Progress in Nuclear Magnetic Resonance Spectroscopy 79 (2014) 48–68

Contents lists available at ScienceDirect

Progress in Nuclear Magnetic Resonance Spectroscopy

journal homepage: www.elsevier.com/locate/pnmrs

NMR of glycans: Shedding new light on old problems

Marcos D. Battistel, Hugo F. Azurmendi, Bingwu Yu, Darón I. Freedberg *

Laboratory of Bacterial Polysaccharides, Center for Biologics Evaluation and Research, Food and Drug Administration, 1401 Rockville Pike, Rockville, MD 20852-1448, United States

Edited by J.W. Emsley and J. Feeney

ARTICLE INFO

Article history:
Received 3 July 2013
Accepted 10 January 2014
Available online 14 February 2014

Keywords:
Glycan NMR
Hydroxyl groups
Hydrogen bonds
Isotopic labeling and 3D Structure
Residual dipolar coupling

ABSTRACT

The diversity in molecular arrangements and dynamics displayed by glycans renders traditional NMR strategies, employed for proteins and nucleic acids, insufficient. Because of the unique properties of glycans, structural studies often require the adoption of a different repertoire of tailor-made experiments and protocols. We present an account of recent developments in NMR techniques that will deepen our understanding of structure–function relations in glycans. We open with a survey and comparison of methods utilized to determine the structure of proteins, nucleic acids and carbohydrates. Next, we discuss the structural information obtained from traditional NMR techniques like chemical shifts, NOEs/ROEs, and coupling-constants, along with the limitations imposed by the unique intrinsic characteristics of glycan structure on these approaches: flexibility, range of conformers, signal overlap, and non-first-order scalar (strong) coupling. Novel experiments taking advantage of isotopic labeling are presented as an option for overcoming spectral overlap and raising sensitivity. Computational tools used to explore conformational averaging in conjunction with NMR parameters are described. In addition, recent developments in hydroxyl detection and hydrogen bond detection in protonated solvents, in contrast to traditional sample preparations in D_2O for carbohydrates, further increase the tools available for both structure information and chemical shift assignments. We also include previously unpublished data in this context. Accurate determination of couplings in carbohydrates has been historically challenging due to the common presence of strong-couplings. We present new strategies proposed for dealing with their influence on NMR signals. We close with a discussion of residual dipolar couplings (RDCs) and the advantages of using ^{13}C isotope labeling that allows gathering one-bond ^{13}C–^{13}C couplings with a recently improved constant-time COSY technique, in addition to the commonly measured ^{1}H–^{13}C RDCs.
Published by Elsevier B.V.

Contents

* Corresponding author.
E-mail address: daron.freedberg@fda.hhs.gov (D.I. Freedberg).

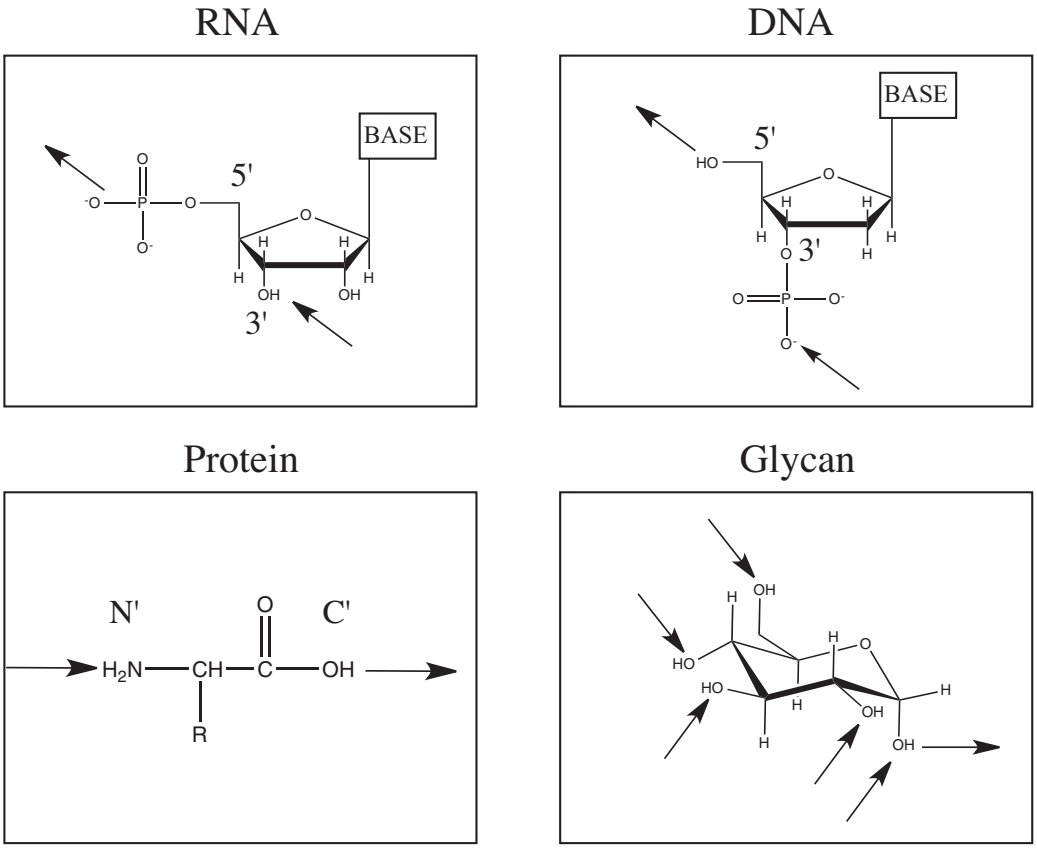
http://dx.doi.org/10.1016/j.pnmrs.2014.01.001
0079-6565/Published by Elsevier B.V.

1. Introduction

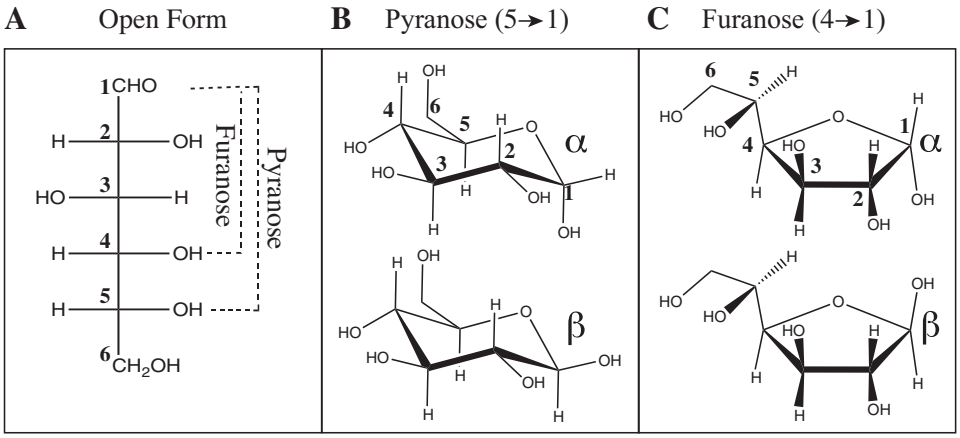
Carbohydrates are the most ubiquitous class of biomolecule in nature, yet we know little about their three-dimensional structures and more importantly, how their structures and flexibility translate into biological functions. In part, this is because the potential of carbohydrates to store chemical information, known as the "carbohydrate code" [1–3], is overwhelming. If each unique oligosaccharide sequence, and its corresponding three-dimensional structure, accounts for a different biological function, the potential for information storage in a glycan sequence is immense: for a linear carbohydrate hexamer with no repeating units, the number of possible combinations rises to $>1 \times 10^9$ sequences [4] compared to 6×10^6 and 6×10^5 for proteins and nucleic acids, respectively. Nucleic acids and proteins elongate unidirectionally, in a template-dependent manner (Scheme 1). Thus, their diversity and growth is restricted by the information contained in the nucleic acid template. In contrast, carbohydrate chain elongation proceeds in a template-independent fashion, in which a glycan unit, chemically reacting with an available hydroxyl group, forms a linkage to the growing chain through its anomeric position. Unlike proteins and nucleic acids, the glycan chain can be branched like glycogen (Scheme 1). Because the template constraint is not present, a leap in diversity occurs.

The carbohydrate code has usually been compared to a dictionary, in which each unit represents a word. In contrast, we propose that the carbohydrate code bears more of a resemblance to hieroglyphs (or glyphs), because each unit's meaning does not only depend on what precedes or follows (context, e.g. primary sequence), but also depends on its 3D shape; the code thus resembles more of an "idea" that can be built upon. This hieroglyph, is not only embedded in the chemical identity of its units but also in the α or β configuration at the anomeric carbon, which influences molecular shape; the possibility of multiple ring sizes (furanose or pyranose); the potential for chain elongation and branching at multiple chemically equivalent hydroxyl groups in different positions; and potential chemical modification of hydroxyl to groups such as amino, O-acetyl, N-acetyl, phosphate, sulfate or even to substitute a hydroxyl group by a hydrogen atom to give a deoxy sugar (Scheme 2). What remains unknown, however, is whether the increase in complexity reflects a functional importance.

Carbohydrates are often found covalently linked, or conjugated, to proteins and lipids. In this manner, the function of the carbohydrate-bearing biomolecule can be modified or modulated. The same carbohydrate fragment may appear in different cellular regions depending on the molecule to which it is conjugated. Because the molecules to which glycans bind can modulate glycan structure and function, their roles may depend on where they are



Scheme 1.



Open, pyranose and furanose forms of an aldose, showing the many equivalent OH groups. The OH groups can be modified by $-PO_4^{-3}$, $-SO_4^{-2}$, $-NH_2$, $-CO-CH_3$, $-NH-CO-CH_3$, -H.

Scheme 2.

localized. For example, glycolipids may form rafts on cell surfaces, and by increasing their local concentration, their interaction with target receptors becomes possible.

Animal cells are covered by glycoproteins and glycolipids, where they serve as mediators for which the glycosylation changes in numerous physiological (like embryogenesis [5] and development [6]) and pathological processes (HIV survival [7] and cancer [8]). Moreover, cells can change glycosylation patterns as a result of stimuli. Other cells can interpret this change in glycosylation by means of carbohydrate binding proteins (lectins). One can envision glycan diversity as a means for cells to respond to the stimuli of different environments. The lectins belonging to other cells in these environments must be able to bind to this diverse set of glycans, so that they can respond appropriately. A notorious example where a cell radically changes its glycosylation pattern is in cancer [9]. This observation shows promise for the development of cancer vaccines [10] and early detection of malignancy [11]. The dynamism and diversity on glycosylation, on one hand enables an overwhelming capacity of encoding information in biological systems, and on the other hand, presents enormous challenges and exciting opportunities for carbohydrate research.

Despite the enormous potential for glycan variability, the actual diversity may be limited. It was estimated from the number of human glycosyltransferases encoded in the genome that there might exist ca. 500 unique glycan structures [12]. This number was clearly underestimated as a simple search in the Consortium for Functional Glycomics webpage (www.functionalglycomics.org) yields 5746 unique glycan structures deposited. Nonetheless, if one considers the potential number of glycan determinants (oligosaccharide composition that can be recognized by a protein), it has been estimated that the number of different sequences in the human glycome is ~7000. This number may also be underestimated, however, it is more feasible and can be approached by chemical, enzymatic or chemoenzymatic synthesis of glycans, if one attempts the generation of the glycome in a test tube [13].

The potential glycan heterogeneity in living systems and in saccharide synthesis, presents difficulties in the production of large quantities of homogeneous samples, required for structural research. Unfortunately, difficulties inherent to carbohydrate research have kept researchers away from the field, which in turn has hampered a faster development of new methods for efficient sample production and for three-dimensional structure determination.

Nevertheless, there have been significant advances in the production of glycans, from the generation of homogeneous glycopeptides [14,15] to "taming" of enzymes for the synthesis of polymers or oligomers [16] and solid phase synthesis of oligosaccharides, recently reviewed by Hsu et al. [17]. This work, has transformed tailored oligosaccharide production, once considered a distant horizon, to an almost tangible goal. The control over the production of the desired oligosaccharide will open the door to the generation of labeled polysaccharide (PS) samples that would enormously aid NMR studies.

We will first summarize the strategies followed in the NMR community for the structural elucidation of proteins and nucleic acids, and subsequently focus on the structure elucidation of glycans and give a perspective on what can we learn, what we have accomplished, and what remains to be done.

1.1. Existing methods for protein solution structure determination

^{13}C,^{15}N protein labeling strategies in minimal media [18] enabled the development of triple resonance NMR experiments, which led naturally to unambiguous assignment of the resonances. Before stable isotope protein-labeling techniques were available, determining protein structure in solution was an unthinkable task. Nowadays, 20 kDa proteins are routinely solved by solution NMR (90 before 2000 and 727 by December 2013, www.rcsb.org). As the number of protein structures in the PDB increased, so did our understanding of structural features that characterize the protein backbone. This understanding led to the development of structure validation protocols (Whatif [19], Prosa [20], Procheck [21,22]), and structure prediction software [23–25]. Computer aided structure prediction tools such as ROSETTA [23,26] are becoming more reliable, as progress was catalyzed by the growth of structures and NMR data available in PDB and BMRB databases (www.pdb.org and www.rcb.org, respectively). As we will discuss later, structural glycobiology is in its infancy, being at the point that protein structural determination was before specific and uniform isotopic labeling technology was widespread.

The methods used to determine protein structures should serve as a reference of what we strive for in glycan structure determination. Although no unique protocol for the generation of a protein structural model exists, general guidelines can be followed (Fig. 1). As displayed in Fig. 1, one cannot proceed in biomolecular structure determination without the unambiguous assignment of resonances. Thus, chemical shift assignment remains the bottleneck for structure determination of all biomolecules via NMR. Once this challenge is overcome, the task to be completed is the interpretation of distance restraints from NOE (nuclear Overhauser effect) correlation peaks, which can be tedious. Presently, the availability of software, like ARIA [27,28], facilitates NOE peak assignment and their utilization for structure calculation, in a semi-automatic manner.

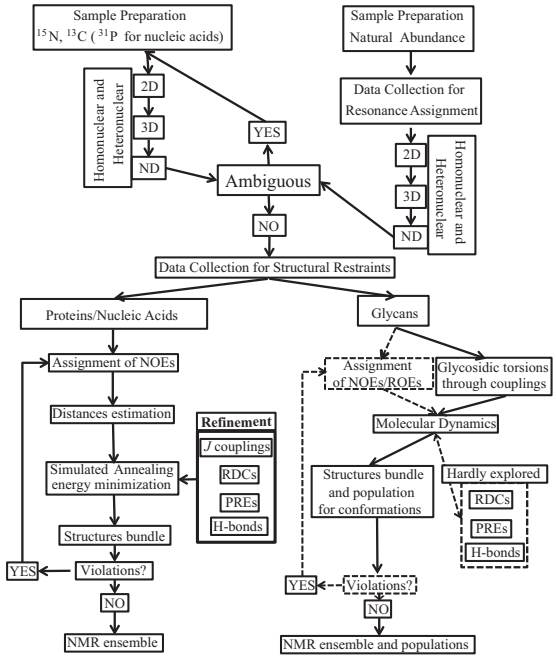


Fig. 1. Flowchart comparing a generalized approach for solution structural determination of biomolecules. Dashed squares identify aspects in the structural determination of glycans that need improvement or are underutilized.

Advances in protein structure determination led to the deposition of more structures (currently there are ~87,000 NMR and X-ray structures in the PDB). The discovery of new folds, common angles adopted by the peptide backbone for different structural elements and tabulation of chemical shifts and other NMR observables (hydrogen bonds, coupling constants, RDCs, etc.), enables the evaluation of a newly determined structure. Thus, as more structures are available, there is more confidence in the validation analysis and therefore in the resulting newly determined structures.

Currently, proteins $\geqslant$25 kDa can be studied with the aid of selective protein perdeuteration [29], which improves the relaxation properties of high molecular weight proteins, and decreases spectral overlap. Nowadays, the challenge for protein structure determination lies in the automation of high throughput structural determination for which a standard procedure is required [30]. As larger proteins and protein complexes of up to ~100 kDa are being solved by NMR [31], new methods are being developed for systems with increased complexity. Researchers are also beginning to tackle highly dynamic systems such as partially folded or "disordered" proteins and to uncover biologically relevant invisible states [32–34]. This change in research focus should yield new methods for analyzing dynamic molecules. It is therefore especially fruitful for flexible glycans and for structural glycobiology in general.

1.2. Existing methods for nucleic acid solution structure determination

As with protein NMR, nucleic acid structural studies are greatly facilitated by the availability of selective labeling (reviewed by Fürtig et al. [35]) and perdeuteration techniques [36]. Nucleic acids suffer from a higher degree of signal overlap when compared to proteins. Nevertheless, structures larger than 25 kDa have been solved, either utilizing selective isotope labeling, perdeuteration, and/or by applying a divide and conquer approach in which small portions of a large molecule are probed separately [37].

Despite the fact that different sets of NMR experiments are required from those used in proteins, NOEs still play a crucial role in structure determination of nucleic acids (Fig. 1) [35,36]. However, because the chemical identity of nucleic acid units and their structural features differ from proteins, new NMR methods were developed to address structural questions of these biomolecules, for instance, determination of sugar backbone torsion angles via HCCH-E.COSY [38] and identification of base pairs, through H-bond (hydrogen bond) detection via $^{2}J_{NN}$ correlations [39]. Therefore, the generation of structural models relies on estimated distances obtained from NOEs in conjunction with sugar backbone angles obtained from homo- and heteronuclear couplings [40].

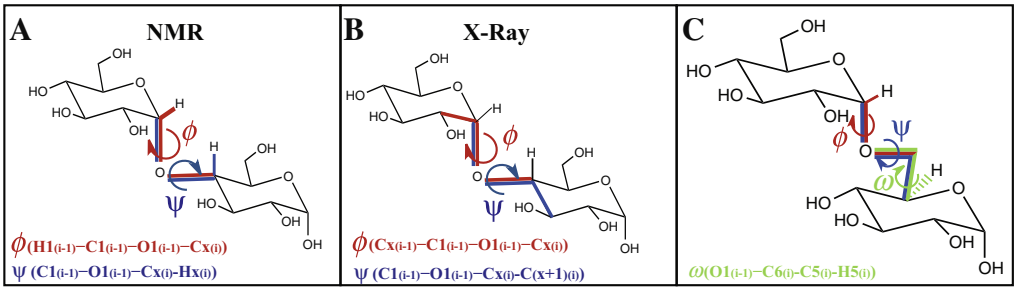
Strategies employed for nucleic acids can be used or adapted for solving glycan structures, because the level of signal overlap observed for nucleic acids approaches that observed in the NMR spectra of glycans. Additionally, as we will see later, H-bond detection could play an important role in glycan structural determination as well.

1.3. Current state of carbohydrate solution structure determination

Glycan structure determination by NMR has some added inherent challenges compared to nucleic acid structure determination. First, glycan structure is dissimilar to nucleic acid and protein structure. In proteins, the backbone is semi rigid, while the side chains are more flexible. In glycans, the opposite is generally the case. Pyranose rings, that compose the oligomer or polymer, are regarded as rigid; furanose rings are potentially flexible, but may be somewhat limited in conformational space as a result of steric or stereo-electronic constraints. Details of these effects and their consequences remain uncharacterized.

To define glycan conformation, therefore, the important parameters are the torsional angles ϕ and ψ, where, in principle, much flexibility is allowed. Depending on the experimental method used, the angles ϕ and ψ are defined slightly differently. In NMR studies, the torsional angles between two residues are generally defined as ϕ ($H1_{(i-1)}$–$C1_{(i-1)}$–$O1_{(i-1)}$–$C_{x(i)}$), and ψ ($C1_{(i-1)}$–$O1_{(i-1)}$–$C_{x(i)}$–$H_{x(i)}$). In X-ray studies, heavy atoms are used instead of hydrogens for the first atom in ϕ and the last atom in ψ (Scheme 3). Rao et al. detail the conversion between NMR and X-ray parameters in their textbook [41]. In the case of (1 → 6) linked glycans, an additional torsion angle (termed χ or ω) is introduced, defined as $O1_{(i-1)}$–$C6_{(i)}$–$C5_{(i)}$–$H5_{(i)}$ (Scheme 3).

The glycosidic (carbon–oxygen–carbon) linker is less rigid than the amide and phosphate backbones in proteins and nucleic acids, respectively. Therefore, unlike proteins and nucleic acids, structure determination of glycans cannot rely solely on distance restraints from NOE/ROE and dihedral angles extracted from coupling constants (Fig. 1). NMR parameters other than NOEs and coupling constants have not been easy to extract, therefore, the energy landscapes of carbohydrates remain poorly characterized. It is generally assumed that the landscape has many valleys, each representing a distinct energy minimum with its corresponding conformation, and that molecules populate different local energy minima in addition to the global minimum. The relative populations among states depend on the energy difference between them. The detailed characterization of multiple conformational states is further complicated by shallow energy barriers between conformations, leading to fast exchange at room temperature. Furthermore, the populations of these conformational states are complicated by solvation, as water can stabilize certain conformations by interacting specifically with different portions of the glycan. In principle, *ab initio* quantum mechanical calculations can provide more accurate energy differences between conformational states. In practice, however, these calculations are difficult to execute with solvation; consequently they are typically carried out without explicit water molecules, which can affect energies of the different conformations. Nevertheless, some *ab initio* quantum mechanical studies on glycans



Scheme 3.

have been reported. One such study on glucose required eight or nine water molecules to correctly reproduce experimental vibrational frequencies in aqueous glucose [42]. As a result, the energy barriers for conformational interconversion also remain unknown. Therefore, researchers overly rely on trans-glycosydic coupling constants and their interpretation via molecular dynamics (MD) simulations (Fig. 1), to extract information on the number of valleys for a given oligosaccharide and their relative population.

Usually, experimental data are available to aid in force field parameterization. However, it is difficult to use experimental solution data to verify these populations for flexible glycans. Thus, NMR parameters such as coupling constants and distances usually reflect a weighted average of the different conformations that are present under determined conditions (fast exchange), rather than separate molecular species (slow exchange) [43]. At the same time, parameterization of force fields for MD, relies in part on NMR observables, whose interpretation is not always clear-cut and may be subject to errors, as we will see in the following sections [44].

Parameterizing force fields is also challenging. Small changes in parameter sets used in MD can significantly impact the energy landscapes [45,46], which in turn will affect the interpretation of the NMR data. To avoid this complication, independent methods for generation of structures and validation protocols are needed for accurate determination of oligosaccharide structures.

Widmalm's group has pioneered efforts to properly parameterize force fields using NMR data [47–49]. In their work, the group routinely synthesizes glycans and measures NOE buildup curves and *J*-couplings to verify and fine-tune carbohydrate force field parameterization. In cases where spectral overlap or relaxation interferes with the accuracy of the measurements, the group has examined deuterated or ^{13}C labeled material to alleviate the problems.

Glycan structure determination can garner much from protein and nucleic acid NMR methods. For example, lessons can be learned spanning everything from how signal degeneracy is handled in NMR of nucleic acids or large proteins, to the information that can be extracted from disordered or partially folded proteins. Disordered proteins have much in common with glycans, including their lack of stable structure, dynamic behavior and the consequent spectral overlap. Nevertheless, new tools, tailored for glycan structure determination must also be developed. Such methods are needed to address flexibility and dynamics, spectral overlap and strong coupling effects [50–57]. Thus, NMR experiments already developed for other biomolecules can be used for glycan structural analysis, however since glycans behave differently from other biomolecules, e.g. different spin scalar couplings, different linkers (ϕ, ψ and ω torsions), relaxation properties and different sets of chemical shifts, the standard protein NMR experiments must be fine-tuned for application to glycans.

In the following section we describe how glycan structure determination may be approached, and review the current methods for conformational analysis.

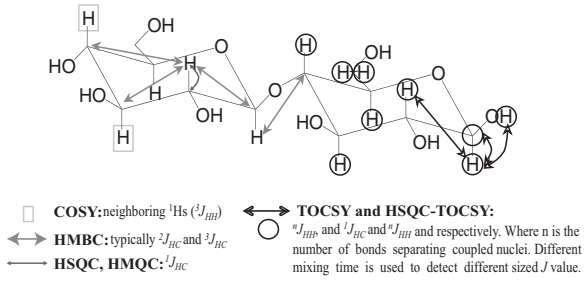
2. Current methods in carbohydrates structural analysis: problems and troubleshooting

Traditionally, three NMR parameters are employed in oligosaccharide structural analysis: (1) chemical shifts; (2) NOEs (or ROEs); (3) coupling constants. We will discuss the topics in the stated order because chemical shifts and NOEs are "stand-alone" parameters, whereas utilization of coupling constants for structural determination of oligosaccharides, while adding to these, cannot on their own define structure for reasons we describe later. As such, coupling constants are highly dependent on MD simulations for their interpretation.

2.1. Chemical shifts

Chemical shift assignment for the system in question is essential for structure determination and is the first step in structural analysis via NMR. Assignments enable us to gather unambiguous structural information at an atomic level, providing that the signals do not overlap. Assignments also provide insight on the chemical environment of a given nucleus, such as the presence of substituents as well as the overall three dimensional structure. ^{13}C chemical shifts can provide information on structural changes and the presence of different functional groups at vicinal carbons [58]. Additionally, variations of chemical shifts as a function of temperature report on transient H-bonds and the presence of conformational averaging [59–61]. ^{1}H and ^{13}C chemical shifts, together with coupling constants can be used for automated glycan sequence determination [62,63]. The observed chemical shifts of interconverting conformations in fast exchange are a population-weighted average of the chemical shifts of each conformation; changes in populations as a function of temperature lead to changes in the observed chemical shift.

Scheme 4 depicts the set of NMR experiments that are used to obtain complete chemical shift assignments of glycans. HSQC experiments [64] can identify directly attached ^{1}H and ^{13}C pairs (or other heteronuclei), whereas HMBC experiments [65] can correlate the atom pairs separated by two or more bonds.

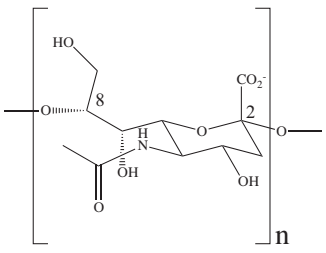


Scheme 4.

Carbohydrate ^{1}H 1D NMR spectra can be divided in two regions, the anomeric region for H1 protons (4.5–6.0 ppm) and the "bulk" region (3.0–4.5 ppm), in which most of the other ring proton resonances appear. Integration of the anomeric signals can be used as a rough estimate of the number of units that constitute the oligomer. In the case of PS, where many repeating units are present, signal overlap will occur. In such cases, the number of proton signals in the anomeric region is usually equal to the number of different types of anomeric protons in a single repeating unit, as opposed to the entire PS. The same behavior is observed for resonances in the bulk region. In the case of α2,8-linked polysialic acid (Scheme 5), a capsular PS of *Neisseria meningitidis* serogroup B of approximately 250 sialic acid units, an HSQC spectrum displays only 10 signals, one for each ^{1}H, ^{13}C pair, similar to what is observed for a single α2,8-linked unit.

In oligosaccharides, units are linked between the anomeric carbon of one residue and virtually any other carbon of the next unit via an oxygen atom (Scheme 1). The anomeric signal isolation makes them useful as NMR "handles". Thus, they provide valuable information for sequential assignment. However, because anomeric proton signals commonly overlap with the water resonance, carbohydrate NMR experiments have traditionally been carried out in D_2O as opposed to H_2O. In early carbohydrate NMR studies, presaturation was used to remove residual ^{1}H signals in the HOD peak. Although these procedures enable the observation of anomeric protons, information on all exchangeable protons in the molecule, such as OHs and HNs, is lost through chemical exchange with the saturated water. As we will see later, these exchangeable protons could otherwise be used to gather structural data [61,66].

The signal isolation in ^{1}H and ^{13}C of anomeric nuclei is useful in 1D selective and 2D-HSQC-TOCSY [67] experiments, where the concept of resonance handles is easy to exploit. These experiments yield data that link all the *J*-coupled protons in a molecule and are often sufficient to carry out a complete assignment for each glycan unit. Thus, each anomeric proton, representative of each glycan unit, will be connected to all other protons in this residue. For simpler spin systems, ^{1}H,^{1}H-COSY [68] and ^{1}H,^{1}H TOCSY may provide the necessary data for the complete assignment [69]. Traditionally, COSY experiments are used for glycan assignments; however TOCSY provides higher resolution and is more efficient in transferring magnetization between coupled protons [70], hence TOCSY is the recommended procedure. Fig. 2 shows results for a phase-sensitive COSY compared to those for a 10 ms mixing time TOCSY collected under the same conditions, except for the number of scans for each t_1 point (32 for COSY and 8 for TOCSY). The TOCSY experiment provides higher sensitivity and resolution in both dimensions because of the absorptive in-phase peaks compared to the dispersive antiphase peaks in COSY. The higher resolution in TOCSY is more evident in the cross-peaks close to the diagonal (Fig. 2A vs. 2B). Although this has been discussed in the literature [70,71], many researchers still use COSY spectra to assign carbohydrate ^{1}H resonances. The experiments mentioned above provide information about the type of linkage between residues and on the sequence



Scheme 5.

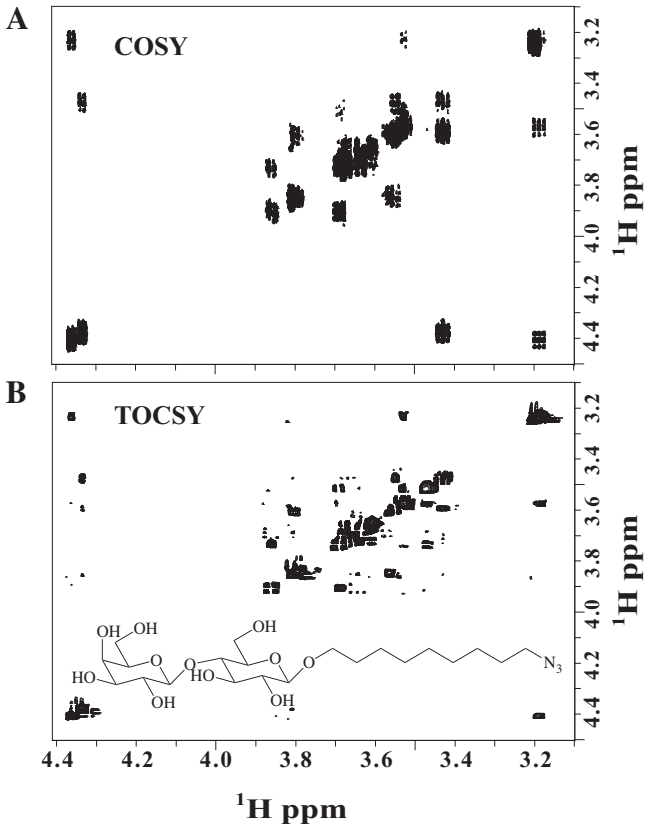


Fig. 2. Homonuclear two-dimensional spectra collected for the compound shown in panel B. Data were collected at 37 °C on a Bruker Avance 700 MHz instrument equipped with a triple gradient QXI probe. (A) Phase-sensitive COSY (Bruker pulse sequence cosyetgpph) with 32 scans per t_1 point (128 total t_1 points). (B) Phase-sensitive TOCSY (Bruker pulse sequence dipsi2etgpph) with 10 ms mixing time and 8 scans per t_1 point (128 total t_1 points). Note the improvement of resolution near the diagonal in TOCSY.

of the units in an oligosaccharide. When labeled material is available, a complete resonance assignment can be carried out via HCCH-COSY [72] and HCCH-TOCSY [73]. For a more extensive review on assignment of carbohydrates see Duus et al. [74]. However, as detailed below, much has changed and the situation has improved since this review appeared.

2.2. NOEs, ROEs and beyond...

One of the most important NMR parameters for structure determination of biomolecules is the estimation of inter-nuclear distances from peak intensities obtained from NOE experiments, whose utility was first demonstrated experimentally by Anet and Bourn [75].

Because NOE experiments and the use of NOE-derived data for inter-nuclear distance estimation are well known and accessible, we only briefly outline the salient features of the nuclear Overhauser effect, and how it may be applied to glycans. We present a brief review of the factors to consider in selecting the "correct" (NOE or ROE) experiment for glycans. Then we discuss how distances can be used to derive three-dimensional models for carbohydrates.

The cross-relaxation rate, σ_{NOE}, depends on the distance between the involved nuclei (r), the molecule's rotational correlation time (τ_c), the Larmor frequency (ω, which depends on the magnetic field strength), and other constants, represented by K, including the Boltzmann constant and the ^{1}H magnetogyric ratio. The resulting expression for σ_{NOE} is [76]

$$\sigma_{NOE} = \frac{K}{r^6}\left(\frac{-\tau_c}{10} + \frac{3\tau_c}{5(1+(2\tau_c\omega)^2)}\right) \quad (1)$$

Nuclear Overhauser enhancements range from +50% on rapidly tumbling molecules (small correlation time) to −100% for very slowly tumbling molecules (long correlation time). For medium size molecules (~700–1200 g/mol), $\omega\tau_c \sim 1$ and the NOE becomes difficult to detect because $\sigma_{NOE} \sim 0$. One alternative is to perform an ROE experiment [77]. In this case, the cross-relaxation rate (σ_{ROE}) is never null, as described by [78]

$$\sigma_{ROE} = \frac{K}{r^6}\left(\frac{\tau_c}{5} + \frac{3\tau_c}{10(1+(\tau_c\omega)^2)}\right) \quad (2)$$

The advantages of this experiment with respect to NOE are that ROE cross-peaks due to dipolar interaction are always positive so that there is no null ROE signal for medium size molecules, and that the effects on ROEs due to chemical exchange can be distinguished from dipolar correlations because of their opposite sign. However, distances between nuclei are harder to estimate than from NOEs, because ROE intensities are also frequency offset dependent, which varies from one resonance to another. Therefore, distances obtained from σ^{Obs}_{ROEij} have to be corrected according to σ^{True}_{ROEij} given by [78,79]

$$\sigma^{True}_{ROEij} = \sigma^{Obs}_{ROEij}\frac{1}{\sin\theta_i \sin\theta_j} \quad (3)$$

where the true ROE between spins i and j depends on the frequency offset of each spin, θ_i and θ_j from the carrier. The ROE also contains a T_2 component.

Moreover, when NOE peaks are not negligible ($\omega\tau_c \not\approx 1$), during the spin-lock mixing time of the ROE experiment, there will be an NOE contribution to the cross-relaxation (σ_{ij}) described by

$$\sigma_{ij} = \sigma_{ROEij}\sin\theta_i \sin\theta_j + \sigma_{NOEij}\cos\theta_i \cos\theta_j \quad (4)$$

This contribution is not easily accounted for and presents a major drawback in the quantitative utilization of ROEs for distance determination.

TOCSY peaks can be observed in ROESY spectra which could complicate data analysis [76], but they can be minimized by using a simultaneously inefficient isotropic mixing sequence for TOCSY yet efficient for ROE [80]. In summary, ROEs can only be quantitative once they have been corrected for resonance offset and when NOEs are negligible under the same experimental conditions.

Therefore, the choice of experiment will depend on the system under study. Even for medium sized molecules, it is not possible to say *a priori* that ROE-based experiments will yield the best results. Under appropriate conditions, NOE experiments can yield good results even for medium size molecules.

The rotational correlation time, τ_c, can be expressed for a simple model as

$$\tau_c = 4\eta\pi a^3/3k_BT \quad (5)$$

where η is the solvent viscosity, a is the molecular radius, T is the temperature of the system, and k_B is Boltzmann's constant. The solvent viscosity and the temperature are directly correlated: the viscosity of water changes by a factor of 6 between 0 and 100 °C [76], therefore, decreasing the temperature, should increase the correlation time and result in observable NOEs that can be more easily interpreted than ROEs. However, in addition to the desired change in viscosity, a change in temperature may also perturb the conformational equilibrium. Eqs. (1)–(5) assume an isotropic tumbling and hence a single correlation time, which may not be the case for many flexible oligosaccharides. For a thorough study of carbohydrate molecular motion refer to the study of Poppe and van Halbeek [81]. When molecular tumbling is partially or fully anisotropic, more correlation times are necessary to fully describe this motion, and relaxation data becomes increasingly difficult to interpret because the internal motion cannot be analyzed separately from the overall tumbling [82,83]. As stated earlier, barrier heights and the overall shape and layout of the potential energy surface of carbohydrates are not well described. Therefore, since internal motion, flexibility and nuclear relaxation in carbohydrates cannot be known in advance, the choice between NOE and ROE-type experiments should be empirically evaluated for each molecule under study based on NOE/ROE buildup curves and the spectra that yields the best data. Molecular weight alone is not a sufficient criterion to select between NOE or ROE experiments. Finally, although long-range NOEs are scarce, short-range average distances can nevertheless be useful in the context of MD simulations to find agreement with experimental data.

2.3. Addressing signal degeneracy and conformational averaging

Once a method for estimating distances has been chosen, two related problems arise: (1) spectral overlap; and (2) conformational averaging.

2.3.1. Spectral overlap

For alleviating spectral overlap others [84–88] and we [89] have developed methods for obtaining fully and partially labeled glycans, which enables heteronuclear resonance editing for resolving assignment ambiguities present in the ^{1}H dimension of oligosaccharides [66,87,90–97]. In addition, ^{13}C isotopic enrichment gives ca. 100-fold signal enhancement compared to natural abundance samples (1.1% for ^{13}C isotope). Labeled material has also proven useful for the observation of carbohydrates in *in vivo* systems and for isolation of carbohydrate signals virtually free of background signals [89,93]. Unfortunately, there is no standard protocol for producing uniformly labeled PS and, at this point, labeling must be approached on a case-by-case basis. Methods for obtaining partially labeled carbohydrates are also available, but these materials are used for specific coupling constants determination and/or residual dipolar coupling studies, not for decreasing signal degeneracy [90,98].

Isotopic enrichment also has a few drawbacks. Replacing all ^{12}C atoms present in the molecule with ^{13}C, results in the ^{13}C nuclei efficiently relaxing each other and leads to short ^{13}C T_2 values, imposing experimental restrictions. Experiments on labeled samples require constant-time evolution in ^{13}C to eliminate ^{13}C–^{13}C multiplets. This presents a disadvantage for ^{13}C acquisition times in the indirect ^{13}C dimension because only multiples of $1/^1J_{CC}$ can be employed. Therefore, to resolve spectral overlap, the constant-time must be increased to achieve high resolution, which may lead to a decrease in or complete loss of signal due to reduced T_2 values.

Without question stable isotopic enrichment provides increased sensitivity and resolution in ^{13}C and ^{15}N-edited NMR experiments. However, it is also important to remember that unlabeled carbohydrates can also efficiently yield isotope-edited spectra. These methods are seldom utilized and should become more widespread.

2.3.2. Conformational averaging

There are three related issues that must be addressed when dealing with conformations of carbohydrates. The experimenter must: (1) define the number of conformations sampled by the oligosaccharide in solution, (2) determine the relative populations of conformers that coexist under given conditions, and (3) determine the rate of inter-conversion. The minimum energy structures are obtained by minimizing the structures' energies under a given force field [46,99–102]. This procedure will normally yield several possible structures as solutions. For proteins and nucleic acids, the

discrimination between more plausible structures or set of structures and their relative populations is usually carried out via restrained MD simulations with NMR derived distance and angular parameters, or unrestrained MD, for which the correlation with experimental data is done *a posteriori*. For a review of the use of MD to analyze carbohydrate structures see Fadda and Woods [45].

For flexible carbohydrates, it is challenging to derive reliable distance constraints based solely on NOE signal intensities. *J*-coupling is used to infer trans-glycosidic torsion angles obtained from empirically parameterized Karplus relations. However the accuracy of the Karplus equation for carbohydrates is in question because solution structures of appropriate model compounds are not available. In addition, because of the potential flexibility in carbohydrates, efforts to fit one single structure to an averaged experimental *J* value may not be appropriate.

A method has been developed to extract populations of conformations in solution from NOEs [103], this task uses a combination of complementary NMR parameters: NOEs (homo- and heteronuclear), ^{3}J-coupling constants ($^{1}H–^{1}H$, $^{1}H–^{13}C$, and $^{13}C–^{15}N$), and residual dipolar couplings (RDCs) [45]. Except for NOEs used to confirm linker connections, other inter-residue NOEs are rare in carbohydrates therefore NOEs alone provide insufficient restraints; they do not lead to a well-defined conformation or set of conformations, because different (ϕ,ψ) values can satisfy similar inter-residue distances. Fig. 3 depicts NMR observables, which in combination can yield better-defined structures or sets of structures. Thus, NOEs often yield the least resolution of the conformational space in the search for existing conformations in solution. Therefore, a complete set of NMR observables is desired if one strives to characterize glycan 3D structures in solution.

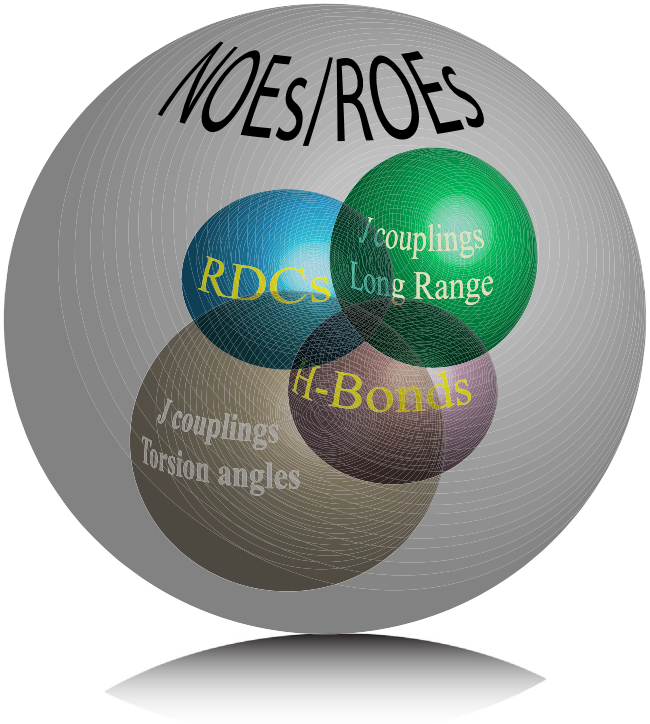


Fig. 3. The conformational space of glycans defined by different NMR observables. Overlapping volumes of the spheres represent a more accurate set of conformations that are described by experimental data.

2.3.3. Hydrogen bonding involving NH and OH groups of glycans

Because of the above-mentioned limitations, our research efforts in recent years have focused on characterizing oligomer conformations by searching for H-bonds, an independent set of NMR constraints. H-bonds are useful because their detection indicates the existence of persistent distinct H-bonded conformations, even if they are in transient equilibrium with other structures in solution.

Transient H-bonds were originally predicted for the pentasaccharide repeating unit of the O-antigen of *Escherichia coli* O142 strain from MD simulations [48]. Later, H-bonds were inferred from the measurement of temperature coefficients for ^{1}H and ^{15}N HSQC cross-peaks of the N-acetyl galactosamine and N-acetyl glucosamine residues. Additionally, temperature variations in $^{1}J_{HN}$ values were monitored and utilized, for the first time, as NMR evidence for H-bonding in conjunction with longer MD simulations [95]. More recent reports also describe a persistent H-bond in the Arixtra pentasaccharide [104,105]. This H-bond was inferred from temperature coefficients and differential exchange rate for H-bonded protons. However, the transient nature of these H-bonds meant that direct detection of scalar coupling through H-bonds remained elusive [104]. The direct detection of $^{3h}J_{NC'}$ and $^{2h}J_{NN}$, through H-bonds, has been reported for nucleic acids [39] and proteins [106], respectively.

H-bonds had not been directly detected for glycans in water until we recently demonstrated their presence in a tetramer of α2-8-linked sialic acid via a direct non-zero value of $^{3h}J_{^{15}N-^{13}C}$ [57]. In this report we showed that three-dimensional models could be built based on H-bonds used in conjunction with heteronuclear coupling constants and NOE restraints, the H-bonds provided previously unavailable information showing that the tetramer formed a helix with two sialic acid residues per turn. When this structure was extended *in silico*, the requirement for decamer binding to its cognate antibody could be explained. The NMR results highlight the importance of working with fully labeled glycans, as they provide unambiguous resonance assignments and the possibility to utilize state-of-the-art NMR experiments to discover new paradigms in oligosaccharide structure [33]. For these glycans studies, it was essential that the NMR data were measured in H_2O to preserve exchangeable protons that provide a wealth of additional information [66,107,108]. Moreover, since the rate of exchange between these protons and the solvent is pH dependent we suggest using buffers for structural studies to avoid large pH changes. However, in some cases buffers may increase the exchange rate of hydroxyl protons and preclude their observation.

Recently, we discovered that hydroxyl protons (OH) of glycans in natural abundance could be readily assigned via HSQC-TOCSY [61]. This strategy enabled us to unambiguously assign OH signals over a wide range of temperatures (−10 to +25 °C) and sample concentrations (50–800 mM), utilizing sucrose and trehalose as model systems. Previously, glycans' OH exchange rate had to be reduced by lowering the sample's temperature in order to enable their assignment [66,107], utilizing high sample concentration [109] or resorting to aprotic or mixed solvents [110–113]. Our recent discovery shows that these conditions are not required for OH detection. Utilizing HSQC-TOCSY we assigned eight common eukaryotic glycans, in α and β configurations (Fig. 4, Table 1). The observed ^{1}H chemical shifts from OH signals ranges from 5.8 to 8.0 ppm. The primary OH groups are more shielded than secondary and tertiary hydroxyl groups. The anomeric hydroxyls are more de-shielded than any other hydroxyl group in a glycan. Similar to anomeric protons, equatorial OH groups are de-shielded compared to axial OH groups by about 0.2 to 0.8 ppm. The same is true for other axial OH protons at other positions in pyranose rings, which are more shielded than their equatorial counterparts, at same position in the ring (compare for instance OH4 for Glc and Gal with the other glycans in Table 1). Observation and assignment of OH increases by twofold the number of glycan ^{1}H resonances available, thereby allowing the OH resonances to be used to increase

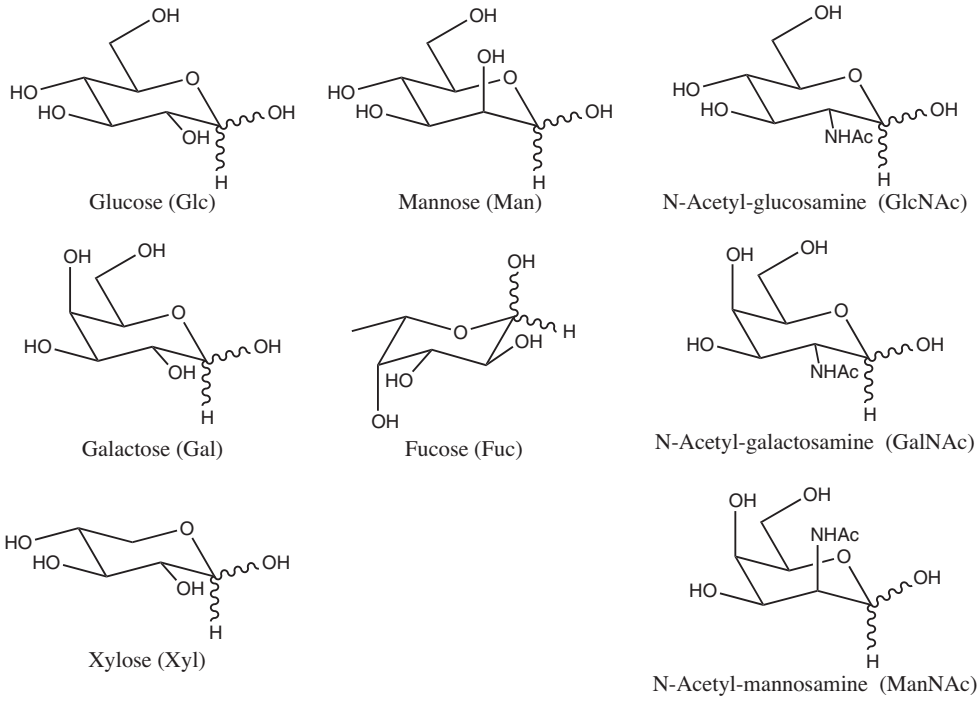


Fig. 4. Glycan chemical structure utilized for OH signal assignment.

resolution and to assist in measuring NOEs and *J-coupling* values as well. Examples of HSQC-TOCSY experiments for OH ^{1}H assignments are shown in Fig. 5. These spectra highlight the distinct hydroxyl fingerprint obtained by epimerizing glucose at C4; the effect of converting an equatorial OH at position 4 to an axial OH is clearly visible.

With the OH assignments given in Table 1, we explored the conditions for their optimal detection (Figs. 6 and 7). Low temperatures decrease the OH exchange rate with water, which in turn facilitates OH observation [66]. Therefore, we decided to perform experiments at 268 K and monitored OH peak intensities as a function of pH (Fig. 6) and salt concentration (with NaCl and $CaCl_2$ from 0 to 150 mM, Fig. 7). For the tested monosaccharides, the observation of ring OH protons, with the exception of anomeric OH, is optimal at a pH range between 6.0 and 6.5, and at low salt concentration as shown in Figs. 6 and 7, respectively. However, anomeric OH protons are more readily observed at lower pH (5.5–6.0).

For all glycans we invariably found that OH signals become less intense with increasing NaCl concentration. Interestingly, the effect that ions increase the OH exchange rates has been previously reported [113,114]. The same general effect was observed with $CaCl_2$. However, Ca^{2+} appears to directly interact with the glycan, for GlcNAc in particular, as evidenced by chemical shift perturbations of hydroxyl signals depending on their atom position. Additionally, Ca^{2+}-induced perturbations were found to be dependent on the anomeric configuration, to the degree that addition of $CaCl_2$, shifted the α/β anomer equilibrium so that the β anomer was favored by 15% over the α when compared to GlcNAc in absence of $CaCl_2$. This finding indicates that Ca^{2+} can selectively interact with GlcNAc and that it may play an important role in modulating glycan conformations in solution.

Optimizing OH detection is the first step towards utilizing these signals as structural probes. Hydroxyl groups provide additional dispersion in the proton dimension that can be used for resonance assignment when aliphatic proton signals overlap. Moreover, OH groups can be used as additional sources of NOE/ROE data. As exemplified for lactose in Fig. 8, hydroxyl signals can yield NOEs to aliphatic protons. Interestingly, we observed NOEs for α-lactose that we did not observe for β-lactose and vice versa, e.g. $OH2_{ter}$, αH5; $OH2_{ter}$, βH2; $OH2_{ter}$, βH3; βH3, $C5_{ter}$; and βH3, $C6_{ter}$ (Fig. 8). This result indicates that the presence of either an axial or equatorial hydroxyl group at the anomeric position may lead to different 3D structures in oligosaccharides, which is consistent with the observation of two sets of chemical shifts (for α-lactose and β-lactose, respectively) [115]. Hydroxyl groups can also provide additional coupling constant and RDC values for structural determination [109], and more importantly, can be used to detect intra- and/or inter-molecular H-bonds to infer preferred glycan conformations [61,107,108]. Therefore, by providing additional dispersion and being the source of previously unavailable data, the utilization of exchangeable protons for glycan structural studies may play a pivotal role in relating the carbohydrate's structure to its function.

Detection of hydroxyl groups enabled us to delineate hydrogen bonding for sucrose in aqueous solution through *J*-coupling (as opposed to by ROE [66,116]) under more dilute conditions than previously reported. We found that by increasing the mixing time in an HSQC-TOCSY experiment, from 10 ms, needed for OH assignment, to 30 ms, we were able to directly detect intramolecular H-bonds implied by X-ray studies. These H-bonds were detected at lower concentrations than previously reported (200 mM). An additional H-bond was determined to arise from intermolecular interaction which is consistent with sucrose dimerization at concentrations ca. 300 mM [61]. The signal to noise of the cross peak for the intramolecular H-bond correlation peak was about 100-fold lower than that for the OH peak at 700 MHz with a conventional room temperature probe. This study will help us better describe the forces that contribute to glycan structural stabilization.

2.4. Coupling constants

Before a detailed discussion on the role of coupling constants in structure determination of glycans structure, we will briefly discuss internal motion within glycans. Coupling constant values

Table 1
Complete ^{1}H, ^{13}C and OH chemical shift assignment for eight common glycans for the α and β pyranosides. Data were collected at 268 K on a Bruker Avance 700 MHz instrument. Chemical shifts are reported using DSS as an internal standard.

	Glc	Gal	Man	Xyl	Fuc	GlcNAc	GalNAc	ManNAc
Hydroxyl ^{1}H chemical shifts (ppm)								
1α	7.20	7.07	7.32	7.07	7.12	7.20	7.12	7.44
2α	6.18	5.98	6.12	6.03	5.94	(HN)8.33	(HN)8.26	(HN)8.259
3α	6.34	5.96	6.08	6.21	5.90	6.30	5.97	6.19
4α	6.34	5.86	6.24	6.22	5.89	6.41	5.97	6.33
6α	5.88	5.97	5.86			5.87	5.95	6.04
1β	7.94	7.84	7.49	7.87	7.76	7.87	7.78	7.87
2β	6.55	6.39	5.89	6.42	6.33	(HN)8.39	(HN)8.33	(HN)8.16
3β	6.48	6.08	6.15	6.36	6.01	6.41	6.08	6.26
4β	6.38	5.88	6.24	6.27	5.89	6.44	6.02	6.33
6β	5.93	5.91	5.92			5.95	6.40	6.01
^{13}C chemical shifts (ppm)								
1α	92.20	92.40	94.16	92.34	96.35	90.93	91.09	93.25
2α	71.60	69.29	70.79	71.61	68.22	54.27	50.43	53.35
3α	72.89	68.47	70.34	72.97	69.38	70.79	67.49	69.04
4α	69.63	69.42	66.92	69.56	72.04	70.09	68.68	66.77
5α	71.42	75.25	72.41	60.88	66.39	71.56	70.61	71.96
6α	60.56	61.36	61.00		15.64	60.58	61.42	60.35
1β	96.02	96.56	93.78	96.73	92.34	95.09	95.57	93.08
2β	74.25	71.99	71.36	74.18	71.82	56.78	53.76	54.27
3β	75.88	71.91	73.17	75.95	73.04	74.09	75.30	72.19
4β	69.67	68.86	66.67	69.35	71.60	69.84	71.30	66.50
5β	76.00	70.54	76.25	65.23	70.98	76.01	67.95	76.35
6β	60.75	61.13	61.05		15.44	60.78	61.17	60.37
Aliphatic ^{1}H chemical shifts (ppm)								
1α	5.23	5.26	5.18	5.08	5.21	5.20	5.23	5.13
2α	3.53	3.79	3.93	3.40	3.76	3.87	4.12	4.34
3α	3.70	3.84	3.84	3.53	3.86	3.77	3.93	4.05
4α	3.40	3.97	3.64	3.48	3.82	3.49	3.99	3.63
5α	3.83	4.08	3.82	3.56	4.21	3.86	4.11	3.87
6α1	3.77	3.76	3.75		1.21	3.81	2.05	2.05
6α2	3.84	3.76	3.88		1.21	3.86		
1β	4.64	4.57	4.90	4.46	4.57	4.71	4.64	5.04
2β	3.23	3.48	3.94	3.10	3.44	3.69	3.89	4.46
3β	3.48	3.64	3.65	3.30	3.65	3.53	3.70	3.84
4β	3.40	3.91	3.56	3.50	3.75	3.47	3.78	3.53
5β	3.46	3.71	3.38	3.20	3.82	3.47	3.93	3.42
5βε				3.80				
6β1	3.72	3.73	3.72		1.25	3.76	2.05	2.10
6β2	3.90	3.73	3.91		1.25	3.92		

obtained from NMR experiments may result from conformational averaging, Bush and co-workers grouped oligosaccharide internal motions into two categories [117]: those belonging to the *first kind*, with internal motion on the ps timescale; and those belonging to the *second kind* with internal motions in the μs–ns time scale. The most common and important are motions of the glycosidic bonds because they change the three-dimensional shape; these motions occur on the μs–ns time scale. Although these types of motions are common, they are difficult to detect because the spectral density functions, and hence the nuclear relaxation rates, are insensitive to motions on this timescale. This is a crucial point because it is the Achilles' heel of carbohydrate solution structure and dynamics determination: motion about the glycosidic linkages is the source of the intrinsic flexibility in many carbohydrates and responsible for their distinct three-dimensional shapes. In order to characterize these shapes, Karplus-like equations were developed. The work was aimed at characterizing torsion angles involving different nuclei including $^{3}J_{H,C}$, $^{3}J_{H,H}$, $^{3}J_{C,C}$, $^{3}J_{H,N}$ in order to gain insight into the conformations adopted around the glycosidic linkage, as well as, N-acetyl and hydroxymethyl groups [98,118–128]. With the same purpose, similar studies have been carried out on selectively ^{13}C labeled samples [97,98]. For an extensive review on coupling constants for carbohydrates and how they relate to torsion angles and their measurement, the reader is referred to Coxon [129]. Since the correlation between coupling constant and dihedral angle was noted [130,131], there have been many successful applications. Hydroxyl groups in glycans can be naturally functionalized or can participate in different glycosidic linkages (α/β 1,2, 1,3, 1,6, etc.). Because these changes in glycans affect the electron distribution in the molecule, it will also affect coupling constants measured. In turn, different parameterizations of Karplus relations, to relate torsion angles to coupling constants, have to be derived. Because new parameterization is needed based on the substitutions on glycans, the use of this type of correlation becomes impractical. Additionally, coupling constants suffer the same averaging problems as NOEs do in carbohydrates. Assuming a Karplus-like equation is accurate in representing a torsion angle, there are still four possible solutions for each coupling constant value. This ambiguity can be partially overcome through the measurement of different scalar coupling constant values involving different sets of atoms that can help define the torsion angle in question, as already suggested [54]. Some conformational space can also be excluded using molecular modeling to find ϕ, ψ regions that do not suffer from steric interactions. However, in many cases, the resulting conformations are not fully determined even when NOEs are considered [115,132]. Labeled carbohydrates are essential to yield easily measured long- and short-range, homo- and hetero-nuclear scalar coupling constants such as, $J_{^{13}C-^{13}C}$, $J_{^{13}C-^{1}H}$, $J_{^{1}H-^{15}N}$

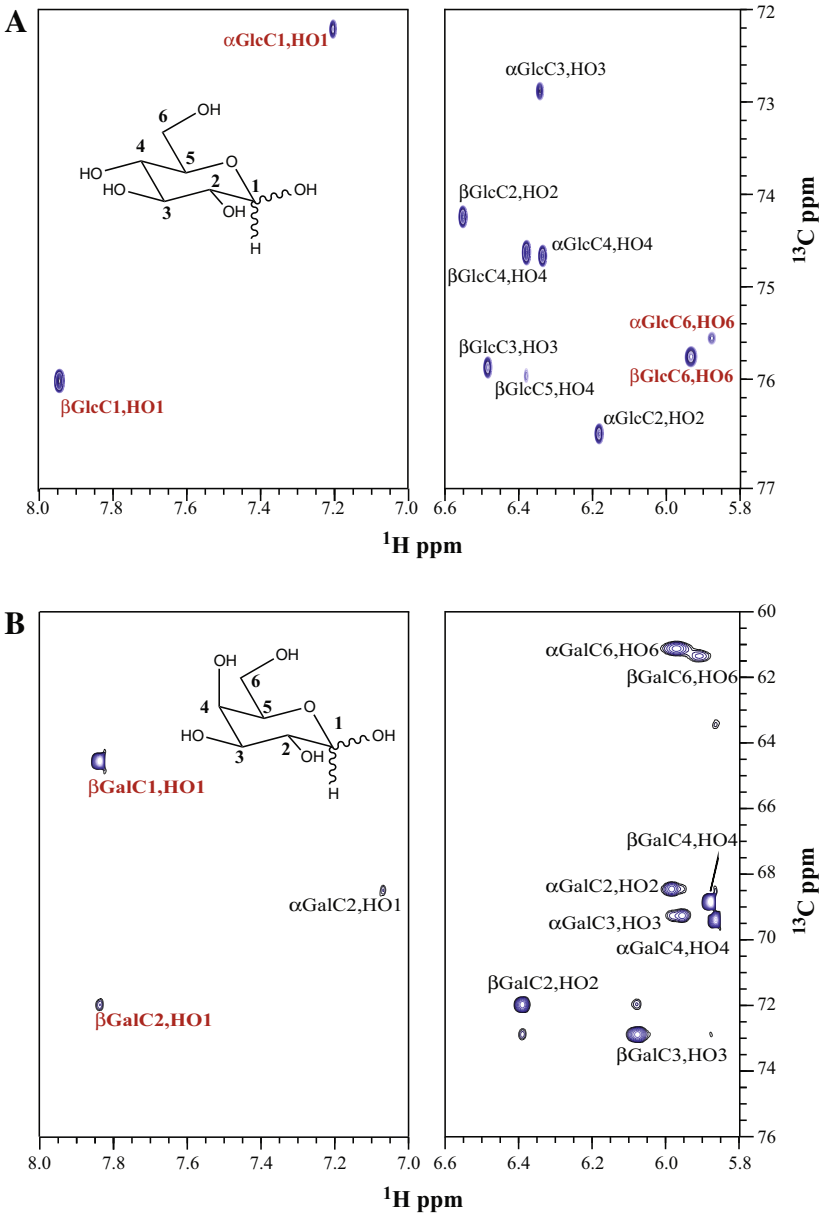


Fig. 5. Hydroxyl proton region of HSQC-TOCSY spectra for glucose (A) and galactose (B). Atom assignments are indicated with black labels for correct chemical shifts whereas bolded labels indicate folded frequencies. The data were collected at 268 K on a Bruker Avance 700 MHz instrument.

and $J_{^{15}N-^{13}C}$. Importantly, these couplings are also helpful in assisting structure determination from RDCs.

Fig. 9 is a compilation of $^{3}J_{H-C-O-C}$ ($^{3}J_{HC}$) and their correlations to angles measured in model compounds. The figure depicts a large variation in $^{3}J_{HC}$ for a given angle, in other words, depending on the parameterization of the Karplus equation, the torsional angle can vary by up to 60° for a single $^{3}J_{HC}$ value. Additionally, the relationship between angle and coupling constant depends on the environment of the coupled nuclei [133], therefore, the residue and identity of the linkage type (e.g., 1 → 4, 1 → 6, etc.) in question has to be incorporated in the parameterization to be able to derive reasonable conformations. Also, chemical substitutions such as sulfate, phosphate, acetyl, or N-acetyl will dictate the parameterization of their own Karplus-like correlations. Consequently, the "quantitative" nature of this type of correlation is brought into question, as pointed out by Karplus himself [133].

Values of $^{3}J_{HC}$ and $^{3}J_{HH}$ coupling constants are routinely used to derive angles (proteins, nucleic acids, small organic compounds). The availability of labeled material facilitates the observation of long-range coupling constants [57,97,134]. Such couplings appear when particular bond geometries are adopted. For example, $^{4}J_{H-H}$ is typically observed when the two long-range-coupled protons, and the three carbon atoms between them, are coplanar and arranged in a "W" geometry (this type of coupling is referred to as "W" coupling) [134–137]. Exploitation of such couplings could further aid in definition of torsions angles in carbohydrates in a Karplus-independent manner.

Coupling constants have often been used to estimate glycosidic torsion angles, however no thorough analysis has been made to determine the accuracy the reported scalar coupling constant values. Without accurate measurement of coupling constants, measurement of RDCs and estimation of torsion angles and Karplus-like equations will be subject to errors. Recent work in our laboratory has uncovered the consequences of strong coupling effects on J values and RDCs [55,56], which will be discussed in detail in the RDC section below.

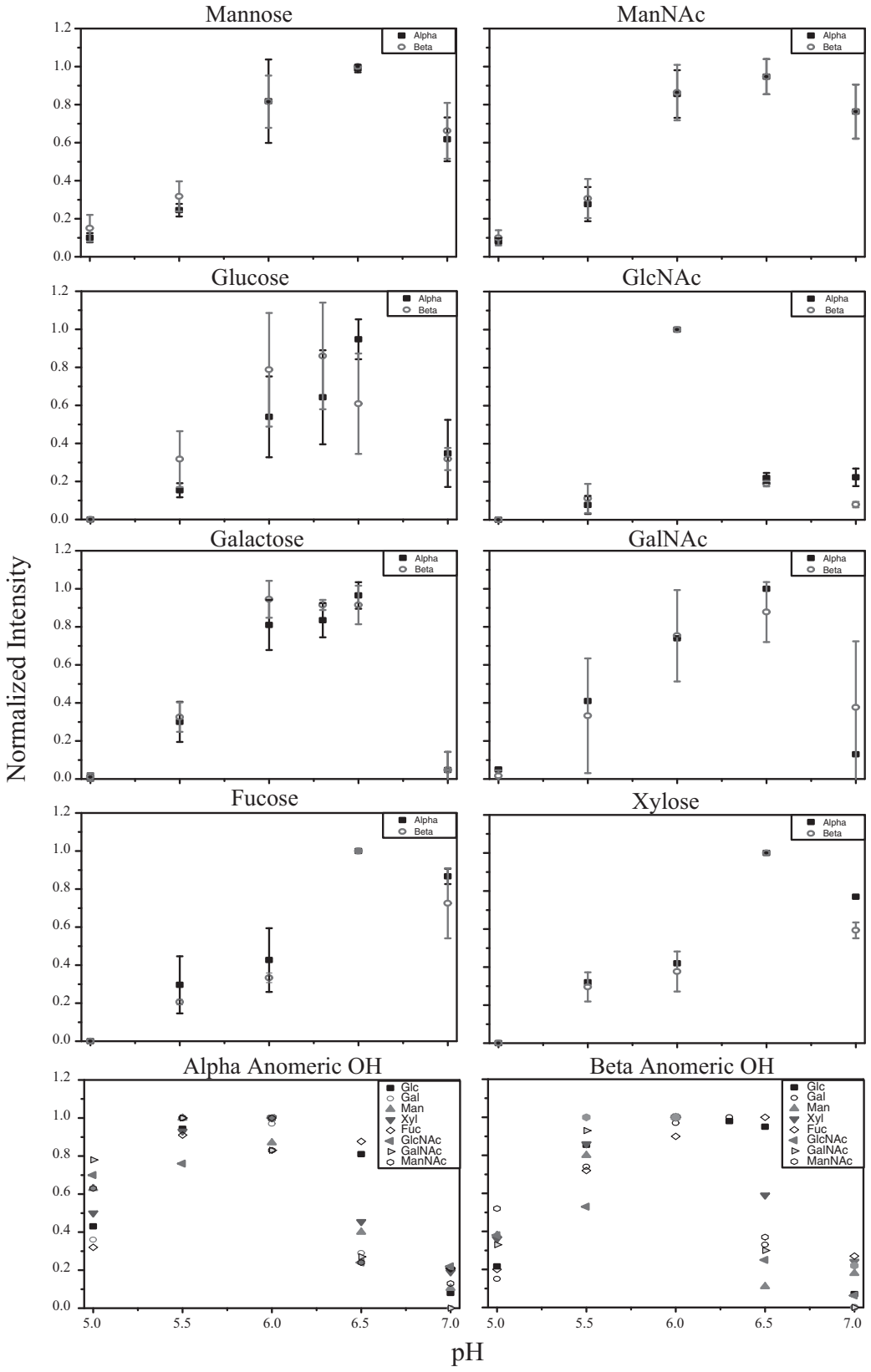


Fig. 6. Optimal pH for OH signal observation for eight common glycans: mannose, N-acetyl mannosamine (ManNAc), glucose, N-acetyl glucosamine (GlcNAc), galactose, N-acetyl galactosamine (GalNAc), fucose and xylose. Each data point represents the average intensity for all hydroxyl ^{1}H signals of each molecule at a given pH, with the exception of the anomeric hydroxyl protons that are plotted separately for all glycans in the bottom panels.

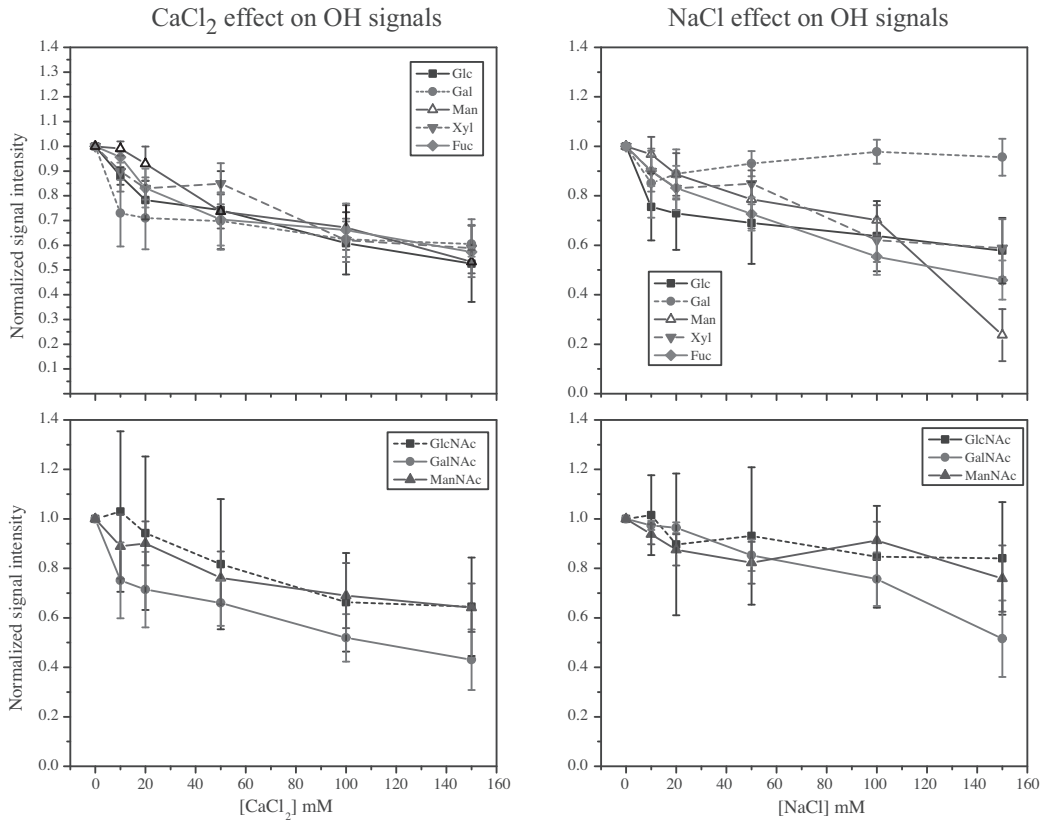


Fig. 7. Salt effect ($CaCl_2$, left panels and NaCl, right panels) on hydroxyl signal observation for the same glycans as Fig. 4. Each data point represents the average intensity of all hydroxyl signals for each molecule, with its respective standard deviation.

PS conformational properties usually depend on molecular size: the larger the molecular size the more flexible the PS becomes and the more difficult it is to characterize. A disaccharide with small variations in ϕ and ψ torsions can be regarded as relatively rigid with well-defined orientations for the sugar rings and thus, defined conformations. However, as the number of units increases the relative orientations of sugar units become more uncertain, even when only two values are possible for each glycosidic torsion angle. When the energy differences between well-defined structures are small, the PS may adopt alternative glycosidic torsions, leading to conformational averaging. Under these circumstances the use of short range or local NMR parameters such as NOEs to define distances between protons and three-bond *J*-couplings to determine dihedral angles become insufficient and they need to be interpreted in conjunction with molecular dynamic simulations.

3. Residual dipolar couplings (RDCs)

As discussed in Section 2.2, the information provided by NOEs and trans-glycosidic *J*-couplings is only short-range or local because the correlated nuclei have to be within a distance of 5 Å (NOEs) or separated by three bonds for torsional angle definition (so called long-range NOEs in proteins refers to amino acid sequence separation). On the contrary, RDCs provide global information, defining the direction of a vector joining two nuclei (either forming a bond or through-space in some cases) relative to a frame defined by global properties of the molecule. The theory for characterizing aligned molecules by NMR was mature by the 1970s [138]. However, the use of the technique for structure and dynamics studies of biomolecules only became widespread with the introduction of weakly aligning media in 1997 [139], because of the advantage of keeping NMR spectra with the same desired characteristics of solution experiments (high resolution and long relaxation times). Early on, RDCs started to be used for glycan 3D structure and dynamics determination [140–142].

As stated earlier, the torsional angles ϕ, ψ, and ω, if applicable, define relative ring orientations, which are important for characterizing oligo- and PS structures. Besides using *J*-couplings, ring orientations may be determined by measuring RDCs, an independent set of NMR observables from NOEs and *J-couplings*. As powerful as the latter are for protein and nucleic acid structure, RDCs are even more important for characterizing glycan structures. The technique has been extensively used for studying both structure, including stereochemistry [143], and dynamics of biomolecules, thus we point the reader to general reviews on the subject available in the literature for detailed descriptions [144–146]. In addition, a recent review on RDCs covers specific issues related to carbohydrates [147]. Here we present a brief account of the latest developments on RDC determinations and analysis, which are further relevant for the carbohydrate research community.

3.1. RDC theory

The dipolar interaction between two nuclear magnetic moments (i, j) in a high magnetic field is described by the truncated dipolar Hamiltonian [148]

$$H_{ij}^{D}(t) = -\left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_i\gamma_j h}{2\pi^2 r_{ij}^3(t)} I_{iz} I_{jz} \frac{(3\cos^2\Theta_{ij}(t) - 1)}{2} \quad (6)$$

where r_{ij} is the distance between the two nuclei, γ_i and γ_j are the magnetogyric ratios of the two spins, $I_{i/jz}$ are the angular momentum spin operators, h is Planck's constant, μ_0 the permittivity of free

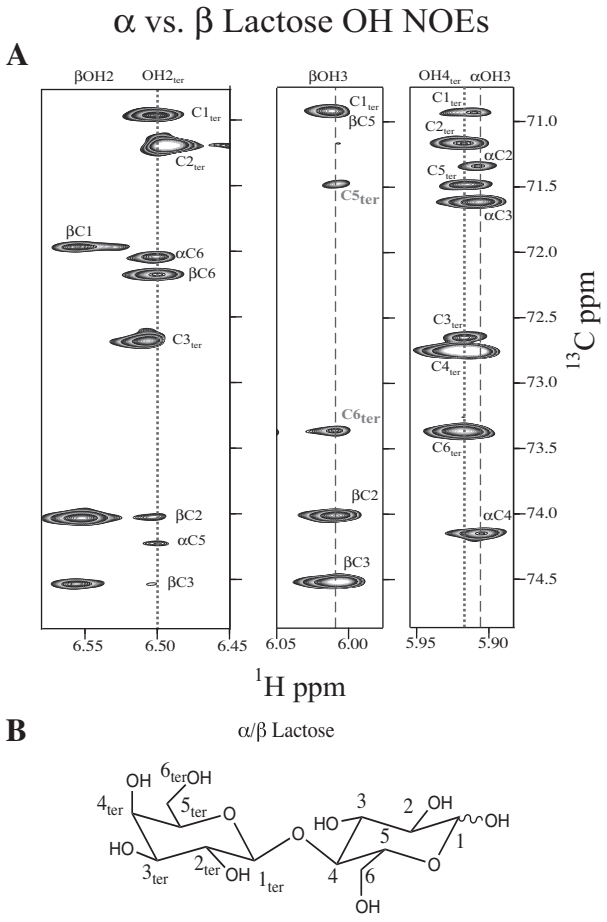


Fig. 8. Strips from HSQC-NOESY spectrum showing NOEs from OH to HC. This figure highlights the use of OHs as structural probes for glycans and potentially other OH bearing molecules. (A) The spectrum was collected at 268 K on 300 mM lactose sample on a Bruker Avance 700 MHz instrument. (B) Lactose chemical structures with numbers indicating atom positions for the reducing and terminal ring.

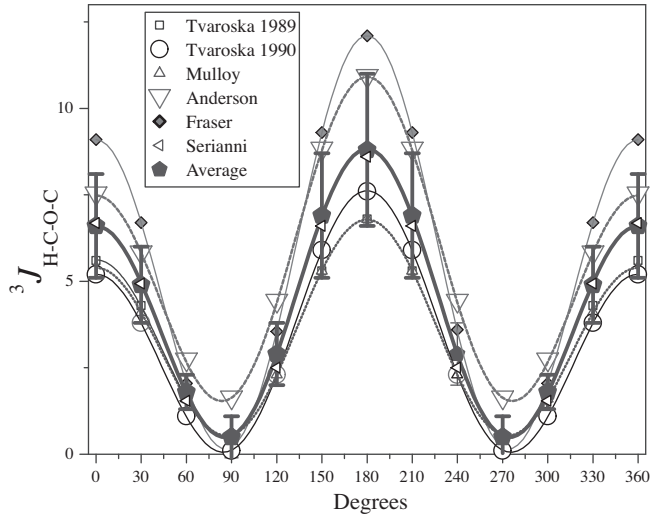


Fig. 9. Karplus-like relations between torsional angles and $^{3}J_{C-O-C-H}$. Each curve was generated utilizing the different parameterized Karplus equations: Tvaroska et al. [121] (□), Tvaroska et al. [119] (○), Mulloy et al. [118] (△), Anderson et al. [120] (▽), Serianni et al. [173] (◁), and the average value for all curves with its corresponding standard deviation (□).

space, and Θ_{ij} is the angle made between the (ij) inter-nuclear vector and the static magnetic field $\mathbf{B_0}$. In solution, molecules normally tumble unconstrained, and because there is no preferred orientation, the time average of $\cos^2\Theta_{ij}$, $\overline{\cos^2\Theta_{ij}} = 1/3$; hence the dipolar interaction described by Eq. (6) is averaged to zero. However, by using media with anisotropic properties such as bicelles [149,150], strained gels [151], or charged viruses [152], the solutes tend to align, taking on preferred orientations due to steric and/or electrostatic interactions that are dependent on properties of the molecule, such as shape, moment of inertia and surface charge distribution. This partial alignment translates into observable through-space dipolar couplings that can be quantified to correlate individual $\mathbf{r}_{ij}$ vector orientations relative to a molecular fixed frame.

The observation of a 'residual' dipolar coupling means that the time average of Eq. (6) for a molecule in solution is not zero, and that the inter-nuclear vector considered has, on average, a preferred orientation relative to the external field ($\overline{D}_{ij} = \langle H^D_{ij}(t)\rangle \neq 0$). If we define the molecular axes as fixed, it is the magnetic field that appears to tumble. From the point of view of the molecule (to which we fix a frame of reference), the probability of finding the external magnetic field (or the average time the field points in a particular molecular direction) depends on the direction we consider [145]. For simplicity, we define the z-direction in Eq. (6) as the direction with maximum probability of being coincident with $\mathbf{B_0}$. We can predict from Eq. (6) that vectors parallel to z ($\Theta = 0$, 180) will have a maximum RDC value. Analogously, we can define the x-direction as the one orthogonal to z with lowest probability to find $\mathbf{B_0}$, and the orthogonal direction to both x and z, as the y-axis. In this frame of reference the probability of finding $\mathbf{B_0}$ in any given direction would be represented by a diagonal tensor $\mathbf{P}$ with principal values (P_{xx}, P_{yy}, P_{zz}), thus this frame of reference is called Principal Axis System (PAS). Normalization requires that the sum of probabilities along all directions be 100%, i.e. $P_{xx} + P_{yy} + P_{zz} = 1$.

Introduction of the probability tensor $\mathbf{P}$ helps to understand the underlying physics. However, it is convenient to calculate RDCs in terms of the traceless alignment tensor $\mathbf{A}$, or the *Saupe matrix* (or *order matrix*) $\mathbf{S}$ [153], which are related by [145]:

$$\mathbf{A} = \mathbf{P} - 1/3\ \mathbf{1};\quad \mathbf{S} = 3/2\ \mathbf{A} = 3/2\mathbf{P} - 1/2\ \mathbf{1} \tag{7}$$

With these definitions, the time average of Eq. (6) under weak alignment conditions results in [145,146,154]:

$$D_{ij}(\theta,\phi) = D_a[(3\cos^2\theta - 1) + \eta\sin^2\theta\cos 2\phi];\ \ D_a = -\left(\frac{\mu_0}{4\pi}\right)\frac{\gamma_i\gamma_j h}{2\pi^2 r_{ij}^3}S_{zz} \tag{8}$$

in which (θ, ϕ) are the polar angles of the $\mathbf{r}_{ij}$ vector in the PAS, and η (called the asymmetry factor) is defined as $(S_{xx} - S_{yy})/S_{zz}$, with the convention for axes definitions that $|S_{zz}| \geqslant |S_{yy}| \geqslant |S_{xx}|$. Because $\mathbf{S}$ is traceless, the values of η are constrained to take on values between 0 and 1. It should be noted that different but equivalent expressions to Eq. (8) are found in the literature but using the alignment tensor $\mathbf{A}$ instead of $\mathbf{S}$, and rhombicity R instead of η [145]. Eq. (8) takes into account only structural characteristics. Local dynamic effects also can be incorporated as a multiplicative factor with the so called generalized order parameter S, whose effect is to scale down the dipolar coupling with increasing mobility [155]. We discuss this in more detail in the next subsection.

Eq. (8) shows that the RDC value will be greater in magnitude for those bonds closely parallel to the direction of greater order (z), with the maximum theoretical value $D_{\max}$ obtained for $\theta = 0$. The factor $D_a = D_{\max}/2$, is called the magnitude of the RDC tensor [154]. The asymmetry factor η, represents the tendency of the molecule for a preferred direction of alignment perpendicular to the z-axis. A cylinder, for instance, would have no preferred alignment direction so defined, and therefore $\eta = 0$. In most cases,

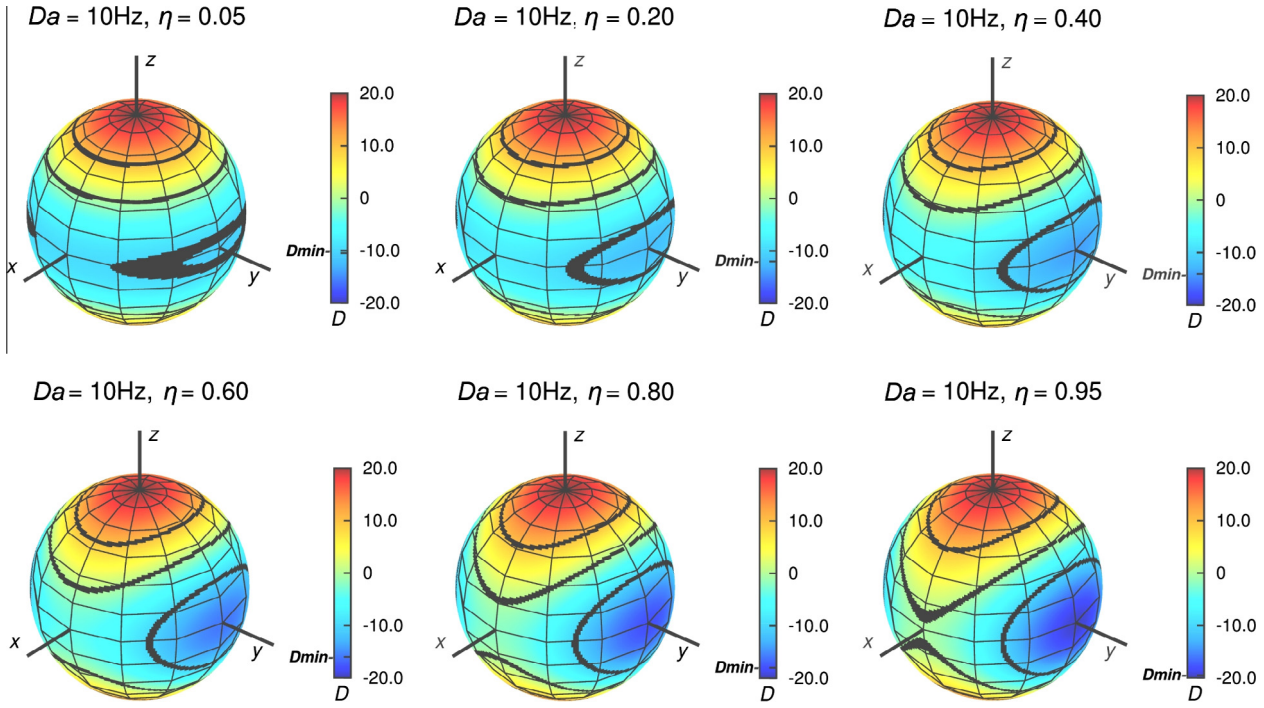


Fig. 10. Color coded representation of the RDC values on the surface of a sphere, derived from Eq. (8), resulting for a fixed D_a value of 10 Hz, and six different values of asymmetry η.

molecules orient asymmetrically and all the axes of the order tensor can be uniquely determined as long as a sufficient number of independent RDCs can be experimentally determined.

To illustrate the behavior of RDCs under different alignment situations, we show in Fig. 10 the values of $D(\theta,\phi)$ color coded for constant D_a and several η values. It can be seen that, when η is large, it is easier to find the orientations for all three axes, which is a relatively common situation for carbohydrates. Notice, however, that there are multiple (θ,ϕ) values consistent with any given D_{ij}. Eq. (8) contains five unknown independent parameters (three angles defining the PAS orientation, plus magnitude and asymmetry. S_{xx}, S_{zz}, S_{xy}, S_{xz}, and S_{yz}, in terms of the Saupe matrix) [156]. Therefore, in the absence of other information, uniquely determining **S** requires measuring RDCs for a set of five independently oriented molecules (i.e. whose PAS are non-parallel), a direct result from Eq. (8). Finding five truly independent molecular alignments is very difficult to achieve experimentally due to limitations on available orienting media. Fortunately we can use our knowledge of molecular structure to constrain the solutions. In fact, many of the early conformational carbohydrate studies using RDC were measured in a single orienting medium [140,142,157–160].

The most common RDC measurement performed for glycans is on C–H bonds, because they are ubiquitous and easy to measure. A nice example of the power of using C–H RDC values in combination with molecular modeling to characterize structures of rigid glycans, is provided by the histo-blood antigens Lewis X (Le^X) and Lewis A (Le^A) [159]. This study also illustrates how molecular alignment reflects on measured RDCs. In the cited study the alignment was achieved with uncharged bicelles that induce molecular alignment through steric forces and the experimental RDC values were well reproduced by (rigid) single models of the glycan antigens. Interestingly, the resulting structures were remarkably similar with an RMSD of 0.145 Å over 20 atoms comprised by the carbons and oxygen atoms of the sugar rings and glycosidic bonds. However, the measured RDC values, although comparable in the case of the GlcNAc residue, were very different for the Fuc and Gal units. Fig. 11 (a modified version of Fig. 7 in Ref. [159]) shows the relationship between structure and alignment tensor that explains the differences observed among C-H RDCs in both cases due to the asymmetry in mass distribution and shape as direct result of the inverted orientations of GlcNAc in each antigen.

3.2. RDCs in oligosaccharides and flexible glycans

Unlike the cases of Le^X and Le^A, single well-defined structures are the exception rather than the rule in the carbohydrate world, thus the measured RDCs are the result of the ensemble average of Eq. (8) over all conformers. Several strategies have been proposed to address these cases and they have been reviewed elsewhere [144,147]. Importantly, the efficacy of all methods relies on the availability of as many RDC measurements as possible. To avoid virtual conformations arising from flexibility in the linkages, it is prudent to fit each rigid residue separately and compare the best-fit alignment tensors. If the alignment tensors of the individual residues match within experimental error, it can be assumed that the linkage has limited motion. However, if the alignment tensors do not match, motion about the glycosidic torsions is implied. As pointed out before, Eq. (8) dictates that at least five independent RDCs are needed for each rigid residue or conformer to determine a 3D structure. This is particularly difficult in carbohydrates because the C-H bond vectors in pyranose rings are nearly parallel or anti-parallel and therefore are not independent. Consequently, it is important to measure as many RDCs as possible. We therefore focused some of our efforts on improving measurements of *J*-couplings and splittings because accurate coupling constants lead to accurate RDCs.

As a larger number of RDCs became available, they played a more important role in conformational studies. RDCs were used to evaluate models generated from MD simulations, as well as from relaxation data [158]. In some cases, NOE data are not compatible with RDC data, mainly because internal molecular motion averages RDCs differently from NOEs. Reconciling these two types of data remains unsolved [161]. Since there is no reliable predictive force

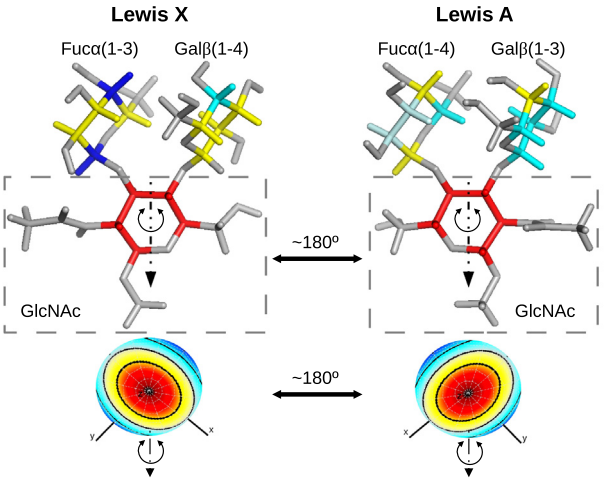


Fig. 11. Relationship between the structures and the alignment tensors for Lewis X and Lewis A antigens when oriented in neutral bicelles. The differences observed among C–H RDCs in both cases is due to the asymmetry in mass distribution and shape as direct result of the inverted orientations of the GlcNAc group in each epitope.

field model for carbohydrates, the glycosidic torsion angles ϕ and ψ of the Lewis antigen of each oligosaccharide were simultaneously scanned over the full range of glycosidic torsions in a four-dimensional grid search over ϕ and ψ with steps of 2.5°. Each rotamer was tested for agreement with NOE and RDC data, and the energy-minimized calculated model for the Lewis antigens. The resulting best-fit single structure model is consistent with a rigid structure [159]. Similarly, Venable et al. found four possible conformations for the five-membered fructose ring in sucrose based on RDC data only, in agreement with previous reports [162]. In addition, there were four energy minima for sucrose. In other studies, RDCs in conformationally flexible molecules were confirmed from MD simulations as in LNF-2, a pentasaccharide containing the Le^X antigen RDCs were complemented with MD simulations because 5 independent C–H RDCs could not be measured for a single residue [140,157].

Since most carbohydrates are flexible, often multiple conformations exist in solution. In such cases, the minimum energy structure may be poorly defined and represented by a broad potential energy well. Characterization of glycans with complex dynamics can be achieved by defining the generalized degree of order (GDO) parameter in terms of Cartesian order tensor elements as follows [142,163],

$$\vartheta = \sqrt{\frac{2}{3}}|\mathbf{S}| \tag{9}$$

Since RDCs are sensitive to motion on all timescales, the GDO is sensitive to dynamic averaging from overall alignment and internal motion effects. With the aid of molecular fragment geometry information (e.g., sugar rings) and RDC values from two or more alignment media for the fragments of interest, it is possible to determine **S** and therefore the GDO for each fragment (ϑ_{frag}). Hence, comparing ratios of $\vartheta_{frag}/\vartheta_{molecule}$ yields information on fragment dynamics. Detailed treatments of more general models can be found in the literature [164,165]. Tian et al. demonstrated the power of this method by studying a flexible mannose trisaccharide [142]. They were able to measure RDCs in two alignment media and define unique Saupe matrix elements for each mannose unit, and compare GDO fragment values for each of them. With these results they could characterize the global dynamic for the oligosaccharide, as well as differentiate the amplitude of motion for each of the residues.

As mentioned in the previous section, if the linkage between two sugar rings is rigid, the order parameters of the two residues should be indistinguishable within experimental error; otherwise, motion around the linkage, results in a different order parameter, which should be observed in the GDOs [142]. Thus, a difference in the GDOs reflects a difference in amplitude of motion between two rigid residues. Because the motion manifested in the GDOs does not always induce nuclear relaxation, RDCs and GDOs can help in more completely delineating motion.

The glycosidic linkage in sucrose was found flexible by fitting a combination of two conformers to experimental RDCs data, yielding a much better fit than a single conformer [162]. Thus, multiple conformations can be determined, though the population of each conformer must remain undetermined. Bush developed a multiple regression analysis technique trying to find possible conformer combination whose calculated RDC was linear combination based on the possible multiple conformers generated by MC (Monte Carlo) or MD simulation. A combination of 5 conformers is the best fit for human milk hexasaccharide LNDFH 1 [166]. Internal motion for this oligosaccharide is proposed because the order parameters for each monosaccharide unit are different. The order parameters are arrived at from a back-calculation of RDCs from the 5 conformers. The authors argue that this method can compensate for the low resolution obtained from MD simulation data, which arises from low energy barriers among carbohydrates conformers. MD simulations cannot reliably predict populations as a broad minimum energy well is a typical feature for multiple conformers with energy barrier estimated lower than 2–4 kCal/mol. It is important to remember however, that each conformation should be represented by at least 5 independent RDCs to avoid an under determined fit to experimental RDCs. Because it is difficult to measure such a large number of independent RDCs in natural isotopic abundance samples, one has to resort to isotopic labeling.

It is also reported that weakly aligned media could change the conformer population ratio or selectively align a particular conformer [167]. Indeed, a reliable method needs to be developed to avoid above possibility.

As stated above, there may be multiple conformations in solution, present in fast-exchange and therefore cannot be directly observed. To avoid complications in data interpretations that arise from different populations in solution in fast exchange, we prefer to measure spectra in isotropic media at the same temperature as those in oriented media.

3.3. Distortion of RDC determinations by strong coupling

All coupled NMR experiments provide splittings due to total spin–spin effects (T^{ss}), therefore RDCs are obtained by taking the difference in couplings between samples in aligned ($T^{ss}_{aniso} = J + D$) and isotropic ($T^{ss}_{iso} = J$) media. Despite the many methods developed to measure T^{ss}, their accuracy is not usually evaluated. The most frequently used method is to measure frequency differences in coupled spectra, e.g., coupled HSQC, where the T^{ss} splittings can be measured either in the ^{1}H or ^{13}C dimension. Splittings are sometimes measured in the ^{1}H dimension of a two-dimensional NMR spectrum because digital resolution in this dimension is typically better than in the ^{13}C dimension without a huge time sacrifice [161]. Other researchers measure T^{ss} in the ^{13}C dimension because ^{1}H–^{1}H coupling was assumed not to impact accuracy of T^{ss}. However we found a large discrepancy (up to 7.0 Hz) between T^{ss} measured by coupled HSQC in the ^{1}H and ^{13}C dimension [55,56]. Because RDCs are calculated as the difference between splittings in oriented media and splittings in isotropic media, the errors in RDCs are the sum of the errors in the isotropic and oriented media. Though 7 Hz is a small fraction of the average

$^1T^{ss}_{CH}$ (ca. 145 Hz) a 7 Hz error is unacceptable for a typical $^1T^{ss}_{CH}$, which typically falls between −15 and 15 Hz for carbohydrates.

Product operator analysis [168] shows that strong ^{1}H–^{1}H coupling leads to appearance of additional observable magnetization [55,56]. This magnetization also gives rise to additional peaks in two-dimensional spectra, sometimes called spurious peaks, affecting the accuracy of frequency differences observed in $^1T^{ss}_{CH}$ coupling. The presence of strong ^{1}H–^{1}H coupling actually alters the eigenfrequencies observed in coupled HSQC spectra, inducing to errors in $^1T_{CH}$ determination in either ^{1}H or ^{13}C dimension.

We also found that long-range $^nT^{ss}_{CH}$ coupling can affect the apparent T_{CH} values determined from splittings in the ^{13}C dimension, while having no effect in the ^{1}H dimension. After fully understanding the effects of ^{1}H–^{1}H strong coupling, we demonstrated the use of full NMR spectrum simulation to yield an accurate $^1J_{CH}$.

In oriented media another problem emerges. Although we can obtain an accurate value of T^{ss} in isotropic media by spectral simulation, it is almost impossible to obtain an accurate value of T^{ss} in weakly aligning media. In isotropic media, spin system for simulation is relatively simple. We only need to include $^3T^{ss}_{CH}$ and $^nT^{ss}_{CH}$, where $^3J_{HH}$ can be measured experimentally; $^2T^{ss}_{CH}$ can be estimated to be between −5 Hz and 10 Hz. But for a spin system in weakly aligned media, many more immeasurable ^{1}H–^{1}H and ^{1}H–^{13}C RDC are introduced, which cannot be estimated accurately. Therefore, it is hard to propose a spin system model that represents the real spin system, without which T^{ss} cannot be accurately determined by spectral simulation. We predicted and demonstrated experimentally that the strong coupling-induced error in T^{ss} can be different for samples in aligned or isotropic media, introducing another source of error in T^{ss} determinations.

To overcome the strong coupling problem, and to avoid the above-mentioned complications in simulations, we developed a new *J*-modulated method, the constant-time INEPT HSQC (CTI-HSQC). This experiment can yield accurate $^1T^{ss}_{CH}$ values in both, isotropic and weakly aligned media [55]. The important feature of this method is to set the INEPT delay to a constant value for all experiments, therefore heteronuclear coupling is selectively modulated during constant-time INEPT but the homonuclear coupling evolution is not affected. As result, the effect of ^{1}H–^{1}H couplings (weak or strong; scalar or dipolar) on peak intensity is identical for all experiments, and ^{1}H–^{1}H couplings do not influence the resulting curve shape. Thus accurate $^1T_{CH}$ values can be determined. This effectively opens the door to probe carbohydrate 3D structure and dynamics using RDCs.

3.4. Accurate determination of $^1J_{CC}$ in ^{13}C enriched glycans

As mentioned earlier, in order to account for possible flexibility, as many RDCs per ring as possible must be measured. One of the values that have not been consistently used in carbohydrates despite its great potential for structural studies is the measurement of one-bond ^{13}C–^{13}C RDC values. One limit of the method is the default requirement of ^{13}C labeling to increase sensitivity, but even when enrichment is available, overlapping signals and strong coupling can hamper the measurement of one-bond ^{13}C–^{13}C couplings. These challenges were discussed in a recent paper and a general

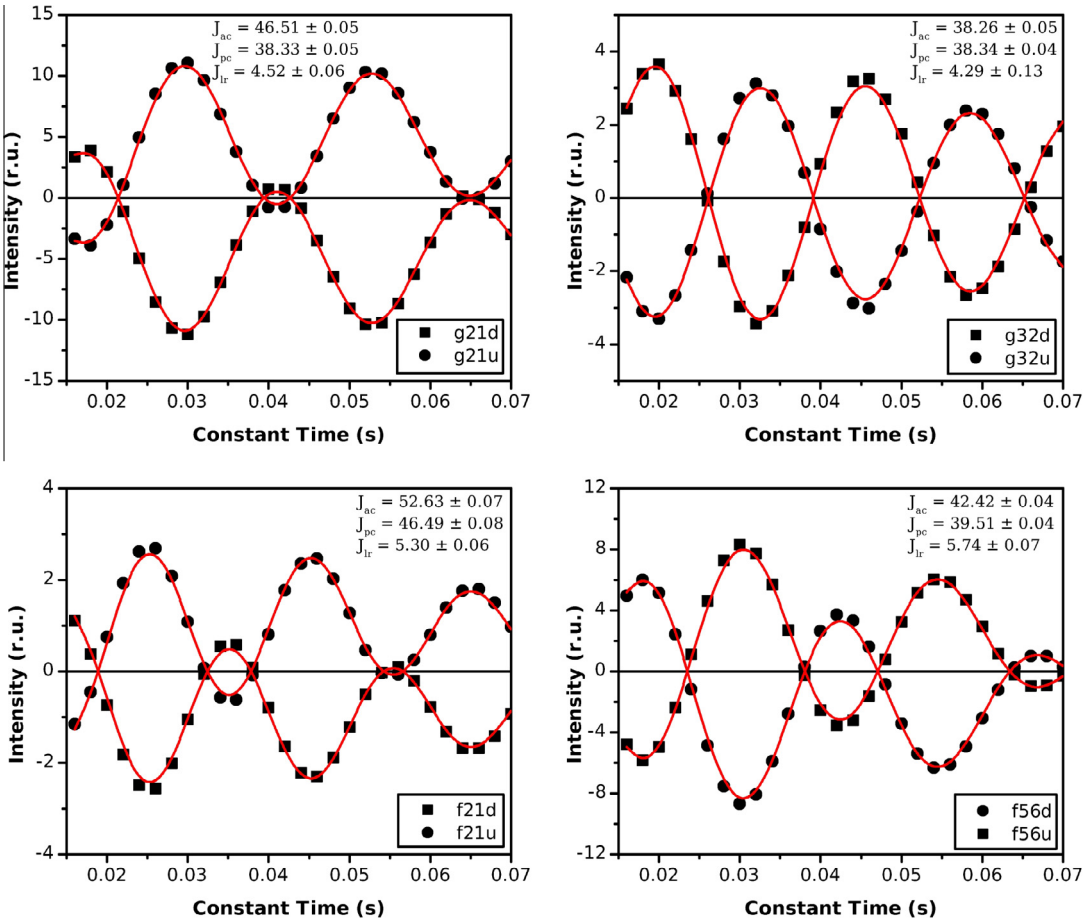


Fig. 12. Representative examples of curve fittings to Eq. (10) for sucrose to obtain $^1J_{CC}$ values from ^{13}C-CT-COSY experiments. Cross-peaks are anti-phase multiplets, from which two intensities are measure: *d* (downfield) or *u* (upfield). The sugar rings are identified by *g* (glucopyranose) and *f* (fructorpyranose).

method proposed that mitigates these problems [169]. In contrast to previous strategies [170–172], the idea is to use the information on active and passive couplings provided by each cross-peak in ^{13}C constant-time COSY experiments using software to filter diagonal peaks and thereby eliminate interference of auto-correlation peaks on cross-peak intensities. Cross-peak intensities can then be approximated by the following expression:

$$I_{pq}(T) \approx I_0 \sin(\pi {}^1T_{pq}^{ac} T) \cos(\pi {}^1T_{pr}^{pc} T) \cos(\pi T^{lr} T) \quad (10)$$

where I_{pq} is the cross-peak intensity as function of the constant-time, I_0 is a proportionality constant, ${}^1T_{pq}^{ac}$ is the one-bond ^{13}C–^{13}C active coupling (*ac*, added for emphasis) between carbon *p* and its neighbor, ${}^1T_{pr}^{pc}$ is the ${}^1T_{CC}$ passive coupling (*pc*) between *p* and its other neighbor, and the last term represents the contribution of all other long-range couplings (*lr*) and any possible transverse relaxation effects on the cross-peak intensity. The strategy is effective in alleviating signal overlap problems. Because the experiment yields cross-peaks that are phased absorptive positive, they do not suffer from the long tails of the dispersive phased cross-peaks typically found in COSY spectra, decreasing signal overlap. In most cases two data points per cross-peak are obtained resulting in increased accuracy derived from multiple coupling determinations for each bond, and potentially allowing determinations not possible by other methods.

Examples of curve fits for sucrose are shown in Fig. 12, where the quality of the approximation by Eq. (10) can be appreciated. Because the experiment is constant-time in the indirect dimension, cross-peaks are effectively decoupled in that dimension, however the signals in the direct dimension are anti-phase doublets or triplets, depending on the presence of passive couplings. Thus, two intensity values are normally obtained from the outer peaks of each multiplet, leading to improved precision on the fitting parameters. In addition, each coupling can potentially be determined from four different cross-peaks, acting as active coupling in two of them and as passive coupling in another two. Overlapping peaks sometimes prevent measuring all potential couplings, but even in this situation, it is usually possible to obtain at least one reliable value, which is not always feasible with alternative methods.

The presence of strong coupling can affect the one-bond coupling values obtained by this method, and can manifest as differences on values obtained for the couplings from different cross-peaks that can be on the order of a few Hz, which may be of concern when determining RDCs. Azurmendi and Freedberg [169] analyzed these effects by simulations of CT-COSY experiments using the software NMR-SIM on a glucose like system (where strong coupling effects are most easily observed), which showed that assuming the orienting media does not change the molecular structure, the presence of strong couplings does not alter the measured ^{13}C–^{13}C RDCs.

4. Future directions

Because glycans can exist in multiple conformations in solution, acquisition of a set of orthogonal NMR observables is preferable to better define the conformational space sampled by a given glycan. Individual NMR observables are insufficient to characterize the 3D shapes of carbohydrates. Inter-residue NOEs/ROEs are usually scarce, or rotation between the glycosidic linkage yield average distances below 5 Å and therefore leads to multiple structures consistent with the data, but cannot always adequately reduce conformational space. Thus, NOEs/ROEs are generally not capable of providing sufficient restraints to define the relative orientation of residues at both sides of the glycosidic linkage. Scalar coupling constants and RDCs on their own are not sufficient either. However, inter-residue H-bonds on their own can significantly reduce the conformational space, yielding a reduced number of conformers that satisfy the constraints. Unfortunately, H-bond detection is difficult because intra-molecular H-bonds are usually competing with H-bonding to water, therefore only a fraction of the molecules in a sample are engaged in detectable H-bonding between residues. Nonetheless, they can be observed under the right conditions. Quantification of H-bonding, either by measurement of through H-bond coupling constant and determination of OH/D exchange as a function of temperature will lead to defining populations of H-bonded species in solution. Presently, direct measurement of these *J*-values is difficult because of large linewidths (~15 Hz). Such measurements will provide an improved view of the conformational space of glycans. A clearer picture of this will have a direct impact on the improvement of force fields for molecular dynamic simulations and 3D structure prediction and structure validation software.

References

[1] R.A. Laine, Information capacity of the carbohydrate code, Pure Appl. Chem. 69 (1997) 1867–1873.
[2] H.J. Gabius, Biological information transfer beyond the genetic code: the sugar code, Naturwissenschaften 87 (2000) 108–121.
[3] R.A. Laine, The information-storing potential of the sugar code, in: Glycosciences, Wiley-VCH Verlag GmbH, 2008, pp. 1–14.
[4] R.A. Laine, A calculation of all possible oligosaccharide isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide: the Isomer Barrier to development of single-method saccharide sequencing or synthesis systems, Glycobiology 4 (1994) 759–767.
[5] R.S. Haltiwanger, J.B. Lowe, Role of glycosylation in development, Annu. Rev. Biochem. 73 (2004) 491–537.
[6] U. Hacker, K. Nybakken, N. Perrimon, Heparan sulphate proteoglycans: the sweet side of development, Nat. Rev. Mol. Cell Biol. 6 (2005) 530–541.
[7] X. Ji, H. Gewurz, G.T. Spear, Mannose binding lectin (MBL) and HIV, Mol. Immunol. 42 (2005) 145–152.
[8] L.G. Durrant, P. Noble, I. Spendlove, Immunology in the clinic review series; focus on cancer: glycolipids as targets for tumour immunotherapy, Clin. Exp. Immunol. 167 (2012) 206–215.
[9] E. Meezan, H. Wu, P.H. Black, P.W. Robbins, Comparative studies on the carbohydrate-containing membrane components of normal and virus transformed mouse fibroblasts. II. Separation of glycoproteins and glycopeptides by Sephadex chromatography, Biochemistry 8 (1969) 2518–2524.
[10] J. Heimburg-Molinaro, M. Lum, G. Vijay, M. Jain, A. Almogren, K. Rittenhouse-Olson, Cancer vaccines and carbohydrate epitopes, Vaccine 29 (2011) 8802–8826.
[11] H. Narimatsu, H. Sawaki, A. Kuno, H. Kaji, H. Ito, Y. Ikehara, A strategy for discovery of cancer glyco-biomarkers in serum using newly developed technologies for glycoproteomics, FEBS J. 277 (2010) 95–105.
[12] K. Drickamer, M.E. Taylor, Glycan arrays for functional glycomics, Genome Biol. 3 (2002). 1034-1031-1034-1034.
[13] R.D. Cummings, The repertoire of glycan determinants in the human glycome, Mol. BioSyst. 5 (2009) 1087–1104.
[14] R.M. Schmaltz, S.R. Hanson, C.-H. Wong, Enzymes in the synthesis of glycoconjugates, Chem. Rev. 111 (2011) 4259–4307.
[15] L.-X. Wang, J.V. Lomino, Emerging technologies for making glycan-defined glycoproteins, ACS Chem. Biol. 7 (2012) 110–122.
[16] H. Hellmuth, S. Wittroc, S. Kralj, L. Dijkhuizen, B. Hofer, J. Seibel, Engineering the glucansucrase GTFR enzyme reaction and glycosidic bond specificity: toward tailor-made polymer and oligosaccharide products, Biochemistry 47 (2008) 6678–6684.
[17] C.H. Hsu, S.C. Hung, C.Y. Wu, C.H. Wong, Toward automated oligosaccharide synthesis, Angew. Chem. Int. Ed. 50 (2011) 11872–11923.
[18] L.P. McIntosh, R.H. Griffey, D.C. Muchmore, C.P. Nielson, A.G. Redfield, F.W. Dahlquist, Proton NMR measurements of bacteriophage T4 lysozyme aided by N-15 isotopic labeling-structural and dynamic studies of larger proteins, Proc. Natl. Acad. Sci. USA 84 (1987) 1244–1248.
[19] G. Vriend, A molecular modeling and drug design program, J. Mol. Graph. Model. 8 (1990) 52–56.
[20] M. Wiederstein, M.J. Sippl, ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins, Nucleic Acids Res. 35 (2007) W407–W410.
[21] V.A. Jaravine, F. Cordier, S. Grzesiek, Quantification of H/D isotope effects on protein hydrogen-bonds by (h3)J(NC′) and (1)J(NC′) couplings and peptide group N-15 and C-13′ chemical shifts, J. Biomol. NMR 29 (2004) 309–318.
[22] R.A. Laskowski, J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, J.M. Thornton, AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR, J. Biomol. NMR 8 (1996) 477–486.
[23] Y. Shen, O. Lange, F. Delaglio, P. Rossi, J.M. Aramini, G. Liu, A. Eletsky, Y. Wu, K.K. Singarapu, A. Lemak, A. Ignatchenko, C.H. Arrowsmith, T. Szyperski, G.T.

Montelione, D. Baker, A. Bax, Consistent blind protein structure generation from NMR chemical shift data, Proc. Natl. Acad. Sci. USA 105 (2008) 4685–4690.
[24] D.S. Wishart, B.D. Sykes, F.M. Richards, The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy, Biochemistry 3 (1992) 1647–1651.
[25] Y. Shen, R. Vernon, D. Baker, A. Bax, De novo protein structure generation from incomplete chemical shift assignments, J. Biomol. NMR 43 (2009) 63–78.
[26] K.T. Simons, R. Bonneau, I. Ruczinski, D. Baker, Ab initio protein structure prediction of CASP III targets using ROSETTA, Proteins: Struct. Funct. Genet. (1999) 171–176.
[27] J.P. Linge, S.I. O'Donoghue, M. Nilges, Automated assignment of ambiguous nuclear Overhauser effects with ARIA, Methods Enzymol. 339 (2001) 71–90.
[28] J.P. Linge, M. Habeck, W. Rieping, M. Nilges, ARIA: automated NOE assignment and NMR structure calculation, Bioinformatics 19 (2003) 315–316.
[29] M.K. Rosen, K.H. Gardner, R.C. Willis, W.E. Parris, T. Pawson, L.E. Kay, Selective methyl group protonation of perdeuterated proteins, J. Mol. Biol. 263 (1996) 627–636.
[30] M. Billeter, G. Wagner, K. Wüthrich, Solution NMR structure determination of proteins revisited, J. Biomol. NMR 42 (2008) 155–158.
[31] L. Kay, Dynamic regulation of archeal proteasome gate opening as studied by TROSY NMR, Science 328 (2010) 98–102.
[32] D.M. Korzhnev, L.E. Kay, Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: an application to protein folding, Acc. Chem. Res. 41 (2008) 442–451.
[33] V. Chevelkov, Y. Xue, D.K. Rao, J.D. Forman-Kay, N.R. Skrynnikov, N-15(H/D)-SOLEXSY experiment for accurate measurement of amide solvent exchange rates: application to denatured drkN SH3, J. Biomol. NMR 46 (2010) 227–244.
[34] A.J. Baldwin, L.E. Kay, NMR spectroscopy brings invisible protein states into focus, Nat. Chem. Biol. 5 (2009) 808–814.
[35] B. Fürtig, C. Richter, J. Whönert, H. Schwalbe, NMR spectroscopy of RNA, ChemBioChem 4 (2003) 936–962.
[36] M.P. Latham, D.J. Brown, S.A. McCallum, P. Arthur, NMR methods for studying the structure and dynamics of RNA, ChemBioChem 6 (2005) 1492–1505.
[37] V. D'Souza, A. Dey, D. Habib, M.F. Summers, NMR structure of the 101-nucleotide core encapsidation signal of the Moloney murine leukemia virus, J. Mol. Biol. 337 (2004) 427–442.
[38] S.J. Glaser, H. Schwalbe, J.P. Marino, C. Griesinger, Directed TOCSY, a method for selection of directed correlations by optimal combinations of isotropic and longitudinal mixing, J. Magn. Reson., Ser. B 112 (1996) 160–180.
[39] A.J. Dingley, S. Grzesiek, Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 2JNN couplings, J. Am. Chem. Soc. 120 (1998) 8293–8297.
[40] J.P. Marino, H. Schwalbe, C. Griesinger, J-coupling restraints in RNA structure determination, Acc. Chem. Res. 32 (1999) 614–623.
[41] V.S.R. Rao, P.K. Qasba, P.V. Balaji, R. Chandrasekaran, Conformation of Carbohydrates, Harwood Academic Publishers, Amsterdam, 1998.
[42] T. Suzuki, T. Sota, Improving ab initio infrared spectra of glucose–water complexes by considering explicit intermolecular hydrogen bonds, J. Chem. Phys. 119 (2003) 10133–10137.
[43] M.L. DeMarco, R.J. Woods, Structural glycobiology: a game of snakes and ladders, Glycobiology 18 (2008) 426–440.
[44] B.L. Foley, M.B. Tessier, R.J. Woods, Carbohydrate force fields 2 (2012) 692–697.
[45] E. Fadda, R.J. Woods, Molecular simulations of carbohydrates and protein–carbohydrate interactions: motivation, issues and prospects, Drug Discov. Today 15 (2010) 596–609.
[46] E.P. Raman, O. Guvench, A.D. MacKerell, CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses, J. Phys. Chem. B 114 (2010) 12981–12994.
[47] R. Eklund, G. Widmalm, Molecular dynamics simulations of an oligosaccharide using a force field modified for carbohydrates, Carbohydr. Res. 338 (2003) 393–398.
[48] C. Landersjö, G. Widmalm, Solution structure of a pentasaccharide representing the repeating unit of the O-antigen polysaccharide from *Escherichia coli* O142: NMR spectroscopy and molecular simulation studies, Biopolymers 64 (2002) 283–291.
[49] E. Säwen, T. Massad, C. Landersjo, P. Damberg, G. Widmalm, Population distribution of flexible molecules from maximum entropy analysis using different priors as background information: application to the ϕ,ψ-conformational space of the α-(1→2)-linked mannose disaccharide present in N- and O-linked glycoproteins, Org. Biomol. Chem. 8 (2010) 3684–3695.
[50] J.H. Prestegard, X.B. Yi, Structure and dynamics of carbohydrates using residual dipolar couplings, in: J.F.G. Vliegenthart, R.J. Woods (Eds.), NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances, 2006, pp. 40–59.
[51] B. Bendiak, T.T. Fang, D.N.M. Jones, An effective strategy for structural elucidation of oligosaccharides through NMR spectroscopy combined with peracetylation using doubly C-13-labeled acetyl groups, Can. J. Chem./Rev. Can. Chim. 80 (2002) 1032–1050.
[52] D. Jeannerat, Computer optimized spectral aliasing in the indirect dimension of ^{1}H–^{13}C heteronuclear 2D NMR experiments. A new algorithm and examples of applications to small molecules, J. Magn. Reson. 186 (2007) 112–122.
[53] G.B.B. Njock, T.A. Bartholomeusz, M. Foroozandeh, D.E. Pegnyemb, P. Christen, D. Jeannerat, NASCA-HMBC, a new NMR methodology for the resolution of severely overlapping signals: application to the study of agathisflavone, Phytochem. Anal 23 (2012) 126–130.
[54] M. Martin-Pastor, C.A. Bush, New strategy for the conformational analysis of carbohydrates based on NOE and ^{13}C NMR coupling constants. Application to the flexible polysaccharide of *Streptococcus mitis* J22, Biochemistry 38 (1999) 8045–8055.
[55] B. Yu, H. van Ingen, D.I. Freedberg, Constant time INEPT CT-HSQC (CTi-CT-HSQC) – a new NMR method to measure accurate one-bond J and RDCs with strong ^{1}H–^{1}H couplings in natural abundance, J. Magn. Reson. 228 (2013) 159–165.
[56] B. Yu, H. van Ingen, S. Vivekanandan, C. Rademacher, S.E. Norris, D.I. Freedberg, More accurate $^{1}J_{CH}$ coupling measurement in the presence of $^{3}J_{HH}$ strong coupling in natural abundance, J. Magn. Reson. 215 (2012) 10–22.
[57] M.D. Battistel, M. Shangold, L. Trinh, J. Shiloach, D.I. Freedberg, Evidence for helical structure in a tetramer of α2-8 sialic acid: unveiling a structural antigen, J. Am. Chem. Soc. 134 (2012) 10717–10720.
[58] P. Soderman, P.-E. Jansson, G. Widmalm, Synthesis, NMR spectroscopy and conformational studies of the four anomeric methyl glycosides of the trisaccharide D-Glcp-(1→3)-[D-Glcp-(1→4)]-α-D-Glcp, J. Chem. Soc., Perk. Trans. 2 (1998) 639–648.
[59] M. Ohnishi, D.W. Urry, Temperature dependence of amide proton chemical shifts – secondary structure of gramicidin and valinomycin, Biochem. Biophys. Res. Commun. 36 (1969) 194–202.
[60] M. Garcia-Viloca, R. Gelabert, À. González-Lafont, M. Moreno, J.M. Lluch, Temperature dependence of proton NMR chemical shift as a criterion to identify low-barrier hydrogen bonds, J. Am. Chem. Soc. 120 (1998) 10203–10209.
[61] M.D. Battistel, R. Pendrill, G. Widmalm, D.I. Freedberg, Direct evidence for hydrogen bonding in glycans: a combined NMR and molecular dynamics study, J. Phys. Chem. (2013).
[62] M. Lundborg, G. Widmalm, Structural analysis of glycans by NMR chemical shift prediction, Anal. Chem. 83 (2011) 1514–1517.
[63] M. Frank, S. Schloissnig, Bioinformatics and molecular modeling in glycobiology, Cell. Mol. Life Sci. 67 (2010) 2749–2772.
[64] G. Bodenhausen, D.J. Ruben, Natural abundance nitrogen-15 NMR by enhanced heteronuclear spectroscopy, Chem. Phys. Lett. 69 (1980) 185–189.
[65] A. Bax, M.F. Summers, Proton and carbon-13 assignments from sensitivity-enhanced detection of heteronuclear multiple-bond connectivity by 2D multiple quantum NMR, J. Am. Chem. Soc. 108 (1986) 2093–2094.
[66] L. Poppe, H. van Halbeek, NMR spectroscopy of hydroxyl protons in supercooled carbohydrates, Nat. Struct. Biol. 1 (1994) 215–216.
[67] G. Otting, K. Wüthrich, Efficient purging scheme for proton-detected heteronuclear two-dimensional NMR, J. Magn. Reson. 76 (1988) 569–574.
[68] W.P. Aue, E. Bartholdi, R.R. Ernst, 2-Dimensional spectroscopy – application to nuclear magnetic-resonance, J. Chem. Phys. 64 (1976) 2229–2246.
[69] L. Braunschweiler, R.R. Ernst, Coherence transfer by isotropic mixing: application to proton correlation spectroscopy, J. Magn. Reson. 53 (1983) 521–528.
[70] A. Bax, Homonuclear magnetization transfer experiments using isotropic and nonisotropic schemes, Isr. J. Chem. 28 (1988) 309–317.
[71] E.D. Becker, High Resolution NMR Spectroscopy, third ed., 2001.
[72] L. Lerner, A. Bax, Application of new, high-sensitivity ^{1}H–^{13}C-N.M.R.-spectral techniques to the study of oligosaccharides, Carbohydr. Res. 166 (1987) 35–46.
[73] A. Bax, G.M. Clore, P.C. Driscoll, A.M. Gronenborn, M. Ikura, L.E. Kay, Practical aspects of proton–carbon–carbon–proton three-dimensional correlation spectroscopy of ^{13}C-labeled proteins, J. Magn. Reson. 87 (1990) (1969) 620–627.
[74] J. Duus, C.H. Gotfredsen, K. Bock, Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations, Chem. Rev. 100 (2000) 4589–4614.
[75] F.A.L. Anet, A.J.R. Bourn, Nuclear magnetic resonance spectral assignments from nuclear Overhauser effects, J. Am. Chem. Soc. 87 (1965). 5250–&.
[76] D. Neuhaus, M.P. Williamson, The Nuclear Overhauser Effect in Structural and Conformational Analysis, second ed., Wiley, New York, 2000.
[77] J.M. Lee, J. Dadok, A.A. Bothner-By, D. Bennett, CAMELSPIN – transverse nuclear Overhauser effect study of erythronolide A and of lysozyme, Fed. Proc. 45 (1986). 1606.
[78] J. Cavanagh, W. Fairbrother, A. Palmer, M. Rance, N. Skelton, Protein NMR Spectroscopy, second ed., Academic Press, San Diego, 2007.
[79] C. Griesinger, R.R. Ernst, Frequency offset and their elimination in NMR rotating-frame cross-relaxation spectroscopy, J. Magn. Reson. 75 (1987) 261–271.
[80] Q. Teng, Structural Biology: Practical NMR Applications, second ed., Springer, US, 2013.
[81] L. Poppe, H. van Halbeek, The rigidity of sucrose: just an illusion?, J Am. Chem. Soc. 114 (1992) 1092–1094.
[82] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity, J. Am. Chem. Soc. 104 (1982) 4546–4559.
[83] G. Lipari, A. Szábo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, J. Am. Chem. Soc. 104 (1982) 4559–4570.

[84] Q.W. Xu, C.A. Bush, Measurement of long-range carbon–carbon coupling constants in a uniformly enriched complex polysaccharide, Carbohydr. Res. 306 (1998) 335–339.

[85] K. Lycknert, G. Widmalm, Dynamics of the *Escherichia coli* O91 O-antigen polysaccharide in solution as studied by carbon-13 NMR relaxation, Biomacromolecules 5 (2004) 1015–1020.

[86] E.W. Sayers, J.L. Weaver, J.H. Prestegard, Hydrogen bonding geometry of a protein-bound carbohydrate from water exchange-mediated cross-relaxation, J. Biomol. NMR 12 (1998) 209–222.

[87] F. Yu, J.H. Prestegard, Structural monitoring of oligosaccharides through ^{13}C enrichment and NMR observation of acetyl groups, Biophys. J. 91 (2006) 1952–1959.

[88] M.A. Macnaughtan, F. Tian, S. Liu, L. Meng, S. Park, P. Azadi, K.W. Moremen, J.H. Prestegard, ^{13}C-sialic acid labeling of glycans on glycoproteins using ST6Gal-I, J. Am. Chem. Soc. 130 (2008) 11864–11865.

[89] H.F. Azurmendi, J. Vionnet, L. Wrightson, L.B. Trinh, J. Shiloach, D.I. Freedberg, Extracellular structure of polysialic acid explored by on cell solution NMR, Proc. Natl. Acad. Sci. USA 104 (2007) 11557–11561.

[90] B.J. Hare, F. Rise, Y. Aubin, J.H. Prestegard, ^{13}C NMR studies of wheat germ agglutinin interactions with N-acetylglucosamine at a magnetically oriented bilayer surface, Biochemistry 33 (1994) 10137–10148.

[91] B.A. Salvatore, R. Ghose, J.H. Prestegard, NMR Studies of a ^{13}C,^{15}N-labeled GM4-lactam glycolipid at an oriented model-membrane interface, J. Am. Chem. Soc. 118 (1996) 4001–4008.

[92] B.A. Salvatore, J.H. Prestegard, Synthesis of a ^{15}N,^{13}C-labeled lactam analog of a GM4-lactone cell-surface glycolipid, Tetrahedron Lett. 39 (1998) 9319–9322.

[93] A. Barb, D.I. Freedberg, M. Battistel, J. Prestegard, NMR detection and characterization of sialylated glycoproteins and cell surface polysaccharides, J. Biomol. NMR 51 (2011) 163–171.

[94] U. Olsson, E. Sawen, R. Stenutz, G. Widmalm, Conformational flexibility and dynamics of two (1→6)-linked disaccharides related to an oligosaccharide epitope expressed on malignant tumour cells, Chem. Eur. J. 15 (2009) 8886–8894.

[95] S.E. Norris, J. Landström, A. Weintraub, T.E. Bull, G. Widmalm, D.I. Freedberg, Transient hydrogen bonding in uniformly ^{13}C,^{15}N labeled carbohydrates in water, Biopolymers (2011).

[96] J. Wu, A.S. Serianni, Isotope-edited 1D and 2D nmr spectroscopy of ^{13}C-substituted carbohydrates, Carbohydr. Res. 226 (1992) 209–218.

[97] T. Klepach, W.H. Zhang, I. Carmichael, A.S. Serianni, C-13-H-1 and C-13-C-13 NMR J-couplings in C-13-labeled N-acetyl-neuraminic acid: correlations with molecular structure, J. Org. Chem. 73 (2008) 4376–4387.

[98] B. Bose, S. Zhao, R. Stenutz, F. Cloran, P.B. Bondo, G. Bondo, B. Hertz, I. Carmichael, A.S. Serianni, Three-bond C–O–C–C spin-coupling constants in carbohydrates: development of a Karplus relationship, J. Am. Chem. Soc. 120 (1998) 11158–11173.

[99] M. Nilges, G.M. Clore, A.M. Gronenborn, Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms circumventing problems associated with folding, FEBS Lett. 239 (1988) 129–136.

[100] P.C. Wendy, D. Cornell, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz Jr., David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, Peter A. Kollman, A second generation force field for the simulation of proteins, nucleic acids, and organic molecules, J. Am. Chem. Soc. 117 (1995) 5179–5197.

[101] R.J. Woods, R.A. Dwek, C.J. Edge, B. Fraser-Reid, Molecular mechanical and molecular dynamic simulations of glycoproteins and oligosaccharides. 1. GLYCAM_93 parameter development, J. Phys. Chem. 99 (1995) 3832–3846.

[102] B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, CHARMM: a program for macromolecular energy, minimization, and dynamics calculations, J. Comput. Chem. 4 (1983) 187–217.

[103] M.J. Forster, B. Mulloy, Rationalizing nuclear Overhauser effect data for compounds adopting multiple-solution conformations, J. Comput. Chem. 15 (1994) 155–161.

[104] D.J. Langeslay, R.P. Young, S. Beni, C.N. Beecher, L.J. Mueller, C.K. Larive, Sulfamate proton solvent exchange in heparin oligosaccharides: evidence for a persistent hydrogen bond in the antithrombin-binding pentasaccharide Arixtra, Glycobiology 22 (2012) 1173–1182.

[105] C.N. Beecher, R.P. Young, D.J. Langeslay, L.J. Mueller, C.K. Larive, Hydroxyl proton hydrogen bonding in the heparin oligosaccharide arixtra in aqueous solution, J. Phys. Chem. B 118 (2014) 482–491.

[106] Y.X. Wang, J. Jacob, F. Cordier, P. Wingfield, S.J. Stahl, S. Lee-Huang, D. Torchia, S. Grzesiek, A. Bax, Measurement of (3h)J(NC′) connectivities across hydrogen bonds in a 30 kDa protein, J. Biomol. NMR 14 (1999) 181–184.

[107] S.Q. Sheng, H. van Halbeek, Evidence for a transient interresidue hydrogen bond in sucrose in aqueous solution obtained by rotating-frame exchange NMR spectroscopy under supercooled conditions, Biochem. Biophys. Res. Commun. 215 (1995) 504–510.

[108] H. van Halbeek, NMR developments in structural studies of carbohydrates and their complexes, Curr. Opin. Struct. Biol. 4 (1994) 697–709.

[109] R.T. Williamson, B.L. Marquez, W.H. Gerwick, K.E. Kover, One- and two-dimensional gradient-selected HSQMBC NMR experiments for the efficient analysis of long-range heteronuclear coupling constants, Magn. Reson. Chem. 38 (2000) 265–273.

[110] B.N. Craig, M.U. Janssen, B.M. Wickersham, D.M. Rabb, P.S. Chang, D.J. O'Leary, Isotopic perturbation of intramolecular hydrogen bonds in rigid 1,3-diols: NMR studies reveal unusually large equilibrium isotope effects, J. Org. Chem. 61 (1996) 9610–9613.

[111] M. Fierman, A. Nelson, S.I. Khan, M. Barfield, D.J. O'Leary, Scalar coupling across the hydrogen bond in 1,3- and 1,4-diols, Org. Lett. 2 (2000) 2077–2080.

[112] T.E. Vasquez, J.M. Bergset, M.B. Fierman, A. Nelson, J. Roth, S.I. Khan, D.J. O'Leary, Using equilibrium isotope effects to detect intramolecular OH/OH hydrogen bonds: structural and solvent effects, J. Am. Chem. Soc. 124 (2002) 2931–2938.

[113] B. Adams, L. Lerner, Observation of hydroxyl protons of sucrose in aqueous solution: no evidence for persistent intramolecular hydrogen bonds, J. Am. Chem. Soc. 114 (1992) 4827–4829.

[114] E. Liepinsh, G. Otting, Proton exchange rates from amino acid side chains – implications for image contrast, Magn. Reson. Med. 35 (1996) 30–42.

[115] D.I. Freedberg, S.O. Ano, S.E. Norris, R. Venable, Carbohydrate structure from NMR residual dipolar couplings: is there a correlation between lactose's anomeric configuration and its three-dimensional structure?, in: J.F.G. Vliegenthart, R.J. Woods (Eds.), NMR Spectroscopy and Computer Modeling of Carbohydrates – Recent Advances, American Chemical Society, Washington, DC, 2006, pp. 220–234.

[116] S.Q. Sheng, H. van Halbeek, Evidence for a transient interresidue hydrogen-bond in sucrose in aqueous-solution obtained by rotating-frame exchange NMR-spectroscopy under supercooled conditions, Biochem. Biophys. Res. Commun. 215 (1995) 504–510.

[117] M. Martin-Pastor, C.A. Bush, New strategy for the conformational analysis of carbohydrates based on NOE and C-13 NMR coupling constants. Application to the flexible polysaccharide of *Streptococcus mitis* J22, Biochemistry 38 (1999) 8045–8055.

[118] B. Mulloy, Long-range carbon–proton coupling constants: application to conformational studies of oligosaccharides, Carbohydr. Res. 184 (1988) 39–46.

[119] I. Tvaroska, Dependence on saccharide conformation of the one-bond and three-bond carbon–proton coupling constants, Carbohydr. Res. 206 (1990) 55–74.

[120] J.E. Anderson, A.I. ljeh, Eclipsed ground-state conformations for methoxycyclohexanes with adjacent methyl-group substitution. An NMR criterion and molecular mechanics calculations, J. Chem. Soc. – Perkin Trans. 2 (1994) 1965–1967.

[121] I. Tvaroska, M. Hricovini, E. Petrakova, 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 (1989) 359–362.

[122] C.A. Podlasek, W.A. Stripe, I. Carmichael, M.Y. Shang, B. Basu, A.S. Serianni, C-13–H-1 spin-coupling constants in the beta-D-ribofuranosyl ring: effect of ring conformation on coupling magnitudes, J. Am. Chem. Soc. 118 (1996) 1413–1425.

[123] R. Stenutz, I. Carmichael, G. Widmalm, A.S. Serianni, Hydroxymethyl group conformation in saccharides: structural dependencies of $^{2}J_{HH}$, $^{3}J_{HH}$, and $^{1}J_{CH}$ spin–spin coupling constants, J. Org. Chem. 67 (2002) 949–958.

[124] H.Q. Zhao, Q.F. Pan, W.H. Zhang, I. Carmichael, A.S. Serianni, DFT and NMR studies of $^{2}J_{COH}$ $^{3}J_{HCOH}$, and $^{3}J_{CCOH}$ spin-couplings in saccharides: C–O torsional bias and H-bonding in aqueous solution, J. Org. Chem. 72 (2007) 7071–7082.

[125] H.Q. Zhao, I. Carmichael, A.S. Serianni, Oligosaccharide trans-glycoside $^{3}J_{COCC}$ Karplus curves are not equivalent: effect of internal electronegative substituents, J. Org. Chem. 73 (2008) 3255–3257.

[126] W. Zhang, H. Zhao, I. Carmichael, A.S. Serianni, An NMR investigation of putative interresidue H-bonding in methyl alpha-cellobioside in solution, Carbohydr. Res. 344 (2009) 1582–1587.

[127] X. Hu, I. Carmichael, A.S. Serianni, N-acetyl side-chains in saccharides: NMR J-coupling equations sensitive to CH–NH and NH–CO bond conformations in 2-acetamido-2-deoxy-aldohexopyranosyl rings, J. Org. Chem. 75 (2010) 4899–4910.

[128] B. Coxon, A Karplus equation for $^{3}J_{HCCN}$ in amino sugar derivatives, Carbohydr. Res. 342 (2007) 1044–1054.

[129] B. Coxon, Developments in the Karplus equation as they relate to the NMR coupling constants of carbohydrates, in: D. Horton (Ed.), Adv. Carbohydr. Chem. Biochem., 2009, pp. 17–82.

[130] M. Karplus, Theory of proton coupling constants in unsaturated molecules, J. Am. Chem. Soc. 82 (1960) 4431–4432.

[131] M. Karplus, Comments on the signs of proton coupling constants, J. Am. Chem. Soc. 84 (1962) 2458–2460.

[132] C. Hervé du Penhoat, A. Imberty, N. Roques, V. Michon, J. Mentech, G. Descotes, S. Pérez, Conformational behavior of sucrose and its deoxy analogue in water as determined by NMR and molecular modeling, J. Am. Chem. Soc. 113 (1991) 3720–3727.

[133] M. Karplus, Vicinal proton coupling in nuclear magnetic resonance, J. Am. Chem. Soc. 85 (1963) 2870–2871.

[134] Q.F. Pan, T. Klepach, I. Carmichael, M. Reed, A.S. Serianni, $^{4}J_{COCCH}$ and $^{4}J_{CCCCH}$ as probes of exocyclic hydroxymethyl group conformation in saccharides, J. Org. Chem. 70 (2005) 7542–7549.

[135] M. Barfield, Conformation and subsituent dependence of long-range H–H couplings over four bonds, J. Am. Chem. Soc. 93 (1971) 1066–1071.

[136] S. Sternhell, Long range H1–H1 spin–spin coupling in Nuclear Magnetic Resonance Spectroscopy, Rev. Pure Appl. Chem. 14 (1964) 15.

[137] M. Barfield, A.M. Dean, C.J. Fallick, R.J. Spear, S. Sternhell, P.W. Westerman, Conformational dependence and mechanisms for long-range hydrogen–

hydrogen coupling constants over four bonds, J. Am. Chem. Soc. 97 (1975) 1482–1492.
[138] J.W. Emsley, J.C. Lindon, NMR Spectroscopy Using Liquid Crystal Solvents, Pergamon Press, London, 1975.
[139] N. Tjandra, A. Bax, Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium, Science 278 (1997) 1111–1114.
[140] M. Martin-Pastor, C.A. Bush, The use of NMR residual dipolar couplings in aqueous dilute liquid crystalline medium for conformational studies of complex oligosaccharides, Carbohydr. Res. 323 (2000) 147–155.
[141] H. Neubauer, J. Meiler, W. Peti, C. Griesinger, NMR structure of saccharose and raffinose by means of homo- and heteronuclear dipolar couplings, Helv. Chim. Acta 84 (2001) 243–258.
[142] F. Tian, H.M. Al-Hashimi, J.L. Craighead, J.H. Prestegard, Conformational analysis of a flexible oligosaccharide using residual dipolar couplings, J. Am. Chem. Soc. 123 (2001) 485–492.
[143] J. Yan, F. Delaglio, A. Kaerner, A.D. Kline, H. Mo, M.J. Shapiro, T.A. Smitka, G.A. Stephenson, E.R. Zartler, Complete relative stereochemistry of multiple stereocenters using only residual dipolar couplings, J. Am. Chem. Soc. 126 (2004) 5008–5017.
[144] M. Blackledge, Recent progress in the study of biomolecular structure and dynamics in solution from residual dipolar couplings, Prog. Nucl. Magn. Reson. Spectrosc. 46 (2005) 23–61.
[145] F. Kramer, M.V. Deshmukh, H. Kessler, S.J. Glaser, Residual dipolar coupling constants: an elementary derivation of key equations, Concepts Mag. Reson. Part A 21A (2004) 10–21.
[146] K. Ruan, J.R. Tolman, NMR residual dipolar couplings as probes of biomolecular dynamics, Chem. Rev. 106 (2006) 1720–1736.
[147] A. Canales, J. Jimenez-Barbero, M. Martin-Pastor, Review: use of residual dipolar couplings to determine the structure of carbohydrates, Magn. Reson. Chem. 50 (Suppl. 1) (2012) S80–85.
[148] R.R. Ernst, G. Bodenhausen, A. Wokaun, Principles of Nuclear Magnetic Resonance in One and Two Dimensions, Oxford Science Publications, Oxford, UK, 1987.
[149] C.R. Sanders, J.H. Prestegard, Orientation and dynamics of beta-dodecyl glucopyranoside in phospholipid-bilayers by oriented sample NMR and order matrix analysis, J. Am. Chem. Soc. 113 (1991) 1987–1996.
[150] C.R. Sanders, J.P. Schwonek, Characterization of magnetically orientable bilayers in mixtures of dihexanoylphophotidylcholine and dimyristoylphosphotidylcholine by solid-state NMR, Biochemistry 31 (1992) 8898–8905.
[151] R. Tycko, F.J. Blanco, Y. Ishii, Alignment of biopolymers in strained gels: a new way to create detectable dipole–dipole couplings in high-resolution biomolecular NMR, J. Am. Chem. Soc. 122 (2000) 9340–9341.
[152] M.R. Hansen, L. Mueller, A. Pardi, Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions, Nature Struct. Biol. 5 (1998) 1065–1074.
[153] A. Saupe, Recent results in the field of liquid crystals, Angew. Chem. Int. Ed. 7 (1968) 97–112.
[154] A. Bax, G. Kontaxis, N. Tjandra, Dipolar couplings in macromolecular structure determination, in: T.L. James, V. Dötsch, U. Schmitz (Eds.), Methods Enzymol, Academic Press, San Diego, 2001, pp. 127–174.
[155] J.R. Tolman, J.M. Flanagan, M.A. Kennedy, J.H. Prestegard, NMR evidence for slow collective motions in cyanometmyoglobin, Nat. Struct. Biol. 4 (1997) 292–297.
[156] J.R. Tolman, A novel approach to the retrieval of structural and dynamic information from residual dipolar couplings using several oriented media in biomolecular NMR spectroscopy, J. Am. Chem. Soc. 124 (2002) 12020–12030.
[157] G.R. Kiddle, S.W. Homans, Residual dipolar couplings as new conformational restraints in isotropically ^{13}C-enriched oligosaccharides, FEBS Lett. 436 (1998) 128–130.
[158] A. Almond, J. Bunkenborg, T. Franch, C.H. Gotfredsen, J.O. Duus, Comparison of aqueous molecular dynamics with NMR relaxation and residual dipolar couplings favors internal motion in a mannose oligosaccharide, J. Am. Chem. Soc. 123 (2001) 4792–4802.
[159] H.F. Azurmendi, M. Martin-Pastor, C.A. Bush, Conformational studies of Lewis X and Lewis A trisaccharides using NMR residual dipolar couplings, Biopolymers 63 (2002) 89–98.
[160] D.I. Freedberg, An alternative method for pucker determination in carbohydrates from residual dipolar couplings: a solution NMR study of the fructofuranosyl ring of sucrose, J. Am. Chem. Soc. 124 (2002) 2358–2362.
[161] P. Berthault, D. Jeannerat, F. Camerel, F. Alvarez Salgado, Y. Boulard, J.C. Gabriel, H. Desvaux, Dilute liquid crystals used to enhance residual dipolar couplings may alter conformational equilibrium in oligosaccharides, Carbohydr. Res. 338 (2003) 1771–1785.
[162] R.M. Venable, F. Delaglio, S.E. Norris, D.I. Freedberg, The utility of residual dipolar couplings in detecting motion in carbohydrates: application to sucrose, Carbohydr. Res. 340 (2005) 863–874.
[163] J.R. Tolman, H.M. Al-Hashimi, L.E. Kay, J.H. Prestegard, Structural and dynamic analysis of residual dipolar coupling data for proteins, J. Am. Chem. Soc. 123 (2001) 1416–1424.
[164] G. Bouvignies, P. Bernado, M. Blackledge, Protein backbone dynamics from N–HN dipolar couplings in partially aligned systems: a comparison of motional models in the presence of structural noise, J. Magn. Reson. 173 (2005) 328–338.
[165] M. Deschamps, I.D. Campbell, J. Boyd, J. Magn. Reson. 172 (2005) 118.
[166] S. Ganguly, J. Xia, C. Margulis, L. Stanwyck, C.A. Bush, Measuring the magnitude of internal motion in a complex hexasaccharide, Biopolymers 95 (2010) 39–50.
[167] H. Desvaux, J.P. Gabriel, P. Berthault, F. Camerel, First use of a mineral liquid crystal for measurement of residual dipolar couplings of a nonlabeled biomolecule, Angew. Chem. Int. Ed. 40 (2001) 373–376.
[168] L.E. Kay, R.E.D. McClung, A product operator description of AB and ABX spin systems, J. Magn. Reson. 78 (1988) 172–177.
[169] H.F. Azurmendi, D.I. Freedberg, Accurate determinations of one-bond ^{13}C–^{13}C couplings in ^{13}C-labeled carbohydrates, J. Magn. Reson. 228 (2013) 130–135.
[170] F. Tian, P.J. Bolon, J.H. Prestegard, Intensity-based measurement of homonuclear residual dipolar couplings from CT-COSY, J. Am. Chem. Soc. 121 (1999) 7712–7713.
[171] Z. Wu, A. Bax, Measurement of homonuclear proton couplings based on cross-peak nulling in CT-COSY, J. Magn. Reson. 151 (2001) 242–252.
[172] M. Martin-Pastor, A. Canales-Mayordomo, J. Jimenez-Barbero, NMR experiments for the measurement of proton–proton and carbon–carbon residual dipolar couplings in uniformly labelled oligosaccharides, J. Biomol. NMR 26 (2003) 345–353.
[173] F. Cloran, I. Carmichael, A.S. Serianni, Density functional calculations on disaccharide mimics: studies of molecular geometries and trans-*O*-glycosidic $^{3}J_{COCH}$ and $^{3}J_{COCC}$ spin-couplings, J. Am. Chem. Soc. 121 (1999) 9843–9851.

Glossary of abbreviations

COSY: correlation spectroscopy
GalNAc: N-acetyl galactosamine
GDO: generalized degree of order
GlcNAc: N-acetyl glucosamine
H-bond: hydrogen bond
HIV: human immunodeficiency virus
HMBC: hetero-nuclear multiple-bond correlation spectroscopy
HSQC: hetero-nuclear single-quantum correlation spectroscopy
μs: microsecond
ManNAc: N-acetyl mannosamine
MC: Monte Carlo
MD: molecular dynamics
NMR: nuclear magnetic resonance
NOE: nuclear Overhauser effect
ns: nanosecond
PDB: Protein Data Bank
PS: polysaccharide
RDC: residual dipolar coupling
ROE: rotating-frame Overhauser effect
TOCSY: total correlation spectroscopy