The extended substrate recognition profile of the dog mast cell chymase reveals similarities and differences to the human chymase

Maike Gallwitz¹, Mattias Enoksson¹,², Michael Thorpe¹, Xueliang Ge¹ and Lars Hellman¹

¹Department of Cell and Molecular Biology, Biomedical Centre, Uppsala University, Box 596, SE-751 24 Uppsala, Sweden
²Present address: Clinical Immunology and Allergy Unit, Department of Medicine, Karolinska Institutet, SE-17176 Stockholm, Sweden

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

Transmitting editor: S. J. Galli

Received 9 December 2009, accepted 10 February 2010

Abstract

Human chymase (HC) constitutes a major granule protease in one of the two human mast cell (MC) types. The main biological role of this haematopoietic serine protease is probably not yet known, although it has been implicated in a large number of functions. Dogs, like humans, have only one chymase. This enzyme is closely related to its human homologue, and the MC subtypes of human and dog appear to be similar as well. Therefore, the functions of the dog chymase (DC) may closely reflect the functions of the HC. Moreover, dogs may serve as good models for studies of human MC functions and MC-related diseases. To reveal functional similarities and differences between the DC and HC, we have determined the extended cleavage specificity of the DC by substrate phage display. This method allows the simultaneous permutation of primed and unprimed substrate positions. The DC was found to have very similar preferences to its human counterpart for substrate positions P₁, P₃, P₄ and P₃’, whereas their preferences differ at positions P₂, P₁’ and P₂’. Therefore, the HC and DC may have co-evolved with a substrate where positions P₁, P₃, P₄ and P₃’ are conserved between dogs and humans, whereas positions P₂ and P₁’ are not and P₂’ differs to a minor extent. The differences observed between these two enzymes suggest that results obtained from dog models cannot be directly extrapolated to human clinical settings but need to be evaluated carefully concerning potential differences in substrate preferences.

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

Introduction

Mast cells (MCs) are important players in several common pathological conditions in both dogs and humans, such as atopy, asthma and mastocytomas (1–3). A protective function for MCs in bacterial and helminth infections has also been documented (4–7). However, the positive physiological role of these cells may not yet be fully recognized.

Many of the functions carried out by MCs depend on two groups of granule-stored serine proteases, chymases and tryptases. These abundant granule proteases are stored in tight complexes with negatively charged proteoglycans and are released into the extracellular environment in response to immunological and neuronal stimuli. Human MCs express either only tryptase or tryptase together with chymase and are referred to as MCₜ (tryptase-positive mast cell) and MCₜc (tryptase- and chymase-positive mast cell), respectively. Also in dogs, MCₜ and MCₜc have been described. In addition to these two MC populations, chymase single-positive MCs are found in dog, MCₜ (8, 9). The human and dog MC subtypes can be found in the same organs, e.g. the skin, lungs, liver, spleen and intestines, albeit in varying proportions (10–12).

Human and dog MCs express only a single chymase, which by phylogenetic analysis is classified as an α-chymase. The α-chymase is encoded from the Cma1 gene in the MC chymase locus on dog chromosome 8 (human chromosome 14). Both the dog chymase (DC) and the human chymase (HC) display chymotryptic activity (13, 14). This specificity is mainly conferred by the residues in positions 189, 216 and 226 [chymotrypsin numbering according to ref. 15 is applied if not indicated otherwise], which in both
species are occupied by amino acids (aa) Ser-Gly-Ala (S-G-A). With the exception of rodents, one or two α-chymases with this identical specificity-conferring triplet have to date been identified in all analysed mammals, including humans, dogs, cattle, sheep and opossums (16–18). The marsupial lineage, as represented by the opossum, diverged from eutherian mammals ~185 million years ago. Therefore, an ancestral α-chymase gene seems to have already existed >185 million years ago (18). The presence of the α-chymase gene this early in the evolution of mammals, and the evolutionary conservation of the specificity-conferring triplet and the proteolytic activity, suggests that the α-chymase exerts an important function (19). This enzyme therefore gained considerable interest and a number of physiological roles have been proposed.

For example, the enzyme has been implicated in allergic symptoms by potentiation of plasma leakage, stimulation of submucosal gland cell secretion and regulation of airway reactivity (20, 21). In addition, chymase has long been associated with modulations of the MC environment by destruction of extracellular matrix proteins, activation of matrix metalloproteases and modulation of cytokine and chemokine activity (22–26). Chymase has also been implicated in the recruitment of leukocytes (27, 28). Moreover, the enzyme seems to be involved in the regulation of blood pressure by conversion of the decapeptide angiotensin I (Ang I) into the active product angiotensin II (Ang II) (29). The HC also has the ability to activate procollagen and initiate collagen fibril formation (30). Although these functions may be important, it is difficult to assign a clear in vivo role, and the question remains whether the most important targets have yet been identified.

In vivo functions of an enzyme can often be addressed with good predictability in animal models, e.g. in the mouse or rat. However, the mouse and rat α-chymases [mouse mast cell protease (mMCP)-5 and rat mast cell protease (rMCP)-5, respectively] differ from most other α-chymases in the aa of the specificity-conferring triplet. mMCP-5 and rMCP-5 hold Asn-Val-Ala (N-V-A) in these positions, instead of S-G-A, and display elastase-like rather than chymase-like activity (31, 32). Several additional chymases are expressed in the MC of most rodents, with the exception of the guinea pig (33). These build the phylogenetically separate β-chymase subfamily, which seem to be absent in non-rodents (18, 34). Several of the β-chymases display chymotryptic activity, thus might hold functions that are carried out by α-chymase in non-rodents (35, 36). The variety of chymases in rodent MC, where human MC hold a single chymase, makes it difficult to apply data from studies in these animals to the human situation.

Here, studies of the DC may yield data, which reflects the role of the HC more closely. DC and HC share a high degree of homology, with 82% aa identity between the two mature enzymes. Both enzymes display a highly positive net charge on +18.4 (DC) and +13.2 (HC) at pH 7 (18). For comparison, the dominant β-chymases in connective tissue MCs from mouse and rat, mMCP-4 and rMCP-1, share <65% aa identity with the HC. The HC and DC are also similar in that they efficiently convert Ang I to Ang II in vitro. In contrast, most rodent β-chymases are inefficient producers of Ang II because they destroy this product by cleavage between the Tyr4-Ile5 bond (29, 37, 38). With regard to general MC biology, more similarities are observed between humans and dogs than between humans and rodents. The human and dog MC subtypes reside in the same tissues, although in varying ratios. In rodents, on the other hand, the two major MC types are physically separated. Rodent connective tissue-type MCs are mainly found in the skin and peritoneum, whereas rodent mucosal MCs reside in the respiratory tract and intestine (39). Moreover, cultured bone marrow-derived MCs (BMMC) from dogs are more like human MCs than rodent BMMC with respect to differentiation requirements and cell phenotype (9).

Despite its potential role as a model enzyme for the HC, the DC has not been studied very intensely to date. Most studies of the DC have focused on its involvement in pathological conditions, e.g. in subacute stages of adverse left ventricular remodelling by MCs and in the pathogenesis of haemodialysis vascular access dysfunction (40, 41). A step to a better understanding of in vivo roles for the DC, and to be able to compare it with the HC, is to determine its substrate specificity. We have produced recombinant DC and determined the preferred substrate residues throughout its substrate-binding site. This analysis was carried out using a library of random nonapeptides displayed on T7 phages. The consensus sequence for the cleavage specificity of the DC from positions P4 to P3’, as obtained from 51 sequenced phages, is presented here.

**Methods**

**Cloning of recombinant DC**

Total RNA was extracted from skin of a Rhodesian Ridgeback dog as previously described (42). Complementary DNA for the coding region of mature DC was obtained from total RNA and amplified by PCR, using oligonucleotide primers dog Cma1–5’ (5’–3’): CACGAAATTCACCATCATCATCATCGGAT-CCGACGATGACGATAAGATCATCGGGGGCACAGAGTCCAAG and dog Cma1–3’ (5’–3’): CAGCTCGAGTGGCAGGCTC-ATCGGCGTTTTACG. Oligonucleotide dog Cma1–5’ contains a sequence encoding six histidine residues (His6-tag) for purification purposes and an enterokinase (EK)-susceptible peptide (Asp-Asp-Asp-Asp-Lys). The EK cleavage site enables subsequent removal of the histidine residues and therefore activation of the protease. The PCR product was sub-cloned into the pCEP-Pu2 vector (43, 44), and the correct nucleotide sequence was confirmed using an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems, Foster City, CA, USA) with vector-specific primers.

**Production and collection of recombinant DC**

Human embryonic kidney cells, HEK 293-EBNA, were transfected with the pCEP-Pu2/DC constructs as previously described (43, 44). Selection was initiated after 7 days by adding 1.5 μg ml⁻¹ puromycin to the cell culture medium (DMEM supplemented with 5% FCS, 50 μg ml⁻¹ gentamycin and 5 μg ml⁻¹ heparin). After 14 days, the puromycin concentration was decreased to 0.5 μg ml⁻¹. Conditioned medium was collected and centrifuged at
3000 r.p.m. for 10 min to remove cell debris. The superna-
tant was saved, and Ni\textsuperscript{2+}-nitrilotriacetic acid (Ni-NTA) agar-
ose beads (Qiagen GmbH, Hilden, Germany) were added to
a final concentration of 0.5 ml l\textsuperscript{-1}. 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
(Treflab, 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 imidazol and 0.2% Triton X-100 in PBS).

Protein purity and concentration were estimated by
separation on 10% SDS–PAGE gels (Fig. 1). Samples were
prepared with reducing sample buffer, containing
β-mercaptoethanol to a final concentration of 5%. To visual-
ize protein bands, gels were stained with Coomassie Brilliant
Blue and de-stained with de-staining buffer [7.5% (v/v) ace-
tic acid and 25% (v/v) methanol in ddH\textsubscript{2}O].

**Activation and purification of DC**

Recombinant DC was digested for 1 or 5 h at 37°C with
0.6 μg EK (Roche Diagnostics, Mannheim, Germany) per
25 μg protease. To remove EK and other impurities, EK-
digested DC was purified by affinity chromatography over
heparin–sepharose columns as previously described (35).

Briefly, a 10-ml PolyPrep chromatography column containing
0.3 ml heparin–sepharose (Amersham Pharmacia Biotech)
was equilibrated with PBS (pH 7.2). Approximately 40 μg of
EK-cleaved DC in PBS (pH 7.2) was applied to the column,
followed by four washes with PBS. Bound protein was eluted
in three steps (pre-eluate, eluate and post-eluate) with 250,
500 and 300 μl 0.8 M NaCl, respectively. The protein content
of the flow-through, wash and eluted fractions was moni-
tored on SDS-PAGE gels.

**Determination of chymase activity using chromogenic
substrates**

Chymase activity was measured against the chromogenic
substrates S-2586 (MeO-Suc-Arg-Ala-Tyr-pNA) (Chromogenix,
Mölndal, Sweden) and S-7388 (N-succinyl-Ala-Ala-Pro-Phe-
pNA) (Sigma, St Louis, MO, USA). Standard measurements
were performed in 96-well microtiter plates with 0.18 mM
S-2586 or 0.20 mM S-7388 in 200 μl PBS. Substrate hydroly-
sis was monitored spectrophotometrically at 405 nm in a Multi-
scan MCC/340 spectrophotometer (Labsystem, Helsinki,
Finland). Control samples without enzyme or with the
trypsinase-susceptible substrate S-2288 (H-D-Ile-Pro-Arg-pNA;
0.2 mM) (Chromogenix) were included.

**Determination of the cleavage recognition profile using
a library of phage-displayed nonapeptides**

A library of T7 phage-displayed non-americ peptides was
used as previously described (32, 35). In this library, the T7
capsid protein 10 is manipulated to contain a nine-aa-long
random peptide (X\textsubscript{9}) followed by a His\textsubscript{6}-tag at the
C-terminus, anchoring the phage to the bead (35). Statisti-
cally, one His\textsubscript{6}-tagged capsid protein is expressed on every
phage to every 10th phage. This means that in a preparation,
the number of His\textsubscript{6}-tagged phages varies from only 1 in 10
phages to that every phage has one His\textsubscript{6}-tag. Upon addition
of protease, phages with a cleavage-susceptible random
nonapeptide are released from the Ni–NTA matrix, which
can be removed and subsequently amplified. Amplified
phages then enter a new selection round (bio-panning)
until a good enrichment of phages expressing protease-
susceptible nonapeptides is obtained. In general, this is
achieved after five bio-pannings.

For the first selection round, an aliquot of \~10\textsuperscript{9} phages
(plaque-forming unit) was allowed to bind to 100 μl Ni–NTA
agarose beads for 1 h while rotating gently at 4°C. Un-
bound phages were removed by 10 washes with 1.5 ml
1 M NaCl, 0.1% Tween 20 in PBS (pH 7.2) and 2 subse-
quent washes with 1.5 ml PBS. The beads were then
re-suspended in 1 ml PBS. Activated heparin–sepharose-
purified DC or PBS control was added. In different trials,
the amount of enzyme varied between 0.1 and 1 μg. Prote-
ase digestion was allowed to proceed overnight at room
temperature (RT) with gentle agitation. Ni–NTA agarose
beads were then pelleted by centrifugation in a tabletop
centrifuge and the supernatant, containing the released
phages, was recovered. To retrieve the maximum possible
number of released phages, the beads were re-suspended
in 100 μl PBS (pH 7.2) and re-pelleted. The supernatant
was then added to the first recovered phages. To remove

---

**Fig. 1.** Purification and activation of recombinant DC. Recombinant
DC containing an N-terminal His\textsubscript{6}-tag and EK site was purified on
Ni–NTA agarose as described in Methods. The purified protease was
activated by removal of the His\textsubscript{6}-tag through EK digestion. EK-
digested protease was purified over heparin–sepharose columns to
remove EK and imidazole. Samples of DC purified on Ni–NTA agarose
(−EK) and following EK digestion (+EK) were analysed by separation
on SDS-PAGE (12.5% polyacrylamide) and visualized by Coomassie Brilliant Blue staining.
released phages, on which the His₆-tag had not been cleaved off, 15 µl of fresh Ni–NTA agarose beads was added to the phage suspension and the mixture was agitated for 15 min followed by centrifugation to recover the supernatant. Thirty microlitres of the supernatant was used to determine the titre of phages detached in each round of selection. A control elution with 100 µl of 100 mM imidazole, releasing all phages still bound to the beads, concluded that at least 10⁸ phages were attached to the matrix in each selection round.

The remaining supernatant was added to a 10-ml culture of *Escherichia coli* BLT5615 [optical density (OD); 0.5] which had been induced to produce T7 phage capsid protein by adding 100 µl of 100 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) 30 min before. The bacteria lysed ~75 min after phage addition. The lysate was centrifuged to remove bacterial debris, and 900 µl of the supernatant, constituting the new phage sub-library, was added to 100 µl of fresh Ni–NTA beads. In the second and subsequent selection rounds, the protocol was as before; however, the Ni–NTA beads were washed 15 times instead of 10. After five rounds of selection, 75 plaques were randomly isolated from Luria Broth (LB) plates and each dissolved in phage extraction buffer (100 mM NaCl and 6 mM MgSO₄ in 20 mM Tris–HCl, pH 8.0). The plaques were vigorously shaken for 30 min in order to allow a substantial fraction of the phages to be in the soluble phase. Phage DNA was 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 GmbH). Purified PCR fragments were then sequenced on an ABI PRISM® 3700 DNA Analyzer (Applied Biosystems).

**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), which was programmed to identify patterns of at least four aa without bias. 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 very similar results. The presented alignment of 51 sequences was derived from Consense, with manual adjustment of seven sequences containing an aromatic aa that had not been aligned in P1 (Fig. 2). Another three sequences were adjusted to reflect the detected preference for arginine in position P2. aa with similar characteristics were grouped as follows: aromatic aa (phenylalanine, tyrosine and tryptophan), negatively charged aa (aspartate and glutamate), positively charged aa (arginine and lysine), small aliphatic aa (glycine and arginine), larger aliphatic aa (valine, leucine, isoleucine and proline) and hydrophilic aa (cysteine, methionine, serine, threonine, histidine, asparagine and glutamine). The nomenclature by Schechter and Berger (45) was adopted to designate the aa in the substrate cleavage region, where P1–P1’ corresponds to the scissile bond.

**Fig. 2.** Chymase-susceptible sequences from phages selected randomly after five selection rounds. After five rounds of selection, individual phages were isolated and sequenced in the random nonameric region of the capsid protein. The general structure of the aa sequence in the phage clone is PGX(X)₉H₆, where X indicates the randomized region. Fifty-one sequences were aligned into a P₄–P₃’ consensus where the cleavage occurs between positions P1 and P1’. Aromatic aa are shown in green, negatively charged aa (Asp and Glu) in red, positively charged aa (Arg and Lys) in yellow, small aliphatic aa (Gly and Ala) in dark blue and larger aliphatic aa (Leu, Val and Ile) in light blue. The P₄–P₃’ region is boxed in black. Arrows indicate two individual phage sequences with almost identical hexapeptides.
Generation of recombinant substrates for the analysis of the cleavage specificity

A new type of substrates was developed to verify the results obtained from the phage-display analysis. Two copies of the *E. coli* thioredoxin gene were inserted in tandem into the pET21 vector for bacterial expression (Fig. 4A). In the C-terminal end, a His$_{6}$-tag was inserted for purification on Ni$^{2+}$ immobilized metal ion affinity chromatography (IMAC) columns. In the linker region, between the two thioredoxin molecules, the different substrate sequences were inserted by ligating double-stranded oligonucleotides into two unique restriction sites, one *Bam*H1 and one *Sal* site (Fig. 4A). The sequences of the individual clones were verified after cloning by sequencing of both DNA strains. The plasmids were then transferred to the *E. coli* Rosetta gami strain for protein expression (Novagen; Merck, Darmstadt, Germany). A 10-ml overnight culture of the bacteria harbouring the plasmid was diluted 10 times in LB + Amp and grown at 37°C for 1–2 h until the OD (600 nm) reached 0.5. IPTG was then added to a final concentration of 1 mM. The culture was then grown at 37°C for an additional 3 h under vigorous shaking, after which the bacteria were pelleted by centrifugation at 3500 r.p.m. for 12 min. The pellet was washed once with 25 ml PBS + 0.05% Tween 20. The pellet was then dissolved in 2 ml PBS and sonicated 6 × 30 s to open the cells. The lysate was centrifuged at 13 000 r.p.m. for 10 min and the supernatant was transferred to a new tube. Five hundred microlitres of Ni–NTA agarose slurry (50:50) (Qiagen GmbH) was added and the sample was slowly rotated for 45 min at RT. The sample was then transferred to a 2-ml column and the supernatant was allowed to slowly pass through the filter leaving the Ni–NTA beads with the bound protein in the column. The column was then washed four times with 1 ml of washing buffer (PBS + 0.05% Tween 20 + 10 mM imidazole + 1 M NaCl). Elution of the protein was performed by adding 150 μl elution buffer followed by five 300 μl fractions of elution buffer (PBS + 0.05% Tween 20 + 100 mM imidazole). Each fraction was collected individually. Ten microlitres from each of the eluted fractions was then mixed with one volume of 2× sample buffer and 1 μl β-mercaptoethanol was added and the sample was heated for 3 min at 80°C. The samples were analysed on a SDS–bis–tris 4–12% PAGE gel and the second and third fractions that contained most protein were combined. The protein concentration of the combined fractions was determined by Bio-Rad DC Protein assay (Bio-Rad Laboratories, Hercules, CA, USA). Approximately 60 μg of recombinant protein was added to each 120-μl cleavage reaction (in PBS). Twenty microlitres from this tube was removed before adding the enzyme, the 0 min time point. The active enzyme was then added (~30 ng DC and 35 ng HC), and the reaction was kept at RT during the entire experiment. Twenty-microtitre samples were removed at the indicated time points (15, 30, 45, 60 and 150 min) and stopped by addition of one volume of 2× sample buffer. One microlitre of β-mercaptoethanol was then added to each sample and the samples were heated for 3 min at 80°C. Twenty microlitres from each of these samples was then analysed on 4–12% precast SDS–PAGE gels (Invitrogen, Carlsbad, CA, USA). The gels were stained overnight in colloidal Coomassie staining solution and de-stained for several hours according to previously described procedures (46).

Results

Production and purification of recombinant DC

DC was expressed as a fusion protein with an N-terminal His$_{6}$-tag and an EK-susceptible peptide (Asp-Asp-Asp-Asp-Lys). The fusion protein was purified on Ni–NTA agarose beads and then EK digested to remove the His$_{6}$-tag along with the EK site, thereby activating the protease. We observed that prolonged cleavage with EK resulted in loss of the protein. This indicates that the DC has a relatively strong autocatalytic activity and that it degrades itself upon incubation at 37°C for several hours. The EK we used is very specific and has not shown any tendency to cleave at other sites than DDDDK with numerous other recombinant enzymes tested under these conditions. This result also gave a strong indication for a relatively low extended cleavage specificity of the DC compared with other MC chymases previously analysed.

Further purification from contaminating proteins as well as imidazole and removal of EK were achieved by passing activated DC over a heparin–sepharose column. We obtained a virtually pure DC that migrates at an apparent size of 32 kDa (Fig. 1). This size is somewhat larger than the theoretical size of 25.5 kDa, probably due to glycosylation at two potential N-glycosylation sites, N113 and N147 (chymotrypsin numbering). The inactive protease migrates as a 34-kDa band (Fig. 1), well in proportion to the calculated contribution of the His$_{6}$-tag and EK site to the molecular weight, 1.5 kDa.

Activity analysis of recombinant DC with chromogenic substrates

The DC has previously been shown to display chymotryptic activity (13). Therefore, we tested the enzymatic activity of recombinant DC against the chymase-susceptible chromogenic substrates S-2586 (MeO-Suc-Arg-Ala-Tyr-pNA) and S-7388 (N-succinyl-Ala-Ala-Pro-Phe-pNA). Activated recombinant DC showed good activity against both substrates (data not shown). Some other enzymes encoded from the chymase locus, e.g. human cathepsin G and cattle duodenase BDMD1, have been shown to possess dual chymotryptic and trypsic activity (47). However, DC displayed no activity against the trypsinase-susceptible substrate S-2288 (H-D-Ile-Pro-Aro-pNA) (data not shown).

Extended substrate specificity of the DC as determined by substrate phage display

To determine the extended substrate specificity of the DC, we used a T7 phage library where one of the capsid proteins contains a random nonapeptide. The nonapeptide is followed by a His$_{6}$-tag that serves to anchor the phages to an Ni–NTA matrix. This library was subjected to DC or a PBS control over five rounds of selection. In the best of 13 trials, the ratio of phages released from enzyme-treated samples, as compared with controls, increased steadily and
reached eight after five rounds of selection. Nonapeptide-coding inserts were then sequenced from 75 individually sampled phages. Fifty-one of the obtained sequences were aligned (Fig. 2). Seventeen background sequences that had also been observed in analyses of other enzymes were excluded. Eleven of these contained the insert WCQVSQSCA and six the insert TLMVPRTGS. Four sequences contained stop codons. The remaining three sequences contained repetitive sequences and did not contain aromatic aa. Although the frequency of background sequences was higher than expected, these were easily identified and therefore probably do not considerably affect the quality of the presented alignment.

Based on the alignment shown in Fig. 2, we determined the prevalence of aa in each position from P4 to P3′ (Fig. 3). In addition, we calculated the total over- or under-representation of each aa in the collected sequences, between P4 and P3′. For this calculation, the total number of an aa counted in positions P4–P3′ of the 51 sequences was compared with the randomly expected number, with respect to the codon frequency for this residue (35). Our analysis revealed a clear over-representation of all three aromatic aa and of all aliphatic aa except Ile. Among the negatively charged aa, Glu but not Asp was over-represented. Met and His were also found more often than randomly expected, whereas all remaining aa were under-represented. The strongest under-representation was seen with Lys.

All residues in position P1 were aromatic aa, with Phe and Tyr being somewhat preferred over Trp. In total, Trp was more over-represented than Tyr, but it aligns more frequently outside the P1 position (Figs 2 and 3). In position P2, the single most frequent residue is Arg (37%). This clear pattern was observed despite an under-representation of Arg in total. Interestingly, no Lys was found in position P2, although Arg and Lys display similar biochemical properties. Other than Arg, aliphatic residues were well tolerated in position P2. Aliphatic residues were also the preferred residues in positions P3 and P4, with Gly and Val dominating both positions. Alanine is less frequent, and the larger Leu and Ile are almost excluded. However, the non-randomized phage region contributes 15 of the 17 Gly in position P4, and Gly might not be truly preferred by the DC. (Also Pro in position P4 is most often contributed from the non-randomized region.)

The primed positions feature mostly aliphatic residues as well. This is very clear in position P1′, where Leu is preferred (35%), followed by Val and Ala. Interestingly, Leu is the only aliphatic P1′ aa that is found adjacent to a negatively charged P2′ residue, Asp or Glu (Fig. 2). Asp and Glu occur with a frequency of 8 and 10%, respectively, in position P2′. The occurrence of all aliphatic aa, except Ile, is in the same range, between 8 and 14%. However, the single most frequent aa is Ser (18%). In position P3′, all aliphatic aa are represented along with His, Arg, Glu and Asp. However, His is in this position exclusively contributed from the non-randomized region. The truly most preferred aa in position P3′ is probably Leu.

Interestingly, two individual phages within the panel display almost identical hexapeptides, Trp-Gly-Val-Tyr-Leu-Asp (WGVYLD) and Trp-Gly-Val-Tyr-Leu-Glu (WGVYLE) (Fig. 2, arrows). However, the consensus representing the potentially

Fig. 3. Distribution of aa present in positions P4–P3′ in random phage sequences cleaved by the DC. Upon alignment of phage-displayed random sequences cleaved after five selection rounds (Fig. 2), the percentage of each aa in positions P4–P3′ was calculated. For clarity, aa are displayed in functional groups, starting to the left with aromatic aa and ending with acidic aa to the right.
most susceptible peptide for the DC, as inferred from the collected 51 sequences, is from P4 to P3: (Gly)/Val—Gly/Val—Arg/Gly—Phe/Tyr/frp—Leu—Ser/Leu > Glu/Asp—Leu/Val.

Verifying the consensus sequence by the use of a new type of recombinant protein substrate

In order to estimate how much the different residues of the consensus site contribute to the specificity, a new type of substrate was developed. The consensus sequence obtained from the phage-display analysis was inserted in the linker region between two *E. coli* thioredoxin molecules, by ligating a double-stranded oligonucleotide encoding the actual sequence into a *Bam*HI site and a *Sal*I site of the vector construct (Fig. 4A). For purification purposes, a His$_6$-tag was added to the C-terminal of this protein (Fig. 4A). A number of related and unrelated substrate sequences were also produced with this system by ligating the corresponding oligonucleotides into the *Bam*HI–*Sal*I sites of the vector. All these substrates were expressed as soluble proteins in a bacterial host, *E. coli*, and purified on IMAC columns to obtain a protein with a purity of 90–95%. These recombinant proteins were then used to study the preference of the DC for these different sequences (Fig. 4B–E). The result shows that the DC very efficiently cleaves the DC consensus (VVRFSLLL) and also the HC consensus sequence (VVLFESEVL) (48). Based on the intensity of the different bands on the gel, we can see that the DC actually prefers the HC consensus site over the DC consensus site by a factor of 5 (Fig. 4B). By changing the glutamic acid residue in the P2 position of the HC consensus sequence into a glycine (VVLFGSGL), the efficiency in cleavage by the DC drops by a factor of 5–8 (Fig. 4B). This latter experiment shows that the DC has a preference for negatively charged over ali- phatic aa in its con- sensus sequence with glycine, causing the rate of cleavage to drop also by a factor of 5 (Fig. 4E). HC cleaved the DC consensus sequence seven to eight times less efficient than the HC consensus sequence (Fig. 4E). Except for the negatively charged aa in the P2 position, the major difference between the HC and the DC consensus sequences is the arginine in the P2 position of the DC consensus sequence. The different rate of cleavage of the DC and the HC is therefore most likely due to the specificity in the P2 position, where the DC accepts and potentially even prefers positively charged aa, whereas the HC does not show such a preference.

A few additional substrates were also included in this study. The optimal sequence for cleavage by the opossum chymase has recently been determined (19). Compared with the HC and the DC, this enzyme was found to have a preference for tryptophan over phenylalanine and tyrosine in the P1 position. When analysing the cleavage of this sequence (VGLWLDVR), we could observe that the DC cleaves this se-
Fig. 4. Analysis of the cleavage specificity by the use of new types of recombinant protein substrates. Panel (A) shows the overall structure of the recombinant protein substrates used for analysis of the efficiency in cleavage by DC and the HC. In these substrates, two thioredoxin molecules are positioned in tandem and the proteins have a His6-tag positioned in their C-termini. The different cleavable sequences are inserted in the linker region between the two thioredoxin molecules by the use of two unique restriction sites, one BamHI site and one Sall site, which are indicated in the bottom of panel (A). Panels (B–D) show the cleavage of a number of substrates by DC. The name and sequence of the different enzyme consensus substrates (cons. seq.) are indicated above the pictures of the gels. The time of cleavage is also indicated above the corresponding lanes of the different gels. The un-cleaved substrates have a molecular weight of ~25 kDa, and the cleaved substrates appear as two closely located bands of a size of 12–13 kDa. Panels (E and F) show cleavage of the same substrates as for the DC in panels (B–D) but now instead after cleavage with the HC.
the DC to self-digest (auto-cleave itself). We observed that prolonged activation cleavage with EK resulted in the DC totally disappearing from the sample. This in marked contrast to other MC chymases that we have studied with the same technology, indicating that the DC is less specific than the other enzymes.

In the phage-display analysis of the DC, statistically one of eight phages was expected to represent a non-specifically released phage (background phage). However, an even higher ratio of approximately one in three background phages was identified. These phages were easy to discriminate from truly cleaved phages due to their lack of aromatic aa and due to their repeated occurrence in different experiments or because they contained stop codons. Therefore, the presented alignment of sequences is likely to provide a rather accurate picture of sequences cleaved by DC, despite the low enrichment.

The DC has previously been shown to display chymotryptic activity (13, 29). In line with this, aromatic aa are strongly over-represented in the sequenced phages and align mostly to the P1 position. Phenylalanine is the preferred P1 residue (43%) and Tyr is almost equally frequent (37%), whereas Trp is found at a lower frequency (19%). A preference for Phe and Tyr has also previously been reported for the HC (50), and this preference has recently been confirmed with substrate phage display (48). Also, mMCP-4 prefers P1 aromatic aa in the same order as the DC in substrate phage display (51). rMCP-1 prefers Phe as well, but here, Trp is favoured ahead of Tyr in the P1 position (51). Thus, preference for Phe seems to be a common feature for the four chymases.

On the other hand, regarding position P2, the four enzymes fall into two groups: Arg is the single most frequent P2 aa for the DC and the second most frequent aa for rMCP-1, whereas the HC and mMCP-4 prefer aliphatic aa, admitting Arg only rarely (48, 51). Interestingly, Pro is not found frequently in this position for any of the enzymes, although Pro is the P2 residue in the substrate Ang I, which is cleaved well by HC, DC and mMCP-4 in vitro (48, 51). These three enzymes display a preference for the Phe$_{18}$-His$_{19}$ bond in Ang I, which leads to generation of the active product Ang II. rMCP-1, on the other hand, is known to destroy Ang II by cleavage of the Tyr$_{4}$-Ile$_{5}$ bond (29, 37, 38).

In positions P3 and P4, the four profiles are again similar to each other, with a frequent occurrence of aliphatic aa. A preference for Val in position P3 is largely conserved between the four chymases, and Val in P3 is also preferred by rMCP-4, an additional $\beta$-chymase in rat (35). Specific interactions with P3 residues distinguish the enzymes from the chymase locus, e.g. chymases, from other chymotrypsin-like serine proteases (52–54). The conserved preference for Val in position P3 may thus be functionally significant. The obtained consensus for the P4 positions may not be as reliable because Gly residues from the non-randomized phage region contribute to some degree.

Also in the positions C-terminal of the scissile bond (primed positions), both similarities and differences between the four enzymes were revealed. In position P’1, the DC clearly prefers Leu (35%), followed by Val (18%). Leu is also preferred by mMCP-4 in this position and is the second preferred residue of rMCP-1 (51). In contrast, the HC prefers Ser, which is found only in 4% of the P’1 positions in the dog sequences (48).

The preference in position P’2 also differs to some extent between the four chymases. Interestingly, the enzymes fall into the same two groups with regard to their preference in the position P2. Namely, rMCP-1 (which frequently admits Arg in P2) and to some extent DC prefer Ser and Gly in position P’2, whereas the HC and mMCP-4 (which admit Arg only infrequently) prefer the negatively charged residues Glu and Asp (48, 51). Preference for Arg in position P2 and preference for Glu/Asp in P’2 thus appear to be exclusive of each other to some degree. Notably, a number of Glu and Asp are also found in position P’2 among the phage sequences cleaved by DC but only in two cases together with an Arg in position P2 (Fig. 2). Thus, the DC seems to select substrates with either Arg in position P2 or Glu/Asp in P’2, which indicates that substrate residues in one position may affect residues in other positions. The finding that Leu is the exclusive aliphatic P’1 residue found adjacent to acidic P’2 residues further supports this notion. Synergistic–antagonistic interactions may also explain why two very similar hexapeptides, Trp-Gly-Val-Tyr-Leu-Asp (WGVYLD) and Trp-Gly-Val-Tyr-Leu-Glu (WGVYLE) (Fig. 2, arrows), were among the sequenced phages, although they do not represent the consensus as determined from the collected sequences. Indications for synergistic/antagonistic effects between substrate residues have also been obtained from the phage-display analysis of the HC (48). Substrate phage display is a unique tool for the detection of such interactions because residues on both sides of the scissile bond can be studied at the same time. Position P3 is dominated by aliphatic residues for all four chymases, with Val being among the most preferred residues for all four enzymes.

In order to validate the results obtained from the phage-display analysis, we studied the cleavage of a number of custom-designed recombinant substrates. This analysis confirmed that both the DC and HC strongly prefer Phe and Tyr over Trp in the P1 position as they cleave the HC consensus

<table>
<thead>
<tr>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P’1</th>
<th>P’2</th>
<th>P’3</th>
</tr>
</thead>
<tbody>
<tr>
<td>rMCP-1</td>
<td>Val</td>
<td>Val</td>
<td>Leu &gt; Arg</td>
<td>Phe &gt; Trp &gt; Tyr</td>
<td>Val &gt; Leu/Ser</td>
<td>Ser &gt; Leu/Gly</td>
</tr>
<tr>
<td>mMCP-4</td>
<td>Val</td>
<td>Val</td>
<td>Leu</td>
<td>Phe &gt; Tyr &gt; Trp</td>
<td>Leu &gt; Met</td>
<td>Glu/Asp</td>
</tr>
<tr>
<td>HC</td>
<td>Gly</td>
<td>Leu</td>
<td>Ala/Val/Leu</td>
<td>Phe/Tyr</td>
<td>Ser/Ala</td>
<td>Glu/Asp</td>
</tr>
<tr>
<td>DC</td>
<td>Gly</td>
<td>Val</td>
<td>Arg &gt; Gly</td>
<td>Phe &gt; Tyr &gt; Trp</td>
<td>Leu &gt; Val</td>
<td>Ser &gt; Glu/Asp &gt; Leu</td>
</tr>
</tbody>
</table>
sequence much more efficiently than the opossum chymase consensus sequence (10 and 50 times more efficiently, respectively). The HC was also found to be considerably more restrictive for other positions and favoured the HC consensus over that of the DC by a factor of six to eight times. In line with this, we found no activity of the HC against the thrombin sequence.

The DC, in contrast, was shown to be less restrictive in its extended cleavage specificity. It cleaved the consensus sequence obtained from the phage-display analysis (VVRFLSLL) efficiently. However, it cleaved the HC consensus sequence (VVLFSEVL) even better. This indicates that a negatively charged aa in the P2’ position is favoured also by the DC. The DC also cleaved the thrombin sequence to a low extent. The information from several positions in the substrate sequence thereby clearly shows that the DC is less restrictive in its extended specificity than the HC.

In summary, our study predicts the following consensus, from P4 to P3’, to represent peptides that are very susceptible to cleavage by the DC: (Gly)Val—Gly/Val—Arg—Phe > Tyr > Trp—Leu—Ser/Leu > Glu/Asp—Leu/Val. This consensus is similar to that of the HC in positions P1, P3, P4 and P3’, whereas clear differences are observed in positions P2 and P1’ and to a lesser extent in P2’. As the DC and HC are highly similar on the aa level, these differences were quite unexpected. The DC also appears to be substantially less specific than its human counterpart and shows relatively strong tendencies of auto-cleavage, a characteristic that is observed to a much lesser extent in the human enzyme. Interestingly, the extended cleavage specificity of DC is very similar to that of the rat β-chymase rMCP-1, although DC efficiently converts the substrate Ang I to Ang II, whereas rMCP-1 is mainly Ang II destroying. The in vivo relevance of these observations remains to be investigated. Tentatively, the four compared chymases may cleave a substrate where positions P1, P3, P4 and P3’ are conserved between the species, whereas positions P2 and P1’ and to a lesser extent P2’ differ.

Bioinformatic screenings for such substrates are currently being conducted. However, it may prove relatively difficult to identify new in vivo substrates and also to draw any conclusions concerning the substrates already proposed due to the relatively low specificity of the DC. Prolonged incubation apparently leads to the cleavage of a very broad range of potential substrates including the enzyme itself. The DC may cleave very many different proteins if given sufficient time and concentration of the enzyme. It is therefore important to take into account factors other than the extended cleavage specificity that determine the function of the DC in its biological setting. The extended cleavage specificity primarily predicts substrate sequence having an advantage in time over other less favoured substrates. These substrates have a chance of being cleaved in the short time window before the enzyme is auto-cleaved or inactivated by protease inhibitors—if they are present in the local MC area. It is then also of major importance how well exposed the potential sites are in the actual protein. Surface-exposed sites in a flexible region will probably be rapidly cleaved, whereas hidden sites in a more rigid structure will not or only rarely, regardless the fact that they contain a theoretically apt sequence. Finally, it needs to be determined whether cleavage of these substrates in the local MC area will actually lead to a biological effect.

However, although difficult, the combination of such criteria and the now better defined extended cleavage specificity will be used to intensify the screening for biologically significant substrates for these key enzymes in MC biology.

**Funding**

Swedish National Research Council (VR-NT) (621-2008-3248).

**References**


Extended cleavage specificity of the dog chymase


53 Hof, P., Mayr, I., Huber, R. et al. 1996. The 1.8 A crystal structure of human cathepsin G in complex with Suc-Val-Pro-Phe(OMe)2–a Janus-faced proteinase with two opposite specificities. EMBO J. 15:5481.