

Diffraction from nonperiodic models of cellulose crystals

Yoshiharu Nishiyama · Glenn P. Johnson ·
Alfred D. French

Received: 13 December 2011 / Accepted: 3 January 2012 / Published online: 20 January 2012
© Springer Science+Business Media B.V. (outside the USA) 2012

Abstract Powder and fiber diffraction patterns were calculated for model cellulose crystallites with chains 20 glucose units long. Model sizes ranged from four chains to 169 chains, based on cellulose $I\beta$ coordinates. They were subjected to various combinations of energy minimization and molecular dynamics (MD) in water. Disorder induced by MD and one or two layers of water had small effects on the relative intensities, except that together they reduced the low-angle scattering that was otherwise severe enough to shift the $1\bar{1}0$ peak. Other shifts in the calculated peaks occurred because the empirical force field used for MD and minimization caused the models to have small discrepancies with the experimental intermolecular distances. Twisting and other disorder induced by minimization or MD increased the breadth of peaks by about $0.2\text{--}0.3^\circ$ $2\text{-}\theta$. Patterns were compared with experimental results. In particular, the calculated fiber patterns revealed a potential for a larger number of experimental diffraction spots to be found for

cellulose from some higher plants when crystallites are well-oriented. Either that, or further understanding of those structures is needed. One major use for patterns calculated from models is testing of various proposals for microfibril organization.

Keywords X-ray · Fibril · Elementary · Microfibril · Amorphous

Introduction

Many technologies have contributed to our knowledge of cellulose structure such as optical, electron and atomic force microscopy, as well as infrared, Raman, and nuclear magnetic resonance spectroscopy. Despite these other methods and the fact that cellulose fibers were first placed in an X-ray beam nearly 100 years ago, experimental diffraction patterns continue to be of interest. Diffraction patterns are important because they directly reflect the time- and spatially-averaged organization of atoms in the sample. However, fibrous samples present greater challenges than do the individual crystals of small molecules that are on the scale of hundreds of microns larger. Fibers usually contain large numbers of nanometer-scale crystals. Ideally, those crystallites are aligned with their long axes parallel to the fiber axis but with random rotational orientations about their long axes. Compared to powders, fiber diffraction patterns allow unambiguous

Y. Nishiyama (✉)
Centre de Recherche sur les Macromolécules Végétales
(CERMAV), Joseph Fourier University of Grenoble,
BP 53, 38041 Grenoble Cedex 9, France
e-mail: yoshi@cermav.cnrs.fr

G. P. Johnson · A. D. French (✉)
United States Department of Agriculture, Southern
Regional Research Center (SRRRC), 1100 Robert E. Lee
Boulevard, New Orleans, LA 70124, USA
e-mail: Al.French@ars.usda.gov

determination of one of the three repeated dimensions in the crystallites, making definition of the remaining two dimensions less difficult and generally reducing overlap of the diffracted intensity. However, some samples of cellulose are powders with more or less random orientation of the crystallites in all three directions. Powder diffraction patterns are often obtained more conveniently and can furnish valuable insights on the sample in question. Still, their interpretation can usually be informed by the more detailed information from fibrous samples.

Enough information is available to determine a crystal structure at atomic resolution if highly oriented fiber specimens with large crystallites are analyzed, as was the case for a series of cellulose allomorphs (Nishiyama et al. 2002, 2003a; Langan et al. 2001; Wada et al. 2004, 2009) and related crystals (Nishiyama et al. 2011). Those specimens gave sharp diffraction spots whose intensities could be extracted rather accurately. Although many diffraction spots still overlap due to the cylindrical averaging in fibers, structural features such as primary alcohol group orientation could be directly obtained by applying classical crystallographic approaches.

On the other hand, most practical samples of cellulose have smaller crystallite sizes, complicated textures (supramolecular structures) that result in lower orientation, and structural defects or paracrystallinity. Some percentage of water may be present, with an unknown effect. As a consequence, there is insufficient data to solve for more than a few variables so the atomic coordinates cannot be obtained directly from diffraction data. Instead, the structure is often analyzed by trial-and-error methods that compare experimental diffraction data with diffraction intensities calculated from a series of atomistic molecular models. The model giving the best agreement is the most plausible one. Once the crystal structure is thought to be known from samples with relatively large crystals, however, then the crystallite size, texture and disorder can be studied with trial and error approaches. These traits are thought to also be important for understanding the various properties of cellulose from various sources.

Interpretation of diffraction data to understand the additional structural aspects is not trivial (Fernandes et al. 2011), being somewhat dependent on the models that are employed to compensate for such deviations from traditional crystals. Interpretations of such

diffraction patterns are typically based on simple analytical expressions. For example, because greater widths of the diffraction peaks correspond to smaller numbers of crystallographic planes, and hence diminished crystallite size, that size can be assessed with the Scherrer equation (Azároff and Buerger 1958; Alexander 1969). For example, this relationship was used to evaluate the peak widths in Fig. 1. Also, diffraction intensities have often been assumed to arise from a combination of sharp peaks arising from the crystalline component and from smooth scattering from amorphous regions. However, that thinking may be problematic.

Observations by microscopy, hypotheses on the morphogenesis, and molecular modeling often lead to structural models that do not necessarily fit this simplistic, two-phase view. For example, cellulose is thought to crystallize in proximity to the polymerization site where many chains are simultaneously produced and deposited, leading to a continuous, fine but crystalline filament. In these filaments, no amorphous regions have been observed, and disordered, or strained regions seem to be very small (Nishiyama et al. 2003b). Yet the notion of “degree of crystallinity” is still often used to interpret the diffraction profile to provide the relative masses of crystalline and amorphous material in a given sample. Based on NMR

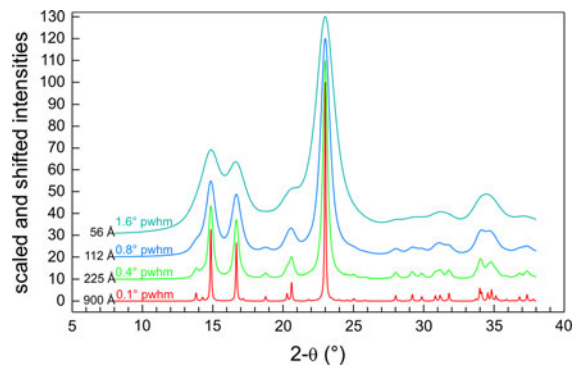


Fig. 1 Powder diffraction patterns calculated with Mercury software (Macrae et al. 2008) based on the cellulose β crystal structure (Nishiyama et al. 2002) and specified peak widths at half maximum (pwhm) ranging from 0.1° to 1.6°. Crystallite sizes are indicated, based on the Scherrer equation [(5), below] constant K of 1.0. In this case the high background at $2\text{-}\theta = 18.5^\circ$ for the upper curve (with 1.6° peak width at half height) arises only from the overlapping of broad peaks; no “amorphous” material is present. There are no peaks in the range of 8–13°, because of the implicit assumption of infinite crystallite size, unlike the patterns shown below in Fig. 3

experiments (Newman 1998), a kind of irreversible aggregation was proposed to occur during the pulping process, leading to a “paracrystalline” region that corresponds to a different notion from the paracrystallinity defined for polymer diffraction analysis (Hosemann 1962).

X-ray, neutron and electron diffraction structure refinements typically depend on defining a unit cell. That cell, with its constituent atoms, repeats on a lattice that extends to infinity, as a working approximation of the arrangement of atoms in the sample. This assumption is employed even for molecules such as polymers in which the crystal cross-sections are measured in nanometers. However, in such small systems, the surface molecules constitute a relatively large portion of the total number of molecules in the crystal and should therefore also contribute to the wide angle X-ray scattering or diffraction.¹ Because the surface molecules of crystals have considerably different environments than those on the interior, they are likely to deviate substantially from the structure of the crystalline core. That should affect the diffraction pattern. In this context, we felt a necessity to relate experimental diffraction data to different computer models proposed based on different techniques.

Approaches involving computerized molecular models are appealing. With just the known unit cell dimensions and crystal symmetry, even small models can successfully predict the hydrogen bonding and other packing details of cellulose crystal structures (Ford et al. 2005). Atomistic computer models of polysaccharide crystals can now be constructed that, at least in cross section, equal or exceed the apparent crystallite sizes in cellulose samples. These models can initially be based on repeated translations of the unit cell coordinates provided by determination of the crystal structure, or on other, more hypothetical arrangements. Ideally, a molecular dynamics (MD) simulation would reproduce the thermal motion of the atoms and give representative arrangements to the surface molecules and thermal disorder inside the crystals. Various MD studies (Heiner et al. 1995; Hardy and Sarko 1996; Kroon-Batenburg and Kroon 1997; Baker et al. 2000; Mazeau and Heux 2003; Yui

et al. 2010), have been reported (see Bergenstråhle et al. (2007) and Bellesia et al. (2010) for more thorough reviews) but the behavior of model cellulose crystals is quite force-field dependent. Each force field is rationalized and justified based on fits to experimental properties and/or quantum mechanics calculations, and should in principle predict the correct behavior of a molecule. However, because of the absence of a consensus among the results from different force fields, it is important to compare the outcome of molecular simulation with experimentally available data.

Besides the selected force field, another factor in the outcome is the type of model crystal. In the present work, we employ the “mini-crystal” method, in which the complete, finite crystal is described. Another approach, the “infinite crystal” method, relies on a typically smaller model that is repeated at periodic boundaries to calculate bulk effects. The latter approach can compensate for the relatively short chain length of the models (in our case 20 glucose residues) and avoids twisting and “uncontrolled edge effects” (Mazeau and Heux 2003). In fact, size, twisting, and edge effects are of major interest in the present work. Regarding length, many of the characteristics observed in the present models were similar to those observed with model chains of only eight residues (Nishiyama et al. 2008), so very long or infinite models were not deemed necessary for the present work. Long (not infinite) models would be needed, of course, if we were concentrating on the disorganized regions along the molecular length that give rise to a regular spacing (Nishiyama et al. 2003b).

Software for calculating diffraction patterns such as shown in Fig. 1 from the atomic coordinates within the repeating unit cell of periodic structures is readily available (Spek 2008; Macrae et al. 2008), but computer models lose their repeating unit cell and their periodicity, when subjected to energy minimization or MD. Therefore, software for calculating the diffraction pattern from finite, mini-crystal models must account for each atom’s position in the aperiodic computer model. The stumbling block was that such software was not available. Subsequently, suitable software for powder patterns became available (Wojdyr 2011), and one of us (YN) has developed software that can simulate the more-informative fiber patterns. These all-atom programs explicitly change the diffraction patterns to account for different values of the crystal length and width.

¹ In a crystal composed of a 10×10 array of polymeric molecules, 36 would be on the surface (36%). In a conventional, sub-millimeter size single-crystal for diffraction, the fraction of molecules on the surface is much less than one percent.

Herein we explore issues of crystallite size, modeling method, and the effects of water molecules on the calculated patterns. Our approach is mostly qualitative, in that it is based on visualization of calculated diffraction patterns and comparison with a few experimental patterns. It is in the trial-and-error spirit, extended from likely atom positions in the unit cell, to an explicit inclusion of all atoms in the proposed structure. In addition to learning more about crystalline cellulose, the software for non-periodic materials also provides the opportunity to better understand the nature of “amorphous” cellulose.

Materials and methods

Samples for diffraction

Highly crystalline cotton cellulose had been previously prepared from desized, scoured and bleached fabric. It was cut to 20 mesh in a Wiley mill and subjected to 2.5 N HCl at reflux for 40 min, a procedure that was optimized to give samples with the most crystalline diffraction pattern (Rowland et al. 1971). Scanning electron microscopy showed that some of the fiber cell wall structure remained.

Diffraction methods

The highly-crystalline powder sample from hydrolyzed cotton was mixed with a fine powder of CaCO₃ (as calibrant) crystals milled using an agate mortar and pestle. The mixture was put in a soda-glass capillary with 0.5 mm outer diameter. The sample capillary was placed on a collimated sample holder in a simple Warhus vacuum camera. A point focused CuK α radiation from an X-ray tube filtered with a Ni foil was used as the incident beam with various exposure times. The diffraction was recorded on a Fujifilm Imaging plate that was read with an Image plate reader BAS 1800II. Diffraction intensity profiles were obtained from the two-dimensional data by circularly integrating the intensities after correcting for the detector tilt using in-house software. The sample-to-detector distance was determined from the peak position of calcite at $d = 3.055 \text{ \AA}$. No background corrections were made but the intensity data were scaled as needed for optimal comparison with the calculated patterns. A diffraction pattern of a ramie

fiber bundle from ADF’s archives had been recorded with a precession camera and CuK α radiation. Also, a previously published (Nishiyama et al. 2002) synchrotron diffraction pattern from the crystal structure determination of tunicate cellulose I β was incorporated in a comparison figure.

Modeling methods

Models were constructed, typically with the Mercury software (Macrae et al. 2008), based on repetition of the dominant fraction in the cellulose I β unit cell (Nishiyama et al. 2002). For this work, the model crystals were built so that the (1 $\bar{1}$ 0) and (1 1 0) planes form the surfaces, widely thought to be a realistic representation (Nishiyama 2009). They were named “original” models. Both energy minimization (fully minimized in vacuum, $\epsilon = 1.0$) and MD simulations were performed using Amber9 or 10, and the Glycam04 (Kirschner and Woods 2001a; Basma et al. 2001; Kirschner and Woods 2001b) or Glycam06 (Kirschner et al. 2008) parameters (we did not find significant differences in the results for the various combinations.) MD simulations used the TIP3P water molecule (Jorgensen et al. 1983). Initially, water was added and the energy was minimized with the cellulose molecules held fixed (group 1). After minimization, the system was heated over 20 ps to 300 K with the cellulose tightly restrained (equilibrated at constant volume) and then equilibrated at constant pressure (groups 2 and 3). MD models were then simulated without restraints for 100 ps (group 4) although substantial alteration of the structure occurred almost instantly. After simulating in the presence of solvent, the cellulose was removed along with 0, 1 or 2 solvation shells. To clarify, the only difference within each of the four groups of MD models with a given number of cellulose chains was the number of solvent shells that were included when calculating the patterns. Including the model that retained the original coordinates, the minimized version thereof and the models from the MD runs, there were 14 different models for which patterns were calculated.

Calculated diffraction patterns

$$I(S) = \sum_m \sum_n f_m f_n \frac{\sin S r_{mn}}{S r_{mn}} \quad (1)$$

Powder patterns

The Debye software (Wojdyr 2011) that we used is based on the well-known (Alexander 1969) Debye scattering equation (1), where I is the intensity as a function of S . S is related to the scattering angle 2θ by $S = 4\pi(\sin\theta)/\lambda$, f_m and f_n are the atomic scattering factors for the m th and n th atoms, and r_{mn} is the distance separating them. Since the intensity is averaged out over all directions, the intensity profile is equivalent to a powder diffraction trace after Lorentz-polarization corrections. Alexander's 1969 book states that "in practice, this procedure is obviously limited to rather small atomic aggregates, even when electronic computers are used", referring to calculations such as those carried out by the Debye program. Forty years later, an ordinary desktop computer requires only a few minutes to calculate a pattern equivalent to a powder diffractometer trace for a large model crystal.

Fiber patterns

Fiber diffraction patterns were calculated with the programs, Calcdiff and Convolute. Convolute takes the output from Calcdiff and further distributes the calculated intensity to account for the crystallite orientation distribution angle. Since the models with which we are concerned do not have rigorous periodicity in any direction, as is the case for all mini-crystal approaches, we calculate the intensity in three-dimensional Cartesian reciprocal space. Then the average is spread over concentric circles so as to account for texture with fiber symmetry, instead of over concentric spheres that would be used in the case of powdered samples.

In Calcdiff, the scattering intensity of a group of atoms in scattering vector \mathbf{k} is expressed as (2)

$$I(\mathbf{k}) = |F(\mathbf{k})|^2 = \left| \sum_m f_m \exp(-i\mathbf{k} \cdot \mathbf{r}_m) \right|^2 \quad (2)$$

where F is the structure factor and f_m is the atomic scattering factor for atom m . Since we assume a fiber texture, the three-dimensional diffraction pattern will be cylindrically averaged around the fiber axis. This is done by converting the Cartesian scattering vector $\mathbf{k}(x, y, z)$ into cylindrical polar coordinates $\mathbf{k}'(R, Z, \phi)$, where

$$R = \sqrt{x^2 + y^2}$$

$$Z = z$$

$$\phi = \arccos(x/R)$$

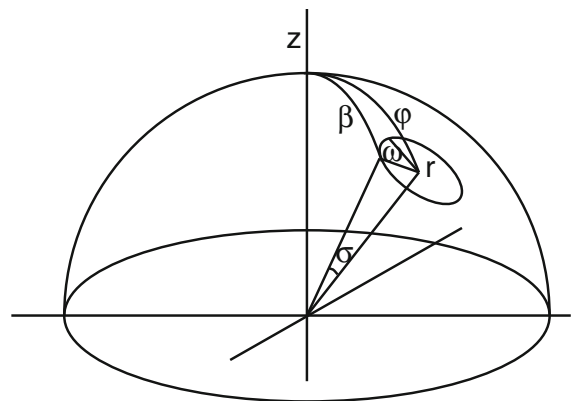
and the relative intensity

$$I(R, Z) = \int_0^\pi I(R, Z, \phi) d\phi$$

was calculated using the QAG algorithm that is implemented in the gnu scientific library (Galassi et al. 2009).

The fiber texture can be expressed as a probability, $P(\phi)$, to find the unique axis of a non-periodic object making an angle ϕ with respect to the macroscopic fiber axis. The Convolute program considers the orientation distribution of the object around the unique axis to be flat. The intensity intrinsic to the individual object $I(R, Z)$ is expressed in polar coordinates as $I(\sigma, r)$, where σ is the angle with respect to the chain axis and r is the distance from the center (see Scheme 1). Now the diffraction intensity at position \mathbf{r} in reciprocal space of the experimental frame due to the macroscopic fiber will be the contribution of intensity $I'(\sigma, r)$ of objects whose orientation makes an angle σ with vector \mathbf{r} . Introducing the angular parameter ω to describe the rotation around the vector \mathbf{r} , the intensity due to the ensemble of crystallites in the fiber can be described as (3)

$$I(\beta, r) = \int I(\sigma, r) \int P(\phi; \beta, \sigma, \omega) d\omega d\sigma \quad (3)$$



Scheme 1 Position of a reflection in reciprocal space due to a crystal inclined by angle ϕ with respect to the fiber axis (Z). The reflection makes an angle σ with the main axis of the crystal (r)

Integration over ω was conducted independent of r and the result was stored in memory as a function of β and σ . When the diffraction pattern is calculated in polar coordinates, the integration becomes a simple matrix operation. For a given resolution r , the array of intensity as a function of the azimuthal angle β will be (4)

$$I_{\beta r} = \sum_{\sigma} P_{\beta\sigma} I'_{\sigma r} \quad (4)$$

The program Convolute first calculates the matrix $P_{\beta\sigma}$ using an arbitrary orientation function (in this particular case, a Gaussian distribution function centered at the fiber axis z with a standard deviation of 5°) and stores it in memory. Then Convolute converts the cylindrically averaged individual object diffraction pattern into polar coordinates and applies the above matrix operation. Finally the intensities are remapped back to two-dimensional Cartesian coordinates.

Graphical representations

Powder diffraction intensities from experiment or the Mercury (Fig. 1) or Debyer (Figs. 3, 6) calculations were plotted with Slidewrite Plus (<http://www.slidewrite.com>). Fiber diffraction data from Calcdiff and Convolute were rendered with ImageJ (Rasband 2011) to resemble diffraction patterns collected on film or image detectors. Atomistic images of the model crystals were prepared with the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081) (Pettersen et al. 2004).

Results

Models

Figure 2 shows examples of different models used in the present report. The smallest crystal model involving four chains is featured for this purpose because it can be shown in greatest detail. The amount of disorder in the cellulose for the equilibrated model is typical of all models that had undergone that procedure, but the amount of twist in the fully minimized and in the MD models is exaggerated compared to that

in the larger models because of the small crystal size. Larger models with 6×6 and 12×12 chains are also shown, without hydrogen atoms or water.

Powder patterns

Figure 3 shows the powder diffraction patterns for models with 4, 16, 36, 64, 100, and 144 cellulose chains, calculated with the Debyer program. All of these patterns were normalized to have the same maximum intensity and then shifted by different amounts to allow presentation of the curves without excessive overlap. The upper set of patterns is for the bare original models, the middle set is for the models equilibrated at constant volume with one solvation shell, and the lower is for the MD models with two solvation shells. In particular the original models show significant low-angle scattering between $2-\theta = 10^\circ$ and $2-\theta = 13^\circ$. It dominates the pattern for the smallest models, producing opposite trends in the diffraction intensity for the 2×2 (4-chain) model compared to the other curves. Even the largest models are affected, with smaller oscillations obvious in the $2-\theta = 10\text{--}15^\circ$ range. These oscillations affect the apparent positions of the peaks, especially the $1\bar{1}0$ peak just before $2-\theta = 15^\circ$.

The curves in Fig. 3 for the models equilibrated at constant volume with one solvation shell indicate much less of the low-angle scattering than was seen for the upper set from original models. Therefore, the peak positions of the equilibrated models were taken as the standard ones. Examination of the other models in the set showed that the small-angle scattering was substantially diminished when two conditions were met: water had to be present, and there had to be some deviation of the structure from the original coordinates. Models with minimized water but the original cellulose coordinates retained the scattering above $2-\theta = 10^\circ$, as did the fully minimized, equilibrated and MD models with no water present. Because the water in the equilibrated models had little apparent effect on the pattern except for the reduction of the low-angle scattering, most of the focus of this report relies on those models and patterns in preference to the bare original models.

We do not fully understand the low-angle scattering or more precisely, the reduction of it by the combination of surrounding water and reduced order in the cellulose. It may be that the water provides a “matrix

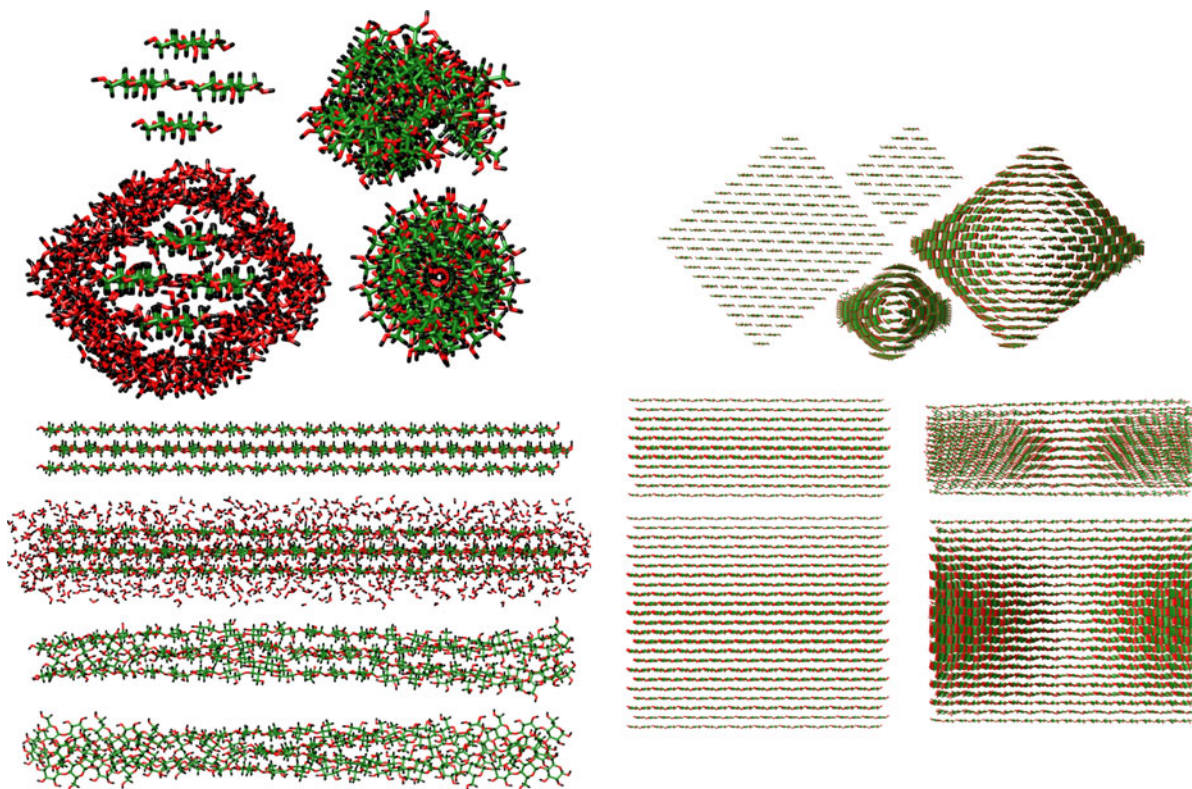


Fig. 2 Model crystallites used to calculate diffraction patterns. In the cluster of four models on the upper left are views looking down the chain axes of four-chain models. In that cluster, clockwise, from the *upper left*: a 2×2 (four-chain) model based on the original coordinates; a model resulting from MD simulation; an energy-minimized model, and an equilibrated restrained model with two solvation shells. Note the regularity

effect,” allowing a gradual decrease in electron density between the cellulose and the vacuum surrounding the model. We found similar low-angle scattering and reduction for hexagonal models also created in this project, on both the powder and fiber patterns. These peaks are clearly not Bragg peaks, as indicated by their absence in Fig. 1.

The bottom set of curves in Fig. 3 indicates that substantial changes have taken place because of MD simulation. Curves for the energy minimization models with no water were nearly identical, as were MD curves with one water (not shown). As seen in Fig. 2, the most obvious changes in the models are the twisting, especially extensive in the smaller models, and the surface disorder in the models subjected to MD. The effect of twist on the patterns is discussed further below. Despite the visually dramatic changes in these models from being subjected to the Glycam

of the energy-minimized model compared to the MD model. On the *lower left*, the four models show the same models perpendicular to the chain axes, with the energy-minimized structure last. The *upper right* four images are of 6×6 and 12×12 models, viewed down the chain axis, before and after energy minimization, and the four images at the *lower right* are views perpendicular to the molecular axes

force field, the effects on the diffraction patterns are relatively subtle. The 2×2 model curve is most affected and could be described as having more exaggerated peak broadening than when the 2×2 model was well-ordered. The other curves are also slightly broader, with the breadth differences diminishing for the larger crystals (see an analysis of peak breadths, below).

In the bottom group of Fig. 3, other than the substantially changed curve for the 2×2 model, the clearest differences from the top and middle groups of curves are shifts of the peak positions, including the $2\ 0\ 0$ peak at about $2\text{-}\theta = 23^\circ$. This results from changes in the intermolecular spacings. The contraction in spacing perpendicular to the fiber axes can be seen in Fig. 2 for the 6×6 and 12×12 models. In Nishiyama et al. (2008), distances were measured for a small model crystal that had undergone 10 ns of MD. Those

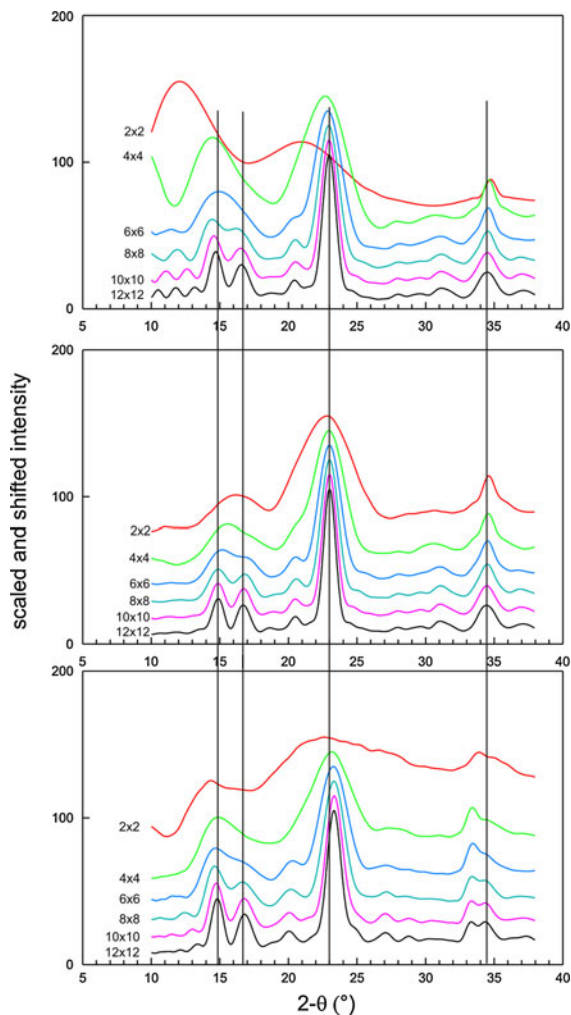


Fig. 3 Powder diffraction patterns calculated with Debye software based on models with 4, 16, 36, 64, 100, and 144 chains. Models in the upper set are from original $I\beta$ coordinates, the central set includes restrained equilibrated models with one solvation shell, and the lower set is for molecular dynamics models with two solvation shells. As these models progressed in size, there was a complete new outer layer of molecules

distances corresponded to what would have been the unit cell dimensions if periodic order had been preserved. Based on those dimensions, the 19-chain MD model in that work would have had a 200 peak at $2-\theta = 23.91^\circ$ instead of 23.01° . The movement of the 200 peak to larger $2-\theta$ values indicates a shorter distance between the (200) planes. A third manifestation is the intensity change in the range of $2-\theta = 33\text{--}35^\circ$. On the upper and middle sets of patterns, the peak at about $2-\theta = 34.5^\circ$ is actually a composite (see also Fig. 1 in the $33\text{--}35^\circ$ $2-\theta$ region).

Interpretation of these changes is aided substantially by the fiber diffraction patterns in Fig. 5. Here, on these powder patterns, 004 moves to the left by more than a degree. This is best detected on the powder patterns for the smallest models, where the other peaks are not well developed. The smaller $2-\theta$ values for the 004 peaks indicate a slightly longer chain length, also detectable by measuring the models in Fig. 2. This is also in agreement with Nishiyama et al. (2008), where the corresponding change was from the original 004 peak at $2-\theta = 34.56^\circ$ to the MD peak of $2-\theta = 33.89^\circ$. Because the a - and c -unit cell dimensions are shrinking and expanding, respectively, the contributors to the composite peak will move in opposite directions and break it up.

Powder peak profiles

The profiles of the 200 peak as calculated by the Debye program were assessed, both for the unscaled maximum peak height, and for the peak width at half height. Models with 100 chains in a 10×10 diamond array and 91 chains in a hexagonal array were examined. The 91-chain model consisted of one central (200) sheet of 11 chains, and two sheets of six to 10 chains arranged in decreasing size on either side of the central sheet. It had been constructed starting with an 11×11 diamond (similar to the models in Fig. 2) and the top and bottom 15 chains were removed.

Peak breadths can provide information on both the crystallite size and on the disorder. The Scherrer equation (5) (Patterson 1939) relates the minimum crystallite size τ to the breadth.

$$\tau = \frac{K\lambda}{\beta \cos \theta} \quad (5)$$

In (5), K is the “shape factor”, λ is the wavelength, β is the peak width at half height ($2-\theta$, expressed in radians), and θ is the position of the peak (half of the $2-\theta$ value). The value of K is often given values of about 0.9 but can range from 0.6 to 2.08. Other traits of the crystals that can broaden diffraction peaks include strain, faults and dislocations. In the present work, it is of interest to learn how much the disorder introduced by MD or minimization broadens the peaks.

Table 1 shows that the original models had the greatest peak heights, and the fully minimized and MD models had the lowest heights. Because the 91-chain model only has 11 rows of (200) sheets and the

Table 1 Raw heights and peak widths at half height

91-Chain hexagonal models			100-Chain diamond models		
Model ^a	Raw peak height	Peak width at half height 2- θ (°)	Model ^a	Raw peak height	Peak width at half height 2- θ (°)
orig	78.0	2.06	orig	104.4 ^b	1.43
min-ws0	77.2	2.06	min-ws0	103.5	1.43
min-ws1	76.0	2.04	min-ws1	100.8	1.43
eqp-ws0	75.5	2.05	eqp-ws0	100.6	1.42
eqv-ws0	74.2	2.06	eqv-ws0	99.3	1.43
eqp-ws1	71.0	1.98	eqp-ws1	90.7	1.42
eqv-ws1	69.9	1.99	min-ws2	90.5	1.45
min-ws2	64.1	2.23	eqv-ws1	89.4 ^b	1.42
eqp-ws2	62.6	2.08	eqp-ws2	82.4	1.44
eqv-ws2	61.5	2.09	eqv-ws2	81.4	1.44
md-ws0	59.4	2.25	md-ws0	77.5	1.57
full min	58.9	2.27	full min	75.5	1.62
md-ws1	55.4	2.16	md-ws1	69.5	1.56
md-ws2	47.9	2.33	md-ws2	62.9 ^b	1.59

^a Model name codes “orig” and “full min” refer to the model based on original coordinates and then fully minimized, respectively. “min” refers to models based on original coordinates but with minimized water, and “md” to models subject to unrestrained MD. “eqp” and “eqv” refer to models with restrained cellulose and unrestrained water, equilibrated at constant pressure and volume, respectively. The number of water shells is designated by “ws#” where # is either 0, 1 or 2. Peak widths are based on CuK α radiation, 1.5418 Å

^b Diffraction pattern displayed in Fig. 3

100-chain model has 19, the peak heights are higher for the 100-chain model. The ranking of the peak heights from the two models are mostly the same except for a switch of the ranking of the min-ws2 and the eqv-ws1 models. It is not easy to determine absolute peak heights from experiment.

The fully minimized and MD models generally had the greatest peak widths as well, with the exception for the 91-chain min-ws2 model. The peak widths were assessed by scaling the diffraction data so that the 2 0 0 peak maxima were 100 and then finding the locations on the 2- θ axis where the calculated intensities on either side of the peak were closest to 50. The data values before and after that point were used in a linear interpolation to provide the 2- θ values at 50% maximum height. The water shells generally decreased the maximum intensities, especially the second one. Besides disorder, another reason for the broader peaks for the MD and minimized models is their smaller crystallite size perpendicular to the (2 0 0) planes; the distance is compressed because of the Glycam force field. According to the Scherrer equation with a K of 1.0, the peak

width of 1.43° for the original 100-chain crystal corresponds to a size in the direction perpendicular to the (2 0 0) planes of 63 Å, and the size of the fully minimized crystal would be 56 Å. For the comparable 91-chain models, the values are 44 and 40 Å, respectively. The size (63 Å) from the original 100-chain crystal in the direction perpendicular to the (2 0 0) planes is smaller than the expected distance of 19 planes times the interplanar spacing of 3.9 Å = 74 Å, perhaps because some of the planes have very few cellulose molecules in them. At the top and bottom of the diamond (see Fig. 2, upper right) there are rows with one, two, or three chains for example. The agreement for the 91-chain model, with a minimum of six chains in its (2 0 0) planes, is closer ($11 \times 3.9 = 42.9$ Å).

Fiber diffraction patterns

Figure 4 shows a pattern calculated with Calcdiff, along with labels that indicate the equator, some of the layer lines (hk1, hk2), the meridian, as well as some of the important diffraction spots. Also indicated is low-

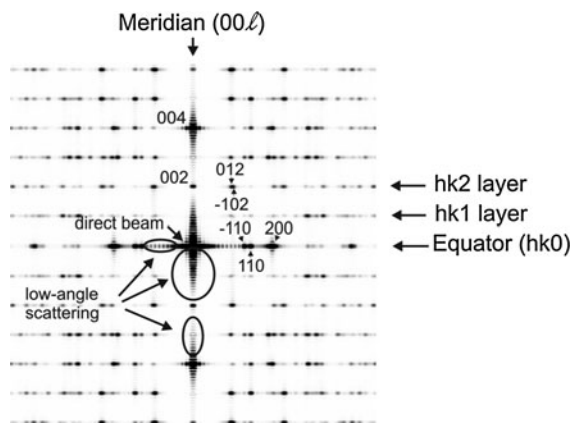


Fig. 4 Calculated diffraction pattern based on the 12×12 model with original $I\beta$ coordinates, using the Calcdiff program. The equator and meridian are labeled, as are the hk1 and hk2 layer lines. Low-angle scattering is indicated, as are some of the prominent reflections

angle scattering along the meridian, parallel to the fiber axis. The spacing in that low-angle scattering is reciprocal to the length of the crystal, some 104 \AA . There is also low-angle scattering along the equator. The output from Calcdiff that gave this pattern serves as input to a second program, Convolute, which adds the effect of the crystallite orientation distribution, as shown in the patterns of the following figure.

Figure 5 shows an array of calculated fiber diffraction patterns, each with a direct representation of reciprocal space. The patterns were calculated to appear as if they were made with a precession camera. Therefore, the layer lines are not curved as they would be if they were in “flat plate” patterns made with a conventional camera and Cu radiation. Also, the meridional reflections are all brought into the sphere of reflection. All were made with the fiber axes assumed to be vertical, with a standard deviation for the crystallite orientation of 5° , leading to the short arcs for the diffraction spots. The maximum intensity was adjusted in ImageJ to be 5% of the maximum value found in the region of the 200 reflection except for the patterns for the 2×2 models. Their maximum intensity was 10% of the maximum in the 200 region. Again, the models have 4, 16, 36, 64, 100, and 144 cellulose chains.

Each of the six individual images (Fig. 5a–f) is composed of two calculated patterns. The left sides are for the MD models, with one solvation shell, and the right sides are for the better-ordered, equilibrated models with one solvation shell. As mentioned above,

the 004 spots on the MD patterns are closer together than on the right side, indicating a longer 004 spacing and longer model crystal, since each model has 20 glucose residues.

One of the main features of interest in this series of calculated diffraction patterns is the development of crystallinity as the model size increases. The increased sharpness of the individual spots for the equilibrated models, compared to the MD models, is also apparent, especially for the smaller models. All of the patterns, except perhaps the one from the 2×2 MD model, have well developed layer lines that result from the 20-residue long model crystals. That length, more than 100 \AA , is longer than distances perpendicular to the molecular axis in all but the biggest of our models. Because the chains remained in an extended conformation and near periodicity is enforced in that direction by covalent bonds, the layer lines are reasonably well-developed. It is instructive to contrast the amount of information from the fiber pattern in Fig. 5a with that of the powder pattern in Fig. 3 for the same 2×2 models.

An important question about the crystallite surfaces regards how the surface molecules contribute to the diffraction pattern. How would the apparent crystallinity of an equilibrated model, with just four chains (our 2×2 model) compare with the MD model with 16 chains (our 4×4 model)? To an extent, the equilibrated 2×2 model could be taken as a crystalline core for the 4×4 MD model, with disordered surface molecules. As seen in Fig. 5, the left side of Fig. 5b (the MD model with 16 chains) is better resolved into individual reflections than is the right side of Fig. 5a (the equilibrated 2×2 model). As the model size increases, this effect seems to diminish, but the addition of surface chains enhances the resolution of the diffraction pattern instead of detracting from it. Quantitative comparisons are more difficult because of the discrepancies in spot positions and overlap because of the dimensional changes induced in the MD models by the force field.

Another point of interest concerns the development of discernable splitting of the $1\bar{1}0$ and 110 spots on the equator. In the powder patterns in Fig. 3, there is some separation in all three types of 8×8 and larger models. This is borne out in the fiber patterns as well. The splitting is one indicator of the size of cotton crystals, the diffraction patterns of which typically have clear separation of these two diffraction maxima.

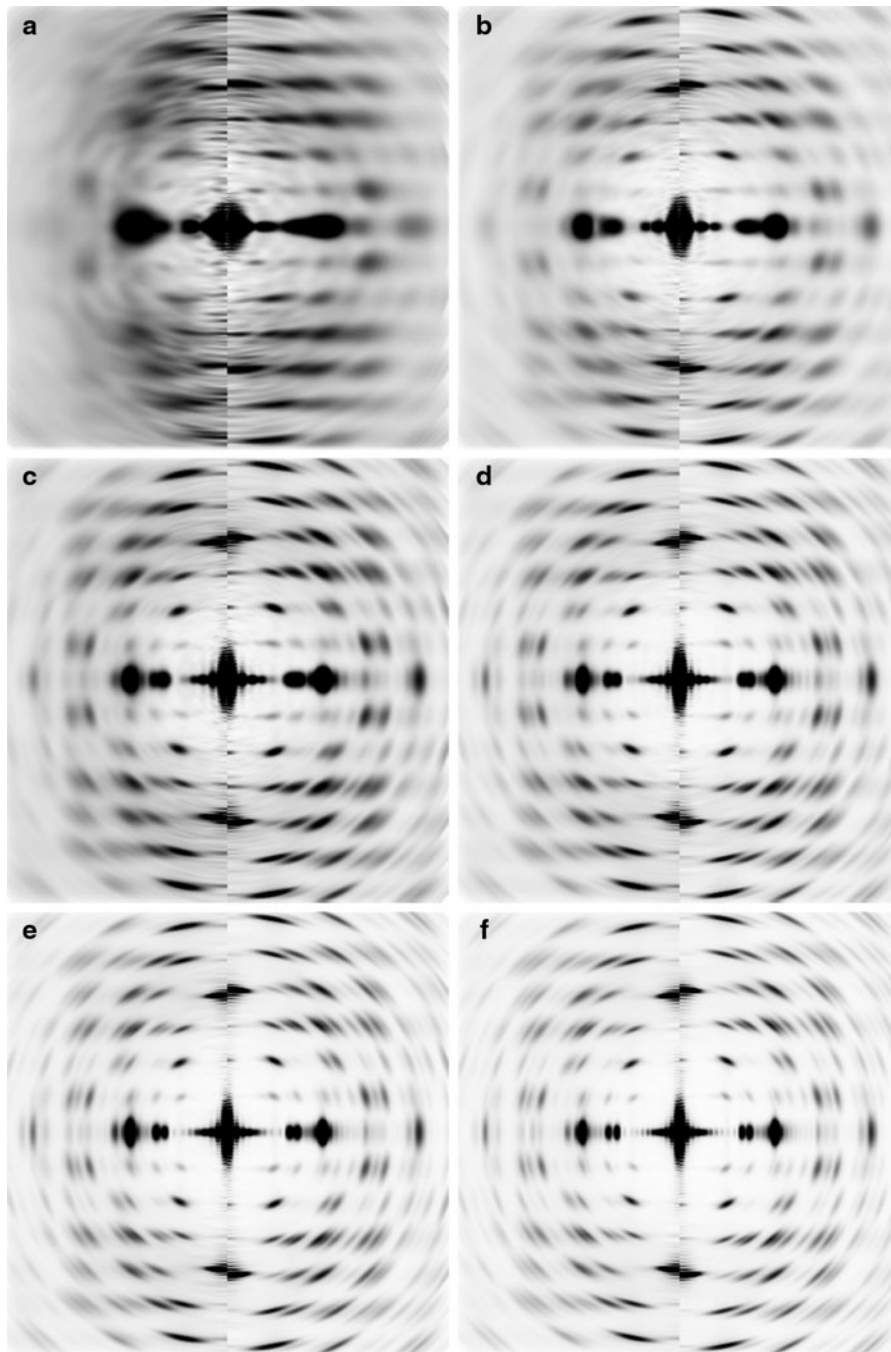


Fig. 5 a–f are calculated diffraction patterns after the use of the Convolvute program. The models have 4, 16, 36, 64, 100, and 144 chains, respectively. The *left half* of each *figure* is from an unrestrained MD run with one solvation shell, and the *right half* is for the same size model with restrained equilibration MD. By placing the pattern for the restrained model (e.g., the *right half* of

c) next to the next larger unrestrained MD pattern (the *left half* of d), the effect of adding a layer of surface chains to the crystal is visualized. Each *left-side* model, being under the control of the Glycam force field, has a longer *c*-axis spacing so the distances between its layer lines are slightly shorter

This observation alone indicates that the cotton cellulose crystallites, which give patterns with visibly diminished intensity between the two reflections, are at least as big as an 8×8 model. However, when drawing that conclusion, it must be remembered that any reduction of the monoclinic angle γ would increase the overlap (reduce the splitting) of these two reflections (Fernandes et al. 2011).

Comparison with experiment

Figure 6 compares an experimental powder pattern from microcrystalline cellulose of cotton with a pattern from a 10×10 model that was equilibrated and includes one solvation shell. The experimental pattern was recorded with CaCO_3 (calcite) powder for calibration, and the peak at 29.23° $2\text{-}\theta$ (3.055 \AA) is from that structure. The discrepancy of the 200 peak positions is due to the small difference in d -spacing, and thus the lattice parameters, between tunicin cellulose used for the model and cotton cellulose, which furnished the experimental pattern. The 200 d -spacing of cotton is reported to be about 1% larger than that of tunicin (Wada et al. 1997). This might be due to small disorder inside the crystal or due to limited crystal size that reduces the attractive long-range

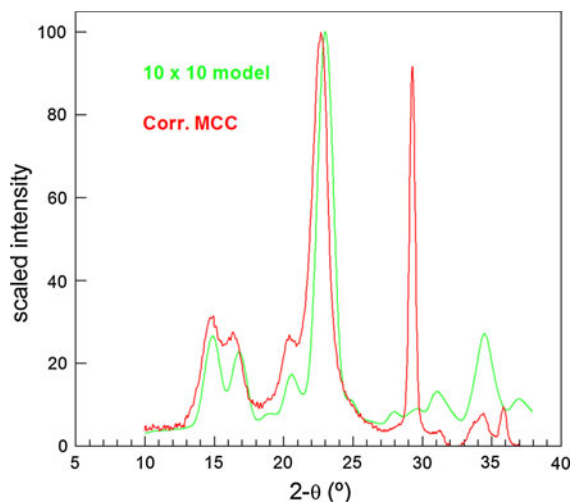


Fig. 6 Comparison of a cotton cellulose powder diffraction pattern of high crystallinity (Corr. MCC) with the pattern calculated for a 10×10 model based on tunicate cellulose coordinates. Also on the experimental pattern is the peak for the CaCO_3 calibration. The discrepancy in positions of the cellulose peaks is due to the difference in unit cell parameters for the two kinds of cellulose structures

London dispersion interactions. The observed and calculated peak widths at half-height for d_{200} are comparable. This is notably without correction for instrumental line-broadening, so the true cotton diffraction pattern's peak width must be somewhat narrower. Therefore, a 10×10 model should be the minimum size.

Figure 7 shows an archived experimental diffraction photograph from scoured and bleached ramie fibers, captured with a precession camera. Inset on this experimental pattern, in the lower right-hand quadrant, is a calculated pattern from an equilibrated 10×10 model with two solvation shells. The typical characteristics for cellulose I are present on the experimental pattern, such as the three very strong distinct spots on the equator ($1\bar{1}0$, 110 and 200) and the 004 meridional reflection that correspond to the peaks on the powder patterns in Fig. 1 at about $2\text{-}\theta = 14.5^\circ$, 16.5° , 23.0° and 34.5° , respectively. Wherever spots appear on the experimental pattern, they agree well with calculated ones. The experimental pattern is somewhat underexposed but shows a distinct separation between the $1\bar{1}0$ and 110 spots. It will be interesting to learn whether more modern diffraction technology can increase the number of

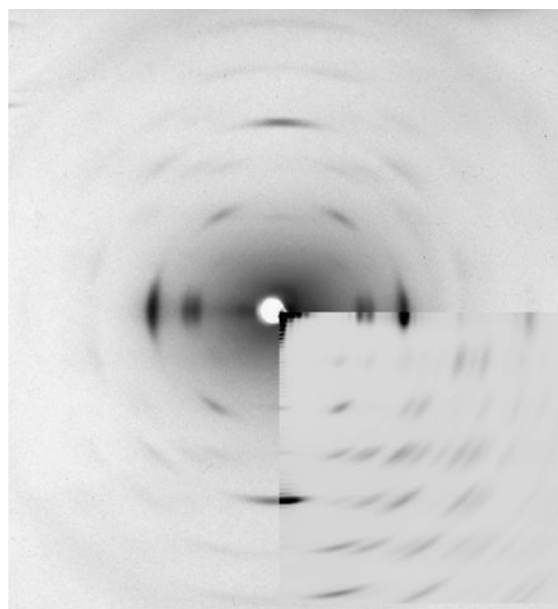


Fig. 7 Experimental ramie cellulose diffraction pattern taken with a precession camera. The *inset lower-right* quadrant is from a 10×10 restrained equilibrated model with two solvation shells

diffraction spots that are visible from such higher-plant cellulose samples, to be more comparable to the calculated diffraction pattern. One hope for better resolution of the weaker spots is the use of synchrotron radiation. However, the number of spots on a synchrotron X-ray pattern for ramie (Paul Langan, personal communication) is very similar to the number of spots in Fig. 7, and a reasonably important element could be missing from our models.

Figure 8 combines a quarter (upper left) of a synchrotron diffraction photograph of tunicate cellulose from the work in Nishiyama et al. (2002) and a quarter of a pattern (lower left) from that work based on the processed experimental data. The right half is a calculated pattern from our computer model based on a 13×13 array of chains that was equilibrated along with one solvation shell. There is substantial resemblance among the patterns although there are also differences. Not only do the left-side patterns extend to longer distances,² but the spots are better resolved, indicating that the computer model is too small. As in Figs. 4 and 5, the fine spacings on the equator and meridian of the calculated pattern are from low-angle scattering that arises from the finite size of the model and sharp cutoff. The left-side patterns exhibit some variation in the observed intensities for the equivalent upper and lower layer lines. For example, see the first few spots from the meridian on the upper and lower first and second layer lines. In any case, exact agreement with the calculated pattern cannot be expected because the discrepancy index (*R* factor) based on the structure factors (the intensity square roots) is 18.6% (Nishiyama et al. 2002).

Twisting

In the present work, the patterns calculated from models that had twisted, either from energy

² The range of the calculated diffraction was selected after consideration of the number of atoms in the model and the number of pixels for which the intensity must be calculated. These factors determine the computer time required for the calculation. Some of the software used for projects other than reported in this paper was limited in the sizes of data arrays that could be handled. Considering that some models were as large as 94,700 atoms and 18 different size crystals were modeled, each with 14 variations of water content, energy minimization and molecular dynamics, the selected step sizes of 0.003 S out to 0.597 S were considered adequate for the present purposes. Larger calculated patterns are definitely possible.

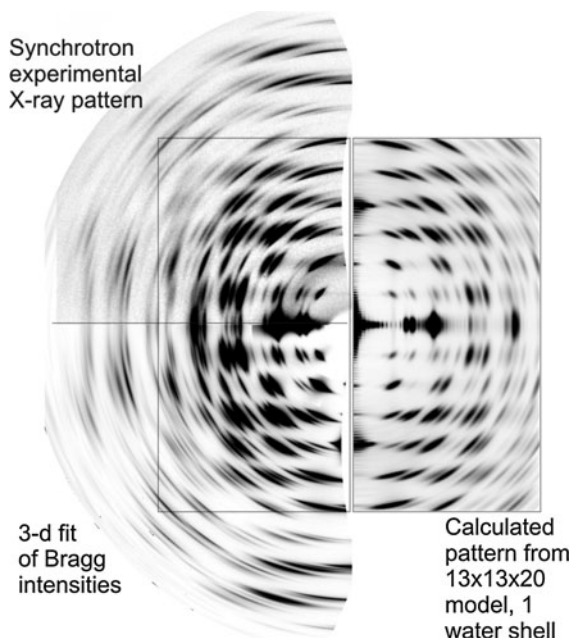


Fig. 8 Comparison of a calculated diffraction pattern from a 13×13 model (*right*) with experimental results for tunicate cellulose (*left*). The *upper left* quadrant is for the experimental data, and the *lower left* is for the processed experimental data. The *thin-line box* on the *left side* corresponds to the area of the calculated pattern on the *right*

minimization or MD, were not decidedly different from the original, untwisted models in their overall character. Any effect from twisting must be separated from the other effects of the Glycam force field, which, we note retains many details of the original crystal structure such as the hydrogen bonding system at room temperature. Just looking at the character of the spots regarding their sharpness or peak widths, there is little to distinguish the patterns from an untwisted model from those of a twisted model one size larger. The peak positions and relative intensities are, at least potentially, too affected by the small discrepancies in the model for making any decision on twisting. Any judgment based solely on breadth of the reflections would require exact knowledge of the number of chains in the crystal.

One difference that may ultimately be useful in resolving this question is shown in Fig. 9. It shows only lower right quarters of full diffraction patterns (no water present) before convolution, such as in Fig. 4. The image on the left is from an untwisted, 12×12 original model, whereas the image on the right is from a fully minimized and twisted 12×12

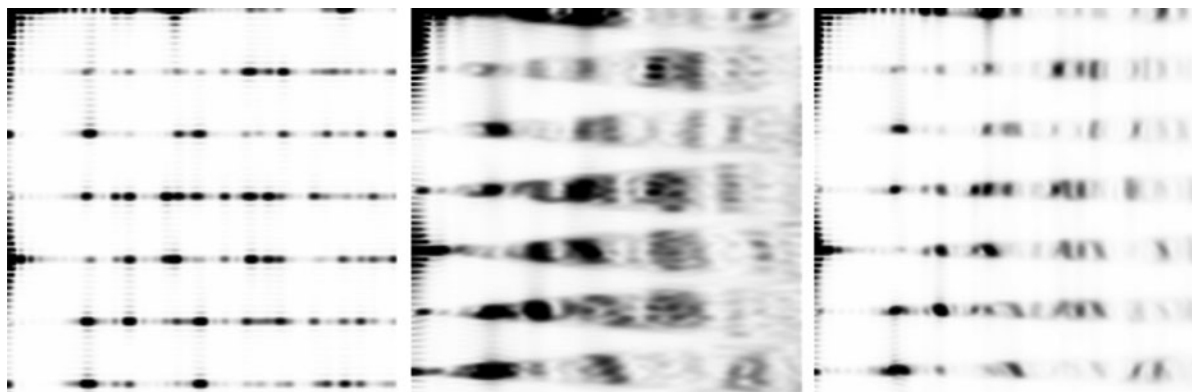


Fig. 9 Lower right quadrants of three diffraction patterns from Calcdiff, without Convolute. *Left*: the lower right quadrant of Fig. 4, a 12×12 model with original coordinates. *Right*: the pattern from a minimized 12×12 model, showing a slight

model. While the layer lines on the pattern from the untwisted model on the left have a constant height, the layer lines on the right have a slice-of-pie shape, with a decided slant to some of the spots. A similar pattern from a 6×6 fully minimized model (Fig. 9, center) has wedges with larger angles because the smaller crystal model is more twisted. Careful analysis of the shapes of arced reflections on experimental patterns could reveal whether the intensity distribution is resulting from a normal distribution of single fibril orientations or is a flatter composite of adjacent centers.

Proposal of Matthews et al

Besides the continual twisting of the model of Matthews et al. (2006), the unit cell dimensions and conformation of the central chain were substantially different from the accepted crystal structure. Namely, their CSFF force field had reduced the monoclinic angle from 96.5° to 90° and changed the *a*- and *b*-unit cell dimensions from 7.784 to 8.47 Å and from 8.201 to 8.112 Å, respectively, the former in response to a change in the O6 conformation on the central chain from the otherwise rarely observed *tg* orientation to the *gg* conformation. The *gg* conformation is often observed in crystals of related small molecules.

Their proposed structure was sketched with standard bond lengths and angles according to the published molecular drawings and geometrical data. The unit cell was propagated into a model 7×7 crystal otherwise similar to the original models herein.

slice-of-pie shape to the individual layer lines. *Center*: a pattern from a minimized 6×6 model (see Fig. 2). Its slice-of-pie shape is more exaggerated because the 6×6 model is more twisted than the 12×12 model

Its calculated diffraction pattern (Fig. 10a, right half) was compared with one from a 7×7 array based on the published $I\beta$ crystal structure (Fig. 10a, left side). The two sides of Fig. 10a are qualitatively and quantitatively different. Consider especially the differences in the meridional spots for the fifth and sixth layer lines that are strong on the pattern from our version of the structure of Matthews et al., but absent on the $I\beta$ pattern. Figure 10b shows patterns resulting from energy minimization with the Glycam force field of the two models, with subsequent crystal twisting. Again the patterns are quite different in detail.

Discussion

As mentioned in the introduction, the ability to calculate diffraction patterns is not novel. What moves the present work into a rare category is the calculation of diffraction patterns from models of cellulose that do not have a conventional unit cell, one that is periodically reproduced in all three dimensions to infinity (or at least thousands of Ångstroms.) In a recent effort on cellulose crystal size and shape, only the equatorial diffraction intensities were calculated (Newman 2008). Newman's pattern also showed substantial low-angle scattering. More recently, a paper that calculated a powder pattern based on the Debye scattering equation became available (Driemeier and Calligaris 2011). It concerns an elegant determination of the degree of crystallinity. The authors encountered several of the same technical issues such as preferred

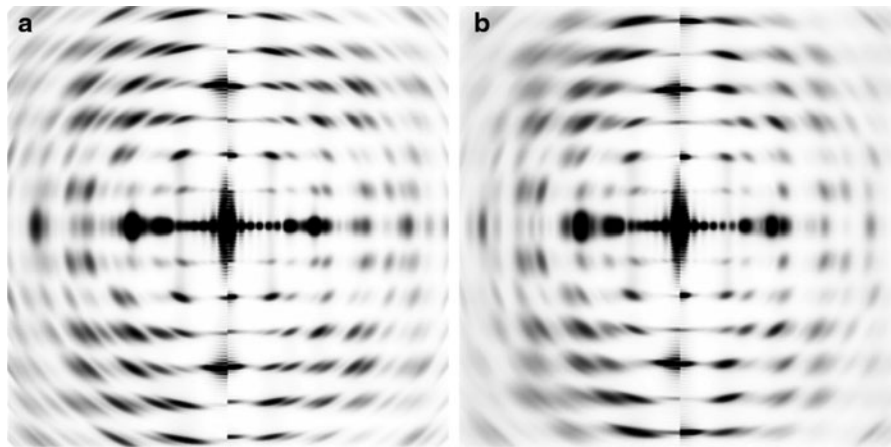


Fig. 10 Comparisons of patterns from 7×7 models based on the Nishiyama et al. (2002) $I\beta$ structure and the structure of Matthews et al. (2006). **a** The left side is from the original $I\beta$ coordinates, and the right side from coordinates from a sketch of

the Matthews et al. structure, incorporating the published geometry but using standard bond lengths and angles. Both models are periodic. **b** The energy-minimized, twisted (non-periodic) versions of the structures in **a**

orientation in the experimental samples, and low-angle scattering from the models (they used a model with propagated original coordinates and no water).

Models described in the present study were almost all based on the simple, nearly square “diamond” models that have surfaces parallel to the $(1\ 1\ 0)$ and $(1\ \bar{1}\ 0)$ diagonals of the unit cell, the accepted shape for some types of cellulose (Elazzouzi-Hafraoui et al. 2008). Other models tested in the present project were truncations of the models shown, with removal of equal numbers of chains near the top and bottom corners, parallel to the $(2\ 0\ 0)$ planes. Chain removal gives a hexagonal shape to the crystal cross section. Only one of those models, with 91-chains, contributed to the data reported herein. As a group, however, powder patterns of hexagonal structures showed lower, wider $2\ 0\ 0$ peaks relative to the diamond model crystals having similar total numbers of chains, perhaps to the extent that experimental patterns can be compared with the calculated ones for further guidance on this point. Other, more elaborate models for the crystallite size and arrangements have been proposed. In one (Ding and Himmel 2006), the microfibrillar unit is composed of seven 36-chain crystallites, each with six totally crystalline core chains, covered by two progressively less crystalline layers of chains. The tools used in the present effort to calculate diffraction patterns, or their successors, could be applied to test those and other more

complicated model structures. However, from the calculated patterns based on the Matthews et al. structure, it seems likely that the basic molecular structures will be more similar to the current $I\beta$ structure than the radically different models of Matthews et al. (2006).

A special case of complexity of the computer models is the possibility of twisting and bending the crystallites. In the present work, the twisting was a consequence of applying a molecular mechanics force field that also caused other minor changes in the structure. Further work is anticipated where the models will undergo the dimensional changes by the force field but not allowed to twist. Patterns from those models could be compared with patterns from twisted models. It may be that the slice-of-pie-shaped layer lines from twisted models can also help sort this problem out. Another approach would be to determine and refine the structure based on the calculated patterns from the twisted models. Would such a structure, based on the conventional assumption of periodicity in all three dimensions, fit the calculated pattern as well as fits of conventional structures to the experimental patterns?

Another concern is for the quality of the experimental results that will be combined with calculated diffraction patterns in future work. Many of the questions for such studies will depend on fairly subtle distinctions, and that will require highly accurate, well-resolved experiments. Consider our assessment

of the impact of the MD and energy-minimized models on the breadths of the 2 0 0 peaks. Additional broadening of the MD and minimized models compared to the original coordinate models was present, but in the rather small range of $0.2\text{--}0.4^\circ 2\text{-}\theta$.

For most conventional natural fiber samples, the microfibril organization is complex inside a (biological) cell that is typically tens of micrometers wide. Standard X-ray sources having a beam diameter of a few hundred μm will average out all orientations of the many fibers and thus result in a less-resolved fiber pattern. Advanced synchrotron sources with microfocus beam size are becoming more and more available. When highly oriented polymer segments can be probed using such beams, it will be possible to obtain more informative diffraction patterns. Even if the microfibril diameters were small, the diffraction pattern would only be blurred in the direction perpendicular to the microfibrils. Such highly oriented structures would give patterns that would be rich in information, as can be seen in Fig. 5a, b. Likewise, although the force field that we used is current, it would be helpful in these matters if the unit cell dimensions for the minimized and MD models were closer to those of the experiments.

Conclusions

Diffraction studies and computer modeling continue to be important methods in the study of polymers. In the case of cellulose, the details of the size and shape of the crystallites are especially important subjects. This work links these two realms. Powder patterns based on conventional software can almost instantly provide illustration of the relationship between crystallite size and diffraction for cellulose, requiring only the unit cell information. However, that software does not provide a way to evaluate non-periodic computer models for their diffraction patterns. Instead, software that considers as input all of the atoms in an entire computer model is needed. Some previous work with such powder patterns was acknowledged, but we believe that the simulated fiber diffraction patterns for cellulose models are without precedent. Likewise, these are the first, as far as we know, patterns based on model structures that were energy-minimized or subjected to MD.

The major focus of the present paper, other than to introduce the calculations of diffraction patterns from non-periodic structures, was to show the development of crystallinity in collections of cellulose chains. Although the sizes for crystallites from different sources have been determined previously, they are controversial, and the present approach allows explicit consideration of realistic environmental factors such as water. The MD models were more crystalline than the equilibrated models of the next smaller size. This argues against the proposed two-phase fibrillar structure with a crystalline core and completely disordered surface. On the other hand, the fact that the present calculations support crystallite sizes larger than the 36-chain structures that are widely considered to result from biosynthesis may be more of a reminder of the diversity of cellulose sources. Fernandes et al. (2011) concluded, based on NMR and infrared spectroscopic methods as well as low-angle neutron and wide-angle X-ray scattering that the crystallites in spruce wood consisted of only 24 chains. Their diffraction patterns have considerably broader peaks than the cotton, ramie, and tunicate patterns in the present work, however.

An important finding, based on the powder pattern line profile analysis, is that the distortions from MD or minimization of the model crystallites broaden the diffraction peaks and therefore indicate somewhat smaller crystallites than their models. In the case of the 100-chain diamond model, the apparent crystallite size was diminished by about 7 Å, roughly equivalent to two layers of cellulose chains. The 91-chain diamond model was similarly reduced, by about 4 Å, or one layer of chains. Also, the shape factor, $K = 1.0$ was apparently more appropriate for the hexagonal crystal because its calculated dimension was much closer to the Scherrer value than the diamond crystal's.

One of the major issues in the present work was the impact of low-angle scattering on the calculated diffraction patterns. To avoid substantial disruption of the pattern, we found that it was necessary to include both water molecules and some minor disorder in the atomic positions. This probably compensates for the fact that our computer models are essentially isolated single entities, with or without the water, whereas the experiments are done on many crystallites adjacent to each other.

Acknowledgments We thank Paul Langan and Henri Chanzy for helpful comments on a draft of this paper. Jodi Hadden contributed thoughts on twisted crystals.

References

- Alexander LE (1969) X-ray diffraction methods in polymer science. Wiley-Interscience, New York, p. 44 and Appendix 1, T-14
- Azáróff LV, Buerger MJ (1958) The powder method in X-ray crystallography. McGraw-Hill, New York, p 254
- Baker AA, Helbert W, Sugiyama J, Miles MJ (2000) New insight into cellulose structure by atomic force microscopy shows the I α crystal phase at near-atomic resolution. *Bio-phys J* 79:1139–1145
- Basma M, Sundara S, Calgan D, Vernali T, Woods RJ (2001) Solvated ensemble averaging in the calculation of partial atomic charges. *J Comput Chem* 22:1125–1137
- Bellesia G, Asztalos A, Shen T, Langan P, Redondo A, Gnana-karan S (2010) In silico studies of crystalline cellulose and its degradation by enzymes. *Acta Crystallogr D Biol Crystallogr* 66:1184–1188
- Bergensträhle M, Berglund LA, Mazeau K (2007) Thermal response in crystalline I β cellulose: a molecular dynamics study. *J Phys Chem B* 111:9138–9145
- Ding S-Y, Himmel ME (2006) The maize primary cell wall microfibril: a new model derived from direct visualization. *J Agric Food Chem* 54:597–606
- Driemeier C, Calligaris GA (2011) Theoretical and experimental developments for accurate determination of crystallinity of cellulose I materials. *J Appl Cryst* 44:184–192
- Elazzouzi-Hafraoui S, Nishiyama Y, Putaux JL, Heux L, Dubreuil F, Rochas C (2008) The shape and size distribution of crystalline nanoparticles prepared by acid hydrolysis of native cellulose. *Biomacromolecules* 9:57–65
- Fernandes AN, Thomas LH, Altaner CM, Callow P, Fosyth VT, Apperley DC, Kennedy CJ, Jarvis MC (2011) Nanostructure of cellulose microfibrils in spruce wood. *Proc Natl Acad Sci USA* 108:E1195–E1203
- Ford ZM, Stevens ED, Johnson GP, French AD (2005) Determining the crystal structure of cellulose IIII by modeling. *Carbohydr Res* 340:827–833
- Galassi M, Davies J, Theiler J, Gough B, Jungman G, Alken P, Booth M, Rossi F (2009) GNU scientific library. Network Theory Ltd., United Kingdom or see <http://www.gnu.org/s/gsl/>
- Hardy BJ, Sarko A (1996) Molecular dynamics simulations and diffraction-based analysis of the native cellulose fibre: structural modelling of the I- α and I- β phases and their interconversion. *Polymer* 37:1833–1839
- Heiner AP, Sugiyama J, Teleman O (1995) Crystalline cellulose I α and I β studied by molecular dynamics simulation. *Carbohydr Res* 273:207–223
- Hosemann R (1962) Crystallinity in high polymers, especially fibres. *Polymer* 3:349–392
- Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 79:926–935
- Kirschner KN, Woods RJ (2001a) Solvent interactions determine carbohydrate conformations. *Proc Natl Acad Sci USA* 98:10541–10545
- Kirschner KN, Woods RJ (2001b) Quantum mechanical study of the nonbonded forces in water-methanol complexes. *J Phys Chem A* 105:4150–4155
- Kirschner KN, Yongye AB, Tschampel SM, González-Outeirio J, Daniels CR, Foley BL, Woods RJ (2008) GLY-CAM06: a generalizable biomolecular force field, carbohydrates. *J Comput Chem* 29:622–655
- Kroon-Batenburg LMJ, Kroon J (1997) The crystal and molecular structures of cellulose I and II. *Glycoconj J* 14:677–690
- Langan P, Nishiyama Y, Chanzy H (2001) X-ray structure of mercerized cellulose II at 1 Å resolution. *Biomacromolecules* 2:410–416
- Macrae CF, Bruno IJ, Chisholm JA, Edgington PR, McCabe P, Pidcock E, Rodriguez-Monge L, Taylor R, van de Streek J, Wood PA (2008) Mercury CSD 2.0—new features for the visualization and investigation of crystal structures. *J Appl Crystallogr* 41:466–470
- 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
- Mazeau K, Heux L (2003) Molecular dynamics simulations of bulk native crystalline and amorphous structures of cellulose. *J Phys Chem B* 107:2394–2403
- Newman RH (1998) Evidence for assignment of ^{13}C NMR signals to cellulose crystallite surfaces in wood, pulp and isolated celluloses. *Holzforschung* 52:157–159
- Newman RH (2008) Simulation of X-ray diffractograms relevant to the purported polymorphs cellulose IV $_1$ and IV $_2$. *Cellulose* 15:769–778
- Nishiyama Y (2009) Structure and properties of the cellulose microfibril. *J Wood Sci* 55:241–249
- 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 (2003a) 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, Kim UJ, Kim DY, Katsumata KS, May RP, Langan P (2003b) Periodic disorder along ramie cellulose microfibrils. *Biomacromolecules* 4:1013–1017
- 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
- Nishiyama Y, Noishiki Y, Wada M (2011) X-ray structure of anhydrous β -chitin at 1 Å resolution. *Macromolecules* 44:950–957
- Patterson A (1939) The Scherrer formula for X-ray particle size determination. *Phys Rev* 56:978–982
- Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25:1605–1612

- Rasband WS (2011) ImageJ. U. S. National Institutes of Health, Bethesda, Maryland, USA. <http://imagej.nih.gov/ij/>
- Rowland SP, Roberts EJ, Bose JL (1971) Availability and state of order of hydroxyl groups on the surfaces of microstructural units of crystalline cotton cellulose. *J Polym Sci A-1* 9:1431–1440
- Spek AL (2008) PLATON, a multipurpose crystallographic tool. Utrecht University, Utrecht, The Netherlands. <http://www.cryst.chem.uu.nl/platon/pl000000.html>
- Wada M, Okano T, Sugiyama J (1997) Synchrotron-radiated X-ray and neutron diffraction study of native cellulose. *Cellulose* 4:221–232
- Wada M, Chanzy H, Nishiyama Y, Langan P (2004) Cellulose III_I crystal structure and hydrogen bonding by synchrotron X-ray and neutron fiber diffraction. *Macromolecules* 37:8548–8555
- Wada M, Heux L, Nishiyama Y, Langan P (2009) X-ray crystallographic, scanning microprobe X-ray diffraction, and cross-polarized/magic angle spinning ¹³C NMR studies of the structure of cellulose III_{II}. *Biomacromolecules* 10:302–309
- Wojdyr M (2011) <http://www.unipress.waw.pl/debyer> and <http://code.google.com/p/debyer/wiki/debyer>
- Yui T, Okayama N, Hayashi S (2010) Structure conversions of cellulose III_I crystal models in solution state: a molecular dynamics study. *Cellulose* 17:679–691