CMLS, Cell. Mol. Life Sci. 59 (2002) 288–306
1420-682X/02/020288-19 $ 1.50 + 0.20/0

CMLS Cellular and Molecular Life Sciences

Review

Pepsinogens, progastricsins, and prochymosins: structure, function, evolution, and development

T. Kageyama

Center for Human Evolution Modeling Research, Primate Research Institute, Kyoto University, Inuyama 484-8506 (Japan), Fax +81 568 63 0085, e-mail: kageyama@pri.kyoto-u.ac.jp

Received 25 May 2001; received after revision 27 August 2001; accepted 30 August 2001

Abstract. Five types of zymogens of pepsins, gastric digestive proteinases, are known: pepsinogens A, B, and F, progastricsin, and prochymosin. The amino acid and/or nucleotide sequences of more than 50 pepsinogens other than pepsinogen B have been determined to date. Phylogenetic analyses based on these sequences indicate that progastricsin diverged first followed by prochymosin, and that pepsinogens A and F are most closely related. Tertiary structures, clarified by X-ray crystallography, are commonly bilobal with a large active-site cleft between the lobes. Two aspartates in the center of the cleft, Asp32 and Asp215, function as catalytic residues, and thus pepsinogens are classified as aspartic proteinases. Conversion of pepsinogens to pepsins proceeds autocatalytically at acidic pH by two different pathways, a one-step pathway to release the intact activation segment directly, and a stepwise pathway through a pseudopepsin(s). The active-site cleft is large enough to accommodate at least seven residues of a substrate, thus forming S_4 through S_3' subsites. Hydrophobic and aromatic amino acids are preferred at the P_1 and P_1' positions. Interactions at additional subsites are important in some cases, for example with cleavage of κ-casein by chymosin. Two potent naturally occurring inhibitors are known: pepstatin, a pentapeptide from *Streptomyces*, and a unique proteinous inhibitor from *Ascaris*. Pepsinogen genes comprise nine exons and may be multiple, especially for pepsinogen A. The latter and progastricsin predominate in adult animals, while pepsinogen F and prochymosin are the main forms in the fetus/infant. The switching of gene expression from fetal/infant to adult-type pepsinogens during postnatal development is noteworthy, being regulated by several factors, including steroid hormones.

Key words. Pepsinogen; pepsin; progastricsin; gastricsin; prochymosin; chymosin; aspartic proteinase.

Introduction

Pepsinogens are precursors of pepsins, gastric digestive proteinases belonging to a family of aspartic proteinases. They are synthesized in the gastric mucosa of vertebrates and converted to pepsins in the acidic environment of gastric juice. Pepsinogens and pepsins have a long history [reviewed in refs. 1, 2]. Indeed, the term 'pepsin' was given to a specific ferment in gastric juice in 1836 by Theodor Schwann. Since pepsins have high proteolytic activity, and of sufficient amounts can be isolated by simple techniques such as ammonium sulfate fractionation because levels of pepsinogens in gastric mucosa and pepsins in gastric juice are high [3], they have been studied extensively from early in the 20th century. Thus a great deal is known about their enzymology and protein chemistry, including their mode of action, and primary and tertiary structures associated with function [4, 5]. Pepsinogens and pepsins have also been studied from biological and medical aspects as good models of secretory proteins [6, 7], and suitable molecular markers of gastric cell development [8] and gastric disease [7, 9]. To date,

several types of pepsin family enzymes are known to be present in the stomach. Since the term 'pepsinogen' and its active form 'pepsin' are historical names, to use them as generic names for zymogens and their active forms, respectively, is appropriate. In this review, recent progress is summarized, focusing especially on structure, function, molecular evolution, and development. Studies from other aspects including clinical, medical, and industrial implications have appeared in other reviews [7, 9, 10].

Multiplicity of pepsinogens

Nomenclature

There are five groups of pepsinogens: pepsinogens A, B, and F, progastricsin, and prochymosin, whose primary structures differ significantly (table1). They are precursors of pepsins A, B, and F, gastricsin, and chymosin, respectively, given EC numbers by the IUBMB of EC 3.4.23.1 to 3.4.23.4, except for pepsin F, without one. Isozymogens are frequent for each type of zymogen, especially pepsinogen A. For example, there are five, four, and six pepsinogen A isozymogens in humans [11], the Japanese monkey (fig. 1A) [12], and the rabbit (fig. 1B) [13], respectively. Numbers have been shown to be extremely large in apes, e.g. 9 and 18 in the gibbon and orangutan, respectively [14], and these isozymogens are thought to be the products of different genes [15, 16], although some of them are generated by post-translational modifications such as phosphorylation [17] and glycosylation [18, 19]. With progastricsin and prochymosin, although phosphorylation, glycosylation, and amino acid mutations have been found occasionally, there is no evidence for multiple genes [20, 21]. Arabic numbers are recommended to distinguish isozymogens in each group, like pepsinogens A-1 and A-2, according to their electrophoretic mobility to the anode. Exceptionally, a traditional isozymogen nomenclature is used for bovine prochymosin, that is prochymosins A, B, and C [21]. Identification of the type of pepsinogen is convincing when information about primary structures is available. Analyses of amino acid composition and enzymatic prop-

Table 1. Classification of pepsinogens.

Group	Active form	EC number	Other names[a]
pepsinogen A [3, 27]	pepsin A	3.4.23.1	pepsinogen, pepsinogen I[b] [11], pepsinogens 1–5 (for human isozymogens [11]
pepsinogen B [27]	pepsin B	3.4.23.2	
progastricsin [139]	gastricsin	3.4.23.3	pepsinogen C [27], pepsinogen II[b] [11], pepsinogens 6 and 7 (for human isozymogens) [11]
prochymosin [99]	chymosin	3.4.23.4	prorennin[c] [99], prochymosins A, B, and C (for bovine isozymogens) [21]
pepsinogen F [37]	pepsin F		pregnancy-associated glycoprotein [109]

[a] Other nomenclatures still in use are given. Nomenclatures used tentatively in some literature including those numbered mainly based on chromatographic and electrophoretic behaviors are not listed.
[b] Used mainly in clinical and medical studies.
[c] Old name. The name 'rennin' appears rarely today in some fungal milk-clotting aspartic proteinases like *Mucor* rennin.

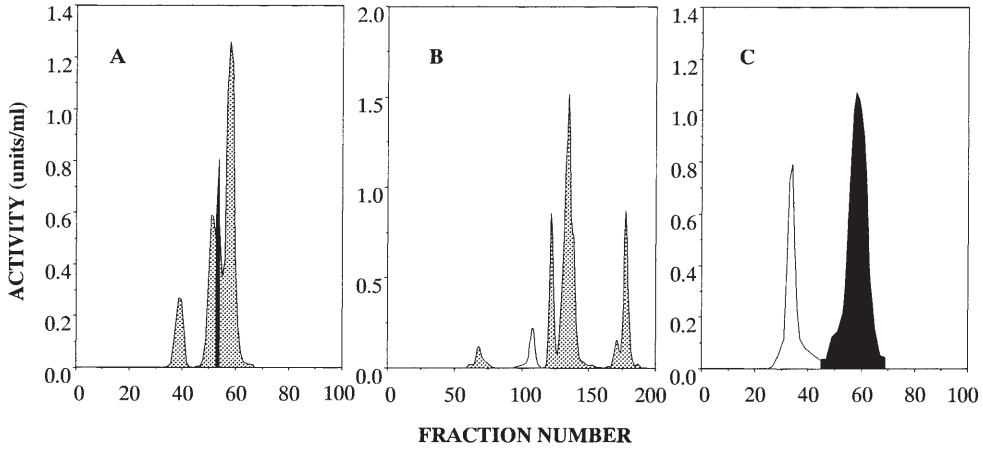


Figure 1. Multiplicity of pepsinogens of the gastric mucosa of the adult Japanese monkey (*A*) [12], rabbit (*B*) [37], and guinea pig (*C*) [33], analyzed by DEAE-Sephacel chromatography at pH 7. Extracts from 1.4, 3.5, and 4.5 g of the fundic part of the stomach, respectively, were applied. Elution was carried out with a linear gradient of NaCl from 0 to 0.5 M. The potential pepsin activity of pepsinogen of each fraction is shown, assayed by the hemoglobin digestion method [26]. ▒, pepsinogen A; ■, progastricsin; □, procathepsin E.

erties, and application of immunochemical methods [2, 22] are also useful for specification. There are problems, however, with non-mammalian vertebrate pepsinogens such as those in fish which appeared much earlier in evolution, resulting in significant diversity in their structures and enzyme properties [23].

In the early days of pepsinogen studies, various nomenclatures were applied [2], for example, featuring roman numerals as for pepsinogens I and II, like arabic numerals mainly based on the order of chromatographic elution or electrophoretic mobility. Some of these nomenclatures are still in use, especially for human cases (table 1). In humans, five pepsinogen A and two progastricsin isozymogens have been detected in gastric mucosa, and these molecular species have been named pepsinogens 1 through 7 [11]. The classification into pepsinogen groups I and II, based on immunological properties, although these correspond to pepsinogen A and progastricsin, respectively, is also widely applied in clinical studies, due to its convenience [24]. In this review, the IUBMB nomenclature is employed.

Expression differences and tissue distribution

Pepsinogens are generally purified from gastric mucosa by combinations of anion exchange chromatography and gel filtration (fig. 1) [2]. Separation of isozymogens whose structures are similar can be achieved by high-performance liquid chromatography on an anion exchange resin [25]. The potential pepsin activity of a pepsinogen is determined by conventional hemoglobin digestion methods [26], but since prochymosin and pepsinogen B are less active against this substrate, other methods such as the milk-clotting assay may be necessary [21, 27]. Some chromogenic peptide substrates are also useful for assaying pepsin activity [28]. Pepsinogen A and progastricsin are mainly expressed in adult vertebrates, although their relative levels differ greatly between animals (fig.1). Both types of pepsinogen are found in primates [12], insectivores [29], and even-toed ungulates [27, 30], although type A pepsinogens predominate in most of these mammals. In the rabbit [13] and carnivores such as the dog [31], pepsinogens A are nearly exclusive components, and progastricsins have not been isolated from these mammals. However, since a cDNA for progastricsin mRNA has been obtained from rabbit gastric mucosa, small amounts might be present [31]. In contrast, in rodents such as the rat [32] and guinea pig [33], there are only progastricsins. Since no pepsinogen A has been detected or isolated from rodents to date, the pepsinogen A gene might have been inactivated. However, expression of a significant amount of procathepsin E in gastric mucosa is known [34] and this is an aspartic proteinase close to pepsinogens [35], so that it could compensate for the lack of pepsinogen A. Prochymosin [36] and pepsinogen F [37] are pepsinogens specific for fetal/infant stages, as described in detail below. Pepsinogen B has been isolated only from pig [27, 38]. Although it may be more widely distributed in vertebrates, it has not been studied extensively, probably because it is difficult to detect by routine assay procedures [27].

The total pepsinogen levels in gastric mucosa vary in adult mammals, presumably correlating with food habits. Values appear to be higher in herbivorous mammals including monkeys [12] and rabbit [13] than in omnivores and carnivores [30, 32, 33]. This may mean that mammals ingesting low-protein or herbaceous foods need more pepsinogens, probably to digest food proteins efficiently. The lowest levels are, however, found in ruminants such as goat [30]. Although ruminants are typical herbivores, their true stomach (abomasum) is adapted for digestion of high-protein foods, i.e., microorganisms raised in the rumen.

The tissue distribution of pepsinogens is primarily restricted to the stomach, pepsinogen A and prochymosin being typical. Progatricsin, however, is also found in some other tissues, including the intestine, lung, seminal vesicles, pancreas, and prostate, as clarified mainly for the human case [39, 40]. High expression of progastricsin in the esophagus is known in the frog [22]. These findings suggest that progastricsin still retains some characteristics of an ancestral housekeeping tissue proteinase. Small amounts of gastric pepsinogens are also known to be secreted into serum, and the determination of serum pepsinogen levels is advantageous for diagnosis of stomach diseases [7, 9, 24, 41]. Although not present in normal tissues, pepsinogen A and progastricsin have been found in some human malignancies [41–43].

Primary structures

Complete primary structures have been determined for more than 50 pepsinogens to date. Typical results are given in fig. 2. Tang et al. [17] first reported the complete amino acid sequence of porcine pepsin A in 1973. Subsequently, the sequence of bovine prochymosin was determined by Foltmann et al. [44] and later that of monkey progastricsin was described by Kageyama and Takahashi [45]. These three sequences, and those of chicken [46] and monkey pepsinogens A [47] were obtained with techniques of protein chemistry. Amino acid sequences deduced from the nucleotide sequences of cDNAs or genomic DNAs identify three regions: the signal peptide (prepeptide), the activation segment (propeptide), and the active enzyme moiety (pepsin moiety), in that order. The number of amino acids is 15, 44, and 326, respectively, for porcine pepsinogen A. The prepeptide is highly hydrophobic, as is commonly observed for secretory proteins, and is removed during pepsinogen synthesis. Na-

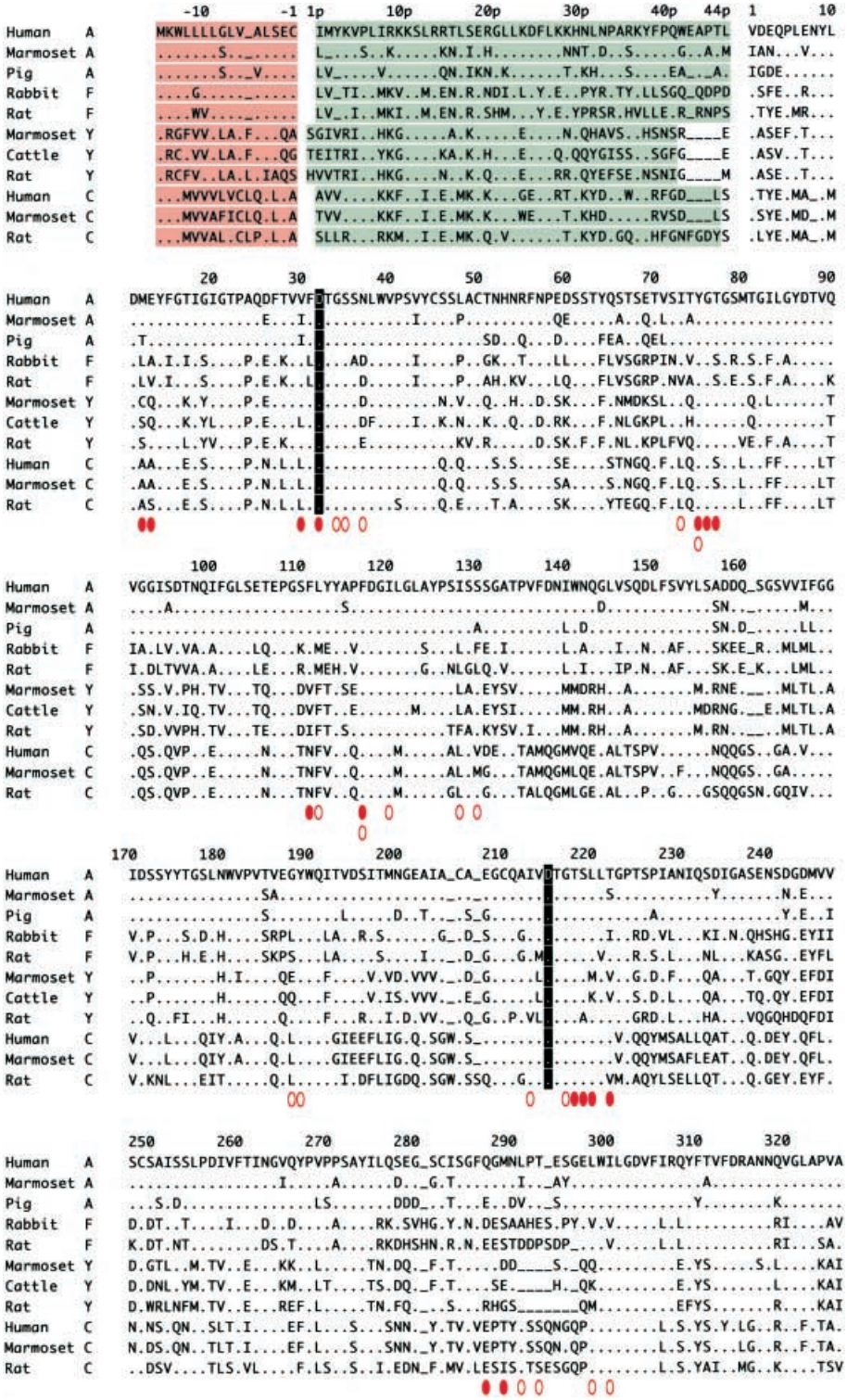


Figure 2. Comparison of the primary structures of typical mammalian pepsinogens A and F, progastricsins (abbreviated as C), and prochymosins (abbreviated as Y). The sequences are deduced from cDNA or genomic DNA sequences of human pepsinogen A [112] and progastricsin [113], common marmoset pepsinogens [92], pig pepsinogen A [49], rat pepsinogens [96], rabbit pepsinogen F [37], and bovine prochymosin [116]. Dots indicate that nucleotides are identical to the human sequence. Deleted residues are shown by bars. The numbering system for porcine pepsin A is applied [17]. The residues in the signal peptide are indicated by negative numbers, and those in the activation segment are characterized by the suffix p. These residues are shaded in light red and green, respectively. Two catalytic-site aspartates are shown in filled boxes. The filled and open red ellipses under the aligned sequences indicate the positions of residues consisting of S_4 through S_1 subsites, and S_1' through S_3' subsites, respectively, of porcine pepsin A, that contact with residues of an ideal heptapeptide substrate [55].

tive pepsinogen thus consists of the activation segment and the pepsin moiety. The activation segment contains relatively high amounts of basic residues (9 Lys, 2 Arg, and 2 His in the case of porcine pepsinogen A) relative to acidic residues (1 Asp and 1 Glu), giving this segment a basic character. In contrast, the pepsin moiety is characteristically highly acidic in nature having many more acidic than basic residues (28 Asp and 13 Glu compared to 2 Arg, 1 Lys, and 1 His). Because the carboxyls of the pepsin moiety are negatively charged at neutral pH, the positive charge of lysine and arginine of the activation segment is essential to compensate for the carboxyl charge and thus stabilize the pepsinogen molecule [48].

Catalytic-site residues are Asp32 and Asp215 (numbering for porcine pepsin A), and residues around these aspartates are well conserved among different types of pepsinogen. Replacement of these aspartates by other residues, as found in the Asp32Ala mutant, inactivates the enzyme [49]. Apart from the sequence similarity around these two aspartates, many topologically equivalent residues are present in the N- and C-terminal halves. Tang et al. [50] suggested that the similarity between the two lobes is a consequence of gene duplication and fusion.

Post-translational modifications such as phosphorylation and glycosylation are occasionally found. For example, Ser68 of porcine [17] and monkey pepsinogens A [47] is phosphorylated. N-glycosylation has been shown to occur at Asn26 and Asn68 of Japanese monkey pepsinogen A-4 [18] and chicken pepsinogen A [46], respectively. The glycosylation site might be generated by chance as a result of mutation, since, although potential N-glycosylation sites can be found in some other pepsinogen molecules, the sites are not uniform [31]. Since post-translational modifications generate a variety of pepsinogen species with more or less different catalytic properties [19], the resulting heterogeneity might be advantageous for effective digestion of the various food proteins in the stomach.

Tertiary structure

Three-dimensional structures have been determined for porcine pepsinogen A [51, 52] and pepsin A [53–56], human pepsinogen A [57] and progastricsin [58], and bovine chymosin [59, 60]. Canadian, British and Russian groups leaded by James, Blundell, and Andreeva, respectively, contributed significantly to these structural analyses. X-ray crystallographic data revealed the tertiary structures of the active enzyme moieties of these zymogens to be very similar and common to those of other en-

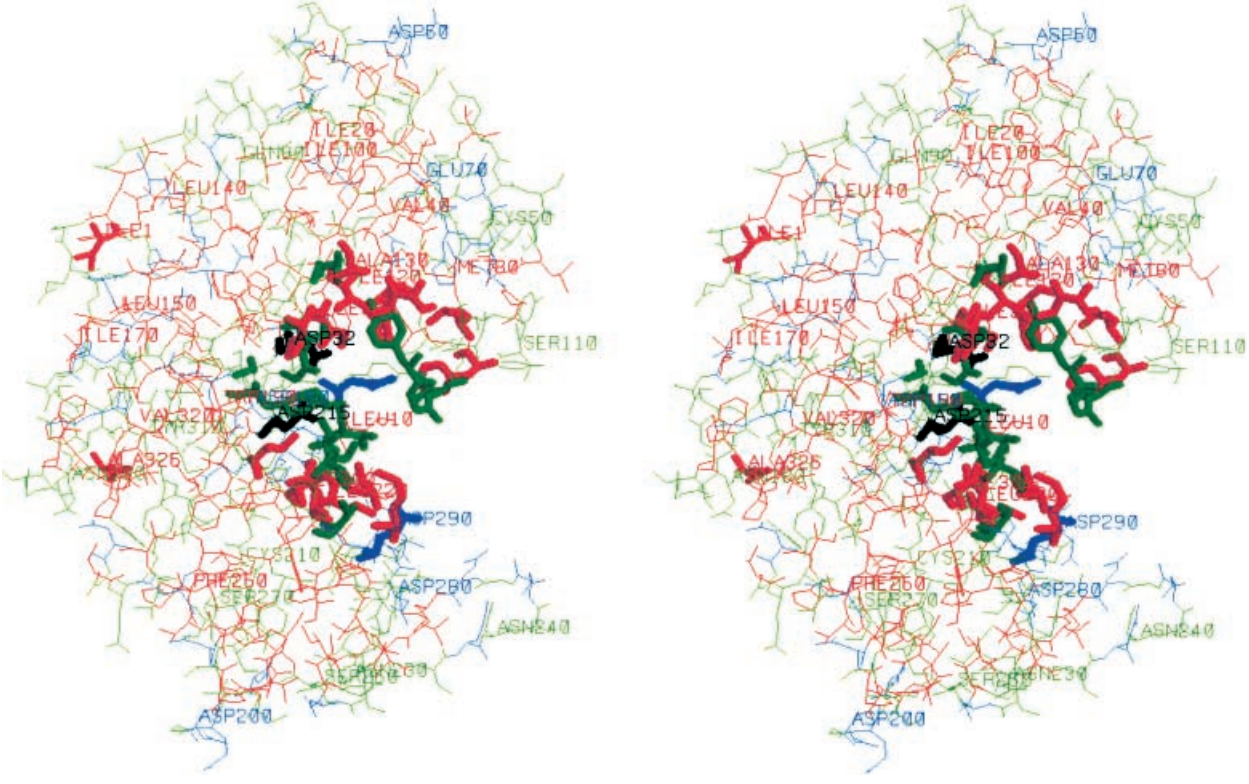


Figure 3. Stereoview of the tertiary structure of porcine pepsin A, obtained from the three-dimensional structure database at the National Center for Biotechnology Information (MMDB Id: 3198, PDB Id: 4PEP; deposited by Andreeva et al. in 1989). The picture was generated using RASMOL ver. 2.5. Residues are shown with wireframes, distinguished by red, green, and blue for non-polar hydrophobic, polar but uncharged, and charged residues, respectively. Catalytic-site Asp32 and Asp215 are colored black. Enzyme residues in contact with residues of an ideal heptapeptide substrate (see fig. 2) [55], Asp32, Asp215, N-terminal Ile1, and C-terminal Ala326 are shown with thick wireframes.

zymes belonging to the aspartic proteinase family (fig. 3). Pepsins consist largely of β sheets and the fold of the protein is stabilized by a large number of hydrogen bonds. Pepsins and other aspartic proteinases share a bilobal symmetry, and, in the case of porcine pepsin A, residues 1–175 form the N-terminal lobe while residues 176 to 326 constitute the C-terminal lobe [54]. As noted above, aspartic proteinases are suggested to have evolved by gene duplication and fusion of an ancestral protein of the size of one lobe of pepsin [50]. Recombinant half-sized N- and C-terminal fragments of porcine pepsin A have been shown to generate active homodimers, mimicking hypothetical ancestral pepsin molecules [61]. To date, however, such an ancestral-type half-sized aspartic proteinase has not been found in living organisms. Exceptionally, the occurrence of a half-sized aspartic proteinase acting as a dimer is known in retroviruses such as the human immunodeficiency virus [62]. Although there is a hypothesis that retroviral proteinases might be fossils of an ancestral dimeric aspartic proteinase [63], the proteinase in question has been suggested to have evolved from a cellular gene by deletion events [64]. Twofold symmetry in each lobe of pepsins and other aspartic proteinases including retroviral proteinases is known [54] and the presence of such intradomain symmetry and repeating sequences of the structural units suggests the possibility that the pepsin structure evolved by double duplication of a gene encoding a quarter-size pepsin molecule. However, alternatively, the intradomain symmetry could be the consequence of a requirement to form a compact domain during the chain-folding process [62].

An extended substrate-binding cleft is located between the two lobes, containing two catalytic-site aspartates, Asp32 and Asp215. The cleft is sufficiently long to accommodate seven residues in a substrate, constituting S_4 through S'_3 subsites (fig. 3) [53–56] and pepsins preferentially cleave peptide bonds between large hydrophobic or aromatic amino acids [5]. This preference can be explained by the hydrophobic character of the S_1 and S'_1 subsites of the enzyme. The S_1 subsite in particular has been shown to be a primary determinant of substrate specificity [65]. As clarified for porcine pepsin A, the S_1 subsite features a flexible loop, Leu71–Gly82, commonly known as a flap, formed by several hydrophobic residues in the neighborhood of the catalytic Asp32 and residues in the flap, including Ile30, Tyr75, Thr77, and Phe111. The S'_1 subsite is formed by residues Tyr189, Ile213, Val291, Thr293, Leu298, and Ile300, and is more exposed (fig. 3) [53–56]. With accommodation in the active-site cleft, a number of hydrogen bonds, such as those from hydroxyl groups of Tyr75 and Thr77, to substrate residues are essential for proper alignment for catalysis [53–56]. The role of some of these bonds has been analyzed by site-directed mutagenesis, such as Thr77Val, and lowered catalytic efficiency was observed [66]. Interactions of pepsins with large peptides or proteins may also occur at the edge or outside the cleft, and for these cases, more substrate-binding pockets such as S_8 through S_5 subsites are anticipated [67]. Some residues, such as Lys319, far from the cleft, have been shown to be involved in catalysis by sustaining the flexibility of the enzyme [68]. Since the overall tertiary structures, including the active-site cleft, of a pepsin family are very similar, the differences in catalytic activities between pepsin types must result from subtle variation in their structures.

Zymogen activation

Pepsinogen is converted to pepsin under acidic conditions with release of the NH_2-terminal activation segments. This process is so-called 'activation' [reviewed in refs. 69, 70]. In zymogen molecules at neutral pH, the nascent binding cleft of the enzyme is occupied by the activation peptide segment at the N-terminal part of the pepsin molecule [51, 52]. Electrostatic, hydrogen-bonding, and hydrophobic interactions between these peptides and the pepsin moiety, including the interactions of the Tyr37p (the suffix 'p' is used for the residues in the activation segment), Lys36p, Tyr9 triad with the two catalytic aspartates, may be critical for stabilizing the pepsinogen structure. In acidic media, most carboxyls of pepsinogen are protonated resulting in the weakening of electrostatic interactions, and large conformational changes occur in the activation segment plus the first 13 residues of the pepsin moiety [51, 52, 55]. Spectroscopic studies have shown that there are conformational changes within 5–100 ms of activation [71, 72]. The active site is then exposed and the pepsinogen molecule cleaves off its own activation segment autocatalytically.

There are two major cleavage sites on the activation segments of various pepsinogens as summarized in figure 4 [69]. One is the connecting peptide bond between the activation segment and the pepsin moiety, and the other is the bond in the middle of the activation segment. These two major bonds might be positioned near the active site after conformational changes [51, 52]. Either or both P_1 and P'_1 residues of these bonds are hydrophobic and aromatic amino acids, which are preferred by pepsins. Cleavage of the former bond generates the active pepsin molecule directly [73, 74], and of the latter results in pseudopepsin, an intermediate form between pepsinogen and pepsin [74–78]. Further cleavage is then necessary to generate the active pepsin. The bond connecting the activation segment and pepsin moiety is cleaved nearly exclusively in monkey pepsinogen A-1, resulting in one-step generation of pepsin A-1 [73]. On the other hand, the bond in the middle of the activation segment is cleaved first in most mammalian progastricsins, resulting in stepwise generation of gastricsin through pseudogastricsin

[77, 79]. With some pepsinogens such as porcine [74] and human [80] pepsinogens A, the two major sites are cleaved at very similar rates, resulting in the simultaneous occurrence of one-step and stepwise activation pathways. The positions of the two major cleavage sites are not uniform between pepsinogens (fig. 4). The pseudopepsin-generating site is located in the region of the 16th ~ 26th residues [13, 76, 78, 80], and the slight variation in cleavage sites between pepsinogens A might be due to differences in the primary structures in this region and the subsite specificity of pepsins A. Since the occurrence of a basic residue at the P_3 position has been shown to exclude the possibility of cleavage of potentially very susceptible bonds such as Leu-Leu and Phe-Leu [65], Arg/Lys at the 14th and 24th residues might hinder cleavage of the 16th-17th and 25th-26th bonds, respectively [81], so that the different distribution of basic residues might result in generation of different-sized pseudopepsins. Although a unique example, the generation of a pseudopepsin by Leu30p-Ala31p cleavage has been observed for rabbit pepsinogen A (isozymogen II-1) [13]. This might be due to the replacement of Pro by Leu at the 30th position, since Pro at the P_1 position, in contrast to Leu, is disliked by pepsin A [65].

Apart from two different activation pathways, two different cleavage reactions are possible for activation: in-

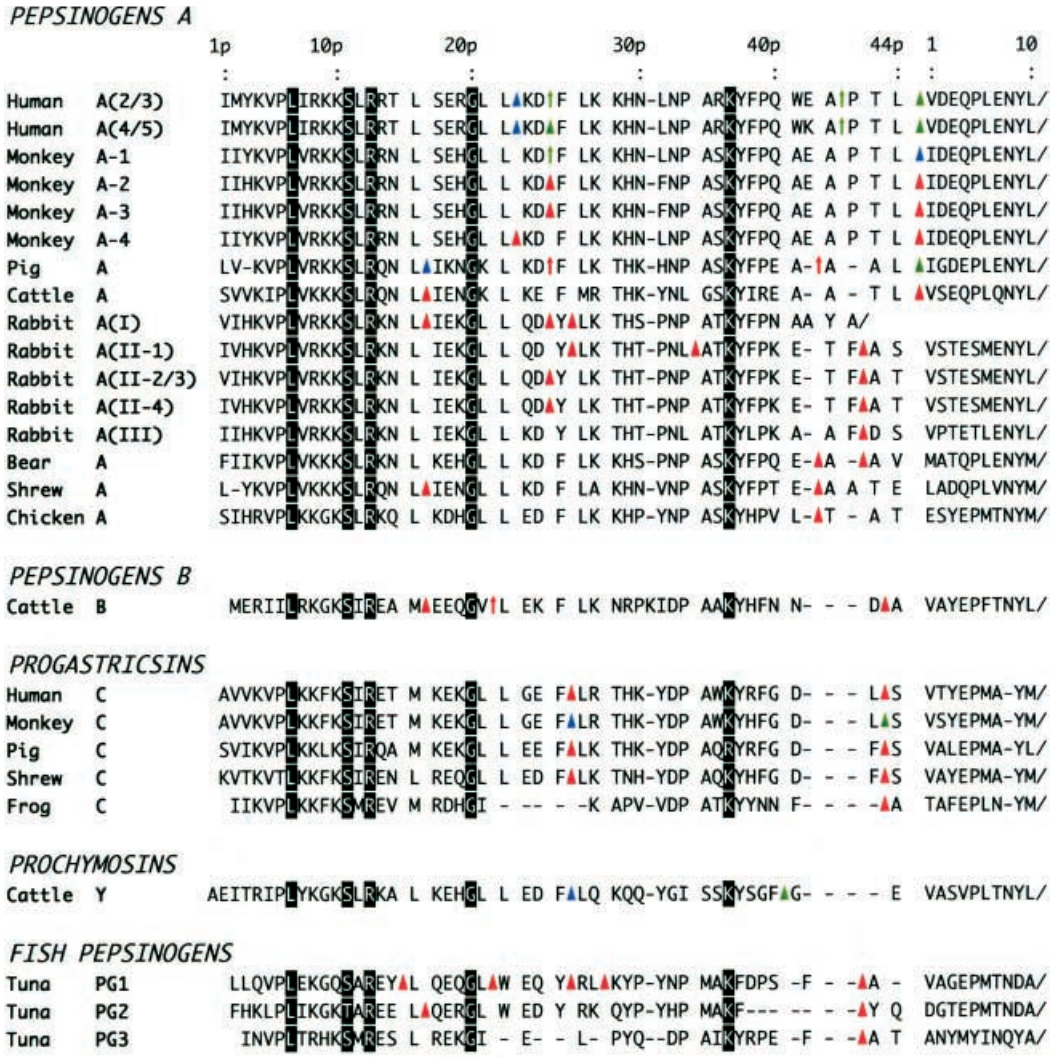


Figure 4. The primary cleavage sites upon activation of various pepsinogens are summarized based on the results of pepsinogens A from human [80, 87], monkey [12, 73, 83], pig [74, 76], cattle [142], bear [69], house musk shrew [29], rabbit [13], and chicken [81]; cattle pepsinogen B [38]; progastricsins from human [77], monkey [12, 83], pig [143], house musk shrew [29], and frog [22]; cattle prochymosin [78]; and tuna pepsinogens [144]. The residues in the activation segment are numbered according to the porcine pepsinogen A sequence and characterized by the suffix p. Residues in filled boxes are conserved in these sequences. The major and minor cleavage sites are indicated by an arrow head and an arrow, respectively. The blue and green colors show that cleavage occurs mainly by intramolecular and intermolecular reactions, respectively, whereas red signifies that their roles have not been clarified. The designations in parentheses are the nomenclatures of the pepsinogen A isozymogens in the original reports. Tuna pepsinogens cannot be classified into known types because their protein sequences have only been partially determined.

tramolecular (unimolecular) and intermolecular [82]. Estimating their respective contributions under different conditions is rather difficult but earlier kinetic studies showed that activation occurs intramolecularly below pH 3, and predominantly intermolecularly at a weakly acidic pH like pH 4 [82]. Activation of immobilized pepsinogen is thought to be suitable to clarify intramolecular reactions. With Sepharose-bound porcine pepsinogen A, cleavage has been shown to generate pseudopepsin A [75, 76], while pepsin A is generated with immobilized monkey pepsinogen A [83]. The involvement of intramolecular reactions might also be estimated in terms of excess substrate since intermolecular reactions between pepsinogens or between pepsinogen and pepsin may be hindered under this condition [84–86]. With monkey and porcine pepsinogen A, the results were essentially the same as with activation of Sepharose-bound zymogens. Only pseudogastricsin was generated from monkey progastricsin suggesting mediation by intramolecular reactions. In contrast, intermolecular activation of pepsinogen A could be observed in the presence of pepstatin. Since the latter rapidly binds to the active site of pepsinogen A that is exposed after conformational change in acidic medium, intramolecular activation cannot proceed [73–75]. Pepstatin-bound pepsinogen A, however, can be attacked intermolecularly by pepstatin-unbound enzyme. With human pepsinogen A, although two pseudopepsins are generated by cleavage of the Leu22p-Lys23p and Asp24p-Phe25p bonds in the absence of pepstatin, ultimate cleavage of the Asp24p-Phe25p bond was observed in its presence, showing that this is due to intermolecular reactions [80, 87]. On the other hand, the Leu22p-Lys23p bond appears to be cleaved intramolecularly.

Proteolytic specificity

Pepsins are endopeptidases and maximally active at acidic pH. They attack native proteins such as hemoglobin and albumin maximally around pH 2. The denaturation of substrate proteins at low pH has been shown to contribute to their efficient hydrolysis [88] and the enzyme has been shown to act on various synthetic peptides similarly at acidic pH, although the optimum varies widely from pH 2 to 5 depending on substrate [5, 89]. Pepsins are obviously adapted to the gastric digestive environment, which is extremely acidic due to HCl secreted from parietal cells of the gastric mucosa. At low pH, most carboxyls of pepsin must be protonated and thus electrostatic repulsion of charged carboxyls is avoided, resulting in a stable conformation and high enzymatic activity [48, 55]. Pepsin hydrolyzes peptide bonds by a general base-assisted nucleophilic mechanism [4]. Asp215 may serve as a general base, abstracting a proton from a water molecule bound between Asp32 and Asp215, since one carboxyl group of these catalytic-site aspartates, held in close proximity in the tertiary structure, would be negatively charged at low pH [55]. Activities of various pepsins against typical protein and peptide substrates and their cleavage specificities for the oxidized B chain of insulin are summarized in table 2 and figure 5, respectively.

Pepsin A specifically cleaves peptide bonds involving various amino acids, especially hydrophobic and aromatic forms. Earlier work with synthetic peptides showed that the most favorable combinations of dipeptides are Phe-Trp, Phe-Tyr, and Phe-Phe, as summarized by Fruton [5]. An NH_2-terminal three-residue and COOH-terminal two-residue extension of the dipeptide greatly enhances the efficiency of the hydrolysis, suggesting that interactions between hexapeptide and pepsin occur at some distance from the catalytic site [90]. Powers et al. [65] have analyzed hydrolysis of peptide bonds in 177 proteins by porcine pepsin A (fig. 6). As was the case with synthetic peptides, the majority of the peptide bonds with high cleavage probabilities have aromatic or hydrophobic amino acids as both P_1 and P_1' residues. The S_1 subsite of pepsin is thought to be the primary determinant of specificity, since the cleavage probability is much higher here than at any other site. Significant differences in probabil-

Table 2. Activity against typical substrates and affinity of inhibitors[a].

Enzyme	Hydrolytic activity (%)[b]			Milk clotting activity (%)[c]	Ki (nM)[d]	
	Hemoglobin	APDT	*Z*Y-L		Pepstatin	*Ascaris* inhibitor
Porcine pepsin A	55–59	70	< 8	25–66	0.05–1	0.1–2.6
Porcine pepsin B	1–3	**100**		0.4–1		
Human gastricsin	**100**	0	**100**	43	100	26
Bovine chymosin	7–14	6		**100**	70	> 700

[a] The variation in some values is due to differences within the literature. Data not available for blank areas.
[b] Hemoglobin-digestive activity [21, 38, 139], and hydrolytic activities against N-acetyl-L-phenylalanyl-L-diiodotyrosine (APDT) [19, 27, 93] and carbobenzoxyl-L-tyrosyl- L-leucine (*Z*Y-L) [93] were determined at pH 2.0. The highest value in each column is taken as 100%.
[c] Assay at pH 6.3 [21, 139]. The value of porcine pepsin B is estimated from a 'casenogram' assay [141]. The highest value is taken as 100%.
[d] Ki values for pepstatin [103, 140] and *Ascaris* inhibitor [106, 140] were determined at pH 3–4.

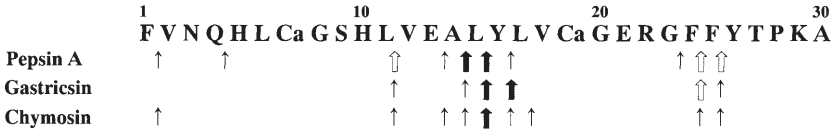


Figure 5. Cleavage sites of oxidized insulin B chain by three different types of pepsin at pH 2. Although the cleavage sites are those of porcine pepsin A [145], gastricsin [146], and bovine chymosin [36], similar results have been obtained with other mammalian pepsins [29, 30, 147]. The primary (⬆) and secondary (⇧) and other minor (↑) cleavage sites have been determined in monkey enzymes [92]. In the case of chymosin, the secondary site is not shown because the cleavage of the bonds other than Leu15-Tyr16 was very slow.

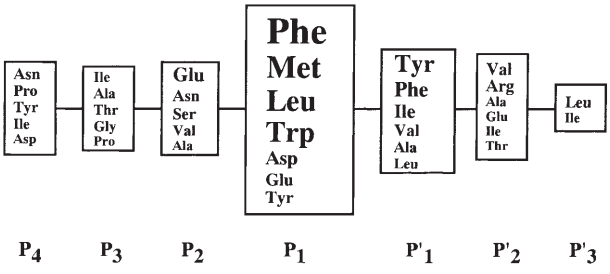


Figure 6. Amino acid residues preferred at P_4 through P'_3 positions, as clarified by Powers et al. [65] analyzing cleavage of 177 proteins or peptides by porcine pepsin A reported in various publications. Amino acid residues with cleavage probabilities larger than the mean are shown, the size of the letters being proportional to the probability. Phe at the P_1 position showed the highest cleavage probability of 0.51. The value of Thr of P'_2 position is 0.18. While the lowest among the listed amino acids, it is slightly higher than the mean value of 0.148.

ity were obtained at the P_4 through P'_3 positions, indicating that pepsin A has a seven-subsite binding site. This size is in agreement with that expected from the hydrolysis of synthetic peptides [90], and was confirmed by the tertiary structural analysis of pepsin A [53–56]. Some chromogenic peptides such as Lys-Pro-Ala-Glu-Phe-Phe(4-NO_2)-Arg-Leu based on the above prediction have been demonstrated to be good substrates for pepsin A and related enzymes [28, 91]. With the advent of solid-phase synthesis, the predicted preference of Pro at the P_4 position and negative effects of basic amino acids such as Arg and Lys at the P_3 position have also been verified by a series of synthetic peptide substrates [91].

Like pepsin A, gastricsin has high proteolytic activity at acidic pH. In several mammalian species, it is known to be maximally active around pH 3 against hemoglobin, a slightly higher pH value than that observed for pepsin A, and its specific activity is about twofold higher [29, 30, 92]. This proteolytic specificity has not been studied as extensively as that of pepsin A, but comparison of human pepsin A and gastricsin showed that although the latter has generally similar substrate specificity, it prefers Tyr at the P_1 position and thus cleaves Tyr-X bonds preferentially (fig. 5) [93].

Chymosin also has maximum proteolytic activity around pH 4 with similar specificity to pepsin A (fig. 5) [94]. The activity, however, is usually weaker in mammals such as the pig, cow, sheep, cat, and rat (table 2) [21, 95, 96]. Exceptionally, New World monkey chymosins have comparable proteolytic activity with pepsin A and are maximally active at pH 2.5 [92]. The roles of some residues in the catalytic activity of chymosin have been estimated by site-directed mutagenesis of the bovine enzyme. A positively charged residue uniquely found in bovine chymosin, Lys220, might have a facilitating effect on proteolytic activity at higher pH since a shift of the optimal pH from 4 to 3.5 was found with the Lys220Leu mutant [97]. A reverse shift to pH 4.4 was evident with the Asp303Ala mutant, suggesting that a negative charge at this position allows maintenance of the correct geometry of the active-site aspartates [98]. The S'_1 subsite of chymosin differs from those of pepsin A and gastricsin, with some deletions in the large loop of residues 289–297 in the C-terminal region (fig. 2). Although these deletions make the S'_1 pocket more open in chymosin than in pepsin A and gastricsin, a significant effect on catalytic activity has not been specified [59, 60].

Of note is that chymosin has the highest milk-clotting activity among the pepsin family enzymes (table 2) [21, 99]. This characteristic is in line with the predominant expression in milk-feeding infants. Bovine chymosin cleaves the Phe105-Met106 bond of κ-casein of bovine milk, resulting in clotting [100]. A negatively charged region near the edge of the active-site cleft, consisting of Glu244, Asp246, Asp248, and Asp250, is involved in this specific cleavage [67], and Asp246 and Asp248 are common to all chymosins except for the chicken enzyme, being replaced by non-charged residues in pepsins A and gastricsins (fig. 2). The negatively charged region in the putative $S_8 \sim S_4$ subsites has electrostatic interactions with the positively charged His-Pro cluster of κ-casein, which is close to the scissile bond, increasing the efficiency of hydrolysis. Addition of a negative charge to this region by a Gly243Asp mutation increased the catalytic activity against a synthetic κ-casein peptide [101]. Species specificity in milk-clotting activity has been established [21, 95]. Although bovine chymosin has similar activity against bovine and porcine milk, porcine chymosin has higher milk-clotting activity against porcine milk [102]. Such species specificity is expected for other chymosins since the primary structure of κ-casein is highly variable between mammals, for example between bovine and rat

κ-caseins (GenBank/EMBL/DDBJ accession numbers M36641 and K02598, respectively).

Information about the specificity of pepsin B is much less extensive than for the three enzymes described above, except that pepsin B is very active against a Phe-Tyr dipeptide (table 2) [27, 38]. The specificity of pepsin F is not known.

Pepsin inhibitors

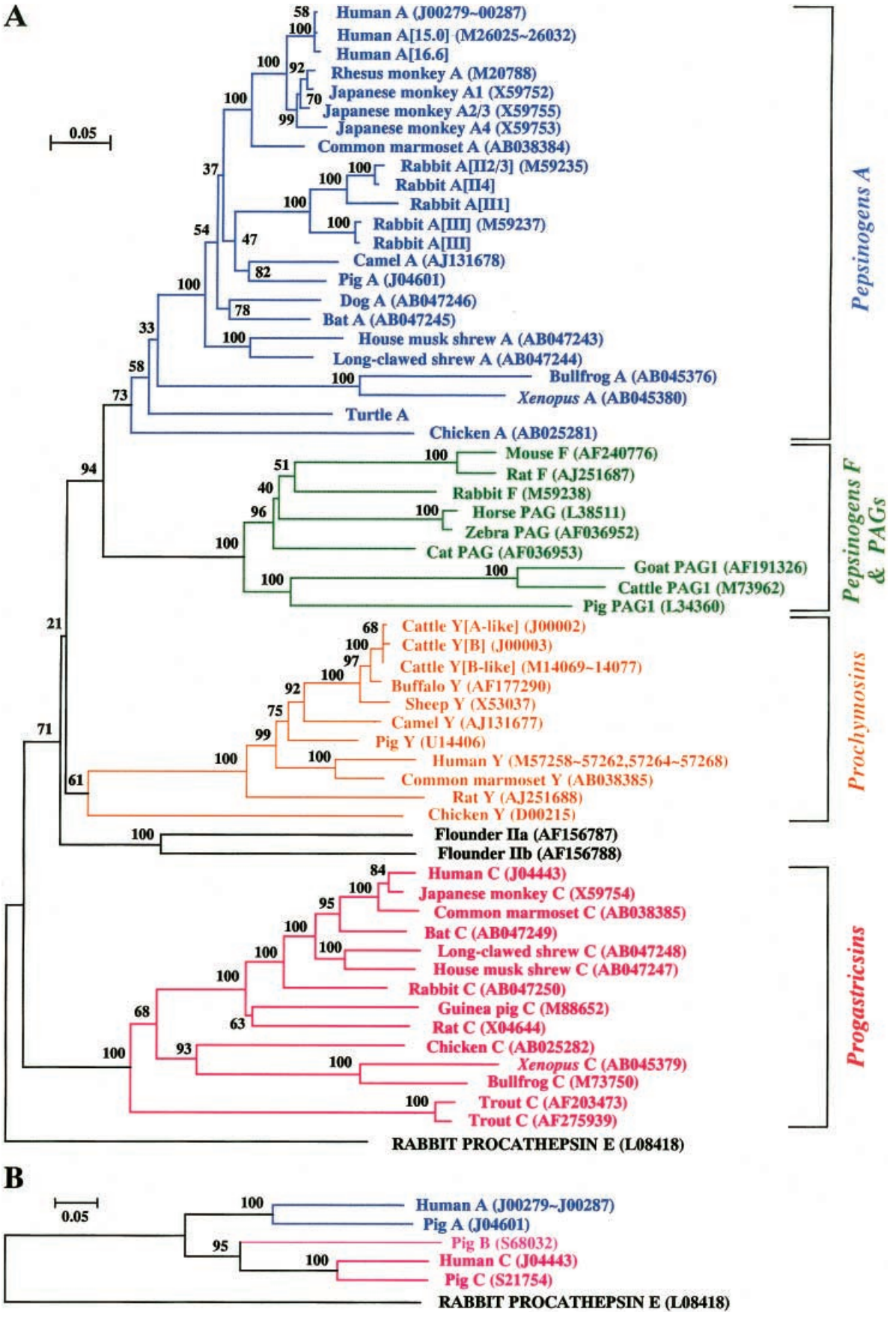
Some naturally occurring inhibitors are known, in addition to synthetic inhibitors such as substrate analogues. The most potent is pepstatin, isolated from *Streptomyces* species, which contains two residues of a novel amino acid, statine, and strongly inhibits pepsin A with a Ki value of 4.6×10^{-11} M [103]. The inhibitor binds to an extended conformation in the active-site cleft, and the first statyl hydroxyl oxygen interacts strongly with the catalytic aspartate via hydrogen bonding to the essential carboxyl groups [4]. Due to the extremely high affinity of pepstatin, the activity of pepsin A is inhibited completely in the presence of equimolar amounts. Pepstatin is effective against aspartic proteinases in general, although its affinity differs between enzymes. Gastricsin and chymosin are less susceptible, and about 100- and 10-fold molar excesses, respectively, are needed for complete inhibition (table 2) [92]. The different susceptibilities between mammalian pepsins are convenient to specify their typing [29, 30, 92], while the different susceptibilities of non-mammalian pepsins such as frog pepsin A and gastricsin are not clear [104].

Ascaris, a parasitic nematode, is also known to secrete a 149-residue proteinous inhibitor with high affinity for pepsin A (Ki values of $10^{-10} \sim 10^{-11}$ M) [105]. cDNAs for this *Ascaris* inhibitor and homologous proteins from other nematodes including *Onchocerca* and *Brugia* have been cloned [106]. Although the intracellular aspartic proteinase cathepsin E is also inhibited efficiently, the effect of the *Ascaris* inhibitor is rather specific for pepsin A and action against gastricsin is over 100-fold weaker. Inhibition of chymosin and other aspartic proteinases is minimal (table 2) [106]. The wide distribution of similar inhibitors in nematodes is in agreement with a role in protecting the worms against gastric digestion. The *Ascaris* inhibitor-pepsin A complex has been clarified by X-ray studies [107]. The high affinity is mainly due to the hydrogen-bonding interactions between the β strand of N-terminal residues 4–8 of the inhibitor and the active-site flap (residues 70–74) of pepsin A. The N-terminal Gln1-Phe-Leu tripeptide residues of the inhibitor occupy the $S_1' \sim S_3'$ subsites of the active-site cleft resulting in inhibition of enzyme activity. The low affinity for gastricsin and chymosin might be due to the slightly different structures of the active site flap of these proteinases. Since another potent proteinous inhibitor has been isolated from roots of *Anchusa strigosa*, with a Ki value of 2×10^{-8} M against pepsin A [108], plants might be sources of new pepsin inhibitors.

Molecular evolution

The different types of pepsinogen are thought to have evolved from a common ancestral aspartic proteinase and a phylogenetic tree based on the nucleotide sequences clearly shows the existence of four clusters, i.e., pepsinogens A and F, progastricsins, and prochymosins (fig. 7A). Progastricsin appears to have diverged first followed by prochymosin, and pepsinogens A and F are most closely related. The grouping of flounder pepsinogens is unresolved but they seem to be different from progastricsin, and might be close to a common ancestor of prochymosin, and pepsinogens A and F. Separation of the latter three forms likely occurred in tetrapod animals. Since the occurrence of progastricsin in trout is established, there might be at least two types of pepsinogen in ancestral fish. Pepsinogen A isozymogens in human [15, 16], monkey [12], and rabbit [37], and bovine prochymosin isozymogens [21] are very close, suggesting that gene duplication events may have occurred independently two or more times in the genes for these pepsinogens. Pepsinogen B is not included in the phylogenetic tree because its structure has only been partly clarified in the NH_2-terminal 67-residue amino acid sequence [38]. From comparisons with known NH_2-terminal amino acid sequences, it is

Figure 7. Molecular evolution of pepsinogens. (*A*) A phylogenetic tree based on the nucleotide sequences of the entire protein-encoding regions, i.e., those encoding signal peptides, activation segments, and pepsin moieties. The sequences are cited from the cDNA or genomic DNA sequences and the accession numbers of GenBank/EMBL/DDBJ databases are given in parentheses. A, F, C, Y, and PAG stand for pepsinogens A and F, progastricsin, prochymosin, and pregnancy-associated glycoprotein, respectively. Numbers or letters in brackets show nomenclatures in the original reports. A and B in the brackets of cattle prochymosin represent prochymosin isozymogens A and B, respectively. Sequences that have not been deposited in databases are cited from references, including those for human pepsinogen A16.6 [15] and turtle pepsinogen A [148]. The sequences for some rabbit pepsinogens are unpublished ones from our laboratory. The tree was constructed by the neighbor-joining method of the program ClustalW with corrected distances. Rabbit procathepsin E was used as an outgroup. The branch lengths are proportional to the evolutionary distances, given as the average numbers of substitutions per site. The scale is shown in the upper part of the figure. The numbers beside nodes represent the bootstrap values, as percentages, obtained for 1,000 replicates. (*B*) A neighbor-joining phylogenetic tree based on the amino acid sequences of the activation segment and N-terminal 23 residues of the pepsin moiety.


A
0.05
Human A (J00279~00287)
Human A[15.0] (M26025~26032)
Human A[16.6]
Rhesus monkey A (M20788)
Japanese monkey A1 (X59752)
Japanese monkey A2/3 (X59755)
Japanese monkey A4 (X59753)
Common marmoset A (AB038384)
Rabbit A[II2/3] (M59235)
Rabbit A[II4]
Rabbit A[II1]
Rabbit A[III] (M59237)
Rabbit A[III]
Camel A (AJ131678)
Pig A (J04601)
Dog A (AB047246)
Bat A (AB047245)
House musk shrew A (AB047243)
Long-clawed shrew A (AB047244)
Bullfrog A (AB045376)
Xenopus A (AB045380)
Turtle A
Chicken A (AB025281)
Pepsinogens A
Mouse F (AF240776)
Rat F (AJ251687)
Rabbit F (M59238)
Horse PAG (L38511)
Zebra PAG (AF036952)
Cat PAG (AF036953)
Goat PAG1 (AF191326)
Cattle PAG1 (M73962)
Pig PAG1 (L34360)
Pepsinogens F & PAGs
Cattle Y[A-like] (J00002)
Cattle Y[B] (J00003)
Cattle Y[B-like] (M14069~14077)
Buffalo Y (AF177290)
Sheep Y (X53037)
Camel Y (AJ131677)
Pig Y (U14406)
Human Y (M57258~57262,57264~57268)
Common marmoset Y (AB038385)
Rat Y (AJ251688)
Chicken Y (D00215)
Prochymosins
Flounder IIa (AF156787)
Flounder IIb (AF156788)
Human C (J04443)
Japanese monkey C (X59754)
Common marmoset C (AB038385)
Bat C (AB047249)
Long-clawed shrew C (AB047248)
House musk shrew C (AB047247)
Rabbit C (AB047250)
Guinea pig C (M88652)
Rat C (X04644)
Chicken C (AB025282)
Xenopus C (AB045379)
Bullfrog C (M73750)
Trout C (AF203473)
Trout C (AF275939)
Progastricsins
RABBIT PROCATHEPSIN E (L08418)
B
0.05
Human A (J00279~J00287)
Pig A (J04601)
Pig B (S68032)
Human C (J04443)
Pig C (S21754)
RABBIT PROCATHEPSIN E (L08418)

probable that pepsinogen B diverged from progastricsin and evolved as an independent group (fig. 7B).

The relationship between pepsinogen F and pregnancy-associated glycoprotein (PAG) deserves mention. Pepsinogen F was found first by Kageyama et al. [37] as an infant-specific pepsinogen in rabbit stomach and later in rat stomach [96]. Pregnancy-associated glycoprotein was found in placenta of artiodactyles and perissodactyles such as cattle and horse [109]. The phylogenetic tree shows that both belong to the same cluster, so that pepsinogen F might have evolved for expression in placenta of cattle and horse due to the physiological requirements of these animals, presumably serving as a carrier protein during pregnancy [110]. PAG in ruminants constitutes a large family of recently duplicated genes like the pepsinogen A family in primates and the rabbit [111].

The phylogeny of pepsinogens is considered to be useful for estimating that of vertebrates, especially mammals [31]. Progastricsin might be the most suitable molecular marker for this purpose, since multiplication of its genes has not been found, and thus vertebrate progastricsins might be products of orthologous genes.

What is the evolutionary origin of pepsinogens? To date, the genomic or cDNA structures of many aspartic proteinases have been clarified, although for some genes, whether the encoded proteins are functional is uncertain. A typical phylogenetic tree constructed based on the nucleotide sequences of active enzyme moieties shows that several major divergences occurred in the past in animal aspartic proteinases (fig. 8). Vertebrate forms including pepsins belong to the same cluster. Since cathepsin E is an intracellular aspartic proteinase distributed in various tissues, and some of its enzymatic properties such as maximal activity at low pH are similar to those of pepsins [35], pepsin (pepsinogen) may be derived from a cathepsin-E-like intracellular proteinase. Although the time of divergence is unclear, it might have occurred in invertebrates like arthropods. If an invertebrate proteinase close to such an ancestral pepsin (pepsinogen) exists, however, it remains to be detected. Although some invertebrate or plant aspartic proteinases are occasionally named with the suffix 'pepsin', like strongyloidespepsin and endothiapepsin, they are clearly very distant from vertebrate pepsins.

Genes and expressional regulation

Complete gene structures have been clarified for eight pepsinogens including human pepsinogen A [112], progastricsin [113] and prochymosin (pseudogene) [114], rat progastricsin [115], bovine prochymosin [116], and chicken pepsinogen A [117], progastricsin [117], and prochymosin [118]. Each consists of nine exons separated by eight introns. Takahashi [119] has reported that the positions of the introns which interrupt the coding sequence of the preproprotein are conserved among pepsinogens, and are similar to those for other mammalian aspartic proteinases, including cathepsins D and E, and renin, again suggesting evolution from a common ancestral gene. The occurrence of multiple genes has

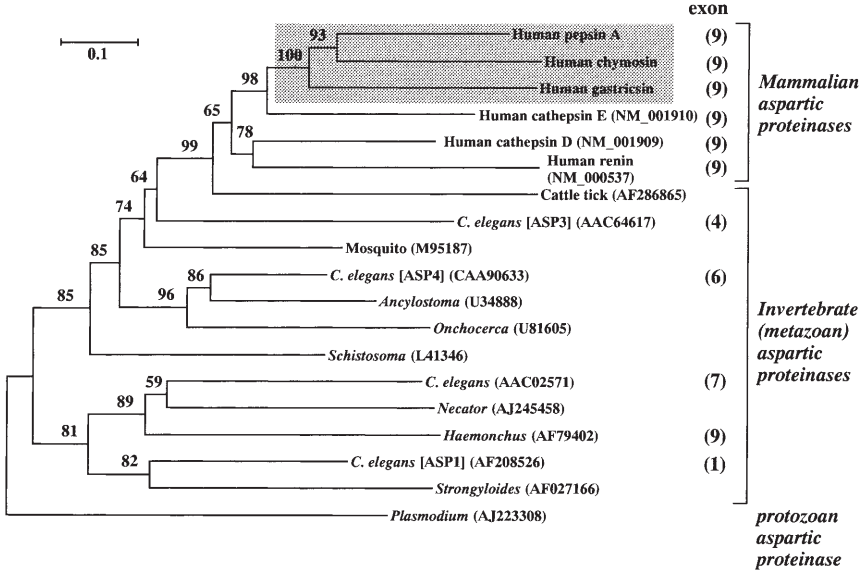


Figure 8. Molecular evolution of animal aspartic proteinases. The neighbor-joining phylogenetic tree was constructed based on the nucleotide sequences of the pepsin moieties of pepsinogens and the active moieties of other animal aspartic proteinases. The number of exons is indicated in parentheses for proteinases for which genomic structures have been clarified. Other notations of the figure are as for fig. 7. The cluster of pepsins is shaded.

been found in the human genome [15] and four pepsinogen A genes, ranging in size from 12.0 to 16.6 kb are arranged in a head-to-tail orientation in the centromeric region of chromosome 11 [16]. Three of these genes are thought to encode pepsinogens 3, 4, and 5 (human pepsinogen A nomenclature). Since at least five human pepsinogen A isozymogens are known, the occurrence of other A genes is conceivable. Genomic fragments derived from several different pepsinogen A genes have also been obtained from a cattle genomic library [120]. In the progastricsin case, the occurrence of multigenes is not known. The human progastricsin gene is located on chromosome 6 and there is only a single copy in the human genome project sequence data. Two protein bands detected by electrophoresis might be products of different alleles of this single gene locus [20].

Since the pepsinogen family is strongly expressed for digestion purposes, the regulatory mechanisms might be amenable to elucidation. Several potential sequences specific for binding of transcriptional factors have been found in the 5′-flanking region of pepsinogen genes and a TATA box is common to all but the chicken progastricsin gene. Although sequences specific for transcriptional factors such as AP-2, HIS4, and MSP are occasionally found, their involvement remains to be clarified [119]. In the chicken prochymosin gene, GATA-and Sox-binding sequences have been found in the 5′-flanking region and shown to be essential for stomach-specific expression, with binding of cGATA-5 protein causing upregulation [121].

The 5′-flanking region of the gene for porcine pepsinogen A has been isolated, and the essential region for promotor activity identified [122]. Several sites from −246 to +21 act synergistically in transcriptional regulation of the gene. The promotor is active in isolated porcine chief cells but not in other cell lines, consistent with the fact that pepsinogen A is synthesized only in chief cells of the gastric mucosa. The similar 5′-flanking sequence of the human pepsinogen A gene, however, showed quite low activity in cultured porcine chief cells, suggesting species-specific regulation. Regulatory protein(s) have yet to be isolated to explain the cell specificity.

The extent of DNA methylation may be involved in regulating the expression of pepsinogen genes, since they are less methylated in the stomach than in other tissues [123, 124]. Pepsinogen is synthesized in gastric chief cells, with abundant rough endoplasmic reticulum, which are thought to have differentiated from undifferentiated mucous neck cells in the middle of the gastric glands. Actions of hormones such as glucocorticoids, and growth factors such as hepatocyte growth factor [125], and interaction with mesenchyme have been shown to be necessary for the differentiation of chief cells [126]. Glucocorticoids also induce the expression of pepsinogen genes, thus acting as important regulators of both function and differentiation of chief cells [127]. Direct actions of the hormones are unknown, but glucocorticoids may act indirectly on pepsinogen genes via other factors, as suggested by the fact that mesenchyme is necessary for enzyme induction [128].

Development

Change in expression from fetal/infant to adult-type pepsinogens

The occurrence of multiple pepsinogen types is correlated with development. Pepsinogen A and progastricsin predominate in adult animals, and prochymosin and pepsinogen F are expressed mainly in the fetus and infant. The developmental change of pepsinogen expression in mammals has been studied extensively by Foltmann, Andrén, and their colleagues [21, 129]. The levels of bovine prochymosin are appreciable in undifferentiated chief and mucous neck cells at about the 70-day-old fetal stage (full term is about 280 days) [130]. Its level reaches a maximum in the newborn, 5 to 10-fold higher than that of pepsinogen A, then decreases gradually during postnatal development and finally becoming negligible in adults (fig. 9). Fetal synthesis of pepsinogen A and progastricsin begins at about 140 and 180 days, respectively, and increases so that their summed level is nearly the same as that of prochymosin around weaning, 6~10 weeks after birth, before reaching a maximum in adult cattle [131]. The results show that the expressional switching from prochymosin to pepsinogen A and progastricsin occurs around weaning although milk-fed calves at older ages, such as 3 months, retain significant synthesizing potential for prochymosin [132]. Hormone-like factor(s) included in the milk might be involved in stimulation of prolonged gene expression.

Switching of pepsinogen gene expression during development has also been demonstrated in other vertebrates, especially in mammals. In the pig [102] and rat [96], prochymosin is produced in the fetal stage and decreases gradually after birth, being replaced by pepsinogens A and/or progastricsin. Although the occurrence of prochymosin has not been demonstrated in rabbit, another fetal/infant-specific pepsinogen, the form F, is transcribed predominantly in infants with replacement by pepsinogens A in adults [37]. Expressional change pursued by Northern analysis in rat stomach showed similar switching around weaning (fig. 10) while in the chicken and quail it occurs in embryos before hatching [133]. Chicken prochymosin is thus expressed strongly in the proventriculus of chicken embryo, showing a peak at day 15, and is replaced by pepsinogen A and progastricsin prior to hatching. Various factors including hormones such as glucocoticoids and thyroxin, hormone-related substances, and transcriptional factors might play roles [123, 127, 128]. Of interest in this context is that while fetal/in-

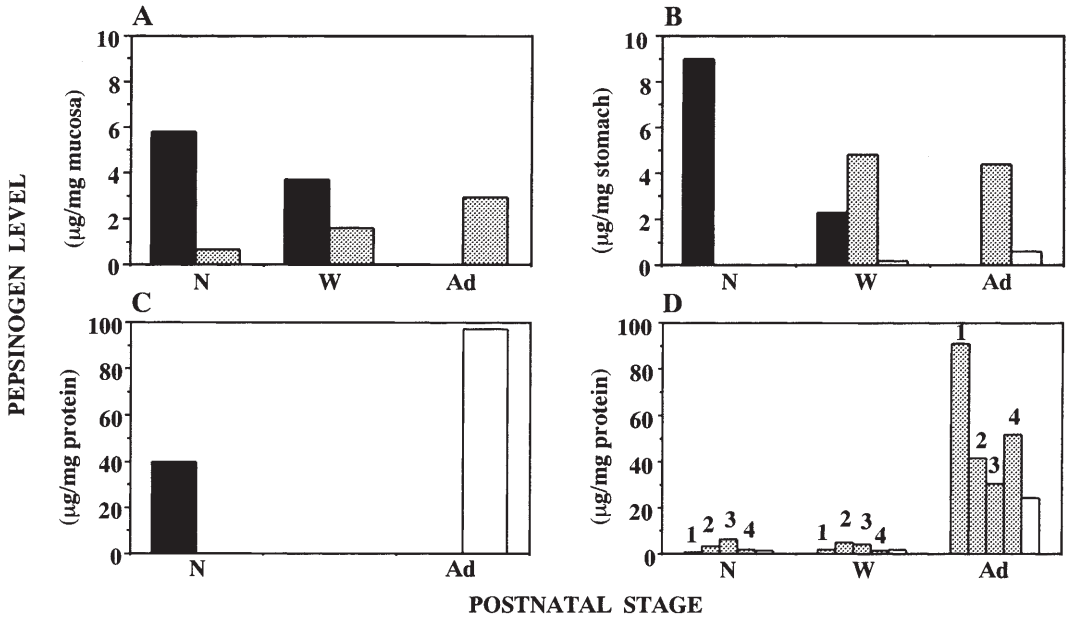
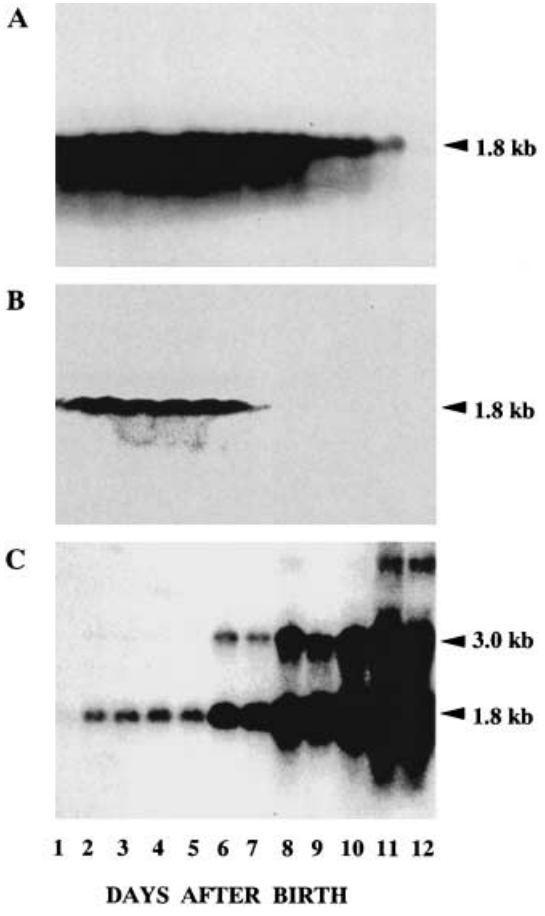


Figure 9. Postnatal change in levels of gastric prochymosin, pepsinogen A, and progastricsin in the cow (*A*) [21, 129], pig (*B*) [102], rat (*C*) [96, 127], and Japanese monkey (*D*) [12]. N, W, and Ad, newborn, around weaning, and adult, respectively; ■, prochymosin; ▒, pepsinogen A; □, progastricsin. The numbers 1 through 4 for Japanese monkey indicate pepsinogens A-1 through A-4, respectively. Data for the level of bovine progastricsin are not available but may be less than 10% of the total in the adult [19]. In the rat, pepsinogen A is not detectable throughout development, presumably due to inactivation of the gene.



fant pepsinogens are distributed widely in mammals and birds, they have not been found in other vertebrates including reptiles, amphibia, and fish [134].

Stage-specific pepsinogens presumably have physiological significance. Foltmann [2, 12] has discussed the contributions of prochymosin in fetus and infant. The active form, chymosin, is thought to be essential for the effective digestion of milk proteins such as κ-casein, as mentioned already in the section on proteolytic specificity. Chymosin in the avian embryo might be involved partly in the digestion of yolk proteins. It has also been shown to be a prerequisite for transfer of immunity in mammals. Since the colostrum is rich in related substances required by the newborn, the low proteolytic activity of chymosin might be thought of as an adaptation allowing effective adsorption of these substances by the gastrointestinal tract. The transfer of maternal antibodies to the newborn is known to occur entirely after birth via the colostrum in artiodactyls such as cattle and pig, and mainly after birth

Figure 10. Cange in expression from fetal/infant to adult-type pepsinogens in rat gastric mucosa during postnatal development, analyzed by Northern blot hybridization [96]. After electrophoresis of total RNA, the blotted membrane was hybridized with cDNAs for rat prochymosin (*A*), pepsinogen F (*B*), and progastricsin (*C*). The major mRNA for each zymogen has a length of around 1.8 kb. The 3.0-kb band and the large-molecular-mass band for progastricsin mRNA are thought to be transcripts in the process of splicing, containing introns as found in monkey pepsinogen A [12].

in rodents [135]. Since maternal antibodies are known to be transferred to the chicken embryo in the yolk, the occurrence of prochymosin in these animals is consistent with the 'immune transfer' hypothesis.

Expressional change in primates

Primates, including human and monkeys, are unique with respect to their milk composition, with characteristically low levels of proteins and high lactose compared with other mammals [136]. Transfer of immunity in primates from mother to newborn occurs mainly before birth through the placenta which might make prochymosin less essential [135]. Indeed, the prochymosin gene is inactivated by mutation in humans [114] and the enzyme is not demonstrable in the stomach of newborn Japanese monkeys [12] or humans [137]. The level of gastric pepsinogens in Japanese monkey is in fact much lower in the infant than in the adult, presumably due to the absence of prochymosin (fig. 9). However, the results in New World monkeys are confusing since they have prochymosin and, moreover, express it throughout postnatal development, resulting in the co-existence of three types of pepsinogens, i. e., prochymosin, pepsinogen A, and progastricsin, in the adult [92]. The active prochymosin gene may have been present in ancestral primates and is retained in New World monkeys, but inactivated in human beings and Old World monkeys. Since New World monkeys eat a variety of foods, including plants and insects, chymosin might be required for the effective digestion of various food proteins. The high proteolytic activity of New World monkey chymosin supports this hypothesis.

There are several pepsinogen A isozymogens in humans [11, 15, 16] and the Japanese monkey [12], whose expression during postnatal development has been shown to proceed sequentially. Of the four pepsinogen A isozymogens in the Japanese monkey, A-3 is expressed predominantly in the fetal/infant stage, A-2 mainly at the 4-month stage, and A-1 and A-4 are found in the juvenile and adult stages. This change is not due to simple switching of gene expression from pepsinogen A-3 through A-2 to A-1/A-4, since the specific activity of each pepsinogen increased with development [12]. Proteolytic enzymes are needed in higher amounts with exposure to various food proteins at later stages and similar alteration in the expression of genes for type A pepsinogens is evident in humans [138]. From chromatographic or electrophoretic analyses, it is also expected to occur in other vertebrates including rabbit [37] and chicken [133].

Perspectives

Pepsinogens and pepsins have diverged significantly in vertebrates, resulting in species specificity, developmental-stage specificity, and substrate specificity, indicating adaptation for a variety of gastric functions. Although typical pepsinogens such as porcine pepsinogen A and bovine prochymosin have been extensively analyzed, pepsinogen diversity and multiplicity still remain to be clarified from molecular and cellular aspects. Further investigations of the structure, function, and expressional regulation of forms such as pepsinogens B and F are necessary for comprehensive understanding of gastric digestive proteinases.

1 Taylor W. H. (1968) Biochemistry of pepsins. In: Handbook of Physiology, Sec. 6, Alimentary Canal, vol. 5, pp. 2567–2587, Code C. F. (ed.), Williams & Wilkins, Baltimore
2 Foltmann B. (1981) Gastric proteinases – structure, function, evolution and mechanism of action. Essays Biochem. **17:** 52–84
3 Herriott R. M. (1938) Isolation, crystallization, and properties of swine pepsinogen. J. Gen. Physiol. **21:** 501–540
4 Davies D. R. (1990) The structure and function of the aspartic proteinases. Annu. Rev. Biophys. Biophys. Chem. **19:** 189–215
5 Fruton J. S. (1970) The specificity and mechanism of pepsin action. Adv. Enzymol. Relat. Areas Mol. Biol. **33:** 401–443
6 Hirschowitz B. I. (1967) Secretion of pepsinogen. In: Handbook of Physiology, Sec. 6, Alimentary Canal, vol. 2, pp. 889–918, Code C. F. (ed.), Williams & Wilkins, Baltimore
7 Gritti I., Banfi G. and Roi G. S. (2000) Pepsinogens: physiology, pharmacology, pathophysiology and exercise. Pharmacol. Res. **41:** 265–281
8 Yasugi S. (1994) Regulation of pepsinogen gene expression in epithelial cells of vertebrate stomach during development. Int. J. Dev. Biol. **38:** 273–279
9 Plebani M. (1993) Pepsinogens in health and disease. Crit. Rev. Clin. Lab. Sci. **30:** 273–328
10 Mohanty A. K., Mukhopadhyay U. K., Grover S. and Batish V. K. (1999) Bovine chymosin: production by rDNA technology and application in cheese manufacture. Biotechnol. Adv. **17:** 205–217
11 Samloff I. M. (1969) Slow moving protease and the seven pepsinogens: electrophoretic demonstration of the existence of eight proteolytic fractions in human gastric mucosa. Gastroenterology **57:** 659–669
12 Kageyama T., Tanabe K. and Koiwai O. (1991) Development-dependent expression of isozymogens of monkey pepsinogens and structural differences between them. Eur. J. Biochem. **202:** 205–215
13 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
14 Narita Y., Oda S. I., Takenaka O. and Kageyama T. (2000) Multiplicities and some enzymatic characteristics of ape pepsinogens and pepsins. J. Med. Primatol. **29:** 402–410
15 Evers M. P., Zelle B., Bebelman J. P., Beusechem V. van, Kraakman L., Hoffer M. J. et al. (1989) Nucleotide sequence comparison of five human pepsinogen A (PGA) genes: evolution of the PGA multigene family. Genomics **4:** 232–239
16 Zelle B., Evers M. P., Groot P. C., Bebelman J. P., Mager W. H., Planta R. J. et al. (1988) Genomic structure and evolution of the human pepsinogen A multigene family. Hum. Genet. **78:** 79–82
17 Tang J., Sepulveda P., Marciniszyn J. Jr, Chen K. C., Huang W. Y., Tao N. et al. (1973) Amino-acid sequence of porcine pepsin. Proc. Natl. Acad. Sci. USA **70:** 3437–3439

18 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
19 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
20 Pals G., Azuma T., Mohandas T. K., Bell G. I., Bacon J., Samloff I. M. et al. (1989) Human pepsinogen C (progastricsin) polymorphism: evidence for a single locus located at 6p21.1-pter. Genomics **4:** 137–148
21 Foltmann B. (1992) Chymosin: a short review on foetal and neonatal gastric proteases. Scand. J. Clin. Lab. Invest. **52 (Suppl 210):** 65–79
22 Yakabe E., Tanji M., Ichinose M., Goto S., Miki K., Kurokawa K. et al. (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
23 Gildberg A. (1988) Aspartic proteinases in fishes and aquatic invertebrates. Comp. Biochem. Physiol. B **91:** 425–435
24 Miki K., Ichinose M., Ishikawa K. B., Yahagi N., Matsushima M., Kakei N. et al. (1993) Clinical application of serum pepsinogen I and II levels for mass screening to detect gastric cancer. Jpn. J. Cancer Res. **84:** 1086–1090
25 Defize J., Pals G., Pronk J. C., Frants R. R., Rimmelzwaan G., Westerveld B. D. et al. (1985) Purification of the pepsinogen A isozymogens by means of high resolution ion-exchange chromatography: evidence for post-translational modifications. Scand. J. Clin. Lab. Invest. **45:** 649–655
26 Anson M. L. and Mirsky A. E. (1932) The estimation of pepsin with hemoglobin. J. Gen. Physiol. **16:** 59–63
27 Ryle A. P. (1970) The porcine pepsins and pepsinogens. Methods Enzymol. **19:** 316–336
28 Dunn B. M., Jimenez M., Parten B. F., Valler M. J., Rolph C. E. and Kay J. (1986) A systematic series of synthetic chromophoric substrates for aspartic proteinases. Biochem. J. **237:** 899–906
29 Narita Y., Oda S., Moriyama A., Takenaka O. and Kageyama T. (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. J. Biochem. **121:** 1010–1017
30 Suzuki M., Narita Y., Oda S., Moriyama A., Takenaka O. and Kageyama T. (1999) Purification and characterization of goat pepsinogens and pepsins. Comp. Biochem. Physiol. B **122:** 453–460
31 Narita Y., Oda S., Takenaka O. and Kageyama T. (2001) Phylogenetic position of Insectivora inferred from the cDNA sequences of pepsinogen A and C. Mol. Phylogenet. Evol. **21:** 32–42
32 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
33 Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O. et al. (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
34 Muto N., Yamamoto M. and Tani S. (1987) Comparative studies of two types of acid proteases from rat gastric mucosa and spleen. J. Biochem. **101:** 1069–1075
35 Kageyama T. (1995) Procathepsin E and cathepsin E. Methods Enzymol. **248:** 120–136
36 Foltmann B. (1966) A review on prorennin and rennin. C. R. Trav. Lab. Carlsberg **35:** 143–231
37 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
38 Nielsen P. K. and Foltmann B. (1995) Purification and characterization of porcine pepsinogen B and pepsin B. Arch. Biochem. Biophys. **322:** 417–422
39 Moriyama A., Kageyama T. and Takahashi K. (1983) Identification of monkey lung procathepsin D-II as a pepsinogen-C-like acid protease zymogen. Eur. J. Biochem. **132:** 687–692
40 Szecsi P. B. (1992) The aspartic proteases. Scand. J. Clin. Lab. Invest. **52 (Suppl 210):** 5–22
41 Diamandis E. P., Nadkarni S., Bhaumik B., Abdelrahman A., Melegos D. N., Borchert G. et al. (1997) Immunofluorometric assay of pepsinogen C and preliminary clinical applications. Clin. Chem. **43:** 1365–1371
42 Sánchez L. M., Freije J. P., Merino A. M., Vizoso F., Foltmann B. and López-Otin C. (1992) Isolation and characterization of a pepsin C zymogen produced by human breast tissues. J. Biol. Chem. **267:** 24725–24731
43 Merino A. M., Vazquez J., Rodriguez J. C., Fernandez R., Quintela I., Gonzalez L. O. et al. (2000) Pepsinogen C expression in tumors of extragastric origin. Int. J. Biol. Markers **15:** 165–170
44 Foltmann B., Pedersen V. B., Jacobsen H., Kauffman D. and Wybrandt G. (1977) The complete amino acid sequence of prochymosin. Proc. Natl. Acad. Sci. USA **74:** 2321–2324
45 Kageyama T. and Takahashi K. (1986) The complete amino acid sequence of monkey progastricsin. J. Biol. Chem. **261:** 4406–4419
46 Baudyš M. and Kostka V. (1983) Covalent structure of chicken pepsinogen. Eur. J. Biochem. **136:** 89–99
47 Kageyama T. and Takahashi K. (1986) The complete amino acid sequence of monkey pepsinogen A. J. Biol. Chem. **261:** 4395–4405
48 Andreeva N. S. and James M. N. (1991) Why does pepsin have a negative charge at very low pH? An analysis of conserved charged residues in aspartic proteinases. Adv. Exp. Med. Biol. **306:** 39–45
49 Lin X. L., Wong R. N. and Tang J. (1989) Synthesis, purification, and active site mutagenesis of recombinant porcine pepsinogen. J. Biol. Chem. **264:** 4482–4489
50 Tang J., James M. N., Hsu I. N., Jenkins J. A. and Blundell T. L. (1978) Structural evidence for gene duplication in the evolution of the acid proteases. Nature **271:** 618–621
51 Sielecki A. R., Fujinaga M., Read R. J. and James M. N. (1991) Refined structure of porcine pepsinogen at 1.8 Å resolution. J. Mol. Biol. **219:** 671–692
52 Hartsuck J. A., Koelsch G. and Remington S. J. (1992) The high-resolution crystal structure of porcine pepsinogen. Proteins **13:** 1–25
53 Andreeva N. S., Zdanov A. S., Gustchina A. E. and Fedorov A. A. (1984) Structure of ethanol-inhibited porcine pepsin at 2-Å resolution and binding of the methyl ester of phenylalanyl-diiodotyrosine to the enzyme. J. Biol. Chem. **259:** 11353–11365
54 Cooper J. B., Khan G., Taylor G., Tickle I. J. and Blundell T. L. (1990) X-ray analyses of aspartic proteinases. II. Three-dimensional structure of the hexagonal crystal form of porcine pepsin at 2.3 Å resolution. J. Mol. Biol. **214:** 199–222
55 Sielecki A. R., Fedorov A. A., Boodhoo A., Andreeva N. S. and James M. N. (1990) Molecular and crystal structures of monoclinic porcine pepsin refined at 1.8 Å resolution. J. Mol. Biol. **214:** 143–170
56 Abad-Zapatero C., Rydel T. J. and Erickson J. (1990) Revised 2.3 Å structure of porcine pepsin: evidence for a flexible subdomain. Proteins **8:** 62–81
57 Bateman K. S., Chernaia M. M., Tarasova N. I. and James M. N. (1998) Crystal structure of human pepsinogen A. Adv. Exp. Med. Biol. **436:** 259–263

58 Moore S. A., Sielecki A. R., Chernaia M. M., Tarasova N. I. and James M. N. (1995) Crystal and molecular structures of human progastricsin at 1.62 Å resolution. J. Mol. Biol. **247:** 466–485
59 Gilliland G. L., Winborne E. L., Nachman J. and Wlodawer A. (1990) The three-dimensional structure of recombinant bovine chymosin at 2.3 Å resolution. Proteins **8:** 82–101
60 Newman M., Safro M., Frazao C., Khan G., Zdanov A., Tickle I. J. et al. (1991) X-ray analyses of aspartic proteinases. IV. Structure and refinement at 2.2 Å resolution of bovine chymosin. J. Mol. Biol. **221:** 1295–1309
61 Lin X. L., Lin Y. Z., Koelsch G., Gustchina A., Wlodawer A. and Tang J. (1992) Enzymic activities of two-chain pepsinogen, two-chain pepsin, and the amino-terminal lobe of pepsinogen. J. Biol. Chem. **267:** 17257–17263
62 Wlodawer A. and Gustchina A. (2000) Structural and biochemical studies of retroviral proteases. Biochim. Biophys. Acta **1477:** 16–34
63 Pearl L. H. and Taylor W. R. (1987) A structural model for the retroviral proteases. Nature **329:** 351–354.
64 Rao J. K., Erickson J. W. and Wlodawer A. (1991) Structural and evolutionary relationships between retroviral and eucaryotic aspartic proteinases. Biochemistry **30:** 4663–4671
65 Powers J. C., Harley A. D. and Myers D. V. (1977) Subsite specificity of porcine pepsin. Adv. Exp. Med. Biol. **95:** 141–157
66 Okoniewska M., Tanaka T. and Yada R. Y. (1999) The role of the flap residue, threonine 77, in the activation and catalytic activity of pepsin A. Protein Eng. **12:** 55–61
67 Gustchina E., Rumsh L., Ginodman L., Majer P. and Andreeva N. (1996) Post X-ray crystallographic studies of chymosin: the existence of two structural forms and the regulation of activity by the interaction with the histidine-proline cluster of κ-casein. FEBS Lett. **379:** 60–62
68 Cottrell T. J., Harris L. J., Tanaka T. and Yada R. Y. (1995) The sole lysine residue in porcine pepsin works as a key residue for catalysis and conformational flexibility. J. Biol. Chem. **270:** 19974–19978
69 Takahashi K. and Kageyama T. (1985) Multiplicity and intermediates of the activation mechanism of zymogens of gastric aspartic proteinases. In: Aspartic Proteinases and Their Inhibitors, pp. 265–282, Kostka V. (ed.), de Gruyter, Berlin
70 Richter C., Tanaka T. and Yada R. Y. (1998) Mechanism of activation of the gastric aspartic proteinases: pepsinogen, progastricsin and prochymosin. Biochem. J. **335:** 481–490
71 Auer H. E. and Glick D. M. (1984) Early events of pepsinogen activation. Biochemistry **23:** 2735–2739
72 Glick D. M., Auer H. E., Rich D. H., Kawai M. and Kamath A. (1986) Pepsinogen activation: genesis of the binding site. Biochemistry **25:** 1858–1864
73 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
74 Kageyama T. and Takahashi K. (1983) Occurrence of two different pathways in the activation of porcine pepsinogen to pepsin. J. Biochem. **93:** 743–754
75 Kay J. and Dykes C. W. (1977) The first cleavage site in pepsinogen activation. Adv. Exp. Med. Biol. **95:** 103–127
76 Christensen K. A., Pedersen V. B. and Foltmann B. (1977) Identification of an enzymatically active intermediate in the activation of porcine pepsinogen. FEBS Lett. **76:** 214–218
77 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
78 Pedersen V. B., Christensen K. A. and Foltmann B. (1979) Investigations on the activation of bovine prochymosin. Eur. J. Biochem. **94:** 573–580
79 Khan A. R., Khazanovich-Bernstein N., Bergmann E. M. and James M. N. (1999) Structural aspects of activation pathways of aspartic protease zymogens and viral 3C protease precursors. Proc. Natl. Acad. Sci. USA **96:** 10968–10975
80 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
81 Pichová I., Pohl J., Štrop P. and Kostka V. (1985) Activation of chicken pepsinogen and chicken pepsin-propart peptide (p1–p42) complex. In: Aspartic Proteinases and Their Inhibitors, pp. 301–308, Kostka V. (ed.), de Gruyter, Berlin
82 al-Janabi J., Hartsuck J. A. and Tang J. (1972) Kinetics and mechanism of pepsinogen activation. J. Biol. Chem. **247:** 4628–4632
83 Kageyama T. and Takahashi K. (1987) Activation mechanism of monkey and porcine pepsinogens A: one-step and stepwise activation pathways and their relation to intramolecular and intermolecular reactions. Eur. J. Biochem. **165:** 483–490
84 Bustin M. and Conway-Jacobs A. (1971) Intramolecular activation of porcine pepsinogen. J. Biol. Chem. **246:** 615–620
85 Marciniszyn J. Jr, Huang J. S., Hartsuck J. A. and Tang J. (1976) Mechanism of intramolecular activation of pepsinogen: evidence for an intermediate delta and the involvement of the active site of pepsin in the intramolecular activation of pepsinogen. J. Biol. Chem. **251:** 7095–7102
86 Kageyama T. (1988) Analysis of the activation of pepsinogen in the presence of protein substrates and estimation of the intrinsic proteolytic activity of pepsinogen. Eur. J. Biochem. **176:** 543–549
87 Athauda S. B., 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
88 Schlamowitz M. and Peterson L. U. (1959) Studies on the optimum pH for the action of pepsin on 'native' and denatured bovine serum albumin and bovine hemoglobin. J. Biol. Chem. **234:** 3137–3145
89 Dunn B. M., Valler M. J., Rolph C. E., Foundling S. I., Jimenez M. and Kay J. (1987) The pH dependence of the hydrolysis of chromogenic substrates of the type, Lys-Pro-Xaa-Yaa-Phe-(NO_2)Phe-Arg-Leu, by selected aspartic proteinases: evidence for specific interactions in subsites S3 and S2. Biochim. Biophys. Acta **913:** 122–130
90 Sampath-Kumar P. S. and Fruton J. S. (1974) Studies on the extended active sites of acid proteinases. Proc. Natl. Acad. Sci. USA **71:** 1070–1072
91 Dunn B. M. and Hung S. (2000) The two sides of enzyme-substrate specificity: lessons from the aspartic proteinases. Biochim. Biophys. Acta **1477:** 231–240.
92 Kageyama T. (2000) New world monkey pepsinogens A and C, and prochymosins: purification, characterization of enzymatic properties, cDNA cloning, and molecular evolution. J. Biochem. **127:** 761–770
93 Tang J. (1970) Gastricsin and pepsin. Methods Enzymol. **19:** 406–421
94 Guillou H., Miranda G. and Pelissier J.-P. (1991) Hydrolysis of β-casein by gastric proteases. I. Comparison of proteolytic action of bovine chymosin and pepsin A. Int. J. Pept. Protein Res. **37:** 494–501.
95 Jensen T., Axelsen N. H. and Foltmann B. (1982) Isolation and partial characterization of prochymosin and chymosin from cat. Biochim. Biophys. Acta **705:** 249–256
96 Kageyama T., Ichinose M., Tsukada-Kato S., Omata M., Narita Y., Moriyama A. et al. (2000) Molecular cloning of neonate/infant-specific pepsinogens from rat stomach mucosa and their expressional change during development. Biochem. Biophys. Res. Commun. **267:** 806–812

97 Suzuki J., Sasaki K., Sasao Y., Hamu A., Kawasaki H., Nishiyama M. et al. (1989) Alteration of catalytic properties of chymosin by site-directed mutagenesis. Protein Eng. **2:** 563–569

98 Pitts J. E., Dhanaraj V., Dealwis C. G., Mantafounis D., Nugent P., Orprayoon P. et al. (1992) Multidisciplinary cycles for protein engineering: site-directed mutagenesis and X-ray structural studies of aspartic proteinases. Scand. J. Clin. Lab. Invest. **52 (Suppl 210):** 39–50

99 Foltmann B. (1970) Prochymosin and chymosin. Methods Enzymol. **19:** 421–436

100 Jollès J., Alais C. and Jollès P. (1968) The tryptic peptide with the renin-sensitive linkage of cow's κ- casein. Biochim. Biophys. Acta **168:** 591–593

101 Williams M. G., Wilsher J., Nugent P., Mills A., Dhanaraj V., Fabry M. et al. (1997) Mutagenesis, biochemical characterization and X-ray structural analysis of point mutants of bovine chymosin. Protein Eng. **10:** 991–997

102 Foltmann B., Jensen A. L., Lønblad P., Smidt E. and Axelsen N. H. (1981) A developmental analysis of the production of chymosin and pepsin in pigs. Comp. Biochem. Physiol. B **68:** 9–13

103 Rich D. H. and Sun E. T. (1980) Mechanism of inhibition of pepsin by pepstatin: effect of inhibitor structure on dissociation constant and time-dependent inhibition. Biochem. Pharmacol. **29:** 2205–2212

104 Ikuzawa M., Inokuchi T., Kobayashi K. and Yasumasu S. (2001) Amphibian pepsinogens: purification and characterization of *Xenopus* pepsinogens, and molecular cloning of *Xenopus* and bullfrog pepsinogens. J. Biochem. **129:** 147–153

105 Abu-Erreish G. M. and Peanasky R. J. (1974) Pepsin inhibitors from *Ascaris lumbricoides*: pepsin-inhibitor complex: stoichiometry of formation, dissociation, and stability of the complex. J. Biol. Chem. **249:** 1566–1571

106 Kageyama T. (1998) Molecular cloning, expression and characterization of an *Ascaris* inhibitor for pepsin and cathepsin E. Eur. J. Biochem. **253:** 804–809

107 Ng K. K., Petersen J. F., Cherney M. M., Garen C., Zalatoris J. J., Rao-Naik C. et al. (2000) Structural basis for the inhibition of porcine pepsin by *Ascaris* pepsin inhibitor-3. Nat. Struct. Biol. **7:** 653–657

108 Abuereish G. M. (1998) Pepsin inhibitor from roots of *Anchusa strigosa*. Phytochemistry **48:** 217–221

109 Xie S. C., Low B. G., Nagel R. J., Kramer K. K., Anthony R. V., Zoli A. P. et al. (1991) Identification of the major pregnancy-specific antigens of cattle and sheep as inactive members of the aspartic proteinase family. Proc. Natl. Acad. Sci. USA **88:** 10247–10251

110 Green J. A., Xie S. and Roberts R. M. (1998) Pepsin-related molecules secreted by trophoblast. Rev. Reprod. **3:** 62–69

111 Green J. A., Xie S., Quan X., Bao B., Gan X., Mathialagan N. et al. (2000) Pregnancy-associated bovine and ovine glycoproteins exhibit spatially and temporally distinct expression patterns during pregnancy. Biol. Reprod. **62:** 1624–1631

112 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

113 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

114 Örd T., Kolmer M., Villems R. and Saarma M. (1990) Structure of the human genomic region homologous to the bovine prochymosin-encoding gene. Gene **91:** 241–246

115 Ishihara T., Ichihara Y., Hayano T., Katsura I., Sogawa K., Fujii-Kuriyama Y. et al. (1989) Primary structure and transcriptional regulation of rat pepsinogen C gene. J. Biol. Chem. **264:** 10193–10199

116 Hidaka M., Sasaki K., Uozumi T. and Beppu T. (1986) Cloning and structural analysis of the calf prochymosin gene. Gene **43:** 197–203

117 Sakamoto N., Saiga H. and Yasugi S. (1998) Analysis of temporal expression pattern and cis-regulatory sequences of chicken pepsinogen A and C. Biochem. Biophys. Res. Commun. **250:** 420–424

118 Hayashi K., Yasugi S. and Mizuno T. (1988) Isolation and structural analysis of embryonic chicken pepsinogen gene: avian homologue of prochymosin gene. Biochem. Biophys. Res. Commun **152:** 776–782

119 Takahashi K. (1992) Gene structures of pepsinogens A and C. Scand. J. Clin. Lab. Invest. **52 (Suppl 210):** 97–110

120 Lu Q., Wolfe K. H. and McConnell D. J. (1988) Molecular cloning of multiple bovine aspartyl protease genes. Gene **71:** 135–146

121 Sakamoto N., Fukuda K., Watanuki K., Sakai D., Komano T., Scotting P. J. et al. (2000) Role for cGATA-5 in transcriptional regulation of the embryonic chicken pepsinogen gene by epithelial-mesenchymal interactions in the developing chicken stomach. Dev. Biol. **223:** 103–113

122 Meijerink P. H., Bebelman J. P., Oldenburg A. M., Defize J., Planta R. J., Eriksson A. W. et al. (1993) Gastric chief cell-specific transcription of the pepsinogen A gene. Eur. J. Biochem. **213:** 1283–1296

123 Ichinose M., Miki K., Furihata C., Tatematsu M., Ichihara Y., Ishihara T. et al. (1988) DNA methylation and expression of the rat pepsinogen gene in embryonic, adult, and neoplastic tissues. Cancer Res. **48:** 1603–1609

124 Ichinose M., Miki K., Wong R. N., Tatematsu M., Furihata C., Konishi T. et al. (1991) Methylation and expression of human pepsinogen genes in normal tissues and their alteration in stomach cancer. Jpn. J. Cancer Res. **82:** 686–692

125 Fukamachi H., Ichinose M., Tsukada S., Kakei N., Suzuki T., Miki K. et al. (1994) Hepatocyte growth factor region specifically stimulates gastro-intestinal epithelial growth in primary culture. Biochem. Biophys. Res Commun. **205:** 1445–1451

126 Fukuda K., Ishii Y., Saiga H., Shiokawa K. and Yasugi S. (1994) Mesenchymal regulation of epithelial gene expression in developing avian stomach: 5′-flanking region of pepsinogen gene can mediate mesenchymal influence on its expression. Development **120:** 3487–3495

127 Furihata C., Kawachi T. and Sugimura T. (1972) Premature induction of pepsinogen in developing rat gastric mucosa by hormones. Biochem. Biophys. Res. Commun. **47:** 705–711

128 Tsukada S., Ichinose M., Yahagi N., Matsubara Y., Yonezawa S., Shiokawa K. et al. (1998) Induction of precocious pepsinogen synthesis by glucocorticoids in fetal rat gastric epithelium in organ culture: importance of mesenchyme for epithelial differentiation. Differentiation **62:** 239–247

129 Andrén A. (1992) Production of prochymosin, pepsinogen and progastricsin, and their cellular and intracellular localization in bovine abomasal mucosa. Scand. J. Clin. Lab. Invest. **52 (Suppl 210):** 59–64

130 Miura H., Yamada J., Kitamura N., Yamashita T. and Andrén A. (1988) An immunohistochemical study of prochymosin- and pepsinogen-immunoreactive cells in the abomasal mucosa of cattle fetuses. Z. Mikrosk. Anat. Forsch. **102:** 101–110

131 Andrén A., Björck L. and Claesson O. (1982) Immunohistochemical studies on the development of prochymosin- and pepsinogen-containing cells in bovine abomasal mucosa. J. Physiol. **327:** 247–254

132 Andrén A. and Björck L. (1986) Milk-feeding maintains the prochymosin production in cells of bovine abomasal mucosa. Acta. Physiol. Scand. **126:** 419–427

133 Yasugi S., Mizuno T. and Esumi H. (1979) Changes in molecular species of pepsinogens in the development of the chick. Experientia **35:** 814–815

134 Inokuchi T., Kobayashi K. I. and Horiuchi S. (1991) Acid proteinases of the fore-gut in metamorphosing tadpoles of *Rana catesbeiana*. Comp. Biochem. Physiol. B **99:** 653–662

135 Morris I. G. (1968) Gamma globulin absorption in the newborn. In: Handbook of Physiology, Sec. 6, Alimentary canal, vol. 3, pp. 1491–1512, Code C. F. (ed.), Williams & Wilkins, Baltimore
136 Farrell J. H. M. (1989) Milk, composition and synthesis. In: Encyclopedia of Reproduction, vol. 3, pp. 256–263, Knobil E. and Neill J. D. (eds), Academic Press, San Diego
137 Reid W. A., McGechaen K., Branch T., Gray H. D., Thompson W. D. and Kay J. (1989) Immunolocalisation of aspartic proteinases in the developing human stomach. J. Dev. Physiol. **11:** 299–303
138 Hirsch-Marie H., Loisillier F., Touboul J. P. and Burtin P. (1976) Immunochemical study and cellular localization of human pepsinogens during ontogenesis and in gastric cancers. Lab. Invest. **34:** 623–632
139 Tang J., Wolf S., Caputto R. and Trucco R. E. (1959) Isolation and crystallization of gastricsin from human gastric juice. J. Biol. Chem. **234:** 1174–1178
140 Valler M. J., Kay J., Aoyagi T. and Dunn B. M. (1985) The interaction of aspartic proteinases with naturally-occurring inhibitors from actinomycetes and *Ascaris lumbricoides*. J. Enzyme Inhib. **1:** 77–82
141 Foltmann B., Szecsi P. B. and Tarasova N. I. (1985) Detection of proteases by clotting of casein after gel electrophoresis. Anal. Biochem. **146:** 353–360.
142 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
143 Foltmann B., Drøhse H. B., Nielsen P. K. and James M. N. (1992) Separation of porcine pepsinogen A and progastricsin: sequencing of the first 73 amino acid residues in progastricsin. Biochim. Biophys. Acta **1121:** 75–82
144 Tanji M., Kageyama T. and Takahashi K. (1988) Tuna pepsinogens and pepsins: purification, characterization and amino-terminal sequences. Eur. J. Biochem. **177:** 251–259
145 Sanger F. and Tuppy H. (1951) The amino acid sequence in the phenylalanyl chain of insulin. Biochem. J. **49:** 481–490
146 Ryle A. P. and Porter R. R. (1959) Parapepsins: two proteolytic enzymes associated with porcine pepsin. Biochem. J. **73:** 75–86
147 Matsuzaki O. and Takahashi K. (1988) Improved purification of slow moving protease from human gastric mucosa and its action on the B chain of oxidized bovine insulin. Biomed. Res. **9:** 515–523
148 Hirasawa A., Athauda S. B. and Takahashi K. (1996) Purification and characterization of turtle pepsinogen and pepsin. J. Biochem. **120:** 407–414