

- Szundi, I., & Stoeckenius, W. (1989) *Biophys. J.* 56, 369-383.
- Terao, T., Maeda, S., Yamabe, T., Akagi, K., & Shirakawa, H. (1984) *Chem. Phys. Lett.* 103, 347.
- Terner, J., & El-Sayed, M. A. (1985) *Acc. Chem. Res.* 18, 331-338.
- Thompson, L. K., McDermott, A. E., Farrar, M. R., Griffin, R. G., Winkel, C., Lugtenburg, J., Brown, R. S., & Herzfeld, J. (1990) *Biophys. J.* 57, 364a.
- Tsuda, M., Glaccum, M., Nelson, B., & Ebrey, T. G. (1980) *Nature (London)* 287, 351-353.
- Veeman, W. S. (1981) *Philos. Trans. R. Soc. London A299*, 629-641.
- Viswanatha, V., & Hruby, V. J. (1979) *J. Org. Chem.* 44, 2892-2896.
- Walker, T. E., Matheny, C., Storm, C. B., & Hayden, H. (1986) *J. Org. Chem.* 51, 1175-1179.
- Wilbur, D. J., & Allerhand, A. (1976) *J. Biol. Chem.* 251, 5187-5191.
- Winkel, C., Aarts, M. W. M. M., van der Heide, F. R., Buitenhuis, E. G., & Lugtenburg, J. (1989) *Recl. Trav. Chim. Pays-Bas* 108, 139-146.
- Zundel, G. (1988) in *Transport through Membranes: Carriers Channels, and Pumps* (Pullman, A., Ed.) pp 409-420, Kluwer Academic Publishers, Dordrecht, The Netherlands.

## Structure of the Oligosaccharide of Hen Phosvitin As Determined by Two-Dimensional $^1\text{H}$ NMR of the Intact Glycoprotein<sup>†</sup>

Ron L. Brockbank and Hans J. Vogel\*

Division of Biochemistry, Department of Biological Sciences, University of Calgary, Calgary, Alberta, Canada T2N 1N4

Received May 1, 1989; Revised Manuscript Received February 15, 1990

**ABSTRACT:** The major form of the oligosaccharide of hen phosvitin was studied with two-dimensional  $^1\text{H}$  NMR of the intact glycoprotein. Its structure was determined from an analysis of the chemical shifts of the structural reporter groups, and it was further confirmed by comparison to several related model oligosaccharides. The oligosaccharide is N-linked and is present in a 1:1 stoichiometry to the protein. It has a complex type 1 triantennary structure with two NeuAc $\alpha$ 2,6Gal $\beta$ 1,4GlcNAc $\beta$ 1,2 arms linked to the Man-4 and Man-4' and a third Gal $\beta$ 1,4GlcNAc $\beta$ 1,4 arm attached to the Man-4. The oligosaccharide contains the common core sequence which is present in all N-linked glycoproteins [Man $\alpha$ 1,3(Man $\alpha$ 1,6)-Man $\beta$ 1,4GlcNAc $\beta$ 1,4GlcNAc $\beta$ 1,N]. In the course of this study, we have found that unique spin systems for the GlcNAc and NeuAc are obtained for spectra recorded in 90%  $\text{H}_2\text{O}$ . Their NH peaks were assigned at low pH, and these assignments proved useful for confirming the identity of cross-peaks in the anomeric region. In addition, the protons of GlcNAc-1 could be correlated to the NH of the asparagine link. The cross-peak patterns determined in phase-sensitive 2D experiments for the H1,H2 protons have a different appearance for each type of monosaccharide, and this information was also used for making first-order assignments. A comparison with model compounds suggests that the solution conformation of the oligosaccharide is not affected by its attachment to the protein.

**P**hosvitin is a highly phosphorylated glycoprotein which is found in the lipovitellin complex in the egg yolks of fish and birds. In response to hormones, the protein is synthesized in the liver as part of the large vitellogenin precursor (Tata, 1976). The precursor is glycosylated at a number of sites and is phosphorylated in the liver, from where it is transported via the bloodstream to the developing embryo (Taborsky, 1983). The vitellogenin complex is subsequently taken up into the oocytes, and here it is cleaved by specific proteases into a number of proteins including phosvitin which contains a single glycosylated site. The characterization of a glycosylated peptide (Shainkin & Perlmann, 1971a,b) from hen phosvitin was carried out almost 20 years ago at a time when no extensive knowledge of the biosynthesis of these compounds was available (Kornfeld & Kornfeld, 1985). In view of our present understanding of N-linked oligosaccharide sequences, the structure obtained from this initial study contains a highly

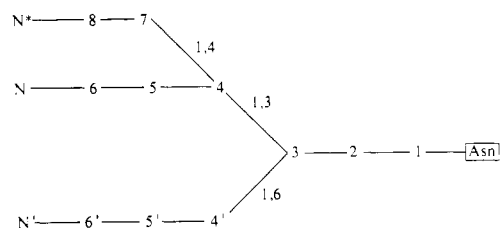
unusual core sequence, and no successful attempts have since been made to reexamine this structure. The amino acid sequence of the major form of phosvitin has been deduced from the DNA sequence (Byrne et al., 1984), and this confirmed that the protein sequence at the glycosylation site was -Asn-X-Ser- which is common to all N-linked glycoproteins.

We have been working on the assignment of the proton nuclear magnetic resonance (NMR)<sup>1</sup> spectra of phosvitin using two-dimensional NMR techniques in an attempt to learn more about the nature of this unusual polyelectrolyte protein. In the course of this work, we detected a number of well-resolved cross-peaks which could be attributed to the oligosaccharide portion of the protein. Vliegthart et al. (1983) have developed a proton NMR method for determining the compositions, sequences, and anomeric linkages of complex oligo-

<sup>†</sup>Supported by the Medical Research Council of Canada. The NMR spectrometer used was purchased with funds provided by the Alberta Heritage Foundation for Medical Research (AHFMR). R.L.B. and H.J.V. are holders of a studentship and a scholarship, respectively, from the AHFMR.

<sup>1</sup> Abbreviations: COSY, correlated spectroscopy; NMR, nuclear magnetic resonance; DQF, double quantum filter; TQF, triple quantum filter; RECSY, relayed correlated spectroscopy; HOHAHA, homonuclear Hartmann-Hahn; NOESY, nuclear Overhauser enhancement spectroscopy; GlcNAc, N-acetylglucosamine; Man, mannose; Gal, galactose; NeuAc, neuraminic acid (sialic acid); H1, anomeric proton; H3a, axial H3 proton; TSP, sodium 3-(trimethylsilyl)propionate-2,2,3,3-d<sub>4</sub>.

Scheme 1



saccharides. Briefly, the characterization of oligosaccharide structures using  $^1\text{H}$  NMR involves measuring the chemical shifts of the "structural reporter groups", those resonances which lie outside of the crowded region of the  $^1\text{H}$  NMR spectrum of carbohydrates between 3.5 and 4 ppm. These reporter groups include the anomeric protons, the H2 and H3 of mannose, the H3a and H3e of NeuAc, and the *N*-acetyl methyl groups in terms of chemical shifts. Their chemical shifts are sensitive to changes in the oligosaccharide sequence (Vliegthart et al., 1983). The application of two-dimensional NMR extends this analysis to the rest of the carbohydrate protons using the scalar coupling network (Homans et al., 1983, 1984, 1987). The increased number of assigned protons provides a more complete framework to initiate studies concerning the three-dimensional structure and dynamics of these compounds. The additional use of nuclear Overhauser and multiple quantum filter experiments has proved useful for allowing a more detailed structural analysis of a limited number of oligosaccharides (Brisson & Carver, 1983a; Homans et al., 1986; Dabrowski et al., 1987). However, even high-field, two-dimensional NMR experiments do not always give a complete assignment of complex carbohydrates because the resonances of certain anomers, such as  $\beta$ -galactose groups, remain unresolved.

The application of two-dimensional NMR methods should allow the analysis of NMR spectra in which the complex oligosaccharide is attached to a more reasonable facsimile of the glycoprotein structure. The current method for structural determination involves separating and purifying the oligosaccharide from the core protein prior to analysis; this necessitates extensive proteolysis or hydrazinolysis of the glycoprotein. This procedure allows the microheterogeneity of the oligosaccharide to be assessed by chromatography and increases the purity of the oligosaccharide sample for spectroscopy, but leaves open the question whether the three-dimensional structures obtained in this manner are identical with intact glycoprotein structures. Since the application of proteolytic techniques to phosvitin has resulted in poor yields, we chose to assign the structural reporter groups for the intact glycoprotein. To assign the carbohydrate portion of phosvitin, the spin systems of the monosaccharide units need to be differentiated from those of the amino acid residues in the protein. We have used the chemical shifts, scalar coupling patterns, and cross-peak fine structure for this purpose. Our data show that proton NMR can be useful to elucidate the oligosaccharide structures for intact glycoproteins. Throughout this work, the oligosaccharide sequence has been labeled according to the convention of Vliegthart as is shown in Scheme I. We believe that this is the first report of this type in which an oligosaccharide sequence has been determined by 2D  $^1\text{H}$  NMR for an intact glycoprotein of this size (35 000 daltons).

## MATERIALS AND METHODS

**Sample Preparation.** Phosvitin purified according to the method of Sundararajan et al. (1960) was obtained from Sigma Chemical Co. This preparation is heterogeneous

(Wallace & Morgan, 1986), and the major form of phosvitin was further purified by FPLC (Pharmacia) using a 30 cm Superose-12 gel filtration column. The fractions containing the major peak were pooled and lyophilized. The protein was dissolved in a 50 mM ammonium bicarbonate buffer (pH  $\approx$  9.0) and passed through a 10-cm Chelex-100 column equilibrated in the same buffer to remove any bound paramagnetic ions. After lyophilization, the sample was redissolved at a  $\approx$  2 mM concentration in either 99.98%  $^2\text{H}_2\text{O}$  (MSD Isotopes) for exchanged experiments or 90:10  $\text{H}_2\text{O}/^2\text{H}_2\text{O}$  for nonexchanged experiments, and potassium chloride was added to a final concentration of 50 mM to maintain ionic strength. The pH of the solution was adjusted with small volumes of concentrated  $^2\text{HCl}$  or  $\text{KO}^2\text{H}$  and checked before and after each experiment with a standardized pH meter equipped with a glass electrode. We found that the protein had sufficient buffering capacity to maintain a constant pH in the course of our experiments. No adjustments of pH values for the presence of  $^2\text{H}_2\text{O}$  have been made. For the peroxide oxidation of phosvitin, 10  $\mu\text{L}$  of 5%  $\text{H}_2\text{O}_2$  freshly exchanged to the  $^2\text{H}_2\text{O}_2$  form was added to 500  $\mu\text{L}$  of a 1 mM phosvitin solution prepared as above; the protein was incubated for 6 h at room temperature.

**NMR Spectroscopy.** All spectra were recorded by using a 5-mm proton probe on a wide-bore Bruker AM 400-MHz spectrometer. They were acquired and processed by using standard Bruker software packages. The phase-sensitive experiments were acquired with time-proportional phase increments (TPPI) as described by Marion and Wüthrich (1983). The HDO resonance was reduced with a bilevel presaturation scheme; however, proton exchange between the solvent and the protein reduces the effectiveness of this scheme, giving rise to residual  $t_1$  noise. For each 2D experiment, multiplication with a nonshifted sine bell was applied to enhance the appearance of the spectra. The spectra were acquired with a sweep width of 4000 Hz and up to 512  $t_1$  data points by 2048  $t_2$  data points. The  $\omega_2$  and  $\omega_1$  dimensions were each zero-filled before transformation, providing a final resolution of  $\approx$  2 and 4 Hz/point, respectively. We have found that the acquisition of data at two different temperatures is helpful to resolve the accidental saturation of cross-peaks found near the water signal. The DQF-COSY acquisition (Piantini et al., 1982) recorded at 15  $^\circ\text{C}$  used 144 scans, and the DQF-COSY at 40  $^\circ\text{C}$  used 208 scans. The NOESY spectrum (256 scans) was recorded with a mixing time of 100 ms (Bodenhausen et al., 1984). The HOHAHA acquisition (160 scans) was obtained with the MLEV-17 mixing sequence with trim pulses of 2.5 ms before and after the mixing sequence (Bax & Davis, 1985); the  $\beta_1$  transmitter operating in the low-power mode produced  $\approx$  2.4 W generating a  $90^\circ$  pulse of 90  $\mu\text{s}$  and a mixing time of 41.6 ms. The RECSY (256 scans) total mixing time (26 ms) was optimized for  $\approx$  8-Hz coupling (Wagner, 1983). The TQF-COSY (168 scans,  $864 \times 2048$  data points) (Piantini et al., 1982) was recorded in the absolute value mode. Five microliters of a 100 mM TSP stock solution was added as an internal reference; the chemical shifts presented in Table I have been further corrected for the pH dependence of TSP (De Marco, 1977). Cross-peak positions are presented as  $\omega_2, \omega_1$  ppm corresponding to the horizontal and vertical axes in the figures.

## RESULTS

The proton spectrum of phosvitin recorded in  $^2\text{H}_2\text{O}$  is shown in Figure 1. Its unusual amino acid composition is readily apparent; the large peak at 4.04 ppm is due to the  $\beta$  protons of the phosphoserine core of the protein. Because this peak

Table I: Chemical Shifts (ppm) of the Reporter Protons of the Oligosaccharide of Hen Phosvitin<sup>a</sup>

residue	H1	H2	H3	NH	methyl		
Asn-169 ( $\delta$ NH)				8.56			
GlcNAc-1	5.02	3.81	3.74	8.09	1.998		
GlcNAc-2	4.62	3.75		(8.08) <sup>a</sup>	2.069		
Man-3	4.78	4.20	3.78				
Man-4	5.14	4.18	3.90				
Man-4'	4.92	4.07	3.86				
GlcNAc-5	4.62	3.79		(8.08)	(2.054)		
GlcNAc-5'	4.62	3.71		(8.08)	(2.056)		
Gal-6,6'	4.44	3.55	3.74				
GlcNAc-7	4.61	3.77		(8.08)	(2.061)		
Gal-8	4.47	3.54	3.73				
residue	H3a	H3e	H4	H5	H6	NH	methyl
NeuAc,N'	1.71	2.69	3.69	3.84	3.75	7.92	2.030

<sup>a</sup> Parentheses indicate those assignments which may be interchanged or cannot be distinguished due to overlap.

<sup>a</sup> Parentheses indicate those assignments which may be interchanged or cannot be distinguished due to overlap.

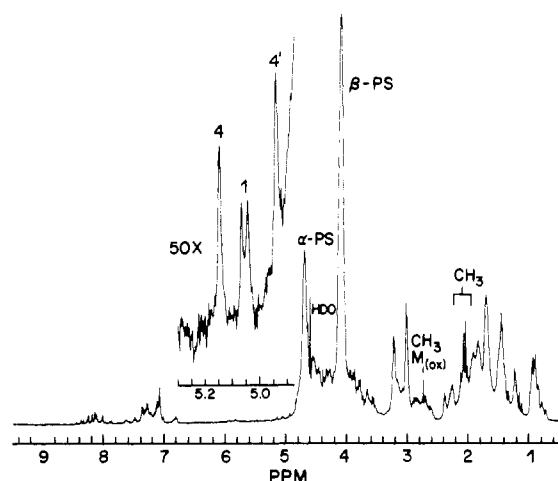


FIGURE 1: Proton spectrum of phosvitin recorded at 40 °C and pH 7.0. The expansion above the spectrum highlights the three anomeric protons which can be resolved with one dimension. The single methionine has been oxidized to selectively shift its methyl resonance away from the oligosaccharide methyl protons.

is over 200-fold more intense than those of the carbohydrate protons, an increased number of scans is required to detect these protons. Furthermore, the contour levels used for plotting the oligosaccharide cross-peaks in the two-dimensional experiments are quite close to the base line. As a result, the  $t_1$  noise along the HDO ( $^1\text{H}^2\text{HO}$ ) and phosphoserine frequencies appears worse in the two-dimensional plots than would normally be found in other samples. The three resolved anomeric protons (Man-4, GlcNAc-1, and Man-4') are plotted in the expansion above the spectrum. The large expansion required to equate the scale of these oligosaccharide peaks to the phosphoserine peak reduces the apparent signal to noise ratio. We have integrated the intensities of the resolved anomeric protons and also the *N*-acetyl methyl protons found near 2 ppm in the 1D NMR spectrum. The ratio of the GlcNAc-1 H1 area to the C4 proton of the single tryptophan residue is 1:1 as is the ratio between the methyl protons of GlcNAc-1 and the single methionine residue. This indicates a one to one stoichiometry of the protein and oligosaccharide. It is also apparent that a one-dimensional analysis of the spectrum is not sufficient to determine the composition and structure of the oligosaccharide because of the overlap of resonances.

A two-dimensional HOHAHA spectrum of the aliphatic region of phosvitin is provided in Figure 2. This spectrum has been divided up into regions in which one would expect to see certain cross-peaks resulting from the oligosaccharide portion of the glycoprotein. In region a, cross-peaks will appear arising from correlations to the H1 protons of GlcNAc, Man,

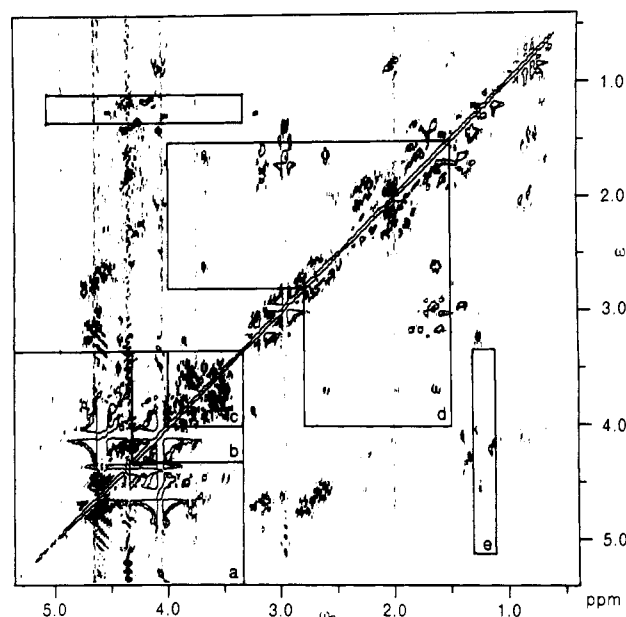


FIGURE 2: Two-dimensional assignment of the oligosaccharide portion of the protein is simplified by dividing the spectrum into specific regions. Regions a-e are chemical shift boundaries within which certain oligosaccharide correlations will be found (details in text). In this HOHAHA spectrum, correlations (H1,H2) to the three resolved anomeric protons can be found in region a. The H1,H2 cross-peaks of Man-4 (5.14,4.18 ppm), GlcNAc-1 (5.02,3.81 ppm), and Man-4' (4.92,4.07 ppm) are found above the diagonal and the corresponding H2,H1 cross-peaks below the diagonal.

Gal, and fucose residues. In region b, correlations to the H2 protons of Man and the H3 protons of Gal ( $\beta$ -1,3) can be found. In region c, the bulk of the oligosaccharide cross-peaks resulting from the H2, H3, H4, H5, and H6 protons which are not found in the other spectral regions is observed. In region d, the cross-peaks to the H3a and H3e protons of NeuAc residues are found, and finally, region e contains the correlations to the H6 methyl protons of fucose groups. Because the carbohydrate is exposed on the glycoprotein surface and in phosvitin is not near aromatic rings which may affect magnetic shielding, we expect that there will be little difference between the chemical shifts of these protons and those of similar model oligosaccharides. Using the expected cross-peak positions, we can first assign the anomeric cross-peaks to their respective sequence positions. A number of the structural reporter cross-peaks may not be resolved from each other; in these cases, we must determine if the pattern of overlapping cross-peaks is consistent with other information that is available such as the chemical composition of the oligosaccharide structure. For phosvitin, the chemical composition has been

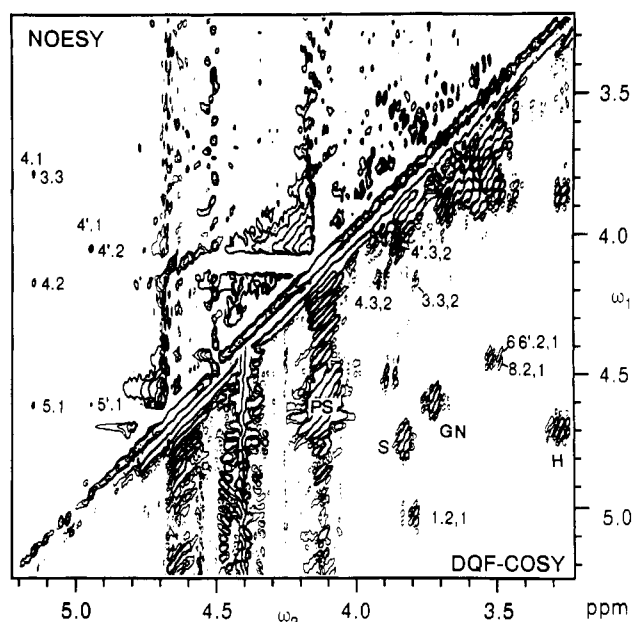


FIGURE 3: DQF-COSY (below the diagonal) and a NOESY (above) spectrum recorded on the same sample under similar conditions (40 °C and pH 4.8) of regions a-c. The correlation peaks are labeled according to Scheme 1, and the oligosaccharide protons (e.g., H<sub>2</sub>,H<sub>1</sub> = 2,1) which give rise to the cross-peak follow the cross-peak position  $\omega_2, \omega_1$  ppm. The NOE's to the Man-4.1 and Man-4'.1 are labeled in the NOESY spectrum. GN represents the overlapping GlcNAc-2, -5, -5', and -7 protons. A few of the amino acid cross-peaks (PS, phosphoserine; H, histidine; S, serine) have been identified.

previously determined by chemical analysis (Shainkin & Perlmann, 1971a) to be five GlcNAc, three Man, three Gal, and two NeuAc residues. The assignment of the oligosaccharide of phosvitin is aided by the unusual amino acid composition of phosvitin in which  $\approx 50\%$  of the 217 amino acids are phosphoserine residues.

The DQF-COSY spectrum of the carbohydrate regions a, b, and c of phosvitin recorded at 40 °C is shown in the lower portion of Figure 3. The application of the double quantum filter prevents dispersion signals found in the diagonal peaks of COSY spectra from interfering with the carbohydrate cross-peaks located near the diagonal. It is apparent that there are a number of resolved and partially resolved cross-peaks in the oligosaccharide region of the spectrum. The cross-peak multiplet appearance is useful for confirming the assignment of resolved reporter protons and for resolving partially overlapping cross-peaks. The *J* coupling or scalar coupling network determines the multiplet pattern; it also determines the overall intensity of the cross-peak. Cross-peak appearance has been used previously for structural analysis of complex oligosaccharides (Berman, 1987; Dabrowski et al., 1987). In Figure 4, an expansion of the appropriate area from a DQF-COSY recorded at 15 °C is presented showing several cross-peaks. The methods to analyze active and passive couplings have been described for correlation experiments (Neuhaus et al., 1985; Müller et al., 1987). The active coupling between the protons H<sub>a</sub>-H<sub>b</sub> appears as anti-phase splitting while the additional passive couplings H<sub>i</sub>-H<sub>a</sub> and H<sub>b</sub>-H<sub>j</sub> appear as in-phase splittings of the active coupling along the frequency axis of the proton to which it is passively coupled (Neuhaus et al., 1985). In Figure 4, we have placed the approximate spin pattern we expect for each monosaccharide based on the coupling data of model oligosaccharides (Vliegthart et al., 1983; Bock & Thøgersen, 1982) beside the one we observed. The contributions to the splitting pattern observed in the H<sub>1</sub>,H<sub>2</sub> cross-peak in <sup>2</sup>H<sub>2</sub>O result mainly from the passive

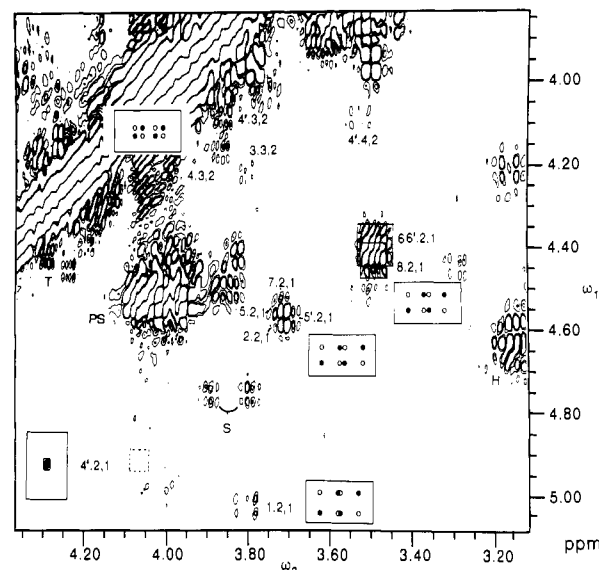


FIGURE 4: DQF-COSY recorded at 15 °C and pH 7.3. The simulated multiplet pattern based on three-bond couplings is placed in the rectangle next to the appropriate oligosaccharide cross-peaks. The active coupling results in the basic anti-phase square while the passive coupling reproduces the anti-phase square along the dimension of the passively coupled proton. The negative components of the simulated cross-peaks are presented as solid circles. The simulated cross-peaks are labeled as follows: a,  $\beta$ -GlcNAc-1,2,1; b, one of  $\beta$ -GlcNAc-2, -5, -5', -7,2,1; c,  $\alpha$ -Man-2,1; d,  $\alpha$ -Man-3,2; e,  $\beta$ -Gal-1,4,2,1. Threonine (T)  $\alpha, \beta$  cross-peaks have been identified near the diagonal in addition to the amino acid cross-peaks labeled in Figure 3.

splitting of H<sub>2</sub> by H<sub>3</sub> as four-bond or longer coupling constants are very small in comparison with the normally observed two- and three-bond coupling constants. This gives the H<sub>2</sub>,H<sub>1</sub> cross-peak of GlcNAc a rectangular rather than square appearance as the active 9.8-Hz coupling is further split by H<sub>3</sub> coupling to H<sub>2</sub> (9.9 Hz) along the H<sub>2</sub> frequency axis. In the other GlcNAc residues (-2, -5, -5', -7), the active splitting is slightly smaller, resulting in less cancelation of the middle components of the cross-peak; however, these patterns are disrupted by the overlap of these cross-peaks. Glucose and Man are epimers at position 2; the small couplings between H<sub>1</sub>,H<sub>2</sub> ( $\alpha$ , 1.8 Hz;  $\beta$ , 0.9 Hz) and H<sub>2</sub>,H<sub>3</sub> (3.8 Hz) protons of the mannose result in the low or vanishing intensity in the DQF experiment as the line width is larger than the active coupling creating an autocancelation of the cross-peak. These cross-peaks can be more easily found in other types of two-dimensional experiments, for example, NOESY as in the upper portion of Figure 3, or in HOHAHA experiments (Figure 1). The Man-4, -4' H<sub>3</sub>,H<sub>2</sub> cross-peak appearance is due to the large passive H<sub>4</sub> splitting (10 Hz). The skewed Man-3 H<sub>3</sub>,H<sub>2</sub> is due to virtual coupling caused by the accidental overlapping of the H<sub>3</sub> and H<sub>4</sub> peaks. Galactose is a 4-position epimer of glucose; the H<sub>1</sub>,H<sub>2</sub> cross-peak of Gal should be similar in appearance to that of the GlcNAc. The overlap of the Gal-6 and Gal-6' distorts the appearance of the H<sub>2</sub>,H<sub>1</sub> cross-peak in Figure 4; at 40 °C (Figure 3), the Gal cross-peaks revert to the expected appearance. This situation is similar but reversed for the cross-peak labeled S in these spectra; the cross-peak is distorted at 40 °C because of the accidental overlap of  $\beta$  and  $\beta'$  cross-peaks at this temperature which are resolved at the lower temperature in Figure 4.

To differentiate the oligosaccharide reporter cross-peaks from those of amino acid AMX ( $\alpha\beta\beta'$ ) spin systems, a number of observations regarding their behavior in two-dimensional experiments are helpful. RECSY or HOHAHA experiments will generate new cross-peaks to the structural reporter protons

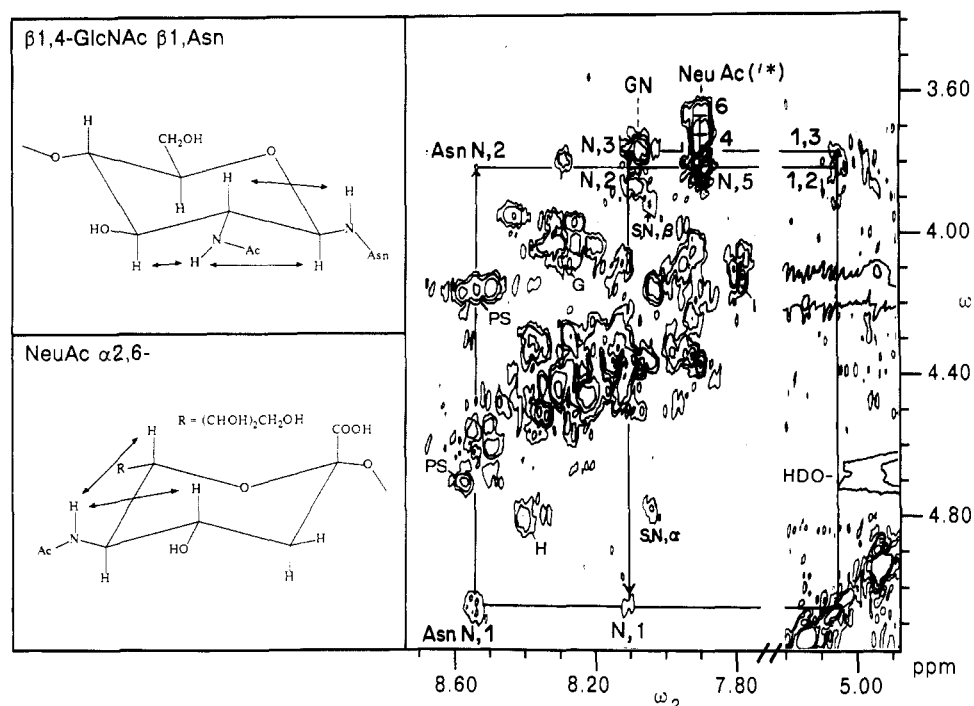


FIGURE 5: RECSY spectrum of phosvitin recorded in 90%  $\text{H}_2\text{O}$  at 40 °C and pH 5.0. Correlations to the GlcNAc-1 and NeuAc protons are outlined. The S cross-peak correlates to the  $\alpha,\beta$  cross-peaks labeled S in Figures 3 and 4. The direction of the relay is indicated by a small arrow and was determined from a COSY spectrum recorded on the same sample.

which also correlate into the crowded spectral region c. Triple quantum filter (TQF) experiments are useful for glycoproteins since this experiment isolates spin systems with mutual three-bond or larger coupling networks, leaving only the cross-peaks of H5,H6,H6'-like protons of hexoses remaining in the spectrum (Homans et al., 1986). TQF spectra will remove the anomeric and H2 structural reporter cross-peaks while leaving the AMX protein spin systems. One should be wary of long-range (four-bond) coupling which may allow these cross-peaks to transverse the triple quantum filter (Homans et al., 1986). We have included experiments in water to assign the exchangeable NH protons on the carbohydrate. The differences in the scalar connectivity patterns found for amino acid NH protons and the oligosaccharide NH protons in nonexchanged samples allow the oligosaccharide NH protons to be unambiguously identified and provide a method for checking the oligosaccharide assignments made in  $\text{H}_2\text{O}$ . The assignment of through-space NOE's should be consistent with the anomeric linkages in the oligosaccharide. We also assume that for small proteins (<100 000 daltons), the local flexibility of the oligosaccharide may allow detection of reporter protons even in proteins where short  $T_2$  relaxation or spin diffusion broadens many of the amino acid resonances below detection.

We will begin our assignment with GlcNAc-1 and proceed out to the distal end of the oligosaccharide chain combining certain groups, such as Gal-6, -6', and -8, where applicable. The labeling of the carbohydrate sequence follows the convention of Vliegthart et al. (1983) and is shown in Scheme I. The basis of our assignment comes from the comparison of several oligosaccharides of similar composition for which structural assignments have been reported previously in the literature. The work by Vliegthart et al. (1983) also provided the initial chemical shifts for a number of closely related complex oligosaccharides as well as the complicated empirical analysis of perturbations in chemical shifts of the reporter protons which result from alterations in the monosaccharide sequence. The closest previously assigned reporter protons of oligosaccharides we found [Vliegthart et al., 1983 (structure

34); Joziassse et al., 1987] have similar anomeric linkages and are capped with  $\alpha 2,6$  NeuAc's. Related structures assigned by Brisson and Carver (1983a), Bhattacharyya et al. (1984), Homans et al. (1984), and Berman et al. (1988) contained additional assignments helpful for fixing the cross-peak positions of the reporter protons and sorting out assignments of other structural possibilities.

The anomeric proton of GlcNAc-1 is assigned at 5.02 ppm and is coupled to H2 at 3.81 ppm (3.81,5.02 ppm). The apparent H2,H1 coupling constant and chemical shift are characteristic of the  $\beta$ -anomer attachment to an asparagine group. The assignment of H3 at 3.74 ppm can be made from a relay to H1. We can unambiguously assign this spin system to GlcNAc-1 even in the presence of many amino acid systems from the protein by recording COSY and RECSY spectra in  $\text{H}_2\text{O}$ . In water, the GlcNAc-1 spin system will include connectivities to the Asn and to the *N*-acetyl NH protons. A low pH is needed to acquire these spectra to reduce the exchange broadening of these NH protons with the solvent. In Figure 5, a RECSY spectrum recorded in water highlighting GlcNAc-1 is shown; GlcNAc-1 has a unique pattern of cross-peaks in this spectrum as the magnetization of the NH group of the oligosaccharide can be connected via the H2 and H1 protons to the NH of the asparagine link. The cross-peak at 8.56, 5.02 ppm is due to the Asn NH-GlcNAc-1 H1 correlation. There is a very weak relay from the Asn NH to H2. A COSY cross-peak from the *N*-acetyl NH and H2 appears at 8.09, 3.81 ppm; the relay peak from this NH to H1 is clearly visible while the relay to H3 is partially under the *N*-acetyl NH's of the remaining GlcNAc's. The NH resonances provide a second reference proton for assignment of the remaining protons in the saccharide ring and help to simplify the assignment of the anomeric region. The anomeric region is complicated by the appearance of large numbers of cross-peaks arising from the spin systems of amino acid residues, such as serine, threonine, and glycine, in the same region. A simple example is highlighted in the RECSY spectra of Figure 5; the cross-peak labeled S in Figure 4 at 3.84,4.76 ppm lies in an

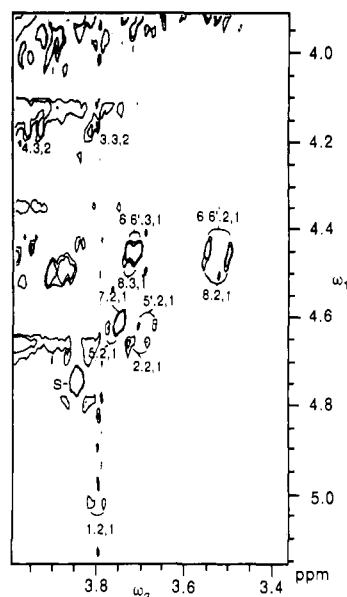


FIGURE 6: Expanded portion of the HOHAHA spectrum presented in Figure 2 providing the assignments of the crowded GlcNAc-2,-5,-5',-7 H2,H1 protons. There are slight negative distortions surrounding each peak; this may account for the nondoublet appearance of the GlcNAc-7 pattern.

area which could initially be assigned to either mannose, fucose, *N*-acetylglucosamine, or amino acid residues. The amino acid spin systems will generally show an upfield relay from the NH- $\alpha$ H to the NH- $\beta$ H proton while in GlcNAc the relay from NH-H2 to NH-H1 will be downfield. The S cross-peak in the previous HOHAHA spectrum has a corresponding NH resonance which relays upfield from 8.05,4.76 ppm to 8.05,3.84 ppm and can therefore be assigned to an amino acid residue rather than to a carbohydrate group. The NH cross-peaks of GlcNAc are easily resolved in this manner from those of the NH- $\alpha$ H protein resonances as well as from those of NeuAc. This should provide a generally applicable method to noninvasively assess the presence of *N*-acetylated carbohydrates on glycoproteins. In previous NMR analyses of oligosaccharides, the amide region has not been utilized to our knowledge.

The GlcNAc-2 H2,H1 cross-peak overlaps with the H1,H2 cross-peaks of GlcNAc-5, -5', and -7. This unresolved pattern at  $\approx$ 3.75,4.6 ppm is expected for the type of triantennary complex oligosaccharide we find in phosvitin which does not contain a bisected GlcNAc group. The severe overlap is expected to cause phase cancellations and shifts in the apparent resonance positions of these cross-peaks. For this reason, the chemical shift values of these cross-peaks in Table I are in parentheses. The cancellation effects of anti-phase peaks are reduced in the HOHAHA spectrum (Figure 6) where the absorption phasing produces positive cross-peaks. Expected relay cross-peaks to the H3 protons are weak or not observed for the GlcNAc residues in this HOHAHA spectrum; this may be the result of the low power used for this experiment and the small chemical shift difference of the H2 and H3 protons. This turned out to be advantageous for the assignment of these GlcNAc H2,H1 cross-peaks. A doublet pattern is generated in the  $\omega_2$  dimension for both GlcNAc and Gal H2,H1 cross-peaks. Using the cross-peak pattern for GlcNAc-1 and the slight differences in the anomeric chemical shifts found in model oligosaccharides, we have identified the approximate positions of the cross-peaks in Figure 6. The GlcNAc-2 H1 is found just downfield and the H2 upfield of the respective GlcNAc-5,-5' protons. The low chemical shift dispersity of

the GlcNAc H2 resonances is confirmed by the overlap of the NH-H2 cross-peaks in the H2 dimension in Figure 5. The relay peaks from the NH protons of these overlapping groups to their respective H1 protons have been saturated by the presaturation scheme required to remove the strong water signal in the spectra.

The H1 of Man-3 generally lies underneath the water signal, and in phosvitin, the H1,H2 cross-peak is further obscured by the presence of the strong  $\alpha,\beta$  cross-peaks from the phosphoserine residues. A tentative assignment from comparing NOESY and HOHAHA spectra can be made for H1,H2 at 4.78,4.20 ppm. The H3,H2 cross-peak of mannose can also be used as a structural reporter group as this cross-peak lies downfield of the crowded spectral region. The Man-3 H3,H2 cross-peak is found at 3.78,4.20 ppm; the initial assignment was made difficult due to the difference in the cross-peak fine structure and intensity between the Man-3 and the two more easily assigned Man-4 and Man-4' H3,H2 cross-peaks. This skewed Man-3 cross-peak has been found in model oligosaccharides by Homans et al. (1986) and Berman (1987); the lack of any apparent passive splitting of the H3 resonance by H4 was attributed to the identical chemical shifts of H3 and H4. We must stress that though this cross-peak is weak and difficult to detect due to the virtual coupling of H3 and H4, the cross-peak appears in a number of phosvitin spectra under different conditions and from different preparations of the protein. Further conformation of the Man-3 H3,H2 comes from NOESY spectra recorded with the same sample under similar conditions to the DQF experiment. An NOE across the Man4 $\alpha$ 1,3Man3 linkage is observed (5.18,3.78 ppm; Figure 3) between the H1 of Man-4, and the H3 of Man-3 can be traced to the Man-3 H3,H2 cross-peak in the DQF-COSY experiment.

Man-4 and -4' are well resolved from H1 to H3. The Man-4 chemical shifts for H1, H2, and H3 are 5.14, 4.18, and 3.90 ppm, respectively, while the Man-4' shifts are 4.92, 4.07, and 3.86 ppm, respectively; these shifts are characteristic of branching mannose groups in complex oligosaccharides. In the H3,H2 cross-peak, the passive coupling to H4 is larger than the active H3,H2 coupling; this pattern, while characteristic of most mannose H3,H2 cross-peaks, is not unique to mannose and also may be found in amino acid AMX spin systems. The  $\alpha,\beta$  cross-peak of threonine, for example, at 4.42,4.48 ppm in Figure 4, has a similar appearance and is found in the same area of the spectrum. The conformation of the assignment of Man-4 and Man-4' requires the correlation to the H1,H2 cross-peak. In the Man-4', there is four-bond coupling between H2 and H4 which is also resolved under certain conditions. Homans et al. (1986) have also reported the resolution of Man H4,H2 couplings in the spectra of model core pentasaccharides; the chemical shift of this Man-4' H4 from the H4,H2 cross-peak is 3.57 ppm. The GlcNAc $\beta$ 1,4 linkage to Man-4 deshields the H4 proton as compared to the chemical shifts found in biantennary oligosaccharides; the H4,H3 cross-peak (3.63,3.90 ppm) is partially resolved in the crowded region. The  $\beta$ 1,2 linkages to GlcNAc-5,-5' put the H1 protons of the GlcNAc residues in proximity to the respective Man-4 and Man-4' H1 protons (Figure 3, NOESY); similar NOE's across the  $\beta$ 1,2 linkages in model oligosaccharides have been found by Cumming et al. (1989).

The cross-peak patterns of the three mannose groups appear similar to 2D spectra published by Homans et al. (1983) and Bhattacharyya et al. (1984). This information shows that phosvitin does contain the common core structure [Man $\alpha$ 1,3(Man $\alpha$ 1,6)Man $\beta$ 1,4GlcNAc $\beta$ 1,4GlcNAc $\beta$ 1,N] and

not the unusual structure [GlcNAc(Man),Man,Man,GlcNAc] originally proposed by Shainkin and Perlmann (1971a,b). Further, the chemical shifts of the mannose branching structure are sensitive to further substitutions on the biantennary structure (Vliegthart et al., 1983) and can be used to determine the linkage of the extra GlcNAc and Gal determined in the chemical analysis by Shainkin and Perlmann but not assigned in the oligosaccharide sequence. The chemical shifts of the Man-4' H1 at 4.92 ppm and the Man-3 H2 at 4.20 ppm further suggest that the third antennary structure is linked as Gal $\beta$ 1,4GlcNAc $\beta$ 1,4 to Man-4. It is possible that small populations of heterogeneous structures, like those found in fetuin, exist and that these minor forms do not present enough signal intensity to be resolved in our experiments. Any major heterogeneity is likely to be confined to the NeuAc position or possibly to limited degradation of the Gal residues. However, we see no evidence for other structures, such as high mannose, a bisecting GlcNAc, a 4'-linked arm, fucose residues, or 1-3-linked galactose, in our spectra.

The GlcNAc-5, -5', and -7 are assigned by using the approach taken for GlcNAc-2. Unfortunately, the chemical shifts of the H1's and H2's overlap strongly in this type of oligosaccharide. Further, the H3 peaks are unresolved from the H2 peaks, reducing the possibilities of assignment via relay correlations. Using the expected chemical shifts from model compounds together with the cross-peak pattern of GlcNAc-1, we have boxed in approximations of the H2,H1 cross-peaks in Figure 6 which would yield a similar overall result. The GlcNAc-5 H2,H1 and GlcNAc-5' H2,H1 are found at 3.79,4.62 and 3.71,4.62 ppm, respectively. The H1's of GlcNAc-5 and -5' overlap with the  $\alpha$  protons of the phosphoserine groups; for this reason, the NOE assignment (Figure 3) is only tentative though it corroborates the other chemical shift data. The assignment of the GlcNAc-7 H2,H1 (3.76,4.60 ppm) is based on the slight upfield position of the H1 proton in model oligosaccharides. For the structural assignment of these GlcNAc's, the additional information available from the methyl resonances (discussed below) is helpful to confirm these assignments.

The Gal-6 and Gal-6' have similar chemical shifts for H2,H1 (3.55,4.44 ppm); this has also been reported for similar triantennary structures. There are strong relay cross-peaks to H3 (3.74,4.44 ppm) in the HOHAHA spectrum, as expected when H4 is weakly coupled, but the chemical shift dispersion of H3 is also small. The H1, H2, and H3 chemical shifts of the nonsialated galactose were found to be 4.47, 3.54, and 3.73 ppm, respectively. It is possible that there could be some heterogeneity in the distal oligosaccharide residues; this type of heterogeneity is difficult to detect using the reporter proton chemical shifts. No H3 cross-peaks associated with  $\beta$ 1,3-linked galactose residues were detected. The chemical shifts of the H3 protons are sensitive to the NeuAc linkage, and only cross-peaks associated with  $\alpha$ 2,6 NeuAc were found.

The NeuAc H3a,H3e,H4 protons make a distinctive cross-peak pattern which is not filtered out in TQF-COSY experiments and appears very intense because of the two-bond coupling present (Figure 7). The fine structure of the NeuAc cross-peaks can be differentiated from those of the amino acid resonances found in this region (Pro, Met, Glx, Arg, and Lys). In cases where overlap is suspected, a more difficult four quantum filter experiment could be tried to selectively remove the NeuAc residues. The H3a's and H3e's have chemical shifts of 1.72 and 2.76 ppm, respectively, characteristic of the NeuAc $\alpha$ 2,6Gal $\beta$ 1,4 linkage. Again, no chemical shifts associated with  $\alpha$ 2,3-linked NeuAc's (1.80,2.67 ppm) were ob-

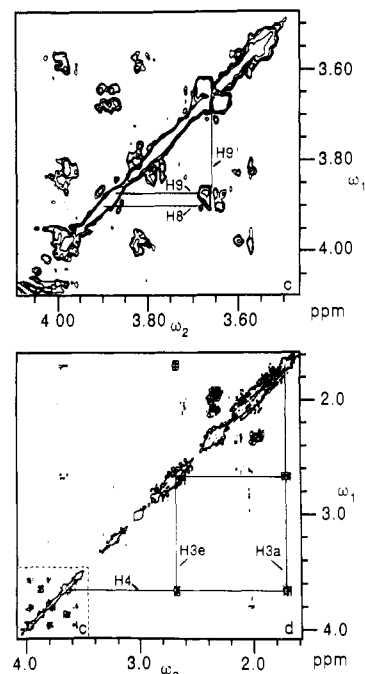


FIGURE 7: TQF-COSY spectrum recorded in 90% H<sub>2</sub>O at 40 °C and pH 4.85 showing the NeuAc cross-peaks. Only spin systems composed of at least three protons which are mutually coupled will pass through the filter. In the lower portion, the H3a,H3e correlations are highlighted; the Met  $\beta$ , $\gamma$  cross-peak centered at 2.10,2.62 ppm and Pro  $\gamma$ , $\delta$  (2.04,3.80 ppm) cross-peaks can be differentiated from the NeuAc groups by comparing the scalar correlations. Above in the expansion of region c, the application of the filter significantly reduces the number of cross-peaks in the crowded region (compare Figure 3). The NeuAc-8,-9,-9' cross-peaks have been traced out as an example.

served in the spectra of phosphatidylcholine. The chemical shift of H4's is 3.77 ppm, and there are relays to the H5's at 3.84 ppm and to the H6's at 3.70 ppm from the H3 protons (data not shown). The assignment of H5 is confirmed by the spectrum recorded in H<sub>2</sub>O (Figure 5) where the NH-H5 cross-peak (7.92,3.84 ppm) and its relay to H4 (7.92,3.77 ppm) and H6 (7.92,3.69 ppm) are also observed. The NeuAc cross-peaks in general appear to be narrower and better resolved than the corresponding GlcNAc cross-peaks. One possible reason for the enhanced line shapes is the negative charge which will be strongly repelled by the polyanionic protein. This will result in enhanced local motion in the terminal portion of the oligosaccharide branches. The NH proton of NeuAc, which is coupled one relay step closer to the H6-H9 protons, appears to be an effective resonance for assigning multiple relay spectra of sialated carbohydrates as an alternative to the H3a,H3e protons.

In the TQF spectra (Figure 7), the H8,H9,H9' protons of NeuAc will also pass through the filter in addition to the H5,H6,H6' protons of the hexose carbohydrates. The application of the triple quantum filter greatly simplifies the crowded oligosaccharide region c. Using the chemical shifts reported for an  $\alpha$ 2,6 NeuAc in a <sup>13</sup>C-<sup>1</sup>H heteronuclear experiment (Berman et al., 1988), we have tentatively traced out the chemical shifts of the NeuAc H8,H9,H9' protons in the TQF expansion of the crowded oligosaccharide region. From a HOHAHA experiment recorded with a 81-ms mixing time, we have been able to obtain a number of relays to H5 protons (GlcNAc-1, 3.56 ppm; Man-3, 3.66 ppm; Man-4, 3.78 ppm; Man-4', 3.64 ppm; GlcNAc-5 or -5', 3.61 ppm). The H5,H6,H6' cross-peaks which are found in region c appear to be from the more distal portion of the oligosaccharide with GlcNAc-1, GlcNAc-2, and Man-3 apparently missing in the



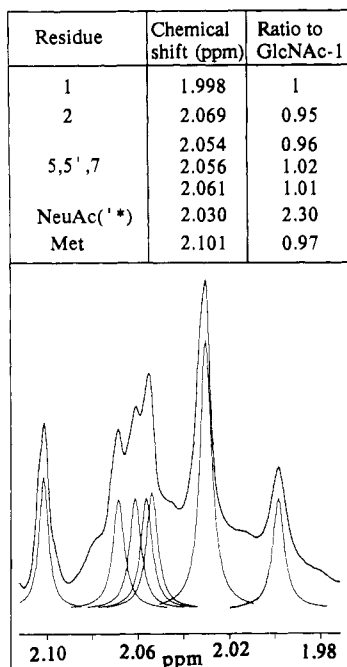


FIGURE 8: Expansion of the 1D proton spectrum for the methyl protons in a nonoxidized sample. These spectra were obtained without decoupling the water signal, and a relaxation delay of 4 s was used to ensure accurate intensities of these signals. The overlapping methyl signals were determined from a line simulation of the spectrum. The simulated peak positions and intensities appear below the spectrum, and the values are listed in the table above.

spectra. In contrast to the anomeric protons, presently there is insufficient data to empirically assign these cross-peaks from assignments of related model compounds.

The methyl region of the carbohydrate spectrum is shown in Figure 8; this region also includes the methyl signal from the single methionine at 2.101 ppm; the methionine methyl may be selectively assigned by gentle peroxide oxidation of the sulfide to give the sulfoxide (see Figure 1). In addition, there are a number of unresolved protons from prolines, glutamates, and other amino acid resonances which contribute to the uneven base line in this region. The relative peak areas were obtained by using a peak simulation routine in which the base line has been corrected for the presence of the underlying protein peaks. The chemical shifts of methyl protons can be used to further characterize the positions of the GlcNAc and NeuAc groups. The methyl of GlcNAc-1 has been assigned from its characteristic upfield chemical shift at 1.998 ppm. All of the GlcNAc methyls in phosvitin are shifted upfield (lower ppm) slightly (0.01 ppm) from their reported positions in model oligosaccharides. The other GlcNAc methyl groups fall between 2.05 and 2.07 ppm. As previously mentioned, the methyl peak at 2.101 is due to Met-39, and only a small peak (0.1 intensity ratio) at 2.098 ppm remains after mild peroxide treatment. This is a valuable piece of information since in model compounds the chemical shift of the sialated GlcNAc-7 is downfield of GlcNAc-2 very near the Met-39 position while the asialated GlcNAc-7 shift is found upfield of the GlcNAc-2 methyl resonance. Thus, the very small remaining peak at 2.098 ppm under the Met-39 methyl could represent some sialated GlcNAc-7 fraction. However, the low peak intensity may simply be due to an underlying amino acid resonance. The GlcNAc-2 methyl is found at 2.069 ppm, and the intensity of the three other methyl groups (GlcNAc-7, -5, and -5') is found just upfield, suggesting that the GlcNAc-7 is asialated in the major form of the phosvitin oligosaccharide. The relative ratio of four methyls compared to the GlcNAc-1

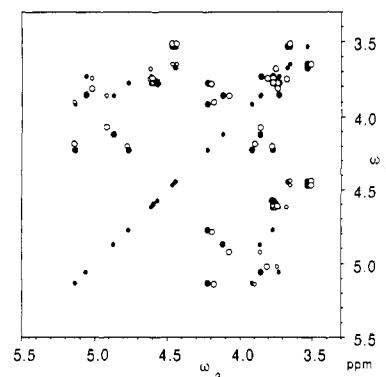


FIGURE 9: Comparison of the chemical shifts found for model oligosaccharides to those found for the phosvitin oligosaccharide. The open circles represent the phosvitin correlations (large circles, COSY type; small circles, RECSY from H1 to the H3 protons), and the solid circles indicate the data from model oligosaccharides.

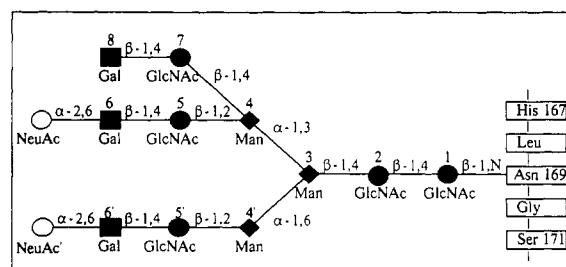


FIGURE 10: Major form of the oligosaccharide of hen phosvitin as determined by two-dimensional NMR in this study.

methyl was obtained, indicating a total of five GlcNAc's which agrees with the chemical analysis of Shainkin and Perlmann (1971a). The NeuAc methyl at 2.030 ppm has a relative area of 2.30 methyls as compared to GlcNAc-1.

In Figure 9, we highlight an overall comparison of the chemical shifts obtained for the structural reporter cross-peaks of the phosvitin oligosaccharide to the chemical shifts obtained from related model oligosaccharides. Our interpretation of the idealized carbohydrate structure of the glycoprotein phosvitin is presented in Figure 10. The pattern of the chemical shifts found in Figure 9 clearly indicates that phosvitin contains a type 1 complex oligosaccharide. It is interesting to note that some slight differences in reporter shifts do appear in the core region. This could be due to a number of factors. For example, the carbohydrate may interact in subtle ways with the intact protein, or, alternatively, different experimental conditions such as local ionic strength could preferentially affect these shifts. Further two-dimensional  $^1\text{H}$  NMR studies of glycoproteins in which the oligosaccharide portion can be assigned both in the intact form and following proteolytic digestion will allow this question to be examined. The average deviation of the anomeric resonances from their model counterparts is 0.016 ppm. As the chemical shifts of these structural reporter protons are quite sensitive to the shielding effects of nearby protons, the general agreement in chemical shifts suggests the overall three-dimensional structure of the triantennary oligosaccharide on phosvitin is quite similar to those found in model glycopeptides. The similar coupling patterns observed for many of the cross-peaks provide further support for this conclusion.

## DISCUSSION

The composition of the oligosaccharide of hen phosvitin had previously been determined by chemical methods as 5.0 GlcNAc's, 6.0 hexose sugars (3.0 Man and 3.0 Gal), and 2.0 NeuAc groups (Shainkin & Perlmann, 1971a). Our present



NMR analysis is quite similar; we found 5 GlcNAc's, 3 Man, 2.3 NeuAc's, and an unspecified amount of galactose. The Smith degradation that was originally performed correlates with the sequence we found through the first two cycles, but the remainder of the structure reported was in error. The chemical shifts obtained by two-dimensional NMR for the structural reporter protons of the phosvitin oligosaccharide are self-consistent with the data obtained by other researchers for model compounds. The chemical shifts of the cross-peaks found for the branching mannose groups in the phosvitin spectra indicate that the core sequence Man $\alpha$ 1,3(Man $\alpha$ 1,6)-Man $\beta$ 1,4GlcNAc $\beta$ 1,4GlcNAc $\beta$ 1,N is maintained in the oligosaccharide structure of this protein. The NeuAc $\alpha$ 2,6Gal $\beta$ 1,4GlcNAc $\beta$ 1,2Man arms are commonly found (Kornfeld & Kornfeld, 1985) in many glycoproteins. The position of the third arm was determined by comparing the chemical shifts found for 4 and 4' linkages (Vliegthart et al., 1983; Bhattacharyya et al., 1984). The results were consistent with a Gal $\beta$ 1,4GlcNAc $\beta$ 1,4 connected to the mannose 4 group. The chemical shifts of the GlcNAc-5, -5', and -7 methyl protons indicate that the two  $\alpha$ 2,6 NeuAc groups are positioned on the 6 and 6' galactose groups.

With the extensive knowledge we now have concerning the biosynthetic pathways for the glycosylation of proteins (Kornfeld & Kornfeld, 1985), the structure we propose would be one of the most likely given the original composition data (Shainkin & Perlmann, 1971a). The specificity of the biosynthesis is consistent with the anomeric linkages we have obtained. The NMR data also show that NeuAc's are connected as  $\alpha$ 2,6 and not  $\alpha$ 2,3, both of which may be found in various combinations in other glycoproteins. We are unaware of studies on the specificity of the chicken sialyltransferase, but a survey on the specificity of a number of  $\alpha$ 2,6 sialyltransferases reports that there is a  $6 \gg 6' > 8$  order of preference (Joziassse et al., 1987). This specificity agrees with the oligosaccharide structure obtained in this study. The incomplete sialation indicates that either partial synthesis or degradation of the oligosaccharide may have taken place as the complete capping by NeuAc is generally considered a completed oligosaccharide structural unit. It would be interesting to examine the bloodstream and liver preparations of phosvitin to determine whether this processing occurs and if it has any relevance to the biological function of this protein. During vitellogenesis, the protein synthesis machinery in the liver becomes almost totally dedicated to the production of vitellogenin (Tata, 1976); this could have an effect on the specificity of the biosynthesis, and also the quantity of precursors may need to increase to keep up with the demand. Once the protein is completely phosphorylated, ionic effects may inhibit any further proteolytic or glycolytic degradation of the protein. Thus, the lack of heterogeneity found in the glycosylation site of phosvitin may not be totally representative of other glycoproteins in this respect.

In addition to the chemical shift data, the cross-peak fine structures are consistent with our proposed structure. Even though the monosaccharide units are essentially all seven-member spin systems, the linearity of the system has the effect of removing the long-range correlations or passive couplings from the cross-peak patterns. The presence of small coupling constants in the ring also affects the degree of passive coupling, and our results clearly indicate that these patterns are useful for assigning the identity of the cross-peak. The inclusion of spectra recorded in 90% H<sub>2</sub>O/10% <sup>2</sup>H<sub>2</sub>O allowed us to obtain NH correlations, and this has provided useful information for the assignment of the GlcNAc and NeuAc systems.

GlcNAc-1 can be unambiguously assigned from its "unique" cross-peak pattern in H<sub>2</sub>O. Correlations to NH protons can be used to confirm the assignments of the ring protons made in <sup>2</sup>H<sub>2</sub>O, and in the case of NeuAc, they provided new assignments (H6) which had not been previously determined. The NH resonances provide potential probes to study structure and hydrogen exchange properties of oligosaccharides similar to the manner in which they are studied in proteins (Woodward & Tüchsen, 1985). Greater knowledge regarding the effects of pH and temperature on the shift and exchange rates of these amide protons would be useful. The assignment of protein backbone NH's is a matter of normal practice in 2D protein NMR spectroscopy, but the NH protons are generally not studied during carbohydrate structural analysis. Although sequence-specific assignment of the amino acid groups is not possible in the case of phosvitin due to its unusual polymeric nature, for other proteins it might be possible to obtain a starting point for sequence-specific assignment using an NOE from the GlcNAc-1 into the Asn linkage site.

We have shown that it is possible to analyze the major oligosaccharide form of phosvitin by 2D proton NMR while it is part of the native protein. This may also be true for other glycoproteins though each case will have to be determined on its own merit. For example, if a protein contains multiple oligosaccharide units that are heterogeneous, such as in fetuin (Green et al., 1988), this approach would be less successful, unless individual components could be prepared by lectin affinity chromatography, for example. Proteins in which there is considerable overlap of the oligosaccharide resonances will also present a more difficult challenge. Protein denaturation could be applied to remove overlapping protein cross-peaks from the structural reporter group resonances of oligosaccharides. Denaturation of the protein not only further exposes the oligosaccharide to the solvent (Tarentino & Plummer, 1982) but also shifts the  $\alpha$ - $\beta$  cross-peaks of the amino acid residues into their random coil positions (Wüthrich, 1986), similar to the chemical shifts found for phosvitin (Vogel, 1983; Brockbank and Vogel, unpublished results).

The dynamic motion of model oligosaccharides has been described as joint rigidity/flexibility. The core oligosaccharide is essentially rigid; the Man-4  $\alpha$ -1,3 linkage fluctuates between two low-energy conformers while the Man-4'  $\alpha$ -1,6 linkage is more fluid (Homans et al., 1986). As expected, there appears to be additional rotational freedom in the more distal residues as determined by the decrease in line width of the NeuAc and to a certain extent the Gal residues. As the line width, to a first approximation, is determined by the efficiency of spin-spin relaxation, local conformational flexibility will result in narrower lines even on large molecules. We note that flexibility for oligosaccharides on glycoproteins has been inferred from carbon-13 NMR relaxation studies (Dill et al., 1979; Berman et al., 1981; Goux et al., 1982) and from <sup>1</sup>H NMR (Cumming & Carver, 1987). Hence, it is unlikely that the mobility of the phosvitin oligosaccharide is caused solely by the unusual structure of this protein (Vogel, 1983), and therefore the same approach of analyzing oligosaccharides that remain attached onto a protein, by <sup>1</sup>H two-dimensional NMR, should be applicable to other glycoproteins as well. On the basis of the similarity in chemical shifts and the coupling patterns observed in this study, the single oligosaccharide of phosvitin appears to have a three-dimensional structure which closely resembles that observed in purified glycopeptides. Further studies with other glycoproteins will have to be performed to find out whether this is a general phenomenon or if interactions between the protein and its attached oligo-

saccharide(s) can cause perturbations in either of the structures or if the protein and its oligosaccharide behave simply as independent entities.

Two-dimensional  $^1\text{H}$  NMR spectra provide information on essentially the entire protein structure. Assignment of all of the resonances creates a large data base from which subsets of relevant information can be extracted. Here, we have assigned the oligosaccharide resonances, which creates both a chemical and a spatial subset of resonances which can then be used for further studies of this glycoprotein. For example, the degradation of the oligosaccharide chains by glycosidases may be monitored in vitro and in vivo situations to provide relevant information regarding the biochemistry and physiology of these reactions.

## REFERENCES

- Bax, A., & Davis, D. G. (1985) *J. Magn. Reson.* **65**, 355.
- Berman, E. (1987) *Eur. J. Biochem.* **165**, 385–391.
- Berman, E., Walters, D. E., & Allerhand, A. (1981) *J. Biol. Chem.* **256**, 3853–3857.
- Berman, E., Dabrowski, U., & Dabrowski, J. (1988) *Carbohydr. Res.* **176**, 1–15.
- Bhattacharyya, S. N., Lynn, W. S., Dabrowski, J., Trauner, K., & Hull, W. E. (1984) *Arch. Biochem. Biophys.* **231**, 72–85.
- Bock, K., & Thøgersen, H. (1982) *Annu. Rep. NMR Spectrosc.* **13**, 1–57.
- Bodenhausen, G., Kogler, H., & Ernst, R. R. (1984) *J. Magn. Reson.* **58**, 370–388.
- Brisson, J. R., & Carver, J. P. (1983a) *Biochemistry* **22**, 3671–3680.
- Brisson, J. R., & Carver, J. P. (1983b) *Biochemistry* **22**, 3680–3686.
- Byrne, B. M., van het Schip, A. D., van de Klundert, J. A. M., Arnberg, A. C., Gruber, M., & AB, G. (1984) *Biochemistry* **23**, 4275–4279.
- Cumming, D. A., & Carver, J. P. (1987) *Biochemistry* **26**, 6664–6676.
- Cumming, D. A., Hellerqvist, C. G., Harris-Brandts, M., Michnick, S. W., Carver, J. P., & Bendiak, B. (1989) *Biochemistry* **28**, 6500–6512.
- Dabrowski, J., Ejchart, A., Kordowicz, M., & Hanfland, P. (1987) *Magn. Reson. Chem.* **25**, 338–346.
- De Marco, A. (1977) *J. Magn. Reson.* **26**, 527–528.
- Dill, K., & Allerhand, A. (1979) *J. Biol. Chem.* **254**, 4524–4531.
- Goux, W. J., Perry, C., & James, T. L. (1982) *J. Biol. Chem.* **257**, 1829–1835.
- Green, E. D., Adelt, G., Baenziger, J. U., Wilson, S., & van Halbeek, H. (1988) *J. Biol. Chem.* **263**, 18253–18268.
- Homans, S. W., Dwek, R. A., Fernandes, D. L., & Rademacher, T. W. (1983) *Biochim. Biophys. Acta* **760**, 256–261.
- Homans, S. W., Dwek, R. A., Fernandes, D. L., & Rademacher, T. W. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6286–6289.
- Homans, S. W., Dwek, R. A., Boyd, J., Mahmoudian, M., Richards, W. G., & Rademacher, T. W. (1986) *Biochemistry* **25**, 6342–6350.
- Homans, S. W., Dwek, R. A., Soffe, B. N., & Rademacher, T. W. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 1202–1205.
- Joziasse, D. H., Schiphorst, W., Van den Eijnden, D. H., Van Kuik, J. A., Van Halbeek, H., & Vliegthart, J. F. G. (1987) *J. Biol. Chem.* **262**, 2025–2033.
- Kornfeld, R., & Kornfeld, S. (1985) *Annu. Rev. Biochem.* **54**, 631–664.
- Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **113**, 967–974.
- Müller, N., Ernst, R. R., & Wüthrich, K. (1986) *J. Am. Chem. Soc.* **108**, 6482–6492.
- Neuhaus, D., Wagner, G., Vasák, M., Kägi, J. H. R., & Wüthrich, K. (1985) *Eur. J. Biochem.* **151**, 257–273.
- Niemann, H., Dabrowski, J., Dabrowski, U., Geyer, R., Keil, W., Klenk, H. D., & Stirm, S. (1985) *Eur. J. Biochem.* **146**, 523–532.
- Piantini, U., Sørensen, O. W., & Ernst, R. R. (1982) *J. Am. Chem. Soc.* **104**, 6800–6801.
- Shainkin, R., & Perlmann, G. E. (1971a) *Arch. Biochem. Biophys.* **145**, 693–700.
- Shainkin, R., & Perlmann, G. E. (1971b) *J. Biol. Chem.* **246**, 2278–2284.
- Sundararajan, T. A., Sampath Kumar, K. S. V., & Sarma, P. S. (1960) *Biochim. Biophys. Acta* **38**, 360–362.
- Taborsky, G. (1983) *Adv. Inorg. Biochem.* **5**, 235–279.
- Tarentino, A. L., & Plummer, T. H. (1982) *J. Biol. Chem.* **257**, 10776–10760.
- Tata, J. R. (1976) *Cell* **9**, 1–14.
- Tüchsen, E., & Woodward, C. (1985) *J. Mol. Biol.* **185**, 405–419.
- Vliegthart, J. F. G., Dorland, L., & van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41**, 209–374.
- Vogel, H. J. (1983) *Biochemistry* **22**, 668–674.
- Vuister, G. W., de Waard, P., Boelens, R., Vliegthart, J. F. G., & Kaptein, R. (1989) *J. Am. Chem. Soc.* **111**, 772–774.
- Wagner, G. (1983) *J. Magn. Reson.* **55**, 151–156.
- Wallace, R. A., & Morgan, J. P. (1986) *Biochem. J.* **240**, 871–878.
- Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.