PNAS
www.pnas.org

Supporting Information for

Nanomechanics of Cellulose Deformation Reveal Molecular Defects that Facilitate Natural Deconstruction

Peter N. Ciesielski,[a*] Ryan Wagner,[b] Vivek S. Bharadwaj,[a] Jason Killgore,[b] Ashutosh Mittal,[a]

Gregg T. Beckham,[c] Stephen R. Decker,[a] Michael E. Himmel,[a] and Michael F. Crowley[a*]

*Corresponding Authors: Dr. Peter Ciesielski, Dr. Michael Crowley
Email: peter.ciesielski@nrel.gov , michael.crowley@nrel.gov

This PDF file includes:

Materials and Methods
Figs. S1 to S6
References for SI reference citations

www.pnas.org/cgi/doi/10.1073/pnas.1900161116

Materials and Methods

Isolation of Cladophora Cellulose Nanofibrils. Cellulose nanofibrils used in this study were purified from green alga Cladophora sp. according to the method reported previously(1). Briefly, the cellulose was purified by repeated boiling in 0.2 N NaOH for 6 h, rinsing and immersing in 0.1 N HCl at room temperature overnight followed by thorough washing to neutral pH. The resulting material was further purified by a treatment at 80°C for 2 h in 0.25 Wt % $NaClO_2$, washed and autoclaved in 1 % H_2SO_4 at 121 °C for 1 h followed by thorough washing to neutral pH. The purified cellulose sample obtained was freeze dried stored at room temperature for further use.

Atomic Force Microscopy. A 0.1 wt % aqueous suspension of isolated *Cladophora* cellulose nanofibrils was drop cast onto plasma cleaned Silicon <111> and track etched polycarbonate (TEPC) substrates. The TEPC consists of 200 nm diameter pores that extend though the substrate. A commercial AFM (Cypher, Asylum Research, Santa Barbara, CA) was used for tapping-mode AFM imaging, lithographic nanomanipulation, and force displacement (F-Z) measurements on the cellulose. Nanomanipulation experiments were performed with aluminum-coated silicon cantilevers (NCLR, Nanosensors, Switzerland) with nominal spring constant k_L = 48 N/m, length L = (225 ± 10) µm, and first free flexural resonance frequency of $f_{1,free}$ = 190 kHz. Force versus displacement measurements were performed with aluminum-coated silicon cantilevers (SEIHR, Nanosensors, Switzerland) with nominal spring constant k_L = 15 N/m, length L = (225 ± 10) µm, and first free flexural resonance frequency of $f_{1,free}$ = 130 kHz. The measured value of k_L for the specific cantilever used in each experiment was determined from the Sader method with hydrodynamic correction and the optical lever sensitivity was determined with the corrected thermal method(2). All uncertainties are reported to the 95 % confidence interval. Nanolithographic manipulation of the cellulose was performed by imaging the surface in tapping mode to identify a fibril of interest, then bringing the tip into quasistatic-contact (i.e. contact mode) with the substrate near the cellulose and imposing a defined linear tip-path across the cellulose while maintaining static force of 2.1 µN ± 0.2 µN with a feedback loop. Path lengths varied from 100 nm to 1 µm (indicated by blue arrows in SI Appendix Fig. S1 and SI Appendix Fig. S2), with a tip velocity of 100 nm/s. The resultant cellulose conformation was imaged again in tapping mode after each manipulation pass. Typically, the initial manipulation trajectory does not cross the entirety of the fiber and serves to loosen the cellulose fiber form the substrate, allowing less-constrained modification in subsequent steps.

Force versus distance spectroscopy experiments were performed by first surveying the TEPC substrate in tapping mode to identify single cantilevered segments of cellulose. Such segments were best identified in the phase imaging channel. Next, a series of 40 force versus distance (F-Z) curves were acquired along the length of the cellulose starting before the supported edge and continuing towards the free end with 2 nm to 5 nm spacing between curves. After each pass, the fibril was reimaged, and the maximum force was increased up to a maximum of 100 nN if kinking was not observed. From the low-force F-Z data, the stiffness versus position response of the cellulose TEPC system was fit to a static beam model for a single cantilevered beam with support compliance given by the indentation stiffness of the substrate-bound fibril. The F-Z response is given by

$$\frac{1}{k_{FZ}} = \begin{cases} \frac{1}{k_b} + \frac{(L_{FZ} - L_{0,\mathrm{fit}})^3}{3EI} & L > L_{0,\mathrm{fit}} \\ \frac{1}{k_b} & L < L_{0,\mathrm{fit}} \end{cases},$$

where k_{FZ} is the equivalent stiffness of the FZ curve, k_b is the stiffness of the substrate-bound fibril, L_{FZ} is the length measured from the start location of the FZ curve series, $L_{0,\mathrm{fit}}$ is the location of the support edge, E is the bending modulus of the cellulose, and I is the cross-section area moment of inertia. We have assumed that the cross section of the cellulose is a rectangle with height h and width b, thus $I = \frac{1}{12} b h^3$. The value of b cannot be measured directly with AFM due to convolution of the tip shape, thus we have assumed that the width to height ratio of the cellulose cross section is a constant $b/h = 1.9 \pm 0.2$ based on statistical analysis of AFM height data and TEM width data in Extended Data Figure 3. The value of $L_{0,\mathrm{fit}}$ is difficult to determine from visual inspection of the AFM data because of the irregular pore profile. Therefore, we fixed all other parameters in the model and used our experimental data to calculate $L_{0,\mathrm{fit}}$. The goodness of the beam model fit varied significantly for individual fibrils, with subsequent stress analysis carried out only for those fibrils exhibiting strong agreement with the model, a reliable flat background for height measurement, and a boundary stiffness consistent with cellulose resting on a homogeneous polycarbonate substrate (i.e. the indentation stiffness was not so low that it suggested a contaminant or cavity beneath the cellulose). From the fit of the elastic static bending response, the effective location of the support was calculated. The kinking stress σ_{kink} in the fibril at the supported end is then given by

$$\sigma_{kink} = \frac{F_{\max} h L}{2I},$$

where $F_{\max}$ is the max force before kinking and slipping and $L = L_{FZ} - L_{0,\mathrm{fit}}$ is the distance from the support location to the location where $F_{\max}$ was applied. Furthermore, there was a strong tendency for the AFM tip to slip off the cellulose at extreme bending angles when the force was applied at the free end of the fibril. To achieve larger kinking stresses at the supported-end of the fibril, without slipping, the load was applied successively from the supported- towards the free-end. Subsequently, the max force was incremented higher, and the successive loading positions were repeated. The subsequent higher-force pass was continued until the F-Z curve exhibited a characteristic deviation from linear elastic response as seen in the portion of the loading curve exceeding $F_{\max}$ in Figure 2F of the main text. Intermittent contact imaging after each force step allowed slipping without kinking to be differentiated from slipping with kinking. Of the 8 kinked data sets, 2 exhibited sufficient goodness of fit to beam theory and low enough uncertainty in height to warrant subsequent analysis of kinking stress. As shown in Figure 2E, the last series of force curves that did not induce kinking was used to perform an analysis of the elastic bending properties of the cellulose. Measuring the height of the cellulose against the TEPC, determining a cross-sectional aspect ratio from statistical analysis of AFM and TEM images of fibrils, and assuming a Young's modulus E = 150 GPa, the effective support location $L_{0\mathrm{fit}}$ was determined.

Transmission Electron Microscopy. Samples were diluted to 0.1 wt % solids and drop-cast onto carbon-coated copper TEM grids (VWR) that were freshly plasma-cleaned. After one minute, the excess solution was removed, and the grid was post-stained with 2 wt % aqueous uranyl acetate (Sigma Aldrich) for 2 min. Excess solution was removed by blotting with filter paper and the grid was allowed to dry in ambient conditions prior to imaging. Images were obtained with a four mega-

pixel GatanUltraScan 1000 camera (Gatan, Pleasanton, CA) on a FEI Tecnai G2 20 Twin 200 kV LaB_6 TEM (FEI, Hillsboro, OR) using Digital Micrograph image capture software.

Cel7A Production and Purification. *Trichoderma reesei* Cel7A (Cellobiohydrolase I (CBH1)) was produced and purified as previously described(3). Briefly, a *cbh1*-delete strain of *T. reesei* QM6a, (AST1116) was used as the expression host after transforming with the native *T. reesei cbh1* gene under the control of the native constitutive *eno* promoter, generating strain JLT102A. The JLT102A strain was grown at 8 L scale on Mandels-Andreotti media containing 5% glucose in a CSTR operating at 28°C, 500-800 rpm, pH 4.8, and 1.0 vvm air/O_2 mix to achieve 20% dissolved O_2. Mycelia were removed by filtration through a 5 um nominal pore size polyethylene industrial bag filter (McMaster-Carr) and the filtrate was diafiltered into 20mM Bis-tris pH 6.5 buffer and concentrated ~20X using a 5,000 MWCO hollow fiber cartridge (GE Lifesciences). After addition of $(NH_4)_2SO_4$ to 1.6M and 0.45 um filtration, the concentrated culture broth was loaded onto a 7.8 mL Source15-Phenyl Tricorn FPLC column (GE Lifesciences) and eluted with a descending 1.6-0.0M $(NH_4)_2SO_4$ gradient in 20 mM Bis-Tris pH 6.5. Active fractions were pooled, diafiltered into 20mM Bis-Tris pH 6.5 to remove $(NH_4)_2SO_4$ and loaded onto a 7.8 mL Source15-Q FPLC column, eluted with an increasing 0.0-0.5M NaCl gradient in 20 mM Bis-Tris pH 6.5. Active fractions were pooled, adjusted to 1.6M $(NH_4)_2SO_4$ by dilution with an equal volume of 3.2M $(NH_4)_2SO_4$ in 20 mM Bis-Tris pH 6.5 and loaded onto a 7.8 mL Source15-Iso column, eluted with a descending 1.6-0.0M $(NH_4)_2SO_4$ gradient in 20 mM Bis-Tris pH 6.5. Active fractions were pooled, concentrated, and diafiltered into 20mM sodium acetate containing 100mM NaCl, then loaded onto a Superdex75 10/60 Prep Grade size exclusion FPLC column eluted with an isocratic gradient in the same buffer. Active fractions were pooled and concentrated and protein concentration was determined by A_{280} using an extinction coefficient of 86760 M^{-1} cm^{-1} (assuming all cysteines form cystines). Purity was determined as a single band by SDS-PAGE. For all purification steps, activity was monitored by hydrolysis of 2.0 mM p-nitrophenyl-β-D-lactobioside in 50 mM sodium acetate pH 5.0 buffer (10 uL fraction + 90 uL substrate, 15 min at 45°C) by measuring absorbance at 405 nm after addition of an equal volume of 1.0 M $NaHCO_3$.

Enzymatic hydrolysis. Enzymatic hydrolysis of purified Cladophora sample was performed in duplicate in 2 ml polypropylene cryovials (Simport – T308-2A) at 0.2% solids loading, at 50°C for a digestion period of 2 h, primarily for imaging purposes. Cel7A was added at 50 mg protein per gram of glucan which was supplemented with B-glucosidase at 1 mg protein per gram of glucan to reduce the physical barrier of hemicelluloses and enhance cellulose accessibility to cellulases. The total volume of the saccharification slurries after adding enzymes and 50 mM citrate buffer (pH 4.8) was 2 mL. Digestions were continuously mixed by inversion at 10–12/min intervals to avoid the separation of solids from liquid. After completion of the 2 h digestion period, an aliquot was taken and stored at 4°C for imaging analysis.

Molecular Simulation. A nanofibril of cellulose I-β degree of polymerization 100 was constructed with the crystal parameters specified in Langan, et al(4-6). A 36-chain model was constructed such that the cross-sectional geometry resembled a diamond as shown in Extended Data Figure 5. All the simulations were conducted using CHARMM program(7) with the CHARMM carbohydrate force field, all36_carb(8, 9). The DOMDEC fast parallel CHARMM method(10) was used in the molecular dynamics. All simulations used a 2 fs timestep and SHAKE(11) to keep the length of bonds to hydrogen fixed. All simulations are performed in vacuum and used a nonbonded cutoff

of 13 Å, with a 2 Å list buffer and heuristic list update using a 1 Å switching function for both dispersion and electrostatics. Systems were equilibrated with initial minimization and 2 ns unrestrained dynamics at 300 K with a Langevin thermostat. All bending simulations were done without a thermostat but were checked at 1 ps intervals and velocities reset to a Boltzmann distribution for 300 K if the temperature was further than 10 K away from 300K.

Glycosidic bond breakage

Bond distance vs. energy profiles for stretching across the C1-C4 atoms were obtained by performing scans of the C1-C4 bond distance using the unrestricted M06-2X level of theory and CBSB7 basis set(12, 13). The unrestricted molecular orbital calculations enable the description of homolytic cleavage of the covalent C-O bonds during the C1-C4 scan. The C1-O4 and C4-O4 bond distances were monitored along with the system energies during these scans. The C1-O4 bond is observed to have a well at 1.39 Å while the C4-O4 bond is observed to have its well at 1.41 Å, which is consistent with the glycosidic bond distances observed in cellulose. The stretching behavior of the 2 bonds diverges beyond 1.55 Å, with the C4-O4 stretch plateauing at 1.58 Å and the C1-O4 stretch continuing upto at 1.88 Å. This is indicative of the glycosidic oxygen remaining bound to C4 and the C1-O4 bond breaking. This is consistent with observations in enzymatic hydrolysis by cellobiohydrolases, which cleave the C1-O4 bond.

Nanofibril Deformation

Restrained bending simulations were designed to mimic the AFM two-point bending experiments by restraining one end of the nanofibril as if it were on a surface and applying a force to the center of mass of a region on the other end of the fibril in a specific direction as shown in SI Appendix Fig. S5 and Fig. S6. The restraining forces on the surface end affected sugar residues one through fifty in order to mimic the substrate surface, as if the cellulose was stationed on flat surface with a shelf edge at residue 50. The restraints are on the first 50 residues in the chains depicted as black in Extended Data Figure 6, on each side, only one side is visible in the figure. Harmonic restraints are applied to the center of mass of each of the first fifty sugars in eight chains, four on each side, preventing movement horizontally and twisting of the fibril on the surface. The bending force was applied to the center of mass of the last 15 sugars of all 36 chains. All other atoms moved without applied restraints.

Deformation was implemented by applying a constant downward force, using the CHARMM "pull" method, on the center of mass point. The biasing force was applied to the center of mass, and thusly distributed among all the atoms in the last 15 sugars of all chains, avoiding severe disruption in a localized region of the fibril. A sequential series of simulations were performed with biasing forces of 0.07 to 5.0 piconewtons. Each simulation was performed to exceed the time needed to come to equilibrium under the applied load, and required between 2 and 30 ns to reach equilibrium. At regular 100 ps intervals, all glycosydic bonds were examined to determine which bonds exceeded the breaking criterion for at least 10 ps. Those meeting the criterion were replaced with hydrolyzed glucan chain ends and the system was energy-minimized for 1000 steps to remove the excess energy from the released bond strain. The objective of these simulations was to determine the equilibrium state at each applied force including deflection, crystalline structure disruption, and bond breakage. It was not an objective to simulate accurately the dynamic process of breaking, which would have been far to computationally expensive and likely not have yielded

any significant insight. However, this can be a focus of future simulations with higher-accuracy non-newtonian approaches.

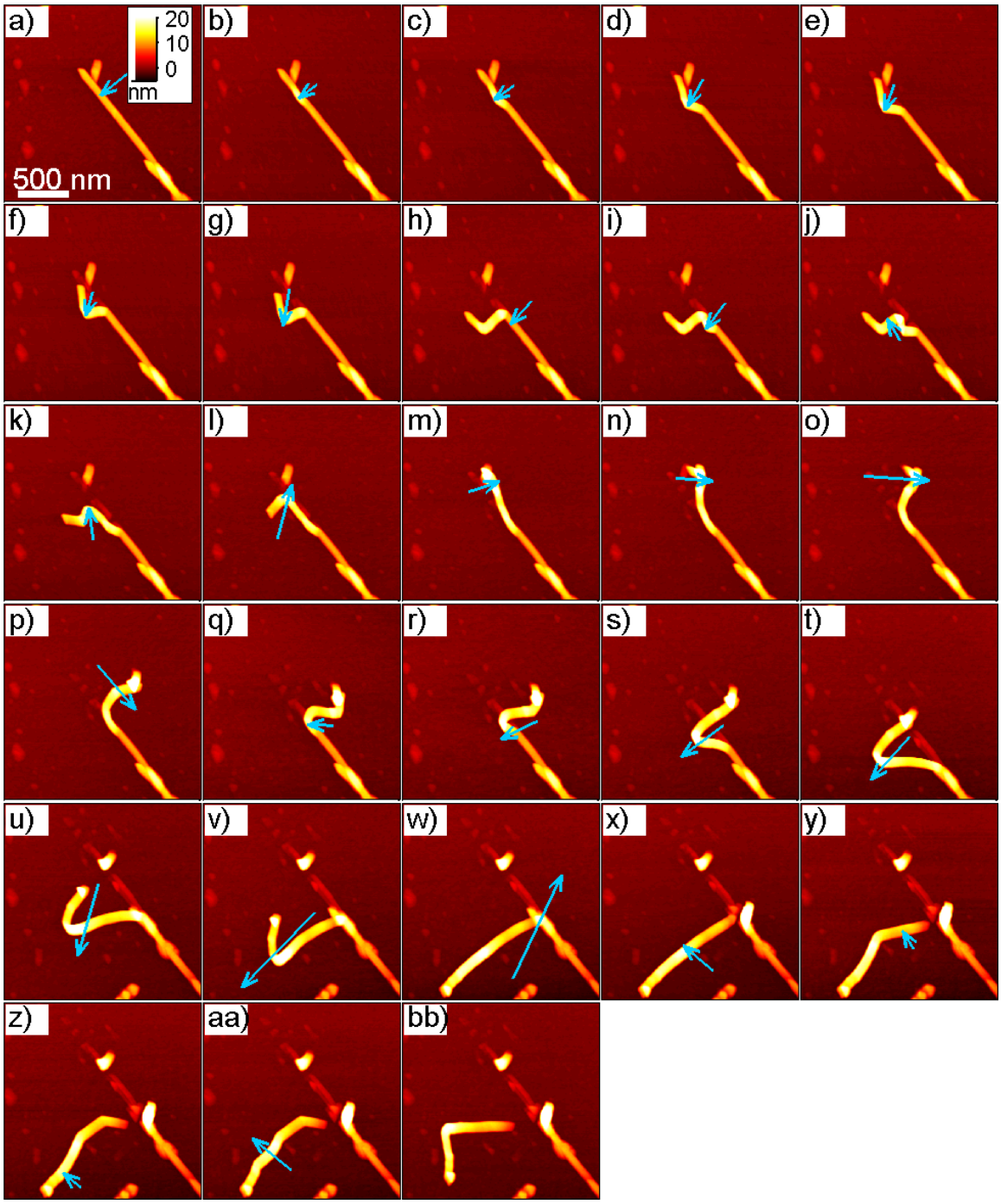

Fig. S1. Complete series of manipulations for Figure 2A and Figure 2C in the main text. The blue arrow indicates the trajectory of the AFM tip.

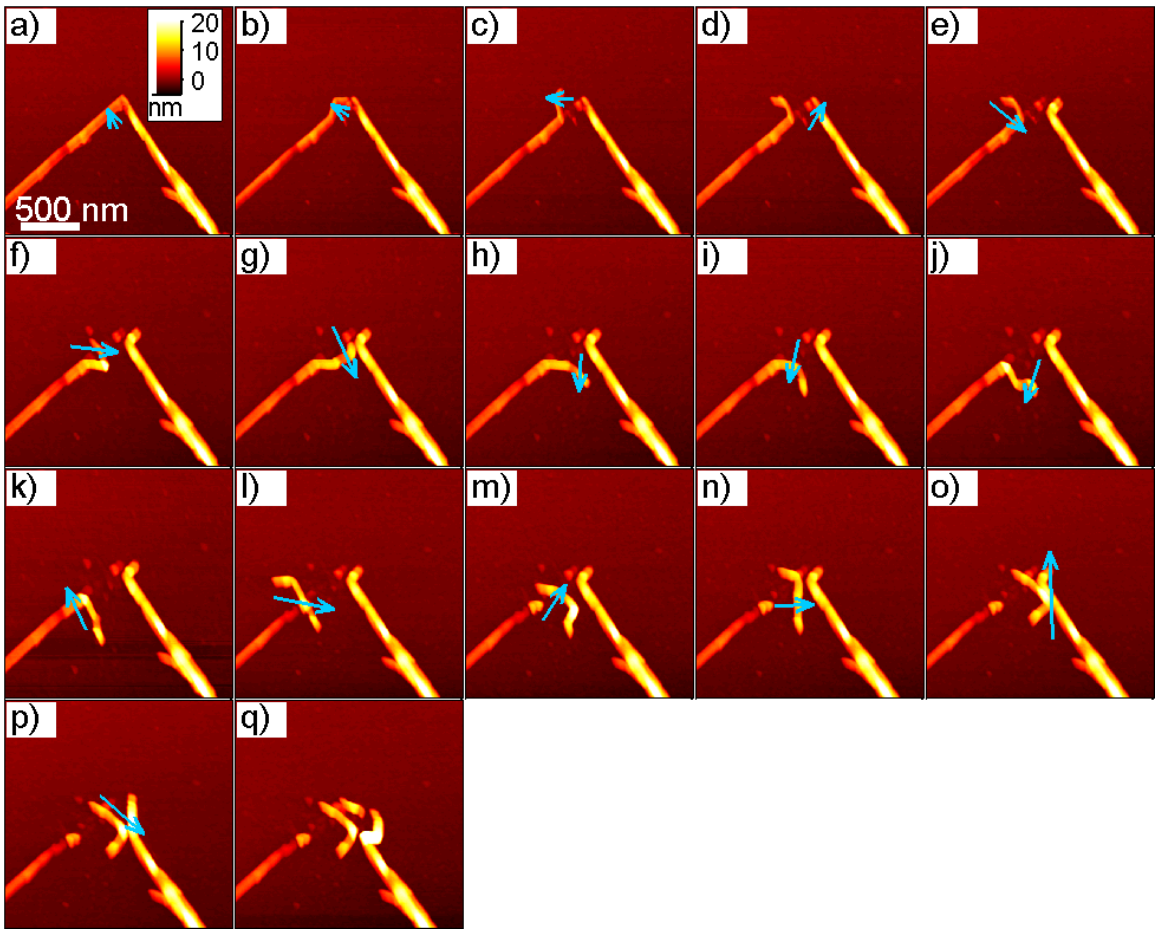

Fig. S2. Complete series of manipulations for Figure 2B in the main text. The blue arrow indicates the trajectory of the AFM tip.

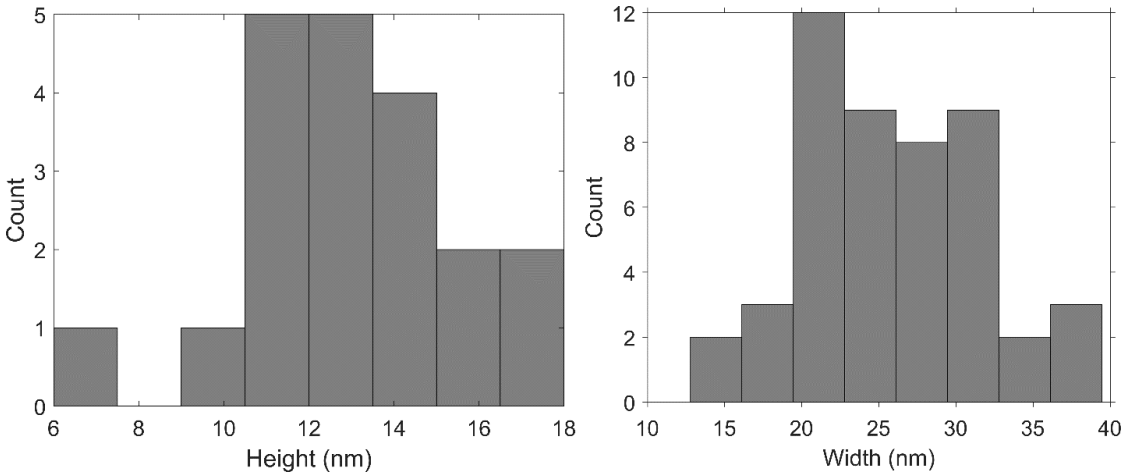

Fig. S3. AFM height data (left) and TEM width data (right) from cellulose fibrils. From these data and the assumption of a constant fibril aspect ratio, we estimate a width to height ratio of 1.9.

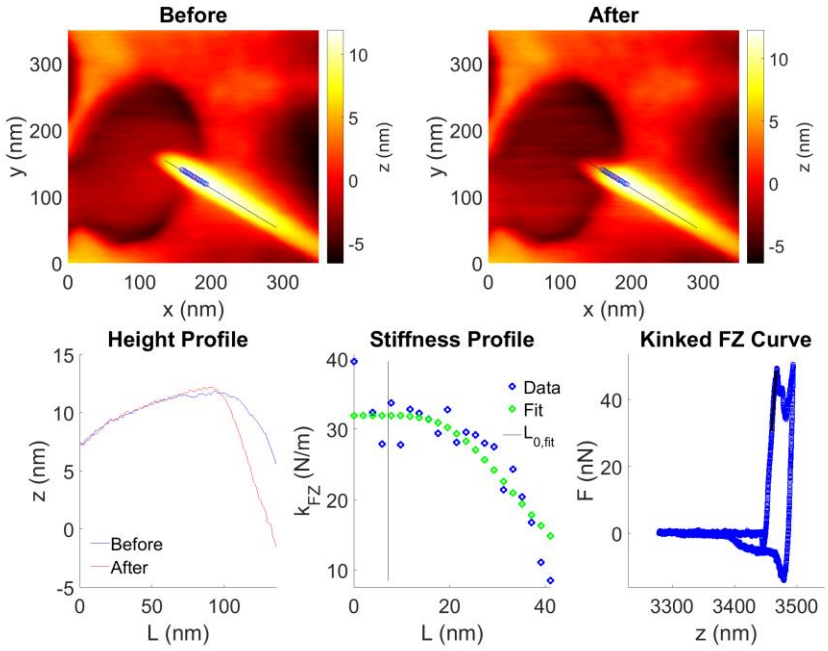

Fig. S4. Alternate two-point bend tests data set. σ_max=3.2±1.1 GPa.

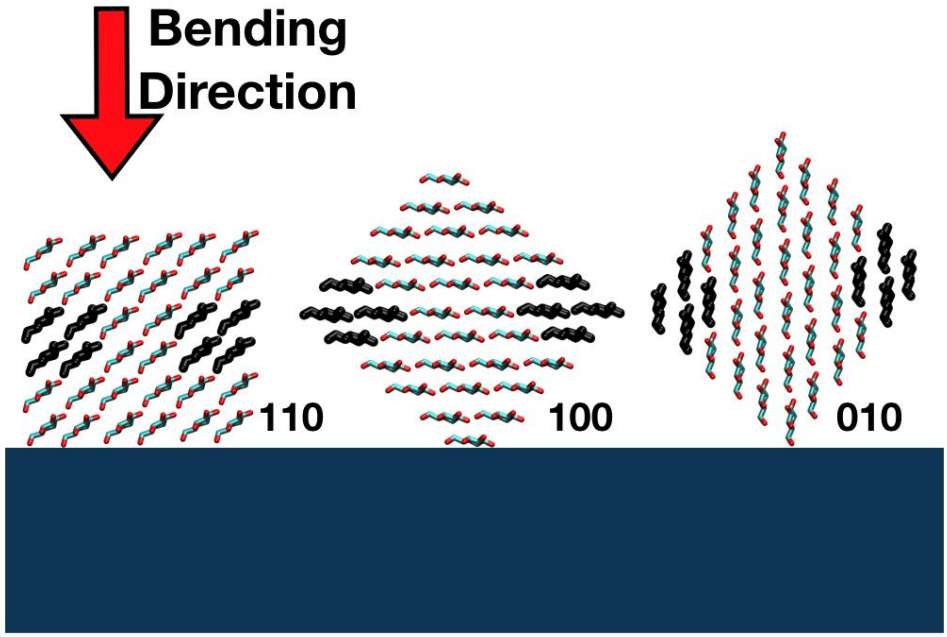

Fig. S5. Cross sections of 36-chain fibrils used in bending simulations. The black rectangle represents the imaginary surface fibrils are deposited on. Fibrils were oriented to mimic bending perpendicular to each of the crystallographic planes: 110, 100, and 010. The restrained chains are highlighted in black for each of the fibril orientations.

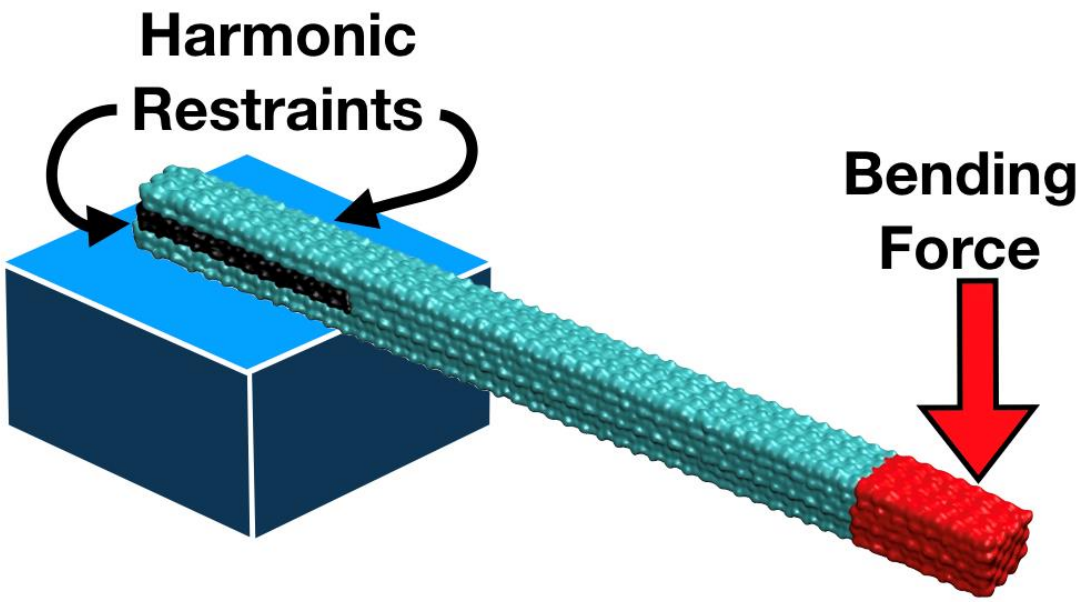

Fig. S6. Bending simulation setup. Blue box is the imaginary platform that the harmonic restraints mimic, the black are the restrained sugars, and the bending force is applied to the center of mass of the red region of the fibril.

References

1. Sugiyama J, Persson J, & Chanzy H (1991) Combined infrared and electron diffraction study of the polymorphism of native celluloses. *Macromolecules* 24(9):2461-2466.
2. Proksch R, Schäffer T, Cleveland J, Callahan R, & Viani M (2004) Finite optical spot size and position corrections in thermal spring constant calibration. *Nanotechnology* 15(9):1344.
3. Linger JG, *et al.* (2015) A constitutive expression system for glycosyl hydrolase family 7 cellobiohydrolases in Hypocrea jecorina. *Biotechnol Biofuels* 8(1):45.
4. Nishiyama Y, Langan P, & Chanzy H (2002) Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction. *J Am Chem Soc* 124(31):9074-9082.
5. Langan P, Sukumar N, Nishiyama Y, & Chanzy H (2005) Synchrotron X-ray structures of cellulose Iβ and regenerated cellulose II at ambient temperature and 100 K. *Cellulose* 12(6):551-562.
6. 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 Iβ. *Biomacromolecules* 9(11):3133-3140.
7. Brooks BR, *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* 30(10):1545-1614.
8. Guvench O, Hatcher E, Venable RM, Pastor RW, & MacKerell Jr AD (2009) CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *J Chem Theory Comput* 5(9):2353-2370.
9. Guvench O, *et al.* (2011) CHARMM additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate–protein modeling. *J Chem Theory Comput* 7(10):3162-3180.
10. Hynninen AP & Crowley MF (2014) New faster CHARMM molecular dynamics engine. *J Comput Chem* 35(5):406-413.
11. Kräutler V, Van Gunsteren WF, & Hünenberger PH (2001) A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. *J Comput Chem* 22(5):501-508.
12. Montgomery JA, Frisch MJ, Ochterski JW, & Petersson GA (1999) A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *J Chem Phys* 110(6):2822-2827.
13. Zhao Y, Schultz NE, & Truhlar DG (2006) Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J Chem Theory Comput* 2(2):364-382.