Identification of O-Glycosylation Sites and Partial Characterization of Carbohydrate Structure and Disulfide Linkages of Human Insulin-like Growth Factor Binding Protein 6[†]

Gregory M. Neumann,[‡] Joe A. Marinaro,[§] and Leon A. Bach*[,§]

School of Biochemistry, La Trobe University, Bundoora, Victoria 3083, Australia, and Department of Medicine, Austin & Repatriation Medical Centre (Austin Campus), University of Melbourne, Heidelberg, Victoria 3084, Australia

Received November 25, 1997; Revised Manuscript Received January 29, 1998

ABSTRACT: The actions of insulin-like growth factors (IGFs) are modulated by a family of high-affinity binding proteins (IGFBPs), including IGFBP-6, which preferentially binds IGF-II and is O-glycosylated. Glycosylated and nonglycosylated recombinant human IGFBP-6, expressed in Chinese hamster ovary cells and *Escherichia coli*, respectively, were purified using IGF-II affinity chromatography and reverse-phase medium-pressure chromatography. Electrospray ionization mass spectrometry (ESMS) of glycosylated IGFBP-6 revealed considerable heterogeneity of carbohydrate composition. Major glycoforms contained 8−16 monosaccharides, including *N*-acetylhexosamine, hexose, and *N*-acetylneuraminic acid. Glycosylation sites of IGFBP-6 were identified as Thr^{126}, Ser^{144}, Thr^{145}, Thr^{146}, and Ser^{152} by using a combination of ESMS and Edman sequencing of tryptic fragments separated by reverse-phase high-pressure liquid chromatography. One oligosaccharide chain contained 5−6 monosaccharides, whereas the others contained 2−4 monosaccharides. Glycosylated IGFBP-6 exhibited greater resistance to proteolysis by chymotrypsin and trypsin than nonglycosylated IGFBP-6. Native disulfide bond positions in IGFBP-6 were localized by means of observed disulfide-linked tryptic fragments, revealing that there are two disulfide-linked subdomains within each of the N- and C-terminal regions and confirming a previous suggestion that the latter regions are not interconnected. A model of IGFBP-6 is developed in which these distinct domains are separated by a central region which is O-glycosylated.

Insulin-like growth factors (IGFs)[1] are important mediators of pre- and postnatal physiological growth (*1*). IGFs are insulin-like in that they are structurally homologous to proinsulin and share metabolic properties with insulin. In addition, IGFs promote mitogenesis and differentiation of a large number of cell types.

IGF actions are modulated by a family of high-affinity IGF binding proteins (IGFBPs), one of which (IGFBP-6) is the focus of this present study. IGFBPs 1−6 have highly conserved sequences in the N- and C-terminal regions and have been extensively characterized (*2*−*4*). IGFBPs 1−5 have 10 and 8 conserved cysteines in their N- and C-terminal domains, respectively. The central regions that connect the conserved N- and C-terminal domains of the IGFBPs are not homologous with each other. Recently, further putative IGFBPs with homology to the N-terminal domains of the six IGFBPs have been described (*5*).

IGFBP-6 differs from the other IGFBPs in a number of important respects. Human and rat IGFBP-6 respectively lack 2 and 4 of the N-terminal cysteines that are conserved in IGFBPs 1−5 (*6*). IGFBP-6 binds IGF-II with the highest affinity of the IGFBPs, and unlike IGFBPs 1−5, IGFBP-6 binds IGF-I with 20−100-fold lower affinity (*7*−*10*). IGFBP-6 is extensively O-glycosylated (*11*), whereas IGFBP-3 (*12*) and IGFBP-4 (*13*) are N-glycosylated, and IGFBP-5 is O-glycosylated to a limited extent (*14*). Human, but not rat, IGFBP-6 has a potential site for N-glycosylation (*6*), but this site is not glycosylated (*11*). Enzymatic deglycosylation of IGFBP-6 has no significant effect on IGF binding (*11*). Proposed roles other than ligand binding for the carbohydrate component of glycoproteins include facilitation of cellular processing, protection from proteolysis, modulation of binding to structural biomolecules including membrane and extracellular matrix components, and retardation of in vivo clearance (*15*, *16*).

N-Glycosylation occurs only on an asparagine that is part of the consensus sequence Asn-X-Ser/Thr, where X is any amino acid other than proline. In contrast, there is no consensus amino acid sequence that determines whether an individual serine or threonine residue will be O-glycosylated (*17*). Prediction algorithms based on empirical observations have been devised (*18*−*21*) but are of limited reliability when

[†] This work was supported by grants from the National Health and Medical Research Council of Australia and the Austin Hospital Medical Research Foundation.

* Address correspondence to this author. Tel: 61 3 9496 3581. Fax: 61 3 9457 5485. E-mail: bach@austin.unimelb.edu.au.

[‡] La Trobe University.

[§] University of Melbourne.

[1] Abbreviations: IGF, insulin-like growth factor; IGFBP, insulin-like growth factor binding protein; ESMS, electrospray ionization mass spectrometry; HPAEC, high-performance anion-exchange chromatography; HPLC, high-performance liquid chromatography; V_{OR}, mass spectrometer ion source orifice voltage; CHO, Chinese hamster ovary; MEM, minimal essential medium; WGA, wheat germ agglutinin; TFA, trifluoroacetic acid; HexNAc, *N*-acetylhexosamine; Hex, hexose; NeuAc, *N*-acetylneuraminic acid.

S0006-2960(97)02894-8 CCC: $15.00
Published on Web 04/16/1998

applied to specific glycoproteins. The aim of the present study was therefore to determine experimentally the glycosylation sites of human IGFBP-6 and to further characterize the carbohydrate composition of the O-linked oligosaccharide chains. To this end, we employed a combination of electrospray ionization mass spectrometry (ESMS) (*22*) and Edman N-terminal amino acid sequencing of intact glycoprotein and tryptic fragments purified by microbore high-performance liquid chromatography (HPLC). Similar approaches have previously enabled rapid structural characterization of proteins (*23*−*25*) and identification of posttranslational modifications including glycosylation (*26*, *27*). Additional information regarding IGFBP-6 disulfide bond linkages was obtained from an analysis of disulfide-linked tryptic peptides, leading to clarification of the domain structure of IGFBP-6. Finally, phosphorylation is another posttranslational modification which may be important for protein function, including that of IGFBP-1, which has increased binding affinity for IGFs when phosphorylated (*28*). Metabolic labeling and ESMS were therefore used to study the presence of this posttranslational modification of IGFBP-6.

EXPERIMENTAL PROCEDURES

Expression of Recombinant Glycosylated Human IGFBP-6 in Chinese Hamster Ovary (CHO) Cells. CHO cells were grown to 50% confluence in α-MEM/10% fetal calf serum (growth medium). phBP6-E3, the eukaryotic expression vector encoding human IGFBP-6 (1 μg) (*29*), was mixed with 6 μL of lipofectamine (Gibco-BRL, Mt. Waverley, Australia) for 30 min at room temperature. CHO cells were then incubated for 6 h with the plasmid/lipofectamine solution diluted to 1 mL with serum-free medium. After incubation in 2× growth medium for 16 h at 37 °C, geneticin-resistant clones were selected by incubation for 14 days in growth medium containing 400 μg/mL geneticin (Gibco-BRL). A number of geneticin-resistant colonies were expanded, and colonies producing the highest levels of IGFBP-6 [as determined by charcoal adsorption assay (*30*) and Western ligand blotting with [^{125}I]IGF-II as described below] were selected and recloned. The single clone producing the greatest amount of IGFBP-6 was then expanded for production of recombinant protein.

Purification of Recombinant Glycosylated Human IGFBP-6. Medium (1 L) conditioned by CHO cells overexpressing IGFBP-6 was applied to an IGF-II affinity column at 1 mL/min for 16 h at 4 °C. The column was washed extensively with Dulbecco's phosphate-buffered saline/0.5 M NaCl, and IGFBPs were eluted with 0.5 M acetic acid (*11*). The eluate was neutralized and applied onto a wheat germ agglutinin (WGA)−agarose affinity column (Sigma) at 1 mL/min for 4 h at 4 °C. After washing with Dulbecco's phosphate-buffered saline, glycosylated proteins containing glucosamine and/or sialic acid monosaccharides, including IGFBP-6 (*11*), were eluted with 0.3 M *N*-acetylglucosamine (Sigma). Proteins in the eluate were further purified by reverse-phase medium-pressure chromatography (ProRPC 5/10 on an FPLC system, Pharmacia, Piscataway, NJ) using a 16−40% acetonitrile/0.1% trifluoroacetic acid (TFA) gradient. Initially, the identity of IGFBP-6 was confirmed by immunoblotting and Western ligand blotting as described below.

Approximately 120 μg of IGFBP-6 was purified from 200 mL of conditioned medium. This IGFBP-6 preparation bound IGF-II with the characteristic 30-fold greater affinity than IGF-I (results not shown), and the amount of IGF-II binding was consistent with the estimated amount of protein based on UV absorbance at 280 nm (results not shown). These results indicate that the expressed IGFBP-6 was biologically active and that significant misfolding is unlikely to have occurred.

Expression and Purification of Recombinant Nonglycosylated Human IGFBP-6 in Escherichia coli. The cDNA for human IGFBP-6 was modified using the polymerase chain reaction to introduce a *Bam*HI site to the 5′ end and to commence translation at Arg^{28}, since this is the predominant form of IGFBP-6 in biological fluids (*31*). Primer sequences were 5′-AATGGATCCCGGTGCCCAGGCTGCGGGCAA-3′ and 5′-TTTTCTAGAGAATTCTTAGCCGCTACTCCCAGTGGGGCAGGA-3′. The modified cDNA was subcloned into the *E. coli* expression vector, pGEX-2T (Pharmacia, Piscataway, NJ), which then encoded a fusion protein of IGFBP-6 separated from the C-terminus of glutathione-S-transferase by a thrombin cleavage site. Following cleavage, an additional Gly-Ser dipeptide from the thrombin cleavage site sequence remains on the amino terminus of IGFBP-6.

E. coli (HB101) were transformed and screened by restriction digestion, and the correct coding sequence of a positive clone was confirmed. Transformed *E. coli* were then grown in 1 L of Luria−Bertani broth, and protein expression was induced with 1 mM isopropyl β-D-thiogalactopyranoside (Sigma) for 3 h at 37 °C. Cells were sonicated in 50 mM Tris-HCl (pH 8.0)/0.1% Triton X-100. Following centrifugation (10000*g*, 30 min, 4 °C), post-sonicate lysate was recycled over a glutathione−agarose column (Pharmacia) at 1 mL/min for 16 h at 4 °C. The column was washed extensively with 50 mM Tris-HCl (pH 8.0)/0.1 M NaCl and equilibrated in the same buffer containing 2.5 mM $CaCl_2$. IGFBP-6 was cleaved from the column with thrombin (20 units, Sigma) for 1 h at 37 °C, following which cleavage was terminated by the addition of 2.5 mM EDTA/1 mM phenylmethanesulfonyl fluoride. Nonglycosylated IGFBP-6 was then purified by IGF-II affinity chromatography and reverse-phase medium-pressure chromatography as described above. Initially, the identity of nonglycosylated IGFBP-6 was confirmed by immunoblotting and Western ligand blotting as described below.

Immunoblotting and Western Ligand Blotting (11). Proteins were separated by sodium dodecyl sulfate−polyacrylamide (12 or 15%) gel electrophoresis under nonreducing conditions and transferred to nitrocellulose membranes, which were subsequently blocked with 5% nonfat skim milk (for immunoblotting) or 1% bovine serum albumin (for Western ligand blotting). For immunoblotting, membranes were probed with an anti-human IGFBP-6 polyclonal antiserum (kindly provided by Dr. M. Kiefer, Chiron Corp., Emeryville, CA), and detection was performed using enhanced chemiluminescence (Super Signal, Pierce, Rockford, IL). Blots were exposed to X-ray film for 1−10 min. For Western ligand blotting, membranes were probed with [^{125}I]-IGF-II and exposed to X-ray film for 1−5 days.

Tryptic Digestion and Reverse-Phase HPLC for Structural Analysis. Prior to tryptic digestion, 20 μg of IGFBP-6 glycoprotein was incubated for 1 h at 37 °C in a total volume of 4 μL of 100 mM ammonium bicarbonate, 6 M urea, and

10 mM dithiothreitol (DTT), intended to induce disulfide bond cleavage (however, as noted in Results below, all disulfide bonds remained intact under the above conditions, which subsequently proved advantageous in allowing information regarding disulfide linkages to be obtained). The urea concentration was reduced to 0.8 M by increasing the volume to 30 μL with 100 mM ammonium bicarbonate, 10 mM DTT, and 5 mM $CaCl_2$, and 0.5 μg of trypsin was added to give a trypsin:protein ratio of 1:40 (w/w). Tryptic digestion was carried out for 18 h at 22 °C and terminated by acidification (0.5% TFA), after which the digest was subjected to reverse-phase HPLC (see below) and ESMS analysis. Reduction and alkylation of approximately 100 pmol of disulfide-linked peptide complexes contained in 20 μL of HPLC-purified fraction 7 was carried out in 200 μL of 5 M guanidine-HCl, 10 mM DTT, and 100 mM iodoacetamide (90 °C, 20 min), followed by acidification (0.2% TFA) and further reverse-phase HPLC.

Samples for reverse-phase HPLC were acidified (0.2% or 0.5% TFA), filtered [0.45 μm poly(vinylidine difluoride) syringe filter], and subjected to HPLC on a C18 microbore column (Vydac 218TP51, 1 mm × 250 mm, 5-μm particle size; 300-Å pore size) attached to a Hewlett-Packard HP-1090 liquid chromatograph. The column was protected by a 0.5-μm prefilter and a C18 guard column (Alltech Macrosphere 300-C18-5u). After column equilibration in 0.1% TFA at a flow rate of 40 μL/min, 250 μL of sample was loaded via a 500-μL loop, the column was washed with 0.1% TFA for 30 min, and the loop was switched out of the flow line to prevent gradient distortion. Peptides were then eluted from the column at the same flow rate, using a linear applied gradient of 0−50% acetonitrile in 0.1% TFA over 150 min, followed by 50−70% acetonitrile in 0.1% TFA over a further 20 min. Fractions of 20−100 μL were collected manually while UV absorbance was monitored at 215 nm. Absorbance data for 215 and 280 nm were simultaneously collected by computer from the diode array detector.

Electrospray Ionization Mass Spectrometry (ESMS). ESMS was carried out using a Perkin-Elmer Sciex API-300 triple quadrupole mass spectrometer fitted with a micro-ionspray ion source and operated in the Q1 scan mode. Reverse-phase HPLC fractions (2−4 μL) were diluted 1:1 with 70% acetonitrile containing 0.2% acetic acid prior to analysis. Samples were infused through 20 μm i.d. fused silica tubing using a 10-μL Hamilton gastight syringe driven by a Harvard syringe pump at a flow rate of 0.2 μL/min. Signal-averaged mass spectra were obtained from 50−100 scans over 5−10 min using a scan range of *m*/*z* 100−3000 and a constant peak width at half-height of 0.6 amu. Source collisional fragmentation spectra were obtained by increasing the source orifice voltage (V_{OR}) from a low value of typically 25−50 V to a high value of 100−200 V, depending on peptide mass. Proteins required a V_{OR} of 50−60 V for satisfactory signals with minimal fragmentation of protein and minimal apparent loss of terminal *N*-acetylneuraminic acid groups from glycoproteins. Partial oxidation of methionine-containing peptides, leading to a 16-Da mass shift, was identified and reversed by incubation of 5−10 μL of each HPLC fraction with 2% (140 mM) 2-mercaptoethanol for 2 h at 37 °C; however, the same treatment was ineffective in breaking disulfide bonds (intermolecular, but not intramolecular, disulfide bonds were broken using this treatment by raising the temperature to 90−100 °C for 5 min). Prior to the recording of spectra, the mass spectrometer mass scale was calibrated at eight points spaced evenly over the 3000 u mass range, to an accuracy equivalent to ±0.01%, using singly charged poly(propylene glycol) ions. Mass spectra were processed by signal averaging with smoothing only where necessary, manual mass determination, and transformation using PE-Sciex BioMultiView software.

MTPHRLLPPL LLLLALLLAA SPGGALARCP GCGQGVQAGC PGGCVEEEDG[50]
GSPAEGCAEA EGCLRREGQE CGVYTPNCAP GLQCHPPKDD EAPLRALLLG[100]
RGRCLPARAP AVAEENPKES KPQAGTARPQ DVNRRDQQRN PGTSTTPSQP[150]
NSAGVQDTEM GPCRRHLDSV LQQLQTEVYR GAQTLYVPNC DHRGFYRKRQ[200]
CRSSQGQRRG PCWCVDRMGK SLPGSPDGNG SSSCPTGSSG[240]

FIGURE 1: Amino acid sequence of human IGFBP-6. The sequence is based on the entry IBP6_HUMAN in the SWISS-PROT database (accession number P24592). The 15 serines and 9 threonines in the mature protein (L^{26}−G^{240}) are shown in bold, and the 16 cysteines are underlined.

Amino Acid Sequencing. N-Terminal amino acid sequences were obtained by sequential Edman degradation using a Hewlett-Packard G1005A automated protein sequencing system, calibrated with PTH-amino acid standards prior to each sequencing run.

Limited Proteolysis of IGFBP-6. Glycosylated or nonglycosylated IGFBP-6 (200 ng) was incubated with 0−10 ng of trypsin (Sigma) or chymotrypsin (Koch-Light Laboratories, Colnbrook, U.K.) for 2 h at 37 °C in 50 mM Tris-HCl (pH 8.0). Samples were analyzed by immunoblotting as described above.

Metabolic Labeling of IGFBP-6 with [^{32}P]Orthophosphate (32). To determine whether IGFBP-6 is phosphorylated, CHO cells overexpressing IGFBP-6 were grown to confluence and incubated in phosphate-free Dulbecco's MEM containing [^{32}P]orthophosphate (50 μCi/mL, Amersham) for 16 h at 37 °C. IGFBP-6 was immunoprecipitated using the anti-human IGFBP-6 polyclonal antiserum described above. The effectiveness of immunoprecipitation was verified by Western ligand blotting. Phosphorylation was analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis of the immunoprecipitate followed by autoradiography.

RESULTS

Expression and Purification of IGFBP-6. IGFBP-6 was purified previously from human cerebrospinal fluid and media conditioned by transiently overexpressing COS-7 cells using IGF-II affinity chromatography and reverse-phase medium-pressure chromatography (*11*, *29*). However, when this purification procedure was applied to media conditioned by IGFBP-6-overexpressing CHO cells, there was substantial contamination by an endogenously synthesized IGFBP. This IGFBP migrated at 24 kDa by Western ligand blotting and was not immunoreactive with an antiserum to IGFBP-6 (results not shown). Quantitative N-terminal Edman sequencing of this initial preparation revealed that 80−85% of the protein content was IGFBP-6 [major and minor sequences R(C)PG(C)GQG and LAR(C)PG(C)G, respectively, with the expected cysteines (Figure 1) in parentheses to indicate absence of a significant PTH-amino acid signal increase in the corresponding Edman cycle], whereas the

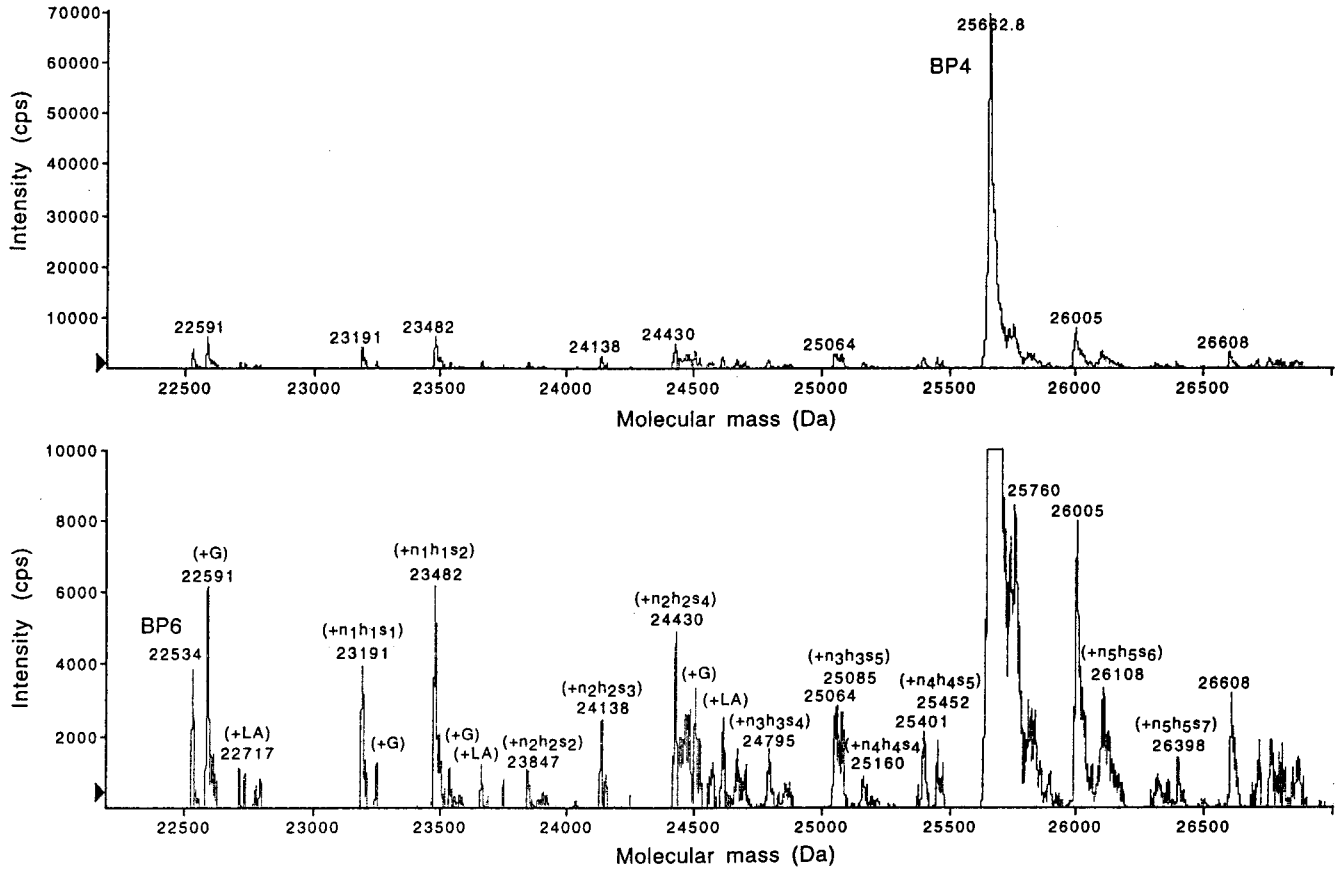

FIGURE 2: ESMS mass spectra of glycosylated IGFBP-6 purified without lectin chromatography. The glycosylated IGFBP-6 preparation (1 μg, 40 pmol) containing 15−20% IGFBP-4 contaminant (as described in Results) in a volume of 1.5 μL (30 μM concentration) was analyzed by ESMS as described in Experimental Procedures. Charge states of 13^+−26^+ were observed in the raw mass spectrum (not shown), which was subsequently transformed to a true mass scale with baseline subtraction and smoothing (three times with three point moving average, weighted 1:2:1). Upper spectrum: mass spectrum (transform) showing a single major peak corresponding to the nonglycosylated IGFBP-4 contaminant. Lower spectrum: same spectrum amplified in the vertical scale to show the much smaller peaks due to the more abundant but heterogeneously glycosylated IGFBP-6. Peaks are labeled with mass in daltons and inferred carbohydrate content in terms of numbers of *N*-acetylhexosamine, hexose, and *N*-acetylneuraminic acid (sialic acid) monosaccharides, denoted by n, h, and s, respectively.

remaining 15−20% had a sequence commencing DEAIH-(C)PP, which is identical with that of rat and human IGFBP-4. ESMS analysis of this preparation (Figure 2) showed a single major peak that was attributed to nonglycosylated hamster IGFBP-4, with a molecular mass of 25 662.8 ± 1.8 Da, compared with calculated masses of 25 678.2 and 25 954.3 Da for rat and human IGFBP-4, respectively. Also present were peaks of much lower intensity at molecular masses including 22 534, 22 591, 23 482, and 24 430 Da, the first two corresponding to nonglycosylated IGFBP-6 with and without the C-terminal Gly (as described below) and the second two subsequently shown to represent IGFBP-6 with varying degrees of glycosylation (Figure 3; Table 1). Although IGFBP-4 was present at a concentration ~5-fold lower than glycosylated IGFBP-6 as determined by Edman sequencing, the observed ESMS signal intensities were an order of magnitude higher. This can be attributed partly to carbohydrate heterogeneity leading to multiple IGFBP-6 mass components compared to a single IGFBP-4 component, and partly to the lower sensitivity of ESMS toward glycoproteins and glycopeptides compared with nonglycosylated polypeptides, as consistently noted in this study. Comparing the ESMS peak sizes representing nonglycosylated IGFBP-4 (which constitutes 15−20% of the total protein content based on Edman sequencing) and IGFBP-6, and assuming equal sensitivity by ESMS, the proportion of IGFBP-6 that is nonglycosylated is estimated to be <3%.

Since there was an appreciable level of apparently nonglycosylated IGFBP-4 detected by Western ligand blotting and ESMS of the above sample and little evidence of 28-kDa glycosylated IGFBP-4 detected by Western ligand blotting of media conditioned by untransfected CHO cells (results not shown), an additional purification step involving lectin chromatography with wheat germ agglutinin (WGA) was added after IGF-II affinity chromatography. WGA binds glycoproteins containing glucosamine and sialic acid residues, including IGFBP-6 (*11*, *29*), and would not be expected to bind nonglycosylated IGFBP-4. Following the addition of this step to the purification procedure, the 24-kDa band was no longer seen by Western ligand blotting or Coomassie staining (results not shown), and Edman sequencing of this preparation demonstrated that the IGFBP-6 and IGFBP-4 contents were >97% and <3% respectively. The strong 25 663-Da peak previously observed by ESMS (Figure 2) and attributed to nonglycosylated IGFBP-4 was weak or absent (reduced by >100-fold) in ESMS mass spectra of the more highly purified preparation at a similar but slightly lower glycoprotein concentration (Figure 3), indicating that the level of contaminating nonglycosylated IGFBP-4 was <0.3%. On this basis, most of the residual IGFBP-4 detected

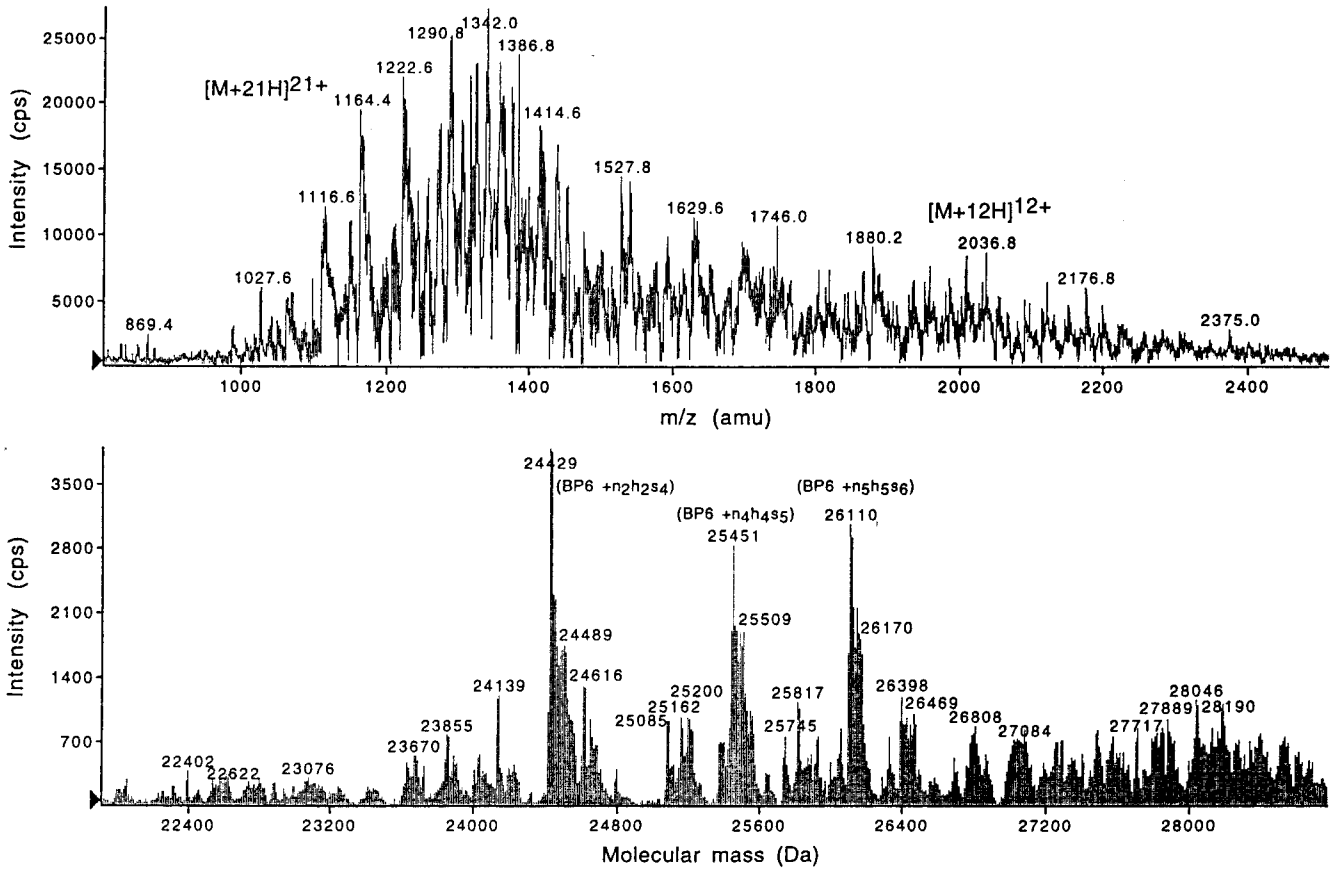

FIGURE 3: ESMS mass spectra of glycosylated IGFBP-6 purified using WGA lectin chromatography. Glycosylated rhIGFBP-6 (1 μg, 40 pmol) in a volume of 2 μL (20 μM concentration) was analyzed by ESMS as described in Experimental Procedures. Upper spectrum: raw mass spectrum showing peaks due to ions with charge states of 10^+-24^+. Lower spectrum: transform from *m/z* to true mass scale with baseline subtraction and smoothing (three times with three point moving average, weighted 1:2:1), showing multiple mass components corresponding to IGFBP-6 with various degrees of glycosylation, as indicated for the major peaks. Peaks are labeled with mass in daltons and inferred carbohydrate content in terms of numbers of *N*-acetylhexosamine, hexose, and *N*-acetylneuraminic acid (sialic acid) monosaccharides, denoted by n, h, and s, respectively (see Table 1 for further analysis).

Table 1: Major Glycoproteins Identified by ESMS of IGFBP-6[a]

av glycoprotein mass (Da)		inferred monosaccharide composition		
obsd	calcd	HexNAc	Hex	NeuAc
24 138.5 ± 3.2	24 137.4	2	2	3
24 429.3 ± 2.5	**24 428.7**	**2**	**2**	**4**
25 085.1 ± 3.0	25 085.2	3	3	5
25 162.0 ± 2.9	25 159.3	4	4	4
25 450.6 ± 3.8	**25 450.5**	**4**	**4**	**5**
25 744.6 ± 3.9	25 741.8	4	4	6
25 816.5 ± 3.6	25 815.9	5	5	5
26 110.2 ± 3.3	**26 107.1**	**5**	**5**	**6**
26 398.2 ± 3.3	26 398.4	5	5	7
26 469.4 ± 4.0	26 472.4	6	6	6

[a] Observed masses are ESMS masses with ±95% confidence limits determined statistically from observed multiple charge states (Figure 3), typically 11^+-23^+. Calculated masses are based on a calculated mass of 22533.0 Da for oxidized IGFBP-6, plus the following monosaccharide average molecular masses: HexNAc (203.195 Da), Hex (162.142 Da), NeuAc (291.258 Da). Results for the three major observed components (Figure 3) are shown in bold. Results are not shown for additional minor components corresponding to the major components but with an additional C-terminal Gly^{240} or N-terminal $Leu^{26}-Ala^{27}$ (see Figures 2 and 3). HexNAc = *N*-acetylhexosamine, Hex = hexose, and NeuAc = *N*-acetylneuraminic acid.

by Edman sequencing is likely to be glycosylated. This more highly purified IGFBP-6 preparation was used for subsequent experiments.

Microheterogeneity of the Amino Acid Sequence of IGFBP-6. Edman N-terminal sequencing of recombinant human IGFBP-6 demonstrated that ~80% of the protein commenced at Arg^{28} with the sequence R(C)PG(C)GQG and most of the remainder (~18% of total protein) commenced at Leu^{26} with the sequence LAR(C)PG(C)G (amino acid numbering based on the sequence of the precursor protein, IBP6_HUMAN, in the SWISS-PROT database, accession number P24592; Figure 1). As inferred from ESMS data, forms of IGFBP-6 with and without the expected C-terminal Gly^{240} residue (differing in mass by 57 Da) were present (Figures 2 and 3). The forms lacking this amino acid appeared to be the more abundant, as confirmed by ESMS analysis of C-terminal tryptic fragments of IGFBP-6, as described below. It is concluded that the predominant form of glycosylated rhIGFBP-6 is Arg^{28}-Ser^{239}. By contrast, nonglycosylated rhIGFBP-6 exhibited a single major ESMS mass of 22 737.6 ± 3.3 Da (Figure 4), corresponding to the expected full-length sequence $G^1SRCP\cdots GSSG^{215}$ (calculated mass, 22 735.1 Da), and a component lacking the C-terminal Gly was not observed.

Microheterogeneity of the Carbohydrate Composition of IGFBP-6. ESMS analysis of recombinant human IGFBP-6 (Figure 3) revealed three major and a considerable number of minor components with masses greater than that calculated for IGFBP-6 by 2−4 kDa or more. The masses of the three

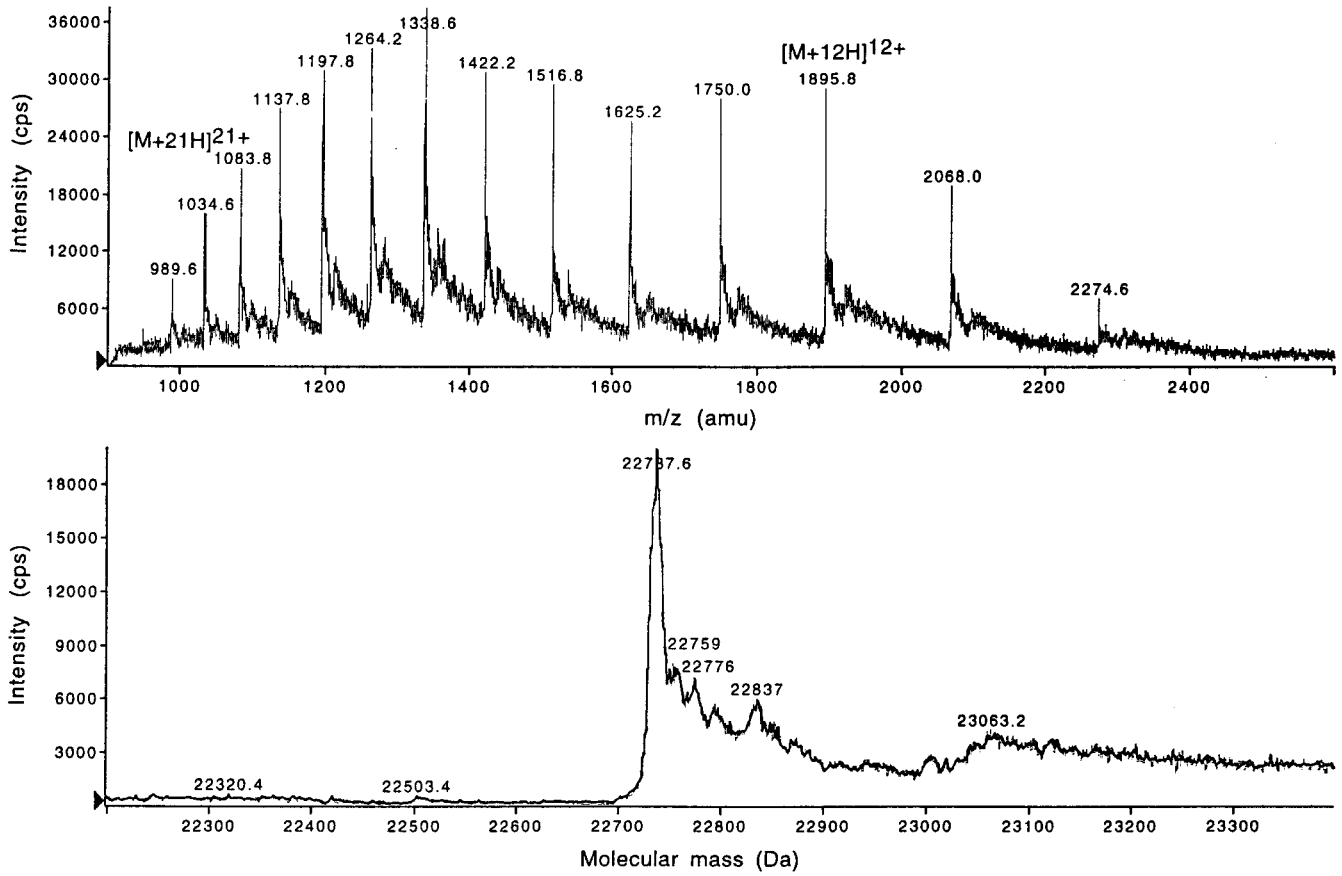

FIGURE 4: ESMS mass spectra of nonglycosylated IGFBP-6. Nonglycosylated IGFBP-6 (1 μg, 40 pmol) in a volume of 2 μL (20 μM concentration) was analyzed by ESMS as described in Experimental Procedures. Upper spectrum: raw mass spectrum showing peaks due to ions with charge states of 10^+-24^+. Lower spectrum: transform from *m/z* to true mass scale without baseline subtraction or smoothing, showing a single major mass component corresponding to the expected sequence GSRCPG...SSG^{215} with 8 disulfide linkages (expected mass, 22 735.1 Da; observed mass, 22 737.6 $\pm$ 3.3 Da).

major components correspond to that of the IGFBP-6 sequence with the following numbers of monosaccharide units attached: *N*-acetylhexosamine (2, 4, or 5 HexNAc, respectively); hexose (2, 4, or 5 Hex, respectively); and the most common sialic acid, *N*-acetylneuraminic acid (4, 5, or 6 NeuAc, respectively) (Table 1). The masses of minor peaks on either side of the major peaks can be related to those of the major peaks through addition or subtraction of the mass of either a NeuAc monosaccharide or a Hex-HexNAc disaccharide (Table 1). The mass distributions of IGFBP-6 purified with and without WGA lectin chromatography differed in that the latter contained glycoforms with lower molecular masses, including nonglycosylated IGFBP-6, the peak intensity for which was at least as great as those for glycosylated forms (Figures 2 and 3). Since nonglycosylated IGFBP-6 was not detected in immunoblots of IGFBP-6 purified without WGA lectin chromatography (results not shown), this indicates considerable ESMS mass discrimination in favor of species with lower degrees of glycosylation. With WGA lectin-purified IGFBP-6, however, this effect may have been at least partly countered by the likely discrimination of lectin chromatography in favor of species with a higher degree of glycosylation. Subject to these considerations, the total carbohydrate content of WGA-purified IGFBP-6, on the basis of the ESMS data (Figure 3), lies predominantly in the range of 7−18 monosaccharide units (Table 1), including 2−6 HexNAc units.

Tryptic Digestion of IGFBP-6. After digestion of 20 μg (approximately 800 pmol) of IGFBP-6 with trypsin, the resulting peptide components were separated by microbore reverse-phase HPLC (Figure 5) and subjected to ESMS analysis. Peptide fragments were identified by comparing their observed masses (Table 2) with those of tryptic fragments predicted from the IGFBP-6 amino acid sequence (Figure 1), and by using additional Edman N-terminal sequence data in selected cases. The presence of the aromatic amino acids Trp and/or Tyr was confirmed by UV absorbance peaks at 280 nm (Figure 5). A number of peak fractions were found to contain disulfide-linked tryptic peptide complexes (Table 2), which were subsequently separated by reduction at elevated temperatures (90−100 °C for 5 min in 2% 2-mercaptoethanol) and then reanalyzed by ESMS with or without HPLC separation; however, none of the intramolecular disulfide bonds (fractions 7, 9, and 11, Figure 5 and Table 2) were cleaved by this treatment. All 16 of the IGFBP-6 cysteines were uniquely accounted for by the observed disulfide-linked peptides; none of the cysteines were involved in more than one complex, and (in particular) no disulfide-linked homodimers were observed, as might be expected if disulfide bonds had been broken and reformed after tryptic digestion (*23*). This evidence strongly implies that the attempted reduction of IGFBP-6 in 10 mM DTT and 6 M urea at 37 °C did not result in disulfide bond cleavage. Moreover, tryptic digestion of a similar quantity of IGFBP-6 under the same conditions, but with

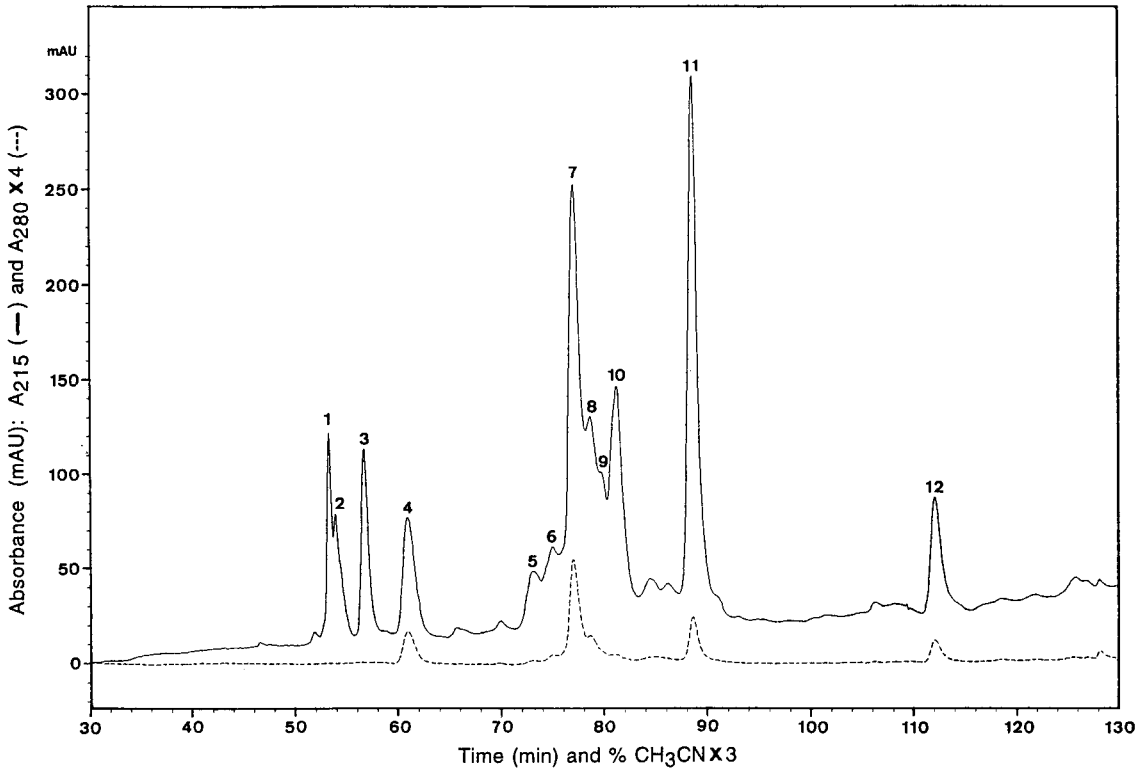

FIGURE 5: Reverse-phase HPLC of a tryptic digest of glycosylated IGFBP-6. Glycosylated IGFBP-6 (20 μg, 800 pmol) digested with trypsin was subjected to reverse-phase HPLC as described in Experimental Procedures. Solid trace: A_{215}. Dashed trace: A_{280}, amplified 4-fold vertically for clarity. Numbered (and other) peaks were subsequently analyzed by ESMS (Table 2) and/or Edman N-terminal sequencing (Table 4). Peak fractions 1, 2, 5, 6, 7, and 9 were found to contain glycopeptides (Table 3).

the addition of iodoacetamide as alkylating agent (5-fold molar excess over DTT), did not result in cysteine alkylation, the major entities identified by ESMS of reverse-phase HPLC fractions being the same disulfide-linked peptides and peptide complexes described below (results not shown). An aliquot of the same batch of iodoacetamide was subsequently used to successfully alkylate disulfide-linked peptides in HPLC fraction 7 (Table 2) using 5 M guanidine and elevated temperatures, thereby confirming its alkylating activity. We conclude that the disulfide bonds identified in the present study are highly likely to be the native linkages of IGFBP-6.

Determination of O-Linked Glycosylation Sites of IGFBP-6. Not all expected peptide fragments of IGFBP-6 were recovered on reverse-phase HPLC and identified by ESMS; however, 88% of the IGFBP-6 sequence was identified, including 22 of the 24 serines or threonines (Figure 1). A small predicted tryptic fragment (203−208) SSQGQR, containing two serines, was not located (Table 2). Apart from the peptide containing amino acids 119−134 and the disulfide-linked peptides containing amino acids 140−165 and 181−193, the mass and sequence data excluded O-linked glycosylation of all other peptides. The absence of O-glycosylation of amino acids 221−240, the Ser-rich C-terminal region of IGFBP-6, was confirmed by ESMS of HPLC-separated peptides obtained by reduction and Cys-alkylation of HPLC fraction 7 (Table 2). The carboxyl-terminal peptide also contains a potential N-glycosylation site at Asn^{229} (*6*); therefore, the ESMS data confirm previous observations, based on carbohydrate composition and lack of digestion by *N*-glycanase (*11*), that this site is not glycosylated.

Mass spectra of HPLC fractions 1 and 2 (Figure 6A,C) exhibited single major mass components corresponding uniquely to a tryptic fragment containing amino acids 119−134 of IGFBP-6 with four and three attached monosaccharide units, respectively (Table 3). The mass difference between the two glycopeptides corresponds precisely to the mass of a single NeuAc. Ion source fragmentation spectra of fractions 1 and 2 (Figure 6B,D) exhibited a ladder of sequential peaks consistent with the carbohydrate sequences $(NeuAc)_2$-Hex-HexNAc, which may represent a linear or branched sequence, and NeuAc-Hex-HexNAc, respectively (Table 3). Because the various HexNAc or Hex isomers are isobaric, their identities cannot be determined from mass alone; however, galactose-*N*-acetylgalactosamine is the most common core disaccharide of O-linked carbohydrate chains (*17*). The relative peak areas of fractions 1 and 2 indicated that the proportions of $(NeuAc)_2$-Hex-HexNAc and NeuAc-Hex-HexNAc were ∼70% and ∼30%, respectively.

There are two potential O-glycosylation sites (Ser^{120} and Thr^{126}) in the fraction 1 glycopeptide. Edman sequencing of the first 10 amino acids of the peptide in fraction 1 revealed the sequence ESKPQAG(−)AR, corresponding to amino acids 119−128 of IGFBP-6, but with no significant PTH-amino acid signal increase in the position for Thr^{126}. Since glycosylated amino acids are not detected under Edman sequencing conditions (*33*), this indicates that Thr^{126} is highly likely to be the glycosylation site. It is assumed that the glycosylation site on the fraction 2 glycopeptide is the same as that on the fraction 1 glycopeptide, given that these differ only by one NeuAc unit (Figure 6; Table 3). Despite careful screening of HPLC-purified fragments (Figure 5;

Table 2: ESMS Masses and Sequence Assignments of Tryptic Fragments of IGFBP-6[a]

HPLC fraction	obsd charge states	average molecular mass		IGFBP−6 sequence assignment	
		obsd	calcd	position no.	sequence
				26, 27	LA[b]
8b	3, 4	3633.8 ± 0.9	3633.9	28−65	RCPG···GCLR[c]
8a	3, 4	3790.7 ± 0.8	3790.1	28−66	RCPG···CLRR[c]
10b	2, 3	3477.6 ± 0.6	3477.8	29−65	CPGC···GCLR[c]
10a	2−4	3634.2 ± 0.5	3633.9	29−66	CPGC···CLRR[c]
11	2−5	3680.1 ± 0.5	3680.1	**67−95**	**EGQE···APLR[d]**
11	2−5	3680.1 ± 0.5	3680.1	**104−108**	**CLPAR**
11 (red.)	3, 4	3123.4 ± 0.8	3123.5	67−95	EGQE···APLR[e]
				96−103	ALLLGR/GR[b]
3	1, 2	1025.0 ± 0.3	1025.1	109−118	APAVAEENPK
1	2−4	2701.7 ± 0.8	2701.7	119−134	ESKPQAGT(*)ARPQDVNR
2	2−4	2410.5 ± 0.7	2410.5	119−134	ESKPQAGT(*)ARPQDVNR
				135−139	R/DQQR[b]
5, 6, 7b, 9	3−6	*g*	*g*	**140−165**	**NPGT···PCRR[f]**
5, 6, 7b, 9	3−6	*g*	*g*	**181−193**	**GAQTLYVPNCDHR**
12	1−3	1829.1 ± 0.3	1829.0	166−180	HLDSVLQQLQTEVYR
4	1, 2	541.5 ± 0.3	541.6	194−197	GFYR
				198−199	KR[b]
7a	2−4	3186.0 ± 0.7	3186.5	**200−202**	**QCR[d]**
7a	2−4	3186.0 ± 0.7	3186.5	**209−217**	**RGPCWCVDR**
7a	2−4	3186.0 ± 0.7	3186.5	**221−239**	**SLPG···TGSS**
7c (minor)	2−4	3244.0 ± 0.8	3243.5		as above, with extra G^{240}
				203−208	SSQGQR[b]
				218−220	MGK[b]
		Cys-Alkylated (S-cam) Tryptic Fragments Isolated from HPLC Fraction 7			
7/32	2	1204.6 ± 0.8	1205.4	209−217	RGPC(cam)WC(cam)VDR
7/21	2	1750.2 ± 0.8	1750.8	221−239	SLPG···C(cam)PTGSS
7/22 (minor)	2	1807.2 ± 0.8	1807.8	221−239	SLPG···C(cam)PTGSSG

[a] Tryptic fragments of IGFBP-6 isolated by HPLC were analyzed by ESMS as described in Experimental Procedures. Observed molecular masses are averages for the indicated observed charge states, with ±95% confidence limits based on small sample statistics. Calculated masses are based on the assigned sequences indicated. Multiple mass components present in a single HPLC fraction are designated a, b, c. Components 10a and 8b, which are isobaric, have been given different assignments on the basis of differences in HPLC elution time and observed charge states. Assigned sequences of peptide complexes, containing two or three disulfide-linked tryptic fragments, are shown in bold. HPLC fractions containing glycopeptides (fractions 1, 2, 5, 6, 7 and 9) are indicated in italics, and T(*) indicates glycosylated T^{126} (see Table 3 for detailed glycopeptide mass data). Not shown are mass data for additional minor mass components of HPLC fractions 5, 6, 7 (7b), and 9 which are 156 Da lower (or, in the case of fraction 2 trailing edge, 156 Da higher) than the major components shown and which are characteristic of partially missed tryptic cleavages at an RR sequence element (as with fractions 8a/8b and 10a/10b). Additional mass data for fraction 11 (red.) were obtained after reduction with 2-mercaptoethanol, and for fraction 7 after Cys-alkylation and further HPLC (yielding fractions 7/21, 7/22, and 7/32), as described in Experimental Procedures. C(cam) = *S*-carboxyamidomethylcysteine. [b] Peptide not located. [c] Contains 3 intramolecular disulfide bonds. [d] Oxidized complex containing 2 disulfide bonds. [e] Partially oxidized, containing 1 disulfide bond. [f] Oxidized complex containing 1 disulfide bond. [g] See Table 3.

Table 2), this peptide was not observed in the nonglycosylated peptide state, implying that Thr^{126} is always glycosylated in IGFBP-6.

Analysis by ESMS of HPLC fractions 5, 6, 7, and 9 (Figure 7; Tables 2 and 3) coupled with N-terminal Edman sequencing of fractions 5−7 indicated the presence in each fraction of the same peptide complex consisting of tryptic fragments 140−165 and 181−193 linked by a single disulfide bond (Cys^{163}−Cys^{190}) and glycosylated to varying degrees. In each case, glycosylation was evidenced by characteristic ESMS mass differences between adjacent major component peaks corresponding to NeuAc (291 Da), Hex-HexNAc (365 Da), and Hex (162 Da) moieties (Table 3), as observed with fractions 1 and 2 (Figure 6). Additional evidence for the presence of glycopeptides was obtained by examining the low mass ($< m/z$ 500 u) regions of spectra recorded using high source collision conditions ($V_{OR} = 100$ V or more) in order to detect carbohydrate-specific fragment ion peaks such as those at m/z 204 and 366, due to the oxonium ions of HexNAc and Hex-HexNAc, respectively (*34*). The presence of carbohydrate in fractions 5, 6, 7, and 9 (and fractions 1 and 2), but not other fractions, was indeed confirmed by observation either of the m/z 204 and 366 diagnostic peaks or of the same peaks shifted to m/z 226 and 388, respectively, due to sodium adduction; to m/z 242 and 404, respectively, due to potassium adduction; and to m/z 203 and 365, respectively, due evidently to loss of neutral sodium or potassium from the adduct ions (results not shown). Because of this complication, identification of glycopeptides was based primarily on the mass data recorded under low ion source collision conditions (Table 3).

Consistent with the presence of the above peptide complex, Edman N-terminal sequencing (17 cycles) of fraction 5 revealed two major sequences that were resolved by comparison with the IGFBP-6 sequence (Figure 1) to yield the sequences NPG(T)(S)(−)(−)(P)SQPN(−)AGVQDTE and GAQ(T)LYV(P)N(C)DH, where parentheses indicate either a missing expected Ser or Thr (indicative of glycosylation), a missing expected cysteine as indicated, or ambiguity due to overlapping signals (see below) or signal carryover. Also present was a partial minor sequence identified as (S)L(P)-(G)(S)PDG, which corresponds to the C-terminal (221−239) tryptic fragment that is not glycosylated, as detailed above. The presence or absence of expected Thr or Ser signals in the above two major sequences from fraction 5 indicates that Thr^{145}, Thr^{146}, and Ser^{152} are highly likely to be glycosylated, whereas Ser^{148} and Thr^{158} are not. An ambiguity arising from the presence of both the 140−165 and 181−193 peptide

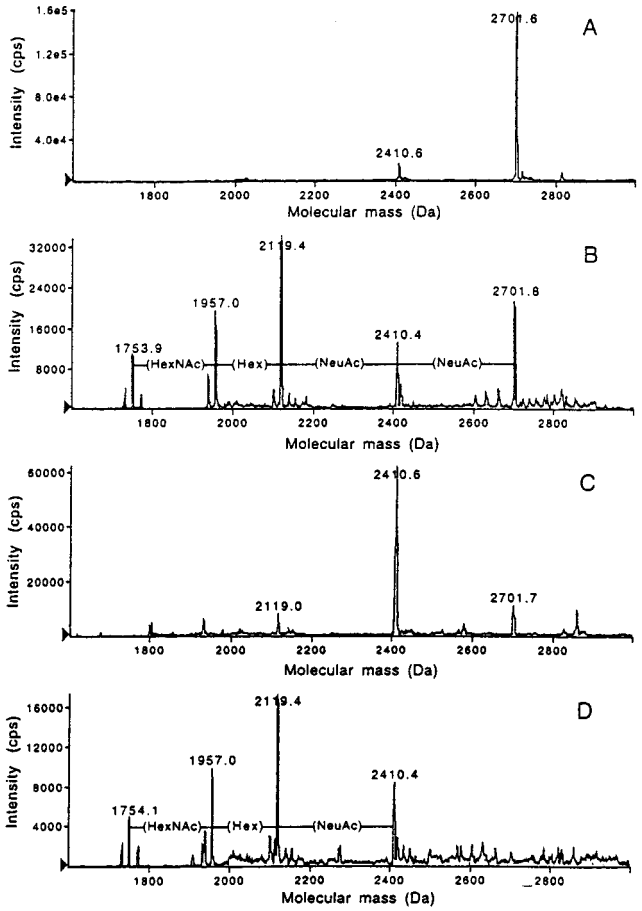

FIGURE 6: ESMS mass spectra of glycopeptides in HPLC fractions 1 and 2. Glycopeptides contained in HPLC fractions 1 and 2 (approximately 10 and 3 pmol, respectively; Figure 5) were analyzed at concentrations of approximately 3 and 1 μM, respectively, by ESMS as described in Experimental Procedures. Raw mass spectra (not shown) exhibiting charge states of 2^+-4^+ were recorded under low and high source collision conditions (V_{OR} = 25 and 150 V, respectively) and transformed from *m*/*z* to a true mass scale, with baseline subtraction but no smoothing. Spectra: A, fraction 1 (V_{OR} = 25 V); B, fraction 1 (V_{OR} = 150 V); C, fraction 2 (V_{OR} = 25 V); D, fraction 2 (V_{OR} = 150 V). Peaks are labeled with mass in daltons and inferred differences in carbohydrate content as indicated (see Table 3).

sequences relates to Thr[143] and Thr[184], both of which could contribute to the observed Thr signal in cycle 4 of the Edman sequencing data. However, quantitation of this signal in cycle 4 from three separate sequencing analyses (fractions 5, 6, and 7) indicates that the signal was too large to derive from either of these sequences alone (Table 4), implying that neither Thr[143] nor Thr[184] is glycosylated. A similar ambiguity arising from the presence of both the 140−165 and 221−239 peptide sequences relates to Ser[144] and Ser[225], both of which could contribute to the observed Ser signal in cycle 5 of the Edman sequencing data. This signal for cycle 5 of Edman sequencing of fraction 5 was too low to be consistent with that expected from both Ser[144] and Ser[225], the latter of which is not glycosylated, but too high to be due to Ser[225] alone (Table 4), indicating that Ser[144] is highly likely to be partially O-glycosylated. In contrast, the Ser signal in cycle 5 of Edman sequencing of fractions 6 and 7 was too large to be accounted for by either Ser[144] or Ser[225] alone, suggesting that neither is glycosylated in these fractions (Table 4).

Edman sequencing of fraction 6 (data not shown) revealed essentially the same major and minor sequences observed with fraction 5 (as detailed above) except that a definite Thr signal was observed in Edman cycle 6. This indicates that, whereas Thr[146] was glycosylated in the fraction 5 glycopeptides, it is predominantly nonglycosylated in the fraction 6 glycopeptides.

On the basis of ESMS (Table 3), glycopeptides in fraction 6 were indeed glycosylated to a lesser extent than those in fractions 5 and 7. The difference in mass (Table 3) between the most intense ESMS peak in fraction 6 (6131 Da) and that in fractions 5 and 7 (6787 Da) corresponds to the combined mass of 1 HexNAc, 1 Hex, and 1 NeuAc monosaccharide, and it is likely that this is the composition of the oligosaccharide chain attached to Thr[146]. Considerably less carbohydrate was associated with the glycopeptides in fraction 9 (Table 3), the mass of the most intense ESMS peak being consistent with the presence of a single $(NeuAc)_2$-Hex-HexNAc oligosaccharide chain. The glycosylation site in this glycopeptide was not determined by Edman sequencing due to excessive sample heterogeneity but is likely to be Thr[145] or Ser[152]. Since the total carbohydrate content of oligosaccharide chains attached to these two amino acids is 3 HexNAc, 3 Hex, and 3 NeuAc monosaccharides (based on the total carbohydrate content of the major glycoforms of fractions 5 and 6, their known glycosylation sites, and the composition of the chains attached to the other amino acids), the composition of the other chain is NeuAc−$(Hex)_2$-$(HexNAc)_2$, although it was not possible to determine whether this chain is attached to Thr[145] and the $(NeuAc)_2$-Hex-HexNAc chain to Ser[152], or the reverse is the case. Ser[144] is only partly glycosylated in fraction 5 (Table 4) and is therefore likely to be associated with a minor additional Hex-HexNAc or NeuAc−Hex-HexNAc carbohydrate chain responsible for the 7152- and 7443-Da components in fraction 5 (Table 3).

From the information derived from the peptide results, it is possible to infer in part the glycosylation sites and carbohydrate responsible for the major glycoforms of IGFBP-6 (Table 5). In the minor glycoform of mass 26 469 Da, all five sites are glycosylated. All sites except the variably glycosylated Ser[144] are glycosylated in the major IGFBP-6 glycoform with mass 26 110 Da (containing 16 monosaccharides; Table 1). In the glycoform with mass 25 451 Da (containing 13 monosaccharides, Table 1), Thr[126], Thr[145], and Ser[152] are glycosylated. The difference in mass between the 25 451 and 24 429 Da glycoforms is 1022 Da, corresponding in mass to the larger oligosaccharide attached to either Thr[145] or Ser[152]; therefore, the glycosylation sites in the 24 429-Da glycoform are Thr[126] and whichever of Thr[145] or Ser[152] does not have the larger oligosaccharide chain attached (Table 5).

Partial Characterization of the Disulfide Bond Linkages of IGFBP-6. Since disulfide bonds of IGFBP-6 were not reduced by the conditions used prior to trypsin digestion, the reverse-phase HPLC-purified peptides included peptide complexes linked by native disulfide bonds, allowing additional information to be obtained regarding the localization of disulfide bond connections in IGFBP-6. Thus, the amino-terminal peptide fragment 28−65 contains six cysteine residues (cysteines 29, 32, 40, 44, 57, and 63) involved in three internal disulfide linkages of undetermined connectivity. Peptide fragment 67−95, with three cysteine residues (cysteines 71, 78, and 84), was found to be disulfide linked to

Table 3: ESMS Masses of Tryptic Glycopeptides of IGFBP-6[a]

HPLC fraction	obsd charge states	av molecular mass obsd	av molecular mass calcd	IGFBP-6 sequence assignment and carbohydrate composition
				(119−134) ESKPQAG**T**ARPQDVNR
1*	2−4	1753.9 ± 0.5	1753.9	HexNAc(**0**);Hex(**0**);NeuAc(**0**)
1*	2−4	1957.0 ± 0.5	1957.1	HexNAc(**1**);Hex(**0**);NeuAc(**0**)
1*	2−4	2119.4 ± 0.4	2119.2	HexNAc(**1**);Hex(**1**);NeuAc(**0**)
1*	2−4	2410.5 ± 0.5	2410.5	HexNAc(**1**);Hex(**1**);NeuAc(**1**)
1	2−4	**2701.7 ± 0.4**	2701.7	HexNAc(**1**);Hex(**1**);NeuAc(**2**)
				(119−134) ESKPQAG**T**ARPQDVNR
2*	2−4	1754.1 ± 0.4	1753.9	HexNAc(**0**);Hex(**0**);NeuAc(**0**)
2*	2−4	1957.0 ± 0.4	1957.1	HexNAc(**1**);Hex(**0**);NeuAc(**0**)
2*	2−4	2119.2 ± 0.4	2119.2	HexNAc(**1**);Hex(**1**);NeuAc(**0**)
2	2−4	**2410.5 ± 0.5**	2410.5	HexNAc(**1**);Hex(**1**);NeuAc(**1**)
				(140−165) NPGTSTTPSQPNS−TEMGPCRR
				(181−193) GAQTLYVPNCDHR
9*	3−6	4160.5 ± 0.6	4160.5	HexNAc(**0**);Hex(**0**);NeuAc(**0**)
9	4, 5	4817.6 ± 0.8	4817.1	HexNAc(**1**);Hex(**1**);NeuAc(**1**)
9	4, 5	**5108.5 ± 0.7**	5108.3	HexNAc(**1**);Hex(**1**);NeuAc(**2**)
				(140−165) NPGTS**T**TPSQPN**S**−TEMGPCRR
				(181−193) GAQTLYVPNCDHR
6	4, 5	5473.4 ± 0.8	5473.7	HexNAc(**2**);Hex(**2**);NeuAc(**2**)
6	4, 5	5764.6 ± 0.7	5764.9	HexNAc(**2**);Hex(**2**);NeuAc(**3**)
6	4, 5	**6131.0 ± 0.9**	6130.2	HexNAc(**3**);Hex(**3**);NeuAc(**3**)
6	4, 5	6422.0 ± 0.8	6421.5	HexNAc(**3**);Hex(**3**);NeuAc(**4**)
				(140−165) NPGT**STT**PSQPN**S**−TEMGPCRR
				(181−193) GAQTLYVPNCDHR
5, 7b	5, 6	6495.6 ± 0.8	6495.5	HexNAc(**4**);Hex(**4**);NeuAc(**3**)
5, 7b	4−6	**6786.9 ± 0.7**	6786.8	HexNAc(**4**);Hex(**4**);NeuAc(**4**)
5, 7b	4−6	7078.6 ± 0.7	7078.1	HexNAc(**4**);Hex(**4**);NeuAc(**5**)
5, 7b	5, 6	7152.0 ± 0.9	7152.1	HexNAc(**5**);Hex(**5**);NeuAc(**4**)
5, 7b	5, 6	7369.4 ± 0.8	7369.3	HexNAc(**4**);Hex(**4**);NeuAc(**6**)
5, 7b	5, 6	7443.4 ± 0.9	7443.4	HexNAc(**5**);Hex(**5**);NeuAc(**5**)

[a] ESMS and calculated masses were determined as for Table 2, using the indicated HPLC fractions. The masses of the most intense peaks observed under normal operating conditions are shown in bold. Masses beside fractions marked with an asterisk (*) were observed using high source collisional fragmentation conditions (V_{OR} = 150−200 V) but were weak or absent under normal operating conditions (V_{OR} = 25 V). Fraction 7 contained minor glycopeptide components (7b) with the same masses, charge states, and relative intensities as the glycopeptides in fraction 5; observed masses for these fractions have been averaged and are presented together. Fractions 9, 6, 5, and 7b contained the same disulfide-linked peptides, as indicated, with differing carbohydrate content as indicated in bold. Also shown in bold are glycosylation sites, where determined, based on Edman N-terminal sequencing data of fractions 1, 5, 6, and 7. Ser^{144} was apparently glycosylated in some fraction 5 glycopeptides, but not those in fraction 6 (Table 4). HexNAc = *N*-acetylhexosamine, Hex = hexose, and NeuAc = *N*-acetylneuraminic acid.

peptide fragment 104−108, containing Cys^{104}. Peptide fragment 140−165 was disulfide-linked to peptide fragment 181−193, showing unambiguously that Cys^{163} is linked to Cys^{190}. Finally, peptide fragments 200−202 and 221−239, each containing one cysteine residue (Cys^{201} and Cys^{234}, respectively), were disulfide-linked to peptide fragment 209−217, which contains two cysteine residues (Cys^{212} and Cys^{214}). It is therefore apparent that the native disulfide linkages in IGFBP-6 occur between cysteines that are close together in the primary sequence, forming four sequential "subdomains" with internal disulfide linkages (Figure 8). Notably, none of the inferred disulfide linkages connect the N-terminal region to the C-terminal region.

O-Glycosylation Protects IGFBP-6 from Proteolytic Degradation. A proposed role of glycosylation is protection of the protein from proteolysis (*15*). To investigate whether glycosylation of IGFBP-6 fulfilled this function, chymotrypsin was incubated with nonglycosylated or glycosylated IGFBP-6. Nonglycosylated IGFBP-6 was cleaved to a greater extent than glycosylated IGFBP-6 (Figure 9). For example, 1 ng of chymotrypsin completely cleaved intact nonglycosylated IGFBP-6 to lower molecular weight forms, whereas 10 ng of chymotrypsin was required to completely cleave intact glycosylated IGFBP-6. Glycosylation of IGFBP-6 similarly inhibited proteolysis by trypsin (results not shown).

IGFBP-6 Is Not Phosphorylated. IGFBP-1, -3 and -5 may be serine-phosphorylated, and it has been suggested that IGFBP-6 has numerous potential Ser/Thr acceptor sites, including Ser^{52}, Thr^{75}, Thr^{126}, Thr^{146}, and Thr^{225} (*28*). However, following immunoprecipitation of IGFBP-6 from media conditioned by stably transfected CHO cells incubated with [^{32}P]orthophosphate, no metabolically labeled IGFBP-6 could be identified, suggesting that IGFBP-6 is not phosphorylated in this cell system (results not shown). By ESMS, no evidence of mass increases of 80 Da (or multiples thereof), indicative of phosphorylation, was observed for intact IGFBP-6 or tryptic fragments, including those containing potential phosphorylation sites, thereby confirming a lack of significant phosphorylation.

DISCUSSION

The present study demonstrates that recombinant human IGFBP-6 is an O-linked glycoprotein displaying considerable heterogeneity with respect to extent of glycosylation and minor heterogeneity at the N- and C-termini of its amino acid sequence. Although it was suggested initially that the first amino acid of mature human IGFBP-6 was Ala^{25} (*6*), the predominant form found in the present study commenced at Arg^{28}. This is consistent with sequences of IGFBP-6 purified from human serum (*31*), cerebrospinal fluid (*11*), and media conditioned by transformed human fibroblasts (*7*),

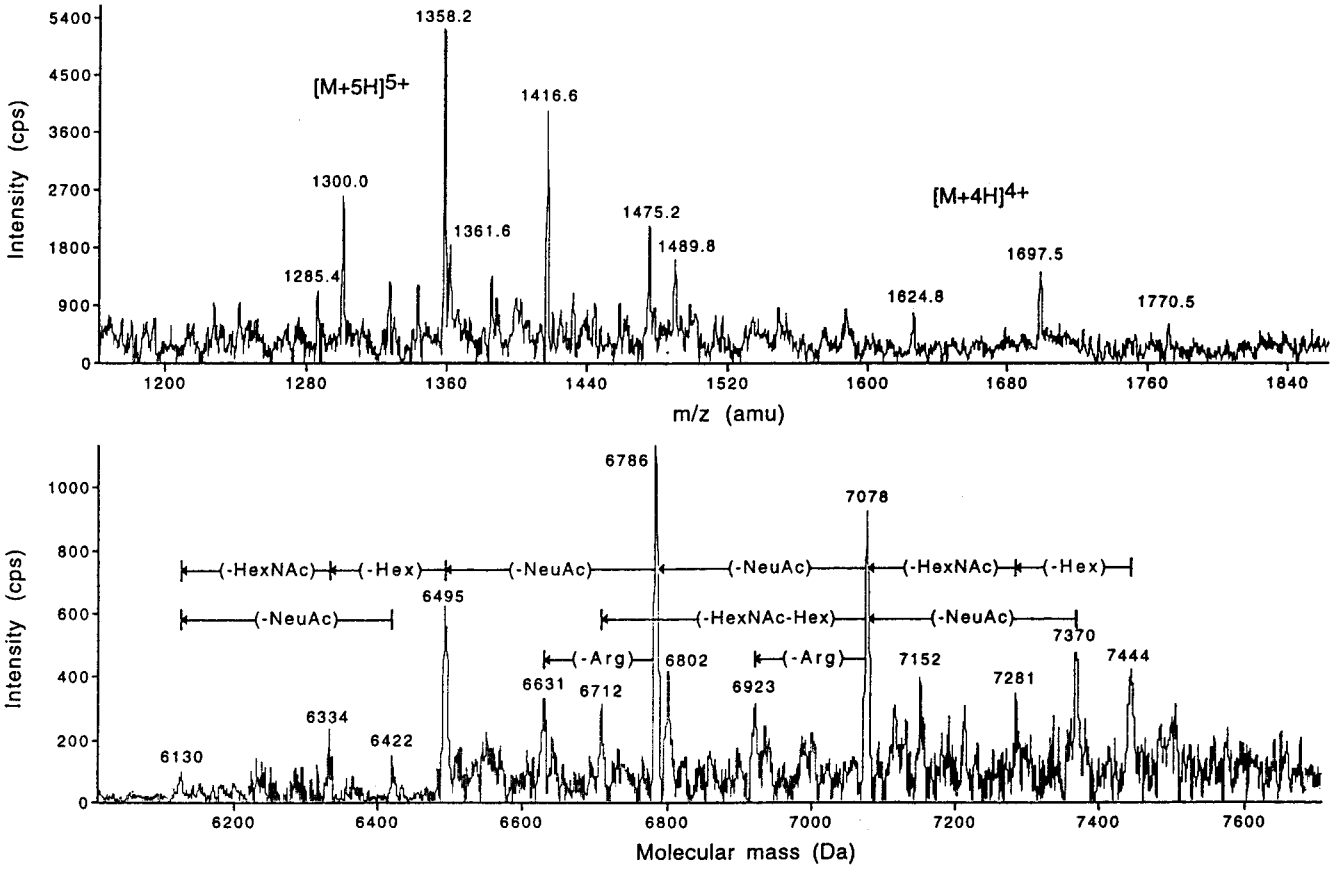

FIGURE 7: ESMS mass spectra of glycopeptides in HPLC fraction 5. Approximately 0.3 pmol of glycopeptide contained in HPLC fraction 5 (Figure 5) was analyzed at a concentration of approximately 0.1 μM by ESMS as described in Experimental Procedures. Upper spectrum: raw mass spectrum showing multiple components with charge states of 4^+ and 5^+. Lower spectrum: transform from *m*/*z* to true mass scale, with baseline subtraction and smoothing (three times with three point moving average, weighted 1:2:1). Peaks are labeled with mass in daltons and inferred differences in carbohydrate content as indicated (see Table 3).

Table 4: Quantitative Analysis of Glycopeptide Edman Sequencing Data[a]

HPLC fraction/Edman cycle no. (expected amino acid)	expected yield from			total expected yield	obsd yield
	NPGTS...	GAQTL...	SLPGS...		
fraction 5/cycle 4 (Thr)	2.8	2.9	0	5.7	5.6
fraction 6/cycle 4 (Thr)	3.4	3.2	0	6.6	6.5
fraction 7/cycle 4 (Thr)	4.6	6.7	0	11.3	11.4
fraction 5/cycle 5 (Ser)	1.4	0	0.3	1.7	0.9
fraction 6/cycle 5 (Ser)	1.7	0	1.1	2.8	2.5
fraction 7/cycle 5 (Ser)	2.3	0	7.2	9.5	10.1

[a] To assess possible glycosylation of Thr^{143}, Ser^{144}, and Thr^{184} of IGFBP-6, Edman sequencing data was quantitated with respect to cycles 4 and 5 for HPLC fractions 5, 6, and 7 (Figure 5; Table 2), each of which contained glycovariants of the same peptide complex composed of tryptic fragment 140−165 (NPGTS) disulfide-linked to tryptic fragment 181−193 (GACTL), as well as varying amounts of nonglycosylated tryptic fragment 221−239 (SLPGS) and other components not containing Ser or Thr (Table 2). Observed PTH-amino acid yields (in pmol) were background corrected using data from the previous cycle where clear of interference, or from several previous cycles where necessary. Expected yields (in pmol) for Thr and Ser in the indicated Edman cycles were calculated using maximal expected recoveries of 40% and 20%, respectively, typically obtained with the Hewlett-Packard G1005A sequencer (*48*), multiplied by the estimated yield from each peptide. The latter was calculated by averaging observed yields in each cycle for the amino acid appropriate to the particular peptide sequence, after background correction, correction for expected recovery, and extrapolation to the appropriate cycle (4 or 5) using a repetitive yield of 97% (*48*). The observed Thr and Ser yields are those expected if Thr^{143}, Ser^{144}, and Thr^{184} were not glycosylated, except for fraction 5, where the observed yield was lower than expected, suggesting partial Ser^{144} glycosylation.

human keratinocytes (*35*), and prostate cancer cells (*36*). A lesser proportion (approximately 20%) of IGFBP-6 in the present study commenced at Leu^{26}. Similar microheterogeneity of the N-terminus was also identified in IGFBP-6 purified from human serum (*31*) and media conditioned by human keratinocytes (*35*). Additionally, the predominant form of glycosylated recombinant human IGFBP-6 in the present study lacked the carboxyl-terminal Gly residue. Five O-glycosylation sites (Ser^{126}, Ser^{144}, Thr^{145}, Thr^{146}, and Ser^{152}) were identified in human IGFBP-6. Of these, Thr^{126} appears to be glycosylated in all of the major glycoforms, whereas the other sites are variably glycosylated.

By inference from masses of different glycopeptide and glycoprotein glycoforms, IGFBP-6 may contain ~7−19 monosaccharide units, with the major glycoprotein components observed by ESMS containing 8−16 monosaccharide units. We have previously shown that IGFBP-6 purified from cerebrospinal fluid has an apparent carbohydrate content

Table 5: Glycosylated Amino Acids and Carbohydrate Chain Composition in Predominant Glycoforms of IGFBP-6[a]

glycoprotein mass	Thr^{126}	Ser^{144}	Thr^{146}	Thr^{145}/Ser^{152}	fraction no. (mass)
24 429	NHSS	n-g	n-g	NHSS	9 (5108.5)
25 451	NHSS	n-g	n-g	NHNHS/NHSS	6 (6131.0)
26 110	NHSS	n-g	NHS	NHNHS/NHSS	5, 7b (6786.9)
26 469	NHSS	NH	NHS	NHNHS/NHSS	5, 7b (7152.0)

[a] Deduced from the carbohydrate content of glycoforms of IGFBP-6 (Table 1) and tryptic glycopeptides from IGFBP-6 (Table 3), as well as the O-glycosylation sites determined by Edman sequencing. One of either Thr^{145} or Ser^{152} is glycosylated in the 24 429 and 25 451 glycoforms, whereas both are glycosylated in the more massive glycoforms. Additional less abundant glycoforms differing from those shown by one extra or one fewer *N*-acetylneuraminic acid and for which there is mass evidence (Table 3) are not shown. Shown in the right-hand column are HPLC fractions containing tryptic peptide 140−165 with the same glycosylation of Ser^{144}, Thr^{145}, Thr^{146}, and Ser^{152} (Table 3) as that deduced for the glycoprotein in the same row. Masses are shown in daltons (Da). n−g = nonglycosylated, N = *N*-acetylhexosamine, H =hexose, and S = *N*-acetylneuraminic acid (sialic acid).

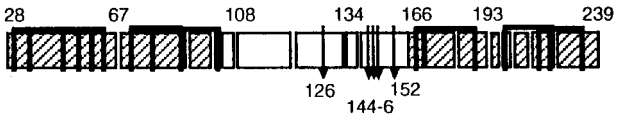

FIGURE 8: Schematic diagram of IGFBP-6 structure. Numbers at the top represent amino acid residues based on the sequence IBP6_HUMAN (accession number P24592) in the SWISS-PROT database as shown in Figure 1. Tryptic fragments are shown as boxes. The conserved N- and C-terminal domains are shaded, and the nonconserved central domain is not. Cysteine residues are marked as vertical lines with potential linkages shown as horizontal lines. O-Glycosylation sites are indicated by arrows and numbered below the diagram.

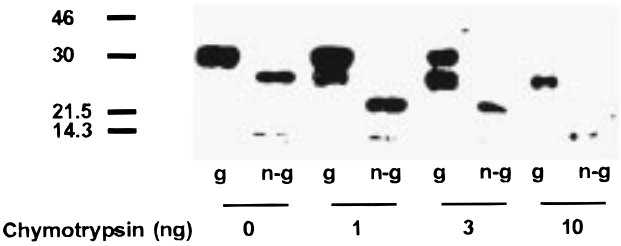

FIGURE 9: Cleavage of glycosylated (g) and nonglycosylated (n-g) IGFBP-6 by chymotrypsin. Glycosylated or nonglycosylated IGFBP-6 (200 ng) was incubated with 0−10 ng of chymotrypsin for 2 h at 37 °C. Samples were then separated by sodium dodecyl sulfate−12% polyacrylamide gel electrophoresis under nonreducing conditions and transferred to a nitrocellulose membrane. Immunoblotting with an anti-human IGFBP-6 polyclonal antiserum was then performed using enhanced chemiluminescent detection. Migration of molecular mass markers (kDa) under reducing conditions is shown on the left.

of 20−32 monosaccharide units (*11*), based on high-performance anion-exchange chromatography (HPAEC) with pulsed amperometric detection of hydrolyzed carbohydrate. It is possible that this apparent discrepancy may to some extent reflect true differences in the degree of glycosylation of IGFBP-6 from these other sources and CHO cells. For example, CHO cells do not express α(2,6)sialyltransferase (ST6N) (*37*) or α(1,3)fucosyltransferase (*38*), although the relevance of these specific enzymes to O-glycosylation of IGFBP-6 has not been demonstrated. However, recombinant human IGFBP-6 expressed in CHO cells migrates on SDS−PAGE analysis with similar apparent molecular mass to that

```
Human    122 PQAGTARPQDVNRRDQQRNPGTSTTPSQPNSAGVQ 156
Bovine    35 PQAGTARSQDVNRRDQQRNSGTSTTPSRSNSGGVQ  69
Mouse    121 PQGGASRSRDTNHRDRQKNPRTSAAPIRPNP--VQ 150
Rat      112 PHGGASRPRD---RDRQKNPRTSAAPIRPSP--VQ 138
```

FIGURE 10: Comparison of amino acid sequences of IGFBP-6 from various nonhuman species with the region surrounding the O-linked glycosylation sites of human IGFBP-6. Actual glycosylation sites of human IGFBP-6 and corresponding amino acids in IGFBP-6 from other species are shown in bold. Amino acid numbering is based on the IGFBP-6 entries IBP6_HUMAN (accession number P24592), IBP6_BOVIN (accession number Q05718), IBP6_MOUSE (accession number P47880) and IBP6_RAT (accession number P35572) in the SWISS-PROT database.

expressed in COS-7 cells (results not shown), and the latter has a carbohydrate content similar to that of native IGFBP-6 from cerebrospinal fluid (*29*). Therefore, it is unlikely that there is a large difference in carbohydrate content between IGFBP-6 from these sources. The apparent discrepancy between the ESMS and HPAEC estimates of carbohydrate content could be due to problems of quantitation with either or both of these techniques. With ESMS, there could be reduced sensitivity toward more heavily glycosylated glycoproteins or some loss of terminal sialic acid units due to collisions in the ion source, although the latter process did not appear to be a major problem at the orifice voltages (50−60 V) used in analyzing IGFBP-6, based on the minimal change in mass spectra observed with increased ion source orifice voltage (results not shown). The relatively small amount of material analyzed in previous studies by HPAEC may have decreased the accuracy of quantitation by that method.

Rat and mouse IGFBP-6 are also O-glycosylated, although apparently to a lesser extent than human IGFBP-6 (*39, 40*). Comparison of these sequences around the sites of O-glycosylation of human IGFBP-6 reveals only limited homology (Figure 10). In particular, four of the five O-glycosylated amino acids in human IGFBP-6 are not conserved in the rat or mouse sequences, which may underlie the lower carbohydrate content of the corresponding proteins. In contrast, all five glycosylated amino acids in the human sequence are conserved in bovine IGFBP-6, and it would therefore be of considerable interest to ascertain the extent of glycosylation of IGFBP-6 from that species. A tryptic fragment of porcine IGFBP-6 also contains the nonglycosylated GAQTLY sequence found in the human protein, and sequencing of the porcine fragment demonstrated that the Thr residue is also not glycosylated in that species (*6*).

In contrast to N-glycosylation, there is no consensus amino acid sequence for O-glycosylation (*17*). Nevertheless, a number of prediction algorithms for O-glycosylation have been constructed by using the available data in different ways. NetOGlyc is a prediction algorithm based on neural network analysis of a large number of empirically determined O-linked glycosylation sites in 131 glycoproteins (*21, 41*). Version 2.0 of this algorithm applied to human IGFBP-6 correctly predicted that Ser^{144}, Thr^{145} and Thr^{146} would be glycosylated. However, the same algorithm incorrectly predicted that Thr^{126} and Ser^{152} would not be glycosylated. Of the nonglycosylated amino acids of IGFBP-6, the NetOGlyc program predicted that Thr^{143}, Ser^{148}, Ser^{232}, and Thr^{236} would be glycosylated. Therefore, this program accurately predicted 60% of the determined glycosylation sites and 76% of the determined nonglycosylated sites

[excluding the two Ser residues not identified in tryptic fragments and of undetermined glycosylation status, although it should be noted that, according to the four prediction algorithms presented here, these two serines (Ser^{203}, Ser^{204}) are unlikely to be glycosylated].

An increased frequency of proline residues at positions −1 and +3 relative to single O-glycosylation sites has been noted previously (*18*). A proline residue is found in IGFBP-6 at the +3 position relative to the Thr^{126} and Ser^{144} glycosylation sites, but this method failed to predict the other sites. Additionally, three nonglycosylated residues were predicted to be glycosylated (Ser^{148}, Ser^{232}, Thr^{236}). This algorithm therefore accurately predicted 40% of the determined glycosylated residues and 82% of determined nonglycosylated residues.

Elhammer et al. (*19*) found that proline, serine, and threonine residues were found with increased frequency in positions −4 to +4 surrounding an O-glycosylated amino acid, and they made use of this observation to create an algorithm for predicting O-glycosylation sites. When applied to IGFBP-6, this algorithm correctly predicted glycosylation of Thr^{126}, Ser^{144}, Thr^{145}, and Thr^{146}. However, it predicted that Ser^{152} would not be glycosylated. Additionally, it predicted that Thr^{143}, Ser^{148}, Ser^{232}, Ser^{233}, Thr^{236}, Ser^{238}, and Ser^{239} would be glycosylated. This algorithm therefore accurately predicted 80% of the determined glycosylated residues and 59% of determined nonglycosylated residues.

A sequence-coupled vector-projection model based on the frequency distribution of the nine amino acids surrounding the O-glycosylation site (*20*) correctly predicted each of the sites identified in IGFBP-6. However, it additionally predicted that Ser^{52}, Ser^{120}, Ser^{148}, Thr^{176}, Ser^{221}, and Ser^{225} would be glycosylated. This algorithm therefore accurately predicted 100% of the determined glycosylated residues and 65% of determined nonglycosylated residues. Prediction methods for O-glycosylation sites are therefore only moderately successful. Further empirical determinations of O-linked glycosylation sites are clearly necessary to enable refinement of these algorithms.

Glycosylation of proteins has been implicated in the modulation of a wide range of their properties (*15, 16*). In the present study, nonglycosylated IGFBP-6 was found to be more susceptible to cleavage by trypsin and chymotrypsin than glycosylated IGFBP-6, indicating that the carbohydrate component imparts to the protein some degree of protection from proteolytic cleavage (*42*). Since proteolysis is a potentially important means by which IGFs are released from complexes with IGFBPs (*43*), this may be a critical function of the glycosylation of IGFBP-6. It is noteworthy that there are relatively few reports of proteases for IGFBP-6 compared with the other IGFBPs, although this may also reflect that IGFBP-6 has been studied to a lesser extent. Glycosylation of proteins has also been implicated in prolonging their circulating half-lives, partially due to inhibition of proteolysis and partially due to decreased clearance by specific receptors (*15*). This potential role has not been studied for IGFBP-6. Glycosylation of some proteins alters the affinity of ligand binding by these proteins (*15*). However, enzymatic deglycosylation of human IGFBP-6 has no effect on IGF binding (*11*), so it would appear that glycosylation has no significant role in IGF-II binding by IGFBP-6. Another role for the carbohydrate component of glycoproteins is alteration of interactions with other biomolecules, including extracellular matrix components (*16*). IGFBP-6 is not significantly cell-associated (*29, 44, 45*), and glycosylation may therefore have an as yet unstudied involvement in the lack of pericellular localization of IGFBP-6. The potential significance of this possibility lies in the role that cell association of IGFBPs may play in determining whether they potentiate or inhibit IGF actions (*3*).

All cysteine residues of IGFBP-3 are involved in disulfide bonds (*46*). Natural fragments of IGFBPs containing either their C-terminal or N-terminal domains have been described (*2*), and consequently it has been assumed that disulfide bonding occurs within each domain rather than between domains. A model of IGFBP structure has been proposed in which the conserved C-terminal and N-terminal domains, both of which are involved in IGF binding, are joined by a flexible nonconserved domain (*47*). The data from the present study confirm and extend the above observations, as we have been able to define four internally disulfide-linked subdomains of human IGFBP-6, two within each of the C-terminal and N-terminal domains.

In conclusion, a considerable amount of information regarding the structure of IGFBP-6 has been obtained by using experimental approaches that require relatively small amounts of glycoprotein and do not depend on specialized mass spectrometric techniques such as on-line LC/MS or LC/MS/MS (*34*), although the latter are clearly invaluable in structural analyses of posttranslationally modified proteins such as glycoproteins (*26, 34*). The approach used here may be particularly useful for the study of glycoproteins that are available in limited quantities. The results indicate that the O-glycosylation sites of IGFBP-6 are clustered in the nonconserved midregion of the molecule, whereas disulfide linkages may constrain the conserved N- and C-termini into independent domains. Phosphorylation sites of IGFBP-1 and -3 are also found in the nonconserved midregion of these molecules (*28*), as are the N-glycosylation sites of IGFBP -3 and -4 (*2*). The conserved regions of the IGFBPs are thought to contain the major determinants of high-affinity IGF binding, whereas posttranslational modifications of the nonconserved middle domains have a relatively minor role; in support of this contention, glycosylation of IGFBP-6 has no effect on IGF-II binding. However, posttranslational modification of the nonconserved region of the IGFBPs may underlie some of their observed properties, such as the relative resistance of glycosylated IGFBP-6 to proteolysis. Further experiments are required to comprehensively define the effects of the attached oligosaccharide chains on the properties and actions of this glycoprotein.

ACKNOWLEDGMENT

We are grateful to Rosemary Condron (School of Biochemistry, La Trobe University, Bundoora, Australia) for Edman N-terminal sequencing.

REFERENCES

1. Humbel, R. E. (1990) *Eur. J. Biochem. 190*, 445−462.
2. Rechler, M. M. (1993) *Vitam. Horm. 47*, 1−114.
3. Bach, L. A., and Rechler, M. M. (1995) *Diabetes Rev. 3*, 38−61.
4. Jones, J. I., and Clemmons, D. R. (1995) *Endocr. Rev. 16*, 3−34.

5. Swisshelm, K., Ryan, K., Tsuchiya, K., and Sager, R. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 4472−4476.
6. Shimasaki, S., Gao, L., Shimonaka, M., and Ling, N. (1991) *Mol. Endocrinol. 5*, 938−948.
7. Martin, J. L., Willetts, K. E., and Baxter, R. C. (1990) *J. Biol. Chem. 265*, 4124−4130.
8. Forbes, B., Ballard, F. J., and Wallace, J. C. (1990) *J. Endocrinol. 126*, 497−506.
9. Roghani, M., Lassarre, C., Zapf, J., Povoa, G., and Binoux, M. (1991) *J. Clin. Endocrinol. Metab. 73*, 658−666.
10. Bach, L. A., Hsieh, S., Sakano, K., Fujiwara, H., Perdue, J. F., and Rechler, M. M. (1993) *J. Biol. Chem. 268*, 9246−9254.
11. Bach, L. A., Thotakura, N. R., and Rechler, M. M. (1992) *Biochem. Biophys. Res. Commun. 186*, 301−307.
12. Zapf, J., Born, W., Chang, J.-Y., James, P., Froesch, E. R., and Fischer, J. A. (1988) *Biochem. Biophys. Res. Commun. 156*, 1187−1194.
13. Ceda, G. P., Fielder, P. J., Henzel, W. J., Louie, A., Donovan, S. M., Hoffman, A. R., and Rosenfeld, R. G. (1991) *Endocrinology 128*, 2815−2824.
14. Conover, C. A., and Kiefer, M. C. (1993) *J. Clin. Endocrinol. Metab. 76*, 1153−1159.
15. Varki, A. (1993) *Glycobiology 3*, 97−130.
16. Parekh, R. B. (1991) *Curr. Opin. Struct. Biol. 1*, 750−754.
17. Pan, Y. T., and Elbein, A. D. (1990) in *Progress in Drug Research* (Jucker, E., Ed.) pp 163−207, Birkhauser Verlag, Basel.
18. Wilson, I. B. H., Gavel, Y., and von Heijne, G. (1991) *Biochem. J. 275*, 529−534.
19. Elhammer, A. P., Poorman, R. A., Brown, E., Maggiora, L. L., Hoogerheide, J. G., and Kedzy, F. J. (1993) *J. Biol. Chem. 268*, 10029−10038.
20. Chou, K.-C. (1995) *Protein Sci. 4*, 1365−1383.
21. Hansen, J. E., Lund, O., Rapacki, K., and Brunak, S. (1997) *Nucleic Acids Res. 25*, 278−282.
22. Smith, R. D., Loo, J. A., Edmonds, C. G., Barinaga, C. J., and Udseth, H. R. (1990) *Anal. Chem. 62*, 882−899.
23. Neumann, G. M., Condron, R., Thomas, I., and Polya, G. M. (1996) *Biochim. Biophys. Acta 1295*, 23−33.
24. Neumann, G. M., Condron, R., and Polya, G. M. (1996) *Int. J. Pept. Protein Res. 47*, 437−446.
25. Neumann, G. M., Condron, R., and Polya, G. M. (1996) *Biochim. Biophys. Acta 1298*, 223−240.
26. Strub, J. M., Sorokine, O., Van Dorsselaer, A., Aunis, D., and Metz-Boutigue, M.-H. (1997) *J. Biol. Chem. 272*, 11928−11936.
27. Linsley, K. B., Chan, S.-Y., Chan, S., Reinhold, B. B., Lisi, P. J., and Reinhold, V. N. (1994) *Anal. Biochem. 219*, 207−217.
28. Coverley, J. A., and Baxter, R. C. (1997) *Mol. Cell. Endocrinol. 128*, 1−5.
29. Bach, L. A., Hsieh, S., Brown, A. L., and Rechler, M. M. (1994) *Endocrinology 135*, 2168−2176.
30. Bach, L. A., and Rechler, M. M. (1996) *Biochim. Biophys. Acta 1313*, 79−88.
31. Zapf, J., Kiefer, M., Merryweather, J., Masiarz, F., Bauer, D., Born, W., Fischer, J. A., and Froesch, E. R. (1990) *J. Biol. Chem. 265*, 14892−14898.
32. Jones, J. I., Busby, W. H., Jr., Wright, G., Smith, C. E., Kimack, N. M., and Clemmons, D. R. (1993) *J. Biol. Chem. 268*, 1125−1131.
33. Gooley, A. A., and Williams, K. L. (1997) *Nature 385*, 557−559.
34. Huddleston, M. J., Bean, M. F., and Carr, S. A. (1993) *Anal. Chem. 65*, 877−884.
35. Kato, M., Ishizaki, A., Hellman, U., Wernstedt, C., Kyogoku, M., Miyazono, K., Heldin, C.-H., and Funa, K. (1995) *J. Biol. Chem. 270*, 12373−12379.
36. Srinivasan, N., Edwall, D., Linkhart, T. A., Baylink, D. J., and Mohan, S. (1996) *J. Endocrinol. 149*, 297−303.
37. Svensson, E. C., Soreghan, B., and Paulson, J. C. (1990) *J. Biol. Chem. 265*, 20863−20868.
38. Goochee, C. F., Gramer, M. J., Anderson, D. C., Bahr, J. B., and Rasmussen, J. R. (1991) *Biotechnology 9*, 1347−1355.
39. Bach, L. A., Tseng, L. Y.-H., Swartz, J. E., and Rechler, M. M. (1993) *Endocrinology 133*, 990−995.
40. Claussen, M., Buergisser, D., Schuller, A. G. P., Matzner, U., and Braulke, T. (1995) *Mol. Endocrinol. 9*, 902−912.
41. Hansen, J. E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J. O., Hansen, J.-E. S., and Brunak, S. (1995) *Biochem. J. 308*, 801−813.
42. Bernard, B. A., Yamada, K. M., and Olden, K. (1982) *J. Biol. Chem. 257*, 8549−8554.
43. Rechler, M. M. (1995) in *Molecular Endocrinology: Basic Concepts and Clinical Correlations* (Weintraub, B. D., Ed.) pp 155−181, Raven Press, New York.
44. Booth, B. A., Boes, M., Andress, D. L., Dake, B. L., Kiefer, M. C., Maack, C., Linhardt, R. J., Bar, K., Caldwell, E. E. O., Weiler, J., and Bar, R. S. (1995) *Growth Regul. 5*, 1−17.
45. Bach, L. A., Leeding, K. S., and Leng, S. L. (1997) *J. Endocrinol. 155*, 225−232.
46. Sommer, A., Maack, C. A., Spratt, S. K., Mascarenhas, D., Tressel, T. J., Rhodes, E. T., Lee, R., Roumas, M., Tatsuno, G. P., Flynn, J. A., Gerber, N., Taylor, J., Cudny, H., Nanney, L., Hunt, T. K., and Spencer, E. M. (1991) in *Modern Concepts of Insulin-Like Growth Factors* (Spencer, E. M., Ed.) pp 715−728, Elsevier, New York.
47. Spencer, E. M., and Chan, K. (1995) *Prog. Growth Factor Res. 6*, 209−214.
48. Coligan, J. E., Dunn, B. M., Ploegh, H. L., Speicher, D. W., and Wingfield, P. T. (1997) *Current Protocols in Protein Science*, John Wiley & Sons, New York.

BI972894E