NMR Spectroscopy in the Study of Carbohydrates: Characterizing the Structural Complexity

WILLIAM A. BUBB

School of Molecular and Microbial Biosciences, University of Sydney, New South Wales 2006, Australia

ABSTRACT: The combination of structural diversity at several levels and limited chemical shift dispersion ensures that NMR spectra of carbohydrates are relatively difficult to interpret. This introduction to applications of NMR spectroscopy for the study of carbohydrates provides guidelines for interpretation of their 1- and 2-D spectra against a background of their tautomeric, configurational, and conformational equilibria in solution and consideration of their biosynthetic diversity. The influence of structural features on chemical shifts and coupling constants is illustrated by the consequences for both homo- and heteronuclear 2-D NMR spectra. Some applications of NMR spectroscopy for studies of carbohydrate metabolism are briefly considered. © 2003 Wiley Periodicals, Inc. Concepts Magn Reson Part A 19A: 1–19, 2003

KEY WORDS: carbohydrate structure; tautomeric equilibria; structural reporter group; 2-D NMR; carbohydrate metabolism; carbohydrate conformation

INTRODUCTION

The Carbohydrate Niche

On even cursory consideration, carbohydrates provide to biology: ribose derivatives, the glue that cements the genetic code; glucose, a major cellular fuel; cellulose a glucose polymer, the most abundant organic substance on this planet; oligosaccharides, key cell recognition antigens in the immune system; more complex carbohydrate polymers that have key roles in the pathogenicity of the microorganisms implicated in many human diseases. Yet, judged by recent citations in highly ranked journals, carbohydrates might be considered bit players in contemporary science. Thus, a search of journal abstracts and titles for 2001 yielded just 2 citations in *Nature* for the keyword "carbohydrate" compared with 276 citations for the keyword "protein"; for *Science*, there were 6 and 249 citations for the respective keywords. Given that more than half of all proteins in nature are predicted to be glycosylated (*1*, *2*) and that, contrary to previous belief, the carbohydrate residues are known to constitute more

Received 23 December 2002; revised 2 April 2003; accepted 2 April 2003

Correspondence to: W.A. Bubb; E-mail: b.bubb@mmb.usyd.edu.au

Concepts in Magnetic Resonance Part A, Vol. 19A(1) 1–19 (2003)

Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/cmr.a.10080

© 2003 Wiley Periodicals, Inc.

than "mere decoration" (*3*), this lack of attention is unlikely to persist.

The rich diversity of structures and biologic contexts in which carbohydrates are found have not been readily adaptable to the formula-driven approach that characterizes the recent major advances in genomics and proteomics. Compared with the 4 bases of DNA and 20 amino acids of proteins, for example, more than 100 different monosaccharides and approximately 50 nonsugar components have been identified in bacterial polysaccharides (*4*). And, for reasons outlined below, even a small number of monosaccharide components can lead to great structural diversity.

Carbohydrate NMR in Context

NMR spectroscopy provided most of the experimental data that enabled the complex equilibria of interconverting forms of reducing sugars to be unraveled (*5*, *6*). As the sensitivity of NMR methods has improved, in conjunction with the availability of ^{13}C-labeled compounds, even minor components of those equilibria have become observable (*7*), and the conformational space occupied by carbohydrates has become better defined (*8*). Conversely, when NMR experiments are used to determine the structures of carbohydrates it is necessary to be aware of all possible molecular permutations to interpret the spectra. Further, the absence of template-directed biosynthesis of saccharide chains often leads to complex mixtures that involve a variety of substitution patterns and chain lengths. The NMR spectra of carbohydrates must therefore be interpreted as a representation of a kaleidoscope of chemical and biologic diversity.

Both compilations of NMR data and experimental techniques have been recorded in a large number of reviews on applications of NMR spectroscopy to carbohydrate research; a comprehensive recent review (*9*) includes a list of 18 reviews published during the 1990s that are related to carbohydrates and NMR. Although an introduction to a subject must inevitably canvas its scope, it is not the goal of this article to provide a detailed summary of this extensive literature. Rather, the objective is to provide a framework for the interpretation of the 1- and 2-D spectra of naturally occurring carbohydrates, with emphasis on the outcomes from their unique structural features. Because of space limitations in an article such as this, the examples are illustrative rather than exhaustive. Carbohydrates include all polyhydroxy aldehydes or ketones, or their derivatives such as the polyols (*10*), but the general impact of their structural features on NMR spectra is evident from consideration of a relatively small number of representative molecules.

ORIGINS OF STRUCTURAL DIVERSITY

Interpretation of the NMR spectra of carbohydrates requires consideration of the numerous structural permutations that might give rise to resonances in their NMR spectra. To put this issue in perspective, compared with just 27 peptides that may be produced from any 3 amino acids the same number of hexoses can theoretically yield 38,016 different trisaccharides (*11*). Only the major structural issues are summarized here as extensive details are readily available in the IUPAC-IUBMB rules for the nomenclature of carbohydrates (*12*) (http://www.chem.qmw.ac.uk/iupac/2carb/).

Configuration: D or L

Monosaccharides are assigned to the D or L series according to the configuration (relative to that of D- or L-glyceraldehyde) of the highest numbered chiral center, which for D-glucose (**1**; Fig. 1) is C-5. Most naturally occurring monosaccharides have the D-configuration although there are notable exceptions (*12*). While an NMR experiment on the free saccharide in an achiral solvent cannot distinguish between enantiomers, NMR spectra of diastereomeric derivatives may enable the configuration to be determined (*13*).

Tautomerism

Monosaccharides are in general stabilized by cyclization to form a hemiacetal or hemiketal, but all six possible tautomers of the D-[1-^{13}C]aldohexoses in $^{2}H_2O$ solutions have been detected by ^{13}C NMR at 30°C (*7*). For D-glucose these tautomers and their reported relative concentrations are shown in Fig. 1. Here, it is evident that for most purposes only the pyranose tautomers, but both anomers (see below), need to be considered in the analysis of NMR spectra. For some substituted monosaccharides, a septanose form has also been observed (*5*, *6*). NMR studies of the tautomeric equilibria in $^{2}H_2O$ for both aldopentoses at 28°C (*14*) and aldohexoses at 30°C (*7*) revealed less than 0.1% acyclic anomer, except for D-idose, in which the open chain form comprised 0.8%.

Most NMR studies of carbohydrates involve aqueous solutions in which there is an equilibrium mixture of tautomers. Of the aldohexose and aldopentose

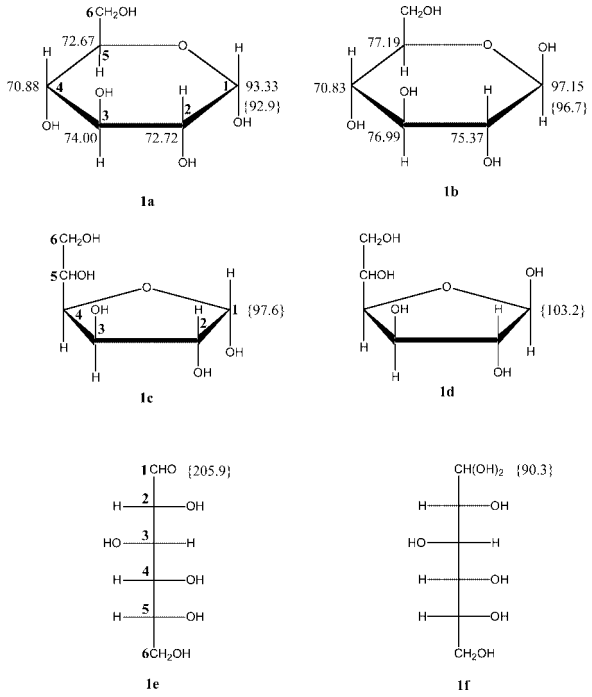
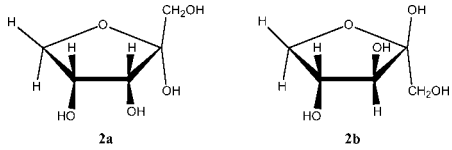


Figure 1 Tautomers of D-glucose; relative concentrations detected in ^{13}C NMR spectra at 150 MHz: α- (**1a**; 37.64%) and β-pyranose (**1b**; 61.96%), α- (**1c**; 0.11%) and β-furanose (**1d**; 0.28%), together with the open-chain aldehyde (**1e**; 0.004%) and its hydrate (**1f**; 0.0059%). Chemical shifts of C-1 for each tautomer relative to α-D-[1-^{13}C]-mannopyranose at δ 95.0 (*7*) are given in parentheses. The additional chemical shifts provided for **1a** and **1b** are from (*49*), and are adjusted to acetone ($^{13}CH_3$) at δ = 30.5.

hemiacetals, the pyranose tautomers predominate and their concentrations exceed 90% for all but D-ribose, D-altrose, D-idose, and D-talose, in which the concentrations of furanose tautomers vary between 21% and 32%. Notwithstanding the usually higher stability of six-membered rings over their five-membered counterparts, the greater tendency of secondary over primary hydroxyl groups to form intramolecular acetals enhances furanose formation for the aldopentoses because the pyranose tautomer must involve a primary hydroxyl group (*5*). Aqueous solutions of the tetroses do contain significant amounts of the free aldehyde and its hydrate, in addition to the α- and β-furanose tautomers (*15*); the hydrates are readily distinguished by the chemical shift of C-1, which for both erythrose and threose hydrates is approximately 9 ppm to lower frequency than the shifts of their furanose counterparts (*16*). For the aldopentoses (*14*) and aldohexoses (*7*) this difference is not as great but C-1 in hydrates (~δ 90–δ 91) is still more shielded than in cyclic tautomers.



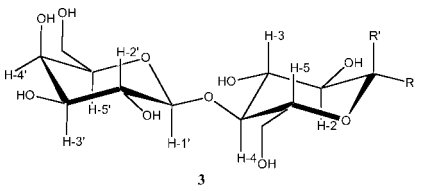
Structure 1

Configuration at the Anomeric Carbon: α or β

Cyclic hemiacetal formation generates a new center of asymmetry, at the anomeric carbon C-1. The configuration of the OH group that is attached to the anomeric carbon is defined by its relationship in a Fischer projection to the oxygen attached to the anomeric reference atom (in general the atom that defines the configuration D or L): α if the oxygen atoms are formally *cis* (as in **1a** and **1c**) and β if they are formally *trans* (as in **1b** and **1d**). For the pyranoses, the usual energetically preferred equatorial orientation of substituents in six-membered rings is opposed by interactions between lone-pair electrons of the endocyclic oxygen and electronegative substituents at C-1 that favor axial orientation of the latter—the so-called anomeric effect (*17*, *18*). Thus, the appreciable concentration of the α-pyranose tautomers of the aldohexoses, including α-D-glucose noted above.

In five-membered rings the steric constraints imposed by the substituents are the dominant influences on the proportion of each anomer (*5*). For example, of the ketopentoses (which can only cyclise as furanoses) a ^2H_2O solution of D-*erythro*-2-pentulose (D-ribulose) at 32°C contains predominantly the α-anomer [**2a** (Structure 1); 60.9%; δ(^{13}C-2), 104.0], with minor amounts of β-furanose [20.4%; δ(^{13}C-2), 107.0] and keto [18.7%; δ(^{13}C-2), 213.9] forms. On the other hand, a solution of D-*threo*-2-pentulose (D-xylulose) at 35°C contained predominantly β-anomer [**2b**; 62.3%; δ(^{13}C-2), 104.4], with minor amounts of α-furanose [17.5%; δ(^{13}C-2), 107.2] and ketone [20.2%; δ(^{13}C-2), 214.4]; that is, steric interactions are minimized by a *cis* arrangement of the 2- and 3-hydroxyl groups in each case (*19*). Note that this stereochemistry is also associated with the lower frequency chemical shift of C-2 in each pair of compounds (*20*).

As the α- and β-anomers are diastereomeric, they have, in principle and in general in practice, distinguishable NMR spectra. In fact, the proportions of anomers and, in particular, the influences of both solvent and temperature on this equilibrium are readily determined by NMR. Thus, the distribution of pyranose anomers for solutions of D-glucose at 35°C was found by ^1H NMR to be (α:β) 35:65 in ^2H_2O and 45.4:54.6 in dimethyl sulfoxide (DMSO)-d_6. For the



(**A**) R = H-1, R' = OH (**B**) R = OH, R' = H-1

Structure 2

^2H_2O solution, the proportion of α-anomer increases to 39% at 80°C (*21*).

Even at relatively low magnetic field strengths, resonances that are characteristic of each anomer are commonly observed for each proton or carbon of a reducing monosaccharide. For oligosaccharides at higher fields, the influence of the anomeric distribution can be manifest in the resonances of nonreducing residues. For example, the ^{13}C resonances for the galactose ring of lactose {β-D-Gal*p*-(1→4)-D-Glc; **3** (Structure 2)} are coincident for each anomer at 20 MHz (*22*), but at 150 MHz the ^{13}C spectra of α-lactose (**3a**) and β-lactose (**3b**) are fully resolved, with the exception of the galactose C-5 (Fig. 2).

In addition to improved signal dispersion at higher magnetic fields, it is important to recognize that the effect of the anomeric configuration on both ^1H and ^{13}C chemical shifts is variable. The practice of "counting carbons" must therefore be undertaken with caution when interpreting the NMR spectra of carbohydrates. To emphasize the origin of each pair of signals, the spectrum shown in Fig. 2 was obtained with full relaxation and no nuclear Overhauser effect (NOE) so that, linewidths being approximately equal, the intensity of each carbon resonance reflects the proportion of the relevant anomer. The relative intensities of different pairs of galactose resonances are much more variable in lactose spectra obtained with partial relaxation and NOE enhancement. Resolution-enhanced ^1H NMR 1-D total correlation spectroscopy (TOCSY) spectra of lactose at 500 MHz have been reported to

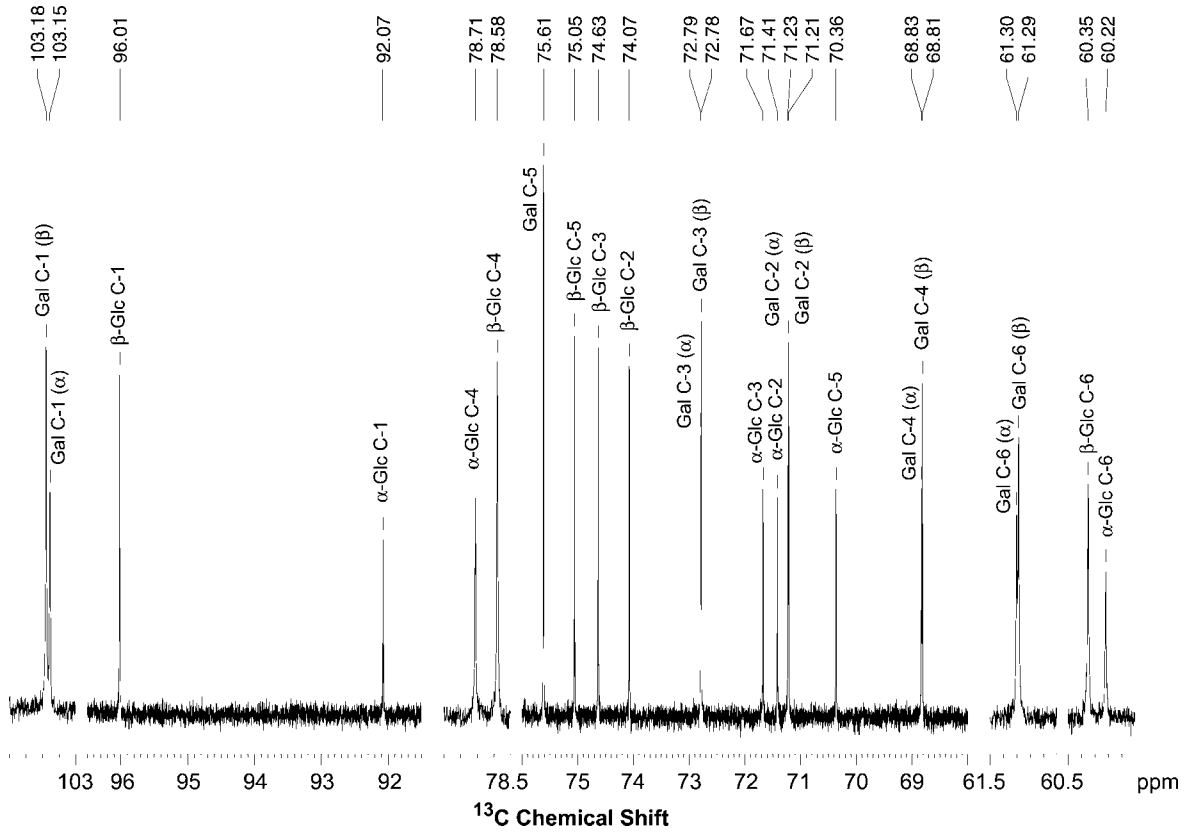
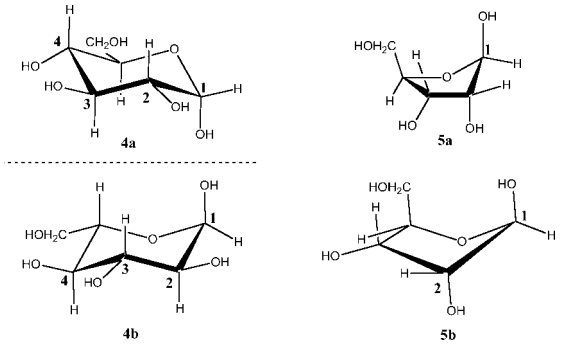
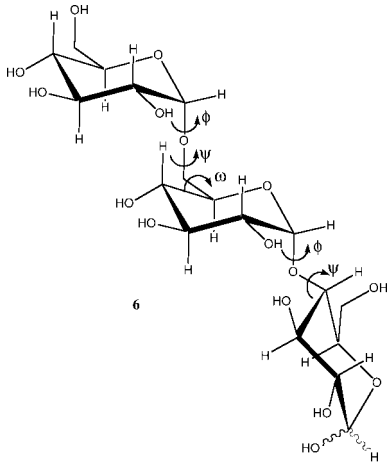


Figure 2 Fully relaxed (pulse sequence recycle time, 20 s) 150.9-MHz ^{13}C NMR spectrum of a 120-mM solution of lactose in ^2H_2O. WALTZ-16 decoupling was applied only during the 2-s acquisition time to eliminate differential NOEs. Chemical shifts (relative to acetone $^{13}CH_3$, δ = 30.5) are given above the assignments to confirm the resolution of resonances in the galactose ring and for comparison (see text) with those for **1a** and **1b**.



Structure 3



Structure 4

reveal similar effects, although only for the galactose H-2 and H-3 (*23*).

Monosaccharide Conformations

The conformations of pyranoses are essentially fixed on the NMR timescale (*2*, *24*) and the two stable chairs, denoted $^4C_1$ for the D-series and $^1C_4$ for the L-series, are in general dominant for naturally occurring monosaccharides (*12*). In this notation, the reference plane for the chair is defined by two parallel sides that exclude the lowest numbered carbon, and the reference side of that plane is that from which the ring numbering is clockwise. The leading superscript then defines the atom that lies above the plane and the trailing subscript the atom lying below the plane. This can be seen by comparison of the $^4C_1$ chair [**4a** (Structure 3)] of α-D-glucopyranose with the $^1C_4$ chair (**4b**) of α-L-glucopyranose (both numbered clockwise when viewed from above). In each case the more stable conformer has the greater number of equatorial hydroxyl groups.

Conformational minima of furanose monosaccharides are represented by the envelope (E) in which a single atom projects out of a plane formed by the remaining four, or the twist (T) in which a plane formed by three adjacent atoms is selected so that the remaining two atoms project either side of the plane. Again, a leading superscript identifies atoms above the reference plane and a trailing subscript atoms that are below this plane. Thus, the ^1E and $^1T_2$ conformations of β-D-ribofuranose are represented by **5a** and **5b**, respectively. Because the barriers to interconversion in furanose rings are considerably lower than those for pyranose rings, it is unlikely that a furanose sugar will have a single preferred conformation. Rather, the conformational equilibrium is described by a pseudorotation of alternating envelope and twist conformations, in which successive atoms adopt exoplanar positions (*25*). As such, the relationship between calculated conformations and the observed magnitudes of ^1H–^1H and ^{13}C–^1H couplings in furanose rings remains ill-defined (*26*).

Conformational Space of Oligo- and Polysaccharides

For oligo- and polysaccharides, glycosyl linkages entail additional conformational flexibility, defined by torsion angles ϕ and ψ as shown for the α1→4 linkage, or ϕ, ψ, and ω as in the α1→6 linkage, of panose {α-D-Glc*p*-(1→6)-α-D-Glc*p*-(1→4)-D-Glc; **6** (Structure 4)}. Although there are strict limitations on the precision with which these conformational variables can be defined, in particular for the more flexible 1→6 linkage, conformations in which protons attached to the carbons involved in the glycosidic linkage are spatially close, as depicted for the α1→4 link of panose, appear to be more heavily populated (*27*). This has been rationalised as a manifestation of the exoanomeric effect by which a substituent (aglycon) attached to a sugar (glycosyl) residue adopts a conformation that maximizes orbital overlap between lone-pair electrons of an exocyclic electronegative substituent (in general oxygen) and nonbonding electrons (σ*) associated with the bond between the anomeric carbon and endocyclic oxygen (*17*, *28*). Due to flexibility around the glycosidic linkage and an in general limited number of NMR-derived restraints, the conformations of oligo- and polysaccharides cannot be defined from NMR data alone (*2*, *29*).

Linkages

The multiplicity of possible linkages between monomeric units adds a further layer of structural complex-

→3)-α-D-Glc*p*-(1→2)-β-D-Gal*f*-(1→6)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→3)-β-D-Gal*f*-(1→

7a

→4)-α-L-Gal*p*NAcA-(1→3)-β-D-Qui*p*NAc4NAc-(1→2)-β-D-Qui*p*3NDAlaAc-(1→4)-α-D-Gal*p*NAcA-(1→4)-α-D-Gal*p*2,6Ac_2-(1→

7b

Structure 5

ity to carbohydrates because they may be linked between the anomeric hydroxyl group of one monosaccharide and any other hydroxyl-bearing carbon in another monosaccharide, yielding for aldohexoses a total of four possible linkages (excluding links between an anomeric hydroxyl and carbon that can only lead to the formation of disaccharides). By contrast with the two options for an amino acid to form a peptide bond that must contribute to a linear (or, rarely, cyclic) assembly, carbohydrates can and commonly do join together in extensively branched structures.

The determination of linkages is often the primary objective of an NMR study of a carbohydrate obtained from natural sources. Even with this apparently simple objective, the perennial problem of limited chemical shift dispersion may require supplementary data from enzymatic and chemical degradation and methylation analysis. In the latter, all free hydroxyl groups of the native saccharide are tagged with a methyl group (*30*) prior to hydrolysis and reductive acetylation; this yields a series of partially methylated alditol acetates that are identified by gas chromatography or gas chromatography–mass spectrometry. The nonmethylated positions in the characterized alditol acetates correspond to positions that were involved in linkages in the native saccharide. The number of possible linkages to be considered in the NMR study is thereby considerably reduced.

Structural Complications of Polysaccharide Biosynthesis

Apart from the spectral complications that arise from the structural permutations that are outlined above, the absence of template-directed synthesis ensures that many carbohydrates are intrinsically inhomogeneous. Whereas most heteropolysaccharides are arranged in regular repeating units, this is not normally the case for homopolysaccharides (*10*). A polysaccharide with a regular repeating unit yields one set of NMR resonances for each monosaccharide residue. The polysaccharide **7a** (Structure 5) (*31*), for example, displayed resonances for only three galactose and two glucose residues. In other words, the asymmetry implicit in the finite lengths of these molecules, whose size ranges from tens to hundreds of kilodaltons, is not evident in their NMR spectra. Other complications such as irregular substitution, including acetylation of hydroxyl groups, do arise, however, and in these cases NMR analysis may only yield an average or approximate structure.

The issue of structural diversity in polysaccharides is best illustrated by consideration of the repeating units of some bacterial polysaccharides. For example, structures with both furanose and pyranose tautomers of the same monosaccharide, as for the galactose residues in **7a**, are by no means unusual, and while the occurrence of the same monosaccharide with different configurations, such as the *N*-acetylgalacturonic acids of **7b** (*32*) is uncommon, it is clearly not impossible.

PRACTICAL NMR ISSUES

Solvents and Chemical Shift Reference Standards

Although NMR data referenced to tetramethylsilane have been reported for solutions of carbohydrates in deuteriochloroform, pyridine-d_5, and dimethyl sulphoxide-d_6, spectra are most commonly obtained for solutions in ^2H_2O; for the latter the sample is often exchanged and lyophiliszed several times to reduce the intensity of the signal from residual $^1HO^2H$, which may obscure those from the saccharide of interest. Such rigor remains important when preparing samples for 1-D ^1H NMR analysis but it is less critical for multidimensional spectra due to the excellent solvent suppression that is now achieved by coherence selec-

tion with pulsed-field gradients (*33*, *34*). Due to their high rate of exchange with solvent water, individual hydroxyl protons are only observable under conditions where that exchange is reduced, as for solutions in a dry organic solvent such as DMSO-d_6 (*35*) or in supercooled water (*36*).

For aqueous solutions, ^{1}H chemical shifts are commonly referenced to internal acetone at δ 2.225 and ^{13}C shifts to the same substance; however, literature values vary from δ 30.5 (most common) to δ 32.9. *p*-Dioxane (δ 67.4) and methanol (δ 49.6) are also used as references for ^{13}C (*37*) and external tetramethylsilane (TMS) and 4,4-dimethyl-4-silapentane-1-sulphonate (DSS) at δ 0.0 for ^{1}H spectra.

For samples that contain ionizable functional groups, such as phosphate and sulphate, chemical shifts in general vary with pH, which accordingly should be specified. The anomeric (^{1}H/^{13}C) resonances of α-D-glucose (δ 5.233/92.9) are often used as chemical shift references for biologic studies (*38*) in which it is in general undesirable to add extraneous substances. The chemical shift difference between the ^{1}H resonances for solvent water and the anomeric proton of α-D-glucose has also been reported as a reliable indicator of sample temperature (*39*).

Isotopes in Carbohydrate NMR

The ^{13}C chemical shifts of mono- and oligosaccharides have been shown to be exquisitely sensitive to the nature of the isotope of hydrogen in the hydroxyl group (*40*). Apart from the obvious outcome that carbons bearing substituted hydroxyl groups (with no possibility of ^{1}H/^{2}H exchange) are readily identified, the consistency of long-range effects of deuterium substitution allow other assignments to be made. In another approach, partially deuterated oligosaccharides in organic solvents yield spectra with multiplet patterns that are diagnostic of the molecular environment (*41*), but since the advent of 2-D NMR techniques these particular effects are rarely exploited.

More recently, ^{13}C- and ^{2}H-labeled monosaccharides have been used to establish the concentrations of minor tautomers (*7*) and provide additional structural restraints (*42*). As there is no single, in general applicable procedure for the preparation of isotopically labeled carbohydrates, specific chemical or enzymatic procedures are required for each problem (*43*). On the other hand, some (especially ^{13}C) isotopically labeled carbohydrates are commercially available and have been used extensively in biologic studies; some examples of these are given below.

Relaxation

Solutions of most monosaccharides yield ^{1}H- and ^{13}C-NMR spectra that are representative of relaxation in the extreme narrowing limit. The dominant mechanism for longitudinal relaxation of ^{1}H (*44*) and ^{13}C (*45*) is intramolecular dipole–dipole relaxation by all proximate protons and directly attached protons, respectively. Proton relaxation in pyranose tautomers is strongly influenced by the availability of common relaxation pathways for the chair conformation, namely, 1,3-diaxial, as between H-1′, H-3′ and H-5′ in **3**, and vicinal–gauche interactions, as between H-1 and H-2 in **3a**. Consequently, the relaxation rate of the axial anomeric proton of β-D-glucopyranose is almost twice that of the equatorial anomeric proton of α-D-glucopyranose. Similarly, in disaccharides composed of glucose residues the relaxation rates of anomeric protons in nonreducing residues are higher than those of protons with the same configuration at the reducing terminus. This occurs because the former receive relaxation contributions from protons in both rings (*44*).

Prior to the widespread availability of multidimensional NMR techniques, the structural dependence of ^{13}C relaxation rates proved a useful assignment tool for carbon atoms in oligosaccharides. For example, the ^{13}C resonances of the terminal galactose residue in stachyose {α-D-Gal*p*-(1→6)-α-D-Gal*p*-(1→6)-α-D-Glc*p*-(1→2)-α-D-Fru*f*} were distinguished from those of the internal galactose through their longer T_1 values (*46*). Although it is now uncommon for relaxation data, per se, to be obtained for the determination of the primary structures of carbohydrates, the effect of different relaxation rates on responses in multidimensional spectra may be diagnostically useful, as discussed below.

The rotational correlation times of many oligosaccharides are such that zero- and double-quantum contributions to dipole–dipole relaxation of the protons are approximately equal. This has important consequences for the determination of contributions to the NOE. For relatively low-molecular-weight polysaccharides, progressively slower molecular tumbling leads to broader NMR signals, but for higher molecular weights segmental motion dominates. Consequently, the interior carbons of the polysaccharides amylose, $\{\rightarrow 4)$-α-D-Glc*p*-$(1\rightarrow\}_n$, and linear dextran, $\{\rightarrow 6)$-α-D-Glc*p*-$(1\rightarrow\}_n$, have T_1, T_2, and NOE values that are within 10% of those for oligomers of their respective building blocks with just eight repeating units (*45*). Hence, NMR spectra with acceptable line widths can be recorded for large polysaccharides, albeit commonly at elevated temperatures that are typically 60–80°C.

STRUCTURAL INFORMATION FROM 1-D NMR SPECTRA

One-dimensional techniques that are based on highly predictable chemical shifts for specific molecular environments have been used extensively for the determination of carbohydrate structures. Assignments confirmed by various techniques, including NMR experiments on samples prepared with selective isotopic enrichment, selective proton decoupling experiments, and relaxation studies (*47*), provided the foundation for a number of empirical rules that remain especially useful for commencing the assignment of 2-D spectra and sometimes for the complete determination of the structure of an unknown substance.

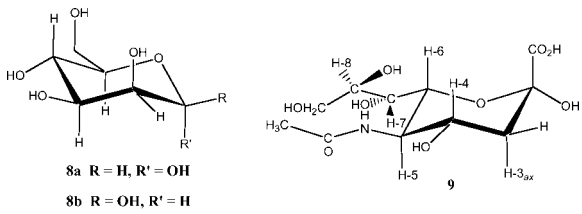
Basic ^{1}H NMR Data

Despite the fact that most resonances are clustered between ~δ 3.4 and ~δ 4.0 (*47–49*), ^{1}H spectra of carbohydrates do contain some well-resolved signals, including those of anomeric protons (δ 4.4–δ 5.5), acetyl (~δ 2.0– δ 2.1) and methyl (~δ 1.2) groups, and other protons that are influenced by specific functionality, including amino groups, phosphorylation, sulphation, glycosylation, and acetylation, or the lack of functionality as in deoxysugars.

Apart from this diagnostic chemical shift data, comparison of the integrated intensities of anomeric protons can reveal the number of monosaccharide residues. Splittings not anticipated from homonuclear coupling may result from proximity to the anomeric center or to another NMR active nucleus such as ^{31}P [e.g., $^3J_{P,H\text{-}1}$ = 7.5 Hz and $^4J_{P,H\text{-}2}$ = 1.8 Hz for α-D-glucose 1-phosphate at pH 8.0 (*50*)]. Note that, because of the limited dispersion of chemical shifts of nonanomeric protons, ^{1}H spectra of carbohydrates are often not first order, in which case line separations do not represent coupling constants. Of perhaps more significance for many contemporary NMR studies, the effects of strong coupling (*51*) may be evident in 2-D spectra. Finally, line widths may be indicative of the relative mobility of the local molecular environment.

Structural–Reporter Groups

Vliegenthart et al. (*52*) introduced the concept of "structural–reporter groups" to describe a procedure in which the chemical shifts, splitting patterns, and line widths of clearly resolved resonances in ^{1}H spectra can be used to assign the primary structures of complex oligosaccharide chains, principally those associated with glycoproteins. Thus, 72 *N*-linked oligosaccharides were characterized from the resonances of the following reporter groups (*53*): anomeric protons (H-1), mannose H-2 and H-3, sialic acid H-3,3′, fucose H-5 and CH_3 groups, galactose H-4 and H-5, and *N*-acetylglucosamine and *N*-acetylneuraminic acid CH_3. The method is based on monitoring the incremental chemical shifts associated with minor structural changes in an extensive library of compounds. Chemical shifts of the structural–reporter groups are remarkably sensitive to primary and secondary structural factors. As suggested above, even the chemical shifts of *N*-acetyl protons, despite their narrow overall range, are diagnostic of the chain position of the monosaccharide to which they are attached. Anomeric configurations are assigned from the magnitude of $J_{1,2}$, with values of 7–9 Hz for the diaxial coupling associated with a β-configuration and 2–4 Hz indicative of the equatorial–axial coupling of α-anomers. For D-mannose, which has an equatorial H-2, typical $J_{1,2}$ values are 1.6 Hz for coupling between diequatorial protons [α-anomer; **8a** (Structure 6)] and 0.8 Hz for the axial–equatorial coupling of the β-anomer (**8b**). This approach was the basis of the elucidation of the structures of highly branched oligosaccharides containing more than 20 monosaccharide residues.



Structure 6

The structural–reporter group approach has also been applied to other classes of carbohydrates that have been examined as extensions of basic building blocks, such as milk oligosaccharides derived from lactose. Thus, linkage of an α-*N*-acetylneuraminic acid residue (**9**) to the 3′ position (Gal ring) of lactose (**3**) is associated with a change in the chemical shift of H-3′ from δ 3.62 in (**3**) to δ 4.11 in the product of this transformation, α-*N*-acetylneuraminyl-(2→3)-lactose (3′-sialyllactose; α-Neu5Ac-(2→3)-β-D-Gal-(1→4)-D-Glc). The position of *O*-acetylation in a diacetyl derivative of 3′-sialyllactose was then determined as *O*-4″ by identification of resonances in the α-*N*-acetylneuraminic acid residue that were shifted to high frequency: H-4″ (Δδ 1.27 ppm), H-5″ (Δδ 0.26 ppm), and H-3*ax* (Δδ 0.13 p.p.m.) (*54*). As further illustration of the effect, discussed above, of the anomeric configuration on chemical shifts in oligosaccharides, distinct resonances for each anomer were observed for

H-1 and H-3 of the Gal residue and H-3*eq* of the α-*N*-acetylneuraminic acid residue of this diacetyl compound, 4-*O*-acetyl-α-*N*-acetylneuraminyl-(2→3)-lactose.

Assignment Methods Based on ^{13}C NMR Data

At a given magnetic field strength, ^{13}C resonances of saccharides are much more dispersed than their ^{1}H counterparts. Guidelines for assignment of the spectra of oligosaccharides (*47*, *55*, *56*), and in general applicable to ^{13}C data for polysaccharides, may be summarized as follows:

1. In the absence of structural modification, a sugar residue in an oligosaccharide has ^{13}C chemical shifts that are usually within ±0.3 ppm of the corresponding resonances in the free monosaccharide.
2. Glycosylation leads to high frequency shifts of 4–10 ppm for the carbons at the anomeric and linked positions. [Note, in this respect, that the anomeric carbon resonances of lactose (Fig. 2) embrace a chemical shift range that is comparable to that of all other carbons excluding C-6].
3. Resonances of carbon atoms that are adjacent to each carbon in the glycosyl linkage are in general shifted to slightly lower frequency.
4. The chemical shift of an anomeric carbon in an α-linkage (~δ 97–δ 101) is *in general* less than that of an anomeric carbon in a β-linkage (~δ 103–δ 105), but the exceptions to this rule recommend determination of $^{1}J_{C\text{-}1,H\text{-}1}$ as the most reliable indication of anomeric configuration.
5. ^{13}C nuclei in furanose sugars are less shielded than in their configurationally related pyranose counterparts.

The application of rules (*1*)–(*3*) can be verified by comparison of literature data (*49*) for the endocyclic carbons of D-glucose given for **1a** and **1b** (Fig. 1) with the chemical shifts for lactose given in Fig. 2. Notwithstanding in general lower chemical shifts for D-glucose, which may reflect differences in sample and measurement conditions, the appreciably larger changes at the glycosylated [rule (*2*)] and adjacent carbons [rule (*3*)] are readily apparent. Although extant 2-D NMR techniques can yield unequivocal assignments, these empirically derived rules still provide useful starting points and confirmatory checks of 2-D NMR data.

Computerized Assignments

The predictability of carbohydrate chemical shifts outlined above has encouraged the development of computer-based assignment procedures. For example, because the chemical shifts of the majority of monosaccharide residues within those oligosaccharides that contain an *N*-acetyllactosamine moiety depend only on the type of monosaccharide and immediate linkages, interrogation of a library of ^{1}H chemical shifts for structural–reporter groups provides reliable predictions of oligosaccharide structures (*57*). This approach has been extended to include oligosaccharides based on a number of other disaccharide cores (*58*).

The enhanced dispersion of ^{13}C NMR data allows all chemical shifts to be incorporated in data collections that, together with the information provided by methylation analysis, have been used to predict the structures of a limited number of regular polysaccharides (*59*, *60*). One of these programs, CASPER (Computer-Assisted SPectrum Evaluation of Regular polysaccharides) has recently been extended to provide predictions of multibranched structures, including those of glycopeptides (*61*). CASPER generates theoretical ^{13}C spectra from inputs of the constituent monosaccharides and their linkage positions and compares these spectra with experimental data to provide predictions of structure.

Artificial neural networks have been used for evaluation of the composition of binary mixtures of alditols (*62*) and determination of the structures of oligosaccharides (*63*). Appropriately trained neural networks obviate the need for scanning of a database to establish the presence of a particular substance (*64*) but the method has not been widely adopted. Similarly, data collections of carbohydrates have not been developed to nearly the same extent as those for proteins and nucleic acids. A summary of approaches to the establishment of carbohydrate data banks is provided with details of one of the most recent developments, SWEET-DB (http://www.dkfz-heidelberg.de/spec/sweetdb/) (*65*).

STRUCTURAL ASSIGNMENT FROM 2-D NMR SPECTRA

Because of limited dispersion of ^{1}H chemical shifts, 2-D NMR approaches to determination of all but the simplest carbohydrate structures in general require data from both homo- and heteronuclear experiments. For oligo- and polysaccharides, the task is to fully assign the subspectra of the constituent monosaccha-

rides, to establish the manner in which they are linked, and confirm the locations of any nonsugar substituents.

For geometrically restricted arrangements of spins such as those in pyranose and furanose monosaccharides, the NMR responses in 2-D experiments are acutely dependent on stereochemistry. Accordingly, any discussion of the use of 2-D NMR for the determination of the structures of carbohydrates must include consideration of the stereoelectronic dependence of coupling pathways. While this topic might be examined independently, viewing the effect of the stereochemical relationship between spins on the intensities of responses in 2-D spectra is a more practical framework for its consideration.

Homonuclear Through-Bond Correlations

Whereas anomeric protons in general provide the reference point for homonuclear correlations in a monosaccharide, other well-resolved resonances (see above) may also be used. In fact, for monosaccharides such as α-*N*-acetylneuraminic acid (**9**), which lacks an anomeric proton, or α- (**8a**) or β-mannose (**8b**), for which the small coupling constants limit TOCSY transfer with the anomeric proton, other reference points may be unavoidable. Note that here the issue is a reference point for the spin system and not correlations per se. For some carbohydrates, such as the polyols, there may be no distinct resonance to which homonuclear correlations can be referred, in which case heteronuclear data (*66*) are in general required for reliable structural assignments.

Vicinal proton–proton coupling constants in pyranose rings can be predicted by modifications to the Karplus equation that allow for substituent effects (*67*). However, estimation of $^{3}J_{H,H}$ in furanose rings is complicated by the problem of conformational flexibility that in general applies to five-membered rings (*16*). Small values of $^{3}J_{H,H}$ in monosaccharide rings impose a serious impediment to the standard procedure for making sequential assignments from TOCSY spectra with increasing mixing times. This is evident in the blocking (at H-4) of magnetization transfer from the anomeric proton in the galactose but not the α- or β-glucose residue in a TOCSY spectrum of lactose (Fig. 3). Magnetization that is transferred from H-3 to H-4 ($J_{3,4}$ = 3.3 Hz) is not further transferred to H-5 ($J_{4,5}$ = 1.6 Hz). It is some measure of the difficulty of evaluating ^{1}H spectra of even the simplest carbohydrates that the latter value has been determined only recently (*68*). Inefficient TOCSY transfer may also be encountered among acyclic resonances, as in the side-chains of neuraminic acid derivatives [$J_{H\text{-}6,H\text{-}7}$ = 1.68 Hz in 3′-sialyllactose; (*69*)]. Correlations whose intensities are reduced by small values of $^{3}J_{H,H}$, and even correlations that represent coupling over more than three bonds may often be observed in a gradient-enhanced correlation spectroscopy (g-COSY) experiment (*70*). Resonances that are not identified from through-bond correlation data must be assigned from through-space or heteronuclear correlation experiments.

Heteronuclear Through-Bond Correlations

Proton-detected ^{1}H/^{13}C correlation experiments provide the greater dispersion of ^{13}C chemical shifts at the much higher sensitivity of ^{1}H and are extensively used for the determination of carbohydrate structures. The utility of these experiments is well illustrated by the fact that, despite their structural similarity, each of the commonly occurring natural polyols has a unique heteronuclear single quantum coherence (HSQC) spectrum (*66*). An HSQC spectrum of lactose is shown in Fig. 4. For such a small molecule, it is evident that many of the correlations can be assigned directly from alignment of ^{1}H resonances with those obtained in a TOCSY spectrum (e.g., Fig. 3). When combined with the basic rules outlined above for assignment of resonances in 1-D spectra, together with consideration of the characteristic multiplet structures displayed by the cross-peaks, this approach enables tentative assignments to be established for many resonances in the HSQC spectra of complex saccharides. These assignments may then be confirmed, and gaps filled, from heteronuclear multiple bond correlation (HMBC) and through-space correlation experiments. An artefact due to strong homonuclear coupling between β-Gal H-2 and H-3 is identified in Fig. 4 by dotted arrows; a weaker one (unmarked) is due to homonuclear coupling between β-Gal H-5 and H-6,6′. These attributes of the sensitivity-enhanced version of the gradient HSQC experiment (*71*) may sometimes be used, as here, to confirm assignments, but may obviously also introduce unwanted complexity.

For most structural fragments in carbohydrates, correlations in HMBC experiments are strongly dependent on stereochemical factors; due to their importance for carbohydrate NMR, these are considered in some detail. Given the range of $^{n}J_{C,H}$ (∼1–∼6 Hz) that is commonly encountered in aldopyranosides (*72*), a delay for evolution of long-range couplings of ≥80 ms (= 1/2 × 6.25 Hz) might be expected to provide a suitable compromise for optimal magnetization transfer in the HMBC experiment; however, efficient transverse relaxation in larger molecules of-

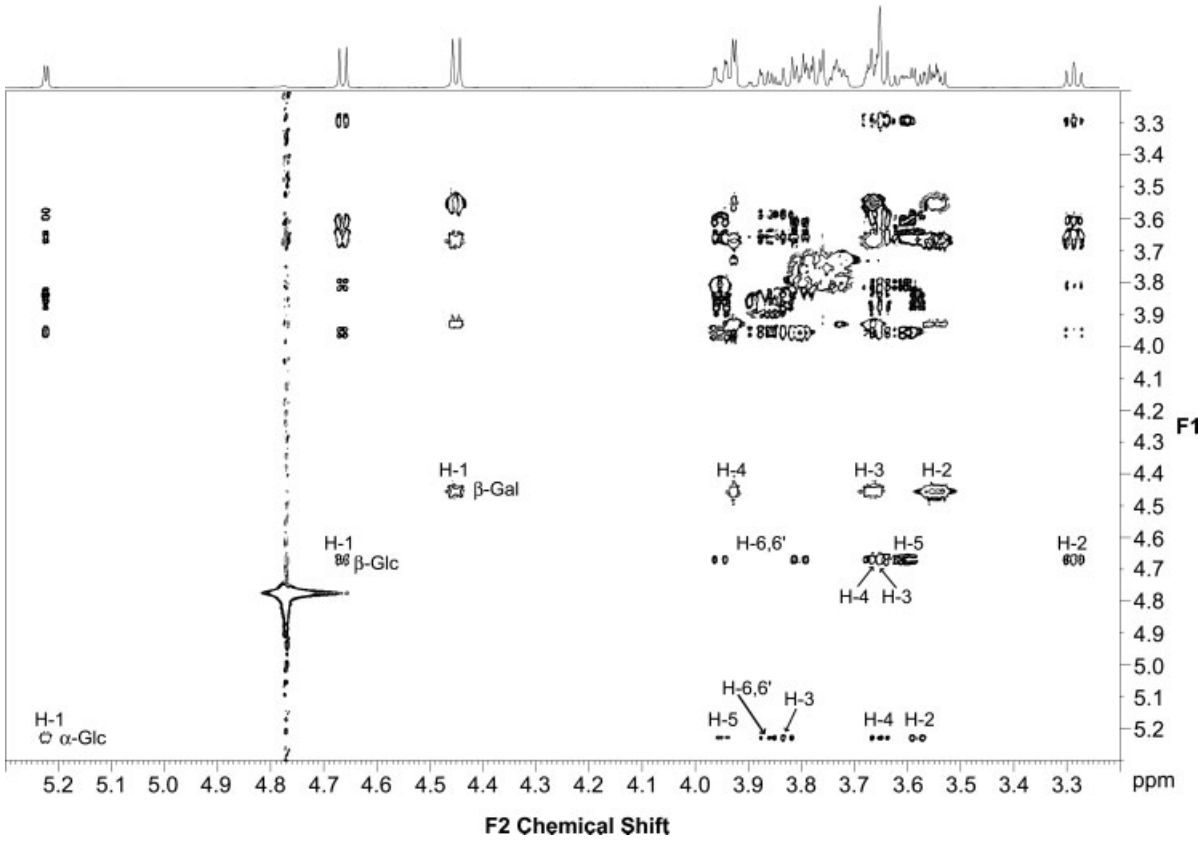


Figure 3 Gradient-enhanced TOCSY spectrum (mixing time, 120 ms; mixing pulse, 29 μs) at 600.1 MHz of a 120-mM solution of lactose in ^2H_2O. The spectrum was acquired for 512 t_1 increments of 8 scans and processed with a cosine window function in each dimension and one level of zero filling in F1. A 1-D ^1H spectrum is shown above the contour plot.

ten dictates the use of delays as short as 40 ms (*73*). While long-range correlations are invaluable for determining a number of assignments, they are most readily illustrated for the relatively uncrowded regions involving the anomeric protons [Fig. 5(a)] and carbons [Fig. 5(b)].

$^1J_{C,H}$. Because one-bond CH couplings in carbohydrates embrace a narrow range of values (~145 Hz for nonanomeric nuclei and 160–170 Hz for the anomeric positions) (*74*), the HSQC experiment in general provides satisfactory responses for the $1/2J_{C,H}$ set within this range. Commonly encountered functional groups such as the methyl groups in acetates or in certain deoxy sugars have ^{13}C–^1H couplings outside of this range but their HSQC responses are enhanced by the increased number of protons that contribute to the signal.

An HSQC experiment obtained without ^{13}C decoupling enables assignment of the anomeric configuration but this information may also be available from the HMBC experiment. Most variants of the HMBC experiment incorporate a filter that minimizes responses from one-bond couplings (*75*). Optimizing this filter for the crowded region of the spectrum that involves responses associated with the nonanomeric nuclei ($^1J_{CH} \approx 145$ Hz) provides a reasonable chance that one-bond couplings for anomeric resonances will be observable in the standard undecoupled HMBC experiment, as shown in Fig. 5(b). A value of ~170 Hz for $^1J_{C\text{-}1,H\text{-}1}$ indicates an equatorial proton at C-1 (e.g., α-D-glucose), while $^1J_{C\text{-}1,H\text{-}1}$ ~160 Hz indicates an axial proton (e.g., β-D-glucose); a number of mechanisms that rationalize this difference in terms of stereoelectronic factors have been advanced (*74*).

$^2J_{C,H}$. Values of $^2J_{C,H}$ vary from approximately −6 to +8 Hz and have been related to the orientation of specific electronegative substituents in the coupling pathway (*76–78*). In the more general treatment that includes all such substituents (*77*), the contribution of each oxygen-bearing substituent is assessed by viewing the coupling pathway from ^{13}C to ^{12}C in a Newman projection. The projections **10a** (Structure 7) and

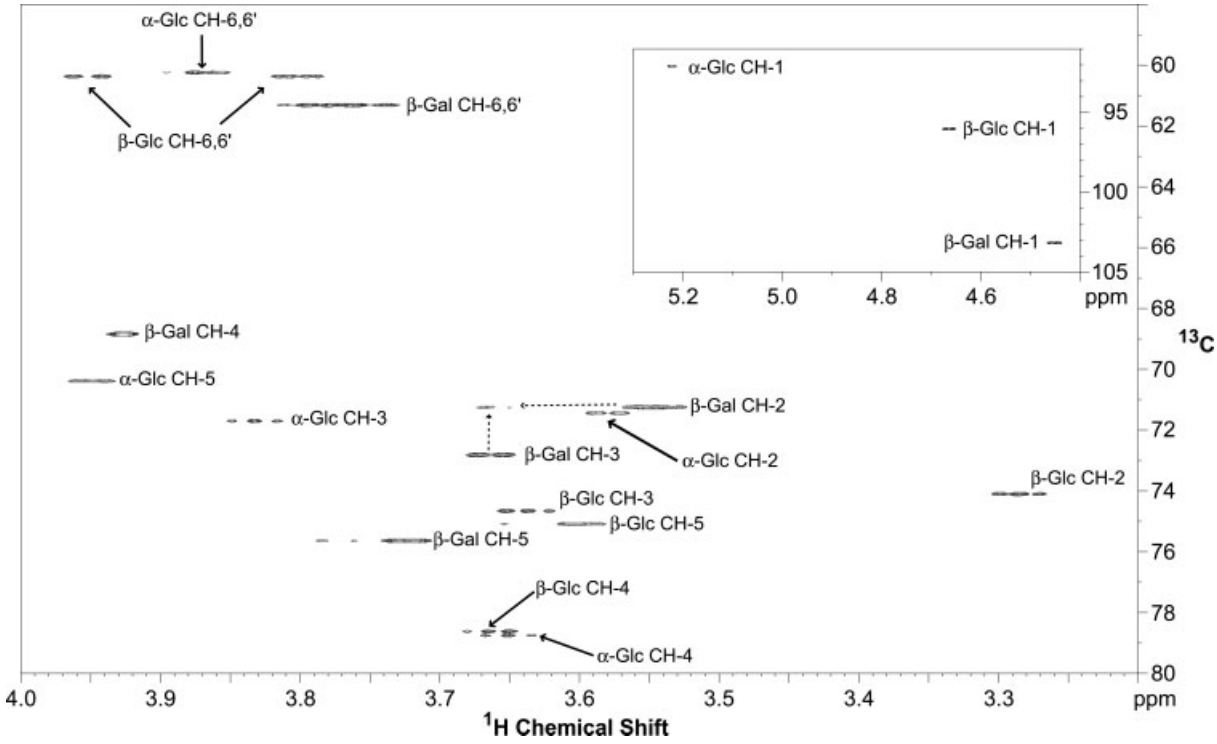
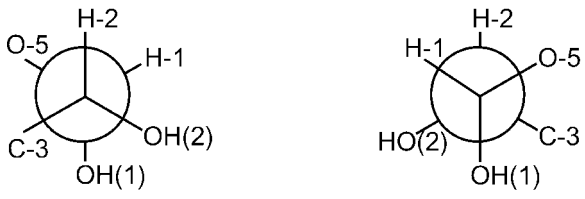


Figure 4 Gradient-selected, sensitivity-enhanced HSQC spectrum ($1/2^{1}J_{C,H}$, 3.45 ms; J = 145 Hz) at 600.1 MHz of a 120-mM solution of lactose in $^{2}H_{2}O$. The spectrum was acquired for 1024 t_1 increments of 4 scans and processed with a cosine window function in each dimension and one level of zero filling in F1.

10b can be used to evaluate the effect of these substituents on $^{2}J_{C\text{-}2,H\text{-}1}$ and $^{2}J_{C\text{-}1,H\text{-}2}$, respectively, for α-D-glucose (**4a**). Contributions to the projection sum are calculated relative to a value of +1 for an axis that is *trans* to the coupled proton. Thus, in **10a** the bonds to the anomeric hydroxyl group, (C-1)—(OH-1), and endocyclic oxygen, (C-1)—(O-5), each have values of +0.5 (cos 60°), while (C-2)—(OH-2) contributes −0.5 to give a projection sum of 0.5. For **10b** the contributions to $^{2}J_{C\text{-}1,H\text{-}2}$ are +0.5 for (C-2)—(OH-2), +1 for (C-1)—(OH-1), and −0.5 for (C-1)—(O-5) to give a projection sum of +1. The reader is left to determine that the projection sums for $^{2}J_{C\text{-}2,H\text{-}1}$ and $^{2}J_{C\text{-}1,H\text{-}2}$, for β-D-glucose are +0.5 and −0.5, respectively. A linear relationship that incorporates a projection sum of −0.5 associated with a large negative $^{2}J_{CH}$, a value of +0.5 associated with $^{2}J_{CH} \approx 0$, and a value of +2 linked to a large positive $^{2}J_{CH}$ was proposed (*77*). By interpolation, a projection sum of +1 implies a small positive $^{2}J_{CH}$.



10a **10b**

Structure 7

From these calculations, the only strong HMBC correlations representing two-bond couplings that involve the anomeric protons [Fig. 5(a)] or carbons [Fig. 5(b)] of lactose should be those between H-1 and C-2 for the β-D-Gal/Glc structural fragments. As Fig. 5(b) shows, these correlations are indeed strong while the remaining correlations are absent or weak except for the β-D-Gal H-1 to C-2 correlation [Fig. 5(a)], which appears to be anomalously intense. However, it has been suggested that factors other than the orientation of electronegative substituents on the coupling pathway contribute to $^{2}J_{C,H}$ because a range of values is observed for apparently identical coupling pathways (*72*). With regard to models for the correlation in question, $^{2}J_{C\text{-}2,H\text{-}1}$ is 0 Hz in methyl β-D-galactoside (*72*), but the equivalent coupling in the β-D-galactose residue of α-*N*-acetylneuraminyl-(2→6)-lactose (6′-sialyllactose; α-Neu5Ac-(2→6)-β-D-Gal-(1→4)- D-Glc) has a value of 2.2 ± 0.7 Hz (*79*).

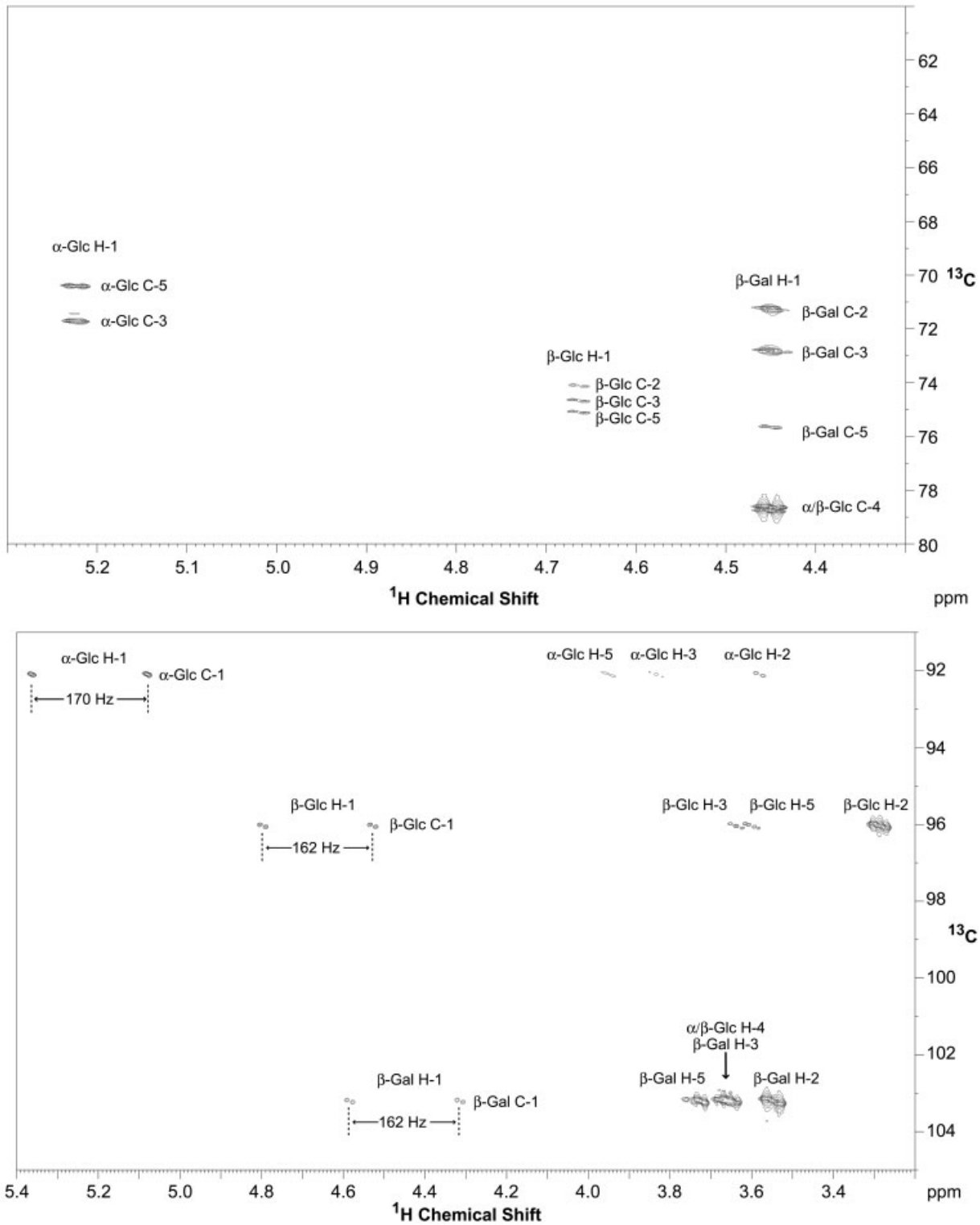


Figure 5 Gradient-selected HMBC spectrum (1/2($^{n}J_{C,H}$)), 80 ms; J = 6.25 Hz) at 600.1 MHz of a 120-mM solution of lactose in $^{2}H_{2}O$. The spectrum was acquired for 1024 t_1 increments of 8 scans and processed with a cosine window function in each dimension and one level of zero filling in F1, and is presented in magnitude mode. (a) Region showing correlations to anomeric protons. (b) Region showing correlations to anomeric carbons.

$^3J_{C,H}$. Values of $^3J_{C,H}$ in carbohydrate rings follow a Karplus-type dependence, with a maximum of ~6 Hz for a dihedral angle of 180° (*74*, *80*). Accordingly, $^3J_{C,H}$ involving C-3 or C-5 should be larger for an α- than a β-anomeric proton and the response in a typical HMBC experiment correspondingly more intense. For lactose this is reflected in the relative intensities of correlations for the α- and β-glucose residues [Fig. 5(a)] and H-5 to C-1 correlation of the β-galactose residue [Fig. 5(b)]. An intense correlation that might be assigned to the β-Gal H-3 to C-1 coupling is also observed in Fig. 5(b) but it cannot be unequivocally distinguished from the putative correlation across the glycosidic linkage (α-/β-Glc H-4 to β-Gal C-1, i.e., H-4 to C-1′). Even with the added dispersion of the ^{13}C dimension, these frustrating ambiguities are commonplace in the NMR spectra of carbohydrates.

In methyl aldopyranosides with the *gluco* or *galacto* configuration, $^3J_{C\text{-}1,H\text{-}3}$ and $^3J_{C\text{-}1,H\text{-}5}$ are ~1 Hz and 2–2.5 Hz, respectively (*72*), while in the β-D-galactose residue of 6′-sialyllactose $^3J_{C\text{-}3,H\text{-}1}$ and $^3J_{C\text{-}5,H\text{-}1}$ are, respectively, 2.5 ± 0.8 and 1.0 ± 0.2 Hz (*79*). By comparison, the normal *gauche* orientation of residues across the glycosidic linkage is associated with $^3J_{C,H}$ of ~4 Hz, which is in principle of sufficient magnitude to yield correlations in an HMBC experiment that confirm the structural identity of the linkage. For lactose, the *trans*-glycosidic couplings are 4.0 ± 0.1 and 5.1 ± 0.3 Hz for $^3J_{C\text{-}4,H\text{-}1'}$ and $^3J_{C\text{-}1',H\text{-}4}$, respectively (*81*). Consideration of all of these couplings suggests that $^3J_{C\text{-}1',H\text{-}4}$ should make the major contribution to the ambiguous correlation in Fig. 5(b) but the experimental evidence remains uncertain. However, an intense cross-peak associated with $^3J_{C\text{-}4,H\text{-}1'}$ (β-Gal H-1 to Glc C-4) confirms the position of the glycosidic linkage [Fig. 5(a)].

Notwithstanding the above focus on long-range correlations involving the anomeric proton or carbon, other HMBC correlations are often useful for completing the assignment of carbohydrate spectra. For the β-Gal residue of lactose, for example, strong correlations between H-6 and C-5, and between H-3/H-5 and C-4, complete the sequence of assignments around the ring. Note also that for pedagogical purposes the spectra shown in Fig. 5 were obtained for a concentrated solution at high resolution; with typical experimental conditions only the more intense cross-peaks might be observable in a realistic measurement time.

Combination Experiments. Provided sufficient sample is available, combination of heteronuclear correlation experiments with those that allow controlled transfer of ^1H magnetization, as in the HSQC-TOCSY or heteronuclear RELAY experiments, provides optimal dispersion of chemical shifts for assignment of the spectra of carbohydrates (*69*). However, samples derived from natural sources are rarely available in sufficient quantity for these techniques to be routinely applied.

Through-Space Correlations

Apart from the ambiguity, noted above, associated with conformational flexibility about the glycosidic linkage, a number of other caveats apply to the exploitation of the NOE in the study of carbohydrates. From a practical experimental perspective, the most important consideration is that correlation times for interacting spins in many small oligosaccharides correspond to the crossover between the positive and negative NOE regimes so that enhancements are either absent or significantly attenuated (*27*). This problem can be circumvented by rotating frame experiments [rotating frame Overhauser effect spectroscopy (ROESY)]. It has been suggested that ROESY experiments are in general more productive for tri- to decasaccharides, with mixing times reduced from 300 to 100 ms as the molecular size increases (*82*). The ROESY experiment may also be useful for identification of cross-peaks (through their changed sign) due to three-spin effects that are common in oligosaccharides (*24*). Because of the dominance of local motion in longer saccharide chains, both NOESY and ROESY experiments provide useful data. Thus, whereas NOESY experiments yield a number of enhancements [Fig. 6(a)] for the (relatively constrained) backbone residues of the heteropolysaccharide repeating unit [**11** (Structure 8)], only ROESY enhancements are observed [Fig. 6(b)] for the much more mobile terminal β-D-Gal residue (**G**) that is both at the end of a side-chain and attached through a more flexible linkage (C-6) (*73*); the boxed antiphase correlation in Fig. 6(a) is probably due to COSY-type transfer between H-1 and H-2 of residue **G** (*51*). Other features to note are the limited number of correlations in both spectra, even when employing relatively long mixing times that may give rise to multistep transfers (*51*, *73*), and the paucity of long-range (i.e., between nonadjacent residues) correlations. Note also that to obtain adequate resolution the data depicted in Fig. 6 were acquired at 70°C, yet there is still a degree of signal overlap that prevents unambiguous assignment of all correlations.

Especially for the axially oriented anomeric protons of β-pyranosides, 1-3 diaxial correlations with H-3 and H-5 (*24*) may fill gaps in assignments where TOCSY transfer is attenuated. Note that these corre-

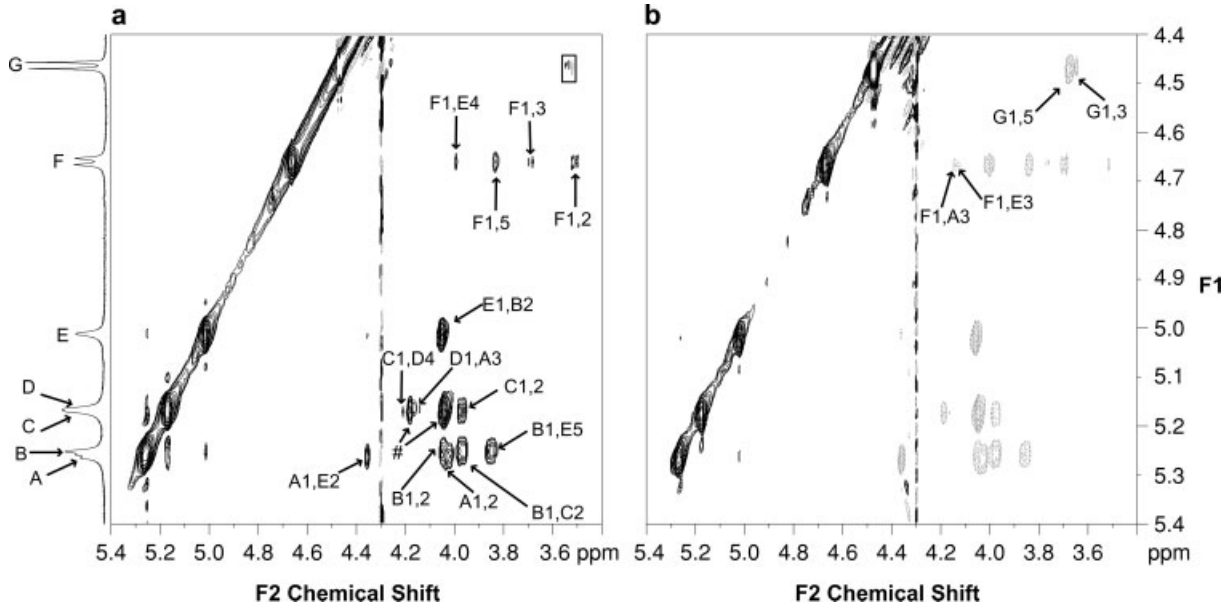


Figure 6 Parts of 600.1-MHz spectra of the polysaccharide (**11**) in $^{2}H_2O$ at 70°C, selected to depict correlations to anomeric protons. (a) NOESY spectrum acquired for 500 t_1 increments of 64 scans and a mixing time of 100 ms. (b) ROESY spectrum acquired for 512 t_1 increments of 48 scans and a mixing time of 140 ms. A 1-D ^{1}H spectrum is shown to the left of the NOESY contour plot. Both 2-D spectra were processed with a cosine window function in each dimension and one level of zero filling in F1. The dark and light (broken lines) contours represent positively and negatively phased peaks, respectively, and the intensity scaling of each contour plot was adjusted to yield comparable levels of thermal noise. Correlations are identified by the lettering scheme depicted in structure (**11**), intraresidue cross-peaks with a single letter (e.g., A1,2), and those between protons on different residues with the relevant residue letters (e.g., B1,E5); correlations observed in both spectra are only labeled in (a). The correlations associated with the anomeric protons of residues C/D marked # could not be unambiguously assigned (*73*); the box delineates an antiphase signal.

lations thus complement the HMBC correlations to C-3 and C-5 that should in general be more intense for the equatorially oriented anomeric protons of α-pyranosides. Due to conformational averaging, the in general limited number of observable short-range NOEs [at least one but not more than three (*8*)], and the almost universal absence of supporting long-range correlations, *trans*-glycosidic NOEs should be interpreted with caution (*2*, *24*). Although heteronuclear NOEs can in principle supplement homonuclear data, they are only applicable to isotopically enriched materials (*83*).

New Developments for Defining Secondary Structure

Because unlabeled samples in isotropic solution do not yield sufficient data to define the 3-D structures of oligosaccharides, additional restraints have been sought. Samples with ^{13}C enrichment enable the determination of additional scalar coupling restraints (*84*) and provide access to heteronuclear NOEs (*83*), while measurements of residual dipolar couplings in anisotropic media (*85–87*) afford long-range restraints that are not available from the measurement of

β-D-Gal*p*-(1→6)-β-D-Gal*p*-(1→4)
G **F** |
→2)-α-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→3)-α-D-Gal*p*-(1→3)-α-L-Rha*p*-(1→2)-α-L-Rha*p*-(1→
C **D** **A** **E** **B**

11

Structure 8

NOEs or coupling constants. The virtues and limitations of these techniques have been well summarized in a recent review (*8*). Because the carbohydrate-specific issues have more to do with synthesis and sample preparation than NMR spectroscopy per se, they are not further discussed here.

CARBOHYDRATE METABOLISM

The central role of glucose as a cellular fuel from archae bacteria to mammals, and the commercial availability of a number of ^{13}C-enriched isotopomers, including D-[U-^{13}C]glucose, form the basis of numerous applications of NMR spectroscopy to the study of metabolic pathways. Space limitations preclude their detailed consideration here but the following examples provide an indication of applications of 1- and 2-D techniques. With appropriate calibration to compensate for the effects of differential relaxation and NOEs, time courses of 1-D ^{13}C NMR spectra enable the monitoring of the fluxes of key intermediates in the glycolytic pathway (*88*). When a mixture of D-[U-^{13}C]glucose and its counterpart containing carbon at natural isotopic abundance are used as a substrate, the structures of cross-peaks in a ^{13}C COSY experiment reveal details of the integrity of carbon backbones in derived metabolites (*89*, *90*). Each cross-peak provides information about a three-carbon unit—the actively (antiphase) coupled pair and the extent of ^{13}C labeling on an adjacent carbon through the presence of passive (in-phase) coupling. Because the multiplets represent a superposition of responses from all isotopomers, the fragments that they represent and their metabolic origin can in principle be deduced.

CONCLUSIONS

The perspective of the relatively low scientific popularity of carbohydrates that was outlined in the Introduction was also presented in a recent series of articles entitled *Carbohydrates and Glycobiology*, where the authors comment that, "The chemistry and biology of carbohydrates has been a Cinderella field: an area that involves much work but, alas, does not get to show off at the ball with her cousins, the genomes and proteins" (*91*). A renaissance in carbohydrate research is being fostered by a combination of new synthetic methods and developments in analytic techniques. Because of its capacity to provide intimate structural details of biomolecular interactions, NMR spectroscopy is certain to play a prominent role in the carbohydrate "revival."

An informed knowledge of the complexities of the chemical and conformational equilibria, and biosynthetic diversity, of carbohydrates is required for reliable interpretation of their NMR spectra. These issues, together with the limited dispersion of chemical shifts, raise problems that are not encountered in the use of NMR to study other classes of molecules. Many of the challenges that arise can be addressed by recently developed heteronuclear NMR techniques whose sensitivity is sufficient to provide comprehensive data for the limited quantities of natural carbohydrates that are usually available. The rapid spread of high-field spectrometers with high-sensitivity cryoprobes will further encourage applications of NMR spectroscopy to the study of these ubiquitous molecules.

ACKNOWLEDGMENTS

The author thanks Prof. Philip Kuchel of the School of Molecular and Microbial Biosciences at the University of Sydney for many helpful comments and criticisms of this article. The author also thanks him and Prof. Tadasu Urashima of Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Japan, for long-term collaborations that have stimulated and sustained a fascination with applications of NMR to the study of the structure and metabolism of carbohydrates.

REFERENCES

1. Apweiler R, Hermjakob H, Sharon N. 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim Biophys Acta 1473:4–8.
2. Wormald MR, Petrescu AJ, Pao Y-L, Glithero A, Elliott T, Dwek RA. 2002. Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. Chem Rev 102:371–386.
3. Schmidt K. 2002. Sugar rush. New Scientist 176:34–38.
4. Lindberg B. 1998. Bacterial polysaccharides: Components. In: Dumitriu S, ed. Polysaccharides. New York: Marcel Dekker; p 237–273.
5. Angyal SJ. 1984. The composition of reducing sugars in solution. Adv Carbohydr Chem Biochem 42:15–68.
6. Angyal SJ. 1992. The composition of reducing sugars in solution: Current aspects. Adv Carbohydr Chem Biochem 49:19–35.
7. Zhu Y, Zajicek J, Serianni AS. 2001. Acyclic forms of [1-^{13}C]aldohexoses in aqueous solution: Quantitation

by ^{13}C NMR and deuterium isotope effects on tautomeric equilibria. J Org Chem 66:6244–6251.
8. Homans SW. 1998. New tricks with isotopically enriched oligosaccharides. Biochem Soc Trans 26:551–560.
9. Duus J, Gotfredsen CH, Bock K. 2000. Carbohydrate structural determination by NMR spectroscopy: Modern methods and limitations. Chem Rev 100:4589–4614.
10. Robyt JF. 1998. Essentials of Carbohydrate Chemistry. New York: Springer-Verlag; 399 p.
11. Laine RA. 1997. Information capacity of the carbohydrate code. Pure Appl Chem 69:1867–1873.
12. McNaught AD. 1997. Nomenclature of carbohydrates (recommendations 1996). Carbohydr Res 297:1–92.
13. York WS, Hantus S, Albersheim P, Darvill AG. 1997. Determination of the absolute configuration of monosaccharides by ^{1}H NMR spectroscopy of their per-*O*-(*S*)-2-methylbutyrate derivatives. Carbohydr Res 300:199–206.
14. Drew KN, Zajicek J, Bondo G, Bose B, Serianni AS. 1998. ^{13}C-labeled aldopentoses: Detection and quantitation of cyclic and acyclic forms by heteronuclear 1D and 2D NMR spectroscopy. Carbohydr Res 307:199–209.
15. Serianni AS, Pierce J, Huang S-G, Barker R. 1982. Anomerization of furanose sugars: Kinetics of ring-opening reactions by ^{1}H and ^{13}C saturation-transfer NMR spectroscopy. J Am Chem Soc 104:4037–4044.
16. Serianni AS, Barker R. 1984. [^{13}C]-Enriched tetroses and tetrofuranosides: An evaluation of the relationship between NMR parameters and furanosyl ring conformation. J Org Chem 49:3292–3300.
17. Lemieux RU, Koto S. 1974. The conformational properties of glycosidic linkages. Tetrahedron 30:1933–1944.
18. Juaristi E, Cuevas G. 1992. Recent studies on the anomeric effect. Tetrahedron 48:5019–5087.
19. Wu J, Serianni AS, Vuorinen T. 1990. Furanose ring anomerization: Kinetic and thermodynamic studies of the D-2-pentuloses by ^{13}C-n.m.r. spectroscopy. Carbohydr Res 206:1–12.
20. Vuorinen T, Serianni AS. 1991. Synthesis of D-*erythro*-2-pentulose and D-*threo*-2-pentulose and analysis of the ^{13}C- and ^{1}H-n.m.r. spectra of the 1-^{13}C- and 2-^{13}C-substituted sugars. Carbohydr Res 209:13–31.
21. Franks F, Lillford PJ, Robinson G. 1989. Isomeric equilibria of monosaccharides in solution: Influence of solvent and temperature. J Chem Soc, Faraday Trans 1 85:2417–2426.
22. Berry JM, Hall LD, Wong KF. 1977. Concerning the tumbling motion of disaccharides in aqueous solution. Carbohydr Res 56:C16–C20.
23. Platzer N, Davoust D, Lhermitte M, Bauvy C, Meyer DM, Derappe C. 1989. Structural analysis of five lactose-containing oligosaccharides by improved, high-resolution, two-dimensional ^{1}H-N.M.R. spectroscopy. Carbohydr Res 191:191–207.
24. Homans SW. 1990. Oligosaccharide conformations: Application of NMR and energy calculations. Prog Nucl Magn Reson Spectrosc 22:55–81.
25. Altona C, Sundaralingam M. 1972. Conformational analysis of the sugar ring in nucleosides and nucleotides. A new description using the concept of pseudorotation. J Am Chem Soc 94:8205–8212.
26. Podlasek CA, Stripe WA, Carmichael I, Shang M, Basu B, Serianni AS. 1996. ^{13}C-^{1}H spin-coupling constants in the β-D-ribofuranosyl ring: Effect of ring conformation on coupling magnitudes. J Am Chem Soc 118:1413–1425.
27. Neuhaus D, Williamson MP. 2000. The Nuclear Overhauser Effect in Structural and Conformational Analysis. 2nd ed. New York: Wiley-VCH; 619 p.
28. Batchelor RJ, Green DF, Johnston BD, Brian PO, Pinto BM. 2001. Conformational preferences in glycosylamines. Implications for the exo-anomeric effect. Carbohydr Res 330:421–426.
29. Rutherford TJ, Partridge J, Weller CT, Homans SW. 1993. Characterization of the extent of internal motions in oligosaccharides. Biochemistry 32:12715–12724.
30. Hakomori S. 1964. A rapid permethylation of glycolipid and polysaccharide, catalyzed by methylsulfinyl carbanion in dimethyl sulfoxide. J Biochem (Tokyo) 55:205–208.
31. Vanhaverbeke C, Bosso C, Colin-Morel P, Gey C, Gamar-Nourani L, Blondeau K, Simonet J-M, Heyraud A. 1998. Structure of an extracellular polysaccharide produced by *Lactobacillus rhamnosus* strain C83. Carbohydr Res 314:211–220.
32. Hanniffy OM, Shashkov AS, Senchenkova SN, Tomshich SV, Komandrova NA, Romanenko LA, Knirel YA, Savage AV. 1999. Structure of an acidic O-specific polysaccharide of *Pseudoalteromonas haloplanktis* type strain ATCC 14393 containing 2-acetamido-2-deoxy-D- and -L-galacturonic acids and 3-(*N*-acetyl-D-alanyl)amino-3,6-dideoxy-D-glucose. Carbohydr Res 321:132–138.
33. Schleucher J, Schwendinger M, Sattler M, Schmidt P, Schedletzky O, Glaser SJ, Sorensen OW, Griesinger C. 1994. A general enhancement scheme in heteronuclear multidimensional NMR employing pulsed field gradients. J Biomol NMR 4:301–306.
34. van Zijl PCM, Johnson MO, Mori S, Hurd REH. 1995. Magic-angle-gradient double-quantum-filtered COSY. J Magn Reson A 113:265–270.
35. Gillet B, Nicole D, Delpuech J-J, Gross B. 1981. High field nuclear magnetic resonance spectra of hydroxyl protons of aldoses and ketoses. Org Magn Reson 17:28–36.
36. Poppe L, van Halbeek H. 1994. NMR spectroscopy of hydroxyl protons in supercooled carbohydrates. Nat Struct Biol 1:215–216.
37. Bock K, Pedersen C. 1983. Carbon-13 nuclear magnetic resonance spectroscopy of monosaccharides. Adv Carbohydr Chem Biochem 41:27–66.
38. Nicholson JK, Foxall PJD, Spraul M, Farrant RD, Lin-

don JC. 1995. 750 MHz ^{1}H and ^{1}H-^{13}C NMR spectroscopy of human blood plasma. Anal Chem 67:793–811.
39. Farrant RD, Lindon JC, Nicholson JK. 1994. Internal temperature calibration for ^{1}H NMR spectroscopy studies of blood plasma and other biofluids. NMR Biomed 7:243–247.
40. Pfeffer PE, Valentine KM, Parrish FW. 1979. Deuterium-induced differential isotope shift ^{13}C NMR. 1. Resonance reassignments of mono- and disaccharides. J Am Chem Soc 101:1265–1274.
41. Reuben J. 1985. Molecular structure as reflected in the ^{13}C NMR spectra of oligosaccharides with partially deuterated hydroxyls. J Am Chem Soc 107:1747–1755.
42. Bose B, Zhao S, Stenutz R, Cloran F, Bondo PB, Bondo G, Hertz B, Carmichael I, Serianni AS. 1998. Three-bond C-O-C-C spin-coupling constants in carbohydrates: Development of a Karplus relationship. J Am Chem Soc 120:11158–11173.
43. Live D, Silks LA III, Schmidt J. 2001. ^{13}C isotopic enrichment for nuclear magnetic resonance studies of carbohydrates and glycoconjugates. Meth Enzymol 338:305–319.
44. Dais P, Perlin AS. 1987. Proton spin–lattice relaxation rates in the structural analysis of carbohydrate molecules in solution. Adv Carbohydr Chem Biochem 45: 125–168.
45. Dais P. 1995. Carbon-13 nuclear magnetic relaxation and motional behavior of carbohydrate molecules in solution. Adv Carbohydr Chem Biochem 51:63–131.
46. Allerhand A, Doddrell D. 1971. Strategies in the application of partially relaxed Fourier transform nuclear magnetic resonance spectroscopy in assignments of carbon-13 resonances of complex molecules. Stachyose. J Am Chem Soc 93:2777–2779.
47. Bock K, Thögersen H. 1982. Nuclear magnetic resonance spectroscopy in the study of mono- and oligosaccharides. Annu Rep NMR Spectrosc 13:1–57.
48. De Bruyn A, Anteunis M, Van Beeumen J. 1977. Chemical shifts of aldohexopyranoses revisited and application to gulosylglucose. Bull Soc Chim Belg 86: 259–265.
49. Hobley P, Howarth O, Ibbett RN. 1996. ^{1}H and ^{13}C NMR shifts for aldopyranose and aldofuranose monosaccharides: Conformational analysis and solvent dependence. Magn Reson Chem 34:755–760.
50. O'Connor JV, Nunez HA, Barker R. 1979. α- and β-Glycopyranosyl phosphates and 1,2-phosphates. Assignments of conformations in solution by ^{13}C and ^{1}H NMR. Biochemistry 18:500–507.
51. Hull WE. 1994. Experimental aspects of two-dimensional NMR. In: Croasmun WR, Carlson RMK, eds. Two-Dimensional NMR Spectroscopy: Applications for Chemists and Biochemists. New York: VCH; p 67–456.
52. Vliegenthart JFG, van Halbeek H, Dorland L. 1981. The applicability of 500-MHz high-resolution ^{1}H-NMR spectroscopy for the structure determination of carbohydrates derived from glycoproteins. Pure Appl Chem 53:45–77.
53. Vliegenthart JFG, Dorland L, van Halbeek H. 1983. High-resolution, ^{1}H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. Adv Carbohydr Chem Biochem 41:209–374.
54. Kamerling JP, Dorland L, van Halbeek H, Vliegenthart JFG, Messer M, Schauer R. 1982. Structural studies of 4-*O*-acetyl-α-*N*-acetylneuraminyl-(2→3)-lactose, the main oligosaccharide in echidna milk. Carbohydr Res 100:331–340.
55. Bradbury JH, Jenkins GA. 1984. Determination of the structures of trisaccharides by ^{13}C-N.M.R. spectroscopy. Carbohydr Res 126:125–156.
56. Coxon B. 1980. Carbon-13 nuclear magnetic resonance spectroscopy of food-related disaccharides and trisaccharides. In: Lee CK, ed. Developments in food carbohydrate, Vol. 2. London: Applied Science; p 351–390.
57. Hounsell EF, Wright DJ, Donald ASR, Feeney J. 1984. A computerized approach to the analysis of oligosaccharide structure by high-resolution proton n.m.r. Biochem J 223:129–143.
58. van Kuik JA, Hård K, Vliegenthart JFG. 1992. A ^{1}H NMR database computer program for the analysis of the primary structure of complex carbohydrates. Carbohydr Res 235:53–68.
59. Lipkind GM, Shashkov AS, Knirel YA, Vinogradov EV, Kochetkov NK. 1988. A computer-assisted structural analysis of regular polysaccharides on the basis of ^{13}C-N.M.R. data. Carbohydr Res 175:59–75.
60. Jansson P-E, Kenne L, Widmalm G. 1987. CASPER—a computerised approach to structure determination of polysaccharides using information from N.M.R. spectroscopy and simple chemical analyses. Carbohydr Res 168:67–77.
61. Stenutz R, Jansson P-E, Widmalm G. 1998. Computer-assisted structural analysis of oligo- and polysaccharides: An extension of CASPER to multibranched structures. Carbohydr Res 306:11–17.
62. Amendolia SR, Doppiu A, Ganadu ML, Lubinu G. 1998. Classification and quantitation of ^{1}H NMR spectra of alditols binary mixtures using artificial neural networks. Anal Chem 70:1249–1254.
63. Radomski JP, van Halbeek H, Meyer B. 1994. Neural network-based recognition of oligosaccharide ^{1}H-NMR spectra. Nat Struct Biol 1:217–218.
64. Corne SA. 1996. Artificial neural networks for pattern recognition. Concepts Magn Reson 8:303–324.
65. Loss A, Bunsmann P, Bohne A, Loss A, Schwarzer E, Lang E, von der Lieth C-W. 2002. SWEET-DB: An attempt to create annotated data collections for carbohydrates. Nucleic Acids Res 30:405–408.
66. Bubb WA, Wright LC, Cagney M, Santangelo RT, Sorrell TC, Kuchel PW. 1999. Heteronuclear NMR studies of metabolites produced by *Cryptococcus neoformans* in culture media: Identification of possible virulence factors. Magn Reson Med 42:442–453.

67. Altona C, Haasnoot CAG. 1980. Prediction of *anti* and *gauche* vicinal proton–proton coupling constants in carbohydrates: A simple additivity rule for pyranose rings. Org Magn Reson 13:417–429.
68. Zwahlen C, Vincent SJF. 2002. Determination of ^{1}H homonuclear scalar couplings in unlabeled carbohydrates. J Am Chem Soc 124:7235–7239.
69. Lerner L, Bax A. 1987. Application of new, high-sensitivity, ^{1}H-^{13}C-N.M.R.-spectral techniques to the study of oligosaccharides. Carbohydr Res 166:35–46.
70. Otter A, Bundle DR. 1995. Long-range ^{4}J and ^{5}J, including *inter*glycosidic correlations in gradient-enhanced homonuclear COSY experiments of oligosaccharides. J Magn Reson B 109:194–201.
71. Turner CJ, Connolly PJ, Stern AS. 1999. Artifacts in sensitivity-enhanced HSQC. J Magn Reson 137:281–284.
72. Podlasek CA, Wu J, Stripe WA, Bondo PB, Serianni AS. 1995. [^{13}C]-Enriched methyl aldopyranosides: Structural interpretations of ^{13}C-^{1}H spin-coupling constants and ^{1}H chemical shifts. J Am Chem Soc 117: 8635–8644.
73. Bubb WA, Urashima T, Fujiwara R, Shinnai T, Ariga H. 1997. Structural characterization of the exocellular polysaccharide produced by *Streptococcus thermophilus* OR 901. Carbohydr Res 301:41–50.
74. Tvaroska I, Taravel FR. 1995. Carbon–proton coupling constants in the conformational analysis of sugar molecules. Adv Carbohydr Chem Biochem 51:15–61.
75. Bax A, Summers MF. 1986. ^{1}H and ^{13}C assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR. J Am Chem Soc 108:2093–2094.
76. Schwarcz JA, Cyr N, Perlin AS. 1975. Orientation effects and the sign of two-bond ^{13}C–^{1}H coupling. Can J Chem 53:1872–1875.
77. Bock K, Pedersen C. 1977. Two- and three-bond ^{13}C–^{1}H couplings in some carbohydrates. Acta Chem Scand B31:354–358.
78. Cyr N, Hamer GK, Perlin AS. 1978. Some stereochemical aspects of two-bond ^{13}C–^{1}H spin–spin coupling. Sign determination in natural abundance ^{13}C nuclear magnetic resonance spectra. Can J Chem 56:297–301.
79. Zhu G, Renwick A, Bax A. 1994. Measurement of two- and three-bond ^{1}H–^{13}C J couplings from quantitative heteronuclear J correlation for molecules with overlapping ^{1}H resonances, using t_1 noise reduction. J Magn Reson A 110:257–261.
80. Tvaroska I, Hricovini M, Petrakova E. 1989. An attempt to derive a new Karplus-type equation of vicinal proton–carbon coupling constants for C-O-C-H segments of bonded atoms. Carbohydr Res 189:359–362.
81. Poppe L, van Halbeek H. 1991. Selective, inverse-detected measurements of long-range ^{13}C, ^{1}H coupling constants. Application to a disaccharide. J Magn Reson 93:214–217.
82. Hounsell EF. 1995. ^{1}H NMR in the structural and conformational analysis of oligosaccharides and glycoconjugates. Progr Nucl Magn Reson Spectrosc 27:445–474.
83. Kiddle GR, Harris R, Homans SW. 1998. Heteronuclear Overhauser effects in carbohydrates. J Biomol NMR 11:289–294.
84. Serianni AS, Bondo PB, Zajicek J. 1996. Verification of the projection resultant method for two-bond ^{13}C–^{13}C coupling sign determinations in carbohydrates. J Magn Reson B 112:69–74.
85. Martin-Pastor M, Bush CA. 2000. The use of NMR residual dipolar couplings in aqueous dilute liquid crystalline medium for conformational studies of complex oligosaccharides. Carbohydr Res 323:147–155.
86. Tian F, Al-Hashimi HM, Craighead JL, Prestegard JH. 2001. Conformational analysis of a flexible oligosaccharide using residual dipolar couplings. J Am Chem Soc 123:485–492.
87. Kiddle GR, Homans SW. 1998. Residual dipolar couplings as new conformational restraints in isotopically ^{13}C-enriched oligosaccharides. FEBS Lett 436:128–130.
88. Mulquiney PJ, Bubb WA, Kuchel PW. 1999. Model of 2,3-bisphosphoglycerate metabolism in the human erythrocyte based on detailed enzyme kinetic equations: *In vivo* kinetic characterization of 2,3-bisphosphoglycerate synthase/phosphatase using ^{13}C and ^{31}P NMR. Biochem J 342:567–580.
89. Jones DNM, Sanders JKM. 1989. Biosynthetic studies using ^{13}C-COSY: The *Klebsiella* K3 serotype polysaccharide. J Am Chem Soc 111:5132–5137.
90. Berthon HA, Bubb WA, Kuchel PW. 1993. ^{13}C n.m.r. isotopomer and computer-simulation studies of the non-oxidative pentose phosphate pathway of human erythrocytes. Biochem J 296:379–387.
91. Hurtley S, Service R, Szuromi P. 2001. Cinderella's coach is ready. Science 291:2337.

BIOGRAPHY

William (Bill) Bubb, B.Sc.(Hons), Ph.D., D.I.C., is a Senior Research Fellow in the School of Molecular and Microbial Biosciences, where he manages the NMR laboratory. His Ph.D. at the University of Sydney, under the guidance of Sev Sternhell, included investigations of correlations between NMR spectral parameters and structural fragments in small organic molecules. Following postdoctoral work with the late Sir Derek Barton on the development of reagents for organic synthesis, and a period in industrial R&D, he returned to the University of Sydney in 1980. His current research is focused on the use of NMR spectroscopy for the characterization of biologic systems under stress and for structural studies of biologically significant carbohydrates.