S.C., and James, M.N.G. (1995) Crystal structure of human pepsin and its complex with pepstatin. *Protein Sci.* **4**, 960-972
44. Kageyama, T. and Takahashi, K. (1986) The complete amino acid sequence of monkey pepsinogen A. *J. Biol. Chem.* **261**, 4395-4405
45. Kageyama, T. and Takahashi, K. (1986) The complete amino acid sequence of monkey progastricsin. *J. Biol. Chem.* **261**, 4406-4419
46. Tsukagoshi, N., Ando, Y., Tomita, Y., Uchida, R., Takemura, T., Sasaki, T., Yamagata, H., Udaka, S., Ichihara, Y., and Takahashi, K. (1988) Nucleotide sequence and expression in *Escherichia coli* of cDNA of swine pepsinogen: involvement of the amino-terminal portion of the activation peptide segment in restoration of the functional protein. *Gene* **65**, 285-292
47. Ichihara, Y., Sogawa, K., Morohashi, K., Fujii-Kuriyama, Y., and Takahashi K. (1986) Nucleotide sequence of a nearly full-length cDNA coding for pepsinogen of rat gastric mucosa. *Eur. J. Biochem.* **161**, 7-12

Downloaded from http://jb.oxfordjournals.org/ at Peking University on October 2, 2012