Cloning of the Gene and cDNA for Human Heart Chymase*

Hidenori Urata, Akio Kinoshita, Dianne M. Perez, Kunio S. Misono, F. Merlin Bumpus, Robert M. Graham, and Ahsan Husain

From the Department of Heart and Hypertension Research, Research Institute of the Cleveland Clinic Foundation, Cleveland, Ohio 44195-5069

We have recently identified and characterized a chymotrypsin-like serine proteinase in human heart (human heart chymase) that is the most catalytically efficient enzyme described, thus far, for the cleavage of angiotensin I to yield angiotensin II and the dipeptide His-Leu. Compared to other chymases, this enzyme also has an unusually high degree of specificity for the substrate angiotensin I. We report here the molecular cloning and nucleotide sequence of the gene and cDNA encoding human heart chymase, and determination of its entire deduced amino acid sequence. These data indicate that human heart chymase is highly homologous to other members of the chymase subfamily of chymotrypsin-like proteinases and, most likely, all evolved from a common ancestral gene. Potential regulatory elements found in the S'-untranslated region of other chymases are also found in the human heart chymase gene. However, this gene lacks mast cell-specific sequences found in the S'- and S'-untranslated regions of the rat chymase II gene. In addition, human heart chymase contains clusters of unique amino acid sequences located at key positions likely involved in substrate binding, which may contribute to its high substrate specificity. These contrasting features of the human heart chymase gene and cDNA, and the potential determinants of its primary structure that underlie its unique functional characteristics are considered.

Mammalian chymases are chymotrypsin-like serine proteinases found in the secretory granules of mast cell and are most likely involved in neurogenic inflammation (Brain and Williams, 1988), submucosal gland secretion (Sommershoff et al., 1989), parasite expulsion (King and Miller, 1984), lipoprotein and extracellular matrix catabolism (Seppa et al., 1979; Vartio et al., 1981), and control of vasoactive peptide metabolism (Reilly et al., 1982; Wintroub et al., 1984; Caughey et al., 1988a; Franconi et al., 1989; Urata et al., 1990b). The primary structures of mammalian chymases have been identified and characterized for the rat (Benfey et al., 1987; Remington et al., 1988), mouse (Le Trong et al., 1989; Serafin et al., 1990, 1991), and dog (Caughey et al., 1990). The structures of chymases show extensive homology to a group of chymotrypsin-like serine proteinases including neutrophil cathepsin G and granzymes (Salveson et al., 1987; Jenne et al., 1989) and cytotoxic cell proteases (Lobe et al., 1986; Meier et al., 1990). Their catalytic and physicochemical properties differ markedly with respect to substrate specificity, catalytic efficiency, net charge, and solubility (Woodbury et al., 1978; Yoshida et al., 1980; Powers et al., 1985; Le Trong et al., 1987; Caughey et al., 1988b; Urata et al., 1990b). It is known that certain catalytic and physicochemical characteristics of chymases such as RMCP II, compared to those of α-chymotrypsin, are due to unique features including the absence of a disulfide bond (Cys63-Cys611 in α-chymotrypsin), the number of charged residues, and differences in the substrate-binding site (Remington et al., 1988).

In humans, chymase-like proteinases have been isolated from the skin (Schechter et al., 1983, 1986, 1990; Sayama et al., 1987), lung (Wintroub et al., 1986), and heart (Urata et al., 1990b). Human skin chymase is found in the granules of the mast cell and is also located at the dermo-epidermal junction, where it probably binds to the heparan sulfate proteoglycans of the basement membrane (Sayama et al., 1987). Human lung chymase is partially characterized and probably also located in the mast cell secretory granules (Wintroub et al., 1986).

During the course of studies to examine the localization of Ang II receptors and the biochemical mechanisms of Ang II formation in the human heart (Urata et al., 1989, 1990a), an Ang II-forming serine proteinase, human heart chymase, was found and purified from the human left ventricle (Urata et al., 1990b). This chymase shows extensive similarities to other chymases in terms of its N-terminal amino acid sequence, immunological reactivity, inhibition properties, and charge. However, its catalytic efficiency and substrate specificity differ from those of the other chymases. Human heart chymase shows a high degree of catalytic efficiency and substrate specificity for the formation of Ang II and His-Leu from Ang I (Urata et al., 1990b). But human skin chymase appears to be less specific since it not only forms Ang II but also degrades bradykinin (Reilly et al., 1982, 1985).

Since the complete amino acid sequence of human chymase was unknown, it was unclear whether these differences in the human chymases resulted because they are different gene products or because they are the same protein being subjected to different post-translational modifications. The purpose of

* This work was supported in part by a grant from the Reinberger Foundation and by National Institutes of Health Grant HL35713.

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.

Dedicated to the memory of Irvine H. Page.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EMBL Data Bank with accession number(s) M69136 and M69137.

1 Recipient of a fellowship from the American Heart Association, North East Ohio Affiliate.

2 To whom correspondence should be addressed. Tel.: 216-444-8150; Fax: 216-444-9263.
Cloning of Human Heart Chymase

this study, therefore, was to determine the entire deduced amino acid sequence of human heart chymase, to investigate its genomic organization, and to obtain further insights into the determinants of its specific functional characteristics. We report here the cloning of the gene and cDNA for human heart chymase, which have allowed its nucleotide sequence to be established, and its genomic organization to be determined. Based on the deduced primary structure of human heart chymase, the potential determinants of its unique functional characteristics are considered as compared to those of leukocyte chymotryptsin-like serine proteinases.

EXPERIMENTAL PROCEDURES

Trypsin Cleavage and Amino Acid Sequencing—Human heart chymase (0.1 µg) was purified as described previously (Urona et al., 1990b) and was digested with C1-reverse-phase HPLC column chromatography using 0.1% trifluoroacetic acid and a gradient of acetonitrile from 0 to 80% over 90 min. The eluted protein peak was collected and evaporated to dryness. The residue was oxidized by the addition of 50 µl of performic acid at 22°C for 15 min, and then dried. The residue was washed and resuspended in 50 µl of 0.1 M ammonium bicarbonate buffer, pH 8.0, containing 0.1 M CaCl2. The mixture was incubated with p-tosyl-L-phenylalanine chloromethyl ketone-treated trypsin (final trypsin/chymase ratio was 1:50, w/w) at 37°C for 4 h. At the end of the incubation 5 µl of 50% acetic acid was added to the mixture, which was then diluted with 5% acetic acid for the chromatographic separation using a reverse-phase HPLC column, as mentioned above. Isolated peptides were evaporated to dryness and sequenced using an Applied Biosystems model 470A gas-phase sequenator (Foster City, CA). The phenylthiobutyloantoin-derivatives were analyzed using an on-line phenylthiobutyloantoin analyzer (Applied Biosystems model 120).

cDNA Cloning—A human genomic ADASH library (Strategene, La Jolla, CA) was screened with a partial RMCPII cDNA probe (Alu I—HindIII fragment containing the 5'-half of the coding region from nucleotide +80 to +368 from the amino acid sequence of two distinct overlapping fragments was determined bidirectionally using bacteriophage M13mp8 M13mp18 HindIII fragments containing the 5'-half of the coding region from nucleotide +80 to +368 and the 3'-half of the coding region from nucleotide +620 to +1880 (Benfey et al., 1982). Clones that cross-hybridized with degenerate oligonucleotide probes (based on the amino acid sequence of two distinct fragments of human heart chymase (NT and T′1; Figs. 1 and 2) and sequenced using an Applied Biosystems model 120).

The amplified DNA sequences were analyzed using an on-line sequencer (Applied Biosystems model 120).

RESULTS AND DISCUSSION

Protein Sequence of Human Heart Chymase

As shown in Fig. 1, about nine major peptides could be identified following tryptic cleavage and HPLC fractionation of human heart chymase. Approximately 70% of the entire primary structure was determined by direct amino acid sequence analysis of these peptides, and in all cases the amino acid sequences were found to be encoded in an open reading frame after isolation and nucleotide sequencing of the human heart chymase cDNA and gene (Figs. 2 and 3). The identity under high stringency (0.1 × SSC, 0.1% SDS, 60°C, wash conditions), after the fragmented DNA had been transferred to nitrocellulose membrane (Biotrace NT, Gelnstate Science, Ann Arbor, MI).

FIG. 1. Chromatographic separation of tryptic peptides of human heart chymase. The peptide mixture obtained by digestion of oxidized human heart chymase (2 nmol) with N-(p-tosyl)-L-phenylalanine chloromethyl ketone-trypsin was eluted from a Vydac C4 reverse-phase HPLC column (0.46 × 25 cm) at a 0–80% acetonitrile gradient. The purified peptides (T′5 to T′23) shown in Fig. 2 were sequenced.

FIG. 2. Deduced amino acid sequence of human heart chymase determined from the nucleotide sequence of its cDNA. The positions and amino acid sequences of 10 peptides (NT, T′5 to T′23) derived from the purified protein are indicated by the overhead arrows. The wavy lines under certain residues indicate those amino acids in the corresponding cDNA sequence that could not be identified by peptide sequencing. The predicted signal peptide, the N-terminal dipeptide of the proenzyme and the mature enzyme comprise amino acids −21 to −3, −2 to −1, and +1 to 226, respectively. The trid of amino acids, His4, Asp11, Ser15, essential for catalysis by all serine proteinases, are indicated by open triangles. Two consensus sites for N-linked glycosylation are shown by the asterisks above Asn41 and Asn84. Sequences corresponding to the primers used for PCR in cDNA cloning are indicated by the double lines at the 5′ and 3′ regions of the cDNA. Numbers at right, amino acids; numbers in brackets nucleotides.
Cloning of Human Heart Chymase

Thirty clones were identified by screening of the ADASH human genomic library under low stringency with the partial RMCP II cDNA probe. Eight of these clones were further processed for secondary screening and were analyzed for cross-hybridization with two degenerate oligonucleotide probes. These probes were synthesized based on regions of the amino acid sequences of the tryptic fragments that were distinct from those of the other members of the chymase subfamily. The differences were used as a basis for designing degenerate but unique oligonucleotide probes for the cloning of human heart chymase.

Gene Cloning, Structure, and Organization

Thirty clones were identified by screening of the ADASH human genomic library under low stringency with the partial RMCP II cDNA probe. Eight of these clones were further processed for secondary screening and were analyzed for cross-hybridization with two degenerate oligonucleotide probes. These probes were synthesized based on regions of the amino acid sequences of the tryptic fragments that were distinct from those of the other members of the chymase subfamily. The differences were used as a basis for designing degenerate but unique oligonucleotide probes for the cloning of human heart chymase.
family. Two of these eight clones cross-hybridized both oligonucleotide probes and were plaque purified. The inserts of these two clones were isolated and subjected to restriction mapping and nucleotide sequence analysis. An 8-kg BamHI fragment of the 15-kb insert, isolated from one of these clones (HC7), contained the entire gene. A 5.0-kb EcoRI fragment of the 16-kb insert of the other clone (HC10) was truncated at the 3' end such that it contained all four intervening sequences but only four of five coding blocks.

The nucleotide sequence of the entire human heart chymase gene is shown in Fig. 3. The gene is approximately 3 kb in length and has five coding blocks and four intervening sequences. The 5'-untranslated region contains a CAAT and a TATA box (Figs. 3 and 4). A consensus polyadenylation motif, AATAAA, which is followed by a cleavage site motif, consisting of a CA sequence located 15 bp downstream along with a GT cluster (Birnstiel, 1985), is found in the 3'-untranslated region (Fig. 3). The first coding block is 58 bp in length; it encodes, in an open reading frame, the first 19 amino acids (–21 to –3) of the preprochymase. A short first coding block similar to the above is present in the genes of other serine proteinases including human cathepsin G (Hohn et al., 1989), human neutrophil elastase (Takahashi et al., 1988), RMCP II (Benfey et al., 1987), and human cytotoxic cell protease (Meier et al., 1990). The second, third, fourth, and fifth coding blocks of the human heart chymase gene encode amino acids –2 to 49, 50–94, 95–179, and 180–226, and are 151, 136, 255, and 141 bp in length, respectively. The five coding blocks are separated by four intervening sequences of 672, 742, 186, and 368 bp. A single in-frame stop codon is present at the end of the fifth coding block.

The overall organization of the human heart chymase gene is similar to that of several other serine proteinases. It is likely, therefore, that these genes are all derived from a single ancestral gene (Fig. 5). These proteinases all have their coding regions located on five coding blocks separated by four intervening sequences, with active-site histidine, aspartic acid, and serine residues located on the second, third, and fifth coding blocks, respectively. The relative positions of these residues within these coding blocks are highly conserved among the chymase subfamily (Fig. 5). A slightly longer second intervening sequence is observed in the human heart chymase gene compared to the sequences in the genes of other serine proteinases. All of the intervening sequences have consensus donor and acceptor splice sites (Breathnach and Chambon, 1981), and all contain potential lariat acceptor sites at a location upstream from the 3' acceptor splice sites (Rogers and Maniatis, 1985). The splicing phases at the border of each intervening sequence are conserved for all chymases including human heart chymase (Fig. 5). For all of these genes, the intervening sequence splice phase is as follows: 1) type I for the first intervening sequence boundary (the intron interrupts the first and second bases of the codon); 2) type II for the second intervening sequence boundary (the intervening block interrupts the second and third bases of the codon); and 3) type 0 for intervening sequences III and IV (these intervening sequences occur between codons) (Rogers, 1985).

The 5'-Untranslated Region of Human Heart Chymase—Sequences in the 5'-untranslated region of the human heart chymase gene contain a TATA box (Corden et al., 1989) at position 201–206 and a CAAT box (Myers et al., 1986) located at position 120–124 (Fig. 3). The 5'-untranslated region from position 1–258 does not contain consenus elements for the binding of AP-1, AP-2, AP-3, AP-4, or consensus responsive elements for heat shock, glucocorticoids, cadmium, or serum (Lewin, 1990). A comparison of the 5' region of the gene with the same location of the RMCP II gene reveals an homologous sequence (CAGTTCCTGTGTTT) at positions 154–166 (Benfey et al., 1987; Sarid et al., 1989) (Fig. 3). This region is homologous to a sequence which confers pancreas-specific gene expression (Boulet et al., 1986). However, the 5' and 3'-untranslated regions of the human heart chymase gene do not contain mast cell-specific DNA sequences (Avraham et al., 1989). Two conserved motifs are also found in the 5'-untranslated regions of both the human heart chymase and the cathepsin G genes (Hohn et al., 1989). These motifs are: CCCTTCTAG (in human heart chymase, 68–76), CCCTTCTAG (in cathepsin G, –106), and CAGCCTTG (in human heart chymase 170–177, cathepsin G, –124). The existence of an enhancer sequence homologous with those in the 5'-untranslated regions of the RMCP II and the cathepsin G genes suggest that a similar enhancer mechanism may participate in the expression of the human heart chymase gene. Interestingly, a sequence (GGGAACCT TC) that is partially homologous to the \(\beta\)-binding site (GGGACTTTC) is found in the 5'-untranslated region of the human heart chymase gene. Binding of the nuclear protein factor \(\kappa\)-B, by the \(\beta\)-binding site promotes the transcription of the immunoglobulin light chain genes by B lymphocytes (Sen and Baltimore, 1986). Furthermore, this putative \(\beta\)-binding site in the human heart chymase gene is located immediately 5' to a sequence that is partially homologous (TGA_TCA) to the phorbol ester reactive element (TGACTCA). Interestingly, phorbol esters, which activate protein kinase C, induce \(\kappa\)-B synthesis, not only in B lymphocytes but also in human T cell lines and in HeLa cells. Thus, the co-localization of a putative \(\beta\)-binding site and a phorbol ester reactive element suggests that a protein kinase C mechanism may be involved in the control of human heart chymase gene expression. Finally, the 5'-untranslated region of the human heart chymase gene contains a sequence (CCTCTCT) that is partially homologous to the putative ribosome-binding site (CCTTCCG) (Hagenbuchle et al., 1978).

Southern Blot Analysis of Human Genomic DNA—To investigate whether human chymase is encoded by one or more genes, genomic DNA isolated from the leukocytes of a normal healthy donor was examined by Southern blot analysis. Three
distinct hybridizing species are apparent in a SacI digest and one in an EcoRI digest (data not shown). Since the gene contains two SacI sites and no EcoRI sites, this suggests that, unlike rodent chymases (Bennet et al., 1987; Serafin et al., 1990, 1991), there is a single gene for human heart chymase. This conclusion is also supported by the fact that no additional bands were observed, even under lower stringency.

cDNA Cloning

The cDNA for human heart chymase was obtained using PCR. Based on the DNA sequence of the gene, two specific primers corresponding to the sense strand immediately 5' to the ATG start site (+244 to +267 bp) and to the antisense strand immediately 3' to the stop codon (+2980 to +3056 bp) (Fig. 3) were synthesized and used to amplify single-stranded cDNA derived from human heart poly(A)+ mRNA. A PCR product of the correct predicted length (769 bp) was obtained and isolated after fractionation of the PCR mixture by electrophoresis. Digestion of the 769- bp PCR product with BamHI resulted in two fragments, both of which were of the correct predicted size (±530 and 240 bp). Similarly, digestion with HindIII resulted in fragments of approximately 490 and 280 bp in length. The entire PCR product was subcloned into pBS II KS and subjected to nucleotide sequencing.

The nucleotide sequence from the cDNA and the genomic DNA, as well as the deduced amino acid sequence of human heart chymase, are shown in Fig. 2. The nucleotide sequence of the cDNA contained a single ATG codon at the 5' end followed in an open reading frame by 741 bp encoding 247 amino acids. The amino acid sequences of all nine tryptic fragments were found in-frame in this open reading frame of the cDNA.

**Deduced Amino Acid Sequence of Human Heart Chymase**

*Signal Peptide and Prosequence*—Selection of the ATG nearest the 5' end of the open reading frame as the site of the initial methionine predicts that a 21-residue prepropeptide precedes the recognized N terminus of the mature active enzyme (Urata et al., 1990b). The first 19 residues of preprohuman heart chymase are hydrophobic and likely represent a signal peptide (von Heijne, 1986). This hydrophobic presequence is similar in size to signal peptides predicted for other mammalian chymases and neutrophil chymotrypsin-like proteinases (Fig. 6). It has been proposed that the sequence Ala-Xaa-Ala predicts the signal peptide cleavage site (Carne and Scheele, 1985). Since this sequence occurs at the end of the proposed signal peptide in human heart chymase, we believe that the human heart chymase signal peptide is cleaved at the Ala3-Gly2 bond. The prossequences of leukocyte chymotrypsin-like proteinases are a pair of acidic residues, usually Glu-Glu (for Refs., see Fig. 6). However, human heart chymase and human neutrophil cathepsin G contain a Gly-Glu sequence at this position (Fig. 6). The similarity of the prossequences in these human leukocyte enzymes suggests a common enzymatic activation step.

*Glycosylation Site*—After the N-terminal Gly-Glu is cleaved, the mature human heart chymase is predicted to be 226 residues in length. The calculated mass of the mature enzyme is 25,000 daltons. The difference between this calculated mass and the M, of the purified enzyme (±30,000) (Urata et al., 1990b) may be due to glycosylation. Human heart chymase contains two consensus N-linked glycosylation sites at Asn61 and Asn66 (Fig. 6). The asparagine at position 59 is almost definitely glycosylated, since a blank cycle was obtained in sequencing the tryptic peptide (T-12) containing this residue (Figs. 2 and 6). The further observation that purified human heart chymase binds to concanavalin A-Sepharose and wheat germ agglutinin-Sepharose (Urata et al., 1990b) supports the possibility that human heart chymase is a glycoprotein. The location of these N-linked glycosylation sites are different from those in dog chymase. Only a single consensus N-linked glycosylation site occurs at Asn61 in mouse mast cell protease I, whereas RMCP I and II contain no glycosylation sites (Fig. 6).

*Active Site and the Extended Substrate-binding Site*—The primary structure of human heart chymase is homologous to those of the other members of the leukocyte chymotrypsin-like serine proteinase family: ≈80% homology with dog mastocytes cell chymase, ≈60% with rodent chymases, and ≈50% with human cathepsin G and cytotoxic cell proteinases (Fig. 6). Residues essential for the catalytic activity of these serine proteinases, His44, Asp48, and Ser199, are conserved in all members, including human heart chymase.

There are several clusters of unique amino acid sequences in the primary structure of human heart chymase including: VTSGPSKF (residue 21–29), TEE (residue 61–69), TSTLH (residue 83–87), SHFRD (residue 154–158), and RS (residue 200–201) (Fig. 6). Interestingly, these unique sequences in human heart chymase are all contained in the regions putatively involved in substrate binding (Remington et al., 1990; Murphy et al., 1988) (Fig. 6). Residues 21–29, 62, 83–87, 154–158, and 200–201 of human heart chymase form a part of the S′, S′, S, S′, and S, substrate-binding sites, respectively. Other regions within human heart chymase form a part of the S, residues 176 and 207), S1 (residue 179), S′1 (residue 20), and Cys21-Cys25), S1 (residue 60), S1 (residue 128), and S1 (residue 157) substrate-binding site. These regions were conserved, compared to those of other chymases and the other chymotrypsin-like leukocyte serine proteinases (Table I).

Our recent studies have demonstrated an unusually high substrate specificity for human heart chymase, as compared to that of other chymases, for the conversion of Ang I to Ang II (Urata et al., 1990b). For example, the Phe2-Phe bond in the undecapeptide substance P and the Tyr2-Leu2 bond in the 28-amino acid peptide vasoactive intestinal peptide are readily hydrolyzed by dog mastocytesoma chymase (Caughley et al., 1988a). But these bonds are either not hydrolyzed or poorly hydrolyzed by human heart chymase (Urata et al., 1990b). More recently, using peptide analogs of Ang I, Kinoshita et al. (1991) have identified several unique features in the substrate specificity of human heart chymase. This study suggests that there are several unique determinants in the substrate for the high specificity of human heart chymase, including a Pro at the P, position and the presence of a dipeptide C-terminal leaving group containing no prolines. Both a P, Pro and a dipeptide C-terminal leaving group containing no prolines are present in Ang I. In RMCP II, residues 21–29 are involved in the formation of a deep cleft near the active site (Remington et al., 1988). This cleft seems to enable the P′ residues of the substrate to interact with the S′ substrate-binding sites of RMCP II. Residues 21–29 comprise the longest unique sequence in human heart chymase, a sequence that is markedly different from those in similar regions of other chymases with regard to charge, hydrophobicity, and side chain length. Thus, residues 21–29 of human heart chymase may be important contributors to the high substrate specificity of human heart chymase.

Six cysteines in the human heart chymase at positions 30, 46, 123, 154, 167, and 188 are all conserved with those in the leukocyte serine proteinase family (Fig. 6). Based on the RMCP II (Woodbury et al., 1978), the cysteines responsible for the formation of the intramolecular disulfide bond are the...
Cloning of Human Heart Chymase

**Fig. 6.** Comparison of the amino acid sequence of human heart chymase with those of other leukocyte chymotrypsin-like proteinases. The amino acid sequence of human heart chymase (hHC) is compared with that of dog mastocytoma cell chymase (dMC) (Caughey et al., 1990), mouse mast cell protease I (mMCP1) (Lobe et al., 1986), mouse mast cell protease II (mMCP2) (Serafin et al., 1990), rat mast cell protease I (rMCP1) (Le Trong et al., 1987), rat mast cell protease II (rMCP2) (Benfey et al., 1987), human neutrophil cathepsin G (hCaG) (Salvesen et al., 1987), and human cytotoxic cell protease I (hCPI) (Meier et al., 1990). Human heart chymase was aligned manually with other chymases and leukocyte chymotrypsin-like proteinases to optimize sequence homology. Aligned residues identical with those of human heart chymase are boxed. The triad of amino acids involved in catalysis by serine proteinase, His45, Asp172, and Ser21, is indicated by the open triangle. The consensus sites for N-linked glycosylation in human heart chymase are indicated by open diamonds. Cysteines that are not conserved among the aligned proteinases are indicated by asterisks.

**TABLE I**

**Important amino acids in the extended substrate-binding site of leukocyte chymotrypsin-like serine proteinases**

| Enzymes                     | S1  | S1' | S2   | S2' | S3   | S3' | S4   | S4' | S5   | S5' | S6   | S6' | S7   | S7' |
|-----------------------------|-----|-----|------|-----|------|-----|------|-----|------|-----|------|-----|------|-----|
| Human heart chymase         | Lys | Ser | Leu  | Ser | Ala  | Ile | Lys  | Tyr | Phe  | Ile | Trp  | Phe |
| Dog chymase                 | Lys | Asn | Leu  | Ser | Ala  | Ile | Ala  | Tyr | Phe  | Ile | Trp  | Ile |
| RMCP1                       | Lys | Gly | Asn  | Ser | Ala  | Ile | Ala  | His | Ser  | Ile | Tyr  | Ile |
| RMCP1F                      | Met | Pro | Asn  | Ser | Ala  | Ile | Val  | Cys | His  | Ile | Val  | Tyr |
| Human neutrophil cathepsin G | Lys | Ser | Ile  | Ala | Glu  | Ile | Cys  | Tyr | Arg  | Ile | Trp  | Ile |

Following: Cys10-Cys66, Cys122-Cys188, and Cys14-Cys167. All chymases, including human heart chymase, lack the disulfide bond Cys10-Cys188 (chymotrypsin numbering) which is present in α-chymotrypsin. As predicted from the three-dimensional structure of RMCP II, the deletion of this disulfide bond leads to the formation of an extended binding site that enhances the substrate specificity of RMCP II (Remington et al., 1988). In addition to these cysteine-forming intramolecular disulfide bonds, another cysteine at position 7 is contained in the mature form of human heart chymase. It is not known, however, whether Cys7 is involved in forming an intramolecular disulfide bond, or whether it remains free. Studies of the purified Hannuka factor (Pasternak et al., 1986) indicate that homodimers of the Hannuka factor may develop through the formation of intermolecular disulfide bridges at Cys7 (chymotrypsin numbering). Similarly, we have observed homodimer formation in human heart chymase after its overnight incubation at a basic pH.

**Evolutionary Relations**

Since insertions and deletions (gaps) of residues in the sequences of homologous proteins occur less frequently than amino acid substitutions (Dayhoff et al., 1972), the evolutionary relation between homologous proteins may be more reliably estimated from the number and position of inferred gaps.

---

1. H. Urata and A. Husain, unpublished observations.
arising when the sequences are optimally aligned (De Haen et al., 1975). Based on this criterion, the human heart chymase gene may have arisen sometime after the gene for trypsin, but before those for elastase and chymotrypsin (Woodbury et al., 1978). Thus, the human heart chymase gene is closely related to those of other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases, such as cathepsin G and cytotoxic cell protease (Fig. 6). A comparison of differences in their genomic organization. The human heart chymase, probably a single gene product, was obtained from the human heart chymase, several of which reside in the mammalian chymases appear to have a broad substrate specificity (Rogers, 1985). In addition, the nature and location of the intervening sequence splice phases are identical in all chymotrypsin-like leukocyte serine proteinase genes, including that of human heart chymase, but are different from those found in the trypsin and the chymotrypsin gene (Reilly et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings show that the human heart chymase is more closely related to other mammalian chymases than to other types of human leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases.

Summary and Conclusions

Human heart chymase is the most efficient specific Ang II-forming enzyme described (Urata et al., 1990). Other mammalian chymases appear to have a broad substrate specificity (Reilly et al., 1985; Caughey et al., 1988a; Le Trong et al., 1987). The complete primary structure of human heart chymase, probably a single gene product, was obtained from the cloning of the gene and cDNA. The high degree of homology between the amino acid sequence and the gene structure of human heart chymase with that of other mammalian chymases suggests that these enzymes all evolved from a common ancestral gene. It is tempting to suggest that unique structures in human heart chymase, several of which reside in the extended substrate-binding site, are responsible for its high degree of specificity for Ang II formation.

Acknowledgments—We are most grateful to Dr. Philip Leder, Harvard Medical School, for the kind gift of the RMCP I and trypsin gene but distinct from that found in the chymotrypsin gene (Rogers, 1985). In addition, the nature and location of the intervening sequence splice phases are identical in all chymotrypsin-like leukocyte serine proteinase genes, including that of human heart chymase, but are different from those found in the trypsin and the chymotrypsin gene (Fig. 5). These results support the hypothesis that in its evolution human heart chymase is more closely related to other mammalian chymases than to other types of human leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases. These findings note that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987). This hypothesis is further supported by the findings produced by a comparison of the mammalian structure of human heart chymase with that of other serine proteinases. These findings indicate that the highest homology occurs with dog mastocytoma chymase (80%), less with the rodent chymases (60%), even less with other chymases but is distinct from those of other chymotrypsin-like leukocyte serine proteinases such as trypsin (Vandensel et al., 1990), elastase (Takahashi et al., 1988), and cathepsin G (Salvesen et al., 1987).