5-14-2018

Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR

Pyae Phyo
Massachusetts Institute of Technology

Tuo Wang
Massachusetts Institute of Technology

Yu Yang
Massachusetts Institute of Technology

Hugh O'Neill
Oak Ridge National Laboratory

Mei Hong
Massachusetts Institute of Technology

Recommended Citation

Phyo, P., Wang, T., Yang, Y., O'Neill, H., & Hong, M. (2018). Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR. *Biomacromolecules, 19* (5), 1485-1497. https://doi.org/10.1021/acs.biomac.8b00039

Biomacromolecules

Article

Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using 1H Polarization Transfer Solid-State NMR

Pyae Phyo, Tuo Wang, Yu Yang, Hugh M. O'Neill, and Mei Hong

Biomacromolecules, **Just Accepted Manuscript** • DOI: 10.1021/acs.biomac.8b00039 • Publication Date (Web): 21 Mar 2018

ACS Publications

Biomacromolecules is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Direct Determination of Hydroxymethyl Conformations of Plant Cell Wall Cellulose Using ^{1}H Polarization Transfer Solid-State NMR

Pyae Phyo [1], Tuo Wang [1$], Yu Yang [1#], Hugh O'Neill [2], and Mei Hong [1]*

[1] Department of Chemistry, Massachusetts Institute of Technology, 170 Albany Street, Cambridge, MA 02139

[2] Center for Structural Molecular Biology, Oak Ridge National Laboratory, Oak Ridge, TN 37831

Revised for ***Biomacromolecules***

February 23, 2018

* Corresponding Author: Mei Hong, E-mail: meihong@mit.edu

[$] Current address: Tuo Wang, Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803.

[#] Current address: Yu Yang, Department of Electronic Science, Xiamen University, Fujian, China, 361005.

Abstract

In contrast to the well-studied crystalline cellulose of microbial and animal origins, cellulose in plant cell walls is disordered due to its interactions with matrix polysaccharides. Plant cell wall (PCW) is an undisputed source of sustainable global energy, therefore it is important to determine the molecular structure of PCW cellulose. The most reactive component of cellulose is the exocyclic hydroxymethyl group: when it adopts the *tg* conformation, it stabilizes intra-chain and inter-chain hydrogen bonding, while *gt* and *gg* conformations destabilize the hydrogen-bonding network. So far, information about the hydroxymethyl conformation in cellulose has been exclusively obtained from ^{13}C chemical shifts of monosaccharides and oligosaccharides, which do not reflect the environment of cellulose in plant cell walls. Here, we use solid-state NMR spectroscopy to measure the hydroxymethyl torsion angle of cellulose in two model plants, by detecting distance-dependent polarization transfer between H4 and H6 protons in 2D ^{13}C-^{13}C correlation spectra. We show that the interior crystalline portion of cellulose microfibrils in *Brachypodium* and *Arabidopsis* cell walls exhibits H4-H6 polarization transfer curves that are indicative of a *tg* conformation, whereas surface cellulose chains exhibit slower H4-H6 polarization transfer that is best fit to the *gt* conformation. Joint constraints by the H4-H6 polarization transfer curves and ^{13}C chemical shifts indicate that it is unlikely for interior cellulose to have a significant population of the *gt* and *gg* conformation mixed with the *tg* conformation, while surface cellulose can adopt a small percentage of the *gg* conformation. These results provide new constraints to the structure and matrix interactions of cellulose in plant cell walls, and represent the first direct determination of a torsion angle in an important non-crystalline carbohydrate polymer.

Introduction

Cellulose, a polysaccharide made up of β-1,4-linked glucan chains, has a deceptively simple chemical structure that belies its complex three-dimensional structure. When produced in highly crystalline forms by microbial and animal sources, the glucosyl residues in each chain are stabilized by sequential O3-H···O5, O2-H···O6 and O2-H···O1 hydrogen bonds. Multiple chains align in parallel by lateral O6-H···O3 and O6-H···O2 hydrogen bonds to form a sheet. Multiple sheets stack up, with inter-sheet separations of ~5 Å, stabilized by van der Waals interactions and weak C-H···O hydrogen bonds [1-4]. Cellulose-producing bacteria, algae and marine tunicates assemble into fibrils with diameters of 8–20 nm, which are sufficiently ordered to produce microcrystals that diffract to atomic resolution. In comparison, plant cell walls (PCWs), which are responsible for making cellulose the most abundant biopolymer on earth, produce elementary microfibrils with smaller diameters of 3-4 nm [5], because PCWs also contain significant amounts of non-cellulosic polymers that interact with cellulose [6-8]. The resulting disorder makes PCW cellulose much less understood than highly crystalline cellulose of microbial and animal origins [9].

Solid-state NMR has played a central role in revealing the cellulose molecular structure [10]. Long before the determination of crystal structures of cellulose, observation of chemical shift doubling of C1 and C4 carbons established the presence of two cellulose allomorphs, Iα and Iβ [11-13]. Cellulose Iα, found in bacteria and algae, is characterized by a C1 singlet at 105 ppm and a C4 doublet at 90 and 89 ppm, while cellulose Iβ, found in marine tunicates, exhibits a C1 doublet at 106 and 104 ppm and a C4 doublet at 89 and 88 ppm [14-16]. Consistently, ^{1}H chemical shifts also differ between the Iα and Iβ allomorphs. Crystal structures confirmed the presence of two non-equivalent structures in each allomorph, showing that cellulose Iα consists of two alternating glucopyranose conformations within each chain while cellulose Iβ consists of two distinct chains in different sheets [1, 2].

However, to what extent these structural features of microbial crystalline cellulose apply to PCW cellulose is so far unknown. In plant primary cell walls, cellulose interacts with hemicellulose and pectins on a sub-nanometer scale, as shown by intermolecular cross peaks in 2D and 3D ^{13}C correlation spectra for both dicot and grass plants such as *Arabidopsis thaliana* and *Brachypodium distachyon* [8, 17-21]. In plant secondary cell walls, significant disorder of cellulose has also been reported [22-27]. Even for the interior of cellulose microfibrils, the chain packing and hydrogen bonding appear to be not identical to those of Iα and Iβ allomorphs [28]. The potential difference is suggested by recently reported high-resolution 2D ^{13}C-^{13}C correlation NMR spectra measured at high magnetic fields [29]. These spectra resolved seven sets of C1-C4 and C4-C6 cross peaks for primary PCW cellulose: five sets can be assigned to interior cellulose while two sets can be assigned to surface cellulose, but none of these cross peaks have identical chemical shifts to those of Iα and Iβ cellulose. In addition, the C2 and C3 chemical shifts of PCW cellulose also have non-negligible differences from the Iα and Iβ values, suggesting that the hydrogen-bonding networks in PCW cellulose may differ from those in Iα and Iβ cellulose. Quantification of the relative intensities of the interior and surface cellulose C4 peaks showed that the number of cellulose chains in PCW microfibrils is larger than 18 but smaller than 36, with an average value around 24 [21], which further supports the notion that PCW cellulose may not be a simple mixture of Iα and Iβ cellulose.

An important structural feature of cellulose is the conformation of the hydroxymethyl group. This exocyclic group is the most reactive component of cellulose [4], with multiple hydrogen-bonding possibilities in Iα and Iβ cellulose [1, 2] and significant mobilities in PCW cellulose [8, 17, 19, 20, 23, 30]. The hydroxymethyl conformation is defined by the torsion angle χ = O5-C5-C6-O6 (**Fig. 1a**) and is traditionally denoted by a two-letter code: the first indicates the O5-C5-C6-O6 torsion angle while the

second indicates the C4-C5-C6-O6 angle. In other words, the *tg* conformation denotes a χ angle of 180˚, which places O5 and C6 *trans* to each other. The *gt* conformation corresponds to χ = +60˚ while the *gg* conformation corresponds to χ = -60˚ (**Fig. 1b**). The different O6H positions impact the hydrogen bonding of cellulose chains. Cellulose Iα and Iβ predominantly show the *tg* conformation, with a χ angle of 158˚ to 170˚, which establishes intra-chain O2-H…O6 and inter-chain O6-H…O2 hydrogen bonds (**Fig. 1b**) [1, 2]. To obtain information about the exocyclic conformations of the less crystalline PCW cellulose, molecular dynamics (MD) simulations have been conducted to probe how decreasing microfibril size affects the hydroxymethyl conformation [31]. These simulations found that a 36-chain Iβ microfibril preserved the *tg* conformation, while 24- and 18-chain Iβ microfibrils adopt significant populations of *gt* and *gg* conformations in both interior and surface chains. ^{13}C spin-lattice T_1 relaxation revealed substantial dynamics in the surface chains, suggesting a loss of O2-H…O6 hydrogen bonds due to conformational differences from the crystal structures [32]. Deconvolution of ^{13}C NMR spectra of algal cellulose that has been subject to layer-by-layer peeling using TEMPO oxidation and alkali extraction suggested that the exterior-facing C6 groups of the outermost surface chains may adopt the *gg* conformation, while the interior-facing C6 groups may adopt the *gt* conformation [33]. However, the intensities of these surface cellulose peaks in these unlabeled samples are low, and the chemical perturbations to the algal cellulose makes these results difficult to interpret.

So far the main experimental observable of the hydroxymethyl conformation in plant cell walls is the ^{13}C NMR chemical shift. Horii and coworkers measured the C6 chemical shift of ten glucose-containing mono-, di-, and tri-saccharides with known crystal structures and found a linear correlation with the exocyclic conformation: a C6 chemical shift of 60.0-62.6 ppm correlates with the *gg* conformation, 62.5-64.5 ppm correlates with the *gt* conformation, while 65.5-66.5 ppm correlates with the *tg* conformation [34]. However, chemical shifts are sensitive not only to local torsion angles but also to hydrogen bonding and interactions with solvents and other molecules, and the environment of small oligosaccharides differs significantly from that of polysaccharides in complex cell walls. There are so far few directly measured distances or torsion angles in carbohydrate polymers to correlate with chemical shifts. DFT calculations have been recently used to provide the missing link between chemical shifts and the cellulose hydroxymethyl conformation [35-38]. These calculations found that C4 and C6 chemical shifts are affected mainly by the χ angle, and to a lesser extent by water hydrogen-bonding on the microfibril surface, glycosidic (ϕ□ □ψ) torsion angles, and distance between HO3 and H4 [38].

The paucity of directly measured distances and torsion angles in cellulose and most carbohydrate polymers limits molecular-level understanding of the structures of this class of important biomolecules [39, 40]. As a first step towards establishing the link between NMR chemical shifts and torsion angles, we have now investigated the hydroxymethyl conformation in PCW cellulose by measuring H4-H6 distances. We detect these distances indirectly through ^{13}C using the 2D solid-state NMR technique, CHHC [41-44]. We first investigate the validity and accuracy of this method using the crystalline bacterial cellulose, which exhibits the *tg* conformation. We then apply this technique to grass and dicot PCWs, and show, for the first time, that interior cellulose predominantly adopts the *tg* conformation while surface cellulose predominantly exhibits the *gt* conformation.

Material and Methods

Plant materials

Uniformly ^{13}C-labeled bacterial cellulose was produced from *Acetobacter xylinus* sub sp. *sucrofermentans* (ATCC 700178) and was purified using a previously published procedure [45-47]. Briefly, cellulose pellicles were grown for 2 weeks at room temperature in liquid culture containing

uniformly ^{13}C-labeled D-glucose as the sole carbon source, then frozen and ground to a slurry. The bacterial debris was removed by washing in 1% NaOH, and cellulose was neutralized by washing with water until the pH reached ~7.

^{13}C-labeled *Brachypodium* and *Arabidopsis* primary cell walls were obtained from seedlings that were grown in the dark for 14 days in liquid culture containing ^{13}C-labeled glucose as the sole carbon source [17, 19, 20, 48]. Starch and intracellular proteins were removed, and the cell walls were never dried during processing. In addition to intact cell walls, we also studied a partially extracted *Arabidopsis* cell wall sample where pectins and hemicellulose were substantially removed using CDTA, Na_2CO_3, XEG, and Cel12A [48]. The final *Brachypodium* cell wall sample contained about 40% water while the *Arabidopsis* samples were 70-80% hydrated.

Solid-State NMR experiments

Magic-angle-spinning (MAS) solid-state NMR (SSNMR) spectra were measured on an 800 MHz (18.8 Tesla) Bruker spectrometer using a 3.2 mm MAS probe. Typical radiofrequency (rf) field strengths were 50 kHz for ^{13}C and 62.5 - 83 kHz for ^{1}H. All ^{13}C chemical shifts were referenced to the adamantane CH_2 peak at 38.48 ppm on the TMS scale.

2D CHHC correlation spectra were measured under 10.5 kHz MAS at 298 K using the pulse sequence shown in **Fig. 2a** [41, 42]. The experiment contains three ^{1}H-^{13}C cross-polarization (CP) steps. The first CP step establishes ^{13}C transverse magnetization, which evolves during the t_1 period to encode ^{13}C chemical shifts. The ^{13}C magnetization is then transferred back through a second CP step to its directly bonded ^{1}H. The ^{1}H magnetization is allowed to diffuse to other protons during a mixing time t_m, after which a third ^{1}H-^{13}C CP step detects the transferred ^{1}H magnetization through its directly bonded ^{13}C. For the PCW samples, we used a contact time of 45-100 μs for the first ^{1}H-^{13}C CP step to preferentially select the rigid cellulose signals while minimizing the signals of mobile matrix polysaccharides. 70-80% of cellulose signals were retained at these contact times compared to the equilibrium intensity measured using 1 ms CP. The second and third CP steps used even shorter contact times of 45 μs, to ensure that the observed C4-C6 cross peaks mainly reflect polarization transfer between the H4 and H6 protons of the same glucosyl residue while minimizing polarization transfer from and to other neighboring protons. Due to the three CP steps, the sensitivity of the CHHC experiment is much lower than that of ^{1}H-driven ^{13}C spin diffusion experiments. However, uniform ^{13}C labeling of the plant cell walls and bacterial cellulose makes this experiment still feasible. ^{1}H spin diffusion mixing times ranged from 0 to 350 μs (**Table S1**). At the longest mixing times, multi-bond intramolecular ^{13}C-^{13}C cross peaks such as C1-C4 and intermolecular cellulose-matrix polysaccharides cross peaks are observed.

The 2D CHHC spectral widths were 248 ppm (50 kHz) for the ω_2 dimension and 118-134 ppm (24-27 kHz) for the ω_1 dimension. A total of 220 t_1 increments were measured, with a maximum ^{13}C evolution time of 4.6 ms. The number of scans per t_1 slice ranges from 128 to 224 for the two intact PCW samples, and were 32 or 64 for bacterial cellulose and the digested *Arabidopsis* cell wall sample. Recycle delays of 1.7 – 2.5 s were used, and the experimental time for each 2D spectrum ranged from 4 to 28 hours (**Table S1**). The 2D spectra were processed using a Gaussian window function with GB = 0.05 and LB = -10 in the TOPSPIN software, and were plotted using lev0 = 4, toplev = 40-50, and 12–16 contour levels.

We conducted ^{13}C-detected water to polysaccharide ^{1}H spin diffusion experiments on *Brachypodium* and digested *Arabidopsis* samples to investigate the locations of different cellulose polymorphs from the microfibril surface based on their water accessibilities. The experiments were

conducted under 10 kHz MAS at 293 K. The polysaccharide ^{1}H magnetization was first suppressed using a ^{1}H T_2 filter of 1.2 ms for the digested *Arabidopsis* cell wall and 3.2 ms for the *Brachypodium* cell wall. Subsequently a ^{1}H spin diffusion mixing period of 0–121 ms was applied to allow water-polysaccharide ^{1}H spin diffusion, after which a ^{1}H-^{13}C CP contact time of 100 μs was used to detect the ^{13}C signals.

Quantification of the CHHC buildup curves

The CHHC buildup curves were constructed by dividing the measured integrated area of each cross peak by the total area of the corresponding row cross section in the 2D spectrum. This ratio was then normalized by the maximum intensity ratio found among all the mixing times in a series. We found that the integrated area of each cross section was constant over the mixing time range of 0 to 350 μs, indicating that there is no significant ^{1}H T_1 relaxation during these mixing times. Therefore, the normalized CHHC buildup intensities can also be obtained simply by comparing the cross peak intensities in different spectra, without normalization by the area of the cross section. This allowed us to propagate the uncertainty (ε) in the buildup intensities from the signal-to-noise ratios (sino) of the cross peaks according to $\varepsilon = S/S_0 \cdot \left[\left(sino_S \right)^{-2} + \left(sino_{S_0} \right)^{-2} \right]^{1/2}$, where S denotes the cross peak height at a particular mixing time while S_0 denotes the maximum cross peak height in the series. The signal-to-noise ratios were reported in TopSpin with 2σ, which corresponds to the 95% confidence interval.

Simulations of the CHHC buildup curves

To extract the χ torsion angle from the CHHC buildup curves, we simulated ^{1}H spin exchange using the SpinEvolution software [49], with a representative input code given in **Table S2**. The coordinates of 12 protons were extracted from five cellulose structural models, four of which from neutron diffraction data [1, 2] and one from DFT calculations (**Tables S3, S4**) [38]. The ^{1}H chemical shift anisotropy (CSA) was neglected in the simulation since CSAs of less than 40 ppm have little effect on spin exchange [42]. Each simulation generates time-dependent polarization transfer from the initially polarized H4 proton to the two H6 protons in the presence of nine other protons. The simulated buildup curves for different χ angles were compared to the experimental data to determine the hydroxymethyl conformation.

Results

Experimental 2D CHHC spectra of bacterial cellulose and plant cell wall cellulose

The hydroxymethyl conformation that defines the orientation of the O6H group relative to the glucopyranose ring can be described by the two distances between the H4 proton and the two H6 protons (**Fig. 1c)**. The *tg* conformation found in crystalline Iα and Iβ cellulose has a long H4-H6 distance (3.5-3.7 Å) and a very short distance of 2.3-2.5 Å. The *gt* conformation, which is found in methanol-crystallized α-D-glucose [50], has two intermediate distances of 2.5-2.7 Å and 2.8-3.1 Å. The *gg* conformation, such as that found in α-D-glucose • H_2O, has two relatively long H4-H6 distances of 3.1-3.3 Å and 3.6-3.7 Å. The different distance distributions suggest that ^{1}H-^{1}H spin exchange should be able to distinguish the different hydroxymethyl conformations. Since dipolar coupling is inversely proportional to distances to the third power, a very short distance is expected to significantly speed up spin diffusion [51]. Therefore, we expect the *tg* conformation to manifest the fastest H4-H6 polarization transfer while the *gg* conformation should have the slowest H4-H6 polarization transfer.

Fig. 2b shows a representative 2D CHHC spectrum, measured on the *Brachypodium* cell wall using a ^{1}H-^{1}H mixing time of 250 μs. Three intramolecular C6-C4 cross peaks are resolved. The (62.7, 84.7) ppm C6-C4 cross peak was previously assigned to surface cellulose type-*f*, the (61.6, 83.7) ppm

cross peak was assigned to surface cellulose type-*g*, while the (65.0, 89.0) ppm peak can be assigned to interior crystalline cellulose [29]. This interior cellulose cross peak contains three sub-types, *a*, *b* and *c*, which have been resolved in previous 2D ^{1}H-driven ^{13}C spin diffusion spectra but cannot be resolved in the CHHC spectra due to their lower sensitivity. Therefore, in this study we focus on the dominant interior cellulose structure and two surface cellulose structures. At ^{1}H-^{1}H mixing times of 250 μs and longer, intermolecular cross peaks are also observed, such as between surface and interior cellulose, and between cellulose and arabinose (Ara) in matrix polysaccharides.

In total, we measured the 2D CHHC spectra of four samples, including bacterial cellulose and three PCWs (**Table S1**). **Fig. 3a** compares the C6-C4 regions of these samples. The bacterial cellulose spectra are dominated by two resolved crystalline cellulose C6-C4 peaks: the (65.0, 89.1) ppm C6-C4 cross peak can be assigned to residue A of Iα cellulose while the (65.0, 90.0) ppm cross peak can be assigned to residue A' based on the literature [16]. In addition, a weak surface cellulose C6-C4 cross peak is observed at (61.6, 83.7) ppm. These chemical shifts match those of type-*g* surface cellulose in PCWs. Among the three PCW samples, the *Brachypodium* cell wall spectrum shows two partially resolved surface cellulose cross peaks, with the type-*f* (62.4, 84.5) ppm cross peak having higher intensity than the type-*g* (61.6, 83.7) ppm cross peak. The two *Arabidopsis* cell wall samples do not have sufficient sensitivity for the type-*g* cross peak, thus we confine our analysis to the type-*f* cross peak. **Fig. 3b** shows representative CHHC spectra of the *Brachypodium* cell wall as a function of the ^{1}H-^{1}H mixing time. It can be seen that the interior cellulose cross peak reached maximum intensity at shorter mixing times (~125 μs) than the surface cellulose peaks, indicating qualitatively that the H4-H6 distances are shorter in interior cellulose than in surface cellulose.

To quantify the H6-H4 cross peak intensities as a function of the ^{1}H-^{1}H mixing time, we divided the integrated intensities of the cross peaks in the C6 and C4 ω_1 cross sections by the integrated areas of each cross section to take into account T_1 relaxation effects. These intensity ratios plateaued by the longest mixing time of 350 μs for all four samples, which allowed us to further normalize the intensity ratios by the maximum ratio of each sample. At vanishing mixing time, we observed non-zero cross peak intensities, which can be attributed to the presence of ^{1}H-^{1}H spin diffusion during ^{1}H-^{13}C CP. Since this intensity is constant for all mixing times [42], we subtracted the normalized intensities of the 0.1 μs 2D spectrum from all data points, giving a corrected initial intensity of zero. The C6-C4 and C4-C6 cross peaks were averaged to give the H4-H6 intensity buildup, except for bacterial surface cellulose, whose C6-C4 cross peak intensity was too weak to be reliable, therefore only the C4-C6 peak intensity was used in the analysis. The CHHC data of all four samples (**Fig. 4**) show faster buildup for the crystalline cellulose C6-C4 cross peak than the surface cellulose C6-C4 peak, with the largest difference found in the *Brachypodium* cell wall, indicating that the hydroxymethyl conformation differs between interior and surface cellulose.

Simulation of the CHHC buildup curves for χ-angle determination

To extract the χ torsion angle, we simulated H4-H6 polarization transfer curves using SpinEvolution [49]. A total of 12 protons were included in the simulation to account for not only polarization exchange between H4 and H6 protons but also the influence of neighboring protons. Twelve protons are the maximum number (*n*) that can be inputted in the SpinEvolution program since the dimension of the density matrix scales with 2^n. Due to the inclusion of nine additional protons, we considered five structural models of cellulose (**Fig. 5**), including four neutron diffraction structures of cellulose Iα residue 1 and residue 2 (CSD Entry: JINROO06) [2], cellulose Iβ center chain and origin chain (CSD Entry: JINROO02) [1], and a DFT energy-minimized cellulose Iβ structure [38]. The nine additional protons include six (H1, H2, H3, OH3, H5, OH6) in the same glucosyl residue that contains the hydroxymethyl group of interest and three from the adjacent glucopyranose ring. For the two

cellulose Iα structures and the cellulose Iβ origin chain, we used H1', OH2', and H5', while for cellulose Iβ center chain, we used H2' in place of H5' due to its shorter distance to H4 (**Tables S3, S4**). All calculated CHHC buildup curves reach an equilibrium value of 0.08 by ~3 ms, as expected for a 12-spin system (**Fig. S1**). Since the maximum experimental ^{1}H-^{1}H mixing time was 350 μs, we normalized all simulated spin diffusion curves by 0.08.

Fig. 5 shows the simulated H4-H6 buildup curves for the three canonical hydroxymethyl conformations using the five starting cellulose structures. For all structures, the *tg* conformation shows the fastest buildup while the *gg* conformation has the slowest buildup, consistent with the prediction (**Fig. 1**). The *gt* conformation has intermediate buildup, and the degree of similarity of the *gt* buildup curve to the *tg* curve, as well as the shape of all buildup curves, depend on the cellulose structure. For the cellulose Iα residue 2 and cellulose Iβ center chain , the difference in buildup rates among the three hydroxymethyl conformations is the largest (**Fig. 5a, d**). For cellulose Iα residue 1 and cellulose Iβ origin chain (**Fig. 5b, c**), the *tg* and *gt* buildup curves are less distinct. For the DFT energy-minimized Iβ structure (**Fig. 5e**), the *tg* and *gt* buildup curves show a significant intensity overshoot at 100-200 μs. Moreover the *gg* buildup rises rapidly and exceeds the *tg* and *gt* intensities by 350 μs. The different buildup trajectories indicate that the H4-H6 polarization transfer rate is sensitive to the locations of the surrounding protons, and multiple cellulose structures should be considered when fitting the experimental H4-H6 buildup curves. **Fig. 5f** compares the simulated *tg* (χ = 166°) buildup curves for the five cellulose structural models, and superimposes them with the measured CHHC intensities of crystalline bacterial cellulose. The simulated curves based on the two cellulose Iα structures agree very well with the experimental data. Since crystalline bacterial cellulose is known to adopt the Iα allomorph, the good agreement validates the CHHC technique. The best-fit χ angle is 166°–180° within experimental uncertainty (**Fig. 6a**). The cellulose Iβ origin-chain structure also gives reasonable agreement with the experimental data, but the center-chain simulation does not.

Since the simulated H4-H6 buildup curve for the Iα *tg* conformation shows good agreement with the measured bacterial cellulose data, we proceeded to measure the hydroxymethyl conformation of structurally unknown cellulose. **Fig. 6** shows the experimental CHHC data of the PCW cellulose and simulations using the cellulose Iβ structures. Alternative simulations based on the cellulose Iα structures are given in **Fig. S2**. For bacterial surface cellulose, the measured buildup curve is best fit to the *gt* conformation and does not agree with the *gg* conformation (**Fig. 6a**). Importantly, this result differs from the chemical shift prediction [52] that the 61.5-ppm C6 chemical shift should be correlated with the *gg* conformation. Therefore, the small-molecule derived correlation does not adequately reflect the local conformation of glucan chains on the surface of bacterial cellulose. This discrepancy may result from differences in the hydrogen-bonding network and water interactions between small molecules and cellulose microfibrils, and underscores the importance of direct distance and torsion angle measurements for structure determination.

In the three PCW samples, all interior crystalline cellulose C6-C4 cross peaks show rapid intensity buildup that is best fit to the *tg* conformation in the single χ-angle fits (**Fig. 6b-d**). The origin-chain *tg* simulation yields excellent agreement with the data, while the center-chain *tg* simulation exhibits an intensity overshoot at short mixing time, which is absent in the experimental data. To determine whether the interior cellulose CHHC data also agree with a mixture of two significant populations of hydroxymethyl conformations, we simulated the *Brachypodium* CHHC data using a 50 : 50 mixture of *tg* and *gg* conformations as well as a mixture of *tg* and *gt* conformations (**Fig. 6b**). The *tg*/*gg* mixture can be readily ruled out because the simulated buildup is much slower than the measured buildup rate (**Fig. 6b**). In comparison, the *tg*/*gt* mixture fit cannot be ruled out, although it shows worse agreement with the experimental data compared to the single-component *tg* fit. Quantum chemical

calculations indicate that the *gt* conformation has distinct C4 and C6 chemical shifts from the *tg* conformation [29, 38], thus it is unlikely for an equimolar mixture of *tg* and *gt* conformations to manifest the same (C4, C6) chemical shifts of (89, 65) ppm [38]. Taken together, these results indicate that the *tg* conformation most likely dominates in the interior cellulose of *Brachypodium* and *Arabidopsis* cell walls (**Fig. 6c, d**).

Compared to interior cellulose, the surface cellulose CHHC buildup rates are slower in all three PCWs. In single-conformation fits, both type-*f* and type-*g* surface cellulose data in the *Brachypodium* cell wall are best fit by the *gt* conformation, with a χ angle of 60-80˚ (**Fig. 6b**). If two conformations are considered, then a *gt*/*gg* mixture with 20% *gg* component can also fit the *Brachypodium* surface cellulose data (**Fig. S3b**). This result is consistent with the DFT calculation that the *gt* conformation is more dominant than the *gg* conformation in surface cellulose [38].

The surface cellulose in the two *Arabidopsis* cell walls (**Fig. 6c, d**) show trends similar to that of the *Brachypodium* surface cellulose. In single-conformation fits, the data are well fit by the *gt* but not the *gg* conformation. When a two-component model is considered, a *gt*/*gg* mixture is also possible (**Fig. S3c**). The upper limit of the *gg* percentage is 50% for the intact *Arabidopsis* cell wall but decreases to 20% for the digested *Arabidopsis* cell wall (**Fig. S3d**). These two cell walls differ substantially in the amounts of the matrix polysaccharides: the intact cell wall contains a significant concentration of xyloglucan while the digested cell wall does not [20, 48]. Based on the resolved xylose C4-C5 cross peak intensity, we estimate that the xyloglucan concentration is about 80% of the concentration of surface cellulose in the intact cell wall [17, 21]. Since the C4 and C6 chemical shifts of the glucose backbone in xyloglucan overlap with the C4 and C6 chemical shifts of surface cellulose, if the xyloglucan backbone hydroxymethyl adopts the *gg* conformation, it would bias the measured CHHC curve at the overlapped C4-C6 cross peak toward a higher percentage of the *gg* conformation. Alternatively, xyloglucan may stabilize the *gg* conformation of surface cellulose. We also simulated the CHHC buildup curves using the DFT-optimized Iβ structure (**Fig. S4**). The simulated *tg* curve significantly exceeds the measured interior cellulose intensities at short mixing times, suggesting that this structure does not represent the PCW interior cellulose well.

Taken together, these CHHC data indicate that interior cellulose in both grass and dicot primary cell walls predominantly adopts the *tg* conformation, without significant contributions of *gt* and *gg* conformations. Surface cellulose hydroxymethyl groups mainly exhibit the *gt* conformation, which precludes intra-chain O2-H…O6 hydrogen bonds and opens the possibility of intra-chain O3-H…O6 hydrogen bonds. A significant population of the *tg* conformation in these surface chains can be ruled out, but the *gg* conformation cannot be ruled out by the CHHC data alone, and may account for up to 20-50% of all surface chains.

Water accessibility of surface cellulose chains

To quantify the relative amounts of type-*f* and type-*g* surface cellulose, we deconvoluted the C6 region of the ^{13}C CP-MAS spectra of *Brachypodium* and digested *Arabidopsis* cell walls (**Fig. 7a**) using the software DMfit [53]. These spectra were measured using a short CP contact time of 100 μs to minimize the intensities of the mobile matrix polysaccharides. Compared to the quantitative ^{13}C spectra, the matrix polysaccharide intensities are reduced by 2-fold. We do not consider the intact *Arabidopsis* cell wall in this deconvolution since this intact cell wall contains a significant amount of arabinan, galactan and xylose signals even with short CP, which these signals overlap with the surface cellulose C6 signals at ~62 ppm. Based on the C6 peak deconvolution, the type-*f* to type-*g* ratio is 2.4 for the digested *Arabidopsis* cell wall and 1.4 for *Brachypodium* cell wall.

We next conducted ^{13}C-detected water-to-polysaccharide ^{1}H spin diffusion experiments to determine the water accessibilities of the different types of cellulose in these PCWs [48, 54] (**Fig. 7b**). In the digested *Arabidopsis* cell wall, both type-*f* and type-*g* surface cellulose show much faster polarization transfer from water than interior cellulose (**Fig. 7c**), consistent with previous results [48]. In addition, the 61-ppm type-*g* peak exhibits higher initial intensities at short mixing times than the 62-ppm type-*f* peak, indicating that type-*g* surface cellulose is slightly more accessible to water than types-*f* surface cellulose. This difference can reflect different interactions of the two surface cellulose chains with water and matrix polysaccharides and/or different hydroxymethyl conformations. Since the difference in the buildup rates is relatively small, it is unlikely that type-*f* surface cellulose exists between cellulose microfibrils within a bundle because these regions are expected to be water-inaccessible. The *Brachypodium* cell wall exhibits the same trend, with slightly higher water accessibility for type-*g* than type-*f* surface cellulose. In this cell wall, the interior cellulose C6 peaks at 65-66 ppm have partial overlap with the 63.8-ppm C5 peak of xylan, which is likely well exposed to water, thus the water polarization transfer intensities for this peak do not accurately reflect the interior cellulose water accessibility and are not shown here.

Discussion

The data shown above represent the first direct experimental measurement of an important torsion angle in cellulose, without relying on small-molecule-derived chemical shift correlations with crystal structures. The CHHC technique has been employed in protein structure determination by NMR [55-57] because of its ability to probe relatively long distances compared to ^{13}C-^{13}C dipolar couplings. Using the structurally known bacterial cellulose as a model compound, we verified that the CHHC buildup curve yields a χ angle of 166-180˚, in excellent agreement with the known *tg* conformation of bacterial cellulose. Comparison of simulated CHHC curves using the proton coordinates from different cellulose structures show that the Iα-based simulations agree well with the data, while the Iβ-based simulations give only partial agreement with the data, with the origin-chain structure fitting the data well while the center-chain structure does not. Thus, the CHHC data is partly sensitive to the proton positions around the H6 and H4 protons.

In both grass and dicot cell walls, the interior cellulose CHHC buildup curves are best fit to the *tg* conformation, while surface cellulose CHHC buildup curves are best fit to the *gt* conformation. However, since the CHHC curves for the three canonical hydroxymethyl conformations have similar shapes and differ only in buildup rates, in principle multiple χ torsion angles can fit the same data. In addressing this possibility, we consider only a bimodal distribution of χ angles, and focus on distributions with significant population of each component, since the NMR spectral sensitivity precludes us from considering a minor component that represents less than ~20% of the total intensity. Within these constraints, the measured interior cellulose CHHC data rule out a significant population of the *gg* conformation mixed with *tg*, but do not rule out a significant population of the *gt* conformation mixed with *tg*. This reflects the similar H4-H6 polarization transfer rates between the *tg* and *gt* conformations. Similarly, the surface-cellulose CHHC data can be fit with a *gt*/*gg* mixture, with the *gg* conformation accounting for 20-50% of the total population. However, the cell wall that can have up to 50% potential *gg* conformation is the intact *Arabidopsis* cell wall (**Fig. S3c**), which has the highest amount of xyloglucan in all cell walls studied. The glucose C6-C4 cross peak of xyloglucan overlaps with the C6-C4 cross peak of surface cellulose. Thus, the intact *Arabidopsis* cell wall CHHC data at this cross peak most likely reflect a mixture of the xyloglucan and surface cellulose hydroxymethyl conformations. Indeed, in the digested *Arabidopsis* cell wall and the *Brachypodium* cell wall, where little xyloglucan is present, the surface cellulose CHHC data indicate a predominantly *gt* conformation.

What is the likelihood that two distinct hydroxymethyl conformations occur with the same C6 and C4 chemical shifts and hence are unresolved in the 2D spectra? Here, quantum chemical calculations provide important additional constraints. Recent DFT calculations showed that the *tg*, *gt* and *gg* conformations have well resolved C4 and C6 chemical shifts [21, 38]: the *gt* and *gg* conformations differ in C4 chemical shifts by ~5 ppm while the *gt* and *tg* conformations differ in C6 chemical shifts by ~3 ppm. The measured C4 and C6 linewidths in the 2D CHHC spectra (**Fig. 3**) are 1.0 ppm for crystalline bacterial cellulose, 2.0 ppm for interior cellulose, and 1.4-1.6 ppm for surface cellulose in the three plant cell walls. Deconvolution of the C6 region of 1D ^{13}C spectra (**Fig. 7a**) indicates even narrower linewidths of 0.7–1.3 ppm for surface cellulose. These narrow linewidths compared to the chemical shift differences indicate that it is unlikely for the *tg*, *gt* and *gg* conformations to have overlapping C6-C4 cross peaks. Therefore, the CHHC data and ^{13}C chemical shifts together support a predominantly *tg* conformation for interior cellulose and rule out a significant percentage of *gt* or *gg* conformations co-mixed with the *tg* conformation. For surface cellulose, the combined CHHC and chemical shift evidence supports a predominantly *gt* conformation, while *gt*/*gg* and *gt*/*tg* mixtures with similar percentages of each conformation can be ruled out. However, a minor population of the *gg* conformation mixed with the main *gt* conformation is possible within experimental uncertainty.

These CHHC results provide new insights into cellulose structures in plant cell walls. The *tg* conformation of interior cellulose indicates that intra-chain O2-H…O6 hydrogen bonds are retained, similar to those in crystalline cellulose, while the *gt* conformation for surface chains implies alternative hydrogen bonds, which may occur between surface chains and water molecules or between surface cellulose and matrix polysaccharides. While the CHHC data alone does not exclude the presence of the *gg* conformation in surface cellulose, this conformation occurs at a low percentage, implying that the exocyclic group does not alternate in a 1 : 1 fashion between an inward-facing *gt* conformation and an outward-facing *gg* conformation or vice versa [31].

Combining these H4-H6 polarization transfer results, the C4 and C6 chemical shifts, the relative amounts of different cellulose types, and the water accessibility of the different cellulose types, we propose the following model of the hydroxymethyl conformations in PCW cellulose microfibrils (**Fig. 8**). Assuming a 24-chain hexagonal arrangement of the microfibril [31, 38], 10 of the 24 chains would reside in the microfibril interior while 14 chains would lie on the surface. The interior ten cellulose chains exhibit the *tg* conformation, without significant presence of either *gt* or *gg* conformations. Six of the surface chains would lie on the hydrophobic surface while eight chains would be located on the hydrophilic surface. Since the NMR spectra show higher type-*f* intensities than type-*g* intensities, we assign the type-*f* chains with the (84.5, 62.4) ppm chemical shifts to the hydrophilic surface and the type-*g* chains with the (83.7, 61.6) ppm chemical shifts to the hydrophobic surface (**Fig. 8a**). This assignment is consistent with the observation that type-*g* cellulose has higher water accessibility than type-*f* cellulose. The assignment of type-*f* and type-*g* surface chains would still apply for a hexagonal 18-chain model [58-60], in which 4 chains would lie on the hydrophobic face and 8 chains would lie on hydrophilic surface. Our data indicate that cellulose chains on both surfaces predominantly adopt the *gt* conformation, but a small percentage of the *gg* conformation may also be present. Finally, the relative intensities of the (85, 62) ppm cross peak and the (89, 65) ppm cross peak in the CHHC spectra indicate that it is not possible to assign the interior-facing exocyclic groups of surface chains to the *tg* conformation and the exterior-facing exocyclic group to the *gt* conformation. In the 24-chain microfibril, this would result in a *tg* : *gt* intensity ratio of 1.4 : 1, with higher intensity for the (89, 65) ppm peak, which contradicts the experimental result that the (85, 62) ppm peak has 30% higher intensity (**Fig. 7**) [21]. In other words, the experimental data indicate a higher population of *gt* than *tg* hydroxymethyl groups in PCW cellulose. When an 18-chain microfibril model is considered,

the same conclusion applies, where assigning inward-facing exocyclic groups of surface chains to the *tg* conformation would result in intensity ratios that are inconsistent with the NMR spectra.

Conclusion

We have measured the hydroxymethyl conformation of cellulose in the primary cell walls of two grass and dicot model plants using a 2D ^{13}C-resolved ^{1}H-^{1}H polarization transfer NMR experiment. The H4-H6 CHHC buildup curves show that the water-inaccessible interior cellulose chains in PCWs predominantly adopt the *tg* conformation while the water-accessible surface cellulose chains predominantly adopt the *gt* conformation. Joint consideration of both the CHHC data and ^{13}C chemical shifts indicates that the interior chains do not have a significant population of *gt* or *gg* conformation mixed with the *tg* conformation, while the surface chains cannot have a significant population of the *tg* and *gg* conformations mixed with the *gt* conformation. But a small percentage of the surface chains may adopt the *gg* conformation. These conclusions hold whether the cellulose Iα or Iβ structures are used in the simulations. The fact that the CHHC buildup curve of type-*g* cellulose with a C6 chemical shift of 61.6 ppm corresponds to the *gt* instead of *gg* conformation indicates that empirical ^{13}C chemical shifts of oligosaccharides do not always apply to PCW cellulose, which interact extensively with water and matrix polysaccharides. To further elucidate whether a small population of a second hydroxymethyl conformation exists on the surface of cellulose microfibrils, the sensitivity of the CHHC technique needs to be significantly enhanced, for example by using dynamic nuclear polarization [61].

Supporting Information Available:

Additional figures and tables include:
Simulated CHHC curves for the *tg* conformation based on different cellulose structures (Figure S1)
Simulations of measured CHHC curves using the cellulose Iα structure (Figure S2).
Simulations of measured CHHC curves using two-conformation models (Figure S3).
Simulations of measured CHHC curves using DFT energy-minimized cellulose structure (Figure S4).
2D CHHC experimental conditions (Table S1).
Input code of SpinEvolution program for CHHC simulations (Table S2).
Atomic coordinates and distances of protons in the CHHC simulations (Tables S3, S4).

Acknowledgements

This research was supported by the Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award # DE-SC0001090. The authors thank Paul White for some of the initial measurements, Dr. Matthias Roos for help with SpinEvolution simulations, and Dr. Daniel Oehme for insightful discussions and for providing the coordinates of a 24-chain Iα-like cellulose structural model [31].

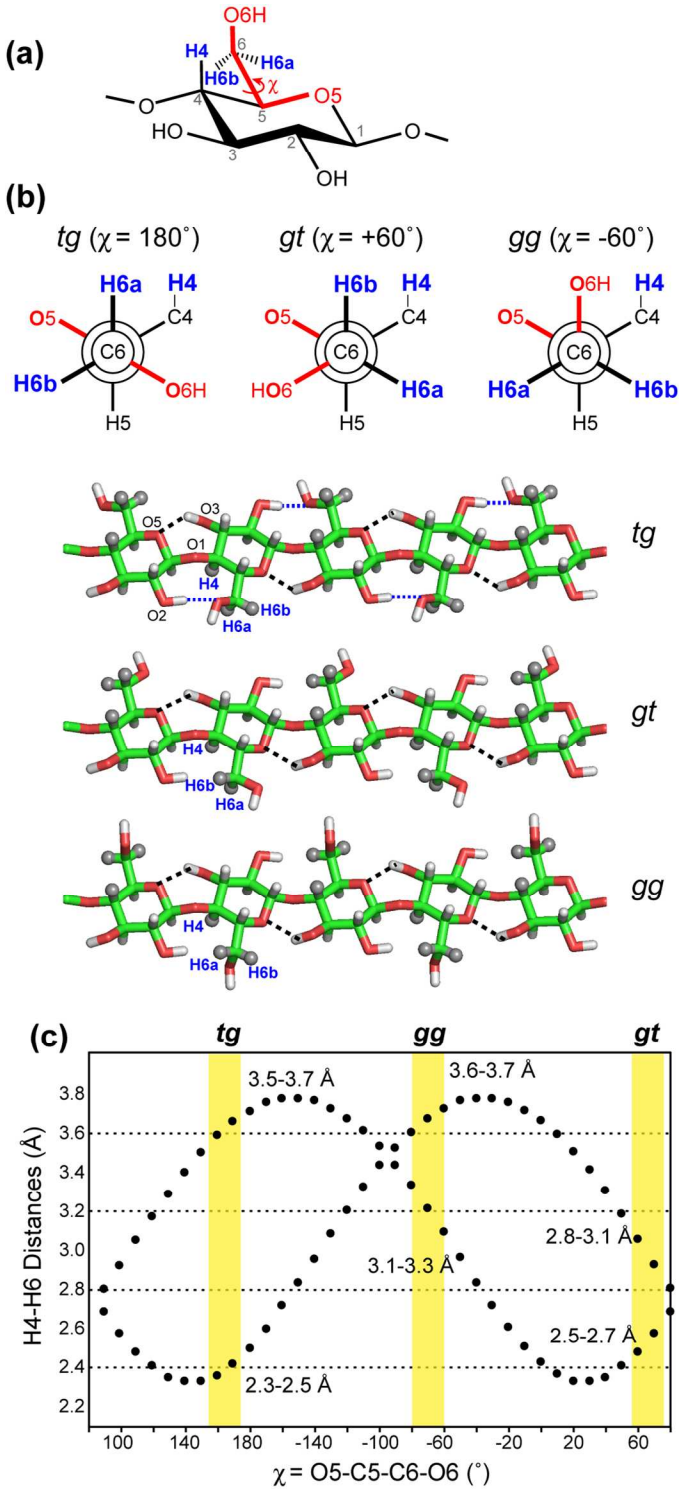


Figure 1. Definition of the hydroxymethyl conformation and its relation to H4-H6 distances. (a) The χ torsion angle is defined as the O5-C5-C6-O6 angle. (b) Newman projection and side view of *tg*, *gt* and *gg* conformations in cellulose Iβ origin chain. The *tg* conformation establishes intra-chain O2-H…O6 hydrogen bonds (blue dashed lines), which are absent in *gt* and *gg* conformations. Intra-chain O3-H…O5 hydrogen bonds are shown as black dashed lines. (c) H4-H6 distances as a function of the χ angle. Yellow shaded areas indicate the range of torsion angles in the crystal structures of model compounds.

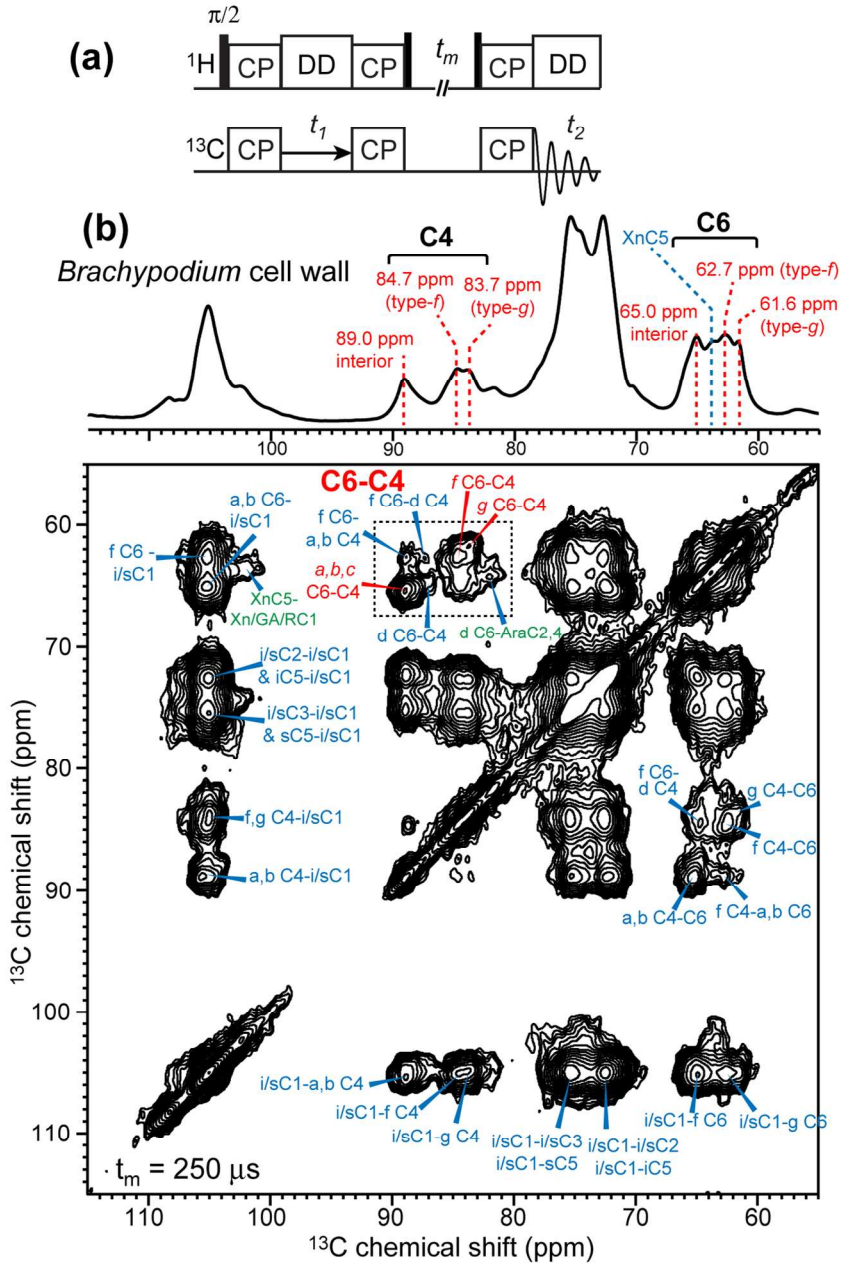


Figure 2. Representative 2D CHHC data. (a) CHHC pulse sequence, involving three ^{1}H-^{13}C CP steps and a ^{1}H-^{1}H mixing time t_m. (b) 2D CHHC spectrum of the *Brachypodium* cell wall, measured with a ^{1}H-^{1}H mixing time of 250 µs. A 1D ^{13}C spectrum measured with a CP contact time of 100 µs is shown at the top to indicate the resolved C4 and C6 peaks. Dashed rectangles indicate the C6-C4 region, which is amplified in Fig. 3.

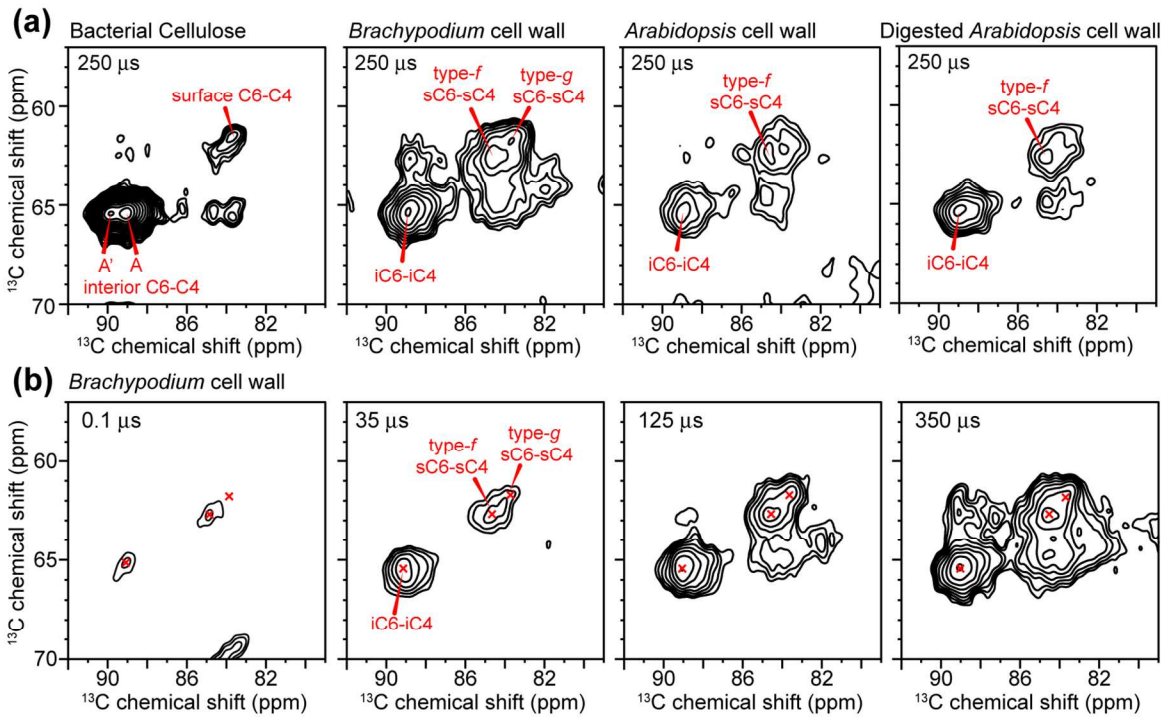


Figure 3. Cellulose C6-C4 region of representative 2D CHHC spectra. (a) Spectra of bacterial cellulose, *Brachypodium* cell wall, *Arabidopsis* cell wall and digested *Arabidopsis* cell wall at 250 µs mixing. Two crystalline cellulose C6-C4 cross peaks and one surface cellulose cross peak are observed in bacterial cellulose. The *Brachypodium* spectrum shows interior cellulose and surface *f* and *g* cellulose cross peaks, while the *Arabidopsis* spectrum shows interior cellulose and type-*f* surface cellulose cross peaks, but lacks the type-g cross peak due to low sensitivity. (b) C6-C4 region of the 2D CHHC spectra of the *Brachypodium* cell wall as a function of ^{1}H-^{1}H mixing times.

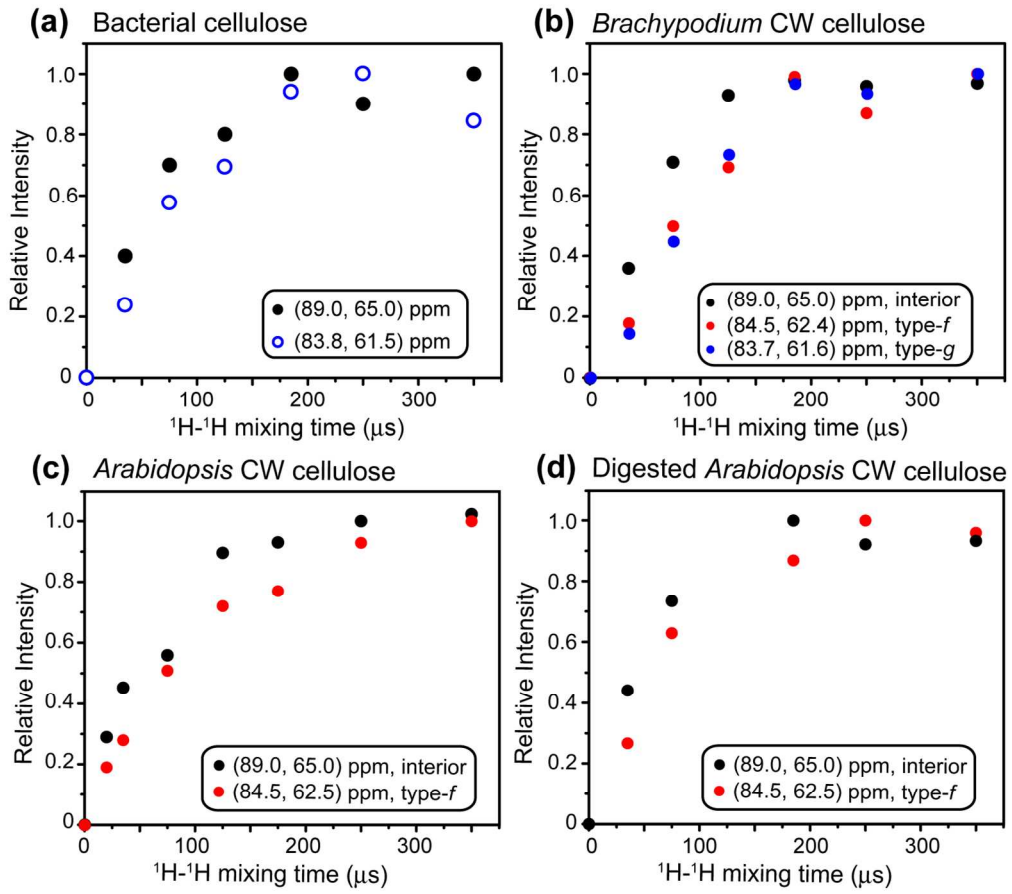


Figure 4. Measured CHHC buildup curves of (a) Bacterial cellulose, (b) *Brachypodium* cell wall cellulose, (c) *Arabidopsis* cell wall cellulose, and (d) Digested *Arabidopsis* cell wall cellulose. Except for bacterial surface cellulose (open circles), all data are obtained from the average intensities of the H4-H6 and H6-H4 cross peaks (filled circles).

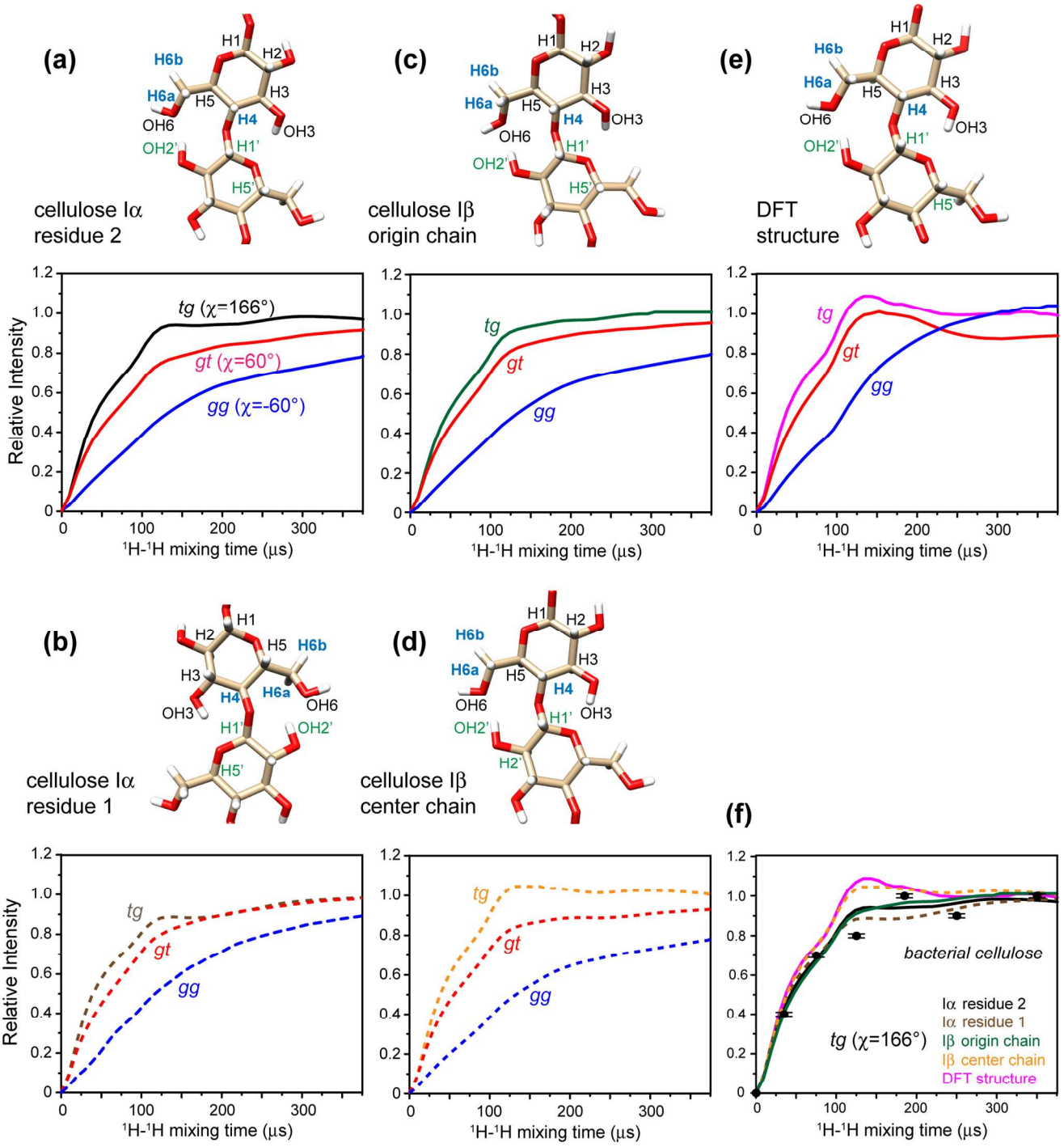


Figure 5. Simulated CHHC buildup curves for *tg*, *gt*, and *gg* hydroxymethyl conformations based on five cellulose structural models. (a) Cellulose Iα residue 2 structure. (b) Cellulose Iα residue 1 structure. (c) Cellulose Iβ origin chain structure. (d) Cellulose Iβ □□□□□□ □□□□□□ structure. (e) DFT-optimized cellulose Iβ □□□□□□□□□. (f) Overlay of the simulated *tg* curves for the five structural models, together with the measured crystalline bacterial cellulose data.

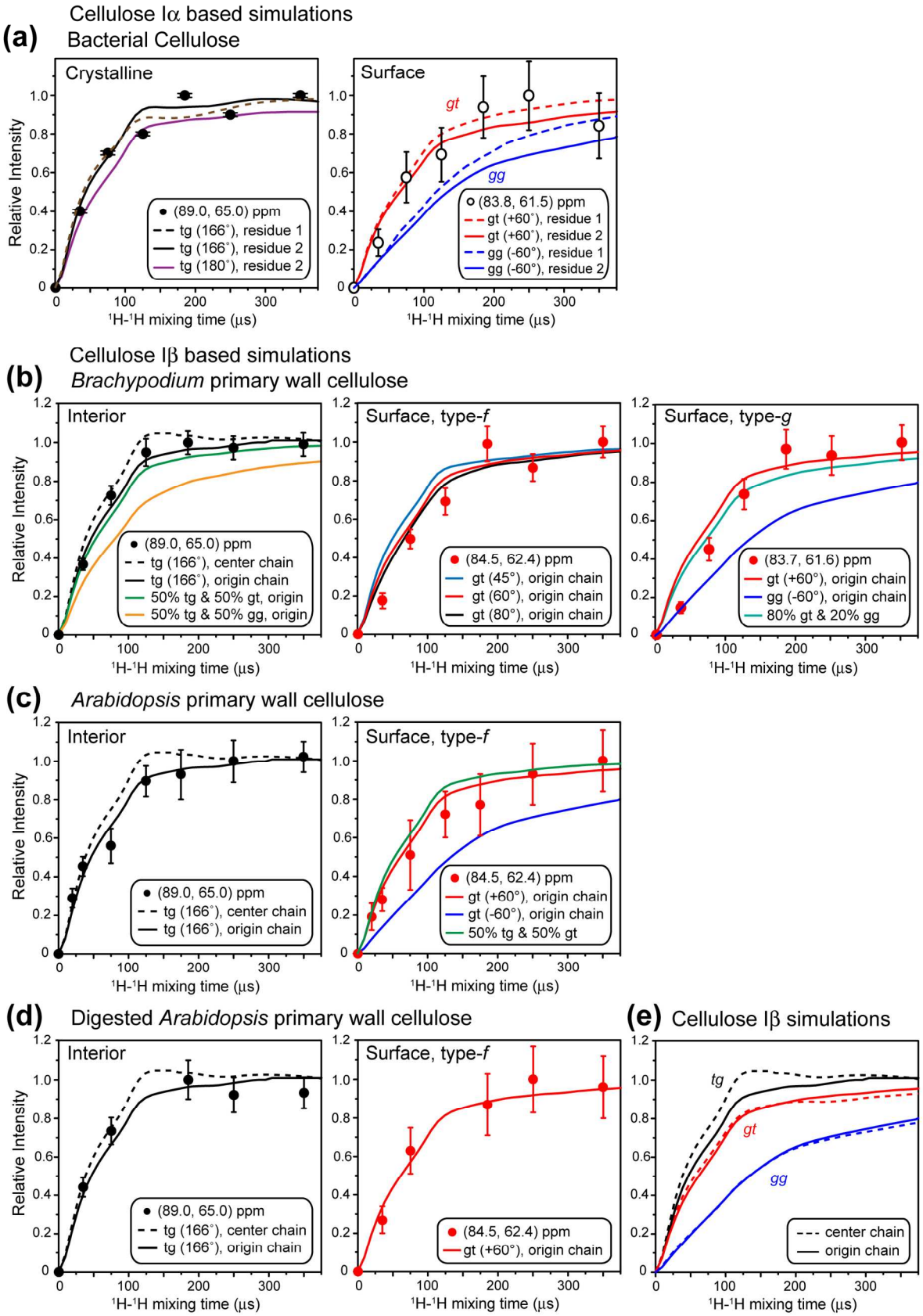


Figure 6. Experimental CHHC data of four cellulose samples, overlaid with simulated CHHC curves based on neutron diffraction structures of cellulose. (a) Bacterial cellulose data, simulated using the

cellulose Iα structures. (b) *Brachypodium* primary wall cellulose data. (c) *Arabidopsis* primary wall cellulose data. (d) Digested *Arabidopsis* primary wall cellulose data. The plant cell wall data are fit using cellulose Iβ structures. Both interior and surface cellulose data are shown. Simulated CHHC curves for different χ angles as well as mixtures of two χ angles are considered. (e) SpinEvolution simulation curves for the *tg*, *gt* and *gg* conformations based on the cellulose Iβ center chain and origin chain structures.

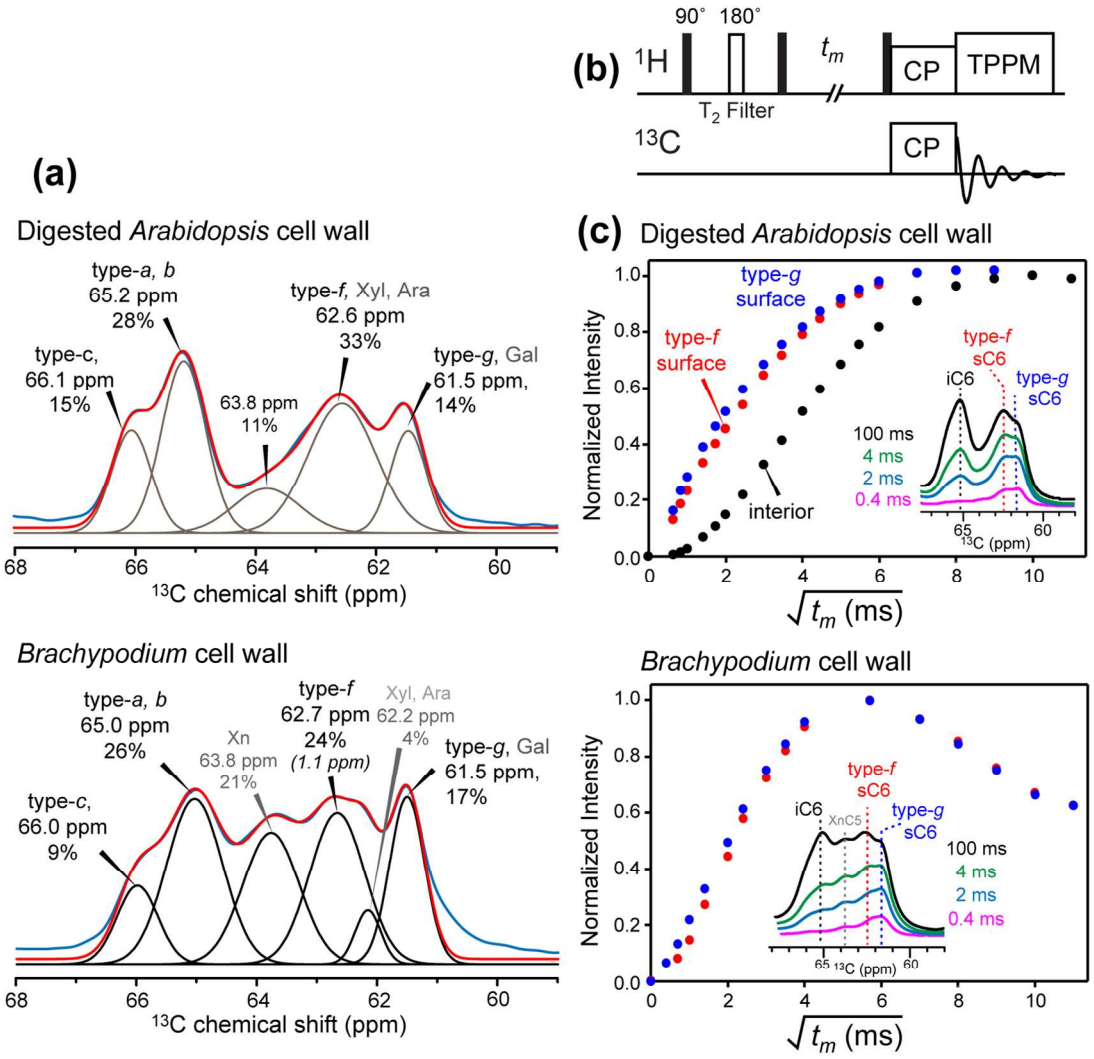


Figure 7. Water accessibilities of surface and interior cellulose in plant cell walls. (a) Deconvolution of the C6 region of the ^{13}C spectra of digested *Arabidopsis* and *Brachypodium* cell walls using Gaussian lineshapes. The spectra were measured using a 100 μs CP to minimize matrix polysaccharide signals. (b) Water-to-cellulose ^{1}H spin diffusion pulse sequence. (c) ^{1}H spin diffusion buildup curves of digested *Arabidopsis* and *Brachypodium* cell walls. The buildup of interior cellulose is slower than those of surface cellulose in digested *Arabidopsis* cell wall. Type-*g* surface cellulose has a slightly faster buildup than type-*f* surface cellulose in both cell walls.

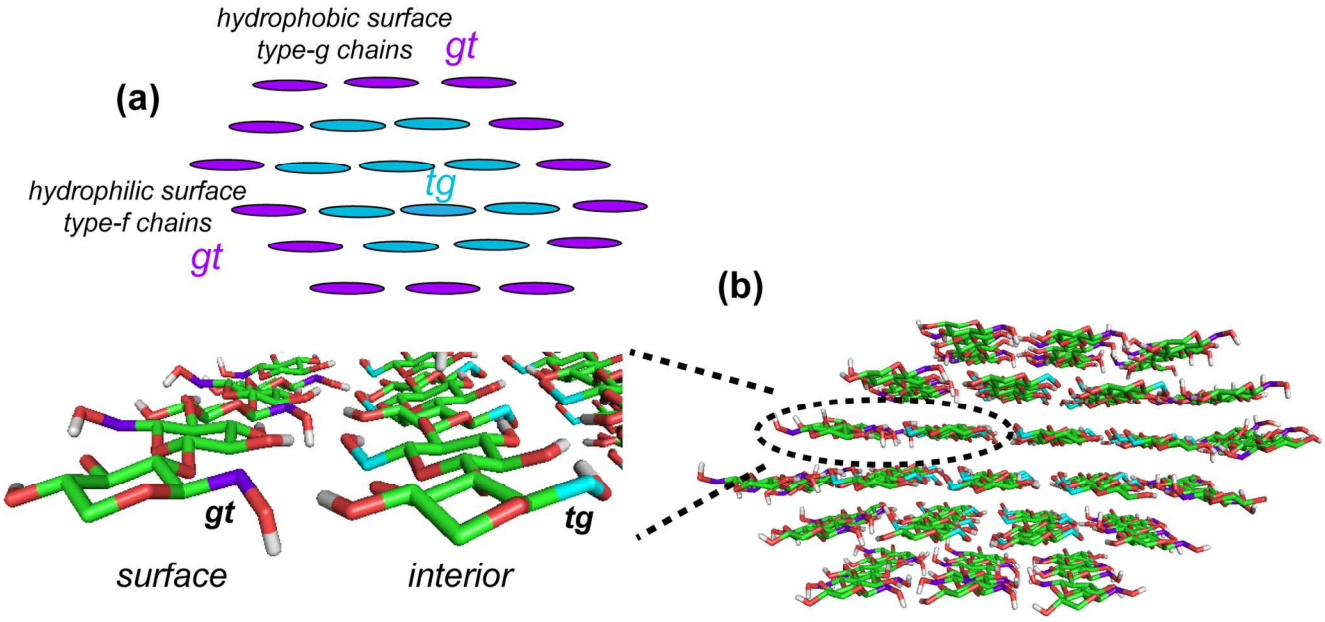


Figure 8. Hydroxymethyl conformations of PCW cellulose determined from the CHHC experiment. (a) Schematic model of a 24-chain cellulose microfibril. The 10 interior chains mainly adopt the *tg* conformation while the 14 surface cellulose chains mainly exhibit the *gt* conformation. (b) Model of a 24-chain hexagonal-shaped cellulose microfibril [31], to illustrate the different hydroxymethyl orientations. Each cellulose chain has the C6 hydroxymethyl group on both sides of the chain axis.

References

1. Nishiyama, Y.; Langan, P.; Chanzy, H., Crystal structure and hydrogen-bonding system in cellulose I_b from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **2002,** 124, (31), 9074-9082.

2. Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P., Crystal structure and hydrogen bonding system in cellulose Ia, from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **2003,** 125, (47), 14300-14306.

3. Jarvis, M., Chemistry: cellulose stacks up. *Nature* **2003,** 426, (6967), 611-612.

4. Pérez, S.; Samain, D., Structure and engineering of celluloses. *Adv. Carbohydr. Chem. Biochem.* **2010,** 64, 25-116.

5. Newman, R. H.; Hill, S. J.; Harris, P. J., Wide-angle x-ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiol.* **2013,** 163, (4), 1558-67.

6. Cosgrove, D. J., Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* **2014,** 22C, 122-131.

7. Fernandes, A. N.; Thomas, L. H.; Altaner, C. M.; Callow, P.; Forsyth, V. T.; Apperley, D. C.; Kennedy, C. J.; Jarvis, M. C., Nanostructure of cellulose microfibrils in spruce wood. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** 108, (47), E1195-E1203.

8. Wang, T.; Phyo, P.; Hong, M., Multidimensional solid-state NMR spectroscopy of plant cell walls. *Solid State Nucl. Magn. Reson.* **2016,** 78, 56-63.

9. Guerriero, G.; Fugelstad, J.; Bulone, V., What do we really know about cellulose biosynthesis in higher plants? *J. Integr. Plant Biol.* **2010,** 52, (2), 161-75.

10. Atalla, R. H.; Vanderhart, D. L., The role of solid state ^{13}C NMR spectroscopy in studies of the nature of native celluloses. *Solid State Nucl. Magn. Reson.* **1999,** 15, 1-19.

11. Earl, W. L.; Vanderhart, D. L., High-resolution, magic angle sample spinning C-13 NMR of solid cellulose-I. *Journal of the American Chemical Society* **1980,** 102, (9), 3251-3252.

12. Atalla, R. H.; Vanderhart, D. L., Native cellulose: a composite of two distinct crystalline forms. *Science* **1984,** 223, (4633), 283-285.

13. Vanderhart, D. L.; Atalla, R. H., Further ^{13}C NMR evidence for the coexistence of two crystalline forms in native celluloses. *ACS Symposium Series* **1987,** 340, 88-118.

14. Kono, H.; Yunoki, S.; Shikano, T.; Fujiwara, M.; Erata, T.; Takai, M., CP/MAS ^{13}C NMR study of cellulose and cellulose derivatives. 1. Complete assignment of the CP/MAS ^{13}C NMR spectrum of the native cellulose. *J. Am. Chem. Soc.* **2002,** 124, (25), 7506-7511.

15. Kono, H.; Erata, T.; Takai, M., Determination of the through-bond carbon-carbon and carbon-proton connectivities of the native celluloses in the solid state. *Macromolecules* **2003,** 36, (14), 5131-5138.

16. Kono, H.; Numata, Y., Structural investigation of cellulose I_a and I_b by 2D RFDR NMR spectroscopy: determination of sequence of magnetically inequivalent D-glucose units along cellulose chain. *Cellulose* **2006,** 13, 317-326.

17. Dick-Perez, M.; Zhang, Y. A.; Hayes, J.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and interactions of plant cell wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* **2011,** 50, (6), 989-1000.

18. Wang, T.; Zabotina, O.; Hong, M., Pectin-cellulose interactions in the *Arabidopsis* primary cell wall from two-dimensional magic-angle-spinning solid-state nuclear magnetic resonance. *Biochemistry* **2012,** 51, (49), 9846-9856.

19. Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Structure and dynamics of *Brachypodium* primary cell wall polysaccharides from two-dimensional ^{13}C solid-state nuclear magnetic resonance spectroscopy. *Biochemistry* **2014,** 53, (17), 2840-2854.

20. Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Cellulose-Pectin Spatial Contacts Are Inherent to Never-Dried Arabidopsis thaliana Primary Cell Walls: Evidence from Solid-State NMR. *Plant Physiol.* **2015,** 168, (3), 871.

21. Wang, T.; Hong, M., Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J Exp Bot* **2016,** 67, (2), 503-14.

22. Dupree, R.; Simmons, T. J.; Mortimer, J. C.; Patel, D.; Iuga, D.; Brown, S. P.; Dupree, P., Probing the Molecular Architecture of Arabidopsis thaliana Secondary Cell Walls Using Two- and Three-Dimensional ^{13}C Solid State Nuclear Magnetic Resonance Spectroscopy. *Biochemistry* **2015,** 54, (14), 2335-2345.

23. Simmons, T. J.; Mortimer, J. C.; Bernardinelli, O. D.; Pöppler, A. C.; Brown, S. P.; deAzevedo, E. R.; Dupree, R.; Dupree, P., Folding of xylan onto cellulose fibrils in plant cell walls revealed by solid-state NMR. *Nat. Commun.* **2016,** 7, 13902.

24. Bardet, M.; Emsley, L.; Vincendon, M., Two-dimensional spin-exchange solid-state NMR studies of ^{13}C-enriched wood. *Solid State Nucl. Magn. Reson.* **1997,** 8, (1), 25-32.

25. Cadars, S.; Lesage, A.; Emsley, L., Chemical shift correlations in disordered solids. *J. Am. Chem. Soc.* **2005,** 127, (12), 4466-4476.

26. Lesage, A.; Bardet, M.; Emsley, L., Through-bond carbon-carbon connectivities in disordered solids by NMR *J. Am. Chem. Soc.* **1999,** 121, 10987-10993.

27. Sakellariou, D.; Brown, S. P.; Lesage, A.; Hediger, S.; Bardet, M.; Meriles, C. A.; Pines, A.; Emsley, L., High-resolution NMR correlation spectra of disordered solids. *J. Am. Chem. Soc.* **2003,** 125, 4376-80.

28. Sturcova, A.; His, I.; Apperley, D. C.; Sugiyama, J.; Jarvis, M. C., Structural details of crystalline cellulose from higher plants. *Biomacromolecules* **2004,** 5, (4), 1333-1339.

29. Wang, T.; Yang, H.; Kubicki, J. D.; Hong, M., Cellulose Structural Polymorphism in Plant Primary Cell Walls Investigated by High-Field 2D Solid-State NMR Spectroscopy and Density Functional Theory Calculations. *Biomacromolecules* **2016,** 17, (6), 2210-22.

30. Dick-Perez, M.; Wang, T.; Salazar, A.; Zabotina, O. A.; Hong, M., Multidimensional solid-state NMR studies of the structure and dynamics of pectic polysaccharides in uniformly ^{13}C-labeled Arabidopsis primary cell walls. *Magn. Reson. Chem.* **2012,** 50, (8), 539-550.

31. Oehme, D. P.; Downton, M. T.; Doblin, M. S.; Wagner, J.; Gidley, M. J.; Bacic, A., Unique aspects of the structure and dynamics of elementary Ibeta cellulose microfibrils revealed by computational simulations. *Plant Physiol.* **2015,** 168, (1), 3-17.

32. Vietor, R. J.; Newman, R. H.; Ha, M. A.; Apperley, D. C.; Jarvis, M. C., Conformational features of crystal-surface cellulose from higher plants. *Plant J.* **2002,** 30, 721-731.

33. Funahashi, R.; Okita, Y.; Hondo, H.; Zhao, M.; Saito, T.; Isogai, A., Different conformations of surface cellulose molecules in native cellulose microfibrils revealed by layer-by-layer peeling. *Biomacromolecules* **2017,** 18, (11), 3687-3694.

34. Horii, F.; Hirai, A.; Kitamaru, R., Solid-state C-13-NMR study of conformations of oligosaccharides and cellulose *Polym Bull* **1983,** 10, (7-8), 357-361.

35. Kubicki, J. D.; Mohamed, M. N. A.; Watts, H. D., Quantum mechanical modeling of the structures, energetics and spectral properties of I alpha and I beta cellulose. *Cellulose* **2013,** 20, (1), 9-23.

36. Kubicki, J. D.; Watts, H. D.; Zhao, Z.; Zhong, L. H., Quantum mechanical calculations on cellulose-water interactions: structures, energetics, vibrational frequencies and NMR chemical shifts for surfaces of I alpha and I beta cellulose. *Cellulose* **2014,** 21, (2), 909-926.

37. Watts, H. D.; Mohamed, M. N. A.; Kubicki, J. D., A DFT study of vibrational frequencies and ^{13}C NMR chemical shifts of model cellulosic fragments as a function of size. *Cellulose* **2013,** 21, (1), 53-70.

38. Yang, H.; Wang, T.; Oehme, D.; Petridis, L.; Hong, M.; Kubicki, J. D., Structural factors affecting 13C NMR chemical shifts of cellulose: a computational study. *Cellulose* **2018,** 25, 23-36.

39. Wang, T.; Hong, M., Structure and Dynamics of Polysaccharides in Plant Cell Walls from Solid-State NMR. In *NMR in glycoscience and glycotechnology*, Kato, K.; Peters, T., Eds. Royal Soc. Chem.: 2017; Vol. 10, pp 290-304.

40. Schubert, M., Insights into Carbohydrate Recognition by 3D Structure Determination of Protein-Carbohydrate Complexes Using NMR. In *NMR in glycoscience and glycotechnology*, Kato, K.; Peters, T., Eds. 2017; Vol. 10, pp 101-122.

41. Lange, A.; Luca, S.; Baldus, M., Structural constraints from proton-mediated rare-spin correlation spectroscopy in rotating solids. *J Am Chem Soc* **2002,** 124, (33), 9704-5.

42. Aluas, M.; Tripon, C.; Griffin, J. M.; Filip, X.; Ladizhansky, V.; Griffin, R. G.; Brown, S. P.; Filip, C., CHHC and ^{1}H-^{1}H magnetization exchange: Analysis by experimental solid-state NMR and 11-spin density-matrix simulations. *J. Magn. Reson.* **2009,** 199, (2), 173-187.

43. Mulder, F. M.; Heinen, W.; Duin, M. v.; Lugtenburg, J.; Groot, H. J. M. d., Spin diffusion with 13C selection and detection for the characterization of morphology in labeled polymer blends with MAS NMR. *J. Am. Chem. Soc,* **1998,** 120, 12891–12894.

44. Wilhelm, M.; Feng, H.; Tracht, U.; Spiess, H. W., 2D CP/MAS 13C isotropic chemical shift correlation established by 1H spin diffusion. *J. Magn. Reson.* **1998,** 134, 255–260.

45. Bali, G.; Foston, M. B.; O'Neill, H. M.; Evans, B. R.; He, J.; Ragauskas, A. J., The effect of deuteration on the structure of bacterial cellulose. *Carbohydr Res* **2013,** 374, 82-8.

46. Gelenter, M. D.; Wang, T.; Liao, S. Y.; O'Neill, H.; Hong, M., (2)H-(13)C correlation solid-state NMR for investigating dynamics and water accessibilities of proteins and carbohydrates. *J Biomol NMR* **2017,** 68, (4), 257-270.

47. He, J. H.; Pingali, S. V.; Chundawat, S. P. S.; Pack, A.; Jones, A. D.; Langan, P.; Davison, B. H.; Urban, V.; Evans, B.; O'Neill, H., Controlled incorporation of deuterium into bacterial cellulose. *Cellulose* **2014,** 21, (2), 927-936.

48. White, P. B.; Wang, T.; Park, Y. B.; Cosgrove, D. J.; Hong, M., Water-polysaccharide interactions in the primary cell wall of *Arabidopsis thaliana* from polarization transfer solid-state NMR. *J. Am. Chem. Soc.* **2014,** 136, (29), 10399-10409.

49. Veshtort, M.; Griffin, R. G., SPINEVOLUTION: a powerful tool for the simulation of solid and liquid state NMR experiments. *J Magn Reson* **2006,** 178, (2), 248-82.

50. Brown, G. M.; Levy, H. A., Alpha-D-Glucose: Precise Determination of Crystal and Molecular Structure by Neutron-Diffraction Analysis. *Science* **1965,** 147, (3661), 1038-9.

51. Dumez, J. N.; Emsley, L., A master-equation approach to the description of proton-driven spin diffusion from crystal geometry using simulated zero-quantum lineshapes. *Phys. Chem. Chem. Phys.* **2011,** 13, (16), 7363-7370.

52. Horii, F.; Hirai, A.; Kitamaru, R., Solid-state high-resolution C-13 NMR studies of regenerated cellulose samples with different crystallites. *Polym Bull* **1982,** 8, 163-170.

53. Massiot, D.; Fayon, F.; Capron, M.; King, I.; Le Calve, S.; Alonso, B.; Durand, J. O.; Bujoli, B.; Gan, Z. H.; Hoatson, G., Modelling one- and two-dimensional solid-state NMR spectra. *Magn Reson Chem* **2002,** 40, (1), 70-76.

54. Hediger, S.; Emsley, L.; Fischer, M., Solid-state NMR characterization of hydration effects on polymer mobility in onion cell-wall material. *Carbohydr. Res.* **1999,** 322, (1-2), 102-112.

55. Bertini, I.; Bhaumik, A.; De Paëpe, G.; Griffin, R. G.; Lelli, M.; Lewandowski, J. R.; Luchinat, C., High-resolution solid-state NMR structure of a 17.6 kDa protein. *J. Am. Chem. Soc.* **2010,** 132, 1032-1040.

56. Loquet, A.; Laage, S.; Gardiennet, C.; Elena, B.; Emsley, L.; Böckmann, A.; Lesage, A., Methyl proton contacts obtained using heteronuclear through-bond transfers in solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **2008,** 130, 10625-10632.

57. Wälti, M. A.; Ravotti, F.; Arai, H.; Glabe, C. G.; Wall, J. S.; Böckmann, A.; Güntert, P.; Meier, B. H.; Riek, R., Atomic-resolution structure of a disease-relevant Aβ(1-42) amyloid fibril. *Proc. Natl. Acad. Sci. U. S. A.* **2016,** 113, E4976-E4984.

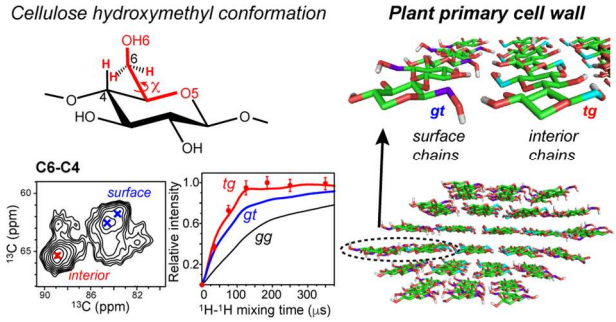
58. Jarvis, M. C., Structure of native cellulose microfibrils, the starting point for nanocellulose manufacture. *Philos. Trans. A Math. Phys. Eng. Sci.* **2018,** 376, 20170045.

59. Nixon, B. T.; Mansouri, K.; Singh, A.; Du, J.; Davis, J. K.; Lee, J. G.; Slabaugh, E.; Vandavasi, V. G.; O'Neill, H.; Roberts, E. M.; Roberts, A. W.; Yingling, Y. G.; Haigler, C. H., Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci. Rep.* **2016**, 28696.

60. Vandavasi, V. G.; Putnam, D. K.; Zhang, Q.; Petridis, L.; Heller, W. T.; Nixon, B. T.; Haigler, C. H.; Kalluri, U.; Coates, L.; Langan, P.; Smith, J. C.; Meiler, J.; O'Neill, H., A Structural Study of CESA1 Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers. *Plant Physiol.* **2016,** 170, 123-135.

61. Kobayashi, T.; Slowing, I. I.; Pruski, M., Measuring Long-Range C-13-C-13 Correlations on a Surface under Natural Abundance Using Dynamic Nuclear Polarization-Enhanced Solid-State Nuclear Magnetic Resonance. *J. Phys. Chem. C* **2017,** 121, 24687-24691.

TOC Figure


Cellulose hydroxymethyl conformation
Plant primary cell wall
OH6
6
H
H
χ
O5
4
HO
O
O
OH
gt
tg
surface chains
interior chains
C6-C4
surface
interior
tg
gt
gg
Relative intensity
1.0
0.6
0.2
60
65
90
86
82
0
100
200
300
¹³C (ppm)
¹³C (ppm)
¹H-¹H mixing time (μs)