# Calculation of single chain cellulose elasticity using fully atomistic modeling

XIAWA WU, ROBERT J. MOON, AND ASHLIE MARTINI

**ABSTRACT:** Cellulose nanocrystals, a potential base material for green nanocomposites, are ordered bundles of cellulose chains. The properties of these chains have been studied for many years using atomic-scale modeling. However, model predictions are difficult to interpret because of the significant dependence of predicted properties on model details. The goal of this study is to begin to understand these dependencies. We focus on the investigation on model cellulose chains with different lengths and having both periodic and nonperiodic boundary conditions, and predict elasticity in the axial (chain) direction with three commonly used calculation methods. We find that chain length, boundary conditions, and calculation method affect the magnitude of the predicted axial modulus and the uncertainty associated with that value. Further, the axial modulus is affected by the degree to which the molecule is strained. This result is interpreted in terms of the bonded and nonbonded contributions to potential energy, with a focus on the breaking of hydrogen bonds during deformation.

***Application:*** This study lays the groundwork for understanding the predictions of atomistic models and to help transition their use from an investigative method to a predictive tool useful for cellulose-based application design.

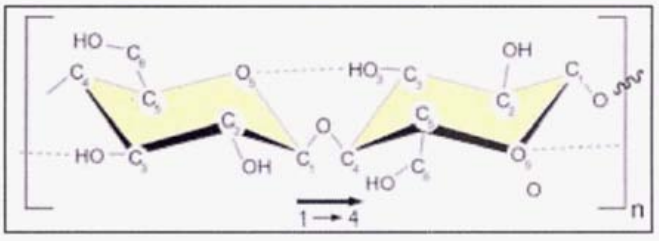
Atomistic modeling of cellulose has been used to complement experimental measurements of cellulose nanocrystals by providing the atomic level detail necessary to predict structural, energetic, and mechanical characteristics, and to gain a fundamental understanding of the atomic-scale origins of these characteristics. Such models frequently are used to predict elastic properties because the simulation methods are relatively simple and results can be compared to experimental data. Predictions of elasticity in the axial (chain) direction for the la or Iß polymorphs of crystalline cellulose have been reported from atomistic model-based studies for more than 20 years [l]. Unfortunately, quantitative comparison of elasticity results from model-to-model or to experimental measurements has been difficult because of the significant effect of variations in the simulation methods, atomic interaction models, and configuration of the modeled structures. For example, the axial modulus predicted using a model of 1x1x10 Iß unit cells was ~11% smaller than that predicted using a model of 4x4x10 Iß unit cells [2]. In another study, a difference in the initial equilibrium unit cell length of 0.1% caused a 22% variation in the axial modulus prediction [3].

The goal of this paper is to clearly illustrate some of these effects such that the meaning of predictions made using atomistic models can be understood more fully. We report results of a molecular dynamics- and molecular mechanics-based study focusing on the most basic structure, a single cellulose molecule, to understand the effects of some variables on model predictions.

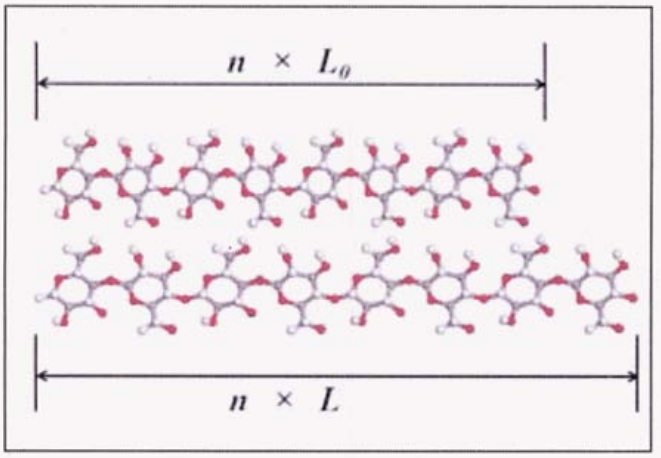
Many chains of different lengths coexist in a bulk polymer, and so the variation of structure, energetic, and elastic properties with chain length is insignificant. However, at the nanoscale, properties are likely to be influenced by so-called end effects, where the molecule is short enough that the contribution the ends of the chains make is comparable to that of the main body of the chain. To characterize this effect, we focused this study on model cellulose chains with different lengths and having both periodic and nonperiodic boundary conditions. These models are used to predict elasticity in the axial direction with three commonty used calculation methods. Results are discussed in terms of the effects of chain length, boundary conditions, calculation method, and strain on the model-predicted axial modulus, and the uncertainly associated with those predictions.

## METHODS

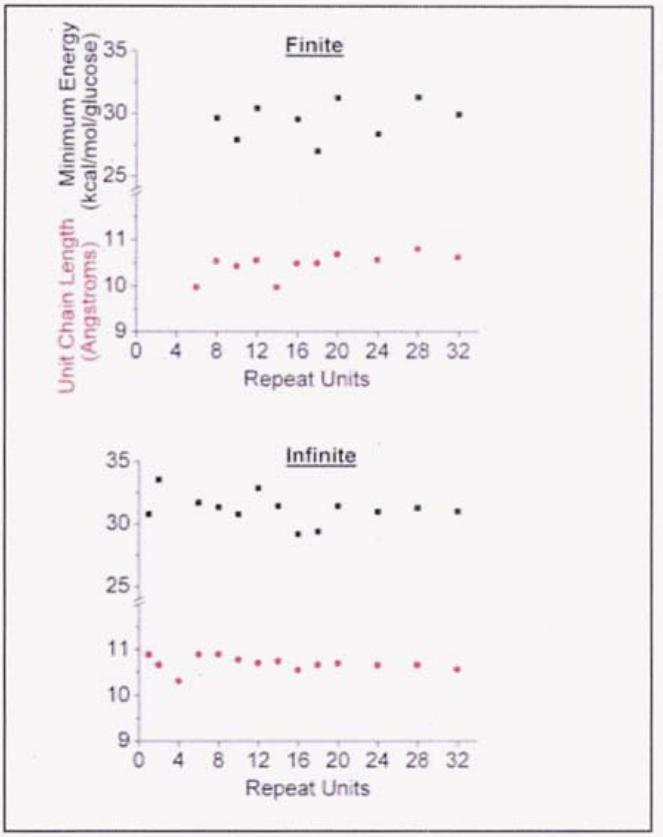
We created single chain cellulose molecules with lengths from one to 32 repeat units, where one unit contains two glucose rings and has a length of $L_0$ **(Fig. 1). Figure 2** shows an example consisting of four repeat units. To create chains of finite length, a single molecule was placed in the middle of a simulation box with dimensions 50x50x ( $n$ x $L_0$ + 50) Å$^3$. infinite




*1. Schematic of single cellulose chain repeat unit, showing the directionally of the 1→4 linkage and intramolecule hydrogen bonding (dotted lines).*




*2. Snapshots of a model cellulose chain at its equilibrium and strained lengths where n is the number of repeat units. Atoms represented as grey (carbon), red (oxygen), and white (hydrogen)spheres.*



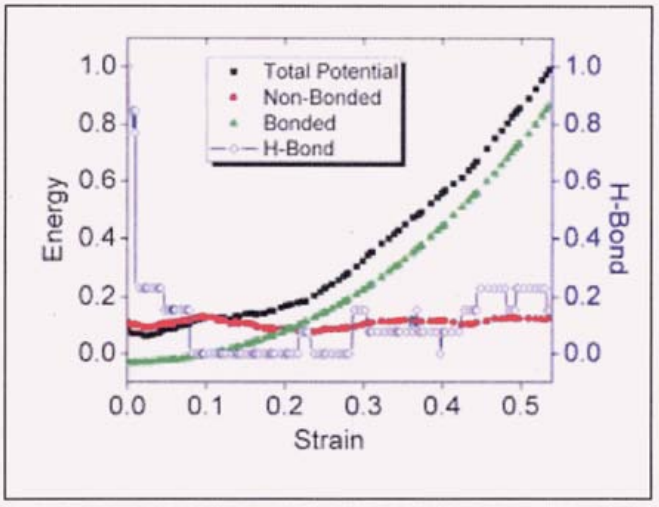

*3. Equilibrium unit cell length (red circles) and potential energy (black squares) for finite and infinite chains as functions of chain length. Each data point represents the average value of 10 repeated simulations.*

length molecules were obtained by adding a covalent bond across the periodic boundary linking the head carbon and tail oxygen atoms. In this case, the molecule is infinitely long in its axial direction and isolated In the other two directions, such that the simulation box is 50x50x ( $n$ x $L_0$) Å$^3$. The lateral dimensions of the simulation box were chosen to be 50 Å to ensure that the molecule did not interact with periodic images of itself during the simulation, taking into consideration all possible bending, rotating, and other movement of the molecule. Atomic interactions were described by the COMPASS force field [4] with all interaction cutoff distance of 10 Å. This force field was chosen because it has been applied succes fully to model cellulosic materials previously [2,3,5].

Initial positions and charges for each atom were computed, and then converted to a LAMMPS (large-scale atomic/molecular massively parallel simulator) input file. All subsequent simulations were performed using LAMMPS [6]. Each initial configuration was equilibrated using molecular dynamics in the NVT ensemble (constant number of atoms, volume, and temperature) at a temperature of 300 K for 10 ps before production runs using molecular mechanics. To evaluate the statistical significance of calculated results, each simulation case (chain length and boundary condition) was run 10 times with same initial atomic structure but different initial velocities (randomized momenta).

To enable calculation of the axial modulus, each model was repeatedly deformed in small increments in the chain direction, as illustrated in Fig. 2. The equilibrium length $L$ of a repeat unit and the corresponding minimum potential energy $U$ were calculated using molecular mechanics with the conjugate gradient energy minimization algorithm. Parameters that quantify clastic properties, with and without the assumption of constant area, have been differentiated as "chain stiffness" and "chain modulus," respectively [3]. However, this is not a significant issue for a single chain in a much larger simulation box where the cross-sectional area is taken to be its occupied area in a unit cell, a value that changes little during the deformation. Therefore, we applied strains up to 0.5. In the cellulose Iß structure, one unit cell consisting of two molecules has a cross-sectional area of 63.42 Å$^2$. Thus, an approximate area for a single molecule is half of this value, or 31.71 Å$^2$. This is the value used in the elasticity calculations. We also note that in the physical system, intramolecular covalent bonds would break at some strain. However, few force fields, including the one used here, can account for breaking of covalent bonds. Introducing this capability into the model is an active area of research.

Initially, all finite and infinite chains have the same unit cell length, 10.38 Å [7]. However, the equilibrium length calculated during the deformation process was found to exhibit some variation between 10 Å and 11 Å **(Fig. 3).** This is consistent with a previous observation that an isolated single cellulose molecule will have a larger equilibrium length than that of a cellulose nanocrystal [3]. Figure 3 also illustrates that the potential energy of finite and infinite molecules is 25-35 kcal/mol/glucose. Thus, we observe that the equilibrium unit cell length and potential energy are not strong functions of boundary conditions or the number of repeat units. In addition, significant variation in the data for short chains indicate that at least six repeat units should be used for calculations based on




*4. A representative plot of potential energy (left y-axis; normalized by 6202 kcal/mol) and corresponding number of hydrogen bonds (right y-axis; normalized by 13) as functions of strain. Data from a simulation of an infinite chaing with eight repeat units.*

model-predicted length and energy. In the finite chain case, error in short chain calculations is the result of molecule end effects [8] and the configuration being dependent on fewer atoms. The latter effect is the primary source of error observed for infinite chains with a small number of repeat units.

## RESULTS

**Figure 4** shows a representative plot of the change in potential energy with strain (black squares). As expected, energy increases rapidly with strain. The total potential energy is the result of the contributions of nonbonded (van der Waals, Coulomb, hydrogen bonds) interactions and bonded interactions (i.e., covalent bonds). Figure 4 shows that the energy associated with nonbonded interactions (red circles) is larger than that of the bonded interactions (green triangles) at small strains (less than ~0.2). However, the opposite is true when the strain is large. Thus, the total potential energy is dominated by nonbonded interactions at low strain and bonded interactions at high strain. Model chains of all lengths and boundary conditions exhibited similar behavior.

The role of hydrogen bonding was evaluated by counting the number of hydrogen bonds present during deformation. A hydrogen bond is identified when the hydrogen bonding distance is less than or equal to 0.35 nm and the donor-hydrogen-accepter angle is less than or equal to 30°. As shown in Fig. 4, hydrogen bonding decreases rapidly within the first 5% strain. The hydrogen bond is modeled by a combination of van der Waals and Coulombic energies, and so contributes to the nonbonded energy. The higher nonbonded energy at small strain primarily results from the existence of hydrogen bonds, as suggested from the plot. As the strain increases, the hydrogen bonds break very quickly, and contribute less to the response of the molecule. For small strains, removing the effect of hydrogen bonds can cause predicted elastic properties to decrease by 50% to 60% [1]. Another study found that removal of hydrogen bonds resulted in a decrease of between 14% and 26% in Iα chain stiffness, while for Iß, removal caused either a 15% decrease or a 7% increase in chain stiffness, depending on the initial atomic coordinates [3]. In the low strain regime, our observations are consistent with these findings. However, our results also indicate that the dependence of elastic behavior on hydrogen bonding should be less for elasticity calculated at large strain rates.

The axial modulus (units: Pa) can be obtained using three methods. Although the functional forms of the methods are equivalent, which parameters are obtained from simulation differs, and so the source of calculation error may differ as well. Note that all three methods are in some way dependent on an estimation of area which, as discussed previously, is an approximation. The first method [3] can be expressed as:

$$\frac{F}{A} = E\left(\frac{L - L_0}{L_0}\right) \quad (1)$$

where the force $F$ (units: N) and molecule equilibrium length $L_0$ (units: m) are obtained by fitting a polynomial to potential energy $U$ (units: J); molecule length $L$ (units: m) is obtained from the simulation; and the cross-sectional area $A$ (units: m$^2$) for a single molecule must be specified.

The second method [2] can be expressed as:

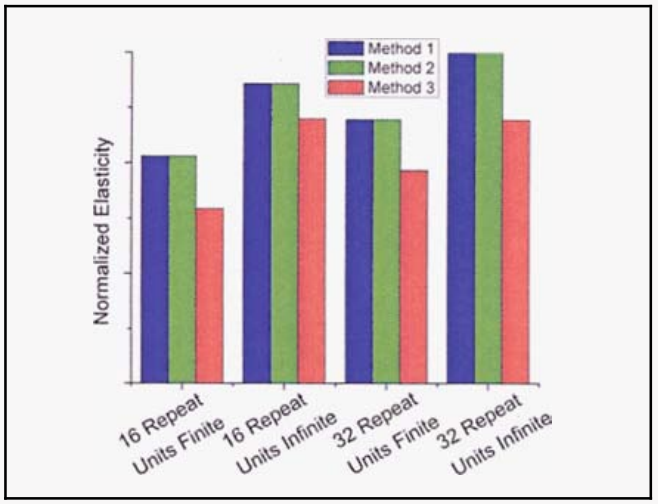
$$\frac{U}{A \times L} = E\frac{1}{2}\left(\frac{L - L_0}{L_0}\right)^2 \quad (2)$$

where U and $L$ are obtained from the simulation and and must be specified The error associated with this method is dependent on the accuracy of the specified equilibrium length. The equilibrium length of a single chain is likely to be greater than that of a crystal. Therefore, error might be introduced into the process if there is a disconnection between the model and the source of its equilibrium length. However, if the equilibrium length is calculated using a fit to the $U$ and $L$ data obtained from the simulation, we find that this method yields the same result as the first method within statistical significance **(Fig. 5).**

The third method [9] can be expressed:

$$\sigma = E\left(\frac{L - L_0}{L_0}\right) \quad (3)$$

where stress o (units: Pa) and length $L$ are obtained from the simulation $L_0$ and must be specified. At first glance, this method appears to be the simplest, and so would potentially have the last error. However, the stress is calculated from the so-called pressure virial, which contains error in itself. First, the pressure virial is a function of area, so that approximated pa-

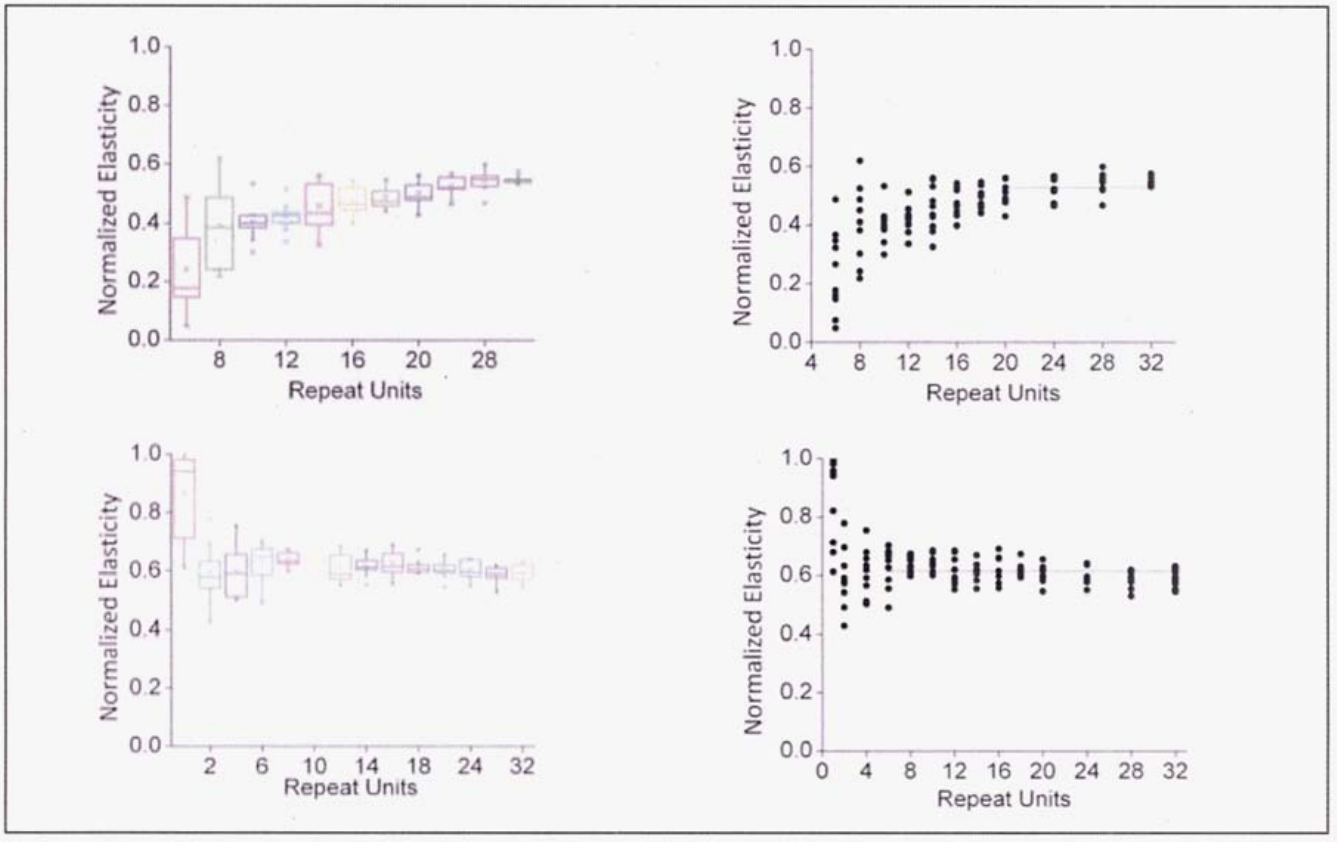



*5. Comparison of three methods for calculating axial elasticity in large strain. Data normalized by 122 GPa. Methods 1 (blue) and 2 (green) yield the same result for molecules with different lengths and boundary conditions, while method 3 (red) predicts lower values.*

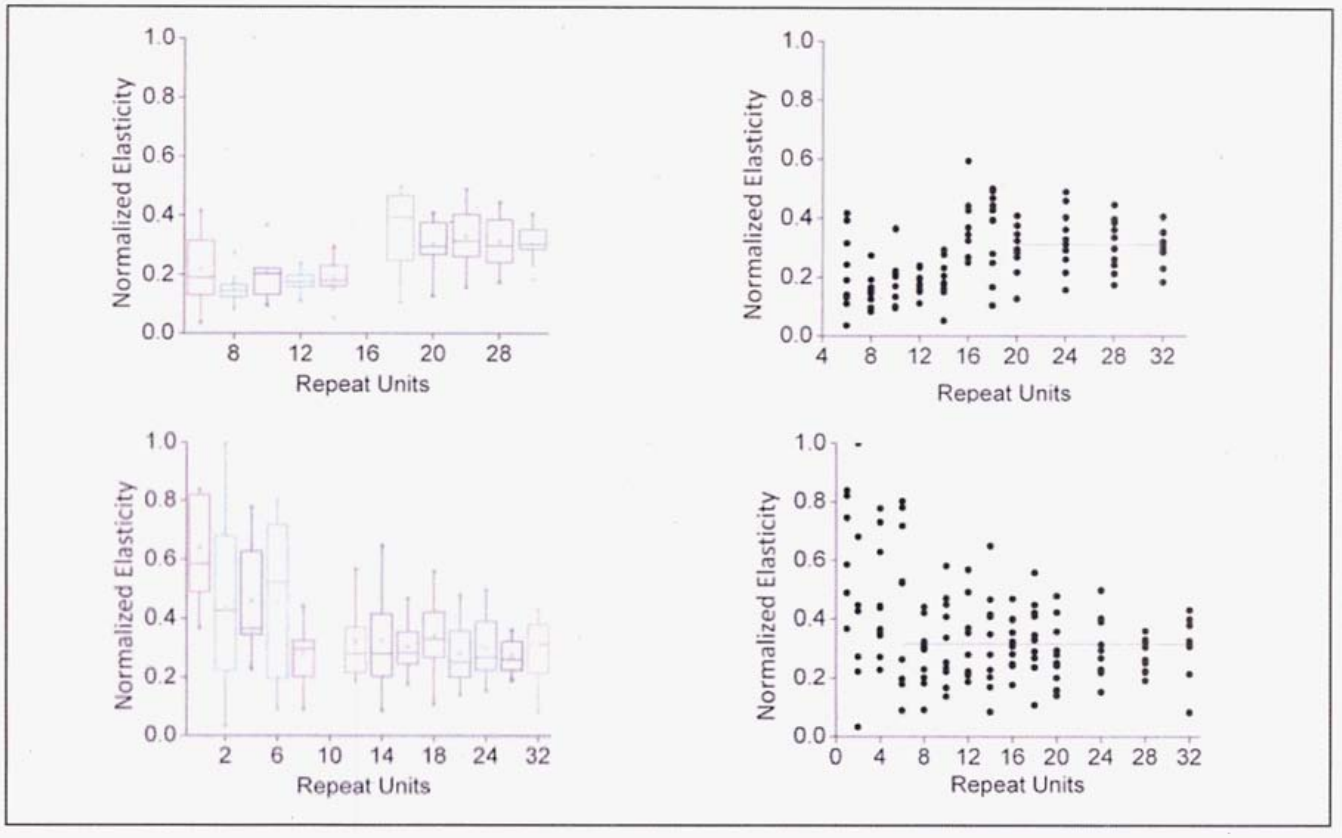
rameter is being specified implicitly. Second, the virial expression is not equivalent to the mechanical Cauchy stress [10]. As shown in Fig. 5, we find using this method yields axial moduli that differ from those predicted by the first two methods by up to 22% for infinite chains and 39% for finite chains. We therefore do not advocate use of this method and will not discuss it further here.

The axial modulus was calculated for both finite and infinite chains with different numbers of repeat units from data taken in the large and small strain regimes. **Figures 6** and **7** show the results, where the plots on the tight show the raw data and plots on the left provide a statistical representation. The statistical plots illustrate the inter-quartile range (large rectangle), mean (point inside the rectangle), median (line inside the rectangle), and maximum and minimum values (highest and lowest point, respectively). The data are normalized in these figures to highlight the fact that the trends ate not specific to the force field used.

Figure 6 illustrates the results obtained from the large strain (>0.2) regime. For both finite and infinite chains, the uncertainly in the data decreases as the chain length increases. For example, the standard deviation in the finite molecule data is 0.14 with six repeat units, and 0.01 with 32 repeat units; for infinite molecules, the standard deviation is 0.15 with six repeat units while that with 32 repeat units is 0.03 (calculations based on normalized data). In addition, the data do not appear to trend in any particular way for the infinite chains, while the axial modulus of the finite chains increases with chain length and asymptotes to the same value as that pre-




*6. Finite (top) and infinite (bottom) molecules at large strain; axial modulus normalized by 177 GPa (the highest value from calculation). The plots on the right show the raw data with a solid line indicating convergence, and left ones provide a statistical representation. Ten independent simulations were run for each case.*




7. Finite (top) and infinite (bottom) molecules at small strain; axial modulus normalized by 802 GPa (the highest value from calculation). The plots on the right showe the raw data with a solid line indicating convergence and left ones provide a statistical representation. Ten independent simulations were run for each case.

dicted for the infinite chain.

Figure 7 shows the results obtained from the small strain (<0.2) regime. As with the large strain regime, the data uncertainly decreases with increasing chain length. However, even with the longest chain the uncertainty is six times that obtained from the same model in the large strain regime. In addition, the magnitude of the axial modulus is much larger in the small strain regime, converging to a value two to three times that of large strain in both finite and infinite cases.

These observations can be understood in terms of the role of potential energy and hydrogen bonds. As shown in Fig 4, the energetics of the low strain regime is dominated by non-bonded interactions associated with the presence of hydrogen bonds. These hydrogen bonds are extremely strong and so contribute to a high axial modulus. In the high strain regime, where most hydrogen bonds have broken, the energetics is dominated by the role of bonded interactions, and we observe a correspondingly lower axial modulus.

Comparison of the axial moduli reported here to those in the literature poses a challenge because, as shown here, results of a molecular simulation differ significantly with model parameters and can be expected to vary even for a fixed set of conditions. The axial modulus predictions in Figs. 6 (large strain) and 7 (small strain) approach 109.4 Gpa and 254.8 GPa, respectively. Molecular simulation-based predictions of the axial modulus of cellulose Iß are typically in the range of 110-180 GPa [9]. So, although there is certainly overlap between our results and those reported elsewhere, it is difficult to reach definitive conclusions. In terms of experimental results, the comparison is even more difficult to make. For example, the axial modulus for cellulose Iß obtained from x-ray scattering was reported to be 220±50 GPa [11]. Further, the elasticity of a single chain will likely differ from that of a crystal because of the role of inter-chain hydrogen bonding. For example, we have found the mean axial elasticity for a 4x8x8 cellulose crystal to be 155 GPa, whereas that of the corresponding eight unit single chain is 113 GPa. Complicating the issue further is the cooperative hydrogen bonding, such that omitting interchain hydrogen bonding (as is necessarily the case for a single chain) will affect intrachain hydrogen bonding [12]. However, while the results shown in this paper cannot be directly extrapolated to cellulose nanocrystals, the trends observed both in the magnitude and the uncertainty of a model-predicted axial modulus are generally applicable.

## CONCLUSION

This study highlights the effects of chain length, strain regime, and calculation method on axial modulus prediction using atomistic simulation and the uncertainly associated with that prediction. All three of these considerations affect the model predictions and need to be taken into careful consideration when using atomistic modeling or interpreting data ob-

tained from it. The limitation of this study is that it was performed with single chains modeled using only one force field (COMPASS). However, a comprehensive study that includes characterization of multiple force fields and the effect of model size, including crystal structure, is underway. These types of studies are critically necessary to enabling interpretation of model data. Only by understanding how "what goes in" affects "What comes out" can atomistic simulations transition from an investigative method to a predictive tool useful for cellulose-based application design. **TJ**

## LITERATURE CITED


1. Moon, R., Martini, A., Nairn, J.A., et al., Chem. Soc. Rev., in press, 2011, DOI:10.1039/C0CS00108B,.
2. Tanaka, F. and Iwata, T., Cellulose 13(5): 509(2006).
3. Eichhorn, S.J. and Davies, G.R., Cellulose 13(3): 291(2006).
4. Sun, H., Ren, P., and Fried, J.R., Comput Theor. Polym. Sci. 8(1-2): 229(1998).
5. Miyamoto, H., Umemura, M., Aoyagi, T., et al., *Carbohydr. Res.* 344(9): 1085(2009).
6. Plimpton, S., *J. Comput Phys.* 117(1): 1(1995).
7. Nishiyama, Y., Langan, P., and Chanzy, H., *J. Am. Chem. Soc.* 124(31): 9074(2002)
8. Mazeau, K. and Heux, L., *J. Phys. Chem. B* 107(10): 2394(2003).
9. Bergenstråhle, M., Berglund, L.A., and Mazeau, K., *J. Phys. Chem. B* 111(30): 9138(2007).
10. Zhou, M., *Proc. R. Soc. Lond. A* 459(2037): 2347(2003).
11. Diddens, I., Murphy, B., Krisch, M., et al.. *Macromolecules* 41(24): 9755(2008).
12. Qian, X., *Mol. Simul.* 34(2): 183(2008).


## ABOUT THE AUTHORS


With the increasing availability and use of atomistic simulations to characterize organic materials, we felt it was important to step back and consider, for the very simple system of a single cellulose chain, how parameters that define a model affect its predictions. Significant work in studying cellulose using atomistic simulation has been conducted since the mid-1480s. and the models have progressed in complexity, size, and accuracy. However, results are typically reported without considering variations due to small changes in model parameters or simply because of the statistical nature of the models themselves. Therefore, this work complements previous studies by providing a reference point and hopefully setting a precedentfor future research to report the subtle, yet important, model details that affect predictions and the uncertainty in those predictions.

The most significant challenge in this study was isolating individual model parameters to determine their effects independently. Atomistic models are extremely complex and we could only hope to make quantitative statements about the roles of model parameters by defining a problem simple enough that individual effects could be identified, yet complex enough to approach something experimentally measurable. We chose to focus on the axial modulus of individual cellulose chains with which we could explicitly control characteristics such as chain length and strain, and where there is complementary (albeit not directly comparable) experimental data available for cellulose nanocrystals. We were very surprised to discover just how much variation was inherent in these models, even when all input parameters were constant. However, along with the realization that uncertainty exists, came the understanding that we could quantify it and so take its effect into account when reporting atomistic predictions. The critical analyses of model predictions in this study are an important first step, and will form the basis for more comprehensive characterizations, including multiple force fields and various crystal sizes and structures. This work will help support future research in developing new engineered nonwoven materials and composites that are based on cellulose nanoparticles.

*Wu is a graduate research assistant and Martini is assistant professor with the School of Mechanical Engineering, Purdue University, West Lafayette, IN, USA. Moon is a materials research engineer with the U.S. Forest Service, Forest Products Laboratories, and adjunct assistant professor with the School of Materials Engineering, Purdue University, Birck Nanotechnology Center, West Lafayette, IN, USA. Email Martini at a-martini@purdue.edu.*

TAPPI JOURNAL
WWW.TAPPI.ORG
THE PAPER AND PACKAGING INDUSTRIES' TECHNICAL RESOURCE
APRIL 2011 | VOL. 10 NO. 4



SPECIAL NANOTECHNOLOGY ISSUE