Posttranslational Modifications of Human Inter-α-Inhibitor: Identification of Glycans and Disulfide Bridges in Heavy Chains 1 and 2[†]

Eva H. N. Olsen,[‡] Henrik Rahbek-Nielsen,[§] Ida B. Thøgersen,[‡] Peter Roepstorff,[§] and Jan J. Enghild*[,‡]

Department of Pathology, Duke University Medical Center, Durham, North Carolina 27710, and Department of Molecular Biology, University of Odense, Odense, Denmark

Received May 15, 1997; Revised Manuscript Received September 29, 1997[⊗]

ABSTRACT: Inter-α-inhibitor (IαI) is a serine proteinase inhibitor found in high concentrations in human plasma. The protein is composed of a light inhibitory chain called bikunin and two heavy chains of unknown function. The three polypeptide chains are covalently assembled via a carbohydrate cross-link [Enghild, J. J., Salvesen, G., Hefta, S. A., Thøgersen, I. B., Rutherfurd, S., & Pizzo, S. V. (1991) *J. Biol. Chem. 266*, 747−751]. The aim of this study was to complete the primary structure by characterizing additional covalent posttranslational modifications of the heavy chains. Analysis revealed three N-linked oligosaccharides located on Asn^{251} and Asn^{554} of heavy chain 1 and on Asn^{64} of heavy chain 2: all these were complex biantennary structures composed of (*Asn*)-$GlcNAc_2$-Man-(Man-GlcNAc-Gal-SA)$_2$. In addition, the IαI heavy chains carried several O-linked glycans located on Thr^{619} of heavy chain 1 and a cluster of four O-linked oligosaccharides on Thr^{612}, Ser^{619}, Thr^{621}, and Thr^{637} of heavy chain 2. The oligosaccharides were short (*Ser/Thr*)-GalNAc-Gal-SA trisaccharides. The IαI heavy chains contain nine Cys residues, of which eight are involved in disulfide bridges. The unpaired Cys residue residing on heavy chain 1, Cys^{26}, appears to be modified by dihexosylation. The other Cys residues exclusively form intrachain disulfide bridges. In heavy chain 1 the two disulfide bonds are formed between Cys^{210} and Cys^{213} and between Cys^{234} and Cys^{506}, and in heavy chain 2, between Cys^{207} and Cys^{210} and between Cys^{596} and Cys^{597}. Interestingly, three of these four disulfides are formed between Cys residues that are either adjacent or only two amino acid residues apart.

Bikunin is a tandem kunitz-type proteinase inhibitor found in blood and other body fluids of humans (*1*, *2*). In plasma, the protein is associated with one or two distinct but homologous heavy chains (HC1, HC2, or HC3).[1] Three combinations of bikunin and heavy chains have been identified: (i) the predominant form, inter-α-inhibitor (IαI), is composed of HC1, HC2, and bikunin; (ii) pre-α-inhibitor (PαI) is composed of HC3 and bikunin; and (iii) HC2/bikunin is composed of bikunin and HC2 (*3*) (for recent review see ref *4*). Although the bikunin proteins all contain more than one subunit, they resist dissociation in reduced SDS−polyacrylamide gel electrophoresis. This unusual stability is due to a recently discovered protein cross-link that covalently joins the subunits of the bikunin proteins through a chondroitin 4-sulfate glycosaminoglycan (GAG) chain (*5*, *6*). The GAG chain originates from Ser^{10} of bikunin, and the heavy chains are covalently linked to it via an ester bond between the α-carbonyl of their C-terminal Asp residues and the carbon-6 of an internal *N*-acetylgalactosamine of the chondroitin 4-sulfate chain. We have named this unique cross-link a protein−glycosaminoglycan−protein cross-link (PGP cross-link) (*5*, *6*).

The complete cDNA sequences as well as the chromosomal location of the genes encoding HC1, HC2, HC3, and bikunin have been determined (*7*−*12*). The heavy-chain cDNAs encode a 30-kDa C-terminal extension, not present in the mature proteins, as well as putative N-terminal propeptides (*7*, *8*). A homologous plasma protein, called inter-α-trypsin inhibitor family heavy chain related protein (IHRP), was recently discovered (*13*, *14*). However, IHRP differs distinctly from the heavy chains in that it is not bound to bikunin and the 30-kDa C-terminal extension is present in the mature protein (*15*). The bikunin cDNA encodes two tandemly arranged in-frame proteins, α_1-microglobulin (α_1m) and bikunin (*16*). This gene structure has been conserved in both fish (*17*, *18*) and mammals (*11*, *19*−*21*). It is common for coexpressed proteins to be involved in the same physiological process, but as the function of both α_1m and bikunin is unclear, the significance of their coexpression remains ambiguous.

[†] This work was supported by the National Institutes of Health (NIH) Grant HL49542 (J.J.E.) and by the Danish Research Academy (E.H.N.O.).

* To whom correspondence should be addressed at Box 3712, Duke University Medical Center, Durham, NC 27710. Tel. 919-684-2872; Fax 919-684-2920; e-mail enghi001@mc.duke.edu.

[‡] Duke University Medical Center.

[§] University of Odense.

[⊗] Abstract published in *Advance ACS Abstracts,* December 15, 1997.

[1] Abbreviations: 4HCCA, α-cyano-4-hydroxycinnamic acid; α_1m, α_1-microglobulin; Cya, cysteic acid; DTT, dithiothreitol; FACE, fluorophore-assisted carbohydrate electrophoresis; Fuc, fucose; GAG, glycosaminoglycan; Gal, galactose; GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; HC, heavy chain; IαI, inter-α-inhibitor; IHRP, inter-α-trypsin inhibitor family heavy-chain related protein; Man, mannose; PαI, pre-α-Inhibitor; PGP cross-link, protein−glycosaminoglycan−protein cross-link; PNGase F, peptide *N*-glycohydrolase F; SA, sialic acid; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

S0006-2960(97)01137-9 CCC: $15.00
Published on Web 01/06/1998

Recent investigation of the biosynthesis of the bikunin proteins, performed in primary human hepatocytes, determined that the GAG chain is added to Ser^{10} of bikunin before the proteolytically mediated dissociation of α_1m/bikunin (*11*, *22*−*25*). Subsequent to the addition of the GAG chain, the bikunin proteins are assembled intracellularly in a process that appears to involve at least two steps: a proteolytic release of the 30 kDa C-terminus of the heavy chains, followed by the addition of the GAG chain to the α-carbonyl of the newly accessible C-terminal Asp residue. The cleavage of the α_1m/bikunin protein and the formation of the PGP cross-link appear to occur in the latter part of the biosynthetic pathway, most likely within the secretory vesicles (*22*).

No congenital deficiency of IαI family proteins has been reported to date, suggesting that the absence of this gene family is lethal. Human bikunin is known to inhibit trypsin, chymotrypsin, neutrophil elastase, plasmin, and cathepsin G *in vitro*, but since other far more efficient proteinase inhibitors are abundant in plasma, a primary role as a proteinase inhibitor *in vivo* appears unlikely (*2*, *26*). The function of the heavy chains is similarly unclear. Recent reports have suggested that IαI may play a role in cell proliferation and cell migration through binding to hyaluronic acid (*27*−*32*). This activity appears to be unrelated to the proteinase inhibitory activity of bikunin (*33*). Furthermore, Wisniewski and co-workers (*34*) have demonstrated that the arthritis-associated hyaluronic acid binding protein, TSG-6, forms a stable complex with IαI. This activity involves a displacement of IαI HC1 and formation of an apparently covalent complex between TSG-6 and the IαI GAG chain.

To advance understanding of the biological role of the bikunin proteins, a complete structural characterization of their components is important. The posttranslational modifications of the bikunin part of IαI have been thoroughly characterized (*35*, *36*). The glycan on Asn^{45} is a complex biantennary structure; the GAG chain originating from Ser^{10} has been determined to be 17 chondroitin sulfate disaccharide units long (*35*), and three disulfide bridges have been assigned to each of the two kunitz domains based on homology with other kunins (*36*). By contrast, the heavy chains have remained uncharacterized. In this report, we describe the characterization of the posttranslational modifications of HC1 and HC2 in human IαI.

EXPERIMENTAL PROCEDURES

Materials. Jacalin−agarose was obtained from Vector Laboratories Inc., Burlingame, CA. Neuraminidase, *β*-galactosidase, peptide *N*-glycohydrolase F (PNGase F), and O-glycosidase were from Boehringer Mannheim, Frankfurt, Germany. Fluorophore-assisted carbohydrate electrophoresis (FACE) monosaccharide composition kits were purchased from Glyco, Inc., Novato, CA. Sequence-grade porcine trypsin was from Promega. BNPS-skatole [2-(2′-nitrophenylsulfenyl)-3-methyl-3-bromoindolenine] and Reacti-gel (1,1′-carbonyldiimidazole-activated cross-linked 6% agarose) were purchased from Pierce, Inc., Rockford, IL. Sequencing reagents were from Applied Biosystems, Inc., Foster City, CA. Clostripain, concanavalin A, methyl α-D-glycopyranoside, galactose, and all other reagents were from Sigma, St. Louis, MO. Inter-α-inhibitor was purified from human plasma as previously described (*3*). Anhydrotrypsin was prepared according to ref *37* and coupled to Reacti-gel as described by the manufacturer.

SDS−Polyacrylamide Gel Electrophoresis. Polypeptides were separated by SDS−PAGE performed in 5−15% linear gradient gels using the glycine/2-amino-2-methyl-1,3-propanediol/HCl system described by Bury (*38*) and were electroblotted onto poly(vinylidine difluoride) (PVDF) membrane according to Matsudaira (*39*). The membranes were stained briefly with Coomassie Brilliant Blue and extensively washed with water.

Monosaccharide Composition Analysis. The composition of the carbohydrates attached to the IαI heavy chains was determined by FACE analysis. The polypeptide chains were dissociated by gentle treatment with 50 mM NaOH for 15 min at 23 °C (*5*) and separated by SDS−PAGE, followed by transfer to PVDF membrane. The bands containing HC1 and HC2 were excised and cut into small pieces (1 × 1 mm), and their monosaccharide compositions were determined. Briefly, three different hydrolysis reactions were performed on each PVDF-bound heavy chain: (i) 0.1 N trifluoroacetic acid (TFA) at 80 °C for 1 h to obtain sialic acid, (ii) 2 N TFA at 100 °C for 5 h to obtain neutral sugars, and (iii) 4 N HCl at 100 °C for 3 h to obtain amine sugars. The hydrolyzed samples were dried down and resuspended in fluorophore labeling reagents and incubated at 37 °C for 14 h. Samples were loaded on MONO composition gels at 5 °C and run at 30 mA for 1 h. Gel images were obtained using the FACE imaging system.

Isolation of the Noninhibitory Heavy Chains. To simplify the structural analysis, the heavy chains were separated from bikunin. Ten milligrams of IαI was submitted to gentle NaOH treatment as described above (*5*). The dissociated IαI polypeptides were dialyzed into 50 mM Tris-HCl and 100 mM NaCl, pH 7.8, and applied on a preequilibrated anhydrotrypsin−Reacti-gel column (1.5 × 10 cm) at 23 °C. The flowthrough containing the heavy chains was collected, and the retained bikunin was eluted using 100 mM glycine, pH 2.5. The procedure was repeated and the heavy chain preparation was tested for remaining bikunin using the trypsin inhibitor counterstain gel assay (*40*).

Isolation of Cysteine Peptides. The heavy chains were digested with trypsin using an enzyme/substrate ratio of 1:50 (w/w) for 4 h at 37 °C. Upon digestion, the peptides were immediately acidified (0.2% TFA final concentration) and separated by cation-exchange chromatography using a polysulfoethyl aspartamide column (Poly LC, Life Science) connected to a Beckman System Gold 126/166 HPLC system. The column was equilibrated in 5 mM KH_2PO_4 and 25% acetonitrile, and the peptides were eluted using a linear gradient from 0 to 250 mM KCl. Peptides were detected at 220 nm and collected manually. Further separation was achieved by reverse-phase HPLC using a combination of columns including Aquapore RP-300 (2 mm × 220 mm, Brownlee) and Vydac C_{18} (2 mm × 220 mm, Vydac) connected to either the above system or a modified Applied Biosystems 130A HPLC system. The columns were equilibrated with 0.1% TFA and developed with linear gradients to 90% acetonitrile. Aliquots of each peptide pool were treated with performic acid and the cysteic acid content was accessed by amino acid analysis (see below). Cys-containing peptides were further analyzed by Edman degradation and

mass spectrometry. The peptide-containing HC1 Cys^{26} was obtained by BNPS-skatole fragmentation of the heavy chains (*41*), followed by subdigestion with clostripain at a 1:10 (w/w) ratio in 50 mM Tris-HCl, 100 mM NaCl, and 2 mM $CaCl_2$, pH 7.8, at 37 °C for 10 h. Clostripain was activated with 5 mM DTT for 15 min at 23 °C. Before use the clostripain was desalted on a PC 3.2/10 fast desalting column (Pharmacia Biotech. Inc.) connected to a Pharmacia SMART system. The reaction products were separated, identified, and characterized by a combination of the techniques described above.

Isolation of Glycosylated Peptides. Purified heavy chains were digested with trypsin as described above, and the reaction was stopped with 200 μM DCI. The sample was adjusted to 100 mM NaCl, 2 mM $CaCl_2$, 2 mM $MnCl_2$, and 50 mM Tris-HCl, pH 7.8, and applied to a preequilibrated concanavalin A−Sepharose column (1 cm × 5 cm) at 23 °C. N-Glycosylated peptides were eluted with 0.5 M methyl-α-D-glycopyranoside. The peptides were further purified by reverse-phase HPLC as described above, and the major peaks were analyzed by Edman degradation and mass spectrometry.

The concanavalin A flowthrough was adjusted to 175 mM Tris-HCl, pH 7.5, and applied to a preequilibrated jacalin−agarose column (1 cm × 5 cm) from which O-glycosylated peptides was eluted with 0.8 M galactose (*42*, *43*). Separation and characterization of the major peaks were performed as described for the N-glycosylated peptides.

Amino Acid Composition Analysis. Peptides were hydrolyzed in 6 N HCl for 24 h at 110°C (*44*). The hydrolysates were dried in a Speed-Vac concentrator (Savant), and the amino acid compositions were determined in a Beckman 6300 amino acid analyzer. For cysteic acid (Cya) detection, samples were treated with performic acid prior to hydrolysis (*45*).

Amino Acid Sequence Analysis. Automated Edman degradation was carried out in an Applied Biosystems 477A protein sequencer with on-line phenylthiohydantoin analysis using an Applied Biosystems 120A HPLC. Samples were applied to biobrene-treated precycled glass fiber filters.

Enzymatic Carbohydrate Cleavage. In order to determine the oligosaccharide structures, the purified glycopeptides were lyophilized and 150 pmol of peptide was redissolved in 10 mL of 50 mM NH_4HCO_3 and incubated with 3 milliunits of neuraminidase for 12 h at 37 °C. N-Linked carbohydrates were then treated with 0.2 milliunits of β-galactosidase for 12 h at 37 °C followed by a 12 h digestion with 0.1 unit of PNGase F at 37 °C. After neuraminidase treatment, O-linked carbohydrates were digested by 0.1 unit of O-glycosidase for 12 h at 37 °C. Aliquots were removed for mass spectrometry analysis before addition of each new enzyme.

Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry. Mass spectra were recorded in a Bruker reflex TOF mass spectrometer (Bruker-Franzen Analytic) operated in linear mode. A 0.7 mL sample solution and an α-cyano-4-hydroxycinnamic acid solution (4HCCA) (20 mg/mL in 98% acetone) were gently mixed (1:1) on a target precoated with 4HCCA (15 mg/mL in 70% acetonitrile) and dried at ambient temperature. Each MALDI spectrum was accumulated from data collected from 20−60 laser shots. The MALDI spectra were internally calibrated using the protonated dimers of α-cyanocinnamic acid, giving a mass accuracy better than 0.1% (*46*). However, glycosylated peptides often show a slightly lower mass accuracy due to broader peaks.

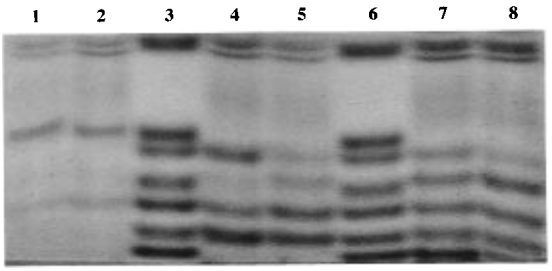

FIGURE 1: Monosaccharide composition of heavy chains 1 and 2. IαI heavy chains 1 and 2 were analyzed by fluorophore-assisted carbohydrate electrophoresis (FACE) in order to get a qualitative impression of their glycosylations. Lane 1, sialic acid hydrolysis of HC1; lane 2, sialic acid hydrolysis of HC2; lane 3, monosaccharide standard; lane 4, neutral sugar hydrolysis of HC1; lane 5, neutral sugar hydrolysis of HC2; lane 6, monosaccharide standard; lane 7, amine sugar hydrolysis of HC1; lane 8, amine sugar hydrolysis of HC2. The monosaccharide standard (lanes 3 and 6) is composed of, from the top, *N*-acetylgalactosamine (GalNAc), sialic acid (SA), mannose (Man), fucose (Fuc), glucose (Glc), galactose (Gal), and *N*-acetylglucosamine (GlcNAc). The purification and separation of the IαI heavy chains as well as the FACE analysis was performed as described in Experimental Procedures.

Mass Spectrometry Analysis of Cys Peptides. The disulfide bridges were reduced by incubation with 33 mM DTT for 10 min. The reaction was stopped by addition of 2 volumes of 2% TFA. The HC2 peptide-containing adjacent Cys residues were further analyzed by alkylation of both native and DTT-reduced peptide. Peptide (20 pmol) in 50 mM NH_4HCO_3 was alkylated with 45 mM 4-vinylpyridine for 10 min, and the reaction was stopped by addition of 2 volumes of 2% TFA. The Cys peptides were analyzed by mass spectrometry as described above.

RESULTS

Monosaccharide Composition of the HC1 and HC2 Glycans. A common posttranslational modification of plasma proteins involves the addition of oligosaccharides (*47*−*49*). The composition of oligosaccharides attached to the individual heavy chain was determined by FACE analysis and provided some initial qualitative information on the monosaccharide compositions (Figure 1). Lanes 1, 4, and 7 show the monosaccharide composition of the HC1 glycans. The HC2 glycans are shown in lanes 2, 5, and 8. The presence on both heavy chains of both *N*-acetylglucosamine (GlcNAc) and mannose (Man) is consistent with the presence of N-glycans, whereas the presence of *N*-acetylgalactosamine (GalNAc) usually signifies O-linked glycans. Thus, each heavy chain contains both N- and O-linked glycans.

Isolation of N-Glycosylated Peptides. All Asn-linked glycans have a core oligosaccharide structure composed of two GlcNAc and three Man residues (Figure 2A) and always occur in the consensus sequence Asn-Xaa-Ser/Thr/Cys- (*50*). Pro or Asp residues in the Xaa position, or a Pro residue flanking this consensus, inhibit carbohydrate attachment (*51*). According to these rules, each of the human IαI heavy chains contain two potential N-linked glycosylation sites. To determine if these sites were indeed glycosylated, heavy-chain peptides containing N-linked oligosaccharides were purified by concanavalin A affinity chromatography followed

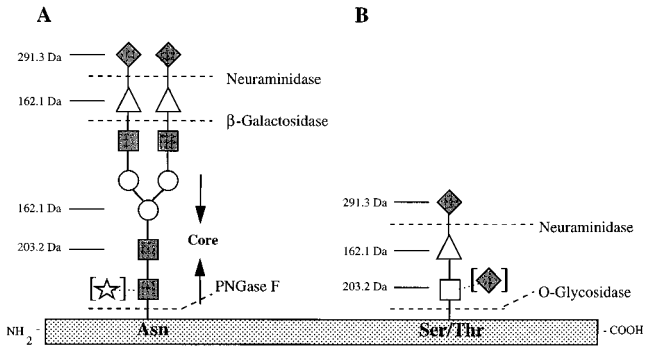

FIGURE 2: Structure of the N- and O-linked oligosaccharides on the heavy chains of human IαI. The structure of the N-linked (A) and O-linked (B) oligosaccharides was determined by carbohydrate sequencing with different glycosidases. The cleavage sites of the glycosidases are indicated with dotted lines. Monosaccharide masses are marked to the left of the individual monosaccharides which are indicated with symbols: filled squares, *N*-acetylglucosamine (GlcNAc); open squares, *N*-acetylgalactosamine (GalNAc); open circles, mannose (Man); open triangles, galactose (Gal); filled diamonds, sialic acid (SA); and open stars, fucose (Fuc). Monosaccharides enclosed in brackets are only present in some of the glycans.

by RP-HPLC. Edman degradation of the purified glycopeptides identified Asn^{251} and Asn^{554} of HC1 and Asn^{64} of HC2 as the sites of glycosylation. The N-glycosylated peptides were recovered with yields similar to those obtained for the unglycosylated peptides, indicating full utilization of these three N-glycosylation sites.

Structure of the HC1 Asn^{251}-Linked Oligosaccharide. The oligosaccharide on the $Phe^{249}-Lys^{257}$ glycopeptide containing Asn^{251} was digested with specific glycosidases in a sequential manner. First, neuraminidase was employed to remove terminal sialic acid residues (SA). Then terminal galactose residues (Gal) were removed by β-galactosidase digestion, and finally the remaining carbohydrate was cleaved from the peptide by PNGase F. The glycopeptide was analyzed by mass spectrometry after each glycosidase treatment, and the structure of the attached oligosaccharide was deduced from the changes in the mass after each glycosidase digestion. The mass of the PNGase F-treated peptide, $Phe^{249}-Lys^{257}$ (1279.9 Da), correlates with the calculated mass of the unglycosylated peptide (Table 1). Thus, all glycans have been removed with PNGase F. The mass difference between the PNGase F-stripped peptide and the β-galactosidase-treated peptide describes the mass of the carbohydrate without the terminal galactose and sialic acid residues. This mass (1300.6 Da) is equivalent to the mass of the core structure (892.8 Da) plus two GlcNAc residues (203.2 Da) (Table 1; Figure 2A). The mass difference (324.4 Da) between the β-galactosidase-treated and the neuraminidase-treated peptide correlates with the attachment of two galactose residues to the terminal GlcNAc residues (Table 1; Figure 2A). Finally, the mass difference between the neuraminidase-treated and the untreated peptide (585.3 Da) indicates that the carbohydrate structure is terminated by two sialic acid residues (Table 1). This demonstrates that the oligosaccharide attached to HC1 Asn^{251} is a complex biantennary structure composed of (Asn)-$GlcNAc_2$-Man-(Man-GlcNAc-Gal-SA)$_2$ (Figure 1A).

Structure of the HC1 Asn^{554}-Linked Oligosaccharide. The second potential N-linked glycosylation site in HC1, Asn^{554}, is also glycosylated, and the oligosaccharide structure was determined as described above. The mass of the PNGase F-treated peptide, $Ala^{553}-Arg^{561}$ (961.4 Da), correlates with the calculated mass of the unglycosylated peptide (Table 1); hence no other modifications are present on this peptide. The difference between the calculated mass of the unglycosylated peptide and the mass of the β-galactosidase-treated peptide (1299.4 Da) corresponds to the attachment of the core structure and two GlcNAc residues (Table 1). The mass difference between the β-galactosidase-treated and the neuraminidase-treated peptide (324.5 Da) indicates an attachment of two galactose residues to the terminal GlcNAc residues (Table 1; Figure 2A). Finally, the mass difference between the neuraminidase-treated and the untreated peptide (584.2 Da) indicates that the carbohydrate structure is terminated by two sialic acid residues (Table 1; Figure 2A). Thus the glycosylation on HC1 Asn^{554} is also a complex biantennary structure similar to the one identified on HC1 Asn^{251}. A subpopulation (less than 10%) of the Asn^{554} glycans had an additional mass of 146 Da, indicating the attachment of a fucose residue (Fuc) at the innermost GlcNAc residue in these molecules (Figure 2A).

Structure of the HC2 Asn^{64}-Linked Oligosaccharide. HC2 contains two potential N-glycosylation sites, of which only Asn^{64} is glycosylated. The lack of posttranslational modifications of HC2 Asn^{391} was confirmed by Edman degradation and mass spectrometry of the purified tryptic peptide. The tryptic peptide $Gly^{59}-Lys^{72}$ containing the glycosylated Asn^{64} was purified, and the carbohydrate structure was analyzed as described above. The observed mass of the PNGase F-treated peptide (1475.6 Da) corresponds to the calculated mass of the unglycosylated peptide (Table 1), and the mass difference between the PNGase F-stripped peptide and the β-galactosidase-treated peptide (1300.7 Da) corresponds to the attachment of the core structure and two GlcNAc residues to HC2 Asn^{64} (Table 1; Figure 2A). The mass difference between the β-galactosidase-treated and the neuraminidase-treated peptide (327.0 Da) correlates with two galactose residues, and the difference in mass between the neuraminidase-treated and the untreated peptide (584.4 Da) corresponds to the addition of two sialic acid residues to the carbohydrate structure (Table 1). Thus, the HC2 Asn^{64}-linked oligosaccharide appears to be a complex biantennary structure identical to the two N-glycans identified on HC1 of human IαI (Figure 2A).

Isolation of O-Glycosylated Peptides. O-Linked oligosaccharides are attached to the hydroxyl group of Thr and Ser residues. In contrast to the N-linked glycans, no strict consensus sequence for O-linked glycosylation has been identified. Only a few rules of thumb have been formulated for O-glycosidic linkage (*52*−*54*). It is therefore less predictable to assess the number of putative O-linked glycosylation sites present on a protein. O-glycosylated peptides of the IαI heavy chains were purified by jacalin affinity chromatography and separated by RP-HPLC. The major peaks were subjected to Edman degradation. An O-glycosylated C-terminal peptide was identified from both heavy chains 1 and 2, and the structure of these carbohydrates was determined by the carbohydrate sequencing approach described above.

Analysis of HC1 O-Glycosylations on Thr^{619}. Edman degradation of the C-terminal HC1 peptide, $Thr^{608}-Arg^{627}$,

Table 1: Identification of N-Linked Carbohydrate Structures[a]

peptide	following fragmentation with	molecular mass (Da) observed	calculated	difference	proposed carbohydrate structure	calcd mass (Da)
HC1 Phe^{249}−Lys^{257}	PNGase F	1279.9	1278.5	1.4		
	β-Galactosidase	2579.1	1278.5	1300.6	core + 2(GlcNAc)	1299.2
	neuraminidase	2902.1	2577.7	324.4	2× galactose	324.3
		3487.3	2902.0	585.3	2× sialic acid	582.5
HC1 Ala^{553}−Arg^{561}	PNGase F	9061.4	960.1	1.3		
	β-galactosidase	2259.5	960.1	1299.4	core + 2(GlcNAc)	1299.2
	neuraminidase	2583.8	2259.3	324.5	2× galactose	324.3
		3167.8	2583.6	584.2	2× sialic acid	582.5
HC2 Gly^{59}−Lys^{72}	PNGase F	1475.6	1474.7	0.9		
	β-galactosidase	2775.4	1474.7	1300.7	core + 2(GlcNAc)	1299.2
	neuraminidase	3100.9	2773.9	327.0	2× galactose	324.3
		3682.6	3098.2	584.4	2× sialic acid	582.5

[a] N-Glycosylated tryptic peptides of IαI HC1 and HC2 were isolated by concanavalin A chromatography and further separated by RP-HPLC. Subsequently the carbohydrate was digested with neuraminidase, followed by β-galactosidase, and finally PNGase F. Aliquots were removed following each enzymatic digestion and analyzed by mass spectrometry as described in Experimental Procedures. The carbohydrate structures were deduced from the mass changes.

Table 2: Identification of O-Linked Carbohydrate Structures[a]

peptide	following fragmentation with	molecular mass (Da) observed	calculated	difference	proposed carbohydrate structure	calcd mass (Da)
HC1 Thr^{608}−Arg^{627}	O-glycosidase	2159.5	2159.4	0.1		
	neuraminidase	2524.8	2159.4	365.4	GalNAc-Hex	365.3
		2816.6	2524.7	291.9	sialic acid	291.3
HC2 Val^{607}−Arg^{644}	O-glycosidase	4012.9	4012.6	0.3		
	neuraminidase	5474.3	4012.6	1461.7	4(GalNAc-Hex)	1461.2
		6628.6	5474.0	1154.6	4× sialic acid	1165.2
		6918.7	5474.0	1444.7	5× sialic acid	1456.5

[a] O-Glycosylated tryptic peptides of IαI HC1 and HC2 were isolated by jacalin−agarose chromatography and further separated by RP-HPLC. Subsequently the carbohydrate was sequenced with neuraminidase and O-glycosidase. Aliquots were removed following each enzymatic digestion and analyzed by mass spectrometry as described in Experimental Procedures. The carbohydrate structures were deduced from the mass changes.

identified Thr^{619} as the site of O-glycosylation. To determine the oligosaccharide structure, the carbohydrate was first digested with neuraminidase, which removes terminal sialic acid residues, and then with O-glycosidase, which cleaves the Gal-GalNAc disaccharide from the Thr residues. The mass of the O-glycosidase-treated peptide corresponds with the calculated mass of the unglycosylated peptide (Table 2), and O-glycosylation is thus the only modification of the peptide. The mass difference between the O-glycosidase-treated peptide and the neuraminidase-treated peptide (365.4 Da) is equivalent to the mass of one GalNAc (203.2 Da) and one Gal (162.1 Da) residue (Table 2; Figure 2B). The mass difference between the neuraminidase-treated and the untreated peptide (291.9 Da) indicates that the carbohydrate structure is terminated by a sialic acid residue (Table 2). Thus the carbohydrate attached to HC1 Thr^{619} is a trisaccharide composed of (Thr)-GalNAc-Gal-SA (Figure 2B).

Analysis of the O-Glycosylations on HC2. Carbohydrate sequencing and mass spectrometry of the jacalin−agarose-retained C-terminal HC2 peptide, Val^{607}−Arg^{644}, indicate the presence of a cluster of four O-glycans. Edman degradation of the peptide identified Thr^{612}, Ser^{619}, Thr^{621}, and Thr^{637} as the points of glycosylation. The HC2 Thr^{637} glycan has previously been identified as a HexNAc-Hex-SA trisaccharide (*6*). The mass of the O-glycosidase-treated Val^{607}−Arg^{644} peptide correlates with the calculated mass of the unglycosylated peptide, and hence O-glycosylation is the only posttranslational modification of this peptide (Table 2). The mass difference between the calculated mass of the unglycosylated peptide and the neuraminidase-treated peptide (1461.7 Da) corresponds to the attachment of four GalNAc-Gal disaccharide units (4 × 365.3 Da), and thus four O-glycosylated sites (Table 2). Two major populations exist in the untreated peptide. The mass difference between these and the neuraminidase-treated peptide (1154.6 Da/1444.7 Da) indicates one population (∼45%) where each of the four O-linked glycans is terminated by a sialic acid residue and another population (∼45%) where an additional sialic acid is attached to the GalNAc residue of one of the Thr^{612}, Ser^{619}, or Thr^{621} trisaccharides (Table 2; Figure 2B). Furthermore, small subpopulations (∼10%) with only three GalNAc-Gal-SA trisaccharides were also detected in addition to the dominant cluster of four O-glycosylated sites (data not shown).

Isolation of Disulfide- and Cysteine-Containing Peptides. It has previously been determined by SDS-PAGE before and after reduction that IαI does not contain any interchain disulfide bridges (*3*, *55*). The disulfide bridge pattern of bikunin have previously been assigned on the basis of sequence homology with other kunins (*36*). However, the status of the Cys residues in the heavy chains is unknown. Initial experiments suggested that denatured non-reduced IαI did not incorporate radioactivity when subjected to alkylation with [^{14}C] iodoacetic acid. Consequently, IαI does not contain free thiol groups (data not shown). To determine the nature of the Cys modification, the heavy chains were digested with trypsin and the Cys-containing peptides were purified by HPLC. Cys-containing peptides were selected

Table 3: Identification of Cys Modifications by Mass Spectrometry[a]

peptide	Cys	molecular mass (Da)			
		calculated[b]	observed (−DTT)	observed (+DTT)	difference
HC1 Ser^{21}−Arg^{31}	Cys^{26}	1234.5	1557.8	1557.8	0
HC1 Gly^{196}−Lys^{223}	Cys^{210}, Cys^{213}	3072.5	3072.0	3073.7	1.7
			3070.6	3072.5	1.9
			3070.5	3072.8	2.3
			3070.7	3072.9	2.2
			3070.4	3073.2	2.8
HC1 Asp^{231}−His^{245}	Cys^{234}	1710.9	4375.6	1710.8	2664.8
HC1 Ala^{491}−Lys^{514}	Cys^{506}	2665.9	4375.6	2667.3	1708.3
HC2 Ile^{206}−Arg^{211}	Cys^{207}, Cys^{210}	676.8	676.3	678./4	1.9
			674.0	676.0	2.0
			674.4	676.0	1.6
			674.0	675.8	1.8
			673.9	676.0	2.1
HC2 Met^{585}−Lys^{606}	Cys^{596}, Cys^{597}	2272./5	2286.5	2288.2	1.7
			2286.4	2288.0	1.6
			2286.6	2288.3	1.7
			2286.5	2288.2	1.7
			2287.0	2288.2	1.5

[a] Cys-containing peptides of IαI HC1 and HC2 were analyzed by mass spectrometry either in their native form or following reduction with DTT as described in Experimental Procedures. The mass differences indicated in the table are the differences between +/− DTT-treated peptides. [b] Calculated masses are derived from the amino acid sequence of the peptides.

by the presence of Cya in amino acid composition and identified by Edman degradation. The yield of the Cys-containing peptides was similar to that of other peptides.

Analysis of HC1 Cys^{26}. The HC1 peptide, Ser^{21}−Arg^{31}, containing Cys^{26} was generated by BNPS-skatole fragmentation followed by clostripain digestion and identified as described above. The observed mass of the peptide (1557.8 Da) was 323.3 Da higher than the calculated mass (Table 3), implying that the peptide was modified. The modification is most likely attached to Cys^{26}, since no PTH amino acid was detected following Edman degradation and all other residues in the peptide were positively identified. Moreover, rigorous reduction and alkylation of the peptide did not lead to any change in the mass (Table 3). Taken together, these data suggest that Cys^{26} of HC1 could be hexosylated with either two galactose or two glucose residues, both with a calculated theoretical mass of 324 Da (*56*−*58*).

Analysis of HC1 Cys^{234} and Cys^{506}. A peptide with a mass of 4375.6 Da was purified. Edman degradation of the peptide revealed two sequences corresponding to HC1 Asp^{231}−His^{245} and HC1 Ala^{491}−Lys^{514}. These peptides contain HC1 Cys^{234} and Cys^{506}, respectively. Reduction of the 4375.6 Da fragment followed by mass spectrometry resulted in two new peptide masses (1710.8 and 2667.3 Da) that correlate with the calculated masses of HC1 Asp^{231}−His^{245} and HC1 Ala^{491}−Lys^{514} (Table 3). Thus HC1 Cys^{234} forms a disulfide bridge with HC1 Cys^{506}.

Analysis of the Adjacent Cys Residues HC1 Cys^{210} and Cys^{213}, HC2 Cys^{207} and Cys^{210}, and HC2 Cys^{596} and Cys^{597}. The following peptides—HC1 Pro^{202}−Lys^{223}, containing Cys^{210} and Cys^{213}; HC2 Ile^{206}−Arg^{211}, containing Cys^{207} and Cys^{210}; and, HC2 Met^{585}−Lys^{606}, containing Cys^{596} and Cys^{597}—were identified by amino acid analysis, Edman degradation, and mass spectrometry. The observed mass of these peptides (Table 3) indicates that neither of the Cys residues is involved in interpeptide disulfide bridges, e.g., to other peptides or glutathione. We note that the observed mass of the HC2 peptide Met^{585}−Lys^{606}, containing Cys^{596} and Cys^{597}, was approximately 14 Da higher than the calculated mass (2272.5 Da) (Table 3). This increase is too small to account for any form of interpeptide disulfide bond and is probably due to some modification elsewhere in the peptide, e.g., oxidation of Met^{585}. As IαI does not contain free sulfhydryl groups, the three pairs of Cys residues most likely form disulfide bridges with each other (HC1 Cys^{210} with HC1 Cys^{213}, HC2 Cys^{207} with HC2 Cys^{210}, and HC2 Cys^{596} with HC2 Cys^{597}). To confirm this, the three peptides were treated with DTT and the masses of the reduced peptides were determined. Each set of experiments was repeated five times, and each time a mass increase of approximately 2 Da was observed, consistent with the reduction of a disulfide bridge (Table 3). We therefore conclude that HC1 Cys^{210} and Cys^{213} share a disulfide bond and that HC2 contains two disulfide bridges one between Cys^{207} and Cys^{210} and another between Cys^{596} and Cys^{597}.

DISCUSSION

In general, plasma proteins are glycoproteins composed of one or an even number of identical subunits and are often cross-linked by interchain disulfide bridges. IαI deviates from this generalization in several ways. The protein is a proteoglycan, it is a heterotrimer, and the three subunits are assembled by the unique PGP cross-link (*5*, *6*, *59*). These features, the high concentration in the plasma, and the apparent lack of a functional significance of these intriguing properties make this protein an interesting candidate for further structural analysis.

In recent years it has become apparent that N-linked oligosaccharides play a role in the conformational maturation of most glycoproteins. These carbohydrate structures are attached in the endoplasmatic reticulum before the folding of the polypeptide chain and may thus affect later biosynthesis events. However, individual glycoproteins depend on their oligosaccharide side chains to varying degrees and in different ways (*60*). N-Linked glycosylations are central in facilitating the folding, transport, cell surface expression, and secretion of glycoproteins (*60*, *61*). In many cases, N-linked carbohydrates also play an important role in regulating the

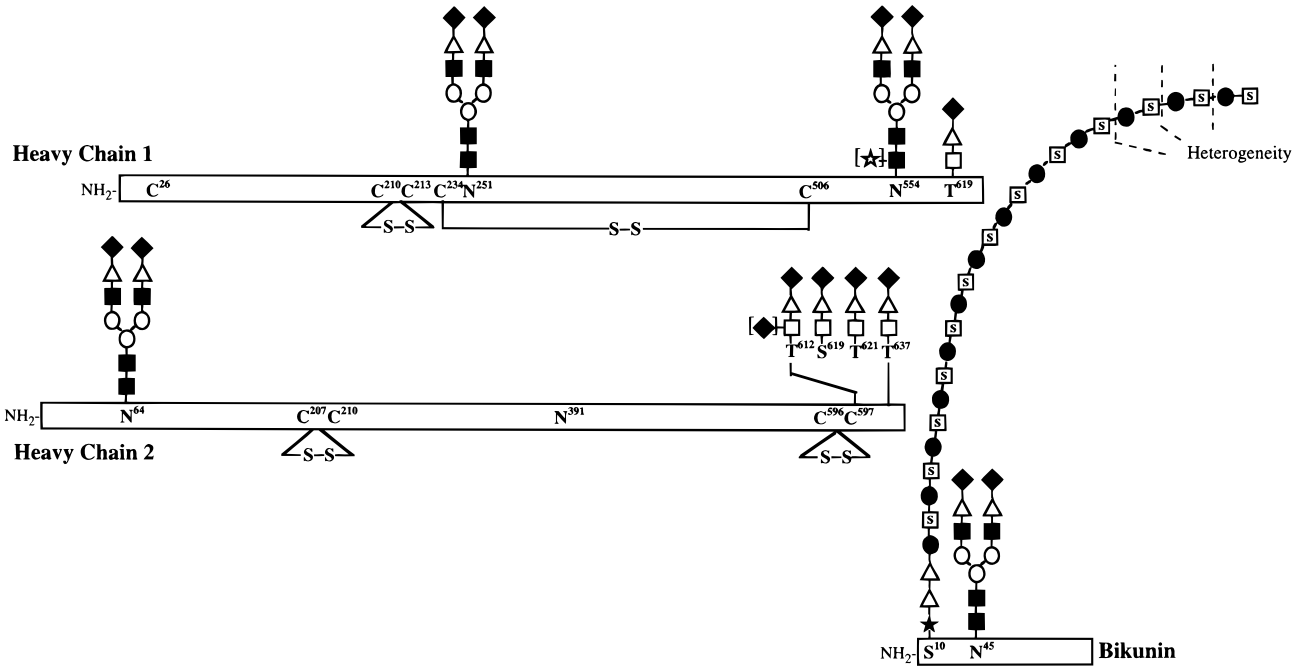

FIGURE 3: Structural overview of human IαI: schematic representation of disulfide bridge pattern and the main forms of N- and O-linked glycosylation of human IαI. The heavy chains are connected to the GAG chain via ester bonds between their C-terminal Asp residues and an internal *N*-acetylgalactosamine residue of the GAG chain (*5*, *6*). The exact position of the heavy chains on the GAG chain is presently unknown. Disulfide bridges are marked with solid lines. The six disulfide bridges in bikunin are omitted for simplicity. The individual monosaccharides are indicated with symbols: filled squares, *N*-acetylglucosamine; open squares, *N*-acetylgalactosamine; open squares marked with S, 4-sulfo-*N*-acetylglucosamine; open circles, mannose; filled circles, glucuronic acid; open triangles, galactose; filled diamonds, sialic acid; open stars, fucose; and filled stars, xylose. Monosaccharides enclosed in brackets are only present in some of the glycans. The N-linked glycosylation on bikunin has previously been described (*35*).

activity, stability, and antigenicity of mature proteins (*62*). O-linked glycans, which are attached after the folding of the protein is completed, may play a critical role in modulating cell adhesion, and recent findings indicate that some glycoproteins carry ligands that interact with clusters of O-glycans on the protein (*63*).

Although various proteins carry more than one N-linked glycan, they often show different carbohydrate structures of the individual glycans (*47*, *63*, *64*). The overall oligosaccharide structures determined for the three N-glycans on the IαI heavy chains are identical to the recently characterized (*Asn*)-GlcNAc$_2$-Man-(Man-GlcNAc-Gal-SA)$_2$ oligosaccharide attached to Asn^{45} of bikunin (*35*) (Figure 3). This biantennary structure (with or without the additional fucosylation) is among the most commonly identified N-linked oligosaccharide structures (*63*). Although some degree of glycan heterogeneity is usually observed, surprisingly little heterogeneity was observed in the IαI N-glycans, which with the exception of the partial fucosylation of HC1 Asn^{554} appear to be very homogeneous. In addition to the N-glycans, the IαI heavy chains also carry 5 O-linked trisaccharides: one attached to HC1 Thr^{619} and a cluster of four near the C-terminus of HC2 (Figure 3). All the IαI carbohydrate structures are terminated by sialic acid residues, as seen with other plasma proteins, which masks the underlying galactose and *N*-acetylglucosamine residues and protects the proteins from rapid clearance via receptors in the liver (*64*, *65*).

IαI does not, as previously mentioned, depend on interchain disulfide bridges to maintain its subunit composition. Examination of the posttranslational modifications of the nine Cys residues in the heavy chains showed that the single unpaired Cys residue in IαI, HC1 Cys^{26}, carries a 323.3 Da reduction-resistant modification. The fact that the modification is resistant to reduction and that the mass of the modification (323.3 Da) correlates with the mass of two hexose residues (324.3 Da) indicate that HC1 Cys^{26} is glycosylated with either two glucose or two galactose residues. Plasma proteins rarely contain free sulfhydryl groups, and unpaired Cys residues are usually modified by the attachment of cysteine or glutathione. Hexosylation of Cys residues has previously been reported in both urinary peptides and erythrocyte membrane peptides (*56*−*58*). However, the functional significance of the Cys glycosylation in these peptides and the glycosylation of Cys^{26} in IαI HC1 remains unclear.

The other IαI heavy chain Cys residues form intrachain disulfides. One of these bridges is formed between HC1 Cys^{234} and Cys^{506} (Figure 3). The other three disulfides are, however, formed between Cys residues that are either adjacent or only two amino acids apart. Such short loop disulfides are predicted to interrupt any ordinary form of secondary structure in their immediate proximity, and it is unlikely that they influence overall protein stability. Disulfide bridges between two sequentially adjacent Cys residues, such as the one identified between HC2 Cys^{596} and Cys^{597}, require special consideration. Theoretical studies (*66*) and crystallographic data (*67*) indicate that formation of such a disulfide requires a *cis* peptide bond between the adjacent Cys residues; although this is possible, such a configuration is definitely not common (*68*, *69*). One of the few other examples known is the presence of a conserved disulfide bridge between adjacent Cys residues found in all nicotinic receptors. The conformation of the Cys^{192}-Cys^{193} bridge, in

the acetylcholine receptor binding site of *Torpedo californica*, affects agonist-induced conformational changes at the receptor's binding site and has been suggested to function as a molecular switch (*70*).

The last two disulfide bridges identified in the human IαI heavy chains are formed between Cys^{210} and Cys^{213} in HC1 and between Cys^{207} and Cys^{210} in HC2 (Figure 3). They are found in the same place in the sequences in HC1 and HC2, and their sequence motifs are practically identical: Cys-Pro-Thr-Cys in HC1 and Cys-Pro-Ser-Cys in HC2. In fact, the Cys-Pro-Thr-Cys motif is also found in position 203−206 in HC3 of PαI, whereas no consensus sequence is found in the heavy-chain homologue IHRP. Such Cys-Xaa-Xaa-Cys disulfides are found in proteins such as triglyceride lipase and *Rhizopus* aspartic proteinase (*71*, *72*), as well as in the reactive centers of several proteins with thioredoxin-related functions (*73*−*76*). In this class of proteins, the redox-active disulfide bond is involved in a variety of redox processes, for example; assisting in the formation of disulfide bonds required for correct protein folding, adding of glutathione to substrates, and reducing disulfides essential for the enzymatic activity (e.g., reduction of ribonucleotide reductase) (*76*, *77*).

Interestingly, all known crystal structures of such disulfides between Cys residues separated by two residues share the same topography [J. Richardson, personal communication]. The Cys-Xaa-Xaa-Cys disulfide is likely to be found at the beginning of an α-helix. They form a fairly planar square structure, where the backbone of the two residues separating the Cys residues is on the surface of the protein, whereas the disulfide is shielded from the surrounding solvent by the side chains of the adjacent residues. It is thus intriguing to speculated that the unusual disulfide bridge pattern in IαI offers a catalytic function for the heavy-chain moiety of IαI. We are currently investigating the possibility of a redox-active function of this intriguing plasma protein.

ACKNOWLEDGMENT

We thank Dr. Peter Højrup for useful comments and suggestions throughout the process of this study and Dr. Dave Rubenstein for helpful suggestions to the manuscript. Furthermore, we thank Dr. Jane S. Richardson for helpful discussions on the structural implications of the disulfide bridges.

REFERENCES

1. Shulman, R. N. (1955) *J. Biol. Chem. 213*, 655−671.
2. Gebhard, W., and Hochstrasser, K. (1986) in *Proteinase Inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 389−401, Elsevier/North-Holland Biomedical Press, Amsterdam.
3. Enghild, J. J., Thøgersen, I. B., Pizzo, S. V., and Salvesen, G. (1989) *J. Biol. Chem. 264*, 15975−15981.
4. Salier, J.-P., Rouet, P., Raguenez, G., and Daveau, M. (1996) *Biochem. J. 315*, 1−9.
5. Enghild, J. J., Salvesen, G., Hefta, S. A., Thøgersen, I. B., Rutherfurd, S., and Pizzo, S. V. (1991) *J. Biol. Chem. 266*, 747−751.
6. Enghild, J. J., Salvesen, G., Thøgersen, I. B., Valnickova, Z., Pizzo, S. V., and Hefta, S. A. (1993) *J. Biol. Chem. 268*, 8711−8716.
7. Gebhard, W., Schreitmüller, T., Hochstrasser, K., and Wachter, E. (1988) *FEBS Lett. 229*, 63−67.
8. Gebhard, W., Schreitmüller, T., Hochstrasser, K., and Wachter, E. (1989) *Eur. J. Biochem. 181*, 571−576.
9. Diarra-Mehrpour, M., Bourguignon, J., Sesboüé, R., Matteï, M.-G., Passage, E., Salier, J.-P., and Martin, J.-P. (1989) *Eur. J. Biochem. 179*, 147−154.
10. Diarra-Mehrpour, M., Bourguignon, J., Bost, F., Sesböué, R., Muschio, F., Sarafan, N., and Martin, J.-P. (1992) *Biochim. Biophys. Acta 1132*, 114−118.
11. Kraumeyer, J. F., Polazzi, J. O., and Kotick, M. P. (1986) *Nucleic Acids Res. 14*, 7839−7850.
12. Kozak, M. (1987) *Nucleic Acids Res. 15*, 8125−8132.
13. Saguchi, K.-I., Tobe, T., Hashimoto, K., Sano, Y., Nakano, Y., Miura, N.-H., and Tomita, M. (1995) *J. Biochem. (Tokyo) 117*, 14−18.
14. Nishimura, H., Kakizaki, I., Muta, T., Sasaki, N., Pu, P. X., Yamashita, T., and Nagasawa, S. (1995) *FEBS Lett. 357*, 207−211.
15. Choi-Miura, N.-H., Sano, Y., Oda, E., Nakano, Y., Tobe, T., Yanagishita, T., Taniyama, M., Katagiri, T., and Tomita, M. (1995) *J. Biochem. (Tokyo) 117*, 400−407.
16. Åkerström, B., and Lögdberg, L. (1990) *Trends Biochem. Sci. 15*, 240−243.
17. Leaver, M. J., Wright, J., and George, S. G. (1994) *Comp. Biochem. Physiol. 108*, 275−281.
18. Hanley, S., McInerney, J. O., and Powell, R. (1995) *Mol. Marine Biol. Biotechnol. 4*, 105−109.
19. Ide, H., Itoh, H., and Nawa, Y. (1994) *Biochim. Biophys. Acta 1209*, 286−292.
20. Gebhard, W., Hochstrasser, K., Fritz, H., Enghild, J. J., Pizzo, S. V., and Salvesen, G. (1990) *Biol. Chem. Hoppe-Seyler 371*, 13−22.
21. Lindqvist, A., Bratt, T., Altieri, M., Kastern, W., and Åkerström, B. (1992) *Biochim. Biophys. Acta 1130*, 63−67.
22. Thøgersen, I. B., and Enghild, J. J. (1995) *J. Biol. Chem. 270*, 18700−18709.
23. Barr, P. J. (1991) *Cell 66*, 1−3.
24. Bratt, T., Olsson, H., Sjöberg, E. M., Jergil, B., and Åkerström, B. (1993) *Biochim. Biophys. Acta 1157*, 147−154.
25. Bratt, T., Cedervall, T., and Åkerström, B. (1994) *FEBS Lett. 354*, 57−61.
26. Albrecht, G. I., Hochstrasser, K., and Schönberger, Ö. L. (1983) *Hoppe-Seyler's Z. Physiol. Chem. 364*, 1697−1702.
27. Chen, L., Mao, S. J. T., and Larsen, W. J. (1992) *J. Biol. Chem. 267*, 12380−12386.
28. Chen, L., Mao, S. J. T., McLean, L. R., Powers, R. W., and Larsen, W. J. (1994) *J. Biol. Chem. 269*, 28282−28287.
29. Huang, L., Yoneda, M., and Kimata, K. (1993) *J. Biol. Chem. 268*, 26725−26730.
30. Castillo, G. M., and Templeton, D. M. (1993) *FEBS Lett. 318*, 292−296.
31. Jessen, T. E., Ødum, L., and Johnsen, A. H. (1994) *Biol. Chem. Hoppe-Seyler 375*, 521−526.
32. Zhao, M., Yoneda, M., Ohashi, Y., Kurono, S., Iwata, H., Ohnuki, Y., and Kimata, K. (1995) *J. Biol. Chem. 270*, 26657−26663.
33. Kobayashi, H., Gotoh, J., Hirashima, Y., and Terao, T. (1996) *J. Biol. Chem. 271*, 11362−11367.
34. Wisniewski, H.-G., Burgess, W. H., Oppenheim, J. D., and Vilcek, J. (1994) *Biochemistry 33*, 7423−7429.
35. Rahbek-Nielsen, H., Thøgersen, I. B., Chen, F., Fransson, L. Å., Roepstorff, P., and Enghild, J. J. (1997) (submitted for publication).
36. Warne, N. W., and Laskowski, M., Jr. (1990) *Biochem. Biophys. Res. Commun. 172*, 1364−1370.
37. Sayers, C. A., and Barrett, A. J. (1980) *Biochem. J. 189*, 255−261.
38. Bury, A. (1981) *J. Chromatogr. 213*, 491−500.
39. Matsudaira, P. (1987) *J. Biol. Chem. 262*, 10035−10038.
40. Uriel, J., and Berges, J. (1968) *Nature 218*, 578−580.
41. Tarr, G. E. (1986) in *Methods of Protein Michrocharacterization* (Shively, J. E., Ed.) pp 155−194, Humana Press Inc., Clinton, NY.
42. Hsi, K.-L., Chen, L., and Yuan, P.-M. (1993) in *Techniques in Protein Chemistry IV* (Angeletti, R. H., Ed.) pp 143−151, Academic Press, Inc., San Diego, CA.

43. Hortin, G. L., and Trimpe, B. L. (1990) *Anal. Biochem. 188*, 271−277.
44. Meltzer, N. M., Tous, G. I., Gruber, S., and Stein, S. (1987) *Anal. Biochem. 160*, 356−361.
45. Ozols, J. (1990) *Methods Enzymol. 182*, 578−601.
46. Vorm, O., and Mann, M. (1994) *J. Am. Soc. Mass Specrom. 5*, 955−958.
47. Paulson, J. C. (1989) *Trends Biochem. Sci. 14*, 272−275.
48. Hayes, B. K., and Hart, G. W. (1994) *Curr. Opin. Struct. Biol. 4*, 692−696.
49. Rudd, P. M., Woods, R. J., Wormald, M. R., Opdenakker, G., Downing, A. K., Campbell, I. D., and Dwek, R. A. (1995) *Biochim. Biophys. Acta 1248*, 1−10.
50. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem. 54*, 631−664.
51. Opdenakker, G., Rudd, P. M., Ponting, C. P., and Dwek, R. A. (1993) *FASEB J. 7*, 1330−1337.
52. Elhammer, Å. P., Poorman, R. A., Brown, E., Maggiora, L. L., Hoogerheide, J. G., and Kézdy, F. J. (1993) *J. Biol. Chem. 268*, 10029−10038.
53. Hansen, J. E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J. O., Hansen, J. E. S., and Burnak, S. (1995) *Biochem. J. 308*, 801−813.
54. Roth, J., Wang, Y., Eckhardt, A. E., and Hill, R. L. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 8935−8939.
55. Morii, M., and Travis, J. (1985) *Biol. Chem. Hoppe-Seyler 366*, 19−21.
56. Lote, C. J., and Weiss, J. B. (1971) *FEBS Lett. 16*, 81−85.
57. Weiss, J. B., and Lote, C. J. (1971) *Nature New Biol. 234*, 25−26.
58. Krishna, R. G., and Wold, F. (1993) in *Methods in Protein Sequence Analysis* (Imahori, K., and Sakiyama, F., Eds.) pp 167−172, Plenum Press, New York.
59. Morelle, W., Capon, C., Balduyck, M., Sautiere, P., Kouach, M., Michalski, C., Fournet, B., and Mizon, J. (1994) *Eur. J. Biochem. 221*, 881−888.
60. Helenius, A. (1994) *Mol. Biol. Cell 5*, 253−265.
61. Shakin-Eshleman, S. H., and Spitalnik, S. L. (1993) *Trends Glycosci. Glycotechnol. 5*, 355−368.
62. Rademacher, R. W., Parekh, R. B., and Dwek, R. A. (1988) *Annu. Rev. Biochem. 57*, 785−838.
63. Fukuda, M. (1994) in *Molecular Glycobiology* (Fukuda, M., and Hindsgaul, O., Eds.) pp 1−52, Oxford University Press, New York.
64. Baenziger, J. U. (1984) in *The Plasma Proteins* (Putnam, F. W., Ed.) Vol. 4, pp 271−315, Academic Press, New York.
65. Ashwell, G., and Harford, J. (1982) *Annu. Rev. Biochem. 51*, 531−554.
66. Chandrasekaran, R., and Balasubramanian, R. (1969) *Biochim. Biophys. Acta 188*, 1−9.
67. Capusso, S., Mattia, C., and Mazzarella, L. (1977) *Acta B. Struct, Crystallogr. 33*, 2080−2083.
68. Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261−287.
69. Zhang, R., and Snyder, G. H. (1989) *J. Biol. Chem. 264*, 18472−18479.
70. Kao, P. N., and Karlin, A. (1986) *J. Biol. Chem. 261*, 8085-8088.
71. Suguna, K., Padlan, E. A., Smith, C. W., Carlson, W. D., and Davies, D. R. (1987) *Proc. Natl. Acad. Sci. USA 84*, 7009−7013.
72. Derewenda, U., Brzozowski, A. M., Lawson, D. M., and Derewenda, Z. S. (1992) *Biochemistry 31*, 1532−1541.
73. Katti, S. K., LeMaster, D. M., and Eklund, H. (1990) *J. Mol. Biol. 121*, 167−184.
74. Eklund, H., Ingelman, M., Soderberg, B. O., Uhlin, T., Nordlund, P., Nikkola, M., Sonnerstam, U., Joelson, T., and Petratos, K. (1992) *J. Mol. Biol. 228*, 596−618.
75. Ji, X., Zhang, P., Armstrong, R. N., and Gilliland, G. L. (1992) *Biochemistry 31*, 10169−10184.
76. Martin, J. L., Bardwell, J. C. A., and Kuriyan, J. (1993) *Nature 365*, 464−468.
77. Nikitovic, D., and Holmgren, A. (1996) *J. Biol. Chem. 271*, 19180−19185.

BI971137D