Characterization of the Cyanogen Bromide Fragments of the β Chain of Human Haptoglobin†

Alexander Kurosky,* Regine E. Hay, Han-Hwa Kim,‡ Billy Touchstone, Marilyn A. Rasco,§ and Barbara H. Bowman

ABSTRACT: Characterization of the cyanogen bromide (CNBr) fragments of the β chain of human haptoglobin revealed five major fragments resulting from cleavage of four methionyl residues. The fragments were isolated by gel filtration in guanidine-HCl on Sepharose 6B and Bio-Gel P10 and P60. Compositional analyses of the five cyanogen bromide fragments accounted for 248–253 amino acid residues in agreement with the number of residues determined for the intact β chain. Most of the carbohydrate was attached to CNBr II. Automated amino-terminal sequence analysis and carboxyl-terminal hydrolysis with carboxypeptidase of the haptoglobin β chain and cyanogen bromide fragments identified 139 residues, or about 55% of the β-chain molecule. The placement of the fragments within the β-chain molecule was established by sequence analysis of whole β chain and a plasmin cleavage fragment. The position of CNBr V was confirmed by the absence of homoserine or homoserine lactone. Cyanogen bromide reaction of intact haptoglobin 1-1 resulted in the isolation of a β-chain fragment, CNBr III, covalently attached to the intact α^1 chain by a single disulfide bond. The β chain was shown to have primary structural similarities to the chymotrypsin family of serine proteases. Partial sequence analysis of CNBr V established the region which is comparable to the serine-195 active-site region: /Asp-Thr-Cys-Tyr-Gly-Asp-Ala-Gly-Ser-Ala-Phe/ (residues 189–199, chymotrypsinogen A numbering). The active-site serine-195 is replaced by alanine; however, the specificity residue of the trypsin-like enzymes, Asp-189, is preserved. Several minor cyanogen bromide cleavage products were also identified in yields of up to 15%. These minor cleavage products give evidence that tryptophanyl residues in proteins, or glycoproteins, are also susceptible to cyanogen bromide cleavage.

Haptoglobin is a serum glycoprotein that has generated considerable interest in the areas of protein evolution and biochemical genetics. The functional unit of haptoglobin is a tetramer composed of two α and two β chains (Smithies et al., 1962) which are linked by interchain disulfide bonds (Connell et al., 1962; Shim and Bearn, 1964). A single disulfide bridge attaches the α and β chains to form an αβ subunit (Malchy and Dixon, 1973a; Kurosky et al., 1974b). Two such subunits are combined by an α21–α21 disulfide bond to form the Hp[1] 1-1 molecule (Malchy et al., 1973).

In human populations there are three common genetic variants of haptoglobin, designated Hp 1-1, Hp 2-1, and Hp 2-2 (Smithies, 1955). These Hp variations are the result of an almost complete duplication of the DNA coding for the α^1 chain, which gives rise to α^2 chains. The complete sequences of the α^1 and α^2 chains have been reported (Black and Dixon, 1968; Malchy and Dixon, 1973b). The two halves of the duplicated α^2 chain are unchanged by mutation, which indicates that the genetic event probably occurred recently.

The hp β chain was shown to be identical in all three common genotypes by acid–urea gel electrophoresis (Smithies et al., 1962), by immunological studies (Shim et al., 1965), and by identical tryptic peptides and amino acid compositions (Cleve et al., 1967). All of the carbohydrate moieties of Hp are attached to the β chain (Shim and Bearn, 1964). The reported molecular weights of the β chain ranged from 32 950 to 42 582, which included 17.8–23.5% carbohydrate. Recent articles by Sutton (1970), Giblett (1975), and Putnam (1975) review our present understanding of the haptoglobin molecule.

Haptoglobin is especially unique because of its ability to bind hemoglobin with virtually irreversible affinity. Because haptoglobin 1-1 independently binds two αβ subunits of hemoglobin, it has been analogized to immunoglobulin. Some sequence similarity between the α chain of haptoglobin and immunoglobulin light chains was presented to support this comparison (Black and Dixon, 1968). Sequence studies of the β chain, however, revealed considerable similarity to the chymotrypsin family of serine proteases (Barnett et al., 1970, 1972; Kurosky et al., 1974a,b).

We report here the isolation and chemical characterization of the cyanogen bromide fragments of the β chain of human haptoglobin. In addition, cyanogen bromide fragments of intact human haptoglobin 1-1 were isolated to investigate the disulfide arrangement of the α and β chains.

Experimental Section

Materials

Cyanogen bromide, Pth amino acids, iodoacetamide, Cheng Chin polyamide sheets, and Sequanal grade reagents used in the sequencer were obtained from the Pierce Chemical Co. [1-^{14}C]Iodoacetamide, specific activity 49 mCi/mmol, was purchased from Amersham/Searle Corp. Acrylamide and *N*,*N*′-methylenebisacrylamide, Electrophoresis Purity, and Bio-Gel beads were products of Bio-Rad Laboratories. Sepharose 6B and Sephadex products were obtained from Phar-

† From the Division of Human Genetics, Department of Human Biological Chemistry and Genetics, The University of Texas Medical Branch, Galveston, Texas 77550. *Received December 31, 1975.* This investigation was supported by Grant HD 03321 from the National Institute of Child Health and Human Development and by Grant H-378 from the Robert A. Welch Foundation.

‡ Present address: Department of Biology, Catholic Medical College, Seoul, Korea.

§ Present address: Department of Biology, Medical Genetics Section, M. D. Anderson Hospital, Houston, Texas.

[1] Abbreviations used are: Hp, for intact haptoglobin; hp, for polypeptide chains; Pth, phenylthiohydantoin; DMBA, *N*,*N*′-dimethylbenzylamine; HI, hydriodic acid; CNBr, cyanogen bromide, CPA, carboxypeptidase A; Dansyl (Dns), 5-dimethylaminonaphthalene-1-sulfonyl; Tris, 2-amino-2-hydroxymethyl-1,3-propanediol; CM, carboxymethyl.

macia Fine Chemicals. Dialysis tubing (retention >3500 daltons) was a product of Spectrum Medical Industries, Inc., Los Angeles. Carboxypeptidase A, treated with diisopropyl fluorophosphate, was purchased from Worthington Biochemical Corp. Coomasie brilliant blue R was obtained from Sigma Chemical Co.

Methods

Haptoglobin Purification. Haptoglobin was purified from human ascites fluid as described by Kurosky et al. (1974b). Most of the haptoglobin employed in this study was isolated from ascites fluid drawn from a single patient (Mal) whose phenotype was Hp 1-1. Haptoglobin 2-1 and 2-2 phenotypes were also utilized in a few studies. The β chain of haptoglobin was isolated by gel filtration (Kurosky et al., 1974b) after reduction with 2-mercaptoethanol and alkylation with [1-^{14}C]iodoacetamide.

Cyanogen Bromide Reaction. Haptoglobin 1-1 or *S*-carbamoylmethylated hp β chain was dissolved in 70% formic acid (15 mg of protein/ml) for 3–4 h prior to reaction with CNBr (Gross, 1967). A 300-fold molar excess of cyanogen bromide was used based on 4 methionyl residues/mol of β chain. Half of the required cyanogen bromide was added at zero time and the remainder after 12 h. The reaction was allowed to proceed in the dark at room temperature for a total of 24 h. The reaction mixture was then diluted with about 10 volumes of distilled water and lyophilized. Approximately 8 mg of a cyanogen bromide reaction mixture was subjected directly to automated sequence analysis to characterize the amino-terminal residues prior to fractionation.

Isolation of β Chain CNBr Fragments. The dried mixture of CNBr cleavage fragments from β chain was dissolved in 5.2 M guanidine-HCl, 0.1 M Tris-HCl, pH 7.0, at a concentration of 20 mg of protein/ml. Originally, the CNBr fragments were separated by gel filtration on a Sepharose 6B column equilibrated with the guanidine/Tris buffer. However, subsequently, a Bio-Gel P10 column equilibrated with 5.2 M guanidine-HCl, 0.1 M Tris-HCl, pH 7.0, followed by a Bio-Gel P60 column equilibrated with 5.2 M guanidine-HCl, 0.2 M formate, pH 3.5, replaced the use of Sepharose to separate the CNBr fragments. The absorbance of the eluate was monitored at 280 nm and appropriate fractions were pooled. Fragments with molecular weight above 5000 were dialyzed extensively against distilled water. Smaller fragments were desalted on a Sephadex G-10 column (2.5 × 90 cm) eluted with 5% acetic acid. After removal of salts, the fragments were lyophilized.

Isolation of Hp 1-1 CNBr Fragments. The lyophilized cyanogen bromide reaction mixture of whole Hp 1-1 was fractionated on Sephadex G-100 eluted with 30% acetic acid. A selected fraction (B) obtained from this column was rechromatographed under identical conditions and, finally, purified by gel filtration on Sephadex G-75. After reduction with 2-mercaptoethanol and alkylation with iodoacetamide, fraction B was subjected to gel filtration on Sephadex G-75.

Characterization of the Small Molecular Weight CNBr Fragments. In a separate experiment designed to examine the small-molecular-weight components resulting from cyanogen bromide reaction with hp β chain, 225 mg of CNBr-reacted β chain was fractionated by gel filtration on a Bio-Gel P4 column (2.5 × 120 cm) eluted with 70% formic acid. The small-molecular-weight fraction (<3000) was pooled, lyophilized, and subjected to paper chromatography (1-butanol–glacial acetic acid–water, 4:1:5, v/v; Bennett, 1967). Ninhydrin positive bands were further purified by high-voltage electrophoresis (pH 2.0). In addition, fractions from the Bio-Gel P10 column containing CNBr I were further purified by high-voltage electrophoresis (pH 6.5).

Acrylamide Gel Electrophoresis. Haptoglobin homogeneity was investigated by slab gel electrophoresis (Woodworth and Clark, 1967) and by immunoelectrophoresis. Purification and molecular weight estimation of the β chain were investigated by sodium dodecyl sulfate-acrylamide gel electrophoresis according to the method of Weber and Osborn (1969). Purification of CNBr fragments was established by sodium dodecyl sulfate–urea and acid–urea–polyacrylamide gel electrophoresis. The sodium dodecyl sulfate–urea gels contained 12.5% polyacrylamide with a ratio of *N*,*N*′-methylenebisacrylamide to acrylamide of 1:10 (Swank and Munkres, 1971). The acid–urea gels were 15% polyacrylamide and had a ratio of *N*,*N*′-methylenebisacrylamide to acrylamide of 1:150 (Panyim and Chalkley, 1969). Approximately 40 μg of protein was applied to each gel. Acrylamide gels were stained with a filtered solution of Coomasie blue containing 0.125% Coomasie brilliant blue R (dissolved in a small volume of acetone), 10% acetic acid, 10% trichloroacetic acid, and 10% 2-propanol. Gels were destained in a Bio-Rad diffusion destainer with a solution of 10% acetic acid and 10% 2-propanol. Acrylamide gels were also stained for carbohydrate using the periodic acid–Schiff method described by Glossman and Neville (1971).

Amino Acid Analysis. Amino acid compositional analyses were obtained with a Beckman 119 analyzer employing single column methodology on Durrum DC-6A resin and a JEOL analyzer utilizing dual-column methodology on Hamilton HP-AN90 resin. Both instruments employed temperature programming in order to separate homoserine from glutamic acid and to separate glucosamine and galactosamine. Time-course hydrolysis for hexosamine analysis was conducted for 3, 6, and 12 h in 2 N HCl at 100 °C (Marshall and Neuberger, 1972). Time-course hydrolysis of proteins was carried out for 24, 48, and 72 h at 107 °C in 5.7 N HCl. Oxidation of half-cystinyl residues to cysteic acid was achieved with dimethyl sulfoxide according to Spencer and Wold (1969).

Carboxyl-Terminal Analysis. The assignment of carboxyl-terminal sequences was made by the time-course stoichiometric release of residues during enzymatic hydrolysis with carboxypeptidase A at 37 °C as described by Ambler (1967). Reactions were terminated by lowering the pH to 3.0 with formic acid.

Automated Sequence Analysis. Automated amino acid sequence analysis was performed with an updated 890B Beckman sequencer using the DMBA program. Sequencer products, converted to Pth amino acid residues, were confirmed by gas chromatography and by amino acid analysis after back-hydrolysis with HI. Routinely, one-tenth of the sample was analyzed by gas chromatography and the remainder was subjected to HI hydrolysis. Sequencer operation, methods of identification, and quantitation of sequencer products were detailed recently by Kurosky et al. (1976b).

Manual Dansyl–Edman Sequence Analysis. Manual sequence analysis was carried out using essentially the dansyl–Edman method described by Hartley (1970a) and modified by Kurosky (1972). Dansylated amino acid residues were identified by two-dimensional thin-layer chromatography on polyamide sheets (5 × 5 cm).

Results

Purification of Haptoglobin and hp β Chain. Slab acrylamide gel electrophoresis of purified haptoglobin gave typical gel patterns for all phenotypes (1-1, 2-1, and 2-2) after staining with 3,3′-dimethoxybenzidine. The typical polymer formation

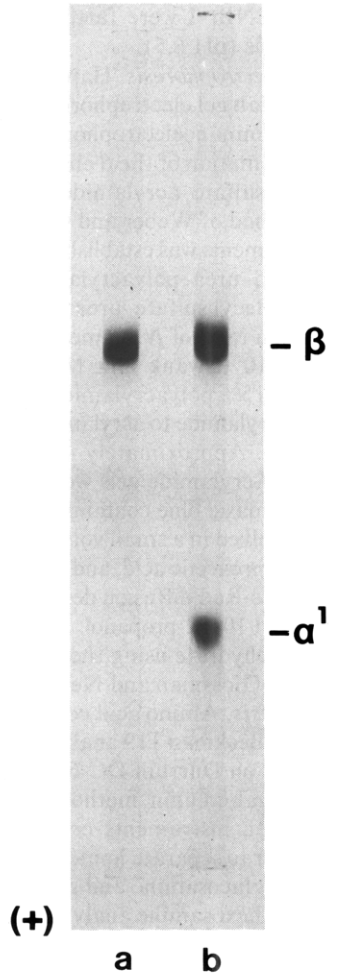

FIGURE 1: Sodium dodecyl sulfate–polyacrylamide gel electrophoresis of purified Hp 1-1 and hp β chain. Samples were treated with 2-mercaptoethanol and stained with Coomasie blue. (a) Carbamoylmethylated human hp β chain; (b) human Hp 1-1.

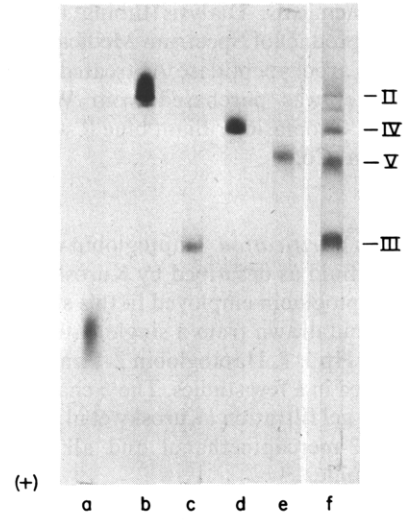

FIGURE 2: Sodium dodecyl sulfate–urea–polyacrylamide gel electrophoresis of the CNBr fragments of hp β chain isolated by gel filtration on Bio-Gel P10 and P60. Samples were treated with 2-mercaptoethanol and stained with Coomasie blue. (a) CNBr I (contains anomalous cleavage products, see text); (b) CNBr II; (c) CNBr III; (d) CNBr IV; (e) CNBr V; (f) unfractionated CNBr reaction mixture. CNBr I diffused out of the gel during the staining and destaining process.

as characterized by Fuller et al. (1973) was observed for the 2-1 and 2-2 phenotypes. Immunoelectrophoresis of purified preparations of haptoglobin indicated a high degree of protein purification with little evidence of other serum proteins. Polyacrylamide gel electrophoretic patterns of Hp 1-1 and hp β chain are given in Figure 1. Further evidence of the purity of the β chain was obtained from amino acid sequence analysis. Only a single amino acid sequence was identified by automated analysis of intact hp β chain. No differences were noted in β chains prepared from phenotypes 1-1, 2-1, or 2-2.

Cyanogen Bromide Reaction. Compositional analysis of the hp β chain indicated 4 methionyl residues/mol of protein (Table I), and, therefore, 5 CNBr fragments were expected. Sodium dodecyl sulfate–urea–polyacrylamide gel electrophoresis of the unfractionated CNBr reaction mixture revealed four major bands (Figure 2f). A stained band corresponding to CNBr I could not be identified on the gel of the unfractionated CNBr reaction mixture. The absence of this fragment was due to loss by diffusion during the staining and destaining process. CNBr I was visualized only when an increased amount was applied to the gel, e.g., Figure 2a. Several minor bands were also evident on the gel of the unfractionated CNBr mixture. These bands probably corresponded to fragments obtained as a result of incomplete cleavages of methionine, as well as minor cleavages of other residues. To confirm the number of CNBr cleavage products, a portion of the unfractionated CNBr reaction mixture was applied directly to the sequencer. The results obtained from three cycles of degradation are given in Table II. Five major amino-terminal sequences could be identified, consistent with the occurrence of 4 methionines and 5 CNBr fragments. The amino-terminal residues identified in the unfractionated cyanogen bromide reaction mixture agreed with the amino-terminal sequences established for the 5 CNBr fragments (see below).

Isolation of β Chain CNBr Fragments. The cleavage products obtained from the hp β chain, after reaction with cyanogen bromide, were fractionated by gel filtration. In one series of experiments, fractionation was achieved on Sepharose 6B, as indicated in Figure 3. Each of the five peaks was identified as indicated. The numbering of the CNBr fragments is according to their actual position in the molecule beginning from the amino terminus (see Discussion). The last unlabeled peak shown in Figure 3 contained salts and small anomalously cleaved peptides, as revealed by amino acid analysis. Fragments I, III, and IV were well resolved but fragments II and V were not. In addition, the methionyl residue at position 18 between CNBr I and II did not cleave completely (5–10% uncleaved) and, consequently, the fragment CNBr I–II co-purified with CNBr II. Incomplete cleavage of methionyl residues by cyanogen bromide has been reported by others (Titani et al., 1972; Olafson et al., 1975) and may be due to methionine oxidation (Gross, 1967). Some difficulty was encountered with the fragments fractionated on Sepharose. In the presence of 5.2 M guanidine-HCl, 0.1 M Tris-HCl, pH 7.0, Sepharose was partially solubilized and contaminated the CNBr fractions with Agarose. After lyophilization, this material appeared very much like protein. To avoid this problem, subsequent CNBr reaction products were fractionated using the more chemically stable Bio-Gel beads (P10 and P60).

Fractionation of CNBr fragments on Bio-Gel P10 resulted in four major peaks, as shown in Figure 4A. The first peak from P10 gel filtration containing CNBr II and CNBr IV was rechromatographed on Bio-Gel P60. The elution profile of this column is given in Figure 4B. Contamination of CNBr II with CNBr I–II was minimized by selecting the center portion of the peak eluted from the P60 column.

TABLE I: Amino Acid Compositions[a] of the Cyanogen Bromide Fragments of Human hp β Chain

Amino Acid	CNBr I[b]	CNBr II[b,e]	CNBr III[c]	CNBr IV[d,e]	CNBr V[b,e]	Total	β Chain[b,f]
Lysine	1.8 ± 0.2 (2)	7.2 ± 0.2 (7)	3.0 ± 0.1 (3)	3.9 ± 0.3 (4)	4.1 ± 0.1 (4)	20	20.0 ± 0.8 (20)
Histidine	0.8 ± 0.1 (1)	3.7 ± 0.1 (4)	1.0 ± 0.04 (1)	2.1 ± 0.2 (2)	1.1 ± 0.1 (1)	9	8.7 ± 0.2 (9)
Arginine		1.2 ± 0.2 (1)	1.7 (2)	0.9 ± 0.2 (1)		4	4.5 ± 0.3 (4–5)
CM-cysteine			0.9 ± 0.01 (1)	1.8 ± 0.2 (2)	1.3 ± 0.2 (2)	5	5.0 ± 0.4 (5)
Aspartic acid	1.1 ± 0.1 (1)	8.6 ± 0.3 (8–9)	4.7 ± 0.1 (4)	5.0 ± 0.1 (5)	6.3 ± 0.2 (6)	24–25	28.4 ± 0.6 (28)
Threonine		7.3 ± 0.2 (7)	1.1 ± 0.04 (1)	4.1 ± 0.3 (4)	4.0 ± 0.2 (4)	16	15.9 ± 1(16)
Serine[g]	0.9 ± 0.01 (1)	4.6 ± 0.1 (5)	1.7 ± 0.1 (2)	3.0 ± 0.3 (3)	7.2 ± 0.2 (7)	18	17.5 ± 0.2 (18)
Glutamic acid	1.0 ± 0.1 (1)	9.0 ± 0.3 (9)	1.6 ± 0.1 (1)	7.0 ± 0.6 (7)	7.5 ± 0.2 (7–8)	25–26	26.5 ± 2 (26–27)
Proline	1.0 ± 0.1 (1)	2.0 ± 0.2 (2)	2.3 ± 0.2 (2)	5.1 ± 0.5 (5)	0.6 ± 0.1 (1)	11	10.4 ± 0.6 (10–11)
Glycine	3.1 ± 0.1 (3)	3.6 ± 0.1 (4)	4.0 ± 0.01 (4)	4.0 ± 0.3 (4)	4.9 ± 0.2 (5)	20	19.1 ± 0.3 (19)
Alanine	2.0 ± 0.1 (2)	5.4 ± 0.01 (5)	2.1 ± 0.1 (2)	3.6 ± 0.2 (3–4)	6.2 ± 0.02 (6)	18–19	21.2 ± 0.1 (21)
Half-cystine				1.8 ± 0.3			
Valine		7.3 ± 0.2 (7)[j]	4.1 ± 0.04 (4)	4.0 ± 0.2 (4)	4.1 ± 0.04 (4)	19	17.4 ± 0.4 (17–18)[j]
Methionine							3.8 ± 0.6 (4)
Isoleucine	0.9 ± 0.1 (1)[j]	3.6 ± 0.3 (4)[j]	1.0 ± 0.1 (1)	2.5 ± 0.3 (2–3)	2.1 ± 0.1 (2)	10–11	8.4 ± 0.3 (8–9)[j]
Leucine	2.0 ± 0.03 (2)[j]	11.2 ± 0.1 (11)[j]	2.0 ± 0.1 (2)	2.7 ± 0.4 (3)	1.9 ± 0.3. (2)	20	21.2 ± 0.6 (21)[j]
Tyrosine		2.3 ± 0.1 (2)	2.8 ± 0.1 (3)	1.9 ± 0.2 (2)	3.5 ± 0.1 (4)	11	10.6 ± 0.2 (11)
Phenylalanine	1.0 ± 0.04 (1)	1.4 ± 0.1 (1–2)	1.8 ± 0.1 (2)	1.9 ± 0.2 (2)	1.8 ± 0.2 (2)	8–9	8.5 ± 0.3 (8–9)
Tryptophan[h]	+ (1)	ND (2)	ND (1)	ND	ND (2)	6	ND (6)
Homoserine[i]	0.8 ± 0.1 (1)	1.1 ± 0.1 (1)	0.9 (1)	1.0 ± 0.2 (1)		4	
Glucosamine[g]		20.0 ± 0.8 (20)		tr.	tr.	20	19.6 ± 0.8 (20)
Galactosamine[g]		2.8 ± 0.1 (3)				3	3.0 ± 0.3 (3)
Total	18	80–82	37	54–56	59–60	248–253	251–257

[a] Values are given in residues per molecule. Values in parentheses are best or nearest integers, except CNBr I and III which are based on final sequence. [b] Duplicate time-course hydrolyses: 24, 48, and 72 h. [c] Average of two analyses. [d] Calculated from seven analyses (18–30 h) before and after reduction and alkylation. [e] Based on 1 homoserine and partial sequence analysis and adjusted to give best integral values. [f] Based on a molecular weight of 38 500 ± 4100 (see Table VII) and adjusted to give best integral values relative to 4 methionyl and 5 half-cystinyl sequences. [g] Adjusted to zero time of hydrolysis. [h] Not determined quantitatively; from sequence analysis. [i] Homoserine + homoserine lactone. [j] Average of only 72-h hydrolyses (duplicate).

TABLE II: Sequence Analysis[a] of Unfractionated hp β Chain Reacted With Cyanogen Bromide.

CNBr Fragment[b]	Cycle 1 Residue (nmol)	Cycle 2 Residue (nmol)	Cycle 3 Residue (nmol)
I[c]	Ile (197)	Leu (194)	Gly (112)
II	Val (112)	Ser (140)	His (86)
III	Pro (119)	Ile (164)	CM-Cys (81)
IV	Leu (166)	Pro (117)	Val (109)
V	Ser (143)	Lys (108)	Tyr (89)

[a] 200 nmol of CNBr reacted β chain analyzed. Pth residues were quantitated by amino acid analysis after HI hydrolysis. [b] Sequences of the fragments are aligned as determined for isolated fragments (see Table IX). [c] Recovery of this sequence relative to the other sequences was greater, probably as a result of incomplete cleavage.

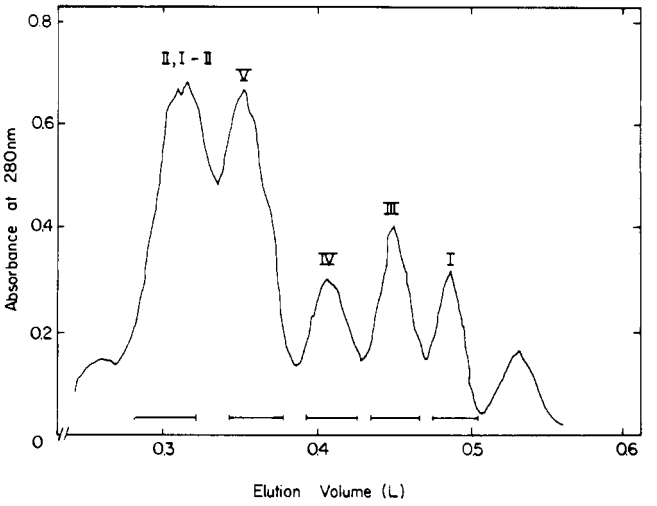

FIGURE 3: Gel filtration of 125 mg of CNBr fragments of hp β chain on Sepharose 6B in 5.2 M guanidine-HCl, 0.1 M Tris-HCl, pH 7.0. The column (2.5 × 140 cm) was eluted at 3 ml/h and 3-ml fractions were collected.

An estimation of the purity of the CNBr fragments after gel filtration on P10 and P60 was obtained by sodium dodecyl sulfate-urea-acrylamide gel electrophoresis (Figure 2a-e). Each CNBr fragment appeared relatively pure, except for CNBr I which stained diffusely and indicated more than one component. The other components in the P10 fraction containing CNBr I were anomalous cleavage products (see below). CNBr I was further purified by high-voltage electrophoresis at pH 6.5. Automated sequence analysis of the purified CNBr fragments indicated no observable contamination, except for CNBr II. This fragment was contaminated with about 5% of CNBr I-II.

Isolation of Hp 1-1 CNBr Fragments. Fragments obtained after reaction of Hp 1-1 with cyanogen bromide were fractionated by gel filtration on Sephadex, as indicated in Figure 5A. One of the CNBr fractions, denoted as fraction B, contained CNBr III covalently attached to the α^1 chain, as indicated by compositional and sequence analyses. After reduction with 2-mercaptoethanol and alkylation with iodoacetamide, the disulfide-linked fragments were separated, as shown in Figure 5B. Compositional analyses of the fragments before and after reduction and alkylation are given in Table III. Sequence analysis of fraction B identified two chains corresponding to CNBr III and the α^1 chain (Table IV). CNBr III and α^1 chain isolated after reduction and alkylation were confirmed by several cycles of automated sequence analysis, as indicated in Table IV. CNBr IV was also purified from the CNBr cleavage mixture of Hp 1-1 and gave an identical composition and partial sequence analysis as CNBr IV purified from CNBr reaction with isolated β chain.

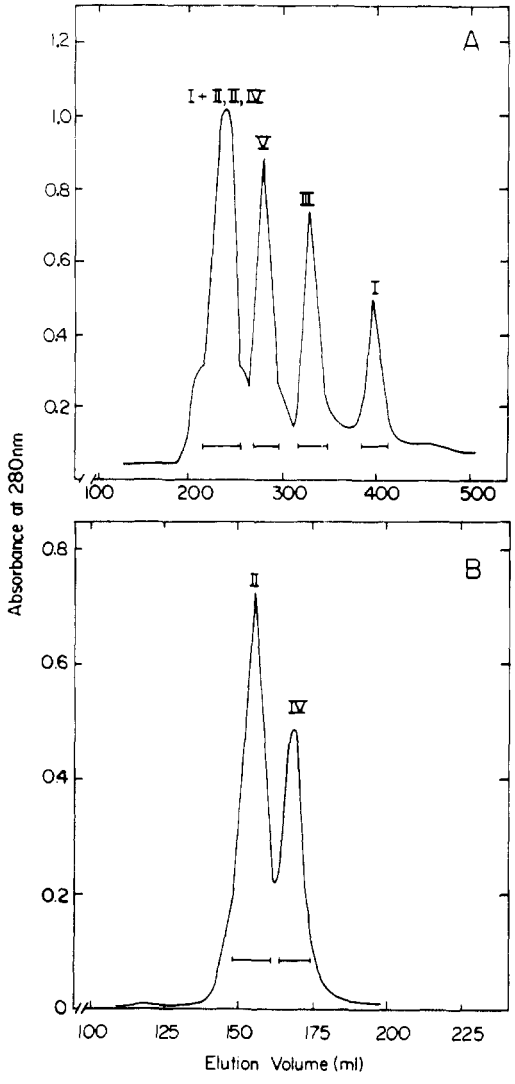

FIGURE 4: Gel filtration of CNBr fragments of hp β chain on Bio-Gel. (A) Separation of 125 mg on a column of P10 (2.5 × 110 cm) eluted with 5.2 M guanidine-HCl, 0.1 M Tris-HCl, pH 7.0, at 3 ml/h. Fractions of 3 ml were collected. (B) Rechromatography of the first peak obtained from the P10 column on a column of P60 (2.5 × 110 cm) eluted with 5.2 M guanidine-HCl, 0.2 M sodium formate, pH 3.5, at 3 ml/h. Fractions of 3 ml were collected. The first fraction obtained from the P10 column (4A) was dialyzed, lyophilized, and redissolved in the guanidine-formate buffer prior to application onto the P60 column.

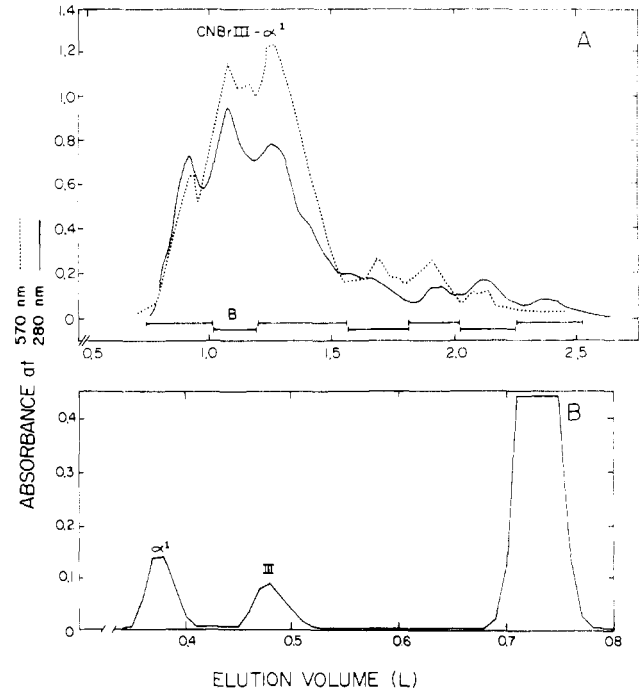

FIGURE 5: Isolation and characterization of CNBr III-α^1 obtained from a CNBr reaction mixture of Hp 1-1. (A) Gel filtration of 480 mg of CNBr-reacted Hp 1-1 on a Sephadex G-100 column (5 × 120 cm) eluted with 30% acetic acid at 20 ml/h. Fractions of 5 ml were collected. Fraction B was pooled and subsequently purified by further gel filtration (see Methods). (B) About 400 nmol of purified fraction B (CNBr III-α^1) was reduced and alkylated and subjected to gel filtration on a column of Sephadex G-75 (2.5 × 140 cm) eluted with 30% acetic acid at 20 ml/h. Fractions of 5 ml were collected. The last peak was the result of salt primarily from the 6 M guanidine-HCl used in the reduction-alkylation reaction.

Anomalous Cyanogen Bromide Cleavage Products, During the course of this study, several anomalous minor cleavage products were identified. The composition and partial sequence analysis of two of these peptides are illustrated in Table V. Peptide CNBr 2C2 represented an anomalous cleavage comprising residues 14–18. The occurrence of tryptophan was not positively established due to low recovery; however, amino-terminal analysis by dansylation gave negative results consistent with the occurrence of a tryptophanyl residue. Moreover, alanine was positively identified after two Edman degradations. A second anomalously cleaved peptide (CNBr Ia) was obtained from the Bio-Gel P10 fraction containing CNBr I. High-voltage electrophoresis (pH 6.5) of this pool resulted in several peptide fractions, two of which were CNBr I (residues 1–18) differing only by deamidation of Gln-15. Peptide CNBr Ia represented the first 20 residues of CNBr III and gave a negative reaction with Ehrlich reagent (Bennett, 1967). Amino acid analysis of CNBr Ia indicated half-cystine in the composition and not CM-Cys as expected. The absence of the carboxymethyl derivative is unexplained. Both of the anomalously cleaved fragments were generated by cleavage of a tryptophanyl peptide bond.

Molecular Weight Estimation. Molecular weights of the CNBr fragments were estimated by sodium dodecyl sulfate–urea–polyacrylamide gel electrophoresis (Figure 2), in addition to gel filtration. The calculated molecular weights of the fragments and whole β chain are summarized in Table VI.

Compositional Analysis. The amino acid compositions of the hp β chain and CNBr fragments are presented in Table I. The total number of residues calculated for hp β chain was 251–257. This range of values is in reasonable agreement with the total number of residues obtained from the CNBr fragments (248–253). The number of residues for the β chain was derived from a preliminary calculation using a molecular weight of 38 500 and a carbohydrate percentage of 19.8 (mean of reported values summarized in Table VII). The composition was then adjusted to best integral values to agree with 4 methionines and 5 half-cystines determined from sequence analyses of fragments of the intact molecule. The β-chain composition presented in Table I is in accord with that reported by Cleve et al. (1969) when adjusted to the same molecular weight. As indicated in Table VII, the reported ultracentrifugational molecular weights of hp β chain ranged from 32 950 to 42 582, giving a mean of about 38 500 ± 4100. The molecular weight of β chain, as calculated from composition, was 35 172 ± 371. This estimation included 19.8% carbohydrate. The molecular weight for intact Hp 1-1 would, therefore, be approximately 88 900 (2(35 170 + 9300)). The molecular weight for the α^1 chain (9300) was obtained from sequence

TABLE III: Amino Acid Composition[a] of Hp 1-1 CNBr III–α^1 before and after Reduction and Alkylation.[b]

Amino Acid	CNBr III–α^1	CNBr III[c]	α^1	α^1[d]
Lysine	11.3	3.0 ± 0.1 (3)	8.3	8
Histidine	3.1	1.0 ± 0.04 (1)	2.2	2
Arginine	3.8	1.7 (2)	2.0	2
CM-cysteine	0	0.9 ± 0.01 (1)	3.5	0
Aspartic Acid	18.6	4.7 ± 0.1 (4)	13.9	15
Threonine	4.3	1.1 ± 0.04 (1)	3.2	3
Serine	4.1	1.7 ± 0.1 (2)	2.2	2
Glutamic Acid	10.7	1.6 ± 0.1 (1)	9.5	9
Proline	9.3	2.3 ± 0.2 (2)	7.4	7
Glycine	11.2	4.0 ± 0.01 (4)	7.0	7
Alanine	7.1	2.1 ± 0.1 (2)	4.9	5
Half-cystine	3.7	0	0	4
Valine	11.4	4.1 ± 0.04 (4)	8.3	8
Methionine	0	0	0	0
Isoleucine	3.7	1.0 ± 0.1 (1)	2.8	3
Leucine	5.5	2.0 ± 0.1 (2)	3.2	3
Tyrosine	6.7	2.8 ± 0.1 (3)	4.5	5
Phenylalanine	2.1	1.8 ± 0.1 (2)	0	0
Homoserine	+	0.9 (1)	0	0
Tryptophan[e]	ND	ND (1)	ND	1
Glucosamine	0	0 (0)	0	0

[a] Values are given in residues per molecule. [b] CNBr III–α^1 after reduction and alkylation was fractionated on G-75 to give CNBr III and α^1 chain. [c] Average of two analyses. Values in parentheses obtained from final sequence (Kurosky et al., 1974a). [d] Residues determined by sequence analysis (Black and Dixon, 1968; Malchy and Dixon, 1973b). [e] Not determined.

TABLE IV: Automated Sequence Analyses[a] of CNBr III–α^1 before and after Reduction and Alkylation.

Cycle	Residue (nmol) CNBr III–α^1[b]		CNBr III[c]		α^1 Chain[c]
1	Pro (157),	Val (147)	Pro	(170)	Val (184)
2	Ile (100),	Asn (0)	Ile	(152)	Asx (75)[d]
3	Half-Cys (ND),	Asp (ND)	CM-Cys	(0)	Asp (ND)
4	Leu (88),	Ser (0)	Leu	(125)	Ser (65)[d]

[a] Results from a single analysis of each fragment. Residue assignments made by gas chromatographic analysis, except as denoted. Gas chromatographic column performance low, Pth-Ser, Pth-Asn, and Pth-CM-Cys were not obtained. [b] 100 nmol applied to reaction cup; however, each mole of CNBr III–α^1 contains 2(CNBr III plus α^1 chain). [c] 250 nmol applied. [d] Determined by amino acid analysis after HI hydrolysis.

TABLE V: Amino Acid Compositions and Sequence Data[a] of Anomalously Cleaved Cyanogen Bromide Peptides of hp β Chain.

Peptide	Analyses (Molar Ratios) and Sequence Data	Residue Position[c]	Recovery (%)
CNBr Ia	Pro–Ile (Cys)[d] Leu–Pro 0.8 0.9 0.9 1.1 0.8 (Ser, Lys, Asx, Tyr, Ala, 0.8 0.9 1.1 0.8 1.0 Glx, Val, Gly, Arg, Val, 1.2 0.9 1.1 0.9 0.9 Gly, Tyr, Val, Ser, Gly) 1.1 0.8 0.9 0.8 1.1	101–120	18[e]
CNBr 2C2	(Trp, Glx) Ala (Lys, Met[f]) ND 1.2 0.8 1.0 1.0	14–18	1.3

[a] By dansyl–Edman. [b] See Methods for purification. [c] According to Figure 8. [d] As half-cystine, see Discussion. [e] Relative to CNBr I after high-voltage paper electrophoresis. [f] As homoserine plus homoserine lactone.

data (Black and Dixon, 1968; Malchy and Dixon, 1973a,b).

Acid hydrolysis of the β chain in 2 N HCl at 100 °C for 12 h released only about 67% of the glucosamine determined after hydrolysis with 5.7 N HCl at 107 °C, as reported in Table I. All of the glucosamine occurred in CNBr II. After sodium dodecyl sulfate–urea and acid–urea–polyacrylamide gel electrophoresis, CNBr II, CNBr IV, and CNBr V stained positively for carbohydrate using the periodic acid–Schiff technique. CNBr IV and V stained faintly, whereas the stained band for CNBr II was distinctly pronounced.

The occurrence of galactosamine in hp β chain was unexpected. Using two-column methodology of amino acid analysis, the hexosamines eluted after phenylalanine. The galactosamine peak in the β chain possessed an A_{440}/A_{570} ratio which was identical to the standard galactosamine ratio. Also, typical for hexosamines, the concentration of the galactosamine decreased with increased hydrolysis time. Almost all of the galactosamine could be accounted for in CNBr II.

Carboxyl-Terminal Analysis. The results obtained from time-course CPA hydrolysis of the CNBr peptides are summarized in Table VIII. CPA hydrolysis of CNBr V released only 0.4 mol of either asparagine, glutamine, or serine/mol of peptide in 24 h with no evidence of homoserine or homoserine lactone. Although these three amino acids were not resolved by the amino acid analyzer conditions routinely employed for protein hydrolyzates, amino acid analysis after acid hydrolysis of the CPA digests indicated aspartic acid. Identical results were obtained from CPA hydrolysis of intact hp β chain.

Amino-Terminal Sequence Analysis. Automated sequence analysis of 36 residues of the intact hp β chain was previously reported by Barnett et al. (1972). These sequence data were obtained from tryptic and chymotryptic peptides, from sequencer analysis of intact hp β chain, and from preliminary sequencer analysis of CNBr II, and indicated that residues 2–9 were reiterated at positions 20–27. However, sequence analysis obtained from β chains, prepared from hp isolated from sera of other animal species, gave no evidence of a sequence reiteration (Kurosky et al., 1976a) and led us to reinvestigate the amino-terminal sequence of human hp β chain. The residue recoveries obtained from automated sequence analysis of intact β chain and the cyanogen bromide fragments are presented in Table IX. Figure 7 is a plot of the recovery data for the β-chain residues from a typical sequencer run. Repetitive yields generally ranged 90–93%. Figure 8 is a summary of the amino-terminal sequence analyses obtained for intact β chain and the cyanogen bromide fragments. As indicated in Figure 8, residues 2–9 are not reiterated and the sequence is corrected as shown. Misinterpretation of earlier data was due to incomplete cleavage of CNBr I and II resulting in CNBr I–II, which readily copurified by gel filtration with CNBr II. The nature of the residues, especially the coincident occurrence of Leu-6 and -27 and the absence of an identifiable residue due to carbohydrate at position-23, led to the interpretation that residues 2–9 were reiterated in human hp β chain. In addition, residues 17, 29, and 33 have been corrected, as indicated in Figure 8.

Further evidence to substantiate the amino-terminal sequence correction and confirm the presence of carbohydrate attached to asparagine-23 was obtained from a tryptic peptide, T2C, isolated from a tryptic hydrolysis of intact *S*-carbamoylmethylated β chain. The tryptic hydrolysate was fractionated by gel filtration on Sephadex G-75, followed by ion-exchange chromatography on Bio-Rad Aminex A5 (Bradshaw et al., 1969). Peptide T2C was comprised of residues 18–41 and

TABLE VI: Molecular Weight Estimates[a] of hp β Chain CNBr Fragments.

Sample	Gel Filtration[b] P10	Gel Filtration[b] P60	Polyacrylamide Gel Electrophoresis: Sodium Dodecyl Sulfate–Urea[c]	Polyacrylamide Gel Electrophoresis: Sodium Dodecyl Sulfate[d]	Composition
CNBr I	2 100 ± 100		2 450 ± 200		1 940
CNBr II		17 900 ± 400	21 600 ± 1500		16 051 ± 131[e]
CNBr III	5 400 ± 300		4 500 ± 30		4 205
CNBr IV		11 600 ± 300	14 600 ± 200		6 029 ± 92
CNBr V	10 900 ± 300		10 300 ± 700		6 572
					34 797 ± 223[e]
hp β Chain				42 000 ± 400	35 172 ± 371[e]
P_2[f]			13 200		13 737 ± 92

[a] Results obtained from three separate analyses. [b] Fish et al. (1969). [c] Swank and Munkres (1971). [d] Weber and Osborn (1969). [e] Includes carbohydrate, 19.8% of the β chain. [f] Plasmin cleavage fragment consisting of residues 129 through 253 inclusive (Kurosky et al., 1974b; Hay, Kurosky and Bowman, to be published).

TABLE VII: Summary of Reported Molecular Weight and Carbohydrate Composition of Human Haptoglobin.

Molecular Weight: Hp 1-1	Molecular Weight: β Chain	Method	Carbohydrate (%)[a]	Reference
84 500	32 950[c]	Ultracentrifugation	17.3[b]	Jayle et al. (1956)
ND	ND		18.6	Schultze et al. (1963)
ND	ND		17.8	Gerbeck et al. (1965)
98 770	40 085[c]	Ultracentrifugation	23.5	Cheftel and Moretti (1966)
ND	38 000	Gel filtration	23.5[d]	Gordon et al. (1967)
96 795	42 582	Ultracentrifugation	19.4	Black et al. (1970)
ND	38 400[e]	Sodium dodecyl sulfate, gel electrophoresis, and ultracentrifugation	ND	Fuller et al. (1973)
ND	42 000	Sodium dodecyl sulfate gel electrophoresis	ND	This report
Mean[f]	38 500 ± 4 100		19.8 ± 2.5	

[a] Percentage of β chain. [b] Does not include fucose or sialic acid. [c] Calculated from Hp 1-1 value using α^1 chain molecular weight of 9300. [d] Not included in mean value; value obtained from Cheftel and Moretti (1966). [e] Calculated from molecular weight difference of Hp 2-2 polymers using α^2 chain molecular weight of 16 100. [f] Ultracentrifugation values only.

TABLE VIII: Carboxyl-Terminal Sequences of the Cyanogen Bromide Fragments.

Fragment	Sequence and Relative Yields[a]
CNBr I	ND[b]
CNBr II	/Val-Hsr[c] 1.0 1.0
CNBr III	/Leu-Lys-Tyr-Val-Hsr[c,d]
CNBr IV	/Ala-Gly-Hsr[c] 0.3 0.3 0.5
CNBr V	/Asn. 0.4
hp β chain[e]	/Trp-Val-Glx-Lys-Thr-Ile-Ala-Glu-Asn·

[a] Mole of amino acid released by CPA per mole of peptide in 24 h. [b] Established by sequence analysis of whole β chain. [c] Homoserine plus homoserine lactone. [d] Established by time-course hydrolysis (see Figure 6). [e] Determined by Barnett et al. (1972) (see Discussion).

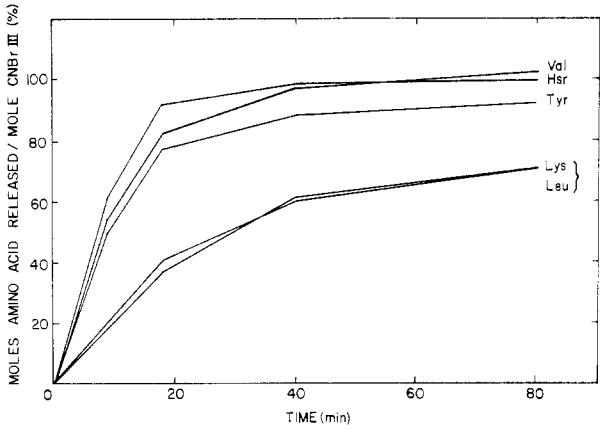

FIGURE 6: CPA hydrolysis of CNBr III.

amino acid analysis confirmed the presence of glucosamine. The amino acid composition and sequence analysis of peptide T2C are summarized in Table X. Hydrolysis of peptide T2C with CPA for 24 h released 1.0 mol of Lys, 1.0 mol of Ala, 2.0 mol of Thr, and 2.0 mol of Leu in agreement with the sequence shown for residues 18–41 in Figure 8. Automated sequence analysis of CNBr II (Table IX) confirmed the sequence determined for tryptic peptide T2C. Sequence analysis of intact β chain, peptide T2C, and CNBr II gave no results for residue 23, presumably because the attachment of carbohydrate at this residue makes it insoluble in the 1-chlorobutane used for extraction of the residues from the reaction cup of the sequencer.

Discussion

Characterization of the cyanogen bromide fragments of the β chain of human haptoglobin presented here is consistent with a composition of 4 methionyl residues and 5 CNBr fragments per molecule. Previous preliminary evidence (Barnett et al., 1972) indicated 6 cyanogen bromide fragments in the hp β chain; however, no evidence for the 6th fragment was

TABLE IX: Automatic Sequence Analyses[a] of the Amino-Terminal Regions of the CNBr Peptides and Whole β Chain.

Edman Cycle	Residue (nmol)					
	β Chain	CNBr I[b]	CNBr II	CNBr III	CNBr IV	CNBr V
0[c]	(200)	(100)	(200)	(400)	(150)	(350)
1	Ile (185)	Ile (47)	Val (179)	Pro (333)	Leu (139)*	Ser (189)
2	Leu (141)	Leu (52)	Ser (138)	Ile (300)*	Pro (96)	Lys (269)
3	Gly (ND)[d]	Gly (22)	His (ND)	CM-Cys (256)	Val (107)*	Tyr (333)
4	Gly (102)	Gly (28)	His (ND)	Leu (315)	Ala (103)	Gln (245)
5	His (130)	His (ND)	Asn (0)[f]	Pro (360)	Asp (63)	Glu (261)
6	Leu (115)	Leu (31)	Leu (149)	Ser (240)	Gln (72)	Asp (186)
7	Asp (81)	Asp (33)	Thr (101)	Lys (267)	Asp (52)	Thr (132)
8	Ala (109)	Ala (17)	Thr (77)*	Asp (243)	Gln (44)	CM-Cys (153)
9	Lys (123)	Lys (32)	Gly (71)	Tyr (266)	CM-Cys (61)	Tyr (176)
10	Gly (61)		Ala (79)	Ala (271)	Ile (57)	Gly (135)
11	Ser (53)		Thr (64)	Glu (317)*	Arg (43)	Asp (119)
12	Phe (59)		Leu (88)*	Val (263)	His (17)	Ala (146)
13	Pro (43)		Ile (91)	Gly (171)	Tyr (24)	Gly (102)
14	Trp (66)[e]		Asn (27)	Arg (ND)	Glu (31)	Ser (93)
15	Glu (57)		Glu (21)	Val (307)	Gly (25)	Ala (131)
16	Ala (66)		Gln (38)	Gly (186)	Ser (32)	Phe (111)
17	Lys (45)		Trp (0)[g]	Tyr (316)	Thr (21)	Ala (123)
18	Met (ND)[f]		Leu (56)*	Val (265)	Val (27)	Val (67)
19	Val (51)		Leu (ND)	Ser (179)	Pro (46)*	His (ND)
20	Ser (35)		Thr (38)	Gly (142)	Glu (21)	Asp (61)
21	His (44)		Thr (37)	Trp (20)*	Lys (14)	Leu (80)
22	His (30)		Ala (30)	Gly (148)	Lys (23)	Glu (48)
23	Asn (0)[g]		Lys (24)	Arg (ND)		Glu (51)
24	Leu (27)		Asn (19)	Asp (65)		Asn (37)
25	Thr (20)		Leu (41)	Ala (74)		Thr (26)
26	Thr (45)			Asn (76)		Trp (0)[h]
27	Gly (ND)			Phe (40)		Tyr (33)
28	Ala (44)			Lys (46)		Ala (46)
29	Thr (30)			Phe (41)		Thr (21)
30	Leu (34)					Gly (29)
31						Ile (24)
32						Leu (24)
33						Ser (31)
34						Phe (16)

[a] Pth residues quantitated by amino acid analysis after HI hydrolysis, except those marked with asterisk which were quantitated by gas chromatography. Results given are for a single analysis; however, except for CNBr I, analysis of fragments and β chain were repeated two to five times on different preparations. [b] After purification by high-voltage electrophoresis. [c] Initial nmol applied to sequencer. [d] Not determined or lost; identified on repeated analysis. [e] Identified by gas chromatography and quantitated from increased Gly and Ala. [f] Confirmed in T2C, see Table X. [g] Carbohydrate attached. [h] Established from chymotryptic peptides (Lee, Kurosky, and Bowman, to be published).

obtained in this study. Since the amino-terminal of CNBr VI reported by Barnett et al. was a tryptophanyl residue, this fragment probably occurred as a result of an anomalous cleavage. Two such anomalous fragments were identified in this study, CNBr Ia and CNBr 2C2, which were the result of cleavage at tryptophan, probably on the amino-terminal side. The yield of CNBr Ia (18%) was relatively substantial, although somewhat exaggerated because of the fact that we could estimate it only relative to the recovery of CNBr I. Anomalous cleavages of protein by cyanogen bromide have also been reported by Sterner et al. (1974). These authors obtained three significant CNBr cleavage products from yeast inorganic pyrophosphatase in yields of about 15%, in addition to several others with yields of less than 5%. Only one of the anomalously generated fragments was characterized with regard to the type of bond cleaved. Their CNBr VII was an anomalous hydrolysis product of an X–Pro bond; however, this cleavage may have resulted from acid catalysis of Asp–Pro, as reported by Piszkiewicz et al. (1970). Our evidence suggests that some tryptophanyl (probably X–Trp) bonds are also susceptible to cleavage by cyanogen bromide, but to a lesser extent than methionyl peptide bonds. Whether or not this cleavage is unique to glycoproteins is an open question.

The adjusted compositional molecular weight estimate of 35 172 ± 371 for the β chain is in reasonable agreement with the molecular weight determined by ultracentrifugation (38 500 ± 4100) shown in Table VII, as well as the sum of the compositional molecular weights of the CNBr fragments (34 797 ± 223) given in Table VI. Ultracentrifugational analysis of the molecular weight of the hp β chain is problematic because of protein aggregation. Attempts to avoid these difficulties by succinylation (Black et al., 1970) may have led to an exaggerated molecular weight. Since the succinylation was not followed by reaction with hydroxylamine, the many hydroxyl groups on the protein and on the carbohydrate portion may have been succinylated as well. It should also be noted that discrepancies in the partial specific volume reported by various authors have raised some doubt concerning the reported molecular weights of Hp 1-1 (Putnam, 1975). The apparent molecular weight obtained by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (42 000) was undoubtedly excessive, due to the occurrence of carbohydrate on the β chain. The presence of covalently bound carbohydrate has been reported to increase molecular weight estimations by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (Segrest et al., 1971).

Molecular weight estimates of the CNBr fragments by gel filtration and sodium dodecyl sulfate–urea–polyacrylamide gel electrophoresis were in reasonable agreement (±18%, Swank and Munkres, 1971) with the molecular weights cal-

TABLE X: Composition and Partial Sequence of Peptide T2C.[a]

Amino Acid	Residue Number[b]	Sequence Analysis[d] Edman Cycle	Residue (nmol)
Lysine	1.0 (1)	1	Met (112)[e]
Histidine	1.7 (2)	2	Val (125)
Arginine	0	3	Ser (ND)
CM-cysteine	0	4	His (85)
Aspartic acid	2.5 (2)	5	His (65)
Threonine	4.3 (5)	6	Asn (0)[f]
Serine	1.3 (1)	7	Leu (77)
Glutamic acid	2.5 (2)	8	Thr (39)
Proline	0	9	Thr (59)
Glycine	1.5 (1)	10	Gly (29)
Alanine	1.9 (2)	11	Ala (36)
Valine	1.1 (1)	12	Thr (27)
Methionine	0.9 (1)	13	Leu (30)
Isoleucine	1.0 (1)	14	Ile (22)
Leucine	3.8 (4)	15	Asn (20)
Tyrosine	0	16	Glu (20)
Phenylalanine	0	17	Gln (21)
Tryptophan	ND (1)[c]	18	Trp (16)[g]
Glucosamine	3.1 (3)	19	Leu (14)
Residues	18–41	20	Leu(21)

[a] Tryptic peptide isolated from a tryptic hydrolysate of *S*-carbamoylmethylated hp β chain (see text). [b] After 24-h hydrolysis. Based on 1 lysine. [c] From a chymotryptic peptide (see Table IX). [d] Automated analysis. Pth residues identified by amino acid analysis after HI hydrolysis, except as indicated. [e] Pth derivative identified by gas chromatography. [f] No residue identified because of carbohydrate attachment to Asn. [g] Calculated from increased Gly and Ala.

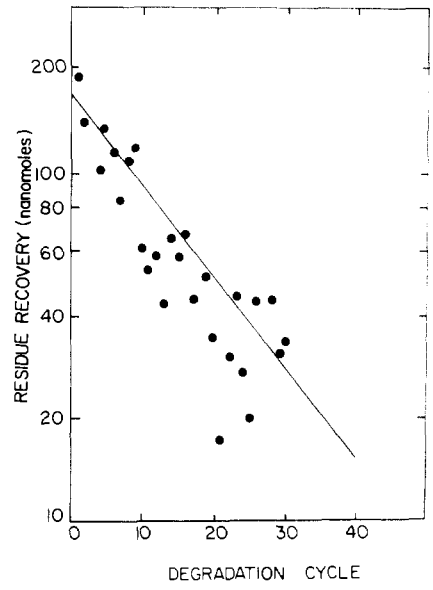

FIGURE 7: Semilogarithmic plot of Pth residue recoveries obtained from analysis of hp β chain. About 8 mg of β chain was applied to the reaction cup. Residue recoveries were calculated from amino acid analysis after HI hydrolysis of the Pth derivatives.

culated from compositional analysis. The exceptions are CNBr IV and CNBr V, which differed 42 and 31%, respectively, from the lowest estimated values. The reason for these differences is presently unknown; however, a number of small polypeptides, e.g., bovine trypsin inhibitor, ACTH, glucagon, and polymyxin, showed similar anomalous behavior (48–83% differences) during polyacryalamide gel electrophoresis (Swank and Munkres, 1971). Our compositional analyses of CNBr IV and V are compatible with the partial sequence data obtained from tryptic and chymotryptic peptides of whole β chain, as well as from CNBr fragments (Kurosky, Lee, Barnett, Hay, and Bowman, to be published). In addition, tryptic peptide maps of CNBr IV and V gave profiles in agreement with the calculated number of lysyl and arginyl residues in these fragments. Finally, a plasmin cleavage fragment, comprising CNBr V, IV, and the last 9 residues of CNBr III (Kurosky et al., 1974b), gave a molecular weight of 13 200 after sodium dodecyl sulfate–urea–polyacrylamide gel electrophoresis, which agrees with the compositional molecular weight of 13 737 ± 92 for this fragment.

The occurrence of a small amount of carbohydrate in CNBr IV and V was probably due to some minor contamination (about 2%) from a portion of CNBr II, or, more likely, may be the result of some nonspecific carbohydrate attachment caused by reaction with cyanogen bromide analogous to cyanogen bromide activation of Sepharose. Further studies are underway to confirm the presence of galactosamine in the β chain, since it is conceivable that, in view of the large amount of carbohydrate, the peak observed as galactosamine may, in fact, be due to some amino acid–sugar complex.

Sequence analysis of the cyanogen bromide fragments (summarized in Figure 8) established the positions of some 139 residues in the hp β-chain molecule corresponding to 55% of the molecule. The alignment of CNBr I and II was achieved by sequence analysis of intact β chain. An overlapping sequence for CNBr III and IV was obtained from a plasmin fragment resulting from cleavage of the Lys–Phe bond at positions 128–129 (Kurosky et al., 1974b). CNBr V must represent the carboxyl-terminal fragment, since it contained no homoserine or homoserine lactone and yielded results (0.4 mol of Asn/mol of CNBr V) after CPA hydrolysis which were identical to CPA hydrolyses of intact β chain. Further evidence for carboxyl-terminal asparagine was provided by Barnett et al. (1970, 1972). In addition, tryptic peptides of intact β chain have also provided overlapping sequences in agreement with those reported in Figure 8 (Kurosky, Rasco, and Bowman, to be published; Barnett, 1971). Other sequence studies of human hp β chain have been reported (Hew, 1970; Malchy, 1970; Malchy and Dixon, 1973a; and Malchy et al., 1973). Hew (1970) characterized a large fragment overlapping CNBr II and III, which, for the most part, agrees with the CNBr alignment presented here. The only differences were residues 104 and 105, which were reported as Pro–Leu (also see Malchy and Dixon, 1973a), whereas we have found Leu–Pro. We have reexamined this region several times and have confirmed our initial findings. Residues 140–146 and 196–201, reported by Malchy (1970), agree with our data, except that we found residue 145 to be glutamine rather than glutamic acid.

The absence of an identifiable residue at position 23, together with the occurrence of glucosamine in peptide T2C, is reasonable confirmation that carbohydrate is attached at this residue. The presence of a threonyl residue at position 25 satisfied the general sequence requirement of Asn–X–Ser/Thr (Hunt and Dayhoff, 1970) for attachment of carbohydrate to Asn.

Cyanogen bromide cleavage of intact Hp 1-1 yielded a fragment after gel filtration which contained the β chain covalently linked by disulfide bonding to CNBr III. The α chains of Hp contain neither methionine nor phenylalanine. Since the amino acid sequence of CNBr III possessed only one half-Cys, the attachment of the α^1 chain must be to Cys-103. This evidence confirmed the identical findings of Malchy and Dixon

```
 1                5                10                15                20          CHO     25
                                                                                   |
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               30                35                40                                  101

Thr-Gly-Ala-Thr-Leu-Ile-Asn-Glu-Gln-Trp-Leu-Leu-Thr-Thr-Ala-Lys-Asn-Leu/ (55 residues) /Val-Met-Pro-

102  α           106               111               116               121                 126
     |
Ile-Cys-Leu-Pro-Ser-Lys-Asp-Tyr-Ala-Glu-Val-Gly-Arg-Val-Gly-Tyr-Val-Ser-Gly-Trp-Gly-Arg-Asp-Ala-Asn-

127              131               136               141               146                 151

Phe-Lys-Phe/Thr,Asx,His/Leu-Lys-Tyr-Val-Met-Leu-Pro-Val-Ala-Asp-Gln-Asp-Gln-Cys-Ile-Arg-His-Tyr-Glu-

152              156                                 193               198                 203

Gly-Ser-Thr-Val-Pro-Glu-Lys-Lys/ (31 residues) /Ala-Gly-Met-Ser-Lys-Tyr-Gln-Glu-Asp-Thr-Cys-Tyr-Gly-

204              208               213               218               223

Asp-Ala-Gly-Ser-Ala-Phe-Ala-Val-His-Asp-Leu-Glu-Glu-Asn-Thr-Trp-Tyr-Ala-Thr-Gly-Ile-Leu-Ser-Phe/ (25
          253
residues) /Asn.
```

FIGURE 8: Alignment of the amino acid sequences of hp β chain determined from the CNBr fragments. The numbering of the fragments was obtained from compositional analysis in Table I. The higher total residue value was chosen for those CNBr fragments in which a range of residues was given.

Protein	Sequences of Serine-195 Region										
	189	190	191	192	193	194	195*	196	197	198	199
hp β-chain	/Asp	Thr	Cys	Tyr	Gly	Asp	Ala	Gly	Ser	Ala	Phe/
Chymotrypsin A or B	/Ser	Ser	Cys	Met	Gly	Asp	Ser	Gly	Gly	Pro	Leu/
Trypsin	/Asp	Ser	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Val/
Elastase	/Ser	Gly	Cys	Gln	Gly	Asp	Ser	Gly	Gly	Pro	Leu/
Thrombin	/Asp	Ala	Cys	Glu	Gly	Asp	Ser	Gly	Gly	Pro	Phe/

FIGURE 9: Comparison of hp β chain to serine proteases in the serine-195 active-site region. Protease sequences were taken from Hartley (1970b). All serine proteases were bovine, except elastase which was porcine. The numbering is that of chymotrypsinogen A and corresponds to hp β-chain residues 199–209 in Figure 8.

(1973a) and Malchy et al. (1973). Two half-cystinyl residues were identified per molecule of CNBr IV and V. The isolation of CNBr IV from whole Hp 1-1 without any modification of the half-cystinyl residues negated any possibility of disulfide linkage between CNBr IV and V. Therefore, possible disulfide combinations are: (1) intrachain, (2) symmetrical interchain, or (3) asymmetrical interchain. The possibility of inter-β-chain disulfide bonds was eliminated by Malchy et al. (1973) who isolated a labeled half-molecule of Hp after ^{35}S sulfitolysis of Hp 1-1. Only one major labeled sequence was obtained, which represented the α-chain sequence about Cys-21. Therefore, the disulfide linkages in CNBr IV and in V are intrachain. Our evidence for the disulfide bond arrangement in human Hp 1-1 (Kurosky et al., 1974b) essentially agrees with that presented by Malchy et al. (1973).

Comparison of hp β-chain sequences to sequences reported for the chymotrypsin family of serine proteases was made elsewhere (Kurosky et al., 1974a,b). A more complete comparative study will be undertaken after completion of the hp β-chain primary structure. However, several relevant comments regarding new data presented here need mentioning. The number of residues calculated for the β chain (251–257) is comparable to the number of residues reported for many of the serine proteases (Dayhoff, 1972). Furthermore, sequence analysis of the amino-terminal region of CNBr V established the hp β-chain sequence, which is positionally analogous to the active site Ser-195 sequence. This comparison is illustrated in Figure 9. Although the hp β chain compares reasonably well to the serine proteases in the Ser-195 region, the most notable exception is the replacement of alanine for Ser-195. Since all of the residues in the β chain analogous to the catalytic residues in the serine proteases are dissimilar, a proteolytic function in haptoglobin comparable to that defined for the serine proteases could not be possible. As indicated in Figure 9, the serine loop (Cys-191 to Cys-220, chymotrypsinogen numbering) is conserved. Interestingly, the specificity residue for the trypsin-like enzymes, Asp-189, is also found in the hp β chain. This emphasizes the possibility that, although the proteolytic function was lost during evolution, hp β chain may be utilizing the binding capacity of the "tosyl-hole" region for its function, which presumably is the binding of hemoglobin. Chemical modification studies will better define the function of haptoglobin and will establish the regions essential for the binding of hemoglobin.

Note Added in Proof

Repeated reexamination of the amino terminus of the hp α chain sequence (Black and Dixon, 1968) indicates that Asp-3 does not exist. Automated sequencer evidence points to the fact that Asn-2 undergoes a concerted deamidation and $\alpha \rightarrow \beta$ rearrangement resulting in only partial cleavage of residue 2 and extensive carry-over of subsequent residues, i.e., serine was found in positions 3 and 4, glycine in 4 and 5, etc. (see Table IV).

Acknowledgments

We are indebted to Mr. Horace D. Kelso, Mr. Terry M. Ward, and Mr. Frank Scott for expert technical assistance. Discussions with Dr. D. R. Barnett were extremely valuable. We thank Dr. Julian Smith, Department of Gynecology, M. D. Anderson Hospital, Houston, Texas, for ascites fluid.

References

Ambler, R. P. (1967), *Methods Enzymol. 11,* 155.
Barnett, D. R. (1971), Ph.D. Dissertation, University of Texas Medical Branch, Galveston.
Barnett, D. R., Lee, T.-H., and Bowman, B. H. (1970), *Nature (London) 225,* 938.
Barnett, D. R., Lee, T.-H., and Bowman, B. H. (1972), *Biochemistry 11,* 1189.
Bennett, C. (1967), *Methods Enzymol. 11,* 330.
Black, J. A., Chan, G. F. Q., and Dixon, G. H. (1970), *Can.*

J. Biochem. 40, 123.
Black, J. A., and Dixon, G. H. (1968), *Nature (London) 218,* 736.
Bradshaw, R. A., Garner, W. H., and Gurd, F. R. N. (1969), *J. Biol. Chem. 244,* 2149.
Cheftel, R.-I., and Moretti, M. J. (1966), *C. R. Hebd. Seances Acad. Sci., Ser. D 262,* 1982.
Cleve, H., Bowman, B. H., and Gordon, S. (1969), *Humangenetik 7,* 337.
Cleve, H., Gordon, S., Bowman, B. H., and Bearn, A. G. (1967), *Am. J. Hum. Genet. 19,* 713.
Connell, G. E., Dixon, G. H., and Smithies, O. (1962), *Nature (London) 193,* 505.
Dayhoff, M. O. (1972), *in* Atlas of Protein Sequence and Structure, Vol. 5, Washington, D.C., National Biomedical Research Foundation, p D-99.
Fish, W. W., Mann, K. G., and Tanford, C. (1969), *J. Biol. Chem. 244,* 4989.
Fuller, G. M., Rasco, M. A., McCombs, M. L., Barnett, D. R., and Bowman, B. H. (1973), *Biochemistry 12,* 253.
Gerbeck, C. M., Rafelson, M. E., Jr., and Bezkorovainy, A. (1965), *Biochim. Biophys. Acta 101,* 229.
Giblett, E. R. (1975), in Structure and Function of the Plasma Proteins, Allison, A. C., Ed., London, England, Plenum Press, p 56.
Glossman, H., and Neville, D. M., Jr. (1971), *J. Biol. Chem. 246,* 6339.
Gordon, S., Bowman, B. H., and Bearn, A. G. (1967), *Am. J. Hum. Genet. 19,* 713.
Gross, E. (1967), *Methods Enzymol. 11,* 238.
Hartley, B. S. (1970a), *Biochem. J. 119,* 805.
Hartley, B. S. (1970b), *Philos. Trans. R. Soc. London Ser. B 257,* 77.
Hew, C.-L. (1970), Ph.D. Dissertation, University of British Columbia.
Hunt, L. T., and Dayhoff, M. O. (1970), *Biochem. Biophys. Res. Commun. 39,* 757.
Jayle, M.-F., Boussier, G., and Tonnelat, J. (1956), *Bull. Soc. Chim. Biol. 38,* 343.
Kurosky, A. (1972), Ph.D. Dissertation, University of Toronto.
Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H., and Bowman, B. H. (1974a), *Biochem. Genet. 11,* 279.
Kurosky, A., Kim, H.-H., Barnett, D. R., Rasco, M. A., Touchstone, B., and Bowman, B. H. (1974b), *Proc. Protides Biol. Fluids 22,* 597.
Kurosky, A., Kim, H.-H., and Touchstone, B. (1976a), *Comp. Biochem. Physiol. 55B,* 453.
Kurosky, A., Markel, D. E., Touchstone, B., and Peterson, J. W. (1976b), *J. Infect. Dis. 133,* 514.
Malchy, B. L. (1970), Ph.D. Dissertation, University of British Columbia.
Malchy, B., and Dixon, G. H. (1973a), *Can. J. Biochem. 51,* 249.
Malchy, B., and Dixon, G. H. (1973b), *Can. J. Biochem. 51,* 321.
Malchy, B., Rorstad, O., and Dixon, G. H. (1973), *Can. J. Biochem. 51,* 265.
Marshall, R. D., and Neuberger, A. (1972), *Glycoproteins 5,* 224.
Olafson, R. W., Jurášek, L., Carpenter, M. R., and Smithie, L. B. (1975), *Biochemistry 14,* 1168.
Panyim, S., and Chalkley, R. (1969), *Arch. Biochem. Biophys. 130,* 337.
Piszkiewicz, D., Landon, M., and Smith, E. L. (1970), *Biochem. Biophys. Res. Commun. 40,* 1173.
Putnam, F. W. (1975), *Plasma Proteins, 2nd ed, 2,* 1.
Schultze, H. E., Haupt, H., Heide, K., and Heimberger, N. (1963), *Clin. Chim. Acta 8,* 207.
Segrest, J. P., Jackson, R. L., Andrews, E. P., and Marchesi, V. T. (1971), *Biochem. Biophys. Res. Commun. 44,* 390.
Shim, B.-S., and Bearn, A. G. (1964), *J. Exp. Med. 120,* 611.
Shim, B.-S., Lee, T.-H., and Kang, Y. S. (1965), *Nature (London) 207,* 1264.
Smithies, O. (1955), *Biochem. J. 61,* 629.
Smithies, O., Connell, G. E., and Dixon, G. H. (1962), *Am. J. Hum. Genet. 14,* 14.
Spencer, R. L., and Wold, F. (1969), *Anal. Biochem. 32,* 185.
Sterner, R., Noyes, C., and Heinrikson, R. L. (1974), *Biochemistry 13,* 91.
Sutton, H. E. (1970), *Prog. Med. Genet. 7,* 163.
Swank, R. T., and Munkres, K. D. (1971), *Anal. Biochem. 39,* 462.
Titani, K., Hermodson, M. A., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1972), *Nature (London) New Biol. 238,* 35.
Weber, K., and Osborn, M. (1969), *J. Biol. Chem. 244,* 4406.
Woodworth, R. C., and Clark, L. G. (1967), *Anal. Biochem. 18,* 295.