The Carbohydrate-Polypeptide Linkages, the Amino Acid Sequences of the Peptides Adjacent to some of these Bonds, and the Composition and Size of the Carbohydrate Units of α_1 -Acid Glycoprotein

By M. SATAKE, T. OKUYAMA, K. ISHIHARA AND K. SCHMID

Department of Biochemistry, Boston University School of Medicine, Boston University Medical Center, Boston, Mass., U.S.A.

(Received 29 July 1964)

1. The glycopeptides derived from a proteolytic digest of sialic acid-free α_1 -acid glycoprotein were separated on a DEAE-cellulose column into five main fractions.
2. The average molecular weight of these glycopeptides was 2400, except for one fraction whose molecular weight was 3100. The average molecular weight of the sialic acid-free carbohydrate units was found to be 2200. From these data and the carbohydrate content of the native protein and the assumed molecular weight of 44000, it was concluded that α_1 -acid glycoprotein probably possesses five carbohydrate units. The sialic acid-containing carbohydrate units of this glycoprotein have an average molecular weight of 3000, except for one unit the molecular weight of which is significantly higher.
3. The *N*-, non-*N*- and *C*-terminal amino acids of the main glycopeptides were determined. Aspartic acid and threonine occur in most peptides. Alanine, glycine, proline, serine and lysine were present in varying amounts. Traces of other amino acids were also found.
4. The amino acid sequence of three main glycopeptides was established and indicated that these glycopeptides are located at different positions of the polypeptide chain of the glycoprotein. These sequences are: Asp(NH₂)-Pro-Lys; Thr-Asp(NH₂)-Ala; Asp(NH₂)-Gly-Thr.
5. From the results of a series of chemical reactions (periodate oxidation, hydrazinolysis, dinitrophenylation, mild acid hydrolysis) it was shown that the hydroxyl group of the *N*-terminal threonine and the ϵ -amino group of lysine are free and that the β -carboxyl group of aspartic acid is present as amide. It was concluded that this amide group is involved in the carbohydrate-polypeptide linkages of at least four carbohydrate units of α_1 -acid glycoprotein.
6. The carbohydrate composition of the sialic acid-free glycopeptides was determined in terms of moles of neutral hexoses, glucosamine and fucose/mole.
7. Fucose, at least to the larger part, is not linked to sialic acid, and its (glycosidic) linkage is significantly more stable toward acid hydrolysis than the bond of the sialyl residues.
8. Heterogeneity of the carbohydrate units of α_1 -acid glycoprotein was found with regard to size and to content of fucose and sialic acid.

In a previous paper on the structure of α_1 -acid glycoprotein a method was described for the preparation of a glycopeptide mixture from this protein (Kamiyama & Schmid, 1962a). Subsequently, a glycopeptide with aspartic acid as the sole amino acid was isolated in low yield (Kamiyama & Schmid, 1962b), suggesting that the β -carboxyl group of aspartic acid is involved in the linkage between the polypeptide and carbohydrate moiety of at least some of these peptides. Further, the composition of the original glycopeptide mixture

seemed to indicate that probably not all polysaccharide units have the same structure (Kamiyama & Schmid, 1962a,b).

With the aim of investigating further the carbohydrate-polypeptide linkages, the carbohydrate units and the amino acid sequences adjacent to the carbohydrate-polypeptide linkages of α_1 -acid glycoprotein, additional glycopeptides were isolated from the mentioned glycopeptide fraction. The partial characterization of these compounds is presented in this paper.

MATERIALS AND METHODS

α_1 -Acid glycoprotein. α_1 -Acid glycoprotein was isolated from pooled outdated normal human plasma in essentially homogeneous form (Schmid, 1953).

Enzymes. Pepsin, papain and carboxypeptidase, twice recrystallized, and chymotrypsin, three-times recrystallized, were purchased from Worthington Biochemical Corp., Freehold, N.J., U.S.A. Pronase (*Streptomyces griseus* protease), from Kaken and Co., Tokyo, Japan, was obtained from Dr S. Kamiyama, Sendai.

Sialic acid-free glycopeptides. The preparation of the sialic acid-free glycopeptides was carried out as follows. A 15 g. portion of α_1 -acid glycoprotein containing approx. 8% of moisture was dissolved in 300 ml. of 0.1 N- H_2SO_4 and heated at 80° for 1 hr. (Warren, 1959). The cooled solution was neutralized with saturated $Ba(OH)_2$, and the $BaSO_4$ was centrifuged off after standing overnight. The washings of the $BaSO_4$ were combined with the supernatant, dialysed exhaustively against water to remove sialic acid and freeze-dried, yielding 12.8 g. of modified glycoprotein. The outside solution was analysed for galactose, mannose, fucose and glucosamine. The subsequent enzymic digestion of the modified glycoprotein was carried out as described by Kamiyama & Schmid (1962a,b), except that the duration of the last step of the proteolysis by Pronase was reduced from 17 hr. to 90 min. The separation of the glycopeptides from amino acids and carbohydrate-free peptides was achieved by partition chromatography on a paper-powder column (Kamiyama & Schmid, 1962a). The yield of glycopeptides was 4.3 g. Because of the presence of a minute amount of carbohydrate-free peptides (80 mg.), as judged by paper chromatography in the two systems indicated below, this particular glycopeptide fraction was rechromatographed on a paper-powder column with the upper phase of the butan-1-ol-acetic acid-water (4:1:5, by vol.) system. The final preparation referred to as starting material appeared homogeneous on paper chromatography and remained at the point of application, as noted by Kamiyama & Schmid (1962a,b).

For the resolution of the glycopeptide fraction, ion-exchange chromatography on DEAE-cellulose (product of Brown and Co., Berlin, N.H., U.S.A.) at room temperature was employed. This ion-exchanger was washed by the customary procedure (Peterson & Sober, 1962), extracted with pyridine-acetate buffer, pH 6.2 and 10.1, and finally equilibrated against pyridine-acetate buffer, pH 6.2 and 10.001. It was observed that the pyridine-acetate buffer of high ionic strength extracted additional material from the washed resin that is largely insoluble in water. The DEAE-cellulose column had diam. 2.5 cm. $\times$ height 61 cm., and had been compressed first at 5 and finally at 10 lb./in.². The column was washed with 2.4 l. of the latter buffer at a flow rate of 30 ml./hr. A 1 g. portion of the above-described starting material, corrected for moisture content, was dissolved in 5 ml. of the mentioned buffer and applied to the column (point A in Fig. 1). The flow rate was kept constant throughout this chromatographic operation and the effluent was collected in 10 ml. fractions. After 2.4 l. of effluent had been collected, a gradient with increasing ionic strength but constant pH of 6.2 was applied (point B in Fig. 1). The gradient was produced by connecting the column with a beaker containing 700 ml. of pyridine-acetate buffer, 10.001, which was stirred. This beaker was

connected with another in which 700 ml. of pyridine-acetate buffer, 10.01, was given. Thereafter, 1300 ml. of pyridine-acetate buffer, pH 6.2 and 10.8, was applied (point C in Fig. 1). Samples (0.2 ml.) of each fraction were used for the hexose determination (Sørensen & Haugaard, 1933). Comparable fractions (fractions 1-10 in Fig. 1) were combined, concentrated and freeze-dried.

Paper chromatography. For paper chromatography of the glycopeptides, peptides and amino acids propanol-ethanol-water-acetic acid (5:15:5:0.74, by vol.) mixture and the upper phase of butan-1-ol-acetic acid-water (4:1:5, by vol.) mixture were utilized. For the chromatographic separation of certain monosaccharides butan-1-ol-pyridine-water (5:3:2, by vol.) mixture was used. After drying of the chromatograms the compounds containing free amino groups were revealed with ninhydrin [0.2% ninhydrin in acetone containing 5% (v/v) of pyridine-acetate buffer, pH 5.0 and 10.2]. The monosaccharides were developed with the silver stain (Trevelyan, Procter & Harrison, 1950).

High-voltage electrophoresis. For the high-voltage electrophoresis a Servonuclear tank filled with Coolant EC-123 (purchased from Savant Instruments Co., Hicksville, N.Y., U.S.A.) and an electrical tension of 2000-4000 v or approx. 60 v/cm. were employed. Whatman no. 3MM paper, washed with 0.03 N-HCl and then with distilled water, was used for analytical and preparative separation, because when wet this paper can be handled more easily than Whatman no. 1 paper. The electrophoresis was carried out at pH 3.5 and 10.025 (pyridine-acetic acid-water, 1:10:490, by vol.) and at pH 6.2 and 10.025 (pyridine-acetic acid-water, 15:0.6:485, by vol.) (Michl, 1951; Ingram, 1959; Dixon, Kauffman & Neurath, 1958).

Carbohydrate analysis. For the determination of the composition of the carbohydrate moiety of the glycopeptides the following procedures were utilized. Neutral hexoses were measured by the orcinol method (Sørensen & Haugaard, 1933) and the anthrone method (Scott & Melvin, 1953), with an equimolar solution of galactose and mannose as standard (Montreuil, Spik, Chosson, Segard & Scheppler, 1963). The latter procedure yields a relatively more stable colour. Hexosamine was determined by the Elson-Morgan method as modified by Masamune & Yosizawa (1955), sialic acid by the method of Werner & Odin (1952) with Ehrlich's reagent and by the Warren (1959) procedure, and fucose by the method of Dische & Shettles (1948) as described by Winzler (1955). To minimize the error in the measurements of the carbohydrate composition of the glycopeptides, appropriate samples of stock solutions of these compounds were utilized for these determinations. Further, artificial mixtures of galactose and mannose at concentrations similar to the experimental tubes were also included in these and the subsequent analyses. The ratio of galactose to mannose was determined according to the procedure of Whistler & BeMiller (1962).

Amino acid analysis. The determination of the N-terminal and non-N-terminal amino acids by the dinitrophenylation procedure was carried out as described previously (Kamiyama & Schmid, 1962a; Fraenkel-Conrat, Harris & Levy, 1955). The C-terminal amino acids were determined according to the method of Akabori *et al.* (1956), with anhydrous hydrazine and hydrazine sulphate as catalyst for the hydrazinolysis (Bradbury, 1956, 1958a,b) and isovaleraldehyde or enanthaldehyde for the separation of the formed hydrazides (Kusama, 1957).

For the identification and determination of the different hydrazides of aspartic acid and glutamic acid as DNP derivatives, the procedure of Akabori *et al.* (1956) was employed. For this purpose the application of an aldehyde must be avoided. After the removal of excess of hydrazine by evaporation *in vacuo* over conc. H_2SO_4 , the hydrazides were separated from hydrazine sulphate by extracting the reaction mixture with methanol. The methanol extract was evaporated to dryness and treated with dinitrofluorobenzene (Schroeder & LeGotte, 1953). The reaction mixture was diluted with water, acidified to pH 1 and extracted with ethyl acetate. For the removal of the acidic monohydrazides of these two amino acids, the ethyl acetate solution was extracted with 2% (w/v) NaHCO_3 . The compounds of the aqueous extract, including the monohydrazide bis-DNP derivatives of aspartic acid and glutamic acid, were isolated and chromatographed on Whatman no. 1 paper by using the Levy I/II system. The bishydrazides of aspartic acid or glutamic acid ($\alpha\beta$ - and $\alpha\gamma$ - respectively), as their tris-DNP derivatives present in the organic solution, can be separated from the other compounds after hydrolysis with 6N-HCl at 100° for 5 hr., followed by chromatography of the *N*-DNP derivatives of aspartic acid and glutamic acid.

The leucine equivalent was determined as described by Hirs, Moore & Stein (1956) and Yemm & Cocking (1955), assuming a ninhydrin colour equivalent to that of leucine.

The amide content was measured as described by Spiro & Spiro (1962), with glucosamine and asparagine as standard and lysozyme and fetuin as controls. Between 40 and 60 mg. of each peptide was hydrolysed in a sealed tube in 4N-HCl at 100° for 4 hr.

Molecular weight. The molecular weights of the glycopeptides were determined by the sedimentation-equilibrium procedure of Yphantis (1960) with the Spinco model E ultracentrifuge equipped with an automatic temperature-control unit (R.T.I.C. type) and schlieren and interference optics. The glycopeptides were dissolved in 0.1 M-NaCl to give a concentration of approx. 1%. The equilibrium runs were performed at 22° at a speed between 12000 and 18000 rev./min. The partial specific volumes of glycopeptides were calculated (Edsall, 1943) to be 0.630. Some of the equilibrium experiments were also carried out in pyridine-acetate buffer, pH 6.2 and 10.1. However, consistently higher values were obtained.

Periodate oxidation. The periodate oxidation (Kamiyama & Schmid, 1962b) of some of the isolated glycopeptides was carried out at pH 4.5 in acetate, 10.02, in the dark and at room temperature for 2½ hr. In a 24 ml. test tube with a glass stopper, 2.5 mg. of glycopeptides (approx. 1 μ mole) was dissolved in 0.5 ml. of water, mixed with 0.5 ml. of 0.1 M- NaIO_4 (50 μ moles) and then with 2.0 ml. of sodium acetate buffer, pH 5.0 and 10.02. After the reaction period the remaining iodate was destroyed by the addition of 2.0 ml. of glycerol (400 μ moles) at 24°. After 3 hr., the pH of the reaction mixture was adjusted to 8.0 with 1N-NaOH. For the reduction of formed aldehydes approx. 800 mg. of NaBH_4 was added and the solution allowed to stand at 2° for 10 hr. (Spiro, 1964). After neutralization with acetic acid the reaction mixture was concentrated to about 3 ml. For the subsequent dinitrophenylation 2 ml. of aq. 2% (v/v) trimethylamine solution and 10 ml. of 5% (v/v) fluorodinitrobenzene in ethanol were added. This reaction, as well as the isolation of the DNP-amino acids, was done according to the conventional procedures (Fraenkel-Conrat

et al. 1955). As a control, a mixture of the same volumes of periodate solution, glycerol and acetate buffer was kept for 3 hr. at room temperature. Thereafter, the same amount of glycopeptide was added and further treated as described above.

RESULTS

Chromatographic and electrophoretic properties and molecular weight of the glycopeptides. The ion-exchange chromatography on a DEAE-cellulose column of the glycopeptide fraction derived from sialic acid-free α_1 -acid glycoprotein resulted in a separation of several glycopeptides from each other as judged by the distribution of the hexose content of the effluent (Fig. 1). Five main fractions were obtained, although a total of ten fractions were separately collected. The weight of all fractions accounted for 93% of the starting material (Table 1). No glycopeptides were isolated in molar ratio

Fig. 1. Fractionation on a DEAE-cellulose column of the sialic acid-free glycopeptides derived from α_1 -acid glycoprotein. On the ordinate the extinction of the colour produced by the orcinol reaction is plotted. Experimental details are given in the text.

Table 1. Recoveries and average molecular weights of the glycopeptides derived from an enzymic digest of sialic acid-free α_1 -acid glycoprotein

The separation of these glycopeptide fractions was achieved by chromatography on a DEAE-cellulose column. Experimental details are given in the text.

Fraction no.	Weight (mg.)	Mol.wt.	Glycopeptide (μ moles)
1	273	2300	119
2	11	—	—
3	120	2400	50
4	99	—	41
5	229	2500	124
6	81		
7	7	—	—
8	19	—	—
9	79	3100	28
10	11	—	—
Total	934		

with respect to the starting material. Paper chromatography established the absence of carbohydrate-free peptides. Fractions 2, 7, 8 and 10, which accounted for a small percentage of the starting material, were not studied.

High-voltage electrophoresis indicated that a better resolution was obtained at pH 3.5 than at pH 6.2 (Fig. 2). Fractions 4, 5 and 6 appeared to be homogeneous, whereas fractions 1 and 3 each contained small amounts of impurities, and fraction 9 consisted of two major components. The electrophoretic mobilities suggested that fraction 1 carried at both pH values a positive net charge and fraction 9 a negative net charge relative to the other fractions (Fig. 2). At pH 11 in phosphate buffer a resolution resulted which was similar to that noted at pH 3.5, except that fractions 3, 4 and 6 each revealed an additional minor component. These minor components, however, might be due to decomposition taking place at pH 11, and were not studied further.

The molecular weight of the sialic acid-free glycopeptides varied between 2300 and 2500, except for that of fraction 9 which was found to be 3100. As shown below, the polypeptide content of the latter fraction is significantly higher than those of the other fractions.

N-, non-*N*- and *C*-Terminal amino acid contents of the glycopeptides. The *N*-, non-*N*- and *C*-terminal amino acid contents of fractions 1, 3, 4, 5, 6 and 9 are listed in Table 2. These results were corrected

for destruction occurring during hydrolysis and recovery of free and DNP-amino acids of the glycopeptides.

Fraction 1 contained as *N*-terminal amino acids aspartic acid and threonine in nearly equal quantities, each amounting to almost 0.5 mole/mole of glycopeptide. Of the non-*N*-terminal amino acids, in addition to aspartic acid, threonine, glycine and serine, the relatively high content of proline and lysine is striking. Lysine and threonine were the *C*-terminal amino acids. After dinitrophenylation and hydrolysis of this glycopeptide, the whole of the former amino acid was isolated as its ϵ -DNP-derivative.

Fraction 3 possessed almost 1 mole of threonine/mole of glycopeptide as *N*-terminal amino acid as well as a minute amount of aspartic acid. There was almost 1 mole of non-*N*-terminal aspartic acid/mole together with smaller amounts of proline and threonine. Threonine, although in relatively low concentration, was the only *C*-terminal amino acid. However, aspartic acid could be isolated as the β -monohydrazide. The recovery of this compound, in agreement with Akabori *et al.* (1956) and Ramachandran & Narita (1958), was very low and difficult to assess. Periodate oxidation of fraction 3 decreased the content of the *N*-terminal threonine from 0.91 to 0.01 mole/mole of glycopeptide.

Fraction 4, as judged by the curve shown in Fig. 1, probably contained small amounts of the glycopeptides present in fractions 3 and 5. As *N*-terminal amino acids only threonine and as non-*N*-terminal amino acids almost 1 mole of aspartic acid/mole and a considerable amount of alanine were isolated. Traces of other amino acids were also noted, a finding observed with the other fractions. The *C*-terminal amino acids were alanine and small amounts of threonine. Periodate oxidation destroyed the *N*-terminal threonine.

Fractions 5 and 6 were notable in that their *N*-, non-*N*- and *C*-terminal amino acids were the same and, moreover, present in approximately the same concentrations. Aspartic acid was the sole *N*-terminal amino acid detected. The *N*-terminal threonine in fraction 5 was probably due to contamination of the glycopeptide of fraction 4. Almost 1 mole of non-*N*-terminal threonine/mole of peptide and about 0.5 mole of glycine/mole were found. As *C*-terminal amino acid only threonine was detected in an amount equal to that of the non-*N*-terminal threonine and to that of the *N*-terminal aspartic acid. A significant quantity of $\alpha\beta$ -bishydrazide of aspartic acid was isolated as the bis-DNP derivative, e.g., 0.6 mole/mole when corrected by the recovery factor given by Ramachandran & Narita (1958). Further, after incubation of fraction 5 with Pronase (weight ratio of enzyme

Fig. 2. High-voltage electrophoresis of the isolated glycopeptide fractions (1 to 9) derived from sialic acid-free α_1 -acid glycoprotein. The experiments were carried out in pyridine-acetate buffer at (a) pH 3.5 and (b) pH 6.2. Experimental details are given in the text.

Table 2. *N*-, non-*N*- and C-Terminal amino acids of certain glycopeptides derived from α_1 -acid glycoprotein

Experimental details are given in the text. The values for *N*-terminal amino acids were corrected for destruction due to hydrolysis, namely 50% for DNP-aspartic acid, 44% for DNP-threonine and assumed 50% for DNP-glutamic acid (Kamiyama & Schmid, 1962a,b). The values for non-*N*-terminal amino acids were corrected for a recovery of 50% as judged by the amino acid composition of some of these fractions obtained by an automatic amino acid analyser. Fractions 3 and 4 contained, in addition, small amounts of tyrosine. The values for C-terminal amino acids were corrected for an estimated recovery of 40%. These fractions contain the expected artifacts (glycine, serine and threonine) due to hydrolysis during hydrazinolysis (Champagne & Smith, 1960).

		Amino acid content (moles/mole of glycopeptide)							
Amino acids		Fraction 1	Fraction 3	Fraction 4	Fraction 5	Fraction 6	Fraction 9	Fraction 9-A	Fraction 9-B
<i>N</i> -Terminal	Asp	0.44	0.02	0.02	0.71	0.64	0.32	—	0.33
	Thr	0.39	0.91	0.47	0.06	—	0.10	—	0.11
	Glu	—	—	—	—	—	0.32	0.33	—
Non- <i>N</i> -terminal	Asp	0.25	0.78	0.82	0.04	0.04	0.20	0.70	0.48
	Thr	0.25	0.10	0.15	0.78	0.91	0.22	0.50	0.28
	Gly	0.18	0.08	—	0.39	0.34	0.22	0.43	*
	Pro	0.36	0.21	0.04	—	—	*	—	—
	Ser	0.10	*	—	—	—	*	—	—
	Glu	*	0.06	*	—	—	0.19	0.38	0.27
	Ala	—	—	0.42	0.08	0.10	*	—	—
	Lys	0.42	—	—	—	—	—	—	—
<i>C</i> -Terminal	Asp†	—	—	—	—	—	—	†	†
	Thr	0.12	0.15	0.05	0.85	0.88	0.15	†	†
	Lys	0.20	—	—	—	—	—	†	†
	Ala	—	—	0.45	—	—	—	†	†
Amide		0.86	1.06	1.02	1.16	1.10	0.92	—	—

* Traces of 0.01 mole of amino acid/mole or less.

† As enanthaldehyde was employed, any hydrazide of aspartic acid escaped detection.

‡ These determinations were not carried out.

to peptide 1:100) for 25 hr. at 40°, no change in the composition of this glycopeptide could be detected.

Fraction 9 contained approximately equimolar amounts of aspartic acid and glutamic acid as *N*-terminal amino acids, and threonine in smaller amounts. Among the non-*N*-terminal amino acids significant quantities of aspartic acid, threonine, glycine, proline and glutamic acid were noted. *C*-Terminal threonine was present in a small amount. From the reaction mixture that was not treated with aldehyde, the DNP derivative of the bishydrazides of aspartic acid could be isolated (0.8 mole/mole of glycopeptide). In addition small amounts of the same derivative of glutamic acid were also found.

Fraction 9 subjected to high-voltage electrophoresis at pH 3.5 was resolved into two main components and one trace component. The latter has not yet been investigated further. One of the main components revealed only glutamic acid as *N*-terminal amino acid, whereas the other was free of this amino acid but possessed *N*-terminal aspartic acid and threonine (Table 2). With regard to the non-*N*-terminal amino acid both fractions probably contained enough aspartic acid to account for the linkage between the carbo-

hydrate and polypeptide moiety. Nevertheless, significant amounts of glutamic acid and threonine were also present. The two main glycopeptides were also notable in that one was essentially free of glycine, whereas the glycine content of the other was about equal to that of each of the other amino acids.

The amide content of these fractions (Table 2) varied between 0.86 and 1.16 moles/mole of peptide.

Digestion with carboxypeptidase. Carboxypeptidase did not liberate any amino acids from these glycopeptides.

Carbohydrate composition of the glycopeptides. The six main glycopeptide fractions were analysed for the content of their carbohydrate constituents (Table 3). Fractions 1, 3 and 4 each contained 7 moles of neutral hexoses/mole, fractions 5 and 6 each contained 8 moles/mole and fraction 9 contained 9 moles/mole. Fractions 1–5 each possessed 5 moles of glucosamine/mole and fractions 6 and 9 each possessed 6 moles of glucosamine/mole. The ratio of neutral hexose to hexosamine varied between 1.3 and 1.5, that of galactose to mannose from 0.9 to 1.4; the fucose content varied between 0.4 and 1.0 moles/mole of peptide. In addition, fraction 9

Table 3. *Compositions of the carbohydrate moieties of certain sialic acid-free glycopeptides derived from α_1 -acid glycoprotein*

Experimental details are given in the text.

Fraction	Neutral hexoses		Hexosamine		Fucose		Sialic acid (%)	Total carbohydrate content (%)	Peptide content (%)	Total weight accounted for* (%)	Ratio of neutral hexoses to glucosamine	Ratio of galactose to mannose
	(%)	(moles/mole)	(%)	(moles/mole)	(%)	(moles/mole)						
1	55	7	38	5	4.4	0.5	0	97	12	109	1.4	1.0
3	51	7	39	5	6.8	1.0	0	96	12	108	1.3	1.2
4	55	7	38	5	3.5	0.5	0	96	12	108	1.4	1.4
5	58	8	38	5	5.8	0.9	0	102	12	115	1.5	1.2
6	57	8	42	6	2.4	0.4	0	101	12	113	1.4	1.2
9	51	9	34	6	4.8	0.9	1.4	91	22	113	1.5	0.9

* These values include the water taken up during hydrolysis.

contained a small amount of sialic acid (0.2 mole/mole). The carbohydrate moiety of these fractions, not corrected for water taken up during hydrolysis, ranged from 96 to 102%, and together with the peptide moiety accounted for 109–115%.

The results obtained from analysing fractions 5 and 6, which apparently contained the same glycopeptide as judged by the evidence presented above, indicate the errors and difficulties in the methods employed. The difference in ratio of neutral hexoses to glucosamine of 0.1 is probably not significant. Further, it was found that the variations in the determination of the fucose content of these glycopeptides are great.

DISCUSSION

For the removal of sialic acid from α_1 -acid glycoprotein relatively strong acid hydrolysis (0.1N-sulphuric acid at 80° for 1hr.) (Svennerholm, 1958; Warren, 1959; Zilliken, Brown & György, 1955) was employed, leading to cleavage of approx. 99% of this acid. Under the conventional conditions (0.025N-sulphuric acid at 80° for 1hr.) a considerably smaller percentage of this compound was split off. These observations suggest that the bonds between the different sialyl residues of α_1 -acid glycoprotein and galactosyl residues, which have been shown to be bound to different carbon atoms of the mentioned hexose (Jeanloz & Closse, 1963), differ greatly in their stability. The linkages of the different neuraminic acids in biological macromolecules vary greatly in resisting hydrolysis (Gibbons, 1963). Further, many neuraminidase preparations are capable of removing the largest part of sialic acid very rapidly, whereas the remaining part is cleaved off at a very much lower rate, if at all. The degree of liberation of the other monosaccharides of α_1 -acid glycoprotein under the conditions utilized for the hydrolysis of sialic acid was also investigated. In addition to the almost quantitative removal of

N-acetylneuraminic acid, approx. 15% of the fucose and perhaps trace amounts of galactose were liberated. No mannose and glucosamine was cleaved off. The acidic solvent used for the subsequent chromatographic separation was found not to hydrolyse off any further sugars.

A proteolytic digest of sialic acid-free α_1 -acid glycoprotein gave five main glycopeptide fractions on DEAE-cellulose chromatography with almost quantitative recovery. The average molecular weights of four main fractions eluted first varied between 2300 and 2500. In no case was a lower molecular weight observed. Our previous report (Kamiyama & Schmid, 1962a) for the average molecular weight of 2600 of the glycopeptide mixture of which only about one-third of the sialic acid had been removed is thus too low. The likely explanation for this discrepancy is that sialic acid renders the glycopeptides highly charged at neutrality, leading to this apparently low average molecular weight (Schachman, 1959). The fraction eluted last had a significantly higher average molecular weight of 3100, a higher polypeptide content (6 rather than 2.5 moles of amino acid residues/mole) and a larger carbohydrate moiety. These findings indicated heterogeneity of the carbohydrate units.

High-voltage electrophoresis, especially at pH 3.5, and the determination of the *N*- and *C*-terminal amino acids revealed the degree of heterogeneity of the glycopeptide fractions. Whereas all fractions except that eluted last appeared largely homogeneous on high-voltage electrophoresis, probably owing to the very similar net charge of the peptides of each fraction, a greater degree of heterogeneity was found by these chemical analyses. No further separation by chromatography on DEAE-cellulose or by high-voltage electrophoresis except for fraction 9 could be achieved under the conditions chosen. Fractions 3 and 5 plus 6 revealed a very high degree of homogeneity, as judged by the

N- and *C*-terminal amino acids. However, their contents of non-*N*-terminal amino acids, primarily proline and glycine, were not identical.

All these glycopeptides contained in small amounts terminal amino acids, as noted before for the unfractionated glycopeptide mixture (Kamiyama & Schmid, 1962*a,b*). These amino acids are present in the glycopeptides probably because of relatively incomplete digestion of the polypeptide chain of the sialic acid-free glycoprotein by Pronase, which, however, is known to possess a very broad specificity. Comparable findings have been reported for ovalbumin (Marks, Marshall, Neuberger & Popkoff, 1962), ovomucoid (Montgomery & Wu, 1963), α_1 -acid glycoprotein (Eylar, 1962), barley albumin (Hochstrasser, 1963) and interstitial-cell-stimulating hormone (Popkoff, 1963). In spite of the broad specificity of this protease only a very small amount of glycopeptide with aspartic acid as the sole amino acid could be isolated (Kamiyama & Schmid, 1962*b*), and extremely long incubation times are required to increase the yield of such a peptide (Eylar, 1962). However, the additional amino acid of which the molar ratio must be expected to be less than one is of great importance, as it marks the position where the glycopeptide belongs in the polypeptide chain. Moreover, because of the partial hydrolysis of certain peptide bonds by Pronase (Bourrillon, Got & Meyer, 1964), the glycopeptides described in the present study could not be isolated in a molar ratio.

The most likely structure of some of these glycopeptides is discussed below.

Fraction 1 was not retained on the DEAE-cellulose column under the conditions selected and showed a positive net charge on high-voltage electrophoresis. Therefore the main peptide of this fraction must be relatively basic. Further, the amino acid composition revealed the presence of lysine, of which the ϵ -amino group was shown to be free. Since the ϵ -amino group is not substituted, the carbohydrate moiety cannot be linked through lysine. Neither is the threonine that is located at the *N*-terminus involved in the carbohydrate-polypeptide linkage, since it is destroyed by periodate oxidation. Therefore the carbohydrate unit is linked mainly through the β -carboxylic group of aspartic acid. Moreover, the amide content of this fraction, which could be shown to be derived from the asparaginyl residue, indicates that it is actually the β -amide group of this amino acid that forms the linkage. Thus, of the several theoretically possible linkages (Neuberger, 1964; Schultze, 1964; Bogdanov, Kaverzneva & Andreyeva, 1964), an aspartylhexosamine bond is excluded and a linkage involving a 2-acetamido-1- β -(*L*-aspart-amido)-1,2-dideoxy-D-glucose structure is most likely.

Table 4. *Amino acid sequence of certain glycopeptides of α_1 -acid glycoprotein*

Strong evidence is presented that the β -amide group of aspartic acid forms the linkage to the carbohydrate unit.

Glycopeptide present in fraction	Amino acid sequence	Type
1	Asp(NH ₂)-Pro-Lys	I
3	Thr-Asp(NH ₂)	
4	Thr-Asp(NH ₂)-Ala	II
5 and 6	Asp(NH ₂)-Gly-Thr	III

From the *N*- and *C*-terminal amino acid analyses, it seems likely that fraction 1 consists of essentially two glycopeptides. It is tentatively suggested that the main component probably possesses the amino acid sequence listed in Table 4. The amount of these two glycopeptides corresponded to approx. 1.5 moles/mole of α_1 -acid glycoprotein. The carbohydrate moiety of these glycopeptides is similar in composition to those of the other fractions and is discussed below.

Fraction 3 consists of one major and a minor glycopeptide, as judged by electrophoresis and chemical analysis. The probable amino acid sequence of the major peptide is shown in Table 5. The hydroxyl group of the *N*-terminal threonine is again free, because this amino acid could be oxidized with periodate. Aspartic acid, which was found as non-*N*-terminal amino acid, could not be detected as *C*-terminal amino acid by the conventional aldehyde procedure but was isolated as the hydrazide. This finding confirms our previous contention that *C*-terminal aspartic acid could not be detected if enanthaldehyde was employed in the hydrazinolysis procedure (Kamiyama & Schmid, 1962*a,b*). This evidence, together with the amide content (approx. 1 mole/mole) of this fraction, indicates that the β -carboxylic group of aspartic acid is not free but is present as amide, which is probably involved in the linkage to the carbohydrate unit. The minor components of this fraction may account for the presence of proline. The *C*- and non-*N*-terminal threonine of fraction 3 is probably due to an additional minor peptide.

Fraction 4 probably contains one glycopeptide that is contaminated with the peptides from fractions 3 and 5. The amino acid sequence of this glycopeptide is *N*-terminal threonine, non-*N*-terminal aspartic acid and *C*-terminal alanine. Further, the amide content of this compound amounted to nearly 1 mole/mole of glycopeptide. Hence it appears very likely that the β -amide group of aspartic acid again links the peptide to the carbohydrate moiety. The presence of proline is probably due to the contaminant of fraction 3. Fractions 1, 3 and 4 revealed an apparently relatively high content of *C*-terminal threonine, which

seems to be an artifact due to hydrolysis during hydrazinolysis (Champagne & Smith, 1960).

Fractions 5 and 6 appear to consist of the same glycopeptide, as judged by the chemical analysis, and show the degree of variation of the results obtained by the procedures employed. The amount of fractions 5 and 6 corresponds to almost 2 carbohydrate units/mole of α_1 -acid glycoprotein. The peptide consists of *N*-terminal aspartic acid, non-*N*-terminal glycine and *C*-terminal threonine. Aspartic acid and threonine were recovered in particularly high yield. The amino acid sequence of this glycopeptide is given in Table 4. Further experiments are needed to explain the low recovery of glycine.

Whereas the amino group of aspartic acid is free, the α - and β -carboxylic groups of this amino acid are substituted since the $\alpha\beta$ -bishydrazide of aspartic acid was isolated. Moreover, on mild acid hydrolysis this compound released nearly 1 mole of amide nitrogen/mole. The hydroxyl group of the *C*-terminal threonine is free for the following reasons. Since threonine is liberated by acid hydrolysis an ether linkage between the hydroxyl group of this amino acid and the carbohydrate moiety is unlikely. Further, hydrazinolysis, which does not split *O*-glycosidic linkages, yields oligosaccharides from α_1 -acid glycoprotein that have been reported to be essentially free of amino acids (Reid, 1964). From results of the degradation of this protein (according to the procedure of Smith), Hughes & Jeanloz (1964) arrived at the same conclusion. Thus these findings strongly indicate that the carbohydrate units of fractions 5 and 6 are attached through the β -amide group of aspartic acid in *N*-glycosidic linkage.

Fraction 9 was eluted at the highest ionic strength employed. It was therefore most tightly bound to the DEAE-cellulose column, and was relatively acidic in nature, a property that was confirmed by high-voltage electrophoresis at pH 6.2 and 3.5. At pH 3.5 two major components were observed, whereas at pH 6.2 this fraction appeared to be almost homogeneous. Since the amount of sialic acid was too small to account for the acidic properties, the high content of acidic amino acid is responsible for the electrophoretic and chromatographic behaviour. The difference in the electrophoretic mobilities of the two main peptides is probably due to their different contents of aspartic acid and glutamic acid. On hydrazinolysis both acidic amino acids yielded bishydrazides. Thus it is possible that the carbohydrate moiety is linked not only through the β -carboxylic group of aspartic acid, but perhaps also in a minor degree through the γ -carboxylic group of glutamic acid to the polypeptide chain, unless this γ -carboxylic group proves to be an unsubstituted amide. The amino acid

composition of the electrophoretically separated peptides was too complex for deduction of any amino acid sequences to be made.

Three different types of amino acid sequences near the carbohydrate-polypeptide linkages are shown in Table 4. Fraction 9 consists of another type characterized by a high content of acidic amino acids. Further experiments are required to demonstrate whether or not these glycopeptides are derived from positions of the polypeptide chain of α_1 -acid glycoprotein that are different from those of types I, II and III (Table 4).

The carbohydrate-polypeptide linkage of which the β -amide group of aspartic acid is an integral part is most likely an *N*-glucosaminide, as indicated (Marks, Marshall & Neuberger, 1963; Tsukamoto, Yamamoto & Miyashita, 1964; Bogdanov *et al.* 1964). The evidence presented above excludes the presence of *N*-acylglucosamine. The presence of imino ester linkages is unlikely because the native protein and the glycopeptides were stable in alkali (Bogdanov *et al.* 1964), but, in view of a report (Bhavanadan, Buddecke, Carubelli & Gottschalk, 1964) on the unexpected stability of *O*-glycosidic linkage between galactosamine and the hydroxyl group of serine, further investigations are needed to evaluate this possibility.

The carbohydrate composition of all fractions is similar and the ratio of neutral hexoses to glucosamine is the same (Table 3). This constant ratio is in contrast with the corresponding values of glycopeptides derived from ovalbumin (Cunningham, Clouse & Ford, 1963) or the heterosaccharides of human colostrum (Got, Front & Bourrillon, 1963). Fraction 9 contains a significantly larger carbohydrate moiety, suggesting heterogeneity of the carbohydrate units of α_1 -acid glycoprotein. The constituents of the glycopeptides, including the peptide moiety, accounted for 108–115%, including water of hydrolysis. All glycopeptides contained a significant percentage of fucose, whereas they were devoid of sialic acid, except for one peptide that carried only a small amount. This finding clearly demonstrates that most of the fucose (which accounts for 1.3% of native α_1 -acid glycoprotein) is not linked to sialic acid.

The number of carbohydrate units of α_1 -acid glycoprotein can be calculated from the molecular weight of the glycopeptides and the total carbohydrate content of this protein, whose molecular weight was assumed to be 44100 (Weimer, Mehl & Winzler, 1950). The average molecular weight of the sialic acid-free glycopeptides is 2400, including 2.5–3.0 moles of amino acid residues/mole. Thus the molecular weight of the sialic acid- and amino acid-free carbohydrate units averages 2200. The total carbohydrate content of α_1 -acid glycoprotein was reported to be 33% (Schmid, 1953) or 15000,

of which the 16 sialic acid residues/mole account for 4500. The sialic acid-free carbohydrate moiety of this protein thus has a molecular weight of 10 500. Hence α_1 -acid glycoprotein probably possesses about 5 carbohydrate units. A comparable value is obtained from the number of moles of glycopeptide isolated and the number of moles of glycoprotein used as starting material for these studies. These results are in fair agreement with our previous report of 5–7 units (Kamiyama & Schmid, 1962a,b) and that by Hughes & Jeanloz (1964) of 3–4 units.

On the basis of the number of carbohydrate units and the number of moles of fucose (3) and sialic acid (16)/mole, heterogeneity of the carbohydrate units must also exist with regard to these sugars.

The average size of the unmodified carbohydrate units, e.g. with the sialic acid and fucose, is about 3000, because they contain on an average 3 sialic acid residues/mole. One of the 5 units may contain 4 sialic acid residues and have a molecular weight as high as 3500.

The authors thank Dr A. G. Maltoltsy for carrying out the amino acid analyses and Dr S. Kamiyama, Sendai, for the gift of Pronase. The expert technical assistance of Miss Alise Polis is gratefully acknowledged. This investigation was supported by grants from The Lilly Research Laboratories, Eli Lilly Co., National Science Foundation (GB-611) and by U.S. Public Health Service Research Grant GM-10374, from the National Institute of General Medical Sciences.

REFERENCES

- Akabori, S., Ohno, K., Ikenaka, T., Okada, Y., Hanafusa, H., Haruna, I., Tsugita, A., Sugae, K. & Matsushima, T. (1956). *Bull. chem. Soc. Japan*, **29**, 507.
- Bhavanandan, V. P., Buddecke, E., Carubelli, R. & Gottschalk, A. (1964). *Biochim. biophys. Res. Commun.* **16**, 353.
- Bogdanov, V. P., Kaverzneva, E. D. & Andreyeva, A. P. (1964). *Biochim. biophys. Acta*, **83**, 69.
- Bourrillon, R., Got, R. & Meyer, D. (1964). *Biochim. biophys. Acta*, **83**, 178.
- Bradbury, J. H. (1956). *Nature, Lond.*, **178**, 912.
- Bradbury, J. H. (1958a). *Biochem. J.* **68**, 475.
- Bradbury, J. H. (1958b). *Biochem. J.* **68**, 482.
- Champagne, M. & Smith, D. B. (1960). *Canad. J. Biochem. Physiol.* **38**, 715.
- Cunningham, L. W., Clouse, R. W. & Ford, D. (1963). *Biochim. biophys. Acta*, **78**, 379.
- Dische, Z. & Shettles, L. B. (1948). *J. biol. Chem.* **175**, 595.
- Dixon, G. H., Kauffman, D. L. & Neurath, H. (1958). *J. biol. Chem.* **233**, 1373.
- Edsall, J. T. (1943). In *Proteins, Amino Acids and Peptides*, p. 375. Ed. by Cohn, E. J. & Edsall, J. T. New York: Reinhold Publishing Corp.
- Eylar, H. E. (1962). *Biochem. biophys. Res. Commun.* **3**, 195.
- Fraenkel-Conrat, H., Harris, J. I. & Levy, A. L. (1955). *Meth. biochem. Anal.* **2**, 359.
- Gibbons, R. A. (1963). *Biochem. J.* **89**, 380.
- Got, R., Front, J. & Bourrillon, R. (1963). *Biochim. biophys. Acta*, **78**, 367.
- Hirs, W. C., Moore, S. & Stein, W. H. (1956). *J. biol. Chem.* **219**, 623.
- Hochstrasser, K. (1963). *Hoppe-Seyl. Z.* **333**, 99.
- Hughes, R. C. & Jeanloz, R. W. (1964). *Abstr. 1st. Meet. Fed. Eur. biochem. Soc., London*, p. 63.
- Ingram, W. M. (1959). *Biochim. biophys. Acta*, **28**, 539.
- Jeanloz, R. W. & Closse, A. (1963). *Fed. Proc.* **22**, 538.
- Kamiyama, S. & Schmid, K. (1962a). *Biochim. biophys. Acta*, **58**, 80.
- Kamiyama, S. & Schmid, K. (1962b). *Biochim. biophys. Acta*, **63**, 266.
- Kusama, K. (1957). *J. Biochem., Tokyo*, **44**, 375.
- Marks, G. S., Marshall, R. D. & Neuberger, A. (1963). *Biochem. J.* **87**, 274.
- Marks, G. S., Marshall, R. D., Neuberger, A. & Popkoff, H. (1962). *Biochim. biophys. Acta*, **63**, 340.
- Masamune, M. & Yosizawa, Z. (1955). *Tohoku J. exp. Med.* **61**, 283.
- Michl, H. (1951). *Mh. Chem.* **82**, 489.
- Montgomery, R. & Wu, Y.-C. (1963). *Biochem. biophys. Res. Commun.* **11**, 249.
- Montreuil, J., Spik, G., Chosson, A., Segard, E. & Scheppler, N. (1963). *J. Pharm. Belg. N.S.* **18**, 529.
- Neuberger, A. (1964). *Abstr. 6th int. Congr. Biochem., New York*, II-105.
- Peterson, E. A. & Sober, H. A. (1962). In *Methods in Enzymology*, vol. 5, p. 3. Ed. by Colowick, S. P. & Kaplan, N. O. New York: Academic Press Inc.
- Popkoff, H. (1963). *Biochim. biophys. Acta*, **78**, 384.
- Ramachandran, L. K. & Narita, K. (1958). *Biochim. biophys. Acta*, **30**, 616.
- Reid, P. E. (1964). *Abstr. Pap. Amer. chem. Soc., 147th Meet., Philadelphia*, p. 5c.
- Schachman, H. K. (1959). *Ultracentrifugation in Biochemistry*, p. 201. New York: Academic Press Inc.
- Schmid, K. (1953). *J. Amer. chem. Soc.* **75**, 60.
- Schroeder, W. A. & LeGotte, J. (1953). *J. Amer. chem. Soc.* **75**, 4612.
- Schultze, H. E. (1964). In *Protides of the Biological Fluids*, p. 288. Ed. by Peeters, H. Amsterdam: Elsevier Publishing Co.
- Scott, T. A. & Melvin, E. H. (1953). *Analyt. Chem.* **25**, 1656.
- Sørensen, M. & Haugaard, G. (1933). *Biochem. Z.* **260**, 247.
- Spiro, M. J. & Spiro, R. G. (1962). *J. biol. Chem.* **237**, 1507.
- Spiro, R. G. (1964). *J. biol. Chem.* **239**, 567.
- Svennerholm, L. (1958). *Acta chem. scand.* **12**, 547.
- Trevelyan, W. E., Procter, D. P. & Harrison, J. S. (1950). *Nature, Lond.*, **166**, 444.
- Tsakamoto, H., Yamamoto, A. & Miyashita, C. (1964). *Biochem. biophys. Res. Commun.* **15**, 1511.
- Warren, L. (1959). *J. biol. Chem.* **234**, 1971.
- Weimer, H. E., Mehl, J. W. & Winzler, R. J. (1950). *J. biol. Chem.* **185**, 561.
- Werner, L. & Odin, I. (1952). *Acta Soc. Med. upsalien.* **57**, 230.
- Whistler, R. L. & BeMiller, J. N. (1962). In *Methods in Carbohydrate Chemistry*, vol. 1, p. 395. Ed. by Whistler, R. L. & Wolfram, M. L. New York: Academic Press Inc.
- Winzler, R. J. (1955). *Meth. biochem. Anal.* **2**, 279.
- Yemm, E. W. & Cocking, E. C. (1955). *Analyst*, **80**, 209.
- Yphantis, D. A. (1960). *Ann. N.Y. Acad. Sci.* **88**, 586.
- Zilliken, F., Brown, G. A. & György, P. (1955). *Arch. Biochem. Biophys.* **54**, 564.