Molecular cloning of a cDNA that encodes the peptide core of a mouse mast cell secretory granule proteoglycan and comparison with the analogous rat and human cDNA

Shalom Avraham*, Richard L. Stevens, Christopher F. Nicodemus, Michael C. Gartner, K. Frank Austen, and John H. Weis

Department of Medicine, Harvard Medical School, and Department of Rheumatology and Immunology, Brigham and Women's Hospital, Boston, MA 02115

ABSTRACT   A cDNA that encodes a mouse secretory granule proteoglycan peptide core was isolated from a cDNA library prepared from nontransformed mouse bone marrow-derived mast cells (BMMC) using as a probe a 280-base-pair fragment of a rat cDNA that encodes the proteoglycan peptide core of rat basophilic leukemia (RBL)-1 cells. Based on the consensus nucleotide sequence and deduced amino acid sequence of the cDNA, the mouse BMMC proteoglycan peptide core is 16.7 kDa and contains a 21-amino acid glycosaminoglycan attachment region consisting of alternating serine and glycine residues. When the predicted amino acid sequence of the mouse BMMC proteoglycan peptide core was compared with the predicted amino acid sequences of the homologous molecules expressed in RBL-1 cells and in human promyelocytic leukemia HL-60 cells, the mouse-derived sequence was more closely homologous to the rat sequence than the human sequence except for the length of the serine-glycine repeat region. The N terminus was found to be a highly conserved region of the molecule in the three species, suggesting that this region is important for the structure, function, and/or metabolism of this family of proteoglycans. Nucleotide sequences within the 5' and 3' untranslated regions of the mouse, rat, and human proteoglycan cDNA were conserved. That similar sequences were also present in the corresponding regions of a cDNA that encodes a rat mast cell protease suggests that particular nucleotide sequences may be important for regulation of expression of those proteins that are destined to reside in secretory granules.

Mouse interleukin 3-dependent, bone marrow-derived mast cells (BMMC) synthesize 200- to 250-kDa highly acidic proteoglycans containing glycosaminoglycans that are almost exclusively chondroitin sulfate E (1). These intracellular proteoglycans are stored in secretory granules (2, 3) bound to basically charged endopeptidases (4) and exopeptidases (5) that are enzymatically active at neutral pH. Chondroitin sulfate E proteoglycans were extracted from BMMC by treatment with Zwittergent and guanidine hydrochloride and were purified to apparent homogeneity by density-gradient centrifugation, ion-exchange chromatography, and gel-filtration chromatography (6). As assessed by NaDodSO$_4$/PAGE of 48-hr [^{3}H]glycine-labeled cells, the peptide core of the mature proteoglycan stored in the secretory granule of mouse BMMC was estimated to be only $\approx$10 kDa (6). Analogous to the peptide core of rat mast cell heparin proteoglycan (7, 8), the peptide core of mouse BMMC chondroitin sulfate E proteoglycan is rich in serine and glycine and has a glycosaminoglycan attachment region resistant to degradation by a wide variety of proteases (1),

even if the glycosaminoglycans are first removed by treatment with chondroitinase ABC (6).

Bourdon and coworkers isolated two partial cDNAs from a rat L2 yolk sac tumor cell cDNA library, which taken together define an $\approx$1000-base-pair (bp) cDNA (9, 10) that encodes an 18.6-kDa proteoglycan peptide core (11). The deduced amino acid sequence of the consensus rat L2 cell-derived cDNA revealed an unusual proteoglycan peptide core containing a 49-amino acid region of alternating serine and glycine residues. We used a gene-specific $\approx$490-bp *Ssp* I 3' fragment of this rat L2 cell-derived cDNA to demonstrate that mouse BMMC, mouse myelomonocytic cells, rat basophilic leukemia (RBL)-1 cells (a transformed rat mucosal-like mast cell), rat serosal cavity-derived mast cells, and rat natural killer cells possess substantial amounts of an $\approx$1.0-kilobase (kb) mRNA that encodes a molecule similar to that expressed in rat L2 cells (12, 13). We then cloned an analogous cDNA, designated cDNA-R4, from RBL-1 cells (14). Using a 280-bp 5' *Bbv* I/*Xmn* I fragment of the RBL-1 cell-derived cDNA-R4, we have now isolated and sequenced a full-length homologous cDNA from a mouse BMMC-derived cDNA library.† When this mouse cDNA was compared with cDNAs that encode secretory granule rat and human proteoglycans and a rat protease, areas of homology were observed in the untranslated and translated regions.

MATERIALS AND METHODS

Construction and Screening of a Mouse BMMC cDNA Library. BMMC ($\approx$10^7 cells), obtained as described (15, 16), were lysed in the presence of 4 M guanidinium isothiocyanate, 0.5% (vol/vol) 2-mercaptoethanol, and 25 mM sodium citrate (pH 7.0), and total RNA was purified by the CsCl density-gradient centrifugation technique of Chirgwin *et al.* (17). The poly(A)$^+$ RNA, obtained by oligo(dT)-cellulose chromatography (18), was converted into cDNA (19). The resulting cDNAs were blunt-ended with T4 DNA polymerase, the internal *Eco*RI sites were methylated, and the cDNAs were ligated to *Eco*RI polylinkers. After selection of cDNAs of >500 bp by Sepharose CL-4B chromatography, the cDNAs were ligated and packaged into λgt11. *Escherichia coli* (strain Y1088) were infected with the recombinant bacteriophages resulting in a cDNA library with a complexity >1 × 10^6. The BMMC-derived cDNA library was probed at 43°C with a 280-bp 5' *Bbv* I/*Xmn* I fragment of the RBL-1 cell-derived cDNA-R4 (ref. 14; see Fig. 1) that had been radiolabeled with [α-^{32}P]dCTP ($\approx$3000 Ci/mmol; 1 Ci = 37 GBq; New England Nuclear) in hybridization buffer [50%

formamide/0.75 M NaCl/75 mM sodium citrate/2× Den-hardt's solution (1× Denhardt's solution = 0.02% bovine serum albumin/0.02% Ficoll/0.02% polyvinylpyrrolidone)/ 0.1% NaDodSO$_4$/1 mM EDTA/salmon sperm DNA carrier (100 μg/ml)/10 mM sodium phosphate]. The filters were washed at 55°C with 30 mM NaCl/3 mM sodium citrate/0.1% NaDodSO$_4$/1 mM EDTA/10 mM sodium phosphate, pH 7.0. Approximately 500,000 recombinants in the library were plated to isolate 18 clones. Three of the clones (designated cDNA-M5, cDNA-M6, and cDNA-M7), which had the larg-est inserts as assessed by agarose gel electrophoresis, were arbitrarily selected for nucleotide sequencing. The individual BMMC-derived cDNAs and their subcloned fragments were inserted into M13mp18 or M13mp19 and sequenced by the dideoxynucleotide chain-termination method of Sanger and Coulson (20). Both strands of cDNA-M6 were sequenced. The nucleotide sequence of cDNA-M6 was compared with the corresponding nucleotide sequences of the rat (9, 10, 14) and human (21) cDNA using the BESTFIT computer program (22).

RNA Analyses. Mouse heart, lung, liver, spleen, kidney, and brain (cerebrum) tissues were frozen in liquid N$_2$. After being pulverized at −80°C, the disrupted tissues were sus-pended in the guanidinium isothiocyanate buffer described above, and the RNA of each was purified. RNA was also isolated from mouse 3T3 fibroblasts (line CCL-92; ATCC), RBL-1 cells (line CRL-1378; ATCC), Kirsten sarcoma virus-immortalized mouse mast cell line KiSV-MCl (23), and mouse BMMC that had been cultured in 50% enriched medium/50% WEHI-3 conditioned medium in the absence (16) and in the presence (24) of mouse 3T3 fibroblasts. Total RNA from each tissue (5 μg per lane) or cell source (2 × 10^6 cell equivalents of RNA per lane) was denatured in formal-dehyde-formamide, electrophoresed in 1% formaldehyde-agarose gels, and transferred to Zetabind (Cuno) (25). The RNA blots were incubated at 43°C for 24 hr in hybridization buffer containing a radiolabeled 641-bp 5′ Acc I fragment or an ≈450-bp gene-specific Acc I 3′ fragment of cDNA-M6. The blots were washed under the above conditions of high stringency, and autoradiography was performed with Kodak XAR film.

RESULTS

Isolation and Nucleotide Sequence of a BMMC-Derived cDNA That Encodes a Mouse Proteoglycan Peptide Core. Three mouse cDNAs (designated cDNA-M5, cDNA-M6, and cDNA-M7) were obtained by screening the BMMC-derived cDNA library with the 280-bp 5′ Bbv I/Xmn I fragment (Fig. 1) of the RBL-1 cell probe cDNA-R4 under conditions of high stringency. After EcoRI digestion of the phage DNA and agarose electrophoresis of the resulting fragments, it was concluded that each mouse cDNA was ≈1.0 kb. The restric-tion map and the DNA fragments of these cDNAs that were sequenced are depicted in Fig. 1. The consensus nucleotide sequence and predicted amino acid sequence are depicted in Fig. 2. cDNA-M6 contained all of the coding information necessary for the peptide core of this mouse mast cell secretory granule proteoglycan. The initiation site for trans-lation is defined by an ATG codon that resides 5′ of a TGA stop codon in the same reading frame. Thirty-eight and >499 untranslated nucleotides were located 5′ and 3′, respectively, of the open reading frame of the nucleotides that encode this proteoglycan peptide core. Based on the position of the 3′ stop codon, the translated proteoglycan peptide core is 16.7 kDa and consists of 152 amino acids. As assessed by the hydropathicity plot (data not shown) and the deduced amino acid sequence (Fig. 2), the mouse proteoglycan peptide core contained a typical hydrophobic signal peptide of 25 amino acids. cDNA-M6 and cDNA-M7 possessed a 21-amino acid
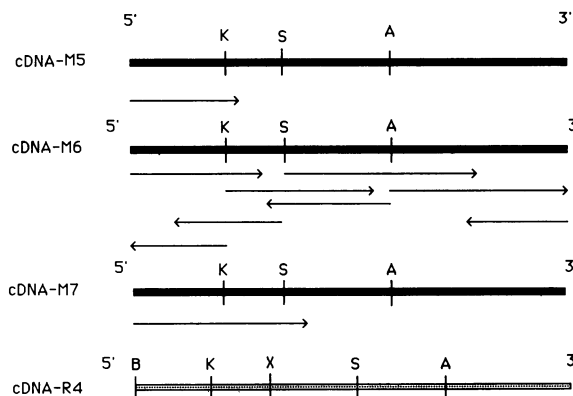

Fig. 1.   Restriction map and nucleotide sequencing strategy of the mouse BMMC-derived cDNA-M5, cDNA-M6, and cDNA-M7, and the RBL-1 cell-derived cDNA-R4. A, B, K, S, and X, sites within the cDNA that are susceptible to the restriction enzymes Acc I, Bbv I, Kpn I, Ssp I, and Xmn I, respectively.

sequence of alternating serine and glycine residues. cDNA-M5 also contained a similar-sized serine-glycine repeat se-quence based on the size of electrophoresed agarose gels of its Kpn I/Ssp I fragment [the region of the cDNA that encodes the serine-glycine repeat sequence (Figs. 1 and 2)] (data not shown). Using a search element of 6 nucleotides and allowing for no mismatches, the most conserved region of the 5′ untranslated region of the mouse (nucleotide residues 31–36; Fig. 2), rat, and human cDNA was the TGGTCA nucleotide sequence that preceded the ATG translation initiation codon for these proteoglycan peptide cores (Table 1). A number of conserved regions were detected in the 3′ untranslated region, two of which are also depicted in Table 1. These conserved 5′ and 3′ nucleotide sequences were similar to those present in the corresponding regions of the cDNA that encodes a rat mast cell protease (26) (Table 1).

A comparison of the deduced amino acid sequences of the mouse and rat proteoglycan peptide cores, as aligned by the BESTFIT computer program for comparing nucleotide se-quences, revealed that 96% of the first 53 N-terminal amino acids and 86% of the last 42 C-terminal amino acids were identical (Fig. 3). The principal difference between the 16.7-kDa mouse proteoglycan peptide core and the 18.6-kDa rat proteoglycan peptide core was the length of their respective serine-glycine repeat regions. Furthermore, as assessed by the BESTFIT computer program analysis of their nucleotide sequences, most of the deduced serine-glycine repeat region of the mouse molecule (amino acid residues 89–106) was more homologous to a region within the serine-glycine repeat region of the rat proteoglycan peptide core (amino acid residues 111–128) than to the beginning or end of this glycosaminoglycan attachment region. When the mouse BMMC proteoglycan peptide core was compared with the analogous 17.6-kDa molecule in the human, the N terminus was found to be a highly conserved region of the molecule in the two species. Seventy percent of the first 53 N-terminal amino acids and 45% of the last 42 C-terminal amino acids were identical. The serine-glycine glycosaminoglycan attach-ment region of the mouse BMMC proteoglycan peptide core was more similar in its length to the analogous proteoglycan peptide core in the human than in the rat. The conservation of the position of the two cysteine residues (at amino acid positions 38 and 47 of the BMMC-derived molecule) in all three species raises the possibility of a disulfide bond in this region of the molecule.

DNA and RNA Blot Analyses. Although four to six frag-ments were detected when a blot of digested mouse genomic DNA was probed under conditions of high stringency with the 5′ Acc I fragment of cDNA-M6, only one genomic

```
1                          TGCTAATCCAGAGGCTGAGTGGAGCCGAGCTGGTCAGG   ATG CAG GTT CCC
                                                                     M   Q   V   P    4

51  GTC GGC AGC AGG CTT GTC CTG GCT CTC GCC TTC GTC CTG GTT TGG GGA
    V   G   S   R   L   V   L   A   L   A   F   V   L   V   W   G    20

99  TCT TCA GTG CAA GGT TAT CCT GCT CGG AGA GCC AGG TAC CAG TGG GTC
    S   S   V   Q   G   Y   P   A   R   R   A   R   Y   Q   W   V    36
                    ↑
147 CGC TGC AAA CCG AAT GGC TTT TTT GCG AAC TGC ATC GAG GAG AAG GGA
    R   C   K   P   N   G   F   F   A   N   C   I   E   E   K   G    52

195 CCA CAG TTT GAC CTA ATA GAT GAA TCC AAT AAC ATC GGC CCT CCC ATG
    P   Q   F   D   L   I   D   E   S   N   N   I   G   P   P   M    68

243 AAT AAT CCT GTT TTG ATG GAA GGA CCC TCA AAA GAT TTC ATC TCC AAT
    N   N   P   V   L   M   E   G   P   S   K   D   F   I   S   N    84

291 TAT GAT GAC TAT GGG TCA GGT TCG GGC TCC GGC TCT GGC TCC GGC TCT
    Y   D   D   Y   G   S   G   S   G   S   G   S   G   S   G   S    100

339 GGC TCG GGT TCC GGC TCC GGA AGT GGC TTC CTA GGT GAC ATG GAA TGG
    G   S   G   S   G   S   G   S   G   F   L   G   D   M   E   W    116

387 GAA TAC CAG CCA ACA GAT GAA AGC AAT ATT GTC TAT TTC AAC TAT AAG
    E   Y   Q   P   T   D   E   S   N   I   V   Y   F   N   Y   K    132

435 CCT TTT GAC AGG ATT CTC ACT GAG CAA AAC CAA GAC CAA CCA GAA GAC
    P   F   D   R   I   L   T   E   Q   N   Q   D   Q   P   E   D    148

483 GAT TTT ATT ATA TGA  ATGTGACGGTCTCTGTCTCCCCACCTCCATGTGGAACAATGTATTCAGTATA
    D   F   I   I   ***                                               152

550 CTTAGTGTACCACGTTTAAATGACCAGTCTCAGGATAAAGAGTTTTACAGAAAATTTAAAATGCCTGGAAAAGACTCT
628 TGAATCCTGTTACCCCTTTCCTCATTAACTCGTAAGGAATTATGCTTTAATGCTGTTACCTATCTTGTTGTTCTGGAA
706 AATGCCTGCATTTATGTGTATTGAATCAACATTTAAGAAATTAACACACACCCCCATTATTATACAATAACTTTCAAA
784 GCCATACTGGTTTTGAAAATTTTAATTTGATAGCAAGTTGATGAACAATCTTTCATACCTAAAGTGTTCAGGAACCCA
862 ACTCGCATTGTGAATTACAAATATATTCCTTTATGTGATTAAAAAGAAAATAAAGTG(A)n>75
```
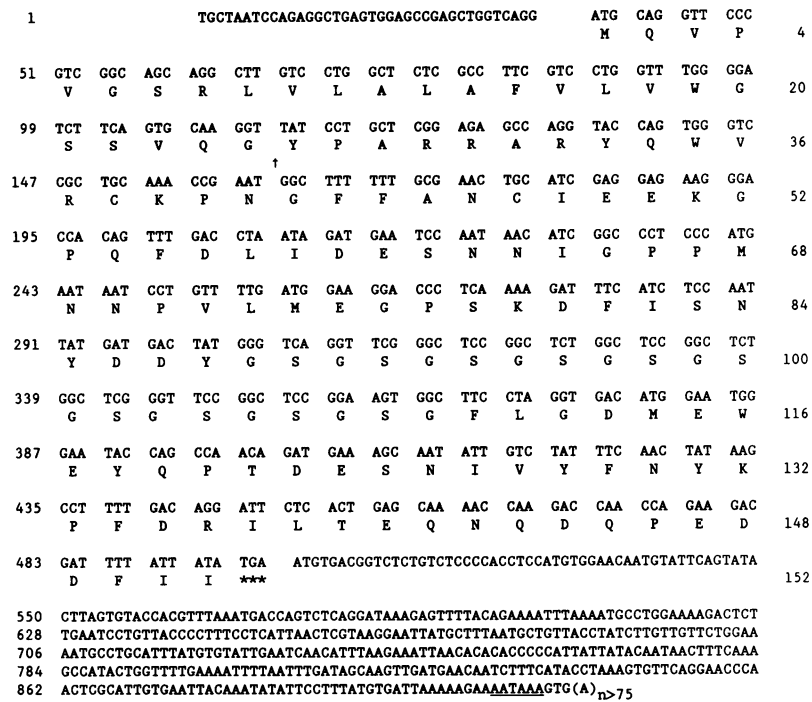
FIG. 2. Nucleotide and amino acid sequence of the BMMC cDNA-M6. ***, Stop codon. The polyadenylylation site is underlined. The numbers on the right and left indicate, respectively, the amino acid (single-letter code) and the nucleotide in the sequence. Arrow indicates the likely cleavage site of the signal peptide.

fragment was detected when the DNA blot was probed with the Acc I 3' fragment of the cDNA (data not shown).

When an RNA blot of total RNA from RBL-1 cells, KiSV-MC1, BMMC, and BMMC cultured in the presence of fibroblasts was probed with the gene-specific Acc I 3' fragment of DNA-M6 under conditions of high stringency (Fig. 4 Left), the mouse probe hybridized intensively in each instance to an ≈1.0-kb species of mRNA and to lesser amounts of two larger-sized mRNAs. In contrast, no hybridization of the probe to mouse 3T3 fibroblast-derived RNA was detected. When RNAs from different mouse tissues were examined, lung tissue had relatively high levels of the ≈1.0-kb transcript (Fig. 4 Right), and heart, spleen, and kidney had lower levels of the 1.0-kb transcript. Although the level of the 1.0-kb transcript was extremely low, small amounts of an ≈5.0-kb transcript were detected in the liver.

The level of the 1.0-kb transcript was extremely low in the cerebrum region of the brain, but it could be detected if the RNA blot was exposed longer of if poly(A)+ RNA was used (data not shown).

DISCUSSION

In the present study, we have screened a cDNA library prepared from mouse BMMC under conditions of high stringency to isolate a cDNA (Figs. 1 and 2) that encodes the peptide core of mouse mast cell chondroitin sulfate E proteoglycan. Although cDNAs that encode secretory granule-like proteoglycan peptide cores have been isolated from rat L2 tumor cells (9, 10), rat basophilic leukemia cells (14), and human promyelocytic leukemia cells (21), this is the first instance in which this type of cDNA has been isolated and

Table 1. Homologous areas within the 5' and 3' untranslated regions of the cDNA that encode three mast cell proteoglycans and a rat mast cell serine protease

	5' untranslated region	3' untranslated regions	
		Region 1	Region 2
BMMC-PG cDNA	TGGTCA(2 bp)ATG	TGA(19 bp)CCC-ACCT	(198 bp)TGTATTG(181 bp)AATAAA
RBL cell-PG cDNA	TGGTCA(2 bp)ATG	TGA(17 bp)CCC-ACCT	(197 bp)TGTATTG(126 bp)AATAAA
HL-60 cell-PG cDNA	TGGTCA(2 bp)ATG	TAA(11 bp)CCC-ACCT	(346 bp)TGGATTG(44 bp)AATAAT
RBL cell-protease cDNA	TGG-CA(2 bp)ATG	TGA(4 bp)CCTGACCT	(142 bp)TGGATTG(18 bp)AATAAA

The cDNA that encodes the mouse BMMC-derived proteoglycan (PG) peptide core was compared to the corresponding sequence in a rat RBL-1 cell-derived proteoglycan cDNA (14), a human HL-60 cell-derived proteoglycan cDNA (21), and a rat RBL-1 cell-derived serine protease cDNA (26). ATG, translation start codon; TAA and TGA, translation stop codons; AATAAA and AATAAT, polyadenylylation sites. The HL-60 cell-derived cDNA reported in our earlier publication (21) described a partial full-length cDNA that was missing part of its 3' untranslated region up to the poly(A)+. Thus, to compare the 3' untranslated region of this human cDNA with the mouse BMMC-derived cDNA, a human genomic fragment containing the HL-60 cell proteoglycan peptide core gene was isolated. The nucleotide sequence of a 3' HindIII/Sal I fragment was determined, and the 271-bp sequence contiguous to the previously published partial full-length cDNA is shown below. Based on the homology of this nucleotide sequence with the corresponding 3' untranslated regions in the rat (14) and mouse proteoglycan cDNA, it is likely that this represents the last 271 bp of the transcript.

```
GAATTCTTAA AGGATTATGC TTTAATGCTG TTATCTATCT TATTGTTCTT GAAAATACCT GCATTTTTTG
GTATCATGTT CAACCAACAT CATTATGAAA TTAATTAGAT TCCCATGGCC ATAAAATGGC TTTAAAGAAT
ATATATATAT TTTTAAAGTA GCTTGAGAAG CAAATTGGCA GGTAATATTT CATACCTAAA TTAAGACTCT
GACTTGGATT GTGAATTATA ATGATATGCC CCTTTTCTTA TAAAAACAAA AAAAAAATAA T
```
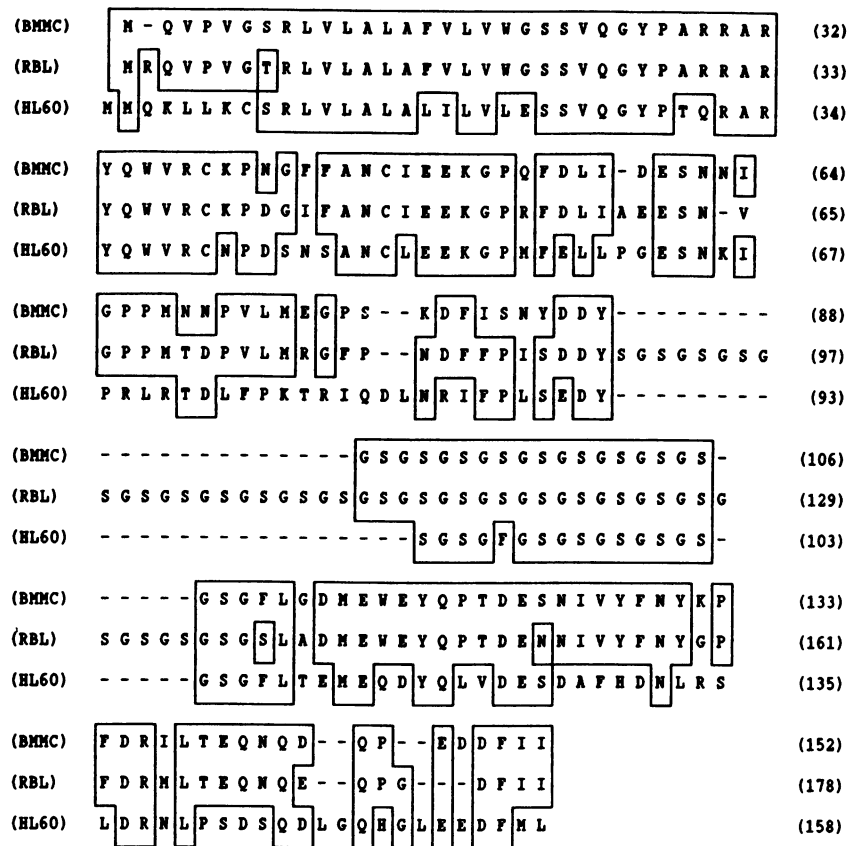
```
(BMMC)  M - Q V P V G S R L V L A L A F V L V W G S S V Q G Y P A R R A R   (32)
(RBL)   M R Q V P V G T R L V L A L A F V L V W G S S V Q G Y P A R R A R   (33)
(HL60)  M H Q K L L K C S R L V L A L A L I L V L E S S V Q G Y P T Q R A R (34)

(BMMC)  Y Q W V R C K P N G F F A N C I E E K G P Q F D L I - D E S N N I   (64)
(RBL)   Y Q W V R C K P D G I F A N C I E E K G P R F D L I A E E S N - V   (65)
(HL60)  Y Q W V R C N P D S N S A N C L E E K G P H F E L L P G E S N K I   (67)

(BMMC)  G P P N N N P V L N E G P S - - K D F I S N Y D D Y - - - - - - - - (88)
(RBL)   G P P N T D P V L N R G F P - - N D F F P I S D D Y S G S G S G S G (97)
(HL60)  P R L R T D L F P K T R I Q D L N R I F P L S E D Y - - - - - - - - (93)

(BMMC)  - - - - - - - - - - - - G S G S G S G S G S G S G S G S G S G S -  (106)
(RBL)   S G S G S G S G S G S G G S G S G S G S G S G S G S G S G S G S G   (129)
(HL60)  - - - - - - - - - - - - - - - S G S G F G S G S G S G S G S -      (103)

(BMMC)  - - - - - G S G F L G D N E W E Y Q P T D E S N I V Y F N Y K P     (133)
(RBL)   S G S G S G S G S L A D N E W E Y Q P T D E N N I V Y F N Y G P     (161)
(HL60)  - - - - - G S G F L T E N E Q D Y Q L V D E S D A F H D N L R S     (135)

(BMMC)  F D R I L T E Q N Q D - - Q P - - E D D F I I                       (152)
(RBL)   F D R N L T E Q N Q E - - Q P G - - - D F I I                       (178)
(HL60)  L D R N L P S D S Q D L G Q H G L E E D F N L                       (158)
```

FIG. 3. Comparison of the deduced amino acid sequences (single-letter code) of the proteoglycan peptide cores that are expressed in mouse BMMC, rat RBL-1 cells, and human HL-60 cells. Numbers indicate the amino acid in the respective sequences.
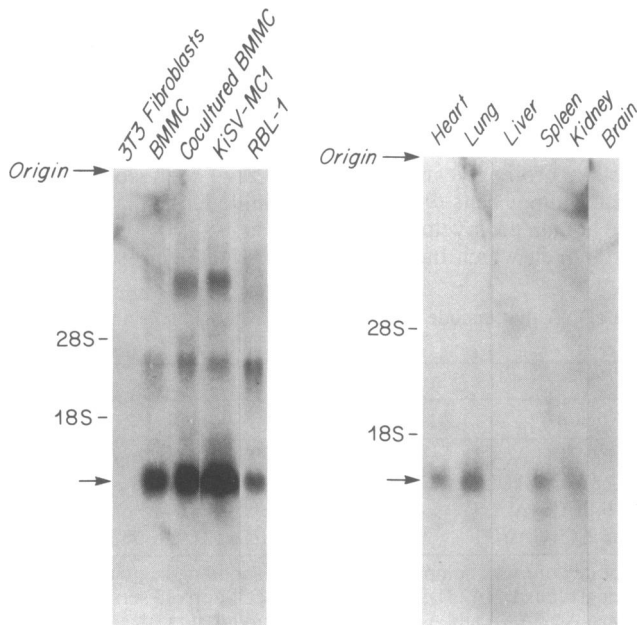
FIG. 4. RNA blot analysis of tissue-derived RNA and cell-derived RNA both probed with the Acc I 3' fragment of cDNA-M6 under conditions of high stringency. (Left) Total RNA was obtained from fibroblasts (lane 3T3 fibroblasts), BMMC cultured in the absence (lane BMMC) and presence (lane cocultured BMMC) of fibroblasts, KiSV-MCl cells (lane KiSV-MCl), and RBL-1 cells (lane RBL-1). (Right) Total RNA was obtained from mouse heart, lung, liver, spleen, kidney, and brain. The migration positions of the 18S (≈2 kb) and the 28S (≈5 kb) ribosomal RNA are indicated. Arrows indicate the 1.0-kb proteoglycan peptide core-specific mRNA.

characterized from a library derived from a nontransformed cell. It is also the first cDNA that has been isolated and characterized that encodes a secretory granule constituent of a mouse mast cell. As assessed by the nucleotide sequence, the 16.7-kDa proteoglycan peptide core expressed in mouse BMMC contains a 21-amino acid glycosaminoglycan attachment region consisting of alternating serine and glycine residues. Based on the -3, -1 rule (27) for cleavage of analogous hydrophobic signal peptides, it is predicted that the first 25 amino acids at the N terminus would be cleaved in the endoplasmic reticulum, resulting in a 14.3-kDa "pro" form of the proteoglycan peptide core. That the peptide core of the mature molecule stored in the secretory granule is only ≈10 kDa (6) indicates that further degradation of the 129-amino acid pro form of the peptide core occurs at its N terminus and/or at its C terminus.

Because secretory granule proteoglycan peptide core is the first constituent of the mast cell's secretory granule that has been characterized at the molecular level in three different species, it was possible to determine whether there are regions of the cDNA that are conserved. When the deduced amino acid sequence of the mouse proteoglycan peptide core was compared with those in the rat and human (Fig. 3), there were conserved sequences within the N-terminal portions of the proteins. A portion of these residues may be important in the targeting of the peptide core through the endoplasmic reticulum and Golgi into the secretory granules. The mouse and rat sequences also showed extensive homology within those residues C terminal to the serine-glycine repeat sequence. This homology was not maintained within the human sequence and may only reflect the closer evolutionary relatedness of the mouse and rat compared with the human.

The comparison of those sequences encoding the serine-glycine repeat region indicates that this region is ≈2 times longer in the rat than in the mouse or human. Since all three

Immunology: Avraham *et al.*

Proc. Natl. Acad. Sci. USA 86 (1989)   3767

predicted proteins are presumably functional within their respective secretory granules, it would appear that the size of the repeat sequence does not influence the biological role of the peptide core. Since there is only one gene in the rat, mouse, and human that encodes these peptide cores, it is likely that the rat gene was altered to encode a duplicated serine-glycine sequence. It is not known whether the duplication of this sequence in the rat occurred as a result of a duplication of a single exon or was due to an unequal crossover event.

One 5' untranslated region that almost immediately preceded the ATG translation-initiation codon and at least two 3' untranslated regions were nearly identical in the mouse (Fig. 2), rat (10, 14), and human (21) proteoglycan peptide core cDNA, and the cDNA that encodes a secretory granule serine protease of RBL-1 cells (26) (Table 1). The nucleotide sequence preceding the translation-initiation codon was not found in the cDNA that encoded the peptide core of a human fibroblast-derived dermatan sulfate proteoglycan (28), and this 5' nucleotide sequence was distinct from the proposed (29) consensus sequence for initiation of translation, which was based on a comparison of 699 vertebrate mRNAs. It is possible that the 5' nucleotide sequence simply represents a conserved linkage sequence with those nucleotides that encode the hydrophobic signal peptide of granule-associated proteins. Although it is not clear what the function of these conserved sequences might be, they may be important for the initiation of translation, for the control of transcription, and/or for the regulation of the turnover of those mRNAs that encode proteins that are destined to reside in secretory granules.

RNA blot analysis of total RNA isolated from different mouse tissues revealed that the lung had relatively high levels of the mRNA that encodes this proteoglycan peptide core (Fig. 4 *Right*). Heart, spleen, and kidney had less of the transcript; the lowest levels were in liver and brain (cerebrum). Because mast cells are present in large numbers in the lung, it is possible that most of these transcripts originate from this cell source. A 5.0-kb transcript that hybridized to the gene-specific probe was detected in the liver of the mouse, and it is possible that this larger-sized transcript in the liver is the result of alternatively spliced RNA. Human tissues have not been examined for the presence of this proteoglycan transcript, but the rat gene has been shown to be also expressed at relatively high levels in the lung of the rat (30). The mouse, however, appears to differ from the rat in that the larger-sized transcript is more prevalent in the rat kidney and rat lung than in the corresponding mouse tissue.

KiSV-MCl, BMMC cultured in the presence of fibroblasts, and BMMC cultured in the absence of fibroblasts are three populations of mouse mast cells that differ substantially in the histochemistry of their secretory granules when stained with alcian blue followed by safranin (23, 24, 31). These cells differ in their relative biosynthesis of ^{35}S-labeled chondroitin sulfate E proteoglycans and heparin proteoglycans from a ratio of 1:2 to 50:1. Despite the nature of the glycosaminoglycans bound to their proteoglycans, RNA blot analysis revealed that all three populations of mouse mast cells contained comparable levels of the $\approx$1.0-kb mouse proteoglycan peptide core mRNA and lesser amounts of two larger-sized transcripts (Fig. 4 *Left*). Because no larger-sized RNA was detected in any of the mouse tissue extracts except for barely detectable amounts in the liver, it is likely that these larger-sized RNAs in the mast cell preparations were incompletely spliced precursor molecules. The presence of comparable levels of the same mRNA in mast cells that synthesize

secretory granule proteoglycans differing in the nature of the glycosaminoglycan side chain bound to their respective peptide core suggests that mast cells can posttranslationally glycosylate this specific proteoglycan peptide core quite differently depending on the microenvironment in which the cells reside.

1. Razin, E., Stevens, R. L., Akiyama, F., Schmid, K. & Austen, K. F. (1982) *J. Biol. Chem.* **257,** 7229–7236.
2. Razin, E., Mencia-Huerta, J.-M., Stevens, R. L., Lewis, R. A., Liu, F.-T., Corey, E. J. & Austen, K. F. (1983) *J. Exp. Med.* **157,** 189–201.
3. Stevens, R. L., Razin, E., Austen, K. F., Hein, A., Caulfield, J. P., Seno, N., Schmid, K. & Akiyama, F. (1983) *J. Biol. Chem.* **258,** 5977–5984.
4. Serafin, W. E., Katz, H. R., Austen, K. F. & Stevens, R. L. (1986) *J. Biol. Chem.* **261,** 15017–15021.
5. Serafin, W. E., Dayton, E. T., Gravallese, P. M., Austen, K. F. & Stevens, R. L. (1987) *J. Immunol.* **139,** 3771–3776.
6. Stevens, R. L., Otsu, K. & Austen, K. F. (1985) *J. Biol. Chem.* **260,** 14194–14200.
7. Yurt, R. W., Leid, R. W., Jr., Austen, K. F. & Silbert, J. E. (1977) *J. Biol. Chem.* **252,** 518–521.
8. Robinson, H. C., Horner, A. A., Höök, M., Ogren, S. & Lindahl, U. (1978) *J. Biol. Chem.* **253,** 6687–6693.
9. Bourdon, M. A., Oldberg, A., Pierschbacher, M. & Ruoslahti, E. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1321–1325.
10. Bourdon, M. A., Shiga, M. & Ruoslahti, E. (1986) *J. Biol. Chem.* **261,** 12534–12537.
11. Oldberg, A., Hayman, E. G. & Ruoslahti, E. (1981) *J. Biol. Chem.* **256,** 10847–10852.
12. Tantravahi, R. V., Stevens, R. L., Austen, K. F. & Weis, J. H. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 9207–9210.
13. Stevens, R. L., Otsu, K., Weis, J. H., Tantravahi, R. V., Austen, K. F., Henkart, P. A., Galli, M. C. & Reynolds, C. W. (1987) *J. Immunol.* **139,** 863–868.
14. Avraham, S., Stevens, R. L., Gartner, M. C., Austen, K. F., Lalley, P. A. & Weis, J. H. (1988) *J. Biol. Chem.* **263,** 7292–7296.
15. Razin, E., Cordon-Cardo, C. & Good, R. A. (1981) *Proc. Natl. Acad. Sci. USA* **78,** 2559–2561.
16. Razin, E., Ihle, J. N., Seldin, D., Mencia-Huerta, J.-M., Katz, H. R., Leblanc, P. A., Hein, A., Caulfield, J. P., Austen, K. F. & Stevens, R. L. (1984) *J. Immunol.* **132,** 1479–1486.
17. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
18. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
19. Okayama, H. & Berg, P. (1982) *Mol. Cell. Biol.* **2,** 161–170.
20. Sanger, F. & Coulson, A. R. (1975) *J. Mol. Biol.* **94,** 441–448.
21. Stevens, R. L., Avraham, S., Gartner, M. C., Bruns, G. A. P., Austen, K. F. & Weis, J. H. (1988) *J. Biol. Chem.* **263,** 7287–7291.
22. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 387–395.
23. Reynolds, D. S., Serafin, W. E., Faller, D. V., Wall, D. A., Abbas, A. K., Dvorak, A. M., Austen, K. F. & Stevens, R. L. (1988) *J. Biol. Chem.* **263,** 12783–12791.
24. Levi-Schaffer, F., Austen, K. F., Gravallese, P. M. & Stevens, R. L. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 6485–6488.
25. Thomas, P. S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5201–5205.
26. Benfey, P. N., Yin, F. H. & Leder, P. (1987) *J. Biol. Chem.* **262,** 5377–5384.
27. von Heijne, G. (1984) *J. Mol. Biol.* **173,** 243–251.
28. Krusius, T. & Ruoslahti, E. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 7683–7687.
29. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8132.
30. Bourdon, M. A., Siga, M. & Ruoslahti, E. (1987) *Mol. Cell. Biol.* **7,** 33–40.
31. Dayton, E. T., Pharr, P., Ogawa, M., Serafin, W. E., Austen, K. F., Levi-Schaffer, F. & Stevens, R. L. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 569–572.