Chapter 20

Informing Saccharide Structural NMR Studies with Density Functional Theory Calculations

Thomas Klepach, Hongqiu Zhao, Xiaosong Hu, Wenhui Zhang, Roland Stenutz, Matthew J. Hadad, Ian Carmichael, and Anthony S. Serianni

Abstract

Density functional theory (DFT) is a powerful computational tool to enable structural interpretations of NMR spin–spin coupling constants (*J*-couplings) in saccharides, including the abundant $^{1}H–^{1}H$ (J_{HH}), $^{13}C–^{1}H$ (J_{CH}), and $^{13}C–^{13}C$ (J_{CC}) values that exist for coupling pathways comprised of 1–4 bonds. The multiple hydroxyl groups in saccharides, with their attendant lone-pair orbitals, exert significant effects on *J*-couplings that can be difficult to decipher and quantify without input from theory. Oxygen substituent effects are configurational and conformational in origin (e.g., axial/equatorial orientation of an OH group in an aldopyranosyl ring; C–O bond conformation involving an exocyclic OH group). DFT studies shed light on these effects, and if conducted properly, yield quantitative relationships between a specific *J*-coupling and one or more conformational elements in the target molecule. These relationships assist studies of saccharide structure and conformation in solution, which are often challenged by the presence of conformational averaging. Redundant *J*-couplings, defined as an ensemble of *J*-couplings sensitive to the same conformational element, are particularly helpful when the element is flexible in solution (i.e., samples multiple conformational states on the NMR time scale), provided that algorithms are available to convert redundant *J*-values into meaningful conformational models. If the latter conversion is achievable, the data can serve as a means of testing, validating, and refining theoretical methods like molecular dynamics (MD) simulations, which are currently relied upon heavily to assign conformational models of saccharides in solution despite a paucity of experimental data needed to independently validate the method.

Key words Density functional theory, NMR spectroscopy, Spin–spin coupling constants, Scalar couplings, *J*-couplings, Saccharides, Conformation, Isotopic labeling, J_{CH}, J_{CC}, Carbohydrate structure

1 Introduction

NMR spectroscopy remains the single most important experimental tool to investigate the conformations and dynamics of saccharides in solution at the molecular level. This stature in the hierarchy of available experimental methods results from the multiple NMR parameters that can reveal the thermodynamic and kinetic

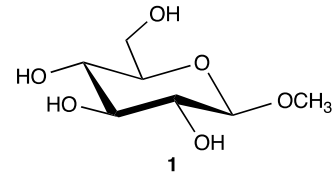
Thomas Lütteke and Martin Frank (eds.), *Glycoinformatics*, Methods in Molecular Biology, vol. 1273,
DOI 10.1007/978-1-4939-2343-4_20,

properties of saccharides, including chemical shifts (δ), indirect nuclear spin–spin coupling constants (J), nuclear relaxation times (T), nuclear Overhauser effects (NOE), residual dipolar couplings (D), chemical shift anisotropy (CSA), and nuclear quadrupolar couplings (NQCC). These parameters can be used to probe molecular behaviors that occur in solution on different time-scales, such as the kinetics of anomerization [1] (milliseconds to seconds), kinetics of side-chain conformational exchange [2] (microseconds to milliseconds), and kinetics of internal *O*-glycosidic bond (linkage) motion [3] (picoseconds to nanoseconds). Despite these prodigious strengths, NMR in the solution state has recognized limitations. A primary limitation is that its measurables are subject to the effects of populational averaging; since NMR samples large ensembles of molecules, its parameters report *averaged* molecular properties. Unlike living cells [4], molecules cannot be synchronized structurally, and thus at any instant in time the many molecules present in solution exist in different structural states and produce a different NMR signature. The resultant signal, that is, the one that is measured, is comprised of individual contributions from each molecule in solution; the exact contribution each makes to the resultant NMR signal depends on factors such as conformer exchange rates and lifetimes. Unless NMR-based methods evolve to permit molecules to be probed individually in solution, this limitation will remain intact for the foreseeable future.

Another limitation of solution NMR, which partly results from its populationally averaged parameters, is the requirement of tools to translate the magnitude of a given NMR parameter into a specific molecular property, such as a molecular torsion angle, internuclear distance, or motion. In many ways, the success of NMR structure determination hinges on the reliability and veracity of this underlying translation, which typically involves at least two components: formulas (equations) that convert the NMR observable (e.g., J-coupling) into a molecular property (e.g., molecular torsion angle), and a model to which these equations are applied (e.g., a three-state model for the rotation about a C–C bond). Herein, we focus on the former problem, and postpone the matter of model selection to future discussion.

While mathematical methods can be used to translate NMR observables, such as nuclear spin-relaxation times [5, 6], into molecular properties, often with inherent and not necessarily validated assumptions, other theoretical tools have evolved to support these efforts. One of these tools is the subject of this review, namely, density functional theory [7] (DFT) and its interface with NMR. Specifically, we address the application of DFT to the molecular interpretation of nuclear (indirect) spin–spin coupling constants (J-couplings), although DFT can also assist in the interpretation of chemical shifts [8], quadrupolar coupling constants [9], and chemical shift anisotropy [10].

Historically, structural analyses of molecules by NMR have focused heavily on the use of three-bond (vicinal) ^{1}H–^{1}H spin–spin coupling constants ($^{3}J_{HH}$) because of their dependencies on H–X–Y–H torsion angles, as originally articulated by Lemieux [11] and Karplus [12]. In saccharides, $^{3}J_{HCCH}$ values play a central role in assigning the preferred conformations of five- (furanose) and six- (pyranose) membered rings in solution. An example of this application is the assignment of a highly preferred $^{4}C_{1}$ (chair) conformation for methyl β-D-glucopyranoside **1** in aqueous solution based on the magnitudes of its intra-ring $^{3}J_{H1,H2}$, $^{3}J_{H2,H3}$, $^{3}J_{H3,H4}$, and $^{3}J_{H4,H5}$ values, all of which are large (≥8 Hz) and consistent with di-axial, vicinally coupled hydrogens (Scheme 1). (For a cogent description of NMR *J*-coupling, *see* ref. 13.) This behavior contrasts with that of α-D-idohexopyranose **2**, where the corresponding $^{3}J_{HCCH}$ values range from 5.0 to 8.1 Hz (Scheme 2) [14]. These $^{3}J_{HCCH}$ suggest that the $^{4}C_{1}$ conformer of **2** cannot be favored in aqueous solution; the observed couplings are considerably larger than expected for di-equatorial, vicinally coupled hydrogens. The observed values suggest a highly preferred $^{1}C_{4}$ or closely related nonplanar form, or a mixture thereof, of **2** in aqueous solution [14].

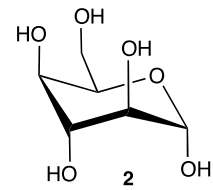


1

(favored $^{4}C_{1}$ conformer)

$^{3}J_{H1,H2}$ = 8.0 Hz
$^{3}J_{H2,H3}$ = 9.5 Hz
$^{3}J_{H3,H4}$ = 9.2 Hz
$^{3}J_{H4,H5}$ = 10.0 Hz

Scheme 1



2

(unfavored $^{4}C_{1}$ conformer)

$^{3}J_{H1,H2}$ = 6.0 Hz
$^{3}J_{H2,H3}$ = 8.1 Hz
$^{3}J_{H3,H4}$ = 7.9 Hz
$^{3}J_{H4,H5}$ = 5.0 Hz

Scheme 2

While $^{3}J_{HCCH}$ values have proven useful in assigning monosaccharide ring conformations in solution, they supply only a fraction of the structural information encoded in *J*-couplings within these structures. Access to this additional, and often redundant, information is helpful when the molecule under investigation is conformationally flexible, which occurs frequently in simple and complex carbohydrates. The greater the molecular flexibility, the greater the potential benefits of accessing the wealth of information available from these additional *J*-couplings. This abundance is illustrated for methyl β-D-glucopyranoside **1**. In **1**, four $^{3}J_{HCCH}$ are available that are sensitive to pyranosyl ring conformation: $^{3}J_{H1,H2}$, $^{3}J_{H2,H3}$, $^{3}J_{H3,H4}$, and $^{3}J_{H4,H5}$ (Scheme 1). Three additional ^{1}H–^{1}H spin-couplings, $^{3}J_{H5,H6R}$, $^{3}J_{H5,H6S}$, and $^{2}J_{H6R,H6S}$, report on the conformation of the exocyclic C6 hydroxymethyl (CH_2OH) fragment (one of these is a two-bond (geminal) $^{2}J_{HCH}$) [15]. Six $^{3}J_{HCOH}$ are also measurable under specific solvent conditions: [16] $^{3}J_{H1,O1H}$, $^{3}J_{H2,O2H}$, $^{3}J_{H3,O3H,}$ $^{3}J_{H4,O4H}$, $^{3}J_{H6R,O6H}$, and $^{3}J_{H6S,O6H}$. These spin-couplings report on the individual H–C–O–H torsions in **1**. Thus, a total of thirteen (13) J_{HH} are available in **1**. In some ring configurations, longer-range (over more than three bonds) ^{1}H-^{1}H spin-couplings (e.g., $^{4}J_{HCCCH}$) can be detected, and their structural and conformational dependencies have been investigated [17]. Long-range $^{4}J_{HCOCH}$ values have also been used as structural constraints in conformational analysis of *O*-glycosicic linkages [18].

v other NMR *J*-couplings are potentially measureable in **1**? There are:

(a) Five $^{1}J_{CC}$: $^{1}J_{C1,C2}$; $^{1}J_{C2,C3}$; $^{1}J_{C3,C4}$; $^{1}J_{C4,C5}$; $^{1}J_{C5,C6}$.

(b) Five $^{2}J_{CC}$: $^{2}J_{C1,C3}$; $^{2}J_{C1,C5}$; $^{2}J_{C2,C4}$; $^{2}J_{C3,C5}$; $^{2}J_{C4,C6}$.

(c) Two $^{3}J_{CC}$: $^{3}J_{C1,C6}$; $^{3}J_{C3,C6}$.

(d) Seven $^{1}J_{CH}$: $^{1}J_{C1,H1}$; $^{1}J_{C2,H2}$; $^{1}J_{C3,H3}$; $^{1}J_{C4,H4}$; $^{1}J_{C5,H5}$; $^{1}J_{C6,H6R}$; $^{1}J_{C6,H6S}$.

(e) Eleven $^{2}J_{CH}$: $^{2}J_{C1,H2}$; $^{2}J_{C2,H1}$; $^{2}J_{C2,H3}$; $^{2}J_{C3,H2}$; $^{2}J_{C3,H4}$; $^{2}J_{C4,H3}$; $^{2}J_{C4,H5}$; $^{2}J_{C5,H4}$; $^{2}J_{C5,H6R}$; $^{2}J_{C5,H6S}$; $^{2}J_{C6,H5}$.

(f) Eleven $^{3}J_{CH}$: $^{3}J_{C1,H3}$; $^{3}J_{C1,H5}$; $^{3}J_{C2,H4}$; $^{3}J_{C3,H1}$; $^{3}J_{C3,H5}$; $^{3}J_{C4,H2}$; $^{3}J_{C4,H6R}$; $^{3}J_{C4,H6S}$; $^{3}J_{C5,H1}$; $^{3}J_{C5,H3}$; $^{3}J_{C6,H4}$.

(g) Five $^{2}J_{COH}$[15]: $^{2}J_{C1,O1H}$; $^{2}J_{C2,O2H}$; $^{2}J_{C3,O3H}$; $^{2}J_{C4,O4H}$; $^{2}J_{C6,O6H}$.

(h) Seven $^{3}J_{CCOH}$[15]: $^{3}J_{C1,O2H}$; $^{3}J_{C2,O3H}$; $^{3}J_{C3,O2H}$; $^{3}J_{C3,O4H}$; $^{3}J_{C4,O3H}$; $^{3}J_{C5,O4H}$; $^{3}J_{C5,O6H}$.

(i) Two dual-pathway J_{CC}[18]: $^{3+3}J_{C1,C4}$; $^{3+3}J_{C2,C5}$.

Conventional NMR *J*-coupling analysis of **1** typically employs seven J_{HH} values (in most studies, *J*-couplings sensitive to C–O torsion angles are ignored). The remaining 61 spin-couplings (excluding those involving the methyl nuclei) are not measured and/or interpreted. Thus, only 7/68 or ~10 % of the full ensemble of *J*-couplings in **1** is used routinely. In conformationally rigid molecules, this may not matter much, but in molecules where

conformational flexibility exists or is suspected, access to a larger group of spin-couplings provides significant redundancy[1] with which to test different conformational models. Access to a larger dataset also makes possible studies of correlated conformations [19], and the development of "functional" *J*-couplings that report on a specific functional property, such as the strength of H-bonds [20]. In acknowledgement of this potential, significant effort has been devoted towards elucidating the structural dependencies of the $^{13}C–^{1}H$ and $^{13}C–^{13}C$ spin-couplings in structures like **1**. This work was comprised of three components: (a) *synthesis* of a diverse set of ^{13}C-labeled compounds to simplify the measurement of J_{CH} and J_{CC} values; (b) *NMR measurements* of J_{CH} and J_{CC} values in solutions of isotopically labeled compounds; and (c) *theoretical calculations* to assist in the structural interpretation of the multiple experimental *J*-couplings. This review addresses component (c), and specifically, the use of density functional theory (DFT) as a tool to inform experimental NMR studies of simple and complex saccharides. The aim of the ensuing discussion is to describe the fundamental features of DFT and describe how it can be used to reveal the complex structural dependencies of NMR *J*-couplings in saccharides. A few examples will be taken from the literature to illustrate the power of the NMR/DFT synergy.

2 Carbohydrate Modeling and Conformational Analysis

Solution NMR data obtained on conformationally flexible molecules yield a time-averaged molecular structure over the time period of the NMR observation (typically microseconds in FT-NMR). An interpretation of these data in terms of a single "average" conformation is, therefore, misleading in that this "virtual" conformation may be poorly represented in solution [21]. For flexible molecules, it can be difficult to deconstruct these time-averaged data into contributions made by individual conformers (i.e., determine conformer populations). To address this problem, computational modeling is used to assist in data analysis [22, 23]. Two methods are commonly applied in carbohydrate modeling: Monte Carlo (MC) sampling [24], and molecular dynamics (MD) simulations [25]. MC sampling is time-independent and randomly generates new conformations. In contrast, MD simulations are time-dependent. In the limit, time-averaged properties and ensemble-averaged properties are identical.

[1] "Redundancy" is defined herein as access to multiple NMR *J*-couplings that report on the same molecular torsion angle, θ. Furthermore, plots of *J*-coupling magnitude as a function of θ should not be superimposable, but rather should be phase-shifted so as to maximize the ability of these multiple couplings to discrimiate between different conformational models. If the plots overlap, the use of redundant *J*-couplings will not improve conformational analyses appreciably.

A fundamental principle, known as the ergodic hypothesis, relates time-dependent MD simulations and time-independent MC sampling. MD is more popular than MC largely because more force fields for carbohydrates have been developed for use with MD. Furthermore, MD methods are based on general biomolecular force fields, which can accommodate the inclusion of explicit solvent and be applied to study carbohydrate–receptor complexes.

Isotopically labeled mono- and oligosaccharides (uniform or site-specific) are not commonly employed in contemporary NMR structural studies, and thus the full power of heteronuclear multidimensional NMR experiments has not yet been brought to bear on structural investigations of these molecules. Although significant advances have been made in the synthesis of oligosaccharides [26], thereby enabling more efficient preparation of isotopically labeled derivatives, the determination of saccharide molecular properties in solution still relies heavily on natural abundance NMR methods and computational modeling. In the NMR studies described below, however, extensive use has been made of ^{13}C-labeled saccharides in order to simplify the measurement and assignment of ^{13}C–^{1}H and ^{13}C–^{13}C spin-couplings, and to make these measurements with greater precision and accuracy than are possible with natural (unlabeled) compounds.

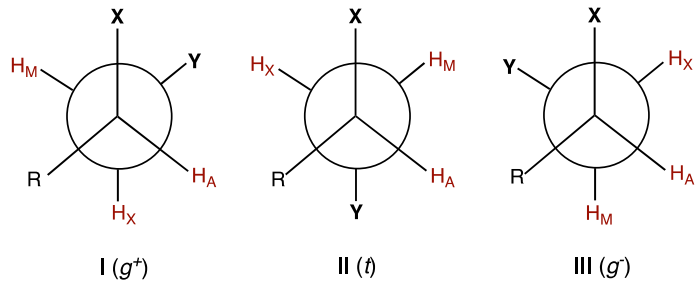
The use of NMR spin-couplings in conventional structural studies of saccharides often requires knowledge of limiting *trans* and *gauche* *J*-couplings to calculate rotamer populations in solution. For example, Eqs. 1–3 are commonly used for a spin-coupled AMX system involving a X–C–C–Y molecular fragment where rotation about the central C–C bond allows conformational exchange between three perfectly

$$^{3}J_{\mathrm{AM}} = \rho_{\mathrm{I}} J_{t} + \rho_{\mathrm{II}} J_{g-} + \rho_{\mathrm{III}} J_{g+} \quad (1)$$

$$^{3}J_{\mathrm{AX}} = \rho_{\mathrm{I}} J_{g+} + \rho_{\mathrm{II}} J_{t} + \rho_{\mathrm{III}} J_{g-} \quad (2)$$

$$\rho_{\mathrm{I}} + \rho_{\mathrm{II}} + \rho_{\mathrm{III}} = 1 \quad (3)$$

staggered rotamers **I**–**III** (X–C–C–Y torsion angles of 60°, –60°, and 180°; Scheme 3). J_t, J_{g+}, and J_{g-} in Eqs. 1–3 are the limiting *J*-couplings in rotamers **I**–**III**, and ρ_{I}, ρ_{II} and ρ_{III} are the fractional



Scheme 3

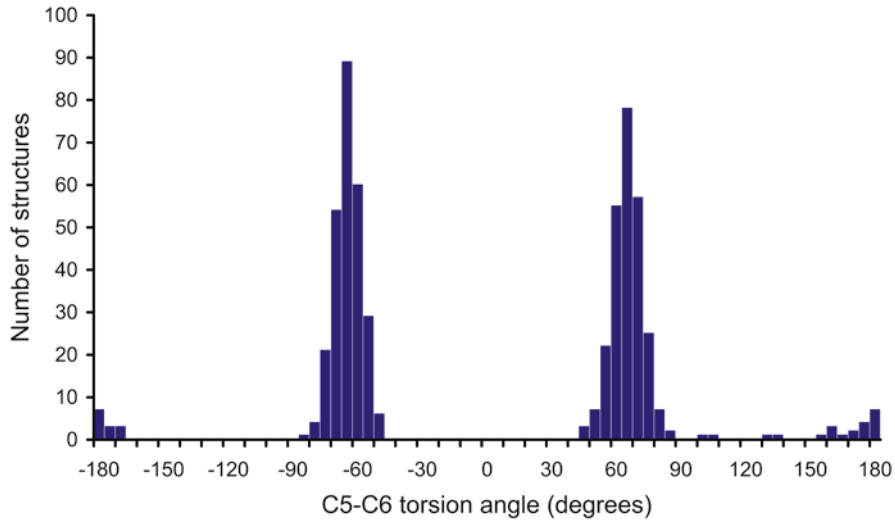


Fig. 1 Distribution of ω torsion angles (O5–C5–C6–O6) in aldohexopyranosyl rings from the CCDC database (www.ccdc.cam.ac.uk) (excluding cyclodextrins and O6-substituted compounds). 551 torsion angles were extracted from 172 different crystal structures

rotamer populations. This three-state model is often invoked to treat *J*-coupling data based on prior study of a crystal structure database of molecules that contain the X–C–C–Y fragment where three staggered (or near staggered) conformations are observed. Such is the case for rotation about the C5–C6 bond in **1** and **2** (exocyclic hydroxymethyl conformation) as shown in Fig. 1, but not for rotation about the ϕ (*phi*) and ψ (*psi*) bonds of internal *O*-glycosidic linkages as shown in Fig. 2. In the latter case, the observed torsion angles deviate significantly from those found in perfectly staggered rotamers, and thus treatment of *J*-couplings across these linkages using a simple three-state model like that shown in Scheme 3 would be unjustified. The three-state model breaks down in cases where perfectly staggered rotamers do not exist in solution, the lifetimes of the rotamers are comparable to the exchange rates between them (intermediates contribute to the NMR signal), and/or more than three conformations contribute to the observed NMR signal. It should also be appreciated that the values of the *gauche* $^3J_{AM}$ and $^3J_{AX}$ in rotamer **III** (Scheme 3) may not be the same despite identical H–C–C–H torsion angles (±60°), since the nature of the substituents X and R *anti* to the coupled hydrogens H_M and H_X, respectively, affect *J*-coupling magnitudes, especially if these substituents are electronegative [34].

Limiting *J*-couplings are often determined experimentally in a reference molecule that is known to exist in a single (or very restricted) conformation in solution. These structurally constrained molecules can be synthesized, although the task is sometimes difficult [35]. Even when possible, however, this approach is not without complications, since the introduction of bonding constraints in

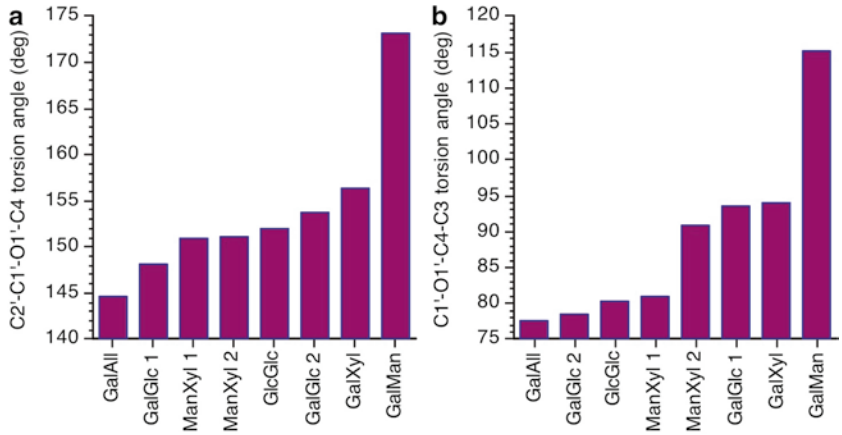


Fig. 2 Values of the *phi* (ϕ) (**a**) and *psi* (ψ) (**b**) torsion angles in several disaccharides determined by x-ray crystallography. GalAll methyl β-D-galactopyranosyl-(1 → 4)-β-D-allopyranoside [27], GalGlc 1, methyl β-D-galactopyranosyl-(1 → 4)-β-D-glucopyranoside; [28], ManXyl methyl β-D-mannopyranosyl-(1 → 4)-β-D-xylopyranoside [29], GlcGlc methyl β-D-glucopyranosyl-(1 → 4)-β-D-glucopyranoside [30], GalGlc 2, methyl β-D-galactopyranosyl-(1 → 4)-α-D-glucopyranoside; [31], GalXyl methyl β-D-galactopyranosyl-(1 → 4)-β-D-xylopyranoside; [32], GalMan methyl β-D-galactopyranosyl-(1 → 4)-α-D-mannopyranoside [33]. ManXyl 1 and ManXyl 2 denote two different conformations of the disaccharide observed in the crystal

the reference molecule can perturb the coupling pathway (e.g., bond length/angle perturbations) and lead to an erroneous limiting *J*-coupling. An alternative approach to obtain limiting *J*-couplings is to develop appropriate Karplus or Karplus-like equations from theoretical calculations. Density functional theory (DFT) can be used to obtain these equations, and from them, limiting *J*-couplings.

3 Introduction to Density Functional Theory

Density Functional Theory (DFT) collapses the complexities inherent in the many-particle wave-function-based approach to the solution of the time-independent Schrödinger equation with its attendant *3n* dimensionality, where *n* is the number of electrons, to an independent-particle treatment of the 3-dimensional electron density. The Hohenberg-Kohn theorem shows that the exact ground-state electronic energy of a molecule is uniquely determined by this electron density [36]. The Kohn-Sham formalism reintroduces an orbital-based approach to the computation of the electronic structure [37], but these orbitals are now those of a hypothetical non-interacting system, which, by construction, has the same electron density as the target system, the correspondence being assured by the adiabatic connection. All the complexity reappears in the construction of the exchange-correlation functional, the central object of DFT.

Early efforts to formulate the exchange-correlation functional relied on a model system, the uniform free-electron gas, certain aspects of which were amenable to exact solution, and led to local(spin)density approximations, L(S)DA, to the full functional. While this treatment proved extremely useful in describing, for example, conduction electrons in metals, it was markedly deficient for molecular systems that possess rapid variations in the electron density over short length scales.

An important step forward was taken with the generalized gradient approximation (GGA), where aspects of nonlocality were introduced by referencing the density gradient in addition to the density at the representative point. An early implementation of this approach was developed by Perdew and Wang (PW86) [38].

Another substantial improvement towards "chemical accuracy" was achieved by the incorporation of elements of exact Hartree–Fock exchange into these gradient corrected functionals. In particular, the three-parameter functional (e.g., Eq. 4) introduced by Becke [39],

$$E_{xc}^{B3} = (1-a)E_x^{LSDA} + aE_x^{HF} + b\Delta E_x^{B88} + E_{corr}^{LSDA} + c\Delta E_{corr}^{GGA} \quad (4)$$

where E_x and E_{corr} are the local exchange and correlation terms, and ΔE_x and ΔE_{corr} are their gradients, has proven extremely popular. The parameters, a, b, and c, are fit to known thermochemical data and depend on the form chosen for ΔE_{corr}^{GGA}, with the prescription of Lee et al. [40] being perhaps the most popular. Despite well-known deficiencies, these so-called B3LYP calculations offer a cost-effective approach to the rapid and satisfactory computation of many molecular properties. Intensive research into both the composition and reliability of density functionals is currently underway.

4 Perturbation Theory in Calculations of NMR Spin–Spin Coupling Constants

Most molecular properties, such as the indirect nuclear spin–spin coupling constant (*J*-coupling), may be defined as the response of the system to a perturbation. For example, for a perturbation of strength λ, the energy can be expanded in a Taylor series (Eq. 5):

$$\boldsymbol{E}(\lambda) = \boldsymbol{E}(0) + \frac{\partial \boldsymbol{E}}{\partial \lambda}\lambda + \frac{1}{2}\frac{\partial^2 \boldsymbol{E}}{\partial \lambda^2}\lambda^2 + \frac{1}{6}\frac{\partial^3 \boldsymbol{E}}{\partial \lambda^3}\lambda^3 + \ldots \quad (5)$$

where $\boldsymbol{E}(0)$ is the energy in the absence of the perturbation. Any nuclear spin I will result in an internal magnetic moment M:

$$\boldsymbol{M} = \gamma \hbar \boldsymbol{I} \quad (6)$$

Using M as the perturbation, Eq. 5 becomes

$$\boldsymbol{E}(\boldsymbol{M}_1, \boldsymbol{M}_2, \ldots) = \boldsymbol{E}(0) + \frac{\partial \boldsymbol{E}}{\partial \boldsymbol{M}_1}\boldsymbol{M}_1 + \frac{1}{2}\frac{\partial^2 \boldsymbol{E}}{\partial \boldsymbol{M}_1 \boldsymbol{M}_2}\boldsymbol{M}_1 \boldsymbol{M}_2 + \ldots \quad (7)$$

The second derivative with respect to two different nuclear spins is the reduced NMR coupling constant, K_{12}. As shown by Pople and coworkers [41–43], the ½ in Eq. 7 disappears since only distinct pairs of nuclei are considered:

$$\boldsymbol{K}_{12} = \frac{\partial^2 E}{\partial \boldsymbol{M}_1 \boldsymbol{M}_2} \tag{8}$$

The normal nuclear indirect spin–spin coupling constant between two nuclei 1 and 2, J_{12}, is then

$$\boldsymbol{J}_{12} = h\frac{\gamma_1}{2\pi}\frac{\gamma_2}{2\pi} \quad \text{and} \quad \boldsymbol{K}_{12} = h\frac{\gamma_1}{2\pi}\frac{\gamma_2}{2\pi}\frac{\partial^2 \boldsymbol{E}}{\partial \boldsymbol{M}_1 \boldsymbol{M}_2} \tag{9}$$

When magnetic fields are involved, the Hamiltonian operator can be complex. Four mechanisms, Fermi-contact (FC), spin-dipole (SD), diamagnetic spin-orbit (DSO) and paramagnetic spin-orbit (PSO), are introduced in Ramsey's expression for the reduced spin–spin coupling constant: [44]

$$\begin{aligned}\boldsymbol{K}_{12} = \left\langle \psi_0 \left| \mathbf{h}_{12}^{\mathrm{DSO}} \right| \psi_0 \right\rangle + 2\sum_{s>0} \frac{\left\langle \psi_0 \left| \mathbf{h}_1^{\mathrm{PSO}} \right| \psi_s \right\rangle \left\langle \psi_s \left| \mathbf{h}_2^{\mathrm{PSO}} \right| \psi_0{}^T \right\rangle}{E_0 - E_s} + \\ +2\sum_t \frac{\left\langle \psi_0 \left| \mathbf{h}_1^{\mathrm{FC}} + \mathbf{h}_1^{\mathrm{SD}} \right| \psi_t \right\rangle \left\langle \psi_t \left| \mathbf{h}_2^{\mathrm{FC}} + \mathbf{h}_2^{\mathrm{SD}} \right| \psi_0{}^T \right\rangle}{E_0 - E_t}\end{aligned} \tag{10}$$

where the first summation is over all singlet states different from the reference state and the second summation is over all triplet states. The terms $\mathbf{h}^{\mathrm{DSO}}$, $\mathbf{h}^{\mathrm{PSO}}$, $\mathbf{h}^{\mathrm{FC}}$, and $\mathbf{h}^{\mathrm{SD}}$ are DSO, PSO, FC, and SD operators that can be found in the literature [44, 45].

In many cases, particularly *J*-couplings involving single-bonded carbon and hydrogen nuclei, the Fermi-contact (FC) term makes the dominant contribution to the calculated spin-coupling. Consequently, early *J*-coupling calculations often recovered only the Fermi-contact term, and ignored the PSO, DSO and SD contributions [45]. However, several current software packages for calculating NMR *J*-couplings now give all four contributions as a matter of course.

5 Incorporation of Solvent Effects in DFT Calculations

Estimates of the effect of solvent on computed spin–spin coupling constants can also be obtained from DFT calculations using two approaches. It is computationally feasible to include a small number of explicit solvent molecules around the molecule under investigation during geometry optimization. However, care must be taken with such a construction to avoid biasing certain molecular conformations. In reality a continuous flux of solvent molecules bathes the periphery of the solute leading to changes in the solute and its environment that may be captured through extensive, and

costly, *ab initio* molecular dynamics simulations. This approach has seldom been pursued.

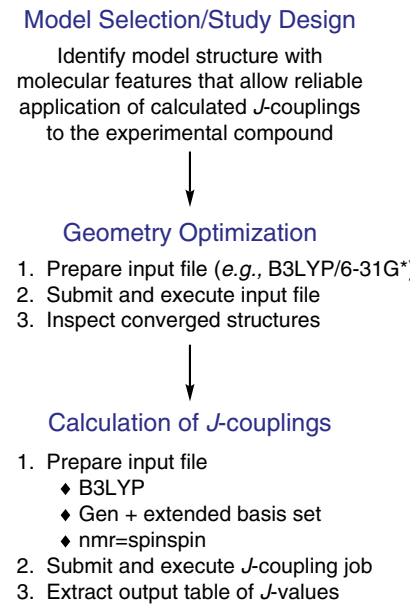
Alternatively, the surrounding medium may be represented as a polarizable continuous dielectric, encapsulating the molecule under scrutiny in a circumscribed cavity. The molecular charge distribution influences the polarization in this environment and the generated reaction field in turn influences the contained solute molecule and its associated electron density. This is the self-consistent reaction field (SCRF) approach. One particularly successful implementation of this model is found in the Integral Equation Formalism (IEFPCM) [46–49]. These techniques and their various shortcomings have been extensively reviewed (for example *see* ref. 50) and are the subjects of ongoing study and improvement.

6 Practical Information for Calculating *J*-Couplings in Saccharides

The general process of calculating NMR *J*-couplings in saccharides can be divided into the three separate and sequential parts summarized in Scheme 4: (1) model structure selection and study design; (2) geometry optimization; and (3) calculation of *J*-couplings. These parts are discussed in the following paragraphs.

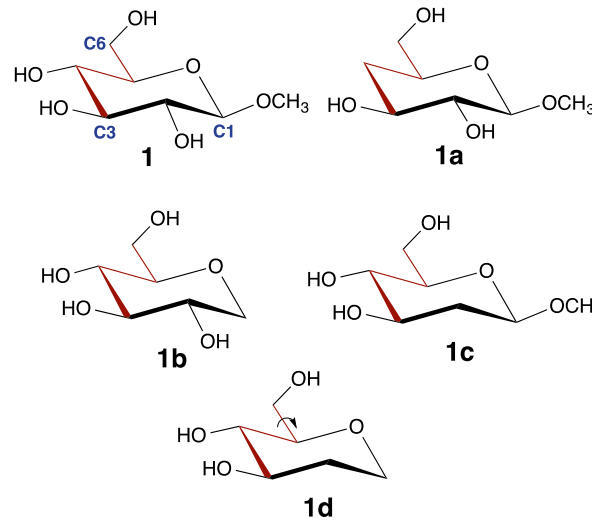
6.1 Model Structure Selection and Study Design

Selection of the model structure for the calculation is critically important and demands careful thought. Several different model structures might be expected to yield reasonable *J*-couplings



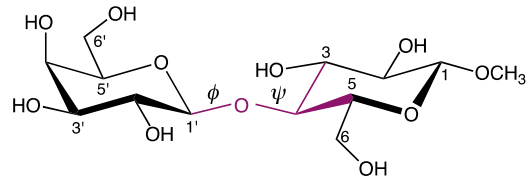
Scheme 4

applicable to the experimental compound. However, various factors may argue against using some of these models. These factors may include distortions in bond lengths and angles caused by the elimination of some ring substituents, the introduction of unwanted H-bonds involving hydroxyl groups, or ring-flipping and/or other undesirable conformational behaviors that do not exist in the experimental compound. For example, in studies of $^{3}J_{C3,C6}$ in **1** (*see* Structure 1), DFT calculations might involve the use of **1a**, **1b**, **1c,** or **1d** as a model structure instead of the fully substituted **1** in order to simplify the structure and reduce the cost of the calculation. Which of these model structures could serve as a valid model? **1a** lacks the in-pathway O4, **1b** lacks the out-of-pathway anomeric OCH_3, **1c** lacks the out-of-pathway O2, and **1d** lacks both the anomeric OCH_3 and O2. An additional concern is the potential influence of exocyclic C5–C6 bond rotation on $^{3}J_{C3,C6}$, as illustrated in **1d**. Model **1d** is probably suitable for DFT studies of $^{3}J_{C3,C6}$, but the effect of rotation about the C5–C6 bond on computed couplings would need to be inspected since this rotation affects $^{3}J_{C3,C6}$ significantly (i.e., while the C3–C4–C5–C6 torsion angle is the prime determinant of $^{3}J_{C3,C6}$, electronegative substituents along the coupling pathway, including those appended to the (terminal) coupled carbons, are also important determinants) [51]. The latter behavior points to the common complication of exocyclic bond rotation on J values in saccharides, especially when the rotation involves bonds to atoms bearing electron lone pairs. For example, rotation of the C3–O3 and C4–O4 bonds in **1** could also affect $^{3}J_{C3,C6}$, despite the fact that the C3–C4–C5–C6 torsion angle is largely fixed by the ring (in this case, C3 and C6 are approximately antiperiplanar). In practice, deciphering the effects of

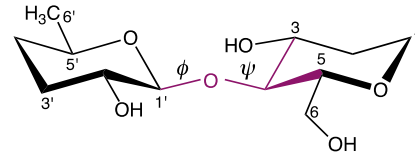


exocyclic C–O bond conformation (mainly hydroxyl groups) is a challenge and a persistent complication when conducting *J*-coupling calculations on saccharides. Often, to simplify the calculations, only one (or a few) combination(s) of C–O torsions is examined, with the realization that only a small portion of the total *J*-coupling hypersurface is being sampled.

Computational cost is a consideration if the study design requires an inspection of multiple conformations or structural analogs. Since computational time is geometrically related to the molecular mass of the model structure on which the calculation is being performed, the time required for calculations on a larger structure containing three or more conformational elements to be scanned can be prohibitive. In this situation, consideration should be given to eliminating heavy atoms from the model structure while retaining the carbon backbone and the hydroxyl groups known to affect the targeted *J*-coupling. In general, attractive model ring structures for saccharide *J*-coupling calculations by DFT are those that include the entire carbon backbone, the ring oxygen of the hemiacetal/acetal or hemiketal/ketal, all glycosidically bonded oxygens between moieties (if relevant), and all electronegative substituents proximal to the pertinent coupling pathways that may affect coupling magnitudes. Heavy atoms not likely to influence the targeted *J*-couplings can be eliminated for the sake of computational efficiency. This means that if the coupling pathway of interest is not affected by the presence of the exocyclic oxygen of the hemiacetal/acetal or hemiketal/ketal, it can be safely deleted from the model. An example of this simplification in studies of two trans-*O*-glycoside $^{3}J_{CCOC}$ values in experimental disaccharide **3** might employ **3a** as a model compound. In **3**, $^{3}J_{C1',C3}$ and $^{3}J_{C1',C5}$ are under investigation (coupling pathways highlighted in red), and model **3a** retains these coupling pathways



β-Gal-(1→4)-β-GlcOCH$_3$ (**3**)



β-Gal-(1→4)-β-GlcOCH$_3$ model (**3a**)

with structural integrity. However, several heavy atoms have been deleted in **3a** relative to **3** to reduce the computation time, with the realistic assumption that these deletions will not affect the computed couplings appreciably. Although the ring oxygen in the right residue in **3** might also be replaced by a carbon, the difference in computation cost is minimal, so there is no serious penalty paid by retaining the ring oxygen. In this case, however, the C1′–O1′–C4–C5 coupling pathway is likely to be influenced by O5 (potential terminal electronegative substituent effect on $^{3}J_{C1',C5}$) and thus it would inadvisable to delete O5 from the model structure.

The following discussion identifies some of the structural factors that influence the magnitudes of representative *J*-couplings in saccharides, which may inform decisions on the selection of a model structure for DFT calculations of *J*-couplings.

6.1.1 Substituent Effects on Homonuclear ^{1}H–^{1}H J-Couplings

Geminal (two-bond) H–C–H *J*-couplings ($^{2}J_{HCH}$) measured in the exocyclic hydroxymethyl (CH_2OH) groups of saccharides (e.g., *see* **1**) are affected by conformation of the C–O bond appended to the exocyclic hydroxymethyl carbon (i.e., rotation of the C6–O6 bond in **1** influences the magnitude of $^{2}J_{H6R,H6S}$) [15]. Vicinal (three-bond) H–C–C–H *J*-couplings are influenced strongly by the relative disposition of hydroxyl groups attached to the two carbons bearing the coupled hydrogens [52–54]. In some pyranosyl ring conformations, long-range ^{1}H–^{1}H *J*-couplings can be observed and their magnitudes depend on H–C–C–C–H coupling pathway geometry [18] and on the relative disposition of hydroxyl groups along the pathway.

6.1.2 Substituent Effects on Homonuclear ^{13}C–^{13}C J-Couplings

One-bond (direct) ^{13}C–^{13}C ($^{1}J_{CC}$) and geminal ^{13}C–C–^{13}C ($^{2}J_{CCC}$) spin-couplings are sensitive to the presence and orientation of hydroxyl groups appended to the coupled carbons, thus requiring their inclusion in a model structure in which these *J*-couplings are to be calculated [55–57]. Unlike $^{1}J_{CC}$ values whose signs are positive in saccharides, $^{2}J_{CCC}$ values can be positive or negative depending on the structure of the C–C–C coupling pathway. For $^{1}J_{CC}$, two factors control its magnitude in HO–C–C–OH fragments: (1) rotation about the central C–C bond, which displays a Karplus-like dependency, and (2) rotation about both C–O bonds, an effect apparently mediated by lone-pair orbital interactions with the C–C bond that modulate its length [55]. The latter C–O bond rotation effects on $^{1}J_{CC}$ appear to be stronger than the C–C bond rotation effect.

$^{2}J_{CCC}$ Values in a HO–C–C(OH)–C–OH coupling pathway in aldopyranosyl rings depend heavily on (a) the relative orientations of hydroxyl groups attached to the terminal coupled carbons (configurational effect) [56], and (b) the orientation of a hydroxyl substituent attached to the central carbon (i.e., a C–O bond conformation effect) [58]. The latter effect is caused by efficient spin-density transfer from one coupled carbon to the other when the intervening hydroxyl proton on the central carbon is oriented

anti to the hydrogen bonded to the same carbon. In this C–O bond conformation, each of the two oxygen lone-pair orbitals is *anti* to one of the two C–C bonds in the coupling pathway, thus allowing efficient overlap between the oxygen lone-pair electrons and the σ* anti-bonding orbitals of the C–C bonds [58].

Geminal $^{2}J_{COC}$ spin-couplings exist within typical saccharide ring systems (intra-ring pathways where one of the coupled carbons is the anomeric carbon) or between saccharide rings (pathways across an *O*-glycosidic linkage involving an anomeric carbon and an aglycone carbon). In the former case, anomeric configuration affects $^{2}J_{COC}$ significantly [57, 58], whereas in the latter, rotation about the phi (ϕ) glycosidic torsion angle is a key determinant [59].

Several vicinal ^{13}C–^{13}C *J*-couplings are found in saccharides involving C–C–C–C or C–C–O–C pathways [51, 60]. These couplings display Karplus-like properties and may be intra-ring (e.g., $^{3+3}J_{C1,C4}$ in aldopyranosyl rings), involve an exocyclic hydroxymethyl carbon (e.g., $^{3}J_{C1,C6}$ in aldohexopyranosyl rings), or exist across an *O*-glycosidic linkage. If intra-ring, two pathways contribute to the observed *J*-coupling (denoted dual-pathway couplings), and their structural interpretation can be complex, although some progress has been made [61]. When calculating $^{3}J_{CC}$ values from DFT, it is prudent to include in the model structure all hydroxyl groups appended to the internal and terminal carbons, since the terminal and internal effects of these groups on $^{3}J_{CC}$ magnitude can be significant (in some cases up to 1–2 Hz).

6.1.3 Substituent Effects on Heteronuclear ^{13}C–^{1}H J-Couplings

One-bond ^{13}C–^{1}H *J*-couplings are sensitive to C–H bond length or, alternatively, to the % *s*-character of the C–H bond [62, 63]. Multiple factors influence this bond length that in turn affect $^{1}J_{CH}$ magnitude. These factors fall into two main groups, namely, those that are "through bond" or hyperconjugative and those that are "through space" or steric. The influence on $^{1}J_{CH}$ of a hydroxyl group attached to the coupled carbon is a major factor affecting coupling magnitude. This effect may be mediated by hyperconjugative oxygen lone-pair donation into the C–H σ^* anti-bonding orbital or through interactions between the aliphatic hydrogen and either the hydroxyl hydrogen or lone-pairs on the oxygen. These mechanisms exert opposing forces on C–H bond length and depend on the C–O bond conformation. C–H bond length can also be affected by the presence of a vicinal hydroxyl group, depending on the relative configuration of the coupled carbon and the neighboring carbon bearing the hydroxyl group. Another factor affecting C–H bond length is H-bonding if the coupled carbon bears an OH group. The effect of this H-bonding on C–H bond length appears to depend on whether the OH group serves as a donor or an acceptor in the H-bond [20].

An additional structural factor that influences C–H bond length and indirectly $^{1}J_{CH}$ magnitude involves 1,3-interactions between axial C–H bonds in aldopyranosyl rings and an axial

hydroxyl group on a carbon two sites removed from the coupled carbon (e.g., the 1,3-interaction between H1 and O3H in an allo-hexopyranosyl ring). The C4–H4 bond in some aldohexopyranosyl rings is a special case of this effect; this bond experiences a 1,3-interaction with the exocyclic C6 hydroxymethyl group regardless of configuration at C4 due to the relatively free rotation about the C5–C6 bond. An axial C4–H4 bond can also experience a 1,3-interaction with the C2 hydroxyl group in some ring configurations (e.g., *manno*).

Geminal ^{13}C–^{1}H *J*-couplings in saccharides show a strong configurational dependence [64], but are also influenced by C–O bond conformation for hydroxyl groups attached to either the coupled or intervening carbon of the HO–C–C(OH)–H coupling pathway. The effect of C–O bond conformation is greater for the C–O bond involving the intervening carbon for reasons similar to those discussed above for $^{2}J_{CCC}$ [65]. These behaviors suggest that both hydroxyl groups should be included in a model structure when calculating $^{2}J_{CCH}$.

Vicinal $^{3}J_{CCCH}$ and $^{3}J_{COCH}$ are not only sensitive to C–C–C–H and C–O–C–H torsion angles, respectively, but are also sensitive to the presence of hydroxyl groups or ring oxygens appended to any carbon in these coupling pathways, necessitating their inclusion in any model structure [66]. In-plane contributions to $^{3}J_{CH}$ values in saccharides from terminal hydroxyl groups have been documented [66].

After a model structure has been selected, the design of the DFT study can be defined. This process includes a decision about which, if any, structural parameters such as bond lengths, bond angles and/or dihedral angles will be either held constant or varied systematically (scanned) in the model structure during the calculation. A specific structural parameter is scanned to determine its effect on a specific *J*-coupling. On the other hand, holding a given structural parameter fixed during a DFT calculation may prevent unwanted H-bonding and/or structural perturbations that might arise during optimization. In general, however, the use of fixed structural constraints should be kept to a minimum, since the model structure becomes increasingly artificial as more of them are applied.

With the study design in place, the model structure is built in silico using a molecular modeling program such as *ChemBio3D* [67], *Spartan* [68] or *GaussView* [69]. The program should allow the user to relax the model structure to obtain reasonable *initial* bond lengths and angles for use in the calculation. The modeling software should also allow certain geometric parameters such as bond or dihedral angles to be defined by the user, and should support the printing or exporting of a table of atomic identities and positions for the model structure. The latter data can take the form of either Cartesian coordinates or a z-matrix (*see* below). In addition to their inclusion in the input file for geometry optimization (discussed below), the starting (initial) coordinates should be recorded in a

spreadsheet created for the project. This spreadsheet will be the repository of all subsequent project data related to the model structure and represents a digital record of the project. Given the large amounts of data that are generated in DFT studies, it is important that a systematic means of retrieving, recording and archiving the data is established before DFT calculations are undertaken.

6.2 Geometry Optimization

Several *ab initio* geometry optimization programs are commercially available to support this step of the process. *GAMESS* (General Atomic and Molecular Electronic Structure System) [70] and *Jaguar* [71] are commonly used programs, but perhaps the most popular program is *Gaussian*, the current version of which is *Gaussian09* [72]. A key advantage of *Gaussian* is that it can be used for designing model systems, geometry optimization, and the calculation of spin–spin coupling constants.

The first step of geometry optimization is to prepare an input file. This file contains information about the model structure to be optimized and how the optimization is to be performed. Specifically for *Gaussian*, input files contain up to five distinct sections. These input files originate as a basic text file, easily made in programs such as Notepad or any generic vi editor. Section one contains the Link 0 commands, allowing for user specification of how many processors to use (%Nproc=#), the amount of dynamic memory (%Mem=#GB), and the location and name of the checkpoint file during optimization (%chk=*filename*). These parameters are always preceded by "%" within this section; additional user-set specifications are available.

Section two delineates the calculation type. Lines containing these commands are preceded by "#" and all parameters can be listed sequentially with a space between each. Selection of the calculation type requires decisions about the force field and basis set that will be employed for geometry optimization. For moderately sized saccharides whose heaviest atom is oxygen, geometry optimization and the calculation of NMR spin–spin coupling constants are conducted using a hybrid functional such as B3LYP [39], which has a mixture of Hartree–Fock exchange and DFT exchange-correlation, and when combined with the 6-31G(d) (commonly referred to as 6-31G*) basis set [73] returns excellent results with reasonable computational efficiency. The 6-31G(d) basis set adds d functions to non-hydrogen atoms to accommodate for distortions in bonding; if p functions on H are also desired, the 6-31G(d,p) basis set should be used. While there are additional functionals available, the B3LYP/6-31G(d) functional produces an optimal balance of speed and accuracy on reasonably small saccharides (several residues or less).

In *Gaussian09*, the default for SCRF is IEFPCM [46] and it is the preferred solvation method for most applications in carbohydrate systems. The keyword "scrf" is included in the routing command line of the input file and is synonymous with "scrf=pcm" and "scrf=iefpcm" from previous implementations. A number of

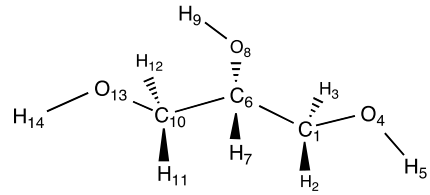
previous weaknesses have been addressed in the most recent implementation, as suggested by Karplus and York [74].

In addition to specifications for the basis set and solvation, the integration grid for the calculation must be delineated. Smaller saccharide structures with many tetrahedral centers perform well with a pruned (99,590) grid, with 99 radial shells and 590 angular points per shell. This grid is selected with the command int = grid = ultrafine, and should be placed on the same line as the calculation level of theory, basis set, and solvation. Larger and smaller grids are available for different levels of structure size and target accuracy.

Section three is brief and contains the title of the calculation under consideration. It lacks any specific symbol (in comparison to "%" and "#" above) and while a title is required, the contents of the title are not utilized by *Gaussian*. The title is for identification of the job once the calculation has been completed.

At this point the identification of initial geometric coordinates and designation of fixed or scanned geometric parameters such as a dihedral angle must be made in accordance with the study design. The "opt = modredundant" keyword must be included in the same input line as the designation of the functional and basis set to allow user specification of fixed or scanned structural parameters. The fourth section of the input file contains the initial coordinates for the model structure obtained from molecular modeling, preceded by the designation of its charge and multiplicity, which are typically zero and one, respectively, for neutral (uncharged) saccharides. Following this designation are the initial geometric coordinates for optimization, which may take the form of a z-matrix or Cartesian coordinates. An input file for the geometric optimization of glycerol **4** at the B3LYP/6-31G* level of theory might take the following form:

```
%NProc=8
%Mem=1GB
%Chk=Glycerol_C-O
# b3lyp/6-31G* opt=modredundant int=grid=ultrafine
scrf=iefpcm
glycerol with rotation about the central C-O bond
0 1
z-matrix for glycerol model structure inserted here
```



4

A z-matrix for **4** that is inserted into this input file at the indicated position might assume the following form:

```
C1
H2 1 B1
H3 1 B2 2 A1
O4 1 B3 2 A2 3 D1
H5 4 B4 1 A3 2 D2
C6 1 B5 4 A4 5 D3
H7 6 B6 1 A5 4 D4
O8 6 B7 1 A6 4 D5
H9 8 B8 6 A7 1 D6
C10 6 B9 1 A8 4 D7
H11 10 B10 6 A9 1 D8
H12 10 B11 6 A10 1 D9
O13 10 B12 6 A11 1 D10
H14 13 B13 10 A12 6 D11

B1 1.0797
B2 1.0826
B3 1.4361
B4 0.9693
B5 1.5248
B6 1.0785
B7 1.4462
B8 0.9700
B9 1.5199
B10 1.0826
B11 1.0828
B12 1.4491
B13 0.9647
A1 109.24
A2 107.47
A3 106.81
A4 108.65
A5 109.98
A6 107.46
A7 106.73
A8 112.49
A9 109.98
A10 111.02
A11 104.02
A12 112.22
D1 -121.35
D2 -162.17
D3 -41.95
D4 -63.57
D5 52.37
D6 82.97
D7 172.00
D8 174.17
D9 52.89
D10 -67.27
D11 -176.04

D 10 6 8 9 -60.0 S 11 30.0
```

The initial lines of this z-matrix provide computer instructions for the atom connectivity in the model structure. For example, the translation of line 7 is: H7 is bonded to C6 with a bond length of B6, H7 makes a bond angle H7–C6–C1 with C1 of A5, and H7 makes a dihedral angle H7–C6–C1–O4 with O4 of D4. The atom numbering follows that shown in **4**. Following these instructions are the initial values of the bond lengths (*B*), bond angles (*A*) and bond dihedral angles (*D*) defined in the z-matrix; these values would have been obtained from prior molecular modeling of the model structure (*see* above). In **4**, thirteen bond lengths, twelve bond angles, and eleven bond torsions are needed to construct the initial model structure in silico.

z-Matrices are used to construct the model structure *in silico* when Cartesian coordinates are unavailable either from a relaxed molecular mechanics calculation or from a crystal structure. Cartesian coordinates can replace the z-matrix in the input file for *Gaussian*; all of the other input lines remain the same. Cartesian coordinates for **4** are as follows:

C1	–2.579782502	0.66126408	–0.077219156
H2	–2.948781214	0.315405941	–1.031158785
H3	–2.943043421	0.006262102	0.704496615
O4	–3.043031954	2.008804077	0.101294541
H5	–2.492309552	2.405261767	0.793503055
C6	–1.055271195	0.661668207	–0.047648089
H7	–0.6707187	1.233655342	–0.877225127
O8	–0.644818415	1.345628087	1.158706244
H9	–0.683272758	0.693843346	1.876114515
C10	–0.473712561	–0.742550936	–0.034344963
H11	0.604773579	–0.697153269	–0.116850502
H12	–0.872769392	–1.335946095	–0.847473391
O13	–0.866265149	–1.273016946	1.255792066
H14	–0.501278666	–2.151672704	1.415314967
D 10 6 8 9 –60.0 S 11 30.0			

If the calculation calls for either a fixed or a scanned structural parameter, the desired constraint is appended to the end of the input file following the atomic coordinates. This addendum constitutes the fifth and final section of the input file. The format is to list the space delimited number of the atomic centers as they appear in the coordinate list followed by the initial value in degrees and a letter designation of "*F*" (for fixed) or "*S*" (for scanned) in subsequent calculations, followed by the number of steps in the scan and the degrees per step.

Some changes in this formatting have occurred in recent updates of *Gaussian*. For example, older versions of *Gaussian* require a dihedral angle specification following the letter designation "*F*" to tell the program the constraint at which to hold the dihedral (e.g., "D 10 6 8 9 -60.0 F" will hold the C10–C6–O8–H9 bond at –60°). Newer versions, such as *Gaussian09*, do not require this dihedral specification; the program assumes the initial dihedral orientation is the same orientation to maintain during the calculation (e.g., "D 10 6 8 9 F" will hold C10–C6–O8–H9 at –60° provided the bond was originally set at –60°). Contact the high-end computing technician at your institution for details on how to format this line and/or submit an input file to the cluster batch queue for execution.

The last line of the example z-matrix and Cartesian coordinates of **4** indicate that a specific bond is to be systematically rotated through 360° during the calculation. In this case, the C10–C6–O8–H9 dihedral angle in **4** will be scanned (S), starting from –60° in 30° increments in 11 steps. The output file will thus contain 12 geometrically optimized structures, with the initial structure containing a C10–C6–O8–H9 dihedral angle of –60°, and 11 additional structures in which this dihedral angle was rotated in 30° increments to a final value of 270°. During the 12 optimizations, all values of *B*, *A*, and *D* in the z-matrix will be optimized (relaxed) at each C10–C6–O8–H9 dihedral angle.

After geometry optimization calculations have been completed, the atomic coordinates of the energetically converged structure or structures are extracted from the output file and imported into a molecular visualization program such as *ChemBio3D* [67], *Spartan* [68], or *GaussView* [69]. If the study includes calculations of multiple structures, either as a single structure scanned about one or more structural parameters or involving two or more constitutional isomers, it is important to inspect the self-consistent field relative energies of each of the converged structures for anomalies. It is also important to inspect the energetically converged structures visually to ensure that they are correct, and that no undesired structural perturbations were introduced during geometry optimization, such as improper dihedral angles, ring inversions, or unwanted H-bonds. If disk space is available for archival purposes, it is advisable to save the output files obtained from geometry optimization since these raw data may prove useful later.

6.3 Calculation of NMR J-Couplings

Theoretical descriptions of the computational process by which NMR spin–spin coupling constants are calculated have been published previously [75]. The present discussion focuses on the practical aspects of the process. Once satisfactory geometrically optimized model structures have been obtained, it is relatively straightforward to calculate NMR spin–spin coupling constants using *Gaussian*. The current release of the program, *Gaussian09* [72], calculates all four contributions to the *J*-coupling, namely, the Fermi-contact,

paramagnetic spin orbit, diamagnetic spin orbit, and spin dipole terms, as discussed above.

An input file for the calculation of *J*-couplings is constructed using the atomic coordinates of the optimized structures. *Gaussian* can retrieve the optimized geometric coordinates from the input file as a z-matrix or set of Cartesian coordinates. However, *Gaussian* allows for the specification of a reference checkpoint file, from which the coordinates can be retrieved. The location of the file must be noted in the Link 0 commands (%chk = *filename*) and "geom = check" must appear with the calculation type specifications. Unlike geometry optimization, *J*-coupling calculations are conducted using the open shell format, and the hybrid functional B3LYP is preferred. An extended basis set optimized for NMR spin–spin coupling constant calculations in saccharides has been developed and is recommended for these calculations [15, 76]. To implement this basis set, the "Gen" keyword is used in the routing command line of the input file. This keyword allows the designation of a unique basis for the calculation. The remainder of the basis set needs to be explicitly designated in the basis set input section following the atomic coordinates. Tight convergence in the SCF procedure is recommended and is the default option in Gaussian09. The "nmr" keyword returns values for the NMR shielding tensors and magnetic susceptibilities in the output file. The "nmr = spinspin" option must be present in the command line to instruct *Gaussian09* to compute NMR spin–spin coupling constants.

An example of a input file that includes the descriptions of the extended basis set for carbon, oxygen and hydrogen, denoted [5s2p1d,3s1p] [15, 76], is as follows:

```
%NProc=4
%Mem=24MW
%chk=filename
# b3lyp/gen nmr=spinspin int=grid=ultrafine scrf=iefpcm
J calculation for model structure X
0 1
coordinates of geometrically optimized model structure X inserted here
```

C	0
S	5 1.00
4232.61	0.002029
634.882	0.015535
146.097	0.075411
42.4974	0.257121
14.1892	0.596555

S	1 1.00
5.1477	1.000000
S	1 1.00
1.9666	1.000000
S	1 1.00
0.4962	1.000000
S	1 1.00
0.1533	1.000000
P	4 1.00
18.1557	0.018534
3.9864	0.115442
1.1429	0.386206
0.3594	0.640089
P	1 1.00
0.1146	1.000000
D	1 1.00
0.56	1.000000
****	
O	0
S	5 1.00
7816.54	0.002031
1175.82	0.015436
273.188	0.073771
81.1696	0.247606
27.1836	0.611832
S	1 1.00
9.5322	1.000000
S	1 1.00
3.4136	1.000000
S	1 1.00
0.9398	1.000000
S	1 1.00
0.2846	1.000000
P	4 1.00

(continued)

(continued)

35.1832	0.019580
7.9040	0.124189
2.3051	0.394727
0.7171	0.627375
P	1 1.00
0.2137	1.000000
D	1 1.00
1.10	1.000000
****	
H	0
S	3 1.00
127.9500	0.01074
19.2406	0.11954
2.8992	0.92642
S	1 1.00
0.6534	1.00000
S	1 1.00
0.1776	1.00000
P	1 1.00
1.0000	1.00000

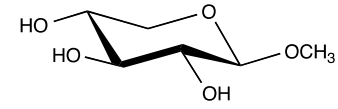
If a solvent continuum is employed, it is common to include it in both the geometry optimization and the J-coupling calculations. If the model structure bears a net charge, this charge must be indicated in the input file immediately prior to the atomic coordinates, along with the multiplicity. For example, the line of the input file containing the charge and multiplicity for the ionized form of *N*-acetyl-neuraminic acid would be "–1 1".

Jobs can be submitted to the batch queue in a manner similar to that for geometry optimization. Depending on the number and size of the model structures, the number of processors available (NProc), and any time constraint considerations, jobs may be run in series or in parallel and with one or more processors used per job. Calculations of NMR spin–spin coupling constants are about twice as computationally costly as calculations of vibrational frequencies, and computational cost increases geometrically with the number of atomic centers in the molecule. If time is a concern, the "nmr = mixed," option can be used with a standard valence

oriented basis set rather than "nmr = spinspin" with the modified basis applied to all four *J*-coupling terms. Users should refer to the *Gaussian09* documentation for a more detailed description of this option. When the calculations have completed, the table of *J*-couplings in the output file can be transferred into the project spreadsheet. The *J*-couplings of interest can then be extracted from this spreadsheet for subsequent analysis.

7 Limitations of DFT Calculations of NMR *J*-Couplings in Saccharides

Computational studies of saccharides are complicated by the presence of multiple exocyclic groups attached to the furanosyl and pyranosyl rings comprising their carbon backbones. The common exocyclic groups include hydroxyl (OH) and hydroxymethyl (CH_2OH) groups, with the former potentially more troublesome than the latter. This complication arises from two sources: (1) the propensity of hydroxyl groups to serve as H-bond donors and acceptors, and (2) the stereoelectronic properties of the electron lone-pair orbitals on the hydroxyl oxygen. In the $^4C_1$ conformer of methyl β-D-xylopyranoside **5**, 3^4 unique conformations exist by virtue of rotation about the four exocyclic C–O bonds, giving 81 different conformers. In principle, these 81 conformers could be geometrically optimized by DFT, and based on the relative energies so determined, the fractional populations of each conformer could be computed. Will these calculated fractional populations accurately reflect those found in solution? Most likely not, because of complications arising from the current DFT treatment of intramolecular H-bonding. Unless the calculations are performed in a manner that replicates solution conditions exactly, that is, solvation by water, intramolecular H-bonding may distort the equilibrium fractional populations of conformers significantly. For instance, in limiting *in vacuo* calculations, these interactions are likely to be much stronger than they would be in the solvated state, thereby stabilizing certain conformations more and others less than would occur in solution. Even when solvation is taken into account using present DFT methodology (*see* above), the computed populations still may not faithfully reproduce those in solution. This situation is due in part to the fact that solvent continuum models typically apply a continuous dielectric to the system to mimic bulk properties. However, the effect of solvent on the H-bonding patterns in



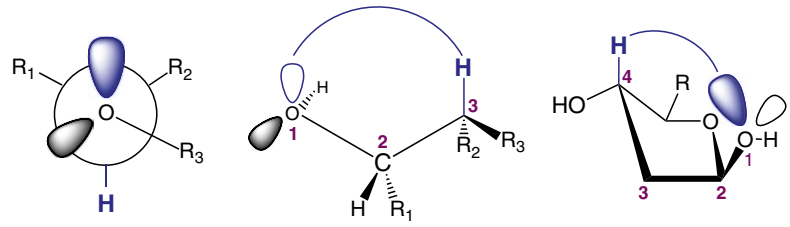
5

a saccharide vary from site to site based on how an explicit solvent molecule may participate in intermolecular H-bonding with the saccharide and thereby disrupt intramolecular H-bonding in the saccharide.

The complication in computing reliable conformer populations from DFT-calculated conformational energies has important implications for efforts to calculate *J*-couplings quantitatively. A DFT-calculated *J*-coupling that accurately reproduces an experimentally measured *J*-coupling in **5** should, in principle, be obtainable by calculating the *J*-coupling in all 81 conformers and weighting them according to the DFT-calculated fractional populations of conformers. However, since the fractional populations cannot be calculated with confidence, efforts to calculate *J*-couplings that accurately reproduce experiment by taking these populations into account are pointless. In addition, the lack of explicit in silico solvent notwithstanding, it is difficult to validate by experiment the fractional C–O rotamer populations calculated by DFT, since NMR determinations of these populations based on analyses of $^{3}J_{HCOH}$ and $^{3}J_{CCOH}$ values are difficult to make in aqueous solution, although some indirect (i.e., *J*-couplings that do not rely on the observation of solvent exchangeable hydroxyl hydrogens) *J*-coupling-based methods appear promising [16]. Even so, NMR can only yield rotamer populations for an *individual* C–O bond in the total population of conformers (that is, the fractional populations of all 81 conformers cannot be determined experimentally), but at a minimum these composite experimental populations could be compared to those computed from the DFT fractional populations as a crude means of validating the latter.

Implicit in the above discussion is the assumption that C–O bond rotations affect individual *J*-couplings in saccharides sufficiently to require that they be taken into account in the calculations. For some *J*-couplings (e.g., $^{3}J_{CCCH}$), C–O bond rotations along the coupling pathway exert little effect, whereas others (e.g., $^{1}J_{CH}$, $^{1}J_{CC}$, $^{2}J_{CCH}$, $^{2}J_{CCC}$) are affected significantly. The latter sensitivities are caused largely by the hyperconjugative or "through-bond" influence of electron lone-pair orbitals on nearby C–C and C–H bond lengths. For example, in addition to bond orientation, vicinal, 1,3-, and 1,4-interactions are "through-space" mechanisms that affect C–H bond lengths and thereby J_{CH} coupling magnitudes (Scheme 5). Since spin-density transfer involving carbon and hydrogen is largely, if not solely, a through-bond phenomenon in saccharides, any perturbation in C–H and C–C bond lengths induced indirectly by exocyclic C–O bond rotations will perturb the *J*-coupling. These effects can be substantial; for example, rotation about the C1–O1 bond in **6** causes substantial changes in $r_{C1,H1}$, and in turn significant changes in both $^{1}J_{C1,H1}$ and $^{2}J_{C2,H1}$ (Fig. 3).

The above discussion considers sources of error in *J*-coupling calculations arising from an inadequate account of conformational

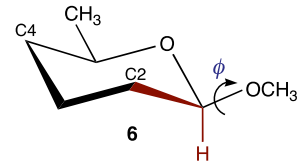


vicinal lone pair effect (lengthens C-H bond)

1,3 lone-pair effect (shortens C-H bond)

1,4 lone-pair effect (shortens C-H bond)

Scheme 5



6

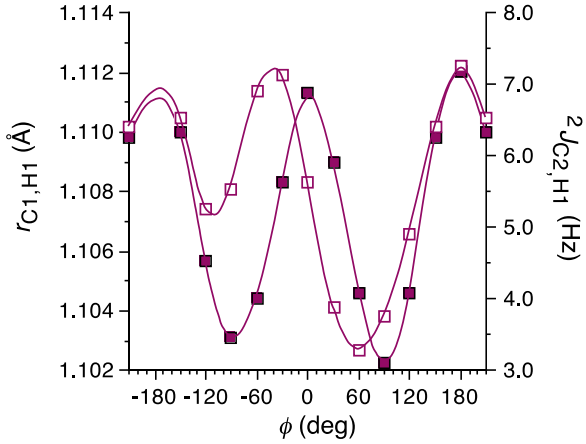


Fig. 3 Effect of phi (ϕ) on $r_{C1,H1}$ (*open squares*) and $^{2}J_{C2,H1}$ (*closed squares*) in **6**. Data have been taken from ref. 65

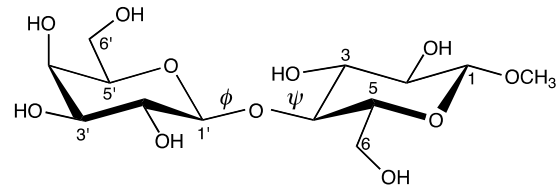
averaging of the calculated *J*-couplings. Superimposed on this source of error are the intrinsic errors of the DFT method itself, which could be considerable in some cases. Estimates have been made of the magnitude of this error by calculating a specific *J*-coupling in a saccharide in which the effects of exocyclic hydroxyl groups are small or absent (e.g., $^{3}J_{CCCH}$). These studies suggest that errors of ±5 % are possible; thus, for large *J*-couplings like $^{1}J_{CH}$, the absolute error could be ±7 Hz, whereas for smaller *J*-couplings like

$^3J_{COCC}$, the absolute error could be ±0.2 Hz. The uncertainties in estimating these errors stem from the fact that the effects of hydroxyl groups cannot be totally eliminated from these determinations. However, since these errors are presumed to be constant, changes in a given calculated *J*-coupling caused by some structural perturbation are likely to be nearly quantitative (i.e., while calculated absolute *J*-couplings may not be quantitative, changes in their calculated values are almost so).

8 Illustrations of *J*-Coupling Studies in Saccharides Assisted by DFT

The application of DFT to structural NMR studies of saccharides has resulted in an improved understanding of *J*-couplings in these biomolecules. While a complete discussion of these findings is not possible here, three examples serve to illustrate the kinds of information that can be obtained from these studies.

A. *Example 1.* The disaccharide, methyl β-D-galactopyranosyl-(1→4)-β-D-glucopyranoside **7** is comprised of two monosaccharides, each containing an exocyclic hydroxymethyl (CH_2OH) group. Two $^3J_{CC}$ values are available within each pyranosyl ring to evaluate ring conformation: $^3J_{C1,C6}$ and $^3J_{C3,C6}$, the former involving a C–O–C–C coupling pathway and the latter a C–C–C–C coupling pathway. Both *J*-couplings are vicinal, and thus both depend primarily on conformation about the central endocyclic bond (C5–O5 and C4–C5, respectively). In ideal $^4C_1$ chair conformers, the dihedral angles between C1 and C6, and between C3 and C6, should be ~180° (*anti*), giving $^3J_{CC}$ values of 3–4 Hz [51, 77, 78]. In $^1C_4$ forms, these angles will be ~±60° (*syn*), giving smaller couplings (<1 Hz) [51]. Four model structures **8**–**11**, all in $^4C_1$ chair conformations, were employed in DFT calculations to investigate two secondary structural effects on $^3J_{C1,C6}$ and $^3J_{C3,C6}$: (1) the effect of exocyclic C5–C6 bond rotation (hydroxymethyl conformation; *ω*), and (2) the effect of *terminal* electronegative (OH) substituents [61]. The coupling pathways for $^3J_{C1,C6}$ in **8** and **9** (and **10** and **11**) differ only in the orientation of O1 (axial in **8**, equatorial in **9**). Likewise, the coupling pathways for $^3J_{C3,C6}$ in



β-Gal-(1→4)-β-GlcOCH$_3$ (**7**)

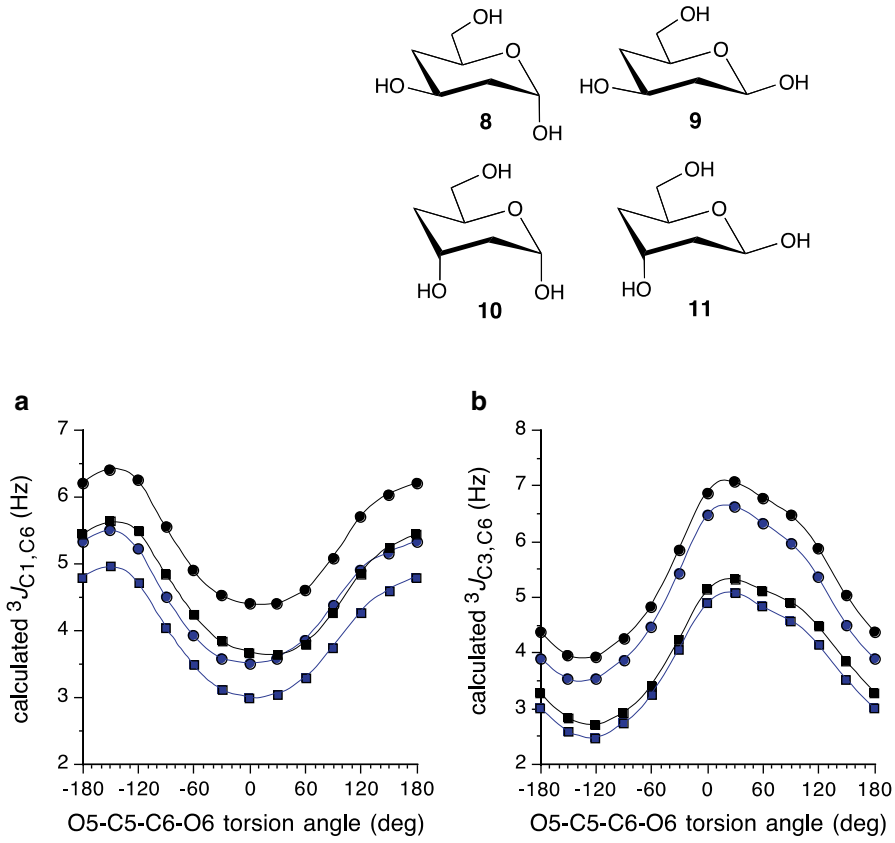


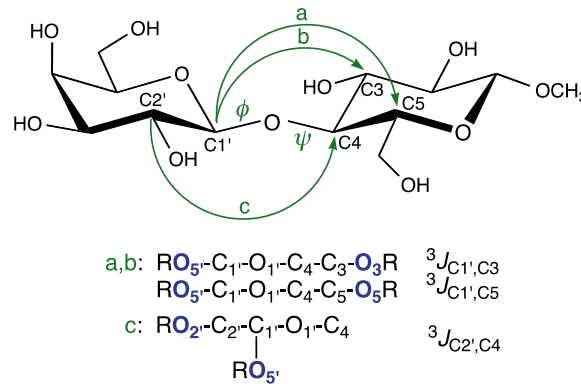
Fig. 4 (**a**) Effect of exocyclic CH_2OH conformation (ω = O5–C5–C6–O6 torsion angle) on calculated $^3J_{C1,C6}$ in **8** (*blue circles*), **9** (*black circles*), **10** (*blue squares*), and **11** (*black squares*). (**b**) Effect of exocyclic CH_2OH conformation on calculated $^3J_{C3,C6}$ (same symbols as in *panel* **a**). Data taken from ref. 61

8 and **10** (and **9** and **11**) differ in the orientation of O3 (equatorial in **8**, axial in **10**). The results of these calculations in which the O5–C5–C6–O6 torsion angle was scanned through 360° in 30° increments are shown in Fig. 4. Ring conformation in **8**–**11** is constrained in silico to $^4C_1$, and thus the calculated $^3J_{C1,C6}$ and $^3J_{C3,C6}$ values should be unperturbed if hydroxymethyl conformation and terminal C–O bond orientation do not affect these *J*-couplings. In both cases, however, the effect of hydroxymethyl conformation (ω) is substantial, with maximal coupling observed at $\omega = 180°$ for $^3J_{C1,C6}$ and near +60° for $^3J_{C3,C6}$. In the latter geometries, O6 lies in the coupling plane, that is, the C1–O5–C5–C6–O6 and C3–C4–C5–C6–O6 molecular fragments are coplanar. The data show that orienting O6 in the plane of the coupling pathway increases the observed coupling by 1–3 Hz (total amplitude of each curve in Fig. 4). The dispersion in *J*-couplings along the *y*-axes of Fig. 4a, b at fixed ω is caused by the effects of C1–O1 bond orientation on

$^3J_{C1,C6}$ (panel a) and of C3–O3 bond orientation on $^3J_{C3,C6}$ (panel b). Thus, at any value of ω, $^3J_{C1,C6}$ is larger in structures containing an equatorial C1–O1 bond, that is, in structures having a coplanar O1–C1–O5–C5–C6 fragment; the enhancement is ~1 Hz. The same effect is observed for $^3J_{C3,C6}$; $^3J_{C3,C6}$ is larger in **9** than in **11**, and is larger in **8** than in **10**, with the observed enhancement being ~1 Hz. The calculations also show that configuration at C3 affects $^3J_{C1,C6}$, and configuration at C1 affects $^3J_{C3,C6}$ values, findings that are in agreement with experimental observations [77, 78].

Data in Fig. 4 show that $^3J_{CC}$ values in saccharides are influenced by the presence of *terminal* electronegative substituents in the coupling pathways, with in-plane orientations enhancing the observed coupling by ~1 Hz if only the three staggered O5–C5–C6–O6 rotamers are considered. These theoretical findings are consistent with experimental observations. $^3J_{C1,C6}$ and $^3J_{C3,C6}$ not only are useful probes of ring conformation (Karplus dependence) *but also report on C5–C6 bond conformation (ω) and on ring configuration at O1 (via $^3J_{C1,C6}$) and C3 (via $^3J_{C3,C6}$)*.

B. *Example 2.* Several *J*-couplings exist across the internal *O*-glycosidic linkage in **7**, including three sensitive to *phi* (ϕ) and three sensitive to *psi* (ψ): for ϕ, $^2J_{C1',C4}$, $^3J_{H1',C4}$, and $^3J_{C2',C4}$; for ψ, $^3J_{C1',H4}$, $^3J_{C1',C3}$, and $^3J_{C1',C5}$. The behavior of these three $^3J_{CC}$ values will be discussed, and specifically a comparison will be made between the *J*-coupling sensitive to ϕ ($^3J_{C2',C4}$) and those sensitive to ψ ($^3J_{C1',C3}$ and $^3J_{C1',C5}$). From the above discussion, it should now be appreciated that *terminal* electronegative substituents in a C–O–C–C coupling pathway significantly affect $^3J_{COCC}$ values. Do *internal* electronegative substituents also affect $^3J_{COCC}$ values, and if so, how does this effect manifest itself? Consider the two coupling pathways under consideration (Scheme 6). In pathways "a" and "b," only *terminal*



Scheme 6 Trans-glycoside $^3J_{COCC}$ coupling pathways in **7** containing only terminal ($^3J_{C1',C3}$ and $^3J_{C1',C5}$), and both internal and terminal ($^3J_{C2',C4}$), electronegative substituents. Note that $^3J_{C1',C3}$ and $^3J_{C1',C5}$ are sensitive to ψ, whereas $^3J_{C2',C4}$ is sensitive to ϕ

electronegative substituents are present, whereas pathway "c" also contains an internal electronegative substituent. To examine the effect of the latter, DFT calculations were conducted on model structures **12** and **13** [79]. Model structure **12** contains a typical *O*-glycoside linkage in which the C2–C1–O1–CH_2 pathway bears an internal electronegative substituent (O5), whereas **13** contains a pathway devoid of this substituent. DFT was used to calculate the dependence of $^3J_{C2,CH2}$ in **12** and **13** on ϕ, which in this case is defined as the C2–C1–O1–CH_2 torsion angle. The results of these calculations are shown in Fig. 5.

The calculated relationships between the C2–C1–O1–CH_2 torsion angle and $^3J_{C2,CH2}$ are Karplus-like, with minima located near 90° and 270°, and maxima located near 0° and 180°. However, the two curves are not superimposable; that for **12** is phase-shifted to the left relative to that for **13.** The maximum for **13** occurs at 180°, while that for **12** occurs at ~165°. DFT studies of other model structures confirmed this shift, which is

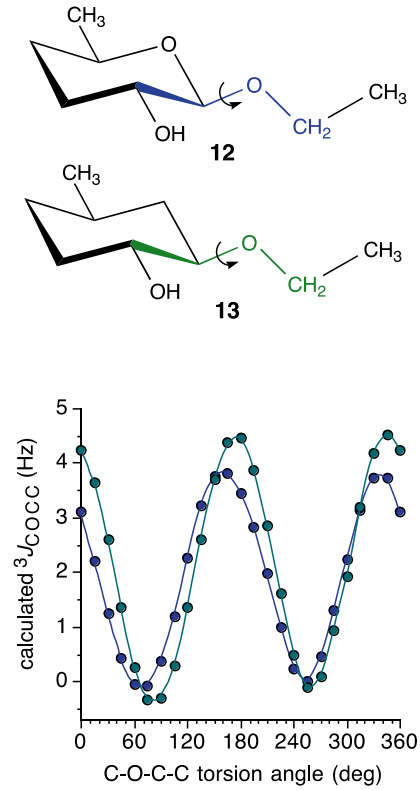
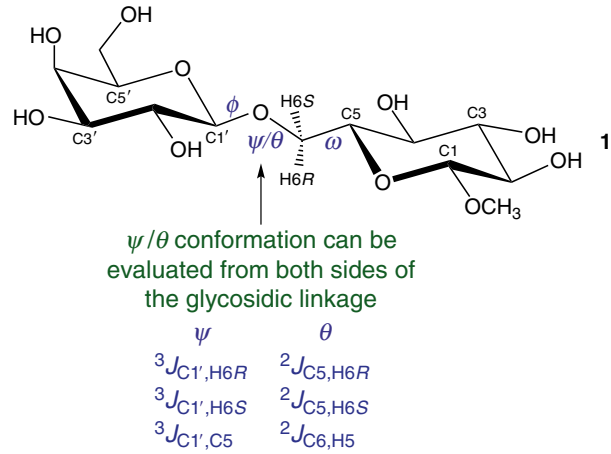


Fig. 5 Calculated $^3J_{COCC}$ values in **12** (*blue symbols*) and **13** (*green symbols*) as a function of the C2–C1–O1–CH_2 torsion angle. The pertinent coupling pathways are highlighted in *blue* in **12** and *green* in **13**. Data were taken from ref. 79

attributed to the effect of an internal electronegative substituent on a C–O–C–C coupling pathway [79]. These results demonstrate quantitatively that pathways "a" and "b" in Scheme 6 differ from pathway "c," necessitating the use of different Karplus equations to treat the two types of *J*-couplings. Earlier DFT studies of the dependence of $^{3}J_{C1',C3}$ and $^{3}J_{C1',C5}$ on ψ, and $^{3}J_{C2',C4}$ on ϕ, in structural mimics of **7** had suggested a small phase shift in the $^{3}J_{C2',C4}$ curve for unknown reasons. Subsequent DFT studies of model structures **12** and **13** confirmed the phase shift and, importantly, provided a structural explanation for the shift.

C. *Example 3*. Examples 1 and 2 illustrate the effect of *terminal* and *internal* substitution of electronegative atoms in C–C–C–C and C–C–O–C coupling pathways on $^{3}J_{CC}$ values. In Example 3, these findings are combined in the treatment of *O*-glycosidic linkage conformation involving an exocyclic hydroxymethyl group. This linkage occurs in the disaccharide, methyl β-D-galactopyranosyl-(1→6)-β-D-glucopyranoside **14** (Scheme 7), and is comprised of three rotatable bonds (ϕ, ψ and ω), in contrast to only two in **7**. It is commonly believed that linkages such as that found in **14** are more flexible than that in **7** due to the increased degrees of freedom, although experimental data supporting this claim is neither abundant nor definitive. In Scheme 7, the ψ bond in **14** is also defined as θ. This duality stems from the fact that ψ and θ specify the same bond *whose torsional behavior can be evaluated from both sides of the linkage*, namely, from three *J*-couplings ($^{3}J_{C1',H6R}$, $^{3}J_{C1',H6S}$, and $^{3}J_{C1',C5}$) on the anomeric carbon side (ψ) (similar to the ψ-sensitive $^{3}J_{COCC}$ in **7**; Example 2), and from three *J*-couplings ($^{2}J_{C5,H6R}$, $^{2}J_{C5,H6S}$, and $^{2}J_{C6,H5}$) on the aglycone side (θ). The three *J*-couplings on the aglycone side are geminal ^{13}C–^{1}H spin-couplings that are



Scheme 7

ordinarily not considered useful *conformational* probes, although they have been applied to make configurational assignments [64]. These geminal *J*-couplings, especially $^2J_{C5,H6R}$, and $^2J_{C5,H6S}$, are attractive because they depend strongly on the O5–C5–C6–O6 torsion angle (ω), as shown in Fig. 6 [19]. The calculated $^2J_{CCH}$ values in Fig. 6 were obtained using model structure **15** in which both ω and θ were rotated in 30° increments through 360°. The change in sign of both $^2J_{CCH}$ values leads to a dynamic range comparable to that observed for $^3J_{CCCH}$ (coupling signs can be determined from 2D ^{1}H–^{1}H TOCSY NMR spectra of ^{13}C-labeled compounds [80]). The spread of $^2J_{CCH}$ values at discrete values of ω reveals their dependence on θ, which is systematic and periodic. The latter behavior is revealed by examining the 3D hypersurfaces of $^2J_{C5,H6R}$, and $^2J_{C5,H6S}$ in which both the ω and θ dependencies are represented (Fig. 7). The surfaces in Fig. 7 can be fit to equations that correlate the magnitudes and signs of $^2J_{C5,H6R}$, and $^2J_{C5,H6S}$ with ω and θ [19].

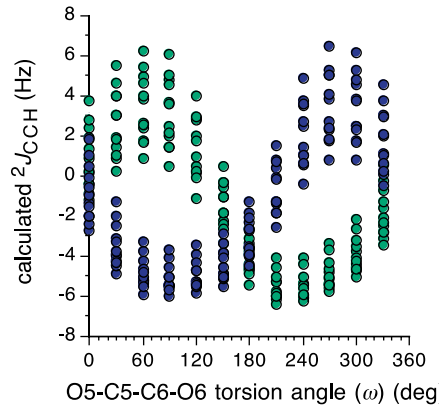
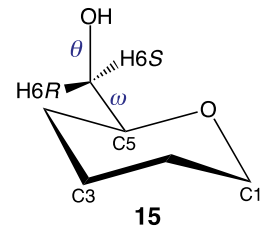


Fig. 6 The dependence of $^2J_{C5,H6R/S}$ on both ω and θ, determined by varying both torsion angles in **15** systematically through 360° in 30° increments. The vertical spread of points at discrete ω values demonstrates the sensitivity of both $^2J_{CCH}$ to θ. *Blue circles*, $^2J_{C5,H6R}$, *green circles* $^2J_{C5,H6S}$. Data were taken from ref. 19



15

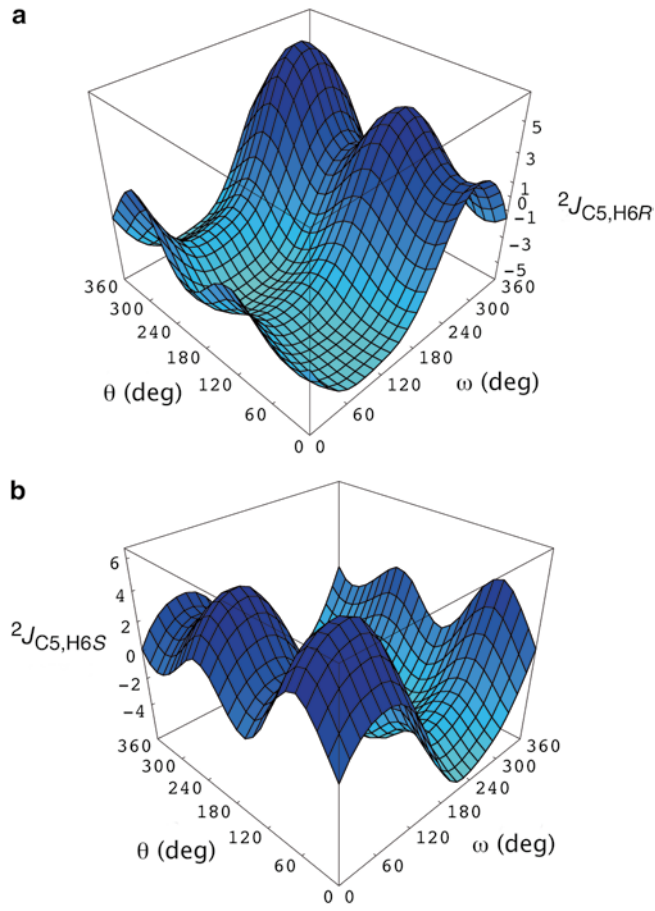
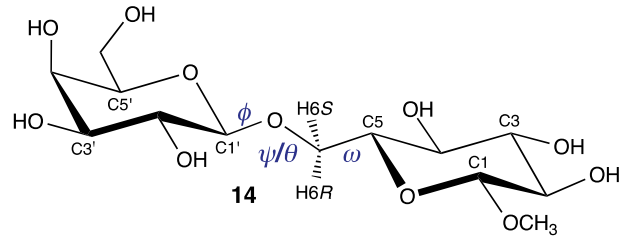


Fig. 7 (**a**) Hypersurface (ω/θ) calculated for $^2J_{C5,H6R}$ in **15**. (**b**) Hypersurface (ω/θ) calculated for $^2J_{C5,H6S}$ in **15**. These data were used to derive Eqs. 11 and 12 in which both ω and θ are variables, allowing determinations of correlated conformational behavior about both bonds. Data were taken from ref. 19

$$^2J_{C5,H6R} = -1.40 + 0.94\cos(\omega) - 4.38\sin(\omega) - 0.79\cos(2\theta) - 1.24\sin(2\theta) \quad (11)$$

$$^2J_{C5,H6S} = -1.32 + 2.24\cos(\omega) + 4.12\sin(\omega) - 0.80\cos(2\theta) + 1.24\sin(2\theta) \quad (12)$$

These *Karplus-like* equations are unconventional in that (a) they involve *two-bond J*-couplings and thus are not true Karplus equations, and (b) they contain two torsion angles as variables instead of one found in typical Karplus relationships. The second characteristic *allows these J-couplings to be used to determine correlated conformation about both ω and θ in solution*, that is, how conformation about ω affects conformation about θ. In one approach, the torsional properties of ω (i.e., the fractional populations of the *gg*, *gt*, and *tg* rotamers; *see* Scheme 3) can be



ϕ	ψ/θ	ω
$^2J_{C1',C6}$	$^2J_{C5,H6R}$	$^2J_{H5R,H5S}$
$^2J_{C2',H1'}$	$^2J_{C5,H6S}$	$^2J_{C5,H6R}$
$^3J_{H1',C6}$	$^2J_{C6,H5}$	$^2J_{C5,H6S}$
$^3J_{C2',C6}$	$^3J_{C1',H6R}$	$^3J_{H5,H6R}$
	$^3J_{C1',H6S}$	$^3J_{H5,H6S}$
	$^3J_{C1',C5}$	$^3J_{C4,H6R}$
		$^3J_{C4,H6S}$

Scheme 8

determined through conventional Karplus analyses of $^3J_{H5,H6R}$, $^3J_{H5,H6S}$, $^3J_{C4,H6R}$, and $^3J_{C4,H6S}$. These ω populations can be used in conjunction with Eqs. 11 and 12 to determine the θ populations. In this fashion, it is possible to establish how θ depends on different values of ω. In the context of disaccharide **14**, this information is particularly useful. Torsion angle ψ can be assessed from the anomeric carbon side of the linkage and also from the aglycone carbon side, permitting significant redundancy in the assessment of ψ/θ. In fact, for $1 \rightarrow 6$ linkages in general, there is considerable redundancy in evaluating the three linkage torsion angles via *J*-couplings, since at least 17 values are available (Scheme 8). This fact suggests that these linkages can be probed in greater detail than is possible for linkages involving only two torsions (e.g., 7). It is worth mentioning that, in the latter case, an effort was made to investigate possible secondary dependencies of the ϕ-dependent *J*-couplings on those sensitive to ψ, and vice versa, but none were found. Thus, there appears to be no opportunity to investigate correlated conformation about ϕ and ψ in these linkages using *J*-couplings.

9 Conclusions

Modern NMR studies of molecular structure increasingly rely on complementary theoretical methods to assist in the interpretation of NMR observables. This partnership is demanded by the intrinsic nature of NMR observables, which are populationally averaged

parameters that require some type of deconstruction into a conformational model. The latter deconstruction cannot be achieved without knowledge and insights supplied by theoretical inputs.

Purists might claim that, if theory were perfect, there would be no need for experiment. Fortunately for experimentalists, we are nowhere near this point at present, but this goal is unachievable without reliable means of experimentally validating predictions made by theoretical tools. An attractive feature of the NMR-DFT synergy is that the weaknesses of both techniques can be identified explicitly, and efforts focused on correcting or bypassing these deficiencies. In this context, the DFT method is not used to predict conformational behavior based on an evaluation of calculated conformational energies, which may seem a logical application. These energies are often unreliable when computed in saccharides due largely to complications arising from H-bonding and other non-covalent interactions that cannot be properly treated at present. Efforts to use DFT-derived energies exclusively to predict fractional conformational populations of saccharides in solution must therefore be considered semiquantitative at best.

When used in conjunction with NMR, however, the power of DFT calculations does not derive from their computed energies, but rather from their use in parameterizing NMR observables. In this capacity, it is possible to experimentally validate the DFT data. For example, validation can be achieved by calculating *J*-couplings in a molecule that exists in a fixed conformation in solution, so that the corresponding experimental *J*-couplings can be directly compared to the DFT predictions. Once these validations are performed on a sufficient number of test cases, the connection between NMR *J*-couplings and molecular structure become more reliable, precise, and quantitative. In other words, the energetic information provided by DFT calculations, which is problematic, becomes irrelevant. The problem shifts to the conformational model used to interpret the *J*-couplings, as eluded to above. Conformational model selection is not a trivial exercise, and indeed other computational tools such as MD, MC, and hybrid QM–MD [81] can help determine which conformational models are reasonable for a given type of conformational exchange.

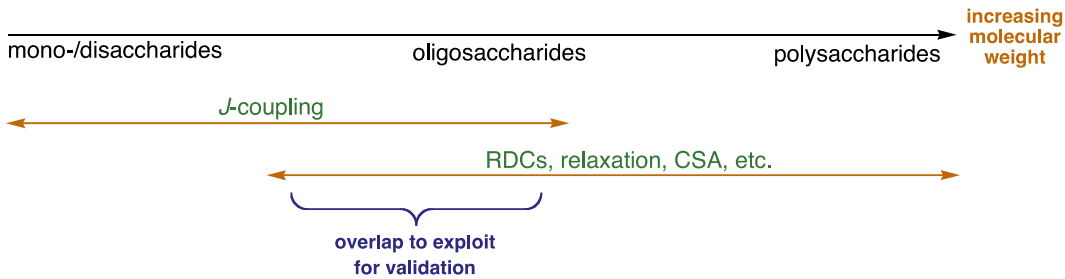
This article has discussed some of the fundamental features of DFT as it relates to the calculation of NMR spin-couplings. With the careful choice of model structures, new insights into the behavior of NMR *J*-couplings can be obtained from DFT calculations, and new applications discovered. The selection of model structures for DFT calculations, however, is a potential source of error, and careful thought needs to be devoted to selecting model structures that capture all of the structural elements believed to be important in determining the value of a particular *J*-coupling.

A distinguishing feature of *J*-couplings is that they report on the electronic (bonding) characteristics of molecules and thus potentially provide useful insights into their reactivities, whereas

residual dipolar couplings (RDCs) report solely on internuclear distances. In this unique capacity, *J*-couplings hold considerable promise not only as conformational (structural) probes, but also as functional probes. Some hints at the latter application have appeared in the literature, such as their use in investigating the strengths of H-bonding in solution [20]. With theoretical support provided by DFT calculations, it is likely that more such applications will be developed.

Accurate experimental measurements of NMR *J*-couplings, especially over multiple bonds, become increasingly difficult as the molecular weight of the molecule increases and its NMR signals broaden (Scheme 9). At some point, size precludes accurate *J*-coupling measurement (i.e., the line-widths are much greater than the line-splittings, rendering the latter unresolvable), and other NMR parameters such as residual dipolar couplings (RDCs) and nuclear spin-relaxation may be needed to assign structure. A molecular weight range exists where both *J*-coupling and RDC measurements can be made in the same molecule (Scheme 9). This overlap can be exploited to validate RDC methodology, which is presently less robust than that for *J*-couplings. Since DFT allows quantitative analysis of *J*-couplings on which to base structural assignments, it follows that these assignments can be compared to those derived from RDC analysis to validate or refine the latter. Thus, *J*-couplings can potentially play a central role in testing and validating structural conclusions drawn from RDC studies so that the latter parameters can be applied with greater confidence in molecules where *J*-coupling studies are precluded.

Parameterization of NMR *J*-couplings with assistance from DFT also permits the testing and validation of results from MD simulations. Using fractional populations calculated by MD, averaged *J*-couplings can be calculated from DFT-parameterized equations and compared to those observed experimentally. This approach is typically unidirectional at present in that the problem cannot be worked in reverse (i.e., generate fractional populations from MD that fit the *J*-couplings exactly) in cases where calculated and experimental couplings differ significantly. Nevertheless, the approach provides a valuable means of evaluating the accuracy of

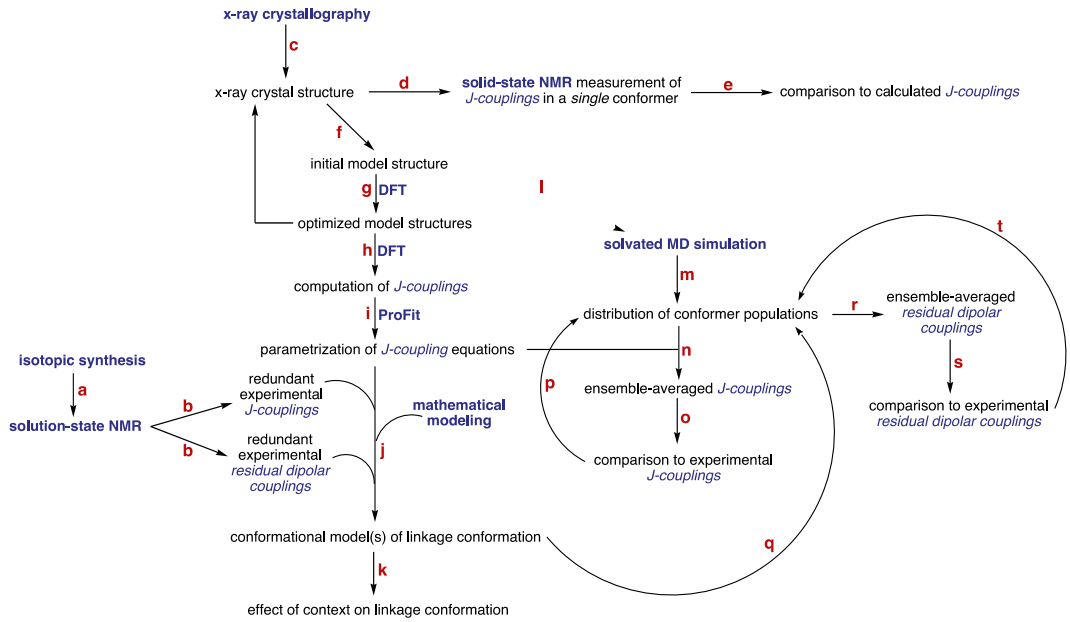


Scheme 9

MD results, which are often assumed to be correct without suitable independent validation. The latter practice is troubling, given that current structural studies of saccharides often depend heavily on MD analyses to assist in the interpretation of experimental data.

NMR *J*-couplings are affected significantly by the nature and number of substituents appended to the atoms of the coupling pathway. For example, three-bond (vicinal) $^{13}C–^{13}C$ spin–spin coupling constants involving C–O–C–C and C–C–C–C coupling pathways are influenced significantly by the presence of terminal and internal electronegative substituents, as shown in Figs. 4 and 5. Prior to the availability of reliable DFT methods, efforts were made to develop generalized Karplus and Karplus-like equations that were applicable to a specific type of coupling pathway decorated in different ways with different types of substituents. An example of this approach is the Haasnoot-Altona equation that is used to interpret $^{3}J_{HCCH}$ values in H–C–C–H coupling pathways containing different substituents on the two carbons [52]. However, continued improvements in CPU speed and efficiencies in computer code will probably render generalized treatments obsolete in the long term. In the future, specific tailor-made equations will be derived rapidly and accurately based on the exact coupling pathway under investigation, thereby eliminating the intrinsic errors often associated with formulating generalized equations.

With the above thoughts in mind, an integrated strategy for modern saccharide structure determination can be envisioned in which NMR, DFT, MD, x-ray crystallography and isotopic labeling play key roles (Scheme 10). Steps a and b involve the



Scheme 10

measurement of redundant *J*-couplings and/or RDCs (and perhaps other NMR observables) by NMR using isotopically labeled substrates. Step c involves x-ray crystallography to obtain a solid-state structure, which can inform the initial in silico structure in DFT calculations and MD simulations (steps f and l) and/or support solid-state NMR measurements of *J*-couplings in a single conformation [82] for comparison to DFT-calculated *J*-couplings (another potential validation tool) (steps d and e). Steps g, h and i were the focus of this review, namely, the use of DFT to parameterize *J*-coupling equations. These equations and the experimental data from Steps b are used to derive an experimentally based conformational model in solution (step j). The conformational element under scrutiny (e.g., a particular type of *O*-glycosidic linkage) could then be investigated in different structural contexts to determine how environment affects geometry (step k). MD simulations with solvent are informed by x-ray crystallography for the initial structure (step l), and produce conformer populations (step m), from which ensemble-averaged *J*-couplings are computed using DFT-parameterized equations (step n). The latter MD-derived *J*-couplings are compared to the experimental *J*-couplings (step o) to validate the MD results. Feedback loop p is envisioned in which an iterative process allows convergence of the MD-derived and experimental *J*-couplings. Depending on the success of step j, it might be possible to compare populations derived solely from experiment to those derived from MD simulations. Likewise, MD-derived populations are used to calculate ensemble-averaged RDCs (step r), which are compared to the experimental RDCs (Step s), with a feedback loop (step t) to allow convergence. Maximum redundancy in the measurement of *J*-couplings and RDCs (steps b) promotes step j; presumably, the larger the pool of measured *J*-couplings and RDCs, the greater the likelihood of deriving conformational models with minimal need for input from theory. This outcome would provide the best validation of the latter methodology (step q).

The connectivities shown in Scheme 10 to develop, test/validate and/or refine conformational models of saccharides, especially those that are flexible, include feedback loops that allow experiment to inform theory, and vice versa. Exploitation of these pathways and interactions should lead to convergence to an internally consistent description of solution behavior that is more reliable than one based on a single technique. The central roles of *J*-couplings, informed by DFT, are apparent (steps b, e, g, h, i, n, o, p, q). If a large number of parameterized *J*-couplings are available, mathematical treatments may permit conformational models to be derived without the need for theoretical input (step i). This goal has inspired and driven many of the studies described in this review, and remains a work in progress.

References

1. Serianni AS, Pierce J, Huang SG, Barker R (1982) Anomerization of furanose sugars: kinetics of ring-opening reactions by ^{1}H and ^{13}C saturation-transfer NMR spectroscopy. J Am Chem Soc 104:4037–4044
2. Hu X, Zhang W, Carmichael I, Seriani AS (2010) Amide *cis-trans* isomerization in aqueous solutions of methyl *N*-formyl-D-glucosaminides and methyl *N*-acetyl-D-glucosaminides: chemical equilibria and exchange kinetics. J Am Chem Soc 132:4641–4652
3. Pereira CS, Kony D, Baron R, Müller M, van Gunsteren WF, Hünenberger PH (2006) Conformational and dynamical properties of disaccharides in water: a molecular dynamics study. Biophys J 90:4337–4344
4. Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE, Matese JC, Perou CM, Hurt MM, Brown PO, Botstein D (2002) Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol Biol Cell 13:1977–2000
5. Kowalewski J, Mäler L, Widmalm G (1998) NMR relaxation studies of oligosaccharides in solution: reorientational dynamics and internal motion. J Mol Liq 78:255–261
6. Almond A, DeAngelis PL, Blundell CD (2005) Dynamics of hyaluronan oligosaccharides revealed by ^{15}N relaxation. J Am Chem Soc 127:1086–1087
7. Parr RG, Yang W (1995) Density-functional theory of the electronic structure of molecules. Annu Rev Phys Chem 46:701–728
8. Taubert S, Konschin HK, Sundholm D (2005) Computational studies of 13C NMR chemical shifts of saccharides. Phys Chem Chem Phys 7:2561–2569
9. Bose-Basu B, Zajicek J, Bondo G, Zhao S, Kubsch M, Carmichael I, Serianni AS (2000) Deuterium nuclear spin-lattice relaxation times and quadrupolar coupling constants in isotopically labeled saccharides. J Magn Reson 144:207–216
10. Bryce DL, Grishaev A, Bax A (2005) Measurement of ribose carbon chemical shift tensors for A-form RNA by liquid crystal NMR spectroscopy. J Am Chem Soc 127:7387–7396
11. Lemieux RU, Kullnig RK, Bernstein HJ, Schneider WG (1958) Configurational effects on the proton magnetic resonance spectra of six-membered ring compounds. J Am Chem Soc 80:6098–6105
12. Karplus M (1959) Contact electron-spin coupling of nuclear magnetic moments. J Chem Phys 30:11–15
13. Günther H (1995) NMR spectroscopy: basic principles, concepts and applications in chemistry, 2nd edn. Wiley, Chichester, pp 69–134
14. Snyder JR, Serianni AS (1986) D-idose: a one- and two-dimensional NMR investigation of solution composition and conformation. J Org Chem 51:2694–2702
15. Stenutz R, Carmichael I, Widmalm G, Serianni AS (2002) 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 67:949–958
16. Zhao H, Pan Q, Zhang W, Carmichael I, Serianni AS (2007) 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: 7071–7082
17. Otter A, Bundle DR (1995) Long-range ^{4}J and ^{5}J, including interglycosidic correlations in gradient-enhanced homonuclear COSY experiments of oligosaccharides. J Magn Reson B109:194–201
18. Barfield M, Dean AM, Fallick CJ, Spear RJ, Sternhell S, Westerman PW (1975) Conformational dependence and mechanisms for long-range hydrogen-hydrogen coupling constants over four bonds. J Am Chem Soc 97:1482–1492
19. Thibaudeau C, Stenutz R, Hertz B, Klepach T, Zhao S, Wu Q, Carmichael I, Serianni AS (2004) Correlated C-C and C-O bond conformations in saccharide hydroxymethyl groups: parametrization and application of redundant ^{1}H-^{1}H, ^{13}C-^{1}H, and ^{13}C-^{13}C NMR *J*-couplings. J Am Chem Soc 126:15668–15685
20. Maiti NC, Zhu Y, Carmichael I, Serianni AS, Anderson VE (2006) 1*J*CH correlates with alcohol hydrogen bond strength. J Org Chem 71:2878–2880
21. Jardetzky O (1980) On the nature of molecular conformations inferred from high-resolution NMR. Biochim Biophys Acta 621:227–232
22. Bukowski R, Morris LC, Woods RJ, Weimar T (2001) Synthesis and conformational analysis of the T-antigen disaccharide β-D-Gal-(1→3)-β-D-GalNAcOMe. Eur J Org Chem 2001:2697–2705
23. Woods RJ, Pathiaseril A, Wormald MR, Edge CJ, Dwek RA (1998) The high degree of internal flexibility observed for an oligomannose

oligosaccharide does not alter the overall topology of the molecule. Eur J Biochem 258:372–386

24. Dowd MK, Kiely DE, Zhang J (2011) Monte Carlo-based searching as a tool to study carbohydrate structure. Carbohydr Res 346:1140–1148
25. Woods RJ (1998) Computational carbohydrate chemistry: what theoretical methods can tell us. Glycoconj J 15:209–216
26. Hsu C-H, Hung S-C, Wu C-Y, Wong C-H (2011) Toward automated oligosaccharide synthesis. Angew Chem Int Ed 50:11872–11923
27. Zhang W, Oliver AG, Serianni AS (2010) Methyl β-D-galactopyranosyl-(1→4)-β-D-allopyranoside tetrahydrate. Acta Crystallogr C C66:o484–o487
28. Stenutz R, Shang M, Serianni AS (1999) Methyl β-Lactoside (methyl 4-*O*-β-D-galactopyranosyl-β-D-glucopyranoside) methanol solvate. Acta Crystallogr C C55:1719–1721
29. Zhang W, Oliver AG, Vu HM, Duman JG, Serianni AS (2012) Methyl 4-*O*-β-D-mannopyranosyl β-D-xylopyranoside. Acta Crystallogr C C68:o502–o506
30. Ham JT, Williams DG (1970) The crystal and molecular structure of methyl β-cellobiosidemethanol. Acta Crystallogr C B26:1373–1383
31. Pan Q, Noll BC, Serianni AS (2005) Methyl 4-*O*-β-D-galactopyranosyl α-D-glycopyranoside (methyl α-lactoside). Acta Crystallogr C C61:o674–o677
32. Zhang W, Oliver AG, Serianni AS (2012) Disorder and conformational analysis of methyl β-D-galactopyranosyl-(1→4)-β-D-xylopyranoside. Acta Crystallogr C C68:o7–o11
33. Hu X, Pan Q, Noll BC, Oliver AG, Serianni AS (2010) Methyl 4-*O*-β-D-galactopyranosyl α-D-mannopyranoside methanol 0.375-solvate. Acta Crystallogr C C66:o67–o70
34. Pachler KGR (1971) Extended Hückel theory MO calculations of proton-proton coupling constants – II: the effect of substituents on vicinal couplings in monosubstituted ethanes. Tetrahedron 27:187–199
35. Galan MC, Venot AP, Glushka J, Imberty A, Boons G-J (2002) Alpha-(2→6)-sialyltransferase-catalyzed sialylations of conformationally constrained oligosaccharides. J Am Chem Soc 124:5964–5973
36. Hohenberg P, Kohn W (1964) Inhomogeneous electron. Gas Phys Rev 136:B864–B871
37. Kohn W, Sham LJ (1965) Self-consistent equations including exchange and correlation effects. Phys Rev 140:A1133–A1138
38. Perdew JP, Wang Y (1986) Accurate and simple density functional for the electronic exchange energy: generalized gradient approximation. Phys Rev B 33:8800–8802
39. Becke AD (1993) Density-functional thermochemistry. III. The role of exact exchange. J Chem Phys 98:5648–5652
40. Lee C, Yang W, Parr RG (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys Rev B 37:785–789
41. Pople JA, McIver JW Jr, Ostlund NS (1967) Finite perturbation theory for nuclear spin coupling constants. Chem Phys Lett 1:465–466
42. Pople JA, McIver JW Jr, Ostlund NS (1968) Self-consistent perturbation theory. I. Finite perturbation methods. J Chem Phys 49:2960–2964
43. Pople JA, McIver JW Jr, Ostlund NS (1968) Self-consistent perturbation theory. II. Nuclear spin coupling constants. J Chem Phys 49:2965–2970
44. Ramsey NF (1953) Electron coupled interactions between nuclear spins in molecules. Phys Rev 91:303–307
45. Helgaker T, Watson M, Handy NC (2000) Analytical calculation of nuclear magnetic resonance indirect spin–spin coupling constants at the generalized gradient approximation and hybrid levels of density-functional theory. J Chem Phys 113:9402–9409
46. Miertus S, Scrocco E, Tomasi J (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of *ab inito* molecular potentials for the provision of solvent effects. Chem Phys 55:117–129
47. Cancés E, Mennucci B (1998) New applications of integral equations methods for solvation continuum models: ionic solutions and liquid crystals. J Math Chem 23:309–326
48. Cancés E, Mennucci B, Tomasi J (1997) A new integral equation formalism for the polarizable continuum model: theoretical background and applications to isotropic and anisotropic dielectrics. J Chem Phys 107:3032–3041
49. Mennucci B, Cancés E, Tomasi J (1997) Evaluation of solvent effects in isotropic and anisotropic dielectrics and in ionic solutions with a unified integral equation method: theoretical bases, computational implementation, and numerical applications. J Phys Chem B 101:10506–10517
50. Tomasi J, Mennucci B, Cammi R (2005) Quantum mechanical continuum solvation models. Chem Rev 105:2999–3093

51. Bose B, Zhao S, Stenutz R, Cloran F, Bondo PB, Bondo G, Hertz B, Carmichael I, Serianni AS (1998) Three-bond C-O-C-C spin-coupling constants in carbohydrates: development of a Karplus relationship. J Am Chem Soc 120:11158–11173
52. Haasnoot CAG, de Leeuw FAAM, Altona C (1980) The relationship between proton-proton NMR coupling constants and substituent electronegativities—I: an empirical generalization of the Karplus equation. Tetrahedron 36:2783–2792
53. Altona C, Ippel JH, Hoekzema AJAW, Erkelens C, Groesbeek M, Donders LA (1989) Relationship between proton-proton NMR coupling constants and substituent electronegativities. V. Empirical substituent constants deduced from ethanes and propanes. Magn Reson Chem 27:564–576
54. Altona C, Francke R, de Haan R, Ippel JH, Daalmans GJ, Hoekzema AJAW, van Wijk J (1994) Empirical group electronegativities for vicinal NMR proton-proton couplings along a C-C bond: solvent effects and reparameterization of the Haasnoot equation. Magn Reson Chem 32:670–678
55. Carmichael I, Chipman DM, Podlasek CA, Serianni AS (1993) Torsional effects on the one-bond 13C-13C spin coupling constant in ethylene glycol: insights into the behavior of ^{1}JCC in carbohydrates. J Am Chem Soc 115:10863–10870
56. Church T, Carmichael I, Serianni AS (1996) Two-bond ^{13}C-^{13}C spin-coupling constants in carbohydrates: effect of structure on coupling magnitude and sign. Carbohydr Res 280:177–186
57. Serianni AS, Bondo PB, Zajicek J (1996) Verification of the projection resultant method for two-bond ^{13}C-^{13}C coupling sign determinations in carbohydrates. J Magn Reson Ser B 112:69–74
58. Klepach T, Serianni AS (unpublished results)
59. Cloran F, Carmichael I, Serianni AS (2000) ^{2}JCOC spin-spin coupling constants across glycosidic linkages exhibit a valence bond-angle dependence. J Am Chem Soc 122:396–397
60. Cloran F, Carmichael I, Serianni AS (1999) 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 121:9843–9851
61. Bose-Basu B, Klepach T, Bondo G, Bondo PB, Zhang W, Carmichael I, Serianni AS (2007) ^{13}C–^{13}C NMR spin-spin coupling constants in saccharides: structural correlations involving all carbons in aldohexopyranosyl rings. J Org Chem 72:7511–7522
62. Müller N, Pritchard DE (1959) C13 splittings in proton magnetic resonance spectra. I. Hydrocarbons. J Chem Phys 31:768–771
63. Serianni AS, Wu J, Carmichael I (1995) One-Bond ^{13}C-^{1}H spin-coupling constants in aldofuranosyl rings: effect of conformation on coupling magnitude. J Am Chem Soc 117:8645–8650
64. Bock K, Pedersen C (1977) Two- and three-bond ^{13}C-^{1}H couplings in some carbohydrates. Acta Chem Scand B 31:354–358
65. Klepach TE, Carmichael I, Serianni AS (2005) Geminal ^{2}JCCH spin-spin coupling constants as probes of the ϕ glycosidic torsion angle in oligosaccharides. J Am Chem Soc 127:9781–9793
66. Podlasek CA, Wu J, Stripe WA, Bondo PB, Serianni AS (1995) [^{13}C]-Enriched methyl aldopyranosides: structural interpretations of ^{13}C–^{1}H spin-coupling constants and ^{1}H chemical shifts. J Am Chem Soc 117:8635–8644
67. ChemBio3D. www.cambridgesoft.com/Ensemble_for_Biology/ChemBio3D/Default.aspx
68. Spartan. www.wavefun.com/products/spartan.html
69. Dennington R, Keith T, Millam J (2009) GaussView, Version 5. Semichem Inc., Shawnee Mission, KS
70. GAMESS. www.msg.ameslab.gov/gamess/
71. Jaguar. www.schrodinger.com/products/14/7/
72. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA Jr, Peralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ (2009) Gaussian 09, Revision A.1. Gaussian Inc., Wallingford, CT
73. Hehre WJ, Ditchfield R, Pople JA (1972) Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. J Chem Phys 56:2257–2261
74. York DM, Karplus M (1999) A smooth solvation potential based on the conductor-like

screening model. J Phys Chem A 103:11060–11079

75. Sychrovsky V, Gräfenstein J, Cremer D (2000) Nuclear magnetic resonance spin-spin coupling constants from coupled perturbed density functional theory. J Chem Phys 113:3530–3547
76. Carmichael I (1993) *Ab initio* quadratic configuration interaction calculation of indirect NMR spin-spin coupling constants. J Phys Chem 97:1789–1792
77. King-Morris MJ, Serianni AS (1987) 13C NMR studies of [1-^{13}C] aldoses: empirical rules correlating pyranose ring configuration and conformation with ^{13}C chemical shifts and ^{13}C-^{13}C spin couplings. J Am Chem Soc 109:3501–3508
78. Wu J, Bondo PB, Vuorinen T, Serianni A (1992) ^{13}C–^{13}C spin coupling constants in aldoses enriched with ^{13}C at the terminal hydroxymethyl carbon: effect of coupling pathway structure on JCC in carbohydrates. J Am Chem Soc 114:3499–3505
79. Zhao H, Carmichael I, Serianni AS (2008) Oligosaccharide trans-glycoside $^{3}J_{COCC}$ Karplus curves are not equivalent: effect of internal electronegative substituents. J Org Chem 73:3255–3257
80. Serianni AS, Podlasek CA (1994) 13C-1H spin-coupling constants in carbohydrates: magnitude and sign determinations via 2D NMR methods. Carbohydr Res 259:277–282
81. Muslim A-M, McNamara JP, Abdel-Aal H, Hillier IH, Bryce RA (2006) QM/MM simulations of carbohydrates. In: NMR spectroscopy and computer modeling of carbohydrates. ACS Symposium Series 2006, vol 930. American Chemical Society. pp 186–202
82. Thureau P, Mollica G, Ziarelli F, Viel S (2013) Selective measurements of long-range homonuclear *J*-couplings in solid-state NMR. J Magn Reson 231:90–94