Supporting Information for

A Novel Supra Coarse-grained Model for Cellulose

Aleksandar Y. Mehandzhiyski,[*1] Nicolas Rolland,[1,2] Mohit Garg,[1] Jakob Wohlert,[3,4] Mathieu Linares,[1,5,6] and Igor Zozoulenko[*1,7]

[1] Laboratory of Organic Electronics, ITN, Linköping University, 60174 Norrköping, Sweden

[2] Division of Theoretical Chemistry and Biology, KTH Royal Institute of Technology, Sweden

[3] Department of Fibre and Polymer Technology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden

[4] Wallenberg Wood Science Centre, KTH Royal Institute of Technology, 100 44 Stockholm, Sweden

[5] Department for Media and Information Technology, ITN, Linköping University, 60174 Norrköping, Sweden

[6] Swedish e-Science Research Centre (SeRC), Linköping University, 60174 Norrköping, Sweden

[7] Wallenberg Wood Science Centre, Linköping University, 60174 Norrköping, Sweden

Email: aleksandar.mehandzhiyski@liu.se; igor.zozoulenko@liu.se

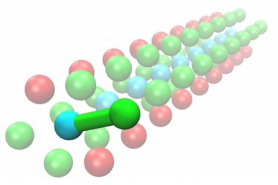
Force field parameters

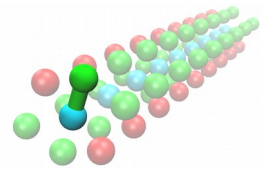
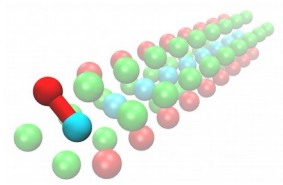
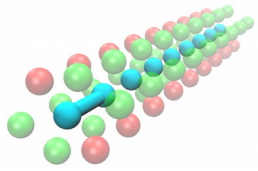
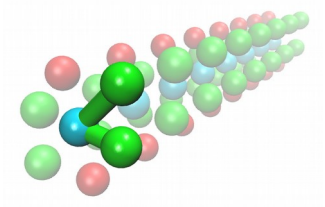
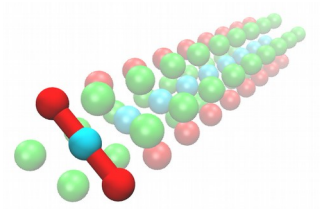
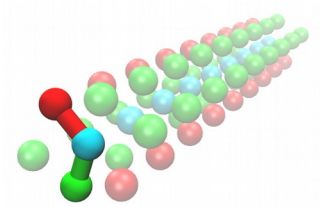
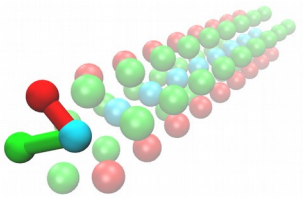
All bonded and non-bonded force field parameters for the sCG model are presented in Table S1 and Table S2, respectively. There are 3 types of beads in the model – one representing the center of the CNC (C), one for the 110 surface (I) and one for the 200 surface (O), which are represented in blue, green and red in Table S1, respectively. There are three main type of bonds – center to 110 surface (CI); center to 200 surface (CO); center to center (CC). The CI bond is separated into two subgroups, due to the different bond length connecting the respective regions. There are six different types of angle bending – 110-center-110 angle (ICI); 200-center-200 angle (OCO); 110-center-200 (ICO-1 and ICO-2); 110-center-center (ICC); 200-center-center (OCC); center-center-

center (CCC). Three dihedral angles are constructed for the sCG model. Two of them are proper dihedrals – 110-center-center-110 (ICCI), 200-center-center-200 (OCCO) and one improper dihdral , which keeps the cross-section of the CNC planar. It should be noted here that some parameters such as the bond between the hydrophilic and hydrophobic part “I-O-bond” and the respective angle “CIO” between the center-hydrophilic-hydrophobic are excluded from the parameterization of the sCG model. The I-O bond for example depends on the CI, CO-bonds and the ICO, ICI and OCO-angles, which already present in the model, see Table S1. The CIO-angle is not included in the model because it is dependent on the ICO, OCO and ICI-angles.

The parameters of the structure (CC-bond, CCC-angle) are fitted to reproduce and match the all-atom distributions. The backbone of the fiber in our model is described by the CC-bond (see Table S1) which in the atomistic model is represented by all 36 chains composing the fiber. The central bead (C) is composed of 6 chains in the cross-section of the fiber and the force constant (k^b) obtained after the force-matching was k^b=90000 kJ/mol/nm^2, which corresponds to a force constant of k^b=15000 kJ/mol/nm^2 per chain. Thus, the bond strength of the cross-section made of 36 chains is k^b_{total}=15000×36=540000 kJ/mol/nm^2. (Note that even though the central bead C is composed of 6 chains, the CC-bond includes interaction of all 36 chains; therefore, its force constant is significantly larger than all other force constants listed in Table S1). Note that in the calculations we use a slightly lower force constant k^b_{total}=500000 kJ/mol/nm^2 in order to account for an inherent disorder, i.e., amorphous regions always presenting in the realistic fiber.

Table S1. Bonded interaction parameters of the supra coarse-grained model for cellulose.

Bonds	**k^b, kJ/mol/nm^2**	**r_0, nm**	**Graphical representation**
CI-1	15000	1.745	

CI-2	21000	1.227	
CO	22000	1.268	
CC	500000	2.068	
Angles	**k^{Θ}, kJ/mol/rad^2**	**Θ_0, deg**	**Graphical representation**
ICI	65000	41.320	
OCO	3000	177.999	
ICO-1	1000	83.170	
ICO-2	1000	55.550	

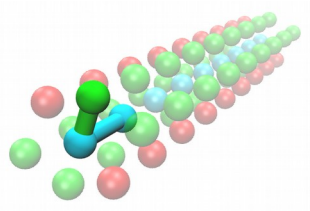
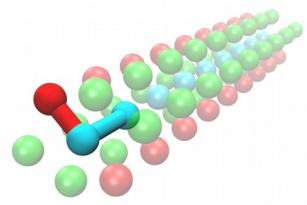
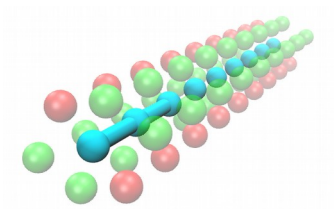
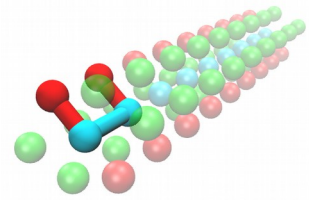
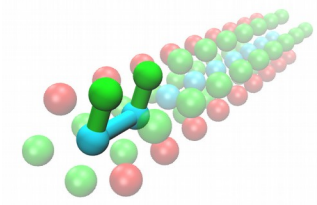
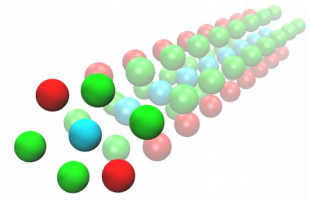
ICC	3000	86.913	
OCC	2500	86.184	
CCC	300000	179.500	
Dihedrals	**k_ϕ, kJ/mol**	**ϕ_s, deg**	**Graphical representation**
OCCO	1200	3.193	
ICCI	1200	3.209	
Improper	**k^ξ, kJ/mol/rad^2**	**ξ_0, deg**	**Graphical representation**
	3000	2.750	

Table S2. Non-bonded interaction parameters of the supra coarse-grained model for cellulose.

	σ, nm	ε, kJ/mol
IC-IC	1.95	25.00
OC-OC	1.35	75.00
IC-OC	1.70	1.00

Mean-square displacement and diffusion coefficients.

The mean-square displacement and the diffusion coefficients for different CNC lengths are presented in Figure S1 and Table S3, respectively.

Table S3. Simulated self-diffusion coefficients for different CNC lengths.

Length, nm	D_t (×10^{-12}) m^2/s
50	16.40
100	7.02
200	4.73
500	1.67
700	1.28
1000	0.99
1200	0.80

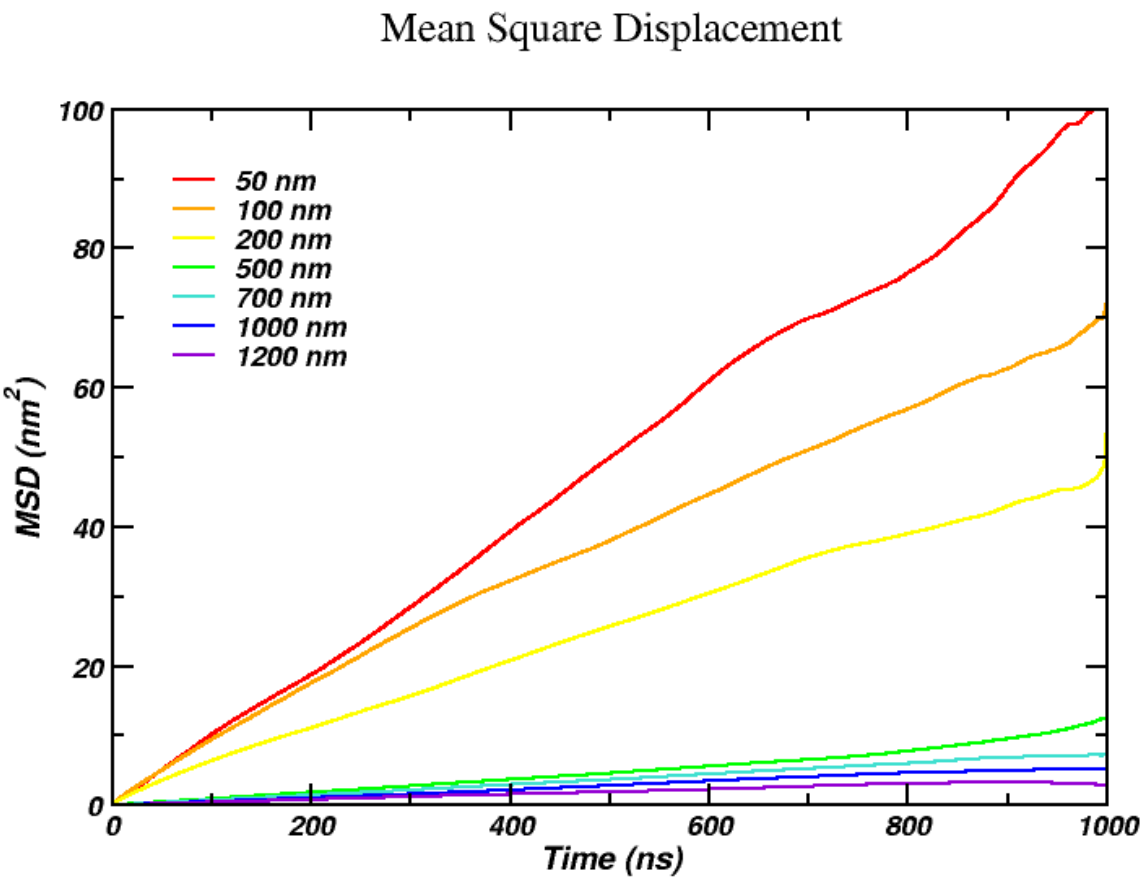


Figure S1. Mean-square displacement of CNCs with different lengths varied between 50 and 1200 nm.

All-atom molecular dynamics simulation details

The all-atom molecular dynamic (MD) simulations needed for the parameterization of the sCG model were performed with the OPLS-AA force field for carbohydrates. [2] Partial charges of the atoms were assigned according to the force field with no further modifications. The fibril cross-section constitutes 36 chains arranged in the cellulose Iβ allomorph and the length of the fibril is 32 glucose units. The fibril is solvated in 68854 water molecules for which we have used the SPC/E

water model compatible with OPLS-AA. The hydrogen bonds were restricted with the LINCS algorithm. The time step in the simulations was set to 2 fs and the equations of motion were integrated with the leap frog method. The AA-MD simulations were performed in the NPT ensemble with the Berendsen barostat and the velocity rescaling thermostat with the relaxation time of 1 and 0.1 ps, respectively. The simulations were first minimized with the steepest descent algorithm and then run for 60 ns, where the last 20 ns were used in the parameterization of the sCG model. The trajectory was saved every 4 ps. The Verlet scheme for the neighbor list search was used with update time of 10 MD steps. The non-bonded interactions cut-off was set to 1.2 nm. The Particle-Mesh Ewald (PME) method was used to calculate the electrostatic interactions.

[1] T. Nishino, K. Takano, K. Nakamae Elastic modulus of the crystalline regions of cellulose polymorphs Journal of Polymer Science: Part B: Polymer Physic, 1995, 33, 1649-1651

[2] W. Damm, A. Frontera, J. Tirado-Rives, W. L. Jorgensen OPLS All-Atom Force Field for Carbohydrates J. Comp. Chem., 1997, 18, 16, 1955-1970