

Quantum mechanical calculations on cellulose–water interactions: structures, energetics, vibrational frequencies and NMR chemical shifts for surfaces of I α and I β cellulose

James D. Kubicki · Heath D. Watts · Zhen Zhao ·
Linghao Zhong

Received: 31 May 2013 / Accepted: 13 August 2013 / Published online: 23 August 2013
© Springer Science+Business Media Dordrecht 2013

Abstract Periodic and molecular cluster density functional theory calculations were performed on the I α (001), I α (021), I β (100), and I β (110) surfaces of cellulose with and without explicit H₂O molecules of hydration. The energy-minimized H-bonding structures, water adsorption energies, vibrational spectra, and ¹³C NMR chemical shifts are discussed. The H-bonded structures and water adsorption energies (ΔE_{ads}) are used to distinguish hydrophobic and hydrophilic cellulose–water interactions. O–H stretching vibrational modes are assigned for hydrated and dry cellulose surfaces. Calculations of the ¹³C NMR chemical shifts for the C4 and C6 surface atoms demonstrate that these $\delta^{13}\text{C4}$ and $\delta^{13}\text{C6}$ values can be upfield shifted from the bulk values as observed without rotation of the hydroxymethyl groups from the bulk *tg* conformation to the *gt* conformation as previously assumed.

Keywords Density functional theory · Nuclear magnetic resonance · Surface · Infrared · Raman · Water

Introduction

Water in plant cell walls (PCWs) is a major factor affecting physicochemical and mechanical properties. The ratio of water to PCW polymers can be between 3:1 and 10:1 with higher ratios occurring depending on the cell type (Jarvis 2011). For example, a major difference between primary and secondary cell walls is the water content. Indeed, one of the reasons the deposition of lignin occurs within the secondary PCW may be to lower the water content (Albersheim et al. 2011). Specialized cells such as seed coatings and root tips appear to have evolved to optimize interactions with free water (Lindberg et al. 1990; Fekri et al. 2008; Iijima et al. 2008; Naran et al. 2008; Jarvis 2011). The variety and nature of water interactions with cellulose and other PCW polymers would require an extensive review, which is beyond the scope of this paper. For our purposes, we note that: “We cannot hope to understand the chemical reactivity of cellulose until we can construct a model for the molecular conformations and hydrogen bonding schemes associated with the cellulose–water interface” (Newman and Davidson 2004).

In spite of the fundamental importance of water in the PCW and the degradation of cellulose, there is a tendency to neglect this component of the system and

Electronic supplementary material The online version of this article (doi:10.1007/s10570-013-0029-x) contains supplementary material, which is available to authorized users.

J. D. Kubicki (✉) · H. D. Watts · Z. Zhao
Department of Geosciences, The Pennsylvania State
University, University Park, PA 16802, USA
e-mail: jdk7@psu.edu

L. Zhong
Department of Chemistry, The Pennsylvania State
University, Mont Alto, PA 17237, USA

focus on the biopolymers instead. The implicit assumption is that the water in the PCW behaves like bulk water and acts as an inert medium rather than active participant in PCW structure, dynamics and function. Much of the work on water–cellulose interaction has focused on extracted cellulose for applied purposes in the fiber, paper, wood, or biofuel industries.

Infrared (IR) spectra of deuterated cellulose has shown that deuteration of cellulose crystalline regions occurs (Mann and Marrinan 1956), but the mechanisms of this isotopic exchange is still a topic of current research (Matthews et al. 2012). Addressing the assumption mentioned above about the water in PCWs behaving like bulk water, Radloff et al. (1996) examined the dynamics of water in extracted cellulose. These authors found three distinct types of H₂O within the cellulose matrix: non-freezable, rigid H₂O amorphous below 270 K, a highly mobile H₂O with isotropic motion below 270 K, and H₂O that could not be removed from the matrix by drying at 370 K. Clearly, the water within PCWs is not likely to have the same properties as bulk water, so the properties of water within the PCW must be studied. Again, reviewing the literature on this subject would require an extensive review, so we refer to this statement: “Water at membrane surfaces affects membrane stability, permeability, and other properties. The properties of water at these and other important interfaces can be very different from those of water in the bulk” (Skinner et al. 2012).

Much of the previous molecular simulation work on cellulose has included water, but often the focus of the simulation is the structure of the cellulose. However, classical molecular dynamics (MD) simulations have examined water densities around cellulose microfibrils to predict how this property is affected by the cellulose surface present (Heiner and Teleman 1997; Heiner et al. 1998; Matthews et al. 2006). These authors concluded that water structuring differs on the (200), (010) and (110) surfaces of I β cellulose and could affect the degradation of cellulose by cellulose degrading enzymes. More recently, Matthews et al. (2011) have proposed a pathway for H–D exchange within the interior of the cellulose microfibril that does not require direct penetration of H₂O (Matthews et al. 2011), unlike the mechanism proposed previously (Mann and Marrinan 1956). Although the effects of water on cellobiose have been modeled with quantum

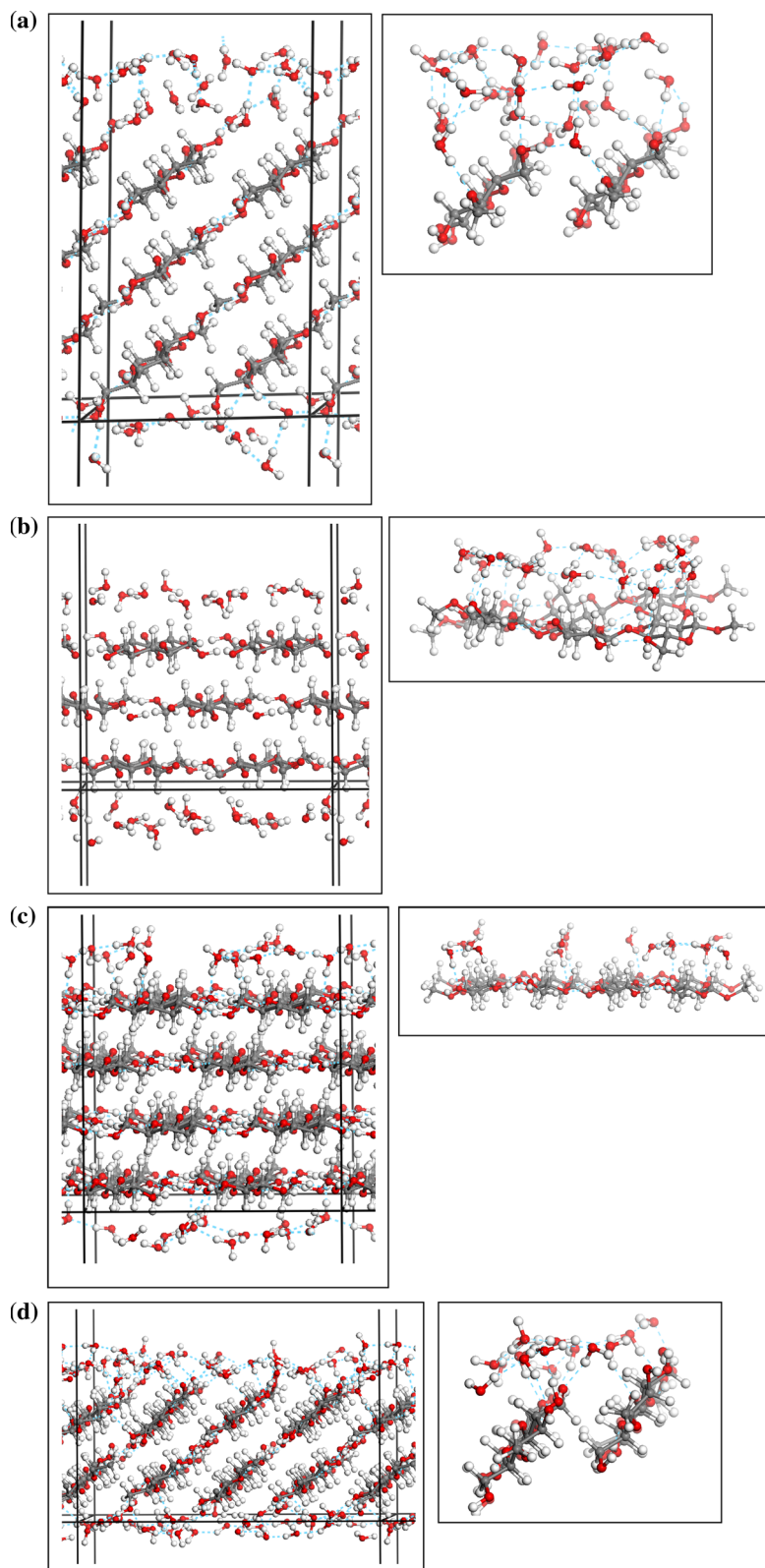
mechanical calculations as a model for cellulose–water interactions (French and Csonka 2011; French et al. 2012), only recently have periodic density functional theory (DFT) calculations been employed to model water interactions with cellulose surfaces (Li et al. 2011). Li et al. (2011) modeled the adsorption of a single H₂O molecule on the I β (100) surface to estimate the relative strengths of H-bonds to various cellulose surface sites. Studies on other surfaces (e.g., α -TiO₂) (Zhang and Lindan 2003) have shown that the addition of multiple H₂O molecules can change the distribution of H-bonds because it is the overall H-bond network strength that determines water adsorption behavior, not the energetics of each individual molecule. Hence, this paper focuses on DFT calculations of multiple H₂O molecules with various cellulose surfaces. To connect the model results with observation, we report calculated vibrational frequencies and ¹³C NMR chemical shifts that result from our modeled structures.

Methods

Surface models of cellulose I α (001) and (021) and I β (100) and (110) were created based on the X-ray and neutron diffraction structures of cellulose (Nishiyama et al. 2002, 2003, 2008). The atomic positions and lattice parameters of these experimentally-derived structures were previously energy minimized with DFT-D2 calculations (Kubicki et al. 2013). The DFT-D2 relaxed structures were then cleaved with the Surface Builder module of Materials Studio 6.0 (Accelrys Inc., San Diego, CA) along the respective Miller index planes. The I α (001) surface was doubled in size along the b-axis direction, the I β (100) was doubled along the b- and c-axis directions, and the I β (110) surface was tripled along the b-axis and doubled along the c-axis. 15 Å of vacuum were incorporated into the simulation cells to create finite slabs with space for addition of H₂O molecules (Fig. 1). The simulation cell dimensions were: I α (001) 10.39 × 13.13 × 30.86 Å³, I α (021) 10.39 × 16.27 × 25.50 Å³, I β (100) 20.79 × 16.37 × 29.18 Å³, and I β (110) 20.88 × 24.36 × 23.60 Å³.

H₂O molecules were added manually to the cells. After partial energy minimizations were performed with the Forcite module of MS 6.0 using the Universal force field (UFF) (Rappé et al. 1992), the same force field was

Fig. 1 Images of periodic (*left*) and cluster (*right*) model cellulose surfaces with H₂O—**a** I α (001), **b** I α (021), **c** I β (100), **d** I β (110). C grey, O red, H white. Dashed blue lines represent H-bonds (i.e., H–O <2.5 Å and O–H–O >90°). Images drawn with Materials Studio 6.0 (Accelrys Inc., San Diego, CA). (Color figure online)



used to run a 10 ps molecular dynamics simulation at 298 K with a 1 fs time step in the N–V–T ensemble. The purpose of these preliminary steps was not to attain an accurate structure but to ensure a reasonable initial structure for the DFT-based energy minimizations. DFT-MD simulations were attempted for the I β (110)-H₂O model but were prohibitively slow.

Periodic DFT-D2 calculations were performed with the Vienna Ab-initio Simulation Package (VASP) (Kresse and Hafner 1993, 1994; Kresse et al. 1994; Kresse and Furthmüller 1996). Projector-augmented plane-wave pseudopotentials were used with the PBE gradient-corrected exchange correlation functional for the 3-D periodic DFT calculations. The choice of electron density and atomic structure optimization parameters were based on published recommendations (Bučko et al. 2011; Li et al. 2011). An energy cut-off of 77,190 kJ/mol (ENCUT = 500 eV) was used with an electronic energy convergence criterion of 9.65×10^{-6} kJ/mol (EDIFF = 1×10^{-7} eV). Atomic structures were relaxed until the energy gradient was less than 1.93 kJ/mol/Å (EDIFFG = -2×10^{-2} eV/Å). 1 k-point samplings were used based on the relatively large size of the simulation cells. All atoms were allowed to relax with the lattice parameters constrained to the values obtained previously for the bulk crystals (Kubicki et al. 2013). The D2 dispersion-correction parameters were 40 Å for the cutoff distance (Bučko et al. 2011) and 0.75 for the scaling factor (s_6) and 20 for the exponential coefficient (d) in the damping function (Grimme 2006).

Frequency analyses were performed on the I α energy minimized surfaces as predicted using VASP (Note: The I β simulation cells were too large to practically perform frequency analyses. Our method for dealing with this issue is described below.). Second derivatives of the potential energy matrix with respect to atomic displacements were calculated using two finite-difference steps (NFREE = 2) and atomic movements of 0.015 Å (POTIM = 0.015). Vibrational modes were analyzed using the registered version of the program wxDragon 1.8.0 (Eck 2012).

Energies of water adsorption (ΔE_{ads}) were calculated using the computational methods above applied to the cellulose–water models and the cellulose surface and water models separately. Both the dry cellulose surface and a periodic model containing n H₂O molecules (where n is the number of H₂O molecules in a given cellulose–water model) were

energy minimized in separate simulations. The total energy of the separate cellulose and water models were then subtracted from the model with water adsorbed onto the surface, i.e.,

$$\Delta E_{\text{ads}} = E(\text{cellulose} - \text{water}) - E(\text{cellulose}) - E(n\text{H}_2\text{O}) \quad (1)$$

In order to evaluate the robustness of the DFT-D2 results, molecular fragments representing the surface polymers and adsorbed H₂O molecules were extracted from the energy-minimized DFT-D2 structures (Fig. 1). The cellulose polymeric chains were terminated with methyl groups manually attached to the O1 atoms using Materials Studio 6.0 (Accelrys Inc., San Diego, CA). The mPW1PW91 exchange–correlation functional (Adamo et al. 1998) and 6-31G(d) basis set (Krishnan et al. 1980; Clark et al. 1983) were used to energy minimize the O and H atoms of the extracted clusters with the C atoms fixed in Gaussian 09 (Frisch et al. 2009). This methodology allows the molecular clusters to retain the basic structure of the periodic system while relaxing the H-bonds. The clusters were then subjected to frequency analyses to obtain the IR and Raman spectra. The calculated vibrational frequencies, IR intensities and Raman activities were used to produce synthetic IR and Raman spectra by assigning full-width at half-maximum values of 1,200 cm⁻¹ to each vibration and summing the intensities all modes. Vibrational modes were visualized using the Molden 4.0 program (Schaftenaar and Noordik 2000).

Gaussian 09 (Frisch et al. 2009) calculations were also performed with the ω B97X-D (Chai and Head-Gordon 2008) functional and the 6-31G(d,p) basis set. The ω B97X-D exchange–correlation functional has proven reliable for reproducing H-bonding (Cirtog et al. 2011); hence, we use this method as a benchmark of the mPW1PW91/6-31G(d) calculations on the larger models. As a test, H-bond distances and dimerization energies were calculated with both ω B97X-D/6-31G(d,p) and mPW1PW91/6-31G(d) for H₂O–H₂O, MeOH–MeOH (MeOH is methanol) and H₂O–MeOH dimers (one where the MeOH is a H-bond donor and one where it is a H-bond acceptor). H-bond lengths were 1.92, 1.88, 1.91 and 1.88 Å, respectively using ω B97X-D, and 1.93, 1.91, 1.90 and 1.92 Å, respectively using mPW1PW91. Dimerization energies were –25, –28, –25 and –28 kJ/mol, respectively, using ω B97X-D, and –32, –35, –33 and –33 kJ/mol, respectively, using mPW1PW91. Thus, the H-bond lengths only

varied by a few hundredths of an Å and the H-bond energies by 7 kJ/mol, which we consider reasonable uncertainties for these parameters. Comparison of synthetic spectra produced for cellulose–water interactions by both methods (Fig. SI-1) shows that the frequencies are consistent to within 50 cm^{-1} . We expect that calculated O–H frequencies in the range $3,000\text{--}3,800\text{ cm}^{-1}$ are only accurate to about 100 cm^{-1} ($\approx 3\%$) compared to experiment, so we conclude that the mPW1PW91/6-31G(d) method provides reasonably accurate results compared to the ω B97X-D/6-31G(d,p) method.

NMR shielding tensor calculations on the finite clusters extracted from the 3-D periodic DFT-D2 calculations without relaxation were carried out using Gaussian 09 (Frisch et al. 2009). Gauge-independent atomic orbitals (GIAO) (Wolinski et al. 1990; Schreckenbach and Ziegler 1995; Cheeseman et al. 1996; Buhl et al. 1999; Karadakov 2006; Wiitala et al. 2006; Lodewyk et al. 2012) were employed with the modified Perdue-Wang exchange–correlation functional mPW1PW91 (Adamo et al. 1998) and the 6-31G(d) basis set (Rassolov et al. 2001). Chemical shifts were calculated relative to methanol because this secondary standard produces $\delta^{13}\text{C}$ in better agreement with experiment (Sarotti and Pellegrinet 2009; Watts et al. 2011) than does a direct comparison of the tensors with the tetramethylsilane standard (Cheeseman et al. 1996). This multi-standard reference method also uses an empirical correction of 49.5 ppm (Gottlieb et al. 1997) for the difference between the $\delta^{13}\text{C}$ in methanol and (TMS) commonly used as an experimental ^{13}C NMR standard (Sarotti and Pellegrinet 2009)

$$\sigma^{13}\text{C}_{\text{calc, MeOH}} + \delta^{13}\text{C}_{\text{exp, MeOH}}$$

This gives an isotropic chemical shielding of 193.0 ppm. To calculate the $\delta^{13}\text{C}$ for any C nucleus i , we used:

$$\delta^{13}\text{C}_i = 193.0\text{ ppm} - \sigma^{13}\text{C}_i$$

Note that for the I α model, the nomenclature for the hydroxymethyl torsion angles χ_1 (O5–C5–C6–O6) and χ_2 (O5–C5–C6–O6), and the glycosidic torsion angles Φ (O5–C1–O–C4), Ψ (C1–O–C4–C5) will be annotated without prime notation for residue 1 (i.e., χ_1 , χ_2 , Φ and Ψ), and with prime notation for residue 2 (i.e., χ_1' , χ_2' , Φ' and Ψ'). For the I β model, unprimed

notation will denote the torsion angle results for the origin chain, while primed notation will denote the torsion angle results for the center chain. This nomenclature is consistent with that employed by Nishiyama et al. (2002, 2003) for cellulose I α and I β . For the “Results” and the “Discussion” sections of this paper that address the relationship between particular NMR chemical shifts (i.e., $\delta^{13}\text{C}4$, $\delta^{13}\text{C}5$, and $\delta^{13}\text{C}6$), the hydroxymethyl torsion angles (i.e., χ_1 and χ_2), and glycosidic torsion angles (i.e., Φ and Ψ), we do not distinguish between the glucose residues for cellulose I α , nor between the origin and center chains for cellulose I β . Therefore, we do not use prime notation for those sections of this paper.

Results

Structural comparison of model versus observed cellulose structure

To demonstrate that the DFT-D2 methodology is capable of representing cellulose structures, we present a comparison of the calculated and experimental structural parameters for I α and I β cellulose. Table 1 compares the lattice parameters, glycosidic torsion angles (Φ and Ψ), hydroxymethyl torsion angles (χ_1 and χ_2), and ring-puckering parameters (θ) determined in Nishiyama et al. (2002, 2008) and Kubicki et al. (2013). Model values are close to experiment in almost every case. This is especially true when one examines the low temperature data (15 K) of Nishiyama et al. (2008) for I β cellulose compared to the DFT-D2 energy minimized structures that represent 0 K (Table 1). Thus, the discrepancies between observed and calculated lattice parameters for I α cellulose may be smaller when considering thermal expansion effects because the room temperature values are systematically larger than the calculated values. Repeat distances (i.e., the distance of the repeating units in a polymeric chain) for various celluloses were found to range from 1.029 to 1.043 nm (Davidson et al. 2004). The variation was ascribed to the size of the microfibril present in each type of cellulose. The repeat distance in the infinite 3-D periodic models for I α and I β in Kubicki et al. (2013) are 1.039 and 1.040 for the most stable *tg*/NetA configurations, which are consistent with Davidson

Table 1 Comparison of observed (Nishiyama et al. 2002, 2008) and calculated (Kubicki et al. 2013) bulk cellulose structural parameters

$I\alpha$	a	b	c	α	β	γ				
Room T	6.717	5.962	10.400	118.08	114.80	80.37				
DFT-D2	6.566	5.664	10.386	116.20	112.30	83.00				
$I\beta$	a	b	c	γ						
Room T	7.76	8.20	10.37	96.62						
Low T	7.64	8.18	10.37	96.54						
DFT-D2	7.55	8.14	10.40	96.40						
	Residue 1					Residue 2				
	Φ	Ψ	χ_1	χ_2	θ	Φ'	Ψ''	χ_1'	χ_2'	θ'
$I\alpha$	−98	−138	167	−75	9.4	−99	−140	166	−74	6.9
DFT-D2	−95	−141	165	−76	1.6	−93	−143	165	−86	0.1
	Origin					Center				
	Φ	Ψ	χ_1	χ_2	θ	Φ'	Ψ''	χ_1'	χ_2'	θ'
$I\beta$ Model A	−98.5	−142.3	170	−70	10.2	−88.7	−147.1	158	−83	6.7
$I\beta$ Model C	−93.9	−143.8	164	−77	3.2	−92.2	−145.5	161	−80	3.2
DFT-D2	−93	−143	168	−74	1.3	−94	−145	165	−76	2.8

Φ = O5–C1–O1–C4; Ψ = C1–O1–C4–C5

χ_1 = O5–C5–C6–O6 torsion, χ_2 = C4–C5–C6–O6 torsion

θ —Puckering parameter (Cremer and Pople 1975, JACS, 97, 1354) relative to unstrained α -D-glycopyranose. Calculated with web-based program supplied by Shinya Fushinobu

et al. (2004). However, the repeat distances for the less stable configurations also fall in this range (i.e., $I\alpha$ /NetB = 1.034, $I\alpha$ /gt = 1.040, $I\alpha$ /gg = 1.038; $I\beta$ /NetB = 1.036, $I\beta$ /gt = 1.040, $I\beta$ /gg = 1.041), so the repeat distance is not highly sensitive to the H-bonding network nor the hydroxymethyl configurations in these model calculations. M05-2X (Zhao et al. 2006) calculations with the 6-31G(d,p) basis set on a $3 \times 3 \times 4$ (4 glucose units with 9 chains) result in repeat distances of 1.047 nm for the central chain to 1.044 nm for the surface chains, so the repeat distance was not sensitive to model size in these calculations either.

One other notable discrepancy was found between the ring-puckering (θ) values for $I\alpha$ cellulose. The experimental θ is significantly larger than that calculated via DFT-D2 energy minimization. This is not an inherent problem with the DFT-D2 method because the model and experimental values are similar for $I\beta$ cellulose (Table 1). The reason for the θ discrepancy in $I\alpha$ is not clear, but it could be due to thermal effects

as noted above, uncertainty in the experimental determination, or other factors.

Energies

A key issue in cellulose–water interactions is the relative energy of the H₂O H-bond with the cellulose surface compared to the H-bond energy between H₂O molecules. The enthalpy of the H-bond between the H₂O molecules and the surface determines whether a surface is hydrophilic or hydrophobic. If the H₂O–surface H-bond is less than −44 kJ/mol (Fubini et al. 1999), then the surface is hydrophilic; if not, then H₂O–H₂O H-bonds are preferred and the surface is hydrophobic. However, the H-bond energy could be less in confined pores (Hiejima and Yao 2004) as may be the case for water in PCWs. DFT methods do not predict H-bond energies extremely accurately (Ireta et al. 2004), so we have used relative H-bond energies calculated with the same method to estimate the ΔE_{ads} of water onto a given surface (see “Methods”).

Table 2 Model energies (in eV) and calculated energies of adsorption (ΔE_{ads} in kJ/mol H_2O)

Surface	Cellulose + $n\text{H}_2\text{O}$	Cellulose	$n\text{H}_2\text{O}$	ΔE_{ads}
I α (001)	−1,976.0581	−1,503.3325	−473.6608	+3.0
I α (021)	−1,978.6477	−1,503.4789	−473.6608	−4.5
I β (100)	−4,485.3690	−4,014.1387	−473.6608	+7.0
I β (110)	−5,172.5618	−3,965.5686	−1,205.2284	−2.1

Table 2 presents the results of our water adsorption energy calculations. First, qualitative agreement with experiment is produced because the hydrophilic surfaces (I α (021) and I β (110); have negative ΔE_{ads} values, whereas the hydrophobic surfaces (I α (001) and I β (100) (Heiner and Teleman 1997) have positive values. We know of no experimentally produced adsorption enthalpy values for individual cellulose surfaces to which we can quantitatively compare our ΔE_{ads} values. In comparison to classical force field results using the CHARMM force field (Guvench et al. 2009), the interaction energies are −65 kJ/mol per glucose for I β (100) and −56 kJ/mol per glucose I β (110) without relaxation of the reactants in Surface + $n(\text{H}_2\text{O}) \rightarrow \text{Surface}-n(\text{H}_2\text{O})$. This comparison exaggerates the difference between DFT and CHARMM because the reactants were not relaxed in the CHARMM calculations, but the CHARMM predictions are still in disagreement for the hydrophobic I β (100) surface with both the DFT results and experimental observation because the interaction energy is more negative than for $\text{H}_2\text{O}-\text{H}_2\text{O}$ and for the hydrophilic (110) surface. The CHARMM force field results show less water H-bonding to the I β (100) surface, so the reason for this discrepancy is not due to forming H-bonds with glycosidic O atoms or interactions with CH groups on the cellulose surface. Furthermore, calculating ΔE_{ads} with the mPW1PW91/6-31G(d) method using the CHARMM structures gives values of −82 and −52 kJ/mol glucose, which are similar to the CHARMM-predicted values but very different from the DFT-D2 results (Table 2).

H-bonding structures

The types and percentages of H-bonds formed at each surface are important because they result in the adsorption energies discussed above and because they provide insight into how cellulose may become disordered. For cellulose surface-water interactions the H_2O and COH groups can form H-bonds either by

donating or accepting H^+ from one another (i.e., H-bonding from the H of the H_2O to the O of the COH or from the H of the COH to the O of the H_2O). In addition, H_2O can form H^+ -donating H-bonds to the glycosidic O atoms on the surface although glycosidic O atoms form weaker H-bonds with water than COH groups (O'Dell et al. 2012). The types of interactions are critical for determining whether interaction with water will disrupt the COH intra- and inter-chain H-bonding of the cellulose microfibril (Nishiyama et al. 2002, 2003) or whether H_2O will have the opportunity to hydrolyze the glycosidic linkage (Gazit and Katz 2013).

Table 3 presents the percentage of each type of H-bond between water and each surface along with the average H-bond length (i.e., the H–O distance in the H-bond). There is no obvious distinction in percentages of H-bond types between the hydrophobic I α (001) and I β (100) and hydrophilic (I α (021) and I β (110) surfaces. Although I β (100) has a high percentage of intrasurface H-bonds and no COH donor H-bonds, as one may expect for a hydrophobic surface, the I α (001) surface, which is also hydrophobic, has a similar distribution of H-bond types as the hydrophilic I β (110) surface (Table 3). In all cases, the percentage of C–OH donor and glycosidic H-bonds are minority percentages. These two types of H-bonds have the potential to disrupt the cellulose bulk structure either by disrupting intra- and inter-chain H-bonding or by hydrolyzing the glycosidic linkage. However, because the disruption of intrasurface H-bonds on these surfaces is minimal, interaction with water on these hydrophobic surfaces does not significantly disorder the cellulose compared to the bulk structure. We note, however, that these structures are based on 0 K energy minimizations and are likely to change if MD simulations of the models were run at finite temperature.

Comparison of the hydrophobic I α (001) and I β (100) H-bond types shows that the latter has no C–OH

Table 3 Percent H-bond types on cellulose surfaces and average H-bond distance (Å)

Surface	C–OH donor	C–OH acceptor	Glycosidic acceptor	Intrasurface
I α (001)	18/1.92	24/1.94	12/2.11	46/1.78
I α (021)	0/–	28/1.82	24/2.11	48/1.76
I β (100)	0/–	24/2.01	2/2.36	74/1.87
I β (110)	18/1.69	42/1.83	5/1.88	35/1.80

donor H-bonds and a larger percentage of intrasurface H-bonds. Considering that structure-disrupting H-bonds may allow for more efficient biodegradation of cellulose by weakening the inter-chain interaction energies (Matthews et al. 2006), it is reasonable to hypothesize that the I α to I β transition could be an effective mechanism for increasing biodegradation resistance of cellulose in the plant cell wall.

Examination of the average H-bond lengths in Table 3 reveals differences between hydrophobic and hydrophilic surfaces, however. The H-bond distance is correlated with H-bond strength (Jeffrey 1997), so the decrease in average H-bond distance between hydrophobic and hydrophilic surfaces explains the calculated ΔE_{ads} values in Table 2. The difference is especially significant for the I β (100) and I β (110) surfaces; the average H-bond distance for each type of H-bond is shorter in the I β (110) case [i.e., I α (001) = 1.88 vs. I α (021) = 1.86 Å; I β (110) = 1.79 vs. I β (100) = 1.91 Å]. Furthermore, I β (110) has a significant percentage (18 %) of strong H-bonds in the C–OH donor category. This type of strong H-bonding may be a problem for classical force fields because force fields are typically parameterized based on average equilibrium configurations and do not accurately model the tails of the distributions around these equilibrium values.

Vibrational frequencies and modes

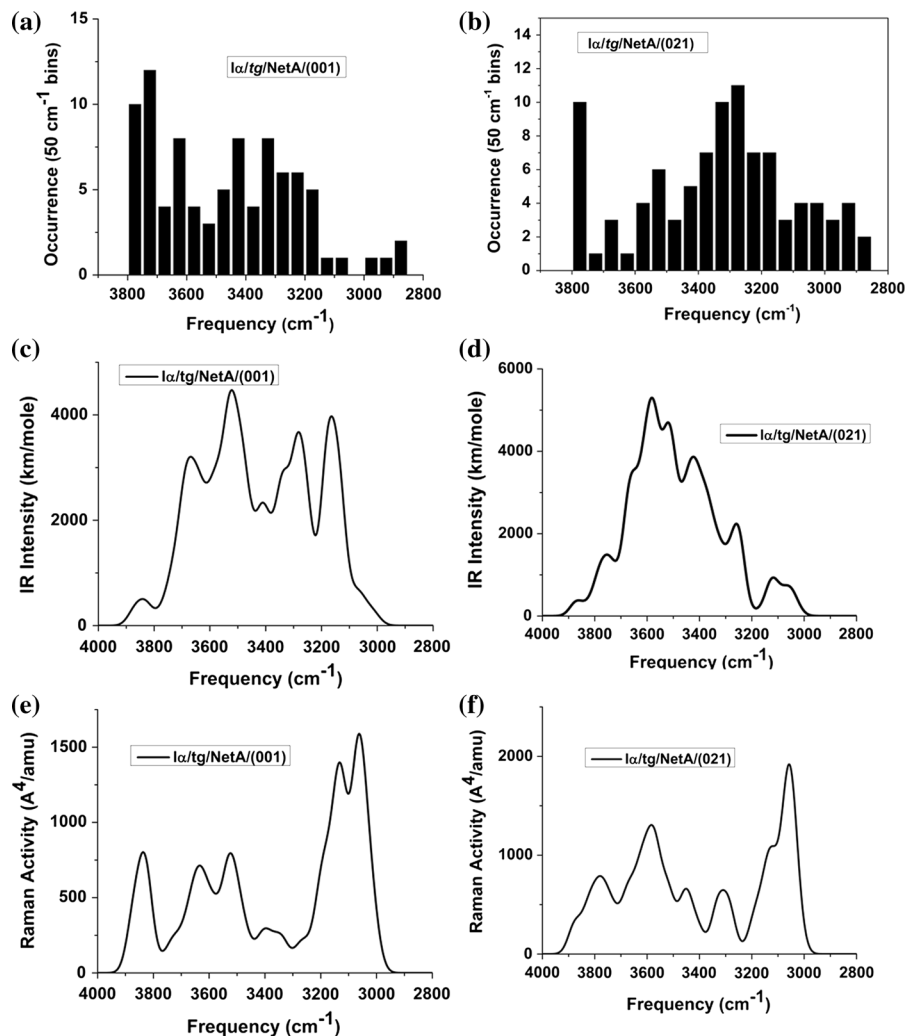
Although the DFT-D2 methodology used in this study does not reproduce observed O–H stretching frequencies more accurately than approximately 100 cm^{−1} (Lee et al. 2013), the relative order of calculated vibrational frequencies matches experiment well (Kubicki et al. 2013). Quantum mechanical calculations of vibrational frequencies provides vibrational modes associated with each frequency which is a significant advantage in interpreting broad spectra

such as those observed for cellulose–water interfaces (e.g., Brizuela et al. 2012). The contribution of each type of H-bond to observed bands can be deconvoluted whereas this is impractical from observed spectra alone because numerous vibrational modes may be contributing to any peak within the IR or Raman spectra (Blackwell 1977; Wiley and Atalla 1987; Maréchal and Chanzy 2000).

Supplemental Information (SI) Table 1 lists all the calculated vibrational frequencies and the associated modes for all the models in this study except for the periodic I β surfaces where frequency calculations were impractical due to the size of the system. However, this problem was circumvented by performing energy minimizations and frequency calculations on molecular clusters extracted from the energy-minimized periodic structures (see “Methods”). The correlation between the periodic DFT-D2 and molecular cluster frequencies and vibrational modes is good even though there is a systematic offset of 100–200 cm^{−1} between the two methods. This could be diminished by scaling the molecular cluster frequencies by an empirical factor of 0.96–0.98 (Alecu et al. 2010), but this was not necessary in the current study because we focus on relative positions of the various O–H stretching frequencies.

Figure 2 illustrates the number of O–H stretching frequencies for I α (001) and (021) as calculated in the periodic DFT-D2 calculations and the model IR and Raman spectra in the 2,800–4,000 cm^{−1} range from molecular cluster calculations. (Note that neither IR nor Raman intensities are calculated within the DFT-D2 methodology used in this study. Gaussian 09 is capable of calculating both IR and Raman intensities which illustrates the advantage of using both approaches in tandem). Figure 3 shows the model IR and Raman spectra for the molecular cluster models of I β (100) and (110) cellulose–water interfaces. Figure 4 represents examples of the vibrational modes used to determine frequency assignments.

Fig. 2 Periodic model frequencies in bins of 50 cm^{-1} (**a, b**) and molecular cluster IR and Raman (**c–f**) spectra of $\text{I}\alpha$ cellulose (001) and (021) surfaces with adsorbed H_2O . Peaks near $3,000\text{ cm}^{-1}$ are dominated by C–H modes; this peak can be used as a marker where O–H modes end and to compare model spectra with observed spectra



Previous researchers have made assignments of cellulose vibrational modes to specific frequencies that can be compared to our results. We caution that all of these studies were performed on non-native cellulose and the vibrational spectra may be altered compared to cellulose in the plant cell wall (Lee et al. 2013). For example, Kalutskaya and Gusev (1981) assigned $3,200$ and $3,560\text{ cm}^{-1}$ bands to adsorbed H_2O molecules on cellulose, and the closest vibrational frequencies in our calculations are due to O–H stretches in H_2O – H_2O interactions although contributions from O–H stretches in COH groups are also probable (SI Table 1). Horikawa et al. (2006) observed polarization effects in bands at $1,317$, $1,337$, $1,355$ and $1,372\text{ cm}^{-1}$ and assigned the first two frequencies to O–H motion perpendicular to the fibril

axis (i.e., in-plane COH bending) while the latter two were thought to be due to motion parallel to the fibril axis. Our calculations produced numerous frequencies in the $1,317$ – $1,372\text{ cm}^{-1}$ range some with motions parallel and other perpendicular to the fibril axis. However, all of these frequencies were due to concerted motions of C and O atoms such as ring-breathing and rocking modes and not due to specific O–H motions such as COH bends. Nakashima et al. (2008) assigned $\text{I}\alpha$ specific bands at 750 and $3,240\text{ cm}^{-1}$ and $\text{I}\beta$ specific bands at 710 and $3,270\text{ cm}^{-1}$. In our model $\text{I}\alpha$, the 750 cm^{-1} band is due to a C3O3H bend and the $3,240\text{ cm}^{-1}$ band is a concerted motion of O2-H/O6-H/O3-H stretches. In our model $\text{I}\beta$, the 710 cm^{-1} band is due to a C3O3H bend in the origin chain and the $3,270\text{ cm}^{-1}$ band is a

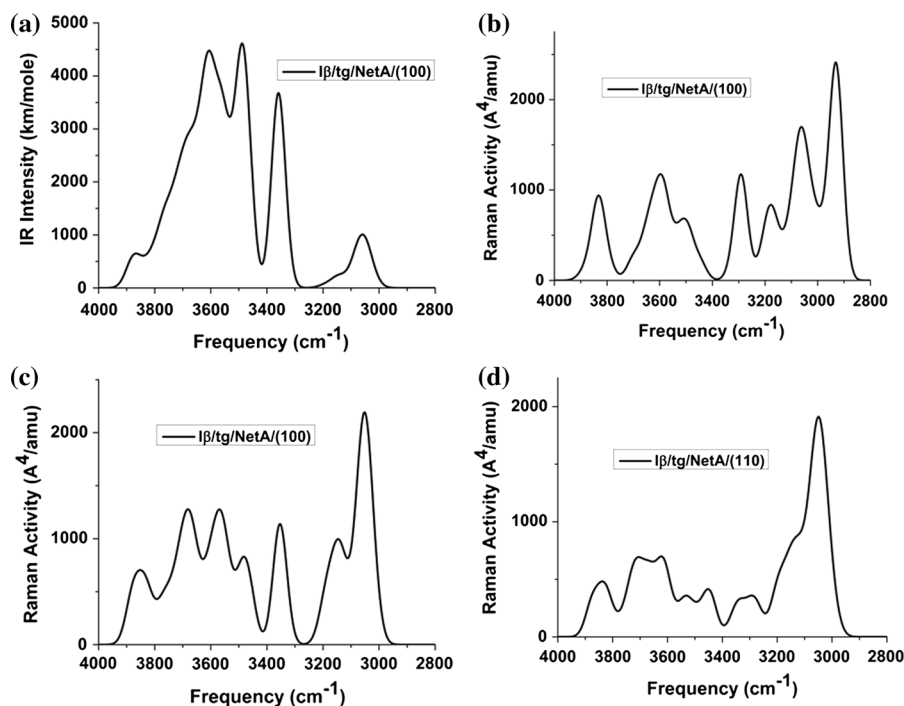


Fig. 3 Molecular cluster IR and Raman (a–d) spectra of Iβ cellulose (100) and (110) surfaces with adsorbed H₂O. Peaks near 3,000 cm^{−1} are dominated by C–H modes; this peak can be

used as a marker where O–H modes end and to compare model spectra with observed spectra

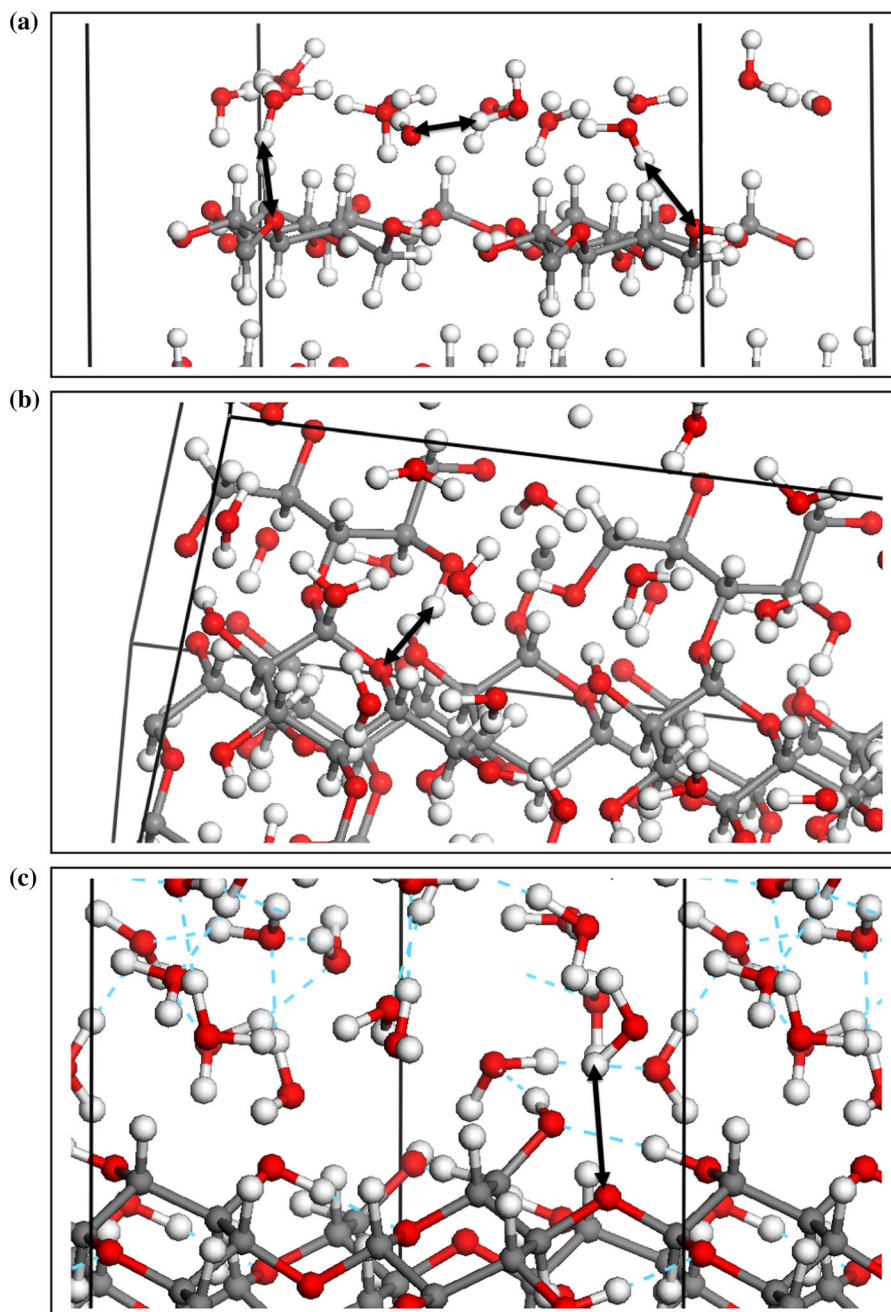
O3–H stretch in the center chain combined with a O6–H stretch because the H-bonding pattern is O6–H–O3–H–O5. The O3–H–O6 H-bond is slightly shorter in Iβ than Iα in our model (1.766 vs. 1.785 Å), so the lower frequency bend in the former is understandable. A simple comparison for the 3,240 and 3,270 cm^{−1} bands cannot be made because the modes have significantly different motions in Iα and Iβ.

The vibrational modes in the Iα periodic DFT-D2 calculations (SI Table 1) do not occur in completely separate frequency ranges, but the O–H stretches of various H-bond types do tend to cluster within representative regions. H₂O–H₂O and H₂O donor O–H stretches are found throughout the 3,200–3,750 cm^{−1} region, so they would be difficult to distinguish as separate peaks within observed spectra. However, the isolated O–H stretches (i.e., those with no associated H-bonding) are only found in the 3,700–3,800 cm^{−1} region consistent with the observation of isolated O–H stretching frequencies at 3,750 cm^{−1}. Weaker H₂O–glycosidic O H-bonds result in frequencies in the 3,550–3,870 cm^{−1} range with stronger C–OH donor H-bonded O–H stretches

fall between 3,250 and 3,550 cm^{−1}. Importantly, COH intrasurface H-bonds occur in the range of 2,950–3,400 cm^{−1}. This result is consistent with the relative strength of C–OH intrasurface H-bonds and could be useful in identifying disruption of the cellulose surface via vibrational spectroscopy. This could explain the observation of Thomas et al. (2013) that IR intensity is retained in the 3,400 cm^{−1} region after deuteration of cellulose surfaces because intrasurface H-bonded groups are less likely to be able to undergo exchange with D₂O than surface OH groups directly H-bonded to D₂O.

As mentioned above, VASP does not calculate IR or Raman intensities, so the calculated frequencies were binned in ranges of 50 cm^{−1} and plotted as bar graphs in Fig. 2a, b. The Gaussian 09 program does allow for calculation of IR and Raman intensities, but the models used in this case are molecular clusters, not the periodic surfaces employed in the VASP calculations. Thus, both methods have their limitations, so we have used both in order to compare the frequencies from the periodic systems to the molecular clusters and then use the calculated IR and Raman intensities

Fig. 4 Example vibrational modes—**a** $I\alpha$ (021) COH acceptor— $3,628\text{ cm}^{-1}$, **b** $I\alpha$ (021) COH intrasurface— $3,400\text{ cm}^{-1}$ and **c** $I\alpha$ (001) H_2O -glycosidic O— $3,622\text{ cm}^{-1}$



from the molecular clusters if the frequencies are similar to the periodic models. Although the molecular cluster frequencies calculated with mPW1PW91/6-31G(d) are systematically higher than the PBE-calculated frequencies for the periodic system by approximately $100\text{--}200\text{ cm}^{-1}$ (e.g., C–H stretches from DFT-D2 occur between $2,800$ and $3,000\text{ cm}^{-1}$ (Fig. 2a, b) and from mPW1PW91/6-31G(d) between

$3,000$ and $3,200\text{ cm}^{-1}$), the range of O–H frequencies is the same and importantly the ordering of the modes is similar with both methods (SI Table 1). We also note that the IR and Raman intensities in Fig. 2c–f do not follow the trend of the number of O–H stretches seen in Fig. 2a, b. This is because the O–H IR intensity increases at lower frequencies as H-bonding strengthens (Paterson 1982).

Isolating or distinguishing the vibrational spectra arising from a particular surface of cellulose microfibrils (CMF) in an experiment is difficult due to the small size of the CMFs. Consequently, models of individual cellulose surface–water interfaces are useful in understanding the differences in cellulose–water interactions among various surfaces (Matthews et al. 2006). For example, comparing Fig. 2c with d and e with f, one can see significant differences in the IR and Raman spectra of the hydrophobic versus hydrophilic $I\alpha$ surfaces. The (001) IR spectrum is evenly distributed in the 3,100–3,700 cm^{-1} (Fig. 2c) whereas the (021) spectrum is highly skewed towards frequencies at 3,600 cm^{-1} . The Raman spectra (Fig. 2e, f) are more similar than the IR spectra, but the (001) surface has greater Raman intensity near 3,200 cm^{-1} . These results may be surprising because the (021) surface is generally considered more hydrophilic, but these synthetic spectra suggest that at least of portion of the H-bonds at the (001)–water interface are stronger than at the (021)–water interface. However, according to the assignment of vibrational modes made above, these stronger H-bonds are actually dominated by COH intrasurface H-bonds. We interpret this result as the intrasurface H-bonding out-competing the potential H-bonds between water and the surface.

Figure 3 contains synthetic IR and Raman spectra for the $I\beta$ molecular fragment surfaces examined in this study. Overall, the spectra of the two cellulose–water interfaces are similar. The $I\beta$ (110) calculated IR spectrum has maximum intensity near 3,450 cm^{-1} compared to 3,650 cm^{-1} for the $I\beta$ (100) surface, that may reflect strong H-bonding in the former. The $I\beta$ (110) Raman spectrum has its lowest frequency O–H overlapping with the C–H stretching region near 3,100 cm^{-1} whereas these two peaks are more distinct in the model $I\beta$ (100) surface spectrum. These distinctions are minor, however; structural and dynamic variations in real systems could mask these subtle differences.

Another way to discern among the various types of H-bonds contributing to the spectra and to compare to experimental spectra is to remove the H_2O molecules from the model and calculate IR and Raman spectra of “dry” surfaces (We note that “dry” or “drying” depends upon the method of drying and that many surfaces dried at ambient conditions may have water adsorbed onto them. In this case, “dry” is completely devoid of water in the model). Supplemental

Information (SI) Figure 2 shows the IR spectra of all four cellulose surfaces. The $I\alpha$ (001) surface loses almost all IR intensity with only small O–H stretching peaks near 3,650 and 3,800 cm^{-1} remaining (in addition to the C–H stretches between 3,000 and 3,100 cm^{-1}). Not only has the removal of water taken away the stronger H_2O – H_2O and H_2O –COH H-bonds, but the intrasurface COH modes have moved from 3,200 to 3,350 (SI Table 1) to 3,600 cm^{-1} (i.e., weaker H-bonding). Similar behavior of the model (021) surface also occurs but the shift is not as strong (i.e., from 3,300 cm^{-1} wet to between 3,400 and 3,500 cm^{-1} dry). The $I\beta$ (100) and (110) spectra change in a similar manner to the $I\alpha$ surfaces (SI Fig. 2), but in the (100) case (SI Fig. 2c), there is a distinctive shift of the COH intrasurface modes from 3,300 to 3,400 cm^{-1} while the rest of the spectrum does not change significantly. This result is consistent with the data and interpretation that were based on analysis of lower frequency bands (750–770 and 950–1,200 cm^{-1}) during dehydration of cotton fibers (Liu et al. 2010).

We conclude that drying cellulose changes the IR spectrum in a more complex manner than simply removing the vibrational modes associated with water; the surface itself changes H-bond strength when water is not present. If this occurs in an experiment, one could assign the higher frequency bands in the cellulose–water interface spectra to COH modes because these peaks remain after drying when the original cellulose–water system would have stronger intrasurface COH H-bonds at lower frequencies. This behavior may have implications for processing cellulose and for the behavior in secondary plant cell walls where varying levels of dehydration occur (Albersheim et al. 2011).

$\delta^{13}\text{C}4$ and $\delta^{13}\text{C}6$ chemical shifts
and hydroxymethyl and glycosidic torsions

^{13}C NMR spectroscopy has been an extremely useful tool for understanding cellulose structure (Erata et al. 1997; Sternberg et al. 2003; Witter et al. 2006). Changes from the bulk ^{13}C chemical shifts ($\delta^{13}\text{C}$) induced by surficial interactions can provide insights into adsorption reactions as well (Dick-Pérez et al. 2011; Fernandes et al. 2011; Harris et al. 2012). Interpreting the structural nature of cellulose with observed $\delta^{13}\text{C}$ changes can be problematic, however,

because rotations between the *tg*, *gt* and *gg* conformations can be rapid compared to ^{13}C NMR relaxation times. This leads to averaging of the $\delta^{13}\text{C}$ values among all these species. Quantum mechanical calculations can be helpful in interpreting the observed shifts because the calculated $\delta^{13}\text{C}$ values are reasonably accurate (i.e., maximum errors of ± 2 ppm) (Kubicki et al. 2013), and the atomic structures giving rise to a given model $\delta^{13}\text{C}$ value are known, and the structures can be manipulated in order to observe induced changes in $\delta^{13}\text{C}$.

Previous work (Wickholm et al. 1998; Newman and Davidson 2004; Malm et al. 2010; Harris et al. 2012) has noted that surficial $\delta^{13}\text{C}_4$ are upfield shifted by approximately 4–5 ppm relative to bulk C4 atoms and observed from 83.3 to 84.9 ppm. The shift in $\delta^{13}\text{C}_4$ value has been ascribed to changes in the χ_1 and χ_2 hydroxymethyl torsion angles even though the δC_5 and δC_6 atoms also involved in this rotation exhibit surface versus bulk shift changes of 2–3 ppm (Horii et al. 1983, 1984; Harris et al. 2012). The relative insensitivity of C5 and C6 NMR chemical shifts to χ_1 and χ_2 torsion angle changes has been previously calculated (Kirschner and Woods 2001; Gonzalez-Outeiriño et al. 2006). In the current study, we focus on the changes to χ_1 and χ_2 torsion angles from bulk cellulose predicted at the cellulose–water interfaces and their effect on $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$. Refer to the Methods section for the nomenclature conventions used herein for Φ , Ψ , χ_1 and χ_2 , and note that we are not using prime notation for the hydroxymethyl torsion angles (χ_1 and χ_2) or the glycosidic torsion angles (Φ and Ψ) results discussed in this section.

Calculated internal $\delta^{13}\text{C}_5$ and $\delta^{13}\text{C}_6$ values agree with observed chemical shifts to within 1 ppm (Kubicki et al. 2013), but the $\delta^{13}\text{C}_4$ values are 1–3 ppm less than the observed values. This error is approximately the difference between internal and surface $\delta^{13}\text{C}_4$ values, so we focus on relative changes in $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$ values between cellulose bulk and surface (i.e., the magnitude of the upfield shift).

Table 4 lists $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$ involved in surficial hydroxymethyl groups of our four model surfaces. Many of the surficial $\delta^{13}\text{C}$ values are indistinguishable from the values calculated for the bulk and would not be observable as surface atoms. Even for the C4 atoms involved in the largest χ_1 and χ_2 distortions (157 and -85 , respectively, for $\text{I}\beta$ (110), the calculated $\delta^{13}\text{C}_4$ is

Table 4 Calculated C4 and C6 $\delta^{13}\text{C}$ values for surface hydroxymethyl groups (sC4 experiment = 83.8 and 84.9 ppm; Newman and Davidson 2004; 84–86 ppm; Harris et al. 2012; sC6 experiment \approx 62 ppm; Harris et al. 2012) and associated hydroxymethyl and glycosidic torsions

Surface	$\delta^{13}\text{C}_4$	$\delta^{13}\text{C}_6$	χ_1 (°)	χ_2 (°)	Φ (°)	Ψ (°)
<i>I</i> α (001)						
Periodic	89.7	66.9	161	-81	-93	-142
Periodic	85.5	64.5	165	-76	-96	-147
Cluster	85.9	64.9	160	-82	-94	-145
Cluster	86.0	65.4	160	-81	-94	-147
<i>I</i> α (021)						
Periodic	88.7	66.4	164	-79	-94	-145
Periodic	87.0	64.6	164	-79	-94	-145
Periodic	88.3	67.1	161	-80	-94	-143
Periodic	87.9	67.4	166	-75	-94	-147
Periodic	88.1	65.1	161	-81	-90	-144
Periodic	88.6	65.9	166	-75	-94	-147
Cluster	84.7	65.8	155	-84	-83	-142
Cluster	80.5	64.2	153	-86	-90	-134
Cluster	89.5	63.7	167	-75	-102	-147
Cluster	80.3	59.8	169	-75	-96	-145
<i>I</i> β (100)						
Periodic	88.1	64.1	162	-80	-93	-139
Periodic	84.7	64.8	163	-79	-93	-139
Periodic	90.0	68.5	158	-84	-91	-142
Periodic	90.9	66.8	158	-84	-94	-143
Cluster	88.5	66.4	157	-85	-92	-142
Cluster	81.6	61.6	159	-83	-94	-140
Cluster	87.1	62.4	158	-82	-92	-139
Cluster	89.5	64.2	159	-82	-94	-139
<i>I</i> β (110)						
Periodic	90.4	65.5	160	-81	-92	-145
Periodic	90.1	66.6	166	-75	-95	-147
Periodic	90.0	65.1	157	-85	-94	-143
Cluster	87.1	64.5	152	-87	-93	-131
Cluster	83.4	61.3	173	-66	-93	-152
Cluster	85.0	61.5	159	-82	-82	-128
Cluster	85.0	65.0	168	-73	-96	-151

“Periodic” results are from the 3-D periodic DFT-D2 structures, and “Cluster” results are from the molecular clusters energy minimized with mPW1PW91/6-31G(d). (*I* α – Res 1 $\chi = 167$, $\chi' = -75$, Res 2 $\chi = 166$, $\chi' = -74$; *I* β “C”–Orig $\chi_1 = -77$, $\chi_2 = 164$, Center $\chi_1 = -80$, $\chi_2 = 161$). In the table below, $\chi_1 = \text{O5–C5–C6–O6}$ and $\chi_2 = \text{C4–C5–C6–O6}$ without distinguishing between residues for *I* α or origin and center chains for *I* β (i.e., not prime notation, as discussed in the “Methods” section)

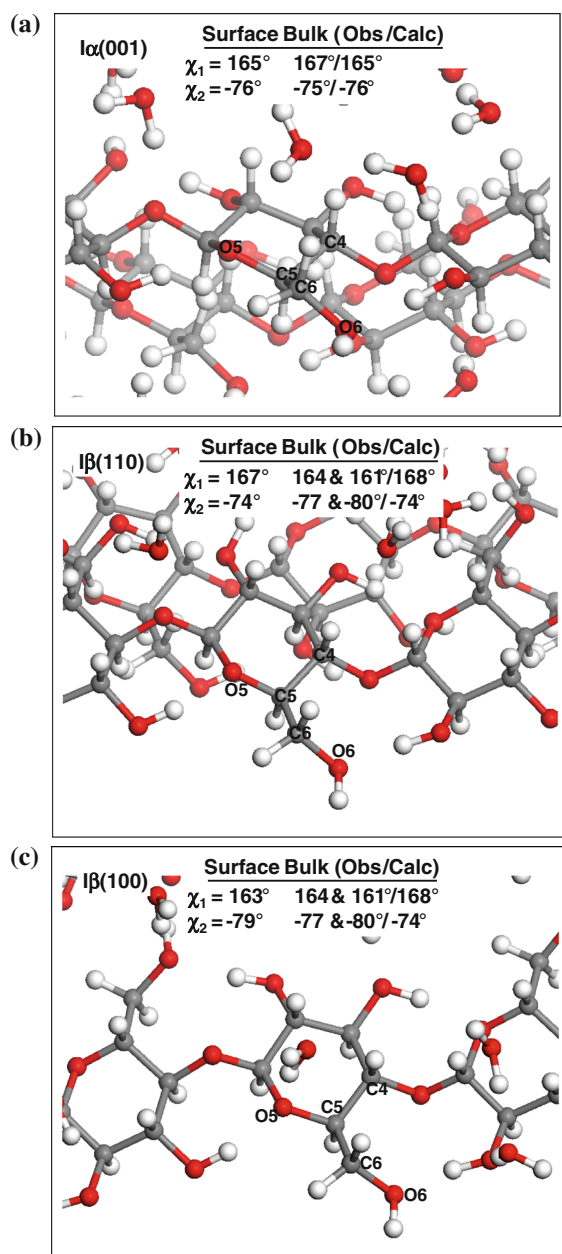


Fig. 5 Rotations of hydroxymethyl groups with C4 $\delta^{13}\text{C}$ values approximating the observed surface C4 atoms

90 ppm (Table 4) which is in the range observed for the bulk). For the $I\alpha(001)$ and $I\beta(100)$ models, there are individual C4 atoms with $\delta^{13}\text{C} < 86$ ppm that correspond to the 4–5 ppm upfield shift observed. The C6 atoms in these hydroxymethyl groups with the lower $\delta^{13}\text{C}_4$ also tend to have lower $\delta^{13}\text{C}_6$ values (Table 4). The hydroxymethyl groups exhibit χ_1 and χ_2 torsion angles that are not significantly different

from bulk observed and calculated values (Table 4; Fig. 5). We also checked the C4 bond distances, C–O–C angles and θ of the rings and found no significant perturbations in these parameters that could explain the change in $\delta^{13}\text{C}_4$.

The $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$ shifts and the associated torsion angles discussed above were based on the periodic DFT-D2 surface-water interface structures without relaxation of the atoms in Gaussian 09. To further investigate the connection between torsions and $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$ values, we relaxed the atoms in the extracted clusters using the mPW1PW91/6-31G(d) method to minimize the model potential energy, and then re-calculated the ^{13}C chemical shifts in the relaxed structures. This will allow for greater relaxation than the periodic system. These molecular cluster ^{13}C chemical shifts and torsion angles are also listed in Table 4. Because the structure is no longer periodic, the χ , Φ and Ψ torsional angles, exhibit a broader range of values. Although none of the hydroxymethyl groups adopt a *gt* or *gg* conformation that could also give rise to the upfield shift (Malm et al. 2010), there are numerous $\delta^{13}\text{C}_4$ and $\delta^{13}\text{C}_6$ values that are upfield shifted from the bulk values and into the range of the observed sC4 and sC6 ^{13}C NMR peaks ($\delta^{13}\text{C}_4$ in the range of 80–86 ppm and $\delta^{13}\text{C}_6$ values in the range 60–64 ppm). As with the DFT-D2 energy-minimized structures, the largest upfield shifts are not necessarily associated with the largest changes in torsion angles (e.g., $I\alpha(001)$ $^{13}\text{C}_4 = 80.3$ ppm and $\delta^{13}\text{C}_6 = 59.8$ ppm with χ_1 , χ_2 , Φ , and $\Psi = 169^\circ$, -75° , -96° and -145° , respectively). We note that calculated values for the $I\beta(110)$ cluster have values of 83.4, 85.0 and 87.1 ppm. The first two could correspond to the deconvoluted peaks at 83.3 and 84.9 ppm of Malm et al. (2010). The other two peaks are not reproduced by this work and may be associated with the (1–10) surface that was not modeled in this study.

These predictions imply that some surficial C atoms may be indistinguishable from the bulk; therefore, quantification of surface interactions via $\delta^{13}\text{C}$ changes may not be possible. Furthermore, in each case where the C4 value is upfield shifted to match the change observed in Harris et al. (2012), the hydroxymethyl group is not H-bonded to an H_2O molecule (Fig. 5). We interpret this result as reflecting changes induced in the $\delta^{13}\text{C}_4$ because H-bonds do not exist similar to the configuration present in the bulk. The adsorbed H_2O molecules are playing a role similar to the

neighboring chains in the bulk; hence, the H-bonded cellulose-water interface is more similar to the bulk structure than the surface without adsorbed water.

H-bonds play a significant role in determining the calculated $\delta^{13}\text{C4}$ and $\delta^{13}\text{C6}$ values because NMR calculations on the molecular clusters listed in Table 4 without the H_2O molecules present (i.e., the “dry” models referred to above) change the model $\delta^{13}\text{C}$ values by up to 6 ppm without changing any atomic positions in the cellulose surface model. To complicate matters, however, these $\delta^{13}\text{C4}$ and $\delta^{13}\text{C6}$ values for the dry surfaces are downfield shifted (e.g., 80.3–86.3 for one C4 on $\text{I}\alpha$ (021)) rather than upfield shifted.

The possibility that C4 and C6 chemical shifts of hydroxymethyl groups in *tg* conformations may be responsible for observed peaks in the 83–84 and 62 ppm regions does not preclude the possibility of hydroxymethyl rotation into a *gt* conformation. For example, C4 atoms in model bulk cellulose with a *gt* conformation resulted in calculated $\delta^{13}\text{C}$ values of 82–84 ppm for the $\text{I}\alpha$ form and 86 ppm for $\text{I}\beta$ (Kubicki et al. 2013). The C6 $\delta^{13}\text{C}$ values were from 59 to 61 and 63 ppm for $\text{I}\alpha$ and $\text{I}\beta$, respectively. Hence, both the non-H-bonded *tg* and the normal *gt* conformations predict $\delta^{13}\text{C}$ values similar to the observed values for C4 and C6 atoms. Resolving this issue will require calculations on surfaces with *gt* conformations and comparison of the results to observed ^{13}C NMR and vibrational spectra.

$\delta^{13}\text{C4}$ could also be influenced by the Φ and Ψ torsional angles defining glycosidic linkage conformations (Horii et al. 1983; Jarvis 1994; Newman and Davidson 2004). Therefore, the model Φ and Ψ values for each surface C4 are included in Table 4. The experimental Φ and Ψ values are -98° to -99° and -138° to -140° for $\text{I}\alpha$ and -89° to -99° and -142° to -147° for $\text{I}\beta$, respectively (Nishiyama et al. 2002, 2003). These bulk torsion angles calculated using the same methodology are -93° to -95° and -141° to -143° for $\text{I}\alpha$ and -93° to -94° and -143° to -145° for $\text{I}\beta$, respectively (Kubicki et al. 2013). Examining the Φ and Ψ values of the three C atoms with $\delta^{13}\text{C}$ values matching the surface C4 atoms observed in Harris et al. (2012), one can see that the glycosidic torsions are not significantly different from the bulk values (Table 4). For these particular atoms, Φ and Ψ values range from -93° to -96° and -139° to -147° , respectively (Table 4). Changes in the model glycosidic

torsions are not responsible for the observed 4–5 ppm upfield NMR chemical shift of the C4 atoms. This is not to say, however, that the glycosidic torsion angles near the surface do not distort from the bulk. The model Φ and Ψ values range from -90° to -99° and -139° to -148° , respectively, for a range of 10° compared to the 2 range calculated for bulk cellulose.

Discussion

One of the advantages of molecular modeling techniques is that they allow one to manipulate the system of interest to isolate particular components that may not be studied separately via experimental means. In this case, our calculations focus on individual surfaces that have not been isolated in the laboratory. This ability is important because water, PCW biopolymers and cellulose binding modules (Lehtiö et al. 2003) may have preferences for specific surfaces. A disadvantage of any modeling is that the model results may be inaccurate or unrealistic. The fact that these DFT-D2 calculations have been capable of reproducing the correct energetics of $\text{I}\alpha$ versus $\text{I}\beta$ cellulose as well as the IR/Raman and NMR spectroscopic signatures of bulk cellulose (Kubicki et al. 2013) lends confidence to the predictions made for cellulose surface-water interactions with the same methods. The prediction of the small negative ΔE_{ads} values for cellulose hydrophilic surfaces and small positive ΔE_{ads} values for hydrophobic surfaces has implications for the formation of cellulose microfibrils as well their interactions with other PCW components. For example, hemicellulose and lignin may prefer to adsorb onto different surfaces based on their relative hydrophilicity/hydrophobicity (Hanus and Mazeau 2006; Petridis et al. 2011).

The prediction that relatively small rotations of the hydroxymethyl groups could give rise to the observed $\delta^{13}\text{C}$ chemical shifts of the surface C4 is in contrast with interpretations of previous experimental studies (Davidson et al. 2004; Šturcová et al. 2004). Interpretation of the ^{13}C NMR and vibrational spectra is key to understanding what is meant by “disordered” cellulose which is often interpreted as “water accessible” cellulose where deuteration occurs not only at the surface but in the interior as well (Nishiyama et al. 2003). Because the surface $\delta^{13}\text{C4}$ values are similar to carbohydrates with *gt* conformations of the hydroxymethyl group (Horii

et al. 1983, 1984), the presence of C4 peaks near 84 ppm and C6 peaks between 62.5 and 64.5 ppm have been interpreted as *gt* conformations on cellulose surfaces (Newman and Davidson 2004).

The current DFT results present an alternative explanation of the observed $\delta^{13}\text{C}$ values. The observed 4–5 ppm decrease in $\delta^{13}\text{C}_4$ of the surface atoms is consistent with the DFT NMR calculations showing particular $\delta^{13}\text{C}_4$ near 86 ppm. The associated $\delta^{13}\text{C}_6$ values (Table 4) are also consistent with the upfield shifted $\delta^{13}\text{C}_6$ (Harris et al. 2012). The fact that in each case these C4 atoms with lower calculated $\delta^{13}\text{C}$ values are not involved with H-bonding to H_2O or other cellulose polymer chains may be the reason for the observed surface $\delta^{13}\text{C}_4$ peaks at 84–86 ppm and $\delta^{13}\text{C}_6$ peaks near 62 ppm. Calculation of the $\delta^{13}\text{C}_4$, $\delta^{13}\text{C}_6$ and the vibrational spectra of cellulose surfaces with the *gt* and *gg* conformations will be a good way to help determine whether or not it is necessary to have hydroxymethyl torsions on cellulose surfaces in order to generate the observed spectra.

Acknowledgments This work was supported by the U.S. Department of Energy grant for the Energy Frontier Research Center in Lignocellulose Structure and Formation (CLSF) from the Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001090. The authors also thank Yoshiharu Nishiyama for suggesting DFT-D2 calculations as a methodology for modeling cellulose. We also acknowledge discussions with Roger Newman and Mike Jarvis as well as numerous CLSF participants that improved the manuscript. Computational support was provided by the Research Computation and Cyberinfrastructure group at The Pennsylvania State University.

References

- Accelrys Inc. (2012) Materials Studio 5.5. San Diego, CA
- Adamo C, Barone V, Introduction I (1998) Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: the mPW and mPW1PW models. *J Chem Phys* 108:664–675
- Albersheim P, Darvill A, Roberts K, Sederoff R, Staehelin A (2011) Biochemistry of the cell wall molecules. In: Plant cell walls: from chemistry to biology. Garland Science, Taylor & Francis Group, LLC, NY, pp 67–118
- Alecu IM, Zheng J, Zhao Y, Truhlar DG (2010) Computational thermochemistry: scale factor databases and scale factors for vibrational frequencies obtained from electronic model chemistries. *J Chem Theory Comput* 6:2872–2887
- Blackwell J (1977) Infrared and Raman spectroscopy of cellulose. In: Aurther J (ed) Cellulose chemistry and technology, ACS symposium series. American Chemical Society, Washington, DC, pp 206–218
- Brizuela AB, Bichara LC, Romano E, Yurquina A, Locatelli S, Brandán SA (2012) A complete characterization of the vibrational spectra of sucrose. *Carbohydr Res* 361:212–218
- Bučko T, Tunega D, Ángyán JG, Hafner J (2011) Ab initio study of structure and interconversion of native cellulose phases. *J Phys Chem A* 115:10097–10105
- Buhl M, Kaupp M, Malkina OL, Malkin VG (1999) The DFT route to NMR chemical shifts. *J Comput Chem* 20: 91–105
- Chai J-D, Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys Chem Chem Phys* 10:6615–6620
- Cheeseman JR, Trucks GW, Keith TA, Frisch MJ (1996) A comparison of models for calculating nuclear magnetic resonance shielding tensors. *J Chem Phys* 104:5497–5509
- Cirtog M, Alikhani ME, Madebène B, Soulard P, Asselin P, Tremblay B (2011) Bonding nature and vibrational signatures of oxirane:(water)_n = 1–3. Assessment of the performance of the dispersion-corrected DFT methods compared to the ab initio results and Fourier transform infrared experimental data. *J Phys Chem A* 115:6688–6701
- Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PVR (1983) Efficient diffuse function-augmented basis sets for anion calculations. III. The 3-21+ G basis set for first-row elements, Li–F. *J Comput Chem* 4:294–301
- Cremer D, Pople JA (1975) A general definition of ring puckering coordinates. *J Am Chem Soc* 97:1354–1358
- Davidson TC, Newman RH, Ryan MJ (2004) Variations in the fibre repeat between samples of cellulose I from different sources. *Carbohydr Res* 339:2889–2893
- Dick-Pérez M, Zhang Y, Hayes J, Salazar A, Zabolina OA, Hong M (2011) Structure and interactions of plant cell-wall polysaccharides by two- and three-dimensional magic-angle-spinning solid-state NMR. *Biochemistry* 50: 989–1000
- Eck B (2012) wxDragon 1.8.0-reg. Copyright 1994–2012 mbB
- Erata T, Shikano T, Yunoki S, Takai M (1997) The complete assignment of the ^{13}C CP/MAS NMR spectrum of native cellulose by using ^{13}C labeled glucose. *Cellul Commun* 4:128–131
- Fekri N, Khayami M, Heidari R, Jamee R (2008) Chemical analysis of flax seed, sweet basil, dragon head and quince seed mucilages. *Res J Biol Sci* 3:166–170
- Fernandes AN, Thomas LH, Altaner CM, Callow P, Forsyth VT, Apperly DC, Kennedy CJ, Jarvis MC (2011) Nanostructure of cellulose microfibrils in spruce wood. *Proc Natl Acad Sci USA* 108:E1195–E1203
- French AD, Csonka GI (2011) Hydroxyl orientations in cellobiose and other polyhydroxyl compounds: modeling versus experiment. *Cellulose* 18(4):897–909
- French AD, Johnson GP, Cramer CJ, Csonka GI (2012) Conformational analysis of cellobiose by electronic structure theories. *Carbohydr Res* 350:68–76
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery Jr JA, Vreven T, Kudin KN, Burant JC et al. (2009) Gaussian 09 Revision B.01. Wallingford, CT
- Fubini B, Zanetti G, Altília S, Tiozzo R, Lison D, Saffiotti U (1999) Relationship between surface properties and cellular responses to crystalline silica: studies with heat-treated cristobalite. *Chem Res Toxicol* 12:737–745

- Gazit OM, Katz A (2013) Understanding the role of defect sites in glucan hydrolysis on surfaces. *J Am Chem Soc* 135: 4398–4402
- Gonzalez-Outeiriño J, Kirschner KN, Thobhani S, Woods RJ (2006) Reconciling solvent effects on rotamer populations in carbohydrates—A joint MD and NMR analysis. *Can J Chem* 84:569–579
- Gottlieb HE, Kotlyar V, Nudelman A (1997) NMR chemical shifts of common laboratory solvents as trace impurities. *J Org Chem* 62:7512–7515
- Grimme S (2006) Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J Comput Chem* 27:1787–1799
- Guvench O, Hatcher ER, Venable RM, Pastor RW, MacKerell AD Jr (2009) CHARMM additive all-atom force field for glycosidic linkages. *J Chem Theory Comput* 5:2353–2370
- Hanus J, Mazeau K (2006) The xyloglucan—cellulose assembly at the atomic scale. *Biopolymers* 82:59–73
- Harris DM, Corbin K, Wang T, Gutierrez R, Bertolo AL, Carloalberto P, Smilgies D-M, Estevez JM, Bonetta D, Urbanowicz BR et al (2012) Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A903V and CESA3T942I of cellulose synthase. *Proc Natl Acad Sci USA* 109:4098–4103
- Heiner AP, Teleman O (1997) Interface between monoclinic crystalline cellulose and water: breakdown of the odd/even duplicity. *Langmuir* 13:511–518
- Heiner AP, Kuutti L, Teleman O (1998) Comparison of the interface between water and four surfaces of native crystalline cellulose by molecular dynamics simulations. *Carbohydr Res* 306:205–220
- Hiejima Y, Yao M (2004) Phase behaviour of water confined in Vycor glass at high temperatures and pressures. *J Phys: Condens Matter* 16:7903–7908
- Horii F, Hirai A, Kitamaru R (1983) Solid-state ^{13}C -NMR study of conformations of the oligosaccharide and cellulose conformation of the CH_2OH group about the exo-cyclic C–C bond. *Polym Bull* 10:357–361
- Horii F, Hirai A, Kitamaru R (1984) CP-MAS C-13 NMR study of spin relaxation phenomena of cellulose containing crystalline and noncrystalline components. *J Carbohydr Chem* 3:641–662
- Horikawa Y, Itoh T, Sugiyama J (2006) Preferential uniplanar orientation of cellulose microfibrils reinvestigated by the FTIR technique. *Cellulose* 13:309–316
- Iijima M, Morita S, Barlow PW (2008) Structure and function of the root cap. *Plant Prod Sci* 11:17–27
- Ireta J, Neugebauer J, Scheffler M (2004) On the accuracy of DFT for describing hydrogen bonds: dependence on the bond directionality. *J Phys Chem A* 108:5692–5698
- Jarvis MC (1994) Relationship of chemical shift to glycosidic conformation in the solid-state ^{13}C NMR spectra of (1 \rightarrow 4)-linked glucose polymers and oligomers: anomeric and related effects. *Carbohydr Res* 259:311–318
- Jarvis MC (2011) Plant cell walls: supramolecular assemblies. *Food Hydrocolloids* 25:257–262
- Jeffrey GA (1997) An introduction to hydrogen bonding. Oxford University Press, New York
- Kalutskaya EP, Gusev SS (1981) An infrared spectroscopic investigation of the hydration of cellulose. *Polym Sci USSR* 22(3):550–556
- Karadakov PB (2006) Ab Initio Calculation of NMR Shielding Constants. In: Webb GA (ed) Modern magnetic resonance. Springer, Netherlands, pp 63–70
- Kirschner KN, Woods RJ (2001) Solvent interactions determine carbohydrate conformation. *Proc Natl Acad Sci USA* 98:10541–10545
- Kresse G, Furthmüller J (1996) Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys Rev B* 54:11169–11186
- Kresse G, Hafner J (1993) Ab initio molecular dynamics for open-shell transition metals. *Phys Rev B* 48:13115–13118
- Kresse G, Hafner J (1994) Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium. *Phys Rev B* 49:14251–14269
- Kresse G, Furthmüller J, Hafner J (1994) Theory of the crystal structures of selenium and tellurium: the effect of generalized-gradient corrections to the local-density approximation. *Phys Rev B* 50:13181–13185
- Krishnan R, Brinkley JS, Seeger R, Pople JA (1980) Self-consistent molecular orbital methods. XX. A basis set for correlated wave functions. *J Chem Phys* 72:650–654
- Kubicki JD, Mohamed MN-A, Watts HD (2013) Quantum mechanical modeling of the structures, energetics and spectral properties of I α and I β cellulose. *Cellulose* 20:9–23
- Lee CM, Mohamed MA, Watts HD, Kubicki JD, Kim S (2013) Sum-frequency-generation (SFG) vibration spectra and density functional theory calculations with dispersion corrections (DFT-D2) for cellulose I α ; and I β . *J Phys Chem B* 117:6681–6692
- Lehtiö J, Sugiyama J, Gustavsson M, Fransson L, Linder M, Teeri TT (2003) The binding specificity and affinity determinants of family 1 and family 3 cellulose binding modules. *Proc Natl Acad Sci USA* 100:484–489
- Li Y, Lin M, Davenport JW (2011) Ab initio studies of cellulose I: crystal structure, intermolecular forces, and interactions with water. *J Phys Chem* 115:11533–11539
- Lindberg B, Mosihuzzaman M, Nahar N, Abeysekera RM, Brown RG, Willison JHM (1990) An unusual (4-O-methyl-D-glucurono)-D-xylan isolated from the mucilage of seeds of the quince tree (*Cydonia oblonga*). *Carbohydr Res* 207:307–310
- Liu Y, Gamble G, Thibodeaux D (2010) Two-dimensional attenuated total reflection infrared correlation spectroscopy study of the desorption process of water-soaked cotton fibers. *Appl Spectrosc* 64:1355–1363
- Lodewyk MW, Siebert MR, Tantillo DJ (2012) Computational prediction of ^1H and ^{13}C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. *Chem Rev* 112:1839–1862
- Malm E, Bulone V, Wickholm K, Larsson P, Iversen T (2010) The surface structure of well-ordered native cellulose fibrils in contact with water. *Carbohydr Res* 345:97–100
- Mann J, Marrinan HJ (1956) The reaction between cellulose and heavy water. *Trans Faraday Soc* 52:481–487
- Maréchal Y, Chanzy H (2000) The hydrogen bond network in I β cellulose as observed by infrared spectrometry. *J Mol Struct* 523:183–196
- Matthews JF, Skopec CE, Mason PE, Zuccato P, Torget RW, Sugiyama J, Himmel ME, Brady JW (2006) Computer simulation studies of microcrystalline cellulose I β . *Carbohydr Res* 341:138–152

- Matthews JF, Bergenstr hle M, Beckham GT, Himmel ME, Nimlos MR, Brady JW, Crowley MF (2011) High-temperature behavior of cellulose I. *J Phys Chem B* 115: 2155–2166
- Matthews JF, Beckham GT, Bergenstr hle-Wohlert M, Brady JW, Himmel ME, Crowley MF (2012) Comparison of cellulose I β simulations with three carbohydrate force fields. *J Chem Theory Comput* 8:735–748
- Nakashima K, Sugiyama J, Satoh N (2008) A spectroscopic assessment of cellulose and the molecular mechanisms of cellulose biosynthesis in the ascidian *Ciona intestinalis*. *Mar Genom* 1:9–14
- Naran R, Chen G, Carpita NC (2008) Novel rhamnogalacturonan I and arabinoxylan polysaccharides of flax seed mucilage. *Plant Physiol* 148:132–141
- Newman RH, Davidson TC (2004) Molecular conformations at the cellulose–water interface. *Cellulose* 11:23–32
- Nishiyama Y, Langan P, Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose I β from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc* 124:9074–9082
- Nishiyama Y, Sugiyama J, Chanzy H, Langan P (2003) Crystal structure and hydrogen bonding system in cellulose I α from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc* 125:14300–14306
- Nishiyama Y, Johnson GP, French AD, Forsyth VT, Langan P (2008) Neutron crystallography, molecular dynamics, and quantum mechanics studies of the nature of hydrogen bonding in cellulose I β . *Biomacromolecules* 9:3133–3140
- O'Dell WB, Baker DC, McLain SE (2012) Structural evidence for inter-residue hydrogen bonding observed for cellobiose in aqueous solution. *PLoS ONE* 7:e45311
- Paterson MS (1982) The determination of hydroxyl by infrared absorption in quartz, silicate glasses and similar minerals. *Bull Min r* 105:20–29
- Petridis L, Pingali S, Urban V, Heller W, O'Neill H, Foston M, Ragauskas A, Smith J (2011) Self-similar multiscale structure of lignin revealed by neutron scattering and molecular dynamics simulation. *Phys Rev E* 83:4–7
- Radloff D, Boeffel C, Spiess HW (1996) Cellulose and cellulose/poly (vinyl alcohol) blends. 2. Water organization revealed by solid-state NMR spectroscopy. *Macromolecules* 29(5):1528–1534
- Rapp  AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J Am Chem Soc* 114:10024–10035
- Rassolov VA, Ratner MA, Pople JA, Redfern PC, Curtiss LA (2001) 6-31G* basis set for third-row atoms. *J Comput Chem* 22:976–984
- Sarotti AM, Pellegrinet SC (2009) A multi-standard approach for GIAO ^{13}C NMR calculations. *J Organ Chem* 74: 7254–7260
- Schaftenaar G, Noordik JH (2000) Molden: a pre- and post-processing program for molecular and electronic structures. *J Comput Aided Mol Des* 14:123–134
- Schreckenbach G, Ziegler T (1995) Calculation of NMR shielding tensors using gauge-including atomic orbitals and modern density functional theory. *J Phys Chem* 99: 606–611
- Skinner JL, Pieniazek PA, Gruenbaum SM (2012) Vibrational spectroscopy of water at interfaces. *Acc Chem Res* 45:93–100
- Sternberg U, Koch F, Prie  W, Witter R (2003) Crystal structure refinements of cellulose polymorphs using solid state ^{13}C chemical shifts. *Cellulose* 10:189–199
-  sturcov  S, His I, Apperley DC, Sugiyama J, Jarvis MC (2004) Structural details of crystalline cellulose from higher plants. *Biomacromolecules* 5:1333–1339
- Thomas LH, Forsyth VT, Sturcov  A, Kennedy CJ, May RP, Altaner CM, Apperley DC, Wess TJ, Jarvis MC (2013) Structure of cellulose microfibrils in primary cell walls from collenchyma. *Plant Physiol* 161:465–476
- Watts HD, Mohamed MNA, Kubicki JD (2011) Comparison of multistandard and TMS-standard calculated NMR shifts for coniferyl alcohol and application of the multistandard method to lignin dimers. *J Phys Chem B* 115:1958–1970
- Wickholm K, Larsson PT, Iversen T (1998) Assignment of non-crystalline forms in cellulose I by CP/MAS ^{13}C NMR spectroscopy. *Carbohydr Res* 312:123–129
- Wiitala KW, Hoyer TR, Cramer CJ (2006) Hybrid density functional methods empirically optimized for the computation of ^{13}C and ^1H chemical shifts in chloroform solution. *J Chem Theory Comput* 2(4):1085–1092
- Wiley JH, Atalla RH (1987) Band assignments in the Raman spectra of celluloses. *Carbohydr Res* 160:113–129
- Witter R, Sternberg U, Hesse S, Kondo T, Koch F-T, Ulrich AS (2006) ^{13}C chemical shift constrained crystal structure refinement of cellulose I α and its verification by NMR anisotropy experiments. *Macromolecules* 38:6125–6132
- Wolinski K, Hinton JF, Pulay P (1990) Efficient implementation of the gauge-independent atomic orbital method for NMR chemical shift calculations. *J Am Chem Soc* 112: 8251–8260
- Zhang C, Lindan PJD (2003) Towards a first-principles picture of the oxide–water interface. *J Chem Phys* 119:9183–9190
- Zhao Y, Schultz NE, Truhlar DG (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J Chem Theory Comput* 2:364–382