

- Rose, M. C., Voter, W. A., Sage, H., Brown, C. F., & Kaufman, B. (1984) *J. Biol. Chem.* 259, 3167-3172.
- Schimmel, P. R., & Flory, P. J. (1968) *J. Mol. Biol.* 34, 105-120.
- Sheehan, J. K., & Carlstedt, I. (1984a) *Biochem. J.* 217, 93-101.
- Sheehan, J. K., & Carlstedt, I. (1984b) *Biochem. J.* 221, 499-504.
- Sheehan, J. K., & Carlstedt, I. (1987) *Biochem. J.* 245, 757-762.
- Sheehan, J. K., Oates, K., & Carlstedt, I. (1986) *Biochem. J.* 239, 147-153.
- Shogren, R. L. (1985) Ph.D. Thesis, Case Western Reserve University.
- Shogren, R. L., Jamieson, A. M., Blackwell, J., & Jentoft, N. (1986) *Biopolymers* 25, 1505-1517.
- Shogren, R. L., Jentoft, N., Gerken, T. A., Jamieson, A. M., & Blackwell, J. (1987) *Carbohydr. Res.* 160, 317-327.
- Slayter, H. S., & Codington, J. F. (1973) *J. Biol. Chem.* 248, 3405-3410.
- Smith, J. A., & Pease, L. G. (1980) *CRC Crit. Rev. Biochem.* 8, 315.
- Stokke, B. T., Elgsaeter, A., Akjak-Braek, G., & Smidsrod, O. (1987) *Carbohydr. Res.* 160, 13-28.
- Svennerholm, L. (1958) *Acta Chem. Scand.* 12, 547-554.
- Tanford, C., Kawahara, K., & Lapanje, S. (1966) *J. Biol. Chem.* 241, 1921-1922.
- Tiffany, M. L., & Krimm, S. (1969) *Biopolymers* 8, 347-359.
- Woodward, H. D., Ringler, N. J., Selvakumar, R., Simet, I. M., Bhavanandan, V. P., & Davidson, E. A. (1987) *Biochemistry* 26, 5315-5322.
- Yang, J. T., Chuen-Shang, C. W., & Martinez, H. M. (1986) *Methods Enzymol.* 130, 208-269.
- Yamakawa, H. (1984) *Annu. Rev. Phys. Chem.* 35, 23-47.
- Yamakawa, H., & Fujii, M. (1973) *Macromolecules* 6, 407-415.

Effects of Glycosylation on the Conformation and Dynamics of O-Linked Glycoproteins: Carbon-13 NMR Studies of Ovine Submaxillary Mucin[†]

Thomas A. Gerken,* Kenneth J. Butenhof,[‡] and Randal Shogren^{†,§}

W. A. Bernbaum Center for Cystic Fibrosis Research, Department of Pediatrics and Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Received September 27, 1988; Revised Manuscript Received February 23, 1989

ABSTRACT: Carbon-13 NMR spectroscopic studies of native and sequentially deglycosylated ovine submaxillary mucin (OSM) have been performed to examine the effects of glycosylation on the conformation and dynamics of the peptide core of O-linked glycoproteins. OSM is a large nonglobular glycoprotein in which nearly one-third of the amino acid residues are Ser and Thr which are glycosylated by the α -NeuNAc(2-6) α -GalNAc- disaccharide. The β -carbon resonances of glycosylated Ser and Thr residues in intact and asialo mucin display considerable chemical shift heterogeneity which, upon the complete removal of carbohydrate, coalesces to single sharp resonances. This chemical shift heterogeneity is due to peptide sequence variability and is proposed to reflect the presence of sequence-dependent conformations of the peptide core. These different conformations are thought to be determined by steric interactions of the GalNAc residue with adjacent peptide residues. The absence of chemical shift heterogeneity in apo mucin is taken to indicate a loss in the peptide-carbohydrate steric interactions, consistent with a more relaxed random coiled structure. On the basis of the ¹³C relaxation behavior (*T*₁ and NOE) the dynamics of the α -carbons appear to be unique to each amino acid type and glycosylation state, with α -carbon mobilities decreasing in the order Gly > Ala = Ser > Thr >> monoglycosylated Ser/Thr \approx disaccharide linked Ser/Thr. The α -carbons of glycosylated Ser and Thr are considerably more constrained than their nonglycosylated counterparts in apo mucin, while the effects of carbohydrate side chain length (i.e., asialo vs native mucin) on the dynamics of the Ser and Thr residues is relatively small. The nonglycosylated Gly residue also exhibits an increase in motion upon removal of GalNAc; thus, the effects of glycosylation extend to residues beyond the amino acids directly bound to carbohydrate. These results are consistent with the changes in molecular dimensions determined by light-scattering techniques for the same series of modified mucins [Shogren et al. (preceding paper in this issue)]. Taken together, these results further demonstrate that mucins possess a highly expanded conformation that is dominated by steric interactions between the peptide core and the O-linked GalNAc residue.

It has been suspected for some time that the presence and nature of the O-linked oligosaccharide side chains in mucous

[†] This work was supported by grants from the Rainbow Chapter of the Cystic Fibrosis Foundation and the United Way of Cleveland, by a Research Grant from the Cystic Fibrosis Foundation, and by NIH Grants DK 39918 and DK 27651.

* Address correspondence to this author.

[‡] Fellow of the Cystic Fibrosis in Pulmonary Disease training grant, NHLBI-HL 07415.

[§] Present address: USDA, Northern Regional Research Center, Peoria, IL 61604.

glycoproteins (mucins) play an important role in governing their viscoelastic properties. Presently, however, very little information is available that describes the effects of O-glycosylation on the conformation and dynamics of O-linked glycoproteins. Likewise, little is known of the mechanism of how the carbohydrate or peptide structure may contribute to the observed physical properties of mucin glycoproteins. To obtain a more complete understanding of the effects of mucin glycosylation on mucin solution structure at both the atomic and macromolecular levels, we have undertaken parallel

carbon-13 NMR and light-scattering (Shogren et al., 1989) studies of native and sequentially deglycosylated ovine submaxillary mucin (OSM).¹ These studies complement our earlier NMR (Gerken & Dearborn, 1984; Gerken, 1986; Gerken & Jentoft, 1987) and light-scattering studies (Shogren et al., 1985, 1986, 1987) with OSM and other more complex mucins.

Mucins make up a class of high molecular weight ($>10^6$) extensively O-glycosylated glycoproteins that are primarily responsible for the protection-rendering viscoelastic properties of mucous secretions, without which the functions of the digestive, respiratory, and urogenital tracts would be severely impaired. Hydrodynamic studies indicate that mucins have random coil structures that are significantly more expanded than denatured globular proteins [for a review see Carlstedt et al. (1985)]. Spectroscopic studies such as NMR and circular dichroism further indicate that mucins are relatively devoid of α - and β -secondary structures and are internally flexible molecules [see Shogren et al. (1989)]. The detailed molecular basis for the expanded structure of mucins remains largely unknown.

OSM is well suited for study by NMR because its carbohydrate side chains are nearly exclusively the disaccharide α -NeuNAc(2-6) α -GalNAc (Gottschalk & Bhargava, 1972). This relatively simple and homogeneous side-chain structure coupled with the ability to sequentially remove the NeuNAc and GalNAc residues with specificity, giving the monosaccharide (asialo) and the deglycosylated (apo) mucin, respectively, makes OSM an ideal system for the study of the effects of carbohydrate on the solution structure of mucins. The OSM peptide core, typical of most mucins, is rich in glycine (19%), serine (18%), threonine (14%), alanine (14%), and proline (11%) with over 75% of the serine and threonine residues containing carbohydrate.

In the preceding paper (Shogren et al., 1989) we reported the results of light-scattering studies on apo, asialo, and native OSM and our success at fitting the data to models describing the solution properties of polyelectrolyte chains. On the basis of the analysis we concluded that mucin glycosylation causes a global conformational change in the peptide that significantly extends the peptide core and increases chain stiffness. The GalNAc residues are shown to produce a ~ 2 -fold increase in chain dimensions compared to apo mucin, while the addition of NeuNAc to the GalNAc residue produces a smaller ~ 1.5 -fold increase in chain dimensions on the basis of the values of the root mean square radius of gyration, $\langle R_g^2 \rangle^{1/2}$, and hydrodynamic radius, R_h .

To confirm and better understand the light-scattering results, we have further examined the conformation and dynamics of the individual carbons of the OSM peptide core as a function of glycosylation using carbon-13 NMR spectroscopy. Because of a lack of sufficient resolution of the peptide core resonances, previous proton NMR studies on apo, asialo, and native OSM (Gerken, 1986) were unable to detect conformational changes in the mucin peptide core as function of O-glycosylation. However, as described in this paper, by ^{13}C NMR it is clear that glycosylation alters the mucin core peptide conformation in a sequence-dependent manner while

reducing the mobility of the glycosylated residues by nearly an order of magnitude. The ^{13}C NMR results from these studies and from earlier work on the more complex pig submaxillary mucin (PSM) furthermore suggest that the peptide-linked GalNAc residue is the major perturbant of the peptide core conformation and dynamics and as such must be a primary determinant of mucin solution structure (Gerken & Jentoft, 1987). These findings are consistent with the changes in molecular dimensions reported for the same series of modified mucins (Shogren et al., 1986, 1987, 1989) and further document that O-linked carbohydrate side chains serve to expand and stiffen the mucin protein core.

MATERIALS AND METHODS

Mucins used for carbon-13 NMR study were purified and modified as described in the preceding paper (Shogren et al., 1989). Prior to NMR study mucins were exhaustively dialyzed against distilled water in the presence of Chelex 100 cation-exchange resin, to remove paramagnetic cations, and lyophilized. Unlike the light-scattering studies described in the previous paper (Shogren et al., 1989), the ^{13}C chemical shift and relaxation parameters of mucins have been shown to be insensitive to molecular weight for mucins with molecular weights over 10^5 (Gerken & Dearborn, 1984). On this basis, the relatively high molecular weight native and modified OSM samples (i.e., enzymatically prepared asialo OSM and hydrolytically prepared asialo OSM) were studied by ^{13}C NMR without molecular weight fractionation and the results compared without reference to molecular weight. The molecular weight ranges of these mucins, determined by light scattering, were $5\text{--}1 \times 10^6$ (Shogren et al., 1987), $8\text{--}2 \times 10^6$, and $1\text{--}0.3 \times 10^6$, respectively (Shogren et al., 1989). In contrast, the molecular weight of the deglycosylated apo mucin is considerably lower ($40\text{--}10 \times 10^3$; Shogren et al., 1989). Thus, to ensure that the differences in molecular weight between the apo mucin and native and asialo mucins could not bias our relaxation data (for example, due to end effects or from an increase in the overall correlation time of the mucin molecule), we obtained for comparison "native" and asialo mucins of molecular weights similar to that of apo mucin. These lower molecular weight mucins were obtained by partial proteolysis of native and asialo OSM by trypsin, producing Trp OSM and Trp asialo OSM as previously described (Shogren et al., 1989). Due to the presence of additional unassigned ^{13}C NMR resonances in a number of these preparations, the protease-treated mucins were fractionated by gel filtration chromatography (Shogren et al., 1989). The higher molecular weight fractions contained only minor contaminating resonances and were subsequently used for our NMR studies. The Trp OSM and Trp asialo OSM fractions studied (Shogren et al., 1989) have average molecular weights of 15 000 and 11 000, respectively, compared to 15 000 for apo mucin. Porcine submaxillary gland mucin (PSM) was isolated and characterized as described by Gerken and Jentoft (1987).

NMR Methods. Natural abundance proton-decoupled 45.3 and 67.9 MHz ^{13}C NMR spectra were obtained on a Bruker WH 180/270 pulsed Fourier transform spectrometer as described earlier (Gerken & Dearborn, 1984; Gerken & Jentoft, 1987). Higher field, 100.6-MHz carbon spectra were obtained by using a Bruker MSL 400 spectrometer. NMR samples were prepared by packing lyophilized mucin into the bottom of the NMR tubes, adding solvent (100 mM KCl, 20% D_2O , 0.04% NaN_3 , and ca. 1% acetonitrile), and allowing the mucin to dissolve over several days. Typical sample concentrations obtained were 80–140 mg/mL. Chemical shift and NT_1 values were insensitive to sample concentration within this range. The

¹ Abbreviations: OSM, ovine submaxillary mucin; PSM, porcine submaxillary mucin; NeuNAc, *N*-acetylneuraminic acid (sialic acid); GalNAc, *N*-acetylgalactosamine; T_1 , spin-lattice relaxation time; NOE, nuclear Overhauser enhancement; ω , NMR observation frequency; τ , rotational correlation time; ϕ , peptide C' , N, C_α , C' dihedral angle; ψ , peptide N, C_α , C' , N dihedral angle; ϕ_g , O-linked glycopeptide dihedral angle C_β , O_γ , C_1 , O_5 .

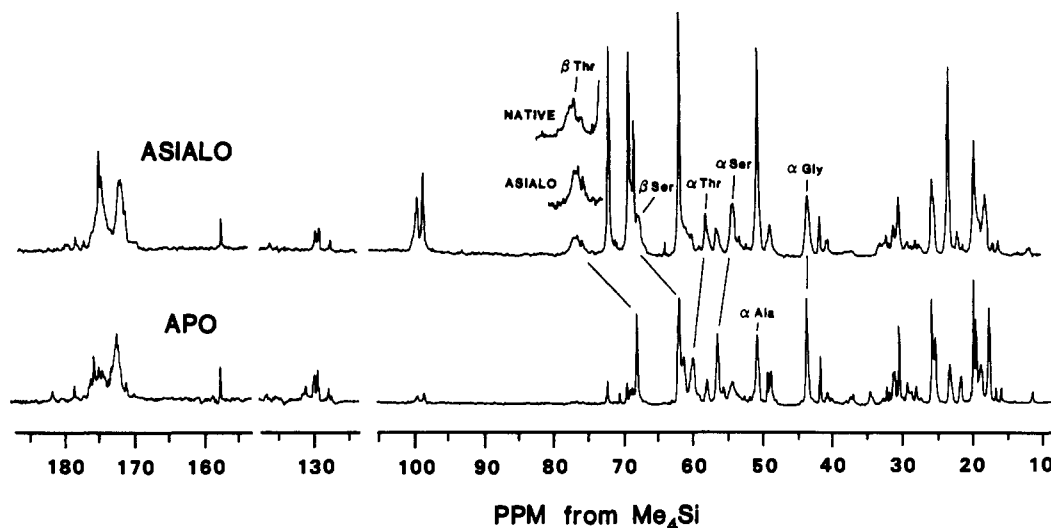


FIGURE 1: 67.9-MHz natural abundance ^{13}C NMR spectra of asialo and apo OSM. Resonances representing carbons of the peptide core discussed in the text are noted. Resonances in the 189–170 ppm region represent amide and carboxylic acid carbonyl carbons, the guanidino carbon of Arg is at 158 ppm, and the ring carbons of Phe appear near 130 ppm. The GalNAc anomeric carbons, which are sensitive to Ser/Thr linkages, appear at ca. 99 ppm, while the remaining GalNAc ring carbons appear as the sharp resonances in the 80–50 ppm region and are assigned to C5, C4, C3, C6, and C2 from low to high field, respectively. Peptide CH and CH_2 carbons appear in the 75–25 ppm region, while CH_3 carbons appear between 25 and 15 ppm. Vertical expansions (above spectrum of asialo mucin) for the Thr β -carbon resonance(s), 70–80 ppm, are shown for native and asialo OSM, respectively. Repetition rates of 2.77 s were used to acquire these spectra, which required 29 000 and 16 000 scans, respectively.

methyl resonance of the acetonitrile, taken as 2.134 ppm from tetramethylsilane, was used as a secondary internal chemical shift reference. Spin-lattice relaxation time experiments were performed by using the fast inversion recovery method of Canet et al. (1975) using 8–11 different τ values (ranging between 1.0 and 0.005 s). Instrument time usage was optimized for obtaining T_1 values of the rapidly relaxing protonated carbons (T_1 s < 0.5 s) as described previously (Gerken & Dearborn, 1984). To improve signal to noise, the broad low-intensity peptide resonances were analyzed after a 10–20-Hz line-broadening function was applied. NMR probe temperature, regulated to $\pm 1^\circ\text{C}$, was measured by using a 4 mm o.d. alcohol thermometer in mucin solvent that was placed into the NMR probe. Temperature determinations were made in the presence of broad-band proton decoupling. The three-parameter nonlinear least-squares method of Kowalewski et al. (1977) was used to calculate T_1 values. Nuclear Overhauser enhancements (NOE) were obtained by comparing the peak intensities of normal and NOE suppressed spectra as previously described (Gerken & Dearborn, 1984). Decoupler on (or off) delay times of 5 s were used. To reduce instrument time requirements, the full NOE spectra were occasionally obtained with one-fourth the number of scans as used for the NOE suppressed spectra and the measured peak heights multiplied by a factor of 2.

Analysis of NMR Relaxation Parameters. The log χ^2 distribution of correlation times model of Schaefer (1973) was used to analyze the multifrequency NMR relaxation data (T_1 and NOE) using a log base, b , of 1000. Calculations were performed by using the program MOLDYN (Craik et al., 1983). Since the minimum regulated probe temperature obtainable for each magnet differed (28 $^\circ\text{C}$ at 67.9 MHz, 30 $^\circ\text{C}$ at 100 MHz, and 34 $^\circ\text{C}$ at 45.3 MHz), we interpolated the 45.3 and 67.9 MHz NT_1 data to 30 $^\circ\text{C}$ by using a least-squares linear fit of the NT_1 data obtained as a function of temperature prior to analyzing the NT_1 data. However, for apo OSM at 45.3 MHz only data at 34 $^\circ\text{C}$ were obtained since the 45.3-MHz magnet irreversibly quenched before a complete multiple temperature data set could be obtained. The 45.3-MHz apo OSM data were therefore corrected to 30 $^\circ\text{C}$ by using the

temperature dependencies obtained for the native and asialo mucin at this field strength.

In addition to NT_1 and NOE data, line width ($1/\pi T_2$) estimates were also included (with low weighting) as a convergence guide. Line widths of 15 and 5 Hz were used for native/asialo mucin and apo mucin, respectively. These values are based on the reported line widths for the structurally similar antifreeze glycoprotein² (Berman et al., 1980). Line widths directly measured from the mucin spectra were not used because the mucin resonances are broadened by chemical shift heterogeneity due to sequence variations. It should be noted, however, that data processed by using line widths of as much as 20–45 Hz resulted in only minor changes in τ compared to the lower line-width fits.

RESULTS AND DISCUSSION

Carbon-13 NMR Resonance Assignments. The proton-decoupled natural abundance ^{13}C NMR spectra of asialo and apo OSM are shown in Figure 1. Peptide core resonances of interest in this study are identified in each spectrum. These and the assignments discussed in the figure legend are based on previous carbon-13 NMR studies of native and asialo OSM (Gerken & Dearborn, 1984). Previous studies have shown that the peptide core conformations of native and asialo mucin are similar since no significant differences in the peptide core resonances were observed between these mucins (Gerken & Dearborn, 1984).³ A comparison of the asialo and deglycosylated apo mucin spectra reveals that upon complete

² The antifreeze glycoprotein is composed of the tripeptide repeat (Ala Ala Thr)_n, where every Thr residue is O-glycosylated by the disaccharide $\beta\text{-Gal}(1\text{--}3)\alpha\text{-GalNAc-O-}$.

³ A more careful analysis of the chemical shifts of native and asialo OSM spectra reveals that the removal of the NeuNAc residue produces a ca. 0.5 ppm downfield shift in the β -Thr resonance(s) and similar changes in the GalNAc C1 resonances. These changes are consistent with the proton NMR studies of native and asialo mucin that reveal changes in the chemical shift of the Thr methyl group upon NeuNAc removal. These shifts are attributed to small torsional changes in the GalNAc glycosidic linkage, especially about the $\text{C}_\beta\text{-O}_\gamma$ bond (Gerken, 1986).

deglycosylation the glycosylated Ser and Thr β -carbons undergo a ca. 8 ppm upfield shift while the α -carbons undergo ca. 2 ppm downfield shift (as indicated in the figure). These shifts are similar in magnitude to those reported previously for the O-glycosylation of Ser and Thr residues in glycoproteins and glycopeptides (Prohaska et al., 1981). The origins of the very broad resonances of the glycosylated Thr and Ser carbons in native and asialo mucin and the subsequent narrowing observed in these resonances upon carbohydrate removal will be discussed in greater detail below. The remaining peptide core carbons appear to be relatively unaffected by deglycosylation except that several resonances become slightly sharper in apo mucin, i.e., the Gly α -carbon and the peptide methyl resonances (17–20 ppm). These findings are consistent with the flexible "random coil" structure proposed for these molecules and confirm that the trifluoromethanesulfonic acid treatment has not chemically altered the peptide core.

Thr/Ser β -Carbon Resonances. A very striking difference is observed in the apparent line widths of the β -carbon resonances of the Thr and Ser residues between the glycosylated and nonglycosylated derivatives. The β -Thr carbons in both asialo and native (Figure 1) OSM appear as broad peaks containing at least three separate resonances (77.49, 46.92, and 76.29 ppm in asialo mucin). Deglycosylation coalesces the β -Thr carbon resonances to a single featureless resonance (68.42 ppm). The same effect is observed for the β -Ser resonances. Similar findings were reported for the effects of O-glycosylation on the ^{13}C NMR spectra of glycoporphin model glycopeptides (Dill et al., 1986). Earlier studies on OSM have attributed the multiplicity of the glycosylated Ser/Thr β -carbon resonances to local peptide sequence effects (Gerken & Dearborn, 1984). These conclusions are supported by the carbon-13 NMR studies (Berman et al., 1980) of the antifreeze glycoprotein, where a single sharp resonance appears for the O-glycosylated Thr β -carbon, and by the known sequence heterogeneity of OSM (Hill et al., 1977). As discussed below, we propose that the chemical shift heterogeneity of the glycosylated β -Ser and Thr carbons is due to the presence of sequence-specific conformations in the peptide core and originates from the conformation dependence of the carbon-13 NMR γ -substituent effect.

The origins of the chemical shift sensitivity of the glycosylated β -Thr/Ser carbons to sequence can readily be understood in terms of the ^{13}C three-bond γ -substituent effect (Tonelli, 1980). The γ -substituent effect, attributed to steric 1–4 interactions, is observed only when 1–4 substituents are in the sterically most hindered gauche (and cis) configurations typically producing (upfield) shifts of –2 to –6 ppm depending on the substituent (Horsley et al., 1970). Because of this, the chemical shifts of peptide β -carbons are sensitive monitors of peptide conformation (Tonelli, 1980; Siemion, 1985), especially since the γ -substituents of the peptide β -carbon are independent of peptide sequence (except for Pro as the $i + 1$ residue) (Tonelli, 1980).⁴

Potential 1–4 gauche interactions that could give rise to γ -effects at the glycosylated β -Thr/Ser carbons in native and asialo OSM can be visualized in Figure 2. Interacting γ -substituents are the C'_{i-1} , the O'_i , and the N_{i+1} of the peptide core (Tonelli, 1980) and C2 and O5 of the α -GalNAc residue. The existence of a range of chemical shifts for the β -carbons of glycosylated Thr and Ser residues indicates that differences

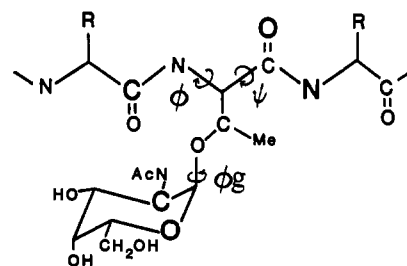


FIGURE 2: Schematic structure of a Ser/Thr glycopeptide unit. Atoms capable of producing γ -chemical shift substituent effects are in bold type. Dihedral angles on which the γ -substituent effect depends are also indicated.

exist in some or all of the dihedral angles described by these interactions, ϕ_i , ψ_i , and ϕ_g (see Figure 2), as a function of peptide sequence. Previous studies on OSM and other O-linked glycoproteins and model building suggest that ϕ_g is constrained to a relatively narrow range of allowed values (Gerken, 1986; Pavia & Ferrari, 1983; Bush & Feeney, 1986; Rao & Bush, 1987). If it is assumed that variations in peptide sequence do not lead to changes in ϕ_g , then differences in the peptide ϕ and ψ dihedral angles must be invoked to explain the data.⁵ This implies that some sequences must have conformations that involve a larger number of 1–4 gauche interactions than others, and this must depend on the nature of the neighboring amino acid residues. The lack of chemical shift heterogeneity in the β -Ser/Thr resonances in apo mucin suggests that conformational changes which generally relieve the 1–4 steric interactions in ϕ and ψ must have occurred, resulting in the formation of a relatively relaxed random coil structure [see Shogren et al. (1989)]. Thus, from this interpretation, it appears that the carbohydrate residues, GalNAc in particular, interact with the peptide core in such a way that steric interactions involving neighboring peptide residues are induced.⁶

Alternate mechanisms are also possible for explaining the chemical shift heterogeneity of the glycosylated β -carbons. For example, long-range chemical shift effects due to peptide and carbohydrate amide carbonyl magnetic anisotropies and δ and ϵ chemical shift substituent effects are likely to be contributors to the observed chemical shifts. Nonetheless, these latter substituent effects, as with the γ -effect, are thought to arise from steric interactions, so that the arguments for conformational changes remain valid. Similarly, the chemical shift effects of the peptide carbonyl anisotropy also depend on peptide conformation (Siemion, 1985). Thus, the overall patterns of the Thr and Ser β -carbon resonances, regardless of mechanism, clearly show that the conformations of the glycosylated residues vary in a sequence-dependent manner.

From the above it appears that the origins of the conformational changes resulting from glycosylation may be steric

⁵ The light-scattering results of the previous paper (Shogren et al., 1989) offer further evidence that the peptide core ϕ and ψ dihedral angles are altered upon O-glycosylation. Furthermore, earlier studies indicate that the Thr-O-GalNAc glycosidic linkage is rather constrained with values for ϕ_g , the most sterically hindered dihedral angle, of approximately 80° (Pavia & Ferrari, 1983; Bush & Feeney, 1986; Gerken, 1986; Rao & Bush, 1987). Such values for ϕ_g , in addition to being sterically favored on the basis of model building, are also consistent with the exoanomeric effect (Lemieux et al., 1979).

⁶ Further evidence for the presence of steric interactions influencing the conformation of Thr O-linked glycopeptides is suggested by the synthetic dipeptide studies of Maeji et al. (1986) and by the conformational analysis of the antifreeze glycoprotein (Bush & Feeney, 1986; Rao & Bush, 1987).

⁴ The longer range, 4- and 5-bond, δ - and ϵ -chemical shift substituent effects, which for peptide β -carbons depend on peptide sequence, are significantly smaller than the γ -effect and are less useful for the study of polypeptide conformation (Tonelli, 1980; Horsley et al., 1970).

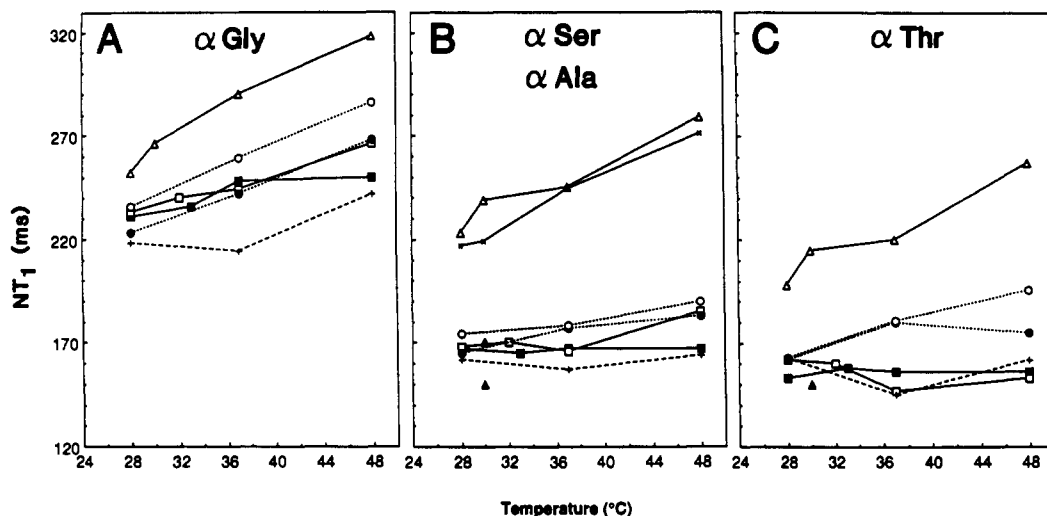


FIGURE 3: 67.9-MHz ^{13}C NMR NT_1 values for the resolved α -carbons in native and sequentially deglycosylated OSM and PSM as a function of temperature. (A) Gly α -carbon NT_1 values for apo OSM (Δ), trypsin-treated asialo OSM (\circ), asialo OSM (\bullet), trypsin-treated OSM (\square), OSM (\blacksquare), and PSM ($+$). (B) Ser α -carbon NT_1 values for the mucins described in (A) plus Ala α -carbon NT_1 values for apo OSM (\times) and antifreeze glycoprotein (\blacktriangle). (C) Thr α -carbon NT_1 values for the mucins described in (A) and for the antifreeze glycoprotein (\blacktriangle). Data for PSM at 30 $^{\circ}\text{C}$ from Gerken and Jentoft (1987) and for the antifreeze glycoprotein from Berman et al. (1980). Data points for PSM are the average of six to eight determinations, points at 28 $^{\circ}\text{C}$ for native and modified OSM are averages of two to three determinations, and all other points represent single determinations.

in nature. However, we must also consider the extent to which hydrogen-bond interactions between the carbohydrate and peptide residues may contribute to the observed changes, especially in light of the recent proton NMR studies of Maeji and co-workers (Maeji et al., 1987). These workers report the formation of two types of amide-based hydrogen bonds between the peptide and carbohydrate residues in several small O-acetylated α -GalNAc-O-Thr glycopeptides in dimethyl sulfoxide. Similar studies in aqueous solvents have not been reported, so it is uncertain if such interactions occur in aqueous solution. A possible role of H-bonds in aqueous solvents has been suggested in relation to the carbohydrate structure of the N-terminus of glycophorin (Dill et al., 1986; Prohaska et al., 1981). However, due to the expanded structure of mucins such H-bond interactions, most likely, will be very short-lived in aqueous solution because of competition with solvent [see Homans et al. (1986)] and, therefore, are probably of minor significance in mucin-type glycoproteins. For example, in the antifreeze glycoprotein, which has a similar expanded conformation, no evidence is found for amide carbohydrate-protein hydrogen-bond interactions (Bush & Feeney, 1986; Rao & Bush, 1987). Likewise, hydrogen bonds involving the peptide core amide protons were not detected in native or modified OSM (Gerken, 1986). The findings that mucin chain dimensions (Shogren et al., 1986, 1989) and carbon-13 NMR spectra are not altered in 6 M guanidine hydrochloride (Gerken & Dearborn, 1984) further support the limited importance of hydrogen-bond-type interactions and emphasize the importance of steric interactions. Thus, sequence-specific steric interactions with neighboring amino acid residues appear to be the source of the peptide conformational changes observed in the glycosylated mucin.

Carbon-13 Relaxation Studies. To examine the effects of carbohydrate on the dynamics of the mucin peptide core, the carbon-13 relaxation time (T_1) and the nuclear Overhauser enhancement (NOE) were obtained for each modified mucin. These measurements were focused on the protonated carbons, whose common ^{13}C - ^1H dipolar relaxation mechanism permits a semiquantitative analysis of each carbon's relative mobility on the basis of NT_1 (where N is the number of attached protons) and NOE. Since the focus of the current studies is

directed toward understanding the effects of glycosylation on the properties of the mucin peptide core, the analysis is limited to the relaxation parameters of the well-resolved, relatively high intensity α -carbons of Gly, Ser, Thr, and Ala.

As described under Materials and Methods, the apo mucin is of significantly lower molecular weight than the native and asialo mucin derivatives. To ensure that the relaxation parameters obtained are not due to changes in molecular weight, low molecular weight derivatives of native and asialo OSM were also studied for comparison to apo mucin. To best evaluate the effects of mucin modification and low molecular weight on the motions of the peptide core, T_1 determinations were originally performed as a function of temperature between 28 and 48 $^{\circ}\text{C}$. The temperature dependence of the 67.9 MHz NT_1 values for native and modified OSM are shown in Figure 3. NT_1 values averaged from several different preparations of PSM, possessing average side-chain lengths varying between 2.5 and 3.3 residues, are also shown in the figure. For further comparison, NT_1 data for the Ala and Thr α -carbons in the antifreeze glycoprotein (Berman et al., 1980), M_r 22 000–11 000, are also shown.

It is apparent from Figure 3 that the α -carbons of Gly, Ser, and Thr in apo mucin have NT_1 values significantly larger than the same carbons in the glycosylated mucins. The non-glycosylated α -Ser and Thr carbons also exhibit a stronger temperature dependence. Unfortunately, the Ala α -carbon resonance is completely masked by the GalNAc C2 resonance in native and asialo OSM and PSM. However, since the glycosylated Thr α -carbon in the antifreeze glycoprotein has an NT_1 value very similar to that observed in native and asialo OSM and even PSM (Figure 3), we can assume that the NT_1 values of the Ala α -carbons in native and asialo OSM will be similar to those in AFGP. [This should be a valid approximation since roughly both OSM (on average) and the antifreeze glycoprotein have a Ser or Thr residue as every third residue.] On this basis, the relaxation time of the Ala α -carbon in apo OSM may also differ from its value in the glycosylated mucins. Since the T_1 values for the high molecular weight mucins are not significantly different from the values for both the low molecular weight trypsin-treated mucins and the antifreeze glycoprotein, it is clear that the elevated NT_1 values

Table I: α -Carbon ^{13}C NMR Relaxation Parameters for Native and Modified OSM and Analysis by $\log \chi^2$ Distribution of Correlation Times Model (30 °C)

	spectrometer frequency ^{a,b}						log χ^2 ^d	<i>p</i>
	45.3 MHz ^c		67.9 MHz ^c		100.6 MHz			
	<i>NT</i> ₁ (ms)	NOE	<i>NT</i> ₁ (ms)	NOE	<i>NT</i> ₁ (ms)			
α -Gly, native	160 (181)	1.9 (2.1)	235 (238)	1.8 (2.1)	326 (317)	0.26	13	
asialo	176 (176)	2.1 (2.1)	232 (232)	1.9 (2.0)	310 (309)	0.29	13	
apo	207 (207)	2.2 (2.3)	262 (261)	1.8 (2.2)	332 (334)	0.18	17	
α -Ser, native	112 (112)	1.7 (1.8)	167 (162)	1.7 (1.7)	257 (240)	1.46	17	
asialo	119 (119)	1.9 (1.9)	169 (169)	1.8 (1.8)	240 (244)	1.12	15	
apo	177 (177)	2.2 (2.2)	231 (231)	2.0 (2.1)	304 (305)	0.27	14	
α -Thr, native	103 (102)	1.7 (1.7)	157 (155)	1.6 (1.6)	255 (240)	2.35	21	
asialo	113 (114)	1.9 (1.8)	167 (164)	1.8 (1.7)	225 (240)	1.33	16	
apo	173 (163)	2.2 (2.1)	207 (215)	2.1 (2.0)	278 (286)	0.33	15	
α -Ala, apo	173 (172)	2.3 (2.1)	222 (225)	2.0 (2.1)	300 (299)	0.29	15	

^a NT_1 and NOE values in parentheses are calculated by using the τ and p values for the $\log \chi^2$ distribution model of Schaefer (1973). ^b Estimated errors in experimental NT_1 values range between 5 and 15 ms depending on relative peak heights and signal to noise ratio. Estimated errors in NOE values are 0.1–0.2. ^c NT_1 values extrapolated to 30 °C as described under Materials and Methods. ^d Estimated variability in τ is $\pm 10\%$ and ± 3 for p .

observed for the apo mucin indeed reflect changes in peptide core dynamics and are not simply due to the apo mucin's lower molecular weight.

Carbon-13 relaxation time studies have been previously reported for OSM and PSM at 28 °C (Gerken & Dearborn, 1984; Gerken & Jentoft, 1987). The results of these studies suggest that the mucin peptide core dynamics is characterized by correlation times (τ_c) somewhat shorter than the correlation time at the T_1 minimum ($\tau_c < 1.6$ ns at 67.9 MHz) in the plot of \log correlation time vs $\log NT_1$, based on the model of a protonated ^{13}C nucleus undergoing isotropic motion with a single correlation time (Doddrell et al., 1972). In this region, the NT_1 values for this model are proportional to mobility; i.e., as motion increases (or τ_c decreases) the NT_1 values increase. Interpreted in this manner, the results in Figure 3 indicate that the deglycosylated apo mucin peptide core is significantly more mobile than the peptide core in native or asialo mucin. Even the nonglycosylated Gly (and by inference Ala) residues show an increase in mobility upon deglycosylation.

From Figure 3 it is clear that each residue's dynamics is unique, with Gly having the greatest mobility and glycosylated Thr the least. In apo mucin the relative α -carbon mobility parallels amino acid side chain size; even the one carbon side chain Ser and Ala residues have nearly identical relaxation behavior. The positive temperature dependencies of Gly, Ala, and the nonglycosylated Ser and Thr α -carbons in apo mucin further indicate that their NT_1 values are roughly proportional to motion (i.e., $\omega^2\tau^2 < 1$). The lack of a significant temperature dependence for the glycosylated Ser and Thr α -carbons in native and asialo mucin, however, suggests that their motions may be associated with a high activation energy (barrier to motion) or that their motions are near the T_1 minimum ($\omega^2\tau^2 = 1$), where actual changes in mobility produce only small changes in NT_1 or even display a negative temperature dependence. Regardless, the temperature dependence of the glycosylated Thr α -carbon suggests that the Thr α -carbon may be more mobile in asialo mucin than in native mucin.

The effects of increasing carbohydrate side chain size on the dynamics of the mucin peptide core have been studied for PSM and OSM (Gerken & Jentoft, 1987; Gerken & Dearborn, 1984). The carbohydrate side chains in PSM range from a single α -GalNAc residue to the sialylated A blood group pentasaccharide (α -GalNAc(1-3)[α -Fuc(1-2)]- β -Gal(1-3)[α -NeuNGl(2-6)]- α -GalNAc-O-Ser/Thr) with average side-chain lengths ranging from 2.5 to 3.3 residues in different

preparations. On the basis of single-temperature relaxation determinations, no correlation between oligosaccharide chain length and peptide core dynamics was observed. The temperature-dependent relaxation data for PSM reported here (Figure 3) concur with these findings; increased side-chain size has no significant effect on the dynamics of the glycosylated Ser/Thr α -carbons compared to native OSM. However, the temperature dependence of Gly indicates that this residue may undergo somewhat more restricted motion in PSM compared to OSM. Thus, the longer and branched side chains of PSM may interact to a greater extent with the peptide residues neighboring the sites of glycosylation. Molecular models of the PSM carbohydrate side chains [see Gerken and Jentoft (1987)] show the 2-6-linked sialic acid and the 1-3-linked β -Gal residues to extend from the peptide-linked GalNAc in opposite directions and demonstrate that both are capable of interacting with the peptide core. As discussed in the previous paper (Shogren et al., 1989), the linkage GalNAc and substituents on C3 and C6 together produce the maximum chain-expanding effects. Recent NMR studies on the O-linked glycopeptides from fetuin further support the possibility of long-range carbohydrate-peptide interactions (Berman, 1987). Thus, on the basis of the carbon-13 relaxation studies on modified OSM and PSM, the peptide-linked GalNAc residue is the chief carbohydrate residue responsible for reducing the mobility of the mucin core protein, while the further addition of carbohydrate residues to GalNAc provide increasingly smaller changes in the peptide core dynamics. These findings are consistent with the light-scattering studies of Shogren and co-workers (Shogren et al., 1986, 1987, 1989).

To more completely evaluate mucin peptide core dynamics, relaxation time data were obtained for native, asialo, and apo OSM at the additional NMR frequencies of 45.3 and 100.9 MHz. The T_1 minimums at these frequencies are shifted to longer and shorter correlation times, respectively, compared to those at 67.9 MHz. The multifrequency NT_1 data obtained for native, asialo, and apo OSM (corrected to 30 °C as described under Materials and Methods) are listed in Table I. NOE data obtained at 45.3 and 67.9 MHz are also listed.

Since, not surprisingly, the multifrequency data for each carbon do not satisfactorily fit the simple model for isotropic motion with a single average correlation time, the data were analyzed by using the two-parameter $\log \chi^2$ distribution of correlation times model of Schaefer (1973). This model has been applied to the study of the internal motions of large

random coiled polypeptides and polymers for which detailed structural information is lacking or difficult to model. The values obtained for the mean correlation time, $\bar{\tau}$, and distribution width parameter, p , and the best-fit NT_1 and NOE values calculated with the $\log \chi^2$ model are also shown in Table I. In this model, $\bar{\tau}$ is the mean correlation time of a $\log \chi^2$ distribution of correlation times and p corresponds to the width of the distribution. The $\log \chi^2$ distribution is asymmetric, with an emphasis on the longer correlation times (lower frequency motions) likely to be present in large random coiled macromolecules as compared to a normal distribution. Small values of p correspond to broader distributions of motions that have a higher proportion of lower frequency long-range "cooperative" motions (Schaefer, 1973; Torchia et al., 1981). As p increases in value, the distribution of correlation times becomes narrower. For very large p the resulting distributions become sufficiently narrow that the calculated results become indistinguishable from those based on a single correlation time model. In general, the plot of NT_1 vs τ_c for the $\log \chi^2$ model is similar to the single average isotropic correlation time model (Doddrell et al., 1972) except the curves are somewhat more shallow and broader depending on the value of p .

The $\bar{\tau}$ values obtained, listed in Table I, for the $\log \chi^2$ model confirm our original interpretation of the single-frequency T_1 data; glycosylation indeed reduces the dynamics of the peptide α -carbons. The dynamics of the Ser and Thr α -carbons are increased nearly an order of magnitude upon complete removal of sugar as shown by values of $\bar{\tau}$ of approximately 1.5–2.4 and 0.27–0.33 ns for native and apo mucin, respectively. Likewise the dynamics of the Gly residue increases by approximately a factor of 1.5 upon complete deglycosylation, from 0.3 to 0.2 ns. Small differences in $\bar{\tau}$ are also observed between native and asialo mucins where the α -carbons of Ser and Thr show a small to moderate decrease in $\bar{\tau}$ upon removal of sialic acid. The Gly α -carbon, however, appears to be unaffected by the removal of sialic acid. Recall that at 30 °C the 67.9-MHz NT_1 data do not reveal significant differences in the α -Thr carbon NT_1 between native and asialo OSM, while differences are observed at elevated temperatures. Thus, both the temperature-dependence studies at fixed frequency and the multifrequency data obtained at fixed temperature reveal true differences in peptide core dynamics between native and asialo mucin. These findings are in agreement with the different values of the persistence length, q , and characteristic ratio, C_∞ , which are measures of chain stiffness, obtained for native and asialo mucin [Table II of the preceding paper (Shogren et al., 1989)]. On the basis of the fit to this model, the relatively flat temperature dependence of NT_1 observed for the glycosylated Ser and Thr α -carbons (Figure 2) can be attributed to $\bar{\tau}$ values near the minimum NT_1 value and not necessarily because of a high energy barrier to rotation.

Interpretations of the calculated p values are less straightforward particularly because no systematic differences (within experimental error) are observed in the value of p with regard to residue type or glycosylation state (see Table I). It is tempting to speculate that the similar p values obtained for all residues regardless of side-chain size suggest that the peptide residues may be undergoing common types of motions (or have similar motional restrictions) although at different $\bar{\tau}$. This would suggest that the types of allowed motions detected for the α -carbon in mucins are not significantly dependent on side-chain size, glycosylation state, or conformation. A more likely situation, however, is that the p values are relatively insensitive and incapable of detecting the differences in the distribution width brought on by the conformational

changes resulting from glycosylation. The relatively small change in p values reported between the random coil, $p = 14$, and nearly α -helical, $p = 16$, form of poly(L-aspartic acid) (Pivcova & Saudek, 1985) suggests that the p values are not highly sensitive to peptide core conformation.

It is of interest to compare the $\log \chi^2$ distribution, $\bar{\tau}$, and p values reported for other random coiled polypeptides and proteins to those obtained for OSM and its derivatives. In general, p values in the range 12–14 and $\bar{\tau}$ values in the range 0.2–0.4 ns are found for most systems (Pivcova & Saudek, 1985; Torchia et al., 1981; Asakura & Muakami, 1985), in good agreement with the values obtained for apo OSM. Pivcova and Saudek (1985) have further obtained data for poly(L-aspartic acid) at the onset of α -helix formation, pH 4, and report p and $\bar{\tau}$ values of 16 and 1.0 ns for the Asp α -carbon, values not dissimilar to those of the glycosylated α -Ser and Thr carbons in the mucins. This comparison provides additional support that the intact mucin peptide core is significantly more restricted and expanded compared to the apo mucin and that the apo mucin has characteristics typical of random coiled polypeptides. Thus, the ^{13}C NMR chemical shift and relaxation data are consistent by indicating that the glycosylated mucin peptide core is considerably more extended and less flexible (i.e., more sterically constrained) than the deglycosylated mucin peptide core and that interactions of the GalNAc residue with the peptide core are the major source of the expanded mucin conformation.

CONCLUSIONS

On the basis of light-scattering studies in the preceding paper (Shogren et al., 1989) we concluded that mucin glycosylation causes a global conformational change in the peptide core that significantly extends chain dimensions and increases chain stiffness. The GalNAc residue causes a ~ 2 -fold increase in chain dimensions compared to apo mucin, while the addition of NeuNAc to the GalNAc residue produces an additional ~ 1.5 -fold increase in chain dimensions. Longer carbohydrate side chains, however, produce no significant additional effects (Shogren et al., 1986, 1987). To confirm and further understand the molecular basis of these findings, we have studied the same series of native and sequentially deglycosylated mucins by carbon-13 NMR spectroscopy.

The results of the carbon-13 NMR studies on native and sequentially deglycosylated OSM closely parallel the findings of the light-scattering studies and offer additional experimental evidence that glycosylation causes both direct and indirect alterations in peptide core conformation and mobility. The ^{13}C NMR chemical shift data clearly show that glycosylation causes significant conformational changes at the Ser and Thr residues and that steric interactions of the GalNAc residue with neighboring amino acid residues are most likely responsible for these changes. In addition, glycosylation decreases the peptide core mobility not only at glycosylated residues but also at residues remote from the site of glycosylation. On the basis of the available data, the length of the carbohydrate side chains, beyond the first few residues, does not appear to be a major factor in altering the mobility of the mucin peptide core. However, the somewhat reduced dynamics of the PSM Gly residue suggest that side-chain size may perturb the mobility of neighboring peptide residues to a limited extent. The NMR and light-scattering results together demonstrate that O-glycosylation of the mucin core protein results in a highly expanded, although internally flexible, mucin molecule. These features along with the mucin's large size account for many of the physical characteristics of these molecules in solution.

The steric origins of the expanded mucin structure suggest that such expanded peptide conformations should be found at multiple Ser/Thr O-glycosidic linkage sites in other glycoproteins. This is indeed found to be the case for the anti-freeze glycoprotein (Bush & Feeney, 1986; Rao & Bush, 1987) and the cell surface glycoproteins, epiglycanin (Slayter & Codington, 1973; Van den Eijnden et al., 1979). In contrast, the more common Asn N-linked glycosidic linkages appear to have less direct effect on the local peptide core conformation [Berman et al., 1981; see Montreuil (1984)]. These differences in behavior may reflect possible different biochemical roles for the O- and N-linked oligosaccharide side chains, where the O-linked side chains may chiefly modify peptide conformation and structure (and provide protection from proteolysis) while the N-linked oligosaccharides may serve as markers, for example, in protein trafficking and recognition (Olden et al., 1985; Kornfeld, 1987).

REFERENCES

- Asakura, T., & Muakami, T. (1985) *Macromolecules* 18, 2614-2619.
- Berman, E. (1987) *Magn. Reson. Chem.* 25, 784-789.
- Berman, E., Allerhand, A., & DeVries, A. (1980) *J. Biol. Chem.* 255, 4407-4410.
- Berman, E., Walters, E., & Allerhand, A. (1981) *J. Biol. Chem.* 256, 3853-3857.
- Bush, C. A., & Feeney, R. E. (1986) *Int. J. Pept. Protein Res.* 28, 386-397.
- Canet, D., Levy, G., & Peat, I. (1975) *J. Magn. Reson.* 18, 199-204.
- Carlstedt, I., Sheehan, J. K., Corfield, A. P., & Gallagher, J. T. (1985) *Essays Biochem.* 20, 40-76.
- Craik, D., Kumar, A., & Levy, G. (1983) *J. Chem. Inf. Comput. Sci.* 23, 30-38.
- Dill, K., Carter, R., Lacombe, J., & Pavia, A. (1986) *Carbohydr. Res.* 152, 217-228.
- Doddrell, D., Glushko, V., & Allerhand, A. (1972) *J. Chem. Phys.* 56, 3683-3689.
- Gerken, T. A. (1986) *Arch. Biochem. Biophys.* 247, 239-253.
- Gerken, T. A., & Dearborn, D. G. (1984) *Biochemistry* 23, 1485-1497.
- Gerken, T. A., & Jentoft, N. (1987) *Biochemistry* 26, 4689-4699.
- Gottschalk, A., & Bhargava, A. (1972) in *Glycoproteins: Their Composition, Structure and Function* (Gottschalk, A., Ed.) Part B, pp 810-829, Elsevier, New York.
- Hill, H. D., Schwyzer, M., Steinmen, H. M., & Hill, R. L. (1977) *J. Biol. Chem.* 252, 3799-3804.
- Homans, S., Dwek, R., Boyd, J., Mahmoudian, M., Richards, W., & Rademacher, T. (1986) *Biochemistry* 25, 6342-6350.
- Horsley, W., Sternlicht, H., & Cohen, J. (1970) *J. Am. Chem. Soc.* 92, 680-686.
- Kornfeld, S. (1987) *FASEB J.* 1, 462-468.
- Kowalewski, J., Levy, G., Johnson, F., & Palmer, L. (1977) *J. Magn. Reson.* 26, 533-536.
- Lemieux, R. U., Koto, S., & Dosin, D. (1979) in *Anomeric Effect, Origins and Consequences* (Szarek, W. A., & Horton, D. H., Eds.) pp 17-29, ACS Symposium Series 87, American Chemical Society, Washington, DC.
- Maeji, N., Inoue, Y., & Chujo, R. (1986) *Carbohydr. Res.* 146, 174-176.
- Maeji, N., Inoue, Y., & Chujo, R. (1987) *Biopolymers* 26, 1753-1767.
- Montreuil, J. (1984) *Biol. Cell.* 51, 115-132.
- Olden, K., Bernard, B. A., Humphries, M. J., Yeo, T., Yeo, K., White, S. L., Newton, S. A., Bauer, H. C., & Parent, J. B. (1985) *Trends Biochem. Sci.* 10, 78-82.
- Pavia, A., & Ferrari, B. (1983) *Int. J. Pept. Protein Res.* 22, 539-548.
- Pivcova, H., & Saudek, V. (1985) *Polymer* 26, 667-672.
- Prohaska, R., Koerner, A., Armitage, I., & Furthmayr, H. (1981) *J. Biol. Chem.* 256, 5781-5791.
- Rao, B., & Bush, C. A. (1987) *Biopolymers* 26, 1227-1244.
- Schaefer, J. (1973) *Macromolecules* 6, 882-888.
- Shogren, R. L. (1985) Ph.D. Thesis, Case Western Reserve University.
- Shogren, R. L., Jamieson, A. M., Blackwell, J., & Jentoft, N. (1986) *Biopolymers* 25, 1505-1517.
- Shogren, R. L., Jentoft, N., Gerken, T. A., Jamieson, A. M., & Blackwell, J. (1987) *Carbohydr. Res.* 160, 317-327.
- Shogren, R. L., Gerken, T. A., & Jentoft, N. (1989) *Biochemistry* (preceding paper in this issue).
- Siemion, I. (1985) *Stud. Org. Chem. (Amsterdam)* 20, 335-346.
- Slayter, H. S., & Codington, J. F. (1973) *J. Biol. Chem.* 248, 4305-4310.
- Tonelli, A. (1980) *J. Am. Chem. Soc.* 102, 7635-7637.
- Torchia, D., Hasson, M., & Hascall, V. (1981) *J. Biol. Chem.* 256, 7129-7138.
- Van den Eijnden, D. H., Evans, N. A., Codington, J. F., Reinhold, V., Silber, C., & Jeanloz, R. W. (1979) *J. Biol. Chem.* 254, 12153-12159.