Cloning and Characterization of a Second Complementary DNA for Human Tryptase

Jeffrey S. Miller, George Moxley, and Lawrence B. Schwartz
Department of Medicine, Medical College of Virginia, Virginia Commonwealth University, Richmond, Virginia 23298

Abstract

A second cDNA for human tryptase, called β-tryptase, was cloned from a mast cell cDNA library in lambda ZAP. Its nucleotide sequence and corresponding amino acid sequence were determined and compared with those of a previously cloned tryptase cDNA, now called α-tryptase. The 1,142-base sequence of β-tryptase encodes a 30–amino acid leader sequence of 3,089 D and a 245–amino acid catalytic region of 27,458 D. The amino acid sequence of β-tryptase is 90% identical with that of α-tryptase, the first 20 amino acids of the catalytic portions being 100% identical. This identity, together with recognition of each recombinant protein by monoclonal antibodies directed against purified tryptase validate the tryptase identity of both α-tryptase and β-tryptase cDNA molecules. Modest differences between the nucleic acid sequences of α- and β-tryptase occurred throughout the cDNA molecules except in the 3′ noncoding regions, which were identical. Although most highly conserved regions of amino acid sequence in the trypsin superfamily are conserved in both tryptase molecules, β-tryptase has one carbohydrate binding site compared to two in α-tryptase, and one additional amino acid in the catalytic sequence. Regions of the substrate binding pocket in β-tryptase (DSCQ, residues 218–221; SWG, residues 243–245) differ slightly from those in α-tryptase (DSCK, residues 217–220; SWD, residues 242–244). The presence of both α- and β-tryptase sequences in each haploid genome was indicated by finding α- and β-tryptase specific fragments after amplification by PCR of genomic DNA in 10 unrelated individuals. Localization of both α- and β-tryptase sequences to human chromosome 16 was then performed by analysis of DNA preparations from 25 human/hamster somatic hybrids by PCR. It is now possible to assess the expression of each tryptase cDNA by mast cells and the relationship of each gene product to the active enzyme. (*J. Clin. Invest.* 1990. 86:864–870.) Key words: mast cell • protease • polymerase chain reaction • chromosome localization • trypsin superfamily

Introduction

Human tryptase is a tetrameric serine protease that is concentrated and stored selectively in the secretory granules of all types of human mast cells, from where it is secreted during mast cell degranulation (1–4). When detected in the circulation or in other biologic fluids, the tryptase level serves as an indication of mast cell activation (5, 6). Subunits of purified tryptase from lung (1, 7), skin (8), and pituitary (9) exhibit diffuse banding patterns in the 31,000–35,000-D range, often with two predominant bands ∼ 2,000-D equivalents apart from one another. The active form of human tryptase has four subunits, each having certain epitopes in common (10), identical NH_2-terminal 20–amino acid sequences (9, 11), and one active enzymatic site (1). On this basis, the size heterogeneity of the subunits was postulated to be due to posttranslational events rather than to different gene products, but the presence of more than one gene for tryptase was not excluded.

Complementary DNA molecules for tryptase were cloned and sequenced from human (11) and dog (12) mast cell cDNA libraries. The corresponding amino acid sequences show 71% identity and include a 30-amino acid leader and 245– and 244–amino acid catalytic portions for the dog and human sequences, respectively. In the study of dog tryptase, cDNA for a related trypsin-like enzyme, called dog protease, also was cloned and sequenced and found to have 53% identity with the dog tryptase amino acid sequence. The current work reports the finding and characterization of a second complementary DNA for human tryptase, called β-tryptase, which is highly homologous to the previously described cDNA sequence for human tryptase, now called α-tryptase. Furthermore, sequences for both α- and β-tryptase are typically present in the haploid genome of normal subjects and have been localized to chromosome 16.

Methods

Materials. The cDNA human mast cell library was constructed in lambda ZAP with purified polyadenylated RNA obtained from mast cells purified from human lung as previously described (11). DNA restriction endonucleases were obtained from Bethesda Research Laboratories (Gaithersburg, MD). [^{32}P]gamma dATP used for 5′ end labeling of synthetic oligonucleotides and [^{35}S]alpha dATP used for dideoxy nucleotide sequencing were obtained from New England Nuclear (Boston, MA). Sequenase polymerase was purchased from U. S. Biochemical Corp. (Cleveland, OH) and Taq polymerase with the polymerase chain reaction (PCR)[1] reagents were purchased from Perkin-Elmer Cetus Corp. (Norwalk, CT). PCR experiments were performed using a thermal cycler (Perkin-Elmer Cetus Corp.) provided by the Medical College of Virginia/Virginia Commonwealth University (MCV/VCU) Nucleic Acid Core Laboratory. Nitroplus 2000 nitrocellulose filters were obtained from Micron Separations, Inc. (Westboro, MA). NuSieve GTG and SeaKem GTG agarose were purchased from FMC Bioproducts (Rockland, ME), and PCRable DNA from human/hamster hybrid cell lines was purchased from Bios Corp. (New Haven, CT).

Oligonucleotides used as primers for DNA sequencing, PCR experiments, and as hybridization probes were synthesized on DNA syn-

1. *Abbreviations used in this paper:* PCR, polymerase chain reaction.

thesizer (Model 380A; Applied Biosystems, Inc., Foster City, CA) and purified by reverse phase HPLC on a Dynamax-300-Å 12-μm (10 mm i.d. × 25 cm l.) C18 reverse phase column (Rainin Instrument Co., Woburn, MA). Two Rabbit HP high-pressure pumps were used to generate an elution gradient from 7.25 to 50% acetonitrile in 0.1 M triethylammonium acetate. Oligonucleotide synthesis and HPLC purification were performed by the MCV/VCU Nucleic Acids Core Laboratory. Alternatively, deblocked oligonucleotides for PCR experiments were obtained from Oligos, Etc. (Ridgefield, CT) and used without further purification.

Screening and sequence analysis of tryptase cDNA molecules. Individual tryptase cDNA clones identified previously were detected using an immunoscreening technique with three antitryptase mAbs (11). In this study, additional clones were detected by screening with a radiolabeled 19-mer oligonucleotide probe corresponding to the β-tryptase sequence (TGCAACTGCGGGAGCAGCA). Duplicate plaque lifts were made from BB4 *Escherichia coli* host cells infected with an amplified portion of the lambda ZAP human mast cell cDNA library. Filters were prehybridized with 5× Denhardt's solution (0.1 g Ficoll 400, 0.1 g polyvinylpyrrolidone, 0.1 g BSA [fraction V] in 100 ml H_2O), 0.5% SDS. The 19-mer oligonucleotide was radiolabeled by the T4 polynucleotide kinase technique to a specific radioactivity of ~ 10^8 cpm/μg. The hybridization mixture contained 6× NET (0.9 M NaCl, 6 mM EDTA, 0.3 M Tris HCl, pH 8.0), 5× Denhardt's solution, 0.1% SDS, 250 μg/ml of heat-denatured yeast RNA, and oligonucleotide (6–8 ng/ml). An overnight hybridization at 52°C was followed by two washes at 37°C and one wash at 52°C with 6× NET/0.5% SDS. Autoradiography was then performed, and positive plaques were selected.

Plaque-derived phage of interest were purified by selecting single plaques containing material hybridizing to tryptase nucleic acid probes and used to infect *E. coli* (XL1 Blue) in the presence of R408 helper phage, which caused the autoexcision of the pBluescript plasmid containing the cDNA inserts of interest. Plasmid preparations of DNA were subjected to digestion with Eco RI to excise the cDNA insert and Eco RV to select inserts likely to contain the β-tryptase sequence. α-Tryptase cDNA has a known Eco RV restriction site at the position corresponding to nucleotides 368/369 of β-tryptase that is not present in β-tryptase cDNA. Each plasmid DNA containing an Eco RV-resistant insert was purified by a technique involving alkaline lysis and used directly for double stranded DNA sequencing with Sequenase (13). Plasmid DNA (10 μl) was diluted to 24 μl with dH_2O and denatured by adding 6 μl of 1 M NaOH and heating at 85°C for 5 min. Ice-cold 5 M ammonium acetate (10 μl) and 80 μl of 100% ethanol were sequentially added and the mixture was then chilled at −70°C for 20 min. The pellet obtained by centrifugation at 12,000 rpm in a centrifuge rotor (Model HB4; Beckman Instruments, Fullerton, CA) was washed in 70% ethanol three times and dried in vacuo. Annealing was performed by resuspending the pellet in 10 μl of sequencing buffer containing primer (10 ng), heating the mixture to 65°C, and then slowly cooling it to room temperature. A standard dideoxy nucleotide–sequencing protocol with Sequenase was then performed using both dGTP and dITP nucleotides to resolve any compression artefacts according to the instructions of the manufacturer (United States Biochemical Corp.). Each cDNA insert corresponding to tryptase was designated GRA-*n*, where *n* is an integer unique to each cloned insert.

PCR fragment evaluation of normal human genomic DNA. Template DNA used for PCR experiments included either α-tryptase or β-tryptase cDNA from plasmid minipreparations prepared as above, human genomic DNA prepared from peripheral leukocytes as described before (14) or buffer alone. Genomic DNA (500–1,000 ng) and a 10^7-fold dilution of plasmid DNA per reaction mixture with 500 ng of each primer were found to be optimal. Denaturation occurred at 95°C for 2 min, annealing at 62°C for 2 min, and extension at 72°C for 3 min, typically over 30 cycles. The optimal $MgCl_2$ concentration was 1.5 mM. Other conditions and commercial reagents were used according to the Perkin-Elmer Cetus instructions. PCR reaction products were concentrated ~ 10-fold on microconcentrators (Centricon 10; Amicon Corp. [Danvers, MA]), diluted back to the original volume

with dH_2O, and then reconcentrated 10-fold. REACT 2 buffer (10×; Bethesda Research Laboratories) was added to the concentrated material (1:10). A mixture consisting of 20 μl of this buffered concentrate and 20 U of Hae III was then incubated for 2 h at 37°C. Agarose gel electrophoresis of digested material was performed with 5% NuSieve GTG/0.5% SeaKem GTG agarose with Tris borate buffer at 100 V for 1–1.5 h. Southern transfers to Zeta Probe-cationized nylon membranes (Bio-Rad Laboratories, Richmond, CA) were performed with 0.4 N NaOH. Hybridizations with 13-mer radiolabeled oligonucleotides were performed at 42°C using the buffer described above except that Denhardt's solution was omitted because of interference with hybridization of small oligonucleotides. DNA sequence data were stored and analyzed using the Genetics Computer Group Sequence Analysis Software Package from the University of Wisconsin (version 6.1; August, 1989) (15).

Chromosome localization. Purified template DNA preparations for PCRs were obtained from 25 human/hamster somatic cell hybrid lines (100 ng/reaction), from normal human cells (50 ng/reaction) and from normal hamster cells (50 ng/reaction) as described by Bios Corp. Buffer controls without template DNA were performed in parallel to assess contamination. Primers were used at 0.2 μM, $MgCl_2$ at 1.5 mM, and other reagents as above in a 100-μl reaction volume under mineral oil. Denaturation was performed at 95°C for 1 min, annealing at 62°C for 1 min, and extension at 72°C for 2 min over 30 cycles. Agarose gel electrophoresis was performed on a medium-size apparatus with 5% NuSieve GTG/0.5% SeaKem GTG agarose with Tris borate buffer at 60 V for 5 h. Amplified material from human chromosome 16 was then subjected to Hae III endonuclease digestion, agarose gel electrophoresis and Southern blot analysis and compared with results obtained with human genomic DNA above for the presence of α-tryptase and β-tryptase sequences in the human genome.

Results

Isolation from a human lung mast cell library of a nucleotide sequence encoding β-tryptase. Of the 100,000 plaques previously analyzed (produced by phage obtained from a mast cell–derived cDNA lambda ZAP library) 10 plaques expressed protein recognized by each of three different murine MAbs directed against human tryptase (11). Seven of these contained inserts of various sizes, with nucleotide sequences identical to the previously reported tryptase cDNA sequence, now called α-tryptase cDNA. Three other clones contained inserts that are highly similar to α-tryptase cDNA, but with a slightly different nucleotide sequence, even though each coded for a protein with tryptase epitopes. The inserts in each of these three clones were identical over the regions sequenced and extended to different degrees toward the 5' and 3' termini. Only one β-tryptase cDNA clone, Gra-1, contained the entire 3' noncoding region and none reached further than nucleotide residue 233 toward the 5' terminus. To obtain the 5' end of the sequence, a 19-mer oligonucleotide complementary to a region toward the 5' end of GRA-1 (nucleotide residues 268–286) and having three nucleotide substitutions as well as three additional nucleotides compared with the corresponding region of α-tryptase was synthesized and used to screen ~ 100,000 new plaques produced with the lambda ZAP library. 10 plaques appeared to be positive and plasmids were prepared from these recombinant phage as described in Methods. The cDNA insert from each corresponding plasmid preparation was subjected to digestion with Eco RV, for which a single cleavage site exists in α-tryptase, but not in the corresponding region of β-tryptase. These digests were analyzed by agarose gel electrophoresis. Seven of the inserts were not digested by Eco RV and were subjected to DNA sequencing. The three Eco RV-sensitive

cDNA inserts were not further analyzed. Three of the seven Eco RV-resistant inserts contained the apparent methionine initiation codon for β-tryptase, but none appeared to contain the entire 3′ noncoding region. Complete sequence data were obtained with GRA-15 from the new cDNA isolates and GRA-1 from the previous cDNA isolates according to the strategy shown in Fig. 1. The nucleotide sequence corresponding to the entire coding and 3′ noncoding regions and a portion of the 5′ noncoding region was obtained and is called β-tryptase cDNA. Fig. 2 shows the entire β-tryptase cDNA and amino acid sequences and also the nucleotide residues of α-tryptase cDNA that differ. Limited sequence analyses performed on the eight other β-tryptase isolates were identical to the sequence illustrated.

The CCAGG sequence preceding the presumed initiator Met is identical in α and β tryptase and corresponds to optimal translation initiation sequences described previously (16, 17). The 1,142-base cDNA sequence is predicted to encode a 30–amino acid leader sequence of 3,089 D and a 245–amino acid catalytic portion of 27,458 D. The 20–amino acid NH₂-terminal sequence of the catalytic portion is identical to the amino acid sequences previously determined from purified tryptase and that predicted from α-tryptase cDNA. Thus, the shared amino acid sequence and the recognition by murine antitryptase MAbs reported previously (11) suggest the identity of this β-tryptase cDNA as an authentic tryptase cDNA.

The complete amino acid sequence of β-tryptase is 90% identical with that of α-tryptase. Highly conserved regions in the trypsin superfamily and in α-tryptase include the IVGG activation site region (residues 31–34), the WVLTAAHC active site histidine region (residues 68–75), the DIALL active-site aspartic acid region (residues 121–125), the M-CAG on the NH₂-terminal side of the active site serine (residues 209–213) and the GDSGGP active site serine region (residues 222–227). These regions also are conserved in β-tryptase. The aspartic acid at position 218 in β-tryptase, like that in α-tryptase,

confers the acknowledged trypsin-like substrate specificity of human mast cell tryptase (18). Putative cystine linkages in β-tryptase between Cys residues at positions 59 and 75, 155 and 230, 188 and 211, and 220 and 248 also correspond to the analogous residues in α-tryptase.

Although α-tryptase cDNA encodes two potential N-linked carbohydrate binding regions and a 274–amino acid protein, β-tryptase cDNA encodes only one potential N-linked carbohydrate binding region (NGT at positions 233–235) and a 275–amino acid protein. The additional amino acid in β-tryptase occurs between the second and third conserved regions (see below) of the serine protease family (QLREQ, residues 89–93 in β-tryptase compared with NSGT, residues 89–92, in α-tryptase). There are also differences in the predicted substrate binding pockets of α- and β-tryptase. The location of these residues is based on extensive homology with other serine proteases, but may need to be revised after x-ray crystallographic data on tryptase are obtained. On one side, DSCQ (residues 218–221) occurs in β-tryptase, DSCK (residues 217–220) in α-tryptase. Another side of the putative binding pocket contains SWG (residues 243–245) in β-tryptase, a typical sequence for serine proteases, as opposed to SWD (residues 242–244) in α-tryptase (11, 19). The Gln in β-tryptase is analogous to trypsin (20), the Lys in α-tryptase to plasma kallikrein (21), and Factor XI (22), each located near the entrance to the substrate binding pocket.

The predicted amino acid sequences for the catalytic portions of α- and β-tryptase molecules, respectively, permitted calculation of net charge values of −4 and −3 and pI values of 6.65 and 6.92 (15). The amino acid composition of β-tryptase is not substantially different from that of α-tryptase (11), both being compatible with previously determined amino acid compositions (1, 7).

Evaluation of the presence of nucleotide sequences coding for α- and β-tryptase in genomic DNA of normal subjects. To determine whether the putative genes for α-tryptase and β-
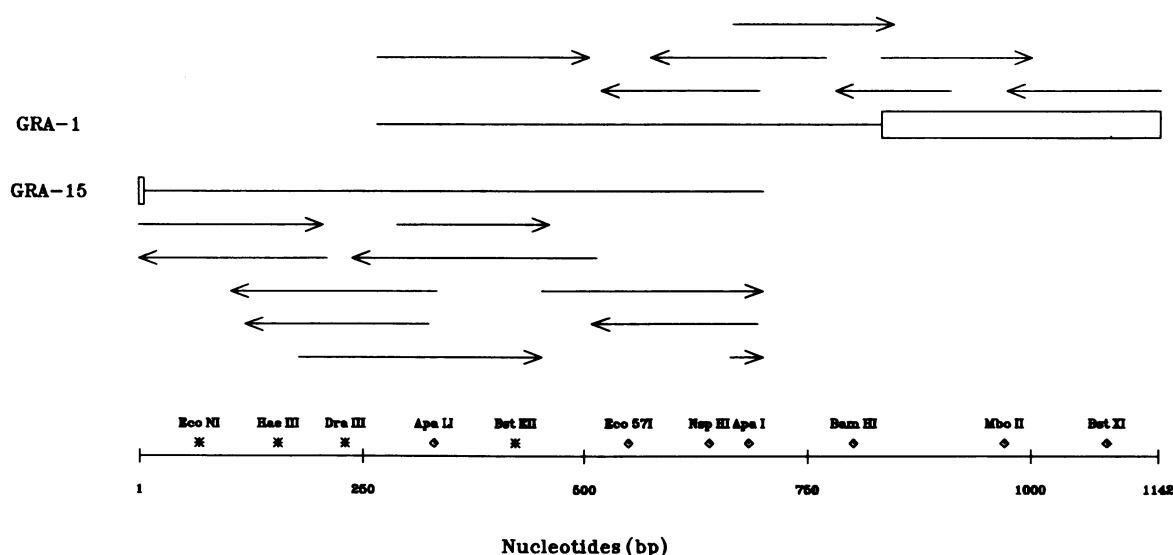
Figure 1. Sequencing strategy and restriction map of β-tryptase cDNA. Hexanucleotide restriction enzymes with one recognition site in β-tryptase cDNA are shown at their sites of cleavage. In addition, the tetranucleotide restriction enzyme Hae III is shown at the cleavage site of interest in the analysis of amplified PCR fragments (see Fig. 3). Restriction sites present in β-tryptase and not α-tryptase are shown with an asterisk, those present in both cDNA molecules are shown with a diamond. Boxed areas of Gra-1 and Gra-15 represent nontranslated regions (see Fig. 2).

```
      MetLeuAsnLeuLeuLeuLeuAlaLeuProValLeuAlaSerArgAlaTyrAlaA   19
CCAGGATGCTGAATCTGCTGCTGCTGGCGCTGCCCGTCCTGGCGAGCCGCGCCTACGCGG   60
               GC

laProAlaProGlyGlnAlaLeuGlnArgValGlyIleValGlyGlyGlnGluAlaProA   39
CCCCTGCCCCAGGCCAGGCCCTGCAGCGAGTGGGCATCGTTGGGGGTCAGGAGGCCCCCA  120
      T           A  C  T     C

rgSerLysTrpProTrpGlnValSerLeuArgValHisGlyProTyrTrpMetHisPheC   59
GGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCT  180
                    G  A  G
                                        ***
ysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAlaHisCysValGlyProA   79
GCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGG  240
              G                C

spValLysAspLeuAlaAlaLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnA   99
ACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGG  300
            A              C      ..CA

spGlnLeuLeuProValSerArgIleIleValHisProGlnPheTyrThrAlaGlnIleG  119
ACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCG  360
           A           G          T AT    CT
    ***
lyAlaAspIleAlaLeuLeuGluLeuGluGluProValLysValSerSerHisValHisT  139
GAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACA  420
        T                  C    CA       G

hrValThrLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThrG  159
CGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTG  480
         TG       G

lyTrpGlyAspValAspAsnAspGluArgLeuProProProPheProLeuLysGlnValL  179
GCTGGGGCGATGTGGACAATGATGAGCGGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGA  540
              C                 C

ysValProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrG  199
AGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGG  600

                                                      ----
lyAspAdpValArgIleValArgAspAspMetLeuCysAlaGlyAsnThrArgArgAspS  219
GAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACT  660
        A                              G  A
--------          ***                              ========
erCysGlnGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnA  239
CATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGG  720
  C  A        T                              A
          ----------
laGlyValValSerTrpGlyGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrA  259
CGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCC  780
        A

rgValThrTyrTyrLeuAspTrpIleHisHisTyrValProLysLysPro*          275
GTGTCACCTACTACTTGGACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAGTCAGGC  840

CTGGGTGTGCCACCTGGGTCACTGGAGGACCAACCCCTGCTGTCCAAAACACCACTGCTT  900

CCTACCCAGGTGGCGACTGCGCCCCCACACCTTCCCTGCCCCGTCCTGAGTGCCCCTTCCT  960

GTCCTAAGCCCCCTGCTCTCTTCTGAGCCCCTTCCCCTGTCCTGAGGACCCTTCCCCATC 1020

CTGAGCCCCCTTCCCTGTCCTAAGCCTGACGCCTGCACTGCTCCGGCCCTCCCCTGCCCA 1080

GGCAGCTGGTGGTGGGCGCTAATCCTCCTGAGTGCTGGACCTCATTAAAGTGCATGGAAA 1140

TCA_n
```

Figure 2. Nucleotide and amino acid sequence of β-tryptase cDNA. The amino acid sequence of β-tryptase is shown first with the number of the last residue shadowed. The entire nucleotide sequence of the composite β-tryptase cDNA is shown on the next line with the unshadowed number representing the last nucleotide on that line. Nucleotide sequence differences in α-tryptase cDNA are shown below the β-tryptase nucleotide sequence. The three nucleotides not present in α-tryptase are indicated with periods. Catalytic triad residues (*****), leader sequence (===), putative substrate binding pocket regions (-----), and the putative asparagine-linked carbohydrate binding site (===) are shown for the amino acid sequence as indicated, and the region from which the oligonucleotide sequence used for selecting LambdaZAP carrying 5' β-tryptase cDNA inserts is underlined. These sequence data are available from EMBL/GenBank/DDBJ under accession numbers M37488 and M30038.

tryptase are allelic or duplicated in each haploid genome, genomic DNA was obtained from 10 subjects and analyzed. An 86-bp nucleotide fragment from either α-tryptase or β-tryptase templates was amplified by PCR using 20-mer (β-tryptase nucleotides 135–154) and 19-mer (β-tryptase nucleotides 202–220) primers. The primers were complementary to identical regions in both α- and β-tryptase cDNA molecules, but the intervening region contained restriction site differences. The β-tryptase product contained a Hae III site; the α-tryptase product did not have this restriction site. Fig. 3 *A* shows that the PCR produced a nucleotide fragment of 86 apparent basepairs with each of the three templates, α-tryptase cDNA, β-tryptase cDNA and genomic DNA. Treatment of the ampli-

fied DNA fragments with Hae III (GG'CC) did not change the electrophoretic mobility of the 86-bp α-tryptase fragment. In contrast, complete cleavage of the 86-bp β-tryptase fragment into 20- and 66-bp portions occurred (Fig. 3 *B*). Treatment of the 86-bp fragments generated from the genomic DNA of each subject yielded fragments of 86, 66, and 20 bp, consistent with the presence of both α- and β-tryptase genetic material in the PCR products from each person. Identical results were obtained with DNA from an additional five unrelated subjects (not shown).

The possibility that the observed pattern of fragments (86, 66, and 20 bp) was due to partial cleavage of the β-tryptase PCR product from genomic DNA was considered. Two 13-
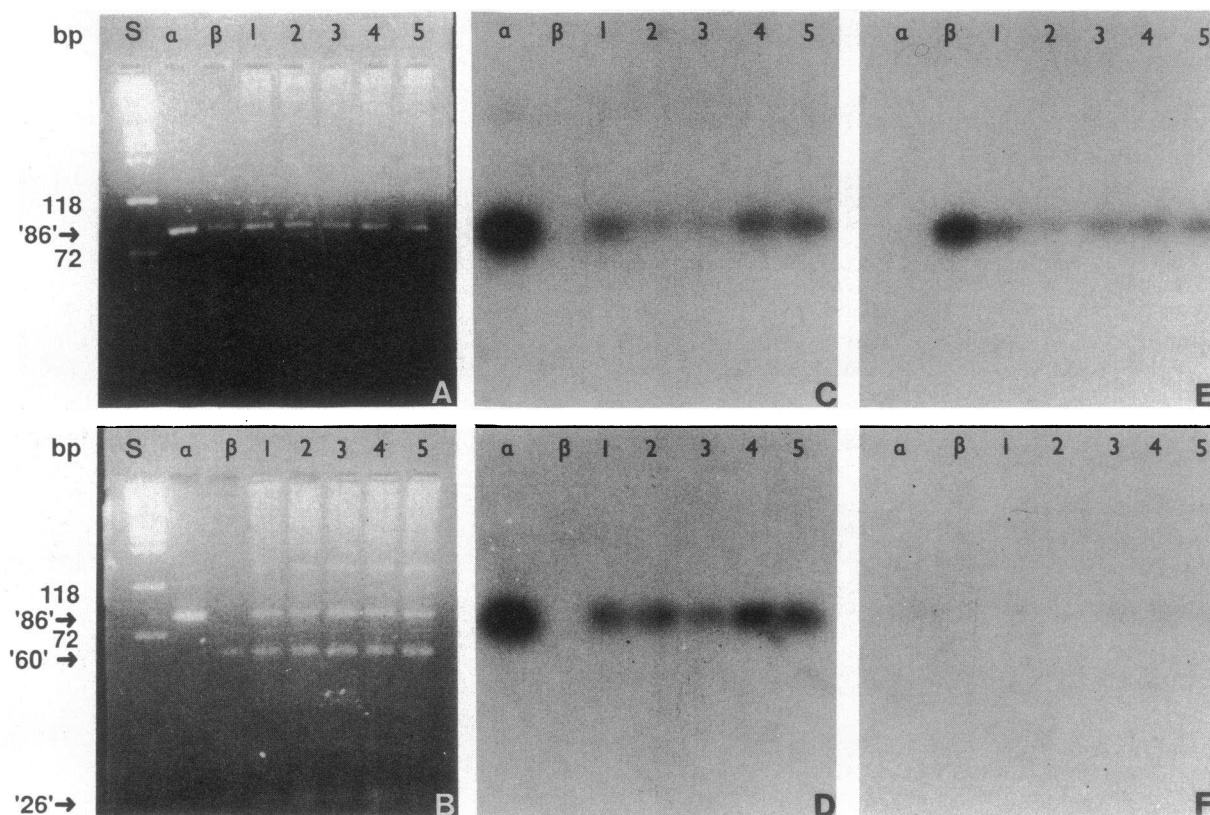
Figure 3. Southern blot analysis for the presence of sequences specific for α- and β-tryptase in the human genome. PCR reactions were performed with α-tryptase cDNA (lane α), β-tryptase cDNA (lane β) and genomic DNA obtained from five unrelated individuals (lanes *1–5*) as templates. Products obtained from the PCR reactions were subjected to electrophoresis in agarose gels either directly (*A*) or after treatment with Hae III (*B*), stained with ethidium bromide and photographed under UV light. Material in each of these gels was blotted and baked onto nitrocellulose paper. The α-tryptase specific 13-mer probe was labeled and hybridized to Hae III-untreated (*C*) and -treated (*D*) material; the blot was then subjected to autoradiography. Radiolabeled hybridized α-tryptase probe was removed by boiling; the filter was then subjected to autoradiography to insure that radioactivity had been removed. Rehybridization with the 13-mer β-tryptase probe was performed with Hae III-untreated (*E*) and -treated (*F*) material. Molecular weight standards are shown in the *S* lane. Arrows show fragment positions with apparent molecular weights of 86, 60, and 26 bp.

mer oligonucleotide probes corresponding to residues 156 to 168 in α- and β-tryptase cDNA, but distinguished from each other by three nucleotide mismatches, were synthesized and labeled with ^{32}P. These probes were then used under conditions of high stringency to identify α- and β-tryptase-specific sequences. The α-tryptase probe recognized the 86-bp amplification product from α, but not from β tryptase cDNA (Fig. 3 *C*); the β-tryptase probe recognized only the 86-bp fragment derived from β-tryptase cDNA (Fig. 3 *E*). Both probes recognized the 86-bp fragment from genomic DNA before Hae III treatment (Figs. 3, *C* and *E*). After Hae III treatment (Figs. 3, *D* and *F*), a positive signal appeared only with the α-tryptase probe, which hybridized to the 86-bp fragment. Hae III cut the β-tryptase PCR fragment within the sequence corresponding to the 13-mer probe, leaving complementary ends of 4 and 9 bp. Hybridization of these fragments with the 13-mer β-tryptase probe does not occur under the conditions used. Thus, PCR products from both α- and β-tryptase DNA were present in the DNA of 10 of 10 persons. The putative genes for α-tryptase and β-tryptase, therefore, are not allelic, but are each present in the normal haploid genome. Although quantitation of genomic sequences based on analysis of PCR products is somewhat uncertain, both sequences appear to be present

there in equal amounts based on the relative quantities visualized of the amplified 86-bp fragments.

Chromosome localization of the nucleotide sequences for α and β tryptase. DNA preparations from the 25 human/hamster somatic hybrids in addition to control human and hamster DNA were used as PCR templates as described in Methods. Only the human DNA control and two of the somatic hybrid-derived DNA preparations yielded a reaction product after the initial 30-cycle amplification reaction, all comigrating at an apparent size of 86 bp. One hybrid had been characterized as having chromosomes 5, 8, and 16, the other had chromosomes 3, 5, 10, 15, and 16. Only chromosome 16 was unique to this pair of heterohybrids and not found in others, indicating that the putative gene(s) for tryptase are on human chromosome 16.

To assess whether the nucleotide sequences for both α and β tryptase were present on human chromosome 16, the PCR product(s) from each of the two positive hybrid lines was treated with Hae III and subjected as above to Southern blot analysis under stringent conditions with the 13-mer probes specific for α- and β-tryptase cDNA. Both probes hybridized to the 86-bp DNA fragments amplified from each of the two hybrid lines. Hae III only cleaved about half of the 86-bp

fragment, indicating that the putative genes for both α-tryptase and β-tryptase are located on human chromosome 16.

Discussion

Two distinct cDNA molecules for human tryptase have now been cloned from a mast cell cDNA library; the one reported previously, now called α-tryptase cDNA, and the one reported currently called β-tryptase cDNA. Verification of each as representing tryptase was based upon complete identity between the NH₂-terminal 20–amino acid sequence obtained with purified authentic tryptase and those obtained from the corresponding region of the cloned cDNA molecules. Recognition of the expressed recombinant protein derived from α- as well as β-tryptase cDNA by three different monoclonal antibodies directed against purified tryptase (11), further validates the tryptase identities of the cloned cDNAs.

Comparison of the corresponding nucleotide sequences for α- and β-tryptase cDNA (Fig. 2) showed 96% identity. Differences occurred throughout the coding region. The 3' nontranslated region of the single β-tryptase cDNA cloned with this region intact and the five nucleotides 5' to the start methionine codon for β-tryptase were identical to the 3' nontranslated region of five different α-tryptase cDNA molecules and the 5' nontranslated region of the single α-tryptase cDNA cloned with this region intact. Conservation of the noncoding regions is unusual among different serine proteases. Perhaps these identical 3' noncoding regions reflect a conserved regulatory function involved with the expression of each tryptase gene or the stability of the corresponding mRNA molecules. Another possibility is that only one copy of this region resides in the genome, shared by the two coding regions, perhaps by an alternative splicing mechanism. Whether the genes for α- and β-tryptase are allelic or both present in the haploid genome can be addressed with the present data. The nucleotide differences observed throughout the coding regions of the α- and β-tryptase cDNAs and the observation that no homozygosity for either α- or β-tryptase PCR products was observed in 10 of 10 unrelated individuals indicate the presence of a distinct gene for each form of tryptase.

The finding that mRNA for both α-tryptase and β-tryptase must have been present for both to be cloned from the cDNA library suggests that mast cells also synthesize the corresponding protein products. However, the capacity of each product to be incorporated into the functional tetrameric enzyme, the relative molar ratios of α- and β-tryptase proteins possible in a functional enzyme and the relative expression of each putative tryptase gene by different types of mast cells are not known. In addition, this study does not rule out the possibilities of polymorphic variations within the tryptase genes or the presence of additional genes for tryptase. At present, the terms α- and β-tryptase should be restricted to refer to the protein products of the corresponding cDNA molecules, whereas tryptase refers to the authentic enzyme.

Whether the expression in vivo of α- and β-tryptase account for the size heterogeneity observed among subunits of authentic tryptase by SDS PAGE is unknown. The calculated sizes of the corresponding mature proteins, 27,423 D for α-tryptase, and 27,458 D for β-tryptase, do not account for the difference of ∼ 2,000 D actually observed between the two dominant protein bands. Carbohydrate differences (two puta-

tive glycosylation sites on α-, one on β-tryptase) may account for much of the size heterogeneity. However, endoglycosidase treatment of lung-derived human tryptase had no effect on the electrophoretic mobility of the subunits by SDS PAGE, whereas treatment of pituitary-derived human tryptase converted subunits of 36,300 and 34,600 D to subunits of 33,400 and 32,400 D (9). Thus, the relative contributions of post-translational events and primary sequence differences to the size heterogeneity of authentic tryptase remain problematic.

The marked similarity in predicted amino acid sequence among the catalytic regions of human α-tryptase (73%) and human β-tryptase (78%) with that of dog tryptase, though not as high as between the two human products (91%), clearly show extensive sequence conservation within this family (Fig. 4). Dog tryptase and human β-tryptase have the identical QLREQ sequence (residues 89–93) and one putative carbohydrate binding site. By comparison, human α-tryptase has NSGT (residues 89–92) and two carbohydrate binding sites. Dog tryptase and human α-tryptase share a carbohydrate region that is lacking in human β-tryptase. In addition, dog tryptase and human β-tryptase each have SWG (residues 243–245) rather than SWD (residues 242–244) residing in the predicted substrate binding pocket (19). They also have a Gln (residue 221) rather than a Lys residue covering the entrance into the substrate binding pocket. Comparisons of the 30–amino acid leader sequence for each of these three tryptase molecules also shows extensive homology, but somewhat less than for the catalytic portions. Leader sequences for α- and β-tryptase show 87% identity, whereas these sequences, respec-
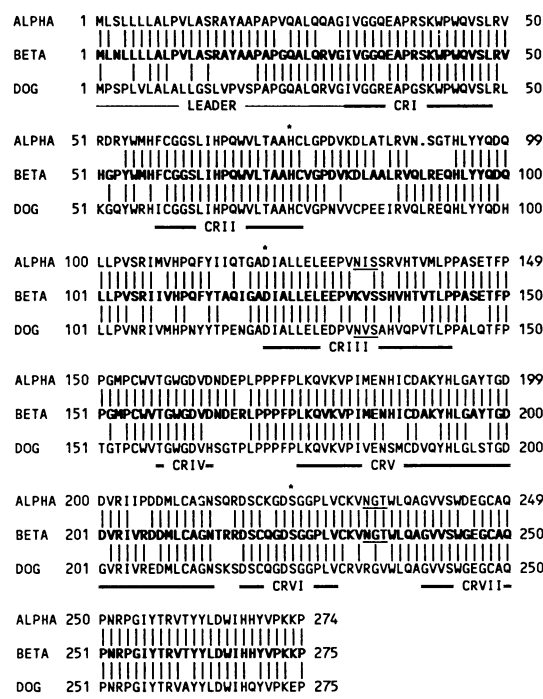

Figure 4. Comparison of the amino acid sequences of human α- and β-tryptase and dog tryptase. The extent of each of the seven conserved regions (CRI–CRVII) of the trypsin superfamily (27) are designated with double lines, the leader with a single line, active site residues with an asterisk and the putative tripeptide carbohydrate binding regions by underlining.

tively, show 53 and 60% identity with the dog tryptase leader. Each leader has a terminal Gly residue, consistent with the novel mode of activation suggested for the dog protease (12). Generally, β-tryptase is more similar than α-tryptase to the corresponding dog enzyme. Whether a second cDNA for dog tryptase, more closely analogous to human α-tryptase, exists is not known at present. The enzyme called dog protease (12) appears to have substantially less homology to dog tryptase than the human enzymes do to one another suggesting that dog protease may not be of the tryptase family.

Tryptase (23) is the name originally used to describe a trypsin-like activity in human and canine mast cells (24) and was later chosen as the name for the corresponding purified enzyme in human (1), canine (25), and rat (26) mast cells. In each species, only mast cells appear to contain substantial amounts of this enzyme. Functional criteria common to these enzymes is their tetrameric structure and their binding to and stabilization by heparin. Conserved sequence homologies among the human and dog enzymes provide a further criterion to define this enzyme family both within and between species. We recommend the term tryptase be reserved for enzymes having these characteristics.

Acknowledgments

References

1. Schwartz, L. B., R. A. Lewis, and K. F. Austen. 1981. Tryptase from human pulmonary mast cells. Purification and characterization. *J. Biol. Chem.* 256:11939–11943.

2. Schwartz, L. B., R. A. Lewis, D. Seldin, and K. F. Austen. 1981. Acid hydrolases and tryptase from secretory granules of dispersed human lung mast cells. *J. Immunol.* 126:1290–1294.

3. Craig, S. S., N. M. Schechter, and L. B. Schwartz. 1988. Ultrastructural analysis of human T and TC mast cells identified by immunoelectron microscopy. *Lab. Invest.* 58:682–691.

4. Irani, A. A., N. M. Schechter, S. S. Craig, G. DeBlois, and L. B. Schwartz. 1986. Two types of human mast cells that have distinct neutral protease compositions. *Proc. Natl. Acad. Sci. USA.* 83:4464–4468.

5. Schwartz, L. B., D. D. Metcalfe, J. S. Miller, H. Earl, and T. Sullivan. 1987. Tryptase levels as an indicator of mast-cell activation in systemic anaphylaxis and mastocytosis. *N. Engl. J. Med.* 316:1622–1626.

6. Schwartz, L. B., J. W. Yunginger, J. S. Miller, R. Bokhari, and D. Dull. 1989. The time course of appearance and disappearance of human mast cell tryptase in the circulation after anaphylaxis. *J. Clin. Invest.* 83:1551–1555.

7. Smith, T. J., M. W. Hougland, and D. A. Johnson. 1984. Human lung tryptase, purification and characterization. *J. Biol. Chem.* 259:11046–11051.

8. Harvima, I. T., N. M. Schechter, R. J. Harvima, and J. E. Fräki. 1988. Human skin tryptase: purification, partial characterization and comparison with human lung tryptase. *Biochim. Biophys. Acta.* 957:71–80.

9. Cromlish, J. A., N. G. Siedah, M. Marcinkiewicz, J. Hamelin,

D. A. Johnson, and M. Chretein. 1987. Human pituitary tryptase: molecular forms, NH-terminal sequence, immunocytochemical localization, and specificity with prohormone and fluorogenic substrates. *J. Biol. Chem.* 262:1363–1373.

10. Schwartz, L. B. 1985. Monoclonal antibodies against human mast cell tryptase demonstrate shared antigenic sites on subunits of tryptase and selective localization of the enzyme to mast cells. *J. Immunol.* 134:526–531.

11. Miller, J. S., E. H. Westin, and L. B. Schwartz. 1989. Cloning and characterization of complementary DNA for human tryptase. *J. Clin. Invest.* 84:1188–1195.

12. Vanderslice, P., C. S. Craik, J. A. Nadel, and G. H. Caughey. 1989. Molecular cloning of dog mast cell tryptase and a related protease: Structural evidence of a unique mode of serine protease activation. *Biochemistry.* 28:4148–4155.

13. Toneguzzo, F., S. Glyn, E. Levi, S. Mjolsness, and A. Hayday. 1988. Use of chemically modified T7 DNA polymerase for manual and automated sequencing of supercoiled DNA. *BioTechniques.* 6:460–469.

14. Moxley, G. 1989. DNA polymorphism of immunoglobulin kappa confers risk of rheumatoid arthritis. *Arthritis Rheum.* 32:634–637.

15. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* 12:387–395.

16. Kozak, M. 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell.* 44:283–292.

17. Avraham, S., R. L. Stevens, C. F. Nicodemus, M. C. Gartner, K. F. Austen, and J. H. Weis. 1989. 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. *Proc. Natl. Acad. Sci. USA.* 86:3763–3767.

18. Steitz, T. A., R. Henderson, and D. M. Blow. 1969. Structure of crystalline alpha-chymotrypsin. *J. Mol. Biol.* 46:337–348.

19. Polgar, L. 1987. Structure and function of serine proteases. *In* Hydrolytic Enzymes. A. Neuberger and K. Broddehurst, editors. Elsevier Science Publishers, Amsterdam. 159–200.

20. Krieger, M., L. M. Kay, and R. M. Stroud. 1974. Structure and specific binding of trypsin: comparison of inhibited derivatives and a model for substrate binding. *J. Mol. Biol.* 83:209–230.

21. Chung, D. W., K. Fujikawa, B. A. McMullen, and E. W. Davie. 1986. Human plasma prekallikrein, a zymogen to a serine protease that contains four tandem repeats. *Biochemistry.* 25:2410–2417.

22. Fujikawa, K., D. W. Chung, L. E. Hendrikson, and E. W. Davie. 1986. Amino acid sequence of human Factor XI, a blood coagulation factor with four tandem repeats that are highly homologous with plasma kallikrein. *Biochemistry.* 25:2417–2424.

23. Lagunoff, D., and E. P. Benditt. 1963. Proteolytic enzymes of mast cells. *Ann. NY Acad. Sci.* 103:185–198.

24. Glenner, G. C., and L. A. Cohen. 1960. Histochemical demonstration of species-specific trypsin-like enzyme in mast cells. *Nature (Lond.).* 185:846–847.

25. Caughey, G. H., N. F. Viro, J. Ramachondran, S. C. Lazarus, D. B. Borson, and J. A. Nadel. 1987. Dog mastocytoma tryptase: affinity purification, characterization and amino-terminal sequence. *Arch. Biochem. Biophys.* 258:555–563.

26. Kido, H., N. Fukusen, and N. Katunuma. 1985. Chymotrypsin and trypsin-type serine proteases in rat mast cells: properties and functions. *Arch. Biochem. Biophys.* 239:436–443.

27. Furie, B., D. H. Bing, R. J. Feldman, D. J. Robinson, J. P. Burnier, and B. C. Furie. 1982. Computer-generated models of blood coagulation factor Xa, factor IXa, and thrombin based upon structural homology with other serine proteases. *J. Biol. Chem.* 257:3875–3882.