Dog Mast Cell Chymase: Molecular Cloning and Characterization†,‡

George H. Caughey,* Wilfred W. Raymond, and Peter Vanderslice
Cardiovascular Research Institute and Department of Medicine, University of California, San Francisco, California 94143
Received December 8, 1989; Revised Manuscript Received February 6, 1990

ABSTRACT: We cloned and characterized a cDNA coding for the complete amino acid sequence of dog mast cell chymase. The cDNA was identified by screening a dog mastocytoma cDNA library with an oligonucleotide probe based on the amino acid sequence of a fragment of dog mastocytoma chymase. The deduced amino acid sequence reveals a putative 21-residue prepropeptide followed by a catalytic domain of 228 residues. The primary structure of the preproenzyme shares features with rat mucosal mast cell chymase (RMCP II), several lymphocyte-associated proteases, and neutrophil cathepsin G. The common characteristics include an apparent activation peptide terminating in glutamic acid, strict conservation of an octapeptide (residues 9–16) in the N-terminal portion of the catalytic domain, and the presence of only six cysteines available for intramolecular disulfide bond formation. However, dog chymase differs in being modified by N-glycosylation. Although the dog chymase catalytic domain exhibits a similar level of sequence identity when compared with both RMCP II and the rat connective tissue mast cell chymase RMCP I (58% and 61%, respectively), the dog enzyme most closely resembles RMCP I in its high predicted net charge (+16) and in the presence of serine at the base of its putative primary substrate binding pocket. The dog chymase differs strikingly from dog mast cell tryptase in the preprosequence and in the structure of the catalytic domain. Therefore, chymase appears not to be closely related to tryptase and may not share a mechanism of activation, even though both enzymes are packaged and released together.

Chymases are mast cell secretory proteases of chymotrypsin-like specificity with suspected roles in neurogenic inflammation (Brain & Williams, 1988), control of vasoactive peptide generation (Reilly et al., 1982; Wintroub et al., 1984; Caughey et al., 1988b; Franconi et al., 1989), submucosal gland secretion (Sommerhoff et al., 1989), parasite expulsion (King & Miller, 1984), and lipoprotein (Kokkonen & Kovanen, 1989) and extracellular matrix catabolism (Seppa et al., 1979; Vartio et al., 1981). The mast cell chymases belong to a broader group of granule-associated serine proteases whose most closely related other members appear to be neutrophil cathepsin G (Salveson et al., 1987) and certain lymphocyte proteases (Jenne & Tschopp, 1988). Mast cell chymases can vary to a striking degree in catalytic and physicochemical properties. These differences have been most extensively documented in comparisons of the two rat chymases [rat mast cell chymase I (RMCP I)[1] and RMCP II] which differ substantially in net charge, solubility, catalytic efficiency, association with proteoglycans, and susceptibility to inactivation by circulating protease inhibitors (Woodbury et al., 1978; Yoshida et al., 1980; Powers et al., 1985; Le Trong et al., 1987). Each rat chymase is produced exclusively by a distinct subset of mast cells. RMCP I is expressed by a population of mast cells that predominate in skin and connective tissues, whereas RMCP II is expressed by mast cells which predominate in mucosal tissues and proliferate in large numbers in response to parasitic infestation (Gibson & Miller, 1986). The basis for these differences in properties and cell of origin is not yet well understood in terms of biological function. Mast cell chymases from several other sources, including mouse (Le Trong et al., 1989), dog (see below), sheep (Huntley et al.,1986), and human (Schechter et al., 1983, 1986; Wintroub et al., 1986; Sayama et al., 1987), have been partially characterized.

A number of investigators have explored chemical and biological features of dog chymase, the focus of the present study. Chymotrypsin-like activity was identified initially in dog mast cells by enzyme histochemical techniques (Glenner et al., 1962). Subsequently, the catalytic properties of small amounts of enzyme isolated from dog skin were compared with those of chymases from other species using chromogenic substrates (Powers et al., 1985). More recently, the purification of larger amounts of chymase from dog skin mastocytomas permitted more extensive characterization (Caughey et al., 1988d; Schechter et al., 1988). Chymase purified from a line of dog mastocytoma cells exhibits marked catalytic subsite preferences when probed with synthetic peptide 4-nitroanilide substrates (Caughey et al., 1988d), removes components of glycocalyx from cultured airway epithelial cells (Varsano et al., 1987), hydrolyzes and inactivates the bioactive airway neuropeptides substance P and vasoactive intestinal peptide (Caughey et al., 1988b; Franconi et al., 1989), and is a highly potent secretagogue for cultured airway submucosal gland serous cells (Sommerhoff et al., 1989). Stimulation of dog mastocytoma cells with ionophore A23187 causes noncytotoxic release of active chymase along with tryptase (Caughey et al., 1988a), the other major neutral serine protease of dog mast cells (Caughey et al., 1987). These studies found no evidence of coupling of A23187-stimulated secretion to activation of either chymase or tryptase from putative inactive precursors. Enzyme histochemical studies appear to localize both chymase and tryptase to mast cell granules (Caughey et al., 1988c).

† Supported by Grants HL-24136 (G.H.C.) and HL-01736 (G.H.C.) from the National Institutes of Health and by the American Lung Association of California Research Fund (G.H.C.). G.H.C. is an RJR Nabisco Research Scholar, and P.V. is a Parker B. Francis Fellow.
‡ The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J02904.
* Correspondence should be addressed to this author at the Cardiovascular Research Institute, University of California, San Francisco, CA 94143-0130.

[1] Abbreviations: RMCP I, rat mast cell chymase I; RMCP II, rat mast cell chymase II.

0006-2960/90/0429-5166$02.50/0

However, variations in apparent amounts of chymase vis-à-vis tryptase in a mixed population of cells in varying stages of differentiation suggest the possibility that the two enzymes may be expressed asynchronously during mast cell development. Recently, a cDNA coding for dog tryptase was cloned from a dog mastocytoma cDNA library (Vanderslice et al., 1989). The deduced sequence of the cloned tryptase suggests a unique mode of activation.

The current study had two principal objectives: (1) to compare the activation sequence and catalytic domain of dog chymase with that of dog tryptase; (2) to clarify the relation between structure and function among mast cell chymases and related leukocyte proteases. Here we report the full primary structure of dog preprochymase deduced from a mastocytoma-derived cDNA and the modification of the catalytic domain by N-glycosylation.

Materials and Methods

Protease Isolation and Purification. Dog mast cell chymase was obtained from the "BR" line of mastocytoma cells (Caughey et al., 1988a,c) that was derived from a dog skin mastocytoma and established as a stable line by serial passage as nodules in the skin of nude mice by Lazarus and co-workers (Lazarus et al., 1986). Active chymase was isolated and purified chromatographically from 2 M NaCl mastocytoma extracts as previously described (Caughey et al., 1988d).

Enzymatic Deglycosylation and Lectin Binding Studies. To investigate the possible N-linked glycosylation of dog chymase, 5 μg of the purified protease was incubated as described (Tarentino et al., 1985) with and without peptide:*N*-glycosidase F (Genzyme, Boston, MA). The samples then were denatured by boiling in the presence of 2-mercaptoethanol and SDS and were subjected to SDS–PAGE. To explore further the nature of the glycans linked to chymase, 5-μg samples subjected to SDS–PAGE (12.5%) were electroblotted to nitrocellulose and were probed with 4 μg/mL biotinylated lens culinaris (lentil) lectin or Con A (EY Labs, San Mateo, CA) in 20 mM Tris-HCl containing 0.5 M NaCl, 0.3% Tween-20, 2 μM $MgCl_2$, and 2 μM $CaCl_2$ for 1 h at 20 °C. The blots were subsequently washed and labeled with a 0.5 μg/mL dilution of avidin–horseradish peroxidase (EY Labs) and were visualized with 4-chloronaphthol/H_2O_2 (Faye & Chrispeels, 1985).

CNBR Cleavage. Purified dog chymase (290 μg) was digested completely with 2 mg of CNBR in 70% formic acid, as assessed by SDS–PAGE. After being dried under N_2, the sample was reconstituted in 0.1% trifluoroacetic acid. Peptide fragments of CNBr-digested chymase were purified by reverse-phase HPLC on a C4 column (Vydac, Hesperia, CA) using a linear gradient of 5–50% acetonitrile in 0.1% trifluoroacetic acid. Selected column fractions corresponding to well-separated chromatographic peaks were collected and reduced in volume under a stream of dry N_2. Automated N-terminal sequence analysis of the purified peptides was carried out by the Biomolecular Resource Center of the University of California, San Francisco, using an Applied Biosystems (Foster City, CA) gas-phase sequenator.

Molecular Cloning of Dog Chymase. To obtain the full sequence of chymase, including its signal sequence and putative activation peptide, a cDNA library prepared in λgt10 from BR mastocytoma poly(A) RNA (Vanderslice et al., 1989) was screened with a degenerate synthetic oligonucleotide probe based on the internal amino acid sequence obtained from a CNBr cleavage peptide. The probe was constructed by the Biomolecular Resource Center of the University of California, San Francisco, using an Applied Biosystems 380B DNA synthesizer and was radioactively phosphorylated with [γ-

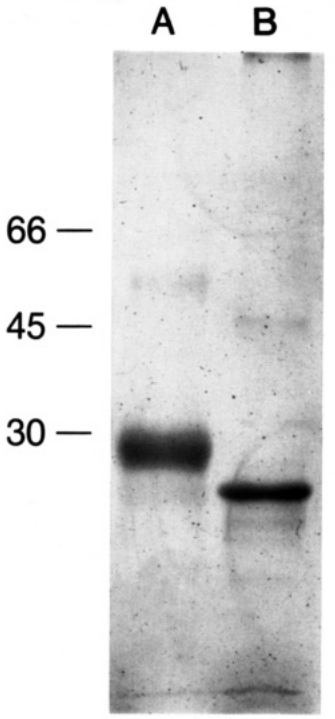


FIGURE 1: SDS–PAGE of dog mast cell chymase. Lane A contains 5 μg of chymase purified from dog BR mastocytoma cells. Lane B contains 5 μg of dog chymase after deglycosylation with peptide:*N*-glycosidase F. The positions of molecular weight marker proteins are shown to the left: BSA (66K), ovalbumin (45K), and carbonic anhydrase (30K). Lanes were stained with Coomassie Blue.

^{32}P]ATP and T4 polynucleotide kinase (Bethesda Research Laboratories, Gaithersburg, MD) (Maniatis et al., 1982) to a specific activity of $\sim 2 \times 10^8$ cpm/μg. Library screening was carried out by plaque hybridization as previously described (Vanderslice et al., 1989) with 20% formamide in the hybridization solution. Recombinants hybridizing to the oligonucleotide probe were plaque-purified and reprobed. Recombinant phage DNA isolated by the plate lysate method (Maniatis et al., 1982) was digested with *Eco*RI (Bethesda Research Laboratories) and was electrophoresed through a 1% agarose gel to determine insert size. One of the largest inserts was cloned into the *Eco*RI site of the plasmid vector pBluescript (Stratagene, LaJolla, CA) which was used to transform *Escherichia coli* XLI-Blue cells (Bullock et al., 1987). Double-stranded DNA was isolated and prepared as described (Birnboim & Doly, 1979). The nucleotide sequence of the cDNA was determined by the dideoxy chain termination method (Sanger et al., 1977) as modified for double-strand sequencing (Chen & Seeburg, 1985) using Sequenase (United States Biochemical) (Tabor & Richardson, 1987). M13 and pBluescript KS primers (Stratagene) were used to obtain nucleotide sequence information near the 3′ and 5′ ends of both strands of the cloned cDNA. A series of synthetic oligonucleotides designed from previously determined sequence was used to prime subsequent sequencing reactions.

Results

Deglycosylation and Lectin Binding. Treatment of purified chymase with peptide:*N*-glycosidase F resulted in sharpening of the Coomassie-stained protein band on SDS–PAGE and a drop in the apparent molecular weight from 29.1K to 26.6K (Figure 1). These findings suggest that dog mast cell chymase is N-glycosylated. This form of posttranslational modification probably accounts for the broader electrophoretic banding and higher apparent molecular weight of dog chymase compared

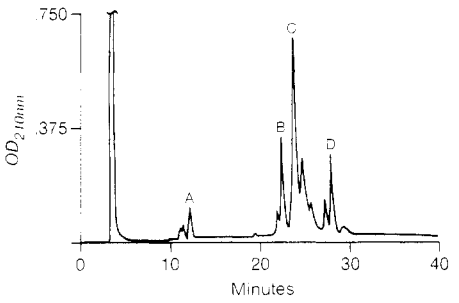


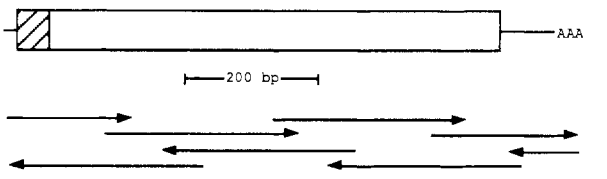
FIGURE 2: Reverse-phase HPLC of dog mast cell chymase CNBr peptides. Cleavage fragments resulting from incubation of purified dog chymase with CNBr were eluted from a Vydac C_4 reverse-phase HPLC column with a linear gradient of CH_3CN in 0.1% trifluoroacetic acid. Peaks A through D were collected and subjected to automated Edman degradation to determine the amino acid sequence.

to the rat chymases RMCP I and RMCP II, which lack consensus N-linked glycosylation sites and are probably not N-glycosylated (Woodbury et al., 1978; Le Trong et al., 1987). Glycosylated chymase transferred to nitrocellulose bound to both lens culinaris lectin and Con A (data not shown). After deglycosylation, chymase binding to both lectins was greatly decreased. These results suggest that dog chymase has been modified posttranslationally through attachment of complex-type N-linked glycans.

CNBr Peptide Sequence. Four HPLC-purified peptides from chymase CNBr digests were subjected to N-terminal sequencing. Peptide A (Figure 2) was a 14-residue fragment corresponding to the N-terminus of the intact protease, confirming the previously determined N-terminal sequence (Caughey et al., 1988d). The sequence of peptide D overlapped with the previously determined N-terminal sequence of the intact protease beginning with Ala^{16} (Figure 3) but yielded clearer data, resulting in revision of the previous assignments at residues 17 and 24. The amino acid sequence determined from these contiguous CNBr fragments agrees with that deduced from the cloned chymase cDNA (see below and Figure 3). The sequence of the CNBr peptides B and C did not overlap with the previously determined sequence (Figure 3). Peptide B corresponds to residues Val^{53} through His^{78} of the deduced sequence, and peptide C corresponds to residues Cys^{123} through Leu^{141}. The 1st residue of CNBr peptide C could not be determined clearly; the 12th cycle yielded a blank followed 2 cycles later by a serine, suggesting a potential glycosylation site.

Molecular Cloning of Dog Chymase. Approximately 9×10^5 plaques from the dog mastocytoma cDNA library were screened with a 44-base oligonucleotide probe based on the amino acid sequence of dog chymase CNBr fragment B. Eight of 63 plaques hybridizing with the probe were plaque-purified, and their DNA was isolated and digested with *Eco*RI. One of the largest inserts was selected for subcloning into the vector pBluescript for sequencing. The 943 base pair cDNA chosen for sequencing reveals an open reading frame of 747 nucleotides terminating with the stop codon TGA (Figure 3). The open reading frame is preceded by nine bases of the apparent 5′ untranslated region. The 3′ end of the cDNA insert following the stop codon contains 184 nucleotides. A consensus polyadenylation signal (AATAAA) begins 18 residues 5′ to a terminal sequence of 10 adenines, which therefore represents the poly(A) tail of the processed transcript.

The 249-residue amino acid sequence deduced from the cDNA open reading frame includes peptides identical with the N-terminus and with the portions of internal sequence of dog chymase identified by Edman sequencing of the intact catalytic domain and of internal fragments generated by CNBr cleavage




```
                                                                  CCTGAGAAG      9

ATG CAT TGT CTT CCT CTC ACC CTG CTG CTC CTT CTC CTA TGT TCC      54
met his cys leu pro leu thr leu leu leu leu leu leu cys ser
-21                                                      -10

AGA GCA GAA GCT GAG GAG ATC ATC GGG GGC ACA GAG TCC AAG CCA      99
arg ala glu ala glu glu ile ile gly gly thr glu ser lys pro
                    -1  +1

CAC TCC CGC CCT TAC ATG GCC CAC CTG GAA ATC CTC ACT CTC CGG     144
his ser arg pro tyr met ala his leu glu ile leu thr leu arg
10                                                      20

AAT CAC CTG GCG TCT TGT GGT GGT TTC CTG ATA AGG CGG AAC TTC     189
asn his leu ala ser cys gly gly phe leu ile arg arg asn phe
                30

GTA CTG ACA GCT GCT CAC TGT GCA GGA AGG TTC ATA ATG GTC ACC     234
val leu thr ala ala his cys ala gly arg phe ile met val thr
40                                                      50

CTG GGA GCC CAT AAC ATA CAA AAG AAA GAA GAC ACG TGG CAG AAG     279
leu gly ala his asn ile gln lys lys glu asp thr trp gln lys
                        60

CTT GAG GTC ATA AAA CAA TTC CCT CAC CCA AAA TAT GAT GAC CTT     324
leu glu val ile lys gln phe pro his pro lys tyr asp asp leu
70                                                      80

ACT CTT CGC CAT GAC ATC ATG CTA CTG AAG TTG AAG GAG AAG GCC     369
thr leu arg his asp ile met leu leu lys leu lys glu lys ala
                    90
 *
AAC CTG ACC CTG GCC GTG GGG ACC CTC CCC CTC TCA CCC CAG TTC     414
asn leu thr leu ala val gly thr leu pro leu ser pro gln phe
100                                                     110

AAC TTC GTC CCA CCT GGG AGA ATG TGC CGG GTG GCT GGC TGG GGA     459
asn phe val pro pro gly arg met cys arg val ala gly trp gly
                    120
                 *
AAA AGA CAA GTC AAT GGA TCA GGC TCT GAC ACT CTG CAA GAG GTG     504
lys arg gln val asn gly ser gly ser asp thr leu gln glu val
130                                                     140

AAG CTG AGG CTC ATG GAC CCC CAG GCC TGT AGA CAC TAC ATG GCT     549
lys leu arg leu met asp pro gln ala cys arg his tyr met ala
                    150

TTT GAC CAC AAC CTC CAG CTG TGT GTT GGC AAT CCC AGG AAG ACA     594
phe asp his asn leu gln leu cys val gly asn pro arg lys thr
160                                                     170

AAA TCT GCA TTT AAG GGA GAT TCA GGG GGC CCT CTT CTG TGT GCT     639
lys ser ala phe lys gly asp ser gly gly pro leu leu cys ala
                    180

GGG GTA GCC CAG GGA ATT GTG TCC TAT GGG CAG AAT GAT GCA AAG     684
gly val ala gln gly ile val ser tyr gly gln asn asp ala lys
190                                                     200

CCC CCT GCT GTC TTC ACC CGA ATC TCC CAC TAC CGG CCC TGG ATC     729
pro pro ala val phe thr arg ile ser his tyr arg pro trp ile
                    210

AAT AAG GTT CTG AAG CAG AAT AAA GCC TGA ATGGAGCCTGGGCAAGCCT    778
asn lys val leu lys gln asn lys ala ***
220                                 228

GAGGGGAAATCTGGAACCAGACCTGAGCAGGCTCTCTGCCACTTACTCTGGAGCTGCCT
CTAGTCTCTACTGAGGCCCCACTACATCCCTCAGACCCAAGAGGGCTCCTCCAGGTCAC
AGAGTCCTCAGTAAACCTCAATAAAGATCCAGCTTCCAAAAAAAAAA               943
```

FIGURE 3: Nucleotide and deduced amino acid sequence of dog mast cell chymase. The upper portion of the diagram depicts the sequencing strategy. The open rectangle shows the portion of the open reading frame coding for the chymase catalytic domain. The hatched rectangle represents the nucleotides corresponding to the proposed preprosequence. The 5′ and 3′ noncoding regions are indicated by lines flanking the open reading frame. Arrows indicate the length and direction of sequence identified by dideoxy sequencing. The lower portion of the diagram shows the full nucleotide sequence and deduced amino acid sequence of the cloned cDNA. The solidly underlined amino acids represent portions of catalytic domain confirmed by sequencing of CNBr peptides. The dashed underlines represent portions of CNBr sequence for which amino acid assignments could not be made. Asterisks mark consensus N-linked glycosylation sites. The essential "catalytic triad" residues common to all serine proteases are boxed. The shaded box identifies the serine associated with serine proteases of chymotrypsin-like substrate specificity. The apparent 3′-polyadenylation signal is overlined.

(Figure 3). The cDNA predicts a catalytic domain of 228 amino acids. His^{45}, Asp^{89}, and Ser^{182} of the deduced chymase sequence correspond to the active-site residues [His^{57}, Asp^{102}, and Ser^{195} in chymotrypsinogen (Figure 4)] essential for the

FIGURE 4: Alignment of protease sequences. The amino acid sequence of dog mast cell preprochymase is compared with that of the rat mucosal mast cell preprochymase RMCP II (rat chymase II) (Benfey et al., 1987), dog mast cell preprotryptase (Vanderslice et al., 1989), and bovine chymotrypsinogen. Rat chymase and chymotrypsinogen residues were aligned according to Remington et al. (1988) based on the proximity of α-carbons in their respective tertiary structures. Dog chymase was aligned with rat chymase, and dog tryptase was aligned manually to optimize sequence similarities with the three other proteases. Aligned residues identical with those of dog chymase are boxed. In the areas of overlap in the catalytic domain, sequence identity for dog chymase with RMCP II, tryptase, and chymotrypsinogen is 58.0%, 31.4%, and 31.4%, respectively.

catalytic activity of all serine proteases. In addition, the cloned dog chymase contains the previously identified octapeptide (amino acids 9–16) shared by chymases, neutrophil cathepsin G, and lymphocyte granzymes (Caughey et al., 1988d), and it contains consensus N-linked glycosylation sites (Asn-X-Ser/Thr) at two sites (Asn^{100} and Asn^{134}). The dog chymase catalytic domain, like that of cathepsin G (Salveson et al., 1987) and other chymases (Woodbury et al., 1978; Le Trong et al., 1987, 1989), contains only six cysteines available for formation of intramolecular disulfide bonds. The 21 amino acid polypeptide preceding the known N-terminus of the catalytic domain is dominated by a highly hydrophobic sequence consistent with a signal peptide and ends C-terminally in a pair of glutamic acids.

Discussion

By sequencing cloned cDNA from a mastocytoma library, we have determined the full primary structure of dog mast cell preprochymase. The amino acid sequence deduced from the major open reading frame of the cloned cDNA confirms the partial sequence obtained from the purified protease.

Selection of the ATG nearest the 5′ end of the open reading frame as the site of the initiator methionine predicts that a 21-residue prepropeptide precedes the recognized N-terminus of the mature, active enzyme. This preprosequence is similar to the 20 amino acid prepropeptide suggested for RMCP II (Benfey et al., 1987). The dog chymase and RMCP II precursor sequences both include a highly hydrophobic sequence and terminate in a pair of acidic residues (GluGlu). Empiric prediction rules (Von Heijne, 1986) suggest that signal peptidase hydrolyzes the prepropeptide on the C-terminal side of one of the two alanines, with Ala^{-3} slightly preferred over Ala^{-5}. The amino acids of the preprosequence remaining after signal peptide removal form a propeptide, possibly functioning as an activation peptide removed at a later step in protein maturation. Comparison of the preprosequences deduced from cloned cDNAs or genes encoding the enlarging group of other serine proteases with apparent prepropeptides ending in glutamic acid

Table I: Comparison of Preprosequences[a]

	↓
Dog chymase	...L C S R A E A E E I I...
Rat chymase (RMCP II) (Benfey et al., 1987)	...L P S G A G A E E I I...
Human cathepsin G (Salveson et al., 1987)	...L P T G A E A G E I I...
Human PMN elastase (Takahashi et al., 1988)	...L G G T A L A S E I V...
Mouse Granzyme C (Jenne & Tschopp, 1988)	...L P L R A G A E E I I...
Mouse Granzyme D (Jenne & Tschopp, 1988)	...L P L R A G A E E I I...
Mouse Granzyme E (Jenne & Tschopp, 1988)	...L P L G A G A E E I I...
Mouse Granzyme F (Jenne & Tschopp, 1988)	...L P L R A G A E E I I...
Mouse CCPI (Lobe et al., 1988)	...L A S R T K A G E I I...
Mouse CCPII (Lobe et al., 1988)	...L P L R A G A E E I I...
CONSENSUS	...L P L R A G A E E I I...

[a] Arrow indicates proposed site of propeptide hydrolysis.

allows formulation of a consensus sequence (LeuProLeu-ArgAlaGlyAlaGluGlu) in the region preceding and including the proposed propeptide (see Table I). Among the preprosequences, the Ala in the −3 position is invariable, and the Ala in the −5 position is highly preferred. Both chymases (dog and RMCP II) and most of the lymphocyte "granzyme" group (Jenne & Tschopp, 1988) contain a C-terminal pair of glutamic acids, whereas the human neutrophil proteases cathepsin G (salveson et al., 1987) and elastase (Takahashi et al., 1988) lack an acidic residue in the −2 position. The similarities in preprosequence in this group of proteases sharing a glutamic acid in the −1 position suggest that they probably also share sites of signal peptidase cleavage and a mechanism of proenzyme activation.

The two cysteines identified in the dog chymase preprosequence appear to be embedded in the signal sequence. Dog chymase thus differs from bovine α-chymotrypsin (Birktoft & Blow, 1972), the best characterized serine protease with a preference for P1 aromatic residues. After hydrolysis of the C-terminal arginine, the 15-residue activation peptide of chymotrypsin remains attached to the active protease via a

disulfide linkage between Cys^{1} of the activation peptide and Cys^{122} of the catalytic domain (see Figure 4) (Birktoft & Blow, 1972).

Although dog chymase and bovine chymotrypsin share the His, Asp, and Ser residues of the catalytic triad and a preference for substrates with P1 residues containing aromatic side chains, they differ in several potentially important respects in the catalytic domain. Purified dog chymase, unlike chymotrypsin, is not subject to autolysis. This has been established by incubation of active chymase for 16 h at 37 °C which does not result in the appearance of lower molecular weight bands on SDS–PAGE under reducing conditions (data not shown). In chymotrypsin, autolytic cleavage after Tyr^{146} and Asn^{148} produces an active two-chain catalytic domain with a deleted dipeptide $Thr^{147}Asn^{148}$. In dog chymase, the residue corresponding to Tyr^{146} in chymotrypsin is Gln or Val (see Figure 4), which are less attractive sites of chymotryptic hydrolysis than Tyr. The residue corresponding to Asn^{148} in chymotrypsin is also Asn in dog chymase, though it differs in being a consensus N-linked glycosylation site. This consensus site appears to be used as such, as indicated by the blank cycle at that position in the sequencing of CNBr peptide C (see Figures 2 and 3). Thus, the glycan attached to chymase Asn^{134} may protect the exposed "autolysis loop" from hydrolysis by chymase and other proteases. Comparison with the corresponding loop identified in crystalline RMCP II (Remington et al., 1988) suggests that the dog chymase Asn^{134}-linked glycan is adjacent to one end of the substrate binding cleft where it may be in a position to influence subsite interactions between enzyme and substrate or inhibitors. The predicted second dog chymase glycan linked to Asn^{100} should be located on the side of the protein opposite to that of the active site. Chymotrypsin and, indeed, RMCP I and RMCP II appear not to be glycosylated at any site (Woodbury et al., 1978; Le Trong et al., 1987; Birktoft & Blow 1972).

Dog chymase, like RMCP I (Le Trong et al., 1987), has a large number of basic residues in its catalytic domain, as reflected by its predicted net charge of +16 at neutral pH. In this respect, therefore, dog chymase is dissimilar to chymotrypsin, RMCP II, and dog tryptase, with a predicted net charge of +4, +3, and −3, respectively (Woodbury et al., 1978; Vanderslice et al., 1989). Resemblances between the structure of dog chymase reported here and the structure of RMCP I perhaps could be anticipated, in that the former was obtained from skin mastocytoma (Caughey et al., 1988d) and the latter is the chymase most prevalent in mast cells of rat skin (Gibson & Miller, 1986). However, notwithstanding the similarities in predicted net charge, dog chymase exhibits virtually the same level of sequence identity with RMCP I (60%) as with RMCP II (58%). The positively charged residues of dog chymase presumably are arranged mainly on the protein surface, where they may play a role in binding to heparin and to other sulfated proteoglycans and glycosaminoglycans of dog mast cell granules (Caughey et al., 1988d; Forsberg et al., 1988). Chymotrypsin, with its low density of positive surface charge, does not appear to bind to heparin at physiological ionic strength (Sayama et al., 1987). Lys^{179} of dog chymase, just three residues away from the catalytically active serine, is likely to be in a position to interact with P2 or P3 residues of chymase substrates. RMCP I and cathepsin G also possess a Lys at this position (Salveson et al., 1987; Le Trong et al., 1987). The corresponding residue in chymotrypsin is Met^{192}, which is also the residue found in RMCP II and in mouse mucosal mast cell chymase (Le Trong et al., 1989). Molecular modeling using the crystal structure of RMCP II (Remington et al., 1988) suggests that Met^{179} (Met^{192} using chymotrypsinogen numbering) helps to form a hydrophobic specificity pocket accommodating a P3 aromatic side chain, perhaps accounting for the highly favorable ~100-fold decrease in the K_m seen when a substrate with a P3 Phe is substituted for a substrate with a P3 Gly (Yoshida et al., 1980). Accurate prediction of the influence of the dog chymase Lys-179 on substrate or inhibitor preferences must await solution of the three-dimensional structure. Clearly, a Lys in this position does not hinder the interaction with large hydrophobic side chains because peptides with P3 Val or Phe are among the best synthetic substrates for both dog chymase and RMCP I (Yoshida et al., 1980; Powers et al., 1985; Caughey et al., 1988d).

Dog chymase shares with both RMCP I and chymotrypsin a Ser at position 189 (chymotrypsinogen numbering). There is an Ala at this site in RMCP II, mouse mucosal mast cell protease, and cathepsin G. The corresponding residue in dog mast cell tryptase, as in trypsin and all serine proteases with specificity toward Arg and Lys residues, is an Asp that is thought to stabilize the cationic P1 residue side chain of the substrate (Krieger et al., 1974). The extent to which substitutions at residue 189 affect substrate preferences and the catalytic efficiency of chymotrypsin-like serine proteases is not yet clear. However, replacement of Asp^{189} of trypsin with Ser or Lys using mutagenesis techniques profoundly alters catalytic activity toward substrates varying in the P1 amino acid (Graf et al., 1987, 1988). By analogy, differences in substrate preferences and catalytic rates among chymotrypsin-like proteases possibly are related to differences in the identity of the residue at position 189.

Dog chymase lacks the Cys^{191}–Cys^{220} disulfide bond found near the active site in chymotrypsin (Birktoft & Blow 1972). This bond, which is present in most serine proteases but is absent in cathepsin G and all chymases described to date, limits the opportunity for substrate subsite interactions by "locking up" the binding cleft on the side of the active site available for binding to substrate residues N-terminal to the site of the scissile bond (Birktoft & Blow, 1972; Remington et al., 1988). The tertiary structure of RMCP II seems to suggest that the absence of the Cys^{191}–Cys^{220} disulfide bond opens up the cleft between the twin β-barrels, thereby accommodating more productive subsite interactions, particularly with substrate P3 residue side chains (Remington et al., 1988). The three main-chain residues ($Ser^{214}Tyr^{215}Gly^{216}$, chymotrypsinogen numbering) thought to form a segment of extended antiparallel β structure with peptide substrates (Segal et al., 1971) are identical in dog chymase and RMCP II and are similar to the $Ser^{214}Trp^{215}Gly^{216}$ segment in chymotrypsin (see Figure 4).

The deduced primary structure of dog chymase, compared to that of RMCP II, contains an identically sized set of additions and deletions in the regions corresponding to residues 61 and 35–41 of chymotrypsin. In RMCP II, these changes in primary structure lead to the most dramatic difference in the tertiary structure of RMCP II compared to that of chymotrypsin, namely, a large shift in the position of a surface loop formed by residues 35–41 (chymotrypsinogen numbering) (Remington et al., 1988). This shift produces a deep cleft which may faciliate interactions with substrate P1′ aromatic residues on the C-terminal side of the scissile bond. Despite the close correspondence in length of segments corresponding to surface loops in dog chymase and RMCP II, there are numerous differences in sequence (see Figure 4). In the segment corresponding to residues 35–41 in chymotrypsin, for example, only 2 of 10 residues in dog chymase and RMCP

II are identical. Until the tertiary structure of dog chymase becomes available, it is hazardous to speculate regarding shared features involving surface loops. However, dog chymase hydrolyzes the neuropeptide substance P exclusively between Phe^7 and Phe^8 (Caughey et al., 1988b), with kinetic parameters comparing favorably to those of extended peptide 4-nitroanilides (Powers et al., 1985; Caughey et al., 1988d). Thus, on kinetic grounds, dog chymase appears to contain a subsite accommodating and favoring a P1′ aromatic residue.

References

Benfey, P. N., Yin, F. H., & Leder, P. (1987) *J. Biol. Chem. 262*, 5377.

Birktoft, J. J., & Blow, D. M. (1972) *J. Mol. Biol. 68*, 187.

Birnboim, H. C., & Doly, J. (1979) *Nucleic Acids Res. 7*, 1513.

Brain, S. D., & Williams, T. J. (1988) *Nature 335*, 73.

Bullock, W. O., Fernandez, J. M., & Short, J. M. (1987) *Biotechniques 5*, 376.

Caughey, G. H., Viro, N. F., Ramachandran, J., Lazarus, S. C., Borson, D. B., & Nadel, J. A. (1987) *Arch. Biochem. Biophys. 258*, 555.

Caughey, G. H., Lazarus, S. C., Viro, N. F., Gold, W. M., & Nadel, J. A. (1988a) *Immunology 63*, 339.

Caughey, G. H., Leidig, F., Viro, N. F., & Nadel, J. A. (1988b) *J. Pharmacol. Exp. Ther. 244*, 133.

Caughey, G. H., Viro, N. F., Calonico, L. D., McDonald, D. M., Lazarus, S. C., & Gold, W. M. (1988c) *J. Histochem. Cytochem. 36*, 1053.

Caughey, G. H., Viro, N. F., Lazarus, S. C., & Nadel, J. A. (1988d) *Biochim. Biophys. Acta 952*, 142.

Chen, E. Y., & Seeburg, P. H. (1985) *DNA 4*, 165.

Faye, L., & Chrispeels, M. J. (1985) *Anal. Biochem. 149*, 218.

Forsberg, L. S., Lazarus, S. C., Seno, N., DeVinney, R., Caughey, G. H., & Gold, W. M. (1988) *Biochim. Biophys. Acta 967*, 416.

Franconi, G. M., Graf, P. D., Lazarus, S. C., Nadel, J. A., & Caughey, G. H. (1989) *J. Pharmacol. Exp. Ther. 248*, 947–951.

Gibson, S., & Miller, H. R. P. (1986) *Immunology 58*, 101.

Glenner, G. G., Hopsu, V. K., & Cohen, L. A. (1962) *J. Histochem. Cytochem. 10*, 109.

Graf, L., Craik, C. S., Patthy, A., Roczniak, S., Fletterick, R. J., & Rutter, W. J. (1987) *Biochemistry 26*, 2616.

Graf, L., Jancso, R. j., Szilagyi, L., Hegyi, G., Pinter, K., Naray-Szabo, G., Hepp, J., Medzihradszky, K., & Rutter, W. J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4961.

Huntley, J. F., Gibson, S., Knox, D., & Miller, H. R. P. (1986) *Int. J. Biochem. 18*, 673.

Jenne, D. E., & Tschopp, J. (1988) *Immunol. Rev. 103*, 53.

King, S. J., & Miller, H. R. P. (1984) *Immunology 51*, 653.

Kokkonen, J. O., & Kovanen, P. T. (1989) *J. Biol. Chem. 264*, 10749.

Krieger, M., Kay, L. M., & Stroud, R. M. (1974) *J. Mol. Biol. 83*, 209.

Lazarus, S. C., DeVinney, R., McCabe, L. J., Finkbeiner, W. E., Elias, D. J., & Gold, W. M. (1986) *Am. J. Physiol. 251*, C935.

Le Trong, H., Parmelee, D. C., Walsh, K. A., Neurath, H., & Woodbury, R. G. (1987) *Biochemistry 26*, 6988.

Le Trong, H., Newlands, G. F. H., Miller, H. R. P., Charbonneau, H., Neurath, H., & Woodbury, R. G. (1989) *Biochemistry 28*, 391.

Loben C. G., Upton, C., Duggan, B., Ehrman, N., Letellier, M., Bell, J., McFadden, G., & Bleackley, R. C. (1988) *Biochemistry 27*, 6941.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning: A Laboratory Manual*, 1st ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Powers, J. C., Tanaka, T., Harper, J. W., Minematsu, Y., Barker, L., Lincoln, D., Crumley, K. V., Fraki, J. E., Schechter, N. M., Lazarus, G. S., Nakajima, K., Nakashino, K., Neurath, H., & Woodbury, R. G. (1985) *Biochemistry 24*, 2048.

Reilly, C. F., Tewksbury, D. A., Schechter, N. B., & Travis, J. (1982) *J. Biol. Chem. 257*, 8619.

Remington, S. J., Woodbury, R. G., Reynolds, R. A., Matthews, B. W., & Neurath, H. (1988) *Biochemistry 27*, 8097.

Salveson, G., Farley, D., Shuman, J., Przybyla, A., Reilly, C., & Travis, J. (1987) *Biochemistry 26*, 2289.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463.

Sayama, S., Iozzo, R. V., Lazarus, G. S., & Schechter, N. M. (1987) *J. Biol. Chem. 262*, 6808.

Schechter, N. M., Fraki, J. E., Geesin, J. C., & Lazarus, G. S. (1983) *J. Biol. Chem. 258*, 2973.

Schechter, N. M., Choi, J. K., Slavin, D. A., Deresienski, D. T., Sayama, S., Dong, G., Lavker, R. M., Proud, D., & Lazarus, G. S. (1986) *J. Immunol. 137*, 962.

Schechter, N. M., Slavin, D., Fetter, R. D., Lazarus, G. S., & Fraki, J. E. (1988) *Arch. Biochem. Biophys. 262*, 232.

Segal, D. M., Powers, J. C., Cohen, G.H., Davies, D. R., & Wilcox, P. E. (1971) *Biochemistry 10*, 3728.

Seppa, H., Vaananen, K., & Korhonen, L. K. (1979) *Acta Histochem. 64*, 64.

Sommerhoff, C. P., Caughey, G. H., Finkbeiner, W. E., Lazarus, S. C., Basbaum, C. B., & Nadel, J. A. (1989) *J. Immunol. 142*, 2450.

Tabor, S., & Richardson, C. C. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 4767.

Takahasi, H., Nukiwa, T., Yoshimura, K., Quick, C. D., States, D. J., Holmes, M. D., Whang-Peng, J., Knutsen, T., & Crystal, R. D. (1988) *J. Biol. Chem. 263*, 14739.

Tarentino, A., L., Gomez, C. M., & Plummer, T. H., Jr. (1985) *Biochemistry 24*, 4665.

Vanderslice, P., Craik, C. S., Nadel, J. A., & Caughey, G. H. (1989) *Biochemistry 28*, 4148.

Varsano, S., Basbaum, C. B., Forsberg, L. S., Borson, D. B., Caughey, G. H., & Nadel, J. A. (1987) *Exp. Lung Res. 13*, 157.

Vartio, T., Seppa, H., & Vaheri, A. (1981) *J. Biol. Chem. 256*, 471.

Von Heijne, G. (1986) *Nucleic Acids Res. 14*, 4683.

Wintroub, B. U., Schechter, N. B., Lazarus, G. S., Kaempfer, C. E., & Schwartz, L. B. (1984) *J. Invest. Dermatol. 83*, 336.

Wintroub, B. U., Kaempfer, C. E., Schechter, N. M., & Proud, D. (1986) *J. Clin. Invest. 77*, 196.

Woodbury, R. G., Katunuma, N., Kobayashi, K., Titani, K., & Neurath, H. (1978) *Biochemistry 17*, 811.

Yoshida, N., Everitt, M. T., Neurath, H., Woodbury, R. G., & Powers, J. C. (1980) *Biochemistry 19*, 5799.