Characterization of the B-chain of Human Plasma $\alpha_2$HS-Glycoprotein

THE COMPLETE AMINO ACID SEQUENCE AND PRIMARY STRUCTURE OF ITS HETEROGLYCAN*

(Received for publication, October 6, 1982)

Fumitake Gejyo‡, Jinq-Lin Chang, Willy Bürgis, Karl Schmid§, Gwynneth D. Offner, and Robert F. Troxler

From the Department of Biochemistry, Boston University School of Medicine, Boston University Medical Center, Boston, Massachusetts 02118

Herman Van Halbeek, Lambertus Dorland, Gerrit J. Gerwig, and Johannes F. G. Vliegenthart

From the Department of Bio-organic Chemistry, University of Utrecht, Crousestraat 79, NL-3522 AD Utrecht, The Netherlands

$\alpha_2$HS-Glycoprotein, a normal human plasma protein, was recently shown to consist of two polypeptide chains. In the present study, we have separated these two chains from one another and have elucidated the complete primary structure of the B-chain. Employing automated Edman degradation, the polypeptide moiety of this chain was shown to consist of 27 amino acid residues with an unequal distribution of the neutral and charged amino acid residues. The first 20 residues are uncharged, whereas the carboxyl-terminal heptapeptide contains all charged residues. Utilizing 300-MHz H-NMR spectroscopy, the carbohydrate unit proved to be a trisaccharide consisting of sialic acid, galactose, and N-acetylgalactosamine O-glycosidically linked to serine (residue 6). The structure of the B-chain was found to be as follows.

\[
\begin{align*}
\text{NeuAc} & \quad \text{Gal} & \quad \text{GalNAc} \\
\text{H}_2\text{N-Thr-Val-Val-Gln-Pro-Ser-Val-Gly-Ala-Ala-Ala-Gly-Pro} & \quad \text{Val-Pro-Pro-Cys-Pro-Gly-Ile-His-Phe-Lys-Val-Val-Gly-Arg} & \quad \text{Arg-Ile-His-Phe-Lys-Val-COOH}
\end{align*}
\]

*This work was supported by a grant from Hoffmann-LaRoche Inc., Nutley, NJ, by Grants GM-10374 and HL-13262 from the National Institute of Health, United States Public Health Service, by Grante PCM 81 05883 and PCM 82 03177 from the National Science Foundation, by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), and by Grant UUKC-OC 79-13 from the Netherlands Foundation for Cancer Research (KWF). 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.

‡ Present address, Second Department of Internal Medicine, Niigata University School of Medicine, Niigata, Japan.
§ Present address, Department of Biochemistry, University of Bern, Bern, and Kantonsspital Aarau, Switzerland.
To whom reprint requests should be addressed at, Boston University School of Medicine, 80 East Concord Street, Boston, MA 02118.

Thus, the molecular weight of the B-chain is 3386. Evaluation of the polypeptide chain by the procedure of Chou and Fasman (Chou, P. Y., and Fasman, G. D. (1979) Adv. Enzymol. 47, 45-148) predicts that the B-chain has two $\beta$-turns. Thereby, the carbohydrate unit which is linked to the Ser residue located in the first $\beta$-turn appears to be directed away from the protein. The second $\beta$-turn probably includes the Cys residue which links the B- to the A-chain. In agreement with the CD analysis, the B-chain lacks $\beta$-conformation but possesses a short $\alpha$-helical region.

$\alpha_2$HS-Glycoprotein, a normal human plasma globulin (1, 2), has been demonstrated to be associated with a considerable number of biological functions. The plasma level of this protein was reported to be significantly decreased in certain malignant (3) and inflammatory diseases (4), malnutrition (5), and Paget's disease (6). Furthermore, in bone, $\alpha_2$HS-glycoprotein was found to be concentrated up to 30-fold with respect to other plasma glycoproteins (7, 8). This glycoprotein also displays opsonic properties (9) and promotes endocytosis (10). Of particular interest is the observation that in certain cancer patients the serum level of this protein correlates with diminished lymphocyte activity, suggesting a relationship between this protein and a depression of cell-mediated immunity (11).

The physicochemical and chemical properties of $\alpha_2$HS-glycoprotein have been described earlier (2, 12) and, pertinent to the present study, revealed two NH$_2$- and two carboxy-terminal amino acids/mol of protein, indicating the presence of two polypeptide chains designated A and B. In this investigation, separation of these two chains from one another and the elucidation of the amino acid and monosaccharide sequences of the B-chain of $\alpha_2$HS-glycoprotein are reported.

EXPERIMENTAL PROCEDURES

Materials—L-1-1-Tosylamido-2-phenylethyl chloromethyl ketone-trypsin (three times recrystallized) was purchased from Worthington. Carboxypeptidase A and S-β-(4-pyridylethyl)-l-cysteine were from Sigma. 4-Vinylpyridine, polybreyne, and D$_2$O (99.96 atom % deuteron) were from Aldrich. Bio-Gel P-4 was from Bio-Rad, and Sephadex G-75 and G-10 were from Pharmacia. DEAE-cellulose was from Whatman. Chemicals for amino acid sequencing were purchased from Beckman. Purine was a gift from Dr. Takashi Okuyama (Seikagaku Kogyo, Tokyo, Japan).

1 The abbreviations used are: $\alpha_2$HS-glycoprotein, $\alpha_2$-glycoprotein discovered by J. F. Heremans and K. Schmid; PTH, phenylthiohydantoin; GalNAc, N-acetylgalactosamine.
Reduction and Alkylation—Human α2HS-glycoprotein was prepared as described earlier (2) and further purified by a method published recently (12). The glycoprotein (28 mg) was prepared with diethiothreitol (20-fold molar excess over the disulfide bonds) in 1.2 ml of 0.13 M Tris-HCl buffer, pH 7.6, containing 6 M guanidine HCl and 0.2% EDTA at room temperature for 4 h. Subsequently, the reduced protein was alkylated with 4-vinylpyridine (3-fold molar excess over the reducing agent) at room temperature for 2 h (13). The reaction mixture was then passed through a Sephadex G-75 column (1.2 × 150 cm), previously equilibrated with 0.1% ammonium bicarbonate.

Molecular Weight Determination—The apparent molecular weight of the B-chain was determined by polyacrylamide gel electrophoresis in 0.1% sodium dodecyl sulfate and 8 M urea (14). Reference proteins or peptides were lysozyme (Mr = 14,600) and cytochrome c fragments I (Mr = 8,279), II (Mr = 6,420), and III (Mr = 2,550) of sperm whale myoglobin. The gels were stained with either Coomassie brilliant blue or periodic acid-Schiff stain.

Sugar Analysis—After methanolation and N-reacetilation and trimethylation of the B-chain, the quantification of each monosaccharide (15) was carried out by capillary gas-liquid chromatography on an SE-30 column (25 m × 0.35 mm inner diameter). The oven temperature was programmed from 130 to 230 °C at 2 °C/min.

Circular Dichroism—Circular dichroism measurements of the B-chain were made with a Cary Model 61 spectropolarimeter. As solvents, 0.1 M NaF, 50% 2-chloroethanol, and 6 M urea were used. The results were expressed as reduced mean residue ellipticities (16).

Amino Acid Analysis—The B-chain and peptides were hydrolyzed with 2 M distilled 6 N HCl in evacuated sealed glass tubes under N₂ at 110 °C for 20 h. Hydrolysates were analyzed with the aid of a Beckman 119 CL automated amino acid analyzer in combination with a Model 126 data system. Values for threonine and serine were not corrected for destruction during hydrolysis and no correction was applied for the partial liberation of valine.

NH₂- and Carboxyl-terminal Amino Acid Analyses—The NH₂-terminal amino acid of the B-chain was determined by the 5-dimethylaminonaphthalene-1-sulfonylethyl technique (17). The carboxyl-terminal sequence was established by digestion with carboxypeptidase A (18). For this experiment, the B-chain (0.3 mg) was treated with 0.1 N NaOH (for explanation, see "Results") in 0.2 M N-ethylmorpholine acetate buffer, pH 8.5, at 25 °C. The liberated amino acids were determined by amino acid analysis.

Tryptic Digestion—The B-chain (4 mg) was digested with 0.08 mg of L-1-tosylamino-2-phenylethyl chloromethyl ketone-trypsin in 0.9 ml of 0.1 M N-ethylmorpholine acetate buffer, pH 8.6, at 25 °C for 6 h. The reaction was terminated by adding 9% formic acid to lower the pH of the solution to 2.5, and the resulting solution was then applied to a Bio-Gel P-4 column (1.2 × 220 cm) equilibrated with 0.1 M formic acid, pH 2.3.

Automated Edman Degradation—Automated Edman degradation was carried out by stepwise degradation with the aid of a Beckman 890C Sequencer equipped with a cold trap using program 121078 with SI wash (19). Samples were applied to the cup in 15% acetic acid (0.3 ml) containing 2 mg of Polybrene using Beckman Sample Application program 02772. PTH-derivatives were obtained as described earlier (20) and identified by high pressure liquid chromatography (21) or, in some cases, by gas-liquid chromatography (22) or by back hydrolysis (23). PTH-S-pyr-dylleucysteine, which is water soluble, was identified by high pressure liquid chromatography.

Preparation of the Glycopeptide Fraction—The B-chain (6 mg) was dissolved in 0.5 ml of water and subjected to exhaustive pronase digestion at pH 8.0. Pronase (0.3 mg) was added three times at 4-h intervals and the solution was incubated at 60 °C. The reaction mixture was then concentrated to 0.4 ml in vacuo and applied to a Sephadex G-10 column (0.9 × 105 cm) equilibrated and eluted with water. The effluent was monitored at 230 nm, and an aliquot of each fraction (0.6 ml) was assayed for neutral sugars (24). The fractions containing sugar were combined, concentrated, and lyophilized.

1H-NMR Spectroscopy—Prior to NMR analysis (25), the glycopeptide preparation was exchanged repeatedly in D₂O employing intermediate lyophilization. For 1H-NMR spectroscopic analysis, a Bruker WM-500 spectrometer (SON Facility, Nijmegen, The Netherlands) was employed operating in the Fourier transform mode at a proton frequency of 500 MHz and equipped with a Bruker AS-2000 computer. Resolution enhancement of the spectrum was achieved by Lorentzian to Gaussian transformation from quadrature phase detection, followed by a complex Fourier transformation (25).

RESULTS

α2HS-Glycoprotein, after reduction and alkylation, yielded two fractions on gel filtration (Fig. 1). The major fraction (A) was eluted near the void volume and contained the A-chain. The minor fraction (B) comprising the B-chain eluted just before the salt peak (S). The B-chain preparations of several experiments were pooled and lyophilized. On sodium dodecyl sulfate-polyacrylamide gel electrophoresis, a single band with an apparent Mr of 5400 was observed. This band stained with both Coomassie brilliant blue and periodic acid-Schiff stain, which demonstrated that this chain contains carbohydrate with sialic acid. NH₂- and carboxyl-terminal amino acid analyses showed the presence of only Thr and Val, respectively, thus confirming the homogeneity of the B-chain preparation.

Circular Dichroism—The circular dichroism spectrum of
the B-chain in 0.1 M NaF showed a negative ellipticity peak at 200 nm, indicating the presence of a small α-helical region and the absence of β-pleated sheet confirmation (Fig. 2). In the presence of 2-chloroethanol, the percentage of α-helical structure increased slightly and in 6 M urea the CD spectrum revealed essentially random conformation.

Evaluation of the B-chain by the method of Chou and Fasman (26) predicts the presence of two β-turns (residues 4 to 7 and 18 to 21) and a short α-helix (residues 7 to 12) and the absence of β-conformation.

**Amino Acid and Carbohydrate Composition of the B-Chain**

**TABLE I**

| Amino Acid and Carbohydrate composition of B-chain from α₂HS-glycoprotein | Residue/mol |
|---|---|
| Amino Acid | Residue/mol |
| Threonine | 0.77 (1)* |
| Serine | 1.06 (1) |
| Glutamic acid | 1.19 (1) |
| Proline | 5.02 (5) |
| Glycine | 2.91 (3) |
| Alanine | 3.23 (3) |
| Valine | 5.48b (6) |
| Isoleucine | 1.00 (1) |
| Phenylalanine | 0.89 (1) |
| Lysine | 1.04 (1) |
| Histidine | 0.96 (1) |
| Arginine | 1.90 (2) |
| Half-Cystinee | 0.82 (1) |
| Total amino acid residue | 26.25 (27)b |
| Monosaccharidef | 3.26 (1) |
| N-Acetylmuramic acid | 1.1 (1) |
| Galactose | 1.0f (1) |
| Galactosamine | 0.9 (1) |

*Numbers in parentheses are determined after sequence analysis.

*This low value is probably due to partial hydrolysis of the Val-Val bond.

*aAsp, Leu, and Tyr were found to be present in small amounts and the corresponding values were 0.18, 0.07, and 0.04, respectively. Met and Trp were not detected.

*bTraces of Man, GlcN, xylose, and Glc were also noted.

flip assumed to be 1.0.

**FIG. 3. Summary of the sequence studies of the B-chain of α₂HS-glycoprotein.** CHO, carbohydrate.

---

**Complete Structure of the B-chain of α₂HS-Glycoprotein**

The amino acid composition (Table I) revealed that Pro and Val comprise approximately 40% of the amino acid residues. It is noteworthy that Asx, Leu, Met, Tyr, and Trp are not present but are contained in the A-chain (12). Assuming the presence of 1 mol of Ile/mol of the B-chain, the total number of amino acid residues was calculated to be 27.

The carbohydrate moiety accounts for 19.4% of the weight of the B-chain. The carbohydrate composition consisting of NeuAc, Gal, and GalNac (Table I) suggests the presence of an O-glycosidic carbohydrate unit. The molecular weight of the B-chain calculated from the amino acid and carbohydrate composition was 3386, in contrast to the apparent M, = 5400 mentioned above. This observation is consistent with the fact that glycoproteins are known to afford erroneously higher molecular weight values on sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

**Amino Acid Sequence**—Automated Edman degradation of the B-chain provided the complete amino acid sequence (Fig. 3 and Table II). Thr, which is located in position 1, was identified as PTH-Thr in the organic phase and obtained in the expected yield. This finding indicates that the oligosaccharide chain is not bound to this amino acid. Cycle 6 yielded unusually small amounts of PTH-serine in the organic phase, suggesting that the carbohydrate unit is linked to this amino acid and that the B-chain contains two species of polypeptide chains, the predominant one glycosylated at residue 6, and a minor one lacking a carbohydrate unit. Furthermore, the single serine residue of this chain must reside in position 6 as the residues in all other positions were positively identified. Digestion with carboxypeptidase confirmed the sequence His-Phe-Lys-Val-COOH (the corresponding data are summarized below).

The entire sequence was confirmed by automated Edman degradation of the tryptic peptides derived from this chain. These peptides were separated from each other by gel filtration affording six fractions (Fig. 4), and their amino acid compositions are given in Table III. The amino acid sequences of the peptides T-1a and T-1b were found to be identical (Table IV and Fig. 3) except that residue 6 in T-1a could not be positively identified. It was concluded that residue 6 in T-1a is a serine O-glycosidically linked to a carbohydrate unit.

---

**TABLE II**

| Residue/mol |
|---|
| CHO, carbohydrate |

---

**TABLE III**

| Residue/mol |
|---|
| CHO, carbohydrate |

---

**TABLE IV**

| Residue/mol |
|---|
| CHO, carbohydrate |

---

**FIG. 3. Summary of the sequence studies of the B-chain of α₂HS-glycoprotein.** CHO, carbohydrate.
This is consistent with the fact that T-1a was eluted before T-1b on gel filtration. Peptide T-5 was the dipeptide Ile-Arg which could be placed at residues 22 and 23. Peptide T-3, the carboxyl-terminal peptide, represents residues 24 to 27. Peptide T-4 could be identified as residues 24 to 26. Fraction T-2 contained only Val (Table III) which was also identified as PTH-Val (Fig. 3). As expected from the carboxyl-terminal sequence - Lys-Val, Val was liberated only to a molar yield of 0.86, 0.52, 0.26, and 0.24 residues/mol of β-chain. However, when the incubation period was extended to 1 h, in addition to Val, Lys, Phe, and His were liberated (0.86, 0.52, 0.26, and 0.24 residues/mol of β-chain, respectively), and after 3 h the corresponding values were 0.91, 0.65, 0.32, and 0.29, respectively. These data are in full agreement with the complete covalent sequence mentioned above.

**Carbohydrate Structure**—The 500-MHz $^1$H-NMR spectrum (Fig. 5) of the glycopeptide fraction appears relatively complex, because many signals stemming from amino acid protons of the relatively large peptide portion occur in the same spectral regions as structural reporter group signals of the oligosaccharide moiety (in the ranges from δ = 2.7 to 3.9 ppm and from δ = 2.8 to 1.7 ppm). However, since the carbohydrate composition of the glycopeptide had been determined (Table I), the carbohydrate structural reporter group signals (Table V) could be recognized and assigned to the primary structure of the carbohydrate units.

### Table II

**Automated Edman degradation of the B-chain of α_{2}HS-Glycoprotein**

| Cycle No. | Amino acid | Yield* nmol
|-----------|------------|--------------
| 1         | Thr        | 23.5         |
| 2         | Val        | 40.0         |
| 3         | Val        | 47.3         |
| 4         | Glu        | 57.8         |
| 5         | Pro        | 27.8         |
| 6         | Ser        | 5.0          |
| 7         | Val        | 59.8         |
| 8         | Gly        | 29.2         |
| 9         | Ala        | 64.3         |
| 10        | Ala        | 67.5         |
| 11        | Ala        | 61.2         |
| 12        | Gly        | 33.7         |
| 13        | Pro        | 19.7         |
| 14        | Val        | 22.6         |
| 15        | Val        | 24.9         |
| 16        | Pro        | 20.0         |
| 17        | Pro        | 15.3         |
| 18        | Cys        | —           |
| 19        | Cys        | —           |
| 20        | Pro        | 10.8         |
| 21        | Gly        | 15.7         |
| 22        | Arg        | 4.8          |
| 23        | Ile        | 33.6         |
| 24        | Arg        | 4.3          |
| 25        | His        | 5.2          |
| 26        | Phe        | 31.9         |
| 27        | Lys        | 4.8          |
| 28        | Val        | 2.4          |
| 29        | ND         | ND           |
| 30        | ND         | ND           |
| 31        | ND         | ND           |

*Numbers in parentheses were determined after sequence analysis.

### Table III

**Amino acid composition of tryptic peptides from B-chain of α_{2}HS-glycoprotein**

| Amino acid | T-1a | T-1b | T-5 | T-3 | T-4 | T-2 |
|------------|------|------|-----|-----|-----|-----|
| Asx        |      |      |     |     |     |     |
| Thr        | 0.55 (1)* | 0.58 (1) |     |     |     |
| Ser        | 1.06 (1) | 1.07 (1) |     |     |     |
| Glx        | 1.20 (1) | 1.17 (1) |     |     |     |
| Pro        | 5.10 (5) | 5.23 (5) |     |     |     |
| Gly        | 3.05 (3) | 3.12 (3) |     |     |     |
| Ala        | 3.18 (3) | 3.25 (3) |     |     |     |
| Val        | 3.95 (5) | 3.75 (5) | 1.04 (1) | 1.00 (1) |
| Cys*       | 0.62 (1) | 0.70 (1) |     |     |     |
| Met        |        |        |     |     |     |
| Ile        |        |        | 1.10 (1) |     |     |
| Leu        |        |        |     |     |     |
| Tyr        |        |        |     |     |     |
| Phe        |        |        | 1.06 (1) | 1.18 (1) |
| Lys        |        |        | 1.00 (1) | 1.00 (1) |
| His        |        |        | 1.01 (1) | 1.19 (1) |
| Arg        | 1.00 (1) | 1.00 (1) | 1.00 (1) |     |     |

**Total number of residues**

|             | 21 | 4 | 3 | 1 |
|-------------|----|---|---|---|
| Position in sequence | 1-21 | 22-23 | 24-27 | 24-26 |

*Numbers in parentheses were determined after sequence analysis.

### Notes

- *Numbers in parentheses were determined after sequence analysis.
- *Low values are probably due to partial hydrolysis of Val-Val bonds.
- *Determined as S-pyridylethylcysteine.
- *PTH derivatives determined by high pressure liquid chromatography.
- *Confirmed by gas-liquid chromatography.
- *Confirmed by amino acid analysis after back hydrolysis.
- *Identified as PTH-S-pyridylethylcysteine.
- *Recoveries not corrected for internal standard recovery.
- *ND. No PTH-amino acid was detected at this or subsequent steps.
the chemical shift of the GalNAc H-1 (δ = 4.920 ppm) in conjunction with the value for its coupling constant J_{1,2} being 4.3 Hz. This anomeric signal is accompanied by a lower intensity doublet at δ = 4.911 ppm having the same coupling constant J_{1,2} of 4.3 Hz, pointing to heterogeneity of the peptide moieties of the glycopeptide fraction. This heterogeneity was also evident from the N-acetyl proton region; GalNAc of the major peptide component possessed its N-acetyl signal at δ = 2.015 ppm, while the minor constituent had this singlet at 2.008 ppm (Fig. 5). These findings are in agreement with the observation that, for glycopeptides with N-glycosidic carbohydrate chains (30, 31), the composition of the peptide moiety influences the chemical shifts of the structural reporter groups of GlcNAc (see Ref. 25) being attached to the peptide. Regarding the position of attachment between Gal and GalNAc, the glycosidic linkage was identified as β(1→3) on the basis of the chemical shift of GalNAc H-4 (δ = 4.230 ppm). Comparison of this chemical shift with that of H-4 of GalNAc (δ = 3.982 ppm) of the model compound GalNAco(1→0)N-acetylserylserine (spectrum not shown) revealed that the signal of H-4 of GalNAc of the carbohydrate unit of the glycopeptide has undergone a significant downfield shift. A similar downfield shift was observed when the chemical shift of the H-4 of GalNAc of chondroitin 6-sulfate (\ldots \rightarrow \rightarrow \rightarrow)GlcUAβ(1→3)GalNAcβ(1→\ldots) (where GlcUA is glucuronic acid) (32) was compared with that of the β-methylglycoside of GalNAc. The chemical shifts of the other GalNAc skeleton proton signals are only moderately affected by this β(1→3) substitution (e.g. the H-1 signal is shifted only from δ = 4.885 to 4.920 ppm; the H-2 signal found at δ = 4.003 ppm and that of H-3 located at δ = 3.98 ppm resolved from the bulk signal stemming from the larger part of the sugar skeleton protons). On the basis of these NMR data, the structure for the carbohydrate unit of the B-chain of α2HS-glycoprotein deduced was NeuAcα(2→3)Galβ(1→3)GalNAc(1→0)Ser.

### Table IV
Automated Edman degradation of the tryptic peptides derived from the B-chain of α2HS-glycoprotein

| Cycle No. | Peptide T-1a | Peptide T-1b |
|-----------|--------------|--------------|
|           | Amino acid   | Yield*       | Amino acid   | Yield*       |
| 1         | Thr          | 38.5 nmol    | Thr          | 19.2 nmol    |
| 2         | Val          | 78.2 nmol    | Val          | 64.9 nmol    |
| 3         | Val          | 55.3 nmol    | Val          | 59.7 nmol    |
| 4         | Gln          | 60.5 nmol    | Gln          | 45.5 nmol    |
| 5         | Pro          | 40.4 nmol    | Pro          | 42.6 nmol    |
| 6         | ND*          | —            | Ser          | 11.7 nmol    |
| 7         | Val          | 62.8 nmol    | Val          | 46.6 nmol    |
| 8         | Gly          | 33.4 nmol    | Gly          | 21.1 nmol    |
| 9         | Ala          | 55.9 nmol    | Ala          | 33.1 nmol    |
| 10        | Ala          | 54.4 nmol    | Ala          | 29.7 nmol    |
| 11        | Ala          | 54.6 nmol    | Ala          | 31.8 nmol    |
| 12        | Gly          | 10.3 nmol    | Gly          | 24.8 nmol    |
| 13        | Pro          | 16.6 nmol    | Pro          | 10.7 nmol    |
| 14        | Val          | 22.5 nmol    | Val          | 13.5 nmol    |
| 15        | Val          | 23.5 nmol    | Val          | 12.0 nmol    |
| 16        | Pro          | 10.0 nmol    | Pro          | 5.7 nmol     |
| 17        | Pro          | 9.2 nmol     | Pro          | 1.3 nmol     |
| 18        | Cys          | —            | Cys          | —            |
| 19        | Pro          | 5.6 nmol     | Pro          | 1.9 nmol     |
| 20        | Gly          | 4.8 nmol     | Gly          | 1.8 nmol     |
| 21        | Arg          | 1.2 nmol     | Arg          | 1.0 nmol     |

* PTH derivatives determined by high pressure liquid chromatography.
* ND, no amino acid was detected.
* Determined as PTH-S-pyridylethylcysteine.

### Table V
'H chemical shifts of structural reporter group protons of constituent monosaccharides for the carbohydrate moiety of the B-chain of α2HS-glycoprotein

| Residue | Reporter group | Chemical shift ppm |
|---------|----------------|-------------------|
| GalNAc  | H-1            | 4.920*            |
|         | H-4            | 4.230             |
|         | NAc            | 2.015*            |
| Gal     | H-1            | 4.564             |
|         | H-3            | 4.070             |
| NeuAc   | H-4            | 3.925             |
|         | H-3eq          | 1.797             |
|         | H-3x           | 2.749             |
|         | NAc            | 2.029             |

* Chemical shift slightly dependent on the nature of the peptide moiety (see text).
Complete Structure of the B-chain of α2HS-Glycoprotein

In the present study, the two polypeptide chains of α2HS-glycoprotein were separated from each other and the complete primary structure of the B-chain of this glycoprotein was elucidated. As to the amino acid sequence of the B-chain, the following is noteworthy. The first 20 residues are uncharged, 5 of the 6 valine (including two Val-Val sequences) and all proline and alanine (Ala-Ala-Ala) residues are within this region. However, this hydrophobic segment carries the carbohydrate unit linked to serine which probably renders this section of the peptide chain relatively hydrophilic. All charged residues are located within the seven amino acids at the carboxyl terminus. Evaluation of the B-chain according to Chou and Fasman (26) indicated that this chain does not assume β-conformation and that it probably possesses a short α-helical region. These findings are in agreement with results obtained by circular dichroism analysis. Two β-turns (33, 34) are predicted in the B-chain with the mentioned helix projected to immediately follow the first β-turn (35). According to this prediction, the carbohydrate would be located on the first β-turn. Similar observations have been reported for other proteins (36, 37). It is also of interest that the second β-turn includes the cysteine residue which links the B- and the A-chain.

The carbohydrate moiety elucidated by high resolution 1H-NMR spectroscopy was found to be comprised of a trisaccharide unit linked to Ser. This procedure affords simultaneously the precise structure of the carbohydrate-peptide linkage, information which is difficult to obtain by chemical techniques. References—We are very much obliged to Dr. Takashi Okuyama, Seikagaku Kogyo Co. Ltd., Tokyo, Japan, for the generous gift of pronase, to Dr. Marshall Elzinga, Brookhaven National Laboratory, Upton, Long Island, NY, for the gift of the cyanogen bromide fragments of myoglobin, to Dr. A. A. Pavia, Laboratoire de Biochimie, Avignon, France, for the reference compound GalNAcO(1→3)N-acetylserine, for NMR investigations, to L. Larsen, Institute of Laboratories, Massachusetts Department of Health, Jamaica Plain, MA, for the generous gift of the supernatant of Cohn fraction V from which α2HS-glycoprotein was isolated, and to George Crombie for performing amino acid analyses.

REFERENCES
1. Heremans, J. F. (ed) (1960) Les Globulines Séroïdes du Système Gamma, pp. 103–105, Arscia, Bruxelles and Masson, Paris
2. Schmid, K., and Bürgi, W. (1961) Biochem. Biophys. Acta 47, 440–453
3. Bradley, W. P., Biasco, A. P., Weiss, J. F., Alexander, J. C., Siverman, N. A., and Chretien, P. B. (1977) Cancer (Phila.) 40, 2264–2272
4. Lebreton, J. P., Joisel, F., Raoul, J. P., Lannuzel, B., Rogez, J. P., and Humbert, G. (1979) J. Clin. Invest. 64, 1118–1129
5. Schell, F. P., Thanankul, O., Supawan, V., and Pongpaw, P. (1980) Br. J. Nutr. 43, 381–383
6. Ashby, B. A., and Smith, R. (1980) Clin. Sci. (Lond.) 58, 435–438
7. Trifitt, J. T., Gebauer, U., Ashton, B. A., Owen, M., and Renolds, J. J. (1976) Nature (Lond.) 262, 226–227
8. Ashton, B. A., Hohling, H. J., and Trifitt, J. T. (1976) Calcif. Tissue Res. 22, 27–33
9. van Oss, C. J., Gillman, C. F., Bronson, P. M., and Border, J. R. (1974) Immunol. Commun. 3, 329–335
10. Lewis, J. G., and André, C. M. (1980) Immunology 39, 317–322
11. Baskies, A. N., Chretien, P. B., Weiss, J. F., Makuch, R. W., Beveridge, R. A., Catalona, W. J., and Spiegel, H. E. (1980) Cancer (Phila.) 45, 3050–3060
12. Gueyo, F., and Schmid, K. (1981) Biochim. Biophys. Acta 671, 76–84
13. Hermosdon, M. A., Ericson, L. H., Neurath, H., and Waish, A. (1973) Biochemistry 12, 3146–3153
14. Swank, R. T., and Munkres, K. D. (1971) Anal. Biochem. 39, 477
15. Kamerling, J. P., Gerwig, G. J., Vliegenthart, J. F. G., and Clamp, J. R. (1975) Biochem. J. 151, 491–495
16. Adler, A. J., Greefield, N. J., and Fasman, G. D. (1973) Methods Enzymol. 27, 675–735
17. Gray, W. R. (1972) Methods Enzymol. 25, 121–138
18. Ambler, H. P. (1972) Methods Enzymol. 25, 143–154
19. Edman, P., and Begg, C. (1967) Eur. J. Biochem. 1, 80–91
20. Pisano, J. J., and Bronzet, T. J. (1969) J. Biol. Chem. 244, 5597–5607
21. Bown, A. S., Mole, J. E., Weisssinger, A., and Bennett, J. C. (1976) J. Chromatogr. 148, 532–535
22. Pisano, J. J. (1979) Methods Enzymol. 26B, 27–44
23. Smithies, O., Gibson, D., Fanning, E. M., Goodflesh, R. M., Gilman, J. G., and Ballantyne, D. L. (1971) Biochemistry 10, 4912–4921
24. Hodge, J. E., and Hofreiter, B. T. (1962) Methods Carbohydr. Chem. 1, 388
25. Vliegenthart, J. F. G., Van Halbeek, H., and Dorland, L. (1981) Pure Appl. Chem. 53, 45–77
26. Chou, P. Y., and Fasman, G. D. (1979) Adv. Enzymol. 47, 45–148
27. Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J. C., Fournet, B., Spik, G., and Montreuil, J. (1978) Eur. J. Biochem. 87, 323–329
28. van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Fiat, A.-M., and Jollès, P. (1980) Biochim. Biophys. Acta 623, 295–300
29. van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Fiat, A.-M., and Jollès, P. (1980) FEBS Lett. 133, 45–50
30. Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Binette, J. P., and Schmid, K. (1978) Biochemistry 17, 5206–5214
31. van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Montreuil, J., Fournet, B., and Schmid, K. (1981) J. Biol. Chem. 256, 5588–5590
32. Welti, D., Rees, D. A., and Wiel, E. J. (1979) Eur. J. Biochem. 94, 505–514
33. Chou, P. Y., and Fasman, G. D. (1979) Biochim. Biophys. Acta 26, 367–383
34. Richardson, J. S. (1981) Adv. Protein Chem. 34, 167–339
35. Chou, P. Y., and Fasman, G. D. (1974) Biochemistry 13, 222–244
36. Montreuil, J. (1980) Adv. Carbohydr. Chem. Biochem. 37, 157–223
37. Aubert, J.-P., Biscette, G., and Lourcheux-Lefebvre, M. H. (1976) Arch. Biochem. Biophys. 175, 410–416
38. Pierce-Cretel, A., Pambianco, M., Strecker, G., Monreuil, J., and Spik, G. (1981) Eur. J. Biochem. 114, 169–178
Characterization of the B-chain of human plasma alpha 2HS-glycoprotein. The complete amino acid sequence and primary structure of its heteroglycan.

F Gejyo, J L Chang, W Bürgi, K Schmid, G D Offner, R F Troxler, H Van Halbeek, L Dorland, G J Gerwig and J F Vliegenthart

*J. Biol. Chem.* 1983, 258:4966-4971.

Access the most updated version of this article at [http://www.jbc.org/content/258/8/4966](http://www.jbc.org/content/258/8/4966)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/258/8/4966.full.html#ref-list-1](http://www.jbc.org/content/258/8/4966.full.html#ref-list-1)