

High-performance liquid chromatographic analysis of carbohydrate mass composition in glycoproteins

M. Kunitani and L. Kresin

Department of Analytical Chemistry, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608 (USA)

ABSTRACT

A method has been developed which characterizes the carbohydrate portion of glycoproteins by its mass composition (mass carbohydrate per mass protein). This size-exclusion HPLC method uses both UV and refractive index detectors in series and requires only a single injection of intact glycoprotein. It has been shown that for calibration purposes any carbohydrate or protein which chromatographs well on the sizing system can be substituted for the carbohydrate or protein moiety of the glycoprotein of interest. The accuracy of this method was demonstrated for RNase B and fetuin. Mass compositions of various monoclonal IgMs were determined with an average intraday precision of approximately 6%; the limit of lot-to-lot resolution. Thus, the primary utility of this simple HPLC method is to identify substantial changes in the carbohydrate moiety among various glycoprotein preparations.

INTRODUCTION

It is known that glycosylation of pharmaceutical proteins can be influenced by many factors, including cell type and cell culture environment, and can result in changes in distribution of glycoforms [1,2]. The carbohydrate moiety can have a significant effect on physical, immunological and biological properties of glycoproteins, including solubility, specific activity and clearance rate from the circulatory system [3,4].

Determination of the exact amount of carbohydrate in a glycoprotein still remains a significant challenge in spite of recent advances in methodology [5]. It usually consists of a difficult series of enzymatic and/or chemical cleavages followed by chromatographic quantitation of the various monosaccharides. The lack of quantitative cleavage of carbohydrate from the protein and chemical degradation of some monosaccharides during oligosaccharide hydrolysis leads to uncertainty as to the

accuracy of the final result. Mass spectrometry and enzymatic techniques are used to yield sequencing information but the precise distribution of the various glycoforms is not determined. In fact, rigorous validation of accuracy for carbohydrate analysis of any complex glycoprotein containing a variety of glycoforms remains unsettled. This paper presents a simple size-exclusion HPLC (SEC) method which determines the carbohydrate mass composition (mass carbohydrate/mass protein) for intact glycoproteins, thus circumventing the potential for incomplete deglycosylation and monosaccharide degradation, weaknesses of traditional compositional analysis. This analysis is not designed to be comprehensive, as the single value of mass composition does not provide detailed characterization of glycoprotein structure. However, it does provide a simple and useful method of comparing the degree of glycosylation in glycoprotein pharmaceuticals.

THEORY

This method is identical to that developed for characterization of polyethylene glycol (PEG)-proteins [6], which uses SEC with UV and refractive

Correspondence to: M. Kunitani, Department of Analytical Chemistry, Chiron Corporation, 4560 Horton Street, Emeryville, CA 94608, USA.

index (RI) detectors in series. The UV detector at 280 nm selectively detects only the aromatic residues in the protein moiety of a glycoprotein, as sugars do not absorb at this wavelength. The RI detector responds to the mass of protein and of the carbohydrate. The protein contribution to the RI signal is calculated from the protein concentration calibration derived from the UV detector data; the remaining RI response is due to the carbohydrate. The fundamental assumption is that the refractive index of the glycoprotein solution (N_{GP}), can be expressed as the sum of the refractive indices of its components, P (protein) and G (carbohydrate):

$$N_{GP} = N_0 + C_P \left(\frac{dn}{dc} \right)_P + C_G \left(\frac{dn}{dc} \right)_G \quad (1)$$

where N_0 is the refractive index of pure solvent; $(dn/dc)_P$ and $(dn/dc)_G$ are the specific refractive index increments of protein and carbohydrate, respectively, and C_P and C_G are their mass concentrations. Since the protein and oligosaccharide moieties are relatively large polymers linked with a single covalent bond, the assumption of eqn. 1 should be accurate. Eqn. 1 can be rearranged to give the working eqn. 2. The detail for this derivation is presented in ref. 6.

$$\frac{W_G}{W_P} = \frac{\frac{\text{RI area glycoprotein}}{\left(\frac{\text{UV area glycoprotein}}{\text{UV area/mg protein}} \right)} - \text{RI area/mg protein}}{\text{RI area/mg carbohydrate}} \quad (2)$$

Determination of glycoprotein mass composition requires calibration of the SEC system with the protein and carbohydrate moieties separately, which will yield the RI and UV peak area/mg terms (calibration slopes). A single injection of the intact glycoprotein will give the UV and RI area for the glycoprotein and allow calculation of the mass composition (W_G/W_P). There are many instances when it is not possible nor practical to calibrate this mass composition analysis with the deglycosylated portion of the glycoprotein. In these instances another non-glycosylated protein can be used as a calibrator. Eqn. 2 then becomes:

$$\frac{W_G}{W_P} = \frac{\frac{\text{RI area glycoprotein}}{\left(\frac{\text{UV area glycoprotein}}{\text{UV area/mg calib. protein} \cdot K} \right)} - \text{RI area/mg calib. protein} \cdot K'}{\text{RI area/mg carbohydrate}} \quad (3)$$

The K and K' in eqn. 3 represent the respective correction factors for differences in UV 280 nm and RI response between the glycoprotein of interest and the calibration protein.

$$K = \frac{E^{0.1\%} \text{ glycoprotein}}{E^{0.1\%} \text{ calib. protein}}$$

and

$$K' = \frac{\left(\frac{dn}{dc} \right)_{\text{glycoprotein}}}{\left(\frac{dn}{dc} \right)_{\text{calib. protein}}} \approx 1 \quad (4)$$

where K is the ratio of extinction coefficients and K' is the ratio the refractive index increment for the glycoprotein of interest to that of the calibration protein. The assumption that $K' \approx 1$ will be evaluated in this paper. The purpose of this work is to establish the validity of the RI/UV method for compositional analysis of glycoproteins and to establish its precision for typical glycoproteins.

EXPERIMENTAL

Materials

Sialic acid, N-acetylglucosamine (GlcNAc), maltose, maltotriose, maltotetraose, ribonuclease A (RNase A), and ribonuclease B (RNase B) from bovine pancreas were obtained from Sigma (St. Louis, MO, USA). Glucose and galactose were from Applied Science Labs. (State College, PA, USA). Fucose was from United States Biochemical Corp. (Cleveland, OH, USA). The dextran 1000 ($M_n = 1010$, $M_w = 1270$) and dextran 6000 ($M_n = 6100$, $M_w = 6700$) were obtained from American Polymer Standards Corp. (Mentor, OH, USA). Bovine serum albumin (BSA) was from National Institute of Standards and Technology (Gaithersburg, MD, USA). IgM samples were from Chiron (Emeryville, CA, USA) and Rockland (Gilbertsville, PA, USA). The source of bovine fetuin was from Gibco Labs. (Grand Island, NY, USA) and was obtained as a gift from the laboratory of Professor R. Townsend [7]. Interleukin-2, a human recombinant protein was produced in *E. coli* (Chiron) and contains no carbohydrate.

Chromatography was performed with system consisting of a SP 8700 pump (Spectra-Physics, San Jose, CA, USA), a WISP 710 B injector (Millipore-Waters, MA, USA), an ERC-7510 RI detector (Erma, Tokyo, Japan), a LC135 UV detector (Perkin-Elmer, Norwalk, CT, USA) set at 280 nm, and a Nelson Analytical 6000 data system (Perkin-Elmer/Nelson, Cupertino, CA, USA). Due to pressure limitations on most RI detector flow cells, the UV detector always proceeds the RI detector in a series configuration.

Methods

Carbohydrates were run on two different HPLC columns based upon their selectivity. Monosaccharide chromatography was performed on an Aminex HPX-87H (300 x 7.8 mm I.D.) column (Bio-Rad, Richmond, CA, USA) with an RI detector. The eluent was 0.005 M H₂SO₄ at a flow-rate of 0.6 ml/min. The column temperature was 45°C and the injection volume was 20 µl. Oligosaccharides were separated from each other and from monosaccharides using an Aminex HPX-42A column (Bio-Rad). The eluent was water at a flow-rate of 0.6 ml/min. The column temperature was 85°C and injection volume was 20 µl.

For the protein and glycoprotein separations, the SEC columns were chosen on their ability to achieve adequate resolution of the analyte proteins from RI system peaks at the included or excluded volume of the column. A Superose 12 column (Pharmacia LKB, Piscataway, NJ, USA) was used to compare the RI response of various proteins. The eluent was 0.1 M sodium sulphate and 0.01 M sodium phosphate (pH 7.0) at flow-rate of 0.5 ml/min. Literature extinction coefficients were used to determine protein concentrations and 100 µl of each protein were injected separately [S-Io]. The chromatographic system used for mass composition analysis of RNase B consisted of two Zorbax GF-250 (25 cm x 9.4 mm I.D.) columns (DuPont, Wilmington, DE, USA) at room temperature using 0.5 M Na₂SO₄, 0.01 M phosphate buffer, pH 7.0, at a flow-rate of 0.5 ml/min. Calibration curves were run for dextran 1000 (as a carbohydrate calibrator) and for RNase A (as a deglycosylated protein calibrator). Repurified RNase B as well as each standard were run in triplicate. The injection volume was 100 µl. A Superose 6 (Pharmacia) column was used for analy-

sis of IgMs using 50 mM (NH₄)₂SO₄, 50 mM phosphate buffer (pH 7.5) at an eluent of a flow-rate of 0.4 ml/min at room temperature. Fetuin was analyzed with a Zorbax GF-250 column using the same eluent at 0.8 ml/min. The injection volume was 100 µl. Dextran 6000 was used as a carbohydrate calibrator and BSA as a protein calibrator for both IgM and fetuin analyses. A generic extinction coefficient of $E^{0.1\%} = 1.4$ was used for IgM analyses while a theoretical value of $E^{0.1\%} = 1.67$ was used for fetuin [5].

RNase B was purified by chromatography at room temperature on a Con A-Sepharose column (4 cm x 2.6 cm I.D.) (Pharmacia, Uppsala, Sweden) equilibrated in 0.1 M Tris-HCl, pH 7.4, containing 1 M NaCl, 1 mM CaCl₂, 1 mM MgCl₂, and 1 mM MnCl₂[11]. The column was loaded with 40 mg of RNase B dissolved in 3.0 ml of eluent, which was eluted at 0.1 ml/min. Following collection of the unretained material (RNase A), a step elution was performed with 10% cc-methyl-D-mannopyranoside in the same eluent. The collected peaks were dialyzed against 50 mM phosphate buffer, pH 6.8. Commercial RNase B and the two peaks collected from the Con-A Sepharose column were analyzed on a pre-cast 10–20% gradient gel (Daiichi, Tokyo, Japan) stained with Coomassie Blue.

For a detailed description of the UV-RI SEC calibration and a analyses procedure see ref. 6. Linear regression analyses were performed on the calibration data (0.2–1.0 mg/ml BSA and dextran 6000) generated daily with each set of samples. If encountered, aggregate shoulders on glycoprotein peaks were included in the integration of analyte peak area.

RESULTS AND DISCUSSION

Carbohydrate calibration

The carbohydrate moiety of any glycoprotein consists of various monosaccharides linked together to form one or more oligosaccharides. Hence the RI response of monosaccharides, oligosaccharides and proteins were independently investigated. The RI increment (dn/dc) is a measure of electron density of a molecule and is not affected by molecular conformation [12]. Since carbohydrates are largely composed of repeating hydroxymethylene (CHOH) functionality, all carbohydrates should have identi-

in carbohydrate chemistry that the RI response of carbohydrates is a measure of mass concentration and varies little among monosaccharides or with the carbohydrate size [13,14]. Phosphorylated or sulphonated carbohydrates have been excluded from consideration in this report. Thus, any carbohydrate with an appropriate retention time can be used as a calibration substitute for the carbohydrate moiety in the UV-RI SEC mass composition analysis of a glycoprotein. This report utilizes commercially available dextran samples as carbohydrate calibrators which are chosen to be baseline resolved from included or excluded volume RI system peaks.

Protein calibration

The most accurate method of protein calibration is to use the deglycosylated protein of interest. However, there are some potential complications with this approach, such as possible insolubility of the deglycosylated protein or lack of complete deglycosylation. For proteins, the main source of electron density and hence the RI response is the amide peptide bond, which is proportional to the number of amino acid residues in the protein. Except for proteins of highly unusual amino acid composition, as a first approximation the number of amino

acid residues is proportional to the protein's molecular mass. Thus, most proteins should have similar RI response per mass [15]. Table III shows RI standard curves for several unglycosylated proteins. The correlation coefficients of the standard curves establish linearity and the slopes are similar, with an R.S.D. of 2.3%, validating the assumption that the RI response of proteins is largely independent of composition and molecular mass, and that $K' \cong 1$ in eqn. 4. Thus, any convenient protein whose extinction coefficient is known may be used for calibration with only a small uncertainty in accuracy. This approach to calibration does not affect assay precision.

Validity of compositional analysis

In order to test the validity of the UV-RI SEC method, the model proteins RNase A and RNase B were used. Both proteins are very well characterized [16]. They have identical amino acid sequences. RNase B contains a single oligosaccharide chain at Asn 35, composed of six mannose and two GlcNAc residues per molecule. RNase A does not have any carbohydrate and represents unglycosylated RNase B. Thus, RNase A can be used as a protein calibrator for RNase B. By this approach the deglycosylation

TABLE III
RI RESPONSE OF PROTEINS

Protein	MW	Concentration (mg/ml)	Peak area · 10 ⁶	R.S.D. (%) (n = 3)	Slope (area/mg I 0 ⁶)	r ²
RNase	13 700	0.16	1.1	—	7.4	0.999
		0.41	2.9	0.6		
		0.81	6.0			
BSA	67 500	0.18	1.1	0.7	7.3	0.999
		0.43	3.0			
		0.93	6.7			
Myoglobin	16 900	0.16	1.0	1.3	7.5	0.999
		0.40	2.9			
		0.81	6.0			
Interleukin-2	15 300	0.22	1.4	—	7.7	0.999
		0.56	4.1			
		1.12	8.5			
Average					7.5	
R.S.D.					2.3%	

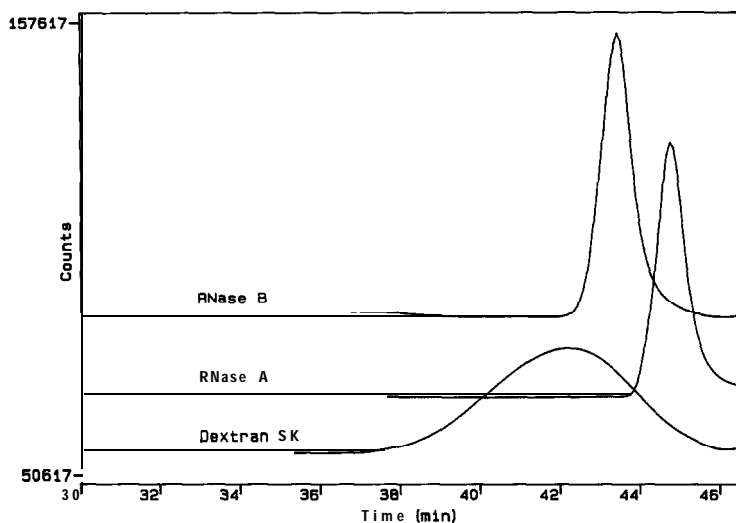


Fig. 3. Refractive index chromatograms for the mass composition analysis of RNase B. RNase A calibration 80 μg , dextran 6000 (6K) calibration 80 μg , purified RNase B analysis 50 μg . HPLC conditions are described in Experimental.

of RNase B and resulting validation of complete carbohydrate cleavage and potential monosaccharide degradation during hydrolysis can be avoided.

Commercial preparations of RNase B contain substantial amounts of RNase A although RNase A preparations are free of RNase B. Thus RNase B was repurified on a Con-A Sepharose affinity column. The RNase A passed through the column, while RNase B was selectively bound by lectin, and was eluted with 10% α -methyl-D-mannopyranoside. The purity of the RNase A and RNase B preparations were verified by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) (not shown). In order to fully separate the relatively small RNase proteins from the included volume, two GF-250 columns in series were used as a sizing system to determine W_G/W_P of RNase B. RNase A and dextran 6000 were used as the protein and the carbohydrate calibrators of the SEC system (Fig. 3). After injection of purified RNase B, its mass composition was determined using eqn. 2. The mass composition (W_G/W_P) was experimentally determined to be 0.11 (R.S.D. = 13%, $n=3$), which is comparable to the literature value of 0.10 [15], demonstrating the accuracy of the UV-RI SEC method in measuring the carbohydrate/protein mass ratio of a glycoprotein.

A problem for the validation of the mass composition method for large, more complex glycoproteins is the lack of a well characterized standard. The same glycoprotein from different sources is likely to contain different types and proportions of oligosaccharide structures, making comparative analyses difficult [5]. Bovine fetuin is one of the most frequently studied glycoproteins with complex oligosaccharide structures and, in spite of the variability of glycosylation in different preparations, it was chosen for study by the UV-RI SEC mass composition method. Mass spectral studies have characterized fetuin as containing three triantennary N-linked and three O-linked oligosaccharides resulting in a theoretical mass composition of 0.29 [17-19]. BSA and dextran 6000 were used as calibrators for the UV-RI SEC mass composition analysis of fetuin, which was found to be 0.18 ± 0.02 ($n=8$) (Fig. 4). Chemical hydrolysis on the identical lot of fetuin followed by high-pH anion-exchange chromatography (HPAEC)-pulsed amperometric detection (PAD) monosaccharide analysis yielded a weight composition of 0.18, comparable to the UV-RI SEC approach [7]. This result is consistent with other HPLC-derived monosaccharide analyses of fetuin, although sialic acid quantitation is generally not reported, making comparable weight com-

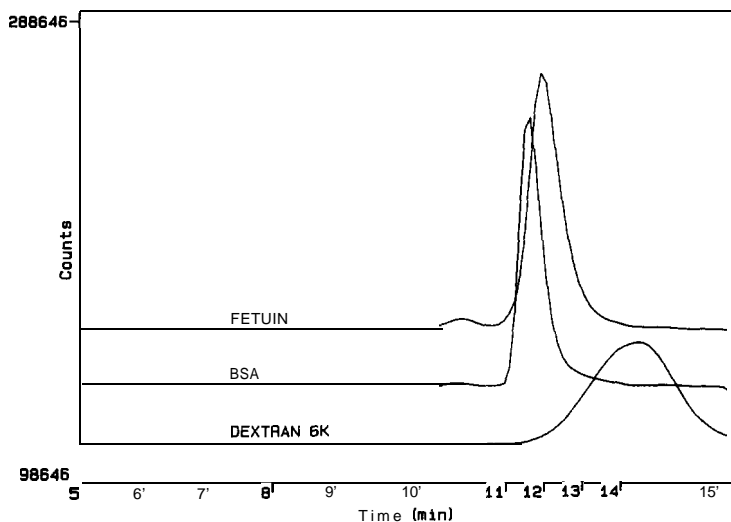


Fig. 4. Refractive index chromatograms for the mass composition analysis of fetuin. BSA protein calibration 48 μg , dextran 6000 calibration 80 μg , fetuin analysis 100 μg . HPLC conditions are described in Experimental.

position calculations impossible [5]. The causes for the discrepancy between theoretical and reported monosaccharide compositions in complex glycoproteins are unknown and are further complicated by variation in glycoprotein preparations and differences among various monosaccharide methods.

Mass composition of IgM

It has been shown that changes in fermentation conditions can alter the distribution and levels of carbohydrates on a glycoprotein produced by mammalian cell culture [1,4]. The UV-RI SEC mass composition analysis was one method which was

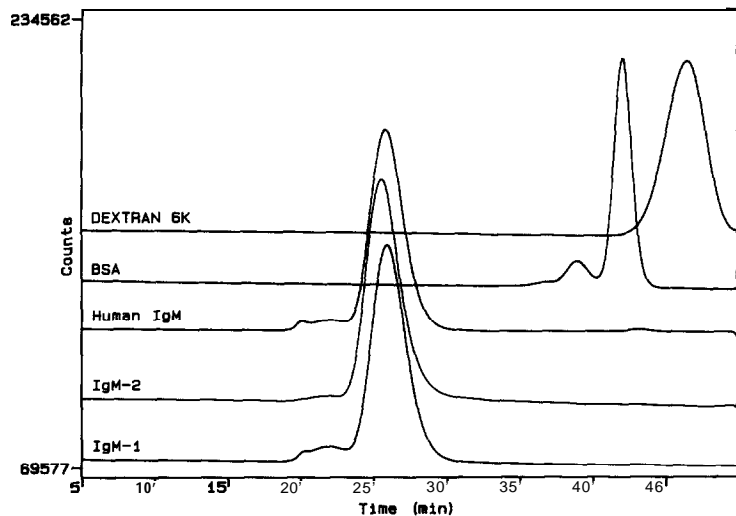


Fig. 5. Refractive index chromatograms for the mass composition analysis of three different IgM antibodies. BSA protein calibration 40 μg , dextran 6000 calibration 70 μg , IgM-1 analysis 90 μg , IgM-2 analysis 100 μg , human IgM analysis 100 μg . HPLC conditions are described in Experimental.

applied to unrelated IgMs produced from cell culture and another from human serum, in order to characterize potential changes in these complex glycoproteins. Calibration of the SEC system was done with dextran 6000 and BSA while a generic $E^{0.1\%}$ for IgM was used with eqn. 3. Although this calibration approach has uncertainties as to the accuracy of the final value, the precision is not affected. The large-pore Superose 6 column was used to chromatograph the large IgM (MW \approx 900 000) glycoprotein between the included and excluded volume of the SEC column (Fig. 5). The results in Table IV shows that the three different IgM antibodies, IgM-1, IgM-2 and commercial human IgM, could easily be distinguished from each other. The higher mass composition IgM-1 over that of IgM-2 correlate with observations such as the lower clearance rate and higher aqueous solubility of IgM-1 compared to that IgM-2 [3]. These results

also show that analyses of different fermentation/purification lots of these IgMs were indistinguishable from each other by the mass composition criteria. Samples 5 and 6 of IgM-1 in Table IV represent, respectively, a 10-fold change in fermentation scale and a different clone, establishing constancy in overall carbohydrate content during these process changes. The average intraday R.S.D. for IgM-1 was 4%, which probably represents the limit of sample-to-sample resolution of the technique. The average intraday R.S.D. for IgM-2 was higher (7%) showing that the method has poorer precision at lower mass composition values. The sensitivity of the mass composition analysis is affected by many factors including the mass composition value, the SEC glycoprotein recovery, the protein extinction coefficient, the SEC column efficiency and RI noise of the HPLC system. Using IgM-1 and IgM-2 as glycoproteins of representative mass composition, the sensitivity was approximately 4 μ g antibody with a limiting RI signal-to-noise ratio of 10. The method is not designed to detect monosaccharide substitution, nor is it precise enough to determine subtle alterations in glycosylation patterns. However, it is precise enough to identify significant changes in mass fraction carbohydrate.

TABLE IV
MASS COMPOSITION OF THREE DIFFERENT IgM ANTIBODIES

Lots 1-6 and 11-16 represent independent fermentation and purification preparations of IgM-1 and IgM-2, respectively. For HPLC conditions, see Experimental.

Fermentation lot	Mass composition W_G/W_P	R.S.D. (%)	<i>n</i>
<i>IgM-1</i>			
1	0.28	11	3
2	0.27	6	3
3	0.28	2	4
4	0.26	4	6
5	0.26	0	3
6	0.26	2	3
Average	0.26	4	-
<i>IgM-2</i>			
11	0.10	10	3
12	0.12	13	4
13	0.11	5	3
14	0.10	6	3
15	0.10	6	3
16	0.13	4	3
Average	0.11	7	
Human IgM	0.33		

CONCLUSIONS

The UV-RI SEC mass composition method for characterizing glycoproteins has been validated for accuracy with ribonuclease and fetuin, two highly studied glycoproteins. The primary utility of this method is to identify substantial changes in the carbohydrate moiety of various glycoprotein preparations. The precision of the method is reasonable, but complications in accuracy may arise in the protein calibration when using the deglycosylated protein moiety. In addition, inaccuracies in the glycoprotein extinction coefficient will cause ambiguity as to the absolute value of carbohydrate mass composition. An asset to this method is its simplicity; requiring only the addition of an RI detector to the usual SEC system. It is not meant to replace monosaccharide analysis, oligosaccharide mapping or mass spectral identification of glycopeptides. However, it will be a useful analysis to complement the sizing information generated on SEC.

ACKNOWLEDGEMENTS

We wish to thank Ira Krull for his suggestions, which initiated work on this application, Gavin Dollinger for helpful discussions and R. Townsend for the fetuin samples and analysis.

REFERENCES

- 1 C. F. Goochee, M. J. Gramer, D. C. Andersen, J. B. Bahr and J. R. Rasmussen, *Bio/Technology*, 9 (1991) 1347.
- 2 R. B. Parekh, R. A. Dwek, J. R. Thomas, G. Opdenakker, T. W. Rademacher, A. J. Wittwer, S. C. Howard, R. Nelson, N. R. Siegel, M. G. Jennings, N. K. Harakas and J. Feder, *Biochemistry*, 28 (1989) 7644.
- 3 S. S. Gauny, J. Andya, J. Thomson, J. D. Young and J. L. Winkelhake, *Hum. Antibod. Hybridomas*, 2 (1991) 33.
- 4 M. Takeuchi, S. Takasaki, M. Shimada and A. Kobata, *J. Biol. Chem.*, 265 (1990) 12127.
- 5 R. R. Townsend, in L. Etre and Cs. Horváth (Editors), *HPLC of Glycoconjugates and Preparative/Process Chromatography of Biopolymers*, American Chemical Society, Washington, DC, in press.
- 6 M. Kunitani, G. Dollinger, D. Johnson and L. Kresin, *J. Chromatogr.*, 588 (1991) 125.
- 7 R. R. Townsend, personal communication.
- 8 T. A. Bewley, *Anal. Biochem.*, 123 (1982) 55.
- 9 C. M. Harris and R. L. Hill, *J. Biol. Chem.*, 244 (1969) 2195.
- 10 T. Arakawa, T. Boone, J. M. Davis and W. C. Kenny, *Biochemistry*, 25 (1986) 8274.
- 11 J. Baynes and F. Wold, *J. Biol. Chem.*, 251 (1976) 6016.
- 12 C. Tanford, *Physical Chemistry of Macromolecules*, New York, 1961, p. 116.
- 13 F. Stolle, *Z. Ver. Dtsch. Zucker Ind.*, 51 (1901) 469.
- 14 L. Tolman and W. Smith, *J. Am. Chem. Soc.*, 28 (1906) 1476.
- 15 P. Kratochvil, *Classical Light Scattering from Polymer Solutions*, Elsevier, Amsterdam, 1987, p. 3 14.
- 16 T. Plummer and C. Hirs, *J. Biol. Chem.*, 239 (1964) 2530.
- 17 R. G. Spiro, *J. Biol. Chem.*, 237 (1962) 382.
- 18 R. R. Townsend, M. R. Hardy, T. C. Wong and Y. C. Lee, *Biochemistry*, 25 (1986) 5725.
- 19 R. R. Townsend, M. R. Hardy, D. Cummmg, J. P. Carver and B. Bendiak, *Anal. Biochem.*, 182 (1989) 1.