Secretory Granule Proteases in Rat Mast Cells. Cloning of 10 Different Serine Proteases and a Carboxypeptidase A from Various Rat Mast Cell Populations

By Claudia Lützelschwab,* Gunnar Pejler,‡ Maria Aveskogh,* and Lars Hellman*

From the *Department of Medical Immunology and Microbiology, University of Uppsala, Biomedical Center, S-751 23 Uppsala, Sweden; and ‡Swedish University of Agricultural Sciences, Department of Veterinary Medical Chemistry, Biomedical Center, S-751 23 Uppsala, Sweden

Summary
Two of the major rat mast cell proteases, rat mast cell protease 1 (RMCP-1) and RMCP-2, have for many years served as important phenotypic markers for studies of various aspects of mast cell (MC) biology. However, except for these proteases only fragmentary information has been available on the structure and complexity of proteases expressed by different subpopulations of rat MCs. To address these questions, cDNA libraries were constructed from freshly isolated rat peritoneal MCs and from the rat mucosal MC line RBL-1. cDNA clones for 10 different serine proteases (RMCP-1-10), and the MC carboxypeptidase A were isolated and characterized. Six of these proteases have not been isolated previously. Based on their protease content, three separate subpopulations of MCs were identified. Connective tissue MCs (CTMCs) from the ear and peritoneum express the chymases RMCP-1 and -5, the tryptases RMCP-6, and -7 and the carboxypeptidase A. However, based on a large difference in the level of expression of RMCP-7, CTMCs of these two organs may be regarded as two separate subpopulations. RMCP-2 and the three closely related proteases of the RMCP-8 subfamily were identified as the major mucosal MC proteases in rat. In contrast to what has been reported for human MCs, no expression of cathepsin G or cathepsin G–like proteases was detected in any of the rat MC populations. To determine mRNA frequencies for the various proteases expressed by normal tissue MCs, an unamplified peritoneal MC cDNA library was screened with a panel of monospecific cDNA probes. These results showed that peritoneal MCs are highly specialized effector cells with mRNA frequencies for the major proteases in the range of several percent of the total mRNA pool.

1 Abbreviations used in this paper: 3H-DFP, (3H)diisopropyl fluorophosphate; aa, amino acid; BMMC, bone marrow-derived mast cell; CPA, carboxypeptidase A; CTMC, connective tissue mast cell; DFP, diisopropyl fluorophosphate; IgERI, the IgE high-affinity receptor; MC, mast cell; MMC, mucosal mast cell; MMCP, mouse mast cell protease; MTC, mastocytoma tumor cell; nt, nucleotide(s); pos., position; RMCP, rat mast cell protease.

13 J. Exp. Med. © The Rockefeller University Press • 0022-1007/97/01/13/17 $2.00
Volume 185, Number 1, January 6, 1997 13–29
Materials and Methods

Purification of MCs. Peritoneal cells from Sprague-Dawley rats (females, 10–15 wk old) were collected by peritoneal washing with 0.025 M Tris-HCl, pH 7.6, containing 0.12 M NaCl and 0.01 M EDTA. MC (connective tissue type) of ~95% purity (the majority of the contaminating cells were red blood cells), as judged by staining with toluidine blue, were prepared by density gradient centrifugation on metrizamide as described by Sterk and Ishizaka (26).

Cell Lines. The RBL-1 cell line was originally established from a rat with basophilic leukemia (27). This cell line, which shows many phenotypic characteristics of rat MCs, was maintained in F-DMEM supplemented with 5% fetal calf serum, 50 μg/ml gentamycin. The cultures were maintained in a 5% CO2 atmosphere at 37°C.

Labeling of MC Proteases with [3H]Disopropyl Fluorophosphosphate. MCs (~2 × 10^6 cells) were solubilized by adding 250 μl of PBS/1 M NaCl/0.5% Triton X-100. 70 μl of the solubilate was mixed with 400 μl of H2O/0.1% Triton X-100, followed by addition of 50 μl of a 0.17 mM solution of [3H]disopropyl fluorophosphosphate (3H-DFP) (6 Ci/mmol; New England Nuclear, Boston, MA). After 1 h incubation, the reaction was terminated by the addition of 250 μl of SDS-PAGE sample buffer. A sample (200 μl) of the incubation mixture was subjected to SDS-PAGE analysis on a 20-cm-long 10–18% gradient gel, according to the method of Laemmli (28). Electrophoresis was performed at 12 mA for ~18 h. After electrophoresis, the gel was stained with Coomassie brilliant blue and subsequently subjected to fluorography. NH2-terminal sequences were determined as previously described (29).

PCR. Degenerate PCR primers directed against conserved regions surrounding the histidine and serine residues characteristic for the catalytic triad of all trypsin-related serine proteases were synthesized. These primers were used in several PCR reactions under different conditions to isolate the central regions of hematopoietic serine proteases from the rat mucosal MC line RBL-1. The primer design and the conditions used for the PCR reactions will be described in a separate communication. In brief, a PCR reaction was performed using RBL-1 cDNA as template. The PCR product was loaded on a 1% low melting point agarose gel, and the DNA from a broad band of ~500 bp in size was isolated and cloned into a plasmid vector. After transformation the resulting colonies were screened on two identical filters with different probes. The filters were either hybridized with a RMCP-2–specific probe or with one of the PCR primers as probe. Clones showing positive hybridization signals with the PCR primer but were negative with the RMCP-2–specific probe were selected for further analysis. The sizes of the inserts were determined by restriction enzyme digestion followed by agarose gel separation, and clones with inserts of appropriate sizes (420–550 bp) were subjected to nucleotide sequence analysis. Inserts containing open reading frames and showing homologies in critical amino acid (aa) residues with other serine proteases were selected for further analysis.

Construction and Screening of a Phage cDNA Libraries. Cultured RBL-1 cells and freshly isolated, purified rat peritoneal MCs were used as starting material for mRNA purification followed by the construction of two separate cDNA libraries (Pharmacia Time Saver cDNA synthesis kit; Pharmacia, Uppsala, Sweden). The cDNA was ligated into the single EcoRI site of the λ-gt 10 vector and the ligated DNA was packaged into phage particles with an in vitro packaging system (Stratagene, La Jolla, CA). A library of ~50,000 recombinants was obtained for the RBL-1 cell line and a library of ~2.5 × 10^5 recombinants was obtained for the rat MCs.

The libraries were spread as a monolayer of the Escherichia coli C-600 Hfl strain with a titer of ~25,000 clones/138-mm plate and the plaques were subsequently transferred to Hybond N + filters (Amersham Int., Amersham, Buckinghamshire, England). The filters were screened with labeled fragments from the cloned PCR products for RMCP-3, -4, -8, -9 and 10 or with fragments from cDNA clones encoding MMCP-4, -5, -6, -7, and mouse MC CPA as probes (10, 13, 31). Filters were washed at low or medium stringency (6× SSC, 0.5% SDS or 2× SSC, 0.5% SDS). Autoradiography was performed for 24–48 h on Kodak Exomat AR film (Eastman-Kodak Company, Rochester, NY). Positive plaques were purified from each screening, phage DNA was prepared and the inserts were subcloned into the pGEM-2 vector. After detailed restriction endonuclease mapping of the cDNA clones, fragments of appropriate sizes were subcloned into the pGEM-blue vector and were sequenced by the dyeoxy chain-termination method (32). All nucleotide sequences presented in this study were established by sequencing both strands of the DNA insert.

Amino Acid Sequence Alignments. Nucleotide or aa sequences for a large panel of mature (NH2-terminally processed) hematopoietic serine proteases and some selected non-hematopoietic serine proteases were compared using the pairwise algorithm in the DNASTAR program (for Macintosh; DNASTAR, Ltd. London, UK). The original references for amino acid sequences are as follows: MMCP-1 (10), MMCP-2 (11), MMCP-4A (12), MMCP-4B (10), MMCP-L (12), MMCP-5 (10, 14), MMCP-6 (13, 31), MMCP-7 (15), RMCP-1 (21), RMCP-2 (22), a rat granzyme like serine protease isolated from rat duodenum which is identical to the RMCP-10 isolated by us (33), mouse granzyme A (34), human granzyme A (35), mouse granzyme B (36), mouse granzyme C (37), mouse granzyme D (38), mouse granzyme E (39), mouse granzyme F (40), human granzyme B (41), human granzyme G (42), rat NK protease 1 (43), the human chymase (44), the human tryptase 1 HMCP 1 (α-trypetase) (45), the human...
tryptase 2 HMCP 2 (β-trypate) (45a), human cathepsin G (46), human azurocidin (47), human protease 3 (myeloblastin) (48), human neutrophil elastase (49), human adipin (complement factor D) (49a), mouse adipin (50), human pancreatic trypsin I (51), human pancreatic trypsin II (51), and human pancreatic chymotrypsin (52).

Construction of Monospecific DNA Probes for the Different MC Proteases. Fragments of sizes from 165 to 220 bp from the 3′ terminal end of the cDNA clones encoding RMCP-1 (R-I), RMCP-2 (R-II), RMCP-3 (R-III), RMCP-4 (R-IV), RMCP-6 (R-VI), and RMCP-7 (R-VII) were excised and inserted into the plasmid vectors pGEM-blue or pGEM-2. The sequences of the inserts were confirmed. The inserts were excised from the plasmid vectors of the different clones by restriction enzyme cleavage. The purified fragments were labeled by random priming and were used as probes for the various phage library screenings and for the Northern blot analysis.

mRNA Frequency Analysis. The specific 3′ fragments for RMCP-1, -2, -3, -4, -6, -7, and the full-length cDNAs coding for RMCP-5, -8, R-CPA and the mouse hepatic core protein (53, 54) were used as probes to determine the mRNA frequencies for these proteins by screening of the unamplified rat peritoneal MC cDNA library. In addition, nearly full-length cDNA probes for the rat α chain of the high-affinity receptor for IgE (IgERI-α), the mouse γ chain of the IgERI, the rat Mac-2 cell surface marker (the IgE binding protein), mouse lysozyme, a partial cDNA clone for the 3′ region of the rat c-kit receptor, a cDNA covering the entire coding region of mouse αβ-microglobulin, and cDNA clones covering the entire coding regions of Mouse GATA-1 and mouse Pu.1 were used as probes in the screening of this cDNA library. These latter cDNAs have been cloned by PCR amplification or direct isolation from cDNA libraries from various mouse and rat hematopoietic cell lines. In addition to these cDNA probes we also used a few oligonucleotide probes, one directed against the 3′ terminal region of the rat IgERI β-chain, and one degenerate oligonucleotide directed against a conserved motif in the zinc finger region of members of the Krüppel-related zinc finger proteins. Positive signals from 4–8 filters were carefully counted and presented in Fig. 8 as number of positive signals in 100,000 plaque plaques. The size of the cDNA library is ~2.5 × 10^6 recombinants and the number of plaques screened with each probe was in the range of 100,000–200,000. Since the cDNA library was constructed with oligo-dT as primers the majority of all clones contain the 3′ ends of the mRNAs. A number of clones from each of the above described screenings and cloning of the various proteases have been analyzed and all clones were found to contain the 3′ end of the mRNAs. Several platings of the unamplified library were made with ~25,000 recombinants/138-mm plate. One additional 10-fold dilution was made and a few plates were plated at this density to determine the actual number of plaques in the screening. Only filters with the high density were used for the screenings, and the filters were used only up to a maximum of four consecutive rounds of hybridizations. The results were reconfirmed by at least two independent experiments.

Northern Blot Analysis. Total cellular RNA was prepared from cultured RBL-1 cells, several organs from Sprague Dawley rats (lung, intestine, spleen, liver, ears, and peritoneal MCs) and from ears of an additional panel of rat strains (Brown Norway, Hooded Lister, Dark Agouti, and Lewis). The RNA was isolated by the guanidinium thiocyanate method (55). Poly(A)^+ RNA was purified from total RNA using oligo dT-coupled magnetic beads following the procedure described by the manufacturer (PolyATract, mRNA Isolation system II; Promega, Madison, WI). RNA was separated on a 1.0% agarose gel containing 0.2 M formaldehyde, and was blotted onto Hybond N+ nylon membranes (Amersham Int.). Membranes were hybridized with the different probes in 6× SSC, 3× Denhardt’s solution, 0.5% SDS, 2 mM EDTA, and 100 μg yeast RNA/ml at 65°C and were washed four times at 65°C during 60 min under high stringency conditions (0.1× SSC, 0.1% SDS).

Genomic Southern Blot Analysis. Genomic DNA was prepared from Sprague Dawley rat liver as previously described (56). 20 μg of genomic DNA was digested to completion with the appropriate restriction enzymes and was fractionated through 0.6% agarose gels. The DNA was transferred to N+ Nylon membranes (Amersham) by blotting in 0.4 M NaOH. Blotting with a Vacuum-Blot unit (Pharmacia, Uppsala, Sweden) was completed in 3 h and the DNA was cross-linked to the membrane by UV-irradiation (GS Gene Linker; BioRad Laboratories, Hercules, CA). Hybridizations were carried out overnight in a solution containing 6× SSC, 5× Denhardt’s solution, 0.5% SDS, 2 mM EDTA, and 100 μg yeast RNA/ml at 65°C. Subsequently, the filters were washed four times at 65°C during 60 min under medium stringency conditions (2× SSC, 0.5% SDS). The filters were hybridized with a 32P-labeled probe consisting of a purified cDNA fragment from the RMCP-8 clone.

Results

Cellular extracts from peritoneal MCs were analyzed by SDS-PAGE (Fig. 1, lane A). A dominant band corresponding to an apparent molecular mass of ~29 kD was detected along with other major bands of ~32 and 17 kD. In addition, a ~38-kD band and several faint bands of ~28, ~31 and ~9–13 kD could be distinguished. Our previous studies have shown that the 29-kD protein gives an NH2-terminal sequence identical to that of mature RMCP-1 (57). Attempts were made to identify additional bands by NH2-terminal aa sequence analysis. Most of the bands yielded mixtures of several sequences, or the NH2–terminals appeared to be blocked. However, the NH2-terminal sequence: Asn-Phe-Tyr-Ser-Asn-Leu-His-Asp-Ile-Met-Leu was obtained for a 13-kD band. This sequence corresponds to an internal portion of RMCP-1 starting at Asn82 (numbering according to Le Trong et al., 1987[21]), indicating a cleavage at the Tyr81-Asn82 bond. Since cleavage after aromatic aa residues is a characteristic feature of MC chymases (29, 58); it appears likely that the cleavage is catalysed by either RMCP-1 (auto catalysis) or by RMCP-5, the other major chymase of rat CTMCs. A ~12-kD band gave the NH2-terminal sequence Ile-Ile-Gly-Gly-Val-Glu, indicating that it is an RMCP-1 fragment containing the intact NH2-terminal. When the MC extracts were radiolabeled with 3H-DFP, major incorporation of radioactivity was observed for the ~32- and ~29-kD bands, indicating that these bands correspond to serine proteases (Fig. 1, lane B). Minor incorporation of label was also observed for a 17-kD polypeptide as well as into polypeptides of ~15 and ~10 kD. Since the latter DFP-binding proteins appear to be too small to represent intact serine proteases it is likely that they correspond to various proteolytic fragments.
Cloning and Structural Analysis of Five Proteases from Rat Serosal MCs

mRNA was isolated from purified peritoneal MCs and used as starting material for the construction of a λ-gt10 cDNA library of a size of ~2.5 × 10⁶ individual recombinants. Fragments originating from cDNA clones encoding the major CTMC proteases in the mouse, MMCP-4, -5, -6, -7, and the mouse MC carboxypeptidase A were used as probes to isolate cDNA clones for the corresponding rat proteases. The isolated rat proteases were numbered according to their mouse counterparts: RMCP-5 (MMCP-5), RMCP-6 (MMCP-6), RMCP-7 (MMCP-7), and rat carboxypeptidase A. However, the nomenclature of the rat counterpart to MMCP-4, RMCP-1 cannot be changed due to historical reasons. The characteristics of the different cDNA clones encoding the five major CTMC proteases and some of the biochemical properties of the corresponding proteases are listed below. A summary of their biochemical characteristics is presented in Table 1.

RMCP-1. A cDNA clone encoding the entire coding region of the major rat MC chymase RMCP-1 was isolated (Fig. 2A, see also Fig. 5A). A comparison was made with the previously published partial sequence determined by PCR amplification (23) and no differences between the nucleotide sequences were found (23). RMCP-1 is translated as a 260-aa primary polypeptide with an 18-aa-long signal sequence, a 2-aa-long NH₂-terminal propeptide (Glu-Glu), and a 13-aa-long COOH-terminal propeptide.

Table 1. Physico-Chemical Characteristics of Various MC Proteases

<table>
<thead>
<tr>
<th></th>
<th>M, (Da)</th>
<th>No. aa</th>
<th>Net charge</th>
<th>No. of cysteines</th>
<th>Isoelectric point*</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMCP-1</td>
<td>25193</td>
<td>227</td>
<td>+19.1</td>
<td>6</td>
<td>9.7</td>
</tr>
<tr>
<td>MMCP-4</td>
<td>25269</td>
<td>227</td>
<td>+19.1</td>
<td>6</td>
<td>9.8</td>
</tr>
<tr>
<td>RMCP-2</td>
<td>25046</td>
<td>227</td>
<td>+5.2</td>
<td>6</td>
<td>8.4</td>
</tr>
<tr>
<td>MMCP-1</td>
<td>24958</td>
<td>226</td>
<td>+4.1</td>
<td>6</td>
<td>8.2</td>
</tr>
<tr>
<td>MMCP-2</td>
<td>24693</td>
<td>224</td>
<td>+7.7</td>
<td>6</td>
<td>8.8</td>
</tr>
<tr>
<td>RMCP-3</td>
<td>25443</td>
<td>226</td>
<td>+9.2</td>
<td>6</td>
<td>9.0</td>
</tr>
<tr>
<td>RMCP-4</td>
<td>24987</td>
<td>226</td>
<td>+12.1</td>
<td>6</td>
<td>9.5</td>
</tr>
<tr>
<td>RMCP-5</td>
<td>25392</td>
<td>226</td>
<td>+14.1</td>
<td>7</td>
<td>9.5</td>
</tr>
<tr>
<td>MMCP-5</td>
<td>25345</td>
<td>226</td>
<td>+13.2</td>
<td>7</td>
<td>9.4</td>
</tr>
<tr>
<td>H-Chymase</td>
<td>25032</td>
<td>226</td>
<td>+13.1</td>
<td>7</td>
<td>9.5</td>
</tr>
<tr>
<td>RMCP-6</td>
<td>27464</td>
<td>245</td>
<td>−4.6</td>
<td>8</td>
<td>6.1</td>
</tr>
<tr>
<td>MMCP-6</td>
<td>27440</td>
<td>245</td>
<td>−3.1</td>
<td>8</td>
<td>6.5</td>
</tr>
<tr>
<td>RMCP-7</td>
<td>27434</td>
<td>245</td>
<td>−4.6</td>
<td>8</td>
<td>6.1</td>
</tr>
<tr>
<td>MMCP-7</td>
<td>27413</td>
<td>245</td>
<td>−7.9</td>
<td>8</td>
<td>5.9</td>
</tr>
<tr>
<td>H-α Tryptase</td>
<td>27425</td>
<td>244</td>
<td>−2.6</td>
<td>8</td>
<td>6.4</td>
</tr>
<tr>
<td>H-β Tryptase</td>
<td>27446</td>
<td>245</td>
<td>−1.3</td>
<td>8</td>
<td>6.7</td>
</tr>
<tr>
<td>RMCP-8</td>
<td>25248</td>
<td>228</td>
<td>+10.6</td>
<td>9</td>
<td>9.0</td>
</tr>
<tr>
<td>RMCP-9</td>
<td>25249</td>
<td>228</td>
<td>+12.3</td>
<td>8</td>
<td>9.4</td>
</tr>
<tr>
<td>R-CPA</td>
<td>35701</td>
<td>308</td>
<td>+18.1</td>
<td>5</td>
<td>9.5</td>
</tr>
<tr>
<td>M-CPA</td>
<td>35822</td>
<td>308</td>
<td>+17.9</td>
<td>5</td>
<td>9.5</td>
</tr>
<tr>
<td>H-CPA</td>
<td>36057</td>
<td>308</td>
<td>+16.9</td>
<td>5</td>
<td>9.5</td>
</tr>
</tbody>
</table>

*pH 7.0
†Terminal processed protein.

Molecular masses (M), no. of amino acids (aa), net charges, no. of cysteines and isoelectric points of the NH₂-terminal processed rat, mouse, and human MC-specific neutral proteases.

The processed mature protease of 227 aa has a molecular mass of 25,193 (excluding potential carbohydrate additions), a net charge of +19.1 (Arg + Lys = 34; Asp + Glu = 16) and contains no potential N-glycosylation sites. (The nucleotide sequence is deposited in GenBank under the accession number U67915.)

RMCP-5. A nearly full-length cDNA clone for the rat chymase RMCP-5 was isolated (Figs. 2A and 5A). The clone contains an open reading frame of 740 bp encoding an 19-aa signal sequence, a 2-aa activation peptide (Gly-Glu) and the entire coding region of the mature protein, but lacks the first A of the initiation codon. The coding region is followed by a 3' untranslated region of 203 bp with the canonical AATAAA polyadenylation signal located at
position (pos.) 921. The mature protein of 226 aa has a molecular mass of 25,392 and has one putative N-glycosylation site (Asn-X-Ser/Thr) at position 58 of the mature protein. The processed proteolytically active RMCP-5 is a basic protein with a net charge of +14.1 (Arg + Lys = 29; Asp + Glu = 16). (The nucleotide sequence is deposited in GenBank under the accession number U67908.) The sequence differs in two positions compared with the previously published sequence of RMCP-5 (two additional nucleotides, T in pos. 810 and a G in pos. 811) (25).

RMCP-6. A cDNA clone containing the entire coding region for the rat tryptase RMCP-6 was isolated. This clone contains an open reading frame of 824 bp encoding a signal sequence of 19 aa, a 10-aa activation peptide (Ala-Pro-Cys-Pro-Val-Lys-Gln-Arg-Val-Gly) and the entire coding region of the mature protein (Figs. 3 and 5C). The coding region is followed by a 39 untranslated region of 246 bp with the non-canonical ATTAAA polyadenylation signal located at position 1061. The mature protein of 245 aa has a molecular mass of 27,464 and contains two putative N-glycosylation sites at positions 75 and 102 of the mature protein. The processed proteolytically active RMCP-6 is an acidic protein with a net charge of −4.6 (Arg + Lys = 19; Asp + Glu = 25). (The nucleotide sequence is deposited in GenBank under the accession number U67909.) The sequence differs in five positions from the previously published sequence of the rat tryptase (A pos. 399, C pos. 432, C pos. 748, T pos. 946, and one base pair deleted at position 1024) (24).
and the entire coding region for the mature protein. The coding region is followed by a 3' untranslated region of 184 bp with the canonical AATAAA polyadenylation signal located at position 1204 (Fig. 3). The mature protein of 245 aa has a molecular mass of 27,434 and two putative N-glycosylation sites located at positions 21 and 102. The processed proteolytically active RMCP-7 is an acidic protein with a net charge of -24.6 (Arg \(1\), Lys \(17\); Asp \(1\), Glu \(23\)). (The nucleotide sequence is deposited in GenBank under the accession number U67910.)

**Rat Mast Cell Carboxypeptidase A.** A nearly full-length cDNA clone for rat MC carboxypeptidase A (R-CPA) was isolated (Figs. 4 and 5). This clone contains an open reading frame of 1,254 bp encoding approximately two-thirds of the signal sequence, a 94-aa activation peptide and the entire coding region of the mature protein, but lacks the initiation codon and the NH2-terminal end of the signal sequence. The coding region is followed by a 3' untranslated region of 178 bp with the canonical AATAAA polyadenylation signal located at position 1399. The mature protein of 308 aa has an molecular mass of 35701 and two putative N-glycosylation sites at positions 127 and 133. The processed proteolytically active R-CPA is a basic protein with a net charge of +18.1 (Arg \(1\), Lys \(46\); Asp \(1\), Glu \(29\)). (The nucleotide sequence is deposited in GenBank under the accession number U67914.)

**Figure 4.** Nucleotide and amino acid sequences for MC carboxypeptidases. The figure shows a comparative analysis of the nucleotide sequences encoding rat (R-CPA), mouse (M-CPA), and human (H-CPA) MC carboxypeptidase A, and the deduced aa sequence for R-CPA. The cleavage sites for the signal sequence and the activation peptide are marked by arrows. The polyadenylation signal is underlined. Nucleotide numbers per line are depicted at the right of the figure. A schematic drawing of the different cDNA clones is shown at the top of the figure where the coding region for the mature protein is indicated by a grey box, and the signal sequence and the activation peptide are indicated by black and white boxes, respectively.
Cloning and Structural Analysis of Five Different Serine Proteases Expressed in the Rat Mucosal MC line RBL-1

Mucosal MCs are difficult to isolate as a pure population. Therefore, we decided to use the mucosal MC line RBL-1 as a source of mRNA for the isolation of rat mucosal MC proteases. The RBL-1 cell line has been widely used as an important in vitro model for studies of rodent MC biology and is considered to be a good representative of rat MMC (59). A RBL-1 cDNA library of ~50,000 independent recombinants was constructed. Only one rat MMC protease, RMCP-2 (the rat homologue to MMCP-1) has previously been identified. Except for MMCP-1, the only additional MMC protease isolated in mouse, MMCP-2, could not be used as probe because of strong cross-hybridization between the different MC chymases. We therefore decided to use a novel technique based on the use of degenerate PCR primers directed against conserved regions surrounding the histidine and serine residues of the catalytic triad characteristic for all proteolytically active members of the large multigene family of trypsin-related serine proteases.

Using this technique, five novel proteases, designated RMCP-3, -4, -8, -9, and -10 were isolated from the RBL-1 cell line. Based on a high degree of sequence identity with previously isolated rat MC chymases, RMCP-3 and -4 were classified as chymases. RMCP-8, -9, and 10 are all highly homologous and may be regarded as a new subfamily of serine proteases including RMCP-8, -9, and -10; (C) the mouse and rat tryptases: RMCP-6, MMCP-6, RMCP-7, and MMCP-7; and (D) three different mammalian MC carboxypeptidases: rat (R-CPA), mouse (M-CPA), and human (H-CPA) carboxypeptidase A. Arrows indicate the cleavage sites for the activation peptides. The grey arrow indicates the autocatalytic initial cleavage involved in the removal of the activation peptide (70). The remaining two aa of the activation peptide are removed by the dipeptidyl peptidase I, leading to the same result as does the single activation step of the MC chymases (70). The positions of the three aa of the catalytic triad are marked with asterisks. The number of aa per line is depicted at the right side of the figure.
ily of MC proteases, more closely related to the T cell granzymes than to the previously characterised MC proteases (see Fig. 6). The cleavage specificities of RMCP-8,-9, and -10 have not yet been determined. To obtain the complete sequences of these proteases, the ~500-bp insert fragments were used as probes in screenings of the RBL-1 cDNA library. Together with a full-length cDNA copy of the previously published RMCP-2, full-length or nearly full-length cDNA clones were isolated for RMCP-3, -4, -8, and -9. No full-length clone was obtained for RMCP-10 (see below). The nucleotide sequences of the RMCP-3 and -4 clones are depicted in Fig. 2 A, and the sequences of the clones encoding RMCP-8,-9, and -10 are shown in Fig. 2 B. The deduced aa sequences of these clones are depicted in Figs. 5, A and B. Fig. 6 shows a comparative analysis of the rat MC serine proteases with other rat, mouse, and human serine proteases. The characteristics of the clones, encoding the five different serine proteases isolated from the RBL-1 cDNA library, and some of the biochemical properties of these proteases are listed below. A summary of their biochemical characteristics is presented in Table 1.

RMCP-2. A cDNA clone containing the entire coding region for RMCP-2 was isolated (Figs. 2 A and 5 A). Only the 5'- and 3'-terminal 250 bp were subjected to nucleotide sequence analysis and were found to be identical to the previously published sequence for RMCP-2 (22). RMCP-2 is translated as a 247-aa primary polypeptide with an 18-aa-long signal sequence and a 2-aa long negatively charged NH2-terminal propeptide (Glu-Glu). The enzymatically active processed protease has a net charge of +5.2 (Arg + Lys = 25; Asp + Glu = 21) and a molecular mass of 25,046. RMCP-2 is also processed in its COOH-terminal end, where 3 aa are removed after translation (22). One putative N-glycosylation site (Asn-X-Ser/Thr) is located in the COOH-terminal end of the protein.
RMCP-3. A clone containing a partial sequence for RMCP-3 was isolated. This clone contains an open reading frame of 717 bp encoding half of the signal sequence, a 2-aa activation peptide (Glu-Glu) and the entire coding region of mature protein (Figs. 2 A and 5 A). The coding region is followed by a 3’ untranslated region of 172 bp with the canonical AATAAA polyadenylation signal located at position 868. The mature protein of 228 aa has a molecular mass of 25,443, and contains no putative N-glycosylation sites. The processed proteolytically active RMCP-3 is a basic protein with a net charge of +9.2 (Arg + Lys = 28; Asp + Glu = 20). (The nucleotide sequence is deposited in GenBank under the accession number U67888.)

RMCP-4. A cDNA clone containing the entire coding region for RMCP-4 was isolated. The clone has an insert size of 254 bp with an open reading frame of 741 bp flanked by 15-bp 5’ and 175-bp 3’ untranslated regions (Figs. 2 A and 5 A). A canonical polyadenylation signal (AATAAA) is located at position 938. The primary translated product has an 18-aa signal peptide and a 2-aa negatively charged activation peptide (Glu-Glu). The enzymatically active 226-aa protease has a molecular mass of 24,987 and no N-linked glycosylation sites. The mature basic protein has a net charge of +12.1 (Arg + Lys = 26; Asp + Glu = 15). (The nucleotide sequence is deposited in GenBank under the accession number U67907.)

RMCP-5. A clone containing the entire coding region for RMCP-5 was isolated (Figs. 2 A and 5 B). The 804-bp cDNA clone has an open reading frame of 717 bp, a 3’ untranslated region of 82 bp and contains a canonical polyadenylation signal at position 783. The unprocessed protein has an 18-aa signal peptide and a 2-aa activation peptide (Gly-Glu). The active basic enzyme has 228 aa with a molecular mass of 25,249, a net charge of +12.3 (Arg + Lys = 27; Asp + Glu = 15), and two putative glycosylation sites at positions 132 and 160. (The nucleotide sequence is deposited in GenBank under the accession number U67913.) Except for a single base pair difference, this sequence was found to be identical to a cDNA previously isolated from rat duodenum by Amerik et al. (33) (Granzyme-like protein II). This single base pair difference may have been caused by the low accuracy of the Taq polymerase used during the isolation of RMCP-10.

### Construction of Monospecific Probes

Several of the MC serine proteases show a high degree of sequence identity, making it difficult to use the entire inserts of the clones as probes in Northern blot, Southern blot, or plaque hybridizations. To solve this problem, the sequences for the different proteases were compared and regions with maximal divergence were selected as regions of interest for the construction of specific probes. Such regions were found in the 3’ non-coding regions of all the different mRNAs. Fragments located within these regions were excised from the cDNA clones and subcloned into plasmid vectors. The probes for RMCP-1, -2, -3, -4, -6, and -7 were all in the range of 165–220 nucleotides in length. For RMCP-5, -8, and CPA the entire inserts of the original cDNA clones were used as probes. To study the specificity of these probes, the insert fragments were purified, labeled and used in a dot blot analysis. All probes, ex-
cept for the RMCP-8 probe, were found to be monospecific (Fig. 7). Because of the high degree of homology shared by RMCP-8, -9, and -10 it has not been possible to excise any fragment of a size large enough to label by random priming, and still maintain mono-specificity (Figs. 2B and 7). Consequently, in all hybridizations performed with the RMCP-8 probe, we detect all three members of this subfamily of serine proteases.

**mRNA Frequency Determination**

Both CTMCs and MMCs store very large amounts of the various proteases in their secretory granules. Based on studies of the protein content in different MC subpopulations, it has been estimated that granule proteases constitute up to 50% of the total cellular protein (60). One important question in MC biology is whether the large amounts of stored proteases is also correlated with a high synthesis rate. One alternative possibility is that the turnover is minimal and that the synthesis rate thereby may be relatively low. In favour of this hypothesis, previous studies of mRNA levels for the various MC proteases in transformed MCs have shown that these levels actually are relatively low (Hellman, L., unpublished observations). However, these results may be questioned as being a phenomenon related to the transformed phenotype of the cell lines and not representative of mature tissue MCs. To address this question, we determined the mRNA frequencies for the various MC proteases, some housekeeping genes and a few cell surface receptor genes expressed by normal tissue MCs. For this purpose, an unamplified cDNA library was screened for positive hybridization signals with a panel of probes. All except one of these probes (the RMCP-8 probe) were monospecific and nearly all contained sequences originating from the 3' ends of the mRNAs (see Materials and Methods for details of the various probes).

The results showed that peritoneal MCs are highly specialized cells, which most likely constantly produce relatively high amounts of the various proteases. The mRNA frequencies of the most abundant proteases were in the range of 0.4–2.5% of the total number of plaques in the screenings (Fig. 8). Actin and GAPDH showed substantially lower frequencies, in the range of 0.03–0.4%. A panel of cell surface markers: the three subunits (α, β, and γ) of the IgE high-affinity receptor (IgERI), the Mac-2, and the c-kit receptor were all present at a level of 0.1% or below. The number of IgE receptors on CTMCs has previously been estimated at \( \sim 300,000/\text{cell} \) (26), a relatively low number compared to the number of stored protease molecules in the granules, which also is reflected by the low number of clones corresponding to the three IgE receptor subunits (0.1, 0.028, and 0.034% for the α, β, and γ chains, respectively). The very low level of c-kit was unexpected. However, growth factor receptors are known to be present at comparably low levels on the surface of most cell types (61). Further, the cDNA library was analysed for the presence of several additional hematopoietic granule proteins. A quite substantial number of clones for lysozyme were detected (235 in 100,000 plaques). 223 clones for the heparin proteoglycan core protein (53, 54), and 34 clones for a deacetylase (62) were detected in 100,000 plaques. Both of these proteins are involved in the synthesis of heparin proteoglycans. No positive signals were observed with cDNA probes against \( \alpha_2 \)-microglobulin (63), TNF-α, IL-4, or IL-5. The mRNA frequencies for a few transcription factors were also determined. We detected 17 and 10 clones in 100,000 plaques for Pu-1 and...
GATA-1, respectively. These values may be slight underestimates, since these probes do not contain sequences from the 3' non-coding regions of the mRNAs and clones shorter than 300–500 bp may thereby escape detection. We also screened the library with a 62-bp degenerate oligonucleotide against conserved regions of the finger and spacer regions of Krüppel-related zinc finger proteins (64). Only 31 positive signals were found in 100,000 plaques. We can from these data conclude that the transcription factors and cell surface markers tested are expressed at much lower levels compared to the most abundant granule components.

To obtain an estimate of what the number of plaques represents in actual mRNA frequencies, the sizes of the inserts in eleven independent clones were determined. 8 out of 11 of the clones had inserts of a size of 600 bp or larger whereas the remaining clones had inserts that were too small to be visualized on a 2% agarose gel (probably ~50 bp). Therefore, these latter clones probably contain only very short adapter fragments. Taking into consideration that ~70% of the clones contain mRNA sequences, this results in estimated mRNA frequencies ~30% higher than the actual number of positive signals in 100,000 plaques.
and consequently, mRNA frequencies for the most abundant proteases in the range of 0.5–3.5% of the total mRNA pool.

**Northern Blot Analysis**

To study the distribution of the different MC proteases in various rat tissues, mRNA was prepared from different organs of Sprague Dawley rats, from ears of several additional rat strains and from purified peritoneal MCs. The purified peritoneal MCs represent an essentially pure population of CTMCs. In the ear, the majority of the MCs present are of the connective tissue subtype, although it should be emphasized that the percentage of MCs in this tissue is relatively low. In mRNA isolated from both ears and from peritoneal MCs, Northern blot analysis showed high expression levels for the chymases RMCP-1 and -5, the tryptase RMCP-6 and the MC carboxypeptidase A (Fig. 9A). In mRNA isolated from the ears of several different strains of rats we detected relatively high levels of RMCP-7 (Fig. 9A). In contrast, peritoneal MCs expressed only low levels of this protease compared with the expression levels for RMCP-1, -5, -6, and CPA. These results are thus in agreement with the results from the mRNA frequency analysis described above (Fig. 8). With respect to their RMCP-7 content, CTMCs of ears and peritoneum may therefore be regarded as two distinct subpopulations of MCs. When comparing different strains of rats, most MC components showed similar patterns of expression. However, the amounts of mRNA for the various MC proteins were different among the various strains, indicating that the number of MCs may differ.

The mucosal MC protease RMCP-2 was undetectable in a Northern blot analysis of RNA from both ears and peritoneum (data not shown). Further, only two clones for RMCP-2 were detected in 100,000 clones of the serosal MC library (Fig. 8), indicating that MC populations in both the peritoneal cavity and the ears are essentially devoid of mucosal MCs. As detected by Northern blot analysis, ear and peritoneal MCs showed only low expression of RMCP-8 (Fig. 9). In addition, only 14 positive clones for RMCP-8 were detected in 100,000 clones of the serosal MC library (Fig. 8). This shows that the RMCP-8 subfamily of proteases is expressed only at a very low level in these two CTMC populations. No hybridization signals for RMCP-3 and -4 were detected in MCs of ear (Fig. 9B) or peritoneum (data not shown), and no positive plaques were found during the screening of the serosal MC library (Fig. 8). From these data we can conclude that the MC populations of these organs most likely completely lack expression of both of these proteases.

Cathepsin G or a cathepsin G–like protease has been detected previously both in human MCs and in a mouse MC tumour (65 and data not shown). To determine whether cathepsin G is expressed also by rat MCs, a cDNA fragment containing the entire coding region of mouse neutrophil cathepsin G was used as probe in a Northern blot analysis of mRNA from ear and peritoneal MCs. No hybridization signals were detected by Northern blot analysis (data not shown) and no positive plaques were found in the serosal MC library after screening under low stringency conditions (Fig. 8). This indicates that mature rat CTMCs completely lack the expression of cathepsin G or cathepsin G–like proteases. In addition, the RBL-1 cell line has previously been analyzed by Northern blots and found to be negative for mRNA encoding cathepsin G or a cathepsin G–like protease (data not shown).

mRNA isolated from the RBL-1 cell line was analyzed, together with mRNA from rat lung, small intestine, ears, and liver for expression of the various MC proteases (Fig. 9B). The RBL-1 cells expressed RMCP-2, -8, and low levels of RMCP-1, -3, and -4, but no RMCP-5 (Fig. 9B). The liver RNA was negative for all of these proteases. In ears we detected the previously identified CTMC proteases RMCP-1 and -5, but no hybridization signals for RMCP-2, -3, and -4. In mRNA isolated from small intestine we found relatively strong hybridization signals for RMCP-2 and the proteases of the RMCP-8 family. A low level of RMCP-1 and quite significant levels of RMCP-5 were also detected (Fig. 9B). In contrast, no visible bands were detected with the RMCP-3- and -4-specific probes (Fig. 9B). In addition to MMCs, the small intestine probably contains a small fraction of CTMCs originating from the sub-mucosa and the muscular layer of this organ. This may explain the detection of RMCP-1 and -5 expression in this tissue. However, the amount of mRNA for RMCP-5 seems too large to originate only from the CTMC population. One possibility, which we presently can not exclude is that rat MMCs also express RMCP-5. Lung is considered to have a mixture of both mucosal and connective tissue MCs. Accordingly, this organ showed expression of both MMC and CTMC proteases (Fig. 9B).

**Genomic Southern Blot Analysis of RMCP-8 and of Closely Related Genes in the Rat Genome**

As previously described, cDNA clones for three closely related proteases of the RMCP-8 subfamily (RMCP-8,-9, and -10) were isolated from the RBL-1 cDNA library. To estimate the total number of closely related members of this subfamily in the rat genome, a Southern blot analysis was performed with genomic DNA from Sprague Dawley and Wistar rats. The blots were hybridized using the full-length cDNA for RMCP-8 as a probe. The pattern of bands on the blots were found to be identical in both strains of rats (Fig. 9C, and data not shown). The results show multiple bands in the size range of 3–20 kb after cleavage with Bam HI, BglII, or EcoRI. The exact number is not possible to determine until information is available from genomic clones for one or several of these proteases. However, a rough estimate indicates three to five potential genes.

**Discussion**

We present in this communication the cloning of eleven different proteases from various rat MC populations. Five
of the proteases have been isolated previously as full-length or partial clones either by conventional cDNA cloning (RMCP-2, and RMCP-10 [Granzyme-like protease II]) or by PCR amplification (RMCP-1, -5, and -6) (22–25, 33). In two of these three proteases, RMCP-5 and -6, several differences in sequence were found between ours and the sequences determined by PCR amplification by Ide et al. (Figs. 2A and 3) (24, 25). We favor the explanation that most of these differences are allelic differences between different strains of rat, as the sequences originate from two different strains, Sprague-Dawley and Lewis (24, 25). However, we can not exclude that some of the differences are caused by the low accuracy of the Taq polymerase used during the PCR amplification. To minimize potential misunderstandings upon comparisons of mouse and rat MC proteases we have suggested a nomenclature which as close as possible harmonises between the corresponding proteases in these two species. This nomenclature has also been used by other investigators in the field (66).

The cloning of the various rat MC proteases has made it possible to construct monospecific cDNA probes for most of these enzymes. These probes have been used as specific tools to estimate mRNA frequencies for the different proteases in normal freshly isolated peritoneal MCs. The screening of the large unamplified cDNA library showed that serosal MCs are highly specialized cells which most likely continuously produce large amounts of the various secretory granule proteases, and with mRNA frequencies for these proteases in the range of several percent of the total mRNA pool. Such high expression levels have previously only been detected in terminally differentiated cells, and can be compared to, e.g., the expression of immunoglobulins in B lymphocyte plasma cells, where the immunoglobulin mRNA may account for 5–10% of the total mRNA pool.

The very high levels of mRNA for the different proteases actualises an important question concerning the role of MCs during normal tissue maintenance and tissue remodelling. If MCs store and only infrequently release the proteases (upon activation by allergen challenge or under inflammatory conditions) there would not be a need for a continuous protease production at such a high level, even if the amount stored in the cell is very high. Possible explanations for this phenomenon are either that MCs continuously release quite substantial amounts of their granule contents, or that a fraction of the proteases are never stored but are secreted without previous passage through the granule compartment. Alternatively, there may be a high rate of intracellular turnover of the proteases. Bone macrophages, osteoclasts, are known to actively participate in the degradation of non-cellular components and thereby help to maintain the elasticity of the bone (67, 68). Possibly, CTMCs have a similar function in the degradation of collagen, proteoglycans, fibronectin and other components of the connective tissue and thereby in the maintenance of the elasticity and function of this tissue.

We have also shown that there is a very tight regulation of the different proteases in the various MC populations. The typical CTMC proteases RMCP-1, -5, -6, and CPA were found at very high frequencies in the serosal MC library, whereas a low number of, or a total lack of clones for the MMC proteases RMCP-2, -3, -4, and the members of the RMCP-8 family were detected in the same library. In contrast, the MMC line RBL-1 and rat small intestine showed strong expression of the typical MMC protease RMCP-2 and the RMCP-8 subfamily, but only weak expression of the CTMC proteases. The separate subpopulations of MCs must thereby be considered as highly specialized cells where the protease genes are tightly regulated and expressed only in one out of several distinct subpopulations.

Mouse and rat MCs show a high degree of similarity in their protease expression. CTMCs of both species express two dominant chymases, RMCP-1/MMCP-4 and RMCP-5/MMCP-5, which are each others homologues. CTMCs of both species also express two tryptases, RMCP-6/MMCP-6 and RMCP-7/MMCP-7, and the MC carboxypeptidase A. In addition, in both species the expression of RMCP-7/MMCP-7 is considerably higher in skin CTMCs than in peritoneal CTMCs. When comparing MMCs of the two species, both similarities and also certain important differences are found. In both species the major MMC proteases are RMCP-2/MMCP-1 and, probably, the RMCP-8 subfamily. We have recently isolated the cDNA for a mouse MC protease, denoted MMCP-8, homologous to the RMCP-8 subfamily (data not shown). In the mouse genome, MMCP-8 is probably the only gene encoding serine proteases of this subfamily. In contrast, studies by genomic Southern blot analysis indicated three or possibly four or five members of the RMCP-8 subfamily in rat (Fig. 9C). Additional differences between MMCs derived from the two species include the fact that no direct homologue of MMCP-2 has yet been isolated in rat, and that no homologues of RMCP-3 and -4 have been isolated from mouse MCs (Fig. 7). One additional difference may also lie in the expression pattern for RMCP-5. We have detected substantial amounts of RMCP-5 mRNA in uninfected intestinal tissue, whereas mouse MMCs lack the expression of MMCP-5 (14). Moreover, Ide et al. observed a variable expression of RMCP-5 during cultivation of the RBL-2H3 cell line and they also detected a slight increase in the amounts of RMCP-5 during parasite infection of rat jejunum (25). In addition, the MC subcellulic specality of RMCP-3 and -4 is not clear. We find low but detectable levels in several different sublines of RBL-1, but no expression in normal uninfected intestine. Possibly, RMCP-3 and -4 may be upregulated during parasite infection, in a similar fashion as has been reported for another MC protease, MMCP-L (12).

Despite the present characterization of cDNAs for three new MC proteases, the RMCP-8 subfamily, none of these enzymes have been identified at the protein level. A possible explanation for this may lie in both extensive and variable glycosylation of these proteases. The members of the RMCP-8 subfamily of proteases have several potential gly-
cosylation sites (see Results). In a two-dimensional gel analysis of rat MMC proteases, Abe et al. (60) showed that freshly isolated MMCs from parasite infected rats gave a number of quite abundant but relatively fuzzy bands in the size range of 24–32 kD. Since none of these bands reacted with an RMCP-2-specific antiserum (60), it is possible that some of them represent differently glycosylated forms of proteases belonging to the RMCP-8 subfamily. This is not unlikely, in view of the fact that the majority of the MC-proteases are present in several differently glycosylated forms, and that these carbohydrates are exclusively N-linked (60, 66, 69).

The expression of cathepsin G or cathepsin G–like proteases in MCs from different mammals still remains an open question. In the present study, no expression of cathepsin G–like proteases was detected in any of the various rat MC populations tested. In contrast, we have previously detected significant levels of cathepsin G in mouse CTMC-like MC tumor cells (data not shown), and further, Schechter et al. have detected a cathepsin G–like protease in human MCs (65). We do not know whether these discrepancies reflect a species difference or if cathepsin G is expressed only in certain populations of immature MCs.

In summary, this investigation has led to a detailed picture of the various proteases and other differentiation markers expressed by different populations of rat MCs. The specific reagents originating from these studies can now be used as important tools in the analysis of the various biological functions associated with these proteases and as specific markers for studies of both early and late events during MC differentiation.

We are grateful to Dr. Anders Karlström for performing NH₂-terminal sequence analysis. We would also like to thank Dr. David Eaker for linguistic revision of the manuscript. This investigation was supported by grants from the Swedish Natural Sciences Research Council.

Address correspondence to Dr. Lars Hellman, Department of Medical Immunology and Microbiology, University of Uppsala, Biomedical Center, Box 582, S-751 23 Uppsala, Sweden.

Received for publication 4 September 1996 and in revised form 21 October 1996.

References


30. Deleted in proof.


