JOURNAL OF MASS SPECTROMETRY, VOL. 32, 33–42 (1997)

Mass Spectrometric Characterization of the *β*-Subunit of Human Chorionic Gonadotropin

ChuanLiang Liu and Larry D. Bowers*
Athletic Drug Testing and Toxicology Laboratory, Department of Pathology and Laboratory Medicine,
Indiana University Medical Center, Medical Science Building A-128, 635 Barnhill Drive, Indianapolis, Indiana 46202-5120,
USA

A high-performance liquid chromatographic/electrospray mass spectrometric (HPLC/MS) technique is described for the characterization of the *β*-subunit of the glycopeptide human chorionic gonadotropin (hCG). The *β*-subunit of hCG was dissociated from the *α*-subunit using 0.1% trifluoroacetic acid (TFA) and separated by reversed-phase HPLC using a 0.1% TFA–acetonitrile gradient. Although reductive alkylation with 4-vinylpyridine allowed direct observation of the intact β-subunit of hCG by HPLC/MS due to the increase in charge, the heterogeneity of the carbohydrate fractions resulted in poor detection limits and extremely complex spectra. After reductive alkylation with either iodoacetate or 4-vinylpyridine, tryptic fragments of either the *α*- or *β*-subunit can be observed using reversed-phase HPLC/MS. HPLC/MS data were consistent with the reported primary sequence, although oligosaccharide attachment sites at both ^{127}Ser and ^{132}Ser could not be documented. Microheterogeneity of the carbohydrate moiety on both N-glycosylation sites on the *β*-subunit could be readily observed. A larger degree of heterogeneity was observed on ^{13}Asn. Differences were also observed in the oligosaccharide distribution in three commerial preparations of hCG. Detection of the C-terminal portion of the *β*-subunit required enzymatic deglycosylation prior to HPLC/MS analysis.

J. Mass Spectrom. **32**, 33–42 (1997)

No. of Figures: 8 No. of Tables: 3 No. of Refs: 34

KEYWORDS: electrospray ionization; glycopeptide; human chorionic gonadotropin; high-performance liquid chromatography/mass spectrometry; vinylpyridine; glycosidases

INTRODUCTION

Human chorionic gonadotropin (hCG) is an important glycoprotein hormone which is monitored in pregnancy testing, cancer detection and doping control in sports. The physiological function of hCG is not completely understood, although it appears to maintain steroid secretion by the *corpus luteum* until the placenta can perform this function during pregnancy.[1,2] Thus it is universally used as an early marker of pregnancy. Ectopic secretion either in an abortive pregnancy or in choriocarcinoma also makes it an important biochemical marker.[3] Its availability as a pharmacological agent and ability to stimulate endogenous steroid production can potentially result in its abuse by athletes either to recover from anabolic steroid cycles or to increase directly testosterone concentration.[4] Current techniques for the detection of hCG and related glycopeptides in urine and blood are exclusively based on immunoassay technology. Although the immunogenic determinants of hCG[5] and the immunoassays developed from these antibodies[6] have been well characterized, immunoassays have not been widely accepted as forensic confirmation of the presence of hCG.

The intact hCG molecule is composed of two noncovalently bound subunits: an *α*-subunit of 92 amino acids and a *β*-subunit of 145 amino acids.[1,2] The *α*-subunit is essentially identical with those of the human pituitary glycoprotein hormones: thyroid stimulating hormone (TSH), follicle stimulating hormone (FSH) and lutropin (LH). The *β*-subunit distinguishes hCG from the other glycoprotein hormones. The hCG molecule forms a tightly folded structure with 11 internal disulfide bonds, six of which are within the *β*-chain.[7] The *β*-subunit reportedly contains two N-linked and four O-linked carbohydrate moieties, while the *α* subunit contains two N-linked polysaccharides.[1,2] The molecular mass of the *β*-chain is about 23 000 (7000 from carbohydrate) and that of the *α*-chain is 15 000 (3000 from carbohydrate). Intact hCG dimer, free *α*- and *β*-subunits, 'nicked' hCG, *β*-core fragment and a carboxy-terminal peptide fragment from the *β*-subunit have been identified in urine samples from different individuals.[8–12] Biochemical characterization of hCG has been accomplished by N-terminal amino acid sequencing for the peptide and monosaccharide analysis combined with different enzymatic digestions for the carbohydrates.[2,13–15]

There are several reasons for developing a mass spectrometric method for characterization and detection of hCG. A rapid, reliable analytical technique capable of characterizing the peptide and carbohydrate portions of hCG and related peptides would be useful for the development of analytical peptide standards. Not only would better characterized standards improve immunoassay calibration, but they may also allow accurate calibration of immunoassays directed against

---

* Correspondence to: L. D. Bowers.

CCC 1076–5174/97/010033–10

Received 15 March 1996
Accepted 10 September 1996

unusual carbohydrate microheterogeneity.[16] Development of a mass spectrometric method for forensic confirmation of hCG in urine would also be beneficial. For athletic drug testing, the analytical cut-off for hCG in urine has been proposed as 10 mIU ml^{-1} (about 50 fmol ml^{-1}).[17] Electrospray ionization mass spectrometry has demonstrated the limits of detection and information content necessary for this application.[18–20] Preliminary work showed that high-performance liquid chromatographic/mass spectrometric (HPLC/MS) detection of hCG at these concentrations from a urine matrix requires immunoaffinity extraction, the details of which are presented elsewhere.[21] To prepare a foundation for the development of an HPLC/MS method for the trace detection and confirmation of hCG in urine,[22] we report here mass spectrometric studies on the peptide and oligosaccharide structure of hCG.

EXPERIMENTAL

Chemicals

hCG (C5297; 3000 IU mg^{-1}), trypsin (treated with 1-tosylamide-2-phenylethyl chloromethyl ketone, TPCK, T8642) and neuraminidase (EC 3.2.1.18, Type X) were purchased from Sigma Chemical (St Louis, MO, USA). O-Glycanase (EC 3.2.1.97, Lot B4036) was purchased from Genzyme (Cambridge, MA, USA). Pregnyl, a pharmaceutical preparation of hCG, was purchased from Organon (W. Orange, NJ, USA) (Lot 0350193315). A standard preparation of hCG β-subunit (CR 121) was a gift from the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) (Bethesda, MD, USA). Sodium iodoacetate, 4-vinylpyridine, trifluoroacetic acid and all other chemicals of ACS reagent grade or better were also obtained from Sigma. High-purity water was obtained from a Millipore (Pleasanton, CA, USA) Ultra-Pure Water system. HPLC-grade acetonitrile was purchased from Baxter (Muskegon, MI, USA).

Preparation of tryptic fragments of the β-subunit of hCG

Purified β-subunit was prepared by incubation of hCG with 0.1% trifluoroacetic acid (TFA) at room temperature for 30 min followed by reversed-phase HPLC separation. The protein was eluted with a 5–50% acetonitrile (containing 0.1% TFA) gradient in 60 min at a flow rate of 1 ml min^{-1}.[23] The β-subunit peak was collected and dried down by a SpeedVac RT100 system (Savant, Farmingdale, NY, USA). The protein was denatured and reduced for 5 h in the presence of 6 mol l^{-1} Guanidine. HCl (GuHCl), 0.5 mol l^{-1} Tris, 2 mmol l^{-1} EDTA (pH 8.3) and dithiothreitol (DTT) followed by alkylation with either iodoacetate or 4-vinylpyridine overnight at 37 °C.[24,25] The reduced and alkylated β-subunit was separated from the reaction mixture by reversed-phase HPLC. The protein fraction was collected and dried. Purified reductively alkylated β-subunit was dissolved in ammonium hydrogencarbonate buffer (50 mM, pH 8.0) followed by the addition of TPCK-treated trypsin at an enzyme-to-substrate ratio of 1:50. The proteolysis reaction was incubated at 37 °C overnight. In some experiments and timed samples were taken during the incubation.

O-Deglycosylation of the reduced and alkylated β-subunit

About 0.5 mg of dry pyridylethylated β-subunit was dissolved in 200 µl of 20 mM sodium phosphatebuffer (pH 6.0) and 250 mIU of neuraminidase in 50 µl of the same buffer were added. The solution was then incubated at 37 °C for up to 20 h to remove sialic acids. In some experiments, timed aliquots were obtained. The protein was purified by HPLC, dried and subjected to trypsin proteolysis. In order to remove all of the oligosaccharide, the β-subunit was treated sequentially with neuraminidase and O-glycanase. A total of 5 mIU of O-glycanase was incubated with the desialylated protein in 20 mM sodium phosphate buffer (pH 6.0) at 37 °C for up to 50 h. In some experiments, timed aliquots were obtained. The protein was purified by reversed-phase HPLC and the dried material was proteolyzed with typsin as above.

Instrumentation

The semi-preparative HPLC system consisted of a Hewlett-Packard 1090L solvent-delivery system (Hewlett-Packard, Little Falls, DE, USA) equipped with a Spectra 100 UV detector (Spectra-Physics, San Jose, CA, USA). The detector was set at 215 nm. A 150 × 4.6 mm i.d. Vydac (Hesperia, CA, USA) C_{18} column (Cat. No. 218TP5415) was used for the protein and peptide purifications.

A PE-Sciex (Thornhill, Ontario, Canada) API-III$^{\mathrm{Plus}}$ quadrupole tandem mass spectrometer equipped with an articulated ionspray interface was used for the mass spectrometric analysis. Intact hCG, intact α- and β-subunits, carboxymethylated and pyridylethylated β-subunits (1 mg ml^{-1}) were directly infused into the mass spectrometer at a flow rate of 10 µl min^{-1} through fused-silica tubing (51 µm i.d.) by a Model 22 syringe pump (Harvard Apparatus, South Natick, MA). The ionization volatge was 4500 V and the orifice potential was set at 65 V. The curtain gas (N_2, 99.999%) flow rate was 1.2 l min^{-1}. The nebulizing air pressure was set at 40 psi. The scan range was from m/z 500 to 1800. The step size was 0.5 Da and the dwell time was 1 ms.

The HPLC portion of the HPLC/MS system consisted of a Beckman (Fullerton, CA, USA) Model 126 solvent-delivery system and Model 166 UV detector. A 5 µl aliquot a tryptic digest was injected through a Rheodyne (Cotati, CA, USA) Model 8125 injector into a 150 × 1 mm i.d. Deltabond C-18 column (Keystone Scientific, State College, PA, USA) and eluted with a linear gradient from 95:5 to 50:50 A–B in 60 min at a flow rate of 50 µl min^{-1}, where solvent A is 0.1% TFA in water and solvent B is 0.1% TFA in acetonitrile. A Valco 'Tee' was used post-column to split the effluent 1:10 with the majority of the effluent going through 127

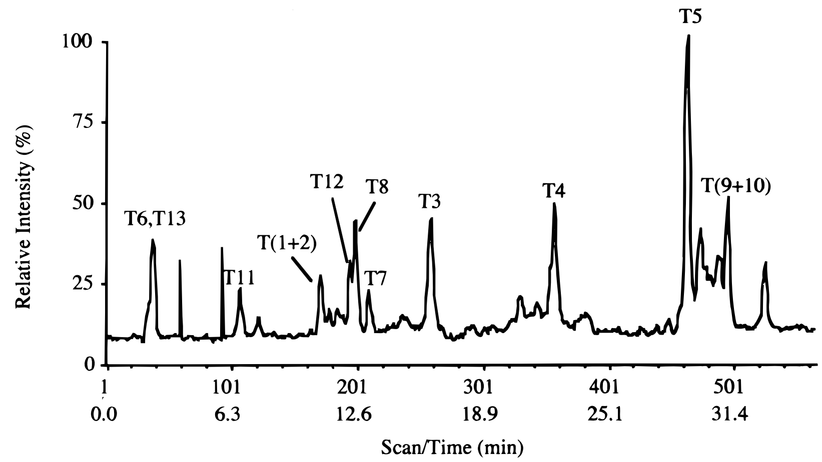

Figure 1. HPLC/MS of the tryptic digest of pyridylethylated β-chain of hCG. The tryptic fragments are indicated by the sequential number from the N-terminal end (βT); definitions can be found in Table 1.

μm i.d. polyether ether ketone (PEEK) tubing to the UV detector and the minor portion flowing through 50 μm i.d. fused-silica tubing to the nebulizing needle of the ionspray interface. About 100 pmol of protein digest from above was injected for each HPLC/MS analysis. Mass calibration was carried out with ammonium adducts of polypropylene glycol. The ionspray voltage was 4500 V with nebulizing zero grade air flowing at 0.6 l min^{-1} at a pressure of 40–50 psi. The curtain gas (N$_2$, 99.999%) flow rate was set at 1.2 l min^{-1}. The orifice potential was held at 120 V during scanning from m/z 150 to 370 and at 65 V during scanning from m/z 370 to 1900. The scan rate was 3.78 s per scan for a step size of 0.5 Da.

Tryptic fragment sequenses, their masses and their relative hydrophobicity and HPLC index were calculated using BioToolBox software (PE-Sciex).

RESULTS

Initial electrospray experiments using direct infusion of intact hCG, dissociated α- or β-subunits, or dissociated, reduced and carboxymethylated α- or β-subunits gave no detectable MS signal. Since the β-subunit sequence contains 12 cysteine residues, reductive alkylation with vinylpyridine could add 12 charge sites to the β-subunit, resulting in a theoretical mass to charge (m/z) ratio of 833 with a total of 30 charges. A large number of ions were observed from the pyridylethylated β-subunit, but the ion abundance was not useful analytically owing to the microheterogeneity of the oligosaccharides. After reduction and alkylation, the β-subunit can be proteolytically cleaved and the peptide fragments analyzed by HPLC/MS. Figure 1 shows an HPLC/MS trace of the tryptic digest of the reduced and pyridylethylated β-subunit. All of the expected tryptic peptide and the N-linked glycopeptide fragments (βT3, βT4; see Table 1) could be identified by their masses as calculated from the multiply charged ions in the spectra and their relative HPLC retentions (Fig. 1).

The spectrum (Fig. 2) of pyridylethylated tryptic fragment βT3 (β9–20), which contains an oligosaccharide at ^{13}Asn, could be identified by the 'collisional-excitation scanning' technique of Huddleston *et al.*[26] The carbohydrate fragment ions at m/z 366, 274 and 204 identified this chromatographic peak as arising from a glycopeptide. Each of the triply charged ions in the m/z 950–1430 region corresponded to a different oligosaccharide attached at ^{13}Asn. The theoretical masses of each of the reported oligosaccharide–peptide structures and the experimental results are in excellent agreement (Table

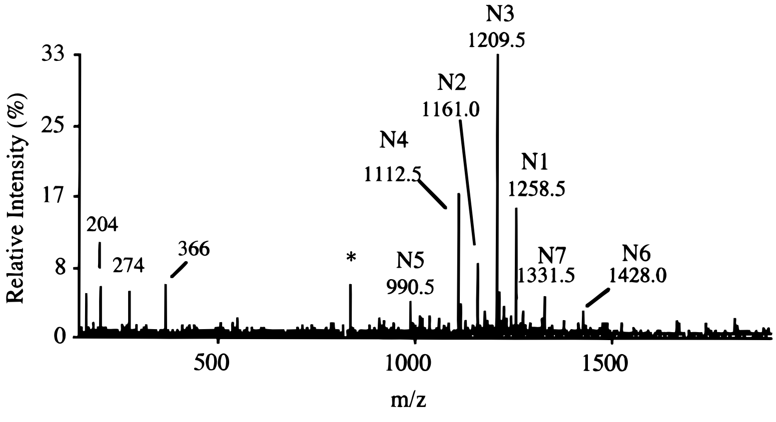

Figure 2. Mass spectrum of pyridylethylated βT3 fragment. Ions in the 200–375 m/z region are characteristic carbohydrate fragment ions. Predicted mass-to-charge ratios: βT3.N1^{3+}, 1257.6; βT3.N2^{3+}, 1160.6; βT3.N3^{3+}, 1208.9; βT3.N4^{3+}, 1111.9; βT3.N5^{3+}, 990.3; βT3.N6^{3+}, 1427.6; βT3.N7^{3+}, 1330.6. The structure of N1–N7 can be found in Table 2. The ion marked with an asterisk is from a co-eluting species.

Table 1. Tryptic fragments of reduced and pyridylethylated hCG β-subunit

Tryptic fragment	Amino acid residues	Peptide mass	Retention time (min)	Oligosaccharide attached[a]	Theoretical glycopeptide mass	Experimental mass
T6	61–63	389.2	2.3	—		389.5
T13	115–122	884.4	2.3	O1	1099.2	1100.0
				O3	771.2	771.5
				O5	808.2	808.5
				O6	625.7	626.0
T11	96–104	970.4	6.7	—		970.5[b]
T1, 2	1–8	982.6	10.7	—		982.5[b]
T8	69–74	747.4	12.3	—		747.5
T12	105–114	1273.6	12.4	—		1274.0
T7	64–68	651.4	13.1	—		651.5
T3	9–20	1418.8	16.2	N1	3769.8	3772.5[b]
				N2	3478.8	3479.5[b]
				N3	3623.7	3625.8[b]
				N4	3332.7	3334.3[b]
				N5	2967.9	2968.8[b]
				N6	4279.7	4281.0[b]
				N7	3988.7	3991.5[b]
T4	21–43	2852.3	22.4	N1	5203.2	5206.3[b]
				N2	4911.5	4915.5[b]
T5	44–60	1974.2	29.0	—		1975.6[b]
T9	75–94	2368.2	31.0	—		2368.8[b]
T14	123–133	1104.6	—[c]	O6	1469.6	1470.0
T15	134–145	1233.7	—[c]	O5	1963.8	1965.0

[a] The structure of the *N*- and *O*-linked oligosaccharides can be found in Tables 2 and 3, respectively.

[b] Mean mass calculated from ions from two or three charge states.

[c] The carboxy-terminal fragments were only observed after treatment with neuraminidase and *O*-glycanase.

1). In addition to the expected ions which correspond to peptide βT3 attached to oligosaccharide structures N1 through N5 (Table 2), there were two triply charged ions at m/z 1331.5 and 1428.0 whose molecular masses were consistent with two proposed triantennary oligosaccharides, N6 and N7.[27]

Alkylation with vinylpyridine (VP) has significant advantages over iodoacetate for MS analysis of cysteine-rich glycopeptides.[28] A good example was the detection of βT4,which has an oligosaccharide attached at ^{30}Asn. The only charge sites on the carboxymethylated glycopeptide were the N-terminal glutamate (^{21}Glu) and the ε-amino group on the C-terminal lysine (^{43}Lys). A doubly charged ion would have an m/z value slightly below the mass detection limit of our instrument (2400 Da). In fact, no peak with a mass corresponding to the βT4 fragment was found in the chromatogram after carboxymethylation. Pyridylethylation of the four cysteine residues in this peptide would increase the predicted maximum charge state of this glycopeptide to 6^+, which could reduce the *m*/*z* to well within the mass range. The mass spectrum (Fig. 3) of βT4 after VP treatment clearly showed the multiply charged (3^+, 4^+ and 5^+) molecular ions.

Microheterogeneity differences between the two N-linked carbohydrates could be demonstrated by comparing the mass spectra of βT3 and βT4. For ^{13}Asn, seven oligosaccharide species were detected with N3 being the main component. For the oligosaccharide attached to ^{30}Asn, only two constituents, N1 and N2, were observed. N1 was the main component with about 25% N2 as the minor component (Table 2), assuming equivalent ionization efficiencies of the two species. No triantennary oligosaccharide was observed on βT4.

In order to determine if HPLC/MS could detect carbohydrate microheterogeneity differences on various hCG preparations, samples from Sigma, Organon and NIDDK were analyzed. For the Sigma and Organon hCG preparations, the microheterogeneity pattern of both ^{13}Asn and ^{30}Asn was similar (Fig. 4). For ^{13}Asn, the peak intensity ratio of N2 (*m*/*z* 1161.0) to N4 (*m*/*z* 1112.5) was 0.42 and 0.52 for the Sigma and Organon preparations, respectively. Five replicates using the same lot of Sigma hCG run on different days gave a ratio ranging from 0.39 to 0.50. NIDDK hCG showed a different pattern. The peak intensity ratio of the ion due to N2 to that of N4 was 1.99. A similar difference was observed for ^{30}Asn (data not shown). The peak intensity ratio of the ion due to N1 (*m*/*z* 1302.5) to that due to N2 (*m*/*z* 1230) was 2.6 and 6.7 for the Sigma and Organon preparations, respectively, while that for NIDDK was 13.7.

Detection of the three C-terminal glycopeptides, βT13, βT14 and βT15, presented a more difficult problem. Pollak *et al.*[29] assigned a broad peak in the region of 80–90 min in their reversed-phase HPLC trace of the tryptic digest of β-hCG as a partial digest of the C-terminal portion, βT (13 + 14 + 15) and βT (14 + 15). Our MS data supported their conclusion since carbohydrate fragment ions were observed in the corresponding elution region from our tryptic digest. The molecular ion portions of the spectra were too complex to be interpreted. In order to confirm the structure of the C-terminal fragment, partial or com-

Table 2. Published structures of oligosaccharides attached to β-hCG

N1	Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 Fucα1 ⟍6 |6 Manβ1–4GlcNacβ1–4GlcNAc ⟋3 Neu5Acα2-3Galβ1–4GlcNAcβ1–2Manα1
N2	Fucα1 Galβ1–4GlcNAcβ1–2Manα1 | ⟍6 6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1
N3	Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 ⟍6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1
N4	Galβ1–4GlcNAcβ1–2Manα1 ⟍6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAβ1–2Manα1
N5	Manα1 ⟍6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1
N6	Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 ⟍6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 4 (or 6) Neu5Aca2–3Galβ1–4GlcNAcβ1⟋
N7	Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 ⟍6 Manβ1–4GlcNAcβ1–4GlcNAc ⟋3 Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1 4 (or 6) Galβ1–4GlcNAcβ1⟋

plete removal of the O-linked carbohydrates was performed both to increase its accessibility to proteolytic attack and to reduce its microheterogeneity.

Careful examination of the first peak in the chromatogram (Fig. 1) revealed ions corresponding to βT13 peptide [$(M + H)^+$, m/z 885.5] and βT13 with oligosaccharides corresponding to O1 [$(M + H)^+$, m/z 1100.0], O3 [$(M + 2H)^{2+}$, m/z 771.5] and O6 [$(M + 2H)^{2+}$, m/z 626.5] attached [Fig. 5(A)]. Co-elution of βT13 and βT6 might be expected since βT13 is a hydrophilic glycopeptide. The microheterogeneity of the oligosaccharides attached to ^{121}Ser is apparent. It was interesting to note that some of the βT13 peptide was apparently not glycosylated. This observation has not been reported previously. Following treatment with neuraminidase, ions corresponding to βT13 peptide, βT13 + O5 (m/z 808.5) and βT13 + O6 (m/z 626.5) were found in the mass spectrum [Fig. 5(B)]. This was consistent with the loss of sialic acid from O1 and O3, respectively (see Table 3). After sequential treatment with neuraminidase and O-

Table 3. Reported structures of *O*-linked oligosaccharides

Label number	Structure
O1	NeuAcα2–3Galβ1–3GalNac— |β1–6 NeuAcα2–3Galβ1–4GlcNAc
O2	NeuAcα2–3Galβ1–3GalNAc— |α2–6 NeuAc
O3	NeuAcα2–3Galβ1–3GalNAc—
O4	NeuAcα2–6GalNAc—
O5	Galβ1–3GalNAc— |β1–6 Galβ1–4GlcNAc
O6	Galβ1–3GalNAc—

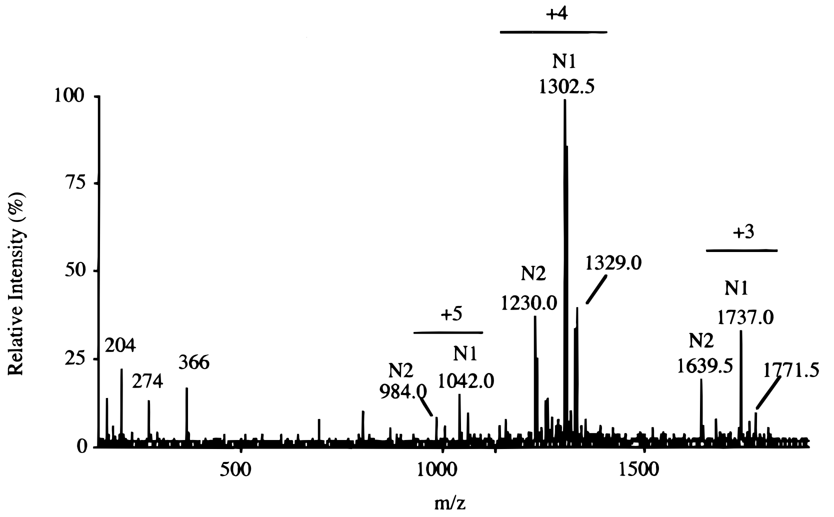

Figure 3. Mass spectrum of pyridylethylated βT4 fragment. Ions in the 200–375 *m/z* region are characteristic carbohydrate fragment ions. Predicted mass-to-charge ratios: βT4 · N1^{5+}, 1041.6; βT4 · N1^{4+}, 1301.8; βT4 · N1^{3+}, 1735.4; βT4 · N2^{5+}, 983.3; βT4 · N2^{4+}, 1228.8; βT4 · N2^{3+}, 1638.1. The structure of N1 and N2 can be found in Table 2. Ions at *m/z* 1329.0 and 1771.5 correspond to the triply and doubly charged ions from an additional pyridylethyl group on alkylated βT4 + N1.

glycanase, βT13 peptide was the main component although a small quantity of βT13 + O5 was still present [Fig. 5(C)]. A time-course study showed that O6 was removed within 4 h of O-glycanase treatment, whereas some O5-containing species remained after 50 h. Genzyme's O-glycanase is reported to cleave specifically the linkage between Gal-GalNAc- and Ser or Thr and substitution on the proximal carbohydrate moiety greatly reduces the rate of enzymatic cleavage.[30] Our data are consistent with a reduction in the rate of enzymatic cleavage due to the presence of a branch Gal at the proximal GalNAc in O5.

Before neuraminidase treatment, no signal corresponding to βT14 or βT15 was found in the HPLC/MS trace of the tryptic digest. After the neuraminidase treatment, an ion at *m/z* 736 could correspond to O6 attached to βT14 [Fig. 6(A)]. This assignment needed further confirmation because two glycosylation sites, ^{127}Ser and ^{132}Ser, have been reported for this peptide sequence.[2] After sequential neuraminidase and O-glycanase (4 h) digestion, the ion at *m/z* 736 was completely replaced by an ion at *m/z* 553.5 [Fig. 6(B)]. The disappearance rate of the *m/z* 736 ion matched the disappearance rate of O6 in βT13 and the newly formed ion at *m/z* 553.5 corresponded to the doubly charged molecular ion of βT14 peptide. It was interesting that we were unable to detect any ions corresponding to βT14 with two carbohydrates attached in any of our

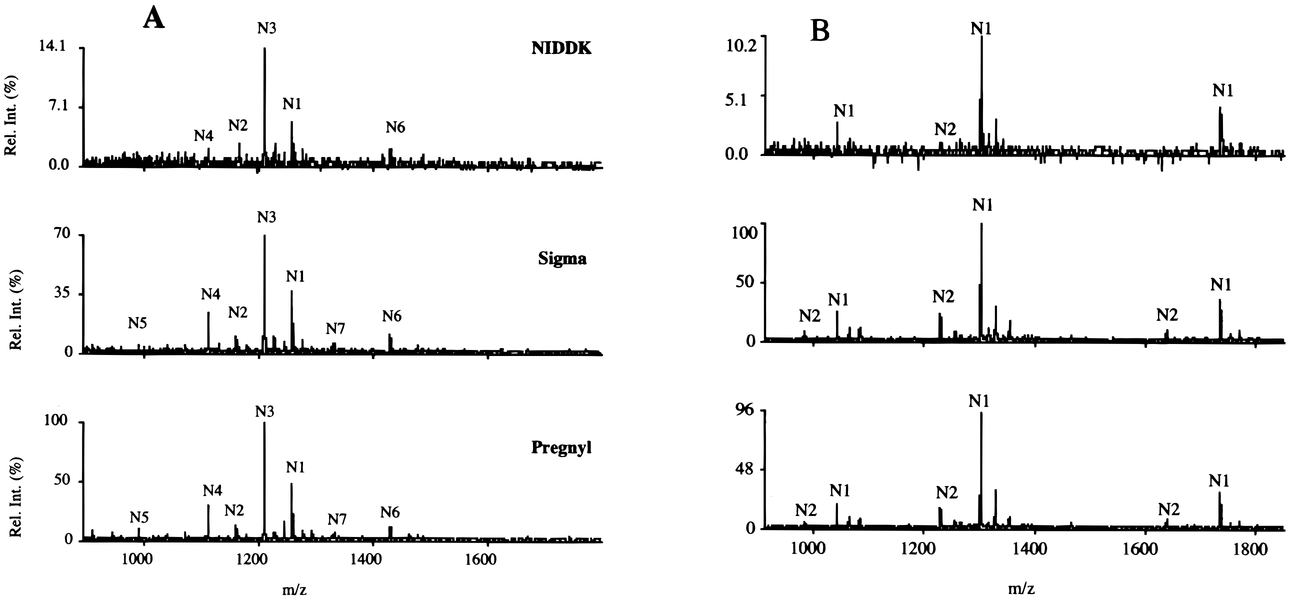

Figure 4. Mass spectral evidence for microheterogeneity on βT3 and βT4 shown from three different sources of hCG. (A) Expanded view of the spectrum of βT3; (B) expanded view of the spectrum of βT4. Upper panel, βT3 and βT4 from NIDDK hCG; middle panel, βT3 and βT4 from Sigma hCG; bottom panel, βT3 and βT4 from Organon hCG.

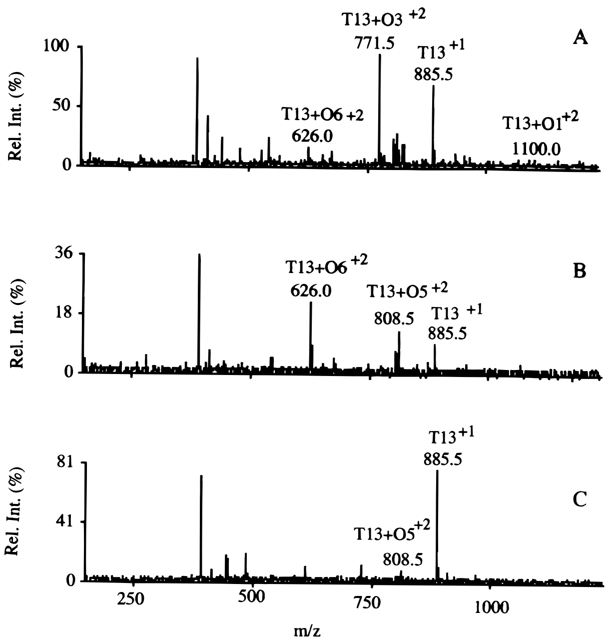

Figure 5. Mass spectra of βT13 after selective carbohydrate removal. (A) Native βT-13; (B) after treatment with neuraminidase; (C) after sequential treatment with neuraminidase and O-glycanse. The structure of O1–O6 can be found in Table 3. The ions in the low-*m*/*z* region came from other co-eluting species from the HPLC peak.

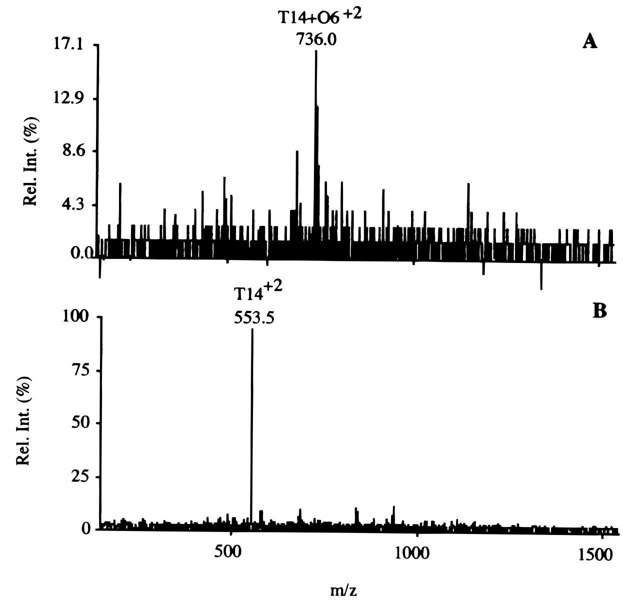

Figure 6. Mass spectra of βT14 after selective carbohydrate removal. (A) after neuraminidase treatment; (B) after seqential treatment with neuraminidase and O-glycanase. The structure of O6 can be found in Table 3.

experiments. Based on the predicted masses for both peptide (MH^+ *m*/*z* 1105.6) and O-linked carbohydrate (maximum molecular mass 748), the doubly charged molecular ion of βT14 with two desialylated carbohydrates attached should have a mass to charge ratio $[(M + 2H)^{2+}$, maximum *m*/*z* 1300.8] well within the mass range of the mass spectrometer.

Even after 20 h of neuraminidase digestion, no ions corresponding to βT15 or its glycosylated forms were observed. The rationale for this is not clear. After neuraminidase treatment and 4 h of O-glycanase digestion, ions corresponding to βT15 peptide $[(M + H)^+$, *m*/*z* 1235.0; $(M + 2H)^{2+}$, *m*/*z* 618.0] and βT15 + O5 $[(M + 2H)^{2+}$, *m*/*z* 983.5) were found (Fig. 7). The abundance of the ion at *m*/*z* 983.5 did not decrease rapidly with O-glycanase digestion, consistent with our other data regarding enzymic cleavage at a branched attachment point.

βT3 and βT4 were used as references to check the completeness of the neuraminidase and O-glycanase digestion reactions. After 20 h of neuraminidase digestion, the absence of sialic acids in the N-linked carbohydrates in βT3 and βT4 (Fig. 8) indicated complete sialic acid removal. For ^{13}Asn, complete removal of sialic acid from N6 resulted in N6–3Neu. Also note that complete removal of sialic acid from N1 and N2 or N3 and N4 resulted in the same product. For ^{30}Asn, N1

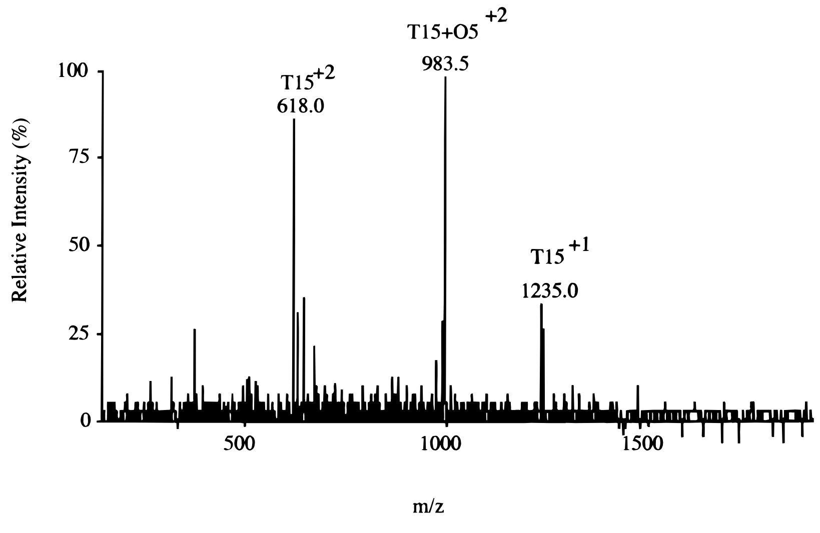

Figure 7. Mass spectra of βT15 after selective carbohydrate removal with sequential neuraminidase and O-glycanase treatment. The structure of O5 can be found in Table 3.

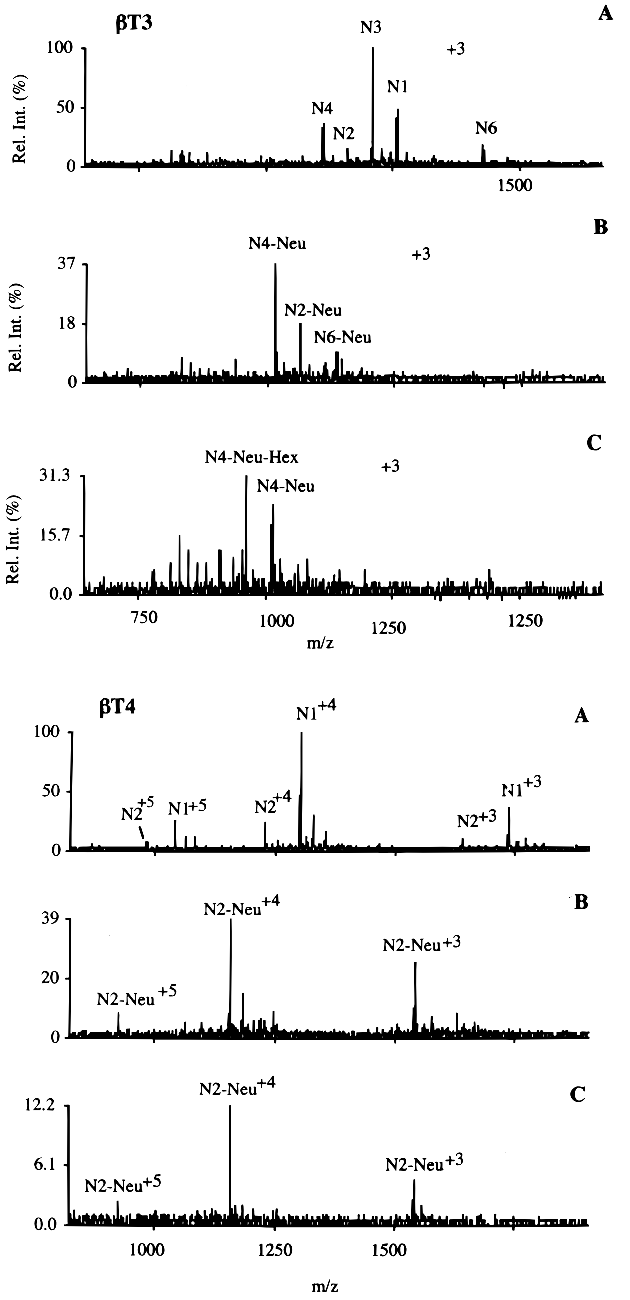

Figure 8. Expanded view of the spectrum of βT3 and βT4 showing the structure change of N-linked carbohydrates after neuraminidase and O-glycanase treatment. (A) Native βT3 and βT4; (B) after neuraminidase treatment; (C) after sequential treatment with neuraminidase and O-glycanase. The structure of labeled carbohydrates can be found in Table 2. Note that complete removal of sialic acid from N1 and N2, or N3 and N4 or N6 and N7 results in identical species.

and N2 changed to N2-Neu after the sialic acid removal. It was interesting that after 50 h of O-glycanase hydrolysis, N2-Neu and N6–3Neu at ^{13}Asn disappeared. A new ion which corresponded to the loss of a galatose from N4-Neu (N4-Neu-Gal) was observed. This indicated that some of the galactoses and all of the fucoses were removed from the oligosaccharide at ^{13}Asn either by the enzyme or by bacteria present in the incubation mixture.

DISCUSSION

Mass spectrometry has played an increasing role in characterization of proteins and glycopeptides. The development of the electrospray interface[31] has allowed efficient ionization of peptides and proteins. Recently, mass spectrometry has proven to be a promising analytical technique for glycopeptide analysis.[23,32] The molecular mass information obtained allows the sensitive detection of both peptide and glycopeptide fragements. HPLC/MS combined with different enzymatic reactions has been used successfully for the characterization of a variety of natural and recombinant glycoproteins.[23,32] In this study, we report the use of mass spectrometry to characterize the β-subunit of hCG. This should provide a foundation for the development of a forensically robust method for hCG detection and for rapid characterization of candidate hCG protein standard preparations.

In initial studies we were unable to observe a signal from either intact hCG or denatured α and β chains directly infused into the mass spectrometer although the predicted mass-to-charge ratio of each of these glycopeptides was well within the mass range of the quadrupole under electrospray conditions. Two possible explanations exist. Since the intact molecule was tightly folded, it was possible that some of the basic amino acids which should be protonated were buried inside the molecule and therefore inaccessible. Therefore, the mass-to-charge ratio would exceed the mass range of the instrument (2400 *m/z*). The microheterogeneity of the carbohydrates could also explain the low signal intensity. As reported by Birken et al.,[2] there are at least five different molecular structures for Asn-linked carbohydrate and two for Ser-linked carbohydrate. Intact hCG has four asparagine-linked carbohydrates and four serine-linked carbohydrates. This carbohydrate 'microheterogeneity' resulting from the many possible combinations would give rise to many molecular ion signals from the same peptide sequence and thus poor detectability. Interestingly, similar difficulty in detection was described for MALDI-TOF analysis of hCG.[33]

Since we were unable to detect native protein or subunit, we focused our attention on the analysis of the β-subunit. The α- and β-subunits of hCG can be readily dissociated with TFA.[23] Experiments using a non-denaturing HPLC mobile phase indicated that TFA dissociation of the chains is complete in less than 5 min. After reductive alkylation with either iodoacetate or vinylpyridine, a small increase in the reversed-phase HPLC retention time was observed compared with that of intact β-subunit, reflecting the increased hydrophobicity. The reduced and alkylated β-subunit was collected for further digestion. We had no difficulty in detecting and identifying picomole quantitites of tryptic peptide fragments (βT1, βT2, βT5-βT12) from hCG based on the pattern of multiply charged ions in each chromatographic peak. Detection of glycopeptide peaks was complicated by the fact that numerous ions arise from the microheterogeneity of the oligosaccharide.

As demonstrated by Huddleston *et al.*,[26] collisional fragmentation can be induced in the interface region resulting in characteristic carbohydrate fragment ions.

At high orifice potentials, Hex-HexNAc$^+$ (*m/z* 366), NeuAc-H_2O^+ (*m/z* 274) and HexNAc$^+$ (*m/z* 204) were observed, where Hex indicates a hexose (glucose, galactose or mannose), HexNAc indicates an N-acetylhexosamine (GlcNAc or GalNAc) and NeuAc indicates N-acetylneuraminic acid. Each oligosaccharide attached to the peptide will give a specific molecular ion in the spectrum. By programming the orifice potential during the elution of a chromatographic peak, oligosaccharide-containing peptides can be identified by the pattern of ions below 370 Da while higher mass ions provide information about the intact oligosacchardies attached to the peptide. Although the structure of the oligosaccharides cannot be deduced from this information, the mass of the oligosaccharides can be obtained. The observed masses of these glycopeptides were consistent with the oligosaccharide structures reported by Birken *et al.*[2] We were also able to detect the presence of triantennary carbohydrate structures, as proposed by Kamerling *et al.*,[27] in several commercial hCG preparations

Significant differences were found between our data on oligosaccharide microheterogeneity and those of an NMR study.[34] For the oligosaccharide attached to ^{13}Asn, Weisshaar *et al.*[34] reported N1 to be the main component with N5 being the only minor component ($<7\%$). We observed the presence of seven oligosaccharides (N1 to N7) with N3 being the main component. For the carbohydrate attached to ^{30}Asn, both techniques suggest that the main oligosaccharide was N1. Our results showed the presence of about 7% N2 as the minor component while NMR indicated the presence of $<3\%$ N5. These quantitative discrepancies could be due to the difference in hCG preparation. The relatively low sensitivity of NMR compared with MS could be another reason for the observed difference. NMR may not be able to detect some minor components present in the mixture. On the other hand, the quantitative nature of the HPLC/MS data assumes equivalent ion production for all of the glycopeptide species. We were also able to distinguish between various preparations of hCG, illustrating the semiquantitative nature of the HPLC/MS technique. The similarity of the Organon and Sigma preparations was expected, since Organon is the source for most of the commercially available hCG preparations. Based on our findings, we conclude that Sigma does relatively little purification of the Organon material. The NIDDK preparation, on the other hand, was purified using anion-exchange and gel permeation chromatography.[2] During the anion-exchange chromatography process, the oligosaccharides which had high negative charges were likely to be enriched and recovered at a relatively high yield. It is not surprising, therefore, that protein species with oligosaccharide chains with higher negative charge (e.g. rich in N-acetylneuraminic acid) would be enriched.

The systematic difference in oligosaccharide heterogeneity between ^{13}Asn and ^{30}Asn was unexpected. Based on the crystallographic studies of Lapthorn *et al.*,[7] the oligosaccharides at ^{13}Asn and ^{30}Asn are spatially close. From these data, it seems unlikely that these differences could reflect enzymatic attack. Nevertheless, O-glycanase was able to cleave some residues on the oligosaccharide attached to ^{13}Asn but not ^{30}Asn, despite the fact that the oligosaccharides had similar structures. This observation suggests that ^{13}Asn is in a more accessible position in solution, which in turn could explain the fact that the carbohydrate at ^{13}Asn is more heterogeneous than that at ^{30}Asn.

Although alkylation with vinylpyridine has advantages over iodoacetate for MS analysis of cysteine-rich glycopeptides, it is not without some anomalies. The low-intensity peaks with a mass difference of 105 from adjacent molecular ions (*m/z* 1329.0 and 1771.5; Fig. 3) appeared to result from the addition of an extra VP to this peptide. These peaks were observed only for selected peptides and with long reaction times. This adduct formation could affect the quantitative analysis by mass spectrometry. Quantitation based on a certain ion could also be affected if the adduct formation is not reproducible. Subsequent work suggests that this mass addition came from either adduct formation (non-covalent association) of vinylpyridine with the peptide or formation of a labile bond, since small changes in orifice voltage or low-energy collision-induced dissociation result in disappearance of the ion (unpublished results).

Confirmation of the reported carboxy-terminal structure of hCG presented the most difficulties. Both free and glycosylated βT13 were observed in the mass spectrum, which has not been reported previously. Based on the orifice voltage required for collisional-excitation scanning, it seems unlikely that the presence of free peptide is an artifact of the method. Neither βT14 nor βT15 could be observed without enzymatic treatment to remove portions of the carbohydrate. The absence of these peptides also occurred for MALDI-TOF,[33] despite the reported superiority of this technique for glycopeptide ionization. Even after removal of the neuraminic acid moieties, we were unable to confirm the presence of two oligosaccharides attached to βT14. Both Birken and Canfield[14] and Kessler *et al.*[15] reported data from radioactive labeling through β-elimination that suggest two equimolar carbohydrate attachment sites at ^{127}Ser and ^{132}Ser on βT14. It is difficult to discount the radioactive labeling data, but a possible explanation consistent with both the MS and classical data is that half of the molecules present are labeled at one or the other site but not both. Work is under way to investigate this hypothesis further. It is also possible that the discrepancy arises from the fact that different hCG preparations were used. Kessler *et al.*[15] concluded that the glycan structure on all of the serine residues is the same, based on similar carbohydrate quantities cleaved from various peptide fractions. This is not supported by our work, which shows significant heterogeneity at the O-linked sites.

CONCLUSIONS

We have been able to characterize tryptic fragments from the β-subunit of hCG using pneumatically assisted electrospray HPLC/MS. We were able to confirm both the primary sequence and the attachment of five oligosaccharide groups. The microheterogeneity of the carbohydrate moieties at ^{13}Asn and ^{30}Asn can be readily observed from the mass spectra of tryptic fragments βT3 and βT4. There is more heterogeneity in the oligosaccharide at ^{13}Asn than that at ^{30}Asn. Three preparations of hCG could be distinguished on the basis of their relative oligosaccharide content. Glycopeptide fragments from the carboxy-terminal sequence could only be observed after neuraminidase or sequential neuraminidase and O-glycanase digestion. Information on individual O-linked glycan heterogeneity was obtained for ^{121}Ser and ^{138}Ser. The glycosylation at both ^{127}Ser and ^{132}Ser could not be confirmed, although the microheterogeneity at the glycosylation site could be documented. Picomole quantities of tryptic fragments could be detected in the scan mode. Several tryptic fragments appear to be candidates for selective mass detection. This supports our hypothesis that this technique could be used to quantify and confirm the presence of trace amounts of hCG in urine.[20]

Acknowledgements

We gratefully acknowledge the donation of a sample of β-hCG from the National Hormone and Pituitary Program, the National Institutes of Diabetes and Digestive and Kidney Diseases, the Nathional Institute of Child Health and Development and the US Department of Agriculture. Portions of this work were presented at the Cologne Workshop on Doping in Sport (April 1994) and the American Society for Mass Spectrometry meeting (May 1995).

REFERENCES

1. J. G. Pierce and T. F. Parsons, *Annu. Rev. Biochem.* **50**, 465 (1981).
2. S. Birken, A. Krichevsky, J. O'Connor, J. Lustbader and R. Canfield, in *Glycoprotein Hormones*, edited by W. W. Chin and I. Boime, p 45. Serono Symposia, Norwell, USA (1990).
3. R. O. Hussa, *The Clinical Marker hCG*. Praeger, New York (1987).
4. A. T. Kickman and D. A. Cowan, *Br. Med. Bull.* **48**, 496 (1992).
5. S. Dirnhofer, S. Madersbacher, J. M. Bidart, P. B. Ten Kortenaar, G. Spottl, K. Mann, G. Wick and P. Berger, *J. Endocrinol.* **141**, 153 (1994).
6. R. Hoermann, P. Berger, G. Spoetti, F. Gillesberger, A. Kardana, L. A. Cole and K. Mann, *Clin. Chem.* **40**, 2306 (1994).
7. A. J. Lapthorn, D. C. Harris, A. Littlejohn, J. W. Lustbader, R. E. Canfield, K. J. Machin, F. J. Morgan and N. W. Isaacs, *Nature (London)* **369**, 455 (1994).
8. T. Endo, R. Nishimura, S. Saito, K. Kanazawa, K. Nomura, M. Katsuno, K. Shii, K. Mukhopadhyay, S. Baba and A. Kobata, *Endocrinology* **130**, 2052 (1992).
9. S. Birken, E. G. Armstrong, M. A. G. Kolks, L. A. Cole, G. M. Agosto, A. Krichevsky, J. L. Vaitukaitis and R. E. Canfield, *Endocrinology* **123**, 572 (1988).
10. D. L. Blithe, A. H. Akar, R. E. Wehmann and B. C. Nisula, *Endocrinology* **122**, 173 (1988).
11. S. Birken, M. A. G. Kolks, S. Amr, B. Nisula and D. Puett, *Endocrinology* **121**, 657 (1987).
12. Y. Kato and G. D. Braunstein, *J. Clin. Endocrinol. Metab.* **66**, 1197 (1988).
13. M. J. Kessler, M. S. Reddy, R. H. Shah and O. P. Bahl, *J. Biol. Chem.* **254**, 7901 (1979).
14. S. Birken and R. E. Canfield, *J. Biol. Chem.* **252**, 5386 (1977).
15. M. J. Kessler, M. Takashi, D. G. Rajendra and O. P. Bahl, *J. Biol. Chem.* **254**, 7909 (1979).
16. L. A. Cole, A. Kardana, F. C. Ying and S. Birken, *Yale J. Biol. Med.* **64**, 627 (1991).
17. P. Laidler, D. A. Cowan, R. C. Hides and A. T. Kickman, *Clin. Chem.* **40**, 1306 (1994).
18. M. Yamashita and J. B. Fenn, *J. Phys. Chem.* **88**, 4451 (1984).
19. P. Kebarle and L. Tang, *Anal. Chem.* **65**, 972A (1993).
20. D. C. Gale and R. D. Smith, *Rapid Commun. Mass Spectrom.* **7**, 1017 (1993).
21. C. Liu and L. D. Bowers, *J. Chromatogr. B* in press (1996).
22. C. Liu and L. D. Bowers, in *Recent Advances in Doping Analysis: Proceedings of the 12th Cologne Workshop on Dope Analysis*, edited by M. Donike, H. Geyer, A. Gotzmann and U. Mareck-Engelke, pp. 235–242. Sport und Buch Strauss, Cologne (1995).
23. T. F. Parsons, T. W. Strickland and J. G. Pierce, *Endocrinology* **114**, 2223 (1984).
24. S. Carr and G. Roberts, *Anal. Biochem.* **157**, 396 (1986).
25. G. E. Tarr, in *Methods of Protein Microcharacterization*, edited by J. E. Shively, p. 155. Humana Press, Clifton, NJ (1986).
26. M. J. Huddleston, M. F. Bean and S. A. Carr, *Anal. Chem.* **65**, 877 (1993).
27. J. P. Kamerling, J. B. L. Damm, K. Hard, G. W. K. Van Dedem, W. D. Boer and J. F. G. Vliegenthart, in *Glycoprotein Hormones*, edited by W. W. Chin and I. Boime, p. 123. Serono Symposia, Norwell, USA (1990).
28. Z. Lam. B. Reinhold and V. Reinhold, in *Proceedings of the 39th ASMS Conference on Mass Spectrometry and Allied Topics*, Nashville, TN, p. 282 (1991).
29. S. Pollak, S. Halpine, B. T. Chait and S. Birken, *Endocrinology* **126**, 199 (1990).
30. J. Umemoto, V. P. Bhavanandan and E. A. Davidson, *J. Biol. Chem.* **252**, 8609 (1977).
31. A. P. Bruins, T. R. Covey and J. D. Henion, *Anal. Chem.* **59**, 2642 (1987).
32. D. A. Lewis, A. W. Guzzetta and W. S. Hancock, *Anal. Chem.* **66**, 585 (1994).
33. P. Laidler, D. A. Cowan, R. C. Hider, A. Keane and A. T. Kicman, *Rapid Commun. Mass Spectrom.* **9**, 1021 (1995).
34. G. Weisshaar, J. Hiyama and A. G. C. Renwick, *Glycobiology* **1**, 393, (1991).