International Immunology, Vol. 21, No. 1, pp. 95–104
doi:10.1093/intimm/dxn128

The extended substrate specificity of the human mast cell chymase reveals a serine protease with well-defined substrate recognition profile

Mattias K. Andersson[1], Mattias Enoksson[2,*], Maike Gallwitz[1,*] and **Lars Hellman[1]**

[1]Department of Cell and Molecular Biology, Uppsala University, Uppsala, The Biomedical Center, Box 596, SE-751 24 Uppsala, Sweden
[2]Department of Medicine, Clinical Immunology and Allergy Unit, Karolinska Institute, SE-17176 Stockholm, Sweden

Keywords: chymase, cleavage specificity, human chymase, mast cell

Abstract

The human chymase (HC) is a major granule constituent of mast cells (MCs) residing in the connective tissue and the sub-mucosa. Although many potential substrates have been described for this important MC enzyme, its full range of *in vivo* substrates has most likely not yet been identified. A major step toward a better understanding of the function of the HC is therefore to determine its extended cleavage specificity. Using a phage-displayed random nonapeptide library, we show that the HC has a rather stringent substrate recognition profile. Only aromatic amino acids (aa) are accepted in position P1, with a strong preference for Tyr and Phe over Trp. Aliphatic aa are preferred in positions P2 to P4 N-terminal of the cleaved bond. In the P1′ position C-terminal of the cleaved bond, Ser is clearly over-represented and acidic aa Asp and Glu are strongly preferred in the P2′ position. In P3′, the small aliphatic aa Ala, Val and Gly were frequently observed. The consensus sequence, from P4 to P3′: Gly/Leu/Val–Val/Ala/Leu–Ala/Val/Leu–Tyr/Phe–Ser–Asp/Glu–Ala/Val/Gly, provides an instrument for the identification of novel *in vivo* substrates for the HC. Interestingly, a very similar cleavage specificity was recently reported for the major chymase in mouse connective tissue mast cells (CTMCs), the β-chymase mouse mast cell protease-4, suggesting functional homology between these two enzymes. This indicates that a rather stringent chymotryptic substrate recognition profile has been evolutionary conserved for the dominant CTMC chymase in mammals.

Introduction

The human chymase (HC) is a serine protease with chymotrypsin-like cleavage specificity that is stored in the granules of mast cells (MCs) of the connective tissue and the sub-mucosa [connective tissue mast cell (CTMC)] (1). In these cells, the HC is one of the most abundant granule-stored proteins, together with tryptase and carboxypeptidase A. Due to their tryptase- and chymase-positive phenotype, the human CTMCs are also named MC_{TC}. Human mucosal MCs, in contrast, lack chymase but express tryptase and are therefore designated MC_T.

The gene encoding the HC is located in the chymase locus on human chromosome 14, together with the genes for neutrophil cathepsin G and granzymes B and H. Both granzyme genes are expressed in the NK/T lineage (2). During mammalian evolution, MC chymase seems to have been the first of the enzymes encoded within the chymase locus to obtain its present cleavage specificity. This conclusion is substantiated by the presence of a clear α-chymase gene in a marsupial, the American opossum (*Monodelphis domestica*) (3). The presence of an ancestral α-chymase gene seems therefore to date back at least 185 million years, to the last common ancestor of marsupials and eutherian mammals (4).

The single HC is an α-chymase. The function of the α-chymase is probably both ancient and important, as suggested by the finding that at least one α-chymase gene is present in all mammals analyzed to date (4–6). This MC enzyme seems to possess evolutionary conserved chymotryptic activity in species where α-chymase is the sole chymase in MCs, i.e. most mammals except for rodents. For example, apart from the HC, the MC chymase in dog and cattle has been shown to display chymotryptic activity (7–9). All three

*These authors contributed equally to this study.

Correspondence to: L. Hellman; E-mail: lars.hellman@icm.uu.se

Received 26 June 2008, *accepted* 10 November 2008

Transmitting editor: S. J. Galli

Advance Access publication 10 December 2008

chymases hold the residues Ser-Gly-Ala (S-G-A) in the specificity-conferring triplet formed by amino acids (aa) 189, 216 and 226 (4). Most of the less well-characterized mammalian α-chymases are very likely to be chymotryptic as well because all mammalian α-chymases analyzed to date, with the exception of rodents, hold the identical specificity-conferring triplet S-G-A (4, 10). In contrast, the mouse and rat α-chymases mouse mast cell protease (mMCP)-5 and rat mast cell protease (rMCP)-5, respectively, possess a differing triplet Asn-Val-Ala (N-V-A) and have in line with this secondarily adopted elastase-like specificity (11, 12). Hamster and gerbil hold α-chymases with very similar triplets (N-V-S and N-V-A, respectively). The hamster, and probably also the gerbil, enzymes have therefore also elastase activity (10, 13). Interestingly, the recently characterized guinea pig α-chymase has a specificity-conferring triplet consisting of S-A-V and is leucine specific (14). However, most rodent MCs are equipped with chymotryptic enzymes. Namely, they express a second subfamily of rodent-specific MC chymases, the β-chymases [reviewed in ref. (15)]. The single exception is the guinea pig, which seems to lack expression of β-chymases (14). Several of the β-chymases have been shown to have chymotryptic activity (16–18). Their presence could thus compensate for the lost chymotryptic activity of α-chymases in rodents (15).

The functional importance of chymotryptic activity in rodent MCs is further underlined by a study of 17 rodent subspecies, where chymase-positive MCs were detected in both the jejunum and tongue of all species, whereas the activity of the second characteristic MC enzyme, tryptase, was absent from intestinal MCs of eight subspecies (19).

This collected information indicates that the chymase has an important and evolutionary conserved function in MC biology. The interest in the identification of *in vivo* substrates for the HC has therefore been considerable, and extensive screenings have identified several potential substrates. The most well-characterized among them is angiotensin (Ang) I, which is converted by HC to Ang II. Although the conversion of Ang I is a prominent function of HC, Ang I is not the only substrate for this enzyme and may not be the most important *in vivo*. Other targets include the connective tissue component, fibronectin, which is cleaved by both HC and its apparent functional homologue in mouse, the prominent CTMC chymase mMCP-4 (20, 21). In addition, HC processes endothelin-1, which possibly has very important effects on blood pressure homeostasis (22, 23). Moreover, HC cleaves cell surface-bound human stem cell factor and thereby induces the release of active growth factor from the producer cell (24). The putative functional homologues for the HC both in dog (dog chymase) and mouse (the β-chymase mMCP-4) have been shown to activate matrix metalloprotease 9 (25, 26). These findings suggest an important role for chymase in extracellular matrix remodeling, cytokine activation and vasoregulation (24, 27–31). However, these results also indicate that MC chymases have a relatively broad spectrum of substrates, making it difficult to assign them a clear *in vivo* role. In addition, the importance of most of these substrates *in vivo* is still not known. Additional substrates are likely to be identified, possibly including a major *in vivo* substrate.

One way to identify novel targets to an enzyme and to thereby reveal new functions is to determine the enzyme's cleavage specificity. This information can then be used to search peptide databases for possible substrates. Different methods have previously been used to elucidate the N-terminal positions of the substrate specificity of HC, including panels of chromogenic peptides or combinatorial libraries (16, 32). In studies using chromogenic peptides, aa residues are varied in one position at a time, whereas combinatorial libraries leave one position fixed, while the others are varied. For example, peptide libraries have been used to map the N-terminal and the C-terminal positions of substrate–HC interactions, although in separate reactions (33). However, none of these methods varies both N- and C-terminal residues simultaneously. In contrast, phage display methodology makes it possible to determine the preference for individual residues throughout the entire substrate-binding region by allowing variation of aa residues in all positions simultaneously. No leaving groups that are foreign to natural peptides are present, and the selection of susceptible peptides is influenced solely by the interactions between neighboring aa residues of the substrate and the enzyme. The best combination of residues can therefore be selected to give reliable representation of the extended substrate specificity of the selected enzyme.

In this communication, we present an analysis of pure recombinant HC using a random nonapeptide library displayed on T7 phages to determine the extended cleavage specificity of the HC. The specificity was analyzed for seven individual positions from P4 to P3′. The HC is shown to display a relatively stringent substrate recognition profile, characterized by aliphatic aa N-terminal of the cleaved bond, high preference for the aromatic aa Phe or Tyr in the P1 position and acidic aa Glu or Asp in the P2′ position.

Materials and methods

Production and purification of recombinant HC

The coding region for the mature HC was obtained by PCR amplification using full-length cDNA as template. For purification purposes, the 5′-primers contained a sequence encoding six His residues and an enterokinase (EK)-susceptible peptide (Asp-Asp-Asp-Asp-Lys). The EK cleavage site enables subsequent removal of the histidine (His_6) tag and thereby activation of the protease. The PCR product was inserted into the pCEP-Pu2 vector (34, 35), and the nucleotide sequence was confirmed using an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) with vector-specific primers. Human embryonic kidney cells, 293-EBNA, were transfected with pCEP-Pu2/HC constructs as previously described (34, 35). Selection was initiated by the addition of 1.5 μg ml^{-1} puromycin to the cell culture medium (DMEM supplemented with 5% FCS, 50 μg ml^{-1} gentamicin and 5 μg ml^{-1} heparin). The level of puromycin was decreased to 0.5 μg ml^{-1} after 7 days of selection.

Conditioned medium was collected and centrifuged to remove cell debris, followed by the addition of 0.5 ml nickel–nitrilotriacetic acid (Ni–NTA) agarose beads (Qiagen, GmbH, Hilden, Germany) per 1 liter of medium. After 3 h of incubation with gentle agitation at 4°C, the beads were

pelleted by centrifugation and transferred to 1.5-ml reaction tubes (TreffLab, Degersheim, Switzerland). Collected Ni–NTA beads were washed three times with washing buffer (1 M NaCl and 0.1% Tween 20 in PBS). Bound protein was then eluted with elution buffer (100 mM imidazole and 0.2% TritonX-100 in PBS). Protein purity and concentration was estimated by separation on 12.5% SDS–PAGE gels. Protein samples were mixed with sample buffer, and β-mercaptoethanol was added to a final concentration of 5%. To visualize protein bands, the gel was stained with Coomassie Brilliant Blue.

Activation and purification of recombinant HC

Recombinant HC was digested for 5 h with EKMax™ EK (Invitrogen), 1 U/10 μg of recombinant protease.

In order to remove EK and other impurities, EK-digested HC was purified by affinity chromatography on heparin–sepharose columns as described previously (18). Shortly, a 10-ml polyprep chromatography column containing 0.2 ml heparin–sepharose (Amersham Pharmacia Biotech) was equilibrated with PBS (pH 7.2). Approximately 30 μg of EK-cleaved HC in PBS (pH 7.2) was applied to the column, followed by washing with 0.3 M NaCl in PBS and subsequent elution with 1 M NaCl in PBS. Enzymatic activity was measured toward the chromogenic substrate S-2586 (MeO-Suc-Arg-Ala-Tyr-pNA) (Chromogenix, Mölndal, Sweden). Standard measurements were performed in 96-well microtiter plates with a substrate concentration of 0.18 mM in 200 μl PBS. S-2586 hydrolysis was monitored spectrophotometrically at 405 nm in a Multiscan MCC/340 spectrophotometer (Labsystem, Helsinki, Finland). The protein content of flow through, wash and eluted fractions was analyzed on SDS–PAGE gels.

Determination of cleavage recognition profile by phage-displayed nonapeptides

A library of phage-displayed nonameric peptides was used as previously described (12, 18, 21). The C-terminus of the capsid protein 10 of the T7 phage is manipulated in to contain a nine-aa-long random peptide followed by a His_6 tag (18). Via this His_6 tag, the phages can be anchored to Ni–NTA beads. Phages with a random peptide that is susceptible to protease cleavage are released from the Ni–NTA matrix and can be removed and subsequently amplified. An aliquot of the amplified phages ($\sim 10^9$ pfu) was allowed to bind to 100 μl Ni–NTA agarose beads for 1 h while rotating gently at 4°C. Unbound phages were removed by washing 10 times with 1.5 ml 1 M NaCl, 0.1% Tween 20 in PBS, pH 7.2, and two subsequent washes with 1.5 ml PBS. The beads were finally re-suspended in 1000 μl PBS. Activated and heparin–sepharose-purified HC (~0.1 μg) was added to the re-suspended beads to start the selection of susceptible peptides. Buffer without protease was used as a control. Protease digestion was allowed to proceed at room temperature overnight with gentle agitation. To recover released phages, the Ni–NTA agarose beads were pelleted by centrifugation in a tabletop centrifuge and the supernatant containing the phages was removed. To ensure that all the released phages were recovered, the beads were re-suspended in 100 μl PBS (pH 7.2), and the supernatant after mixing and centrifugation was added to the first recovered phages. A control elution of the phages still bound to the beads, using 100 μl 100 mM imidazole, concluded that at least 1×10^8 phages were attached to the matrix during each selection. To ensure that the His_6 tag had been hydrolyzed on all phages recovered after protease digestion, 15 μl fresh Ni–NTA agarose beads were added to the phage suspension and the mixture agitated for 15 min followed by centrifugation to recover the supernatant. Thirty microliters of the supernatant was used to determine the amount of phages detached in each round of selection. The remaining 1070 μl of the supernatant was added to a 10-ml culture (optical density ~0.6) of *Escherichia coli* (BLT5615). The bacteria had previously been induced to produce the T7 phage capsid protein by adding 100 μl 100 mM isopropyl β-D-1-thiogalactopyranoside 30 min prior to phage addition. The bacteria lysed ~75 min after phage addition. The lysate was centrifuged to remove cell debris and 900 μl of the phage sublibrary was added to 100 μl fresh agarose beads. After binding and washing the sublibrary, a new round of selection was started. Following five rounds of selection, 50 plaques were arbitrarily isolated from luria broth plates and each dissolved in phage extraction buffer (100 mM NaCl and 6 mM $MgSO_4$ in 20 mM Tris–HCl, pH 8.0) and vigorously shaken for 30 min in order to extract the phages from the agarose. The phage DNA was then amplified by PCR using primers flanking the variable region of the gene encoding the modified T7 phage capsid protein. After amplification, PCR fragments were purified using the QIAquick PCR Purification Kit (Qiagen). Purified PCR fragments were then sequenced using an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems). In order to trace potential EK contamination in the purified HC fraction, a control experiment subjecting the phage library to five biopannings of 0.5 U EK was performed (data not shown). In this analysis, we obtained almost only the characteristic negatively charged aa followed by a lysine or arginine that is favored by the EK. No such sequences were found among the sequences retrieved with the HC.

Generation of a consensus sequence from sequenced phage inserts

Phage insert sequences were aligned by hand assuming a preference for aromatic aa in position P1 or with the bioinformatics tool Consense (A. Kaplan and M. Gallwitz, unpublished results) using an unbiased algorithm searching for patterns of at least three aa in a row. For the manual alignment, sequences with only one aromatic aa were aligned first and sequences with more than one possible cleavage site were then aligned to fit this pattern. Both methods yielded identical alignments. Aa of approximately functional equivalence were for both methods grouped as follows: aromatic aa (Phe, Tyr and Trp); negatively charged aa (Asp and Glu); positively charged aa (Lys and Arg), small aliphatic aa (Gly and Ala); larger aliphatic aa (Val, Leu, Ile and Pro) and hydrophilic aa (Ser, Thr, His, Asn, Gln, Cys and Met). The nomenclature by Schechter and Berger (36) was adopted to designate the aa in the substrate cleavage region, where P1–P1′ corresponds to the scissile bond.

Cleavage of a synthetic peptide based on the consensus sequence

A synthetic peptide containing the consensus sequence was designed and purchased from GenScript (Piscataway, NY, USA). The 10-aa peptide with the sequence Ala-Ala-Gly-Val-Ala-Tyr-Ser-Glu-Ala-Ala and with a purity of 95% was dissolved in Me_2SO. The consensus peptide (0.44 mM) was subjected to cleavage by HC (12.4 nM) at 37°C in PBS and 5% Me_2SO, pH 7.2. Samples for HPLC analysis were taken at time points 15, 30, 60 and 150 min of cleavage. The enzyme was inactivated by the addition of trifluoroacetic acid (TFA) (1% final concentration) and the samples were kept frozen until analysis. The cleavage products were analyzed by HPLC on a Hypersil GOLD Phenyl Column (5 μm, 250 × 4.6 mm, Thermo Scientific). The samples were separated in a linear gradient of acetonitrile (10–90%) and 0.1% TFA over 90 min. The samples were monitored at 229 nm and the fractions corresponding to peaks were collected and analyzed by mass spectrometry.

Results

Production and purification of recombinant HC

The HC was expressed as a fusion protein with an N-terminal His_6 tag and an EK-susceptible peptide (Asp-Asp-Asp-Asp-Lys). Following purification on Ni–NTA agarose beads, the fusion protein was digested by EK to remove the His_6 tag along with the EK site and to activate the protease. Activated HC was further purified from contaminating proteins, EK and imidazole by passing over a heparin–sepharose column. On SDS–PAGE gels, the EK-digested HC has an apparent size of ~33 kDa, somewhat over the theoretical molecular weight of 25 kDa (Fig. 1). This may partly be due to glycosylation at the identified hypothetical N-glycosylation sites N72 and N95 [chymotrypsin numbering according to ref. (37)]. In addition, the HC has a highly positive net charge (+13.2 at pH 7), which may confer an increased apparent molecular weight in SDS–PAGE analysis. The inactive protease migrates as a 35 kDa band, where the His_6 tag and EK site contribute a calculated 1.5 kDa to the molecular weight of the protein (Fig. 1).

Activity analysis of recombinant HC

The HC is known to have chymotrypsin-like cleavage specificity, and the chymotryptic activity of the recombinant protein was confirmed using the chromogenic chymotrypsin substrate S-2586 (MeO-Suc-Arg-Ala-Tyr-pNA, Chromogenix) prior to the phage display analysis. The EK-digested enzyme, but not its proform, showed good proteolytic activity against S-2586 (data not shown).

Substrate recognition profile of HC as determined using a phage-displayed nonapeptide library

The extended substrate specificity of HC was determined by phage display analysis using a T7 phage library, where each phage contains a nine-aa-long random peptide followed by a His_6 tag. This library was subjected to HC or PBS (control) over five rounds of selection. The experiment was performed under physiologic salt and pH conditions, to mimic most *in vivo* conditions. We are, however, aware of that the enzymatic activity and the extended cleavage specificity of the enzyme may be different at other pH and salt concentrations. Such conditions may occur in inflamed hypoxic tissue where, for example pH values may be lower.

During the experiment, the number of released phages after each selection increased steadily, compared with the control, with each selection round (Fig. 2). After five rounds of selection, we observed a 67-fold increase in the ratio of phages released from the library subjected to HC, compared with PBS control. After the fifth selection round, 50 DNA sequences were sampled from individually isolated phage plaques. Forty-eight of these were then aligned

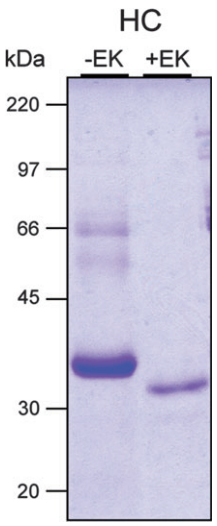

Fig. 1. Purification and activation of recombinant HC. Recombinant HC containing an N-terminal His_6 tag and EK site was purified on Ni–NTA agarose as described in experimental procedures. The purified protease was activated by removal of the His_6 tag through EK digestion. EK-digested protease was purified over heparin–sepharose columns to remove EK and imidazole. Samples of HC purified on Ni–NTA agarose (−EK) and following EK digestion (+EK) were analyzed by separation on SDS–PAGE (12.5% polyacrylamide) and visualized by Coomassie Brilliant Blue staining.

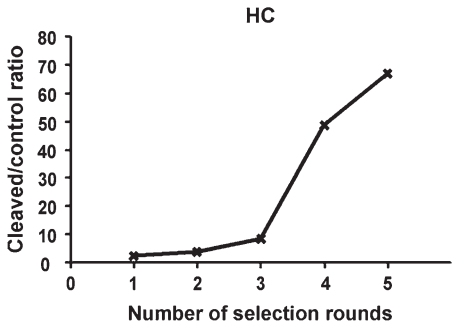

Fig. 2. Ratio of T7 phages released after digestion with HC versus PBS control. A library of T7 phages containing nine-aa-long random peptides in the C-terminus of capsid protein 10 was subjected to cleavage with HC. After cleavage, the released phages were amplified and subjected to the next round of selection. The ratios of released phages compared with a PBS control after each biopanning are shown.

(Fig. 3). Two background sequences were not included in the alignment, one lacking aromatic aa and one containing a stop codon (Leu-Leu-Gly-Val-Gly-Val-Val-Arg-His and Gly-Ser-Arg-Thr-Ala-Ala-STOP-Trp-Arg, respectively). Based on the shown alignment, the prevalence of aa in each position was analyzed (Fig. 4). Moreover, we compared the total number of each aa represented in positions P4 to P3′ with the randomly expected number of this aa with regard to its codon frequency [to eliminate stop codons, only T and G are utilized in the third position of each triplet encoding the random nonapeptide (18)]. This analysis reveals a clear over-representation of Phe and Tyr, but not of the bulky Trp among the aromatic aa (Fig. 5). The small and medium-sized aliphatic aa Gly, Ala and Val are also over-represented, whereas the larger Ile is under-represented. Pro is heavily under-represented, as are Cys, Asn and to a lesser extent Gln. The positively charged Lys and Arg are clearly under-represented, with Lys being completely excluded from the analyzed sequences (four to five Lys would be expected among the retrieved 18 positive aa with respect to the codon frequency of Lys and Arg). On the contrary, the negatively charged Asp and even more pronounced Glu are present in higher numbers than randomly expected.

As to each single position (Figs 3 and 4), position P1 is strongly dominated by Tyr and Phe residues, and these two aa are present in equal proportions. Only 8% of the sequences contain Trp in this position. Interestingly, Trp occurs exclusively in sequences containing several aromatic aa. In addition, only a fraction of the Trp residues found in the sequences (22%) are aligned in the P1 position, whereas this is the case for the majority of the Tyr (79%) and Phe (54%) residues. Thus, Tyr has the strongest preference for the P1 position among the aromates, although Phe displays the highest over-representation in total.

In positions P2, P3 and P4, aliphatic aa are strongly preferred. Position P2 is occupied by aliphatic aa in 58% of the sequences, evenly distributed among Ala, Val and Leu. However, the largest of these three aliphatic aa, Leu, occurs more frequently in sequences containing only one aromatic aa and less frequently in sequences with several aromatic aa.

In the P3 position, Val is the single most preferred aa, and 54% of the aa are aliphatic in total. Negatively charged aa are also tolerated in this position (15%), whereas positively charged aa appear to be excluded.

The P4 position contains aliphatic aa to 73%, with Gly most frequently observed. In contrast to the P3 position, positively charged aa are also accepted in P4 (10%), whereas negatively charged ones are here excluded.

The positions C-terminal of the cleaved bond are less dominated by aliphatic aa. Ser is the single most frequent aa (27%) in position P1′. Small aliphatic aa are also abundant, with Gly and Ala together present in 27% of the sequences as well.

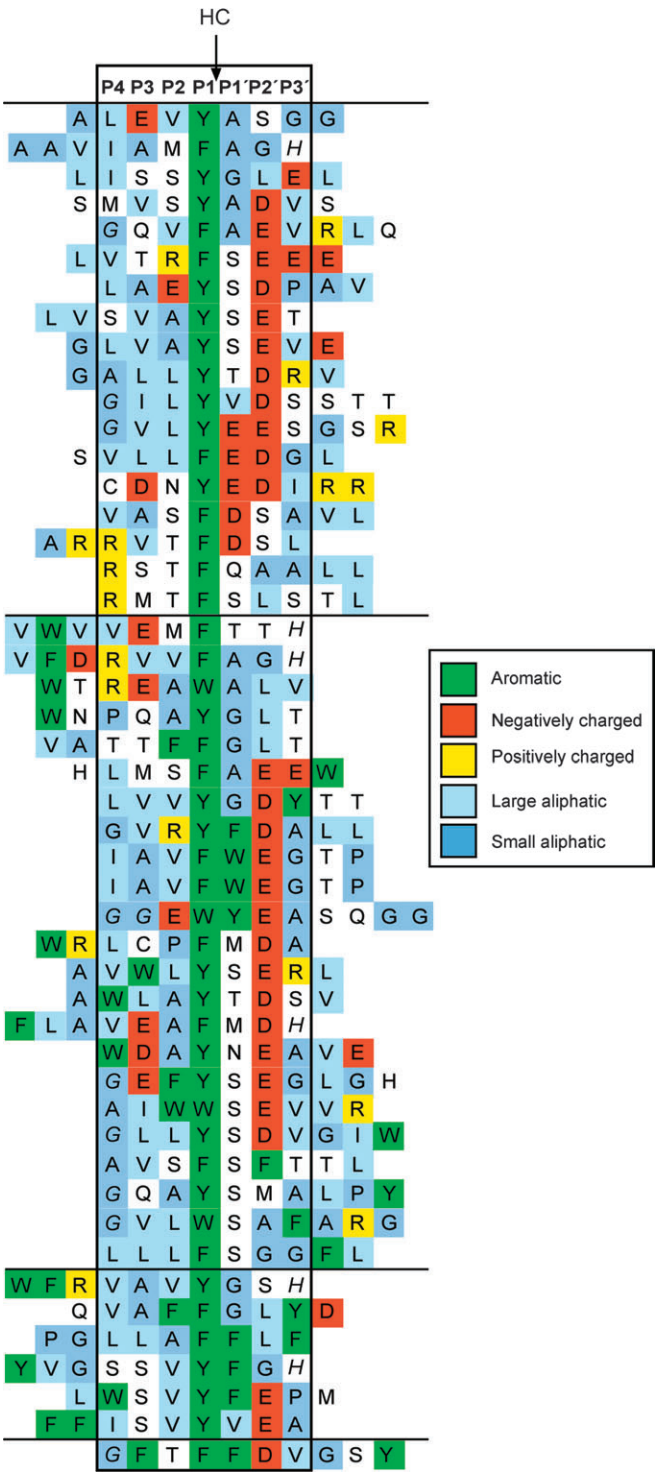

Fig. 3. Alignment of 48 phage-displayed peptide sequences susceptible to cleavage by the HC. After five rounds of selection, individual phages were isolated. The sequence encoding the random nonameric region of the capsid protein of each phage clone was determined. The general structure of the aa sequence in the phage clones is PGG$(X)_9$HHHHHH, where $(X)_9$ indicates the randomized region. The 48 sequences were aligned into a P4 to P3′ consensus where cleavage occurs between positions P1 and P1′. Aromatic aa as potential P1 residues are shown in green, negatively charged aa (Asp and Glu) in red, positively charged aa (Arg; no Lys was present) in yellow, small aliphatic aa (Gly and Ala) in blue, and larger aliphatic aa (Leu, Val and Ile) in light blue. The P4 to P3′ region is boxed in black. Sequences containing one, two, three or four aromatic residues are ordered in groups starting from above.

In the P2′ position, the HC strongly prefers negatively charged aa, which are found in 58% of the sequences. Asp and Glu constitute approximately half of the negatively charged residues, respectively (Fig. 4). Interestingly, among the 18 sequences containing only one aromatic aa (Fig. 3, upper panel), negatively charged aa are present not only in position P2′ but also rather frequently in P1′ or both positions (28%). Altogether, 72% of these sequences hold at least one negative aa in the P1′ or P2′ positions, C-terminal of the presumed cleavage site. This indicates that a negatively charged aa C-terminal of the cleavage site is highly preferred and that the P2′ is the optimal position. In analogy to the aromatic aa in position P1, Glu is the most over-represented negatively charged aa in the substrate region P4 to P3′, but it is Asp that is most specifically confined to positions P1′ and P2′ (88% of the Asp residues, but only 64% of the Glu residues are found in positions P1′ or P2′).

In position P3′, 52% of the aa are aliphatic. Notably, among the 23 sequences containing two aromatic aa, a big proportion of the aliphatic residues (75%) is contributed by the small Gly and Ala.

Overall, the sequence representing the potentially most susceptible peptide to the HC is from P4 to P3′: Gly/Leu/Val–Val/Ala/Leu–Ala/Val/Leu–Tyr/Phe–Ser–Asp/Glu–Ala/Val/Gly.

To verify the results, a synthetic peptide corresponding to the consensus sequence (Ala-Ala-**Gly-Val-Ala-Tyr-Ser-Glu-Ala**-Ala, where aa in bold indicate positions P4–P3′) was synthesized and subjected to cleavage by the HC at 37°C in PBS. The products at different time points were separated by HPLC (Fig. 6). HC was confirmed to exclusively cleave the P1 Tyr–P1′ Ser bond by mass spectrometry. The initial velocity of substrate cleavage was determined to be 0.28 μM s^{-1} at enzyme concentration of 12.4 nM and substrate concentration of 0.44 mM.

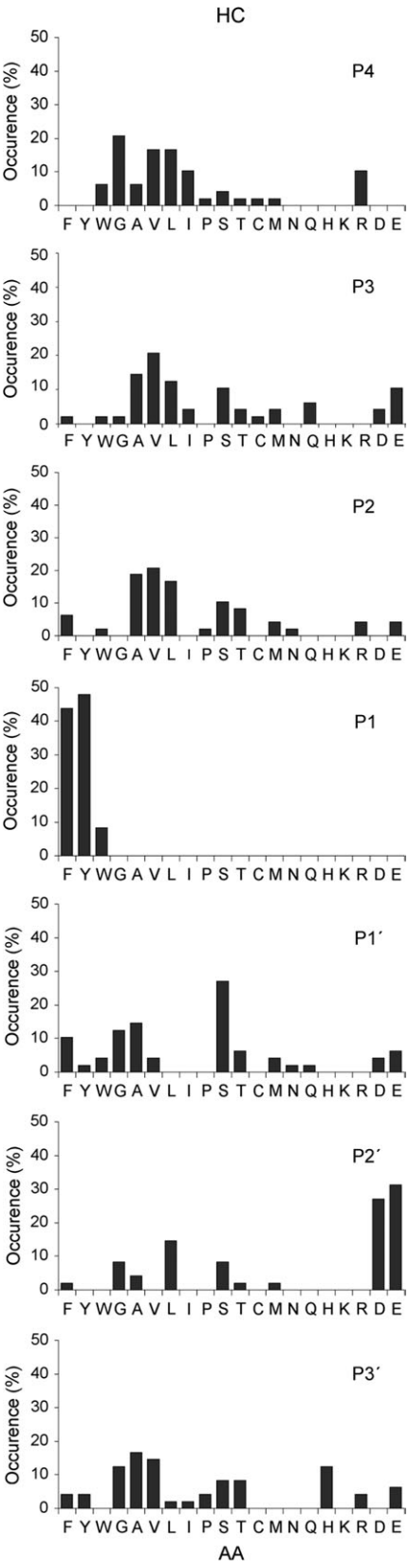

Fig. 4. The distribution of aa present in position P4 to P3′ in random phage sequences cleaved by the HC. Upon alignment of phage-displayed random sequences cleaved after five rounds of selection by HC (Fig. 3), the percentage of each aa in position P4 to P3′ were calculated. For clarity, aa are displayed in functional groups, starting to the left with aromatic aa and ending with acidic aa to the right.

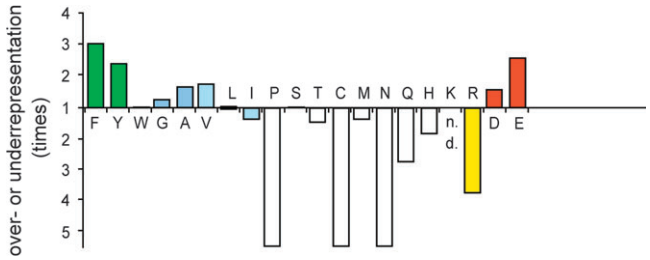

Fig. 5. Over- or under-representation of each aa within the P4 to P3′ region in total. The total number of each aa in the P4 to P3′ region in all 48 aligned sequences (Fig. 3) was counted separately and compared with the randomly expected number with regard to the relative codon frequency of this aa. The result is presented as times over-representation (standing boxes) or under-representation (hanging boxes). Colors correspond to Fig. 3.

Discussion

Phage-displayed random peptide libraries have markedly enhanced the possibility to accurately map the extended cleavage specificity of proteolytic enzymes. The main advantage of this method over other approaches is the independent and simultaneous variation of aa residues in all positions. We have used the phage display technology to determine the preferred aa for all positions from P4 to P3′ for the HC. Our results partly confirm previous results, but also contribute important novel information that can prove valuable for the identification of additional *in vivo* substrates for the HC. In the following, we will discuss the single substrate positions analyzed in this study, starting from P1 to P4 and continuing with the P1′ to P3′ positions.

The S1 pocket of HC, interacting with substrate position P1, has previously been shown to be large and hydrophobic and to accommodate mainly aromatic aa (38, 39). Using the phage display technology, we were able to confirm that Tyr and Phe are preferred over Trp (16, 32, 40). It has been reported that two water molecules are buried at the 'bottom' of the S1 pocket, which should allow placement of aromatic side chains with distal polar groups (39), such as Tyr. In line with this, we observed that among the three aromatic aa, Tyr residues had the strongest preference for the P1 position. This could be due to hydrophilic interactions in the bottom of the S1 pocket conferring a slower K_{off} to Tyr.

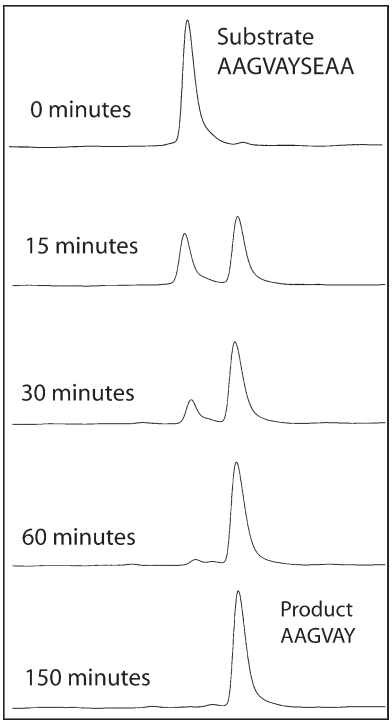

Fig. 6. Cleavage of a synthetic consensus peptide by the HC. A synthetic peptide (0.44 mM) containing the consensus sequence (Ala-Ala-Gly-Val-Ala-Tyr-Ser-Glu-Ala-Ala) was subjected to HC (12.4 nM) for 0, 15, 30, 60 and 150 min. The substrate and cleavage product (Ala-Ala-Gly-Val-Ala-Tyr) were separated by HPLC and detected at 229 nm. Peak content was determined by mass spectrometry.

In position P2, Pro has been shown to fit well into the S2 pocket of HC (16, 32, 33). Pro is also the P2 residue in Ang I (31, 41, 42), which is cleaved well by HC. In crystallization studies, this preference has been explained by the ability of Pro to induce a kink in the peptide chain, inserting the P1 residue into the S1 pocket (38, 39). Our study reveals, however, a general strong selection against Pro residues (Fig. 5), with a Pro in the P2 position of only 1 out of 48 sequences (Fig. 3). Interestingly, this Pro occurs in the context Pro-Phe-Met (P2-P1-P1′), which is identical in Kit ligand, a described substrate of the HC (Fig. 7) (24). The clearest preference is instead observed for the aliphatic aa Ala, Val and Leu in position P2. These three residues were found to be the best P2 residues after Pro by Powers *et al.* (16). Otherwise, the P2 position seems to be rather permissive, with aliphatic, hydrophilic, aromatic, positively and negatively charged aa admitted to some extent.

In the P3 position, the HC prefers Val. This preference is also found for the β-chymases rMCP-4, rMCP-1, mMCP-4 and mMCP-1 (18, 21, 43). The HC also admits negatively charged aa to the P3 position but excludes positively charged aa. This feature has also been observed in previous studies (32, 33) and can be explained by potential interactions of the P3 residue with the positively charged Lys192 and Arg217 in the S3 specificity pocket (39).

In position P4, 73% of the found aa are aliphatic. However, 9 of the 10 Gly residues in this position are part of the non-randomized region of the capsid protein, and Gly may therefore be over-represented. With respect to this, Leu and Val may be equally well admitted as Gly. Nonetheless, even when we exclude sequences with P4 Gly residues originating from the vector, aliphatic P4 residues are still present in

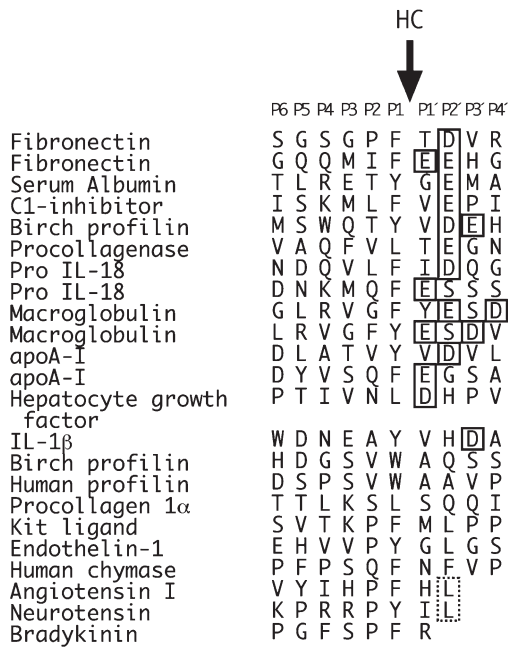

Fig. 7. Cleavage sites in natural substrates of the HC. Alignment of cleavage sites found in substrates of the HC. Acidic aa residues C-terminal of the cleaved bond are boxed. Aa with a negatively charged C-terminal carboxyl group in position P2′ are boxed with dotted line.

67% of the remaining sequences. It has previously been reported that an Ile residue fits very well into the shallow hydrophobic S4 site (39). In agreement with this, P4 is the preferred position for Ile residues in the set of sequences analyzed by us (Fig. 4). Like others, we also observe an acceptance for the positively charged Arg in P4 (16, 32). This can structurally be explained by ionic interaction with Asp175 in the S4 pocket (39). The lack of acidic aa in this position further supports this observation. Notably, three out of the five sequences with Arg in P4 in the present study hold Thr-Phe in positions P2–P1, and one additional sequence holds Val-Phe. This is an indication that synergistic interactions of several substrate residues might influence binding to the enzyme.

C-terminally of the scissile bond, we find a strong preference for Asp and Glu in the P2′ position. Interestingly, this preference is also observed in many natural peptides that are cleaved by the HC *in vitro* (Fig. 7). Within this panel, the only more pronounced preference is that for Phe and Tyr in P1, suggesting an important role of position P2′ for the HC. A preference for acidic aa in the primed positions was also observed in a screening of a peptide library with the HC (33), and recently, the preference for acidic P2′ residues was also reported using a combinatorial peptide library (14). Pereira *et al.* (39) suggest that the S2′ pocket of HC may attract acidic aa by close interactions of Lys40, Arg143 and Lys192 with the P2′ position of a substrate. The authors also indicate that these features could influence peptide binding beyond the S2′ position. This hypothesis is substantiated by our finding that negatively charged aa are overrepresented not only in P2′ but also in the P1′ position, among substrates holding only one aromatic residue (Fig. 3). However, the single most frequent aa in P1′ is Ser, a feature that has previously never been reported for the HC. Moreover, small aliphatic aa are also accepted in this position (Fig. 3).

For the P3′ position, Pereira *et al.* (39) suggest that it must be occupied by a relatively small aa like Gly or Ala in order to avoid distortion of the conformation of neighboring enzyme residues. Our data are clearly in line with this prediction, showing that the preferred P3′ aa are the small aliphatic aa Gly, Ala and Val.

In peptides like Ang I and neurotensin, the C-terminal dipeptide-leaving group is one of several determinants for efficient catalysis by HC (40). For example, introducing Asp^{10} instead of Leu^{10} in the P2′ position of an Ang I analog makes this a poorer substrate than the wild-type peptide (40), even though Asp otherwise is strongly preferred in the P2′ position. A model of Ang I in the active site of HC predicts that the ε-amino group of Lys40 in this case interacts with the C-terminal carboxyl group of Leu^{10}, which seems to be a more favorable than the interaction with an acidic side chain (39). The importance of Lys40 in the selective conversion of Ang I by HC was later confirmed in a study by site-directed mutagenesis (44). Another similar model suggests that one important determinant for the specific cleavage of the Phe^{8}-His^{9} bond of Ang I is a synergetic interaction of Arg143 and Lys192 with Leu^{10} of Ang I (45). Our approach is not sensitive to such interactions since no dipeptide-leaving groups are present, and the consensus presented here does probably not fully apply in these situations.

Taken together, our analysis predicts the consensus sequence, from P4 to P3′, Gly/Leu/Val–Val/Ala/Leu–Ala/Val/Leu–Tyr/Phe–Ser–Asp/Glu–Ala/Val/Gly to be the preferred combination of substrate residues of the HC. Although most of the aa in our consensus have been shown or predicted, by biochemical or structural studies, to fit well into the respective S subsite of the HC, the combined consensus is here presented for the first time. Notably, two independent phage sequences contribute the identical pentapeptide Val-Ala-Tyr-Ser-Glu (from P3 to P2′) to the sequence panel. Identical pentapeptides are randomly found only in 1 out of 3 200 000 (20^5) sequences, making it likely that the Val-Ala-Tyr-Ser-Glu pentapeptide is a very good substrate for the HC. This substrate contains Ala instead of Pro in P2 and Ser in P1′, novel consensus aa revealed in this study. Interestingly, the pentapeptide Val-Ala-Tyr-Ser-Glu is not present in previously identified natural substrates for the HC (Fig. 7). However, we have identified several novel hypothetical targets for the HC with this exact sequence, focusing on membrane proteins or extracellular proteins that are likely to be accessible for the HC. For example, the pentapeptide is present in a number of pathogenicity-conferring proteins from different microorganisms, such as cell wall binding repeat protein from *Clostridium* sp. (SwissEntry Q0TTI2) (46) and fibronectin-binding protein from *Pseudomonas* sp. (SwissEntry Q4ZSR1). By cleaving these targets, HC released from intestinal MCs could tentatively have an important role in gut immunity. The pentapeptide is also retrieved in members of the ABC transporter family, such as multidrug resistance-associated protein from *Apergillus oryzae* (SwissEntry Q2UTW4) and sulfate ABC transporter from the proteobacterium *Shewanella* (SwissEntry Q8E8L0). Cleavage of these targets could maybe inhibit these transporters, leading to starvation of the microorganism. Anti-bacterial properties could also be conferred by cleavage of the flagellar motor switch protein flim from the enterobacterium *Buchnera* (SwissEntry Q89AZ4), and the HC could have anti-viral properties if it cleaves the outer capsid protein vp2 of the bluetongue virus (SwissEntry P05309). This protein is known to be the major target of the host immune response (47) and may be evolutionary conserved. Furthermore, the pentapeptide is present in engorgement factor alpha protein (SwissEntry Q6T255), a protein triggering the engorgement of the female ticks that is required to induce egg laying (48). MCs have previously been implied in resistance against ticks (49). Tentatively, HC could enter the tick during feeding, inactivate the engorgement factor and thus prevent tick propagation.

Interestingly, previous studies have mainly focused on human proteins as substrates for the HC, whereas the Val-Ala-Tyr-Ser-Glu pentapeptide seems to be present mainly in a number of microbial and fungal proteins. However, all the proteins described above are to date merely hypothetical targets. It will be interesting to investigate whether any of them are to be added to the list of natural substrates for the HC. Future studies will now focus on further refining the panel of novel candidate substrates and then testing the candidates *in vitro* and finally *in vivo*.

Moreover, the novel consensus for the HC can be used to estimate the predictability of chymase studies in rodents for the situation in humans. The extended cleavage specificity of the opossum chymase, the elastase-like rMCP-5, and of the β-chymases rMCP-4, rMCP-1, mMCP-4 and mMCP-1 have previously successfully been determined with the phage display system (3, 12, 18, 21, 43). rMCP-1 and mMCP-4 are expressed by rat and mouse CTMCs, which are most similar to the chymase-expressing MCs in humans. Interesting similarities are revealed when comparing the substrate recognition profile of the above β-chymases with that of the HC. In P1, Phe or Tyr is generally preferred. Aliphatic aa are preferred in all positions' N-terminal of the cleaved bond, with the strongest preference for Val in P3. The single exception is rMCP-4, which selects aromatic aa in position P2. Positions P1′ and P3′ are the least restrictive in the β-chymases, but not in the HC. However, both the HC and mMCP-4 strongly prefer negatively charged aa in the very important P2′ position. These similarities in the substrate recognition profiles of HC and mMCP-4 [discussed in (21)] strongly support the functional homology between the two chymases. Therefore, results obtained from studies of mMCP-4 in mice may even shed light on the role of HC in human. An mMCP-4 knockout has recently been developed and interesting and valuable information is appearing from this important animal model (20, 26).

Funding

Swedish National Research Council (621-2005-3897).

Acknowledgements

We wish to thank Lotta Tollstoy Tegler at the Department of Biochemistry and Organic Chemistry, Uppsala University, Sweden, for performing the HPLC analysis.

Abbreviations

aa	amino acids
Ang	angiotensin
CTMC	connective tissue mast cell
EK	enterokinase
HC	human chymase
His_6	histidine
MC	mast cell
mMCP	mouse mast cell protease
N-V-A	Asn-Val-Ala
Ni–NTA	nickel–nitrilotriacetic acid
rMCP	rat mast cell protease
S-G-A	Ser-Gly-Ala
TFA	trifluoroacetic acid

References

1 Irani, A. A., Schechter, N. M., Craig, S. S., DeBlois, G. and Schwartz, L. B. 1986. Two types of human mast cells that have distinct neutral protease compositions. *Proc. Natl Acad. Sci. USA* 83:4464.

2 Caughey, G. H., Schaumberg, T. H., Zerweck, E. H. *et al.* 1993. The human mast cell chymase gene (CMA1): mapping to the cathepsin G/granzyme gene cluster and lineage-restricted expression. *Genomics* 15:614.

3 Reimer, J. M., Enoksson, M., Samollow, P. B. and Hellman, L. 2008. Extended substrate specificity of opossum chymase—implications for the origin of mast cell chymases. *Mol. Immunol.* 45:2116.

4 Gallwitz, M., Reimer, J. M. and Hellman, L. 2006. Expansion of the mast cell chymase locus over the past 200 million years of mammalian evolution. *Immunogenetics* 58:655.

5 Caughey, G. H., Raymond, W. W. and Wolters, P. J. 2000. Angiotensin II generation by mast cell alpha- and beta-chymases. *Biochim. Biophys. Acta* 1480:245.

6 Gallwitz, M. and Hellman, L. 2006. Rapid lineage-specific diversification of the mast cell chymase locus during mammalian evolution. *Immunogenetics* 58:641.

7 Caughey, G. H., Viro, N. F., Lazarus, S. C. and Nadel, J. A. 1988. Purification and characterization of dog mastocytoma chymase: identification of an octapeptide conserved in chymotryptic leukocyte proteinases. *Biochim. Biophys. Acta* 952:142.

8 Kunori, Y., Muroga, Y., Iidaka, M., Mitsuhashi, H., Kamimura, T. and Fukamizu, A. 2005. Species differences in angiotensin II generation and degradation by mast cell chymases. *J. Recept. Signal Transduct. Res.* 25:35.

9 Jolly, S., Detilleux, J., Coignoul, F. and Desmecht, D. 2000. Enzyme-histochemical detection of a chymase-like proteinase within bovine mucosal and connective tissue mast cells. *J. Comp. Pathol.* 122:155.

10 Solivan, S., Selwood, T., Wang, Z. M. and Schechter, N. M. 2002. Evidence for diversity of substrate specificity among members of the chymase family of serine proteases. *FEBS Lett.* 512:133.

11 Kunori, Y., Koizumi, M., Masegi, T. *et al.* 2002. Rodent alpha-chymases are elastase-like proteases. *Eur. J. Biochem.* 269:5921.

12 Karlson, U., Pejler, G., Tomasini-Johansson, B. and Hellman, L. 2003. Extended substrate specificity of rat mast cell protease 5, a rodent alpha-chymase with elastase-like primary specificity. *J. Biol. Chem.* 278:39625.

13 Kervinen, J., Abad, M., Crysler, C. *et al.* 2008. Structural basis for elastolytic substrate specificity in rodent alpha-chymases. *J. Biol. Chem.* 283:427.

14 Caughey, G. H., Beauchamp, J., Schlatter, D. *et al.* 2008. Guinea pig chymase is leucine-specific: a novel example of functional plasticity in the chymase/granzyme family of serine peptidases. *J. Biol. Chem.* 283:13943.

15 Caughey, G. H. 2002. New developments in the genetics and activation of mast cell proteases. *Mol. Immunol.* 38:1353.

16 Powers, J. C., Tanaka, T., Harper, J. W. *et al.* 1985. Mammalian chymotrypsin-like enzymes. Comparative reactivities of rat mast cell proteases, human and dog skin chymases, and human cathepsin G with peptide 4-nitroanilide substrates and with peptide chloromethyl ketone and sulfonyl fluoride inhibitors. *Biochemistry* 24:2048.

17 Newlands, G. F., Knox, D. P., Pirie-Shepherd, S. R. and Miller, H. R. 1993. Biochemical and immunological characterization of multiple glycoforms of mouse mast cell protease 1: comparison with an isolated murine serosal mast cell protease (MMCP-4). *Biochem. J.* 294:127.

18 Karlson, U., Pejler, G., Froman, G. and Hellman, L. 2002. Rat mast cell protease 4 is a beta-chymase with unusually stringent substrate recognition profile. *J. Biol. Chem.* 277:18579.

19 Shimizu, H., Nagakui, Y., Tsuchiya, K. and Horii, Y. 2001. Demonstration of chymotryptic and tryptic activities in mast cells of rodents: comparison of 17 species of the family Muridae. *J. Comp. Pathol.* 125:76.

20 Tchougounova, E., Pejler, G. and Abrink, M. 2003. The chymase, mouse mast cell protease 4, constitutes the major chymotrypsin-like activity in peritoneum and ear tissue. A role for mouse mast cell protease 4 in thrombin regulation and fibronectin turnover. *J. Exp. Med.* 198:423.

21 Andersson, M. K., Karlson, U. and Hellman, L. 2008. The extended cleavage specificity of the rodent beta-chymases rMCP-1 and mMCP-4 reveal major functional similarities to the human mast cell chymase. *Mol. Immunol.* 45:766.

22 Nakano, A., Kishi, F., Minami, K., Wakabayashi, H., Nakaya, Y. and Kido, H. 1997. Selective conversion of big endothelins to tracheal smooth muscle-constricting 31-amino acid-length endothelins by chymase from human mast cells. *J. Immunol.* 159:1987.

23 Maurer, M., Wedemeyer, J., Metz, M. *et al.* 2004. Mast cells promote homeostasis by limiting endothelin-1-induced toxicity. *Nature* 432:512.

24 Longley, B. J., Tyrrell, L., Ma, Y. *et al.* 1997. Chymase cleavage of stem cell factor yields a bioactive, soluble product. *Proc. Natl Acad. Sci. USA* 94:9017.
25 Fang, K. C., Raymond, W. W., Blount, J. L. and Caughey, G. H. 1997. Dog mast cell alpha-chymase activates progelatinase B by cleaving the Phe88-Gln89 and Phe91-Glu92 bonds of the catalytic domain. *J. Biol. Chem.* 272:25628.
26 Tchougounova, E., Lundequist, A., Fajardo, I., Winberg, J. O., Abrink, M. and Pejler, G. 2005. A key role for mast cell chymase in the activation of pro-matrix metalloprotease-9 and pro-matrix metalloprotease-2. *J. Biol. Chem.* 280:9291.
27 Saarinen, J., Kalkkinen, N., Welgus, H. G. and Kovanen, P. T. 1994. Activation of human interstitial procollagenase through direct cleavage of the Leu83-Thr84 bond by mast cell chymase. *J. Biol. Chem.* 269:18134.
28 Okumura, K., Takai, S., Muramatsu, M. *et al.* 2004. Human chymase degrades human fibronectin. *Clin. Chim. Acta* 347:223.
29 Kofford, M. W., Schwartz, L. B., Schechter, N. M., Yager, D. R., Diegelmann, R. F. and Graham, M. F. 1997. Cleavage of type I procollagen by human mast cell chymase initiates collagen fibril formation and generates a unique carboxyl-terminal propeptide. *J. Biol. Chem.* 272:7127.
30 Mizutani, H., Schechter, N., Lazarus, G., Black, R. A. and Kupper, T. S. 1991. Rapid and specific conversion of precursor interleukin 1 beta (IL-1 beta) to an active IL-1 species by human mast cell chymase. *J. Exp. Med.* 174:821.
31 Reilly, C. F., Tewksbury, D. A., Schechter, N. M. and Travis, J. 1982. Rapid conversion of angiotensin I to angiotensin II by neutrophil and mast cell proteinases. *J. Biol. Chem.* 257:8619.
32 Raymond, W. W., Ruggles, S. W., Craik, C. S. and Caughey, G. H. 2003. Albumin is a substrate of human chymase. Prediction by combinatorial peptide screening and development of a selective inhibitor based on the albumin cleavage site. *J. Biol. Chem.* 278:34517.
33 Bastos, M., Maeji, N. J. and Abeles, R. H. 1995. Inhibitors of human heart chymase based on a peptide library. *Proc. Natl Acad. Sci. USA* 92:6738.
34 Vernersson, M., Ledin, A., Johansson, J. and Hellman, L. 2002. Generation of therapeutic antibody responses against IgE through vaccination. *FASEB J.* 16:875.
35 Hallgren, J., Karlson, U., Poorafshar, M., Hellman, L. and Pejler, G. 2000. Mechanism for activation of mouse mast cell tryptase: dependence on heparin and acidic pH for formation of active tetramers of mouse mast cell protease 6. *Biochemistry* 39:13068.
36 Schechter, I. and Berger, A. 1967. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 27:157.
37 Greer, J. 1990. Comparative modeling methods: application to the family of the mammalian serine proteases. *Proteins* 7:317.
38 McGrath, M. E., Mirzadegan, T. and Schmidt, B. F. 1997. Crystal structure of phenylmethanesulfonyl fluoride-treated human chymase at 1.9 A. *Biochemistry* 36:14318.
39 Pereira, P. J., Wang, Z. M., Rubin, H. *et al.* 1999. The 2.2 A crystal structure of human chymase in complex with succinyl-Ala-Ala-Pro-Phe-chloromethylketone: structural explanation for its dipeptidyl carboxypeptidase specificity. *J. Mol. Biol.* 286:163.
40 Kinoshita, A., Urata, H., Bumpus, F. M. and Husain, A. 1991. Multiple determinants for the high substrate specificity of an angiotensin II-forming chymase from the human heart. *J. Biol. Chem.* 266:19192.
41 Wintroub, B. U., Schechter, N. B., Lazarus, G. S., Kaempfer, C. E. and Schwartz, L. B. 1984. Angiotensin I conversion by human and rat chymotryptic proteinases. *J. Invest. Dermatol.* 83:336.
42 Chandrasekharan, U. M., Sanker, S., Glynias, M. J., Karnik, S. S. and Husain, A. 1996. Angiotensin II-forming activity in a reconstructed ancestral chymase. *Science.* 271:502.
43 Andersson, M. K., Pemberton, A. D., Miller, H. R. and Hellman, L. 2008. Extended cleavage specificity of mMCP-1, the major mucosal mast cell protease in mouse-high specificity indicates high substrate selectivity. *Mol. Immunol.* 45:2548.
44 Muilenburg, D. J., Raymond, W. W., Wolters, P. J. and Caughey, G. H. 2002. Lys40 but not Arg143 influences selectivity of angiotensin conversion by human alpha-chymase. *Biochim. Biophys. Acta* 1596:346.
45 Yamamoto, D., Shiota, N., Takai, S., Ishida, T., Okunishi, H. and Miyazaki, M. 1998. Three-dimensional molecular modeling explains why catalytic function for angiotensin-I is different between human and rat chymases. *Biochem. Biophys. Res. Commun.* 242:158.
46 Demarest, S. J., Salbato, J., Elia, M. *et al.* 2005. Structural characterization of the cell wall binding domains of Clostridium difficile toxins A and B; evidence that Ca2+ plays a role in toxin A cell surface association. *J. Mol. Biol.* 346:1197.
47 Ghiasi, H., Fukusho, A., Eshita, Y. and Roy, P. 1987. Identification and characterization of conserved and variable regions in the neutralization VP2 gene of bluetongue virus. *Virology* 160:100.
48 Weiss, B. L. and Kaufman, W. R. 2004. Two feeding-induced proteins from the male gonad trigger engorgement of the female tick *Amblyomma hebraeum*. *Proc. Natl Acad. Sci. USA* 101:5874.
49 Matsuda, H., Nakano, T., Kiso, Y. and Kitamura, Y. 1987. Normalization of anti-tick response of mast cell-deficient W/Wv mice by intracutaneous injection of cultured mast cells. *J. Parasitol.* 73:155.