Sum-Frequency-Generation Vibration Spectroscopy and Density Functional Theory Calculations with Dispersion Corrections (DFT-D2) for Cellulose Iα and Iβ

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JUNE 2013

Impact Factor: 3.3 · DOI: 10.1021/jp402998s · Source: PubMed

5 AUTHORS, INCLUDING:

Christopher M Lee
Pennsylvania State University
31 PUBLICATIONS **265** CITATIONS
SEE PROFILE

Mohamed Naseer ali Mohamed
The New College
15 PUBLICATIONS **149** CITATIONS
SEE PROFILE

Heath Watts
Pennsylvania State University
14 PUBLICATIONS **98** CITATIONS
SEE PROFILE

J. D. Kubicki
University of Texas at El Paso
217 PUBLICATIONS **4,438** CITATIONS
SEE PROFILE

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

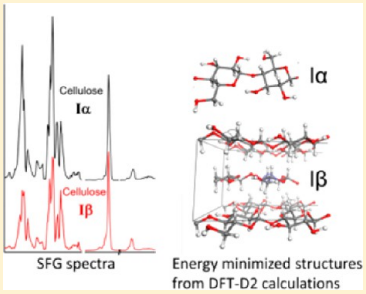
Sum-Frequency-Generation Vibration Spectroscopy and Density Functional Theory Calculations with Dispersion Corrections (DFT-D2) for Cellulose Iα and Iβ

Christopher M. Lee,[†] Naseer M. A. Mohamed,[‡] Heath D. Watts,[‡] James D. Kubicki,[‡] and Seong H. Kim*[,†,§]

[†]Department of Chemical Engineering and Materials Research Institute, and [‡]Department of Geosciences, The Pennsylvania State University, University Park, Pennsylvania 16802, United States

Supporting Information

ABSTRACT: Sum-frequency-generation (SFG) vibration spectroscopy selectively detects noncentrosymmetric vibrational modes in crystalline cellulose inside intact lignocellulose. However, SFG peak assignment in biomass samples is challenging due to the complexity of the SFG processes and the lack of reference SFG spectra from the two crystal forms synthesized in nature, cellulose Iα and Iβ. This paper compares SFG spectra of laterally aligned cellulose Iα and Iβ crystals with vibration frequencies calculated from density functional theory with dispersion corrections (DFT-D2). Two possible hydrogen-bond networks A and B (Nishiyama et al. *Biomacromolecules* **2008**, *9*, 3133) were investigated for both polymorphs. From DFT-D2 calculations the energetically favorable structures for cellulose Iα and Iβ had CH_2OH groups in tg conformations and network A hydrogen bonding. The calculated frequencies of C−H stretch modes agreed reasonably well with the peak positions observed with SFG and were localized vibrations; thus, peak assignments to specific alkyl groups were proposed. DFT-D2 calculations underestimated the distances between hydrogen-bonded oxygen atoms compared to the experimentally determined values; therefore, the OH stretching calculated frequencies were ∼100 cm^{-1} lower than observed. The SFG peak assignments through comparison with DFT-D2 calculations will guide the SFG analysis of the crystalline cellulose structure in plant cell walls and lignocellulose biomass.



INTRODUCTION

Vibration spectroscopy of cellulose has been an important tool in the biological sciences and also sustainable engineering to understanding cellulose structure.[1] Plants use much of the carbon and solar energy captured during photosynthesis to produce cell walls, which are critical for proper growth and development.[2] Plant biomass is processed for use as lumber, textiles, papers, and additives in foods, paints, and films.[3] Recently, global efforts are underway to convert plant cell walls (which are collectively called lignocellulosic biomass) to biofuels for transportation needs.[4−6] Cellulose is the most abundant and important constituent in plant cell walls. Its crystal structure and complex network with other carbohydrate polymers are key factors determining the mechanical strength and degradability of plant cell walls.[7,8] A nondestructive spectroscopic technique that can selectively detect crystalline cellulose imbedded in amorphous carbohydrate polymers is necessary to advance our understanding of the roles of crystalline cellulose in the chemical and physical properties of plant cell walls.

Cellulose is a linear polymer chain of β-D-glucopyranose monomers connected by 1,4 glycosidic linkages. In plants, chains are synthesized by cellulose synthase complexes (CSCs) present at the plasma membrane.[9,10] Shortly after polymerization, several chains crystallize into elementary microfibrils and are deposited into the cell wall.[11] There are two crystal structures that are produced in nature: cellulose Iα, predominantly found in algae and bacterial biofilms and cellulose Iβ, predominantly found in land plants and tunicates.[12] In cellulose microfibrils, the six-membered glucan unit has the arm-chair conformation and the hydroxyl (O2H, O3H) and exocyclic hydroxymethylene ($C6H_2OH$) groups at the equatorial positions, and the methine (CH) functional groups on carbons 1, 2, 3, 4, and 5 at the axial positions (Figure 1). Three conformations of the $C6H_2OH$ groups are possible through the rotation of the C5−C6 bond, the tg conformation is found to be the dominant form in cellulose Iα and Iβ as shown in Figure 1.[13,14] Cellulose chains are linear and assembled side by side into a sheet through networks of intrachain hydrogen bonds (O3H···O5; O2H···O6) and interchain hydrogen bonds (O6H···O2′; O2H···O6′; O6H···O3′) (Figure 2). There are two possible patterns of hydrogen bonding within a sheet which are called networks A and B.[15] In network A (Figure 2a), the hydrogen of O2 is involved in intrachain hydrogen bond to O6 (O2H···O6) and the hydrogen of O6 is involved in interchain hydrogen bonds

Received: March 26, 2013
Revised: May 8, 2013

   dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

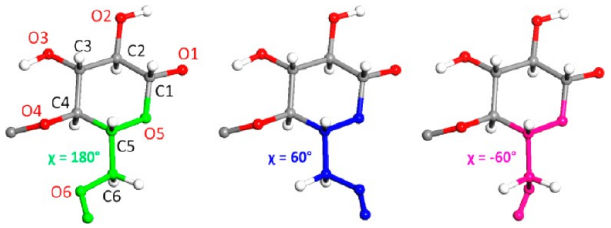


Figure 1. Molecular model of the exocyclic $C6H_2OH$ group in tg (green), gt (blue), and gg (pink) conformations. In the pairwise representation (t = trans and g = gauche), the first letter is the torsion angle of O5−C5−C6−O6 and the second letter the angle of C4−C5−C6−O6. χ is the angle between the C5−O5 and C6−O6 bond axes (C = dark gray; O = red; H = light gray).

(O6H···O3′).[15] In network B (Figure 2b), the hydrogen of O2 is involved in interchain hydrogen bond to O6 (O2H···O6′) and the hydrogen of O6 is involved in intrachain hydrogen bond to O2 (O6H···O2).[15]

These sheets are stacked through van der Waals interactions between the hydrophobic facets of the glucan units to form three-dimensional crystals of cellulose Iα and cellulose Iβ.[13,14] Sheet stacking influences the covalent bond lengths as well as the torsion angles of the glycosidic bonds and the hydroxymethylene groups in cellulose Iα and Iβ.[16] In the cellulose Iα triclinic unit cell, all sheets are identical and the chains in each sheet contain alternating units with slightly different CH_2OH conformations (χ = 166° and 167°, Figure 3b).[14] In the cellulose Iβ monoclinic unit cell, chains in each sheet are identical, but the CH_2OH conformations of the glucose units in two adjacent sheets are slightly different (χ = 158° and 170°, Figure 3c).[13]

The structural models described in Figure 3 are constructed based on X-ray and neutron diffraction studies of aligned nanowhiskers (cylinder-shape nanocrystals) of cellulose Iα isolated from the green alga *Glaucocystis nostochinearum* and cellulose Iβ isolated from the tunicate *Halocynthia roretzi*.[13,14] These species are unique in that they produce large crystals (~20 nm) of a single allomorph. Cellulose crystals in cell walls of vascular plants contain varying ratios of Iα and Iβ forms and are thought to be more disordered than cellulose from *Halocynthia* and *Glaucocystis*.[17] In addition, cellulose in vascular plant cell walls are smaller (~2−5 nm) and interact with hemicellulose, pectin, and lignin components in the cell wall.[11,18]

Determining the native cellulose polymorphism,[19,20] crystallinity,[21] hydrogen bonding,[22−24] and mechanical properties[25] is commonly performed using IR and Raman vibration spectroscopy as well as X-ray diffraction (XRD) and solid-state nuclear magnetic resonance (NMR).[26−30] These techniques are useful for structural analysis of isolated or purified cellulose; but they suffer from spectral interferences from other matrix polysaccharide components when applied to the structural study of cellulose within intact plant cell walls.[31−33] Because crystalline cellulose is more resistant to hydrolysis than amorphous polysaccharides, cellulose can be isolated from plant cell walls. However, this process can alter the crystal structure of native cellulose. In order to study crystalline cellulose structure in situ without separation of cellulose from other matrix components, it is desirable to use a spectroscopic technique that is highly selective to cellulose only.

Recently, sum-frequency-generation (SFG) vibration spectroscopy has been shown to selectively detect crystalline cellulose in plant cell walls and lignocellulosic biomass without interferences from amorphous components such as hemicellulose, pectin, and lignin.[33,34] SFG is a nonlinear optical process in which visible (ω_{VIS}) and infrared (ω_{IR}) photons are combined into a single photon whose frequency is the sum of the two input frequencies, $\omega_{SFG} = \omega_{VIS} + \omega_{IR}$. The output SFG signal $I(\omega_{SFG})$ is expressed as[35−37]

$$I(\omega_{SFG}) \propto |\chi_{eff}^{(2)}|^2 I(\omega_{VIS}) I(\omega_{IR}) \quad (1)$$

where $I(\omega_{VIS})$ and $I(\omega_{IR})$ are the intensities of the visible and infrared laser beams and $\chi_{eff}^{(2)}$ is the effective nonlinear susceptibility:[35]

$$\chi_{eff}^{(2)} = \frac{N \sum_{\alpha,\beta,\gamma} \langle M_{\alpha\beta} A_\gamma \rangle}{\varepsilon_0 (\omega_{IR} - \omega_q - i\Gamma)} \quad (2)$$

Here, N is the number density, ε_0 is the dielectric constant of vacuum, $\langle M_{\alpha\beta} A_\gamma \rangle$ is the angle-average of the product of the Raman and infrared tensors, ω_q is the frequency of a normal vibration mode, and Γ is the damping constant. In the SFG analysis with a tunable IR beam, signals are resonantly

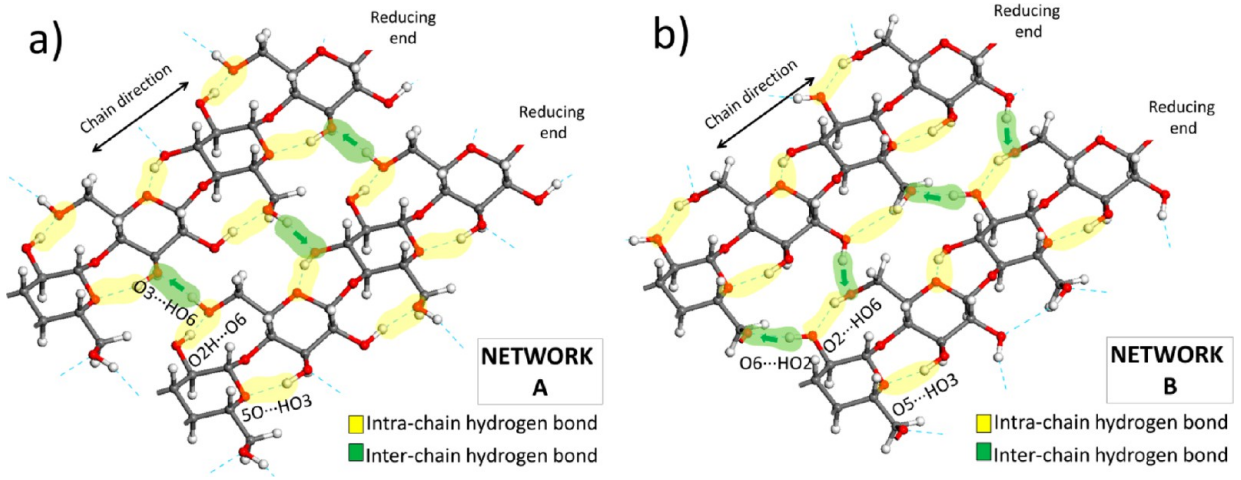


Figure 2. Two possible hydrogen-bonding networks in cellulose Iα and Iβ crystals: (a) network A and (b) network B[15] (C = dark gray; O = red; H = light gray).

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

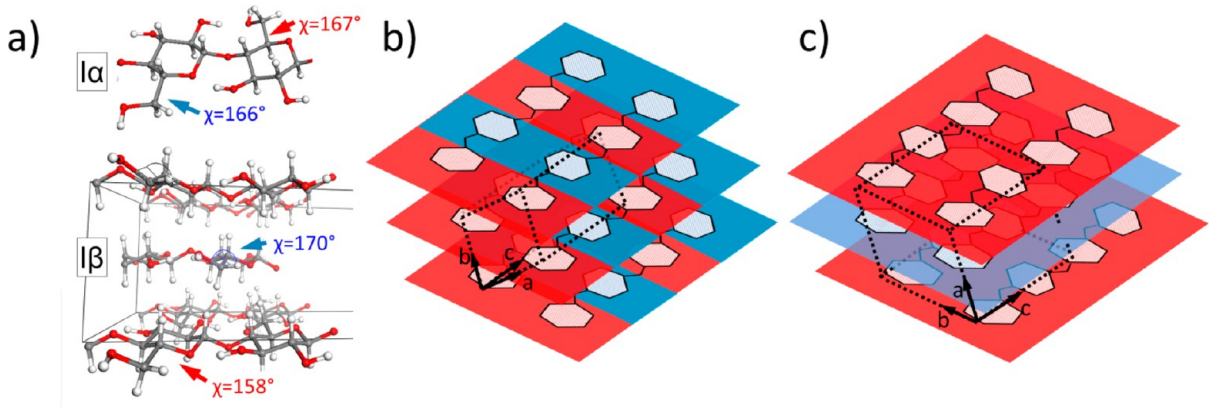


Figure 3. Two conformations of the exocyclic CH_2OH groups in cellulose Iα and Iβ are shown in (a). The χ angles are marked in red and blue colors. Stacking of cellulose chains in (b) cellulose Iα and (c) cellulose Iβ. The dotted boxes are the triclinic unit cell containing a single chain in (b) and the monoclinic unit cell containing two chains in (c). The unit cell axes are marked following the notation from Sugiyama et al.[85] The c-axis in (b) is reversed for clarity.

enhanced as the frequency of the input IR beam (ω_{IR}) approaches ω_q in the sample. Equation 2 also states that the vibration mode must be both IR- and Raman-active. The third-order tensor, $\chi_{eff}^{(2)}$, dictates that the SFG process takes place only when the irradiated medium has no inversion symmetry, otherwise $\chi_{eff}^{(2)}$ is zero.[38] Only certain vibrational modes in crystalline cellulose within plant cell walls meet this non-centrosymmetry.[39] In contrast, other amorphous matrix polymers (hemicellulose and lignin) in plant cell walls do not produce SFG signal even though they are composed of monomers which contain several chiral centers.[33] Without net polar ordering or electronic resonance, the SFG signals from chiral centers are very weak using reflection geometry, which is chosen due to opaque plant cell wall samples.[40] For these reasons, SFG can selectively detect cellulose in lignocellulose biomass.

Peak assignment is challenging, however, because several factors affect the SFG signal generation from crystalline cellulose. The nonlinear susceptibilities of individual vibration modes depend on the polarizations of input and output photons. If a single crystal or molecular thin film of cellulose is available, one could conduct thorough polarization analyses to obtain the $\chi_{eff}^{(2)}$ tensors of various vibration modes.[41,42] Unfortunately, native cellulose chains crystallize into only a few nanometers wide microfibrils and often cellulose Iα or Iβ crystals are present together in polycrystalline samples[18] or are aggregated into bundles.[29,43] Moreover, cellulose is a birefringent material containing many chiral centers in the glucan monomeric unit.[44] Thus, the polarization of the light can be altered while light travels through the sample.[45] Furthermore, SFG intensity analysis is complicated because the phase matching condition is strongly influenced by the spatial distribution of cellulose crystals in the sample over the SFG coherence length.[46] Hieu et al. proposed the $\chi_{eff}^{(2)}$ tensors for cellulose in a single cotton fiber.[39] In this method, it was assumed that the cellulose crystalline axis coincides with the fiber axis in the laboratory coordinates. However, the cellulose microfibril orientation in cotton fibers is known to contain several layers, each layer with alternating spiral arrangements.[47] Thus, using SFG to analyze plant cell walls is complicated because the native cellulose structure could be influenced by the growth process.[11,48] In this situation, theoretical calculations of vibration modes and frequencies could provide valuable information needed for peak assignment of the SFG spectra.

There have been several simulation studies to calculate the vibration frequencies for cellulose to assist peak assignments. Earlier works have used a classical mechanics approach called normal-coordinate analysis to compare calculated vibration modes with the IR and Raman spectra of cellulose from the alga *Valonia ventricosa*.[49,50] Recently, ab initio calculations have been used to evaluate the crystal structures of native cellulose[15,51] and to predict the vibrational spectra using Hartree−Fock[52] and density functional theory (DFT).[53,54] Barsberg et al. used cellulose I single-chain models in DFT calculations to estimate peak positions and intensities. However, these results did not distinguish between Iα and Iβ because only the covalent intrachain interactions were considered. A more advanced method is DFT with dispersion corrections (DFT-D2) that includes an empirical correction for long-range interactions.[55] DFT-D2 has been recently applied to cellulose Iα and Iβ to calculate crystal structures, intermolecular forces, interactions with water, and allomorph conversion.[56,57]

In this study, we report the SFG spectra of highly crystalline and laterally packed cellulose Iα and Iβ nanowhiskers prepared from *Glaucocystis* and *Halocynthia*, respectively, and compare the SFG peaks with the fundamental vibration modes calculated from DFT-D2.[54] The crystal structures of cellulose Iα and Iβ were constructed from the experimentally determined atomic coordinates and optimized through energy minimization processes. The vibration energies obtained from DFT-D2 calculations were rescaled with the empirically known peak positions of specific localized vibration modes. The comparison of the DFT-D2 calculations with the SFG spectra made the peak assignments possible for the exocyclic methylene group in cellulose Iα and Iβ crystal and the hydrogen-bonding network in each allomorph. Being able to detect and identify the vibration modes of crystalline cellulose in plant cell walls and lignocellulosic biomass will greatly assist the structural study for the roles of crystalline cellulose in plant growth and mechanics as well as recalcitrance of biomass to degradation processes.

METHODS

Cellulose Sources. Films of highly aligned cellulose Iα and Iβ nanowhiskers, prepared by H_2SO_4 treatment of *Glaucocystis nostochinearum* and the tunicate *Halocynthia roretzi*, were

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

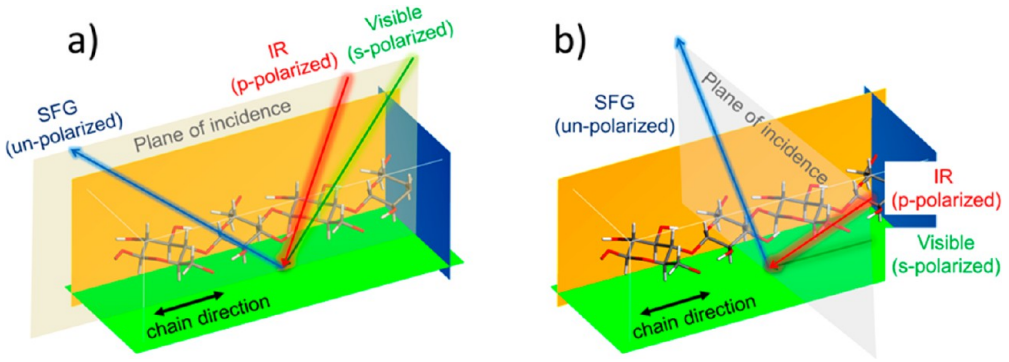


Figure 4. Two laser incidence plane geometries for aligned cellulose Iα and Iβ nanowhiskers. The nanowhisker alignment axis (black arrow) of the cellulose films was positioned (a) parallel (||) and (b) perpendicular ($\perp$) to the laser incidence plane (gray). The polarizations of IR and visible incidence beams were in-plane (p) and out-of-plane (s), respectively, with respect to the laser incidence plane.

generously provided by Yoshiharu Nishiyama, CERMAV–CNRS, France. The details of sample preparation and full characterization of these samples with X-ray and neutron diffraction were previously published.[13−15] In these samples, cellulose chains were roughly aligned with the unit cell *c*-axis along one direction by shearing of an aqueous suspension of nanowhiskers during the drying process.[58] The SEM images and 2D X-ray diffraction data of these samples are shown Figure S1 in the Supporting Information. Avicel PH-101 with ~50 μm particle size microcrystalline cellulose (FMC Biopolymer, CAS-No. 9004-34-6) was pressed into a pellet.

SFG Vibration Spectroscopy. The SFG setup (EKSPLA) used in our laboratory was described in previous studies.[33,34] The 532 nm laser pulse was generated by frequency doubling of the 1064 nm output from a Nd:YAG laser. An optical parameter generator/amplifier (OPG/OPA) using β-BaB_2O_4 and $AgGaS_2$ crystals was pumped with 532 and 1064 nm laser pulses and generated tunable IR pulses in the wavelength range of 2.3−10 μm with <6 cm^{-1} bandwidth. The incidence angles of visible and IR pulses with respect to surface normal were 60° and 56°, respectively. The incident visible and IR beams were s-polarized and p-polarized, respectively, with respect to the laser incidence plane. The SFG signal was collected in the reflection geometry with no polarization selection. A beam collimator was used to enhance the collection efficiency of the SFG signals. SFG spectra were taken at 4 cm^{-1}/step for the 1000−1500 and 2700−3050 cm^{-1} regions, and 8 cm^{-1}/step in the 3096−3800 cm^{-1} region, with averaging 100 shots/step. SFG spectra of cellulose films were taken with the plane of laser incidence parallel and perpendicular to the nanowhisker alignment axis, as shown in Figure 4. Spectral data were fitted with the Lorentzian function to find the peak center position.

Infrared Spectroscopy. IR spectra of cellulose Iα and Iβ films were obtained in a transmission mode using a Nexus 760 FT-IR spectrometer (Thermo Scientific). The sample was fixed normal to the incident beam which was polarized using a wire grid polarizer (Harrick Scientific Products). Spectra were taken with the IR polarization parallel and perpendicular to the nanowhisker alignment axis. All spectra were taken in the region 500−4000 cm^{-1} using 2 cm^{-1}/step and averaged over 100 scans.

Raman Spectroscopy. Cellulose Iα and Iβ films were analyzed with a Nicolet 8700 FT-Raman spectrometer (Thermo Scientific). This system was equipped with a near-infrared (1064 nm) Nd:YAG polarized excitation source, a CaF_2 beam splitter, and a germanium detector. Polarized incident light was normal to the sample and spectra were taken by rotating the sample so that the nanowhisker alignment axis was aligned parallel and perpendicular to the incident beam polarization. All spectra were taken in the region 250−3700 cm^{-1} using 8 cm^{-1}/step and averaged over 1000 scans.

DFT-D2 Calculations. Periodic DFT calculations were performed with the Vienna Ab-initio Simulation Package (VASP).[59−62] The initial crystal structures of cellulose Iα and Iβ were created using Materials Studio 5.5 (Accelrys Inc., San Diego, CA) based on the coordinates determined by Nishiyama et al.[13,14] The simulation cell size contained $C_{24}O_{20}H_{40}$ for cellulose Iα and $C_{48}O_{40}H_{80}$ for cellulose Iβ. These corresponded to two unit cells for cellulose Iα and Iβ, respectively. Cartesian coordinate files generated with Materials Studio 5.5 were converted into VASP 5.2 input files via a Perl script written by A. V. Bandura (St. Petersburg State University). Manipulations of the hydroxymethylene group torsions to three different conformations (tg, gg, and gt) and hydrogen-bonding networks A and B were performed manually in Materials Studio 5.5. Networks A and B were based on two possible hydrogen atom positions centered about 6O suggested by Nishiyama and co-workers.[15]

Projector-augmented planewave pseudopotentials were used with the PBE gradient-corrected exchange correlation function for the 3-D periodic DFT calculations. The choice of electron density and atomic structure optimization parameters was based on previous work.[56,57] An energy cutoff of 800 eV was used with an electronic energy convergence criterion of 1×10^{-7} eV. Atomic structures were relaxed until the energy gradient was less than 0.02 eV/Å at 2 × 2 × 2 k-point samplings. Atoms were first allowed to relax with the lattice parameters constrained to the experimental values, then the atoms and lattice parameters were allowed to relax to obtain reported structures, energies and spectroscopic properties. The dispersion-correction parameters were 40 Å for the cutoff distance[56] and 0.75 for the scaling factor (s_6) and 20 for the exponential coefficient (d) in the damping function.[55]

Frequency analyses were performed on the energy-minimized structures as predicted using VASP. 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 program wxDragon 1.8.0.[63] Representative movies are shown in the Supporting Information.

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

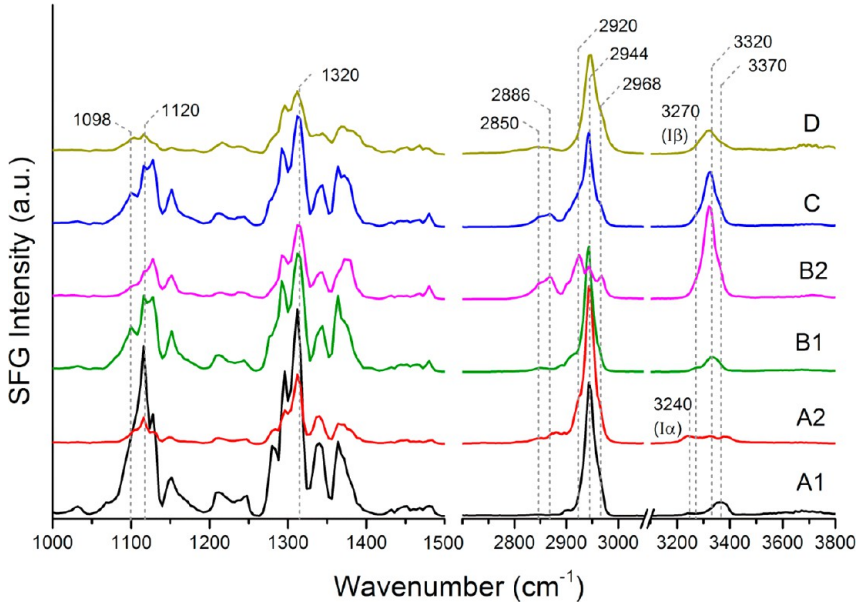


Figure 5. SFG spectra of laterally packed and aligned films of cellulose Iα nanowhiskers (from *Glaucocystis*; A1 and A2) and cellulose Iβ nanowhiskers (from *Halocynthia*; B1 and B2). (C) Linear combination of the B1 and B2 spectra. (D) SFG spectrum of randomly packed powder of cellulose Iβ (Avicel PH-101). Spectra A1 and B1 were taken with the laser incidence plane perpendicular to the chain alignment direction, and A2 and B2 were taken with the laser incidence plane parallel to the chain alignment direction. For the region 2700−3800 cm^{-1} the *y*-axis was enlarged 4 times for clarity. The peaks discussed in the text are highlighted with dotted lines.

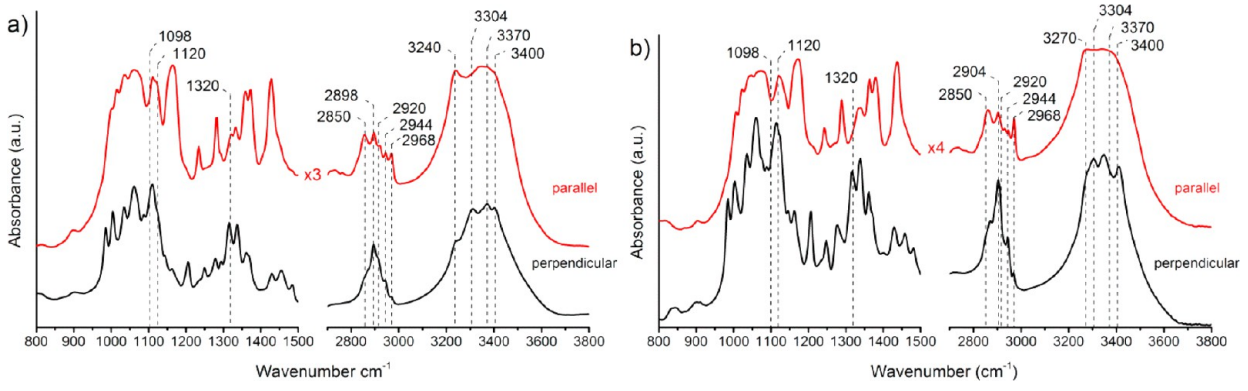


Figure 6. Polarized FT-IR spectra of laterally packed and aligned films of (a) cellulose Iα and (b) Iβ nanowhiskers. Spectra were taken with the IR polarization perpendicular and parallel to the nanowhisker alignment direction. The peaks discussed in the text are highlighted with dotted lines.

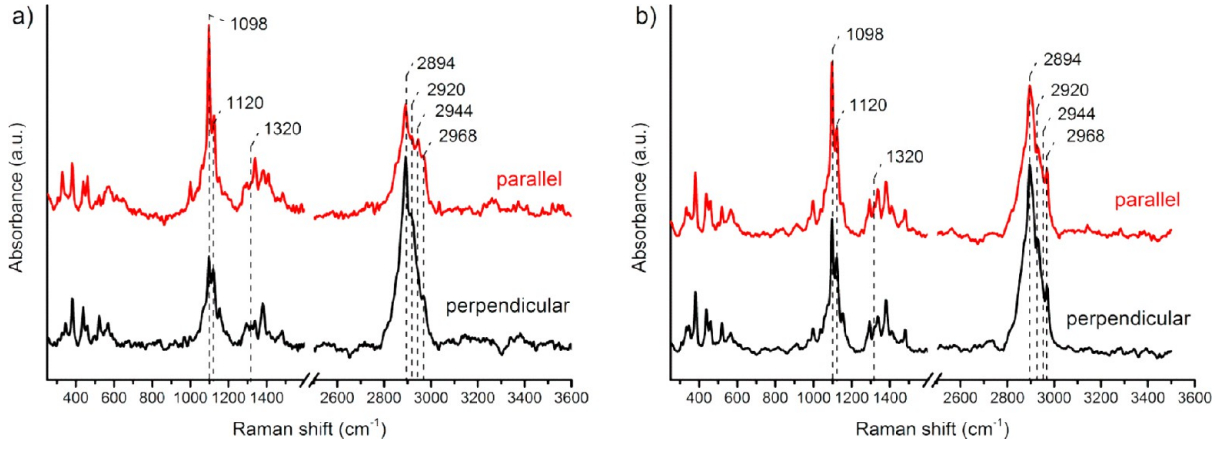


Figure 7. Polarized FT-Raman spectra of laterally packed and aligned films of (a) cellulose Iα and (b) Iβ nanowhiskers. Spectra were taken with the excitation laser polarization perpendicular and parallel to the nanowhisker alignment direction. The peaks discussed in the text are highlighted with dotted lines.

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

RESULTS AND DISCUSSION

I. SFG Spectra of Aligned Cellulose Iα and Iβ. The SFG spectra for the laterally packed and aligned films of cellulose Iα and Iβ nanowhiskers as well as randomly oriented powders of cellulose Iβ (Avicel PH-101) are shown in Figure 5. It is important to note that, although SFG responses can originate from both the bulk and surface species, the cellulose SFG signals are mostly from the crystalline phase. The surface species can easily be detected when the signals from the bulk phase and surface do not overlap. It is known that the OH groups at the cellulose crystal surface are readily converted to OD through OH/OD exchange with D_2O.[64] When the crystalline cellulose samples were treated with D_2O, the OD SFG signal was negligible, although IR showed very strong OD signals.[33,65]

In the case of the aligned samples (A and B spectra in Figure 5), nanowhiskers were aligned along the long axis without specific directionality of the reducing end.[58] Within each nanowhisker the cellulose chains are parallel; i.e., the reducing ends of the glucose units point the same direction. Since cellulose crystals are distributed anisotropically, the SFG spectra of the aligned nanowhisker samples showed some dependence of the polarization of the input IR and visible beams. For comparison, the polarized IR and Raman spectra of the aligned cellulose Iα and Iβ films are shown in Figures 6 and 7, respectively. The peak positions and polarization dependences of SFG, IR, and Raman peaks in the stretching and bending regions of the alkyl groups are shown in Table 11. For comparison, the Raman and IR peak assignments from the literature are also shown in Table 1.[50,66]

Table 1. SFG, Raman, and IR Peaks Observed for Cellulose Iα and Iβ in the Alkyl Stretching and Bending Regions and Empirical Assignments to Vibration Modes

			vibration mode		
SFG (cm^{-1}) (pol,[a] int[b])	Raman (cm^{-1}) (pol,[c] int[b])	IR (cm^{-1}) (pol,[c] int[b])	this work	Raman (Wiley and Atalla[50])	IR and Raman (Blackwell et al.[66])
2968 (∥, m)	2967 (∥, m)	2968 (∥, m)	CH_2 ν_{as}	CH_2 ν_{as}	CH ν
2944 (⊥, vs)	2944 (?, m)	2944 (⊥, m)	CH_2 ν_{as}	CH_2 ν_{as}	CH_2 ν_{as}
2920 (∥, m)	2919 (?,m)	2929 (∥, m)	CH_2 ν_{as}	–	CH ν
–[d]	2894 (⊥, vs)	2898–2904 (⊥, m)	CH ν	CH ν	CH ν
2886 (∥, w)	–[d]	2871 (∥, m)	CH_2 ν_s	CH_2 ν_s	CH ν
2850 (∥, w)	2848 (∥, w)	2855 (∥, m)	CH_2 ν_s	CH_2 ν_s	CH_2 ν_s
1484 (?, w)	1479 (⊥, vw)	1484 (⊥, w)	CH_2 δ	CH_2 δ	CH_2 δ
1468 (?, w)	–[d]	1456 (⊥, w)	CH_2 δ	–	–
1432 (∥, w)	–[d]	1429–1436 (∥, s)	CH_2 δ	CH_2 δ	CH_2 δ

[a]In reflection geometry, the laser incidence plane is either parallel (∥) or perpendicular (⊥) to the chain alignment axis. [b]Intensity: very strong (vs), strong (s), medium (m), weak (w). [c]The polarization of the normal incident light is parallel (∥) or perpendicular (⊥) to the chain alignment axis. Vibration modes: asymmetric stretching (ν_{as}), symmetric stretching (ν_s), bending (δ). [d]Not observed.

Due to the noncentrosymmetric requirement in the SFG process (eq 2), not all peaks seen in IR and Raman spectra are observed in SFG. Note that in the SFG spectra of these cellulose samples there is no peak observed at 2898 or 2904 cm^{-1} (Figure 5) while both Raman and IR spectra show a strong peak in this region (Figures 6 and 7). The strong peaks at 2898 cm^{-1} in Raman and 2904 cm^{-1} in IR are attributed to the stretch vibrations of the methine groups at the axial positions of the six-membered ring of the glucopyranose unit.[50,66] There are equal numbers of methine groups on both sides of the cellulose chain with dipole moments that point in the opposite directions. When chains are stacked during the native crystallization process, the symmetric arrangement of these groups cancels their dipoles across the entire crystal. This may explain the absence of the methine stretch modes in SFG (Figure 5), although they are IR- and Raman-active (Figures 6 and 7). A similar phenomenon has been observed for the SFG spectra of the CH_2 groups in a well-packed self-assembled monolayer.[67−69]

The shoulders at 2920−2968 cm^{-1} in IR and Raman (Figures 6 and 7) are the main peaks observed in SFG where the dominant peak is centered at 2944 cm^{-1} (Figure 5). Minor peaks are observed at 2920 and 2968 cm^{-1}, except when the aligned Iβ nanowhiskers from *Halocynthia* were probed with the laser incidence plane parallel to the nanowhisker alignment direction (B2 in Figure 5). Previously, the 2944 cm^{-1} peak was attributed to the methylene asymmetric stretch modes of the exocyclic CH_2OH groups (Table 1).[33,34] More details of the mode assignment for these peaks will be discussed with DFT-D2 calculations (subsection III).

Comparison of the SFG peaks (Figure 5) in the fingerprint region (<1400 cm^{-1}) with the IR (Figure 6) and Raman (Figure 7) spectra further reveals the complexity of the SFG peak interpretation. For example, a major peak at 1098 cm^{-1} is observed in Raman, but this peak is minor in IR. This might explain the weak intensity of this peak in SFG, since eq 2 dictates that the SFG-active functional groups must be both IR- and Raman-active. But this rule is not as strictly applicable to cellulose as in the case for molecular species at interface.[46,70] Other examples include the 1120 and 1320 cm^{-1} peaks that are observed to be strong in SFG and IR, but are weak in Raman. It is believed that these complications in the SFG peak intensity investigation must be partially due to the birefringence of cellulose crystals[44] and light scattering by crystal aggregates.[34,46]

In the OH stretch region, the SFG spectra of cellulose (Figure 5) show well-resolved discernible peaks but with weak intensity, which contrasts to IR (Figure 6). Because cellulose is hygroscopic, the cellulose OH peaks in IR can sometimes be obscured by signals from water. In contrast, SFG is insensitive to water molecules inside the cellulose sample since they are randomly distributed.[33] Cellulose Iα shows three peaks at 3240, 3320, and ~3370 cm^{-1}, while cellulose Iβ shows two peaks at 3270 and 3320 cm^{-1}. The 3240 and 3270 cm^{-1} peaks are known to be characteristic to cellulose Iα and Iβ, respectively.[71]

The random powders of cellulose Iβ (Avicel PH-101) did not show any polarization dependence. A simple linear combination of the SFG spectra of the aligned cellulose Iβ film taken at the parallel and perpendicular orientations (Figure 5C) produces a spectrum whose relative peak intensities are similar to those of the random sample (Figure 5D). This suggests that the SFG signals from Avicel are the average signals from crystallites which have no preferred orientation with respect to the laser incidence plane.

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

II. Comparison with DFT-D2 Calculations for Cellulose Iα and Iβ. The vibration modes of cellulose Iα and Iβ with two possible hydrogen-bonding schemes (networks A and B, Figure 2) were calculated using the unit cell structure optimized through energy minimization. Table 2 summarizes the lattice parameters determined from XRD analyses and obtained from DFT-D2 calculations. Figure 8 shows that the torsion angles of the hydroxymethylene groups in the energy-relaxed network A structure (165−167°) did not differ significantly from the initial tg conformation (158−170° as shown in Figure 3a), whereas the network B final structure rotated one of the hydroxymethylene groups to a position between tg and gt (86−91°). The network A structure was found to be more stable than the network B structure by approximately 20 and 24 kJ per glucose unit for cellulose Iα and Iβ, respectively. We also examined two other structures for cellulose Iα and Iβ where the initial positions of the hydroxymethylene group were set into the gt or gg conformations (Figure 1); but after energy minimization these structures were significantly unstable compared to those from the *tg* initial conformation. For this reason, the gt and gg models were excluded from consideration.

Table 2. Comparison of Unit Cell Parameters Determined from Experiments from XRD and Calculated from DFT-D2 for Cellulose Iα and Iβ with Networks A and B

	Iα			Iβ		
lattice parameter	expt[14]	net A	net B	expt[13]	net A	net B
a (Å)	6.71	6.55	7.33	7.78	7.54	7.53
b (Å)	5.96	5.65	5.01	8.20	8.14	8.40
c (Å)	10.4	10.38	10.34	10.38	10.39	10.35

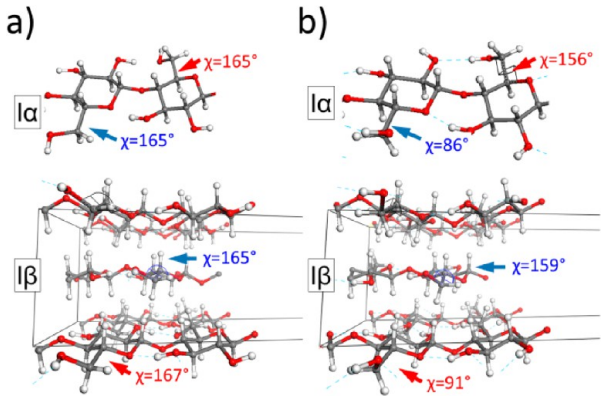


Figure 8. Cellulose Iα (top) and Iβ (bottom) structures obtained after full relaxation (energy minimization) from DFT-D2 calculations: (a) network A and (b) network B (C = dark gray; O = red; H = light gray). Red and blue arrows correspond to the same color units shown as in Figure 3.

To correct for the harmonic approximation in DFT-D2, the calculated vibration modes were rescaled against the experimental SFG data. For this comparison the CH_2 bending (δ), CH stretch (ν), and CH_2 stretch (ν) modes were selected because their normal modes are known to be more isolated from others (this was later confirmed in our DFT-D2 calculations).[50] The CH_2 δ modes are near 1452 and 1484 cm^{-1}.[50] In the alkyl stretch vibration region, the calculated wavenumbers increase in the order of CH ν (delocalized) < CH_2 ν_s < CH ν (localized) < CH_2 ν_{as}.[49,50] As shown in Table 1, the major peak at ~2900 cm^{-1} is assigned to the CH ν mode in the IR, Raman and SFG spectra.[33,49,50,72−74] The most abundant alkyl group in cellulose is the methine group; thus, the strongest peak of the alkyl stretch region in IR and Raman can be assigned to this peak. When the calculated CH ν (localized) frequencies were set to the 2900 cm^{-1} peak and the same scaling factor was applied to other vibration modes, all C−H stretch and bend modes were in good agreement with the wavenumber regions observed in SFG, IR, and Raman spectra. The scaled peak positions for the CH, CH_2, and OH stretch modes are shown in Figure 9. Vibration modes in the fingerprint region (<1400 cm^{-1}) were also calculated, but are not shown here because vibration modes are delocalized over several groups. Thus, the peak assignment of the skeletal modes in the fingerprint region is complex and not critical for the purposes of this study. Movies of the representative vibration modes can be found in the Supporting Information and snapshots of several vibration modes are shown in Figure 10.

II.1. CH_2 Stretching and Bending Modes of Hydroxymethylene (CH_2OH) Group. The methylene ($C6H_2$) stretching and bending modes for cellulose Iα and Iβ are found to be localized and have little coupling with other groups, with only a minor contribution from the C5H stretch (see movies M3A−M6A and M3B−M6B in the Supporting Information). Examples of these ν_s and ν_{as} vibration modes are shown in Figure 10, a and b. The calculated CH_2 stretching frequencies in Figure 9a are 2850−2890 cm^{-1} for ν_s (symmetric) and 2920−2960 cm^{-1} for ν_{as} (asymmetric) for both cellulose Iα and Iβ. These are in good agreement with the observed peaks in SFG (Figure 5 and Table 1). Especially, the calculations for the network A structures of cellulose Iα and Iβ and network B of Iα with the tg conformation predict CH_2 ν_{as} peaks to be at ~2940−2945 cm^{-1}, which is very close to the experimentally observed peak at 2944 cm^{-1} (Figure 5, Table 1).

In the case of the network B structure of cellulose Iβ, the DFT-D2 calculations predicted the CH_2 ν_{as} peak of the corner chain with χ = 91° to be at 2960 cm^{-1} and that of the center chain with χ = 159° to be at 2916 cm^{-1} (Figure 9a). These positions are close to the two peaks observed for the aligned cellulose Iβ sample with the laser incidence plane parallel to the nanowhisker alignment axis (spectrum B2 in Figure 5). However, this may be coincidental because the gt geometry of the CH_2OH group in the network B structure (Figure 8b) is not consistent with the structure determined from the X-ray and neutron diffraction studies (Figure 3a). A previous XRD study suggested that in both polymorphs mixtures of both networks are possible; although network A is more favorable, network B might be associated less ordered regions.[15] NMR studies have suggested that gt or gg conformations are present in regions with reduced order, most likely at the crystal surfaces where the CH_2OH group is not confined to the tg geometry by a well-organized network of intra- and interchain hydrogen bonds.[75] However, SFG signals from crystalline cellulose are insensitive to functional groups that are accessible to water.[33] Thus, it is still uncertain to assign the peaks at 2920 and 2968 cm^{-1} to either the gt conformation or the network B structure.

The strong 2944 cm^{-1} peak could be due to the Fermi resonance of the overtone of the CH_2 δ mode.[39,41,76] In Table 1, the overtones (2δ) of the 1468 and 1484 cm^{-1} peaks are expected to be at 2936 and 2968 cm^{-1}, respectively. Since the CH_2 δ modes are symmetric (Figure 10c), it is unlikely that they will have Fermi resonance with the CH_2 ν_{as} mode. The only possible mode for the Fermi resonance would be the CH_2 ν_s mode. However, the energy difference between the CH_2 ν_s

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

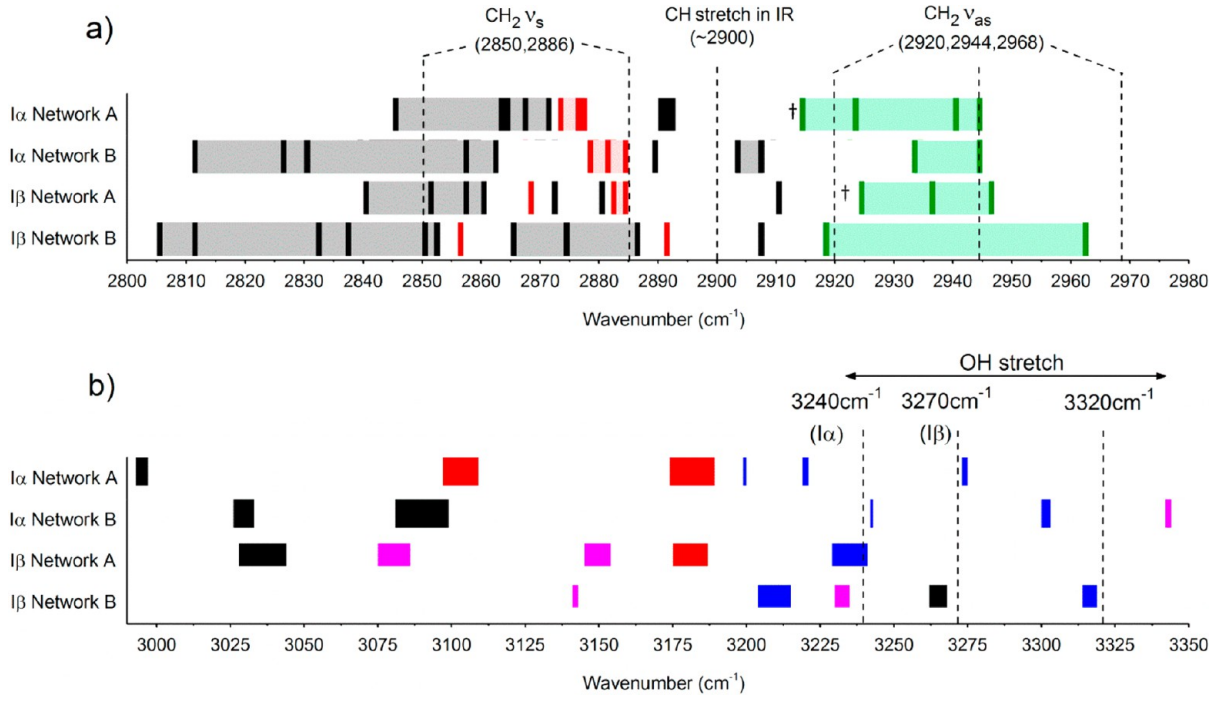


Figure 9. Peak positions predicted by DFT-D2 calculations for (a) CH and CH_2 stretching and (b) dominant OH stretching modes of the energy-minimized structures of cellulose Iα and Iβ crystals with two different hydrogen bond arrangements (networks A and B). In (a), the CH methine stretching regions are marked with black, the CH_2 symmetric stretching regions are red, and the CH_2 asymmetric regions are green. The bands marked with † have some degree of coupling between the CH_2 asymmetric stretching and the 5CH stretching. In (b), the intrachain hydrogen-bonded O2H and O3H stretching groups are marked with black and blue, respectively, and the interchain hydrogen-bonded O2H/O6H and O2H/O3H/O6H groups are marked with pink and red, respectively. The observed SFG peaks for cellulose Iα and Iβ are shown with the dashed line.

mode (2850−2890 cm^{-1}) and the CH_2 δ overtone (2934 and 2968 cm^{-1}) is too large (45−120 cm^{-1}) to have a significant exchange.[77] Additionally, the Fermi resonance peak intensity is expected to be weaker than the fundamental mode intensity. If the 2944 cm^{-1} peak is due to the Fermi resonance between the CH_2 ν_s and CH_2 2δ modes, the intensity would be smaller than the peaks in the 2850−2890 cm^{-1} region. However, the 2944 cm^{-1} peak is always the dominant peak compared to the 2850−2890 cm^{-1} region which is very weak in most cases. Thus, it is unlikely that the 2944 cm^{-1} peak is due to the Fermi resonance. Unfortunately, this could not be confirmed since the Fermi resonance between the fundamental CH_2 ν_s mode and the CH_2 δ overtone could not be calculated in this method.

II.2. CH Stretching from the Methine (CH) Groups at the Carbon 1, 2, 3, 4, and 5 Positions. The stretching vibrations of the methine (CH) groups at the axial positions of the glucose ring on carbons 1, 2, 3, 4, and 5 are the strongest peaks at 2904 and 2898 cm^{-1} in the IR and Raman spectra, respectively (Figures 6 and 7), but are absent in SFG (Figure 5). The methine groups are arranged perpendicular to the chain axis; their peaks in IR and Raman become enhanced when the incident light is polarized perpendicular to the chain axis (Figures 6 and 7). It has been proposed that in SFG these modes are symmetry-canceled because the axial-positioned CH groups are in equal numbers on either side of the cellulose chain (Figure 10, d and e).[33] DFT-D2 analysis shows that the peaks around 2900 cm^{-1} are the stretch modes of the CH groups with opposite displacements (Figure 10d). Below 2850 cm^{-1} there are several CH stretch vibrations predicted from DFT-D2 calculations (Figure 9a). These modes involve some degree of the C−C ring stretch and are delocalized over the entire cellulose chain (Figure 10e, and see movies M8A and M8B in the Supporting Information). The peaks in this region are very weak in IR and Raman (Figures 6 and 7) and negligible in the SFG (Figure 5). Overall, DFT-D2 calculations show that the collective motions of the CH groups are equally distributed in the opposite directions along the [100] direction of the cellulose crystal, which supports the previous explanation for the absence of the CH stretch peak at ~2900 cm^{-1} in SFG.[33,34]

II.3. OH Stretching from Intra- and Interchain Hydrogen Bonding. The peak positions of the OH stretching modes in cellulose Iα and Iβ are sensitive to the intra- and interchain hydrogen bonding networks.[23,66] Because most hydroxyl groups are connected by hydrogen bonds, their stretching vibrations are highly coupled (i.e., as the O−H bond stretches the hydrogen bond donor H atom approaches the central O atom; Figure 10e, and see movies M1,2A and M1,2B in the Supporting Information). The O−H stretching vibration wavenumbers obtained from DFT-D2 calculations (Figure 9b) range from 3000 to 3350 cm^{-1}, which is lower than the experimentally observed values (3200−3400 cm^{-1} in SFG and 3200−3600 cm^{-1} in IR).[78] One possible source for this discrepancy is the use of the rescaling factor obtained from the C−H vibration modes in our method. Different covalent bonds may require different scaling factors since the degree of deviation from the harmonic oscillation approximation can vary from one chemical bond to another.[79]

The vibration wavenumber of the O−H stretching mode could be off by as much as 100 cm^{-1} with a 0.1 Å change in the O···O distance between hydrogen-bonded oxygen atoms.[80] These small changes in the O···O distance can be generated with the small deviations in the torsional angle of the hydroxymethylene groups (Figure 8). As shown in Table 3, many O···O distances calculated by DFT-D2 are smaller than

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

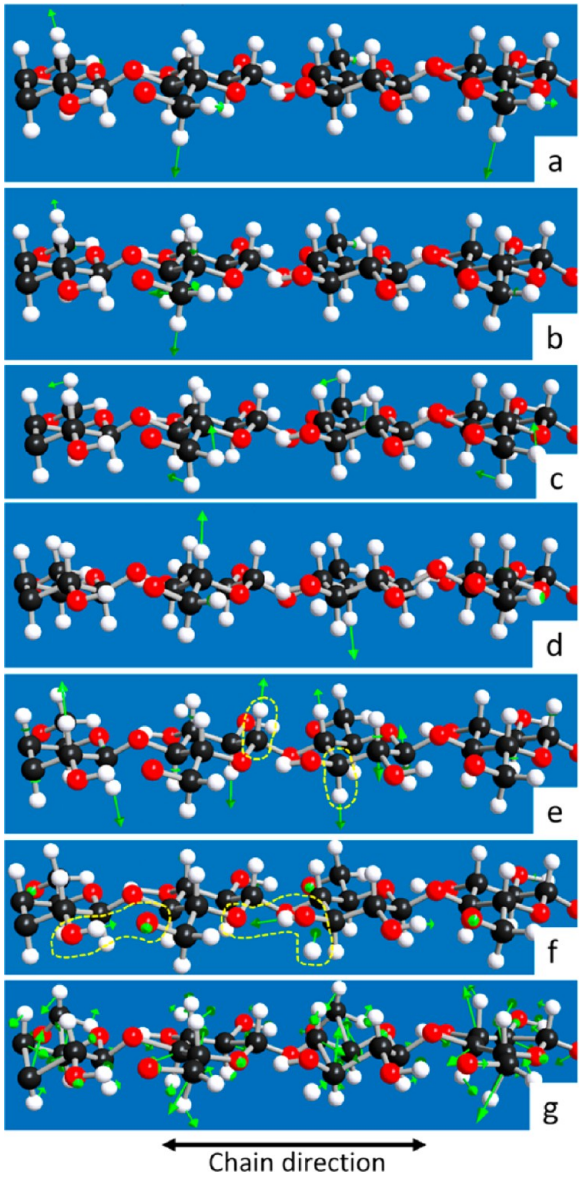


Figure 10. Snapshots of vibrational modes from DFT-D2 calculations for cellulose Iβ network A: (a) CH_2 symmetric stretch at 2867 cm^{-1}, (b) CH_2 asymmetric stretch at 2943 cm^{-1}, (c) CH_2 bend at 1477 cm^{-1}, (d) localized CH stretch at 2910 cm^{-1}, (e) delocalized CH stretch at 2872 cm^{-1}, (f) stretch of intra- and interchain hydrogen-bonded OH groups at 3231 cm^{-1}, and (g) skeletal stretch and bend at 1098 cm^{-1} (C = dark gray; O = red; H = light gray; displacement vectors = green).

those determined by X-ray and neutron diffraction.[13−15] This is likely due to the inaccuracy of the hydrogen bond energy calculation in the DFT-D2 energy minimization.[54,56,57] The dimension of the unit cell *b* axis is about 0.1−0.9 Å smaller than the experimental values, implying that the interchain distances are slightly decreased. The exception is the unit cell of cellulose Iβ with network B in which the *b*-axis dimension is ∼0.2 Å larger than the experimental value (Table 2). This is consistent with the slightly higher OH stretch peak positions for cellulose Iβ with network B, compared to the others (Figure 9b).

Even with these changes in the *b*-axis dimension of the unit cell, it is noteworthy that the stretching vibrations of the O3H group were closer to the experimental values (Figure 9b). The O3H group is connected only by an intrachain hydrogen bond (O3H···O5) along the chain direction, in both network A and B (Figure 3) and would therefore be more sensitive to the *c*-axis dimension. The calculated *c*-axis lattice constant was close to the experimental value (Table 2), and the peak positions for the O3−H stretch mode were in the 3200−3325 cm^{-1} range (blue in Figure 9b), which were within 100 cm^{-1} to the values observed experimentally.

II.4. Skeletal Stretching and Bending Vibrations (1000−1400 cm^{-1}). The skeletal stretch modes of the C−C and C−O bonds and the deformation modes of the C−C−H and C−O−H groups of the glucan chain are found in the 1000−1400 cm^{-1} region.[50,66] There are several well-resolved SFG peaks with relatively high signal intensities in this region (Figure 5). However, it is difficult to assign these peaks to specific vibration modes because each band is a combination of several vibrational modes coupled together. This means that the vibrational motions of atoms are delocalized over several covalent bonds and often over the entire chain.[50] This can be seen in the example shown in Figure 10f (and movies M8,9A and M8,9B in the Supporting Information). Because the major differences between polymorphs are interchain hydrogen bonds, the skeletal vibration motions in this region are similar in cellulose Iα and Iβ.[71]

III. Implications. The comparison of the SFG peak positions of cellulose Iα and Iβ samples with the wavenumbers estimated from DFT-D2 calculations is the first step toward molecular understanding of the SFG activity of cellulose crystals. The cellulose nanowhiskers from *Glaucocystis* and *Halocynthia* are good model systems for the initial peak assignment study since they are highly crystalline, contain mostly single allomorphs, and have been fully analyzed with X-ray and neutron diffraction studies.[13,14] The peak assignment of these well-characterized cellulose Iα and Iβ samples will provide basis for the comparative SFG study of crystal structures of different polymorphs.[65] The SFG intensities of cellulose in plants can vary depending on the spatial orientation and distribution of crystalline cellulose. Thus, SFG can provide information of three-dimensional arrangements or packing of cellulose microfibrils in plant cell walls at various developmental stages.[81] The spatial distribution of cellulose could be imaged if SFG is combined with microscopy.[39,82,83] The SFG peak assignment would greatly assist the interpretation of these SFG microscopic imaging of cellulose crystals distributed in biomass. The digestibility of cellulose in biomass could be strongly influenced by the accessibility of enzymes to cellulose microfibrils, which could be altered or controlled through pretreatment processes.[8,84] Without a molecular understanding of high-quality large crystals, structural deviations in thinner cellulose microfibrils in biomass would not be properly understood. The SFG peak assignment of cellulose Iα and Iβ made through comparison with DFT-D2 calculations in this study will be the foundation for future research on spatial distributions of cellulose microfibrils and cellulose polymorphism in developing plant cell walls and their roles in digestibility of lignocellulose biomass.

■ CONCLUSION

The SFG spectra of uniaxial-aligned cellulose Iα and Iβ crystals were compared with the vibration modes simulated from DFT-D2 calculations for the energy-minimized crystal structures of cellulose Iα and Iβ. According to DFT-D2 calculations, the SFG peaks that can be assigned to specific normal modes were the methylene (CH_2) stretching and bending vibrations. In

dx.doi.org/10.1021/jp402998s | J. Phys. Chem. B XXXX, XXX, XXX−XXX

Table 3. Comparison of the O···O Distance between Hydrogen-Bonded Hydroxyl Groups Determined from XRD and Calculated from DFT-D2 for Cellulose Iα and Iβ with Networks A and B[a]

		Iα			Iβ		
glucose unit[b]	intrachain H-bond[c]	expt[14]	net A	net B	expt[13]	net A	net B
blue	O3−H···O5	2.91	2.74	2.69	2.70	2.74	2.65
	O2−H···O6[d]	2.46	2.67	–	2.86	2.74	–
red	O3−H···O5	2.86	2.74	2.71	2.76	2.73	2.85
	O2−H···O6[d]	2.48	2.74	–	2.76	2.67	–
		Iα			Iβ		
glucose unit[b]	interchain H-bond[c]	expt[14]	net A	net B	expt[15]	net A	net B
blue	O6−H···O3	2.82	2.70	–	2.67	2.66	2.93[e]
	O2−H···O6	–	–	2.63	3.29	–	2.81
red	O6−H···O3	2.77	2.69	–	2.80	2.73	2.72[e]
	O2−H···O6	–	–	2.71	3.62	–	2.66

[a]All distances are in angstroms. [b]Labeling based on the colors scheme in Figures 3 and 8. [c]The O···O distances greater than 3.0 Å were omitted. [d]The O2···O6 distances are shorter in Iα than Iβ; however, the bond angles are less favorable in Iα (127−134°) than in Iβ (159−165°); thus, the strength of the O2−H···O6 hydrogen bond is comparable in both allomorphs. [e]There were no experimental values reported for this hydrogen bond.

particular, the CH_2 asymmetric stretching vibration of the exocyclic CH_2OH group in the tg conformation was predicted to be around 2944 cm^{-1}, which was close to the strongest peak in the alkyl stretch region of the SFG spectra of highly crystalline, uniaxial-aligned cellulose Iα and Iβ. DFT-D2 calculations also showed that the stretch modes of the methine (CH) group at the axial position of the glucopyranose unit were coupled with those on the opposite side of the glucan unit. This supports earlier predictions that, although these modes were dominant in IR and Raman, they are SFG-inactive. The peak ranges calculated for the OH stretch modes were lower than the experimental values due to the inaccuracy in the hydrogen bonding stabilization in DFT-D2. Even with this potential error, it was noteworthy that the peak positions of the intrachain hydrogen-bonded hydroxyl groups (O3H···O5) were close to the observed SFG peaks (3240−3450 cm^{-1}).

■ ASSOCIATED CONTENT

Supporting Information

Movies for the representative vibrational modes in cellulose Iα and Iβ relaxed structures from DFT-D2 which include the following: O2H/O3H/O6H stretch, $6CH_2$ stretching and bending, CH stretch, skeletal stretching and deformation. All movies are from structures with network A intra- and interchain hydrogen bonding. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: shkim@engr.psu.edu.

Present Address

[§]Department of Chemical Engineering, N-323 Millennium Science Complex, University Park, PA 16801.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported as part of The Center for Lignocellulose Structure and Formation, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under award no. DE-SC0001090.

■ REFERENCES

(1) Zugenmaier, P. *Crystalline Cellulose and Derivatives: Characterization and Structures*; Springer Verlag: Berlin, Germany, 2007.

(2) Albersheim, P.; Darvill, A.; Roberts, K.; Sederoff, R.; Staehelin, A. *Plant Cell Walls*; Garland Science: New York, 2011.

(3) Pérez, S.; Samain, D. Structure and Engineering of Celluloses. *Adv. Carbohydr. Chem. Biochem.* **2010**, *64*, 25−116.

(4) Langan, P.; Gnanakaran, S.; Rector, K. D.; Pawley, N.; Fox, D. T.; Cho, D. W.; Hammel, K. E. Exploring New Strategies for Cellulosic Biofuels Production. *Energy Environ. Sci.* **2011**, *4*, 3820−3833.

(5) da Costa Sousa, L.; Chundawat, S. P. S.; Balan, V.; Dale, B. E. 'Cradle-to-Grave' Assessment of Existing Lignocellulose Pretreatment Technologies. *Curr. Opin. Biotechnol.* **2009**, *20*, 339−347.

(6) Himmel, M. E.; Ding, S. Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007**, *315*, 804−807.

(7) Jarvis, M. C.; McCann, M. C. Macromolecular Biophysics of the Plant Cell Wall: Concepts and Methodology. *Plant Physiol. Biochem.* **2000**, *38*, 1−13.

(8) Ding, S.-Y.; Liu, Y.-S.; Zeng, Y.; Himmel, M. E.; Baker, J. O.; Bayer, E. A. How Does Plant Cell Wall Nanoscale Architecture Correlate with Enzymatic Digestibility? *Science* **2012**, *338*, 1055−1060.

(9) Paredez, A. R.; Somerville, C. R.; Ehrhardt, D. W. Visualization of Cellulose Synthase Demonstrates Functional Association with Microtubules. *Science* **2006**, *312*, 1491−1495.

(10) Somerville, C.; Bauer, S.; Brininstool, G.; Facette, M.; Hamann, T.; Milne, J.; Osborne, E.; Paredez, A.; Persson, S.; Raab, T. Toward a Systems Approach to Understanding Plant Cell Walls. *Sci. Signal.* **2004**, *306*, 2206−2211.

(11) Cosgrove, D. J. Growth of the Plant Cell Wall. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 850−861.

(12) Atalla, R. H.; Vanderhart, D. L. Native Cellulose: A Composite of Two Distinct Crystalline Forms. *Science* **1984**, *223*, 283−285.

(13) Nishiyama, Y.; Langan, P.; Chanzy, H. Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.

(14) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-Ray and Neutron Fiber Diffraction. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.

(15) Nishiyama, Y.; Johnson, G. P.; French, A. D.; Forsyth, V. T.; Langan, P. Neutron Crystallography, Molecular Dynamics, and Quantum Mechanics Studies of the Nature of Hydrogen Bonding in Cellulose Iβ. *Biomacromolecules* **2008**, *9*, 3133−3140.

(16) 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*, 7506−7511.
(17) Šturcová, A.; His, I.; Apperley, D. C.; Sugiyama, J.; Jarvis, M. C. Structural Details of Crystalline Cellulose from Higher Plants. *Biomacromolecules* **2004**, *5*, 1333−1339.
(18) 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*, E1195−E1203.
(19) Marrinan, H. J.; Mann, J. Infrared Spectra of the Crystalline Modifications of Cellulose. *J. Polym. Sci.* **1956**, *11*, 301−311.
(20) Nelson, M. L.; O'Connor, R. T. Relation of Certain Infrared Bands to Cellulose Crystallinity and Crystal Lattice Type. Part I. Spectra of Lattice Types, I, II, III and of Amorphous Cellulose. *J. Appl. Polym. Sci.* **1964**, *8*, 1311−1324.
(21) Nelson, M. L.; O'Connor, R. T. Relation of Certain Infrared Bands to Cellulose Crystallinity and Crystal Lattice Type. Part II. A New Infrared Ratio for Estimation of Crystallinity in Celluloses I and II. *J. Appl. Polym. Sci.* **1964**, *8*, 1325−1341.
(22) Mann, J.; Marrinan, H. J. Crystalline Modifications of Cellulose. Part II. A Study with Plane-Polarized Infrared Radiation. *J. Polym. Sci.* **1958**, *32*, 357−370.
(23) Liang, C.; Marchessault, R. Infrared Spectra of Crystalline Polysaccharides. I. Hydrogen Bonds in Native Celluloses. *J. Polym. Sci.* **1959**, *37*, 385−395.
(24) Tsuboi, M. Infrared Spectrum and Crystal Structure of Cellulose. *J. Polym. Sci.* **1957**, *25*, 159−171.
(25) Morikawa, H.; Hayashi, R.; Senda, M. Infrared Analysis of Pea Stem Cell Walls and Oriented Structure of Matrix Polysaccharides in Them. *Plant Cell Physiol.* **1978**, *19*, 1151−1159.
(26) Newman, R. H.; Ha, M. A.; Melton, L. D. Solid-State ^{13}C NMR Investigation of Molecular Ordering in the Cellulose of Apple Cell Walls. *J. Agric. Food. Chem.* **1994**, *42*, 1402−1406.
(27) Dick-Pérez, M.; Zhang, Y.; 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*, 989−1000.
(28) Davies, L. M.; Harris, P. J.; Newman, R. H. Molecular Ordering of Cellulose after Extraction of Polysaccharides from Primary Cell Walls of Arabidopsis Thaliana: A Solid-State CP/MAS ^{13}C NMR Study. *Carbohydr. Res.* **2002**, *337*, 587−593.
(29) Thomas, L. H.; Forsyth, V. T.; Šturcová, A.; Kennedy, C. J.; May, R. P.; Altaner, C. M.; Apperley, D. C.; Wess, T. J.; Jarvis, M. C. Structure of Cellulose Microfibrils in Primary Cell Walls from Collenchyma. *Plant Physiol.* **2013**, *161*, 465−476.
(30) Isogai, A.; Usuda, M.; Kato, T.; Uryu, T.; Atalla, R. H. Solid-State CP/MAS ^{13}C NMR Study of Cellulose Polymorphs. *Macromolecules* **1989**, *22*, 3168−3172.
(31) Sarkar, P.; Bosneaga, E.; Auer, M. Plant Cell Walls Throughout Evolution: Towards a Molecular Understanding of Their Design Principles. *J. Exp. Bot.* **2009**, *60*, 3615−3635.
(32) Park, S.; Baker, J. O.; Himmel, M. E.; Parilla, P. A.; Johnson, D. K. Research Cellulose Crystallinity Index: Measurement Techniques and Their Impact on Interpreting Cellulase Performance. *Biotechnol. Biofuels* **2010**, *3*, 1−10.
(33) Barnette, A. L.; Bradley, L. C.; Veres, B. D.; Schreiner, E. P.; Park, Y. B.; Park, J.; Park, S.; Kim, S. H. Selective Detection of Crystalline Cellulose in Plant Cell Walls with Sum-Frequency-Generation (SFG) Vibration Spectroscopy. *Biomacromolecules* **2011**, *12*, 2434−2439.
(34) Barnette, A. L.; Lee, C.; Bradley, L. C.; Schreiner, E. P.; Park, Y. B.; Shin, H.; Cosgrove, D. J.; Park, S.; Kim, S. H. Quantification of Crystalline Cellulose in Lignocellulosic Biomass Using Sum Frequency Generation (SFG) Vibration Spectroscopy and Comparison with Other Analytical Methods. *Carbohydr. Polym.* **2012**, *89*, 802−809.
(35) Bain, C. D. Sum-Frequency Vibrational Spectroscopy of the Solid/Liquid Interface. *J. Chem. Soc., Faraday Trans.* **1995**, *91*, 1281−1296.
(36) Wang, H.-F.; Gan, W.; Lu, R.; Rao, Y.; Wu, B.-H. Quantitative Spectral and Orientational Analysis in Surface Sum Frequency Generation Vibrational Spectroscopy (SFG-VS). *Int. Rev. Phys. Chem.* **2005**, *24*, 191−256.
(37) Shen, Y. Surface Properties Probed by Second-Harmonic and Sum-Frequency Generation. *Nature* **1989**, *337*, 519−525.
(38) Lambert, A. G.; Davies, P. B.; Neivandt, D. J. Implementing the Theory of Sum Frequency Generation Vibrational Spectroscopy: A Tutorial Review. *Appl. Spectrosc. Rev.* **2005**, *40*, 103−145.
(39) Hieu, H. C.; Tuan, N. A.; Li, H.; Miyauchi, Y.; Mizutani, G. Sum Frequency Generation Microscopy Study of Cellulose Fibers. *Appl. Spectrosc.* **2011**, *65*, 1254−1259.
(40) Fu, L.; Wang, Z.; Yan, E. C. Chiral Vibrational Structures of Proteins at Interfaces Probed by Sum Frequency Generation Spectroscopy. *Int. J. Mol. Sci.* **2011**, *12*, 9404−9425.
(41) Rocha-Mendoza, I.; Yankelevich, D. R.; Wang, M.; Reiser, K. M.; Frank, C. W.; Knoesen, A. Sum Frequency Vibrational Spectroscopy: The Molecular Origins of the Optical Second-Order Nonlinearity of Collagen. *Biophys. J.* **2007**, *93*, 4433−4444.
(42) Gualtieri, E. J.; Haupert, L. M.; Simpson, G. J. Interpreting Nonlinear Optics of Biopolymer Assemblies: Finding a Hook. *Chem. Phys. Lett.* **2008**, *465*, 167−174.
(43) Horikawa, Y.; Sugiyama, J. Localization of Crystalline Allomorphs in Cellulose Microfibril. *Biomacromolecules* **2009**, *10*, 2235−2239.
(44) Cranston, E. D.; Gray, D. G. Birefringence in Spin-Coated Films Containing Cellulose Nanocrystals. *Colloids Surf., A* **2008**, *325*, 44−51.
(45) Nadiarnykh, O.; LaComb, R. B.; Campagnola, P. J.; Mohler, W. A. Coherent and Incoherent SHG in Fibrillar Cellulose Matrices. *Opt. Express* **2007**, *15*, 3348−3360.
(46) LaComb, R.; Nadiarnykh, O.; Townsend, S. S.; Campagnola, P. J. Phase Matching Considerations in Second Harmonic Generation from Tissues: Effects on Emission Directionality, Conversion Efficiency and Observed Morphology. *Opt. Commun.* **2008**, *281*, 1823−1832.
(47) Anderson, D. B.; Kerr, T. Growth and Structure of Cotton Fiber. *Ind. Eng. Chem.* **1938**, *30*, 48−54.
(48) Seagull, R. W. Changes in Microtubule Organization and Wall Microfibril Orientation During in Vitro Cotton Fiber Development: An Immunofluorescent Study. *Can. J. Bot.* **1986**, *64*, 1373−1381.
(49) Cael, J.; Gardner, K.; Koenig, J.; Blackwell, J. Infrared and Raman Spectroscopy of Carbohydrates. Paper V. Normal Coordinate Analysis of Cellulose I. *J. Chem. Phys.* **1975**, *62*, 1145−1154.
(50) Wiley, J. H.; Atalla, R. H. Band Assignments in the Raman Spectra of Celluloses. *Carbohydr. Res.* **1987**, *160*, 113−129.
(51) Qian, X. The Effect of Cooperativity on Hydrogen Bonding Interactions in Native Cellulose Iβ from Ab Initio Molecular Dynamics Simulations. *Mol. Simul.* **2008**, *34*, 183−191.
(52) Heiner, A. P.; Sugiyama, J.; Teleman, O. Crystalline Cellulose Iα and Iβ Studied by Molecular Dynamics Simulation. *Carbohydr. Res.* **1995**, *273*, 207−223.
(53) Barsberg, S. Prediction of Vibrational Spectra of Polysaccharides-Simulated IR Spectrum of Cellulose Based on Density Functional Theory (DFT). *J. Phys. Chem. B* **2010**, *114*, 11703−11708.
(54) Kubicki, J.; Mohamed, M.-A.; Watts, H. Quantum Mechanical Modeling of the Structures, Energetics and Spectral Properties of Iα and Iβ Cellulose. *Cellulose* **2013**, *20*, 9−23.
(55) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(56) Bučko, T.; Tunega, D.; Ángyán, J. G.; Hafner, J. Ab Initio Study of Structure and Interconversion of Native Cellulose Phases. *J. Phys. Chem. A* **2011**, *115*, 10097−10105.
(57) Li, Y.; Lin, M.; Davenport, J. W. Ab Initio Studies of Cellulose I: Crystal Structure, Intermolecular Forces, and Interactions with Water. *J. Phys. Chem.* **2011**, *115*, 11533−11539.

(58) Nishiyama, Y.; Shigenori, K.; Masahisa, W.; Takeshi, O. Cellulose Microcrystal Film of High Uniaxial Orientation. *Macromolecules* **1997**, *30*, 6395−6397.
(59) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169−11186.
(60) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Open-Shell Transition Metals. *Phys. Rev. B* **1993**, *48*, 13115−13118.
(61) Kresse, G.; Furthmüller, J.; Hafner, J. Theory of the Crystal Structures of Selenium and Tellurium: The Effect of Generalized-Gradient Corrections to the Local-Density Approximation. *Phys. Rev. B* **1994**, *50*, 13181−13185.
(62) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal−Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251−14269.
(63) Eck, B. Wxdragon. Available at: www.wxdragon.de.
(64) Nishiyama, Y.; Isogai, A.; Okano, T.; Müller, M.; Chanzy, H. Intracrystalline Deuteration of Native Cellulose. *Macromolecules* **1999**, *32*, 2078−2081.
(65) Lee, C.; Mittal, A.; Barnette, A. L.; Kafle, K.; Park, Y. B.; Shin, H.; Johnson, D. K.; Kim, S. H. Cellulose Polymorphism: Exocyclic CH_2OH Conformation and Chain Orientation Determined by Sum-Frequency-Generation (SFG) Vibration Spectroscopy. *Cellulose* **2013**, *20*, 991−1000.
(66) Blackwell, J.; Vasko, P. D.; Koenig, J. Infrared and Raman Spectroscopy of Cellulose. *Cellul. Chem. Technol.* **1977**, 206−218.
(67) Weeraman, C.; Yatawara, A. K.; Bordenyuk, A. N.; Benderskii, A. V. Effect of Nanoscale Geometry on Molecular Conformation: Vibrational Sum-Frequency Generation of Alkanethiols on Gold Nanoparticles. *J. Am. Chem. Soc.* **2006**, *128*, 14244−14245.
(68) Yatawara, A. K.; Tiruchinapally, G.; Bordenyuk, A. N.; Andreana, P. R.; Benderskii, A. V. Carbohydrate Surface Attachment Characterized by Sum Frequency Generation Spectroscopy. *Langmuir* **2009**, *25*, 1901−1904.
(69) Chen, E. H.; Hayes, P. L.; Nguyen, S. T.; Geiger, F. M. Zinc Interactions with Glucosamine-Functionalized Fused Silica/Water Interfaces. *J. Phys. Chem.* **2010**, *114*, 19483−19488.
(70) Boyd, R. W. *Nonlinear Optics*; Academic Press Elsevier: New York, 2008.
(71) Sugiyama, J.; Persson, J.; Chanzy, H. Combined Infrared and Electron Diffraction Study of the Polymorphism of Native Celluloses. *Macromolecules* **1991**, *24*, 2461−2466.
(72) Kataoka, S.; Cremer, P. S. Probing Molecular Structure at Interfaces for Comparison with Bulk Solution Behavior: Water/2-Propanol Mixtures Monitored by Vibrational Sum Frequency Spectroscopy. *J. Am. Chem. Soc.* **2006**, *128*, 5516−5522.
(73) Snyder, R. G.; Aljibury, A.; Strauss, H. L.; Casal, H.; Gough, K.; Murphy, W. Isolated C−H Stretching Vibrations of N-Alkanes: Assignments and Relation to Structure. *J. Chem. Phys.* **1984**, *81*, 5352−5361.
(74) Fischer, S.; Schenzel, K.; Fischer, K.; Diepenbrock, W. Applications of FT Raman Spectroscopy and Micro Spectroscopy Characterizing Cellulose and Cellulosic Biomaterials. In *Macromolecular Symposia*; Wiley-VCH: Weinheim, Germany, 2005; pp 41−56.
(75) Newman, R. H.; Davidson, T. C. Molecular Conformations at the Cellulose−Water Interface. *Cellulose* **2004**, *11*, 23−32.
(76) Lu, R.; Gan, W.; Wu, B.; Chen, H.; Wang, H. Vibrational Polarization Spectroscopy of CH Stretching Modes of the Methylene Group at the Vapor/Liquid Interfaces with Sum Frequency Generation. *J. Phys. Chem. B* **2004**, *108*, 7297−7306.
(77) Harris, D. C.; Bertolucci, M. D. *Symmetry and Spectroscopy*; Dover Publications: New York, 1978.
(78) Marechal, Y.; Chanzy, H. The Hydrogen Bond Network in Iβ Cellulose as Observed by Infrared Spectrometry. *J. Mol. Struct.* **2000**, *523*, 183−196.
(79) Bauschlicher, C. W.; Langhoff, S. R. The Calculation of Accurate Harmonic Frequencies of Large Molecules: The Polycyclic Aromatic Hydrocarbons, a Case Study. *Spectrochim. Acta, Part A* **1997**, *53*, 1225−1240.
(80) Kubicki, J.; Sykes, D.; Rossman, G. R. Calculated Trends of OH Infrared Stretching Vibrations with Composition and Structure in Aluminosilicate Molecules. *Phys. Chem. Miner.* **1993**, *20*, 425−432.
(81) Park, Y. B.; Lee, C.; Zhang, T.; Koo, B. W.; Park, S.; Cosgrove, D. J.; Kim, S. H. Monitoring Structural Variations of Cellulose in Intact Plant Cell Walls Using Sum Frequency (SFG) Spectroscopy. Submitted for publication.
(82) Chung, C.-Y.; Boik, J.; Potma, E. O. Biomolecular Imaging with Coherent Nonlinear Vibrational Microscopy. *Annu. Rev. Phys. Chem.* **2013**, *64*.
(83) Miyauchi, Y.; Sano, H.; Mirzutani, G. Selective Observation of Starch in a Water Plant Using Optical Sum-Frequency Microscopy. *J. Opt. Soc. Am. A* **2006**, *23*, 1687−1690.
(84) Wang, W.; Chen, B. S.; Donohoe, P. N.; Ceiesielski, A.; Mittal, A.; Katahira, R.; Kuhn, E. M.; Kafle, K.; Lee, C.; Park, S.; Kim, S. H.; Tucker, M. P.; Himmel, M. E.; Johnson, D. K. Effect of Mechanical Forces on the Effectiveness of Three Reactors Used for Dilute Acid Pretreatment of Corn Stover. Submitted for publication.
(85) Sugiyama, J.; Vuong, R.; Chanzy, H. Electron Diffraction Study on the Two Crystalline Phases Occurring in Native Cellulose from an Algal Cell Wall. *Macromolecules* **1991**, *24*, 4168−4175.