

# Evaluation of reactive force fields for prediction of the thermo-mechanical properties of cellulose I<sub>β</sub>

Fernando L. Dri<sup>a</sup>, Xiawa Wu<sup>b</sup>, Robert J. Moon<sup>d</sup>, Ashlie Martini<sup>c,\*</sup>, Pablo D. Zavattieri<sup>a,\*</sup>

<sup>a</sup> Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>b</sup> School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

<sup>c</sup> School of Engineering, University of California Merced, Merced, CA 95343, USA

<sup>d</sup> USDA Forest Service, Forest Products Laboratory, Madison, WI 53726, USA

## ARTICLE INFO

### Article history:

Received 5 April 2015

Received in revised form 24 June 2015

Accepted 26 June 2015

### Keywords:

Cellulose I<sub>β</sub>

Molecular dynamics

Force fields

Thermo-mechanical behavior

Anisotropy

## ABSTRACT

Molecular dynamics simulation is commonly used to study the properties of nanocellulose-based materials at the atomic scale. It is well known that the accuracy of these simulations strongly depends on the force field that describes energetic interactions. However, since there is no force field developed specifically for cellulose, researchers utilize models parameterized for other materials. In this work, we evaluate three reactive force field (ReaxFF) parameter sets and compare them with two commonly-used non-reactive force fields (COMPASS and GLYCAM) in terms of their ability to predict lattice parameters, elastic constants, coefficients of thermal expansion, and the anisotropy of cellulose I<sub>β</sub>. We find that none is able to accurately predict these properties. However, for future studies focused on a given property, this paper presents the information needed to identify the force field that will yield the most accurate results.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Representing about  $1.5 \times 10^{12}$  tons of the total annual biomass production [1], cellulose is considered an almost inexhaustible source of raw material that can potentially meet the increasing demand for environmentally friendly and biocompatible products [2]. During biosynthesis, multiple cellulose chains form bundles, called cellulose fibrils, which have regions where the cellulose chains arrange with a high degree of order (crystalline-like), and regions that are disordered (amorphous-like). The intra- and inter-chain hydrogen bonding network makes cellulose a relatively stable polymer, and gives the cellulose fibrils high axial stiffness and improved thermal stability. Naturally occurring crystalline cellulose co-exists in two polymorphs, I<sub>α</sub> and I<sub>β</sub>, with the latter being the most stable phase [3–5] and the focus of this study. The complicated crystalline structure of cellulose presents significant challenges for experimental characterization due to difficulties in testing, propagation of uncertainties in these tests [6], and variations in the crystalline cellulose being tested (e.g. crystal structure, defects, percent crystallinity) [3]. Atomistic simulation of cellulose can complement experiments by providing

insights into the molecular-scale origins of material properties, some of which cannot be analyzed with current experimental techniques [7,8]. This is particularly true for cellulose nanomaterials whose very small size make it exceedingly challenging to experimentally measure properties [3].

First principles density functional theory (QM-DFT), as well as molecular dynamics (MD) simulations, can be used to probe the crystalline structure and its response to external stresses. Dri et al. [7,8] predicted the temperature dependence of the crystalline structure, its thermodynamic and elastic properties and anisotropy of cellulose I<sub>β</sub>, using QM-DFT with a semi-empirical correction for van der Waals interactions [9]. Although recent computational advances enable larger simulation sizes with QM-DFT, the use of MD with classical potential energy functions (force fields) is often necessary to reach the relevant temporal and spatial scales of many research questions [10]. A force field (FF) is characterized by parameters that are fit to enable the prediction of certain material properties. The fitting process is typically called parameterization and the resultant parameters referred to as a parameter set. The development of a FF and parameterization for a given material is non-trivial [11]. Some simple FFs may be extended to a diverse set of systems (e.g., Lennard-Jones potential) but are not expected to yield quantitatively accurate results. On the other hand, more complex FFs, although limited to a small number of similar materials (e.g. Embedded Atom Method potentials are typically suitable

\* Corresponding authors.

E-mail addresses: [amartini@ucmerced.edu](mailto:amartini@ucmerced.edu) (A. Martini), [zavattieri@purdue.edu](mailto:zavattieri@purdue.edu) (P.D. Zavattieri).

only for metals), may be better able to provide quantitative accounts of molecular structure and conformation [12].

In this study, different reactive FF parameter sets are assessed with respect to their ability to accurately describe the thermo-mechanical behavior and anisotropy of cellulose  $I_\beta$ , and compare them with two commonly-used non-reactive force fields, COMPASS and GLYCAM. The reactive FFs are based on ReaxFF [13], with three different parameter sets (ReaxFF\_Mattsson [14], ReaxFF\_Chenoweth [15] and ReaxFF\_Rahaman [16]). The reactive FF has the ability to simulate bond forming and breaking, which is useful (and in some cases necessary) for simulations where the property being studied is dependent on the formation or breaking of covalent bonds. Unfortunately, none of the available parameterizations was originally generated to model crystalline cellulose, limiting their accuracy and ability to predict some key thermo-mechanical behaviors.

By comparison, non-reactive FFs are more computationally efficient but the bonding state of atoms is not captured such that covalent bonds cannot be broken or formed during the simulation. COMPASS [17] is a Class II force field, used previously to compute elastic parameters for crystalline cellulose [18,19]. GLYCAM [20] is a Class I force field, used in the literature to compute cellulose crystalline structure and thermal expansion [21–23]. There are other (non-reactive) force fields used in the literature, such as CHARMM and GROMOS, that have been used to study the structure and mechanical properties of amorphous and crystalline cellulose [24,25]. A previous study compared the predicted lattice parameters of cellulose microfibrils in water using CHARMM, GLYCAM and GROMOS [10]. With the exception of one lattice angle that was predicted with 27% error with respect to the experimental value, the authors reported that these three non-reactive FFs have similar characteristics. As such, only GLYCAM and COMPASS will be evaluated in our study.

In this work, we evaluate the accuracy of MD simulation with these different reactive and non-reactive FFs and parameter sets by comparing their predictions to QM-DFT calculations and experimental measurements. The FFs are evaluated in terms of their ability to predict a range of different properties, including lattice parameters, elasticity tensors and thermal properties of crystalline cellulose. This comparison will allow us to determine which FF best describes individual properties and assist researchers in selecting the most appropriate FF and parameterization for a given simulation-based study.

## 2. Computational methodology

### 2.1. Crystalline cellulose model

Cellulose ( $[C_6H_{10}O_5]_n$ ) is an organic compound that can be described as a linear chain of glucose rings with a flat ribbon-like conformation. Each chain is formed by one-hundred to over ten-thousand  $\beta$  (1  $\rightarrow$  4) linked D-glucose units; van der Waals (vdW) and intermolecular hydrogen bonds (H-bonds) promote parallel stacking of multiple cellulose chains forming a crystalline structure [3]. The most basic classification method divides cellulose types into 4 basic polymorphs that are identified as I, II, III or IV each one having their own subtypes [3]. Cellulose I, also called native cellulose has a mix of two polymorphs, viz., cellulose  $I_\alpha$ , which is a triclinic structure, and  $I_\beta$ , with a monoclinic structure, that coexist in various proportions depending on its source [26,27]. The  $I_\alpha$  structure is the dominate polymorph for most algae and bacteria, whereas  $I_\beta$  is the dominant polymorph for higher-plant cell wall cellulose and in tunicates [3,5].

The crystal and molecular structure together with the hydrogen bonding system in cellulose  $I_\beta$  has been accurately characterized by

Nishiyama and co-workers [26–30]. In this work, we adopt the atom coordinates for cellulose  $I_\beta$  network A reported by Nishiyama et al. [26]. In order to account for the atom positions inside a unit cell, we take advantage of the symmetry and antisymmetry operations provided by crystallographic space groups. For instance, for cellulose  $I_\beta$ , the space group is commonly accepted to be  $P2_1$  (#4) [31]. Fig. 1 depicts the crystalline structure reported by Nishiyama et al. [26] after the symmetry operations are applied to the original atom coordinates. Cellulose  $I_\beta$  unit cell was generated by arranging two parallel cellulose chains (as opposed to antiparallel), one positioned at the corner (origin chain) and the other at the center of the unit cell (center chain). The center chain is shifted by  $c/4$  relative to the corner chain in the axial direction. Fig. 1 also includes a Cartesian system of coordinates with axes 1, 2 and 3 which will be useful for our discussion. Direction 1 is chosen to be parallel to  $a$ , and direction 3 is parallel to  $c$ . For the monoclinic  $P2_1$  structure,  $b$  is not orthogonal to  $a$ . Therefore, direction 2 is chosen such that it is orthogonal to directions 1 and 3.

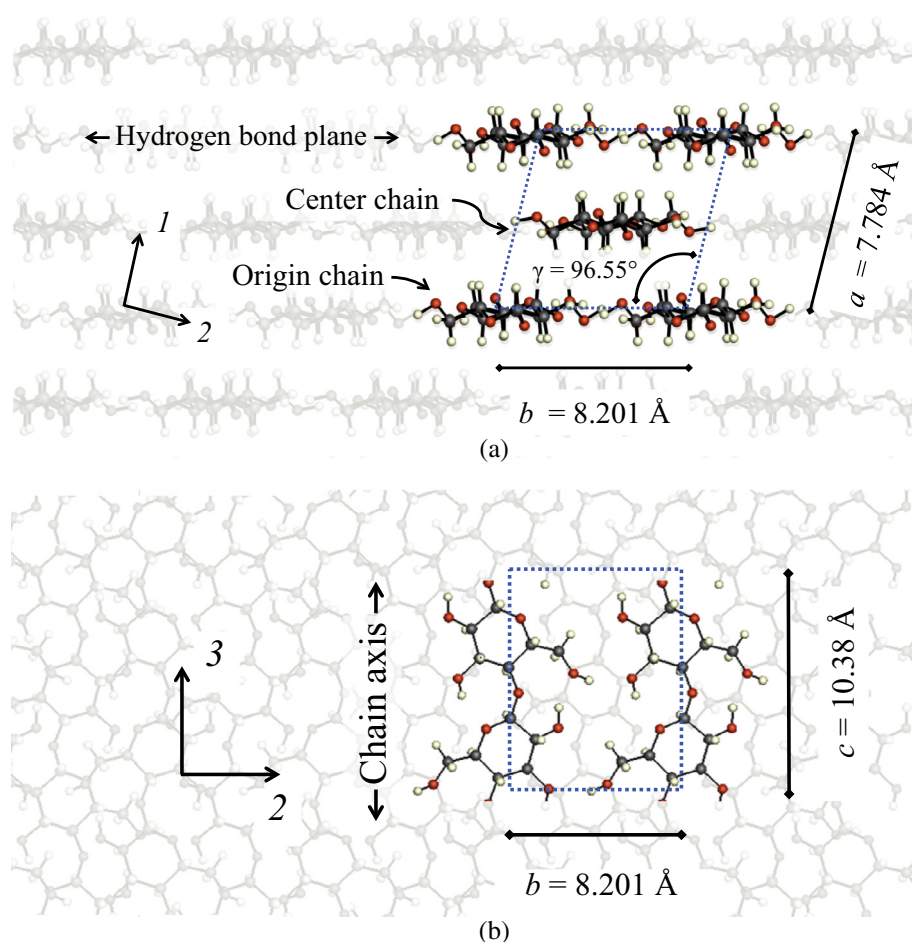
A further classification of cellulose I can be based on the H-bond network patterns, A and B, proposed by Nishiyama [26]. The relative occupancies of the two networks are different according to the polymorph: network A is  $\sim$ 70–80% occupied in  $I_\beta$  but only  $\sim$ 55% occupied in  $I_\alpha$  [27,29]. This work focuses on cellulose  $I_\beta$  with network A since it is believed to be one of the most commonly occurring polymorphs [3]. Fig. 2 provides a schematic representation of the H-bond network A reported in [29,33] for both origin and center chains. For this work, this structure was constructed with Materials Studio [34], and the Crystalline cellulose – atomistic toolkit using the Crystalline cellulose-atomistic toolkit [32] in NanoHUB.org, which is publicly available.

In this work, we modeled a periodic structure. Although use of periodic boundary conditions precludes the simulations from capturing surface effects, we applied them here for three reasons. First, periodic boundary conditions enabled us to effectively model a much larger system. This is particularly important in the chain direction since the crystals are usually 50–500 nm long [3], leading to fully atomistic model that requires extremely long computational times. Second, periodic boundaries provide a numerical means of applying strain without imposing artificial constraints on the chains themselves since the boundaries themselves can be changed to impose strain [35]. Lastly, the periodic model was used because it enabled us to isolate the bulk material response from that due to the crystal surfaces. The effect of surfaces on mechanical properties depends on surface-to-volume ratio and surface chemistry, which in turn depend on the source of the crystal and it is not straightforward to measure experimentally. Therefore, we focus here on the bulk material response, which was achieved through the use of periodic boundary conditions. This may, however, result in differences between model-predicted material properties and those measured experimentally for crystals with a large surface-to-volume ratio.

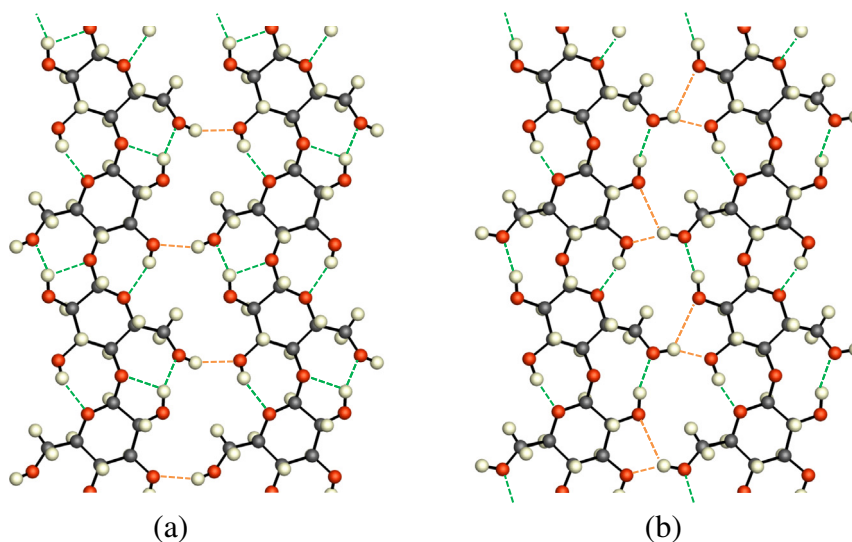
### 2.2. Force fields

LAMMPS simulation software [36] was used to compare the ReaxFF parameter sets with COMPASS and GLYCAM in terms of their ability to accurately represent cellulose  $I_\beta$  under different simulation conditions. All the force fields include bonded and non-bonded interactions, e.g., covalent bonds, covalent bond angles and torsions, vdW and Coulomb interactions. Multiple parameterizations exist for ReaxFF for different materials. In this work we used three different ReaxFF parameterizations (ReaxFF\_Mattsson [14], ReaxFF\_Chenoweth [15] and ReaxFF\_Rahaman [16]).

Previous studies have shown the important role of H-bonding on Cellulose I crystalline stability and properties [10,21,27,30,37,38]. A H-bond is a short-range, angularly dependent



**Fig. 1.** Expanded views of the  $P2_1$  unit cell structure of cellulose  $I_\beta$  network A showing the characteristic layered conformation [32]. Experimental (room temperature) lattice parameters  $a$ ,  $b$ ,  $c$ , from Nishiyama et al. [26] are shown. Red spheres denote oxygen ions, gray spheres represent carbon ions and white spheres represent hydrogen ions. Dotted blue lines denote the unit cell. (a) View along the  $c$ -axis (into the page). Layers of  $I_\beta$  are stacked along the  $a$ -axis. (b) View along the  $a$ -axis direction. Atomic coordinates were obtained after applying symmetry operations to the original structure reported by Nishiyama et al. [26]. Cartesian system coordinates (1, 2, 3) are superimposed for reference. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 2.** Hydrogen-bonding patterns in cellulose  $I_\beta$  network A. Intra- and inter-molecular hydrogen bonds are depicted in green and orange, respectively. (a) Chains at the origin of the unit cell and (b) chains at the center of the unit cell as reported in [29,33]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

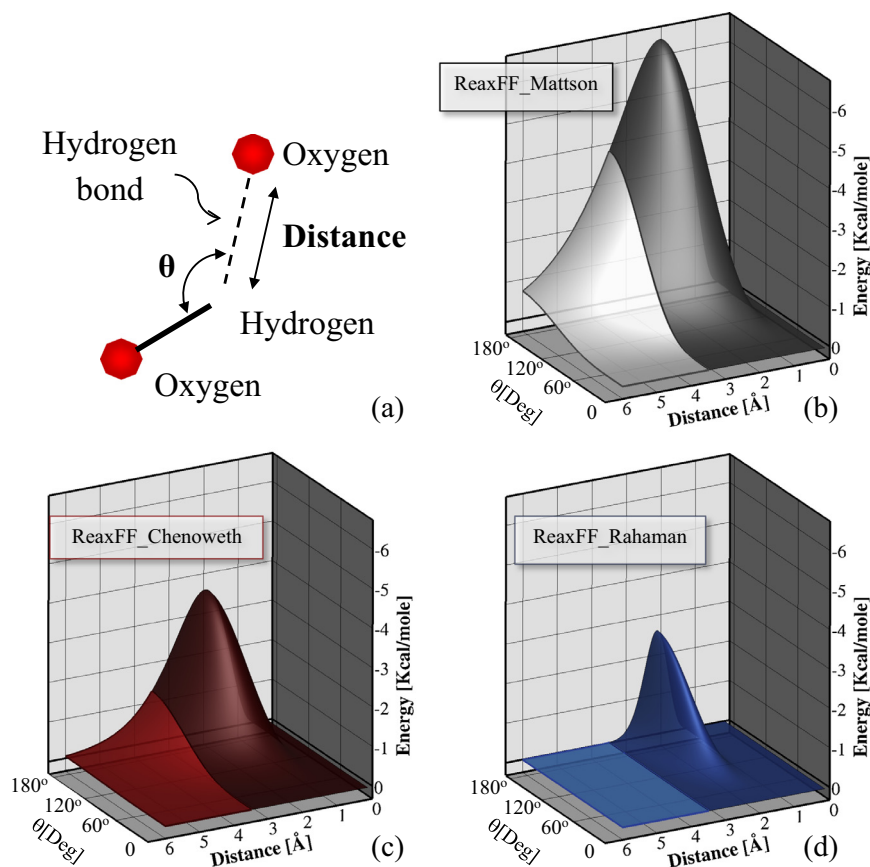
interaction between a small electronegative donor atom (such as oxygen, nitrogen, or fluorine) that has covalently a bonded hydrogen atom and an electronegative acceptor atom. This interaction is mostly polar, but there is a partial covalent character that is strongest when the donor-hydrogen – acceptor angle is nearly linear ( $D-H-A = 180^\circ$ ) [39]. Long range interactions are treated differently by each FF. COMPASS and GLYCAM (and most other similar non-reactive FFs) use an implicit representation of H-bonds where their effect is integrated into the electrostatic and vdW interaction terms. In such case, a common cutoff distance is applied to vdW, Coulomb and H-bonds. This cutoff distance is typically 10 Å, which is the value adopted in this work. In contrast, ReaxFF has an explicit description of H-bonds with input parameters that define this type of interaction. As a result, ReaxFF can provide more information about the intra- and inter-chain hydrogen bonding network in the cellulose crystal but the results are susceptible to the FF parameterization used. Although the H-bond cutoff distance can be specified in ReaxFF, there is still no clear indication of what its value should be. H-bonds are usually defined as having the electronegative donor and acceptor atoms less than 3.5 Å apart and with a  $D-H-A$  angle of greater than  $120^\circ$ . Matthews et al. [39] reported  $D-H-A$  angles greater than  $100^\circ$  and distances up to 4 Å for MD simulations of cellulose  $I_\beta$ . These results appear to be contradicted by experimental data reported by Nishiyama et al. [30] where a H-bond survey of cellulose  $I_\beta$  reveals angles between  $108^\circ$  and  $170^\circ$  and distances between 1.6 and 2.8 Å.

Fig. 3 shows the hydrogen bond energy surface for each of the ReaxFF parameterizations as a function of the distance and the

angle between the hydrogen atom and the acceptor atom. The following two H-bond cutoff distance values were selected after careful examination of the hydrogen bond energy surfaces: (i) A cutoff value of 3.5 Å, which coincides with the standard definition of hydrogen bond interactions [40] and forces the numeric simulation to comply with the experimental results reported in [30]; and (ii) A cutoff value of 6.0 Å, which is considered to be large enough to account for the entire bond energy as shown in Fig 3. It is worthwhile mentioning that the default cutoff distance value for H-bond interactions in LAMMPS is 6.0 Å. However, for ReaxFF\_Rahaman [16], the H-bond interaction is already negligible at 3.5 Å (see Fig. 3c), and as a result, no difference between 3.5 and 6.0 Å cutoff distances should be expected for this particular FF.

### 2.3. Equilibration and lattice constant calculation

An initial equilibration procedure is performed to obtain the crystal structure of cellulose  $I_\beta$ . First, a unit cell is built based on the experimental measurements by Nishiyama et al. [26]. An initial Gaussian velocity distribution is imposed on the atoms to produce an equivalent 300 K temperature. The unit cell is then repeated  $4 \times 4 \times 8$  times in the  $a$ ,  $b$  and  $c$  directions, respectively, to create a simulation cell. Subsequently, this simulation cell is equilibrated in a canonical ensemble for 50 ps with a time step 0.25 fs. Finally the simulation is coupled with a thermal bath at 300 K controlled by the Nosé–Hoover thermostat. This equilibration process allows relaxing inter-atomic stress without changing the size of the simulation box. A second equilibration is conducted in an



**Fig. 3.** Hydrogen bond energy surface as a function of the distance and angle for each of the ReaxFF parameterizations: ReaxFF\_Mattsson [14], ReaxFF\_Chenoweth [15] and ReaxFF\_Rahaman [16]. These surfaces are shown in dark color up to a cutoff distance of 3.5 Å. The light color denotes the continuation of the surface to larger values of the distance. The inset in the upper left shows the definition of distance and angle for an O–H–O H-bond. All the figures are plotted in the same scale. We also note the differences in the H-bond energy assigned to the interaction by each of the parameterizations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

isothermal–isobaric ensemble at temperature 300 K and pressure 1 atm, also controlled by Nosé–Hoover thermostat and barostat methods, for 300 ps with a time step of 0.25 fs. This equilibration process relaxes the simulation box as well as the atomic configurations under 1 atm pressure. The dimensions of the simulation cell are averaged over the last 10 ps in the second equilibration process in order to calculate the lattice constants of cellulose  $I_\beta$ .

#### 2.4. Elasticity calculation

Molecular mechanics are used to predict the stiffness matrix of cellulose  $I_\beta$  with different force fields. After equilibrating the simulation cell in the isothermal–isobaric ensemble, we increase its size in one direction through successive small length steps (e.g., elongate in the 1-direction by 0.2%) while keeping the other two directions fixed. The simulation cell then undergoes an energy minimization process with the conjugate gradient (CG) method to allow it to reach its minimum energy state. The elongation and minimization processes are repeated until the total strain in the extending direction reaches 4%. The strain and stress values at each step are recorded and a linear fit of the strain–stress relationship provides the stress vectors corresponding to the strain. The same procedure is performed in the orthogonal directions, 1, 2 and 3, as well as the shear directions, 12, 13 and 23 (see Fig. 4). After all six simulations, we obtain the stiffness matrix that relates the strain and stress as following:  $\sigma_i = C_{ij}\epsilon_j$  where  $\sigma$  is stress and  $\epsilon$  is strain. The inverse of the stiffness matrix  $C_{ij}$  is the compliance matrix  $S_{ij}$ . The Young's moduli in the 1, 2 and 3 directions can be calculated as  $E_{11} = 1/S_{11}$ ,  $E_{22} = 1/S_{22}$  and  $E_{33} = 1/S_{33}$ . The stiffness matrix is calculated for all reactive and non-reactive force fields.

#### 2.5. Thermal expansion calculation

The thermal expansion coefficients of cellulose  $I_\beta$  in different directions relative to the crystallographic structure are calculated with the simulation cell equilibrated at different temperatures. We study the response of the simulation cell to heating from 200 K to 500 K with a temperature interval of 20 K and a constant pressure of 1 atm. The atoms in the simulation cell are assigned with initial velocities at the desired temperature, and then equilibrated following the same two-step processes as described in the Equilibration section: the simulation cell is first equilibrated in the canonical ensemble and then in the isothermal–isobaric ensemble for 50 ps and 300 ps, respectively, with controlled temperature and pressure. The changes in lattice constants are calculated at each temperature.

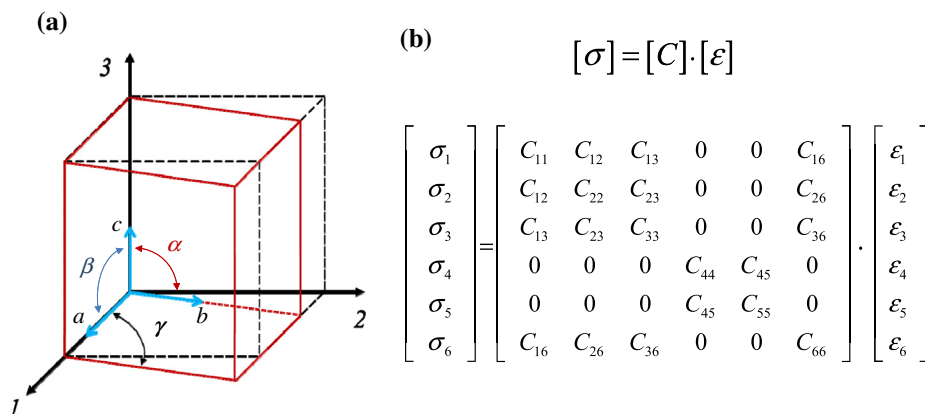
### 3. Results and discussion

#### 3.1. Lattice parameters

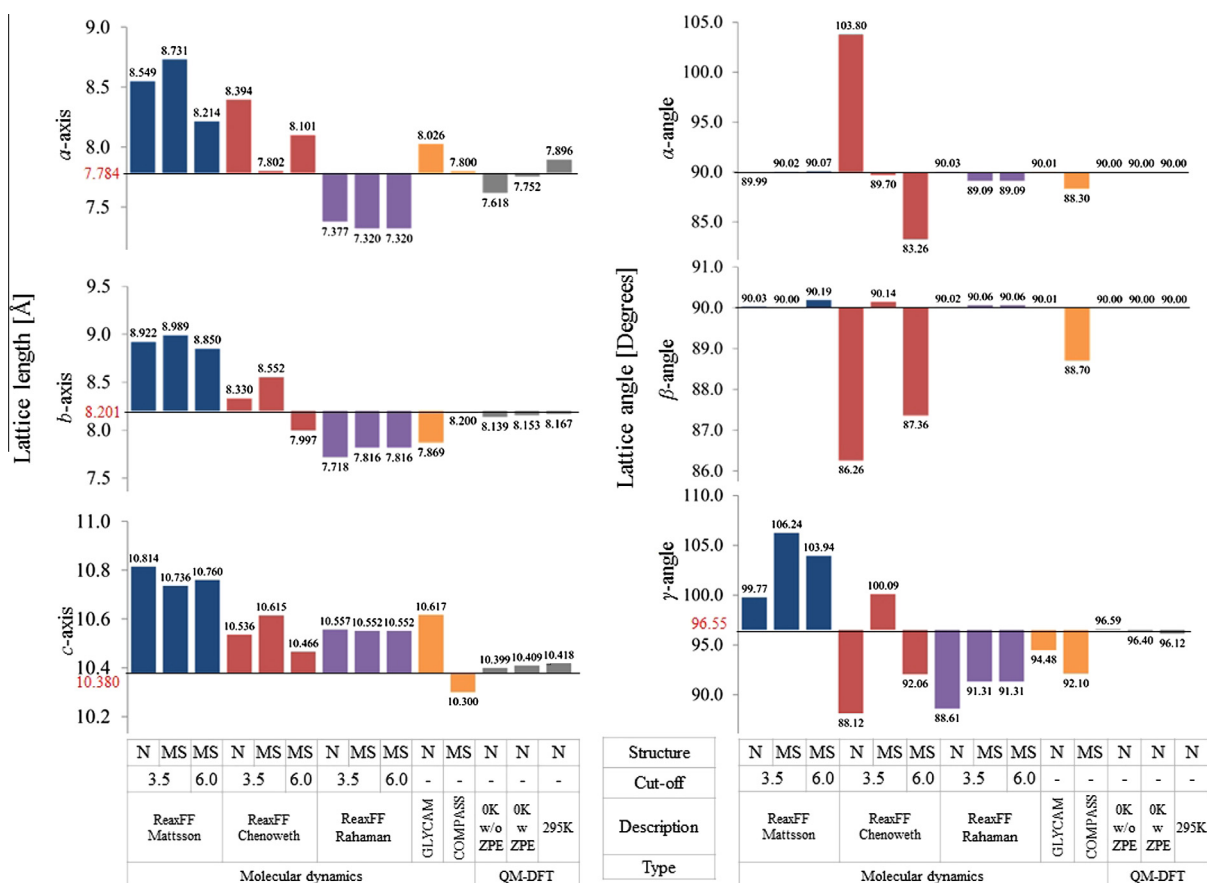
The lattice parameters for cellulose have been measured by several authors [26–31,42–45] using different experimental techniques and crystal sources. For cellulose  $I_\beta$  network A Nishiyama et al. [26] report:  $a = 7.784 \text{ \AA}$ ,  $b = 8.201 \text{ \AA}$ ,  $c = 10.380 \text{ \AA}$ ,  $\alpha = 90^\circ$ ,  $\beta = 90^\circ$ ,  $\gamma = 96.55^\circ$ , Volume =  $658.3 \text{ \AA}^3$  at 293 K. Most of the lattice parameters exhibit variations around 1% over a wide range of temperatures and crystalline sources, except for the lattice parameter  $a$ , which can change significantly with temperature; temperature effects are discussed in a later section.

Fig. 5 summarizes the comparison of simulation predictions with the experimental values reported by Nishiyama et al. [26] for cellulose  $I_\beta$  structure A (measured for crystalline cellulose from tunicates using X-ray and neutron fiber diffraction). Each bar represents the difference between the simulation result and the experimental values. Specifically, smaller bars represent better agreement. Results from previous QM-DFT calculations performed at 0 K (with and without zero-point vibrational energy (ZPE) correction) and 295 K [7,8] are also included as a reference. Of the two non-reactive FFs, COMPASS exhibits the best approximation for the lattice constants, with a total difference smaller than  $0.08 \text{ \AA}$  in each direction. This is less than 0.8% variation with respect to the experimental value. On the other hand, GLYCAM overestimates the  $a$  axis by  $0.42 \text{ \AA}$  (5.4%) and the  $c$  axis by  $0.237 \text{ \AA}$  (2.3%). However, it underestimates the  $b$  axis by  $0.32 \text{ \AA}$  (4.0%). These FFs exhibit the opposite trend for lattice angles: COMPASS underestimate the  $\alpha$  angle by  $1.7^\circ$  (1.9%), the  $\beta$  angle by  $1.3^\circ$  (1.4%) and the  $\gamma$  angle by  $4.45^\circ$  (4.6%), whereas GLYCAM has negligible deviation in the  $\alpha$  and  $\beta$  angle but underestimates the  $\gamma$  angle by  $2.07^\circ$  (2.1%).

Each ReaxFF parameterization exhibits unique behavior, which emphasizes the importance of this comparison. ReaxFF\_Mattsson gives the least accurate approximation for the lattice axis, with maximum deviation that exceeds  $0.947 \text{ \AA}$  (12.2%) in the  $a$ -axis,  $0.788 \text{ \AA}$  (9.6%) in the  $b$ -axis and  $0.38 \text{ \AA}$  (3.6%) in the  $c$ -axis. While the lattice angles  $\alpha$  and  $\beta$  exhibit almost no deviation from the Nishiyama structure, the  $\gamma$  angle is overestimated by  $9.69^\circ$  (10%). On the other hand, ReaxFF\_Chenoweth produces results with the highest angular deviation for  $\alpha$  and  $\beta$ . Moreover, this FF is very sensitive to the initial structure being used. The best approximation is achieved using a H-bond cutoff distance of  $3.5 \text{ \AA}$ . In this particular case,  $\alpha$  and  $\beta$  angles exhibit negligible deviations ( $<0.5\%$ ) whereas the  $\gamma$  angle is overestimated by  $3.54^\circ$  (3.6%). The same structure exhibits good agreement in the  $a$ -axis direction with values



**Fig. 4.** (a) Schematic representation of the cellulose  $I_\beta$  monoclinic unit cell aligned with the Cartesian coordinate system used in this work (red solid lines). A cubic cell (black dashed lines) is used to help visualize the orthogonality between axis  $a$ – $c$  and  $b$ – $c$ , highlighting the non-orthogonal relationship between  $a$  and  $b$ . (b) Stiffness matrix relating stresses and strains in the monoclinic unit cell [41]. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 5.** Predicted cellulose I<sub>β</sub> equilibrium lattice parameters from molecular dynamics (this work) and QM-DFT [7], for the different force-fields, H-bond cutoff distances (3.5 and 6.0 Å) and simulation parameters (vibrational energy and temperature). Reference lines are from the Nishishama et al. [26] network A at 293 K.

comparable to QM-DFT results [7]. The *b* and *c*-axis show results similar to the other parameterizations. Finally, ReaxFF\_Rahaman underestimates both *a* and *b*-axis by less than 0.464 Å (6%), but overestimates the *c*-axis by roughly 0.172 Å (1.7%). The  $\alpha$  and  $\beta$  angles exhibit almost no deviation from experimental values (less than 1°), whereas the  $\gamma$  angle is underestimated by 5.24° (5.4%). It is important to note that ReaxFF\_Rahaman shows no variation with respect to the cutoff distance due to the shape of the H-bond energy surface (see Fig. 3d).

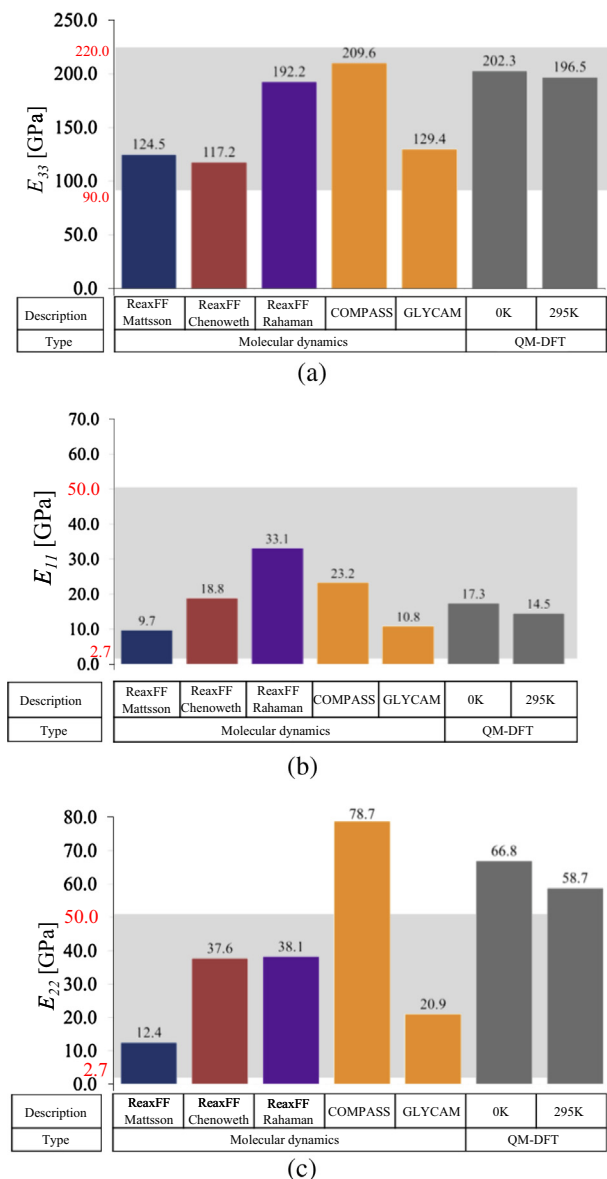
In summary, none of the FFs used in this study is capable of representing all the experimental lattice parameters accurately. Similar limitations were found in a previous analysis of three other force fields (CHARMM, GLYCAM and GROMOS) [10,46]. However, it is possible to achieve a good representation of lattice constants or angles (but not both) by choosing the appropriate combination of FF, parameter set and H-bond cutoff distance.

### 3.2. Young's modulus

For some applications, the lattice parameters and, hence the final shape of the crystalline structure, are not as important as the elastic behavior of the materials. Fig. 6 compares the Young's modulus along the three principal directions according to the Cartesian coordinate system defined in Fig. 1. For the 3-direction (coincident with the *c*-axis), Diddens et al. [47] reported values of  $E_{33} = 220 \pm 50$  GPa using Inelastic X-ray Scattering (IXR) of bleached flax fiber bundles aligned in a parallel fashion. Diddens and coworkers [47] claimed that IXR was not affected by the amorphous zones occurring in natural cellulose, and that the elastic behavior was mostly related to the highly crystalline regions.

Other previous studies used X-ray diffraction on ramie fibers, yielding values of  $E_{33}$  between 90 and 148 GPa [42–44,48,49]. Recently, Dri et al. [7,8] reported QM-DFT simulations for crystalline cellulose I<sub>β</sub> in the range of 200 GPa. As shown in Fig. 6a, all the FFs produce results within the relatively wide range of experimental values reported in [42–44,47–49]. The results for ReaxFF\_Mattsson and ReaxFF\_Chenoweth fall in the range  $117 < E_{33} < 125$  GPa. Fig. 6 only shows those results with a H-bond cutoff distance of 3.5 Å, since we found that the H-bond cutoff distance had a relatively small influence on these results. This might be explained by the weak force produced by H-bond interactions compared to the covalent bonds that govern the mechanical response in the *c*-axis direction. ReaxFF\_Rahaman and COMPASS are the only FFs that produce results on the order of  $E_{33} = 200$  GPa, which was predicted by QM-DFT.

The Young's moduli in the transverse directions, both 1 and 2, exhibit similar trends. Diddens and coworkers [47] reported a value of the Young modulus in the 1–2 plane (the specific direction was uncertain) of  $15 \pm 1$  GPa. Lahiji et al. [50] and Wagner et al. [6] performed atomic force microscopy nanoindentation on cellulose nanocrystals (CNCs), which are the pure crystalline particles extracted from tunicate fibrils through acid hydrolysis to dissolve the amorphous zones. While Lahiji et al. [50] reported values between 18 and 50 GPa, Wagner et al. [6] reported a mean value of 8.1 GPa and a 95% confidence interval of 2.7–20 GPa. As described in [50], these tests consisted of applying the load in the direction perpendicular to the surface of a CNC lying on a flat substrate. However, the specific crystallographic orientation with respect to the loading direction was unknown during the tests due to the variability in particle shape and lack of control of how



**Fig. 6.** Young's modulus computed in three principal directions from molecular dynamics (this work) and QM-DFT [7]. (a)  $E_{33}$ , (b)  $E_{11}$  and (c)  $E_{22}$ . The gray shaded region in the background represents the range of the experimental values reported in [6,42–44,47–50]. Maximum and minimum values are labeled in red. Cellulose  $I_\beta$  network A is used as the initial structure with a H-bond cutoff distance of 3.5 Å and simulation parameters (force field in MD and temperature in QM-DFT) given in the figure. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

they lie on the substrate, leading to large variability in the Young's modulus data. For a more comprehensive discussion see Dri et al. [8] and Wu et al. [35]. As seen in Fig. 6b and c, ReaxFF\_Mattsson produces the smallest estimates of the transverse Young's modulus, barely exceeding the lower limit reported by Wagner et al. [6]. ReaxFF\_Chenoweth produces values of  $E_{11}$  in good agreement with QM-DFT simulations but yields smaller values of  $E_{22}$ . On the other hand, ReaxFF\_Rahaman tend to produce higher values of Young's moduli. While  $E_{11}$  is almost twice the value obtained with QM-DFT, exhibiting almost no effect of the H-bond cutoff distance, it also produces the largest reported value of  $E_{22} = 38.1$  GPa among the reactive FFs. While both non-reactive FFs produce values of  $E_{11}$  between the experimental data, COMPASS is the only one that produces values of  $E_{22}$  that are closer to those obtained with QM-DFT. The discrepancy in the Young's modulus values between model

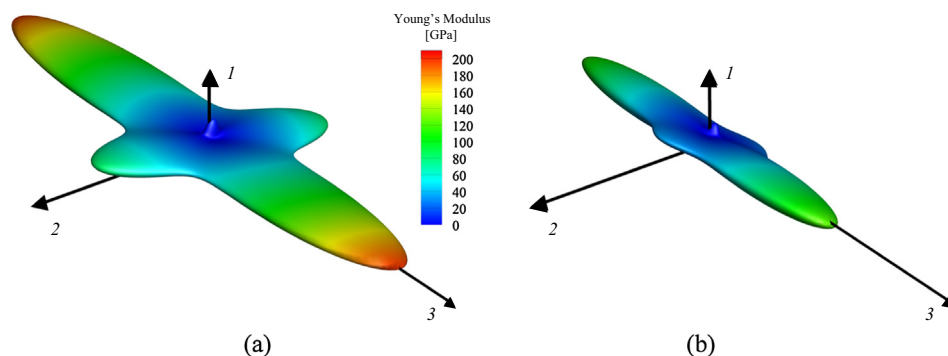
predictions and experimental measurements is mainly due to uncertainties in the shape of the CNCs and in identifying the specific direction of the measurement with respect to the cellulose crystal structure (i.e. along direction 1, 2 or any direction in-between).

### 3.3. Anisotropic elasticity

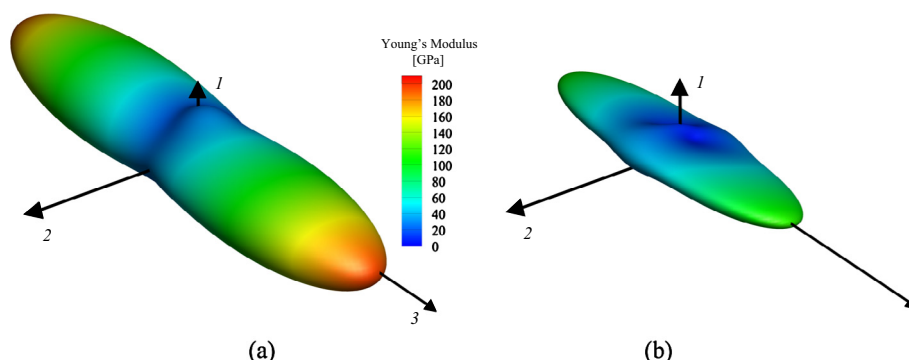
Additional mechanical information can be extracted from the computed compliance matrix by generating a surface contour plot of the variation in Young's modulus with crystallographic direction. A post-processing software, *Anisotropy Calculator – 3D Visualization Toolkit* [51], was used for this purpose. Each point on the surface represents the magnitude of the Young's modulus in the direction of a vector from the origin (i.e. at the intersection of the 1, 2, and 3 axes in the interior of the surface) to a given point on the surface. The color also represents the magnitude of the Young's modulus. These values are obtained based on the calculated stiffness matrix in the 123 system of coordinates (Fig. 4b) and the proper rotation. The shape of this surface is indicative of the elastic anisotropy of cellulose  $I_\beta$ . For instance, the computed Young's modulus surface for a linearly elastic isotropic material is a perfect sphere with a radius equal to the Young's modulus. However, the cellulose  $I_\beta$  surfaces in Figs. 7–9 exhibit extreme variations in the Young's modulus, as denoted by the accentuated contour lobe along the 3-axis (i.e. along the cellulose chains) relative to the smaller lobes along the 1 and 2 directions. In general, these plots bring an intuitive and rich perspective of the anisotropic elastic behavior of cellulose  $I_\beta$  by providing property information well beyond that given only along the 1, 2, and 3-directions. More information about the interpretation of these 3D plots can be found in [7,8]. The 3-D representation of the data provides additional fidelity when comparing the different simulation approaches, and parameters. We expect this information to also highlight the main differences between the FFs and QM-DFT. While the values of  $E_{11}$ ,  $E_{22}$  and  $E_{33}$  obtained with some FFs seem to closely follow those obtained with QM-DFT, it may still be possible for FFs to considerably differ from the QM-DFT values in other directions. Moreover, since the calculation of Young's modulus for all the possible directions involve the entire stiffness matrix, discrepancies in the Young's moduli between FFs may also imply differences in the shear behavior.

Fig. 7a reports the variation of the Young's modulus with respect to the crystallographic direction computed based on QM-DFT results at 300 K [8]. The largest values (red contours) are along the 3-axis, with the smallest values along the 1-axis. For instance, the maximum Young's modulus is 206 GPa, which is comparable to that of steel. Using Fig. 7a as a reference, all the Young's modulus surfaces obtained with the other FFs (7b through 9) were plotted maintaining the same view angle, scale and color contour levels to facilitate comparisons between results. The largest values (red contours) are along the 3-axis, with the smallest values in the 1–2 plane. It is important to remark that the deformation along the 3-direction is governed by covalent bonds that form the cellulose chains, whereas the mechanical behavior in the 1–2 plane is governed by non-bonded interactions. First we will analyze the shape of the surfaces predicted by the various FFs. Then the role of the non-bonded interactions will be examined by focusing on the mechanical response in the 1–2 plane.

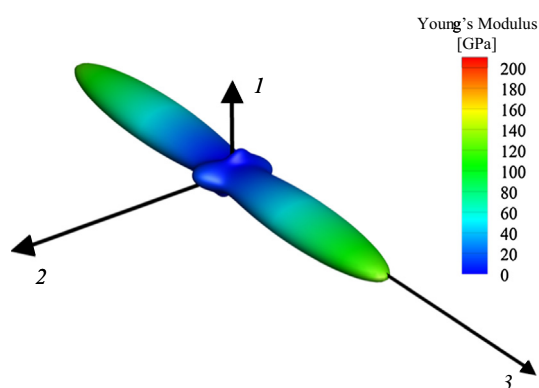
Fig. 7b shows the Young's modulus variation with crystallographic direction based on MD results with ReaxFF\_Mattsson parameterization (with a 3.5 Å cutoff distance for H-bonds interactions). The ReaxFF\_Rahaman parameterization is shown in Fig. 8a and the ReaxFF\_Chenoweth parameterization in Fig. 8b (both with a 3.5 Å cutoff distance). We note that the ReaxFF\_Rahaman parameterization is the only one that produces results close to those obtained from QM-DFT [7,8] in the 3-direction. The results in



**Fig. 7.** (a) Surfaces showing contours of Young's modulus for cellulose  $I_\beta$  from Ref. [8] (computed with QM-DFT at 300 K). Each point on the surface represents the magnitude of Young's modulus in the direction of a vector from the origin to that point. The color contours help to identify the Young modulus variation of cellulose  $I_\beta$  and emphasize its extreme anisotropy. The axes 1, 2 and 3 indicate the primary directions with respect to the crystallographic directions (e.g., Fig. 1). (b) Surfaces showing contours of Young's modulus for cellulose  $I_\beta$  computed using ReaxFF\_Mattsson parameterization (with 3.5 Å cutoff distance for H-bonds). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** Surfaces showing contours of Young's modulus for cellulose  $I_\beta$  computed using (a) ReaxFF\_Rahaman parameterization and (b) ReaxFF\_Chenoweth parameterization (both with 3.5 Å cutoff distance for H-bonds).



**Fig. 9.** Surfaces showing contours of Young's modulus for cellulose  $I_\beta$  predicted with GLYCAM.

Fig. 9 were computed with the non-reactive FF, GLYCAM. The general shape of the Young's modulus surface resembles the one presented in Fig. 7a, but exhibits softer transitions between directions.

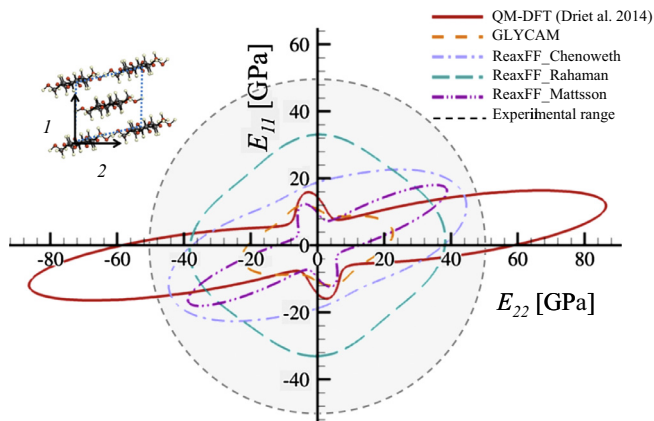
Variations of the *transverse* Young's modulus within a given crystallographic direction in the 1–2 plane predicted by these FFs are shown in Fig. 10. Only non-bonded interaction are present in the 1–2 plane (see Fig. 1a). This makes it an ideal for analyzing how vdW, Coulomb and H-bonds interactions affect the mechanical behavior representation for each FF. Experimental [6] as well as QM-DFT [8] results are superimposed on both figures for reference. In this case, the non-bonded interaction cutoff distance was set to

10 Å for the GLYCAM force field and the H-bond cutoff distance was defined as 3.5 Å for each of the three ReaxFF parameterizations. The four MD simulations produced values within the limits of experimental characterizations [6]. Most of the curves present an oblong shape with smooth variations with orientation. ReaxFF\_Mattsson is the only parameterization that produces a curve resembling QM-DFT results [8] but with a different size (smaller) and orientation.

These large discrepancies between FFs in Young's modulus in the 1–2 plane are an indication of how differences in non-bonded interaction (specifically, the H-bonds) directly affect the sliding between adjacent planes, and thus the shear behavior. While, in most cases, the values of  $E_{11}$ ,  $E_{22}$  and  $E_{33}$  are of primary interest, the differences in shear behavior may significantly affect the prediction of torsional and bending stiffness of cellulose crystals and fibrils.

### 3.4. Thermal expansion

The ability of FFs to predict lattice parameter variations with temperature is critical to predicting the thermal expansion coefficients (TECs),  $\xi_i$ . Fig. 11 shows predicted lattice parameters,  $a$ ,  $b$ ,  $c$  and angle  $\gamma$ , of the cellulose  $I_\beta$  network A as a function of the temperature for the FFs studied in this work. QM-DFT [7] and experimental values [52–55] are also included for comparison. The TEC is calculated as the slope of a linear regression line fit to the lattice parameter vs. temperature data using least squares. As apparent from Fig. 11, QM-DFT predictions are similar in magnitude to and exhibit trends consistent with experimental results. In contrast, all FFs yield highly variable results, with lattice parameter

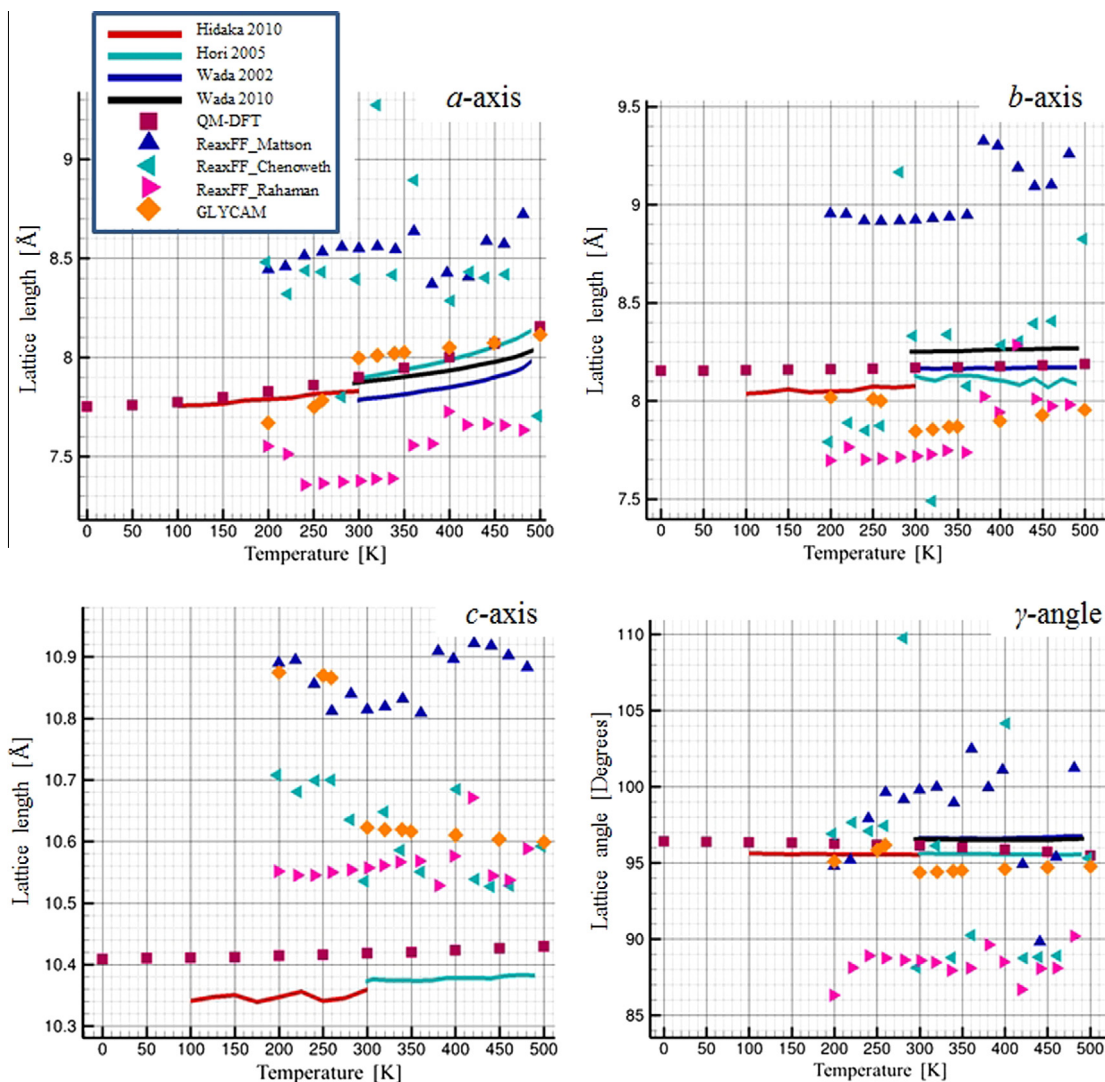


**Fig. 10.** Variation of the transverse Young's modulus (1–2 plane) for different FF parameterizations. QM-DFT results at 300 K [8] and experimental results (shaded in grey with a dashed grey line defining the perimeter) [6] added for reference. The inset in the upper left corner shows the orientation between the original input structure and the Cartesian system of coordinates. The final structure (after analysis) may not be aligned as shown in the inset figure.

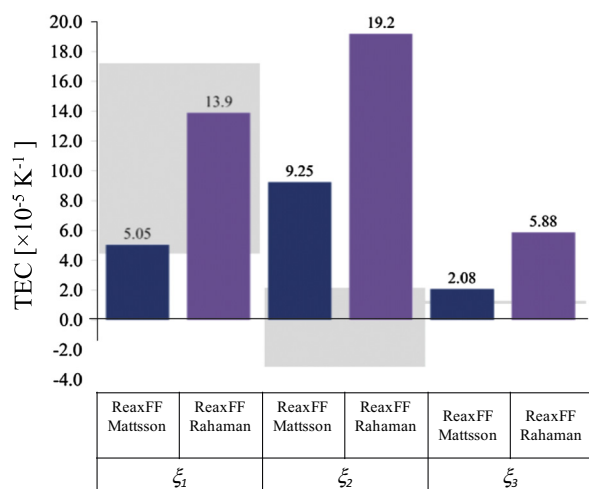
magnitudes quite different from those in experiment and exhibiting non-linear increases with temperature. The non-linearity is particularly noticeable at higher temperatures, which is reasonable since most FFs are fitted at or near ambient temperature and so may be less reliable far from the conditions under which they were fit. This suggests that any estimation of TEC using a FF must be considered cautiously, and certainly TEC cannot be calculated using the entire temperature range.

Therefore, none of the FFs appear to be able to capture the TEC over the entire temperature range in the various directions accurately. However, by using a smaller temperature interval to calculate TEC, it may be possible provide a limited indication of predictive capabilities for some of the FFs. Specifically, for ReaxFF\_Mattsson and ReaxFF\_Rahaman, the data between 250 and 350 K is reasonably linearly. Note that, even within this reduced range, GLYCAM and ReaxFF\_Chenoweth results are too scattered to enable reasonable linear fitting. We therefore limit our analysis to TECs calculated from linear fitting to data between 250 and 350 K for ReaxFF\_Mattsson and ReaxFF\_Rahaman only.

A summary of the TEC results is depicted in Fig. 12. For the  $\alpha$ -axis, ReaxFF\_Mattsson predicted a value of  $\xi_1 = 5.05 \times 10^{-5} \text{ K}^{-1}$  whereas ReaxFF\_Rahaman yields a higher value,  $\xi_1 = 13.9 \times 10^{-5} \text{ K}^{-1}$ , showing reasonable agreement with experimental



**Fig. 11.** Predicted lattice parameters  $a$ ,  $b$ ,  $c$ , angle  $\gamma$  of cellulose  $I_\beta$  network A compared with QM-DFT [7] and experimental data measured by Hidaka et al. using wood cellulose [52], by Hori and Wada using wood cellulose [53], by Wada using tunicate (halocynthia) [54], and by Wada et al. using green algae [55].



**Fig. 12.** Computed values of TEC vs. experimental results (in grey). The experimental values are taken from [52–55].

results of cellulose  $I_\beta$  containing biomass. Hikada et al. [52] and Hori and Wada [53] reported values of  $\xi_1 = 5.2 \times 10^{-5} \text{ K}^{-1}$  and  $13.6 \times 10^{-5} \text{ K}^{-1}$ , respectively, for wood cellulose between room temperature and 500 K. Wada et al. [54] reported values between  $4 \times 10^{-5} \text{ K}^{-1}$  and  $17 \times 10^{-5} \text{ K}^{-1}$  for tunicate cellulose in a range between room temperature and 473 K. Wada et al. [55] reported a value of  $9.8 \times 10^{-5} \text{ K}^{-1}$  using green algae. For the  $b$ -axis, the computed values are  $\xi_2 = 9.25 \times 10^{-5} \text{ K}^{-1}$  for ReaxFF\_Mattsson and  $\xi_2 = 19.2 \times 10^{-5} \text{ K}^{-1}$  for ReaxFF\_Rahaman. For this particular axis, the computed TEC values are an order of magnitude larger than the experimental values. For instance a value of  $\xi_2 = 2.1 \times 10^{-5} \text{ K}^{-1}$  was reported in [52],  $3 \times 10^{-5} \text{ K}^{-1}$  was reported in [53],  $0.5 \times 10^{-5} \text{ K}^{-1}$  was reported in [54], and  $1.2 \times 10^{-5} \text{ K}^{-1}$  was reported in [55]. In the  $c$ -axis direction, a TEC value of  $\xi_3 = 2.08 \times 10^{-5} \text{ K}^{-1}$  is computed using ReaxFF\_Mattsson, whereas ReaxFF\_Rahaman yields a higher  $5.88 \times 10^{-5} \text{ K}^{-1}$ . Hori and Wada [53] reported a value of  $0.6 \times 10^{-5} \text{ K}^{-1}$  while Hikada et al. [52] reported  $0.4 \times 10^{-5} \text{ K}^{-1}$ . The experimental values reported in [52–55] have been included in Fig. 12. It should be noted that the relatively large range of experimental values is mainly due to the variation of TEC with respect to temperature. For instance, Wada et al. [55] reported that  $\xi_1$  increases from  $4.3 \times 10^{-5} \text{ K}^{-1}$  at room temperature to  $17 \times 10^{-5} \text{ K}^{-1}$  at 473 K. However, none of the FFs appear to be able to capture the TEC in the various directions accurately, which makes any quantitative comparison between FFs difficult.

### 3.5. Summary of findings

Taken together, the detailed results shown in the previous sections indicate that (a) QM-DFT is better able to predict thermo-mechanical properties of cellulose  $I_\beta$  than any of the FFs studied here, and (b) while none of the FFs studied here accurately predicted all properties, some of them can be relied upon for accurate prediction of a select properties.

The primary difference between FFs and QM-DFT is that QM-DFT explicitly accounts for the electron exchange and correlation that govern the bonding interaction between atoms, and hence, the cellulose thermo-mechanical behavior. As has been reported in previous work [7,8] the use of QM-DFT with a semi-empirical correction for van der Waals interactions has been shown to yield the best agreement with experimental data in terms of lattice parameters and thermo-mechanical properties. In fact, these models have been tested systematically on different

systems including molecular crystals, crystals and isolated molecules [9]. On the other hand, FFs employ semi-empirical potential which, not surprisingly, cannot properly predict all these properties. However, QM-DFT is not always suitable for a given modeling study because of its computational time constraints, leading to simulation domains that may be orders of magnitude smaller than those typically handled by FFs. This is especially critical for simulations that involve entire cellulose crystals and their surfaces, or slower thermal processes such as thermal conduction. For such cases, it is desirable to use an empirical model, chosen based on the goal of the simulation [8]. Additionally, standard QM-DFT cannot properly describe weak interactions between molecules unless the proper correction is applied [9]. Discussion of the use of QM-DFT for obtaining the thermo-mechanical behavior of cellulose can be found in [7,8].

For comparing FFs, the analyses reported here suggest several specific guidelines. First we discuss prediction of lattice constants, which can be analyzed quantitatively since specific values have been measured. Most of the FFs analyzed yield reasonable predictions, with the average error across all six constants ( $a, b, c, \alpha, \beta, \gamma$ ) being less than 6% for all FFs. In terms of average error, the best predictions are obtained using ReaxFF\_Chenoweth with a 3.5 Å cutoff (1.8%), GLYCAM (1.9%) or COMPASS (1.5%). So, if a reactive potential is needed, the ReaxFF\_Chenoweth potential is the best choice; if not, COMPASS gives the best predictions.

Analysis of the ability of the FFs to predict elastic constants is less direct since there is a wide range of experimentally measured values. All FFs predict  $E_{11}$  and  $E_{33}$  within the experimentally-measured range, and all of them, except COMPASS, predict  $E_{22}$  within the experimentally-measured range. Therefore, we evaluate the FF predictions by comparison to the data from QM-DFT at 295 K. This analysis reveals that, if only the chain direction modulus ( $E_{33}$ ) is of interest, then the best FFs are COMPASS (6.7% error) and ReaxFF\_Rahaman with a cutoff of 3.5 Å (2.2% error). However, these FFs fail to predict the Young's moduli in the transverse direction with any reasonable accuracy. In fact, none of the FFs predict the Young's moduli in the transverse direction with accuracy better than 30%. The best prediction of transverse moduli comes from the ReaxFF\_Chenoweth FF with either H-bond cutoff (~32% error). More dramatic differences between FF were revealed by studying the surfaces of Young's modulus for all directions. While values of  $E_{11}$ ,  $E_{22}$  and  $E_{33}$  may be in the range of experimental values, the Young's modulus for directions other than 1, 2 and 3 show variations that imply significant discrepancies in the way adjacent H-bond planes slide relative to one another due to the way the non-bonded interactions are being described. This observation has important implications for the ability of the FFs to predict bending and twisting of cellulose crystals and fibrils.

Lastly, none of the FFs is able to predict thermal expansion with any reasonable accuracy, and only ReaxFF\_Mattsson and ReaxFF\_Rahaman predict temperature-dependent lattice constants that can be linearly fit to yield CTE values. This is consistent with previous observations that different FFs have a strong influence on the way bonding interaction is described, while QM-DFT can accurately account for electron exchange and correlation. This also suggests that existing empirical models should not be used to predict thermal properties of crystalline cellulose.

## 4. Conclusion

The reactive force field ReaxFF (with three different parameter sets) was tested and compared with two commonly used non-reactive FFs (COMPASS and GLYCAM) to evaluate how accurately they can predict the structure, thermo-mechanical behavior, and property anisotropy of crystalline cellulose  $I_\beta$ . We found that

none of the tested force fields yield results in perfect agreement with experimental data and QM-DFT results for all predicted properties. However, depending on what needs to be studied, a given property may be predicted accurately if an appropriate FF is chosen. In addition, simulations can be used to understand general trends and, depending on the FF, isolate specific effects, such as the role of H-bonds if a reactive FF is used. This work provides guidelines to select a FF or, in the case of reactive FFs, a parameter set, based on the focus of their study. Most importantly, this work highlights the limitations of common force fields used for modeling crystalline cellulose, encouraging future research to parameterize a FF specifically for this material.

## Acknowledgements

The authors are grateful to financial support by the Forest Products Laboratory under USDA grant: 07-CR-1111120-093, the Air Force Office of Scientific Research grant FA9550-11-1-0162 and National Science Foundation through Grant No. CMMI-1131596.

## References

- [1] D. Klemm, B. Heublein, H.P. Fink, A. Bohn, *Angew. Chem. Int. Ed.* 44 (22) (2005) 3358–3393.
- [2] Y. Habibi, L.A. Lucia, O.J. Rojas, *Chem. Rev.* 110 (6) (2010) 3479–3500.
- [3] R.J. Moon, A. Martini, J. Nairn, J. Simonsen, J. Youngblood, *Chem. Soc. Rev.* 40 (7) (2011) 3941–3994.
- [4] T. Bucko, D. Tunega, J.G. Ángyán, J. Hafner, *J. Phys. Chem. A* 115 (35) (2011) 10097–10105.
- [5] R. Parthasarathi, G. Bellesia, S.P.S. Chundawat, B.E. Dale, P. Langan, S. Gnanakaran, *J. Phys. Chem. A* 115 (49) (2011) 14191–14202.
- [6] R. Wagner, R. Moon, J. Pratt, G. Shaw, A. Raman, *Nanotechnology* 22 (45) (2011) 455703.
- [7] F. Dri, L.H. Hector Jr., R.J. Moon, P.D. Zavattieri, *Cellulose* 20 (6) (2013) 2703–2718.
- [8] F. Dri, S. Shang, L.G. Hector Jr., P. Saxe, Z.-K. Liu, R. Moon, P.D. Zavattieri, *Mater. Sci. Eng.* 22 (2014) 085012.
- [9] T. Bucko, J. Hafner, S. Lebegue, J.G. Angyan, *J. Phys. Chem. A* 114 (2010) 11814–11824.
- [10] J.F. Matthews, G.T. Beckham, M. Bergensträhle-Wohlert, J.W. Brady, M.E. Himmel, M.F. Crowley, *J. Chem. Theory Comput.* 8 (2) (2012) 735–748.
- [11] A.R. Leach, *Molecular Modelling: Principles and Applications*, Pearson Education, 2001.
- [12] W.J. Hehre, *A guide to molecular mechanics and quantum chemical calculations: Wavefunction* Irvine, CA, 2003.
- [13] A.C.T. van Duin, S. Dasgupta, F. Lorant, W.A. Goddard III, *J. Phys. Chem. A* 105 (41) (2001) 9396–9409.
- [14] T.R. Mattsson, J.M.D. Lane, K.R. Cochrane, M.P. Desjarlais, A.P. Thompson, F. Pierce, G.S. Grest, *Phys. Rev. B* 81 (5) (2010) 054103.
- [15] K. Chenoweth, A.C.T. van Duin, W.A. Goddard, *J. Phys. Chem. A* 112 (2008) 1040–1053.
- [16] O. Rahaman, A.C.T. van Duin, W.A. Goddard, D.J. Doren, *J. Phys. Chem. B* 115 (2) (2010) 249–261.
- [17] H. Sun, *J. Phys. Chem. B* 102 (38) (1998) 7338–7364.
- [18] S.J. Eichhorn, G.R. Davies, *Cellulose* 13 (3) (2006) 291–307.
- [19] F. Tanaka, T. Iwata, *Cellulose* 13 (5) (2006) 509–517.
- [20] K.N. Kirschner, A.B. Yongye, S.M. Tschampel, J. González-Outeiriño, C.R. Daniels, B.L. Foley, R.J. Woods, *J. Comput. Chem.* 29 (4) (2008) 622–655.
- [21] Z. Qiong, V. Bulone, H. Ågren, Y. Tu, *Cellulose* 18 (2) (2011) 207–221.
- [22] R.J. Maurer, A.F. Sax, *Procedia Comput. Sci.* 1 (1) (2010) 1149–1154.
- [23] T. Yui, S. Nishimura, S. Akiba, S. Hayashi, *Carbohydr. Res.* 341 (15) (2006) 2521–2530.
- [24] K. Kulasinski, S. Ketten, S.V. Churakov, D. Derome, J. Carmeliet, *Cellulose* (2014), <http://dx.doi.org/10.1007/s10570-014-0213-7>.
- [25] R. Sinko, S. Mishra, L. Ruiz, N. Brandis, S. Ketten, *ACS Macro Lett.* 3 (2013) 64–69.
- [26] Y. Nishiyama, P. Langan, H. Chanzy, *J. Am. Chem. Soc.* 124 (31) (2002) 9074–9082.
- [27] Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, *J. Am. Chem. Soc.* 125 (47) (2003) 14300–14306.
- [28] P. Langan, N. Sukumar, Y. Nishiyama, H. Chanzy, *Cellulose* 12 (6) (2005) 551–562.
- [29] Y. Nishiyama, G.P. Johnson, A.D. French, V.T. Forsyth, P. Langan, *Biomacromolecules* 9 (11) (2008) 3133–3140.
- [30] Y. Nishiyama, P. Langan, H. Chanzy, *Acta Crystallogr. Sec. D* 66 (11) (2010) 1172–1177.
- [31] J. Sugiyama, R. Vuong, H. Chanzy, *Macromolecules* 24 (14) (1991) 4168–4175.
- [32] M.G. Zuluaga, R. Moon, F. Dri, P. Zavattieri, *Crystalline Cellulose – Atomistic Toolkit*, 2013, Available from: <<https://nanohub.org/resources/matrix2surface>>.
- [33] A. Šturcová, I. His, D.C. Apperley, J. Sugiyama, M.C. Jarvis, *Biomacromolecules* 5 (4) (2004) 1333–1339.
- [34] I. Accelrys Software, *Material Studio Modeling Environment*, Accelrys Software Inc., San Diego, CA, USA, 2011.
- [35] X. Wu, R.J. Moon, A. Martini, *Cellulose* 21 (2014) 2233, <http://dx.doi.org/10.1007/s10570-014-0325-0>.
- [36] S. Plimpton, A. Thompson, P. Crozier, Available from: <<http://lammps.sandia.gov/>>.
- [37] M. Bergensträhle, L.A. Berglund, K. Mazeau, *J. Phys. Chem. B* 111 (30) (2007) 9138–9145.
- [38] Y. Li, M. Lin, J.W. Davenport, *J. Phys. Chem. C* 115 (23) (2011) 11533–11539.
- [39] J. Matthews, M. Himmel, J. Brady, *Simulations of the Structure of Cellulose*, National Renewable Energy Laboratory (NREL), Golden CO, 2010.
- [40] L. Alenka, D. Chandler, *Nature* 379 (6560) (1996) 55–57, <http://dx.doi.org/10.1038/379055a0>.
- [41] R.M. Jones, *Mechanics of composite materials*, vol. 2. Taylor & Francis London, 1975.
- [42] I. Sakurada, T. Ito, K. Nakamae, *Die Makromolekulare Chemie* 75 (1) (1964) 1–10.
- [43] I. Sakurada, Y. Nukushina, T. Ito, *J. Polym. Sci.* 57 (165) (1962) 651–660.
- [44] M. Matsuo, C. Sawatari, Y. Iwai, F. Ozaki, *Macromolecules* 23 (13) (1990) 3266–3275.
- [45] V.L. Finkenstadt, R.P. Millane, *Macromolecules* 31 (22) (1998) 7776–7783.
- [46] P. Chen, Y. Nishiyama, K. Mazeau, *Cellulose* 21 (2014) 2207, <http://dx.doi.org/10.1007/s10570-014-0279-2>.
- [47] I. Diddens, B. Murphy, M. Krisch, M. Müller, *Macromolecules* 41 (24) (2008) 9755–9759.
- [48] T. Nishino, K. Takano, K. Nakamae, *J. Polym. Sci., Part B: Polym. Phys.* 33 (11) (1995) 1647–1651.
- [49] A. Ishikawa, T. Okano, J. Sugiyama, *Polymer* 38 (2) (1997) 463–468.
- [50] R.R. Lahiji, X. Xu, R. Reifengerger, A. Raman, A. Rudie, R.J. Moon, *Langmuir* 26 (6) (2010) 4480–4488.
- [51] M.G. Zuluaga, F. Dri, P. Zavattieri, R. Moon, *Anisotropy Calculator – 3D Visualization Toolkit*, 2013, Available from: <<https://nanohub.org/resources/matrix2surface>>.
- [52] H. Hidaka, U.-J. Kim, M. Wada, *Holzforchung* 64 (2) (2010) 167–171.
- [53] R. Hori, M. Wada, *Cellulose* 12 (5) (2005) 479–484.
- [54] M. Wada, *J. Polym. Sci. Part B-Polym. Phys.* 40 (11) (2002) 1095–1102.
- [55] M. Wada, H. Ritsuko, K. Ung-Jin, S. Sono, *Polym. Degrad. Stab.* 95 (8) (2010) 1330–1334.