Characterization of Genes Encoding Known and Novel Human Mast Cell Tryptases on Chromosome 16p13.3*

Michele Pallaro, Marlena S. Fejzo, Laleh Shayesteh, John L. Blount, and George H. Caughey‡

From the Cardiovascular Research Institute and Departments of Medicine and Laboratory Medicine, University of California, San Francisco, California 94143-0911

Tryptases are serine proteases implicated in asthma and are very highly expressed in human mast cells. They fall into two groups, α and β. Although several related tryptase mRNAs are known, it is unclear which if any are transcripts of separate haploid genes. The studies described here investigated the nature and number of human tryptases and sought possibly novel members of the family. To this end, two human bacterial artificial chromosome (BAC) clones containing tryptase genes were identified and mapped to chromosome 16p13.3, of which ~2.2 megabases are syntenic with the part of mouse chromosome 17 containing tryptase genes mouse mast cell protease (mMCP)-6 and -7. Sequencing and restriction mapping suggest that the BACs may partially overlap. Sequenced BAC genes correspond to three known β-tryptases (βI, βII, and βIII), an α-like gene, and a pair of novel hybrid genes related partly to α/β-tryptases and partly to orthologs of mMCP-7. βI and βIII, βI and αII, as well as the two mMcP-7-like genes, may be alleles at single loci; in total, there are at least three nonallelic tryptase genes in the isolated BAC clones. DNA blotting and restriction analysis suggest that the BACs include most members of the immediate tryptase family. Thus, chromosome 16p13.3 harbors a cluster of known and previously undescribed members of the tryptase gene family.

Mast cell tryptases are trypsin-like serine proteases whose distinguishing features include secretion as catalytically active, heparin-bound oligomers that resist inactivation by plasma antiproteases (1). Extracellular targets of human tryptases suggested by in vitro studies include neuropeptides, procoagulant proteins, urokinase, stromelysin, and proteinase-activated receptor-2 (2–6). Isolated tissue and in vivo studies in animals (7, 8) and humans (9, 10) suggest a pathogenic role for tryptases in asthmatic airway inflammation and obstruction. Moreover, studies in mice suggest that genetic determinants of asthma-like airway hyperresponsiveness include one or more genes in the vicinity of tryptase loci on chromosome 17 (11).

Known members of the tryptase family are expressed selectively in mast cells and basophils (12–14). Compared with amounts of granule-associated serine proteases in leukocytes, levels of tryptases in human mast cells are exceptionally high, comprising up to 25% of cellular protein (15, 16). These levels are achieved by packaging and storage in secretory granules. Purified human tryptases exhibit size heterogeneity by SDS-polyacrylamide gel electrophoresis and can be chromatographically separated into isoforms which in part may be due to differences in glycosylation (17). Based on differences in immunoreactivity and in amino acid sequence deduced from cloned cDNAs, known human tryptases divide into two groups, α and β (18–20). β-Tryptase appears to be the major type stored in mast cell secretory granules and is the major form isolated from extracts of the richest tissue source, which is the lung (21). β-Tryptase mRNAs are also the main types identified in purified mast cells from human lung and skin (22). On the other hand, by immunossay, α-tryptase is the major isoform in blood in normal subjects (23) and may be the predominant type expressed by basophils (22). The basis for the cell selectivity of tryptase expression and for the exceptionally high expression of β-tryptase in human mast cells is not known.

Four different human tryptase cDNAs (α, βI, βII, and βIII) have been generated from human lung (19, 20) and skin (18) mRNA, and the organization and complete sequence of one human tryptase gene (βI) has been determined (18). In addition, one or more tryptase genes are expressed in cell lines, including Mono Mac 6 cells (α-tryptase) (24), HMC-1 and U-937 cells (βI-tryptase) (24), and KU812 cells (α- and βI-tryptases) (22, 25). In addition to βI-tryptase, HMC-1 cells also transcribe a second very closely related mRNA (here designated βIIb-tryptase), whose product is predicted to differ from βI-tryptase itself by one amino acid (26). Because all of the described β-tryptases are 98–99% identical in amino acid sequence, they have been considered to be possible allelic variants of each other (18, 27). α-Tryptase is less closely related (91% identical to βI) and is more likely to be a product of a separate gene in the haploid genome (19, 20). Both α- and β-tryptases appear to reside on human chromosome 16 based on amplification of DNA from hamster-human hybrid cells (20). However, the number and proximity of haploid α- and β-genes as well as their precise chromosomal location and genetic neighbors are not known.

In mice, two different tryptases (mMCP-6 and -7)1 have been identified (12, 13). Structurally, human α- and β-tryptases are more closely related to one another than to mMCP-6 and -7, which are much more different from each other than any combination of known human tryptases (27). These observations suggest 1) that the ancestors of the known human tryptases diverged from each other after the point in evolution when the

---

1 The abbreviations used are: mMCP, mouse mast cell protease; YAC, yeast artificial chromosome; BAC, bacterial artificial chromosome; FISH, fluorescence in situ hybridization; UTR, untranslated region; PCR, polymerase chain reaction; bp, base pair(s); kb, kilobase(s); contig, group of overlapping clones.

---

* This work was supported in part by Grants HL-54774 and HL-24136 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Cardiovascular Research Inst., University of California, San Francisco, CA 94143-0911. Tel.: 415-476-9920; Fax: 415-476-9749; E-mail: ghc@itsa.ucsf.edu.

‡ This paper is available on line at http://www.jbc.org.

Received for publication, September 22, 1998, and in revised form, November 16, 1998.
known mouse and human tryptases shared a common ancestor, 2) that human α- and β-tryptase therefore are not the equivalents, respectively, of mMCP-6 and -7, and 3) that the human genes that are true orthologs of mMCP-6 and -7 may await discovery. The latter possibility is particularly likely to apply to orthologs of mMCP-7, whose gene structure is distinct in several features from that of the cloned human β-tryptase gene (13). Among serine protease genes as well as other extended gene families, differences in intron-exon organization tend to (13). Among serine protease genes as well as other extended gene families, differences in intron-exon organization tend to

In addition to functional data linking tryptases to the pathogenesis of asthma, recent genetic data suggest that chromosome 16 in some populations contains determinants of susceptibility to asthma (28). In work reported here, to explore a genesis of asthma, recent genetic data suggest that chromosomal gene families, differences in intron-exon organization tend to

**EXPERIMENTAL PROCEDURES**

**Screening of YAC and BAC Libraries—** Human genomic MegaYAC libraries A, B, and C from Center d’Etude du Polymorphisme Humain (29) and a human genomic BAC (30) library were screened by PCR using sets of oligonucleotide primers spanning various portions of known tryptase cDNAs and genes. When MegaYAC screening with these tryptase-specific primers failed to identify positive clones, the BAC library was screened. The primers that identified two BAC clones and isolated two BAC clones amplify a highly conserved 184-bp region in the 3′-UTR of tryptase genes. This region contains repetitive elements (see Fig. 1). These PCR primers (5′-caacacacactcttc and 5′-caggaggggactcggtg) were used at a concentration of 200 nM in buffer F from the Optimizer kit (InVitrogen) containing Taq Gold (Perkin-Elmer). Reactions were cycled as follows: 95 °C for 10 min (one cycle); 95 °C for 1 min, 55 °C for 30 s, 72 °C for 30 s (25 cycles); and 7 min at 72 °C (1 cycle). Products were separated on agarose gels.

**Blotting and Hybridization—** Full genomic and BAC DNA was digested with restriction endonucleases at 37 °C, size-fractionated on agarose gels, transferred to Hybond N-Plus membranes (Amersham Pharmacia Biotech), and then fixed by UV cross-linking or baking in a vacuum oven. Blots were prehybridized for 1 h at 65 °C in 90 mM sodium citrate buffer containing 0.9 M NaCl, 2.5% SDS, 1 g/liter each of Ficoll 400, polyvinylpyrrolidone, and bovine serum albumin, and 100 µg/ml of sheared salmon sperm DNA and then probed with radiolabeled DNA prepared from full-length human tryptase βII cDNA (18) or with a BAC βI gene fragment containing portions of flanking introns as templates (18). After labeling, probes were purified on a G-25 spin column (Amersham Pharmacia Biotech) and hybridized at 2 × 10⁵ to 1 × 10⁶ cpm/ml. After incubation overnight, blots were washed twice in 30 mM sodium citrate buffer containing 0.3 mM NaCl and 0.1% SDS for 15 min at 65 °C followed by two washes in 1.5 mM sodium citrate buffer containing 15 mM NaCl and 0.1% SDS at 65 °C for 15 min. Hybridizing bands were revealed by autoradiography.

**Fluorescence in Situ Hybridization—** BAC clones identified in the present work and a λ clones containing the tryptase βI gene (18) identified in prior work were localized to a specific human chromosomal band by standard FISH techniques, as described previously (31). Briefly, metaphase spreads of human peripheral blood lymphocytes were denatured in 30 mM sodium citrate buffer containing 0.3 mM NaCl and 70% formamide for 5 min at 72 °C and then dried in graded solutions of ethanol. BAC and λ probes incorporating digoxigenin-11-dUTP were prepared by nick translation, and a human chromosome 16-specific centromeric probe was labeled directly by incorporating Texas Red dUTP. 10 µl of hybridization mixture contained labeled probes, and Cot-1-digested DNA in 30 mM sodium citrate buffer containing 0.3 mM NaCl, 50% formamide, and 10% dextran sulfate were applied to each slide. After incubation overnight at 37 °C, slides were washed to remove unbound probes, incubated with fluorescein-labeled anti-digoxigenin antibody, and then counterstained with 0.2 mM 4,6-diamino-2-phenyldole to reveal chromosomal bands.

**Isolation and Sequencing of BAC Tryptase Genes—** The BAC libraries were originally prepared from partial HindIII digests of genomic DNA. To obtain smaller fragments of the BAC inserted genomic sequences, the insertions were completely digested with HindIII, and the resulting fragments were subcloned into pBluescript (SK +) (Stratagene). Briefly, HindIII-digested BAC DNA ligated into pBluescript was electroporated into DH10B E. coli (Life Technologies, Inc.). Bacterial plaques containing cloned tryptase genes were identified by plaque hybridization using labeled tryptase DNA probes. Plasmid DNA from individual clones was purified, digested with HindIII, and compared with the pattern of fragments generated by HindIII digestion of DNA from the parent BAC clones. Relevant portions of tryptase-positive subclones were sequenced by the Biomolecular Resource Center of the University of California at San Francisco using tryptase- or vector-specific oligonucleotide primers.

**BAC Restriction Mapping—** BAC DNA was digested with a variety of endonucleases, including NotI, EcoRI, HindIII, BamHI, BglII, SacI, PstI, AvaII, MscI, and AluI. When necessary because of the generation of large fragments, restriction digests were separated on a 1% agarose gel via inverted field gel electrophoresis on a CHEF Mapper (Bio-Rad). Electrophoresed DNA was blotted and sequentially probed with tryptase- and BAC insertion end-specific probes. BAC end sequences for each clone were determined by direct sequencing using primers specific for the BAC cloning vector arms. BAC end-specific probes were then generated via PCR, cloned into pCR 3.1, and then sequenced and labeled for use as probes.

**PCR Screening of BAC Tryptase Genes—** Because analysis of human tryptase cDNA and gene sequences revealed the presence of nucleotide changes in exon 3, causing differences in susceptibility to SecII digestion, we developed a PCR approach to screen BACs for tryptase genes based on restriction polymorphisms. We generated a 242-bp amplier bracketing restriction SecII restriction sites using the following primers: 5′-caggagcccagcaggtgcg and 5′-caggagggactcggtg. These primers were used at a concentration of 200 mM in buffer containingVent (exo-) DNA polymerase (New England Biolabs, Beverly, MA). Reactions were cycled as follows: 95 °C for 5 min (1 cycle); 95 °C for 30 s, 55 °C for 45 s, 72 °C for 30 s (40 cycles); and then 7 min at 72 °C. Products were separated on agarose gels. Some of the 242-bp amplifiers were sequenced after subcloning into pCR 2.1 (Invitrogen) or pBlue-script (Stratagene).

**Gene Comparisons and Phylogenetic Analysis—** Matrix plots, multiple sequence alignments, and phylogenetic trees were generated with GeneWorks 2.5.1 (Oxford Molecular). All sequences were also aligned with those in the GenBank nonredundant data base and in the expressed sequence tag data base using various BLAST algorithms (32). Repetitive sequences were identified via matrix plots and by analyzing sequences by RepeatMasker2.

**RESULTS AND DISCUSSION**

**Identification of BACs Containing Tryptase Genes—** In matrix analyses of previously cloned tryptase cDNAs (Fig. 1A), we noted that the 3′-UTR of all human tryptases is highly conserved in a region containing several different repeated sequences (Fig. 1B) that are present to a much smaller degree in dog tryptase and are absent in murine tryptase 3′-UTRs by matrix analysis (not shown). These repeated sequences differ from those reported in mouse chymase (mMCP-1, -2, and -4) 3′-UTRs, which modulate mRNA stability (33). However, the human tryptase 3′-UTR repeats may plav a similar role. The portion of the human tryptase 3′-UTR bracketing the repeated sequences is unique to tryptases based on the results of nucleic acid database searches. PCR primers amplifying the repeat region identified two clones from a BAC library but no clones from a YAC library.

**Characterization of BAC Tryptase Genes by DNA Blotting—** To confirm the presence of tryptase-like sequences in the identified BACs, restriction endonuclease-digested BAC DNA was subjected to blotting and probed separately with a full-length tryptase βII cDNA and with a BamHI/AvaII βI gene fragment containing exon 5 and portions of flanking introns (18). The results shown in Fig. 2 in comparison with similar treated blots of human full genomic DNA suggest that both BACs contain more than one tryptase gene. Although the banding patterns of BACs A and B differ, several bands appear

---

1 http://ftp.genome.washington.edu/cgi-bin/RepeatMasker.
The portion of Alu-sparse element of the dimeric family in the human genome. This YAC-poor region is syntenic with a small portion of mouse chromosome 17 containing mMCP-6 and -7 genes (35). This region of synteny corresponds to <0.1% of the human genome.

Characterization of BAC Inserts—By inverted field gel electrophoresis, the BAC A and B insertions are ~74 and 62 kb, respectively. As shown in Fig. 4A, we developed a complete EcoRI, NotI, and HindIII (and partial BgIII) restriction map of both BACs. Double digestes of the BACs with NotI and EcoRI generate a similar ~17.5-kb fragment from both BACs. Multiple HindIII-generated bands also are similar. Furthermore, a probe based on the SP6 end of BAC A hybridizes with a specific HindIII fragment of BAC B and a probe developed from the SP6 end of BAC B hybridizes with a specific small, internal HindIII fragment of BAC A. However, probes prepared from the T7 ends of both BACs hybridize solely with the BAC of origin. BAC A ends in the middle of a 152-bp fragment of a tryptase gene, which has a highly similar (but complete) counterpart “mMCP-7-like gene” (see below) in BAC B. Alignment of the two BACs based on restriction maps and cross-hybridizing end sequences and genes (Fig. 4A) suggest that over half of each BAC is structurally homologous to the other, although the inserted sequences are oriented in opposite directions relative to the SP6 and T7 polymerase sites of the vector. Because the sequences at ho-

![Figure 1](https://example.com/fig1.png)

**FIG. 1** Comparison of α and β genes and identification of conserved repeats in flanks and 3′-UTRs. *A* shows a computer-assisted matrix plot of human αII and β tryptase gene nucleotide sequence. The two genes are highly similar, even in introns and flanking regions, as manifested by the strong, nearly continuous diagonal line of identity. This suggests that the two nonallelic genes arose from recent duplications. The 5′ and 3′ termini of the genes are indicated. Three prominent regions of repeats (designated α–c) appear as lines parallel to the main line of identity. Repetitive region α is near the end of the sequenced 5′ flanks. Repetitive region b is in the 3′-UTR. Repetitive region c, in the 3′ flank, is a short interspersed element of the dimeric Alu family (subfamily Jb) whose relatives are numerous in the human genome. *B* shows the portion of βI 5′ flank and 3′-UTR containing the repeated elements identified above. The sequences are nearly identical in αII. The 5′ flank sequence contains three direct, G-rich, 19-bp repeats, which are underlined. Data base searches suggest that the repeated elements may be unique to tryptases, because they are not detected elsewhere. Because of their location and conservation, they are candidate transcriptional regulatory sites that can lie within tandemly repeated sequences. In the 3′-UTR sequence, two types of repeats are underlined, with or without italics. All known human tryptase cDNAs and genes have these repeats, which appear to be unique and to be present to a much lower degree or not at all in other mammalian tryptase cDNAs, as revealed by further matrix analysis (not shown). These repeated sequences potentially play a role in regulation of human tryptase mRNA stability. Residues in bold print correspond to primers used in PCR amplifications to screen YAC and BAC libraries for genomic clones containing tryptase genes.

identical, suggesting either that the BACs partially overlap or that they contain genes that are sufficiently similar that they yield similar restriction patterns. Together, as shown in Fig. 2, the bands in the two BACs account for all of the bands seen in the full genomic blot, suggesting that the BACs collectively include most or all of the tryptase genes of the genome.

**Chromosomal Location of BAC Inserts and the β Tryptase Gene**—By FISH, both BACs, as well as a previously isolated λ phage clone containing the tryptase βI gene (18), localize to a site near the end of the short arm of chromosome 16 (band 13.3). There is no discernible difference in the location of any of these clones by this technique, suggesting that all map to the same or closely adjacent sites. These data confirm and extend prior data assigning α and β tryptase genes to chromosome 16 using panels of hamster human hybrids (19, 20). Because no YACs containing tryptase genes were identified, it is likely that the βI-tryptase gene and the identical tryptase-rich BACs map to the small portion of chromosome 16p13.3 that is not “covered” by YACs (34). As shown in Fig. 3, part of the 16p13.3 YAC-poor region is syntenic with a small portion of mouse chromosome 17 containing mMCP-6 and -7 genes (35). This region of synteny corresponds to <0.1% of the human genome.
mologous loci are similar but not identical in the two BACs, despite close organizational parallels, we conclude that BACs A and B are most likely to be partially overlapping sections of sister chromatids and that tryptase genes at homologous sites are allelic variants of each other.

Characterization of BAC β-Tryptase Genes—BAC A contains a β-tryptase gene whose exons match previously described βII cDNAs. BAC B contains two identified β-like genes whose exons correspond to cDNAs βI and βIII (18). The organization (see Fig. 6A) of the BAC β genes is identical to that of a previously characterized βI gene (GenBank™ accession number M33494) (18). It should be noted that the BAC B βI gene represents the Glu99 version of tryptase (βIa) rather than the Lys99 version (βIb, which may be another allele at this locus) also transcribed in HMC-1 cells (26). The βIII gene resides at a separate site. The reason for the existence of two loci of highly similar genes is unclear. βIII-Tryptase differs from βIa at only three of 245 catalytic domain amino acids, none of which (based on preliminary modeling) are in sites thought to be critical determinants of function. One potential explanation for the presence of two very similar genes is a recent gene amplification event, perhaps in response to evolutionary pressure to increase production of tryptase mRNA and protein. Related events are observed in connection with other genes. For example, gene amplification is a known genetic adaptation to selective pressure to increase the product of the mammalian dihydrofolate reductase gene and is a mechanism that can occur more readily than regulatory region mutations as a means of increasing transcription (36). Another possible explanation is that the duplication of β genes was remote rather than recent and that near identity at the two loci has been maintained by gene conversion (37), which is a postulated mechanism for maintaining similarity of tandemly arrayed genes. Intriguingly, the three residues by which βIII differs from βI tryptase are all α-like residues, suggesting that there may have been genetic exchange between β and α genes. Our complete βI and βIII tryptase gene sequences (including 5′ and 3′ flanks) and βII partial sequences, including 3′ UTR, are deposited in Gen-

![Genomic DNA blots](image)

**FIG. 2.** Genomic DNA blots. This figure shows autoradiograms generated by hybridization of a radiolabeled fragment of a tryptase gene with blots of full genomic and BAC DNA digested with AvaII, BalI (MscI), or BamHI. The probe is a 345-bp AvaII/BamHI fragment of tryptase βI comprising portions of exons 5 and 6 and all of intron 5. The left and right panels show blots of full genomic and BAC DNA, respectively. The position of size markers in kb are shown to the left. Note the similar size of many bands in the full genomic and BAC blots. Although BAC patterns partly differ from each other, in combination they appear to account for all bands generated by full genomic DNA, suggesting that BAC A and B together may include most of the tryptase genes of the genome. Differences in band intensity arise from disparities in the number of genes contributing to a band of a given size and from differences in the extent of homology to the probe. For example, the BamHI digests of full genomic and BAC DNA appear identical, containing a strong ~0.8-kb band due to multiple αβ tryptase genes and a weak ~1.6-kb band attributable to an mMCP-7-like gene, which contains a ~0.8-kb intron 4 insertion (see Fig. 4A) in comparison to α and β genes. This ~1.6-kb band is absent from the BAC A lanes because the BAC A insert contains only part of the mMCP-7-like gene and lacks the 3′ portion homologous to the probe.

![Chromosomal location of human tryptase genes](image)

**FIG. 3.** Chromosomal location of human tryptase genes. A diagram of human chromosome 16 in metaphase is shown at the top of the figure. The solid arrow indicates the approximate location of the FISH signal generated by probes prepared from of a bacteriophage λ clone, containing the human tryptase βI gene, and from BACs A and B, containing multiple tryptase genes. In all cases, selective hybridization was seen near the end of the short arm on band 16p13.3, suggesting that the DNA cloned within the λ and BAC vectors map to the same or highly similar locations in the genome. The remainder of the figure demonstrates that tryptase-hybridizing human chromosome 16p13.3 is syntenic with a portion of mouse chromosome 17 containing tryptase genes. After identification of the general location of tryptase genes by FISH, we identified a small portion of the telomeric side of band 16p13.3, suggesting that the DNA cloned within the BAC A insert contains only part of the mMCP-7-like gene and lacks the 3′ portion homologous to the probe. One possible explanation for the presence of two very similar genes is a recent gene amplification event, perhaps in response to evolutionary pressure to increase production of tryptase mRNA and protein. Related events are observed in connection with other genes. For example, gene amplification is a known genetic adaptation to selective pressure to increase the product of the mammalian dihydrofolate reductase gene and is a mechanism that can occur more readily than regulatory region mutations as a means of increasing transcription (36). Another possible explanation is that the duplication of β genes was remote rather than recent and that near identity at the two loci has been maintained by gene conversion (37), which is a postulated mechanism for maintaining similarity of tandemly arrayed genes. Intriguingly, the three residues by which βIII differs from βI tryptase are all α-like residues, suggesting that there may have been genetic exchange between β and α genes. Our complete βI and βIII tryptase gene sequences (including 5′ and 3′ flanks) and βII partial sequences, including 3′ flank, are deposited in Gen-
Characterization of BAC α-Tryptase Genes—BAC A contains a gene (αII) with an exon sequence similar but not identical to that reported for α cDNA (here designated αI) (19). The deduced amino acid sequences of αI and II, as shown in Fig. 5, differ in only three residues. The alignment shown in Fig. 4A suggests that the BAC A αII gene and the BAC B βIIa gene may be allelic partners, a possibility further supported by the very high homology between the two genes and their respective extended flanking regions, as revealed by the matrix analysis in Fig. 1A. The gene corresponding to previously reported αI cDNA also may be an allele at this locus or may be the product of another α gene not detected in the BACs. This is the first report of α tryptase gene structure, which, not surprisingly, is similar to that of β genes (Fig. 6A), containing a 6-exon, 5-intron pattern, including the placement of a phase-0 intron immediately upstream of the initiator Met codon, which separates the site of transcription initiation from protein coding sequence. This feature is characteristic of tryptases but is unusual in other genes. The α gene described here also has an 11-bp deletion in intron 4 compared with β-tryptase genes.
Another difference between α and β genes is the presence of a SacII restriction site in exon 3 (see Figs. 4B and 6A). These differences provide ways to measure a versus β DNA in PCR-generated amplimers containing tryptase gene products of mixed heritage. The predicted protein products of both new α-tryptases in this work contain Gln residues in the −3 position of the propeptide (in the −3 position relative to the IVG N terminus of the mature catalytic domain) is a key α-like feature that may limit autolytic trimming of the propeptide by hydrolysis at the basic residue that occupies this position in most tryptases.

Characterization of BAC mMCP-7-like Genes—As noted above, the SP6 end of BAC A contains the 5′ portion of a tryptase gene. This sequence ends at the terminal HindIII site of BAC (the BAC library was generated from HindIII partial digests). BAC B, however, contains the full sequence of the apparent allelic partner of the partial BAC A gene. Although the phase and placement of introns are similar to those of other human tryptase genes (as shown in Fig. 6A), their size varies, particularly in exon 4, which is much larger than the other introns, and is also true of exon 4 in the gene encoding mMCP-7 (13). Among the protein-coding sequences, however, only exon 5 closely resembles mMCP-7. The region of nucleotide sequence similarity to mMCP-7 includes all of exon 5 and contiguous portions of introns 4 and 5. However, there is no discernible similarity between most of the large human intron 4 insertion and the intron of similar size in the mMCP-7 gene. The closest match in searches of the GenBankTM data base (including expressed sequence tags), using exon 5 as the query sequence, are mMCP-7 and the apparent orthologs of mMCP-7 in gerbil (39) and rat (40). No human expressed sequence tags are highly similar. Remarkably, the remaining exons of the mMCP-7-like gene are closely related to α- and β-tryptases. The α-like features include a predicted Gln at the −3 position of the propeptide and a SacII site in exon 3. However, this α/mMCP-7-like gene has a second, novel SacII site in the same exon; it also has a β-like PstI site in exon 6. Finally, a stop codon that is 40 amino acid residues premature relative to that in other tryptase genes is identified. Together, the above findings suggest that this mMCP-7-like gene is a hybrid that may have arisen from one or more homologous recombination events involving α-like genes and the putative ortholog of mMCP-7.

The similarity of exons 1–4 to α-tryptases suggests that the formation of this hybrid was recent on the time scale of mammalian evolution. At present, there is no evidence that other exons of the mMCP-7-homologous gene, which may have become lost following unequal crossing over events, persist in the human genome. A dendrogram summarizing the possible phylogenetic relationship between different exons of the α/mMCP-7-like gene is shown in Fig. 6B, which also suggests that known α- and β-tryptases may be orthologs of another mouse tryptase, mMCP-6. The sequence of the BAC B gene (mMCP-7-like I), encompassing all protein-coding exons and intervening introns, and that of the 5′ portion of the BAC A gene (mMCP-7-like II) up to the HindIII site at the SP6 terminus of the BAC insert, are deposited in GenBank™ (AF098926 and AF098931).

Distinguishing Tryptase Genes by Sac II Digestion of PCR Amplifiers—Sequencing of BAC A and B tryptase genes revealed differences in SacII restriction sites in exon 3 (Fig. 6A). Tryptases II, βI, II, and III have no SacII sites in this region. Therefore, as shown in Fig. 4B, the 242-bp fragment of the gene amplified using the primers described above is unaffected by SacII. The α II gene, however, has a single site, generating two fragments, of which the 113-bp fragment is unique to known α-genes. The mMCP-7-like gene in BAC A has a single SacII site, which differs from the α site and generates fragments of 51 and 191 bp. The mMCP-7-like gene in BAC B has two sites, one the same as the BAC A gene and the other identical to the

Characterization of Human Tryptase Genes

FIG. 5. Amino acid sequence alignment. The deduced amino acid sequence of human prepro-mMCP-7-like tryptase is compared with that of α- and β-tryptases. Regions of sequence identity among all of the tryptases are boxed. Note that the region of greatest sequence disparity corresponds to exon 5 (residues 167–221), which is the mMCP-7-homologous portion of the human mMCP-7-like gene. The remainder of the sequence is nearly equally related to human α- and β-tryptases, although Glns28 of the propeptide (in the −3 position relative to the IVG N terminus of the mature catalytic domain) is a key α-like feature that may limit autolytic trimming of the propeptide by hydrolysis at the basic residue that occupies this position in most tryptases.

| H mMCP-7-like AA | H tryptase beta II AA | H tryptase beta I AA | H tryptase beta III AA | H tryptase alpha II AA | H tryptase alpha I AA |
|------------------|----------------------|---------------------|-----------------------|-----------------------|----------------------|
| Ala Leu Asp Ala | His Leu Asp Ala      | His Leu Asp Ala     | His Leu Asp Ala       | His Leu Asp Ala       | His Leu Asp Ala       |
| Val Asp Asp Ala  | Val Asp Asp Ala      | Val Asp Asp Ala     | Val Asp Asp Ala       | Val Asp Asp Ala       | Val Asp Asp Ala       |
| Phe Asp Asp Ala  | Phe Asp Asp Ala      | Phe Asp Asp Ala     | Phe Asp Asp Ala       | Phe Asp Asp Ala       | Phe Asp Asp Ala       |
| Arg Asp Asp Ala  | Arg Asp Asp Ala      | Arg Asp Asp Ala     | Arg Asp Asp Ala       | Arg Asp Asp Ala       | Arg Asp Asp Ala       |
| 50               | 50                   | 50                  | 50                    | 50                    | 50                    |

| H mMCP-7-like AA | H tryptase beta II AA | H tryptase beta I AA | H tryptase beta III AA | H tryptase alpha II AA | H tryptase alpha I AA |
|------------------|----------------------|---------------------|-----------------------|-----------------------|----------------------|
| HNPAPGTTRV YYLLW1HRH VPKK | HNPAPGTTRV YYLLW1HRH VPKK | HNPAPGTTRV YYLLW1HRH VPKK | HNPAPGTTRV YYLLW1HRH VPKK | HNPAPGTTRV YYLLW1HRH VPKK | HNPAPGTTRV YYLLW1HRH VPKK |
| 275               | 275                  | 275                 | 275                   | 275                   | 275                   |
Characterization of Human Tryptase Genes

α-gene site, resulting in 51-, 62-, and 129-bp fragments. Thus, the 51- and 62-bp fragments are characteristic of mMCP-7-like genes. As predicted, based on the genes we characterized, BAC A amplimers restricted with SacII generate fragments of 242, 191, 129, 113, and 51 bp but are missing the 62-bp fragment found only in the BAC B mMCP-7-like tryptase. BAC B, however, has the 51-bp fragment, as well as all of the BAC A fragments. Similar PCR/SacII analysis of individual HindIII BAC subclones from which αII, βI, βII, and βIII genes were obtained identify only the restriction pattern of the gene already identified in the subclone, suggesting that only one type of gene lies within a given subclone. No α-like genes were identified in the BAC B-derived subclones; however, the 113-bp band in the BAC B digest suggests the possible presence of an α-like gene not yet subcloned. PCR of full human genomic DNA yields all of the bands seen in BAC B. Overall, the PCR results suggest that both BACs contain α, β, and mMCP-7-like genes and are consistent with the hypothesis that the BACs contain overlapping portions of the same locus in distinct haploid chromosomes. An alternative, although less likely interpretation is that each BAC contains separate parts of the genome that are highly homologous because of past, large scale duplication.

Fig. 6. Organization and origins of human tryptase genes. A compares intron-exon structures. Genes so far obtained from BAC clones are of three types: β-, α-, and mMCP-7-like, as shown. Each has a 6-exon (boxes), 5-intron (lines) structure, drawn to scale. UTRs are hatched boxes. Exons encoding prepro- and mature enzyme are lightly shaded and black boxes, respectively. Codons of catalytic His (H) Asp (D), and Ser (S) residues common to serine proteases are indicated. Human gene intron 1 separates 5′-UTR from protein-coding sequence, an idiosyncratic feature of tryptase genes. β- and α-like genes are nearly identical in intron size and phase, except for an 11-bp deletion in intron 4. The mMCP-7-like human gene is distinguished by insertions in introns 4 and 5 and by a 5th exon more closely related to the mouse tryptase mMCP-7 than to any human tryptase, as indicated by stippling of exons. The other exons are closely related to human α- and β-tryptases (slightly more closely to the former than the latter). The 0.8-kb intron 4 insertion is similar in size but not in sequence to that of mMCP-7. Thus, the novel mMCP-7-like human tryptase gene appears to be a chimera arising recently in mammalian evolution from an α-like gene and a relative of mMCP-7, the remainder of which may have been lost to the human genome during recombination. It remains to be seen whether the human mMCP-7-like chimera is expressed and active. B displays dendrograms based on deduced amino acid sequence of tryptase exons 4 and 5. These trees, which were prepared using the unweighted pair group with arithmetic mean (UPGMA) multiple sequence alignment algorithm in GeneWorks, help to clarify the origins of the novel human mMCP-7-like tryptase. The number accompanying each branch is the fraction of mismatched amino acids in pairs of aligned sequences. Human α-tryptases I and II are identical in amino acid sequence in exons 4 and 5, as are βI, II, and III; therefore known α- and β-tryptases appear as single, separate branches. The mMCP-7 branches of the tree are bold lines. The exon 4 tree suggests that the human mMCP-7-like exon 4 is most closely related to the corresponding exon in α-tryptase, whereas exon 5 is most closely related to exon 5 in mMCP-7 and related rodent genes. These data support the proposed hybrid nature of the human mMCP-7-like gene and also suggest that the known human α- and β-tryptases (as well as known dog and cow tryptases) are orthologs of the mouse tryptase mMCP-6.
Characterization of Human Tryptase Genes

Conclusion—In this work we characterize a cluster of mast cell tryptase genes encoding known as well as previously undescribed members of the family extending over ∼50 kb of chromosome 16p13.3. Our findings establish that in the haploid genome there are at least two αβ tryptase genes that appear to be orthologs of mMCP-6. Homologous positions of αII and βIa genes suggest that one or more αβ pairs may be alleles at the same locus. Additionally, we identify a novel gene with one of six exons being closely related to the mouse tryptase, mMCP-7. Duplication of αβ-like genes, via a gene dosage effect, may contribute to the high levels of tryptase expression in human mast cells. Whether the identified unique 3′-UTR repeated sequences and regulatory sequences flanking the α and β genes contribute to differences in gene transcription and whether “locus control” elements exert more global influence over expression of members of the tryptase gene cluster remain to be determined.

REFERENCES

1. Caughey, G. H. (1997) Am. J. Respir. Cell Mol. Biol. 16, 621–628
2. Schwartz, L. B., Bradford, T., and Griffin, J. H. (1985) Biochem. Biophys. Res. Commun. 129, 76–81
3. Tam, E. K., and Caughey, G. H. (1990) J. Clin. Invest. 86, 37–43
4. Stack, M. S., and Johnson, D. A. (1994) J. Biol. Chem. 269, 9416–9419
5. Gruber, B. L., Marchese, M. J., Suzuki, K., Schwartz, L. B., Okada, Y., Nagase, H., and Ramamurthy, N. S. (1995) J. Clin. Invest. 95, 1383–1393
6. Sekizawa, K., Caughey, G. H., Lazarus, S. C., Gold, W. M., and Nadel, J. A. (1989) J. Clin. Invest. 83, 175–179
7. Clark, J. M., Abraham, W. M., Fishman, C. E., Forteza, R., Ahmed, A., Cortez, A., Warner, R. L., Moore, W. R., and Tanaka, R. D. (1995) Am. J. Respir. Crit. Care Med. 152, 2076–2083
8. Johnson, P. R. A., Ammit, A. J., Lin, C. M., Armstrong, C., Basham, J., Caughey, G. H., and Black, J. L. (1997) Eur. Respir. J. 10, 38–43
9. Ma, C., Martin, S., Trask, B., and Hamlin, J. L. (1993) Nature 363, 662–667
10. Reynolds, D. S., Gurlay, D. S., Austen, K. F., and Serafin, W. E. (1991) J. Biol. Chem. 266, 3847–3853
11. McNeil, H. P., Reynolds, D. S., Schiller, V., Ghidyal, N., Gurlay, D. S., Austen, K. F., and Stevens, R. L. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11174–11178
12. Reynolds, D. S., Lewis, R. A., and Austen, K. F. (1981) J. Biol. Chem. 256, 11399–11403
13. Smith, T. J., Hougland, M. W., and Johnson, D. A. (1984) J. Biol. Chem. 259, 11046–11051
14. Little, S. S., and Johnson, D. A. (1995) Biochem. J. 307, 341–346
15. Vandervis, B., Ballinger, S. M., Tom, E. S., Goldstein, S. M., Craig, C. S., and Caughey, G. H. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 3811–3815
16. Miller, J. S., Westin, E. H., and Schwartz, L. B. (1989) J. Clin. Invest. 84, 1188–1195
17. Miller, J. S., Moxley, G., and Schwartz, L. B. (1990) J. Clin. Invest. 86, 864–870
18. Harvima, I. T., Karkola, K., Harvima, R. J., Naukkarinen, A., Neittaanmaki, H., Horsmanheimo, M., and Fraki, J. E. (1990) Arch. Dermatol. Res. 281, 231–237
19. Xia, H. Z., Kepley, C. L., Sakai, K., Cheh, J., Irani, A.-M. A., and Schwartz, L. B. (1995) J. Immunol. 154, 5472–5480
20. Schwartz, L. B., Sakai, K., Bradford, T. R., Ren, S. L., Zweiman, B., Wobore, A. S., and Metcalfe, D. D. (1995) J. Clin. Invest. 96, 2702–2710
21. Huang, R. A., Abrink, M., Gobi, A. E., Nilsson, G., Aveskogh, M., Larsson, L. G., Nilsson, K., and Hellman, L. (1993) Scand. J. Immunol. 38, 359–367
22. Blom, T., and Hellman, L. (1993) Scand. J. Immunol. 37, 203–208
23. Chan, H., Tanaka, R., and Clark, J. (1995) PASEJ 9, 1343 (abstr.)
24. Caughey, G. H. (1995) in Mast Cell Proteases in Immunology and Biology (Caughey, G. H., ed) pp. 305–329, Marcel Dekker, New York
25. Daniels, S. E., Bhattacharyya, S., James, A., Leaves, N. I., Young, A., Hill, M. R., Faux, J. A., Ryan, G. F., Souf, P. N. L., Lathrop, G. M., Musk, A. W., and Cookson, W. O. (1996) Nature 383, 247–250
26. Cohen, D., Chumakov, I., and Weissenbach, J. (1996) Nature 383, 698–701
27. Shizuya, H., Hirren, B., Kim, J. U., Manecke, V., Slepek, T., Tschir, Y., and Simon, M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 8794–8797
28. Kallioniemi, O. P., Kallioniemi, A., Kuriu, T., Thor, A., Chen, L. C., Smith, H. S., Waldman, F. M., Pinck, D., and Gray, J. W. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 5321–5325
29. Altschul, S. F., Gish, W., Miller, W., Myers, E. M., and Lipman, D. J. (1990) J. Mol. Biol. 215, 403–410
30. Xia, Z., Ghidyal, N., Austen, K. F., and Stevens, R. L. (1996) J. Biol. Chem. 271, 8747–8753
31. Doggett, N. A., Goodwin, L. A., Tesmer, G. J., Meinke, L. J., Bruce, D. C., Clark, L. M., Altherr, M. R., Ford, A. A., Chi, H. C., Marrone, B. L., Lengmire, J. L., Lane, S. A., Whitmore, S. A., Lowenstein, M. G., Sutherland, R. D., Mundt, M. O., Knill, E. H., Bruno, W. J., Macken, C. A., Torney, D. C., Wu, J.-R., Griffith, J., Sutherland, G. R., Deaven, L. L., Callen, D. F., and Moyzis, R. K. (1995) Nature 377, (suppl.) 335–385
32. Gurish, M. F., Johnson, K. R., Webster, M. J., Stevens, R. L., and Nadeau, J. H. (1994) Mamm. Genome 5, 656–657
33. Ma, C., Martin, S., Trask, B., and Hamlin, J. L. (1993) Genes Dev. 7, 605–620
34. Kourilsky, P. (1986) Trends Genet. 2, 60–62
35. Sakai, K., Ren, S., and Schwartz, L. B. (1996) J. Clin. Invest. 97, 988–995
36. Murakomo, Y., Ide, H., Itoh, H., Tomita, M., Kobayashi, T., Maruyama, H., Horii, Y., and Nawa, Y. (1995) Biochem. J. 309, 921–926
37. Lutzelschwab, C., Peijer, G., Aveskogh, M., and Hellman, L. (1997) J. Exp. Med. 185, 13–29