PCCP

RSC Publishing

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2013, **15**, 18223

Detecting low-level flexibility using residual dipolar couplings: a study of the conformation of cellobiose†‡

Nicholle G. A. Bell,[a] Graeme Rigg,[a] Sarah Masters,[b] Juraj Bella[a] and Dušan Uhrín*[a]

Received 16th July 2013,
Accepted 3rd September 2013

DOI: 10.1039/c3cp52987h

www.rsc.org/pccp

We have developed novel NMR methods for the measurement of heteronuclear residual dipolar couplings (RDCs) in molecules with severely overlapping NMR resonances. These and other methods enabled us to obtain 31 RDCs for α-D-cellobiose and 24 RDCs for β-D-cellobiose. The interpretation of the data in the approximation of a rigid disaccharide structure, using RDCs and interglycosidic ^3J coupling constants, yielded conformation that is very close to that determined using X-ray crystallography. However, depending on which ring was used to calculate the order parameters, the dihedral angle ψ_H varied up to 30° or 40°, while the ϕ_H angle was always the same. This indicates residual flexibility of the glycosidic linkage between the two monosaccharide rings and was observed for both α- and β-D-cellobiose. The RDC analysis using rigid fragments rather than a complete molecule has thus shown that the glycosidic bond of cellobiose is not completely rigid and exhibits low-level flexibility. The sources of this flexibility are discussed and evidence presented to support a hypothesis that it is associated with the ψ more than the ϕ angle.

Introduction

Residual dipolar couplings (RDCs) have been proven to be very useful in the structural characterisation of small molecules,[1,2] including oligosaccharides.[3–5] It is however evident that the flexibility of molecules complicates the interpretation of RDCs and this issue is the focal point of continuous research efforts.[6–11] Central to the analysis of RDCs is the requirement to characterise the molecular alignment by calculating an alignment tensor. This can be done either in a self-consistent manner, using five or more RDCs for each rigid fragment (*e.g.* a monosaccharide ring), *via* the singular value decomposition (SVD) method,[12] or by estimating the alignment tensor without utilising the experimental RDCs, for example by performing a Monte Carlo search of structures in the vicinity of infinite two-dimensional plates,[13] or using the radius of the gyration tensor[14] or the moment of inertia tensor.[9]

The SVD approach allows rigorous characterisation of the alignment of rigid fragments and is more likely to reveal flexibility, rather than hiding any discrepancies between the experiment and the theory in the "structural noise". Using cellobiose, 4-*O*-(β-D-glucopyranosyl)-β-D-glucopyranose (see the inset to Fig. 1), a disaccharide that has rather "rigid" glycosidic linkage, we present here novel approaches that allow the detection of low-level molecular flexibility.

Cellobiose is the disaccharide repeating unit of cellulose, a key component of plant cell walls. Its conformation has been the subject of continuous research by both theoreticians and experimentalists producing a considerable amount of data. Theoretical studies have identified conformational energy minima of cellobiose,[15–24] which are characterised by certain combinations of glycosidic dihedral angles ϕ and ψ. Using the ϕ_H (H1′–C1′–O4–C4) and ψ_H (C1′–O4–C4–H4) definitions of dihedral angles, the most stable conformer of cellobiose is referred to as *syn*-ϕ_H/*syn*-ψ_H (ϕ_H–ψ_H ~ 60°), while the identified metastable states are *anti*-ϕ_H/*syn*-ψ_H (ϕ_H ~ 180°/ψ_H ~ 60°) or *syn*-ϕ_H/*anti*-ψ_H (ϕ_H ~ 60°/ψ_H ~ 180°). Although all theoretical models show the existence of these minima, their populations vary dramatically depending on the method used for their calculation. For example, molecular dynamics simulations using CHARMM[25] with the Carbohydrate Solution Force Field (CSFF)[19] identified the *anti*-ϕ_H conformer as the most stable, while replica exchange dynamics using the AMBER/GLYCAM force field[18] placed 99% of the molecules at 300 K into the *syn*-ϕ_H/*syn*-ψ_H region and ~1% and ~0% into the *anti*-ψ_H and *anti*-ϕ_H regions, respectively. Similar results were obtained with the GROMOS 45A4 force-field[26] with the artificially enhanced sampling around the glycosidic bond using the local elevation umbrella method.[20]

[a] *EastChem School of Chemistry, University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, UK. E-mail: dusan.uhrin@ed.ac.uk*
[b] *Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand*

† Celebrating 300 years of Chemistry at Edinburgh.

‡ Electronic supplementary information (ESI) available: Sample preparation details, definition of the order parameter and calculations of rmsds, the ^{13}C spectrum of cellobiose, examples of the determination of HH and CC splittings, 2D RDC or *J*-coupling constants-based rmsd maps, and a survey of dihedral angles of cellobiose related carbohydrates. See DOI: 10.1039/c3cp52987h

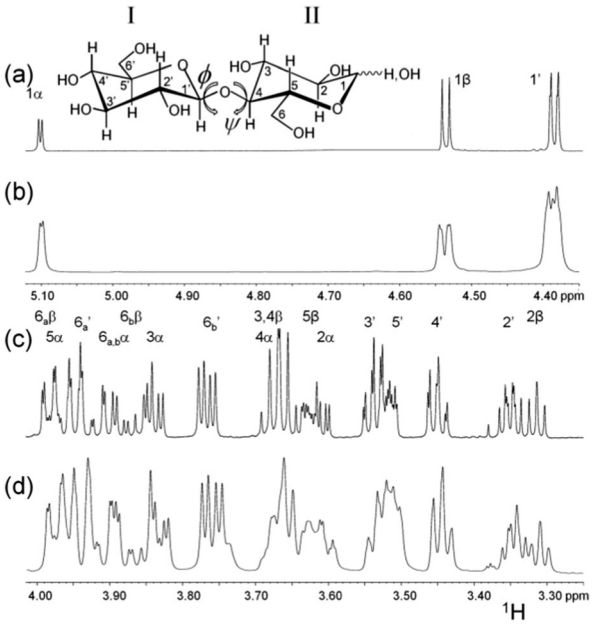

Fig. 1 800 MHz ^{1}H spectra of cellobiose. Expansions of the anomeric and aglyconic signals of the isotropic, (a) + (c), and aligned, (b) + (d), samples. The structure of cellobiose, together with the ring and atom numbering used, is shown in the inset.

However, this time the populations of the *anti*-conformers were approximately reversed. In contrast to these results, quantum mechanical calculations in the gas-phase predicted the dominance of the *anti*-ϕ_H/*syn*-ψ_H conformer.[15,16] The existence of this conformer in the gas phase was later confirmed (in the presence or absence of a single water molecule) using vibrational spectroscopy under molecular beam conditions in conjunction with quantum chemical calculations.[17] Recent DFT calculations using the COSMO solvation model,[21] while still favoring the *anti*-ϕ_H/*syn*-ψ_H conformer, have reduced the energy difference between the *anti*-ϕ_H and *syn*-ϕ_H states. Subsequent use of the SMD solvation model has placed the *syn*-ϕ_H conformer 2 kcal mol^{-1} below the minimum of the *anti*-ϕ_H conformer.[22]

Substantiated by the analysis of inter-residue ROEs and hetero- and homonuclear three-bond interglycosidic coupling constants, experimental NMR studies in water have consistently identified *syn*-ϕ_H/*syn*-ψ_H as the major conformer of cellobiose.[27–31] Investigations by the Widmalm group[28,31] concluded that the *syn*-conformer is present at ~93%, while the minor *anti*-ϕ_H and *anti*-ψ_H states are populated to 0–2% and 5–7%, respectively. Although in a broad agreement with some of the theoretical studies,[18,20] given the very small populations of the minor conformers, one has to be cautious when interpreting the experimental NMR. This is because such interpretation relies on a number of simplifications, such as the two-spin approximation, the assumption of uncorrelated molecular motions or isotropic reorientation, all of which are used in the analysis of proton cross relaxation rates. The Karplus curves parameterisation used for the interpretation of scalar coupling constants is also only an approximation. Most crucially, the analysis of averaged NMR parameters associated with flexible systems relies on the use of computed structural models. Given the variations between the different force fields and solvation methods demonstrated above, the time averaging of theoretical parameters and their ''agreement'' with experimental data cannot therefore be taken as a definite proof of the existence of conformational flexibility. Approaching the problem without any preconception, one has to ask two questions: (i) is the glycosidic linkage of cellobiose rigid or does it exhibit some flexibility? (ii) if flexible, what is the nature of the underlying motion? In this work we have employed RDCs and scalar coupling constants to answer the first question and to hypothesize on the second.

Experimental

All chemicals were purchased from Sigma-Aldrich®. Cellobiose (40 mg, 0.12 mmol) was dissolved in 550 μl of 100% D_2O or in the alignment medium. The alignment media consisted of a ternary mixture[32] of pentaethylene glycol monododecyl ether, $C_{12}E_5$ (59.8 μl, 0.14 mmol), *n*-hexanol (18.5 μl) and D_2O (550 μl). The deuterium splitting was 44 Hz at 297 K. Sample preparation details are given in the ESI.‡

All spectra of aligned and isotropic cellobiose were recorded at 297 K on an 800 MHz Bruker Avance spectrometer equipped with a TCI triple-resonance cryoprobe with z gradients. Twenty 1D intensity modulated ^{13}C spectra were collected for the determination of one-bond splittings using the pulse sequence of Fig. 2a or b in a pseudo 2D manner. The following parameters were used: number of scans 16 × 12 = 192 (12 repetitions of each experiment in a cyclic manner to ensure long-term averaging of spectrometer instabilities), an acquisition time of 0.68 s, and a relaxation time of 1.5 s. The BIRD^{d,X} pulse[33,34] was optimised for $^1J_{CH}$ = 155 Hz by setting $\Delta = 1/2^1J_{CH}$. The evolution interval T was set to 4, 6, 8, 10 ms and then 150 to 165 ms in 1 ms increments. Pulsed field gradients, G_o–G_3, were applied over 500 μs at 33, 13, 18 and 27%, respectively. The total acquisition time was 2.75 hours.

The long-range coupling constants were measured using the pulse sequence of Fig. 2c (or 2d). The following parameters

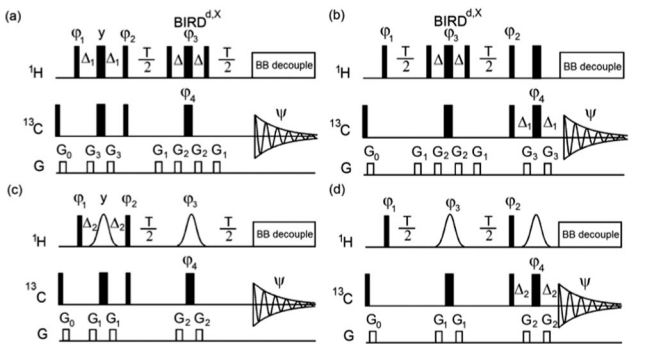

Fig. 2 Pulse sequences of intensity modulated ^{13}C-detected 1D experiments for the measurement of one-bond, (a) and (b), and long-range, (c) and (d), proton–carbon coupling constants. Narrow and wide rectangles represent 90° and 180° rectangular pulses, respectively, with phase *x*, unless stated otherwise. T is the variable evolution interval, $\Delta = 1/2^1J_{CH}$, $\Delta_1 = 1/4^1J_{CH}$, $\Delta_2 = 1/4^nJ_{CH}$. The gradients were applied for a period of 0.5 ms at the following strengths: G_0 = 33%, G_1 = 13%, G_2 = 18%, G_3 = 27%. The following phase cycling was as follows: $\varphi_1 = 8x$, $8(-x)$; $\varphi_2 = y, -y$; $\varphi_3 = 2x, 2(-x)$; $\varphi_4 = 4x, 4(-x)$ and $\psi_{rec} = 4(x, -x), 4(-x, x)$.

were used: number of scans 16 × 60 = 960 (60 repetitions of each experiment in a cyclic manner), an acquisition time of 0.68 s, and a relaxation time of 1.5 s. A 30 ms Gaussian selective 180° ^{1}H pulse was applied. The refocusing interval ($\Delta_2 = 1/4^nJ_{CH}$) was optimised for a 5 Hz coupling constant. The evolution interval, T, was set from 50 to 525 ms in 25 ms increments, G_o–G_3 were applied over 500 μs at 33, 13, 18 and 27%, respectively.

The RDCs were determined as a difference between the splittings measured in the aligned and isotropic media. The sign of RDCs was, in most cases, determined unambiguously based on the known sign of the scalar couplings. Where this was not possible, *e.g.* $D_{H2,H4}$, the sign was implied during the calculation of the alignment tensor. The experimental splittings were determined by a three parameter fit (I_o, K_{CH}, and T_2) of signal intensities derived from the series of intensity modulated spectra using eqn (1):

$$I = I_0 \sin(\pi^1 K_{CH} T)\exp[-(T + 2\Delta)/T_2] \quad (1)$$

where I is the signal intensity at time t, I_o is the maximum signal intensity, and T_2 is the effective spin–spin relaxation time. K_{CH} is either J_{CH} or $J_{CH} + D_{CH}$ and $\Delta = 1/2^1J_{CH}$. The long range splittings were determined by a four parameter (I_o, $^nK_{CH}$, τ_{eff}, and T_2) fit of signal intensities using eqn (2):

$$I = I_0 \sin[\pi^n K_{CH}(T + \tau_{eff})]\exp[-(T + \tau_{180})/T_2] \quad (2)$$

where τ_{eff} is an effective evolution of couplings during the selective inversion pulse and τ_{eff} is its duration.

Calculations were performed using the resources of the National Service for Computational Chemistry Software (NSCCS) (http://www.nsccs.ac.uk) utilizing the Gaussian 03 suite of programs.[35] Coordinates for the molecular structure of cellobiose were taken from a neutron diffraction structure of β-D-cellobiose.[36] For the analysis of α-D-cellobiose, the coordinates of the reducing ring were replaced with those of the neutron diffraction structure of α-D-glucose.[37] The heavy atom coordinates were fixed at these positions. The position of the hydrogen atoms was then optimized using the B3LYP functional[38–40] with the 6-31G* basis set.[41–43]

Results

The mutarotation of the anomeric centre of carbohydrates is slow compared to the time scale of NMR experiments, hence α- and β-D-cellobiose can be treated as independent molecules.[44] Nevertheless, as a consequence of the anomeric equilibrium (30 and 70% of α- and β-D-cellobiose, respectively) ^{1}H and ^{13}C signals of both the non-reducing (**I**) and the reducing (**II**) monosaccharide rings[45] are doubled. These rings are referred to here as $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$ in α-D-cellobiose and $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ in β-D-cellobiose. While the ^{1}H and ^{13}C chemical shifts of rings **II** are sufficiently distinct, there is only a few Hz difference between the chemical shifts of the corresponding nuclei of rings **I**. This creates severe signal overlap as illustrated by the 800 MHz ^{1}H spectra of the isotropic and aligned samples of cellobiose in Fig. 1. In addition, β-D-cellobiose shows higher order ^{1}H spectra even at this high magnetic field. In order to circumvent both these problems, NMR studies of cellobiose are usually carried out using methyl-cellobiosides,[19,27,30,31,46,47] methyl-α-D-cellobioside in particular.[30,46,47] We have decided not to follow this strategy for two reasons. Firstly, it was suggested, based on the analysis of the RDCs of lactose, cellobiose and maltose,[48] that the anomeric configuration of the reducing ring can affect the geometry of the glycosidic bond. Secondly, it was observed that the addition of a small hydrophobic group at the anomeric carbon, such as Me, can induce specific associations with certain types of aligning media.[49] Such interactions can potentially also affect the conformation of the glycosidic linkage. The decision to study unmodified cellobiose however brought several challenges, in particular the accurate measurement of the coupling constants of both anomeric forms.

Measurements of $^{1,n}D_{CH}$, D_{HH}, and $^{1,n}D_{CC}$ RDCs

The minute ^{1}H and ^{13}C chemical shift differences of cellobiose (Fig. 1 and Fig. S1, ESI‡) make it very difficult to obtain accurate coupling constants of α- and β-D-cellobiose using 2D ^{1}H–^{13}C, HSQC-type experiments. In contrast, differences of only a few of Hz provide sufficient resolution of ^{13}C resonances in 1D ^{1}H-decoupled ^{13}C spectra. This was the case for half of the corresponding ^{13}C resonances of the non-reducing rings **I** (Fig. S1, ESI‡). We therefore designed *J*-modulated ^{13}C-detected experiments (Fig. 2) that take the advantage of the superior resolution of ^{13}C spectra and yield the values of heteronuclear splittings.

These experiments can be characterised as 1D refocused INEPT[50] experiments with a BIRD^{d,X} pulse[33,34] applied amid a spin-echo interval. Exclusive inversion of ^{13}C-attached protons during this variable time interval refocuses proton–proton (or long-range proton–carbon) scalar and dipolar couplings, while at the same time preserves the signal modulation by $^1K_{CH}$ ($^1J_{CH}$ or $^1J_{CH} + {}^1D_{CH}$). This modulation can take place while the magnetisation is on the carbon (Fig. 2a) or the proton (Fig. 2b) spins. Fitting the signal intensities of individual carbon resonances according to eqn (1) yields the values of $^1K_{CH}$.

The pulse sequences of Fig. 2a and b can easily be modified for the measurement of long-range proton–carbon splittings ($^nK_{CH}$). By inverting "selectively" a single ^{12}C-attached proton during the evolution and refocusing intervals (pulse sequences of Fig. 2c and d), the signal intensities of all carbons that are long-range coupled to the inverted proton are modulated by $^nK_{CH}$. It may seem that the use of these methods is limited due to the extensive overlap of proton resonances discussed above. However, the selectivity condition only requires the chemical shift of the inverted proton to be sufficiently separated from the chemical shift of other protons with which it shares a mutual coupling. In addition, if more than one inverted proton is long-range coupled to the same carbon, this carbon cannot be analysed using eqn (2). Also, one has to consider the ^{1}H spectra of individual ^{13}C-isotopomers, as large one-bond couplings change the position of resonances. Nevertheless, the limitations of the use of this experiment can often be reduced to the requirement for proton resonances within one monosaccharide ring to be sufficiently separated. While for isotropic samples it

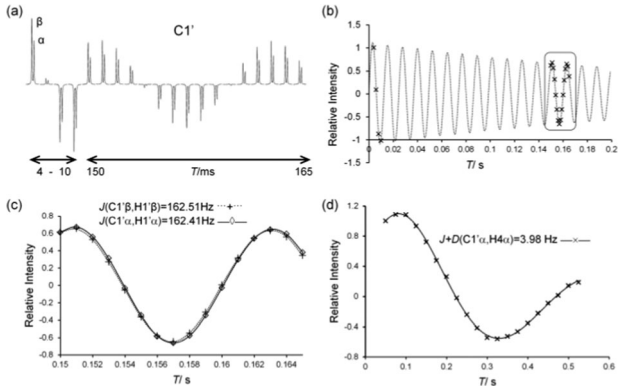

Fig. 3 (a) $C_{1'\alpha}$ and $C_{1'\beta}$ signals from the intensity-modulated spectra of the isotropic sample of cellobiose (pulse sequence of Fig. 2a), as a function of the evolution time *T*; (b) intensity of the $C_{1'\beta}$ signals shown in (a) fitted to eqn (1); (c) expansion of the boxed area from (b) showing both signals; (d) fitting (eqn (2)) of the signal intensities of $C_{1'\alpha}$ acquired on the aligned sample using the pulse sequence of Fig. 2c with the selective inversion of $H_{4\alpha}$.

is straightforward to ascertain if this is the case, care must be taken when this experiment is applied to the aligned samples.

$^1K_{CH}$ splittings were measured using the pulse sequences shown in Fig. 2a and b. Both pulse sequences yielded identical values; the long-range splittings were consequently measured only by the pulse sequence of Fig. 2c. An example of the application of these methods to cellobiose is shown in Fig. 3. As can be seen from Fig. 3a–c, a very small difference between $J_{C1'\beta,\,H1'\beta}$ (162.51 Hz) and $J_{C1'\alpha,\,H1'\alpha}$ (162.41 Hz) is detectable using these methods. Altogether 14 out of 20 $^1D_{CH}$ of CH pairs were measured for the four monosaccharide rings (Table 1). The remaining six RDCs, ($^1D_{C3',H3'}$, $^1D_{C4',H4'}$ and $^1D_{C5',H5'}$ of rings $\mathbf{I}_\alpha$ and $\mathbf{I}_\beta$) were affected by higher order effects in the ^{13}C-satellite spectra and were therefore excluded from the subsequent analysis. Also measured were 13 $^nD_{CH}$ intra-ring couplings and three out of four inter-ring $^3D_{CH}$ couplings across the glycosidic linkage. Unfortunately $^3D_{C1',H4\beta}$ was not accessible because of strong coupling between H_3 and H_4 of ring $\mathbf{II}_\beta$. An example of the analysis of a long-range splitting is shown in Fig. 3d.

The heteronuclear RDCs obtained by these experiments were supplemented by 10 $^nD_{H,H}$, 18 $^1D_{C,C}$ and two $^2D_{C,C}$ RDCs measured using intensity-based selective 1D methods[51,52] and a 2D INADEQUATE experiment[53] described previously. The experimental details of these measurements together with illustrative examples are given in Fig. S2 and S3 (ESI‡).

Order parameters of cellobiose rings are similar but not identical

RDCs of rings **I** and **II** were initially analysed separately. The orientations of individual principal axes frames (PAFs), principal order parameters, generalised order parameters (GDOs) and the rhombicities, η, were calculated using SVD[12] (see ESI‡ for definitions). These calculations are based on 12 RDCs each for rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$, and 11 and 20 RDCs for rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$, respectively. A very good agreement between the experimental and theoretical RDCs was obtained with rmsds between 0.08 and 0.14 Hz (Table 1) indicating the high accuracy of the measured RDCs.

GDO has been proposed as a convenient parameter for assessing the flexibility of moieties connecting rigid molecular fragments.[7] When applied to carbohydrates, GDOs of rigid monosaccharide rings connected by rigid glycosidic linkages are expected to be identical. Our calculations show that the GDOs of rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ differ by 1% while GDOs of rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$ differ by 23%. Therefore based solely on this criterion, β-D-cellobiose appears to be a rigid molecule, while α-D-cellobiose shows signs of flexibility. In order to investigate this apparent discrepancy further, we analysed both molecules in the approximation of a rigid structure. This involved exploring the conformational space described by the glycosidic torsion angles ϕ and ψ in a search for the best fit between the experimental and theoretical RDCs. The results presented below use angles ϕ_r, ψ_r, which are the angles relative to those of the X-ray structure of β-D-cellobiose[36] [$(\phi_r, \psi_r) = (0^\circ, 0^\circ) = (\phi_H, \psi_H) = (+44^\circ, -12^\circ)$].

Conformation of cellobiose: rigid structure approximation

The glycosidic linkage of α- and β-D-cellobiose was rotated in 30° steps for both ϕ and ψ angles generating 121 structures per molecule. The parameters of the order tensors were then calculated using the SVD method for each structure in two different ways. Firstly, all RDCs of rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ (24 intraring + 2 interring) or rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$ (31 intraring + 3 interring) were taken into account. The rmsds between the experimental and theoretical RDCs as a function of ϕ_r, ψ_r angles are shown in the form of a 2D map in Fig. 4. As can be seen, the coordinates of the global minima for both α- and β-D-cellobiose coincided with the torsion angles determined using X-ray crystallography [$(\phi_r, \psi_r) = (0, 0)$]. A more detailed inspection of the global minima in steps of 10° refined their positions and placed them at $(\phi_r, \psi_r) = (-10, 0)$ and $(\phi_r, \psi_r) = (0, -10)$ for α- and β-D-cellobiose, respectively (data not shown). Nevertheless, the minima identified using this method are rather flat and not very distinct from additional local minima; in addition, the highest rmsds seen on these 2D maps are smaller than anticipated.

In the second approach, the order matrices were calculated using the RDCs of ring **I**, while the rmsds were based on the RDCs of ring **II** and *vice versa*. Two sets of rmsds for each of 121 structures of α- and β-D-cellobiose calculated in this way are presented in Fig. 5. A more detailed inspection of the global minima in steps of 10° (Fig. S4, ESI‡) yielded values of (ϕ_r/°, ψ_r/°, rmsd/Hz) = (−10, −30, 1.10) and (−10, 0, 0.25) based on the alignment of ring $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$, respectively. Analogous analysis of β-D-cellobiose generated global minima at (−10, −40, 0.41) and (−10, 0, 0.31), based on the alignment of rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$, respectively. Several observations can be made based on the inspection of these 2D maps. (i) All maps (Fig. 5) show a better definition of minima contained in deeper valleys compared to when RDCs of both rings were considered at the same time (Fig. 4); (ii) general trends are identical for both α- and β-D-cellobiose. In all four cases a global minimum was found at $\phi_r = -10^\circ$, while ψ_r showed two values: $\psi_r = 0^\circ$ (based on the alignment of rings **I**) or $\psi_r = -30^\circ$ (or -40°)

Table 1 RDCs of cellobiose and order matrix parameters

	α-D-Cellobiose				β-D-Cellobiose			
	Ring $\mathbf{I}_\alpha$ (non-reducing)		Ring $\mathbf{II}_\alpha$ (reducing)		Ring $\mathbf{I}_\beta$ (non-reducing)		Ring $\mathbf{II}_\beta$ (reducing)	
Atoms	Exp.	Calc.	Exp.	Calc.	Exp.	Calc.	Exp.	Calc.
One-bond proton–carbon couplings								
C1–H1	12.60	13.13	−9.40	−9.26	12.38	12.80	12.67	12.74
C2–H2	12.65	12.30	13.87	13.88	12.71	12.40	13.05	12.82
C3–H3	[a]		13.96	13.77	[a]		13.18	13.20
C4–H4	[a]		13.19	13.37	[a]		12.67	12.47
C5–H5	[a]		13.65	13.43	[a]		12.66	12.88
Long-range proton–carbon couplings								
C1–H2	0.61	0.51	—		0.26	0.35	1.13	0.89
C2–H3	−0.21	0.06	0.23	0.23	0.1	0.47		
C3–H1			−0.3	−0.26				
C3–H2			0.2	0.06			−0.32	−0.51
C3–H4			0.23	0.40				
C4–H3			0.58	0.49				
C5–H1			−0.63	−0.65				
C5–H4			0.87	0.99				
Proton–proton couplings								
H1–H2	1.84	1.76	−0.63	−0.82	1.48	1.64	1.93	1.90
H1–H3					−6.26	−6.35		
H2–H3	1.04	1.12	1.56	1.46	1.31	1.42		
H2–H4			−5.63	−5.63				
H3–H4			1.34	1.58				
One-bond carbon–carbon couplings								
C1–C2	−0.85	−0.76	0.45	0.38	−0.89	−1.08	0.35	0.19
C2–C3	−0.98	−1.00	[a]	[a]	−0.95	−0.76	[a]	[a]
C3–C4	0.43	0.75	−0.55	−0.41	0.65	0.42	−0.05	−0.10
C4–C5	−0.75	−1.13	0.1	0.35	−1.1	−1.25	0.15	0.04
C5–C6	0.7	0.59	−0.1	−0.29	0.6	0.52	0.1	0.38
Inter-ring couplings								
		Exp.	Calc.			Exp.	Calc.	
C1′–H4		−1.03	−0.73			[a]	[a]	
C4–H1′		−1.2	−1.48			−1.05	−0.99	
C1′–C4		−0.35	−0.45			−0.48	−0.48	
Number of intra-ring RDCs (Number of unique RDCs)								
	11(7)		20(15)		12(9)		12(6)	
Order matrix parameters based on the analysis of individual rings								
10^4 S_{zz}	3.74		4.72		4.07		4.10	
10^4 S_{yy}	−2.95		−3.09		−2.92		−3.00	
10^4 S_{xx}	−0.79		−1.6		−1.15		−1.10	
10^4 GDO/η	3.94 ± 0.10/0.58 ± 0.06		4.80 ± 0.06/0.31 ± 0.02[d] 4.58 ± 0.18/0.41 ± 0.06[e]		4.20±0.04/0.44±0.02		4.24 ± 0.12/0.46 ± 0.05	
rmsd[b]	0.10		0.14[d] 0.14[e]		0.08		0.12	
(ϕ_r, ψ_r, rmsd)	(−10°, −30°, 1.10)		(−10°, 0°, 0.25)		(−10°, −40°, 0.41)		(−10°, 0°, 0.31)	
(ϕ_r, ψ_r, rmsd)	(−10°, −10°, 1.60)		(−10°, 0°, 0.65)		(−10°, −10°, 1.36)		(−10°, 0°, 0.71)	
Order matrix parameters based on simultaneous analysis of both rings								
10^4 GDO/η	4.78 ± 0.05/0.31 ± 0.01				4.30 ± 0.05/0.43 ± 0.02			
rmsd[c]	0.19				0.19			

[a] Not included in calculations because of higher order effects. [b] Using RDCs of individual monosaccharides separately. [c] Using RDCs of both monosaccharides including the interglycosidic RDCs. [d] Based on 20 RDCs of ring **II**. [e] Based on 11 RDCs of ring **II**.

(based on the alignment of rings **II**); (iii) the position of the minima changed compared to when RDCs of both rings were considered simultaneously.

These results indicate that the approximation of a rigid structure is not valid for cellobiose. If cellobiose was completely rigid, identical results would have been obtained regardless of which ring was used to calculate the order matrix parameters. The most interesting result of this analysis is the fact that both α- and β-D-cellobiose show identical trends, *i.e.* one set of dihedral angles based on the RDCs of the non-reducing ring and another set based on those of the reducing ring. This is despite the fact that GDOs of both rings of β-cellobiose were

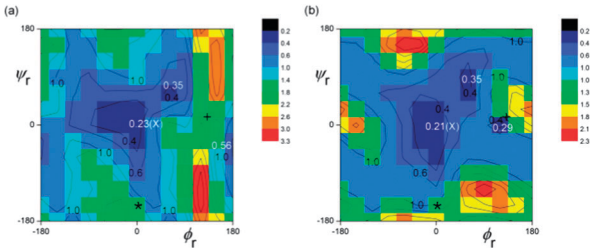

Fig. 4 2D rmsd maps between the experimental and theoretical RDCs based on the simultaneous analysis of RDCs of both rings as a function of ϕ_r, ψ_r angles. (a) α-D-cellobiose, the minima were observed at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 0.23), (60, 60, 0.35) and (180, −30, 0.56); (b) β-D-cellobiose, the minima were observed at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 0.21), (120, 0, 0.29) and (60, 60, 0.35). The conformers identified by the molecular mechanics calculations[20] are indicated here (and in Fig. 5–7) by * (*syn*-ϕ_H/*anti*-ψ_H) and + (*anti*-ϕ_H/*syn*-ψ_H); a white cross at (0, 0) indicates the X-ray structure of cellobiose.[36]

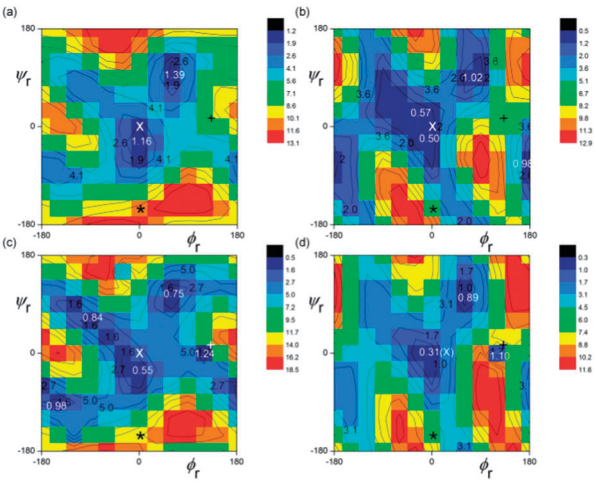

Fig. 5 2D rmsd maps between the experimental and theoretical RDCs as a function of ϕ_r, ψ_r angles. (a) α-D-cellobiose, the rmsds for ring **II** are based on the order parameters of ring **I**. The main minima occur at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, −30, 1.16) and (60, 90, 1.39); (b) α-D-cellobiose, the rmsds for ring **I** are based on the alignment of ring **II**. The minima occur at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 0.50), (60, 90, 1.02) and (180, −60, 0.98); (c) β-D-cellobiose, rmsds for ring **II** are based on the alignment of ring **I**. The main minima occur at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, −30, 0.55), (90, 90, 0.75), (−90, 60, 0.84), (−150, −90, 0.98) and (120, 0, 1.24); (d) β-D-cellobiose, the rmsds for ring **I** are based on the alignment of ring **II**. The minima occur at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 0.31), (60, 90, 0.89) and (120, 0, 1.10).

identical, while those of α-cellobiose differed. It is thus evident that the second treatment of RDCs is preferable. Simultaneous consideration of RDCs of both rings produces an ''average'' structure, characterised by shallow definition of the minima on the RDC rmsd maps. It should be noted that considering the cellobiose rings individually also produces local minima (see Fig. 5). Do these minima have a physical meaning? Can they be interpreted as signs of the flexibility of the glycosidic linkage? It has been observed previously[54] that a random combinations of dihedral angles can yield small rmsds between the theoretical and experimental RDCs. This is a consequence of the mathematical relationship between the order matrix and RDCs and does not necessarily mean that these minima have a physical meaning. Using the approximation of a rigid structure the relevance of these minima can be investigated through the comparison with other experimental parameters, such as scalar inter-residue coupling constants, as illustrated next.

Rigid structure approximation: inclusion of interglycosidic $^3J_{CH}$ and $^3J_{CC}$ coupling constants

The obtained values of the $^3J_{CH}$ and $^3J_{CC}$ coupling constants (Table 2) are in a very good agreement with the literature values for ^{13}C enriched[30] or ^{13}C natural abundance[29,55] cellobiose.

Using the appropriate Karplus type equations[56,57] the combined rmsds between five experimental and theoretical couplings (see ESI‡) were calculated as a function of ϕ_r and ψ_r dihedral angles for α-D-cellobiose. Very similar results were obtained for β-D-cellobiose, although here only one of the two interglycosidic $^3J_{CH}$ coupling was available (data not shown). When only more readily measurable $^3J_{CH}$ couplings are used, a *J*-based 2D rmsd map of α-D-cellobiose (Fig. S5, ESI‡) shows four sets of minima for each dihedral angle. When both $^3J_{CH}$ and $^3J_{CC}$ couplings are considered, a *J*-based 2D rmsd map of α-D-cellobiose (Fig. 6) shows four minima, at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (−10, −6, 0.22), (−181, −6, 0.28), (−10, 163, 0.57) and (−181, 163, 0.63), where the global minimum at (ϕ_r, ψ_r) = (−10, −6°) is very close to the coordinates of the X-ray structure. The appearance of several minima is caused by the periodic nature of the Karplus curves and limited accuracy of their parameterisation. Fig. 6 also shows the positions of the *anti*-ϕ and *anti*-ψ conformers as predicted by molecular modelling. Although distinct, these are close to two local minima shown on this map. Hence the usefulness of *J*-couplings in distinguishing between *syn* and *anti*-conformers is likely limited.

When RDCs based rmsds of α-D-cellobiose (Fig. 5) were combined with *J* based rmsds (Fig. 6), the global minima identified individually by each method reinforced each other and dominated the rmsd landscape (Fig. 7). Their closer inspection in steps of 10° based on the alignment of ring $\mathbf{I}_\alpha$ or $\mathbf{II}_\alpha$ (Fig. S6, ESI‡) positions them at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (−10, −10, 1.60) or (−10, 0, 0.65), respectively. Analogous analysis of β-D-cellobiose yielded similar results (data not shown), but is not as informative because the $^3J_{C1',\,H4}$ coupling constant cannot be measured for this anomeric form. Addition of interglycosidic *J* couplings thus (i) eliminated the secondary minima produced by the analysis of the RDCs only, and (ii) moved the global minima to within 10°of the X-ray structure.

Table 2 Summary of *J* and RDCs [Hz] across the glycosidic linkage of cellobiose

	α-D-Cellobiose		β-D-Cellobiose	
Atoms	*J*	*D*	*J*	*D*
H1′, C4	3.95	−1.20	3.97	−1.05
C2′, C4	3.1	*b*	3.1	*b*
C1′, H4	5.01	−1.03	*a*	*a*
C1′, C3	~0.5	*b*	*b*	*b*
C1′, C5	2.1	*b*	2.1	*b*

[a] Not measured because of higher order effects. [b] Too small to be measured.

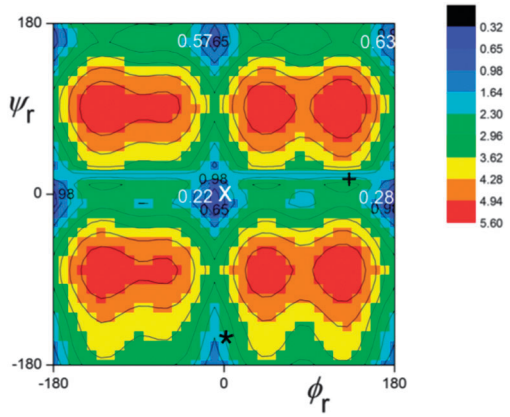

Fig. 6 Combined rmsds between the experimental and theoretical $^3J_{CH}$ and $^3J_{CC}$ coupling constants as a function of ϕ_r, ψ_r angles. The minima are located at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (−10, −6, 0.22), (−181, −6, 0.28), (−10, 163, 0.57) and (−181, 163, 0.63).

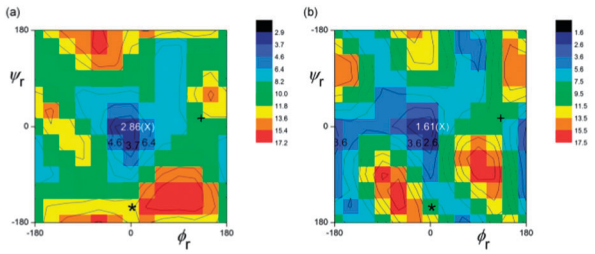

Fig. 7 2D rmsd maps between the experimental and theoretical RDCs and *J* coupling as a function of ϕ_r, ψ_r angles of α-D-cellobiose. (a) The rmsds of ring **II** were based on the alignment of ring **I**. The global minimum was found at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 2.86); (b) the rmsds for ring **I** were based on the alignment of ring **II**. The global minimum was found at (ϕ_r/°, ψ_r/°, rmsd/Hz) = (0, 0, 1.61).

In summary, the interpretation of RDCs and inter-ring *J* couplings *in the approximation of a rigid glycosidic linkage* yielded a solution conformation of cellobiose that is very close to that observed by X-ray in the solid state. Based on this analysis, it is very unlikely that the secondary minima on the RDC rmsd maps have a physical meaning, however, as both RDCs and *J* couplings report on an *average conformation* in solution this analysis does not rigorously exclude such a possibility.

Alignment parameters of rings I and II of cellobiose

We now return to examine closely the alignment of the individual rings of cellobiose. In addition to the strength of alignment, characterised by the GDO parameter discussed above, we analyse the orientations of PAFs of individual order matrices. These are conveniently illustrated using Sauson–Flamsteed plots, which show a superposition of PAFs for rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$ (Fig. 8a) and $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ (Fig. 8b). It is evident for both anomeric forms that not all corresponding eigenvectors point in the same direction. Only the *zz* axes are practically collinear, while the two Cartesian frames, PAF ($\mathbf{I}_\alpha$) and PAF ($\mathbf{II}_\alpha$), or PAF ($\mathbf{I}_\beta$) and PAF ($\mathbf{II}_\beta$), are rotated around the *zz* axis by 27° or 20°, respectively. Fig. 8c and d convey the same information, but in addition also relate the orientation of

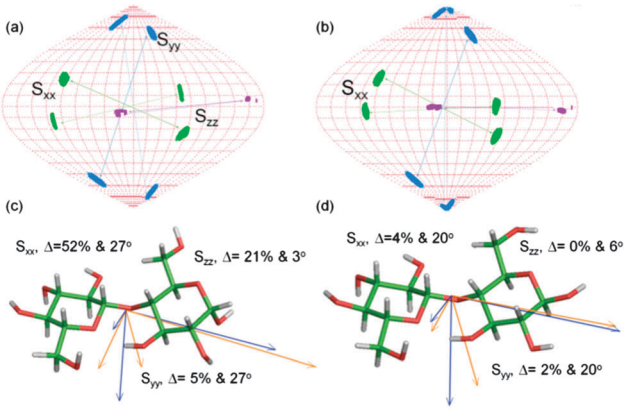

Fig. 8 Sauson–Flamsteed plots for (a) α- and (b) β-D-cellobiose. Full and dotted lines represent the directions of the PAFs for the ring **I** and **II**, respectively. (c) and (d) show the orientation of PAFs based on ring **I** (blue) and ring **II** (orange) in α-D-cellobiose and β-D-cellobiose, respectively. Differences (Δ) in the eigenvalues (%) and orientations (°) between the two corresponding order matrices are also given.

PAFs to the molecular structure of cellobiose. In this figure the relative strength of the alignment is coded by the length of the eigenvectors. It can be seen that the largest eigenvalue (S_{zz}) is associated with the direction along the long axis of the disaccharides. Its orientation within the molecular frames of α- or β-D-cellobiose is practically identical (standard deviation ±2°) regardless of which of the four rings ($\mathbf{I}_\alpha$, $\mathbf{II}_\alpha$, $\mathbf{I}_\beta$, or $\mathbf{II}_\beta$) was used to calculate the *PAF*. The second largest (S_{yy}) ordering direction is approximately perpendicular to the plane defined by the two monosaccharide rings and its orientation differs depending on which ring (**I** or **II**) was used to calculate the PAF. Interestingly, this difference is very similar for α- and β-D-cellobiose.

This analysis clearly shows that despite identical (β-D-cellobiose) or similar (α-D-cellobiose) GDOs of rings **I** and **II**, the alignment of these rings is not identical. This was found to be the case for both α- and β-D-cellobiose and is the reason why the analysis of RDCs in the approximation of a rigid glycosidic bond produced different dihedral angles (Fig. 5) depending on which ring was used to calculate the order matrix. Overall, the above analysis implies that some flexibility is associated with the glycosidic linkage of cellobiose.

Discussion

Measurement of RDCs

As a consequence of the narrow chemical shift range, ^{1}H NMR spectra of carbohydrates often display higher order character. Since RDCs are determined as a difference between the splitting of signals measured in isotropic and aligned samples, strong alignment could change the character of the higher order effects and hence the measured splittings in the aligned samples. This could lead to errors in the measured RDCs. Weak alignment, in combination with methods that can measure the splitting of resonances accurately and precisely, is therefore preferable. We have demonstrated in the past that the intensity-based methods[51,52,58] provide precise and accurate values of splittings

and hence RDCs. Despite severe overlap of ^{1}H resonances, these methods are perfectly adequate for the measurement of ^{1}H–^{1}H RDCs of carbohydrates, as they incorporate some element of signal selection, such as selective excitation,[51] selective polarisation transfers or chemical-shift-selective filters.[52,59] In this work we have proposed 1D methods for the measurement of heteronuclear coupling constants that overcome the limitations associated with severe overlap of ^{1}H resonances. These methods take advantage of a large dispersion of ^{13}C signals and their singlet character and are applicable to the measurement of $^1D_{CH}$ and $^nD_{CH}$ couplings. The increased sensitivity of ^{13}C detection delivered by cryogenic probes makes these experiments a real alternative to their ^{1}H-detected equivalents.

Assessing the accuracy and the precision of RDCs

Accurate and precise RDCs are essential for rigorous interpretation of molecular alignment; however the quality of experimental RDCs can only be assessed by comparing them to their theoretical values calculated using accurate structures of rigid molecular fragments. It is our experience that optimizing positions of protons even in neutron diffraction structures by *ab initio* calculations decreases the rmsds between the experimental and back-calculated RDCs. This approach is mandatory for X-ray structures, which for fundamental reasons have incorrect positions of hydrogen atoms. We have therefore optimised hydrogen positions prior to RDC interpretation using the B3LYP method with the 6-31G* basis set.

An exact solution of an order matrix requires as a minimum five RDCs,[12,60] hence their quality can only be assessed if additional RDCs are available. This exercise is only meaningful if a sufficient number of unique RDCs, *i.e.* measured along non-collinear directions, can be obtained. In hexopyranose chairs of carbohydrates many bond vectors, *e.g.* axial CH vectors, have, at least to a first approximation, identical orientation. When such degenerate RDCs are excluded for rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$, a unique set of 9 (out of 12) and 6 (out of 12) RDCs is obtained, respectively; while 7 (out of 11) and 15 (out of 20) can be classed as unique for rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$, respectively. We therefore have a sufficient number of RDCs to characterise the alignment of α- or β-D-cellobiose rings. In addition, the very small rmsds (0.08 to 0.14 Hz) between the experimental and theoretical RDCs indicate that the obtained RDCs are both accurate and precise.

Unsurprisingly, the largest set of unique RDCs was collected for the $\mathbf{II}_\alpha$ ring. This is due to the equatorial position of the H^α anomeric proton of α-D-cellobiose and a very good dispersion of ^{1}H chemical shifts of this ring. Is this the reason why the *GDO* of this ring is the largest and differs most from *GDO*s of the other rings? In order to answer this question, we have recalculated the order matrix of ring $\mathbf{II}_\alpha$ using a reduced set of RDCs. When only 7 out of 15 unique RDCs, and the same types as available for ring $\mathbf{I}_\alpha$ were used, the GDO parameter changed from 4.82 ± 0.06 to 4.58 ± 0.18. The error bounds of the two GDOs just touch and the GDO based on the reduced set of RDCs is less precise. However, it is still larger than the average GDO of rings $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ (4.22) or the GDO of ring $\mathbf{I}_\alpha$ (3.94). Monte Carlo simulations show a comparable spread of the eigenvector orientations (Fig. 8a and b) for all rings, with an exception of the S_{zz} component of ring $\mathbf{II}_\alpha$. This is smaller due to the largest number of unique RDCs measured. In all instances the spread of eigenvectors is much smaller than the differences between the orientations of the S_{yy} or S_{xx} principle axes for any of the rings. We can therefore conclude that the experimental RDCs are measured accurately and precisely and that the order matrix parameters derived for individual monosaccharide rings are meaningful.

Is the interaction of cellobiose with the aligning media nonspecific?

Freedberg *et al.* have reported 2–3 fold larger axial $^1D_{CH}$ RDCs for the reducing $\mathbf{II}_\alpha$ ring of lactose compared to the reducing ring, $\mathbf{II}_\beta$, measured in DMPC/DHPC or (DMPC/DHPC/CTAB) bicelles.[48] Presumably due to the overlap of $\mathbf{I}_\alpha$ and $\mathbf{I}_\beta$ resonances only one set of RDCs was reported for the non-reducing rings of this disaccharide. The authors interpreted the data by suggesting that the anomeric configuration affects the overall shape of the molecule and thus the conformation of the glycosidic linkage. This could indeed explain the reported difference between the $^1D_{CH}$ RDCs of rings $\mathbf{II}_\alpha$ and $\mathbf{II}_\beta$. The authors also indicated that similar results were obtained for cellobiose and maltose. Our data for cellobiose aligned in $C_{12}E_5$–*n*-hexanol–D_2O medium showed the average axial $^1D_{CH}$ RDCs of rings $\mathbf{II}_\alpha$ and $\mathbf{II}_\beta$ to be 13.7 ± 0.34 Hz and 12.9 ± 0.27 Hz, respectively, *i.e.* they differ on average only by 0.8 Hz. Much larger differences observed by Freedberg *et al.* are therefore very likely caused by a specific interaction of 1 → 4 linked disaccharides with bicellar media, which is presumably different for the α- and β-anomeric configuration of the reducing monosaccharide.

Is there any evidence in our data for a specific interaction of either α- or β-D-cellobiose with the $C_{12}E_5$–hexanol–water aligning medium? Focusing initially on the strength of the alignment, there is practically no difference between the $\mathbf{I}_\beta$ and $\mathbf{II}_\beta$ rings, while larger differences were observed for rings $\mathbf{I}_\alpha$ and $\mathbf{II}_\alpha$, the S_{zz} and S_{xx} components in particular (Table 1). However, as seen in Fig. 8, the orientation of PAFs changes in an identical manner when switching the alignment from one based on the RDCs of the non-reducing to that of the reducing rings for both α- and β-D-cellobiose. The largest principal component does not change its orientation at all. Furthermore this orientation is practically identical for all four rings. At the same time the PAF rotates by about 24° (±4°) around the S_{zz} axis. It is worth noting that $^3J_{H1',C4}$, which is sensitive to the ϕ angle, was identical (3.95 and 3.97 Hz) for both anomeric forms of cellobiose. Unfortunately, $^3J_{C1',H4}$, which is sensitive to the ψ angle could only be measured for α-D-cellobiose due to the higher order between H3 and H4 protons of β-D-cellobiose. The other ψ-dependent coupling constant, $^3J_{C1',C5}$, was small (2.1 Hz) and identical for both anomeric forms of cellobiose. All available corresponding interring RDCs (Table 2) were very similar between α- and β-D-cellobiose indicating that the conformation of the glycosidic linkage is the same for both molecules.

This analysis suggests that (i) the nature of the interaction with the $C_{12}E_5$–*n*-hexanol–D_2O medium is very similar for both

anomeric forms of cellobiose (ii) this interaction either affects the conformation of the glycosidic linkage of both forms equally, or it does not affect it at all (iii) the stronger alignment of the $\mathbf{II}_\alpha$ ring (larger S_{zz} and consequently larger $^1D_{CH}$) is likely caused by steric factors reflecting the configuration of the anomeric center. The C_1 axial OH group of ring $\mathbf{II}_\alpha$ strengthens the alignment relative to that of ring $\mathbf{II}_\beta$, in which the OH group is equatorial.

Conformation of cellobiose in solution

As outlined in the Introduction, all available experimental, and the majority of the theoretical data, suggest that the glycosidic linkage of cellobiose in solution has a strong preference for the *syn* conformation on both dihedral angles. Nevertheless, theoretical calculations and interpretation of experimental data concluded that the *anti*-conformers are present, albeit in very small amounts. What can be said about the conformation of the glycosidic linkage of cellobiose based on our data? Let us start by reiterating that both *J* couplings, interpreted through parameterised Karplus equations, and RDCs, interpreted using the SVD method, report on ''an average conformation''. This is due to much faster dynamics of the glycosidic linkages (ps) compared to NMR sampling times (ms to s). The measured *J* couplings therefore represent a population weighted average. A limited number of interglycosidic couplings, together with a limited accuracy of Karplus curves, does not allow for a straightforward deconvolution of *J* couplings into constituting conformers without using model structures. The situation is even more complex for the interpretation of RDCs using the SVD method. Here, the population averaging of different states using the relative depths of the minima obtained from RDC based rmsd maps, such as those shown in Fig. 5, does not have a physical meaning. In conclusion, *J* couplings and RDCs will, in the case of a flexible glycosidic linkage, provide ''virtual conformers''. This is the case for our combined analysis of RDCs and interglycosidic *J* couplings (Fig. 7), which identified the position of the global minimum at $(\phi_r, \psi_r) = (-10^\circ, -10^\circ/0^\circ)$.

By virtue of its magnitude, a large coupling constant cannot be an average of smaller numbers. In cellobiose, the $^3J_{C2',C4}$ reporting on the ϕ angle is large (3.1 Hz, the maximum on the corresponding Karplus curve is 3.65 Hz) as is the $^3J_{C1',H4}$ reporting on the ψ angle (5.0 Hz, the maximum on a corresponding Karplus curve is 5.6 Hz). These values imply that a dominant conformer of cellobiose (i) either occupies a limited ϕ/ψ space characterized by large interglycosidic coupling constants, *i.e.* the *syn* conformer of the X-ray structure, or (ii) that the glycosidic linkage interconverts between states characterized by large coupling constants across the glycosidic bond. Fig. 6 illustrates the fact that scalar coupling constants cannot fully distinguish between the *syn* and *anti*-conformers as both are characterised by large interglycosidic coupling constants. However, the *anti*-conformers referred to here have dihedral angles $(\phi_r, \psi_r) = (-181, -6)$ or $(-10, 163)$, which differ from those calculated $[(\phi_r, \psi_r) = (+136, 12)$ or $(4, -150)]$ by molecular dynamics simulations.[20] As the corresponding interglycosidic couplings of the computed conformers are small, their population must be low; otherwise lower experimental coupling constants across the glycosidic bond would be expected.

In the approximation of a rigid glycosidic linkage and using RDCs only (Fig. 5), we have shown that $\phi_r = -10^\circ$, while $\psi_r = -30^\circ/-40^\circ$ or 0° depending on which rings are used to calculate the order matrix. It must be emphasized that this interpretation of the experimental data is fundamentally flawed because the underlying assumption it relies on – the rigidity of the glycosidic linkage – is not fulfilled. Therefore the obtained dihedral angles cannot be taken at their face values. We also note that $\psi_r = -30^\circ/-40^\circ$ is not compatible with the large $^3J_{C1',H4}$ (5.1 Hz) measured for α-D-cellobiose. Nevertheless, our analysis of RDCs and interglycosidic *J* couplings allows us to conclude that the prevailing conformation of cellobiose in solution is very similar to that observed in the solid state using X-ray crystallography.

The most important finding of our RDC investigation of cellobiose is the fact that the order parameters calculated based on the RDCs of rings **I** and **II** differed slightly (Fig. 8). Our calculations only assumed rigidity of individual monosaccharide rings. No other assumptions were made; no parameterisation or models were needed to interpret the experimental data. These admittedly small differences are indicative of segmental motions of a partially flexible glycosidic linkage. They can be a consequence of local variations of the glycosidic angles of about $\pm 30^\circ$ often seen in MD simulations of carbohydrates.[28] It is possible that a stochastic averaging of local motions will yield different alignment parameters for individual rings and thus a different set of dihedral angles based on the analysis of RDCs in the approximation of a rigid structure as was the case here. Alternatively, the partial flexibility can also be a consequence of an occasional transition to an *anti*-conformer on either of the glycosidic angles. The detected flexibility can also be a combination of both motions. However, based on the treatment of the experimental data presented here, it is not possible to determine which is the case for cellobiose. We would like to note in passing that experimental parameters that could be potentially very useful in assessing the relative populations of the *syn*- and *anti*-conformers in carbohydrates are $^1J_{CH}$ coupling constants of anomeric and aglyconic carbons.[61] Differences of up to 6 Hz were predicted between the two anomeric forms. However, their current parameterisation is not accurate and therefore these coupling constants cannot be used in conformational analysis of the glycosidic linkage at present, despite the fact that they are easily measureable with high accuracy.

Finally, several approaches were put forward to date for the interpretation of RDCs in flexible carbohydrates. These utilise SVD type analysis[10] or predict the alignment without using the experimental RDCs.[9,62] Studying cellobiose using these methods is, in our opinion, a challenging task due to the limited flexibility of this molecule and the existence of a dominant conformer as demonstrated by others and us. Nevertheless, these approaches can in principle shed more light on the nature of the motion underlying the limited flexibility of cellobiose.

Is the flexibility of cellobiose associated with a particular dihedral angle?

Although it is tempting to interpret the fact that the analysis of RDCs in the approximation of a rigid glycosidic linkage consistently yielded $\phi_r = -10^\circ$, while ψ_r varied between $-30/-40$ and 0, as an indication of enhanced flexibility of theψ angle, such implication cannot be made rigorously. Nevertheless, indirect evidence exists to support this hypothesis. Considering the ring carbons C1, 2, 3 and 5 carrying equatorial hydroxyl groups, the ^{13}C relaxation times of two cellobiose-related disaccharides, Me-β-D-lactoside[63] and maltose,[64] showed that these are 6.5 and 25% shorter, respectively, for the non-reducing ring carbons relative to those of the reducing ring. This translates to lower generalized order parameter,[65] S^2, of the reducing rings and implies higher mobility associated with the ψ angle.

It has been suggested that the stabilizing exo-anomeric effect[66] tends to make the ϕ angle more rigid relative to the ψ angle. There is theoretical and experimental evidence to support this assertion. Tvaroška and Carver have analyzed 2-methoxytetrahydropyrans[67] (ϕ angle) and 1,4-dimethoxytetrahydropyrans[68] (ψ angle) using *ab initio* methods and found a deep and narrow global minimum for the ϕ angle, while a broad double-minimum spanning 80° was identified for the ψ angle. Based on these results the authors concluded that the ψ angle is more flexible than the ϕ angle.

A survey of 21 crystal structures of cellobiose and its related oligosaccharides[69] showed that the variations of ϕ angles (up to 29°) are half of that for ψ angles (up to 58°). It is interesting to note that the acetylated carbohydrates cluster together (Fig. S7, ESI‡) and are separated from the other structures in their ψ_H angles. Nevertheless, even if this group was excluded, the range of ψ angles is substantial (up to 39°). Such spread of ψ angles fits well with a broad double-minimum predicted for the ψ_H values by Tvaroška and Carver.[68] Another interesting observation is that the changes in ψ and ϕ angles in this set of related molecules are correlated (Fig. S7, ESI‡). This may also be the case in solution. In summary, the spread of ψ_H angles frozen out in the solid state indicates a possibility that a larger range of ψ_H could be explored in solution.

Overall, the above data suggest that the flexibility of the glycosidic linkage of cellobiose is more likely associated with the ψ rather than the ϕ dihedral angle. Longer cellulose oligosaccharides will need to be studied to decide if this flexibility can be attributed to the end-ring effects or if it is a genuine property of the ψ angles in cellulose.

Conclusion

Using novel methods for the measurement of heteronuclear RDCs on molecules with severely overlapping NMR resonances, together with other intensity- and frequency-based methods for the measurement of different types of RDCs, we have collected a sufficient number of RDCs to perform a detailed analysis of the conformation of cellobiose. Using the approximation of a rigid glycosidic bond, a combination of RDCs and interglycosidic ^3J coupling constants produced a global minimum corresponding to the *syn–syn* conformer with dihedral angles $(\psi_H, \phi_H) = (+34^\circ, -12^\circ/-22^\circ)$ that are within 10° of the X-ray structure of cellobiose. However, we were also able to show that the non-reducing and reducing rings of cellobiose in weakly aligned medium have slightly different order parameters indicating that the glycosidic bond of cellobiose is not completely rigid.

We have suggested that larger flexibility is more likely associated with the ψ rather than ϕ angle. This could be due to occasional transition from *syn*- to *trans*-conformers on either angles, or stochastic local averaging of the glycosidic dihedral angles. From the perspective of cellulose, the movements detected in cellobiose can be exaggerated because both monosaccharide rings of cellobiose are effectively the ''terminal'' residues of cellulose. In order to establish if this is the case, detailed studies of longer cellulose oligosaccharides will be needed. Such oligosaccharides will likely require partial deuteration and/or ^{13}C labeling in order to provide high quality experimental data suitable for their detailed conformational analysis.

The NMR methodology for accurate and precise measurement of large numbers of very small RDCs developed here is applicable to the conformational study of molecules ranging from simple organic molecules to complex carbohydrates. At the same time, we have shown that the detailed analysis of the alignment parameters of rigid molecular fragments can reveal low-level flexibility of molecules.

Acknowledgements

The authors would like to thank Drs. Tran N. Pham and Andrew P. Herbert for writing computer programs and scripts used in this work. Grant Simpson for his investigation of the alignment of cellobiose. The authors would like to acknowledge the use of the EPSRC UK National Service for Computational Chemistry Software (NSCCS) at Imperial College London in carrying out this work.

Notes and references

1 C. M. Thiele, *Eur. J. Org. Chem.*, 2008, 5673–5685.
2 G. Kummerloewe and B. Luy, *TrAC, Trends Anal. Chem.*, 2009, **28**, 483–493.
3 K. E. Kövér, L. Szilágyi, G. Batta, D. Uhrín and J. Jiménez-Barbero, in *Comprehensive Natural Products II*, ed. M. Lew and L. Hung-Wen, Elsevier, Oxford, 2010, pp. 197–246.
4 J. H. Prestegard and X. B. Yi, in *NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances*, ed. J. F. G. Vliegenthart and R. J. Woods, vol. 930, 2006, pp. 40–59.
5 A. Canales, J. Jiménez-Barbero and M. Martín-Pastor, *Magn. Reson. Chem.*, 2012, **50**, S80–S85.
6 C. M. Thiele, V. Schmidts, B. Boettcher, I. Louzao, R. Berger, A. Maliniak and B. Stevensson, *Angew. Chem., Int. Ed.*, 2009, **48**, 6708–6712.
7 F. Tian, H. M. Al-Hashimi, J. L. Craighead and J. H. Prestegard, *J. Am. Chem. Soc.*, 2001, **123**, 485–492.

8 A. Almond and J. O. Duus, *J. Biomol. NMR*, 2001, **20**, 351–363.
9 H. F. Azurmendi and C. A. Bush, *J. Am. Chem. Soc.*, 2002, **124**, 2426–2427.
10 B. Stevensson, C. Landersjo, G. Widmalm and A. Maliniak, *J. Am. Chem. Soc.*, 2002, **124**, 5946–5947.
11 J. C. Xia and C. Margulis, *J. Biomol. NMR*, 2008, **42**, 241–256.
12 J. A. Losonczi, M. Andrec, M. W. F. Fischer and J. H. Prestegard, *J. Magn. Reson.*, 1999, **138**, 334–342.
13 M. Zweckstetter and A. Bax, *J. Am. Chem. Soc.*, 2000, **122**, 3791–3792.
14 A. Almond and J. B. Axelsen, *J. Am. Chem. Soc.*, 2002, **124**, 9986–9987.
15 G. L. Strati, J. L. Willett and F. A. Momany, *Carbohydr. Res.*, 2002, **337**, 1833–1849.
16 A. D. French and G. P. Johnson, *Can. J. Chem.*, 2006, **84**, 603–612.
17 E. J. Cocinero, D. P. Gamblin, B. G. Davis and J. P. Simons, *J. Am. Chem. Soc.*, 2009, **131**, 11117–11123.
18 T. Y. Shen, P. Langan, A. D. French, G. P. Johnson and S. Gnanakaran, *J. Am. Chem. Soc.*, 2009, **131**, 14786–14794.
19 N. J. Christensen, P. I. Hansen, F. H. Larsen, T. Folkerman, M. S. Motawia and S. B. Engelsen, *Carbohydr. Res.*, 2010, **345**, 474–486.
20 L. Peric-Hassler, H. S. Hansen, R. Baron and P. H. Hunenberger, *Carbohydr. Res.*, 2010, **345**, 1781–1801.
21 F. A. Momany and U. Schnupf, *Carbohydr. Res.*, 2011, **346**, 619–630.
22 A. D. French, G. P. Johnson, C. J. Cramer and G. I. Csonka, *Carbohydr. Res.*, 2012, **350**, 68–76.
23 S. Mendonca, G. P. Johnson, A. D. French and R. A. Laine, *J. Phys. Chem. A*, 2002, **106**, 4115–4124.
24 B. J. Hardy and A. Sarko, *J. Comput. Chem.*, 1993, **14**, 848–857.
25 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.
26 R. D. Lins and P. H. Hunenberger, *J. Comput. Chem.*, 2005, **26**, 1400–1412.
27 L. M. J. Kroonbatenburg, J. Kroon, B. R. Leeflang and J. F. G. Vliegenthart, *Carbohydr. Res.*, 1993, **245**, 21–42.
28 E. A. Larsson, M. Staaf, P. Soderman, C. Hoog and G. Widmalm, *J. Phys. Chem. A*, 2004, **108**, 3932–3937.
29 N. W. H. Cheetham, P. Dasgupta and G. E. Ball, *Carbohydr. Res.*, 2003, **338**, 955–962.
30 U. Olsson, A. S. Serianni and R. Stenutz, *J. Phys. Chem. B*, 2008, **112**, 4447–4453.
31 E. Hatcher, E. Sawen, G. Widmalm and A. D. MacKerell, *J. Phys. Chem. B*, 2011, **115**, 597–608.
32 M. Ruckert and G. Otting, *J. Am. Chem. Soc.*, 2000, **122**, 7793–7797.
33 J. R. Garbow, D. P. Weitekamp and A. Pines, *Chem. Phys. Lett.*, 1982, **93**, 504–509.
34 D. Uhrin, T. Liptaj and K. E. Kover, *J. Magn. Reson., Ser. A*, 1993, **101**, 41–46.
35 M. J. T. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian, Inc., Wallingford CT*, 2004.
36 S. S. C. Chu and G. A. Jeffrey, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1968, **B 24**, 830–838.
37 G. M. Brown and H. A. Levy, *Acta Crystallogr., Sect. B: Struct. Sci.*, 1979, **35**, 656–659.
38 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
39 C. T. Lee, W. T. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
40 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200–206.
41 W. J. Hehre, R. Ditchfie and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.
42 Pc. Harihara and J. A. Pople, *Theor. Chim. Acta*, 1973, **28**, 213–222.
43 M. S. Gordon, *Chem. Phys. Lett.*, 1980, **76**, 163–168.
44 C. E. Perles and P. L. Onofrio Volpe, *Acta Chim. Slov.*, 2009, **56**, 209–214.
45 M. U. Roshind, P. Tahtinen, M. Niemitz and R. Sjhohn, *Carbohydr. Res.*, 2008, **343**, 101–112.
46 B. J. Hardy, A. Gutierrez, K. Lesiak, E. Seidl and G. Widmalm, *J. Phys. Chem.*, 1996, **100**, 9187–9192.
47 W. H. Zhang, H. Q. Zhao, I. Carmichael and A. S. Serianni, *Carbohydr. Res.*, 2009, **344**, 1582–1587.
48 D. I. Freedberg, S. O. Ano, S. E. Norris and R. M. Venable, in *NMR Spectroscopy and Computer Modeling of Carbohydrates: Recent Advances*, ed. J. F. G. W. R. J. Vliegenthart, 2006, vol. 930, pp. 220–234.
49 X. B. Yi, A. Venot, J. Glushka and J. H. Prestegard, *J. Am. Chem. Soc.*, 2004, **126**, 13636–13638.
50 G. A. Morris and R. Freeman, *J. Am. Chem. Soc.*, 1979, **101**, 760–762.
51 T. N. Pham, T. Liptaj, P. N. Barlow and D. Uhrin, *Magn. Reson. Chem.*, 2002, **40**, 729–732.
52 L. Jin, T. N. Pham and D. Uhrin, *ChemPhysChem*, 2007, **8**, 1228–1235.
53 L. Jin and D. Uhrin, *Magn. Reson. Chem.*, 2007, **45**, 628–633.
54 A. Almond, J. Bunkenborg, T. Franch, C. H. Gotfredsen and J. O. Duus, *J. Am. Chem. Soc.*, 2001, **123**, 4792–4802.
55 T. Rundlof, A. Kjellberg, C. Damberg, T. Nishida and G. Widmalm, *Magn. Reson. Chem.*, 1998, **36**, 839–847.

56 I. Tvaroška, M. Hricovini and E. Petrakova, *Carbohydr. Res.*, 1989, **189**, 359–362.
57 B. Bose, S. Zhao, R. Stenutz, F. Cloran, P. B. Bondo, G. Bondo, B. Hertz, I. Carmichael and A. S. Serianni, *J. Am. Chem. Soc.*, 1998, **120**, 11158–11173.
58 T. N. Pham, T. Liptaj, K. Bromek and D. Uhrin, *J. Magn. Reson.*, 2002, **157**, 200–209.
59 P. T. Robinson, T. N. Pham and D. Uhrin, *J. Magn. Reson.*, 2004, **170**, 97–103.
60 A. Saupe, *Angew. Chem., Int. Ed.*, 1968, **7**, 97–112.
61 I. Tvaroška and F. R. Taravel, *J. Biomol. NMR*, 1992, **2**, 421–430.
62 J. Xia, C. J. Margulis and D. A. Case, *J. Am. Chem. Soc.*, 2011, **133**, 15252–15255.
63 M. L. Hayes, A. S. Serianni and R. Barker, *Carbohydr. Res.*, 1982, **100**, 87–101.
64 R. B. Best, G. E. Jackson and K. J. Naidoo, *J. Phys. Chem. B*, 2001, **105**, 4742–4751.
65 G. Lipari and A. Szabo, *J. Am. Chem. Soc.*, 1982, **104**, 4546–4559.
66 H. Thogersen, R. U. Lemieux, K. Bock and B. Meyer, *Can. J. Chem.*, 1982, **60**, 44–57.
67 I. Tvaroška and J. P. Carver, *J. Phys. Chem.*, 1994, **98**, 9477–9485.
68 I. Tvaroška and J. P. Carver, *THEOCHEM*, 1997, **395**, 1–13.
69 Z. Peralta-Inga, G. P. Johnson, M. K. Dowd, J. A. Rendleman, E. D. Stevens and A. D. French, *Carbohydr. Res.*, 2002, **337**, 851–861.

