Demonstration of Heterogeneity of Chick Ovalbumin Glycopeptides Using 360-MHz Proton Magnetic Resonance Spectroscopy†

P. H. Atkinson,* A. Grey, J. P. Carver, J. Hakimi, and C. Ceccarini

ABSTRACT: Ovalbumin glycopeptides AC-C and AC-D at various stages of purification were studied by high-field proton magnetic resonance spectroscopy (^{1}H NMR). In a homogeneous substance, the intensity of the various resonances appears in integral amounts, while subintegral intensities usually denote mixtures of structure. We show how ^{1}H NMR can be used to nondestructively assay the purification of major components from mixtures. In glycopeptide AC-C we have spectroscopic evidence for the four different glycopeptide species, three of which have been described [Shepherd, V., & Montgomery, R. (1978) *Carbohydr. Res. 61*, 147; Tai, T., Yamashita, K., Ito, S., & Kobata, A. (1977) *J. Biol. Chem. 252*, 6687]. However, we did detect a fourth structure not previously reported. In glycopeptide AC-D, we have spectroscopic evidence for five different compounds, only two of which have been previously reported (Tai et al., 1977; Shepherd & Montgomery, 1978).

Ovalbumin is a glycoprotein which can be crystallized from hen's egg white (Neuberger, 1938). Several early studies suggested that ovalbumin contained oligosaccharide components of different sugar composition (Cunningham et al., 1963; Levvy et al., 1966), and five major classes of asparaginyl carbohydrate (AC-A, AC-B, AC-C, AC-D, and AC-E) were separated (Huang et al., 1970). Subsequently, eight unambiguously different species were separated and their structures determined (Tai et al., 1975, 1977; Yamashita et al., 1978). In a further fractionation of asparaginyl carbohydrate from ovalbumin, Montgomery's group reported that glycopeptides AC-C and AC-D each contained two subcomponents (Shepherd & Montgomery, 1978). Both laboratories observed a Man_6-$GlcNAc_2$-Asn structure in glycopeptide D but reported different compositions in the other subcomponents. Conchie & Strachan (1978) confirmed the assignments of Kobata's group for the major subcomponents of glycopeptides AC-C, AC-D, and AC-E.

Much oligosaccharide structural information can be derived directly from ^{1}H NMR spectroscopy. The anomeric hydrogen of each monosaccharide gives a resonance in a characteristic chemical shift range. These appear as a unique subset of shifts whose precise value is dependent on linkage, anomeric configuration, substitution, and the microenvironment of a given hydrogen in secondary structure. Assignments, however, depend on assignments previously made in a variety of model compounds. In asparagine-linked glycans of the *N*-acetyllactosamine and sialyl type, assignments have been made for oligosaccharides found in the urine of patients with GM1 and GM2 gangliosidosis (Wolfe et al., 1974; Strecker et al., 1977), in the asialocarbohydrate units of human serotransferrin (Dorland et al., 1977a), in sialyl oligosaccharides from patients with mucolipidosis I and II (Dorland et al., 1978), in α_1-acid glycoprotein (Fournet et al., 1978), in the carbohydrate–asparagine linkage moiety (Dorland et al., 1977b), in fucosylglycoasparagines in the urine of patients with fucosidosis (Strecker et al., 1978), and in the single glycan of chicken ovotransferrin (Dorland et al., 1979), and the asialo glycopeptides from human α_1-acid glycoprotein have been examined at 500 MHz (van Halbeek et al., 1980). Much less work has appeared on the oligomannosyl glycopeptides. However, early pioneering work on yeast mannans and galactomannans (Gorin et al., 1968, 1969; Gorin & Spencer, 1968) at 100 MHz provided some chemical shift assignments for resonances characteristic of this type of oligosaccharide, and a number more were made at 360 MHz for various glycopeptides (Narasimhan et al., 1980). Cohen & Ballou (1980) reported ^{1}H NMR structural assignments at 180 MHz of some avalbumin glycopeptides and of IgM-derived glycopeptides, as well as for a series of oligosaccharides purified from yeast mannans. A report (Carver et al., 1981) shows how structures can be assigned to various ovalbumin glycopeptides at 360 MHz by comparison of groups of key resonances to the values published for known structures. Another report extends this and generalizes an approach to determining structures of unknown compounds by high-field proton NMR spectral analysis alone (J. P. Carver and A. A. Grey, unpublished results).

In this paper, we show that high-field ^{1}H NMR can be used to nondestructively assay homogeneity in oligomannosyl and "hybrid" (Yamashita et al., 1978) asparaginyl glycopeptides from chick ovalbumin and monitor their subsequent purification. In ovalbumin asparaginyl carbohydrates AC-C and AC-D we can predict a total of nine different glycopeptides, five more than have been previously reported.

Materials and Methods

Preparation of Ovalbumin and Ovalbumin Glycopeptides. Ovalbumin was purified (Kekwick & Canaan, 1936) from eggs purchased at local supermarkets. On polyacrylamide gel electrophoresis a single band was generally observed. However, in several preparations two bands were detected (Figure 1), and further recrystallizations did not alter this pattern. The major band had an apparent molecular weight of 45 000.

Preparation and Fractionation of Ovalbumin Glycopeptides. Ovalbumin glycopeptides were obtained after exhaustive Pronase digestion (Johansen et al., 1969; Nuenke & Cunningham, 1969) and separated by using Dowex AG-50W-X2 (Huang et al., 1970) (200–400 mesh, Bio-Rad Laboratories, Richmond, CA). Figure 2 shows a typical profile of the five

† From the Departments of Developmental Biology and Cancer, Albert Einstein College of Medicine, Bronx, New York 10461, the Department of Medical Genetics, University of Toronto, Toronto, Ontario, Canada M5S 1A8, and the Department of Biology, Hunter College, New York, New York 10021. *Received February 12, 1980; revised manuscript received January 6, 1981.* This work was supported by grants from the National Institutes of Health (CA13402, GM24071, and CA13330) and the National Cancer Institute of Canada and by Grant MRC-MT3732. P.H.A. was an Established Investigator of the American Heart Association during this work, and J.H. is a Postdoctoral Fellow of the Leukemia Society of America.

* Address correspondence to this author at the Albert Einstein College of Medicine.

0006-2960/81/0420-3979$01.25/0

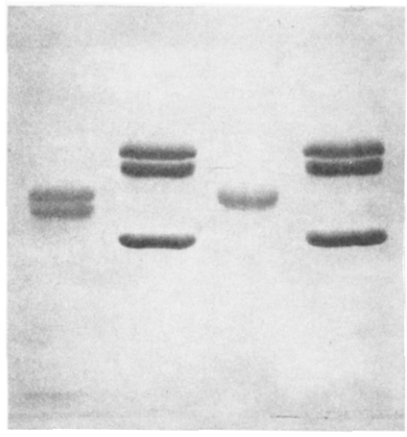


FIGURE 1: $NaDodSO_4$ gel electrophoresis of different ovalbumin preparations: (from left to right) ovalbumin preparation I, 20 μg; purified *Sindbis* virus grown in chick embryo fibroblasts, 65 μg; ovalbumin preparation II, 20 μg; *Sindbis* virus. Proteins were separated on 10% $NaDodSO_4$–polyacrylamide gels as described by Maizel (1970), and bands were identified by Coomassie Blue staining.

major ovalbumin glycopeptides. Each of the major species was checked for amino acid composition. Asparagine was the only significant amino acid present (see below, as previously shown; Lee & Montgomery, 1962). Bio-Gel P2 (–400 mesh, Bio-Rad Laboratories) columns of 0.9 × 170 cm were equilibrated in H_2O containing 1 mM sodium azide and were eluted in a 4 °C room at a flow rate of 1 mL/h. Elution of glycopeptides was followed by the phenol–sulfuric acid method of Dubois et al. (1956).

Preparation of Markers. Oligosaccharide markers were prepared from the major homogeneous components of AC-A, AC-C, AC-D, and AC-E by using endo-β-*N*-acetylglucosaminidase-H. This enzyme was purified and used as described (Tarentino & Maley, 1974). It had an activity of 0.46 unit/mL, and 200 μL was used in a total volume of 400–1000 μL for 3–10 μmol of glycopeptide at 37 °C for 15 h. Oligosaccharides were purified from the digestion mixture by Sephadex G-25 chromatography followed by Bio-Gel P2 chromatography. Compositions of the oligosaccharides were Man_5-Gal-$GlcNAc_4$[oligo(A)], Man_5-$GlcNAc_3$[oligo(C)], Man_6-GlcNAc[oligo(D)], and Man_5-GlcNAc[oligo(E)] as determined by gas–liquid chromatography (GLC).

Carbohydrate Analysis. GLC was performed on a Hewlett-Packard chromatograph, Model 5830A, equipped with a flame ionization detector using a 6 ft × $^1/_8$ in. glass column packed with 3% OV-225 on Supelcoport (Supelco, Inc., Bellefonte, PA) as described by Grimes & Greegor (1976). Glycopeptides were hydrolyzed in 2 N HCl sealed under nitrogen at 100 °C for 2.5 h. Samples were neutralized with NaOH to a neutral red end point and dried in vacuo. Neutral and amino sugars were quantitated as alditol acetates (Albersheim et al., 1967). Dried hydrolysates were resuspended in 0.2 mL of saturated $NaHCO_3$ and re-N-acetylated with 10 μL of acetic anhydride for 10 min at room temperature. Samples were dried, 2-deoxyglucose was added as the internal standard, and the samples were reduced in 1 M NH_4OH containing 2 mg/mL sodium borohydride. Methyl borate was removed by five additions of methanol containing 0.5% acetic acid. Monosaccharides were acetylated with 1.0 mL of redistilled acetic anhydride in sealed tubes under nitrogen at 100 °C for 1–1.5 h and then dried in vacuo. The alditol acetates were extracted from salts with 1 mL of redistilled acetone by using mild sonication for 10 min. Samples were centrifuged, and the supernatants were taken to dryness under N_2 at 20–25 °C. For analysis, the alditol acetates were redissolved in redistilled ethyl acetate, and 2–4 μL was injected. Quantitation of asparagine and glucosamine was kindly performed by Dr. Olga Blumenfeld (Department of Biochemistry, Albert Einstein College of Medicine). Samples were hydrolyzed with 2.9 and 5.8 N HCl in sealed vials under N_2 at 100 °C in a boiling water bath for 4 h or an oven for 24 h, respectively. Glucosamine was chromatographed on a short column in a JEOL amino acid analyzer; asparagine and glucosamine were observed on a Durrum column. Neutral sugars and *N*-acetylglucosamine in aliquots of identical volume were quantitated by GLC analysis. Values obtained in the GLC analysis were normalized to asparagine/glucosamine values determined in the amino acid analysis in order to obtain the asparagine/glucosamine/mannose ratios as reported in Table II.

NMR Methods. (*a*) *Sample Preparation.* Oligosaccharides and glycopeptides not fractionated on Bio-Gel P2 columns were desalted by using a Sephadex G-10 column equilibrated with H_2O containing 1 mM sodium azide. Where substances had been treated with glycosidases, a prior gel filtration on a Sephadex G-25 fine (0.9 × 170 cm) column was performed to remove protein. Divalent heavy metal ions were removed by subsequent column chromatography on Chelex-100 (200–400 mesh, Bio-Rad Laboratories; 0.9 × 10 cm) in the Na^+ form. Elution of substances was followed by absorption at 230 nm. Samples were exchanged twice in deuterium oxide (Merck Sharp and Dohme, Montreal, Canada; 99.7 atom % D) and then stored for several days in a vacuum desiccator over P_2O_5. Samples were redissolved in deuterium oxide (Stohler Isotope Chemicals, Waltham, MA; 100.0 atom % D) in an appropriate volume to give a concentration between 1 and 20 mM and examined in 5-mm tubes or 100-μL microcells (Wilmad Co.), depending on amount of material available. Acetone was added to roughly equimolar concentrations. Carbon tetrachloride or tetrachloroethylene was used as the spectrally transparent medium for suspension of microcells in the 5-mm tubes.

(*b*) *Instrumental Conditions.* Proton magnetic resonance spectroscopy was performed mostly by using the 360-MHz facility at Brookhaven National Laboratories with the kind permission of Dr. Alan McLaughlin. Alternatively, the 360-MHz facility at the Purdue University Biological Magnetic Resonance Laboratory was used. Spectra were obtained by Fourier transformation (FT mode) of accumulated free induction decay (FID) signals after 16–512 pulses, 16K data points, and a delay between pulses of 4–6 s. Spectral widths of 4000 Hz were examined. Samples were examined at 20–24 °C and at a set temperature between 60 and 80 °C, in order to move the residual HDO resonance from the anomeric region. Caution must be employed in comparing chemical shifts at different temperatures since several resonances were found to be quite temperature dependent. Chemical shifts were calculated by setting the resonance from internal acetone at 2.225 ppm [relative to external sodium 4,4-dimethyl-4-silapentanesulfonate (DSS)] at all temperatures. The invariance of the acetone chemical shift has been verified by direct measurement. In addition, a calibration for the chemical shift of HDO as a function of temperature has been determined and used to confirm the temperature of the sample. Many of the preliminary experiments were performed at 220 MHz on the modified Varian HR-220 at the Toronto Biomedical NMR Center.

Results

As originally reported by Huang et al. (1970), ovalbumin

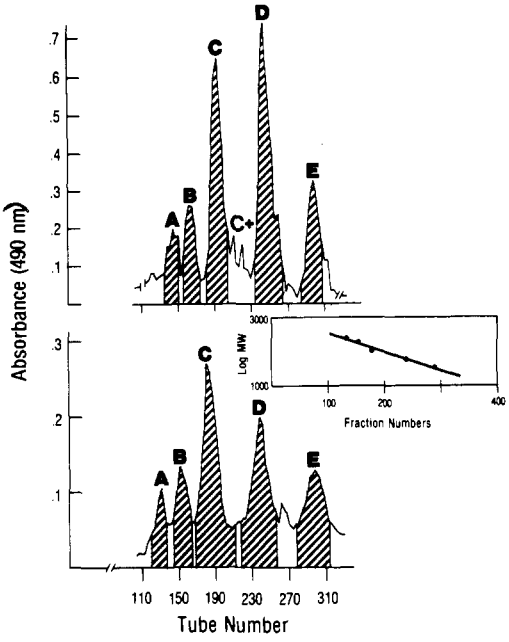


FIGURE 2: Fractionation and pooling of ovalbumin glycopeptides on Dowex AG 50-X2 resin. Two individual crystalline ovalbumin preparations were subjected to exhaustive Pronase digestion followed by Sephadex G-25 chromatography. Approximately 410 mg of neutral sugar from ovalbumin preparation I (see Figure 1) glycopeptides (top) and 660 mg of hexose from ovalbumin preparation II (see Figure 1) glycopeptides (bottom) were chromatographed on a 2.5 × 140 cm column of Dowex AG 50-X2 as described under Materials and Methods. Pooling of the major peaks indicated by the crosshatched areas was relatively narrow (top) or broad (bottom). (Inset) Relationship between the log of the molecular weight and the elution fraction of the major glycopeptide component.

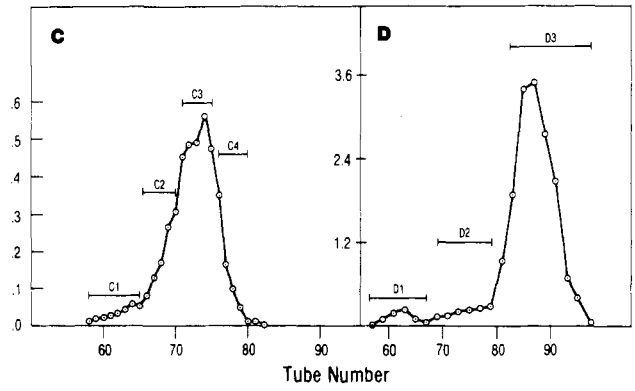


FIGURE 3: Bio-Gel P2 gel chromatography of glycopeptides AC-C and AC-D. Ovalbumin glycopeptides purified from two Dowex columns were applied to Bio-Gel P2 columns, 93.8 and 59.6 mg of neutral sugar for AC-C and AC-D, respectively. Columns were eluted and assayed as described under Materials and Methods. Optical density profile of neutral sugars as assayed by the phenol–sulfuric acid method is shown.

glycopeptides were fractionated into five peaks, and we have adopted the terminology of these authors, namely, AC (asparaginyl carbohydrate)-A, AC-B, AC-C, AC-D, and AC-E. Two separate fractionations are shown where pooling of the major peaks was relatively narrow (Figure 2, top) or broad (Figure 2, bottom). As had been previously reported (Huang et al., 1970), the glycopeptides eluted in order of their average molecular size. There was a linear relationship between the log of the average molecular weight and the elution fraction of the major glycopeptide component (inset, Figure 2). Three size classes were pooled after Bio-Gel P2 chromatography (Figure 3), except for AC-C which was pooled into four fractions. For each Dowex peak, fraction 1 contained 5–12%,

Table I: Subclasses of Major Glycopeptides of Ovalbumin

ovalbumin Dowex[a] glycopeptides	distribution in Bio-Gel P2 subfraction[b] 1	2	3
AC-A (5.1)	11.5 (0.6)	13.0 (0.7)	*75.5* (3.8)
AC-B (11.8)	6.7 (0.8)	13.7 (1.6)	*80.0* (9.4)
AC-C (38.0)	5.0 (1.9)	16.5 (6.3)	*78.5* (29.8)[c]
AC-D[d] (28.3)	6.7 (1.9)	7.7 (2.2)	*85.7* (24.2)
AC-E (16.9)	5.5 (0.9)	10.5 (1.8)	83.5 (24.2)

[a] Ovalbumin glycopeptides were fractionated on a Dowex AG 50X column as described. [b] Each Dowex peak was subfractionated on Bio-Gel P2 as described, and the distribution of neutral sugars in each fraction is shown. Subspecies percentages in AC-B and AC-D are an average of two columns. Numbers in parentheses are the percentage of each subclasses in unfractionated ovalbumin glycopeptides. Neutral hexose determination was by the phenol-sulfuric acid method. Figures in italics are the relatively homogeneous components. [c] Includes C3 and C4.

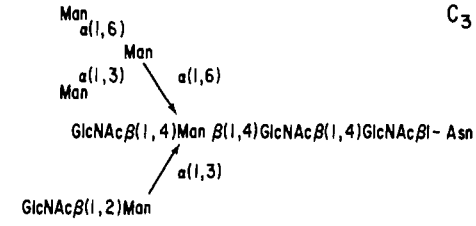
Table II: Gas–Liquid Chromatography Compositions: Comparisons of Ovalbumin Glycopeptide Subspecies

ovalbumin species[a]	ovalbumin subspecies[b] molar ratios[c] (Man/Gal/GlcNAc) subspecies 1	subspecies 2	subspecies 3[d]
AC-C[f]	1:0.128:1.58	1:0.063:1.46	1:<0.05:0.732
AC-D[f]	1:<0.05:0.218	1:<0.05:1.72[e]	1:<0.05:0.248

[a] Dowex chromatography nomenclature. [b] Bio-Gel chromatography nomenclature. [c] All molar ratios are normalized with respect to one mannose residue. [d] All three species are purified subspecies. [e] This ratio is of material from a second Bio-Gel P2 subfractionation of AC-D2. It is the major component of AC-D2 and is called AC-D2B. [f] Amino acid analysis was performed on AC-C3 and AC-D3. Asn was the only amino acid present, giving a composition of Man/GlcNAc/Asn of 5.14:4.3:1 and 6.01:1.83:1 for AC-C3 and AC-D3, respectively.

fraction 2, 7–14%, and fraction 3, 75–85% of the neutral sugars present in parent Dowex glycopeptide (Table I). In this paper we present data concerning the homogeneity and/or heterogeneity of two of the five major ovalbumin glycopeptides, AC-C and AC-D. The carbohydrate composition of the Bio-Gel P2 subfractions of AC-C and AC-D was determined by gas–liquid chromatography (Table II). This analysis confirmed that the only saccharides present in the various subfractions were mannose, galactose, and *N*-acetylglucosamine.

Glycopeptides of AC-C. We examined the homogeneity of the various Dowex and Bio-Gel P2 subfractions by using 360-MHz ^{1}H NMR spectroscopy. Assignments of resonances in the major structure in AC-C (termed "AC-C3B") are made in another report (Carver et al., 1981). AC-C3B has the composition Man_5-$GlcNAc_4$-Asn (Table II), and the structure is deducible from its ^{1}H NMR spectrum:



C$_3$

Other compounds present in AC-C have similar molecular weights but differ in the ratio of Man/GlcNAc present. ^{1}H resonances of mannose linked α1,6, α1,3, α1,2, and β1,4 and also *N*-acetylglucosamine linked β1,4 and β1,2 all fall in distinct chemical shift ranges. In the few cases where overlap

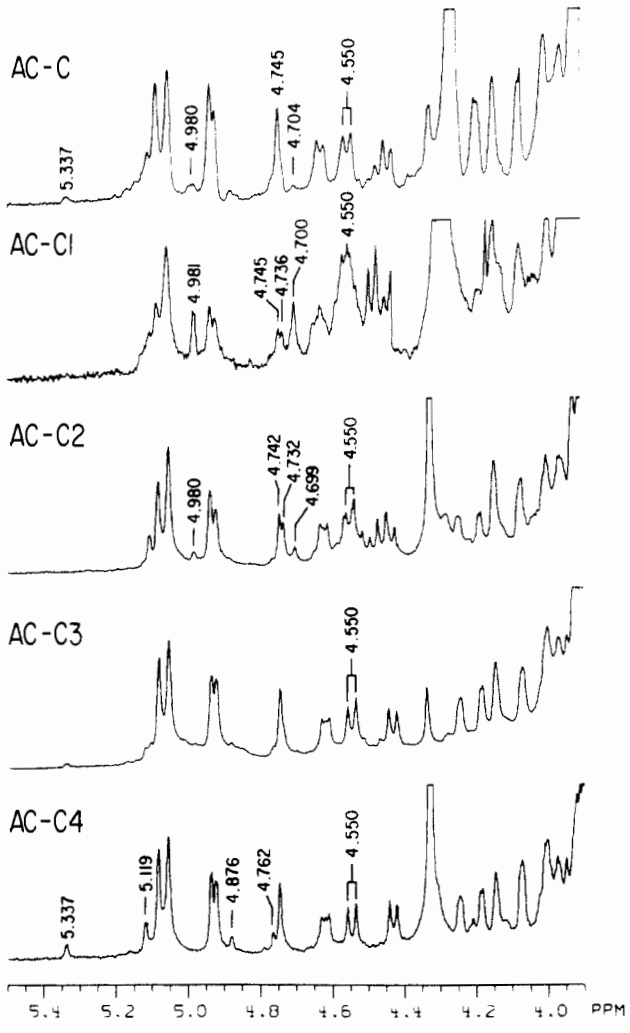


FIGURE 4: ^{1}H NMR spectra of subfractions of AC-C. AC-C, 26 mg of neutral sugar in 1.8 mL of D_2O in a tube, was examined at 80 °C with four scans. AC-C1, 4.3 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined at 72 °C with four scans. AC-C2, 16.3 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined at 70 °C with four scans. AC-C3, 20 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined at 69 °C with 64 scans. AC-C4, 15 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined at 70 °C (46 °C nominal as C1, C2, and C4) with four scans. Water suppression was not used in any of these spectra.

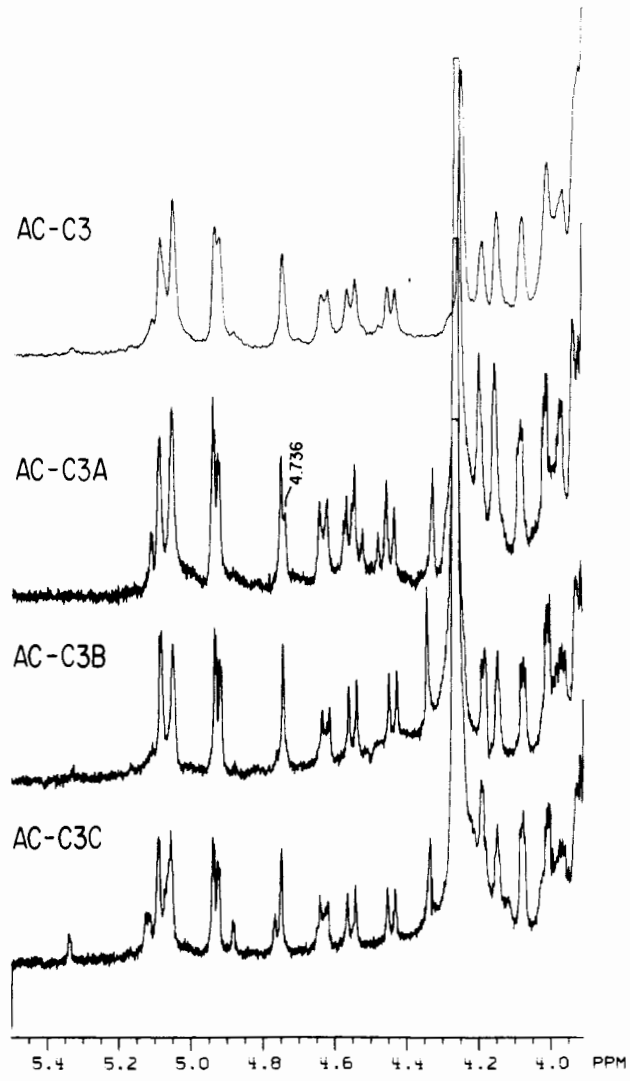


FIGURE 5: ^{1}H NMR spectra of subfractions of AC-C3 at 360 MHz. AC-C3, 20 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined as described in Figure 4 except at a slightly higher temperature. AC-C3A, 3.8 mg in 0.5 mL of D_2O in a tube, was examined at 80 °C with 16 scans. AC-C3B, 0.5 mg in 0.1 mL of D_2O in a microcell, was examined at 80 °C with 16 scans. AC-C3C, 1.0 mg in 0.1 mL of D_2O in a microcell, was examined at 80 °C with 16 scans. Water suppression was not used in any of these spectra.

occurs, the monosaccharide and linkage giving rise to any particular ^{1}H resonance can be distinguished on the C1-H and C2-H pair of chemical shift values. Within the C1-H chemical shift range for any particular linkage, discrete resonances are observed as a result of overall molecular structure (J. P. Carver and A. A. Grey, unpublished experiments). For example, mannose (β1,4) resonances, at 70 °C, are found in the range 4.696–4.770. We can, therefore, deduce that in addition to AC-C3B whose (1→4)-β-mannose gives rise to a resonance at 4.745 ppm in the spectrum of AC-C, there are three other nonidentical compounds whose mannose β1,4 C1 hydrogens have resonances at 4.762, 4.732–4.736, and 4.699–4.704 ppm. Enrichment in intensity of minor resonances characteristic of a given compound can be observed as it is enriched by Bio-Gel P2 chromatographic fractionation. These observations are now more fully detailed.

AC-C1 (Figure 3) is a mixture of glycopeptides, but enriched intensities at ~4.980 and ~4.700 ppm are seen (Figure 4) as minor components in the spectrum of parental AC-C. There is also an increase in intensities at ~4.550 ppm which is characteristic of *N*-acetylglucosamine residues in β linkage external to the chitobiose core. Examination of resonances to high field (~2.05 ppm) due to the hydrogens of the methyl group in GlcNAc confirms that the major component of AC-C1 is enriched in GlcNAc compared to parent AC-C. The presence of a compound similar to that described by Narasimhan et al. (1980) is likely GN-GN(GN) (in their terminology).

AC-C2 is also a mixture. Two components have spectra which resemble the major compounds in AC-C and AC-C1 (Figure 4). However, there is an enriched resonance at 4.732 ppm which arises from a β1,4-linked mannose, while the resonances at ~4.980 and ~4.700 ppm characteristic of the GN-GN(GN)-like component of AC-C1 are markedly reduced in AC-C2. Hence, this resonance at ~4.732 ppm though present in the spectrum of AC-C1 and enriched in AC-C2 must belong to a different compound.

AC-C3 is near homogeneous (Figure 4) but still contains other structures as seen when it is subfractionated into AC-C3A, AC-C3B, and AC-C3C by a further step of Bio-Gel P2 chromatography (C3A elutes in the least volume; data not shown). The spectrum of the major component, AC-C3B, shows that it is more homogeneous than AC-C3 (Figure 5).

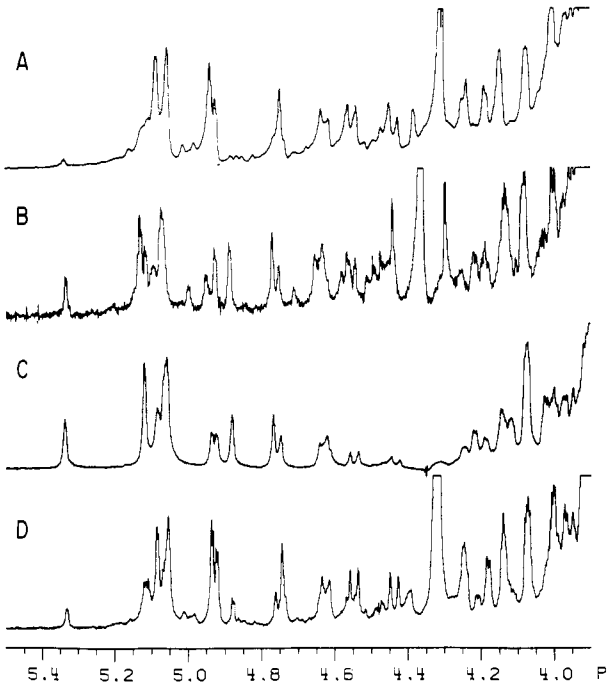
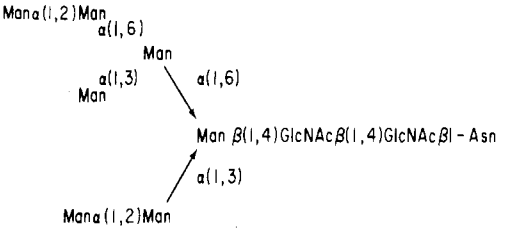


FIGURE 6: ^{1}H NMR spectra of Dowex column pools of AC-C at 360 MHz. (A) Narrowly pooled AC-C (see Figure 2, top), 17.4 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined 71.5 °C. (B) C^+ (region between AC-C and AC-D from Dowex columns; see Figure 2 top), 20.9 mg of neutral sugar in 1.0 mL of D_2O in a tube, was examined at 70 °C. (C) Broadly pooled AC-C (panel D) was rechromatographed twice on Bio-Gel P2, pooling the leading shoulder each time. The final fraction, ~2.0 mg of neutral sugar, was examined in 0.4 mL of D_2O at 70 °C. (D) Broadly pooled AC-C (see Figure 2, bottom), 30 mg of neutral sugar in 0.5 mL of D_2O, was examined at 72 °C. Water suppression was not used in obtaining these species.

AC-C3A is a mixture consisting mostly of the major component, AC-C3B, and a minor component similar to AC-C2 (see a minor (1→4)-β-Man resonance at 4.736 ppm). AC-C3C is ~70% AC-C3B but contains a unique set of minor resonances (Figure 5). A difference spectrum of AC-C3C and AC-C3B yields a spectrum for this minor component which is consistent with a composition of Man_7-$GlcNAc_2$-Asn.

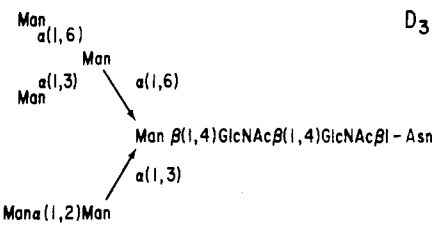
AC-C4 contains mostly the major component of ovalbumin, AC-C3B, but also contains 10–15% Man_7 glycopeptide as detected (Figure 4) by the presence of mannose resonances at 5.337 ppm. If the intensity of this resonance is arbitrarily set at 1, then the other resonances attributable to this compound are those units at 5.119 ppm (two intensities) and obscured resonances at 5.063 ppm (two intensities; these can be seen in the difference spectrum of AC-C3C minus AC-C3B), at 4.876 ppm (one intensity), and at 4.762 ppm (one intensity).

We had noted an overall tendency of compounds containing a higher ratio of Man/GlcNAc to elute in Bio-Gel P2 and Dowex chromatography in a significantly greater volume than compounds of similar molecular weight but with more GlcNAc. We utilized this property to enrich for the Man_7-$GlcNAc_2$-Asn compound by using ^{1}H NMR spectroscopy as an assay. AC-C pooled narrowly (Figure 2, top) showed the presence of resonances consistent with minor amounts of Man_7-$GlcNAc_2$-Asn (Figure 6A). Fractions eluting in a greater volume (i.e., between AC-C and AC-D on Dowex) were pooled separately from the major peaks (referred to as AC-C^+). The spectrum of AC-C^+ was enriched in resonances at 5.337, 5.119, 4.876, and 4.762 ppm (Figure 6B). Similarly, when a broadly pooled AC-C from the Dowex column (Figure 2, bottom) was examined, these resonances were again quite evident (Figure 6D). This material was subjected to Bio-Gel P2 chromatography, and the trailing shoulder was pooled and rechromatographed twice. A fraction was obtained which was ~60% enriched in the resonances consistent with Man_7-$GlcNAc_2$-Asn (Figure 6C). The 40% contaminating structure present is the major glycopeptide of ovalbumin, AC-C3B. Interpretation of the spectrum of this mixture of compounds showed that the resonance at 5.337 ppm arises from the anomeric hydrogen of 2-O-substituted (with mannose) (1→3)-α-mannose. Because the resonance is uniquely in the Man_7 compound, we set its intensity at 0.6. The next resonance to high field integrates to ~1.0 intensity unit and arises from the anomeric hydrogens of 2-O-substituted (by mannose) (1→6)-α-mannose (0.6 unit) in the Man_7 compound and terminal (1→3)-α-mannose from AC-C3B (0.4 unit). The mixture of resonances occurring at ~5.07 ppm arises from the anomeric hydrogens from terminal (1→2)-α-mannose in the Man_7 compound (1.2 units), from internal (1→3)-α-mannose (0.4 unit) and terminal (1→3)-α-mannose (0.4 unit) both in AC-C3B, and from the anomeric hydrogen of the GlcNAc attached to asparagine (1.0 unit). The next doublet to high field arises from terminal and internal (1→6)-α-mannose in AC-C3B (0.8 unit) while the resonance at 4.876 ppm is from the anomeric hydrogen of internal (1→6)-α-mannose (0.6 unit). The (1→4)-β-mannose resonance at 4.762 ppm (0.6 unit) arises from the Man_7 compound and that at 4.745 ppm comes from AC-C3B (0.4 unit). The 8–9-Hz coupled doublet at ~4.62 ppm is from the chitobiose core GlcNAc found in both the Man_7 compound and AC-C3B (1.0 unit). The two 8–9-Hz coupled doublets to high field, at ~4.55 and 4.43 ppm, are from arm GlcNAc (0.4 unit) and bisecting GlcNAc (0.4 unit), respectively, in AC-C3B (Carver et al., 1981). The intensity of the bisecting GlcNAc is somewhat reduced because of the effects of water suppression. From this information and the analysis of Carver et al. (1981) we can deduce a structure of



In all, we have concluded that there are four different glycopeptides which can be partially separated from broadly pooled AC-C in Bio-Gel P-2 and identified by ^{1}H NMR. They are the major component, AC-C3B (see above), the major GN-GN(GN)-like component of AC-C1, the additional component to AC-C3B in AC-C2 and AC-C3A, and the Man_7-$GlcNAc_2$-Asn compound found in AC-C4 and AC-C3C.

Glycopeptides of AC-D. We performed a similar analysis of mixtures of glycopeptides purified from AC-D. The parent AC-D consisted mainly of the compound (AC-D3) (Carver et al., 1981)



However, this fraction showed evidence of contaminating structures, for example, 8–9-Hz coupled doublets due to arm

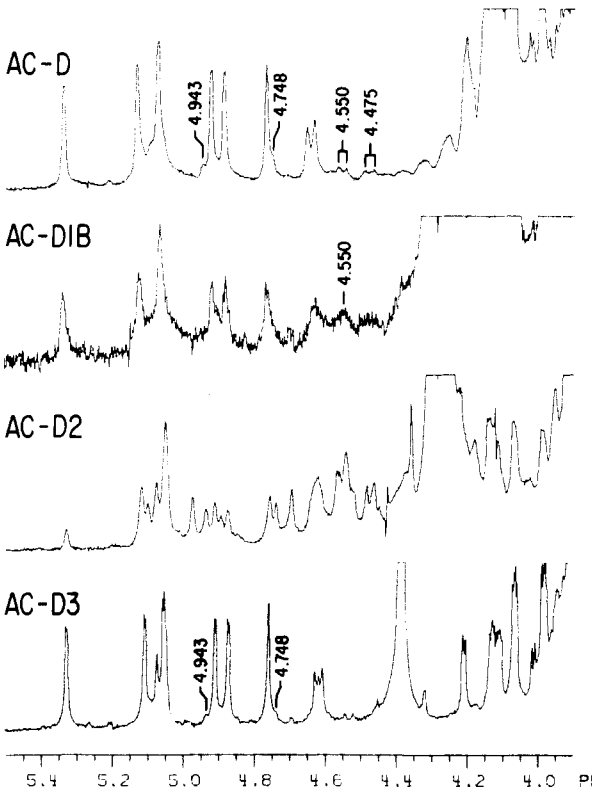


FIGURE 7: Spectra of Bio-Gel P2 subfractions of AC-D at 360 MHz. AC-D, 8.3 mg of neutral sugar in 0.5 mL of D_2O in a tube, was examined at 75 °C with 64 scans. AC-D1B, a subfraction from a different but similar parent AC-D, had 0.52 mg of neutral sugar. This was examined in a microcell in 0.1 mL of D_2O at 70 °C with 256 scans. AC-D2 came from a subsequent subfractionation run of AC-D. It had 13 mg of neutral sugar in 0.8 mL of D_2O and was examined at 70 °C in a tube with 64 scans. AC-D3 was similar in concentration and came from the same series as AC-D2 with four scans.

and bisecting GlcNAc at 4.475 and 4.550 ppm, respectively, and also shoulders at 4.748 and 4.943 ppm (Figure 7).

We subfractionated AC-D by Bio-Gel P2 chromatography as for AC-C in order to enrich and characterize the minor components (Figures 3 and 7). The spectrum of the major component of AC-D1, as in AC-C1, was enriched in 8–9-Hz doublet intensities at 4.55 ppm characteristic of arm β-linked *N*-acetylglucosamine residues (i.e., external to the chitobiose core), though the rest of its spectrum was similar to the major component of AC-C1. The NMR, but not GLC, analysis is consistent with a composition of Man_6-$GlcNAc_5$-Asn. This fractionation behavior and spectrum were reproducible in other fractionations of different starting material of AC-D. AC-D2 showed a mixture of resonances totally different from the major structure of AC-D. It was further fractionated by Bio-Gel P2 into AC-D2A, AC-D2B, AC-D2C, and AC-D2D pools where AC-D2A eluted in the least volume. AC-D2A yielded a spectrum very similar to that of AC-D1B (data not shown). AC-D2B has characteristic resonances at 4.979 and 4.712 ppm and was enriched in intensities at ~4.560 (Figure 8) arising from arm GlcNAc's. As with AC-C1, the presence of an α-mannose resonance to low field (4.979 ppm) and a β-mannose resonance to high field (4.719 ppm) of their respective ranges is GN-GN(GN)-like (Narasimhan et al., 1980). These workers have concluded that their equivalent to AC-D2B from ovalbumin was a compound of composition Man_3-$GlcNAc_5$-Asn (see citation for structure). AC-D2C is characterized (Figure 8) by a strong intensity of (1→6)-α-mannose at 4.949 ppm which is not in any other compounds of AC-C or AC-D. It must therefore belong to a third distinct structure in AC-D. A resonance at 4.767 ppm, a major intensity in AC-D2D, cannot be due to contamination by the latter because of the absence of a resonance at 5.339 ppm in AC-D2C (Figure 8) which is present in the spectrum of AC-D2D. Therefore, AC-D2C contains two compounds different from the major components in AC-D2A, AC-D2B, and AC-D2D. The latter component, containing a high mannose to glucosamine ratio, has a spectrum identical with that of AC-D3 (Figure 8; cf. Figure 7).

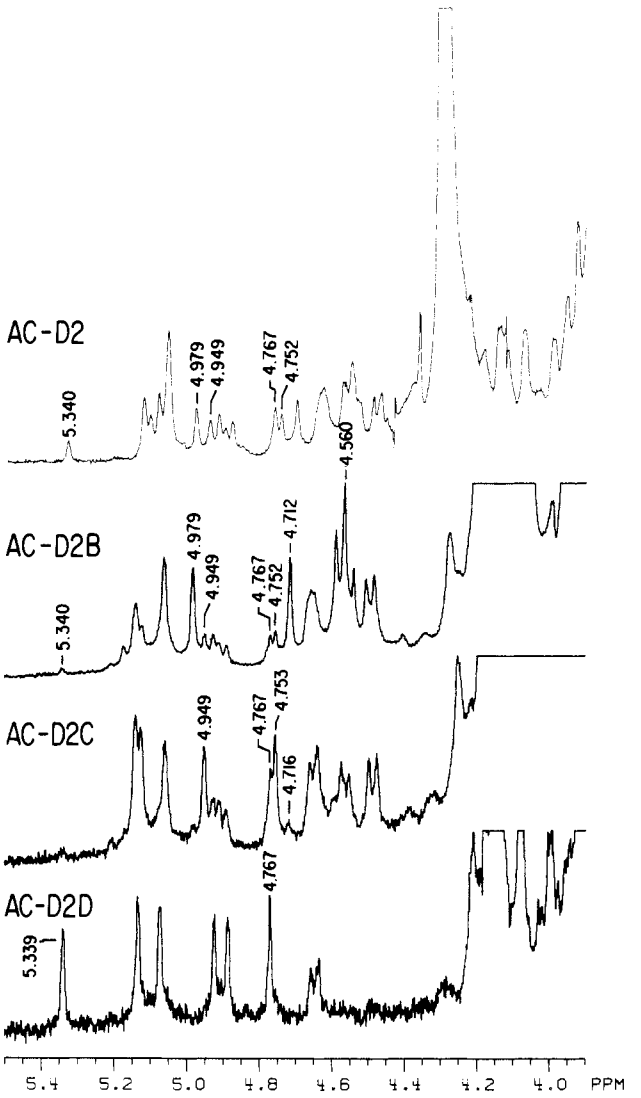


FIGURE 8: ^1H NMR spectra at 360 MHz of Bio-Gel P2 subfractions of AC-D2. AC-D2 was as described in Figure 7. Subfraction AC-D2A were not available in this series. AC-D2B contained 1.84 mg of neutral sugar. AC-D2C contained 0.95 mg of neutral sugar. AC-D2D contained 1.17 mg of neutral sugar. All were examined in 0.4–0.5 mL of D_2O in a tube with 256, 64, and 128 scans, respectively, at 70 °C.

Similar to our experience with AC-C, the mode of pooling from the Dowex column affected the relative amounts of different structures in AC-D. As is obvious from its spectrum, AC-D2B is enriched in GlcNAc. Like similarly enriched compounds in AC-C, it eluted in significantly less volume in Bio-Gel chromatography than other compounds of similar molecular weight but having a higher Man/GlcNAc ratio. Thus, by NMR analysis AC-D2B was found enriched in the Dowex fraction designated "C^+" pooled between AC-C and AC-D (Figure 2, top), more than in the later eluting AC-D.

In conclusion, we have evidence for five different structures in AC-D. These are major components of AC-D1, AC-D2B, two compounds in AC-D2C, and finally AC-D3, the major component of AC-D.

Discussion

The relative intensities of resonances in the high-resolution ^{1}H NMR spectrum reflect the number of protons in that chemical environment. It follows that for any homogeneous glycopeptide preparation the ratios of anomeric proton intensities must be integral. Conversely, in situations such as those shown in Figures 4–8 where many sets of resonances occur in nonintegral ratios of intensity, it can be concluded that the fraction contains a mixture of structures. When the mixture is too complex with inadequate dispersion of resonances, such as in AC-D2C, structures cannot be deduced and further purification is necessary. Classical methods of compositional (GLC) or linkage (methylation) analysis quantitate amounts of component saccharides after hydrolysis. As a result, direct information as to relative positions of components in original structure is lost and must be regained indirectly or by application of other techniques such as sequential glycosidase digestion, acetolysis, and periodate oxidation. Because ^{1}H NMR spectroscopy is nondestructive, this additional information is preserved in the spectra. Thus ^{1}H NMR spectroscopic analysis represents a valuable adjunct to classical analysis because some inference as to types and numbers of component structures in mixtures can be made. Further purification efforts can be appropriately monitored. Many of the structures referred to above occur as a very small fraction of the initial starting material (Table I), and the probability cannot be excluded that such glycopeptides arise from a small amount of contaminating glycoprotein. If it is so, it would modify the arguments above only in that the total number of structures present in ovalbumin would be reduced. However, sodium dodecyl sulfate ($NaDodSO_4$) gel analysis (Figure 1) of start-recrystallized ovalbumin would suggest the glycopeptides analyzed are from ovalbumin, because of the apparent homogeneity of the glycoprotein.

Three structures in AC-C have been described (Tai et al., 1977). These structures "III-A, III-B, and III-C" had a composition of Man_5-$GlcNAc_4$-Asn, Man_7-$GlcNAc_2$-Asn, and Man_4-$GlcNAc_5$-Asn and correspond to AC-C3B, the Man_7 compound, and likely the major component of AC-C2 and AC-C3A. There is a component in AC-C1 which does not correspond to a previously reported structure; we would predict it has a composition of Man_3-$GlcNAc_6$-Asn. These glycopeptides are being further purified so that an unambiguous assignment of structure can be made. Two compounds in AC-D have been described (Shepherd & Montgomery, 1978); namely, Man_4-$GlcNAc_4$-Asn and Man_6-$GlcNAc_2$-Asn. The latter corresponds to that of the major component of AC-D and to the structure reported by Tai et al. (1975). One of the components of AC-D2C appears to correspond to a structure with a composition Man_4-$GlcNAc_4$-Asn. However, there are three other structures in AC-D that have not been previously suspected, and their structures remain to be elucidated.

It is known that asparagine-linked oligosaccharides are processed or trimmed after initial protein glycosylation during translation (Hunt et al., 1978; Chen & Lennarz, 1978; Kornfeld et al., 1978; Hubbard & Robbins, 1979; Spiro et al., 1979). What relationship processing may bear to the eventual expression of glycoprotein oligosaccharides at the cell surface, especially of the oligomannosyl type (Muramatsu et al., 1976), is not known. The microheterogeneous complexity of the latter set of components seems no less than that evident in ovalbumin. Some such compounds serve as recognition structures in teratocarcinoma cells (Grabel et al., 1979), and it seems of importance to have methods for analyzing subtle differences in structure to understand such processes. The existence of independently derived mutants whose defects cause similar subtleties in glycosylation of cell surface molecules (Stanley & Sudo, 1981) and profound differences in glycosylation patterns of cotranslated viral glycoproteins (Hakimi & Atkinson, 1980) would suggest the presentation of microheterogeneous arrays is no accident. ^{1}H NMR analysis should be an invaluable adjunct to understanding the nature and regulation of glycosylation.

Acknowledgments

We gratefully acknowledge the use, by permission of Dr. Alan McLaughlin, of the 360-MHz facility at Brookhaven National Laboratories, under the auspices of the U.S. Department of Energy, and the use, by permission of Dr. John Markley, of the 360-MHz facility at Purdue University, under the auspices of the National Institutes of Health Division of Research Resources. We thank Joyce Tsang Lee and Andrew Marks for excellent technical assistance. We thank Drs. Cohen and Ballou for providing us with their manuscript prior to publication.

References

Albersheim, P., Nevins, D. J., English, P. D., & Karr, A. (1967) *Carbohydr. Res. 5*, 340.

Carver, J. P., Grey, A. A., Winnik, F. M., Hakimi, J., Ceccarini, C., & Atkinson, P. H. (1981) *Biochemistry* (in press).

Chen, W. W., & Lennarz, W. J. (1978) *J. Biol. Chem. 253*, 5780.

Cohen, R. E., & Ballou, C. E. (1980) *Biochemistry 19*, 4345.

Conchie, J., & Strachan, I. (1978) *Carbohydr. Res. 63*, 193.

Cunningham, L. W., Clouse, R. W., & Ford, J. D. (1963) *Biochim. Biophys. Acta 78*, 379.

Dorland, L., Haverkamp, J., Schut, B. L., Vliegenthart, J. F. G., Spik, G., Strecker, G., Fournet, B., & Montreuil, J. (1977a) *FEBS Lett. 77*, 15.

Dorland, L., Schut, B. L., Vliegenthart, J. F. G., Strecker, G., Fournet, B., Spik, G., & Montreuil, J. (1977b) *Eur. J. Biochem. 73*, 93.

Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J. C., Fournet, B., Spik, G., & Montreuil, J. (1978) *Eur. J. Biochem. 87*, 323.

Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Spik, G., Fournet, B., & Montreuil, J. (1979) *Eur. J. Biochem. 100*, 569.

Dubois, M., Gilles, K. A., Hamilton, J. D., Rebers, P. A., & Smith, F. (1956) *Anal. Chem. 28*, 350.

Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Binette, J. P., & Schmid, K. (1978) *Biochemistry 17*, 5206.

Gorin, P. A. J., & Spencer, J. F. T. (1968) *Can. J. Chem. 46*, 2299.

Gorin, P. A. J., Mazurek, M., & Spencer, J. F. T. (1968) *Can. J. Chem. 46*, 2305.

Gorin, P. A. J., Spencer, J. F. T., & Bhattacharjee, S. S. (1969) *Can. J. Chem. 27*, 1499.

Grabel, L. B., Rosen, S. D., & Martin, G. R. (1979) *Cell (Cambridge, Mass.) 17*, 477.

Grimes, W. J., & Greegor, S. (1976) *Cancer Res. 36*, 3905.

Hakimi, J., & Atkinson, P. H. (1980) *Biochemistry 19*, 5619.

Huang, C. C., Mayer, H. E., Jr., & Montgomery, R. (1970) *Carbohydr. Res. 13*, 127.

Hubbard, S. C., & Robbins, P. W. (1979) *J. Biol. Chem. 254*, 4568.

Hunt, L. A., Etchison, J. R., & Summers, D. F. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 754.

Johansen, P. G., Marshall, R. D., & Neuberger, A. (1969) *Biochem. J. 78*, 518.

Kekwick, R. A., & Cannan, R. A. (1936) *Biochem. J. 30*, 227.
Kornfeld, S., Li, E., & Tabas, I. (1978) *J. Biol. Chem. 253*, 7771.
Lee, Y. C., & Montgomery, R. (1962) *Arch. Biochem. Biophys. 97*, 9.
Levvy, G. A., Conchie, J., & Hay, A. J. (1966) *Biochim. Biophys. Acta 130*, 150.
Maizel, J. V., Jr. (1970) *Methods Virol. 5*, 179.
Muramatsu, T., Koide, N., Ceccarini, C., & Atkinson, P. H. (1976) *J. Biol. Chem. 251*, 4876.
Narasimhan, S., Harpaz, N., Longmore, G., Carver, J. P., Grey, A. A., & Schachter, H. (1980) *J. Biol. Chem. 255*, 4876.
Neuberger, A. (1938) *Biochem. J. 32*, 1435.
Nuenke, R. H., & Cunningham, L. W. (1961) *J. Biol. Chem. 236*, 2452.
Shepperd, V., & Montgomery, R. (1978) *Carbohydr. Res. 61*, 147.
Spiro, R. G., Spiro, J. J., & Bhoyroo, V. D. (1979) *J. Biol. Chem. 254*, 7659.
Stanley, P., & Sudo, T. (1981) *Cell (Cambridge, Mass.) 23*, 763.
Strecker, G., Herlant-Peers, M. C., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., & Farriaux, J. P. (1977) *Eur. J. Biochem. 81*, 165.
Strecker, G., Fournet, B., & Montreuil, J. (1978) *Biochimie 60*, 725.
Tai, T., Yamashita, K., Ogata, M., Arakawa, N., Koide, N., Muramatsu, T., Iwashita, S., Inoue, Y., & Kobata, A. (1975) *J. Biol. Chem. 250*, 8569.
Tai, T., Yamashita, K., Ito, S., & Kobata, A. (1977) *J. Biol. Chem. 252*, 6687.
Tarentino, A. L., & Maley, F. (1974) *J. Biol. Chem. 249*, 811.
van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Schmid, K., Montreuil, J., Fournet, B., & Hull, W. E. (1980) *Eur. J. Biochem. 114*, 11.
Wolfe, L., Senior, R. G., & Ng Ying Kin, N. M. K. (1974) *J. Biol. Chem. 249*, 1828.
Yamashita, K., Tachibana, Y., & Kobata, A. (1978) *J. Biol. Chem. 253*, 3862.

Comparative Study of Ribonucleotide, Deoxyribonucleotide, and Hybrid Oligonucleotide Helices by Nuclear Magnetic Resonance†

Arthur Pardi,* Francis H. Martin,‡ and Ignacio Tinoco, Jr.

ABSTRACT: The nonexchangeable base protons and the hydrogen-bonding NH–N imino protons were used to study the conformations and the helix–coil transitions in the following oligonucleotides: (I) $dCT_5G + dCA_5G$, (II) $rCU_5G + rCA_5G$, (III) $dCT_5G + rCA_5G$, (IV) $rCU_5G + dCA_5G$. The first three mixtures all form stable double-helical structures at 5 °C, whereas IV forms a triple strand with an rCU_5G:dCA_5G 2:1 ratio. The chemical shifts of the imino protons in the double strands indicate that I, II, and III have different conformations in solution. For example, the hydrogen-bonded proton of one of the C·G base pairs is more deshielded (a 0.4-ppm downfield shift) in helix I than in helix II or III. This implies a significant change in helical parameters, such as the winding angle, the distance between base pairs, or overlap of the bases. The coupling constants of the H1′ sugar protons show that helix I has 90% 2′-endo sugar conformation, whereas helix III has greater than 85% 3′-endo conformation for the observed sugar rings. The sugar pucker data are consistent with helix I having B-family geometry; III has A-family geometry. The chemical shifts of the nonexchangeable base protons in system I were followed with increasing temperature. The midpoints for the transitions, T_m's, for all the base protons were 28–30 °C; this indicates an all-or-none transition.

Proton nuclear magnetic resonance (NMR)[1] studies of oligonucleotides have greatly increased our understanding of the conformations and physical properties of nucleic acids in solution. The chemical shifts of the NH–N imino resonances, base proton resonances, and the H1′ ribose or deoxyribose resonances have been used, separately or in combination, to deduce RNA A or A′ (Arnott et al., 1975) type conformations for double-helical ribooligonucleotides (Arter et al., 1974; Heller et al., 1974; Borer et al., 1975; Hughes et al., 1978). Double-stranded deoxyribooligonucleotides have shown spectra consistent with a B-form geometry in solution (Cross & Crothers, 1971; Patel, 1974; Kallenbach et al., 1976; Early et al., 1977). These conclusions are based mainly upon comparison of the experimental chemical shifts with those computed from ring current effects for an assumed geometry, as well as the sugar pucker deduced from the H1′ proton coupling constants.

X-ray studies of RNA–DNA hybrid duplexes have shown them to adopt A-type geometries (Milman et al., 1967; Arnott et al., 1975). The only ^1H NMR study of an RNA–DNA duplex shows the structure of the hybrid to be different than that of the DNA–DNA duplex of the same sequence and consistent with an A form in solution (Selsing et al., 1978).

Triple-stranded structures are not uncommon in polynucleotide or oligonucleotide solutions (Bloomfield et al., 1974). Under conditions where triplexes were formed, Geerdes

† From the Department of Chemistry and Laboratory of Chemical Biodynamics, University of California, Berkeley, Berkeley, California 94720. *Received November 21, 1980.* This work was supported by National Institutes of Health Grant GM 10840 and by the Division of Biomedical and Environmental Research of the Department of Energy under Contract No. W-7405-ENG-48. We also thank the Stanford Magnetic Resonance Laboratory (supported by National Science Foundation Grant GP 26633 and National Institutes of Health Grant RR 00711) for the use of the HXS 360-MHz facilities.

‡ Present address: Chemistry Department, University of Colorado, Boulder, CO 80303.

[1] Abbreviations used: NMR, nuclear magnetic resonance; EDTA, ethylenediaminetetraacetate; oligo(A), oligo(riboadenylic acid); oligo(U), oligo(ribouridylic acid); TSP, sodium 3-(trimethylsilyl)propionate-*2,2,3,3-d*$_4$.

0006-2960/81/0420-3986$01.25/0