PCCP

One-Bond ^{13}C-^{13}C Spin-Coupling Constants in Saccharides: A Comparison of Experimental and Calculated Values By Density Functional Theory Using Solid-State ^{13}C NMR and X-Ray Crystallography

Journal:	*Physical Chemistry Chemical Physics*
Manuscript ID	CP-ART-11-2022-005363.R1
Article Type:	Paper
Date Submitted by the Author:	04-Apr-2023
Complete List of Authors:	Serianni, Anthony; University of Notre Dame, Department of Chemistry and Biochemistry Tetrault, Timothy; University of Notre Dame, Department of Chemistry and Biochemistry Meredith, Reagan; University of Notre Dame, Department of Chemistry and Biochemistry Yoon, Mi-Kyung; University of Notre Dame, Department of Chemistry and Biochemistry Canizares, Christopher; University of Notre Dame, Department of Chemistry and Biochemistry Oliver, Allen; University of Notre Dame, Department of Chemistry and Biochemistry Carmichael, Ian; University of Notre Dame, Radiation Laboratory and Department of Chemistry & Biochemistry

One-Bond ^{13}C-^{13}C Spin-Coupling Constants in Saccharides: A Comparison of Experimental and Calculated Values By Density Functional Theory Using Solid-State ^{13}C NMR and X-Ray Crystallography

Timothy Tetrault,[1] Reagan J. Meredith,[1] Mi-Kyung Yoon,[1,4] Christopher Canizares,[1] Allen G. Oliver,[2] Ian Carmichael,[3] and Anthony S. Serianni[1*]

[1]Department of Chemistry and Biochemistry, [2]Molecular Structure Facility, and [3]Radiation Laboratory, University of Notre Dame, Notre Dame, IN 46556-5670 USA; [4]Omicron Biochemicals, Inc., South Bend, IN 46617 USA

*Corresponding author: aseriann@nd.edu

ABSTRACT

Methyl aldohexopyranosides were ^{13}C-labeled at contiguous carbons, crystallized, and studied by single-crystal X-ray crystallography and solid-state ^{13}C nuclear magnetic resonance (NMR) spectroscopy to examine the degree to which density functional theory (DFT) can calculate one-bond ^{13}C-^{13}C spin-coupling constants ($^1J_{CC}$) in saccharides with sufficient accuracy to permit their use in *MA'AT* analysis, a newly-reported hybrid DFT/NMR method that provides probability distributions of molecular torsion angles in solution (Zhang *et al.*, *J. Phys. Chem.* B, 2017, **121**, 3042–3058; Meredith *et al.*, *J. Chem. Inf. Model.*, 2022, **62**, 3135–3141). Experimental $^1J_{CC}$ values in crystalline samples of the doubly ^{13}C-labeled compounds were measured by solid-state ^{13}C NMR and compared to those calculated from five different DFT models: (1) $^1J_{CC}$ values calculated from single structures identical to those observed in crystalline samples by X-ray crystallography (all atom refinement); (2) $^1J_{CC}$ values calculated from the single structures in (1) but after Hirshfeld atom refinement (HAR); (3) $^1J_{CC}$ values calculated from the single structures in (1) after DFT-optimization of hydrogen atoms only; and (4 and 5) $^1J_{CC}$ values calculated in rotamers of torsion angle θ_2 (C1–C2–O2–O2H) or ω (C4–C5–C6–O6) from which either specific or generalized parameterized equations were obtained and used to calculate $^1J_{CC}$ values in the specific θ_2 or ω rotamers observed in crystalline samples. Good qualitative agreement was observed between calculated $^1J_{CC}$ values and those measured by solid-state ^{13}C NMR regardless of the DFT model, but in no cases were calculated $^1J_{CC}$ values quantitative, differing (over-estimated) on average by 4–5% from experimental values. These findings, and those reported recently from solution NMR studies (Tetrault *et al.*, *J. Phys. Chem.* B **2022**, *126*, 9506–9515), indicate that improvements in DFT calculations are needed before calculated $^1J_{CC}$ values can be used directly as reliable constraints in *MA'AT* analyses of saccharides in solution.

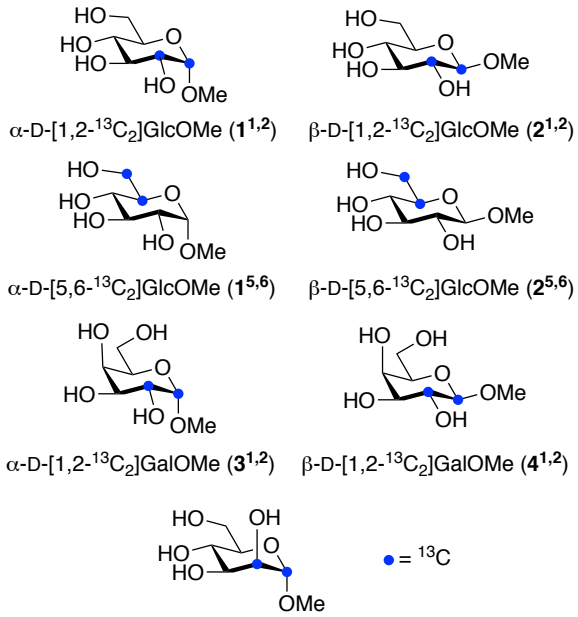
INTRODUCTION

Experimental spin-spin coupling constants (*J*-couplings) measured by NMR spectroscopy are widely used to determine the structural properties of saccharides in solution.[1–3] *J*-Couplings are sensitive to valence bond angles, bond torsion angles, and bond lengths within and adjacent to the coupling pathway. Vicinal (three-bond) *J*-couplings ($^3J_{ab}$) depend strongly on the torsion angle between the two coupled nuclei, and experimental values can be interpreted quantitatively via Karplus or Karplus-like relationships, revealing conformational preferences in solution.[4–10] Geminal *J*-couplings ($^2J_{ab}$) are used in structure determinations less frequently, and are affected by the configurations of the terminal coupled carbons and by the configuration of the central carbon. In the latter regard, $^2J_{CCC}$ values depend strongly on the rotameric properties of the C–O bond involving the central carbon, and less so on the rotameric properties of the C–O bonds involving the coupled carbons.[11–13] Direct (one-bond) *J*-couplings ($^1J_{ab}$) depend strongly on bond length, with larger values associated with shorter bond lengths (greater *s*-character).[1,14–16] Despite suggestions that ^1J values could serve as valuable structural constraints in saccharides,[17,18] detailed conformational interpretations of them have proven challenging.[19–24]

Experimental determinations of saccharide conformation often rely on computational methods such as density functional theory (DFT) or molecular dynamics (MD) simulations.[25–29] Current approaches combine experimental parameters with appropriately restrained calculations to derive a model that best matches the experimental data.[29–34] While this approach has proven valuable, it is not without limitations stemming, in part, from limited experimental validations of the computational methods. The structural dependencies of experimental parameters may not be understood quantitatively and/or the quantity of experimental data used to derive a conformational model may be limited, placing greater reliance on the model obtained from calculations. X-Ray crystallography and solid-state ^{13}C NMR spectroscopy have been used synergistically to determine the structural properties of carbohydrates in solution by addressing some of the limitations discussed above.[19–24]

MA'AT analysis,[35–41] a new experimental method to model molecular conformation in solution, combines DFT calculations and experimental NMR *J*-couplings to give probability distributions of molecular torsion angles in solution. *MA'AT* probability distributions are superimposable on those obtained by MD simulation, allowing direct comparisons of experimental and calculated models.[35–41] The method uses ^{2}J and ^{3}J values primarily to model various conformational elements in saccharides (e.g., *O*-glycosidic linkages,[35–37] ring conformation,[39] side-chain conformation[38,41]). To extend *MA'AT* analysis to multi-state models, it would be useful to employ $^{1}J_{CC}$ and $^{1}J_{CH}$ values as inputs due to their high abundance in saccharides and their attractive conformational properties (*e.g.*, large dynamic ranges).

The study described here builds on recent solution[42] and solid-state[11,43] NMR studies to establish the usefulness of J_{CC} values in *MA'AT* analysis. Solid-state ^{13}C NMR spectroscopy (ssNMR) provides experimental J_{CC} values in conformationally-defined ^{13}C-labeled crystalline compounds that are devoid of effects from motional averaging in solution. Direct comparisons between these experimental values and DFT-calculated $^{1}J_{CC}$ values in corresponding *in silico* structures can be used to evaluate the accuracy of $^{1}J_{CC}$ values calculated by DFT. In this study, crystalline doubly ^{13}C-labeled methyl aldohexopyranosides **1**^{1,2}, **2**^{1,2}, **1**^{5,6}, **2**^{5,6}, **3**^{1,2}, **4**^{1,2}, and **5**^{1,2} (superscripts denote the labeled carbons) (Scheme 1) were used to measure one-bond ^{13}C-^{13}C spin-coupling constants between the contiguously labeled carbons by solid-state ^{13}C NMR. This group of monosaccharides was chosen due to their ease of synthesis and crystallization, their structural simplicity, and their structural diversity (different anomeric configurations, configurations at other ring carbons, and

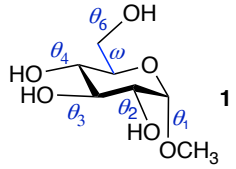


Scheme 1. Chemical structures of doubly ^{13}C-labeled **1**^{1,2}, **2**^{1,2}, **1**^{5,6}, **2**^{5,6}, **3**^{1,2}, **4**^{1,2} and **5**^{1,2}. The one-bond ^{13}C-^{13}C spin-coupling between the two labeled carbons was investigated in each compound. Superscripts on the compound numbers denote the labeled carbons.

C–C coupling pathways). DFT calculations were performed on *in silico* models of **1**–**5** under different conditions, and calculated $^{1}J_{CC}$ values were compared to those obtained by solid-state ^{13}C NMR to determine the extent to which DFT reproduces the experimental values. The results were then compared to those obtained from recent solution NMR studies that addressed the same problem using different but complementary experimental approaches.[42]

EXPERIMENTAL

A. *General.* Reagents, materials and synthetic procedures, X-ray crystallographic data, conditions for solution- and solid-state ^{13}C NMR measurements, and representative solution- and solid-state NMR spectra can be found in the Supporting Information.



Scheme 2. Structure of methyl α-D-glucopyranoside **1** showing four exocyclic C–O torsion angles θ_1 (C2–C1–O1–CH_3), θ_2 (C1–C2–O2–H), θ_3 (C2–C3–O3–H, θ_4 (C3–C4–O4–H), and two side-chain torsion angles ω (C4–C5–C6–O6) and θ_6 (C5–C6–O6–H).

B. NMR Measurements. High-resolution solution 1D ^{13}C{^{1}H} NMR spectra were obtained on **1**–**5** using 5-mm NMR tubes and a Varian DirectDrive 600 MHz FT-NMR spectrometer equipped with a 5 mm ^{1}H-^{19}F/^{15}N-^{31}P AutoX dual broadband probe. 1D ^{13}C{^{1}H} spectra were collected in $^{2}H_2O$ at 20 °C with ~10,000 Hz spectral windows and ~1.5 s recycle times, and were zero-filled twice to give final digital resolutions of ~0.08 Hz per point. Chemical shifts are reported in ppm (δ) relative to the ^{1}H signal of residual D_2O at δ 4.81 ppm. ^{13}C-^{13}C spin-couplings were obtained by analysis of the doublet character of the two intense signals arising from the mutually coupled ^{13}C-labeled carbons in each compound (Figures S1–S5, Supporting Information). Since one of the ^{13}C-labeled carbons in **1**^{1,2}, **2**^{1,2}, **3**^{1,2}, **4**^{1,2} and **5**^{1,2} is an anomeric carbon, non-first-order effects on the measurements of the $^{1}J_{C1,C2}$ values were negligible. Since the chemical shifts of C5 and C6 of the ^{13}C-labeled carbons in **1**^{5,6} and **2**^{5,6} are >1,600 Hz different in chemical shift at 150 MHz, non-first-order effects on the measurements of the $^{1}J_{C5,C6}$ values were negligible.

One-bond ^{13}C-^{13}C spin-coupling constants ($^{1}J_{CC}$) were measured on crystalline samples of ^{13}C-labeled compounds using a JEOL ECX-300 solid-state FT-NMR spectrometer operating at a

^{1}H frequency of 300 MHz and equipped with a 3.2-mm magic angle spinning (MAS) probe, as previously described (Figures S3 and S5, and Table S1, Supporting Information).[11,43] Briefly, 20–50 mg of each ^{13}C-labeled compound (30–60 *w*/*w* %) was mixed with KBr as an internal reference for magic angle calibration. The MAS frequency was set to 15.9 or 16 kHz. To confirm the sample identity and purity, 1D cross-polarization (CP) MAS ^{13}C NMR spectra were obtained on each sample using a 1.97 μs ^{1}H 90° pulse and a 1.5 ms contact time, and 32 scans were collected. For measurement of solid-state ^{13}C-^{13}C spin-couplings, the reference signals (S_o) were obtained using a single-band selective 180° Gaussian function, and the *J*-modulated signals (S) were obtained using a double-band selective 180° Gaussian function multiplied by a cosine wave. The total spin-echo delay was increased from 5 to 56 ms in 3-ms increments. Spectra were collected with spectral windows of 300 ppm, a total of 8 scans, and a 20 s relaxation delay. The data were processed using JEOL Delta v5.0.4.4 NMR processing software. Signal intensities were measured to give the intensity ratios of the *J*-modulated to reference signals (S/S_o), which were plotted against the total echo interval. Three or more measurements were made on each compound.

C. Hirshfeld Atom Refinement (HAR) of X-ray Diffraction Data for Structures ***1****[1,2],* ***2****[1,2],* ***1****[5,6],* ***2****[5,6],* ***3****[1,2],* ***4****[1,2], and* ***5****[1,2]*. Structural parameters for hydrogen atoms in X-ray diffraction data of all compounds were refined using the Hirshfeld atom refinement (HAR) method.[44,45] The diffraction data of **1**–**5** were refined using the NoSpherA2 module[46] and ORCA 5.0.3[47] for DFT calculations in the Olex2[48] software. The B3LYP functional was used in combination with the cc-pVTZ basis set.[49] Integration accuracy was set to low, SCF threshold was set to SloppySCF, and the SCF Strategy was set to EasyConv.

CALCULATIONS

A. $^{1}J_{CC}$ Calculations in Fixed Structures (Set 1). $^{1}J_{CC}$ calculations were performed on single fixed structures **1**^{c}**–5**^{c} (the superscript ‘‘c’’ denotes an *in silico* structure in a conformation identical to that found in the crystal structure of the corresponding compound). X-ray crystal

structures of **1**[1,2], **2**[1,2], **1**[5,6], **2**[5,6], **3**[1,2], **4**[1,2], and **5**[1,2] were obtained (all atoms in these structures were allowed to refine, including the hydrogen atoms) and the Cartesian coordinates of each were used directly in $^1J_{CC}$ calculations. The X-ray structures were not geometry-optimized prior to these $^1J_{CC}$ calculations. $^1J_{CC}$ values were calculated in *Gaussian*16[50] as described previously.[11,43] The Fermi contact,[51–53] diamagnetic and paramagnetic spin–orbit, and spin–dipole[51] terms were calculated using the B3LYP functional and a [5s2p1dl3s1p] basis set,[20,25] and the reported calculated $^1J_{C1,C2}$ values were unscaled. The *J*-coupling calculations included the effect of solvent water, which was treated using the Self-Consistent Reaction Field (SCRF)[54] and the Integral Equation Formalism (polarizable continuum) model (IEFPCM)[55] as implemented in *Gaussian*16.

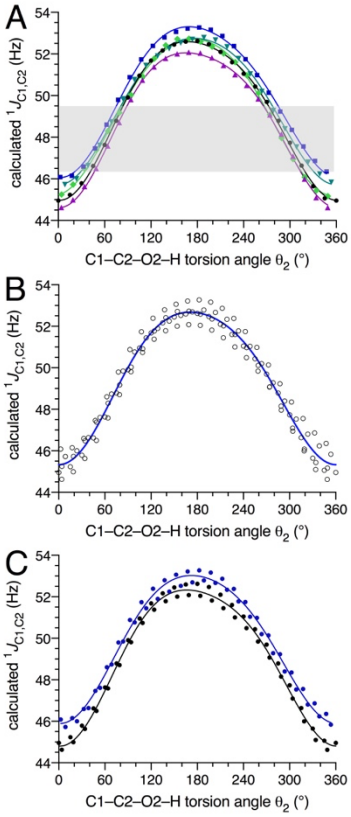


Figure 1. (A) Plot of calculated $^1J_{C1,C2}$ values as a function of θ_2 in **1**[c] (black), **2**[c] (blue), **3**[c] (purple), **4**[c] (green) and **5**[c] (lime). The curves were generated from eqs. [2], [3], [4], [5] and [6], respectively. The shaded grey area shows the range of $^1J_{C1,C2}$ values observed experimentally (Table 1). (B) Data shown in (A) fit to a generalized equation (eq. [9]). (C) Data shown in (A) separated into data for α- (black) (**1**[c] and **3**[c]) and β- (blue) (**2**[c] and **4**[c]) anomers, showing slightly larger calculated $^1J_{C1,C2}$ values in the latter.

B. $^1J_{CC}$ Calculations in HAR Structures (Set 2). Cartesian coordinates obtained from HAR of **1**[1,2], **2**[1,2], **1**[5,6], **2**[5,6], **3**[1,2], **4**[1,2], and **5**[1,2] were used directly in $^1J_{CC}$ calculations, the latter conducted as described for Set 1.

C. Hydrogen Atom Optimization Prior To $^1J_{CC}$ Calculations (Set 3). The Cartesian coordinates of the X-ray crystal structures of **1**[1,2], **2**[1,2], **1**[5,6], **2**[5,6], **3**[1,2], **4**[1,2], and **5**[1,2] (Set 1) were used as input for geometry optimization by DFT. During optimization, the coordinates of all atoms except hydrogen were fixed as well as torsion angle θ_2 (C1–C2–O2–O2H), and all hydrogen

structural parameters were allowed to optimize. $^1J_{CC}$ values were calculated in the resulting optimized structures as described for Set 1.

*D. Calculated $^1J_{C1,C2}$ in Model Structures **1**^{c} and **2**^{c} (Sets 4 and 5)*

D.1. Geometry Optimization. Geometry optimizations were conducted in *Gaussian*16[44] using the B3LYP functional[56] and 6-311+g(d,p) basis set[57,58]. In **1**^{c}, torsion angle θ_1 (C2–C1–O1–CH_3) (Scheme 2) was fixed at the angle observed in crystalline **1**[1,2] (–75.47°). Torsion angle θ_2 (C1–C2–O2–O2H) was set initially at 180° and rotated in 15° increments through 360° to give 24 optimized structures. The remaining exocyclic torsion angles θ_3–θ_6 (Scheme 2) were allowed to rotate freely during geometry optimization. The effect of solvent water was included in these calculations using the Self-Consistent Reaction Field (SCRF)[54] and the Integral Equation Formalism (polarizable continuum) model (IEFPCM)[55] as implemented in *Gaussian*16.

Similar geometry optimizations were conducted on **2**^{c}, with θ_1 (C2–C1–O1–CH_3) fixed at the angle observed in crystalline **2**[1,2] (92.02°), and θ_2 rotated in 15° increments through 360°. The remaining exocyclic torsion angles θ_3–θ_6 were allowed to rotate freely during geometry optimization.

D.2. $^1J_{C1,C2}$ Calculations. $^1J_{C1,C2}$ values in **1**^{c} and **2**^{c} as a function of θ_2 were calculated in *Gaussian*16 as described above.

*D.3. Equation Parameterization of $^1J_{C1,C2}$ as a Function of θ_2 in **1**^{c} and **2**^{c}*. Calculated $^1J_{C1,C2}$ values in the structures were plotted (Figure 1) and the curves were fit to a modified Karplus-like equation (eq. [1]) using R (see the Supporting Information for a brief discussion of eq. [1]).

$$^1J_{CC}\,(\mathrm{Hz}) = k + c_1 \cos\theta_2 + s_1 \sin\theta_2 + c_2 \cos 2\theta_2 + s_1 \sin 2\theta_2 \qquad \text{eq. [1]}$$

The following parameterized equation for **1**^{c} was obtained (eq. [2]).

$$^{1}J_{C1,C2}\,(\text{Hz}) = 49.92 - 3.66\cos\theta_2 + 0.02\sin\theta_2 - 0.58\cos 2\theta_2 - 0.08\sin 2\theta_2$$
RMSD = 0.14 Hz eq. [2]

The following parameterized equation for **2**^{c} was obtained (eq. [3]).

$$^{1}J_{C1,C2}\,(\text{Hz}) = 50.14 - 3.63\cos\theta_2 + 0.09\sin\theta_2 - 0.40\cos 2\theta_2 - 0.15\sin 2\theta_2$$
RMSD = 0.12 Hz eq. [3]

The goodness-of-fit of each equation to the calculated *J*-couplings was expressed as a root mean squared deviation (RMSD).

*E. Calculated $^{1}J_{C1,C2}$ Values in Model Structures **3**^{c}, **4**^{c} and **5**^{c}.* DFT calculations on **3**^{c}–**5**^{c} were performed as described above for **1**^{c} and **2**^{c}. For **3**^{c}, θ_1 was fixed at the angle observed in crystalline **3**^{1,2} (123.09°) and θ_2 was rotated in 15° increments through 360°. For **4**^{c}, θ_1 was fixed at the angle observed in crystalline **4**^{1,2} (98.04°) and θ_2 was rotated in 15° increments through 360°. For **5**^{c}, θ_1 was fixed at the angle observed in crystalline **5**^{1,2} (169.83°) and θ_2 was rotated in 15° increments through 360°. The remaining exocyclic torsion angles θ_3–θ_6 were allowed to rotate freely during geometry optimization. Calculated $^{1}J_{C1,C2}$ values were plotted as a function of θ_2 (Figure 1) and the resulting curves were fit to give eq. [4] for **3**^{c}, eq. [5] for **4**^{c}, and eq. [6] for **5**^{c}.

$$^{1}J_{C1,C2}\,(\text{Hz}) = 49.14 - 3.70\cos\theta_2 + 0.16\sin\theta_2 - 0.56\cos 2\theta_2 - 0.19\sin 2\theta_2$$
RMSD = 0.09 Hz eq. [4]

$$^{1}J_{C1,C2}\,(\text{Hz}) = 48.98 - 3.64\cos\theta_2 + 0.05\sin\theta_2 - 0.32\cos 2\theta_2 - 0.06\sin 2\theta_2$$
RMSD = 0.16 Hz eq. [5]

$$^{1}J_{C1,C2}\,(\text{Hz}) = 49.25 - 3.78\cos\theta_2 + 0.37\sin\theta_2 - 0.32\cos 2\theta_2 + 0.06\sin 2\theta_2$$
RMSD = 0.13 Hz eq. [6]

*F. Calculated $^1J_{C5,C6}$ Values in Model Structures **1**^{c} and **2**^{c} as a Function of ω and Equation Parameterization.* DFT calculations of $^1J_{C5,C6}$ in **1**^{c} and **2**^{c} were performed as described above for $^1J_{C1,C2}$. Torsion angle ω (C4–C5–C6–O6) was rotated in 15° increments through 360° and exocyclic C–O torsion angles θ_1, θ_2, and θ_3 were allowed to rotate freely in all optimizations. Torsion angles θ_4 (C3–C4–O4–H; 61.53° in **1** and 102.63° in **2**) and θ_6 (C5–C6–O6–H; –88.26° in **1** and –58.02° in **2**) were fixed at values observed in the crystal structures. The resulting calculated $^1J_{C5,C6}$ values were plotted as a function of ω (Figure 2) and the resulting curves were fit to give eq. [7] for **1**^{c} and eq. [8] for **2**^{c}.

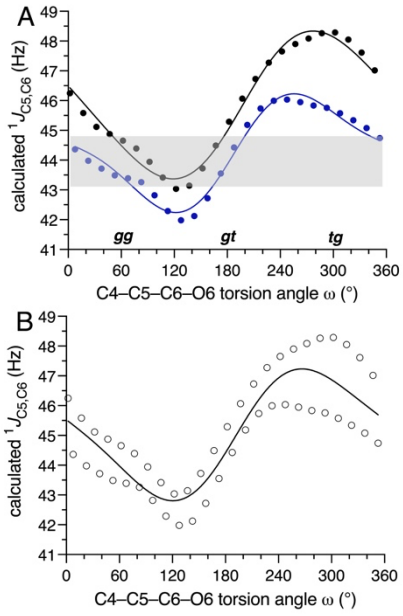


Figure 2. (A) Plots of calculated $^1J_{C5,C6}$ values as a function of ω in **1** (black) and **2** (blue). The curves were generated from eqs. [7] and [8], respectively. The shaded grey area shows the range of $^1J_{C5,C6}$ values observed experimentally (Table 1). Values of ω associated with the *gg*, *gt* and *tg* rotamers are shown. (B) Data in (A) fit to a generalized equation ([eq. [10]).

$$^1J_{C5,C6}\,(\text{Hz}) = 45.84 + 0.73\cos\omega - 2.29\sin\omega - 0.14\cos 2\omega + 0.22\sin 2\omega \quad \text{RMSD} = 0.26\ \text{Hz} \qquad \text{eq. [7]}$$

$$^1J_{C5,C6}\,(\text{Hz}) = 44.20 + 0.35\cos\omega - 1.75\sin\omega - 0.10\cos 2\omega + 0.51\sin 2\omega \quad \text{RMSD} = 0.23\ \text{Hz} \qquad \text{eq. [8]}$$

G. Generalized Equations. The curves in Figure 1 for **1**^{c}–**5**^{c} (from eqs. [2]–[6]) show modest *y*-axis displacements of ~1 Hz or less, allowing the full dataset to be averaged to give a generalized equation relating $^1J_{C1,C2}$ to θ_2 in **1**–**5** (eq. [9]).

$$^1J_{C1,C2}\,(\text{Hz})\,(\text{gen}) = 49.29 - 3.68\cos\theta_2 + 0.14\sin\theta_2 - 0.44\cos 2\theta_2 - 0.03\sin 2\theta_2 \quad \text{RMSD} = 0.49\ \text{Hz} \qquad \text{eq. [9]}$$

Likewise, the plots in Figure 2 reveal very similar curves for **1**^{C} and **2**^{C} (from eqs. [7] and [8]), with modest *y*-axis displacements of ~1 Hz or less, allowing the full dataset to be averaged to give a generalized equation relating $^{1}J_{C5,C6}$ to ω in **1** and **2** (eq. [10]).

$$^{1}J_{C5,C6}\,(\text{Hz})\,(\text{gen}) = 45.02 + 0.54\cos\omega - 2.02\sin\omega - 0.12\cos 2\omega + 0.36\sin 2\omega \qquad \text{RMSD} = 0.94\ \text{Hz} \qquad \text{eq. [10]}$$

RESULTS AND DISCUSSION

*A. C–C Coupling Pathways and $^{1}J_{CC}$ Behavior in **1**–**5**, and Experimental Strategy.* The calculated dependencies of $^{1}J_{C1,C2}$ in **1**^{C}–**5**^{C} on θ_2, and of $^{1}J_{C5,C6}$ in **1**^{C} and **2**^{C} on ω, are shown in Figures 1 and 2, respectively. As θ_2 (C1–C2–O2–H torsion angle) is rotated through 360°, $^{1}J_{C1,C2}$ ranges from ~45 Hz to ~53 Hz (~8 Hz dynamic range). $^{1}J_{C1,C2}$ values of ~47 Hz are observed at θ_2 of 60° and –60°, whereas a value of ~53 Hz is observed at θ_2 = 180°. This behavior is caused, at least partly, by the effects of O2 lone-pair orbitals on C1–C2 bond length, with the two perfectly staggered C2–O2 rotamers (θ_2 = 60° and –60°) that orient one of the lone-pair orbitals *anti* to the C1–C2 bond associated with smaller $^{1}J_{C1,C2}$ values compared to the third rotamer (θ_2 = 180°) in which the O2–H bond is *anti* to the C1–C2 bond. The former orientation is expected to cause C1–C2 bond lengthening relative to the latter orientation, which is consistent with the observed difference in the calculated $^{1}J_{C1,C2}$ values (longer bond, less *s*-character, smaller $^{1}J_{CC}$). This structural explanation is consistent with plots of $r_{C1,C2}$ as a function of θ_2 in **1**^{C}–**5**^{C} (Figure S4A, Supporting Information) *if only the three perfectly staggered conformers are considered*. However, other structural factors are at work, otherwise the curves in Figure S4A would all be U-shaped with the smallest $r_{C1,C2}$ at θ_2 = 180°. Similar lone-pair behavior is expected from rotation of the C1–O1 bond, but only one C1–O1 rotamer (ϕ) was considered in this work, being the most populated based on stereoelectronic factors (*exo*-anomeric effect).[59] α-Anomers **1** and **3** give slightly smaller calculated $^{1}J_{C1,C2}$ values than β-anomers **2** and **4**, consistent with the behavior of experimental $^{1}J_{C1,C2}$ values observed in the anomers of methyl D-glucopyranoside and methyl D-galactopyranoside in aqueous solution (Table 1).[60]

$^1J_{C5,C6}$ values in **1**^{c} and **2**^{c} depend on both ω and θ_6. The effect of ω is shown in Figure 2 in which θ_6 was fixed at the value observed in the crystal structures of **1**^{5,6} (–88.26°) and **2**^{5,6} (–58.02°). Smaller calculated $^1J_{C5,C6}$ values are observed at ω = 60° (*gg* rotamer) and 180° (*gt* rotamer) compared to those at ω = –60° (*tg* rotamer), especially in the α-anomer (Figure 2). A plot of $r_{C5,C6}$ as a function of ω in **1**^{c} and **2**^{c} (Figure S4B, Supporting Information) shows that $r_{C5,C6}$ is larger in the *tg* rotamer than in the *gg* and *gt* rotamers, leading to the expectation that $^1J_{C5,C6}$ should be smaller in the former (less *s*-character). The opposite behavior is observed, indicating that structural factors in addition to $r_{C5,C6}$ must be responsible for the behavior shown in Figure 2.

Table 1. Experimental $^1J_{CC}$ Values Measured in Aqueous Solutions, and in Crystalline Solids, of **1**^{1,2}–**5**^{1,2}, **1**^{5,6} and **2**^{5,6}.

cmpd	$^1J_{CC}$ (solution) (Hz)[a]	$^1J_{CC}$ (crystal) (Hz)[b]	[$^1J_{CC}^{cry}$ – $^1J_{CC}^{sol}$] (Hz)[c]
1^{1,2} (αGlc)	46.5 ± 0.1 (46.7)[d]	48.05 ± 0.13	1.6
2^{1,2} (βGlc)	46.8 ± 0.1 (46.9)	48.57 ± 0.09	1.8
3^{1,2} (αGal)	46.3 ± 0.1 (46.4)	48.82 ± 0.09	2.5
4^{1,2} (βGal)	46.7 ± 0.1 (46.7)	49.07 ± 0.09	2.4
5^{1,2} (αMan)	47.2 ± 0.1 (47.2)	49.42 ± 0.27	2.2
1^{5,6} (αGlc)	43.3 ± 0.1 (43.3)[e]	43.25 ± 0.12	~0
2^{5,6} (βGlc)	43.2 ± 0.1 (43.3)	44.82 ± 0.11	1.6

[a]In ^2H_2O solution, ~25 °C, ~20 m*M* solutions. [b]Average *J*-coupling determined from three measurements (see Table S1, Supporting Information); the indicated error is the standard deviation. [c]Value obtained by subtracting $^1J_{CC}^{sol}$ from $^1J_{CC}^{cry}$. [d]$^1J_{C1,C2}$ values in parentheses were reported previously in either the methyl or ethyl glycosides in aqueous solution (ref. 60). [e]$^1J_{C5,C6}$ values in parentheses were reported previously (ref. 63).

$^1J_{C5,C6}$ depends strongly on both ω and θ_6 making it possible to construct a hypersurface relating both torsion angles to the $^1J_{CC}$ (Figures S6 and S7, Supporting Information) To simplify the calculations in this work, only the dependence on ω was considered, with θ_6 fixed at values observed in the crystals. More discussion of the dependence of $^1J_{C5,C6}$ on ω and θ_6 is provided in the Supporting Information.

Monosaccharides **1**–**5**, which have different configurations at C1 and/or C2, were chosen to determine whether $^1J_{C1,C2}$ values can be calculated reliably in different pyranosyl ring structures. In **1**–**5**, rotation of the C1–C2 bond, which affects $^1J_{C1,C2}$,[21] is constrained by the pyranosyl ring, which highly prefers the $^4C_1$ chair conformation in all compounds. Consequently, only rotations about the C1–O1 and C2–O2 bonds influence $^1J_{C1,C2}$ values in solution (neglecting intermolecular factors such as hydrogen bonding). In crystalline **1**–**5**, the latter rotations are

constrained by the lattice, and crystal structures were obtained to provide these angles with reasonable accuracy (less than ±4°).[11] The same crystalline, doubly ^{13}C-labeled compounds were then used to measure $^{1}J_{CC}$ values by solid-state ^{13}C NMR.[61,62] Since the relevant C–C and C–O torsion angles are known from the crystal structures, *in silico* models of each compound can be built to replicate these structures and $^{1}J_{CC}$ values calculated for comparison to experimental values.

*B. Experimental $^{1}J_{C1,C2}$ and $^{1}J_{C5,C6}$ Values Measured in Aqueous Solution and in Crystalline Solids of **1**^{1,2}–**5**^{1,2}, **1**^{5,6} and **2**^{5,6}.* Aqueous ($^{2}H_2O$) solutions (~20 m*M*) and crystalline **1**^{1,2}–**5**^{1,2}, **1**^{5,6} and **2**^{5,6} gave solution and solid-state ^{13}C NMR spectra, respectively, that contained two intense signals arising from the two ^{13}C-labeled carbons. Each signal was split equally into a doublet due to ^{13}C-^{13}C spin-coupling ($^{1}J_{CC}$) between the labeled carbons (see representative spectra in the Supporting Information). The measured $^{1}J_{CC}$ values are shown in Table 1. The experimental $^{1}J_{C1,C2}$ values ranged from 46.3 to 49.4 Hz, overlapping that of the calculated values (Figure 1A). $^{1}J_{C5,C6}$ values ranged from 43.2 to 44.8 Hz, and were 3–4 Hz smaller than $^{1}J_{C1,C2}$ values, as expected given the smaller number of electronegative atoms appended to the C5–C6 fragment than to the C1–C2 fragment.[64] As found for $^{1}J_{C1,C2}$, the range of experimental $^{1}J_{C5,C6}$ values overlapped that of the calculated values (Figure 2A). $^{1}J_{CC}$ values measured in aqueous solution were, on average, ~2 Hz smaller than corresponding values measured in the crystalline samples (Table 1).

*C. Methods Used To Calculate $^{1}J_{CC}$ Values in **1**^{c}–**5**^{c}.* Five sets of DFT calculations were conducted to enable comparisons between experimental and calculated $^{1}J_{CC}$ values in **1**–**5**.

In Set 1, DFT calculations of J_{CC} values were performed on conformations of **1**^{c}–**5**^{c} identical to those observed in their respective X-ray crystal structures (i.e., the Cartesian coordinates of the crystal structures were used directly in DFT calculations of $^{1}J_{CC}$ values without geometry optimizations). These X-ray structures were obtained by all-atom refinement of the diffraction data, including the hydrogen atoms.

In Set 2, DFT calculations of J_{CC} values were performed on the Hirshfeld Atom Refined (HAR) X-ray structures of **1**^{c}–**5**^{c} (no geometric optimization was performed).

In Set 3, DFT calculations were performed on the X-ray structures of **1**^{c}–**5**^{c} used in Set 1 after DFT geometry optimization of only the hydrogen atoms in each structure.

In Set 4, either torsion angles θ_2 or ω in **1**^{c}–**5**^{c} were rotated in 15^{o} increments, followed by geometry optimization, to give a set of conformers in which either $^{1}J_{C1,C2}$ or $^{1}J_{C5,C6}$ values were calculated and used to parameterize equations (eqs. [2]–[8]) that relate each $^{1}J_{CC}$ to either θ_2 or ω. These equations were then used to calculate $^{1}J_{CC}$ values using torsion angles measured in the crystal structures, and the resulting values were compared to the experimental values (see Calculations for details).

Table 2. Experimental and Calculated $^{1}J_{CC}$ Values in **1**^{1,2}–**5**^{1,2}, **1**^{5,6} and **2**^{5,6} Obtained by Solid-State ^{13}C NMR and DFT.

compound	experiment	calculated				
	$^{1}J_{CC}$ (crystal) (Hz)[a]	$^{1}J_{CC}$ (crystal) (Hz)[b]	$^{1}J_{CC}$ (HAR) (Hz)[c]	$^{1}J_{CC}$ (H-opt) (Hz)[d]	$^{1}J_{CC}$ (specific) (Hz)[e]	$^{1}J_{CC}$ (general) (Hz)[f]
1^{1,2} (αGlc)	48.05 ± 0.13	44.9	49.3	49.3	49.3	48.8
2^{1,2} (βGlc)	48.57 ± 0.09	46.2	51.4	51.0	50.7	50.2
3^{1,2} (αGal)	48.82 ± 0.09	44.9	50.4	50.5	51.4	51.8
4^{1,2} (βGal)	49.07 ± 0.09	43.5	51.2	50.8	50.6	50.6
5^{1,2} (αMan)	49.42 ± 0.27	48.0	52.9	52.2	52.7	52.7
1^{5,6} (αGlc)	43.25 ± 0.12	37.6	46.0	45.7	45.8	45.2
2^{5,6} (βGlc)	44.82 ± 0.11	35.3	43.4	44.3	44.2	44.8

[a]Average J-coupling from solid-state ^{13}C NMR determined from three measurements (see Table S1, Supporting Information); the indicated error is the standard deviation. [b]Calculated $^{1}J_{CC}$ values from Set 1. [c]Average calculated $^{1}J_{CC}$ values from Set 2; HAR was repeated in triplicate. All standard deviations were ≤ 0.1 Hz. [d]Calculated $^{1}J_{CC}$ values from Set 3. [e]Calculated $^{1}J_{CC}$ values from Set 4. [f]Calculated $^{1}J_{CC}$ values from Set 5.

In Set 5, $^{1}J_{CC}$ values were calculated as described for Set 4, but generalized equations [9] (for **1**^{1,2}–**5**^{1,2}) and [10] (for **1**^{5,6}–**2**^{5,6}) were used instead of eqs. [2]–[8].

For Sets 3–5, all geometry optimizations were conducted using the B3LYP functional[53] and 6-311+g(d,p) basis set[57,58]. A limited test of functionals and basis sets gave very similar calculated $^{1}J_{CC}$ values when crystal structures in which the hydrogen atoms were optimized (Set 3) were used in the calculations. For example, calculated $^{1}J_{CC}$ values in structures that were geometry optimized using the B3LYP and wB97XD[65] functionals with the 6-311+g(d,p) basis set

differed by 0.02 Hz or less. Differences in C–H and O–H bond lengths were also small (0.001 Å and 0.004 Å, respectively).

D. Comparisons of DFT-Calculated $^1J_{CC}$ Values to Experimental Values Obtained by Solid-State ^{13}C NMR. Calculated $^1J_{CC}$ values, obtained using different methods (Sets 1–5) were compared to experimental $^1J_{CC}$ values obtained by solid-state ^{13}C NMR (Table 2). $^1J_{CC}$ values calculated directly from the crystal structures (Set 1) ranged from 35.3–48.0 Hz and were consistently *smaller* than the experimental values. This comparison, however, is not meaningful since C–H bond lengths in Set 1 structures are likely to be significantly underestimated, which in

Table 3. C–H and O–H Bond Lengths and DFT-Calculated $^1J_{CC}$ Values in **1**–**5** in X-ray Structures (Set 1), HAR Structures (Set 2) and in X-ray Structures That Were Geometry Optimized for Hydrogen Atoms Only (Set 3).

cmpd	X-ray structure					HAR structure					DFT optimization (hydrogen atoms only)				
	$r_{C1,H1}$ (Å)	$r_{C2,H2}$ (Å)	$r_{O2,H}$ (Å)	$^1J_{C1,C2}$ (Hz)		$r_{C1,H1}$ (Å)	$r_{C2,H2}$ (Å)	$r_{O2,H}$ (Å)	$^1J_{C1,C2}$ (Hz)		$r_{C1,H1}$ (Å)	$r_{C2,H2}$ (Å)	$r_{O2,H}$ (Å)	$^1J_{C1,C2}$ (Hz)	
1	0.954	0.976	0.847	44.9		1.114	1.110	0.945	49.3		1.095	1.093	0.965	49.3	
2	0.987	0.977	0.824	46.2		1.113	1.092	0.944	51.4		1.101	1.095	0.962	51.0	
3	0.950	0.971	0.837	44.9		1.098	1.101	0.947	50.4		1.096	1.092	0.961	50.5	
4	0.947	0.945	0.766	43.5		1.099	1.104	0.935	51.2		1,102	1.096	0.962	50.8	
5	0.953	0.989	0.848	48.0		1.095	1.108	0.941	52.9		1.095	1.097	0.966	52.2	
	$r_{C5,H5}$ (Å)	$r_{C6,H6R}$ (Å)	$r_{C6,H6S}$ (Å)	$r_{O6,H}$ (Å)	$^1J_{C5,C6}$ (Hz)	$r_{C5,H5}$ (Å)	$r_{C6,H6R}$ (Å)	$r_{C6,H6S}$ (Å)	$r_{O6,H}$ (Å)	$^1J_{C5,C6}$ (Hz)	$r_{C5,H5}$ (Å)	$r_{C6,H6R}$ (Å)	$r_{C6,H6S}$ (Å)	$r_{O6,H}$ (Å)	$^1J_{C5,C6}$ (Hz)
1	0.984	0.968	0.950	0.820	37.6	1.112	1.107	1.087	0.968	46.0	1.095	1.095	1.090	0.963	45.7
2	0.974	0.988	0.946	0.854	35.3	1.113	1.111	1.089	0.959	43.4	1.099	1.098	1.090	0.965	44.3

turn affects the calculated $^1J_{CC}$ values (see discussion below). The remaining four methods (Sets 2–5) gave calculated $^1J_{CC}$ values that were consistently *larger* than experimental values. For Set 2, the differences ranged from 1.3–3.5 Hz (2.6–6.8%) with an average difference of ~2.2 Hz (4.6%). For Set 3, the differences ranged from 0.5–2.8 Hz (1.2–5.5%) with an average difference of ~1.8 Hz (3.8%). For Set 4, differences ranged from 0.6–3.3 Hz (1.4–6.4%) with an average difference of ~2.0 Hz (4.1%). For Set 5, differences ranged from 0–3.3 Hz (0–6.4%) with an average difference of ~1.7 Hz (3.6%). Taken collectively, the average percent difference determined by the four methods of calculating the $^1J_{CC}$ values ranged from 3.6–4.6%, and in all cases the calculated values were over-estimated.

The finding that DFT-calculated $^1J_{CC}$ values are over-estimated by 4–5% is consistent with that obtained from solution NMR studies of conformationally constrained compounds.[42] Crystal structures obtained by HAR (Set 2) showed no significant differences with those in Set 1 (all-atom refinement) for heavy atoms but significant differences were observed for the hydrogen atoms, especially in regard to C–H bond lengths. In Set 3, the C–H and O–H bond lengths (which typically cannot be measured reliably by X-ray crystallography) were optimized before $^1J_{CC}$ values were calculated. During geometry optimization, bond lengths involving the heavy atoms were fixed at values found in the crystal structures, which can be measured reliably. C–H and O–H bond lengths in the structures from Sets 1–3 are summarized in Table 3. HAR (Set 2) and geometry optimization of all hydrogen atoms (Set 3) resulted in lengthening of both the C–H and O–H bonds by ~0.1 Å. These changes led to larger calculated $^1J_{C1,C2}$ and $^1J_{C5,C6}$ values by 3–6 Hz (relative to those calculated using Set 1 structures) even though C–C bond lengths were fixed at the same lengths found in the Set 1 crystal structures during calculations to give the Set 2 and 3 structures (Table 2).

The effect of C–H bond lengths on calculated $^1J_{CC}$ values was investigated further computationally to determine the degree of sensitivity. Starting from the crystal structure of **1**[1,2] (Set 1), the C1–H1 and C2–H2 bonds were both lengthened incrementally by 0.02 Å from 1.00 to 1.12 Å with all remaining structural parameters held fixed, and $^1J_{C1,C2}$ values were calculated in the resulting structures (no geometry optimization was performed in these calculations). Calculated $^1J_{C1,C2}$ values increased linearly as both $r_{C1,H1}$ and $r_{C2,H2}$ were increased (Figure 3). From the slope of the plot, a 0.01 Å increase in the C–H bond lengths resulted in a ~0.4 Hz increase in the calculated $^1J_{C1,C2}$. Based on results

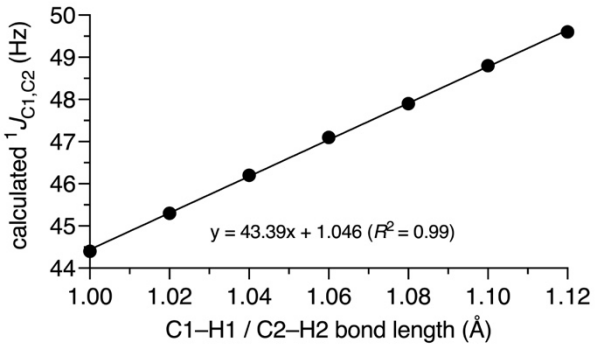


Figure 3. Plot of calculated $^1J_{C1,C2}$ values in **1**[c] as a function of the C1–H1 and C2–H2 bond lengths. Both C–H bond lengths were set initially at 1.0 Å and increased by 0.02 Å. All remaining structural parameters were fixed at values found in the X-ray structure of **1**[1,2] (Set 1).

from a limited study of functionals and basis sets for geometry optimization (see above), differences of ~0.001 Å in calculated C–H bond lengths are expected when different combinations of functionals and basis sets are used, but these small changes give very similar calculated $^1J_{CC}$ values, with the differences in the latter comparable to the errors in the experimental solid-state $^1J_{CC}$ measurements. In an additional test of the effect of the functional and basis set, $^1J_{C1,C2}$ was calculated in **1**^{c} using the coordinates of the heavy atoms obtained from the crystal structure of **1**[1,2] (Set 1) but adjusting the C1–H1 and C2–H2 bond lengths to 1.120 Å. Two structures, one geometry-optimized at the B3LYP/6-311+g(d,p) level of theory and the other at the level of wB97XD/6-311+g(d,p), gave calculated $^1J_{C1,C2}$ values that differed by less than 0.1 Hz. Whether the behavior shown in Figure 3 applies to all types of $^1J_{CC}$ values in saccharides, however, remains to be determined. However, it is clear that the ability to calculate $^1J_{CC}$ values accurately will depend strongly on the *in silico* structures containing accurate C–H bond lengths

CONCLUSIONS

As discussed in prior work,[42] the inclusion of $^1J_{CH}$ and $^1J_{CC}$ values in *MA'AT* analysis[35–41] could significantly improve experimental modelling of multi-state conformational equilibria and dynamics in solution. However, the use of these spin-couplings requires confidence that quantitative values can be computed, and that there is sufficient knowledge of secondary conformational effects, which arise mainly from C–O bond rotations involving the coupled carbon(s), to treat them properly during equation parameterization by DFT. Recent NMR solution studies showed that, under experimental conditions that allow both primary and secondary factors to be reliably recapitulated in DFT calculations, calculated $^1J_{CH}$ values were found to be in very good agreement with experimental values but calculated $^1J_{CC}$ values were not, the latter being over-estimated by ~4–5%.[42] To validate the latter finding using a different experimental approach, the present study was undertaken to compare DFT-calculated and experimental $^1J_{CC}$ values under conditions where the latter could be measured without the complications arising from the time-averaging of secondary conformational factors that occurs in solution. This time-averaging

was eliminated by making experimental measurements on crystalline samples in which C–O bond conformations were fixed by the crystalline lattice and, importantly, could be determined by X-ray crystallography. As a consequence, secondary C–O bond conformational effects could be treated reliably in DFT calculations in order to allow meaningful comparisons of calculated and experimental values.

This work has revealed deficiencies in the ability of DFT, as currently implemented in this laboratory, to calculate $^{1}J_{CC}$ values accurately. The magnitude of the discrepancy between calculated and experimental $^{1}J_{CC}$ values depends on how model structures used in the DFT calculations were obtained, as illustrated by the varying results obtained when five different sets of structures were used (Table 2). However, in general, calculated $^{1}J_{CC}$ values were found to be overestimated by ~4–5%, in good agreement with conclusions drawn from recent solution NMR measurements.[42] The source of the error is unclear, and given the small dataset in this and prior work, it is unlikely that all calculated $^{1}J_{CC}$ values deviate from experiment to the same extent. This being the case, the application of a uniform 4–5% correction factor to all calculated $^{1}J_{CC}$ values regardless of coupling pathway structure is not likely to be a reliable solution.

The percent error in calculated $^{1}J_{CC}$ values is similar to those estimated previously for $^{2}J_{CC}$ and $^{3}J_{CC}$ values, which were < 4%.[54] However, the larger magnitudes of $^{1}J_{CC}$ values produce absolute errors of 4.5–6 Hz. These large absolute errors are likely to reduce the reliability of *MA'AT* models if calculated $^{1}J_{CC}$ were to be included as constraints in the analysis.

A large body of prior work in non-carbohydrate systems has addressed the development and application of various computational methods such as CASTEP[66–69] to calculate NMR parameters in the solid-state, including *J*-couplings[70,71] and chemical shift,[72] and the importance of using structures in these calculations that contain accurate bond lengths involving hydrogen. DFT studies typically examine different functionals and basis sets to determine which combination gives the best fit to the experimental data. It should be appreciated that the functionals and basis sets used for geometry optimization, and those used for the *J*-coupling calculations, determine collectively the accuracy of calculated $^{1}J_{CC}$. In the present work, a limited assessment of

functionals and basis sets was made for geometry optimization, but only one combination was used in the *J*-coupling calculations. Within these constraints, calculated $^1J_{CC}$ values did not differ to any appreciable extent, and it is unlikely that a more thorough survey of functionals and basis sets for geometry optimization will resolve the observed discrepancy between calculated and experimental $^1J_{CC}$. This being the case, possible sources of the error may be in the functional and basis set used for the *J*-coupling calculations, an inherent flaw in the DFT method itself, possible unknown structural factors not fully captured in the DFT calculations, or some combination of these sources. These limitations notwithstanding, it is still remarkable that DFT as applied in this work is able to calculate $^1J_{CC}$ values as well as it does.

An important, although in retrospect perhaps not surprising, result of this work is the observation that calculated $^1J_{CC}$ values in H–C–C–H fragments are very sensitive to the lengths of the C–H bonds, with longer bonds correlating with larger $^1J_{CC}$ values. Prior work had shown that rotation of the C–C bond and rotation of the two C–O bonds in vicinal diol fragments are strong determinants of the $^1J_{CC}$, with C–O bond rotation exerting the greater effect.[21] While it is often assumed that C–C bond length, presumably affected by both C–C and C–O bond rotation, is the root cause of changes in $^1J_{CC}$, the role of C–H bond length is less obvious. Indeed, perhaps the dependency of $^1J_{CC}$ on C–C bond rotation may be traced in part to changes in C–H bond length, which likely vary as steric interactions vary with C–C bond rotation (greater in eclipsed rotamers than in staggered rotamers, for example). This behavior bears directly on the choice of functional and basis set used to obtain geometry optimized structures for use in $^1J_{CC}$ calculations. Functionals and/or basis sets that are known to give accurate C–H and C–C bond lengths should be used when calculating $^1J_{CC}$ values. From an experimental standpoint, C–H bond lengths measured from conventional crystal structures are inappropriate for $^1J_{CC}$ calculations since these lengths are often held constant at 0.998–1.000 Å during refinement to simplify analyses of electron densities, especially for structures obtained at lower resolutions. High-resolution diffraction data (0.7 Å or less) enable all-atom refinement that includes hydrogen, as applied in this work. However, even when all-atom refinement is performed, the resulting C–H bond lengths

are still inaccurate (typically under-estimated), necessitating HAR to give final structures suitable for $^{1}J_{CC}$ calculations.

While the absolute magnitudes of $^{1}J_{CC}$ values calculated by DFT in this work deviated by ~5% from those measured experimentally, DFT is likely to provide accurate overall dependencies of a given $^{1}J_{CC}$ value on a specific molecular torsion angle (i.e., the shape and amplitude of the curve are well determined). To the extent that this is true, a potential solution to the problem of accuracy is to use DFT-parameterized $^{1}J_{CC}$ equations in *MA'AT* analysis but to allow the constants in these equations to optimize during the fitting of the remaining, and presumably more reliable, redundant *J*-couplings (i.e., allow the $^{1}J_{CC}$ curve to shift along the *y*-axis until a best fit of the experimental *J*-values is obtained). This solution is different than applying a uniform correction factor to all $^{1}J_{CC}$ equations, which is not recommended, since the former allows each equation to be optimized independently, that is, the correction factors would be equation-specific. Alternatively, arbitrary values of the constants could be tested empirically to determine which, if any, lead to better fits of the data. Neither of these solutions, however, is satisfying since they don't address the core cause of the discrepancy. Future work is needed to further document the discrepancy and identify its origin.

CONFLICTS OF INTEREST

There are no conflicts of interest to declare.

Electronic Supplementary information (ESI) available: Synthesis of ^{13}C-labeled compounds **1**–**5**; Crystallization and X-ray crystallography of **1**–**5;** Solution ^{1}H and ^{13}C{^{1}H} NMR spectra of $\mathbf{1}^{1,2}$–$\mathbf{5}^{1,2}$, $\mathbf{1}^{5,6}$ and $\mathbf{2}^{5,6}$; Partial ^{13}C{^{1}H} NMR spectrum of $\mathbf{1}^{1,2}$ showing the natural abundance carbons; Partial ^{13}C{^{1}H} NMR spectrum of $\mathbf{1}^{5,6}$ showing the natural abundance carbons; Solid-state ^{13}C NMR spectra of $\mathbf{1}^{1,2}$–$\mathbf{5}^{1,2}$, $\mathbf{1}^{5,6}$ and $\mathbf{2}^{5,6}$; Plots of $r_{C1,C2}$ in $\mathbf{1}^{1,2}$–$\mathbf{5}^{1,2}$ vs θ_2, and of $r_{C5,C6}$ in $\mathbf{1}^{5,6}$ and $\mathbf{2}^{5,6}$ vs ω; Fitting statistics from solid-state ^{13}C NMR measurements of $^{1}J_{CC}$ in $\mathbf{1}^{1,2}$–$\mathbf{5}^{1,2}$, $\mathbf{1}^{5,6}$ and $\mathbf{2}^{5,6}$; Plots of S/S_0 vs τ to determine $^{1}J_{C1,C2}$ or $^{1}J_{C5,C6}$ in $\mathbf{1}^{1,2}$–$\mathbf{5}^{1,2}$, $\mathbf{1}^{5,6}$ and $\mathbf{2}^{5,6}$;

Discussion of equation [1] for *J*-coupling equation parameterization; Additional DFT calculations; Torsional dependencies of $^{1}J_{C5,C6}$ values; Representative Cartesian coordinates for DFT structures **1**^{c}–**5**^{c}; Complete reference 50; References. See DOI: XXXX.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (CHE 2002625 to A. S.) and by Omicron Biochemicals, Inc., South Bend, IN. The Notre Dame Radiation Laboratory is supported by the Department of Energy Office of Science, Office of Basic Energy Sciences, under Award Number DE-FC02-04ER15533. This is document number NDRL 5379.

REFERENCES

1. R. H. Contreras and J. E. Peralta, Angular dependence of spin-coupling constants, *Prog. NMR Spectrosc.,* 2000, **37**, 321–425.

2. C. A. G. Haasnoot, F. A. A. M. De Leeuw, H. P. M. De Leeuw and C. Altona, The relationship between proton-proton NMR coupling constants and substituent electronegativites. II—Conformational analysis of the sugar ring in nucleosides and nucleotides in solution using a generalized Karplus equation, *Org. Magn. Reson.*, 1981, **15**, 43–52.

3. F. Li, J. H. Lee, A. Grishaev, J. Ying and A. Bax, High accuracy of Karplus equations for relating three-bond *J*-couplings to protein backbone torsion angles, *ChemPhysChem.*, 2015, **16**, 572–578.

4. M. Karplus, Contact electron-spin coupling of nuclear magnetic moments, *J. Chem. Phys.*, 1959, **30**, 11–15.

5. M. Karplus, Vicinal proton coupling in nuclear magnetic resonance, *J. Am. Chem. Soc.,* 1963, **85**, 2870–2871.

6. R. U. Lemieux, R. K. Kullnig, H. J. Bernstein and W. G. Schneider, Configurational effects on the proton magnetic resonance spectra of six-membered ring compounds, *J. Am. Chem. Soc.*, 1958, **80**, 6098–6105.

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

8. M. J. Minch, Orientational dependence of vicinal proton-proton NMR coupling constants: The Karplus relationship, *Concepts Magn. Reson.*, 1994, **6**, 41–56.

9. B. Mulloy, T. Frenkiel and D. B. Davies, Long-range carbon-proton coupling constants: Application to conformational studies of oligosaccharides, *Carbohydr. Res.*, 1988, **184**, 39–46.

10. I. Tvaroska and F. R. Taravel, Carbon-proton coupling constants in the conformational analysis of sugar molecules, *Adv. Carbohydr. Chem. Biochem.*, 1995, **51**, 15–61.

11. J. Lin, R. J. Meredith, A. G. Oliver, I. Carmichael and A. S. Serianni, Two-bond ^{13}C-^{13}C spin-coupling constants in saccharides: Dependencies on exocyclic hydroxyl group conformation, *Phys. Chem. Chem. Phys.*, 2021, **23**, 22912–22922.

12. G. E. Maciel, J. W. McIver, N. S. Ostlund and J. A. Pople, Approximate self-consistent molecular orbital theory of nuclear spin couplings. III. Geminal proton-proton coupling constants, *J. Am. Chem. Soc.,* 1970, **92**, 4151–4157.

13. I. Ando and G. A. Webb, *Theory of NMR Parameters*, Academic Press, New York, 1983, pp. 106–107.

14. N. Müller and D. E. Pritchard, ^{13}C splittings in proton magnetic resonance spectra. I. Hydrocarbons, *J. Chem. Phys.*, 1959, **31**, 768–771.

15. P. E. Hansen, Carbon-hydrogen spin-spin coupling constants, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1981, **14**, 175–295.

16. N. Marco, A. A. Souza, P. Nolis, C. Cobas, R. R. Gil and T. Parella, $^{1}J_{CH}$ NMR profile: Identification of key structural features and functionalities by visual observation and direct measurement of one-bond proton-carbon coupling constants, *J. Org. Chem.*, 2017, **82**, 2040-2044.

17. T. Klepach, W. Zhang, I. Carmichael and A. S. Serianni, ^{13}C-^{1}H and ^{13}C-^{13}C NMR *J*-couplings in ^{13}C -labeled *N*-acetyl-neuraminic acid: Correlations with molecular structure, *J. Org. Chem.,* 2008, **73**, 4376–4387.

18. T. J. Church, I. Carmichael and A. S. Serianni, ^{13}C-^{1}H and ^{13}C-^{13}C spin-coupling constants in methyl β-D-ribofuranoside and methyl 2-deoxy-β-D-*erythro*- pentofuranoside: Correlations with molecular structure and conformation, *J. Am. Chem. Soc.*, 1997, **119**, 8946–8964.

19. M. J. Hadad, W. Zhang, T. Turney, L. Sernau, W. Wang, R. J. Woods, A. Incandela, I. Surjancev, A. Wang, M-K. Yoon, A. Coscia, C. Euell, R. Meredith, I. Carmichael and A. S. Serianni, NMR spin-couplings in saccharides: Relationships between structure, conformation and the magnitudes of J_{HH}, J_{CH} and J_{CC} values, *in New Developments in NMR 10: NMR in Glycoscience and Glycotechnology*, ed. T. Peters and K. Kato, Royal Society of Chemistry, 2017, pp. 20–100.

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

21. I. Carmichael, D. M. Chipman, C. A. Podlasek and A. S. Serianni, Torsional effects on the one-bond ^{13}C-^{13}C spin coupling constant in ethylene glycol: Insights into the behavior of $^{1}J_{CC}$ in carbohydrates, *J. Am. Chem. Soc.*, 1993, **115**, 10863–10870.

22. C. Venkata, M. J. Forster, P. W. A. Howe and C. Steinbeck, The potential utility of predicting one bond carbon-proton coupling constants in the structure elucidation of small organic molecules by NMR spectroscopy, *PLoS ONE,* 2014, **9**: e111576. doi:10.1371/journal.pone.0111576

23. J. Powell, D. Valenti, H. Bobnar, E. Drain, B. Elliott, S. Frank, T. McCullough, S. Moore, A. Kettring, R. Luliucci and J. K. Harper, Evaluating the accuracy of theoretical one-bond ^{13}C-^{13}C scalar couplings and their ability to predict structure in a natural product, *Magn. Res. Chem.*, 2017, **55**, 979–989.

24. J. San Fabián, J. M. Garcia de la Vega, R. Suardíaz, M. Fernández-Oliva, C. Pérez, R. Crespo-Otero and R. H. Contreras, Computational NMR coupling constants: Shifting and scaling factors for evaluating $^{1}J_{CH}$, *Magn. Res. Chem.*, 2013, **51**, 775–787.

25. T. Klepach, H. Zhao, X. Hu, W. Zhang, R. Stenutz, M. J. Hadad, I. Carmichael and A. S. Serianni, Informing saccharide structural NMR studies with density functional theory calculations, in *Glycoinformatics: Methods in Molecular Biology*, ed. T. Lütteke and M. Frank, Springer, New York, 2015, pp. 289–331.

26. R. J. Woods, Predicting the structures of glycans, glycoproteins, and their complexes, *Chem. Rev.*, 2018, **118**, 8005–8024.

27. Y. Yu and M. Delbianco, Conformational studies of oligosaccharides, *Chem. Eur. J.*, 2020, **26**, 9814–9825.

28. Y. Zhang, T. Yamaguchi and K. Kato, New NMR tools for characterizing the dynamic conformations and interactions of oligosaccharides, *Chem. Lett.*, 2013, **42**, 1455–1462.

29. T. Watanabe, H. Yagi, S. Yanaka, T. Yamaguchi and K. Kato, Comprehensive characterization of oligosaccharide conformational ensembles with conformer classification by free-energy landscape via reproductive kernel Hilbert space, *Phys. Chem. Chem. Phys.,* 2021, **23**, 9753–9760.

30. Q. R. Johnson, R. J. Lindsay, L. Petridis and T. Shen, Investigation of carbohydrate recognition via computer simulation, *Molecules,* 2015, **20**, 7700–7718.

31. A. Plazinska and W. Plazinski, Comparison of carbohydrate force fields in molecular dynamics simulations of protein-carbohydrate complexes, *J. Chem. Theory Comput.*, 2021, **17**, 2575–2585.

32. S-J. Park, J. Lee, Y. Qi, N. R. Kern, H. S. Lee, S. Jo, I. Joung, K. Joo, J. Lee and W. Im, CHARMM-GUI *Glycan Modeler* for modeling and simulation of carbohydrates and glycoconjugates, *Glycobiology,* 2019, **29**, 320–331.

33. M. Yang and A. D. MacKerell Jr., Conformational sampling of oligosaccharides using Hamiltonian replica exchange with two-dimensional dihedral biasing potentials and the weighted histogram analysis method (WHAM), *J. Chem. Theory Comput.*, 2015, **11**, 788–799.

34. M. W. van der Kamp, K. E. Shaw, C. J. Woods and A. J. Mulholland, Biomolecular simulation and modelling: Status, progress and prospects, *J. R. Soc. Interface,* 2008, **5**, S173–S190.

35. W. Zhang, T. Turney, R. Meredith, Q. Pan, L. Sernau, X. Wang, X. Hu, R. J. Woods, I. Carmichael and A. S. Serianni, Conformation populations of β-(1 →4) *O*-glycosidic linkages

using redundant NMR *J*-couplings and circular statistics, *J. Phys. Chem. B,* 2017, **121**, 3042–3058.

36. W. Zhang, R. Meredith, M.-K. Yoon, X. Wang, R. J. Woods, I. Carmichael and A. S. Serianni, Synthesis and *O*-glycosidic linkage conformational analysis of ^{13}C-labeled oligosaccharide fragments of an antifreeze glycolipid, *J. Org. Chem.,* 2019, **84**, 1706–1724.

37. W. Zhang, R. Meredith, Q. Pan, X. Wang, R. J. Woods, I. Carmichael and A. S. Serianni, Use of circular statistics to model αMan-(1→2)-αMan and αMan-(1→3)-α/βMan *O*-glycosidic linkage conformation in ^{13}C-labeled disaccharides and high mannose oligosaccharides, *Biochemistry*, 2019, **58**, 546–560.

38. T. Turney, Q. Pan, L. Sernau, I. Carmichael, W. Zhang, X. Wang, R. J. Woods and A. S. Serianni, *O*-Acetyl side-chains in monosaccharides: Redundant NMR spin-couplings and statistical models for acetate ester conformational analysis, *J. Phys. Chem. B,* 2017, **121**, 66–77.

39. R. J. Meredith, M. McGurn, C. Euell, P. Rutkowski, E. Cook, I. Carmichael and A. S. Serianni, *MA'AT* analysis of aldofuranosyl rings: Unbiased modeling of conformational equilibria and dynamics in solution, *Biochemistry*, 2022, **61**, 239–251.

40. R. J. Meredith, L. Sernau and A. S. Serianni, *MA'AT*: A web-based application to determine rotamer population distributions in solution from nuclear magnetic resonance spin-coupling constants, *J. Chem. Inf. Model.*, 2022, **62**, 3135–3141.

41. R. J. Meredith, T. Tetrault, M.-K. Yoon, W. Zhang, I. Carmichael and A. S. Serianni, *N*-Acetyl side-chain conformation in saccharides: Solution models obtained from *MA'AT* analysis, *J. Org. Chem.,* 2022, **87**, 8368–8379.

42. Tetrault, R. J. Meredith, W. Zhang, I. Carmichael and A. S. Serianni, One-Bond ^{13}C-^{1}H and ^{13}C-^{13}C Spin-Coupling Constants as Constraints in *MA'AT* Analysis of Saccharide Conformation, *J. Phys. Chem. B*, 2022, ***126***, 9506–9515.

43. W. Zhang, M.-K. Yoon, R. J. Meredith, J. Zajicek, A. G. Oliver, M. Hadad, M. H. Frey, I. Carmichael and A. S. Serianni, ^{13}C-^{13}C spin-coupling constants in crystalline ^{13}C -labeled saccharides: Conformational effects interrogated by solid-state ^{13}C NMR spectroscopy, *Phys. Chem. Chem. Phys.,* 2019, **21**, 23576–23588.

44. D. Jayatilaka and B. Dittrich, X-ray structure refinement using aspherical atomic density functions obtained from quantum-mechanical calculations, *Acta Cryst.*, 2008, A**64**, 383–393.

45. S. C. Capelli, H.-B. Bürgi, B. Dittrich, S. Grabowsky and D. Jayatilaka, Hirshfeld atom refinement, *IUCrJ*, 2014, **1**, 361–379.

46. F. Kleemiss, O. V. Dolomanov, M. Bodensteiner, N. Peyerimhoff, L. Midgley, L. J. Bourhis, A. Genoni, L. A. Malaspina, D. Jayatilaka, J. L. Spencer, F. White, B. Grundkötter-Stock, S. Steinhauer, D. Lentz, H. Puschmann and S. Grabowksy, Accurate crystal structures and chemical properties from NoSpherA2, *Chem. Sci.*, 2021, **12**, 1675–1692.

47. F. Neese, The ORCA program system, *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 2012, **2**, 73–78.

48. O. V. Dolomanov, L. J. Bourhis, R. J. Gildea, J. A. K. Howard and H. Puschmann, OLEX2: a complete structure solution, refinement and analysis program, *J. Appl. Cryst.*, 2009, **42**, 339–341.

49. R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, *J. Chem. Phys.*, 1992, **96**, 6796–6806.

50. M. J. Frisch, *et al.*, *Gaussian 16, Revision B.01*, Gaussian Inc., Wallingford, CT, 2016, https://gaussian.com/gaussian16/.

51. V. Sychrovský, J. Gräfenstein and D. Cremer, Nuclear magnetic resonance spin-spin coupling constants from coupled perturbed density functional theory, *J. Chem. Phys.*, 2000, **113**, 3530–3547.

52. T. Helgaker, M. Watson and N.C. Handy, 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.,* 2000, **113**, 9402–9409.

53. V. Barone, J. E. Peralta, R. H. Contreras and J. P. Snyder, DFT calculation of NMR J_{FF} spin-spin coupling constants in fluorinated pyridines, *J. Phys. Chem. A,* 2002, **106**, 5607–5612.

54. E. Cancès, B. Mennucci and J. Tomasi, A new integral equation formalism for the polarizable continuum model: Theoretical background and applications to isotropic and anisotropic dielectrics, *J. Chem. Phys.,* 1997, **107**, 3032–3041.

55. R. Cammi, B. Mennucci and J. Tomasi, Fast evolution of geometries and properties of excited molecules in solution: A Tamm-Dancoff model with application to 4-dimethyl-aminobenzonitrile, *J. Phys. Chem. A,* 2000, **104**, 5631–5637.

56. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *J. Chem. Phys.,* 1993, **98**, 5648–5652.

57. A. D. McLean and G. S. Chandler, Contracted Gaussian basis sets for molecular calculations. I. Second row atoms, Z = 11-18, *J. Chem. Phys.,* 1980, **72**, 5639–5648.

58. R. B. J. S. Krishnan, J. S. Binkley, R. Seeger and J. A. Pople, Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions, *J. Chem. Phys.,* 1980, **72**, 650–654.

59. E. Juaristi and G. Cuevas, *The Anomeric Effect*, CRC Press, Boca Raton, 1995, pp. 95–111.

60. B. Bose-Basu, T. Klepach, G. Bondo, P. B. Bondo, W. Zhang, I. Carmichael and A. S. Serianni, ^{13}C-^{13}C NMR spin-spin coupling constants in saccharides: Structural correlations involving all carbons in aldohexopyranosyl rings, *J. Org. Chem.*, 2007, **72**, 7511–7522.

61. P. Thureau, G. Mollica, F. Ziarelli and S. Viel, Selective measurements of long-range homonuclear *J*-couplings in solid-state NMR, *J. Mag. Reson.*, 2013, **231**, 90–94.

62. P. Thureau, I. Carvin, F. Ziarelli, S. Viel and G. Mollica, A Karplus equation for the conformational analysis of organic molecular crystals, *Angew. Chem.*, 2019, **131**, 16193–16197.

63. C. Thibaudeau, C., R. Stenutz, B. Hertz, T. Klepach, S. Zhao, Q. Wu, I. Carmichael and A. S. Serianni, Correlated C–C and C–O bond conformations in saccharide hydroxymethyl groups: Parameterization and application of redundant ^{1}H-^{1}H, ^{13}C-^{1}H and ^{13}C-^{13}C NMR *J*-couplings. *J. Am. Chem. Soc.*, 2004, **126**, 15668–15685.

64. T. Bandyopadhyay, J. Wu, W. A. Stripe, I. Carmichael and A. S. Serianni, ^{13}C-^{1}H and ^{13}C-^{13}C Spin Couplings in [2'-^{13}C]2'-Deoxyribonucleosides: Correlations with Molecular Structure, *J. Am. Chem. Soc.*, 1997, **119**, 1737–1744.

65. J.-D. Chai and M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

66. S. A. Joyce, J. R. Yates, C. J. Pickard and F. Mauri, A first principles theory of nuclear magnetic resonance J-coupling in solid-state systems, *J. Chem. Phys.*, 2007, **127**, 204107.

67. T. F. G. Green and J. R. Yates, Relativistic NMR J-coupling with ultrasoft pseudopotentials and the zeroth-order regular approximation, *J. Chem. Phys.* 2014, **140**, 234106.

68. S. A. Joyce, J. R. Yates, C. J. Pickard and S. P. Brown, Density functional theory calculations of hydrogen-bond-mediated NMR J coupling in the solid state, *J. Am. Chem. Soc.*, 2008,**130**, 12663–12670.

69. C. E. Hughes, G. N. M. Reddy, S. Masiero, S. P. Brown, P. A. Williams and K. D. M. Harris, Determination of a complex crystal structure in the absence of single crystals: Analysis of powder X-ray diffraction data, guided by solid-state NMR and periodic DFT calculations, reveals a new 2'-deoxyguanosine structural motif, *Chem. Sci.*, 2017, **8**, 3971–3979.

70. A. L. Webber, J. R. Yates, M. Zilka, S. Sturniolo, A.-C. Uldry, E. K. Corlett, C. J . Pickard, M. Pérez-Torralba, M. A. Garcia, D. Santa Maria, R. M. Claramunt and S. P. Brown, Weak intermolecular CH···N hydrogen bonding: Determination of ^{13}CH-^{15}N hydrogen-bond mediated *J* Couplings by solid-state NMR spectroscopy and first-principles calculations, *J. Phys. Chem. A*, 2020, **124**, 560–572.

71. S. A. Joyce, J. R. Yates, C. J. Pickard and S. P. Brown, Density functional theory calculations of hydrogen-bond-mediated NMR *J* Coupling in the solid state, *J. Am. Chem. Soc.*, 2008, **130**, 12663–12670.

72. R. K. Harris, P. Hodgkinson, C. J. Pickard, J. R. Yates and V. Zorin, Chemical shift computations on a crystallographic basis: Some reflections and comments, *Magn. Reson. Chem.*, 2007, **45**, S174–S186.