Biochem. J. (1976) **155**, 303–315
Printed in Great Britain

The Structures of the Carbohydrate Moieties of the α Subunit of Human Chorionic Gonadotrophin

By JOHN F. KENNEDY and MARTIN F. CHAPLIN
Department of Chemistry, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.

(*Received* 22 *May* 1975)

The α subunit of human chorionic gonadotrophin was reduced with dithiothreitol followed by carboxymethylation with iodoacetic acid. The modified glycoprotein was hydrolysed with trypsin to give various peptides, the identities of which were established, and glycopeptides. The glycopeptides were separated by gel filtration and ion-exchange chromatography; they were subjected to component analysis and were found to represent the two carbohydrate moieties in the parent glycoprotein. Sequential removal with glycoside hydrolases of monosaccharide units from the glycopeptides demonstrated (1) that galactose, mannose, glucosamine (2-amino-2-deoxyglucose) and neuraminic acid (5-amino-3,5-dideoxy-*glycero*-*galacto*-2-nonulosonic acid) residues possess the D configurations, (2) that the glucosamine units are *N*-acetylated and (3) the order of the monosaccharide units in the chain, the neuraminic acid units being furthest from the peptide backbone of the subunit and substituting the D-galactose units. Methylation analysis of the glycopeptides by adaptation of the Hakomori technique demonstrated that: (4) D-galactose, D-mannose and *N*-acetylglucosamine (2-acetamido-2-deoxy-D-glucose) units exist in the pyranose forms; (5) the D-galactopyranose units are linked in the 1 and 6 positions; (6) the D-mannopyranose units exist in several forms, one in a terminal non-reducing position, one as 1,2-linked residues, some as 1,6-linked residues and some as 1,2,6-linked branch points; (7) the *N*-acetylglucosamine units are 1,6-linked. On the basis of the results of methylation and enzymic analysis, structures are proposed for the carbohydrate moieties and the assignments are compared with other data previously obtained by periodate-oxidation studies [Kennedy *et al.* (1974) *Carbohydr. Res.* **36**, 369–377].

Human chorionic gonadotrophin is known to consist of two subunits, α and β (Swaminathan & Bahl, 1970; Butt & Kennedy, 1971; Kennedy, 1973). In common with the structures of human follicle-stimulating hormone (Reichert, 1972), thyrotrophin (Cornell & Pierce, 1973) and luteinizing hormone (Pierce *et al.*, 1971*a*,*b*) it has been found that in terms of biological activity, the β subunit is hormone-specific, whereas the α subunit is common to the other glycoprotein hormones (Kennedy, 1973; Pierce *et al.*, 1971*a*,*b*). The amino acid sequences of both these subunits of human chorionic gonadotrophin α and β have been reported (Bellisario *et al.*, 1973; Carlsen *et al.*, 1973), as have the positions of the disulphide bridges in the similar sheep luteinizing hormone (Ward *et al.* 1973; Chung *et al.*, 1973).

Problems in the elucidation of the carbohydrate structures of glycoprotein hormones have been reviewed (Kennedy, 1971). The carbohydrate compositions of several preparations of native human chorionic gonadotrophin have been reported (Bahl, 1969*a*,*b*; Ashitaka *et al.*, 1970; Bell *et al.*, 1969) and some studies on the linkage (Kennedy *et al.*, 1974) and sequence (Bahl, 1969*b*, 1973; Bahl *et al.*, 1973) of these carbohydrate residues have been undertaken. The carbohydrate moieties in both subunits are known to be specific for the particular hormone, and they vary in content and structure (Bellisario *et al.*, 1973).

Although Bahl (1969*b*) analysed the carbohydrate contents of glycopeptides produced by tryptic hydrolysis of desialylized human chorionic gonadotrophin, the yields were low, and the reported amino acid compositions do not agree with the various sequence studies. The possibility remains that all these glycopeptides were derived from the β subunit. To investigate these discrepancies and to discover the sequence and linkages of the carbohydrate moieties of the hormone, studies on pure α subunit have now been undertaken.

Materials and Methods

Hormone subunit

The α subunit of human chorionic gonadotrophin (hereafter termed 'subunit α') (lot E455 Qua-2) was kindly given by Dr. P. Donini, Istituto Farmacologico Serono, Rome.

General analytical methods

I.r. spectral analysis. I.r. spectra were obtained by analysis of the test material as a KBr disc with a Perkin–Elmer 472 spectrometer.

Conductivity measurements. The presence of salts in column eluates was detected by use of a standard conductivity cell, conductivities being measured in mmhos.

Amino acid analysis. After hydrolysis at 108°C in 5.9M-HCl for 24 or 72h, samples were analysed for amino acids with a Locarte amino acid analyser. The values for individual amino acids were corrected for destruction and incomplete hydrolysis. L-Cysteine was analysed as L-cysteic acid by prior oxidation with performic acid (125μl/mg). L-Tryptophan was determined from the u.v. spectrum (215–320nm) of the intact α subunit and the predetermined L-tyrosine, L-phenylalanine and L-cystine contents using calibration curves constituted for the absorbances of standard solutions of the four amino acids in the same solvent system.

Hexosamine analysis. After hydrolysis at 100°C in 4M-HCl for 6h, samples were analysed for hexosamines (2-amino-2-deoxyhexoses) with a Locarte amino acid analyser.

Neutral monosaccharide analysis. Neutral monosaccharides were determined both by methanolysis (1.5M-HCl in methanol, 16h, 100°C) followed by g.l.c. as their trimethylsilyl ethers on 10% silicone ester 30 on Celite (Chambers & Clamp, 1971) and by hydrolysis with trifluoroacetic acid (1.0M, 3h, 100°C for L-fucose; 2.0M, 4–6h, 100°C for other neutral monosaccharides) followed by analysis with a JEOL JLC-6AH carbohydrate analyser. The values obtained were corrected for destruction during liberation and averaged over the two methods.

Neuraminic acid analysis. Neuraminic acid was determined spectrophotometrically as described by Warren (1959); values were corrected for degradation during hydrolysis. The neuraminic acid was assumed to be the *N*-acetyl (5-acetamido) and the *N*-glycollyl (5-hydroxyacetamido) derivatives, and the *N*-glycollyl content was determined by using Eegriwe's colour reagent for glycollic acid (Klenk & Ullenbruck, 1966).

Mass-spectral analysis. Mass spectrometry was carried out with an AEI MS9 double-focusing mass spectrometer fitted with a heated inlet system and a direct insertion probe, with a source temperature of 180–190°C.

Enzymic activities. The enzymic compositions of the various enzyme preparations were determined by using the phenolic glycosides as substrates (4-nitrophenyl α-D-galactopyranoside, 2-nitrophenyl β-D-galactopyranoside, 4-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside, 4-nitrophenyl α-D-mannopyranoside) and under the conditions in which they were ultimately to be used. The activities of impurities were measured under the same conditions (Chaplin *et al.*, 1970) and expressed in terms of enzymic activity per standard amount of the preparation.

Preparation of enzymes

Neuraminidase. Neuraminidase (acylneuraminyl hydrolase, EC 3.2.1.18, from *Vibrio cholerae*) was obtained from Koch–Light Laboratories Ltd., Colnbrook, Bucks., U.K. (25000units/mg). One unit is the amount of enzyme required to liberate 0.0667μg of *N*-acetylneuraminic acid from a human serum glycoprotein in 1min at 37°C.

Mannosidase. α-D-Mannosidase (α-D-mannoside mannohydrolase, EC 3.2.1.24) was extracted from jack-bean meal (Sigma Chemical Co. Ltd., Kingston-upon-Thames, Surrey, U.K.) by a combination and adaptation (Chaplin, 1970; Chaplin *et al.*, 1970) of the methods of Li (1967) and Snaith & Levvy (1968). The enzyme was finally purified by fractionation on a column (1.8cm×60cm) of Sephadex G-25 (fine grade) by using 0.05M-sodium acetate buffer, pH5.0, containing 1% (v/v) toluene, 0.1M-NaCl and 0.1mM-$ZnSO_4$, at 4°C, as eluent. Fractions of the first (protein) peak containing α-D-mannosidase activity (assay as below) were selected. The α-D-mannosidase (10units/mg) contained small amounts of α-D-galactosidase (α-D-galactoside galactohydrolase, EC 3.2.1.22; 0.1%) and β-D-galactosidase (β-D-galactoside galactohydrolase, EC 3.2.1.23; 0.5%) activities, but no β-*N*-acetylglucosaminidase (β-D-acetamidodeoxyglucosidase, 2-acetamido-2-deoxy-β-D-glucoside acetamidodeoxyglucohydrolase, EC 3.2.1.30) activity. The activity (in units/mg) is defined as the number of μmol of 4-nitrophenol released/min at 20°C per mg of enzyme when a solution of the enzyme (0.1ml) is incubated with a solution of 4-nitrophenyl α-D-mannopyranoside (1mg/ml) in 0.05M-sodium acetate buffer, pH5.0, containing 1% (v/v) toluene, 0.1M-NaCl and 0.1mM-$ZnSO_4$ (1ml). The activities of contaminant enzymes were measured under the same conditions but by using 4-nitrophenyl α-D-galactopyranoside and 2-nitrophenyl β-D-galactopyranoside as substrates.

β-N-Acetylglucosaminidase. β-*N*-Acetylglucosaminidase was extracted from jack-bean meal and purified by a combination of $(NH_4)_2SO_4$ and alcohol precipitations and DEAE-Sephadex A-50 ion-exchange chromatography (Li & Li, 1970). The enzyme was finally purified by fractionation on a column (1.8cm×60cm) of Sephadex G-25 (fine grade) by using 0.1M-sodium citrate buffer, pH6.0, containing 1% (v/v) toluene and 0.05M-NaCl, at 4°C as eluent. Fractions of the first (protein) peak containing β-*N*-acetylglucosaminidase activity (assay as below) were selected. The β-*N*-acetylglucosaminidase (12units/mg) contained α-D-mannosidase

(20%), β-D-galactosidase (0.6%) and α-D-galactosidase (0.2%) activities. The activity (in units per mg) is defined as the number of μmol of 4-nitrophenol released/min at 20°C per mg of enzyme when a solution of the enzyme (0.1 ml) is incubated with a solution of 4-nitrophenyl 2-acetamido-2-deoxy-β-D-glucopyranoside (1 mg/ml) in 0.1 M-sodium citrate buffer, pH 6.0, containing 1% (v/v) toluene and 0.05 M-NaCl (1 ml). The activities of contaminant enzymes were measured under the same conditions by using the above substrates.

Galactosidase. β-D-Galactosidase was prepared from a mixture of β-D-glucosidase (β-D-glucoside glucohydrolase, EC 3.2.1.21) and β-D-galactosidase (from sweet almonds; Koch–Light Laboratories Ltd., marketed as β-D-glucosidase, 250 mg) by ion-exchange chromatography on a column (2 cm × 35 cm) of CM32 (Whatman) to decrease as far as possible the α-D-mannosidase activity. The column was eluted with three linear gradients as follows: 0.01 M-sodium acetate, pH 4.8 (100 ml), plus 0.01 M-sodium acetate, pH 4.8, containing 0.05 M-NaCl (100 ml); 0.01 M-sodium acetate, pH 4.8, containing 0.07 M-NaCl (100 ml) plus 0.01 M-sodium acetate, pH 4.8, containing 0.2 M-NaCl (100 ml); and 0.05 M-sodium acetate, pH 4.8, containing 0.2 M-NaCl (100 ml) plus 0.1 M-sodium acetate, pH 4.8, containing 0.5 M-NaCl (100 ml). Fractions were assayed for protein, β-D-galactosidase, α-D-galactosidase and α-D-mannosidase activity (Fig. 1). After the initial large protein peak of unadsorbed material, the next peak was selected in part for use as the β-D-galactosidase. Subsequent peaks containing β-D-galactosidase and eluted with buffers containing higher concentrations of NaCl were unsatisfactory, since they contained appreciable amounts of α-D-mannosidase activity. The enzyme was finally purified to remove low-molecular-weight material by fractionation on a column (1.8 cm × 60 cm) of Sephadex G-25 (fine grade) with 0.01 M-sodium acetate buffer, pH 4.8, containing 1% (v/v) toluene and 0.05 M-NaCl, at 4°C as eluent. Fractions of the first protein peak containing β-D-galactosidase activity (assay as below) were selected. The enzyme (1.0 units/mg) contained α-D-mannosidase (20%) but no α-D-galactosidase activity. The activity (in units/mg) is defined as the number of μmol of 2-nitrophenol released/min at 20°C per mg of enzyme when a solution of the enzyme (0.1 ml) is incubated with a solution (1 ml) of 2-nitrophenyl β-D-galactopyranoside (1 mg/ml) in 0.01 M-sodium acetate buffer, pH 4.8, containing 1% (v/v) toluene and 0.05 M-NaCl. The activity of the contaminant enzyme was measured under the same conditions by using the above substrate.

Preparation of human chorionic gonadotrophin α subunit glycopeptides

Subunit α (70.3 mg) was dissolved, under N_2, in 3 ml of 0.5 M-Tris/HCl buffer, pH 8.5, made 8 M with urea (4 × recrystallized from methanol) and containing 2% (w/v) EDTA. Dithiothreitol (61.7 mg, approx. tenfold molar excess over thiol groups) was added and the mixture left stirring for 40 min under N_2. Iodoacetic acid (148.8 mg, approx. equal to total thiol content) and NH_3 (0.05 ml, sp.gr. 0.88) were added and the incubation was continued for 30 min in

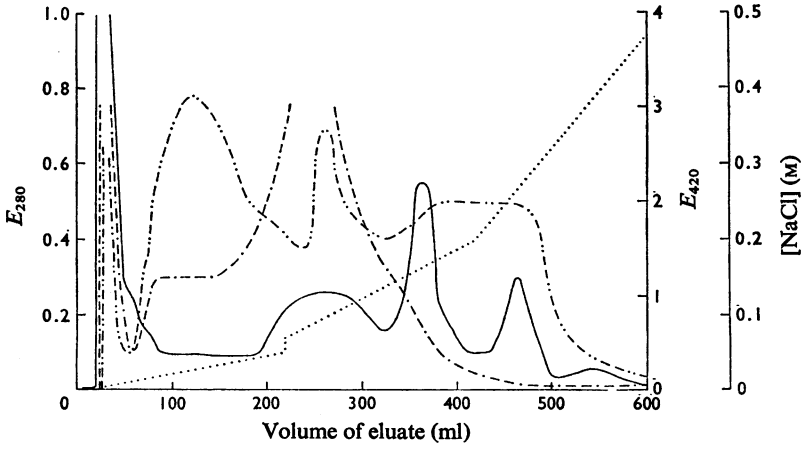

Fig. 1. *Fractionation of glycoside hydrolases from sweet-almond emulsin on a column of CM-cellulose CM32*

——, Protein content (E_{280}); —·····—, β-D-galactosidase activity by spectrophotometric assay (E_{420}); —·—·—, α-D-mannosidase activity by spectrophotometric assay (E_{420}); ·····, concentration of NaCl in eluate.

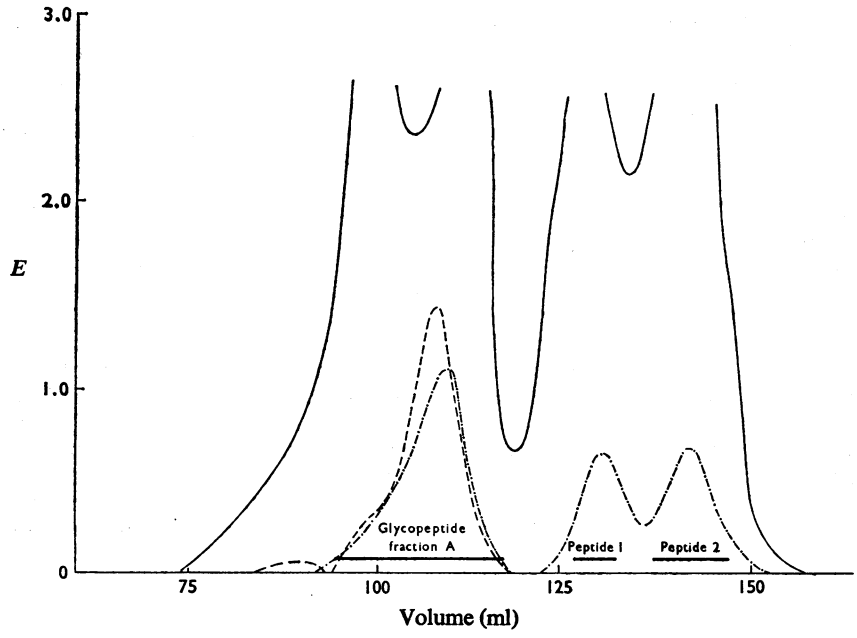

Fig. 2. *Fractionation of the tryptic peptides and glycopeptides from subunit α on a column of Sephadex G-75*

-·-·-·, Tyrosine content (E_{275}); ——, peptide content (E_{225}); – – – –, carbohydrate content by spectrophotometric assay (E_{415}); – – – –, trypsin content by spectrophotometric assay (E_{253}). Fractions pooled are indicated by the solid bars.

the dark. Desalting was accomplished by gel filtration in the dark on a column (1.9cm×55cm) of Sephadex G-75 run at 35ml/h (4°C) in 0.5% (w/v) NH_4HCO_3. The glycoprotein derivative was separated completely from salt and was recovered by freeze-drying. The carboxymethylated subunit α (65mg) was dissolved in 0.1M-NH_4HCO_3 (3ml) and treated, under N_2, with a solution (in 0.001M-HCl, 100μl) of trypsin (EC 3.4.21.4, 2×crystallized; from bovine pancreas, Koch–Light Laboratories Ltd.; 250μg) which had been treated with 1-chloro-4-phenyl-3-L-tosyl-*p*-sulphonamidobutan-2-one (TPCK; Sigma Chemical Co.) to remove chymotryptic activity, as described by Kostka & Carpenter (1964). Three further similar additions of trypsin solution were made at 15min intervals. At the end of 1h the hydrolysis was stopped by freezing and freeze-drying. The resultant mixture of glycopeptides and peptides was initially purified by gel chromatography on a column (1.2cm×150cm) of Sephadex G-75 run at 10ml/h (4°C) in 0.5% (w/v) NH_4HCO_3 (Fig. 2). Fractions were analysed for peptides spectrophotometrically at 225, 245 and 275nm, for carbohydrates by using the cysteine assay on a microscale (Kennedy & Butt, 1969) and for trypsin activity by incubation (of a 100μl sample) with 0.25M-benzoyl-L-arginine ethyl ester as substrate in 0.1M-sodium phosphate buffer, pH7.0 (500μl) at 25°C for 30min; the increase in absorption at 253nm was determined. Fractions were pooled and freeze-dried.

Final purification of the glycopeptides was achieved by ion-exchange chromatography on a column (0.67cm×37cm) of DEAE-Sephadex A-25 by using a stepwise gradient of NH_4HCO_3 (0.04, 0.15, 0.26, 0.38, 0.45 and 0.7M, 4°C; Fig. 3). Fractions were analysed for peptides and carbohydrates as described above (Fig. 3), pooled and freeze-dried.

Methylation analysis of subunit α and its glycopeptides

Methylation of the α subunit and its glycopeptides was performed by the dimethyl sulphinyl carbanion method (Hakomori, 1964) as reported for follicle-stimulating hormone (Kennedy & Chaplin, 1972) by using about 1μmol of glycoprotein or glycopeptide. The dimethyl sulphoxide solutions containing the methylated products were made 50% (w/v) aqueous by the addition of water and desalted on a column (1.8cm×56cm) of Sephadex G-15 run at 30ml/h (4°C) in 0.5% (w/v) NH_4HCO_3. The fractions were analysed for peptides and carbohydrates as described previously and for salt by use of a conductivity cell (Fig. 4 for subunit α, glycopeptides 1 and 2 gave similar protein, carbohydrate and salt elution

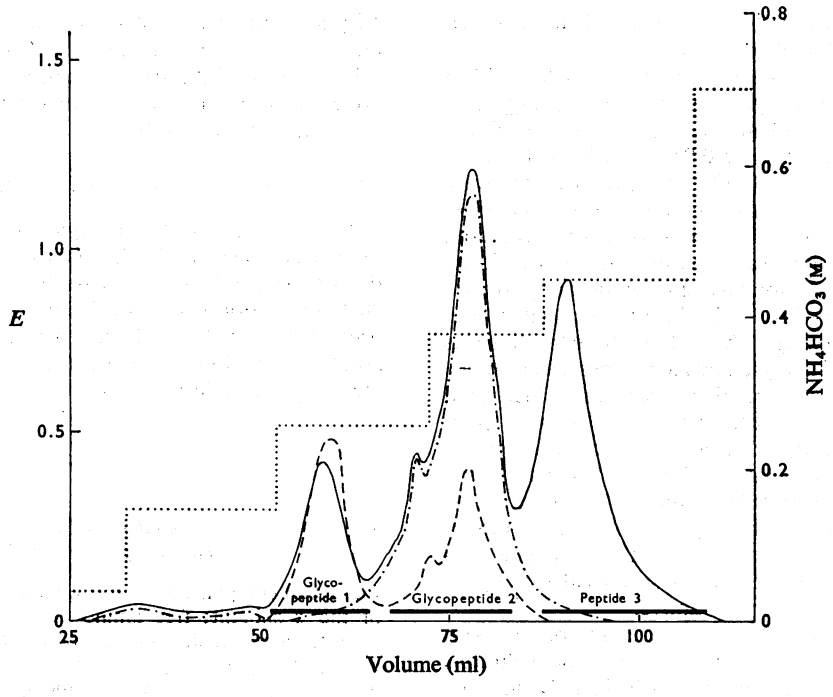

Fig. 3. *Fractionation of glycopeptide fraction A on a column of DEAE-Sephadex A-25*

——, Peptide content (E_{245}); -·-·-·, tyrosine content (E_{275}); ----, carbohydrate content by spectrophotometric assay (E_{415}); ·····, concentration of NH_4HCO_3 in eluate. Fractions pooled are indicated by the solid bars.

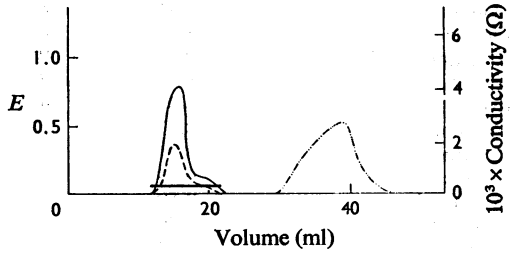

Fig. 4. *Desalting of methylated subunit α on a column of Sephadex G-15*

——, Protein content (E_{275}); ----, carbohydrate content by spectrophotometric assay (E_{415}); -·-·-·-, salt content (conductivity). Fractions pooled are indicated by the solid bar.

profiles). The methylated products were freeze-dried and, after hydrolysis in 90% (w/v) formic acid and 0.125M-H_2SO_4, reduction and acetylation, analysed by g.l.c. and mass spectrometry all as described by Kennedy & Chaplin (1972).

Carbohydrate sequence studies of subunit α and its glycopeptides

The glycopeptides (approx. 1 μmol of each) were desialylized with 0.1M-HCl (1ml, 1h, 80°C). The neuraminic acids and other monosaccharides released were separated from the glycopeptides by fractionation on a column (1.8cm×60cm) of Sephadex G-25 (fine grade), run at 18ml/h (4°C) in 0.5% (w/v) NH_4HCO_3. The fractions were analysed as described above, pooled and freeze-dried. The low-molecular-weight fraction was further analysed for *N*-acetyl- and *N*-glycollyl-neuraminic acid by the spectrophotometric process described above and for monosaccharides by g.l.c. as their trimethylsilyl ethers (Chambers & Clamp, 1971).

The desialylized glycopeptides were individually treated with α-D-mannosidase (1mg) in 0.05M-sodium acetate buffer, pH5.0, containing 0.1M-NaCl and 0.1mM-$ZnSO_4$, for 24h at 37°C with the addition of 1% (v/v) toluene to prevent bacterial growth. The glycopeptides were separated from the enzyme and monosaccharides by fractionation on the column of Sephadex G-25 (fine grade) and the

monosaccharides analysed by g.l.c. as their trimethylsilyl ethers as described above.

The treated glycopeptides were incubated with a mixture of β-*N*-acetylglucosaminidase (1mg) and β-D-galactosidase (3mg) in 0.05M-sodium citrate buffer, pH6.0, containing 0.05M-NaCl and 1% (v/v) toluene, for 24h at 37°C. The glycopeptides were separated from the enzymes and monosaccharides as described above. The monosaccharide fraction, after freeze-drying, was separated from the citrate buffer by ion-exclusion on a chromatography column (1.2cm×40cm) of AG 50WX2 (Li^+ form) (Barker *et al.*, 1969) run in water. The carbohydrate fraction was freeze-dried and analysed by g.l.c. as described above.

The treated glycopeptides were then incubated with a mixture of α-D-mannosidase (2mg), β-*N*-acetylglucosaminidase (2mg) and β-D-galactosidase (1mg) in 0.05M-acetate buffer, pH5.0, containing 0.1M-NaCl, 0.1mM-$ZnSO_4$ and 1% (v/v) toluene, for 24h at 37°C. The glycopeptides and monosaccharides were separated from the enzymes on Sephadex G-25, and the monosaccharides analysed by g.l.c., both as described above.

Finally the glycopeptides were hydrolysed with 2M-trifluoroacetic acid (100°C, 3h), the solutions evaporated and neutralized with NH_4HCO_3 and the monosaccharides separated on Sephadex G-25 as above. The freeze-dried monosaccharide fractions were redissolved in anhydrous methanol (1ml), and acetic anhydride (0.2ml) was added to re-acetylate the 2-amino-2-deoxyhexoses. After 6h at 20°C, water was added, the samples were evaporated to dryness and analysed as their trimethylsilyl ethers as described above.

Reaction of subunit α and its glycopeptides with D-galactose oxidase

Subunit α and its glycopeptides were assayed for D-galactose before and after they were desialylized by acid hydrolysis or by the use of neuraminidase. An enzymic method using D-galactose oxidase (Boehringer Mannheim G.m.b.H., Tutzing, München, W. Germany), peroxidase and ammonium 2,2′-azino-bis(3-ethylbenzothiazoline-6-sulphonate) as described for D-glucose oxidase by Kennedy & Doyle (1973) was developed for the assay. Incubations were carried out at pH7.0, and the amount of D-galactose was calculated from a calibration curve obtained with D-galactose.

Results

Amino acid and carbohydrate analyses of subunit α and its glycopeptides

The total amino acid and carbohydrate content of subunit α (Table 1) accounted for 90% of its weight; the rest was assumed to be bound water. The value for $E^{1\%}_{1cm}$ at 280nm was 4.24. Compositions and yields of the tryptic fragments are given in Table 2. The amino acid analysis of the tryptic peptides could be interpreted in terms of the known cleavage pattern of trypsin on the sequence.

Methylation analyses of subunit α and its glycopeptides

The i.r. spectra of the methylated material showed the marked decrease in the –OH and –NH– absorptions (3360cm^{-1}) and increase in the absorption at 2820cm^{-1} noticed previously for methylated glycoproteins (Kennedy & Chaplin, 1972). The chromatographic retention times of the methyl ethers of alditol acetates derived from the methylated glyco-

Table 1. *Component analyses of the α subunit of human chorionic gonadotrophin*

Individual components were liberated by acid hydrolysis of the α subunit; free amino acids and basic carbohydrates were then determined by ion-exchange chromatography, tryptophan by u.v. adsorption, neutral carbohydrates by g.l.c. as their trimethylsilyl ethers, and neuraminic acid by spectrophotometry; for details, see the text.

Component	Amount of component from (mol/mol of subunit)	
	Component analyses	Sequence studies*
L-Aspartic acid	5.91	6
L-Threonine	7.81	8
L-Serine	8.70	8
L-Glutamic acid	9.99	9
L-Proline	6.78	7
Glycine	4.42	4
L-Alanine	4.83	5
L-Cystine	3.91	5
L-Valine	6.73	7
L-Methionine	2.50	3
L-Isoleucine	1.06	1
L-Leucine	4.04	4
L-Tyrosine	4.17	4
L-Phenylalanine	4.04	4
L-Histidine	2.83	3
L-Lysine	5.96	6
L-Arginine	3.33	3
L-Tryptophan	0.00	0
D-Mannose	6.48	—
D-Galactose	2.83	—
D-Glucose	0.00	—
L-Fucose	<0.02	—
N-Acetylglucosamine	5.55	—
N-Acetylgalactosamine	0.00	—
N-Acetylneuraminic acid	2.48	—
N-Glycollylneuraminic acid	0.16	—

* From Bellisario *et al.* (1973).

Table 2. *Analysis of peptides and glycopeptides from the α subunit of human chorionic gonadotrophin*

Peptides were released from the reduced carboxymethylated α subunit by the action of trypsin and separated by gel filtration and ion-exchange chromatography; for details, see the text. Component analyses as described in the legend to Table 1 were performed. The numbers in parentheses give the expected amino acid composition on the basis of the cleavage pattern: residues 35–36, 45–46, 51–52, 63–64, 68–69, 100%; residues 75–76, 70%; residues 44–45, 50%; residues 91–92, 40%; and residues 42–43, 0%.

	Amount of component (mol/mol)				
	Peptide			Glycopeptide	
Component no. ...	1	2	3	1	2
Peptide fragments ...	T_3, T_4, T_7, T_{10}	T_6, T_9, T_{14}	T_1	T_8	T_{11}, T_{12}, T_{13}
Components					
L-Aspartic acid	0.29 (0.0)	1.00 (1.0)	3.13 (3.0)	0.98 (1.0)	1.03 (1.0)
L-Threonine	1.84 (2.7)	0.04 (0.0)	1.09 (1.0)	2.06 (2.0)	2.13 (2.3)
L-Serine	1.07 (1.0)	1.46 (1.4)	2.51 (2.0)	2.59 (2.0)	1.56 (1.6)
L-Glutamic acid	0.66 (1.0)	0.09 (0.0)	6.15 (6.0)	1.20 (1.0)	1.60 (1.0)
L-Proline	2.00 (2.0)	0.00 (0.0)	4.18 (5.0)	0.00 (0.0)	0.76 (0.0)
Glycine	1.33 (1.4)	0.21 (0.0)	2.63 (2.0)	0.74 (0.0)	1.13 (0.6)
L-Alanine	1.10 (1.0)	0.10 (0.0)	1.98 (2.0)	1.31 (1.0)	1.54 (1.0)
Carboxymethyl-L-cysteine	0.00 (0.0)	0.00 (0.0)	3.66 (5.0)	2.00 (2.0)	2.28 (3.0)
L-Valine	1.46 (2.4)	0.00 (0.0)	1.16 (1.0)	1.96 (2.0)	1.55 (1.6)
L-Methionine	0.78 (1.7)	0.00 (0.0)	0.74 (1.0)	0.00 (0.0)	0.18 (0.3)
L-Isoleucine	0.00 (0.0)	0.00 (0.0)	0.86 (1.0)	0.00 (0.0)	0.26 (0.0)
L-Leucine	1.50 (2.0)	0.03 (0.0)	1.83 (2.0)	0.11 (0.0)	0.31 (0.0)
L-Tyrosine	0.93 (1.0)	0.76 (1.0)	0.15 (0.0)	0.00 (0.0)	1.61 (2.0)
L-Phenylalanine	0.52 (0.7)	0.01 (0.0)	3.08 (3.0)	0.08 (0.0)	0.41 (0.3)
L-Histidine	0.13 (0.0)	0.06 (0.0)	0.20 (0.0)	0.17 (0.0)	2.77 (3.0)
L-Lysine	2.40 (3.2)	0.16 (0.5)	0.05 (0.0)	1.20 (1.0)	1.25 (1.3)
L-Arginine	1.08 (1.0)	1.05 (1.0)	1.00 (1.0)	0.11 (0.0)	0.14 (0.0)
Yield (%)	67	67	61	69	72
D-Mannose				4.06	2.79
D-Galactose				1.11	1.19
L-Fucose				0.00	0.00
N-Acetylglucosamine				3.13	2.86
N-Acetylneuraminic acid				1.10	1.27
N-Glycollylneuraminic acid				0.07	0.09

protein or glycopeptides were corrected for minor variations between runs by reference to the elution position of the 1,5-di-*O*-acetyl-2,3,4,6-*O*-methyl-D-glucitol used as standard. The components of the chromatograms were identified by comparisons of the corrected retention times with those obtained for standard methylated monosaccharides (Table 3). The identifications were confirmed by mass-spectral analysis of the components from preparative runs and reference to patterns obtained for the standard compounds. Certain minor peaks observed in the chromatograms were considered to have arisen from the relatively unstable neuraminic acid units. No neuraminic acid units were recovered in sufficient quantity to be analysed, owing presumably to the lability of such residues to acidic conditions.

Carbohydrate sequence analyses of subunit α and its glycopeptides

The patterns of the sequential release of monosaccharides from the two glycopeptides show distinct differences (Table 4) in the release of D-mannose and *N*-acetylglucosamine.

Reaction of subunit α and its glycopeptides with D-galactose oxidase

The period necessary for full chromophore development in the D-galactose oxidase assay of the glycopeptides (Table 5) was five times the normal period. This was thought to be due to inhibition of the reaction by the carboxymethylcysteine groups. However, the sum of the amounts of D-galactose

Table 3. *Methylation analysis of the α subunit of human chorionic gonadotrophin*

Methylation was performed by using dimethyl sulphoxide (Kennedy & Chaplin, 1972) and the resultant methyl sugars released by acid hydrolysis. Identification and quantitation were achieved by g.l.c. and mass spectrometry of their reduced acetylated derivatives. For details, see the text. Retention times (t_R) are expressed relative to that of the appropriate standard on which the structural assignment was ultimately based. There were no significant differences between the analyses. Amounts of components are expressed as a percentage of total methylated components recovered, as measured from peak areas. Values in parentheses are the molar ratios of the carbohydrate units present.

Component no.	Relative retention times (t_R)	Amount of component in: Methylated subunit α	Methylated glycopeptide 1	Methylated glycopeptide 2	Identification (based on g.l.c. and mass-spectral data)	Linkage situation in subunit α
1	1.00	8 (1)	13 (1)	1	1,5-Di-*O*-acetyl-2,3,4,6-tetra-*O*-methyl-D-mannitol	*O*-D-mannopyranosyl-(1→
2	1.05	7 (1)	14 (1)	2	1,2,5-Tri-*O*-acetyl-3,4,5-tri-*O*-methyl-D-mannitol	→2)-*O*-D-mannopyranosyl-(1→
3	0.96	8 (1)	13 (1)	3	1,5,6-Tri-*O*-acetyl-2,3,4-tri-*O*-methyl-D-mannitol	→6)-*O*-D-mannopyranosyl-(1→
4	0.94	3 (1)	2	16 (1)	1,5-Di-*O*-acetyl-3,4,6-tri-*O*-methyl-2-acetamido-2-deoxy-D-glucitol	2-acetamido-2-deoxy-D-glucosyl-(1→
5	1.05	14 (2)	9 (1)	14 (1)	1,5,6-Tri-*O*-acetyl-2,3,4-tri-*O*-methyl-D-galactitol	→6)-*O*-D-galactopyranosyl-(1→
6	1.03	23 (3)	13 (1)	30 (2)	1,2,5,6-Tetra-*O*-acetyl-3,4-di-*O*-methyl-D-mannitol	→2 } →6 } -*O*-D-mannopyranosyl-(1→
7	1.07	37 (5)	37 (3)	34 (2)	1,5,6-Tri-*O*-acetyl-3,4-di-*O*-methyl-2-acetamido-2-deoxy-D-glucitol	→6)-*O*-(2-acetamido-2-deoxy-D-glucopyranosyl)-(1→
Yield of methylated material (% of that before hydrolysis)		93	83	78		

Table 4. *Carbohydrate sequence studies on the glycopeptides from the α subunit of human chorionic gonadotrophin*

The glycopeptides were treated sequentially with acid and glycoside hydrolases. The monosaccharides released were separated from high-molecular-weight materials by gel filtration, and identified and quantified by g.l.c. as their trimethylsilyl ethers; for details, see the text.

	Carbohydrate released (mol/mol of glycopeptide)							
	Glycopeptide 1				Glycopeptide 2			
Treatment (in chronological order)	Neuraminic acid	D-Galactose	D-Mannose	*N*-Acetyl-glucosamine	Neuraminic acid	D-Galactose	D-Mannose	*N*-Acetyl-glucosamine
1. 0.01M-H_2SO_4, 1h, 80°C	1.10	0.00	0.07	0.00	1.27	0.05	0.05	0.00
2. α-D-Mannosidase	0.00	0.34	0.94	0.00	0.00	1.01	1.59	0.00
3. β-*N*-Acetylglucos-aminidase+β-D-galactosidase	0.00	0.73	0.80	2.22	0.00	0.02	0.24	0.67
4. α-D-Mannosidase+ β-*N*-acetylglucos-aminidase+β-D-galactosidase	0.00	0.05	1.38	0.00	0.00	0.11	0.31	1.15
5. 2M-trifluoroacetic acid, 3h, 100°C	0.00	0.00	0.86	0.91	0.00	0.00	0.59	1.05

Table 5. D-*Galactose oxidase treatment of the α subunit of human chorionic gonadotrophin and its glycopeptides*

The α subunit and its glycopeptides, and modified forms thereof, were treated with D-galactose oxidase and the amount of the D-galactose units oxidized thereby were determined spectrophotometrically; for details, see the text.

Sample	Amount of neuraminic acid released (mol/mol)	Amount of D-galactose oxidized (mol/mol)
Subunit α	0.00	0.00
Subunit α hydrolysed, 0.02M-HCl, 1h, 80°C	2.71	2.67
Subunit α hydrolysed, 0.01M-H_2SO_4, 1h, 80°C	2.42	2.07
Subunit α hydrolysed, neuraminidase, 16h, 37°C	2.78	2.80
Subunit α glycopeptide 1	0.00	0.00
Subunit α glycopeptide 1, hydrolysed 0.01M-H_2SO_4, 1h, 80°C	1.17	1.08*
Subunit α glycopeptide 2	0.00	0.00
Subunit α glycopeptide 2, hydrolysed 0.01M-H_2SO_4, 1h, 80°C	1.36	1.23*

* Colour development period notably prolonged.

derivatives formed in both the desialylized glycopeptides equalled that in the desialylized glycoprotein.

Discussion

The amino acid analysis of the subunit α is in agreement with the reported sequence and the $E^{1\%}_{1cm}$ at 280nm is close to that previously reported (4.3; Morgan *et al.*, 1974). The carbohydrate analysis of subunit α gave slightly lower values than that reported previously by Bahl (1973) but substantially the same relative proportions. In particular, the absence of L-fucose and *N*-acetylgalactosamine are endorsed. There seemed to be random heterogeneity in the amino group of the neuraminic acid residues, 93% being *N*-acetylated whereas the remainder were *N*-glycollylated, the value of 7% for the latter being outside the limits of variation of background colour and of colour produced from pure *N*-acetylneuraminic acid. The subsequent use of the glycosidases, which are specific for the D configurations, demonstrated, by virtue of the cleavages obtained, that the assumptions of the D configurations for the neuraminic acid, D-mannose, D-galactose and *N*-acetylglucosamine units are correct. The susceptibility of the *N*-acetylglucosamine units to β-*N*-acetylglucosaminidase also confirmed the *N*-acetylated state of these units. (It is appreciated that these deductions are not valid for the *N*-acetylglucosamine units actually involved in the glyco-

peptide linkages since they were released by acid hydrolysis.)

The use of trypsin containing chymotryptic activity (Bahl, 1969*b*; Bellisario *et al.*, 1973) generally has yielded complex mixtures of glycopeptides in low yield, which are difficult to fractionate. By use of trypsin freed from chymotrypsin, high yields (T_8, 69%; T_{11}, 72%) of the glycopeptides have been achieved. These yields compare very favourably with those obtained (T_8, 38%; T_{11}, 30%) by Bellisario *et al.* (1973). Carboxymethylations of subunit α gave a very good yield of carboxymethylated subunit α (92%), which on tryptic hydrolysis could be fractionated into five fractions, thus showing that the neuraminic acid present did not interfere with the action of trypsin. The cleavage pattern showed hydrolysis adjacent to L-lysine and L-arginine as expected from the specificity of trypsin but no hydrolysis, due to chymotrypsin contamination, adjacent to L-tyrosine, L-phenylalanine, L-carboxymethylcysteine or L-methionine. It was noticed that there was incomplete hydrolysis between residues 42 and 46 (Fig. 5), where one residue of L-arginine and two of L-lysine lie close together. Two of the fractions obtained from the tryptic hydrolysate consisted of mixtures of low-molecular-weight peptides as deduced from component analysis and sequence data. Peptide fraction 1 (Fig. 1) consisted of a mixture of peptides T_3 (mol.wt. 1051), T_4 (mol.wt. 1180), T_7 (mol.wt. 721) and T_{10} (mol.wt. 839) and peptide fraction 2 (Fig. 1) consisted of a mixture of peptides T_6 (mol.wt. 147), T_9 (mol.wt. 530) and T_{14} (mol.wt. 105). Peptide fraction 3 consisted of the large *N*-terminal peptide, on the basis of amino

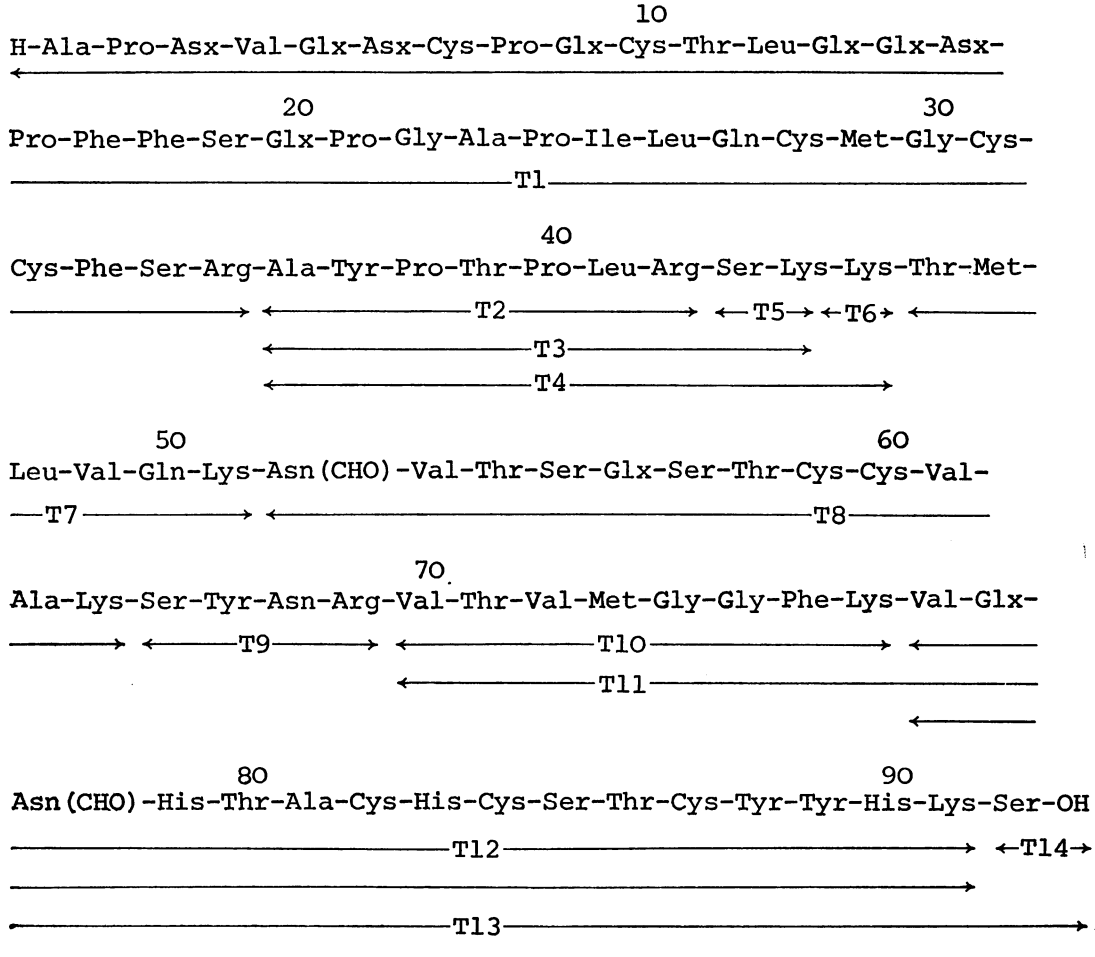

Fig. 5. *Sequence of subunit α (Ballisario et al., 1973) showing breakdown pattern on tryptic hydrolysis*

acid analysis and molecular weight (4099; five carboxymethyl-L-cysteine groups).

Glycopeptides 1 and 2 were eluted in overlapping positions (as glycopeptide fraction A, Fig. 2) on gel chromatography on account of their similar molecular weights (glycopeptide 1; mol.wt. 3060, two carboxymethyl-L-cysteine groups. Glycopeptide 2; mol.wt. 3450, three carboxymethyl-L-cysteine groups; molecular weights calculated from component analyses). The terminal peptide co-eluted with the glycopeptides, but all three were separated by ion-exchange chromatography on the basis of their contents of free carboxylic acid groups. Glycopeptide 2 showed slight heterogeneity on the basis of its amino acid composition and ion-exchange chromatography, but this is explained by non-quantitative tryptic hydrolysis between residues 75–76 and 91–92 and does not affect the separation of the two carbohydrate moieties of subunit α. Carbohydrate and amino acid analyses conducted across the peak for glycopeptide 2 followed one another closely (Fig. 2). The species present all contain the same carbohydrate moiety and all of them are easily separated collectively and quantitatively from glycopeptide 1 fraction and the large *N*-terminal peptide (T_1).

The sum of the amounts of carbohydrates present in the two glycopeptides is congruent with that originally present in the intact α subunit and this, together with the lack of detection of other carbohydrate-containing fractions in the purification process, further indicated that only two carbohydrate chains are present in the subunit.

The methylation technique of Hakomori (1964) has been found to methylate fully both polysaccharides (Hellerqvist *et al.*, 1968; Björndal & Lindberg, 1969) and glycoproteins (Kennedy & Chaplin, 1972). The methylated subunit and its glycopeptides were separated from the methylation media by gel filtration on Sephadex G-15 in good yield (78–93%), similar to that obtained on methylating follicle-stimulating hormone (89%; Kennedy & Chaplin, 1972) and are much greater than those obtained by Hakomori (1964) (15–25%). That incorrect structural assignments could not be made because of incomplete methylation was demonstrated by the singularity of the types of methylated-carbohydrate patterns obtained; incomplete methylation would give rise to a variety of methylated derivatives for a particular carbohydrate unit. The data indicate that the various carbohydrate units are present in the subunit α as the pyranose and not furanose forms. The production of methylated derivatives of *N*-acetylglucosamine which are not *N*-methylated has been noted previously (Kennedy & Chaplin, 1972; Watanabe & Eukimbora, 1974), whereas earlier reports (Björndal *et al.*, 1970) claimed that complete *N*-methylation was always obtained by Hakomori methylation.

From the results of analysis of the methylated samples of both subunit α and the glycopeptides 1 and 2 (Table 3) it is possible to make structural assignments to the carbohydrate units of subunit α. It is immediately apparent that the two glycopeptides have different internal carbohydrate structures based, however, on similar types of linkage.

The sequential enzyme-degradation studies on the two glycopeptides, carried out in parallel under identical conditions, also show this difference in internal carbohydrate structure (Table 4). Although the sequential release of neuraminic acid and D-galactose are similar in both glycopeptides the rates of release of D-mannose and N-acetylglucosamine were different.

Most of the D-galactose residues of subunit α, glycopeptide 1 and glycopeptide 2 (95%, 92% and 91% respectively) were susceptible to D-galactose oxidase after removal of the terminal neuraminic acid residues. Although the exact response of terminal D-galactose residues cannot be assumed to be identical with that of the D-galactose standard, since combined D-galactose gives a molar response in the assay dependent on its substitution position (Schlegel *et al.*, 1968), the results indicate that all of the neuraminic acid residues are linked directly to the D-galactose residues, in agreement with studies on intact human chorionic gonadotrophin (Tsuruhara *et al.*, 1972). The 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methyl-D-hexitol produced from methylated glycopeptide 1 could not therefore be due to the presence of terminal D-galactose residues. Accordingly, it is concluded that the data demonstrate the presence of a terminal non-reducing D-mannose residue, since 1,5-di-*O*-acetyl-2,3,4,6-tetra-*O*-methyl-D-mannitol and -galactitol have similar g.l.c. retention times and identical mass spectra. The remaining D-mannose residues are shown to be situated internally within the glycopeptide structure and linked linearly in the chain or at branch points. *N*-Acetylglucosamine residues are linked 1–6 internally and in glycopeptide 2 an additional residue is terminal non-reducing. The carbohydrate-sequence studies support the assumption from the amino acid-sequence data (Bellisario *et al.*, 1973) that the carbohydrate–peptide linkage is the well-known 2-acetamido-1-(4-L-aspartoyl)-1,2-dideoxy-β-D-glucosylamine.

From these studies structures of glycopeptides 1 and 2 may be proposed. The most likely structures that can be deduced are those shown in Fig. 6; these are subject to the following variations. In glycopeptide 1 it is uncertain which of the D-mannose units in the tri-(D-mannosyl) sequence is branched, and the terminal non-reducing partial chains [*O*-α-D-mannopyranosyl-(1 → 6)-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)-(1→] and [5-acetamido/hydroxyacetamido-3,5-dideoxy-D-*glycero*-D-*galacto*-2-α-nonulosonyl-(2 → 6)-β-D-galactopyranosyl)-1 →]

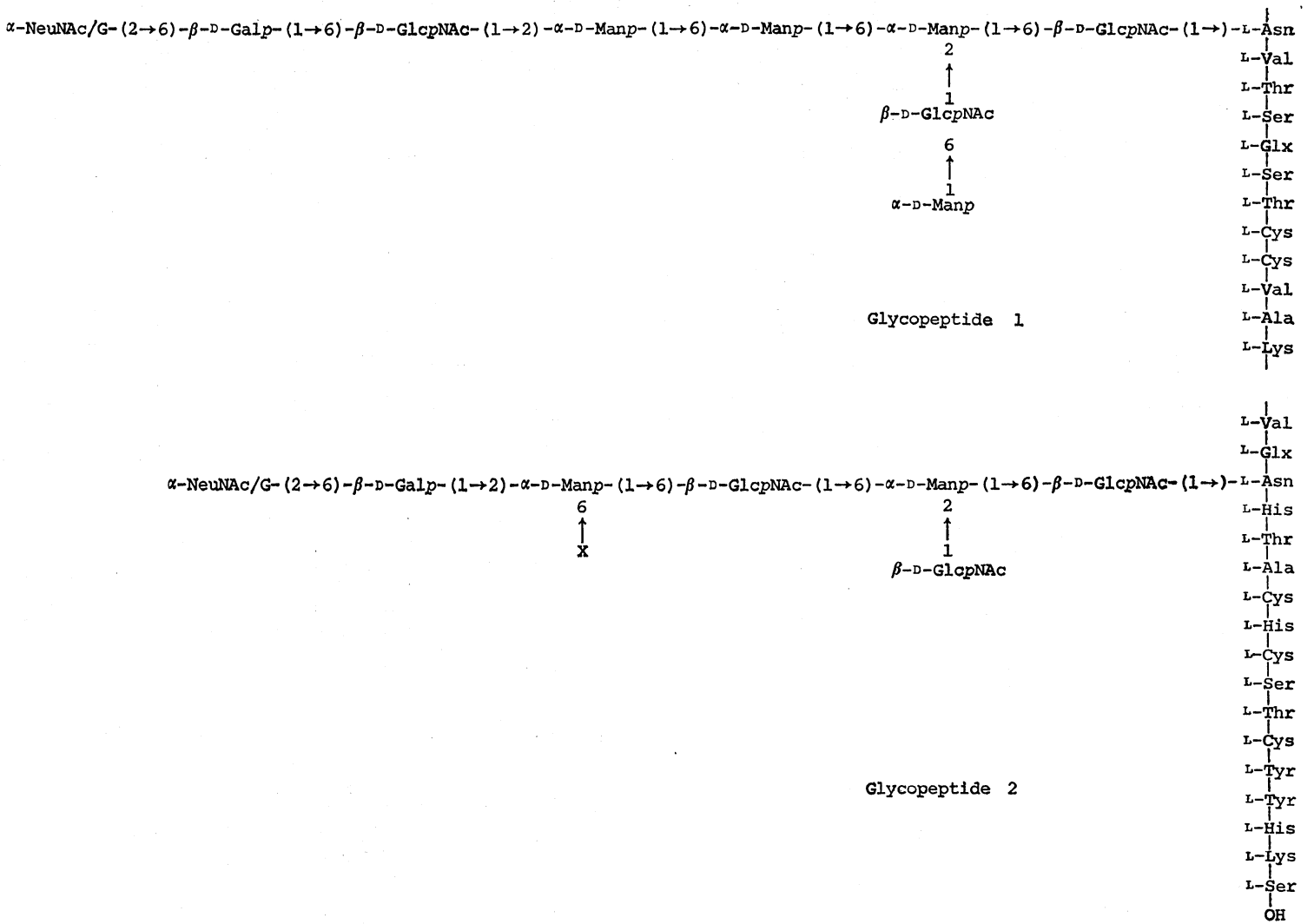

Fig. 6. ***Proposed structures of glycopeptides*** **1** ***and*** **2** ***of the carbohydrate moieties of subunit*** **α**
G, Glycollyl; X, unknown substituent (see the text).

may be interchanged. In glycopeptide 2, the partial terminal non-reducing chains attached to the branched D-mannose residues may be interchanged. The unknown material linked to D-mannose in glycopeptide 2 could be an artifact caused by undermethylation, but this seems unlikely, owing to the similar results obtained in analysis of methylated subunit α. It may be, however, that part of the glycopeptide chain protects, by hydrogen bonding or steric effect, this position on the D-mannose residue from methylation or that it is due to an acetyl, or similar, blocking group. The proposals agree substantially with, and confirm, the earlier periodate-oxidation studies on intact human chorionic gonadotrophin (Kennedy *et al.*, 1974). They also show general but not specific agreement with the type of structure proposed by Bahl (1969*b*) for his, probably impure, glycopeptides derived from desialylized human chorionic gonadotrophin.

On the other hand, the present data provide an interesting comparison with that (Kennedy & Chaplin, 1972) for the related glycoprotein hormone, human follicle-stimulating hormone, in which the D-galactopyranose units are 1,2-linked, the D-mannopyranose units exist in three forms, terminal non-reducing, 1,6-linked and 1,3,4-linked branch points, and the *N*-acetylglucosamine units are 1,6-linked. It therefore appears that in spite of the immunological and peptide sequence similarities of the α-subunits of the gonadotrophins and thyroid-stimulating hormone, the carbohydrate structures not only differ in the monosaccharide sequence but also in the linkage positions between the units.

We thank Professor M. Stacey for his interest in this work. M. F. C. acknowledges receipt of a Birmingham University Faculty of Science and Engineering Research Fellowship.

References

Ashitaka, Y., Tokura, Y., Tane, M., Mochizuki, M. & Tojo, S. (1970) *Endocrinology* **87**, 233–244

Bahl, O. P. (1969*a*) *J. Biol. Chem.* **244**, 567–574

Bahl, O. P. (1969*b*) *J. Biol. Chem.* **244**, 575–583

Bahl, O. P. (1973) *Horm. Proteins Pept.* **1**, 171–199

Bahl, O. P., Carlsen, R. B. & Bellisario, R. (1973) *Proc. Int. Congr. Endocrinol. 4th* 667–672

Barker, S. A., Hatt, B. W., Kennedy, J. F. & Somers, P. J. (1969) *Carbohydr. Res.* **9**, 327–334

Bell, J. J., Canfield, R. E. & Sciarra, J. J. (1969) *Endocrinology* **84**, 298–307

Bellisario, R., Carlsen, R. B. & Bahl, O. P. (1973) *J. Biol. Chem.* **248**, 6796–6809

Björndal, H. & Lindberg, B. (1969) *Carbohydr. Res.* **10**, 79–85

Björndal, H., Hellerqvist, C. G., Lindberg, B. & Svensson, S. (1970) *Angew. Chem. Int. Ed. Engl.* **9**, 610–619

Butt, W. R. & Kennedy, J. F. (1971) in *Structure–Activity Relationships of Protein and Polypeptide Hormones*, Part 1 (Margoulies, M. & Greenwood, F. C., eds.), pp. 115–121, Excerpta Medica Foundation, International Congress Series no. 241, Amsterdam

Carlsen, R. B., Bahl, O. P. & Swaminathan, N. (1973) *J. Biol. Chem.* **248**, 6810–6827

Chambers, R. F. & Clamp, J. R. (1971) *Biochem. J.* **125**, 1009–1018

Chaplin, M. F. (1970) Ph.D. Thesis, University of Birmingham

Chaplin, M. F., Gray, C. J. & Kennedy, J. F. (1970) in *Gonadotrophins and Ovarian Development* (Butt, W. R., Crooke, A. C. & Ryle, M., eds.), pp. 77–98, Livingstone, Edinburgh

Chung, D., Sairam, M. R. & Li, C. H. (1973) *Arch. Biochem. Biophys.* **159**, 678–682

Cornell, J. S. & Pierce, J. G. (1973) *J. Biol. Chem.* **248**, 4327–4333

Hakomori, S. I. (1964) *J. Biochem. (Tokyo)* **55**, 205–208

Hellerqvist, C. G., Lindberg, B., Svensson, S., Holme, T. & Lindberg, A. A. (1968) *Carbohydr. Res.* **8**, 43–55

Kennedy, J. F. (1971) in *Structure–Activity Relationships of Protein and Polypeptide Hormones*, Part 2 (Margoulies, M. & Greenwood, F. C., eds.), pp. 360–361, Excerpta Medica Foundation, International Congress Series no. 241, Amsterdam

Kennedy, J. F. (1973) *Endocrinol. Exp.* **7**, 5–18

Kennedy, J. F. & Butt, W. R. (1969) *Biochem. J.* **115**, 225–229

Kennedy, J. F. & Chaplin, M. F. (1972) *Biochem. J.* **130**, 417–423

Kennedy, J. F. & Doyle, C. E. (1973) *Carbohydr. Res.* **28**, 89–92

Kennedy, J. F., Chaplin, M. F. & Stacey, M. (1974) *Carbohydr. Res.* **36**, 369–377

Klenk, E. & Ullenbruck, G. (1966) *Hoppe-Seyler's Z. Physiol. Chem.* **307**, 266–271

Kostka, V. & Carpenter, F. H. (1964) *J. Biol. Chem.* **239**, 1799–1803

Li, S.-C. & Li, Y.-T. (1970) *J. Biol. Chem.* **245**, 5153–5160

Li, Y.-T. (1967) *J. Biol. Chem.* **242**, 5474–5480

Morgan, F. J., Canfield, R. E., Vaitukailis, J. L. & Ross, G. T. (1974) *Endocrinology* **94**, 1601–1606

Pierce, J. G., Liao, T. H., Carlsen, R. B. & Reimo, T. (1971*a*) *J. Biol. Chem.* **246**, 866–872

Pierce, J. G., Bahl, O. P., Cornell, J. S. & Swaminathan, N. (1971*b*) *J. Biol. Chem.* **246**, 2321–2324

Reichert, L. E. (1972) *Endocrinology* **90**, 1119–1122

Schlegel, R. A., Gerbeck, C. M. & Montgomery, R. (1968) *Carbohydr. Res.* **7**, 193–199

Snaith, S. M. & Levvy, G. A. (1968) *Biochem. J.* **110**, 663–670

Swaminathan, N. & Bahl, O. P. (1970) *Biochem. Biophys. Res. Commun.* **40**, 422–427

Tsuruhara, T., Dafau, M. L., Hickman, J. & Catt, K. J. (1972) *Endocrinology* **91**, 296–301

Ward, D. N., Reichert, L. E., Liu, W.-K., Nahm, H. S., Hsia, J., Lamkin, W. M. & Jones, N. S. (1973) *Recent Prog. Horm. Res.* **29**, 533–556

Warren, L. (1959) *J. Biol. Chem.* **234**, 1971–1975

Watanabe, K. & Eukimbora, T. (1974) *Agric. Biol. Chem.* **38**, 1973–1980