Structure and Dynamics of Porcine Submaxillary Mucin As Determined by Natural Abundance Carbon-13 NMR Spectroscopy†

Thomas A. Gerken* and Neil Jentoft

Cystic Fibrosis Center, Rainbow Babies and Childrens Hospital, Department of Pediatrics and Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Received November 18, 1986; Revised Manuscript Received March 18, 1987

ABSTRACT: Nearly all of the resonances in the ^{13}C NMR spectrum of porcine submaxillary mucin glycoprotein (PSM) have been assigned to the peptide core carbons and to the carbons in the eight different oligosaccharide side chains that arise from the incomplete biosynthesis of 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). By use of these assignments, a nearly complete structural analysis of intact PSM has been performed without resorting to degradative chemical methods. Considerable structural variability in the carbohydrate side chains was observed between mucins obtained from different animals, while no variability was observed between glands in a single animal. The dynamics of the PSM core and carbohydrate side chains were examined by using the carbon-13 nuclear magnetic resonance relaxation times and nuclear Overhauser enhancements of each assigned carbon resonance. The peptide core of PSM exhibits internal segmental flexibility that is virtually identical with that of ovine submaxillary mucin (OSM), whose carbohydrate side chain consists of the α-NeuNAc(2–6)α-GalNAc disaccharide. The longer oligosaccharide side chains of PSM, therefore, have no significant effect on peptide core mobility compared to the shorter side chains of native OSM or asialo-OSM. Although the dynamics of the shorter carbohydrate side chains shared by both OSM and PSM appear to be identical, the A and H blood group structures in PSM have reduced mobilities, indicating that the glycosidic linkages of the terminal sugars in these determinants are relatively inflexible. These results differ from most reports of glycoprotein dynamics, which typically find the terminal carbohydrate residues to be undergoing rapid internal rotation about their terminal glycosidic bonds. The results reported here are consistent with previous studies on the conformations of the A and H determinants derived from model oligosaccharides and further indicate that the conformations of these determinants are unchanged when covalently bound to the mucin peptide core. In spite of their carbohydrate side-chain heterogeneity, mucins appear to be ideal glycoproteins for the study of O-linked oligosaccharide conformation and dynamics and for the study of the effects of glycosylation on polypeptide conformation and dynamics.

Mucins are extensively glycosylated, very high molecular weight proteins (M_r 10^5–10^7) that are primarily responsible for the protection-rendering viscoelastic properties of mucous secretions. To better understand the solution state of these structurally heterogeneous and difficult to study glycoproteins, we have been using both carbon-13 and proton NMR[1] spectroscopy to systematically study the relatively simple submaxillary mucins. Studies of ovine submaxillary mucin (OSM), which has a relatively simple carbohydrate structure, have been reported (Gerken & Dearborn, 1984; Gerken, 1986). In this paper, the results of our carbon-13 NMR studies of the more complex and heterogeneous porcine submaxillary mucin (PSM) will be discussed and compared to those of OSM and other glycoproteins.

Porcine submaxillary mucin was chosen for subsequent study because its carbohydrate side chains are progressively more complex than the α-NeuNAc(2–6)-α-GalNAc-Ser/Thr disaccharide found in OSM. The PSM oligosaccharides consist of several structures derived from the incomplete biosynthesis of the branched A blood group active oligosaccharide: α-GalNAc(1–3)-[α-Fuc(1–2)-]β-Gal(1–3)-[α-NeuNGl(2–6)-]α-GalNAc-Ser/Thr (Carlson, 1968; Van Halbeek et al., 1981). Four neutral carbohydrate side chains (mono to tetra) together with the corresponding sialic acid containing oligosaccharides are expected, as shown in Figure 1. Typically, porcine submaxillary mucin is typed as either A^+ or A^- from the presence or absence of the A blood group tetrasaccharide (α-GalNAc(1–3)-[α-Fuc(1–2)-]β-Gal(1–3)-α-GalNAc-Ser/Thr). In addition, the PSM trisaccharide (α-Fuc(1–2)-β-Gal(1–3)-α-GalNAc-Ser/Thr) is an H blood group structure. The PSM peptide core composition is typical of mucins, containing predominantly glycine, alanine, and proline together with a nearly equal number of serine and threonine residues, most of which are glycosylated. Thus, PSM is an ideal mucin for examining the effect of increased carbohydrate chain size on the conformation and dynamics of the peptide core and carbohydrate side chains of mucous glycoproteins. Furthermore, PSM is an excellent model for the study of the conformation and dynamics of the native A and H blood group structures linked to protein.

Materials and Methods

Materials. Porcine submaxillary glands were originally obtained frozen from Pelfreeze. For later studies pairs of

† This work was supported by grants from the Rainbow Chapter of the Cystic Fibrosis Foundation and the United Way of Cleveland, by New Investigator Research Grant IO35-02 from the Cystic Fibrosis Foundation, and by NIH Grants AM 27651 and HL 32227. The generous support of the Armington Foundation is also acknowledged.

* Author to whom correspondence should be addressed.

[1] Abbreviations: GalNAc, *N*-acetylgalactosamine; NeuNAc, *N*-acetylneuraminic acid (sialic acid); NeuNGl, *N*-glycolylneuraminic acid (sialic acid); Fuc, fucose; Gal, galactose; Man, mannose; GlcNAc, *N*-acetylglucosamine; OSM, ovine submaxillary mucin; HPLC, high-pressure liquid chromatography; NMR, nuclear magnetic resonance; PSM, porcine submaxillary mucin; DIS, deuterium-induced isotope shift; T_1, spin–lattice relaxation time; NOE, nuclear Overhauser enhancement; ω_c, NMR observation frequency; τ_c, rotational correlation time.

0006-2960/87/0426-4689$01.50/0

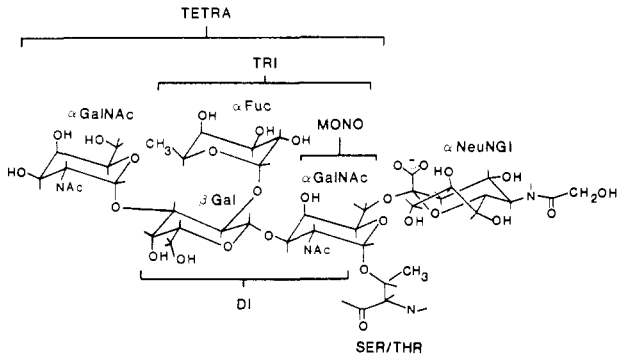


FIGURE 1: Structure of the PSM A blood group active glycopeptide. The tetra- (α-GalNAc(1–3)[α-Fuc(1–2)]β-Gal(1–3)α-GalNAc-Ser/Thr), tri- (α-Fuc(1–2)β-Gal(1–3)α-GalNAc-Ser/Thr), di- (β-Gal(1–3)α-GalNAc-Ser/Thr), and monosaccharides (α-GalNAc-Ser/Thr) are indicated, each of which may have the NeuNGl residue attached to the C6 position of the peptide-linked GalNAc residue. The tetrasaccharide is an A blood group structure and the trisaccharide an H blood group structure.

glands were obtained fresh from a nearby slaughterhouse and frozen for subsequent purification. Neuraminidase type X was obtained from Sigma Chemical Co.; all other chemicals and materials were from commercial sources and of the highest purity available.

Mucin Isolation. PSM was obtained according to the methods of deSalegui and Plonska (1969) as modified by Gerken and Dearborn (1984) and Shogren et al. (1986). Mucin from the 54% ethanol precipitate was exhaustively dialyzed against distilled water and Chelex cation-exchange resin. The solution was clarified by centrifugation and lyophilized. The dry mucin was stored at −20 °C. For our later studies, mucin was isolated from single glands so as to examine the degree of variation within and between individual animals.

Mucin Modification. The sialic acid residue was quantitatively removed from PSM by using neuraminidase from *Clostridium perfringens* as previously described for OSM (Gerken & Dearborn, 1984). The quantitive removal of NeuNGl was confirmed from the carbon-13 NMR spectra of the treated mucins.

PSM-Reduced Oligosaccharides. The removal of the PSM oligosaccharide side chains was performed by the method of Carlson (1968). The neutral reduced tri- and tetrasaccharide alditols were purified by exclusion chromatography on a Bio-Gel P4 (−400 mesh) 2.5 × 90 cm column eluted with 50 mM sodium bicarbonate buffer (pH 8).

NMR Methods. Natural abundance proton-decoupled 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). NMR samples were prepared by packing lyophilized mucin into the bottom of 10-mm-diameter NMR tubes, adding 1.5–2 mL of solvent (100 mM KCl, 20% D_2O, 0.04% NaN_3, and ca. 1% acetonitrile), and allowing the mucin to dissolve over several days. The 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 by obtaining T_1 values for the rapidly relaxing protonated carbons (T_1s < 0.5 s) as described previously for OSM (Gerken & Dearborn, 1984). To improve the signal-to-noise ratio, the broad low-intensity peptide resonances were analyzed after applying a 10–17-Hz line-broadening function. Determinations required 60–80 h of instrument time (13 000–23 000 scans per τ value). Sample temperature, regulated to 28 °C, was measured by a 4-mm o.d. alcohol thermometer that was equilibrated in mucin solvent with the decoupler on. Temperature measurements were made before and after each T_1 determination. 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 from a comparison of the normal and NOE-suppressed spectra using the method of Opella et al. (1976). Decoupler on (or off) delay times of 5 s were used. To minimize instrument time, the full NOE spectra were obtained with one-fourth of the number of scans as the NOE suppressed spectra, and the measured peak heights were multiplied by a factor of 2.

For the pH titration of the sialic acid residue, hydrogen ion concentrations were determined by using a Radiometer PHM 64 meter before and after spectral accumulation. No adjustments were made for the 20% D_2O of the solvent. The pH-dependent NMR shifts were fit to theoretical curves with a three-parameter least-squares nonlinear minimization program. Output parameters were the high and low chemical shift limits and the pK_a value.

Deuterium-induced ^{13}C isotope shifts (DIS) (Pfeffer et al., (1980) were obtained for several preparations of native PSM and asialo-PSM and the alditols. Spectra revealing the DIS were obtained from coaxial NMR tubes (6.5-mm o.d. NMR tube inside a 10-mm o.d. NMR tube) containing samples at identical concentrations in H_2O (6.5-mm tube) and D_2O (10-mm tube). Resolution enhancements were used to increase the resolution of the mucin DIS spectra. All spectra have a data-point resolution of 0.014 ppm.

Spectral calculations were performed by using Symphony (Lotus Development Corp.). Spectra were constructed from input variables of chemical shift, line width, and intensity by using Lorentzian line shapes.

RESULTS

The carbon-13 NMR spectra of several different preparations of native PSM are shown in Figure 2; for comparison, the spectrum of native OSM is also shown. PSM B and D are different preparations using pooled glands, while PSM A and C are mucins isolated from single glands from two different animals. Note the variability in both the 90–110 and 62–80 ppm regions representing carbohydrate anomeric and ring carbons, respectively. This clearly demonstrates oligosaccharide structural heterogeneity, while the minimal changes observed in the 10–60 ppm region (where many peptide carbons appear) suggest little difference in peptide core composition. We have utilized the differences in oligosaccharide distribution between the different preparations of PSM to help assign nearly all of the resonances of the PSM spectra shown in Figure 2 and discussed below.

Oligosaccharide Resonance Assignments. The well-resolved anomeric carbons, 90–110 ppm, were assigned by comparison to the shifts of the anomeric carbons in related structures: OSM (α-NeuNAc(2–6)α-GalNAc-Ser/Thr) (Gerken & Dearborn, 1984), the antifreeze glycoprotein from antartic fish (β-Gal(1–3)α-GalNAc-Thr) (Berman et al., 1980), and the A blood group oligosaccharide and its derivatives (α-Fuc(1–2)-[α-GalNAc(1–3)]-β-Gal-OR and α-Fuc(1–2)β-Gal-OR, R = H, NAc, or hexosamine) (Lemieux et al., 1980; Rosevear et al., 1982). These assignments are shown graphically in Figure 3 (and listed in Tables I–V in the supplementary material). Since several anomeric carbon resonances are unique for specific oligosaccharide sequences, it is possible to determine the oligosaccharide side-chain distribution from the

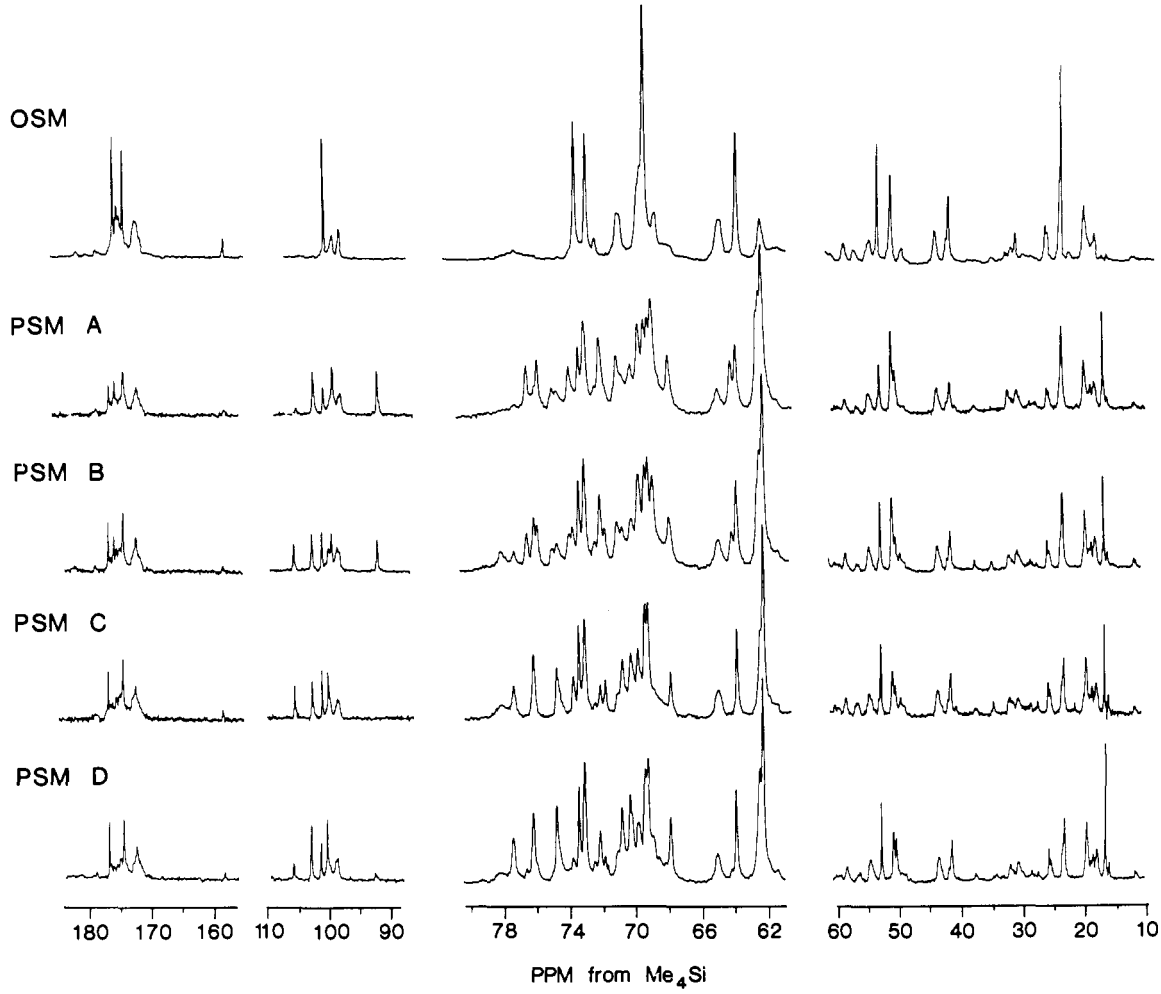


FIGURE 2: 67.9-MHz natural abundance ^{13}C NMR spectra of OSM and several preparations of PSM. Resonance assignments are listed in Tables I–V of the supplementary material. The PSM oligosaccharide side-chain distribution and average chain lengths for each preparation are listed in Table I of the text. Spectra were obtained from 80–100 mg/mL samples of mucin in 0.1 M KCl, 20% D_2O, and 0.04% NaN_3. Sample pH values ranged from 4 to 6. Approximately 50 000–100 000 scans were required for each spectrum with repetition rates of 1.25–2 s.

Table I: Characterization of PSM Oligosaccharide Side Chains by Carbon-13 NMR[a]

PSM preparation	oligosaccharide[b] (mol %)				NeuNGl[c] (mol %)	av length[d]	% Ser and Thr glycosylated[e]
	mono	di	tri	tetra			
PSM A (gland a)	28	8	5	59	38	3.3	76
PSM A (gland b)	37	7	4	53	46	3.2	83
PSM B (pool 1)	36	22	11	31	53	2.9	75
asialo-PSM B	35	21	13	31	0	2.4	69
PSM C (gland a)	43	29	29	0	50	2.4	72
PSM C (gland b)	31	38	31	0	52	2.5	68
asialo-PSM C	34	31	35	0	0	2.0	65
PSM D (pool 2)	29	18	41	12	48	2.8	73
asialo-PSM D	33	16	40	11	0	2.3	70
PSM E (gland a)	39	9	47	4	40	2.6	74
PSM E (gland b)	35	4	58	4	35	2.6	nd

[a] As described in Figure 3 and in text. [b] Oligosaccharide distribution disregarding the presence/absence of NeuNGl (see Figure 1). [c] NeuNGl residues per mole of peptide-linked GalNAc residue. [d] Total sugars per mole of peptide-linked GalNAc residue. [e] Based on areas of the Ser and Thr glycosylated and nonglycosylated α-carbon resonances. No correction was made for the possible resonance overlap of the glycosylated Ser carbon with additional amino acid resonances. nd, not determined due to low signal-to-noise ratio.

integration of this region as described in the legend of Figure 3 and listed in Table I of the text. The integration of this region is valid because no large differences are found in the relaxation times and NOE's of the protonated anomeric carbons whose areas are measured (see below). Furthermore, the ^{13}C NMR derived compositions are in agreement with the chemically determined carbohydrate compositions (data not shown). Note that the oligosaccharide distribution of the NeuNGl residue cannot be readily determined from its anomeric carbon resonance because it shows no long-range sensitivity to oligosaccharide structure. However, on the basis of the pattern of the C4 and C5 resonances of the peptide-linked GalNAc residue (assigned in Table V of the supplementary material), we can estimate that about half of the NeuNGl residue must be attached to α-GalNAc residues which are substituted by β-Gal residues. This 50/50 split in the placement of the NeuNGl residue is confirmed by the HPLC analysis of the benzoylated oligosaccharide alditols

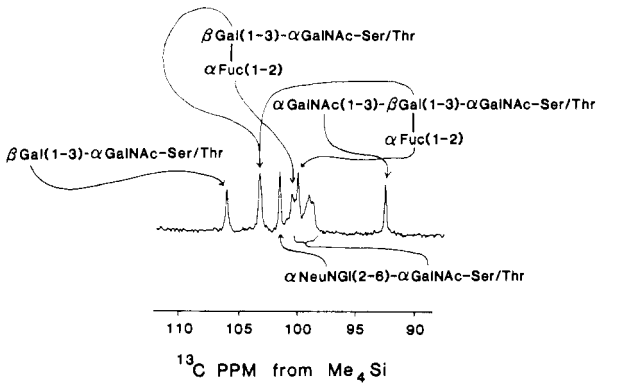


FIGURE 3: Expansion of the PSM B anomeric carbon region with resonance assignments. The arrows identify the anomeric carbon resonances for the indicated carbohydrate residues in each of the possible PSM oligosaccharide side-chain structures. The areas of the well-resolved resonances at 92.5 ppm (terminal GalNAc) and 105.9 ppm (β-Gal in disaccharide) are used to determine the tetra- and disaccharide concentrations, respectively, while the difference in area between the resonance at 103.1 ppm (β-Gal in tetra and trisaccharide) and the tetrasaccharide provides the disaccharide area. The monosaccharide composition is determined from the area of the peptide-linked α-GalNAc resonances (100.5–98.7 ppm) after correcting for both the overlapping α-Fuc resonances (due to the tetra- and trisaccharides) and for the contributions of the tetra-, tri-, and disaccharides to the peptide-linked α-GalNAc resonances that also appear in the same region. The carbohydrate residue composition can be obtained from the anomeric carbon resonances in a similar manner. Note that the peptide-linked and terminal α-GalNAc residues can be quantified separately. After determining the ratio of the total carbohydrate content to the peptide-linked α-GalNAc residue, the average number of carbohydrate residues per side chain can be obtained. Since the unprotonated NeuNGl anomeric carbon is not fully relaxed, the composition of the NeuNGl residue is obtained from the integration of the NeuNGl-linked GalNAc C6 and C9 carbons.

using the method of Daniel et al. (1981) (Jentoft and Gerken, unpublished data).

The resonances of the sialic acid residue were unambiguously identified by comparing carbon-13 NMR spectra before and after treatment with neuraminidase. Spectra of asialo-PSM B and C and the PSM B minus asialo-PSM B difference spectrum are shown in Figure 4. The difference spectrum also identifies the C6, C5, and C4 GalNAc resonances that shift upon the removal of the sialic acid residue. The specific NeuNGl resonance assignments were made by comparison to the NeuNAc residue in OSM and further supported by the observed deuterium isotope shift (DIS) (as shown in Table I of the supplementary material). A maximum DIS will be observed for carbons containing directly bound hydroxyl groups (i.e., β-effect), while smaller or near zero shifts will be observed for nonhydroxylated carbons, depending on the proximity of neighboring hydroxyls (i.e., γ- and δ-effects) (Pfeffer et al., 1980). The NeuNGl residue in PSM D (Figure 1) was also titrated to low pH. The pH dependences of the chemical shifts of the NeuNGl residue (Table I of the supplementary material) are consistent with their assignment and with the previous titration of the NeuNAc residue in OSM. The pK_a value of the NeuNGl residue in PSM is nearly identical with the value reported for the NeuNAc residue in OSM, 2.10 and 2.06, respectively. The low pK_a value of the OSM NeuNAc residue has been shown to be due to a hydrogen bond between the C8–OH and the C2–COO^- groups (Gerken & Dearborn, 1984); thus, this interaction is confirmed for the NeuNGl residue in PSM. Quantitation of the NeuNGl residue is obtained from the integration of the resolved NeuNAc C9, C5, and C3 or (2–6)GalNAc C6 resonances (63.79, 52.87, 41.46, and 65.13 ppm, respectively).

The resonances in the 62–80 ppm region could be assigned by comparison to the chemical shifts of model compounds and by correlation of the variations in oligosaccharide side-chain distribution (determined from the anomeric carbons) to the observed resonance areas. These assignments, and those made for the PSM-derived reduced tetra- and trisaccharide alditols, are compared to model compounds in Tables II–V of the supplementary material for each residue type. The observed deuterium-induced ^{13}C isotope shifts (DIS) that further support these assignments are also shown. DIS values are not reported for the peptide-linked GalNAc residue because its resonances are relatively broad (line widths of 30 Hz compared to 10–17 Hz for the other sugar residues). For example, note the resonances at 78.36 and 65.13, in Figure 2, assigned to the GalNAc C3 in the disaccharide and the GalNAc C6 with attached NeuNGl, respectively. Support for these assignments is further provided by spectral simulation of this region using

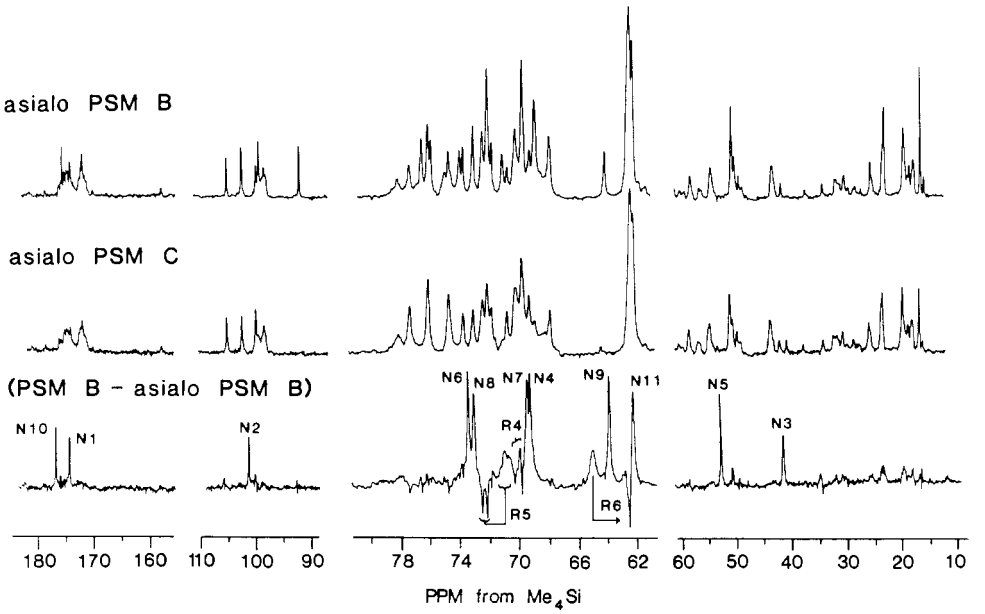


FIGURE 4: 67.9-MHz natural abundance ^{13}C NMR spectra of two asialo-PSM preparations and a native–asialomucin difference spectrum. Spectra were obtained as described in Figure 2: (top) asialo-PSM B; (middle) asialo-PSM C; (bottom) difference spectrum of native PSM B minus asialo-PSM B. The sialic acid (N) and peptide-linked α-GalNAc (R) C4, C5, and C6 resonances are labeled in the difference spectrum.

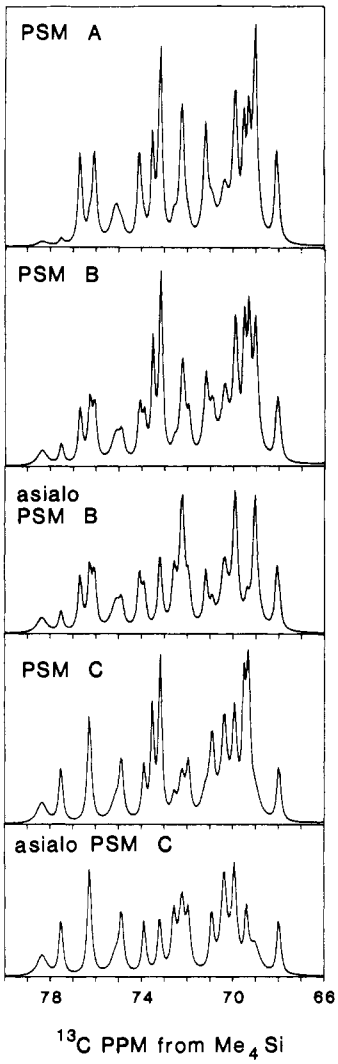


FIGURE 5: Calculated ^{13}C NMR spectra of the carbohydrate ring carbons, 66–80 ppm, for PSM preparations A, B, and C and asialo-PSM preparations B and C. Spectra are based on oligosaccharide compositions averaged for each preparation from Table I and the chemical shifts listed in Tables I–V of the supplementary material. The NeuNGl content of the native mucins was split evenly between the C3-substituted and unsubstituted peptide GalNAc residue. Compare these spectra with those in Figures 2 and 4.

the assigned chemical shifts and the anomeric carbon-derived oligosaccharide distribution. The agreement between the calculated (Figure 5) and experimentally observed spectra (Figures 2 and 4) is excellent. This confirms our basic oligosaccharide assignments but does not exclude the possibility that a few assignments may be interchanged for resonances of similar chemical shift within a given oligosaccharide.

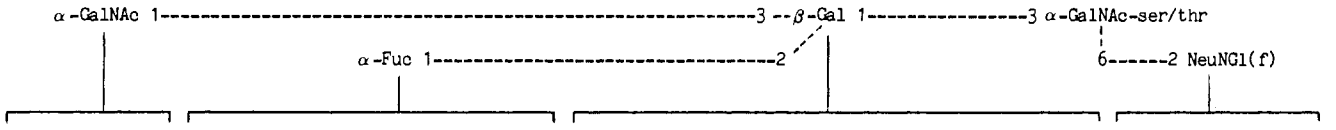
Several unusual resonance shifts deserve comment. The C1 carbon of the terminal GalNAc residue linked (1–3) to β-Gal exhibits an unusually high field shift, 92.5 ppm, for a glycosidically linked anomeric carbon. With this substitution, small (<1 ppm) long-range shifts in the Fuc C1 and C2 carbons are observed in the mucin and model compounds, while 1–2 ppm shifts are observed in the alditols (Table III of the supplementary material). The β-Gal residue is also highly sensitive to substitution and other long-range effects with the C1, C2, C3, and C4 carbons being differentially sensitive to substitution at C2 and C3. In the mucin, substitution of the β-Gal C2 with the Fuc results in the predicted (Rosevear et al., 1982) 5.6 ppm downfield shift at C2, a 2.8 ppm upfield shift at C1, and a 1.0 ppm downfield shift at C3. Further substitution of the β-Gal at C3 by α-GalNAc produces only a 1.8 ppm downfield shift at C3 and larger, 6.1 and 3.4 ppm, upfield shifts at C4 and C2, respectively. Similar shift patterns are observed for the alditols and model compounds. It should be noted, however, that significant chemical-shift differences are observed between the alditols and the nonreduced mucin oligosaccharides.

The substitution of the peptide-linked α-GalNAc residue also is apparent from the ^{13}C NMR spectrum. The attachment of NeuNGl at C6 results in the expected 2.6 ppm downfield and 1.4 ppm upfield shifts at C6 and C5, respectively (Gerken & Dearborn, 1984), with smaller perturbations at C4 as predicted from model compound studies (Prohaska et al., 1981; Berman, 1984; Sabesan & Paulson, 1986). However, a comparison of the chemical shifts of several native PSM and asialo-PSM preparations reveals that, except for the peptide-linked GalNAc residue, there are no long-range chemical-shift perturbations in the remaining carbohydrate residues due to the sialic acid residue. Substitution at the GalNAc C3 by an unsubstituted β-Gal residue gives rise to a 9.3 ppm downfield shift at C3, in agreement with the 9.7 ppm shift in models (Prohaska et al., 1981; Pavia & Ferrari, 1983). However, upon Fuc(1–2) substitution of the β-Gal residue, the GalNAc C3 is observed to shift upfield 3.3 ppm (net substitution shift 6.1 ppm downfield). This substitution shift is confirmed by our spectral calculations (Figure 5). Similar (2–5 ppm) upfield shifts are reported at the C4 or C3 carbons of the GlcNAc, respectively, upon the C2 fucosylation of the β-Gal residue in β-Gal(1–4 or 1–3)β-GlcNAc-OR model compounds (Lemieux et al., 1980; Rosevear et al., 1982; Bush et al., 1985). Much smaller (~0.5 ppm) shifts are observed for the GalNAc C2, C4, and C5 carbons upon substitution at C3 with β-Gal (Prohaska et al., 1981; Pavia & Ferrari, 1983). It is observed that several resonances from carbons in the peptide-linked GalNAc residues (especially C3 and C5) are relatively broad (~30 Hz). This broadening is presumably due to serine and threonine peptide sequence heterogeneity as described for OSM (Gerken & Dearborn, 1984; Gerken, 1986). As in OSM, the removal of the NeuNGl residue causes a narrowing of the line widths of the GalNAc C5 and C6 carbons (see difference spectrum in Figure 4).

Compared to the 60–80 ppm spectral region, the 10–60 ppm region of the various preparations of PSM (Figure 2) are remarkably similar, except for variations in the NeuNGl C5 and C3, GalNAc C2 and C8, and Fuc C6 resonances, 52.9, 41.5, 50.5, 23.4, and 16.6 ppm, respectively, reflecting nearly identical amino acid compositions. This is further demonstrated by the nearly identical amino acid composition obtained for PSM in the present work[2] compared to the composition of PSM reported by deSalegui and Plonska (1969). The amino acid composition of PSM differs from that of OSM predominantly in proline (45% decrease), serine (20% increase), and threonine (10% decrease). These changes are reflected in the areas of the proline α, β, γ, and δ carbons [61, 30, 24, and 49 ppm, respectively (Figure 2)] and the glycosylated serine and threonine α-carbons [54.8 and 58.6 ppm (Gerken & Dearborn, 1984)]. As in OSM, nonglycosylated serine and threonine α-carbons resonances appear [56.7 and 61 ppm, respectively (Gerken & Dearborn, 1984)], from which it can be determined that on average ~25% of the serine and threonine residues are nonglycosylated in PSM (see Table I). The agreement between the areas of the peptide-linked GalNAc C1 carbon (calculated from the anomeric carbon reso-

[2] An amino acid composition obtained for one of our PSM preparations gave Thr 12.4, Ser 21.2, Glx 6.9, Pro 5.8, Gly 19.0, Ala 13.8, and Val 7.5 residues/100 residues. Each of the remaining amino acids residues accounted for less than 3 residues/100 residues.

Table II: ^{13}C NT_1 Values (ms) and NOE for the Terminal α-GalNAc, α-Fuc, β-Gal, and α-NeuNGl Residues in PSM[a,b]



tetra(c)		+GalNAc tetra(c)		-GalNAc tri(d)		+GalNAc +Fuc tetra(c)		-GalNAc +Fuc tri(d)		-GalNAc -Fuc di(e)			
pos	NT_1	pos	NT_1	pos	NT_1	pos	NT_1	pos	NT_1	pos	NT_1	pos	NT_1
1	179 (12)	1	174 (5)	1	183 (9)	1*	179 (10)	1*	184 (7)	1	233 (9)	2	[755] (102)(g)
2*	182 (6)	2+	174 (6)	2	182 (12)	2	172 (15)	2	186 (12)	2	200 (14)	3	208 (6)
3*+	174 (6)	3+	178 (4)	3+	181 (22)	3	172 (5)	3	179 (13)	3	209 (11)	4	180 (2)
4*+	[192] (4)	4*#	171 (0)	4*#	188 (4)	4	166 (15)	4+	171 (8)	4+	204 (17)	5	184 (2)
5*	187 (6)	5*	179 (3)	5*	182 (14)	5	187 (12)	5*	196 (8)	5*	217 (16)	6	182 (5)
6	[280] (6)	6*	[1095] (84)	6*	[1179] (72)	6*+	[340] (28)	6*+	[354] (8)	6*+	[301] (16)	7	178 (2)
8*	[1350](105)											8+	187 (4)
												9	[268] (16)
												11+	[329] (18)
ring avg(h)		ring avg(h)		ring avg(h)		ring avg(h)		ring avg(h)		ring avg(h)		ring avg(h)	
NT_1	177	NT_1	175	NT_1	183	NT_1	175	NT_1	183	NT_1	213	NT_1	187
NOE	1.6	NOE	1.7	NOE	1.5	NOE	1.6	NOE	1.5	NOE	1.7	NOE	1.7

[a] Obtained at 67.9 MHz, 28 °C, 0.1 M KCl, 80–100 mg/mL mucin. [b] Key: *, resonance overlap with another carbon from same residue type in different oligosaccharide; +, resonance overlap with another carbon of unrelated residue; #, data taken from asialomucin. Standard deviation in parentheses; values in brackets not included in ring average. [c] Values from PSM A and B and asialo-PSM B (n = 5). [d] Values predominantly from PSM D and asialo-PSM D (n = 3–8). [e] Values predominantly from PSM B and asialo-PSM B (n = 3–5). [f] Values from PSM A, B, and D (n = 4). [g] Nonprotonated carbon. [h] Average values for ring carbons; NOE determined from PSM A and D (n = 1–2).

Table III: ^{13}C NT_1 Values (ms) and NOE for the Peptide-Linked α-GalNAc and Peptide Residues in PSM[a,b]

peptide-linked α-GalNAc-Ser/Thr								peptide carbons				
+Gal + NeuNAc[c]		+Gal − NeuNAc[d]		−Gal + NeuNAc[c]		−Gal − NeuNAc[d]			PSM		OSM[f]	
pos	NT_1	pos	NT_1	pos	NT_1	pos	NT_1	carbon	NT_1[e]	NOE[h]	NT_1	NOE
1*	170 (2)	1*	162 (9)	1*	170 (2)	1*	162 (9)	α-Gly	212 (10)	1.5	227 (7)	1.7
2*,+	166 (14)	2*,+	163 (12)	2*,+	180 (6)	2*,+	186 (5)	α-Ser OG	162 (7)	1.6	159 (5)	1.7
3*	152 (10)	3*	165 (13)	3*,+	173 (9)	3*,+	173 (6)	α-Thr OG	163 (11)	1.5	153 (7)	1.6
3*[g]	161 (4)	3*[g]	152 (18)	4+	167 (7)	4*,+	196 (1)	α-Ser OH	189 (9)	nd	174 (6)	1.6
4*,+	nd	4*,+	nd	5+	177 (2)	5	193 (9)	γ-Thr/Val	918 (60)	2.2	860 (43)	2.3
5+	174 (7)	5*	171 (1)	6*	[186] (12)	6*,+	[312] (6)	β-Ala	981 (108)	2.4	1083 (117)	2.3
6*	[186] (12)	6*,+	[312] (6)	8*	[1356] (123)	8*	[1411] (154)					
	[1356] (123)	8*	[1411] (154)									
	av 165		av 163		av 173		av 187					
			av GalNAc NOE[h] 1.6									

[a] Obtained at 67.9 MHz, 28 °C, 0.1 M KCl, 80–100 mg/mL mucin. [b] Key: *, resonance overlap with another carbon from same residue type in different oligosaccharide; +, resonance overlap with another carbon of unrelated residue. Standard deviation in parentheses; values in brackets not included in ring NT_1 average. [c] Values from native PSM A, B, and D (n = 3–4). [d] Values from asialo-PSM B and D (n = 2–4). [e] Values from native PSM and asialo-PSM A, B, and D (n = 6–8). [f] Values from native and modified OSM (Gerken & Dearborn, 1984) plus additional determinations (NT_1, n = 6; NOE, n = 3). [g] Fucosylated oligosaccharide. [h] NOE determined from PSM A and D (n = 1–2).

nances, see Figure 3) and the serine and threonine glycosylated α-carbons is excellent (0.97 ± 0.10 to 1.00, 9 determinations). Thus in PSM there appears to be little overlap of the glycosylated serine α-carbon resonance with other amino acid resonances. The glycosylated serine and threonine β-carbon resonances are considerably broader and are not readily observed in PSM due to overlap with the carbohydrate carbon resonances (68 and 77 ppm, respectively). These broad resonances can be observed in OSM in Figure 2. The breadth of these and other peptide carbon resonances is attributed to chemical-shift heterogeneity due to neighboring residue sequence heterogeneity (Gerken & Dearborn, 1984; Gerken, 1986). The α-carbon of glycine is well resolved (43.6 ppm), while the alanine α-carbon overlaps the C2 carbon resonances of both α-GalNAc residues (50.9 ppm). Finally, the peptide methyls of threonine and valine (19.7 ppm), as well as alanine (18.0 ppm), can be observed in the high-field region of the spectrum.

Carbon-13 Relaxation Studies. To examine the dynamics of the PSM carbohydrate side chains and peptide core, the carbon-13 relaxation time, T_1, and the nuclear Overhauser enhancement (NOE) were measured for several preparations of native PSM and asialo-PSM. These measurements were focused on the protonated carbons, whose common ^{13}C–^1H dipolar relaxation mechanism allows comparisons on the basis of the values of the NT_1 (where N is the number of directly attached protons) and NOE. These values are listed in Table II for the carbohydrate residues attached to GalNAc and in Table III for the peptide-linked GalNAc residue and the major peptide residues. Due to the absence of a well-developed motional model and the lack of multifrequency data, we shall interpret the data in a qualitative manner using the trends obtained from a ^{13}C nucleus undergoing dipolar relaxation while undergoing isotropic rotation with a single correlation time, τ_c (Doddrell et al., 1972). For example, at 67.9 MHz with this model the NT_1 has a minimum of ~0.11 s at τ_c = 1.6×10^{-9} s/rad ($\sim\omega_c^{-1}$), with NT_1 values increasing for motions both above and below this minimum. Within this range the NOE changes from a minimum value of ~1.2 ($\tau_c \gtrsim \omega_c^{-1}$) to a maximum value of ~3.0 ($\tau_c << \omega_c^{-1}$). In the absence of chemical-shift heterogeneity, the resonance line width ($1/\pi NT_2$) can also be used as a relative measure of

mobility, with line widths roughly proportional to mobility over a wide range of motions. Thus for motions near the NT_1 minimum, the NOE is ~1.4 and the line width ~4 Hz for this simple model. Using this model and the observed NT_1, NOE, and line widths, we conclude that the motions dominating the ^{13}C NMR relaxation in PSM are on the more mobile side of the NT_1 minimum and are in the range of 40–500 ps. This range of motion spans the relatively rapid internal rotations of methyl groups and the slower segmental motions of the peptide α-carbons, respectively. In the following discussions NT_1 values will be taken as roughly proportional to the degree of motional dynamics.

NT_1 values have been obtained for nearly every carbon in each of the possible PSM carbohydrate side chains and peptide core as shown in Tables II and III. The T_1 values in these tables were obtained from several PSM preparations with different side-chain distributions and represent an average of 3–8 determinations. Errors of less than 10% are estimated for most of the listed values. No significant differences in T_1 were noted for the same carbons in separate PSM preparations with different oligosaccharide distributions. Likewise, the removal of the sialic acid residue produced no detectable change in the relaxation times of the carbohydrate residues attached to the peptide-linked GalNAc residue.

Several key points can be made from the data presented in Tables II and III. For simplicity we will compare the average of the carbohydrate ring NT_1 values in the following discussion. First, the range in NT_1 values for the carbohydrate residues is relatively narrow, suggesting reduced internal dynamics and overall motion comparable to the peptide core. Furthermore, as the side-chain size increases, the side-chain dynamics are reduced as shown by the decrease in ring NT_1 values: 213 ms for the β-Gal ring in the disaccharide; 183 ms for both the β-Gal and α-Fuc rings in the trisaccharide; and 175–177 ms for the β-Gal, α-Fuc, and α-GalNAc residues in the tetrasaccharide. It is especially noteworthy that the terminal residues in the tri- and tetrasaccharide (α-Fuc and α-GalNAc) show identical motional dynamics (i.e., T_1, NOE, and line width) as does the β-Gal residue to which they are attached. Thus there is very little (or no) rotational freedom about the α-Fuc(1–2) and α-GalNAc(1–3) glycosidic linkages to β-Gal. The α-Fuc(1–2)β-Gal glycosidic linkage may be particularly restricted based on the large change in NT_1 of the β-Gal residue upon Fuc substitution. In contrast, motion about the β-Gal(1–3)α-GalNAc glycosidic linkage is shown by the larger NT_1 values of the β-Gal residue compared to those of the peptide-linked α-GalNAc residue (213 vs. ~165 ms in the disaccharide, 183 vs. ~165 ms in the trisaccharide, and 175 vs. ~165 ms in the tetrasaccharide). Motion about the α-NeuNGl(2–6)α-GalNAc glycosidic linkage is also shown (ring NT_1 values of 187 vs. 165–173 ms, respectively). These latter data are consistent with the ^{13}C NMR relaxation data from OSM (α-NeuNAc(2–6)α-GalNAc) and the antifreeze glycoprotein from antarctic fish (β-Gal(1–3)α-GalNAc-O-Thr) which have ring NT_1 values of 192 and 170 and 220 and 170 ms, respectively, at 67.9 MHz and ~30 °C (Gerken & Dearborn, 1984; Berman et al., 1980). From these results and the additional data from PSM, the sialic acid (2–6) glycosidic linkage appears to be less mobile than the unsubstituted β-Gal(1–3) glycosidic linkage. The PSM relaxation data further suggest that the presence of both sugars on a single peptide-linked α-GalNAc residue does not significantly hinder the motion of the other.

The NT_1 data for the peptide-linked GalNAc residue clearly show that its mobility is reduced by the addition of either α-NeuNGl or β-Gal (Table III). The somewhat lower NT_1 values for the GalNAc residue with β-Gal attached are presumably due to additional motional restrictions as the result of further substitutions of the β-Gal residue.

Relaxation times have also been obtained for the PSM neutral tri- and tetrasaccharide alditols. The trisaccharide alditol average NT_1 values are 329, 320, and 317 ms for the α-Fuc, β-Gal, and $GalNAcH_2$ residues, respectively. For the tetrasaccharide alditol, average NT_1 values of 227, 225, 222, and 237 ms are obtained for the α-GalNAc, α-Fuc, β-Gal, and $GalNAcH_2$ residues, respectively. The consistency of the NT_1 values of the sugars within each alditol suggests that they undergo very little (or at least slow) internal rotation about the glycosidic linkages as is the case for the oligosaccharides attached to PSM. Assuming isotropic rotation, correlation times of 15 and 23 ps are obtained for the tri- and tetrasaccharide alditols, respectively.

From the PSM NT_1 values it is clear that the *N*-acetyl and Fuc methyl groups, the Gal and GalNAc C6, and the NeuNGl C11 hydroxymethylene groups undergo rapid internal rotation. As expected on the basis of increased size and hydrogen bonding to solvent, the hydroxymethylene groups are less mobile than the methyl groups.

The relaxation parameters for the major resolved PSM peptide carbons are listed in Table III. Values obtained for native and modified OSM are listed for comparison. Overall the values for PSM compare very well with those obtained from OSM, indicating that for these mucins the length of the carbohydrate side chain, ranging from 1 in asialo-OSM to a mean of over 3 residues in PSM A, has no detectable effect on the peptide core mobility. In addition, the PSM oligosaccharide side-chain distribution (i.e., compare PSM A and PSM D) has no effect on the peptide relaxation parameters within experimental error. From the relaxation data we conclude that the glycosylated Ser and Thr residues are less mobile than the nonglycosylated residues, while the glycine residue possesses a high degree of flexibility. As expected, the peptide methyls undergo rapid internal rotation.

The NOE values listed in Tables II and III are also consistent with the results obtained from OSM (Gerken & Dearborn, 1984). The carbohydrate and peptide residues have NOE values between 1.5 and 1.7, indicating reduced motion, while the more mobile hydroxymethyl and peptide methyls have values of 1.9–2.0 and 2.2–2.4, respectively. Due to the relatively large estimated errors in the NOE determinations (~20%), no distinctions will be made between the different carbohydrate residues on the basis of the observed NOE.

Discussion

These studies were undertaken to compare the structure and dynamics of pig submaxillary gland mucin to the less complex sheep submaxillary mucin (OSM) that we have previously studied (Gerken & Dearborn, 1984; Gerken, 1986). PSM may be said to be a more typical mucin than OSM in that its carbohydrate side chains are structurally heterogeneous and longer in length (see Figure 1) than the side chains of OSM (Carlson, 1968; Gottschalk & Bhargava, 1972). In addition, PSM solutions appear to be more viscous than OSM, a result supported by the generally higher molecular weights that are reported for PSM compared to OSM (Shogren et al., 1987b). A discussion of the structural, conformational, and dynamic information we have obtained from our carbon-13 NMR studies of PSM follows.

Mucin Structure and Heterogeneity. Nearly all of the resonances in the carbon-13 NMR spectrum of PSM have been assigned to specific oligosaccharide and peptide carbons

(see Tables I–V in the supplementary material and Results). This was achieved by comparison to model compound data and by taking advantage of the structural variations observed between different PSM preparations. Thus, for the first time a nearly complete structural analysis of intact PSM has been performed without resorting to degradative chemical methods. Indeed, this is the first publication to describe the oligosaccharide side-chain distribution of a mucin that possesses heterogeneity in its oligosaccharide side-chain structures. The PSM side-chain distribution and carbohydrate residue composition is obtained from the anomeric carbon resonances as shown in Figure 3. The average carbohydrate side-chain length can be determined since the anomeric carbons of the peptide-linked and terminal α-GalNAc residue can be quantitated separately. Finally, the extent of peptide core glycosylation can be estimated from the areas of the glycosylated and nonglycosylated serine and threonine α-carbons.

The carbohydrate side-chain analysis of several different preparations of PSM are listed in Table I. As shown in the table, a range of oligosaccharide distributions is found with the most variability observed in the di- (4–31%), tri- (4–58%), and tetra- (0–59%) saccharide compositions with the least variability observed in the monosaccharide (28–43%) and sialic acid (35–53%) compositions. The absence of the A blood group tetrasaccharide in some animals is consistent with the previous reports of both A^+ and A^- PSM (Carlson, 1968). In spite of the differences in the distribution of the longer side chains, the average side-chain length remains between 2.4 and 3.3 sugar residues, and the overall extent of glycosylation of the serine and threonine residues appears to be on the order of 75%, similar to that observed for OSM (Gerken & Dearborn, 1984). The similarities in the high-field spectral region (0–60 ppm, Figure 2) in all of the PSM preparations further indicate that the PSM peptide core composition does not significantly vary within or between individuals and that it is different from OSM, as indicated by their different amino acid compositions.

The oligosaccharide compositional analyses of the single-gland preparations reveal several interesting trends. First, a comparison of the three single-gland preparations, PSM A, C, and E, suggests that the presence of the tri- or tetrasaccharide (H or A blood group structures, respectively) tends to exclude high concentrations of the other. However, because of the small number of glands studied this point shall not be stressed. The second and more significant finding involving the single-gland preparations is that the two submaxillary glands (labeled a and b) of a given individual produce mucin containing the same carbohydrate side-chain distribution (including the amount of sialic acid). This is clearly shown in Table I. These results indicate that at a given time, the concentrations of the glycotransferases and substrates involved in the posttranslational biosynthesis of the PSM carbohydrate side chains are constant within an individual animal [see Schacter and Williams (1982)].

Mucin Dynamics and Conformation. Despite the very high molecular weight of the mucin [$>10^6$ (Shogren et al., 1986)] and high viscosity, the PSM peptide core and attached carbohydrate side chains appear to be internally flexible as indicated by the resolution of the ^{13}C NMR spectra and the measured ^{13}C T_1 and NOE values (also see Overview of Mucin Dynamics, below). These conclusions are identical with our findings with native and modified OSM, where we concluded that OSM is segmentally flexible with a random coil structure (Gerken & Dearborn, 1984; Gerken, 1986). In fact, the ^{13}C NMR relaxation values (T_1 and NOE) obtained for the PSM peptide core are virtually identical (within experimental error) with those of native OSM, sialic acid modified OSM, and asialo-OSM (see Table III). It is apparent that the mucin peptide core dynamics are relatively insensitive to the size of the carbohydrate side chain in the range of 1 to over 3 sugar residues. This holds for the α-carbons of glycosylated serine and threonine, for the nonglycosylated α-carbons of serine and glycine, and for the peptide side-chain methyls of threonine, valine, and alanine. Similar invariability in mucin flexibility with respect to side-chain size (from two to several residues) has been observed from an analysis of mucin diffusion constant and radius of gyration (Shogren et al., 1986, 1987b; Gerken & Shogren, unpublished data). The latter studies also show that the glycosylated mucins are more expanded and less flexible than denatured (random coil) nonglycosylated proteins (Shogren et al., 1987b). Our relaxation studies of OSM and PSM agree with these findings since the nonglycosylated serine (and threonine) residues are shown to be more flexible than the glycosylated residues [Table III; see also Shogren et al. (1987a)].

In contrast to OSM, however, the longer A and H carbohydrate side chains in PSM appear to have relatively restricted conformations. We find that the α-Fuc(1–2)β-Gal and the α-Fuc(1–2)]α-GalNAc(1–3)]β-Gal portions of the tri- and tetrasaccharides (and their alditols) have virtually constant T_1, NOE values, and line widths, indicating that there is little or no flexibility in the glycosidic linkages of these A and H blood group structures (Table II). Thus these pendant groups may be envisioned as being conformationally fixed, perhaps helping to explain their high antigenicities. However, some motion about the β-Gal(1–3)α-GalNAc glycosidic linkage is found in the tetra- and trisaccharides. This is based on the lower relaxation times of the C3-substituted peptide-linked GalNAc residues compared to those of the terminal Gal, Fuc, and GalNAc residues (NT_1s of ~165 vs. 176–183 ms, Table III). Since the T_1 values for the terminal groups of the tri- and tetrasaccharide are similar (183 vs. 176 ms, respectively), these groups must undergo only slightly different rates of motion about the β-Gal(1–3)α-GalNAc linkage. The NeuNGl residue linked (2–6) to the α-GalNAc residue may also undergo similar rates of motion (NT_1 of 187 ms). In contrast, the unsubstituted β-Gal(1–3) residue in the disaccharide appears to have significantly greater motion about its glycosidic linkage (NT_1 of 213 ms).

It is interesting to note that in PSM the unsubstituted β-Gal residue (1–3) linked to GalNAc is more flexible than the NeuNGl residue (2–6) linked to GalNAc. These findings are in excellent agreement with the carbon-13 NMR relaxation data for the same linkages in the antifreeze glycoprotein and OSM [NT_1 of 220 and 192 ms, respectively (Berman et al., 1980; Gerken & Dearborn, 1984)]. Our studies with the modified sialic acid residue in OSM have shown that the removal of the NeuNAc exocyclic side chain greatly increases its mobility [NT_1 of 192 vs. 220 ms, respectively (Gerken & Dearborn, 1984)], indicating that the reduced mobility of the sialic acid residue is at least partly due to the bulk and interactions of its exocyclic side chain [see Gerken and Dearborn (1984) and Gerken (1986)]. Our studies further show that the NeuNAc(2–6) linkage itself is less mobile than one would expect, with relatively free rotation observed only at the GalNAc C5–C6 bond (Gerken & Dearborn, 1984; Gerken, 1986). Finally, the removal of the sialic acid residue produces no detectable change in the side-chain mobility (or conformation) in the remaining sugars attached to the peptide-linked α-GalNAc residue at C3. Thus the sialic acid residue shows

no intra- or interside-chain interactions with the remaining carbohydrate residues. From the several PSM ^{13}C NMR spectra, we estimate that about half of the sialic acid residues are attached to GalNAc residues substituted at C3; thus, any large effects due to the removal of the sialic acid residue should have been observed in the behavior of the longer oligosaccharides.

The C8 hydroxyl group of the exocyclic side chain of the NeuNAc residue in OSM has been shown to strongly hydrogen bond to its C2 carboxyl group, forming a bicyclic structure (Gerken & Dearborn, 1984). This interaction is based on the unusual ^{13}C NMR pH titration shifts and on the pK_a values of native and modified NeuNAc residues in OSM. Since the NeuNGl residue in PSM titrates with the same pK_a (2.1) and exhibits nearly identical ^{13}C NMR titration shifts as the NeuNAc residue in OSM, we conclude that the exocyclic side-chain interaction and the conformation of the two residues are virtually identical.

The NT_1 values of the peptide-linked α-GalNAc residue in PSM clearly show that its substitution affects its mobility. This residue, when substituted at C3 as in the di-, tri-, and tetrasaccharides, shows the least mobility, as indicated by its low NT_1 value ($\sim$165 ms). This suggests that the C3-substituted GalNAc residue may undergo motional dynamics very similar to that of the glycosylated serine and threonine α-carbons (NT_1 values of 162 and 163 ms, respectively); thus, there may be reduced conformational freedom about the GalNAc-O-Ser/Thr linkage. The α-GalNAc residue when substituted only with a NeuNGl residue shows increased mobility with an NT_1 value of 173 ms, identical with that obtained for native OSM (170 ms). The unsubstituted peptide-linked GalNAc residues in both PSM and asialo-OSM have similarly increased mobilities with NT_1 values of 187 and 189 ms, respectively. Thus the glycopeptide linkage of the unsubstituted peptide-linked GalNAc residue in both OSM and PSM displays the greatest conformational freedom [see Gerken (1986)].

We now will attempt to relate the dynamics of the intact oligosaccharide side chains of PSM to their expected conformations based on the results of previously reported conformational studies of similar oligosaccharides and alditols. This is of interest because no reports of the conformation and dynamics of the intact A or H determinants attached to large polypeptides have appeared. Conformational studies of the A and H type 1 determinants (α-GalNAc(1–3)[α-Fuc(1–2)]β-Gal(1–3)β-GlcNAc-OR and α-Fuc(1–2)β-Gal(1–3)β-GlcNAc-OR, respectively) have been reported (Lemieux et al., 1980; Rao et al., 1985), while the conformation of the ovarian cyst (or pig submaxillary) mucin derived A active tetrasaccharide alditol (α-GalNAc(1–3)[α-Fuc(1–2)]β-Gal-(1–3)GalNAcH$_2$) has recently appeared (Bush et al., 1986). The above workers have employed proton NMR methods (coupling constants and NOE determinations) combined with conformational calculations to arrive at the most likely solution conformation for these oligosaccharides and alditols. Their results indicate that the terminal residues of the A and H determinants have unique, relatively rigid conformations determined primarily on the basis of steric and nonbonded interactions. Our carbon-13 NMR chemical shift and relaxation data on PSM are consistent with these conclusions and offer further insight into the dynamics of these structures.

Space-filling CPK models of the tetra-, tri-, and disaccharide side chains attached to a peptide chain at threonine have been built[3] on the basis of published dihedral angles for the A blood group determinant and its derivatives (Lemieux et al., 1980; Bush et al., 1986; Bush & Feeney, 1986) and the proposed conformation of the OSM glycopeptide (Gerken, 1986). An examination of these models shows that the terminal α-Fuc and α-GalNAc residues are nearly perpendicular to the β-Gal ring. In the model the terminal GalNAc residue H1 and Gal H4 protons are in contact, thus explaining the unusual high-field ^{13}C NMR shifts of these carbons on the basis of steric interactions [see Lemieux et al. (1980) and Kochetkov et al. (1984)]. This interaction is observed in both the model compounds and the intact mucin, on the basis of observed carbon-13 chemical shifts, and is further documented by the large proton NOE observed between these protons (Lemieux et al., 1980; Bush et al., 1986). In agreement with our relaxation data, the α-GalNAc(1–3)β-Gal linkage in the model has essentially no rotational freedom due to interactions with both the Fuc and Gal residues. The α-Fuc(1–2)β-Gal linkage in the model is also hindered primarily due to interactions of the Fuc H1 and H5 protons with the Gal O3 and the peptide-linked GalNAc O3, respectively. Such interactions are also reported in the type 1 A and H oligosaccharides (Lemieux et al., 1980; Bush et al., 1986). Steric interactions between the Fuc and the peptide-linked GalNAc residue are further suggested by the $\sim$3 ppm upfield shift of the peptide-linked GalNAc C3 upon substitution of the β-Gal residue by Fuc. These interactions, which are not altered if the terminal GalNAc residue is removed, may help explain the reduced mobility of the α-Fuc(1–2)β-Gal(1–3) moiety in PSM.

As shown in Tables II–IV of the supplementary material, differences are found between the chemical shifts of the intact PSM oligosaccharide side chains and their respective alditols. The largest differences are found for the trisaccharide Fuc and Gal residues. These findings suggest that the conformation of the trisaccharide is more significantly affected by the nature of the substituent attached to the Gal C1 than is the conformation of the tetrasaccharide.[4] These results suggest that the terminal GalNAc residue of the tetrasaccharide may alter the conformation of the Fuc(1–2) linkage, thereby reducing the sensitivity of the Fuc residue to the Gal C1 substituent. In the molecular model of the tri- and tetrasaccharides [based on the data of Lemieux et al. (1980)], steric interactions are readily observed between both the β-Gal C6 CH_2OH and the α-Fuc C6 CH_3 groups with the peptide-linked GalNAc ring and its C2 *N*-acetyl group. These interactions limit motion about the Gal(1–3)GalNAc glycosidic linkage to a range of approximately 120°. Thus in the tri- and tetrasaccharide models the β-Gal(1–3)α-GalNAc linkage appears to have greater flexibility than the terminal glycosidic linkages. This is confirmed from the carbon-13 NMR NT_1 measurements,

[3] Dihedral angles used to construct the tri- and tetrasaccharide models are ϕ^{Fuc} (O-ring,C1,O1,C2) = −60, ψ^{Fuc}(C1,O1,C2,C1) = 140°, ϕ^{GalNAc}(O-ring,C1,O1,C3) = 50°, ψ^{GalNAc}(C1,O1,C3,C2) = −160°, ϕ^{Gal}(O-ring,C1,O1,C3) = −70°, ψ^{Gal}(C1,O1,C3,C2) = −100°, (Lemieux et al., 1980; Bush et al., 1986); for the disaccharide ϕ^{Gal}(O-ring,-C1,O1,C3) = −70°, ψ^{Gal}(C1,O1,C3,C2) = −140° (Bush & Feeney, 1986); and for the NeuNGl and threonine-linked GalNAc residues χ^1 (N,Cα,Cβ,Oγ) = 40°, χ^2 (Cα,Cβ,Oγ,C1) = 150°, ϕ^{GalNAc}(Cβ,Oγ,-C1,O5) = 80°, ψ^{C5}(GalNAc C5,C6,O6,NeuNGl C2), and ϕ^{C1} (NeuNGl C1, C2, GalNAc O6,C6) = −60° (Gerken, 1986; Bush & Feeney, 1986). Dihedral angles are reported following the IUPAC–IUB recommendations (1970).

[4] The relatively small (−0.3 ppm) chemical-shift difference in the Gal C1 carbon between the trisaccharide side chain and its alditol compared to the larger (1.3 ppm) shift difference between the tetrasaccharide and its alditol is presumably the result of compensating upfield and downfield shifts in the trisaccharide alditol (see Table IV in the supplementary material).

which show slightly greater mobility about this linkage in both the tri- and tetrasaccharides compared to the mobility of the terminal glycosidic linkages (compare Tables II and III). The removal of the Fuc residue eliminates its steric interactions with the peptide-linked GalNAc residue in the model, thus giving the β-Gal(1–3)α-GalNAc linkage significantly increased motional freedom. This is experimentally confirmed by the large NT_1 value obtained for the unsubstituted β-Gal residue (Table III).

In the model of PSM the (2–6)-linked NeuNGl residue attached to GalNAc extends in the opposite direction from the (1–3)-linked residue(s) that may also be attached to GalNAc, not unlike the schematic structure of Figure 1. This conformation is in agreement with our ^{13}C NMR chemical shift and relaxation data which indicate that the (1–3)- and (2–6)-linked side chains are not in contact with each other. The NeuNGl residue is proposed to extend linearly from the GalNAc residue and is expected to have relatively unhindered rotation about the GalNAc C5–C6 bonds and relatively hindered rotation about the GalNAc C6–O6 and GalNAc O6–NeuNAc C2 bonds as indicated for OSM [see Gerken and Dearborn (1984) and Gerken (1986)]. The reduced mobility of the NeuNGl (or NeuNAc) residue compared to the unsubstituted β-Gal residue is therefore due to both the sialic acid's larger size and the hindered motions of the GalNAc C6–O6–NeuNAc C2 linkage, the latter presumably because of steric interactions with the sialic acid's carboxyl group [see Sabesan and Paulson (1986)].

The α-GalNAc-O-Thr linkage has been studied in OSM and the antifreeze glycoprotein (Gerken, 1986; Bush & Feeney, 1986). These studies suggest that this linkage has limited flexibility at the Thr Cβ–Oγ bond but otherwise has a relatively fixed conformation. Similar results are found for PSM, although, for side chains containing the β-Gal substituent, motion about the Thr/Ser linkage is further reduced and the GalNAc, serine, and threonine residues appear to move as one unit.

Overview of Mucin Dynamics. Using a very simple model of carbon-13 relaxation (see Results), we estimate that motions in the relatively narrow range of 40–500 ps dominate the ^{13}C NMR relaxation of PSM. This range of motion encompasses the rapid internal rotations of the Fuc and GalNAc methyl groups as well as the slower motions of the peptide α-carbons. The latter arises from the internal segmental flexibility of the peptide core. In contrast, the ^{13}C NMR relaxation of the conformationally more rigid peptide core of most globular proteins (where there is very little allowed segmental flexibility) is dominated by the much slower rotational correlation time of the entire macromolecule, typically on the order of tens to hundreds of nanoseconds. The relatively narrow range and moderately rapid motions observed for mucins are therefore due in part to the peptide core's internal segmental flexibility and do not reflect the mucin's overall rotational correlation time, which would be very long for such large, extended molecules.

A relatively narrow range of mobilities is observed for the terminal sugars in the A and H tetra- and trisaccharide determinants in PSM. In contrast, most studies of glycoprotein carbohydrate side-chain dynamics indicate that the terminal carbohydrate residues typically undergo relatively rapid motions compared to the internal residues (Dill & Allerhand 1979; Berman et al., 1981; Goux et al., 1982). This is most clearly shown by the ^{13}C NMR relaxation data of Berman et al. (1981) for the carbohydrate side chains of the globular protein ribonuclease B. Under the identical ^{13}C NMR conditions used for PSM (67.9 MHz and 29 °C), the terminal mannose residues of the branched $(Man)_5(GlcNAc)_2$-Asn oligosaccharide have NT_1 values of ~250 ms, while the internal residues have values of ~150 ms. In the same study, the linear trisaccharide side chains of ribonuclease B_M, β-Man(1–4)β-GlcNAc(1–4)β-GlcNAc-Asn, were found to have NT_1 values of 400 and 300 ms for the terminal Man and internal GlcNAc residues, respectively. Comparing these results with the PSM data (listed in Tables II and III) clearly demonstrates that the terminal residues in the A and H tetra- and trisaccharides do not undergo the same type of internal motions. The data suggest that the terminal GalNAc, Fuc, and Gal residues in these oligosaccharides undergo motion as a conformationally rigid group with dynamics not much greater than, and perhaps dominated by, the segmental motions of the peptide core. The reduced mobility is easily understood after a careful analysis of the molecular models of the A and H oligosaccharide side chains, which show that steric interactions can account for their relative inflexibility. These findings further demonstrate that the carbohydrate side chains of glycoproteins and glycoconjugates are not universally highly mobile and that their dynamics are dependent to a large extent on their structure and glycosidic linkages. This may imply that the antigenic nature and other functions of the mucin oligosaccharides arise from their reduced conformational flexibility. Since it is not well understood how carbohydrate side-chain dynamics may relate to their biological and physical chemical functions, further studies of the conformational dynamics of the more complex mucins are in order.

ACKNOWLEDGMENTS

We thank Dr. C. A. Bush and co-workers for sending to us preprints of their manuscripts dealing with the conformations of the antifreeze glycoprotein and the A blood group oligosaccharide.

SUPPLEMENTARY MATERIAL AVAILABLE

Tables I–V listing carbon-13 NMR resonance assignments for the PSM oligosaccharide side chains and chemical-shift comparisons to the PSM alditols and other model compounds (5 pages). Ordering information is given on any current masthead page.

Registry No. Tetrasaccharide, 108864-79-5; trisaccharide, 108773-40-6; disaccharide, 5143-15-7; monosaccharide, 14215-68-0.

REFERENCES

Berman, E. (1984) *Biochemistry 23*, 3754–3759.
Berman, E., Allerhand, A., & De Vries, A. (1980) *J. Biol. Chem. 255*, 4407–4410.
Berman, E., Walters, D. 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.
Bush, C. A., Panitch, M., Dua, V., & Rohr, T. (1985) *Anal. Biochem. 145*, 124–136.
Bush, C. A., Yan, Z., & Rao, B. (1986) *J. Am. Chem. Soc. 108*, 6168–6173.
Canet, D., Levy, G., & Peat, I. (1975) *J. Magn. Reson. 18*, 199–204.
Carlson, D. (1968) *J. Biol. Chem. 243*, 616–626.
Daniel, P. F., DeFuedis, D. F., Lott, I. T., & McCluer, R. H. (1981) *Carbohydr. Res. 97*, 161–180.
deSalegui, M., & Plonska, H. (1969) *Arch. Biochem. Biophys. 129*, 49–56.
Dill, K., & Allerhand, A. (1979) *J. Biol. Chem. 254*, 4524–4531.

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., & Dearborn, D. (1984) *Biochemistry 23*, 1485–1497.
Gottschalk, A., & Bhargava, A. (1972) in *Glycoproteins: Their Composition, Structure, and Function* (Gottschalk, A., Ed.) Elsevier, New York.
Goux, W. J., Perry, C., & James, T. L. (1982) *J. Biol. Chem. 257*, 1829–1835.
IUPAC–IUB Joint Commission of Biochemical Nomenclature (1970) *Biochemistry 9*, 3471–3479.
Jardetzky, O., & Roberts, G. (1981) *NMR in Molecular Biology*, Academic, New York.
Jentoft, N. (1985) *Anal. Biochem. 148*, 424–433.
Kochetkoy, N., Chizhov, O., & Shashkov, A. (1984) *Carbohydr. Res. 133*, 173–185.
Kowalewski, J., Levy, G., Johnson, F., & Palmer, L. (1977) *J. Magn. Reson. 26*, 533–536.
Lemieux, R., Bock, K., Delbaere, L., Koto, S., & Rao, V. (1980) *Can. J. Chem. 58*, 631–653.
Opella, S., Nelson, D., & Jardetzky, O. (1976) *ACS Symp. Ser. No. 34*, 397–417.
Pavia, A., & Ferrari, B. (1983) *Int. J. Pept. Protein Res. 22*, 539–548.
Pfeffer, P., Parrish, F., & Unruh, J. (1980) *Carbohydr. Res. 84*, 13–23.
Prohaska, R., Koerner, T., Armitage, I., & Furthmayr, H. (1981) *J. Biol. Chem. 265*, 5781–5791.
Rao, B., Dua, V., & Bush, C. A. (1985) *Biopolymers 24*, 2207–2229.
Rosevear, P., Neunez, H., & Barker, R. (1982) *Biochemistry 21*, 1421–1431.
Sabesan, S., & Paulson, J. (1986) *J. Am. Chem. Soc. 108*, 2068–2080.
Schacter, H., & Williams, D. (1982) *Adv. Exp. Med. Biol. 144*, 3–28.
Shogren, R., Jamieson, A., Blackwell, J., & Jentoft, N. (1984) *J. Biol. Chem. 269*, 14657–14662.
Shogren, R., Jamieson, A., Blackwell, J., & Jentoft, N. (1986) *Biopolymers 25*, 1505–1517.
Shogren, R., Gerken, T. A., Jentoft, N., Jamieson, A., & Blackwell, J. (1987a) *Biophys. J. 51*, 84a.
Shogren, R., Jentoft, N., Gerken, T., Jamieson, A., & Blackwell, J. (1987b) *Carbohydr. Res. 106*, 317–327.
Van Halbeek, H., Dorland, L., Haverkamp, J., Veldink, G., Vliegenthart, F., Fournet, B., Ricart, G., Montreuil, J., Gathmann, W., & Aminoff, D. (1981) *Eur. J. Biochem. 118*, 487–495.

Oxygen Infrared Spectra of Oxyhemoglobins and Oxymyoglobins. Evidence of Two Major Liganded O_2 Structures†

William T. Potter,‡ Melvin P. Tucker,§ Robert A. Houtchens,‖ and Winslow S. Caughey*
Department of Biochemistry, Colorado State University, Fort Collins, Colorado 80523
Received January 15, 1987; Revised Manuscript Received March 19, 1987

ABSTRACT: The dioxygen stretch bands in infrared spectra for solutions of oxy species of human hemoglobin A and its separated subunits, human mutant hemoglobin Zurich (β63His to Arg), rabbit hemoglobin, lamprey hemoglobin, sperm whale myoglobin, bovine myoglobin, and a sea worm chlorocruorin are examined. Each protein exhibits multiple isotope-sensitive bands between 1160 and 1060 cm^{-1} for liganded $^{16}O_2$, $^{17}O_2$, and $^{18}O_2$. The O–O stretch bands for each of the mammalian myoglobins and hemoglobins are similar, with frequencies that differ between proteins by only 3–5 cm^{-1}. The spectra for the lamprey and sea worm hemoglobins exhibit greater diversity. For all proteins an O–O stretch band expected to occur near 1125 cm^{-1} for $^{16}O_2$ and $^{17}O_2$, but not $^{18}O_2$, appears split by ~25 cm^{-1} due to an unidentified perturbation. The spectrum for each dioxygen isotope, if unperturbed, would contain two strong bands for the mammalian myoglobins (1150 and 1120 cm^{-1}) and hemoglobins (1155 and 1125 cm^{-1}). Two strong bands separated by ~30 cm^{-1} for each oxy heme protein subunit indicate that two major protein conformations (structures) that differ substantially in O_2 bonding are present. The two dioxygen structures can result from a combination of dynamic distal and proximal effects upon the O_2 ligand bound in a bent-end-on stereochemistry.

Binding O_2 reversibly enables hemoglobins and myoglobins to function in O_2 delivery to widespread sites of oxygen utilization (Antonini & Brunori, 1971; Bunn & Forget, 1986). These heme proteins also contribute to oxidant stress by serving as electron donors to O_2 to produce oxy radicals and hydrogen peroxide [e.g., see Wallace et al. (1982), Kawanishi and Caughey (1985), Caughey and Watkins (1985), and Winterbourn (1985)]. The present understanding of how globin structures control O_2 transport and minimize O_2 reduction remains far from complete. The key question of how O_2 binds at heme sites has not been fully answered. The discovery of an oxygen isotope sensitive O–O stretch band near 1100 cm^{-1} in the infrared spectra for solutions of human oxyhemoglobin A and of bovine oxymyoglobin provided the first direct experimental evidence on bond type (i.e., an O–O bond order of 1.5) and a stereochemistry that is "bent end on" (Caughey et al., 1975). This stereochemistry is supported by X-ray and

† This research was supported by U.S. Public Health Service Grant HL-15980.
‡ Present address: Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN 55455.
§ Present address: SERI, Golden, CO 80401.
‖ Present address: Bioproducts—Central Research, Dow Chemical USA, Midland, MI 48640.

0006-2960/87/0426-4699$01.50/0