Journal of Chromatography, 317 (1984) 11–26
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 418

PURIFICATION OF GLYCOPEPTIDES OF HUMAN PLASMA PROTEINS BY HIGH-PERFORMANCE LIQUID CHROMATOGRAPHY

NOBUHIRO TAKAHASHI, YOKO TAKAHASHI, THOMAS L. ORTEL*, JAY N. LOZIER*, NORIAKI ISHIOKA and FRANK W. PUTNAM*

Department of Biology, Indiana University, Bloomington, IN 47405 (U.S.A.)

SUMMARY

The combination of gel permeation chromatography and high-performance liquid chromatography proves to be very effective for the purification of high-molecular-weight glycopeptides containing a single glycan, that have been difficult to separate by other procedures. In order to facilitate comparison of the chromatographic properties of glycopeptides derived from a variety of proteins and having different structures, identical procedures were used for their purification. The method was applied to a series of human plasma proteins, including immunoglobulin D, ceruloplasmin, hemopexin, β-2-glycoprotein I, 3.1S α-2-leucine-rich glycoprotein, and α-1-B-glycoprotein. All the purified glycopeptides were placed in the protein structure of these plasma proteins. In several cases the carbohydrate structure has been determined by collaborating groups. Immunoglobulin D is the first example of a glycoprotein whose entire primary structure has been defined by utilizing a single protein source. Furthermore, hemopexin and 3.1S α-2-leucine-rich glycoprotein were both found to contain GalN oligosaccharide, which had not previously been identified in these proteins. The method was also used to identify the oligosaccharide that is missing in a carbohydrate variant of ceruloplasmin.

INTRODUCTION

Human serum, in which about 100 proteins have been identified, is a particularly rich source of glycoproteins[1]. The carbohydrate moieties of glycoproteins are thought to have important biological roles, including stabilization of protein conformation and regulation of the lifetime of glycoproteins in the circulatory system by governing their uptake by cells[2]. Two different kinds of linkage between carbohydrate moieties and protein have been recognized: (1) N-glycosidic linkage of glucosamine, where asparagine is involved in the carbohydrate–peptide bond, and (2) O-glycosidic linkage of galactosamine, where serine and/or threonine as well as hydroxylysine are involved in the carbohydrate–peptide bond. Because many plasma proteins are glycosylated at multiple sites, for structural analysis of glycoproteins it

---

* Present address: Indiana University School of Medicine, Indianapolis, IN 46223, U.S.A.

0021-9673/84/$03.00 © 1984 Elsevier Science Publishers B.V.

is necessary to isolate pure glycopeptides containing a single oligosaccharide. Most methods for the isolation of carbohydrate for structural analysis involve extensive digestion with non-specific protease, such as pronase, which results in a mixture of glycans with only one or two amino acids attached. However, it is obvious that isolation of glycopeptide with more than four or five amino acid residues is necessary for the identification of the attachment site in the protein structure. But, when glycopeptides are produced by more specific cleavage methods, the purification of an individual glycopeptide is difficult for following reasons: (1) the heterogeneity of the carbohydrate chain; (2) the multiple forms of glycopeptides resulting from incomplete enzymatic digestion due to steric hindrance of the sugar chains; (3) high affinity to separation matrices that are used in conventional methods, such as paper electrophoresis and ion-exchange chromatography. These problems affect the determination of the amino acid sequence of the polypeptide chain and of the sites of glycosylation, and also hinder elucidation of the oligosaccharide structure.

In this report, we describe extensive application of a method involving a combination of gel permeation (GPC) and reversed-phase high-performance liquid chromatography (HPLC) for the preparative isolation of glycopeptides. In order to facilitate comparison of the chromatographic properties of glycopeptides derived from a variety of proteins and having different structures, identical procedures were used for purification. The method was applied to a series of human plasma proteins including immunoglobulin D (IgD), ceruloplasmin, hemopexin, β-2-glycoprotein I, 3.1S α-2-leucine-rich glycoprotein, and α-1-B-glycoprotein, and in each protein the peptide sequence has been established at each site. In several cases the carbohydrate structure has been determined by collaborating groups. It is of interest that one of these, immunoglobulin D, is the first example of a glycoprotein whose entire primary structure has been defined by utilizing a single protein source. Furthermore, hemopexin and 3.1S α-2-leucine-rich glycoprotein were both found to contain GalN oligosaccharide, which had not previously been identified in these proteins. The method was also used to identify the oligosaccharide that is missing in a carbohydrate variant of ceruloplasmin. The behavior of a glycopeptide in reversed-phase chromatography seems to be determined mainly by its amino acid composition, although some glycopeptides may be separated into several components because of the heterogeneity of their carbohydrate structure.

EXPERIMENTAL

Material

Immunoglobulin D (protein WAH) from the plasma of a patient with multiple myeloma was purified by the method of Lin and Putnam[3]. Ceruloplasmin was prepared from normal pooled human plasma, and the carbohydrate variants of the protein were separated by hydroxyapatite chromatography, as described by Noyer *et al.*[4]. Cohn fraction IV-4 of human plasma, provided by Dr. Harold Gallick of Michigan Department of Public Health (Lansing, MI, U.S.A.), was used for the purification of apohemopexin by a modified method of Tsutsui and Mueller[5]. Heme-hemopexin, β-2-glycoprotein I, 3.1S α-2-leucine-rich glycoprotein, and α-1-B-glycoprotein were obtained from Behringwerke Laboratories (Marburg/Lahn, F.R.G.). All of these proteins were pure as judged by various criteria, such as ami-

no-terminal analysis and polyacrylamide gel electrophoresis with sodium dodecyl sulfate.

The sources of enzymes, reagents and columns were: trypsin treated with L-(tosylamido-2-phenyl)ethyl chloromethyl ketone (TPCK), and chymotrypsin, Worthington (Freehold, NJ, U.S.A.); *Staphylococcus aureus* V8 protease, Miles (Elkhart, IN, U.S.A.); cyanogen bromide (CNBr) and trifluoroacetic acid (TFA), Pierce (Rockford, IL, U.S.A.); 1-propanol of chromatographic grade, Burdick & Jackson Labs. (Muskegon, MI, U.S.A.); Sephadex G-50 Superfine, Pharmacia (Uppsala, Sweden); reversed-phase Synchropak RP-P column (25 × 1.0 cm I.D.), Synchrom (Linden, IN, U.S.A.).

Methods for enzymatic digestion and chemical cleavage, amino acid analysis, and sequence determination

The proteins (10–50 mg) were aminoethylated or carboxymethylated and were digested at an enzyme/substrate ratio of 1/50 in 0.1 *M* ammonium hydrogen carbonate at 37°C for 6 h. The alkylated proteins were also cleaved with CNBr in 70% formic acid[6]. Amino acid analysis was performed with the Beckman amino acid analyzer (Model 121M) and sequence determination with the Beckman sequencer (Model 890C), as described[7,8].

Purification of glycopeptides by combination of GPC and HPLC

In order to facilitate comparison of the chromatographic properties of glycopeptides a combination of GPC and HPLC was used. The lyophilized digest was dissolved in 0.1 ml of formic acid and diluted with 0.9 ml of distilled water. The sample was applied to a column (100 × 1.5 cm I.D.) of Sephadex G-50 and was eluted with 0.1% TFA, containing 10% 1-propanol at a flow-rate of 8 ml/min. The eluate fractions were pooled appropriately and were lyophilized. Each lyophilized fraction was dissolved in 0.1 ml of formic acid and diluted with 0.9 ml of distilled water. A 20-μl volume, was taken from these fractions and was analyzed by the amino acid analyzer after hydrolysis with 4 *M* hydrochloric acid at 108°C for 6 h. The fraction containing GlcN or GalN was purified by HPLC. The HPLC system consists of a Beckman controller (Model 421) and pump (Model 110A) and a reversed-phase Synchropak RP-P column (25 × 1.0 cm I.D.). Peptides were eluted at a flow-rate of 4.16 ml/min with a linear gradient from 0 to 30% of 1-propanol containing 0.1% TFA, during 75 min.

RESULTS AND DISCUSSION

Purification of GlcN-glycopeptides of ceruloplasmin (major form Type I)

Human ceruloplasmin is one of the largest glycoproteins known and consists of a single polypeptide chain with 1046 amino acid residues (MW 132,000)[9]. Although Ryden and Eaker[10] purified only three glycopeptides from ceruloplasmin by a conventional method, four attachment sites for GlcN oligosaccharide were identified in the molecule by the determination of the complete amino acid sequence[9]. The purification of the glycopeptides from the MW 67,000 fragment by HPLC was described previously by Tetaert *et al.*[11]. The method described here is modified from our previous method designed to isolate all four glycopeptides from the whole ce-

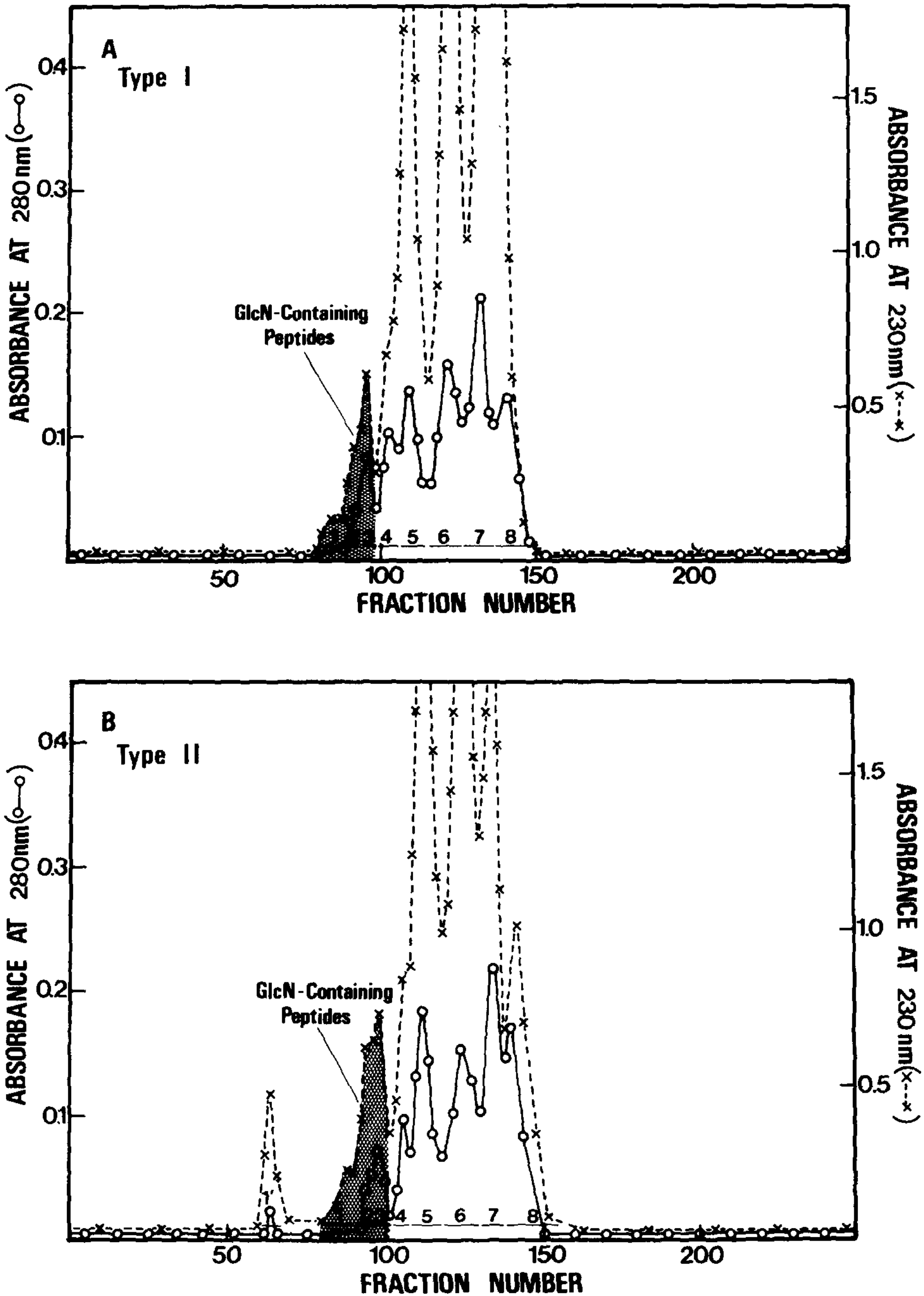

Fig. 1. Initial separation of a tryptic digest of human ceruloplasmin on a Sephadex G-50 column: (A) tryptic digest of Type I; (B) tryptic digest of Type II. The GlcN-containing fractions are shaded.

ruloplasmin molecule, even though the peptide mixture obtained by enzymatic digestion is extremely complex. As was done previously, the first separation of the tryptic digest of the carboxymethylated protein (Type I) was performed by Sephadex G-50 column chromatography to eliminate smaller peptides that lack sugars (Fig. 1A). Although eight peaks were obtained by this method, GlcN was detected only in the first three peaks. The fractions containing GlcN (shown by shading) were purified further by HPLC on a Synchropak RP-P column (Fig. 2A) with a gradient from 0 to 30% of 1-propanol during 75 min, which is a modification of the previous condition[11] designed to improve peptide separation. The column was also changed from analytical size (25 × 0.41 cm I.D.) to preparative size (25 × 1.0 cm I.D.), because more glycopeptide was necessary to analyze both the peptide and carbohydrate structures. The profile of three chromatograms is shown in Fig. 2A as a three-dimensional visualization in order to aid recognition of the overlapping peaks. From the major form (Type I) ceruloplasmin, all four glycopeptides CP1, CP2, CP3 and CP4, were isolated in yields of 46%, 30%, 37% and 36%, respectively[12]. Two peptides, CP1 and CP4, were separated into two and three peaks, respectively (Fig. 2A). The presence of several forms of these peptides is probably due to heterogeneity in the carbohydrate structure and/or to deamidation of glutamine residues during purification. In addition, the glycopeptide CP4 was obtained in two forms depending on whether or not trypsin cleaved after lysine-751[12]. Most of the CP4 peptide obtained had not been cleaved after this residue, probably due to the adjacent aspartic acid residue at position 750[12]. The glycopeptide generated by the cleavage at lysine-751 is designated as *CP4 in Fig. 2A. After the sequence analysis of the glycopeptides, all of them were placed in the complete amino acid sequence of ceruloplasmin reported by Takahashi *et al.*[9]. Ceruloplasmin contains both di- and triantennary N-glycosidic glycans[13], but only the triantennary glycan was suggested to be capable of mediating hepatocyte uptake *in vivo*[14]. However, it is not yet known what the structure of the oligosaccharide is at each of the four attachment sites.

Profile comparison of carbohydrate variants (major form Type I and minor form Type II) of ceruloplasmin

Human ceruloplasmin exists in two different forms that can be separated by hydroxylapatite chromatography[15], and the two forms were reported to differ only in carbohydrate content[16]. Nevertheless, Ryden[17], using a conventional method, isolated only three tryptic glycopeptides from both Type I and Type II in very low yield, and the corresponding peptides were identical in both Type I and Type II. Therefore, our method was also applied to compare the two carbohydrate variants. The two chromatograms of the first separation of the tryptic digests on a Sephadex G-50 column are shown in Fig. 1. Although there is a peak at void volume (fraction number 62) in the chromatogram of Type II, the two profiles are very similar to each other. In both chromatograms only three fractions were detected that contained GlcN by amino acid analysis, *i.e.* fractions 1, 2 and 3. The profile of the chromatogram is shown in Fig. 2 as a three-dimensional visualization. When the two figures are compared, it is obvious that glycopeptide CP2 is missing in Type II. However, the other three glycopeptides are eluted at exactly the same positions and in the same forms by HPLC. In other words, for both Type I and Type II ceruloplasmin, peptides CP1 and CP4 were found in two and three peaks, respectively. Peptide CP4 was

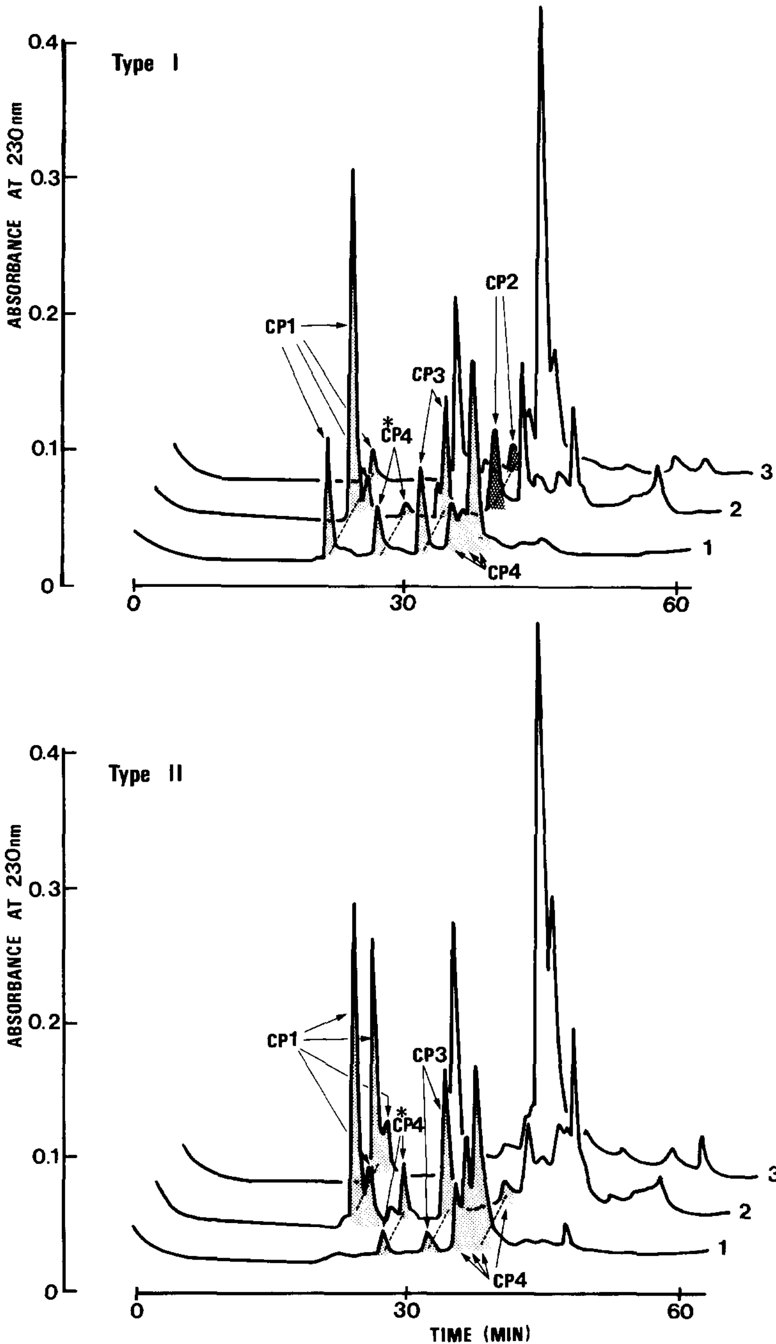

Fig. 2. Three-dimensional visualization of the chromatograms obtained by reversed-phase chromatography on a Synchropak RP-P column of the fractions of Fig. 1 that contain GlcN. The glycopeptides are shaded. Corresponding peptides are connected by a dotted line. Peptide CP2 is cross-hatched in the chromatogram for Type I, but is missing in Type II. The number on the right side of each chromatogram gives the number of the fraction in Fig. 1.

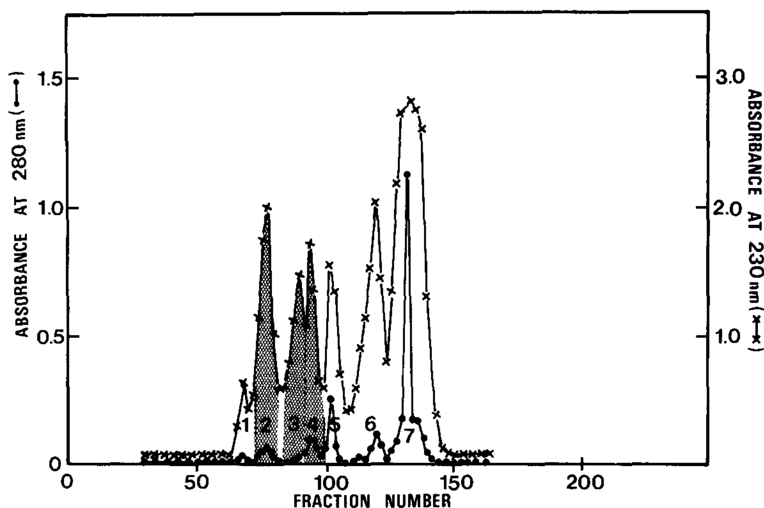

Fig. 3. Separation of the tryptic digest of carboxymethylated citraconylated hemopexin on a Sephadex G-50 column. Fractions containing hexosamine are cross-hatched.

obtained in two forms, depending on whether or not trypsin cleaved after lysine-751, the yield at this site being very similar for both Type I and Type II. The three glycopeptides common to Type I and Type II, CP1, CP3, and CP4, were isolated from Type II in yields of 49%, 41%, and 38%, respectively, and their yields were very similar for Type I. Amino acid sequence analysis showed that the three corresponding peptides from both Type I and Type II had the same sequence[12]. Although the non-glycopeptide corresponding to glycopeptide CP2 has not yet been isolated, the results described, as well as data on other tryptic peptides from Type II, suggest that Type II is not glycosylated at position Asn-339, whereas Type I is[12].

Purification of hemopexin glycopeptides with a GalN oligosaccharide or with multiple sites for GlcN oligosaccharide

Human hemopexin consists of a single polypeptide chain (MW 60,000) and contains about 20% carbohydrate, which is reported to be composed of mannose, galactose, glucosamine, and sialic acid[18]. Therefore, it was expected that all of the carbohydrate was attached to the protein only by N-glycosidic linkage[18].

Because a precipitate formed during tryptic digestion of the carboxymethylated citraconylated hemopexin, the supernatant was applied to a column of Sephadex G-50 (Fig. 3). The peptides were separated into seven fractions, three of which (fractions 2, 3, and 4) contained GlcN. However, GalN was also detected in fraction 3, although GalN had not been identified previously in this protein[18]. Fraction 3 was further purified by HPLC, and pure GalN-containing peptide was obtained as three peaks (Fig. 4B). The multiple peaks are probably due to heterogeneity in the carbohydrate structure and/or to deamidation of glutamine during purification, just as in the case of ceruloplasmin[12]. After we identified GalN in this peptide by amino acid analysis following hydrolysis with 4 *M* hydrochloric acid at 108°C for 6 h, Dr. J. Baenziger of Washington University of St. Louis confirmed its presence and determined the GalN content of the peptide by gas chromatography. Sequence analysis

of hemopexin showed that the GalN-containing peptide is amino-terminal in the protein structure, and sequence and amino acid analysis of the peptide indicated that the GalN oligosaccharide is probably O-linked to the amino-terminal threonine[19].

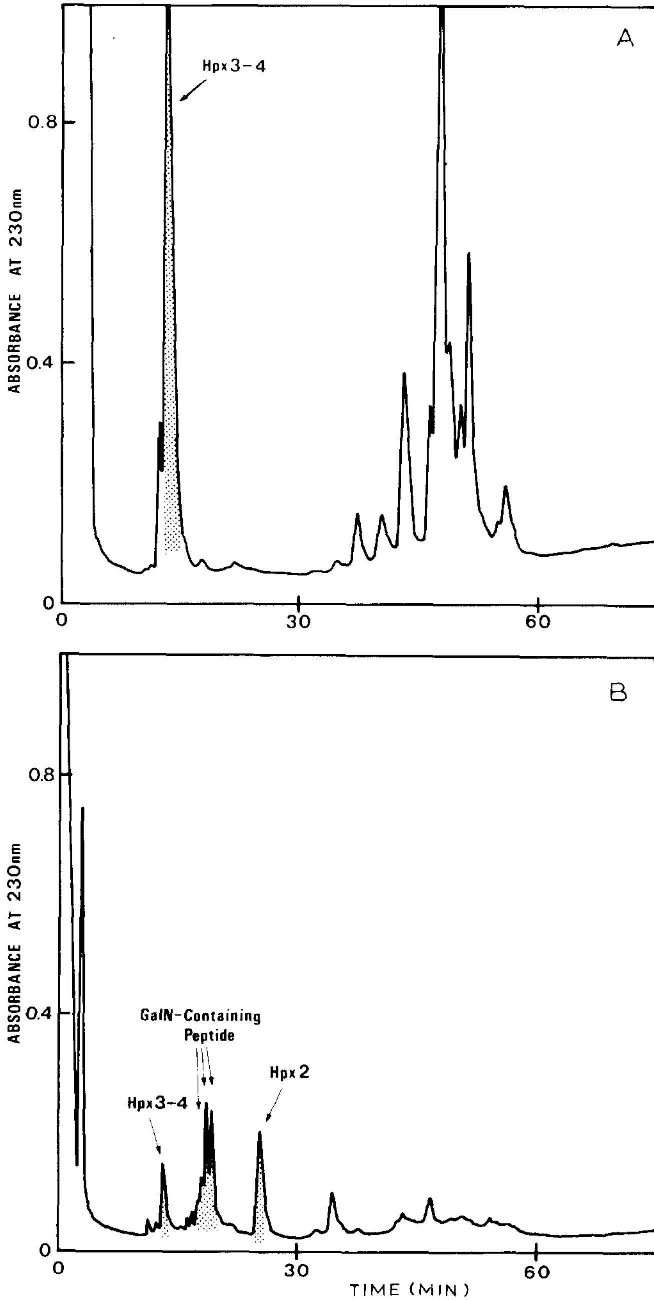

Fig. 4.

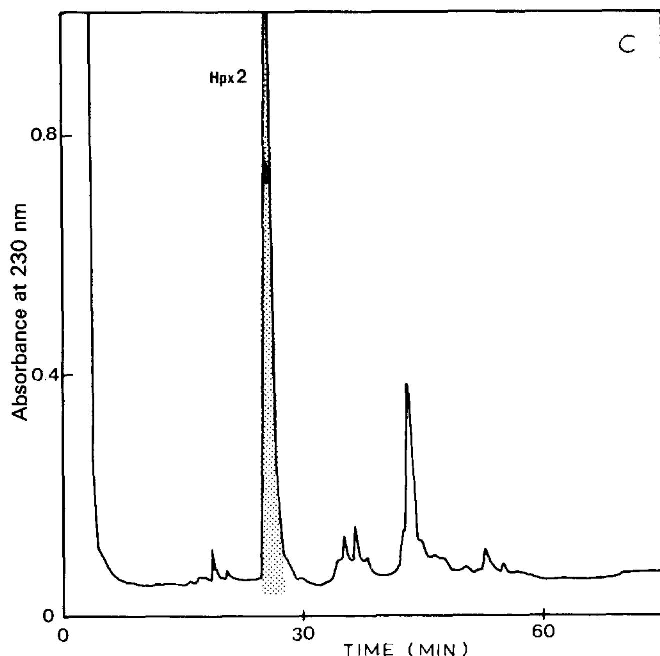

Fig. 4. Purification by HPLC of glycopeptides with a GalN oligosaccharide and with multiple glycosylated sites for GlcN oligosaccharide from fractions of Fig. 3. The glycopeptides are shaded: (A) fraction 2 of Fig. 3; (B) fraction 3; (C) fraction 4.

In addition, although GalN had not previously been identified in 3.1S α-2-leucine-rich glycoprotein[20], we also found a GalN-containing peptide in this protein by use of our method.

The method described is also capable of purifying a glycopeptide with multiple GlcN oligosaccharides. In fact, peptide Hpx3-4, which has two GlcN oligosaccharides, was isolated in a 82% yield from fraction 2 of Fig. 3 by use of HPLC (Fig. 4A). Although this peptide has only 17 amino acid residues, a peptide with 64 amino acid residues was eluted in the same fraction by GPC on a Sephadex G-50 column. Because the contribution of a GlcN oligosaccharide to the molecular weight is about 3000 Daltons, a relatively small peptide with multiple sites for GlcN oligosaccharide such as Hpx3-4, behaves like a large polypeptide on GPC. However, this property is a key to separation of glycopeptides with multiple glycosylated sites from large non-glycosylated peptides, because the elution position of a glycopeptide in reversed-phase column chromatography is determined mainly by its amino acid composition, as described later. Therefore, a peptide glycosylated at multiple sites tends to be eluted early in reversed-phase chromatography, whereas a large non-glycosylated peptide tends to be eluted later, because the large peptide usually is highly hydrophobic. The other hemopexin glycopeptides with a single attachment site for GlcN oligosaccharide were purified from fraction 4 of Fig. 3 (Hpx2 in Fig. 4C) and also from the precipitate (Hpx1 and Hpx5, data not shown). Glycopeptides of hemopexin were also isolated successfully by this method from other digests of the alkylated protein made with *S. aureus* V8 protease, chymotrypsin, and trypsin[19]. The example

illustrated here shows the adaptability of the method to the purification of glycopeptides, even from a peptide mixture obtained after chemical modification of protein. By the same method, a glycopeptide with three GlcN oligosaccharides, produced by CNBr cleavage, was also purified easily from β-2-glycoprotein I[6].

Purification of IgD glycopeptides with multiple sites for GalN oligosaccharide, or with different structural types of GlcN oligosaccharides

The human IgD myeloma protein, designated WAH, for which we have reported the complete amino acid sequence[21], offered an excellent opportunity to test the efficacy of our method. It is the first example of a glycoprotein whose entire primary structure has been defined by using a single source, and it contains several different GalN oligosaccharide structures and several different GlcN structures, all of which were determined on the same protein by a collaborating group[22,23]. A series of four or five GalN oligosaccharides are clustered in the first half of a 64-residue polypeptide, called the hinge region. About 50% of these have the structure Galβ1 → 3 GalNAc, and the remainder also contain one or two residues of N-acetylneuraminic acid. The first GlcN oligosaccharide in the Fc region (here designated D1) is a virtually unprocessed structure of the high-mannose type, whereas the last two GlcN oligosaccharides (D2 and D3) are of the extensively processed complex type and are notable for their extreme degree of microheterogeneity[22]. We previously used HPLC to purify a series of chymotryptic GalN glycopeptides[24] from IgD WAH, and earlier we used conventional methods to purify the GlcN glycopeptides that were used for carbohydrate structural study[22]. In the present work we isolated all of the tryptic glycopeptides of this IgD protein by the new method and have also isolated a nonglycosylated peptide corresponding to D2.

The tryptic digest of the Fc fragment, obtained by limited papain digestion of the IgD protein[3], was separated into five fractions by initial GPC (Fig. 5). The first

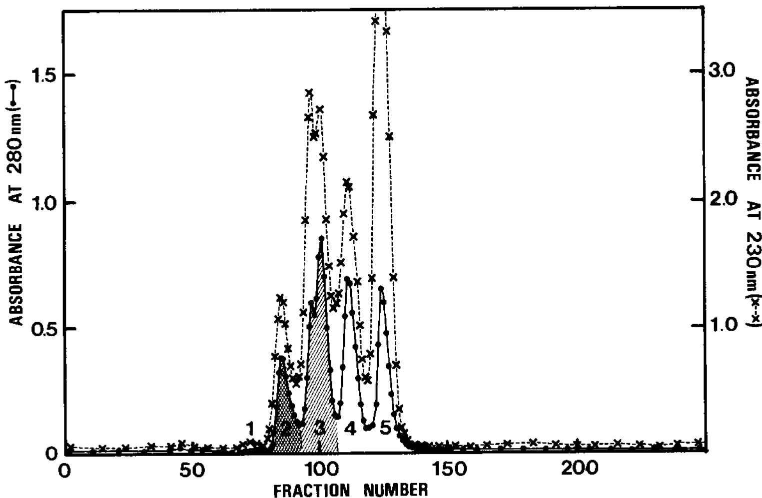

Fig. 5. Separation on a Sephadex G-50 column of a tryptic digest of Fc fragment obtained by limited papain digestion of IgD. Fractions containing hexosamine are cross-hatched; however, fraction 3 is shaded, because it contained less hexosamine.

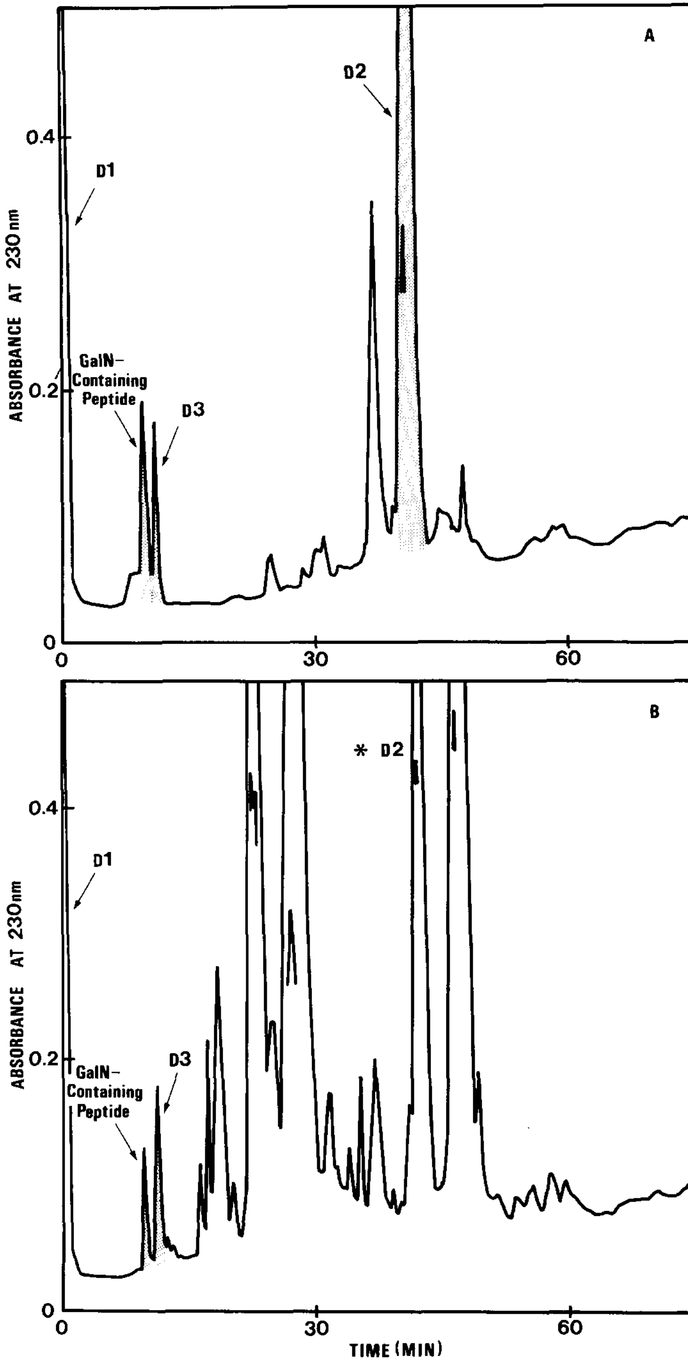

Fig. 6. Purification by HPLC of glycopeptide glycosylated at multiple sites with GalN oligosaccharide and of glycopeptides with high-mannose or complex-type carbohydrate from fractions of fig. 5. The nonglycosylated peptide of D2 is designated *D2. Peaks containing glycopeptide are shaded.

three fractions contained hexosamine. Fraction 1 had no glucosamine and was the pure hinge glycopeptide with all of the GalN oligosaccharides of the protein. The amino acid sequence of this peptide was identical to that of residue 237–266 in the δ-chain[21]. A shorter hinge GalN glycopeptide was isolated from fraction 2 by a second purification step on a Synchropak RP-P column (Fig. 6A). This peptide (residue 249–266) contained three or four attachment sites for GalN oligosaccharides. In previous work[24] the determination of the amino acid sequence of the hinge glycopeptide was very difficult because of the presence of a series of GalN oligosaccharides and the resistance to proteolytic cleavage of the GalN-rich region. The sequence of the tryptic hinge peptide, purified from the protein sample by the new method, was easier to determine; possibly because the protein sample was collected from the serum of the same patient on a different sampling date, the peptide lacked one GalN oligosaccharide (*i.e.* at position Ser-238). The hinge region of IgD seems also to have some heterogeneity in the number and location of attachment sites for GalN oligosaccharides from one protein to another, obtained from different patients with multiple myeloma[25], and even from sample to sample obtained from a single patient on different days. The presence of a GalN glycan at Ser-238 seems to affect the cleavage by papain and also the sequence determination.

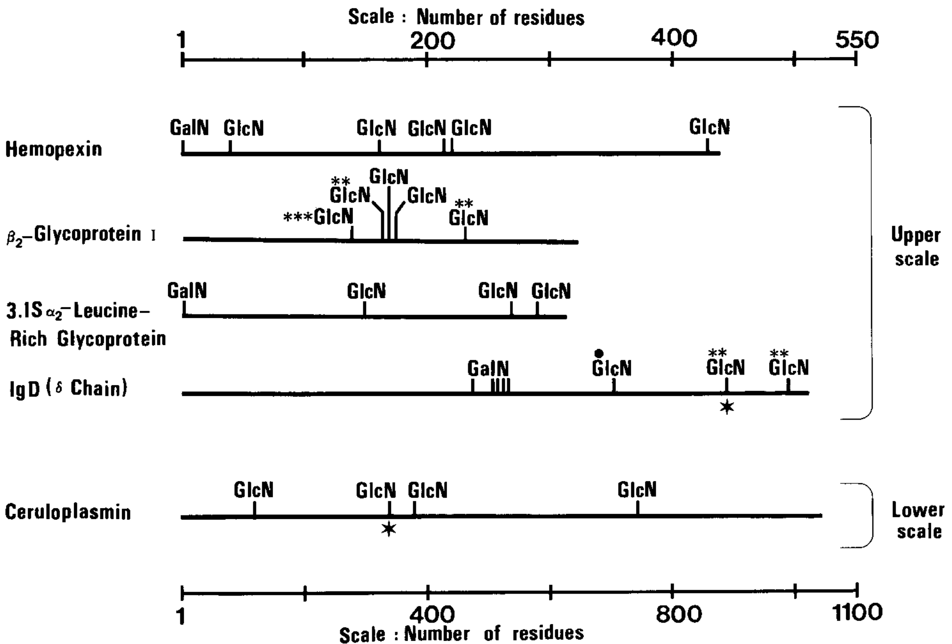

Fig. 7. Linear structural models of human plasma glycoproteins from which glycopeptides were purified. The attachment sites for GalN and GlcN oligosaccharides are indicated by GalN and GlcN. The oligosaccharides for which the structures are known are also distinguished by ** for diantennary, *** for triantennary, and ● for a high-mannose type of GlcN oligosaccharide. If the carbohydrate is missing in some molecules, the site is marked by ★. In the text and in Table I the GlcN-glycopeptides of each protein are numbered in their order from the amino terminus, *e.g.* CP1, CP2, CP3, CP4 in ceruloplasmin. The upper scale for the number of amino acid residues is used for the upper four proteins, and the lower scale is used for ceruloplasmin.

Three glycopeptides, D1, D2 and D3, that contain GlcN oligosaccharide were isolated from fractions 2 and 3 of Fig. 5 by HPLC (Fig. 6). The yields of the purified peptides D1 and D3 were about 70%. The lower yield of D2 (25%) was explained by the isolation of a nonglycosylated peptide in a 25% yield (Fig. 6B). The difference in retention times of the glycosylated peptide (40.5 min) and the non-glycosylated peptide (43.0 min) was 2.5 min (Fig. 6). These peptides could be separated preparatively by GPC on the basis of the large size of the carbohydrate.

CONCLUSIONS

In order to facilitate amino acid sequence analysis of a series of human plasma

TABLE I

RETENTION TIME (t_R), HYDROPHOBICITY, AMINO ACID SEQUENCE, AND CARBOHYDRATE STRUCTURE OF GLYCOPEPTIDES

The hydrophobicity was calculated by the method of Sasagawa *et al.*[27] using their list of non-weighted retention constants. An asterisk (*) indicates the attachment site for GlcN oligosaccharide, and a plus sign (+) for GalN oligosaccharide. The amino acid sequence is given in the standard one-letter code.

Peptide	*t_R (min)*	*Hydrophobicity*	*Sequence*	*Carbohydrate structure*
Ceruloplasmin				
CP1	21.9	114.8	EHEGAIYPDN*TTDFQR	Complex
	23.7			
CP2	37.8	135.8	AGLQAFFQVQECN*K	Complex
CP3	32.1	138.9	EN*LTAPGSDSAVFFEQGTTR	Complex
CP4	35.1	213.6	ELHHLQEQN*VSNAFLDKGEFYIGSK	Complex
	36.5			
	37.5			
Hemopexin				
GalN-peptide	18.0	126.6	T+PLPPTSAHGNVAEGETKPDPDVTER	
	18.6			
	19.2			
Hpx2	25.5	110.8	SWPAVGN*CSSALR	
Hpx3-4	13.2	81.8	N*GTGHGN*STHHGPEYMR	
β_2-Glycoprotein I				
β_2 1	20.4	38.2	KPSAGN*NSLY	Complex, triantennary
β_2 5	25.5	54.9	LGN*W	Complex, diantennary
Immunoglobulin D				
GalN-peptide	9.6	75.6	GSLAKAT+T+APAT+TRNTGR	
D1	0	21.1	HSN*GSQSQHSR	High mannose
D2	40.5	249.3	EVN*TSGFAPARPPPQPGSTTFWAWSVLR	Complex, diantennary
D3	11.1	52.4	TLLN*ASR	Complex, diantennary

glycoproteins, a method combining GPC and reversed-phase HPLC was developed and was used for the purification of the tryptic glycopeptides. Linear structural models of the plasma proteins are shown schematically in Fig. 7, in which the attachment sites for GalN and GlcN oligosaccharides are indicated for each protein. Where the carbohydrate structures are known, they are identified. The method described here was used successfully to purify: (1) glycopeptides with a single attachment site for GlcN or GalN oligosaccharide; (2) glycopeptides glycosylated at multiple sites with either GlcN or GalN oligosaccharides; (3) glycopeptides with microheterogeneity of carbohydrate structures; (4) glycopeptides with high-mannose or complex-type structures of GlcN oligosaccharide. In addition, we could isolate both nonglycosylated and glycosylated peptides from the same region in the δ heavy chain of IgD, and we could identify the oligosaccharide that is missing in a carbohydrate variant of ceruloplasmin. These polymorphisms, which are due only to carbohydrate, were not identified previously by conventional methods. Thus, the heterogeneity in the number and location of attachment sites for both GalN and GlcN oligosaccharides in glycoproteins may be more common than has been recognized.

The retention time (t_R), hydrophobicity, amino acid sequence, and carbohydrate structure of glycopeptides described here are summarized in Table I. Two glycopeptides from β-2-glycoprotein I[26] are added to the table, because their carbohydrate structures are now known. In order to learn about the behavior of glycopeptides in reversed-phase column chromatography, retention times of glycopeptides are plotted against ln $(1 + H)$ in Fig. 8B, in which H indicates the hydrophobicity of the peptide, calculated by the method of Sasagawa *et al.*[27] using their list of non-weighted retention constants. For reference purposes, the retention times of nonglycosylated peptides are also plotted against ln $(1 + H)$ in Fig. 8A. The conditions for reversed-phase chromatography differ from those of Sasagawa *et al.* In our method, a Synchropak RP-P column was used instead of a μBondapak C_{18} column, and the mobile phase modifier was aqueous 1-propanol containing 0.1% TFA instead of aqueous acetonitrile containing 0.07% TFA. However, the retention times of both the non-

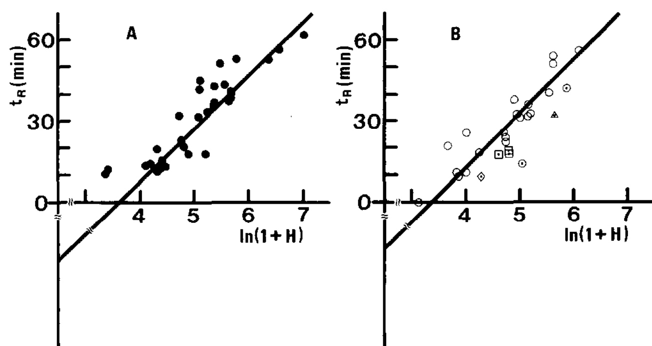

Fig. 8. Relationship between retention time (t_R) and ln $(1 + H)$. The observed retention times are plotted against ln $(1 + H)$, where H indicates the hydrophobicity of the peptide calculated by the method of Sasagawa *et al.*[27] using their list of non-weighted retention constants. (A) Data for nonglycopeptides; (B) data for glycopeptides. ○, Glycopeptides with a single GlcN oligosaccharide; ⊡, glycopeptides with a single GalN oligosaccharide; ⊙, glycopeptides with two GlcN oligosaccharides; △, glycopeptide with a single GalN and a single GlcN oligosaccharide; ◇, glycopeptide glycosylated at multiple sites for GalN oligosaccharide.

glycopeptides and the glycopeptides were linearly related to the natural logarithm of the sum of the retention constant of Sasagawa *et al.*[27] for the constituent amino acids (Fig. 8). The glycopeptides listed in Table I and the other glycopeptides discussed in the text were used for the graph in Fig. 8B, and the non-glycopeptides plotted in Fig. 8A were also obtained from the plasma proteins described here in the process of their sequence analysis.

The behavior of the glycopeptides in reversed-phase chromatography was very similar to that of non-glycosylated peptides, however, the glycopeptides with multiple attachment sites tend to be eluted earlier than expected from the hydrophobicity estimated only from their constituent amino acids (Fig. 8B). Glycopeptide D2 of IgD offers a good example for comparison of the behavior of glycopeptides in reversed-phase chromatography. Although glycopeptide D2 was eluted 2.5 min earlier than the non-glycopeptide *D2, the difference in the elution positions is far less than that expected from the large size of the carbohydrate structure, whose contribution to the molecular weight is about 50% of the total for the glycopeptide. However, even though the effect of the carbohydrate is considerably smaller than expected, it is still possible to use reversed-phase chromatography to separate two peptides that have the same polypeptide structure but differ by the presence or absence of carbohydrate.

ACKNOWLEDGEMENTS

We thank S. A. Dorwin, P. H. Davidson, J. Madison, and K. Huss for their excellent technical assistance. This work was supported by Grants AM 19221 from the National Institutes of Health, IM-2H from the American Cancer Society, and CA08497 from the National Cancer Institute.

REFERENCES

1 F. W. Putnam, in F. W. Putnam (Editor), *The Plasma Proteins*, Vol. 4, Academic Press, New York, 2nd ed., 1984, in press.
2 J. U. Baenziger, in F. W. Putnam (Editor), *The Plasma Proteins*, Vol. 4, Academic Press, New York, 2nd ed., 1984, in press.
3 L.-C. Lin and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 76 (1979) 6572.
4 M. Noyer, F. E. Dwulet, Y. L. Hao and F. W. Putnam, *Anal. Biochem.*, 102 (1980) 450.
5 K. Tsutsui and G. C. Mueller, *Anal. Biochem.*, 121 (1982) 244.
6 J. Lozier, N. Takahashi and F. W. Putnam, *J. Chromatogr.*, 266 (1983) 545.
7 N. Takahashi, R. A. Bauman, T. L. Ortel, F. E. Dwulet, C.-C. Wang and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 80 (1983) 115.
8 Y. Takahashi, N. Takahashi, D. Tetaert and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 80 (1983) 3686.
9 N. Takahashi, T. L. Ortel and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 81 (1984) 390.
10 L. Ryden and D. Eaker, *Eur. J. Biochem.*, 44 (1974) 171.
11 D. Tetaert, N. Takahashi and F. W. Putnam, *Anal. Biochem.*, 123 (1982) 430.
12 T. L. Ortel, *Dissertation*, Indiana University, Bloomington, 1983.
13 K. Yamashita, C.-J. Liang, S. Funakoshi and A. Kobata, *J. Biol. Chem.*, 256 (1981) 1283.
14 J. U. Baenziger and D. Fiete, *Cell*, 22 (1980) 611.
15 L. Broman, *Nature (London)*, 182 (1952) 1655.
16 R. Richterich, A. Temperli and H. Aebi, *Biochim. Biophys. Acta*, 56 (1962) 240.
17 L. Ryden, *Int. J. Protein Research*, 3 (1971) 191.
18 G. Uhlenbruck, I. Reese, P. Vaith and H. Haupt, *J. Clin. Chem. Clin. Biochem.*, 17 (1979) 29.
19 N. Takahashi, Y. Takahashi and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 81 (1984) 2021.

20 H. G. Schwick and H. Haupt, in F. W. Putnam (Editor), *The Plasma Proteins*, Vol. 4, Academic Press, New York, 2nd ed., 1984, in press.
21 N. Takahashi, D. Tetaert, B. Debuire, L.-C. Lin and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 79 (1982) 2850.
22 S. J. Mellis and J. U. Baenziger, *J. Biol. Chem.*, 258 (1983) 11546.
23 S. J. Mellis and J. U. Baenziger, *J. Biol. Chem.*, 258 (1983) 11557.
24 N. Takahashi, D. Tetaert and F. W. Putnam, in M. Elzinga (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1982, p. 463.
25 T. Takayasu, S. Suzuki, F. Kametani, N. Takahashi, T. Shinoda, T. Okuyama and E. Munekata, *Biochem. Biophys. Res. Commun.*, 105 (1982) 1066.
26 J. Lozier, N. Takahashi and F. W. Putnam, *Proc. Nat. Acad. Sci. U.S.*, 81 (1984) 3640.
27 T. Sasagawa, T. Okuyama and D. C. Teller, *J. Chromatogr.*, 240 (1982) 329.