Proc. Natl. Acad. Sci. USA
Vol. 77, No. 6, pp. 3388–3392, June 1980
Biochemistry

Covalent structure of human haptoglobin: A serine protease homolog

(amino acid sequence/plasminogen/prothrombin/protein evolution)

ALEXANDER KUROSKY*, DON R. BARNETT*, TONG-HO LEE*†, BILLY TOUCHSTONE*, REGINE E. HAY*, MARILYN S. ARNOTT*‡, BARBARA H. BOWMAN*, AND WALTER M. FITCH§

*Division of Human Genetics, Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston, Texas 77550; and §Department of Physiological Chemistry, University of Wisconsin Medical Center, Madison, Wisconsin 53706

Communicated by Alexander G. Bearn, February 11, 1980

ABSTRACT **The complete amino acid sequences and the disulfide arrangements of the two chains of human haptoglobin 1-1 were established. The α^1 and β chains of haptoglobin contain 83 and 245 residues, respectively. Comparison of the primary structure of haptoglobin with that of the chymotrypsinogen family of serine proteases revealed a significant degree of chemical similarity. The probability was less than 10^{-5} that the chemical similarity of the β chain of haptoglobin to the proteases was due to chance. The amino acid sequence of the β chain of haptoglobin is 29–33% identical to bovine trypsin, bovine chymotrypsin, porcine elastase, human thrombin, or human plasmin. Comparison of haptoglobin α^1 chain to activation peptide regions of the zymogens revealed an identity of 25% to the fifth "kringle" region of the activation peptide of plasminogen. The probability was less than 0.014 that this similarity was due to chance. These results strongly indicate haptoglobin to be a homolog of the chymotrypsinogen family of serine proteases. Alignment of the β-chain sequence of haptoglobin to the serine proteases is remarkably consistent except for an insertion of 16 residues in the region corresponding to the methionyl loop of the serine proteases. The active-site residues typical of the serine proteases, histidine-57 and serine-195, are replaced in haptoglobin by lysine and alanine, respectively; however, aspartic acid-102 and the trypsin specificity residue, aspartic acid-189, do occur in haptoglobin. Haptoglobin and the serine proteases represent a striking example of homologous proteins with different biological functions.**

Human haptoglobin (Hp) is a plasma glycoprotein composed of two types of polypeptide chains, α and β, that are covalently associated by disulfide bonds. In humans, three major phenotypes, designated Hp 1-1, Hp 2-1, and Hp 2-2 (1–3), involve considerable molecular variation due to an almost complete duplication of the Hp^1 allele that resulted in a Hp^2 allele (4, 5). During gel electrophoresis Hp 2-1 and 2-2 both exhibit a polymeric series mediated by disulfide bonding (6–9) that can be formulated as $(\alpha^2\beta)_n$ (n = 3, 4, 5, . . .) for Hp 2-2 and $(\alpha^1\beta)_2{\cdot}(\alpha^2\beta)_n$ (n = 0, 1, 2, . . .) for Hp 2-1 (10–13). The structure of Hp 1-1 is represented simply as $(\alpha^1\beta)_2$. Comprehensive reviews on Hp are available (14–16).

Interest in Hp is usually related to its unique ability to bind hemoglobin (Hb). An $\alpha\beta$ subunit of Hb binds to each of two independent binding sites on a molecule of Hp 1-1 with an affinity that is too high to measure precisely (17). Binding of Hb $\alpha\beta$ subunits is reported to occur on the β chain of Hp (18).

Partial amino acid sequence analysis of the haptoglobin β polypeptide (hp β) chain indicated it to be chemically similar to the chymotrypsinogen family of serine proteases (19, 20). This suggests that Hp is evolutionarily related to the serine proteases as a result of gene duplication of a serine protease gene (or a precursor thereof) and subsequent divergence to its present structure. This report presents the complete primary structure of human Hp 1-1 (α^1 and β chains), including the disulfide arrangements. We have repeated the sequence analysis of human hp α^1 chain reported by others (21, 22) because our analysis revealed a number of differences in the amino-terminal region. The completed α^1 and β chain sequences of Hp were extensively compared with the chymotrypsinogen family of serine proteases to establish their evolutionary relationship.

METHODS

Hp and its component chains were purified from human ascites fluid and from plasma as described (23). Purification and characterization of the CNBr peptides were reported by Kurosky *et al.* (24). Purified CNBr fragments II, III, IV, and V were subjected to hydrolysis with chymotrypsin, trypsin, staphylococcal protease, and thermolysin. Tryptic and chymotryptic hydrolyses were also performed on whole β chain. The primary structure of the α^1 chain of Hp was established by automated sequence analysis of intact α^1 chain and by characterization of tryptic and staphylococcal protease peptides. In addition, fragments obtained by arginyl hydrolysis by trypsin of succinylated α^1 chain and by prolonged treatment of α^1 chain with CNBr (1000-fold molar excess for 72 hr) were also characterized. Methods of peptide purification and dansyl(5-dimethylaminonaphthalene-1-sulfonyl)-Edman were conducted as described (25). Automated sequence analyses were performed on the Beckman Sequencer (updated model 890B) as reported (23). Limited proteolytic cleavage of Hp 1-1 by plasmin was done as described (26).

Comparison of the primary structures of the α^1 and β chains of Hp with the chymotrypsinogen family of serine proteases was according to the procedures of Fitch (27, 28) or Needleman and Wunsch (29).

RESULTS

The amino acid sequences of the α^1 and β chains of human Hp are given in Fig. 1. Limited proteolytic cleavage of intact human Hp 1-1 by human plasmin gave a major cleavage between residues 130 and 131 (90% yield). A minor cleavage between 161 and 162 (10% yield) was also elicited with use of higher ratios of plasmin to Hp. The disulfide linkages in Hp 1-1 are described in the legend to Fig. 1. A summary representation of the tetrachain arrangement of human Hp 1-1, including the disulfide assignments, is illustrated in Fig. 2.

Comparison of the sequence of the hp β chain with sequences of representative serine proteases from the chymotrypsinogen family is shown in Fig. 3. Computer analysis by the method of Fitch (27, 28) revealed that the probability was $<10^{-5}$ that the chemical similarity between the β chain of Hp and the plas-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: Hp, intact haptoglobin; hp, used in references to haptoglobin polypeptide chains; Hb, hemoglobin.

† Present address: Department of Biochemistry, Catholic Medical College, Seoul, Korea.

‡ Present address: Department of Biology, Environmental Biology Section, M.D. Anderson, Houston, Texas 77030.

```
                   1         5           10          15          20          25
 α1 CHAIN:  ·Val-Asp-Ser-Gly-Asn-Asp-Val-Thr-Asp-Ile-Ala-Asp-Asp-Gly-Cys-Pro-Lys-Pro-Pro-Glu-Ile-Ala-His-Gly-Tyr-
26                   31          35          40          45          50     -Lys-    55
Val-Glu-His-Ser-Val-Arg-Tyr-Gln-Cys-Lys-Asn-Tyr-Tyr-Lys-Leu-Arg-Thr-Glu-Gly-Asp-Gly-Val-Tyr-Thr-Leu-Asn-Asn-Glu-Lys-Gln-
56                   61          65          70          75          80
Trp-Ile-Asn-Lys-Ala-Val-Gly-Asp-Lys-Leu-Pro-Glu-Cys-Glu-Ala-Val-Cys-Gly-Lys-Pro-Lys-Asn-Pro-Ala-Asn-Pro-Val-Gln·

                   1         5           10          15          20      CHO     25
 β CHAIN:    Ile-Leu-Gly-Gly-His-Leu-Asp-Ala-Lys-Gly-Ser-Phe-Pro-Trp-Gln-Ala-Lys-Met-Val-Ser-His-His-Asn-Leu-Thr-
26                   31          35          40          45 CHO      CHO               55
Thr-Gly-Ala-Thr-Leu-Ile-Asn-Glu-Gln-Trp-Leu-Leu-Thr-Thr-Ala-Lys-Asn-Leu-Phe-Leu-Asn-His-Ser-Glu-Asn-Ala-Thr-Ala-Lys-Asp-
56                   61          65          70          75          CHO               85
Ile-Ala-Pro-Thr-Leu-Thr-Leu-Tyr-Val-Gly-Lys-Lys-Gln-Leu-Val-Glu-Ile-Glu-Lys-Val-Val-Leu-His-Pro-Asn-Tyr-Ser-Gln-Val-Asp-
86                   91          95          100         105         110               115
Ile-Gly-Leu-Ile-Lys-Leu-Lys-Gln-Lys-Val-Ser-Val-Asn-Glu-Arg-Val-Met-Pro-Ile-Cys-Leu-Pro-Ser-Lys-Asp-Tyr-Ala-Glu-Val-Gly-
116                  121         125         130         135         140               145
Arg-Val-Gly-Tyr-Val-Ser-Gly-Trp-Gly-Arg-Asn-Ala-Asn-Phe-Lys-Phe-Thr-Asp-His-Leu-Lys-Tyr-Val-Met-Leu-Pro-Val-Ala-Asp-Gln-
146                  151         155         160         165         170               175
Asp-Gln-Cys-Ile-Arg-His-Tyr-Glu-Gly-Ser-Thr-Val-Pro-Glu-Lys-Lys-Thr-Pro-Lys-Ser-Pro-Val-Gly-Val-Gln-Pro-Ile-Leu-Asn-Glu-
176                  181         185         190         195         200               205
His-Thr-Phe-Cys-Ala-Gly-Met-Ser-Lys-Tyr-Gln-Glu-Asp-Thr-Cys-Tyr-Gly-Asp-Ala-Gly-Ser-Ala-Phe-Ala-Val-His-Asp-Leu-Glu-Glu-
206                  211         215         220         225         230               235
Asn-Thr-Trp-Tyr-Ala-Thr-Gly-Ile-Leu-Ser-Phe-Asp-Lys-Cys-Ser-Ala-Val-Ala-Glu-Tyr-Gly-Val-Tyr-Val-Lys-Val-Thr-Ser-Ile-Gln-
236                  241         245
Asn-Trp-Val-Gln-Lys-Thr-Ile-Ala-Glu-Asn·
```

FIG. 1. Amino acid sequences of the α^1 and β chains of human Hp 1-1. Carbohydrate attachment is indicated by CHO. The interchain disulfide linkages are $\alpha^1$15-$\alpha^1$15 and $\alpha^1$72-β105 and the intrachain linkages are $\alpha^1$34-$\alpha^1$68, β148-β179, and β190-β219 (8, 9, 24). A polymorphism of the α^1 chain has either Glu or Lys at position 53 (5).

minogen B chain was due to chance (length of subsequence examined = 15, average minimal base difference = 1.52, and number of sequences compared was >12,750). A similar calculation of the comparison of hp β chain and thrombin B chain gave a probability of $<10^{-8}$. Similarly, comparison of hp α^1 chain with activation peptides of serine proteases is given in Fig. 4. Computer analysis according to the method of Needleman and Wunsch (29), using maximum base matches from the genetic code and a gap penalty of 1.999, identified a significant degree of similarity between hp α^1 chain and the fifth kringle[¶] region of plasminogen ($P < 0.014$ that the similarity was due to chance). The relationship of Hp to the chymotrypsinogen family of serine proteases is represented by the evolutionary tree given in Fig. 5.

DISCUSSION

Amino acid sequence analysis of hp α^1 chain revealed a number of differences when compared to the previously reported sequence (21, 22). Our results indicated one less residue; asparagine at position 2 was not confirmed. Moreover, we have determined residues 15–20 to be Cys-Pro-Lys-Pro-Pro-Glu rather than Gln-Pro-Pro-Pro-Lys-Cys as previously assigned. In addition, we find position 43 to be Glu rather than Gln. Repeated sequence analysis using α^1 chain alkylated with iodo[1-^{14}C]-acetamide clearly identified radiolabeled Cys at position 15 rather than at 20. These same sequence differences were also confirmed in partial sequence analysis of the hp α^2 chain.

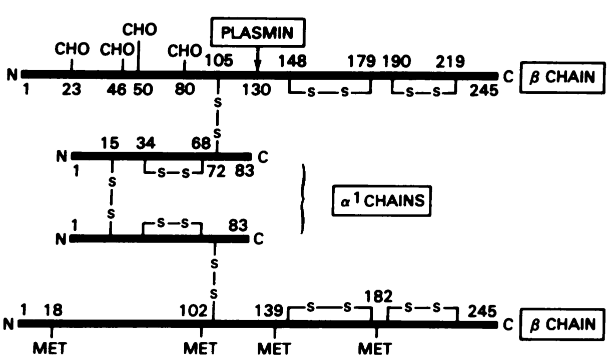

FIG. 2. Summary representation of the major features of the human Hp 1-1 tetrachain structure. CHO indicates carbohydrate attachment to Asn. Plasmin cleaved at Lys-130 of both β chains.

Our sequence studies establish a molecular weight of 9189 for hp α^1 chain and 27,259 for the peptide portion of hp β chain. Because the carbohydrate content of the β chain is 19.4% (38), the total molecular weights of the β chain and intact Hp 1-1 were calculated to be 33,820 and 86,018, respectively. These molecular weight values fall within the ranges reported by ultracentrifugational analysis (reviewed in ref. 24).

Comparison of the hp β chain to the chymotrypsinogen family of serine proteases (Fig. 3) strongly indicates Hp to be a homolog of this family of proteins. The probability is $<10^{-5}$ that the similarity is due to chance. The hp α^1 chain also shows significant chemical similarity to the activation peptides of the serine proteases and especially to the fifth kringle region of plasminogen. The probability is <0.014 that this similarity is due to chance. A charge relay system, as has been described for the serine proteases, could not occur in Hp because the positions corresponding to the proteolytic active site residues, histidine-57 and serine-195, are replaced in Hp by lysine and alanine, respectively. This is consistent with the fact that Hp has no known proteolytic function. Aspartic acid-102 and the trypsin specificity residue aspartic acid-189 are found in Hp as well as aspartic acid-194, which forms an internal ion pair with the α-amino group of the amino-terminal residue in a number of serine proteases (34). Of the three intrachain disulfide loops common to all serine proteases (histidyl, 42–58; methionyl, 168–182; and seryl, 191–220), only the methionyl and the seryl loops are found in the hp β chain. However, the interchain disulfide between the attached activation peptide portions and the enzyme portions in the serine proteases so far characterized aligns precisely with the disulfide attaching hp α and β chains. Taken together, these data further support our earlier proposal (39) that Hp is initially synthesized as a single chain and subsequently cleaved. Typically in eukaryotic multicellular organisms, the serine proteases are synthesized as single-chain precursors and subsequently activated by proteolysis. In some cases, as for example Factor XI (40), the precursor chain structure is further elaborated by disulfide bonding. Strikingly, the tetrachain chain structure of Factor XIa, which is a serine protease, is identical to that of Hp and, like Hp, has two functional sites per tetrachain. It is highly possible that the precursor structure of Hp resembles Factor XI—i.e., two identical chains covalently attached by disulfide bonding. This hypothesis does

¶ Defined in ref. 37.

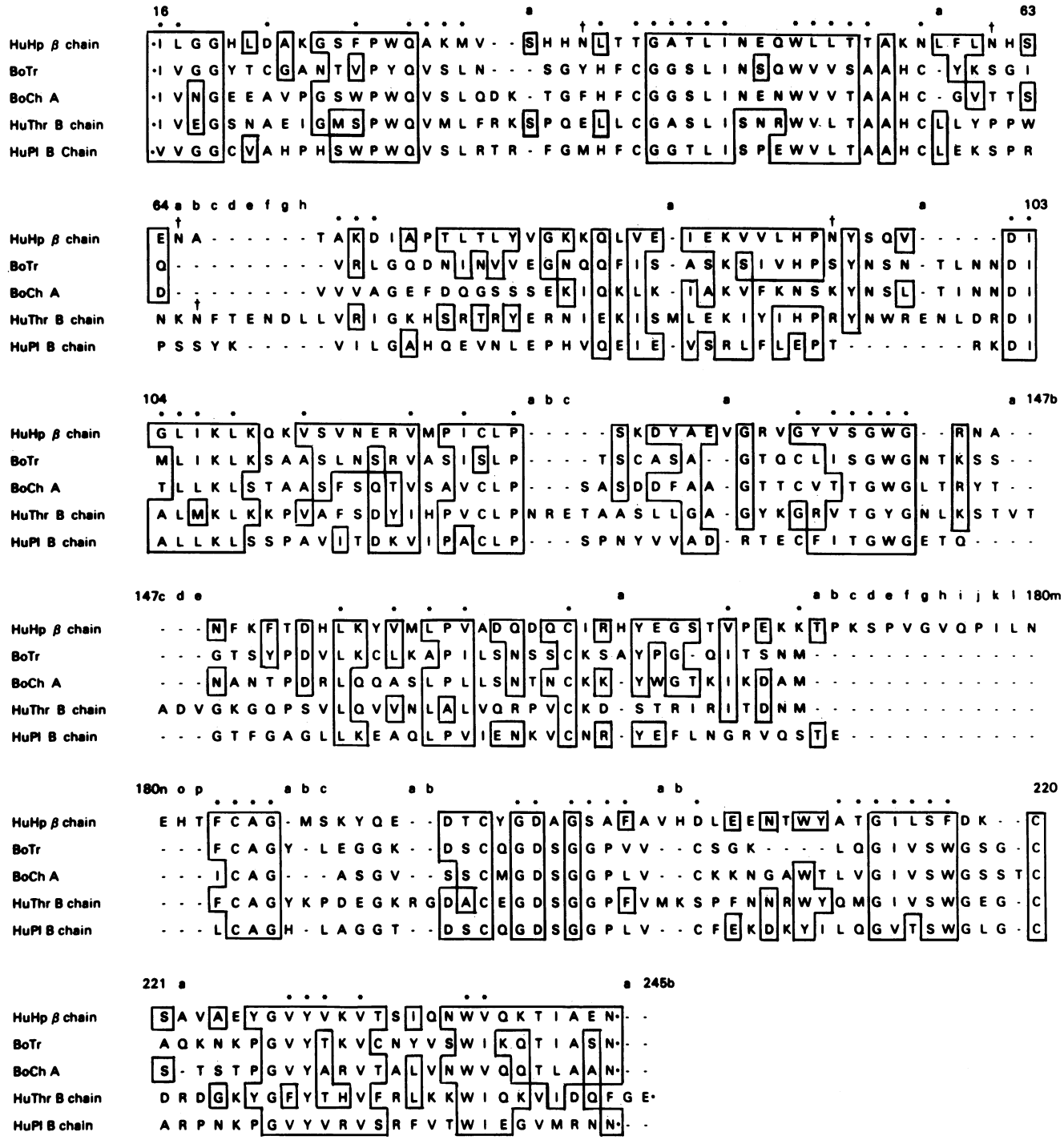

FIG. 3. Comparison of the β chain of human Hp (HuHp) to bovine trypsin (BoTr) (30), bovine chymotrypsin A (BoCh A) (31), human thrombin B chain (HuThr B) (32), and human plasmin B chain (HuPl B) (33). Numbering is that of bovine chymotrypsinogen A with insertions indicated by letters as 62A, 62B, etc. Residues in the serine proteases that are chemically similar to residues in hp β chain are boxed in. Single-letter amino acid abbreviations are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr. Chemically similar residues are: I = L = V, G = A, T = S, Y = F = W, D = E, D = N, Q = N, E = Q, and K = R. *Residues determined to be internal and inaccessible to water molecules in the three-dimensional structures of α-chymotrypsin, trypsin, and elastase (34). †Covalent attachment of carbohydrate.

not agree with the conclusions from a single family study of Hp Bellevue by Javid (41), which suggested that the genes coding for the α and β chains are not linked. Javid's results, however, can also be explained by other genetic mechanisms such as crossing-over.

Homology of Hp to the serine proteases is also supported by comparative model building of Hp to known three-dimensional structures of the serine proteases. Computer fitting of the primary structure of the hp β chain to the three-dimensional structures of trypsin, elastase, and α-chymotrypsin by Greer (42) was in excellent agreement and was significantly better than a similar comparison between microbial and mammalian serine proteases.

Comparison of the hp β chain to various serine proteases (Fig. 3) revealed a fairly consistent degree of identity (29–33%) with no clear indication of a preferred comparison. Comparison of the hp α^1 chain to activation peptides of the serine proteases demonstrated the greatest degree of identity (excluding the

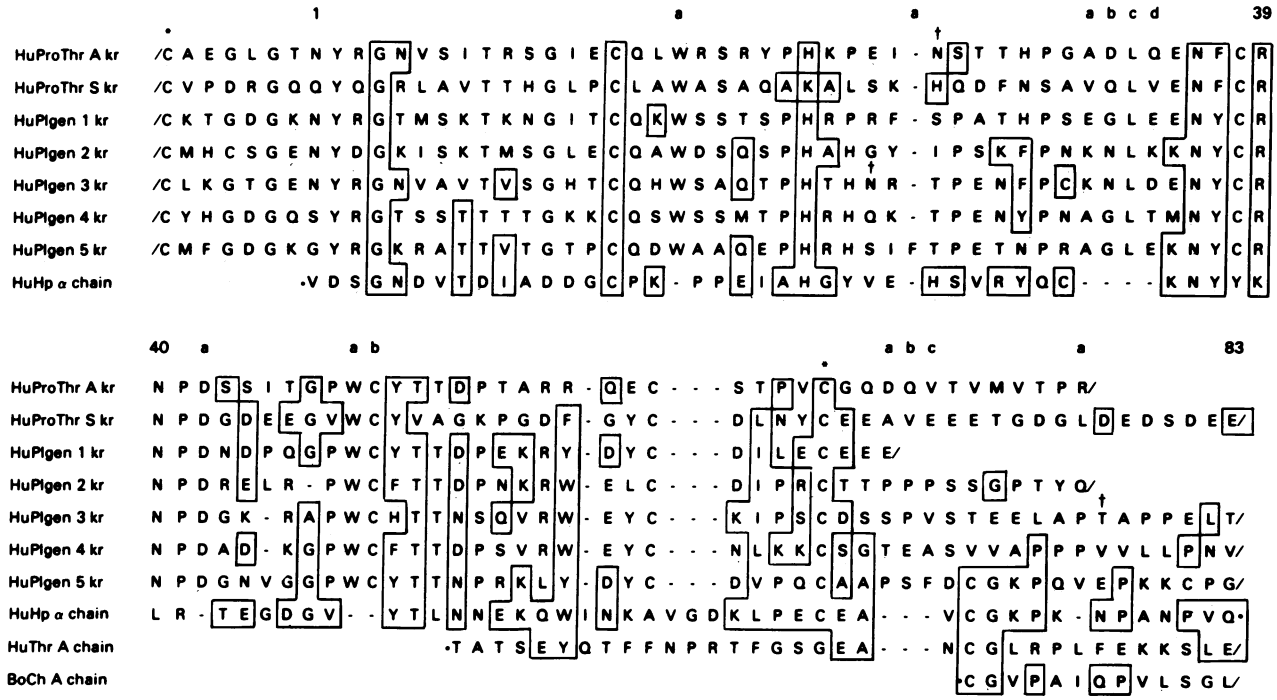

FIG. 4. Comparison of the α^1 chain of human Hp to the activation peptide regions of human prothrombin (HuProThr) (35), human plasminogen (HuPlgen) (36), and bovine chymotrypsinogen A (BoCh A) (31). A kr and S kr are the kringle regions of prothrombin and 1 kr–5 kr are the five kringle regions of plasminogen (36). Residues in the activation peptide regions chemically similar to residues in hp α^1 chain are boxed (see Fig. 3). The numbering is that of human hp α^1 chain with insertions indicated by letters. *Demarcates the actual kringle structures (see ref. 37). †Covalent attachment of carbohydrate.

short A chain of chymotrypsin) with the fifth kringle region of plasminogen (25% identity). This is interesting in view of the fact that, like Hp, plasminogen and prothrombin both reflect partial gene reiteration in their activation peptide region. If one assumes a single-chain structure for Hp, the duplication of human hp α^1 to give rise to α^2 is similar to the partial gene

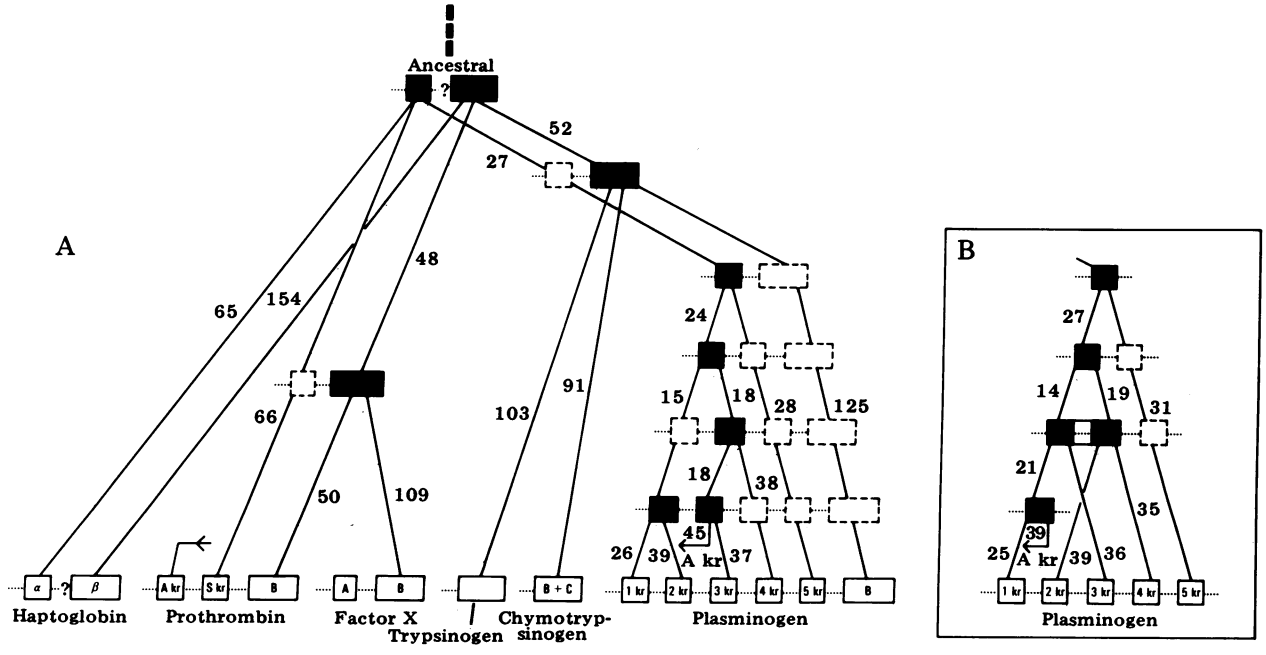

FIG. 5. Two most parsimonious trees (*A* and *B*) relating Hp to the chymotrypsinogen family of serine proteases. Only the kringle region of plasminogen is shown for tree *B* because the trees were otherwise virtually identical. The sequences comprising the kringle regions are as defined in Fig. 4 and are represented by the smaller boxes. The enzyme regions are the sequences given in Fig. 3 and are represented by the larger boxes. Intervening and additional sequences are denoted by a dashed line. Empty solid boxes represent the sequences used for computation. Filled boxes are ancestral elements that undergo a partial gene duplication at that point. The numbers on the legs are the minimal number of nucleotide substitutions during the interval between two solid boxes (empty or filled). Because they are uncorrected for multiple substitutions at the same site, no estimates of time or rate are warranted. Accordingly, node positions are arbitrary. Question mark denotes that the presence of a linkage between two gene fragments is uncertain. Tree *A* required a total of 1222 nucleotide substitutions; tree *B* required 1221. Both trees required 776 substitutions in the enzyme region.

duplication that resulted in the kringle regions in both prothrombin and plasminogen. Moreover, there is also evidence of a triplication of hp α chain (Hp Johnson) in the human population (4, 43).

Two evolutionary trees (*A* and *B*) relating Hp to the serine proteases based upon computer comparison of nucleotide base sequences (Fig. 5) reveal several interesting associations. The most parsimonious phylogenies for the activation peptide region (Fig. 4) and the enzyme region (Fig. 3) were constructed separately, principally because not all sequences have portions belonging to each region, to determine whether the two resultant best trees are mutually consistent. With one exception to be discussed later, the two trees were consistent, and they are presented jointly in Fig. 5. The order of divergence of trypsinogen and chymotrypsinogen from the plasminogen line is not shown because the orders differed by only one nucleotide substitution; however, this tree presents the possibility that the divergence of trypsinogen and chymotrypsinogen (also proelastase) included a deletion of a kringle-like segment. Among the observations of particular note are the partial internal gene duplications of the kringle regions of plasminogen and prothrombin. Strikingly, the A kringle of prothrombin is much more similar to the kringles of plasminogen than to the S kringle. This suggests that the A kringle of prothrombin was derived from a plasminogen kringle long after both gene functions had been established. Finally, the phylogenetic consistency of the two regions examined (activation peptide region and enzyme region) gives further evidence that these sequences have been genetically linked for an extended period of their evolutionary history and suggests that this is also true for the α and β chains of Hp. The only exception to the mutual consistency of the trees for the two regions is kringle A of prothrombin as noted above.

The present evidence strongly indicates that the *Hp* gene resulted from a duplication of a serine protease precursor gene that subsequently diverged, resulting in a loss of proteolytic function and acquisition of a new function. Because Hp has been reported to occur in early vertebrates such as eel (44), it cannot be considered a new protein provided that the eel and human sequences are orthologous. Therefore, why should the structure of Hp have been so well conserved over such a large evolutionary time span? One obvious possibility is that the new function acquired by Hp involves a considerable portion of the molecule, resulting in a low substitution rate. Some evidence of a moderately low substitution rate for the hp β chain was reported (23). The strong binding of Hp to Hb could explain such a molecular constraint and thereby emphasizes Hb binding as an important Hp function.

We thank Dr. Jonathan Greer for allowing us to preview his manuscript. We are indebted to Linda Merryman, Horace D. Kelso, Fu-Mei Lo, Terry M. Ward, and Ronald Niece for excellent technical assistance. We thank Dr. Julian Smith (Department of Gynecology, M.D. Anderson Hospital, Houston, Texas) for ascites fluid. This work was supported by Grant HD 03321 from the National Institute of Child Health and Human Development, by Grant H-378 from the Robert A. Welch Foundation, by Grant CA 17701 from the National Cancer Institute, by the Harris and Eliza Kempner Fund, by the Burkitt Foundation, and by Grant DEB-7814197 from the National Science Foundation.

1. Smithies, O. (1955) *Biochem. J.* **61,** 629–641.
2. Smithies, O. & Walker, N. F. (1955) *Nature (London)* **176,** 1265–1266.
3. Smithies, O. & Walker, N. F. (1956) *Nature (London)* **178,** 694–695.
4. Smithies, O., Connell, G. E. & Dixon, G. H. (1962) *Nature (London)* **196,** 232–236.
5. Black, J. A. & Dixon, G. H. (1968) *Nature (London)* **218,** 736–741.
6. Smithies, O., Connell, G. E. & Dixon, G. H. (1962) *Am. J. Hum. Genet.* **14,** 14–21.
7. Bearn, A. G. & Franklin, E. C. (1958) *Science* **128,** 596–597.
8. Malchy, B. & Dixon, G. H. (1973) *Can. J. Biochem.* **51,** 249–264.
9. Malchy, B., Rorstad, O. & Dixon, G. H. (1973) *Can. J. Biochem.* **51,** 265–273.
10. Ogawa, A. & Kawamura, K. (1966) *Proc. Jpn. Acad.* **42,** 413–417.
11. Ogawa, A., Kagiyama, S., Kawamura, K. & Yanase, T. (1970) *Proc. Jpn. Acad.* **46,** 814–819.
12. Fuller, G. M., Rasco, M. A., McCombs, M. L., Barnett, D. R. & Bowman, B. H. (1973) *Biochemistry* **12,** 253–258.
13. Pastewka, J. V., Ness, A. T. & Peacock, A. C. (1975) *Biochim. Biophys. Acta* **386,** 530–537.
14. Sutton, H. E. (1970) in *Progress in Medical Genetics,* eds. Steinberg, A. G. & Bearn, A. G. (Grune & Stratton, New York), Vol. 7, pp. 163–216.
15. Putnam, F. W. (1975) in *The Plasma Proteins,* ed. Putnam, F. W. (Academic, New York), Vol. 2, 2nd Ed., pp. 1–50.
16. Giblett, E. R. (1974) in *Structure and Function of the Plasma Proteins,* ed. Allison, A. C. (Plenum, New York), Vol. 1, pp. 55–72.
17. Chiancone, E., Alfsen, A., Toppolo, C., Vecchini, P., Finazzi Agrò, A., Wyman, J. & Antonini, E. (1968) *J. Mol. Biol.* **34,** 347–356.
18. Gordon, S. & Bearn, A. G. (1966) *Proc. Soc. Exp. Biol. Med.* **121,** 846–850.
19. Barnett, D. R., Lee, T.-H. & Bowman, B. H. (1972) *Biochemistry* **11,** 1189–1194.
20. Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H. & Bowman, B. H. (1974) *Biochem. Genet.* **11,** 279–293.
21. Black, J. A. & Dixon, G. H. (1970) *Can. J. Biochem.* **48,** 133–146.
22. Malchy, B. & Dixon, G. H. (1973) *Can. J. Biochem.* **51,** 321–322.
23. Kurosky, A., Kim, H.-H. & Touchstone, B. (1976) *Comp. Biochem. Physiol.* **55B,** 453–459.
24. Kurosky, A., Hay, R. E., Kim. H.-H., Touchstone, B., Rasco, M. A. & Bowman, B. H. (1976) *Biochemistry* **15,** 5326–5336.
25. Kurosky, A. & Hofmann, T. (1976) *Can. J. Biochem.* **54,** 872–884.
26. Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H. & Bowman, B. H. (1974) in *Protides of the Biological Fluids, Proceedings of the Colloquium,* ed. Peeters, H. (Pergamon, New York), Vol. 22, pp. 597–602.
27. Fitch, W. M. (1966) *J. Mol. Biol.* **16,** 9–16.
28. Fitch, W. M. (1970) *J. Mol. Biol.* **49,** 1–14.
29. Needleman, S. & Wunsch, C. D. (1970) *J. Mol. Biol.* **48,** 443–453.
30. Mikes, O., Holeysovsky, V., Tomasek, V. & Sorm, F. (1966) *Biochem. Biophys. Res. Commun.* **24,** 346–352.
31. Blow, D. M., Birktoft, J. J. & Hartley, B. S. (1969) *Nature (London)* **221,** 337–340.
32. Butkowski, R. J., Elion, J., Downing, M. R. & Mann, K. G. (1977) *J. Biol. Chem.* **252,** 4942–4957.
33. Wiman, B. (1977) *Eur. J. Biochem.* **76,** 129–137.
34. Stroud, R. M., Kay, L. M. & Dickerson, R. E. (1971) *Cold Spring Harbor Symp. Quant. Biol.* **36,** 125–140.
35. Walz, D. A., Hewett-Emmett, D. & Seegers, W. H. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 1969–1972.
36. Sottrup-Jensen, L., Claeys, H., Zajdel, M., Peterson, T. E. & Magnusson, S. (1978) in *Progress in Chemical Fibrinolysis and Thrombolysis,* eds. Davidson, J. F., Rowan, R. M., Samama, M. M. & Desnoyers, P. C. (Raven, New York), Vol. 3, pp. 191–209.
37. Magnusson, S., Peterson, T. E., Sottrup-Jensen, L. & Claeys, H. (1975) in *Proteases and Biological Control,* Cold Spring Harbor Conferences on Cell Proliferation, eds. Reich, E., Rifkin, D. B. & Shaw, E. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), Vol. 2, pp. 123–149.
38. Black, J. A., Chan, G. F. Q. & Dixon, G. H. (1970) *Can. J. Biochem.* **40,** 123–132.
39. Barnett, D. R., Kurosky, A., Fuller, G. M., Kim, H.-H., Rasco, M. A. & Bowman, B. H. (1974) in *Protides of the Biological Fluids, Proceedings of the Colloquium,* ed. Peeters, H. (Pergamon, New York), Vol. 22, pp. 589–595.
40. Kurachi, K. & Davie, E. W. (1977) *Biochemistry* **16,** 5831–5839.
41. Javid, J. (1967) *Proc. Natl. Acad. Sci. USA* **57,** 920–924.
42. Greer, J. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 3393–3397.
43. Smithies, O. (1964) *Cold Spring Harbor Symp. Quant. Biol.* **29,** 309–319.
44. Kodama, M., Hashimoto, K. & Matsuura, F. (1975) *Bull. Jpn. Soc. Sci. Fish.* **41,** 1015–1019.