Computational Study of the Dissolution of Cellulose into Single Chains: the Role of the Solvent and Agitation

Eivind Bering (eivind.bering@gmail.com)
Norwegian University of Science and Technology: Norges teknisk-naturvitenskapelige universitet
https://orcid.org/0000-0003-0263-2469
Jonathan Torstensen
Western Norway University of Applied Sciences - Stord Campus: Hogskulen pa Vestlandet - Campus Stord
Anders Lervik
Norwegian University of Science and Technology: Norges teknisk-naturvitenskapelige universitet
Astrid S. de Wijn
Norwegian University of Science and Technology: Norges teknisk-naturvitenskapelige universitet

Research Article

Keywords: Pretreatment, MD simulation, Urea, Alkali metals, Tensile strength, Dissolution

Posted Date: August 5th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-679922/v1

Version of Record: A version of this preprint was published at Cellulose on January 6th, 2022. See the published version at https://doi.org/10.1007/s10570-021-04382-9.

Cellulose manuscript No.
(will be inserted by the editor)

Computational study of the dissolution of cellulose into single chains: the role of the solvent and agitation

Eivind Bering · Jonathan Torstensen · Anders Lervik · Astrid S. de Wijn

Received: date / Accepted: date

Abstract We investigate the dissolution mechanism of cellulose using molecular dynamics simulations in both water and a mixture solvent consisting of water with Na^+, OH^- and urea. As a first computational study of its kind, we apply periodic external forces that mimic agitation of the suspension. Without the agitation, the bundles do not dissolve, neither in water nor solvent. In the solvent mixture the bundle swells up with significant amounts of urea entering the bundle, as well as more water than in the bundles subjected to pure water. We also find that the mixture solution stabilizes cellulose sheets, while in water these immediately collapse into bundles. Under agitation the bundles dissolve more easily in the solvent mixture than in water, where sheets of cellulose remain that are bound together through hydrophobic interactions. Our findings highlight the importance of urea in the solvent, as well as the hydrophobic interactions, and are consistent with experimental results.

Keywords Pretreatment · MD simulation · Urea · Alkali metals · Tensile strength · Dissolution

E. Bering (corresponding author)
PoreLab, Department of Physics,
Norwegian University of Science and Technology (NTNU),
NO–7491 Trondheim, Norway,
orcid: 0000-0003-0263-2469,
E-mail: eivind.bering@gmail.com;

J. Torstensen
Department of Mechanical and Marine Engineering,
Western Norway University of Applied Sciences (HVL),
NO-5063 Bergen, Norway,
orcid: 0000-0002-8632-8028,
E-mail: jtor@hvl.no

A. Lervik
Department of Chemistry,
Norwegian University of Science and Technology (NTNU),
NO-7491 Trondheim, Norway,
orcid: 0000-0002-1505-6731,
E-mail: anders.lervik@ntnu.no

A. S. de Wijn
Department of Mechanical and Industrial Engineering,
Norwegian University of Science and Technology (NTNU),
NO–7491 Trondheim, Norway,
orcid: 0000-0003-4664-6811,
E-mail: astrid.dewijn@ntnu.no

1 Introduction

Cellulose is becoming more and more prominent as a component used in development of new materials, cosmetics, as a food additive, biomedical applications and in fabrics. It is typically cited as being the worlds most abundant polymer [1], and for this reason it is an interesting candidate for packaging and clothing applications in particular. In this context, several efforts have been made to dissolve cellulose into single chains in order to either form more elastic films (for packaging applications) or to be able to weave fabrics with specific properties, such as in the lyocell process [2]. Cellulose is a glucose $\beta(1\rightarrow4)$ polymer, where each glucose unit is rotated 180° along the screw axis as illustrated in Fig. 1. Due to this, it is debated whether glucose or cellobiose is indeed the repeating unit in cellulose [3]. Regardless, cellulose is naturally produced as chains that immediately form sheets, then bundles [1]. As of yet, single chained cellulose has not been found in nature, and cellulose seemingly exists predominantly as 6 by 6 chain bundles [1,4], or possibly as conformations of 18 or 24 chains [5,6,7]. Thus, in order to better utilise cellulose for packaging and fabric applications and more, it is desirable to be able to dissolve it into smaller units, especially single chains. A considerable amount of effort has gone into research on dissolving cellulose into single chains [8,9].

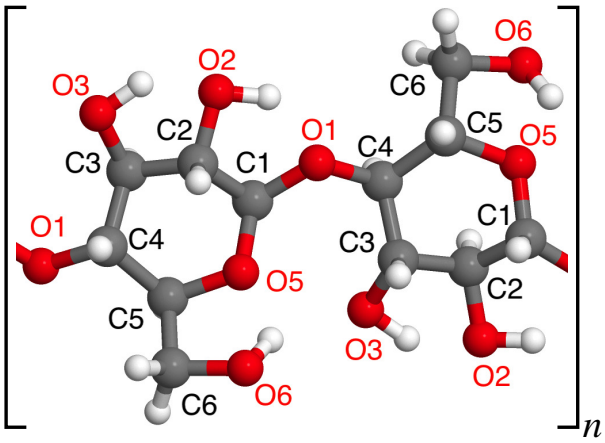


Fig. 1: The structure of cellulose. Carbon atoms are grey, oxygen atoms are red, and hydrogen atoms are white. The numbering of the atoms inside the glucose units is indicated. The O1 of one unit is the O4 of the next, and we label it by O1. For clarity, we designate cellobiose as the repeating unit of cellulose in order to capture the screw axis 180 degree rotation. However, glucose is most likely the cellulose monomer.

1.1 Overview of experimental dissolution of cellulose by alkali solvents

Cellulose is an amphiphilic molecule [9], and its structure is depicted in Fig. 1. Hydrogen bonds connect chains in-plane, while van der Waals bonds act perpendicular to the plane of the molecules. The manifestations and interplay between these interactions and the structure is still not well-understood. Direct evidence of hydrophilicity may be observed through contact angle measurements, which are always $< 90°$, indicating a hydrophilic structure. However, by modifying the orientation of cellulose crystals of a cellulose surface, it is also possible to tune the contact angle between $12°$ and $\sim 37°$ [10]. This is a clear indication of the innate hydrophobicity of cellulose, which has also been demonstrated by the difficulty of removing hydrophobic compounds from cellulose, even after extensive vacuum treatment [11]. When cellulose dissolves, the cellulose chains form a single-chain suspension as opposed to a suspension of bundles.

Cellulose dissolution is a complex process, but, regardless of the chosen method, it appears to always be composed of two steps:

- disruption of interchain bonds
- stabilization of dissolved chains

A complicating issue with disrupting cellulose bonds is that these are both hydrophilic *and* hydrophobic. This has been highlighted by several works, among others by Lindman et al. [8], stating that hydrophobic cellulose interactions are "typically" overlooked as the stronger hydrophilic hydrogen bonds are assumed to be more important. However, experimental results indicate that both these bonds need to be broken and stabilized. Initially, cellulose solvents were derivatizing, meaning that the cellulose chains were chemically modified to achieve dissolution. However, derivatizing sovents do not yield a suspension of native cellulose. One derivatizing method is carboxymethylation, yielding carboxymethyl cellulose, where all or some of the hydroxyl groups are replaced by carboxymethyl groups. These groups loosen chains from the cellulose bundle due to steric effects, and increased osmotic pressure, ensuing dissolution. Carboxymethyl cellulose has several uses within the fabric, food, medical and electronic applications [12]. Non-derivatizing cellulose solvents range from water based to non-aqueous systems as well as ionic liquid systems [9].

Based on the established amphiphilic nature of cellulose, one would expect that a mixed polar/nonpolar solvent is an absolute requirement for dissolving cellulose. However, this is seemingly not the case. At short chain lengths (< 200 anhydrous glucose units), cellulose is readily dissolved in water and sodium hydroxide [18]. Moreover, cellulose may be dissolved in several molten salts [19], but not in the corresponding aqueous salt solutions. This is exemplified by $LiClO_4 \times 3H_2O$ which is an excellent solvent, while $LiClO_4$ with more added water does not dissolve cellulose [19]. In the latter case, water out-competes cellulose in forming hydrogen bonds with the ion, and interchain hydrogen bonds are not disrupted. In several ionic liquids, the ionic nature of the solvent is also considered the reason for cellulose dissolution. It is not yet fully understood if the cation or anion has the leading role in such solvents [9].

Water based alkaline solvent systems are summarized in Table 1. In these solvents, the ions are believed to disrupt cellulose inter-chain hydrogen bonds, by direct binding and also by increasing the osmotic pressure as they penetrate the cellulose structure. It has even been suggested that partial deprotonation of cellulose hydroxyls (pKa = 13.3) may occur, but this effect is assumed to be minor. However, this is a good reason to apply hydroxide salts in the dissolution process as they may deprotonate. The main reason for including hydroxide is that it facilitates water intrusion into the bundle [9]. The role of the ion is thus assumed to be disruption of hydrogen bonds. Larger ions such as K^+ are not as efficient as smaller ions (Li^+), which is typically attributed to a larger hydrodynamic radius not allowing hydrated ions to penetrate the cellulose bundle. The degree of solubility of seems to follow the Hofmeister series [20]. One should note that the mechanism behind this series for protein solvation is still not completely understood.

As pointed out by Medronho and Lindman [9], highly asymmetrical electrolytes, with one large and one small ion, appears to provide the best solubility for non-ionic polymers

Table 1: Overview of non-derivatizing, alkaline, water-based cellulose solvents

Solvent	Hydrophilic component	Ampiphilic component	Temperature	Mechanism
Water / sodium hydroxide [13]	H_2O, Na^+, OH^-	–	4°C	Disruption of hydrogen bonds. Dissolution only of chains with fewer than 200 glucose units.
Water / lithium hydroxide [14]	H_2O, Li^+, OH^-	–	4°C	Disruption of hydrogen bonds.
Water / alkali hydroxide / urea [15]	H_2O, Li^+, Na^+, K^+, OH^-	urea	Optimal at ~ -10°C	Ions penetrate the structure and destabilizes hydrogen bonds. Urea shields cellulose chains from agglomerating. LiOH is more effective than NaOH, which is more effective than KOH.
Water / sodium hydroxide / thiourea / urea [16]	H_2O, Na^+, OH^-	urea / thiourea	Optimal at ~ -8 to -12°C	Cations penetrate the structure and destabilizes hydrogen bonds. Thiourea/Urea shields cellulose hydrophobic interactions. Urea/Thiourea was found to be better compared to only urea or only thiourea.
Water / sodium hydroxide / PEG [17]	H_2O, Na^+, OH^-	PEG	-15°C	Cations penetrate the structure and destabilizes hydrogen bonds. PEG forms hydrogen bonds with cellulose.

such as cellulose. The effects on hydrodynamic radii is also the reason why there are typically stringent concentration requirement for the ionic solvent components. Partial dissolution and re-agglomeration are known to happen in water and alkali systems.

One way to improve the dissolution is to use mixed solvents. Examples are solvents with thiourea/urea as the amphiphilic component added to a water-alkali-hydroxide mixture [15,16]. In these cases, the amphiphilic urea/thioura are assumed to shield cellulose hydrophobic interactions, hindering re-agglomeration after dissolution. For instance, Xiong et al. [21], found evidence of hydrophobic interactions between cellulose and urea. Accumulation of urea at hydrophobic sites of cellulose has also been suggested by others [22]. The alternative inclusion of polyethylene glycol (PEG) instead of thiourea/urea, shows that the amphiphilicity is important in the added solvent component. This has been discussed extensively for PEG in terms of protein dissolution [23]. However, the role of amphiphilic solvent components is still debated, as they are not strictly required to dissolve cellulose [9]. For PEG, hydrogen bonding with cellulose and association with the cellulose chain (possibly due to hydrophobic interactions) have been identified as the functional roles [17].

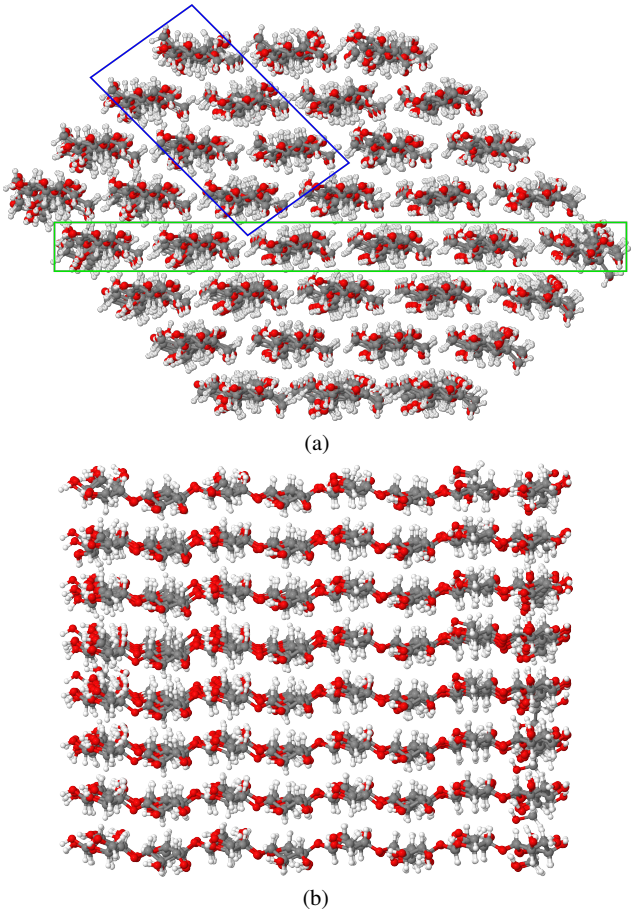
Molecular Dynamics (MD) simulations are applied to verify experimental results and assumption of cellulose dissolution mechanisms. MD simulations have inherent inaccuracies and a vast amount of tunable parameters (cellulose structure, the choice of force field, atom/molecule types, boundary conditions, etc.) which play a role in how these numerical experiments should be interpreted. MD simulations can, on the other hand, provide useful information that is not accessible in experiments about the location and dynamics of the molecules, ions, and atoms. Interactions between cellulose and urea has been studied in order to confirm its role as a cellulose stabilizer [24,21,25,26,27], but the literature seems to diverge on the subject. Wernersson et al. [24] support that hydrophobic cellulose-urea interactions are predominant, supporting observations made by Xiong et al. [21]. Bergenstråle-Wholert et al. argue that urea is also amphiphilic, and compete with water for direct hydrogen bonding with cellulose [25]. Liu et al. [26] described that a reduction in urea concentration at oxygen atoms O1 and O5 (see Fig. 1) was necessary to achieve re-agglomeration. Studies also indicated that ions (i.e. Na^+) most preferably located to the cellulose hydroxyl groups. However, Cai et al [27] suggested that urea bind cellulose hydroxyls through hydrogen bonding. They found a decrease in cellulose-urea bonding from 265 to 283 K (-8 to 10°C), in agreement with temperature requirements for dissolution of cellulose. More pronounced cellulose-urea interactions at lower temperatures, was also found by Wernersson et al. [24].

From a chemical and thermodynamic point, is it reasonable that urea/thiourea binds to hydroxyls and non-polar cellulose groups, even if the presence of both these interactions are not always found in all simulation/experimental studies. An overview of the constituents of the alkaline/hydroxide method and their function is given in Table 2.

Regardless of the exact solvent system composition, low temperature favors dissolution in water based alkaline systems. This is typically attributed to temperature-dependent changes in the cellulose structure. Naively, one would expect that all dissolution processes were favored by increasing temperatures. However, cellulose bundles (and some other polymeric systems) are known to become more hydrophobic with higher temperatures [8]. While the number of in-

trachain hydrogen bonds is dropping with increasing temperature, new interchain and intersheet hydrogen bonds are formed, forming a three-dimensional hydrogen bonding network that stabilize the cellulose structure at high temperatures (425–550 K) [28]. Considering that the hydroxide/urea/water system is based on ion and water intrusion into the cellulose bundle, increased hydrophilicity at lower temperatures favours dissolution.

Imperative to all dissolution methods is vigorous stirring. In this context, stirring supplies drag forces that act opposite of the cellulosic material direction of motion. These forces are believed to increase material swelling, and their nature is described amongst others by Tulus et al. [29]. The type of applied stirring is not believed to decrease polymer chain lengths to a large extent.



(a)

(b)

Fig. 2: Bundle with 36 chains after energy minimization, seen from the side (a) and from the front (b). Smaller bundles are investigated in more detail, with shapes illustrated by the drawn lines: a bundle with 6 chains (blue), and a sheet with 6 chains (green).

2 Methodology

In this paper, we have modeled 6 and 36 chain systems with different geometries, with a basis in a maze configuration [30]. The modelling aims at elucidating dissolution thermodynamics. As such, the most important feature of the models, is that the chains are aligned with in-plane hydrogen bonds and out-of-plane van-der-Waals interactions. Even though the exact bundle geometry are debated for different species [7], this feature is constant.

We simulate bundles of cellulose chains surrounded by either pure water, or a mixture solvent consisting of water, Na^+, OH^-, and urea. We analyze the dissolution by considering the density of the cellulose bundle in the different solvents, as well as the solvent accessible surface area. Besides investigating the effect of the solvent, we also include on agitation by external forces. Below we describe step by step how our system is set up, how the forces are added, and how we analyze the results.

2.1 Simulation setup

The cellulose molecules in our model were made with the GLYCAM carbohydrate builder [31], producing chains with 4 cellobiose units (8 glucose molecules) with glycosidic angles $\phi = 22.6°$ and $\psi = -26.2°$ (from Cellulose-builder [32, 33]). A bundle of 36 chains of 4 cellobiose units each was constructed in a maze configuration [30], with a intraplanar chain separation of 0.82 nm and an interplanar chain separation of 0.389 nm. The tool Intermol (v. 0.1.0) [34] was used to convert the input files to a format compatible with LAMMPS [35], which was used for all molecular dynamics simulations (v. March 2018). Input configurations with solvents was generated with Packmol (v. 20.010) [36], with a weight percentage of about 7% NaOH and 12% urea in the mixture solvent to match experiments [15].

Cellulose was modeled with the GLYCAM06 force field [37], water with the SPC/E model [38], the LJ parameters for the alkali metals were set for SPC/E [39], and the standard GLYCAM06 hydroxyl LJ parameters were used for the hydroxide. Arithmetic combination rules was used for all remaining LJ parameters. The alkali metals have charge $+1e$, in the hydroxide the oxygen had charge $-1.02e$ and hydrogen had a charge of $+0.02e$. The urea molecules were modeled with the improved generalized Amber force field [40]. Nonbonded interactions between the cellulose were modeled by a switching function with an inner cutoff of 7 Å and an outer cutoff of 9 Å. For water and the other solvents, a simple cutoff of 9 Å was used. Long-range non-bonded interactions were obtained by the particle-particle particle-mesh (PPPM) solver with a relative error in forces of 1.0×10^{-4}.

The simulations started with energy minimization of the cellulose bundle without solvent by a steepest descent algo-

Table 2: Overview of non-derivatizing, alkaline, water-based cellulose solvents

Solvent constituent	Constituent function	Method	Reference
Urea	Interact directly with cellulose through hydrogen bonding	MD, NMR	[25]
Urea	Accumulate at cellulose hydrophobic facets, aid alkali intrusion. No direct cellulose-urea interactions	XRD, DSC	[22]
Urea	Interact directly with cellulose through hydrogen bonding	MD	[27]
Urea	Predominantly solubilize hydrophobic portions of cellulose	MD	[24]
Hydroxyl-ion	Possible deprotonation of cellulose hydroxyls	(review)	[9]
Alkali cations	Cation hydrates bind to cellulose	XRD, DSC	[22]
Water	Hydrogen bonding with cellulose	MD, NMR	[25]

Table 3: Simulation overview. The size denotes the number of chains in the system. Each chain is composed of 4 cellobiose units.

Size	Shape	Environment	Water [#]	NaOH [#]	Urea [#]	Box dimensions [nm]
36	bundle	water, NPT	31822			$10 \times 10 \times 10$
36	bundle	mixture, NPT	27352	1064	1215	$9.8 \times 9.8 \times 9.8$
6	sheet	water, NPT	14438			$7.6 \times 7.6 \times 7.6$
6	sheet	mixture, NPT	13437	524	600	$7.6 \times 7.6 \times 7.6$
6	sheet	mixture, NVT	13437	524	600	$7.6 \times 7.6 \times 7.6$
6	bundle	water, NPT	7348			$8.4 \times 5.2 \times 5.2$
6	bundle	mixture, NPT	6213	242	276	$8.4 \times 5.2 \times 5.2$

rithm with about 5000 iterations, followed by 1 ns of molecular dynamics at temperature $T = 1$ K with a Langevin thermostat [41]. Solvent was then added to the systems with Packmol [36]. The simulations were then run at $T = 1$ K with constant volume for 10 ps, before switching to a Nosé-Hoover style isothermal-isobaric simulations and linearly increasing the temperature to $T = 258$ K. We use a time step of 0.1 fs during the initialisation, and 1 fs during the rest of the simulations. These NPT simulations were initially run with a relaxation time of 10 ps for the barostat and a relaxation time of 50 fs for the thermostat while increasing the temperature to the target temperature over 0.1 ns. In the continued simulations the relaxation time for the thermostat was set to 0.1 ps and the relaxation time for the barostat was set to 1 ps. Table 3 presents an overview of the systems under investigation with the number of solvent molecules and the size of the simulation box.

2.2 Agitation

We have added periodic agitation of the sample to our MD simulations. In the experiments, agitation of the system induces shear stresses on the bundle, most likely through stresses from the moving solvent. A direct implementation of this share rate in the solvent is complicated, and it is not considered to be necessary. Due to their geometry, the largest forces on the bundles will be in the direction of the long axis of the bundle.

We therefore implement agitation of our simulated bundles by periodically stretching and compressing them. This is implemented in two different ways, with length-controlled and force-controlled harmonic oscillations. These are illustrated in Fig. 3. In simulations with length-controlled oscillations, we control the position of the outermost carbon atom in each chain along the long axis, and the chains are free to move in the axial plane. In the simulations with force-controlled oscillations, a force is added to the outermost carbon atom in each chain along the longest axis, and the chains are otherwise free to move. The frequency of the oscillatory stretching and compression is 2 ns^{-1} for both modes of control. The length-controlled oscillations have an amplitude of 0.8 nm, and the force-controlled oscillations adds a force

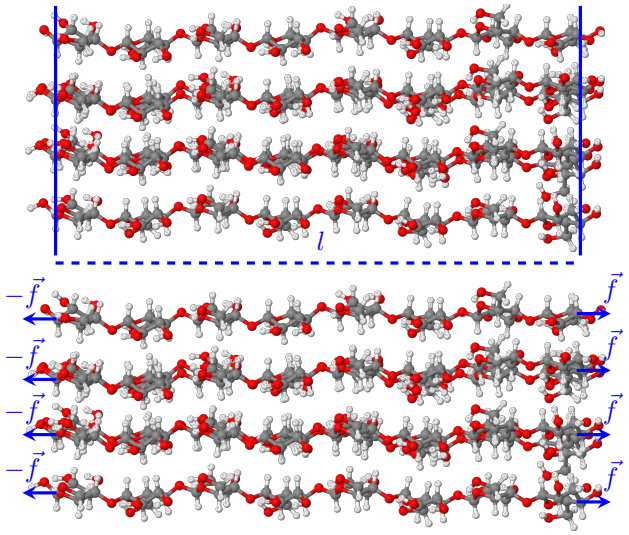


Fig. 3: Illustration of the length-controlled (left) and force-controlled (right) simulation mode. With length controlled oscillations, the positions of the outermost carbon atoms in all the chains are controlled along the longest axis, and the chains are free to move in the axial plane. In the force-controlled case, a force is added to the outermost carbon atoms in each chain, and the chains are otherwise free to move.

with amplitude 0.35 nN per chain. All chains remain intact throughout the simulations.

2.3 Modeling data analysis

The volume of the fibril was estimated by the volume of convex hull of the atoms (treated as point particles) in the fibril, obtained by `scipy.spatial.ConvexHull` as implemented in SciPy [42]. These convex hulls were also used to determine the absorbed molecules: for each solvent molecule an additional convex hull was created with the center of mass of the solvent molecule as an additional node. The solvent molecule is then defined to be absorbed if the addition of the solvent molecule did not change the convex hull. Note that the absorbed molecules are not included in the mass of the bundle for the densities presented later. The solvent accessible surface area (SAS) was estimated with MDTraj [43], which uses the Shrake and Rupley algorithm [44].

The number of clusters was found by calculating the average distances between identical atoms in neighbouring chains, if this distance is less than 5 Å they are considered to be in the same cluster.

3 Results

We first explore the system without agitation, shown in Fig. 2 after energy minimization. This system will serve as our reference. To gain further insight into the role of the solvents, we have also considered smaller bundles in more detail since this allows us to reach longer time scales. The configurations of these smaller bundles are illustrated in Fig. 2 with the (blue and green) outlines. We then show results with agitation. Finally, we investigate the effects of the length by repeating some simulations with bundles of 6 chains with 9 cellobiose units rather than 4.

3.1 Bundles with 36 chains

36 chains of 4 cellobiose units each were placed in a maze configuration, which is shown in Fig. 2 after energy minimization. The bundle was then solvated in water or in the mixture and allowed to equilibrate for 10 ns as described above. In the time span of our simulations (40–45 ns), the bundles do not dissolve in water nor in the mixture solvent. The density of the cellulose in water remains stable around 1.5 g/cm^3, while the density in the mixture solvent stabilizes between 1.3 and 1.35 g/cm^3 (see Table 4).

To better understand what role the different molecules in the solvent play, we investigate how much of the different species is absorbed by the bundle. This is shown as a function of time in Fig. 5 for the mixture solvent. We see that urea is absorbed the most, with a concentration in the bundle that exceeds that of the bulk. Some water molecules are absorbed as well, and the concentration in the bundle is approximately half that of the bulk. The fraction of absorbed water in the bundle in the water-only solvent was significantly lower than that in the mixture solvent. The Na^+ and OH^- ions are not absorbed into the bundle. The fact that smaller ions penetrate less deeply into the bundle could be considered counter-intuitive. Most likely this is related to the geometry of the molecules and placement of charges.

Bundles of cellulose are known to be very stable without agitation, a picture we have largely confirmed in our simulations described above. Nevertheless, we can already see that cellulose behaves differently in the mixture solvent than in pure water. The lower density of the cellulose in the mixture solvent is almost certainly due to the high concentration of urea in the bundle, which increases the separation between the cellulose molecules. In these bundles without any agitation, the lower density of cellulose and high penetration of urea into the bundle do not yet lead to dissolution. Nevertheless, this may provide an indication as to why cellulose is more easily dissolved in the mixture solvent, and gives insight on the role of urea in dissolution.

Table 4: Density (ρ) of cellulose and Solvent Accessible Surface Area (SAS) relative to the SAS of the same number of fully dissolved chains. NPT= constant pressure and temperature, NVT= constant volume and temperature.

Size	Shape	Environment	ρ [g/cm^3]	Average SAS	Smallest SAS	Largest SAS
36	bundle	water, NPT	1.51 ± 0.02	0.182 ± 0.001	0.046 ± 0.002	0.44 ± 0.01
36	bundle	mixture, NPT	1.33 ± 0.02	0.192 ± 0.002	0.042 ± 0.004	0.46 ± 0.01
6	sheet	water, NPT	0.9 ± 0.1	0.54 ± 0.01	0.28 ± 0.02	0.71 ± 0.01
6	sheet	mixture, NPT	0.86 ± 0.07	0.56 ± 0.01	0.45 ± 0.01	0.73 ± 0.03
6	sheet	mixture, NVT	0.89 ± 0.03	0.58 ± 0.01	0.45 ± 0.01	0.72 ± 0.01
6	bundle	water, NPT	1.31 ± 0.06	0.427 ± 0.008	0.26 ± 0.02	0.59 ± 0.02
6	bundle	mixture, NPT	1.47 ± 0.03	0.420 ± 0.004	0.24 ± 0.01	0.56 ± 0.01

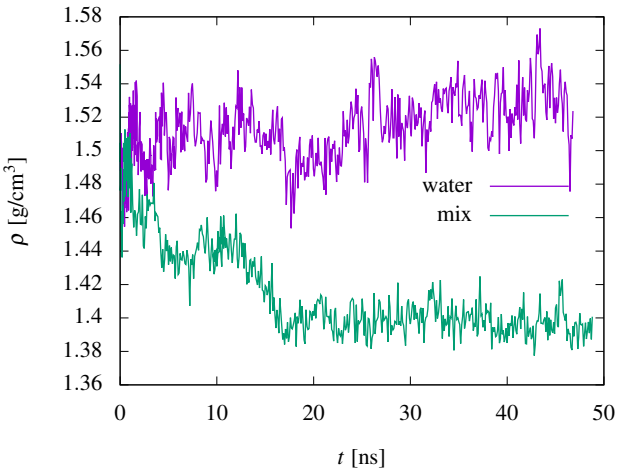


Fig. 4: Density of cellulose as a function of time for a bundle of 36 chains at $T = 258$ K, in pure water and the mixture solvent.

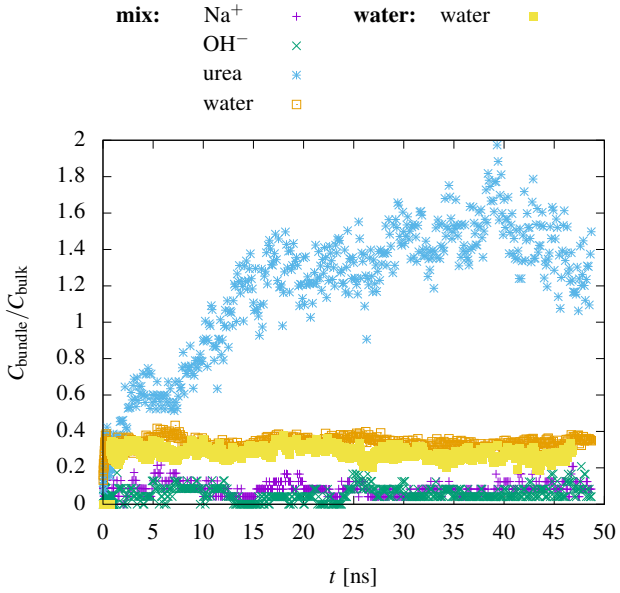


Fig. 5: Concentration ratio of solvents absorbed in a bundle with 36 chains of 4 cellobiose units per chain at $T = 258$ K, compared to the bulk.

3.2 Smaller systems: 6 chains

In order to get a better understanding of the system on a longer time-scale, we study bundles and sheets with fewer chains, as these are far less computationally expensive and allow for more detailed investigations. In contrast to the bundles, sheets of cellulose are inherently unstable, and are not found in nature. The rationale for including sheets are that they may represent an intermediate stage of dissolution. Moreover, sheet dissolution has been suggested to occur spontaneously on timescales accessible by MD simulations [26].

We study bundles and sheets composed of 6 chains of cellulose, with 4 cellobiose units per chain. The initial configurations of these systems are illustrated in Fig. 2 with different colour outlines. We start simulations from the initial configurations and equilibrate in water or the mixture solvent. The density and SAS after equilibration are shown in Table 4.

The density of these small bundles in both mixture and water is similar to that of 36-chain bundles. Sheets, however, have lower density. The sheets are not perfectly flat, and probably less stiff than bundles. Consequently, when they bend, there is more extra volume in the convex hull than there is for bundles.

To investigate the surfaces of the bundles and sheets in more detail, we calculate the SAS compared to that of a set of fully dissolved chains. This is included in table Table 4 and also shown in Fig. 6 for water and Fig. 7 for the mixture solvent. The SAS of the chain with the highest and lowest areas for each system is shown with dashed lines. The systems with sheets as initial configuration have the highest SAS, both for the averages and the lowest SAS. This is to be expected, since the surface of a sheet is larger than that of a

bundle. Note that the system with a sheet in water has at least one chain with a very low SAS, on par with that of bundles. This indicates that the sheet is forming a bundle. We can see this in Fig. 6, where the lowest SAS clearly decreases with time from around the typical average value in a sheet to the typical lowest level for a bundle. This does not happen in the mixture solvent, shown in Fig. 7. Snapshots of the initial sheet and resulting bundle in water are also included. We have found that in simulations without any solvent, i.e. in NVT in vacuum, these sheets also rapidly reconfigure to bundle-like configurations.

For the SAS of the sheet in mixture solvent, the average value is slightly higher than in water, but in contrast to the same system in water, none of the chains have a very low SAS in this case. In Fig. 6, it can also be seen that the smallest SAS does not drop down to the level of a bundle in the mixture solvent. This indicates that the sheet remains stable as a sheet in the mixture solvent, rather than collapsing into bundle. This increased SAS and more stable open geometry could possibly aid dissolution in the mixture solvent.

It has previously been reported that similar sheets of 6 chains actually dissolve in MD simulations in a similar mixture solvent at 261 K on a timescale of ~ 20 ns [26]. We have therefore performed simulations with the same configuration of the sheets with the same number of solvent molecules as those authors. The simulations were repeated twice in different ensembles, two times NPT and one NVT, as Liu et al. [26]. However in our simulations the sheets always remained intact. The main difference between our simulations and those of Liu et al., that we suspect may account for this different behaviour, is the choice of force field. Liu et al. used the general purpose CHARMM36 force field, while we use the more specialized GLYCAM06. It is quite likely that the dissolution mechanism is very sensitive to the force field parameters and especially the partial charges on the different atoms in the cellulose, since crystalline structure is also sensitive to it [45]. We illustrate this by computing the radial distribution function (RDF) between the various solvent molecules and two specific oxygen atoms in the cellulose, O1 and O5. This is shown in Fig. 8. While there are peaks, some of them in similar locations, our RDF look qualitatively different from those in [26], shown there in figure 6.

3.3 Agitation

The mixture solvent at 258 K is known experimentally to dissolve the cellulose bundles, but in the lab it is necessary to agitate the samples for dissolution to occur [15]. We therefore implement agitation of our simulated bundles by periodically stretching them. We have tested two types of agitation, length- and force-controlled oscillations, in both water and mixture.

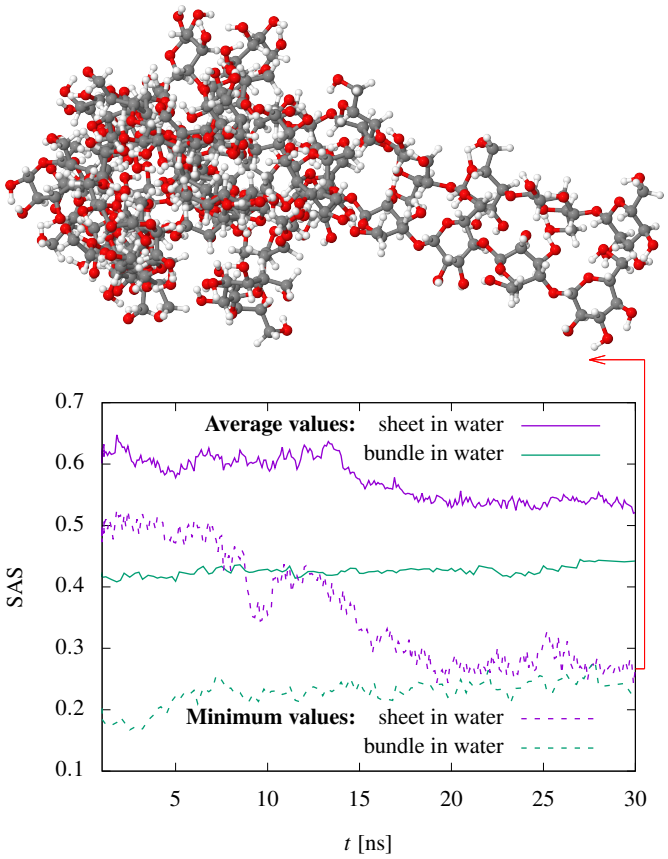


Fig. 6: SAS per chain of cellulose, 6 chains in sheet and bundle at $T = 258$ K, water. The color indicates the system, as specified in the legend. The solid lines are averages, and the dashed lines show the SAS of the chain with the lowest value for each system. The configuration after 30 ns is shown above, confirming that the sheet is turning in to a bundle.

The cross sections of bundles with 36 chains after 30 ns of length controlled oscillations and 10 ns of force controlled oscillations in the mixture solvent and water can be seen in Fig. 9. These times correspond approximately to the earliest times where the cellulose density in both the mixture and water solvents were ≤ 0.5 g/cm^3. In all cases, we see signs of dissolution. However, there are clear differences. Dissolution is slower and more incomplete in water.

We see a significant dependence on whether the simulations were force-controlled or length-controlled. The snapshots of the force-controlled simulations show the most disorder. In the water solvent, there still remains a crystalline-like core, while in the mixture solvent, the bundle has fully dissolved. In water, the chains are dissolving from the outside-in, one by one, and the core has not yet fully dissolved. This indicates that the dissolution in water is slower than in the mixture solvent.

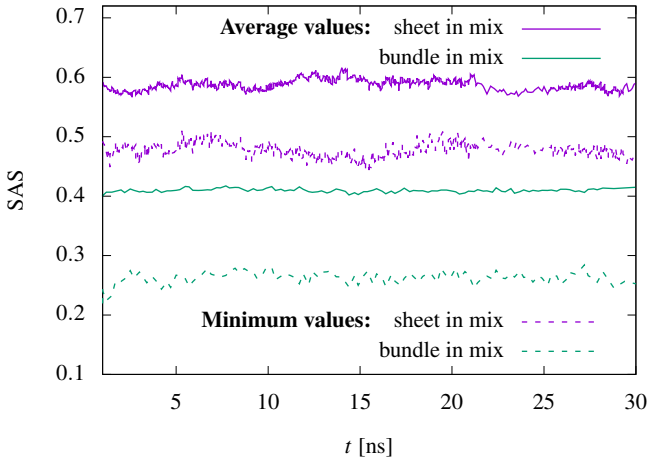
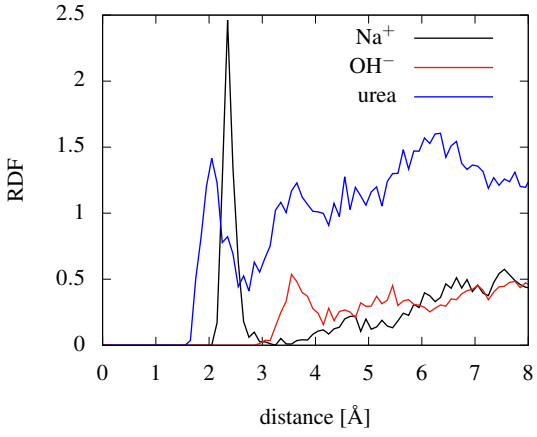
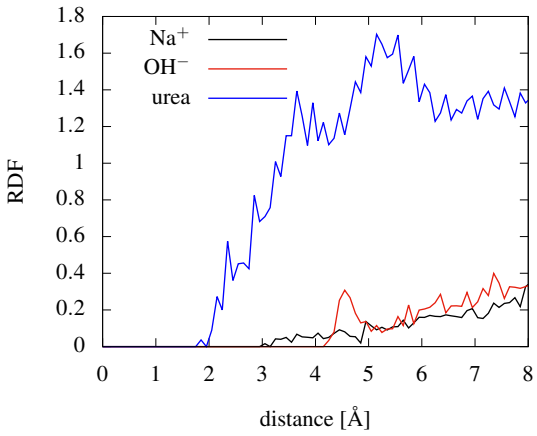


Fig. 7: SAS per chain of cellulose, 6 chains in sheet and bundle at $T = 258$ K, mix. The color indicates the system, as specified in the legend. The solid lines are averages, and the dashed lines show the SAS of the chain with the lowest value for each system.



(a) RDF for O1



(b) RDF for O5

Fig. 8: RDF for the oxygen atoms O1 and O5 for a 6-chain sheet in the mixture solvent, to be compared with [26], Figure 6. The location of these atoms in the glucose units is indicated in Fig. 1.

The length-controlled oscillation produces sheet-like structures in both water and the mixture solvent, but in water this is much more pronounced and the sheets are larger and more ordered. The geometry of the sheets that are formed in the length-controlled simulations is shown in Fig. 10. They are held together by hydrophobic interactions, rather than hydrophilic interactions. The fact that these structures are more present in water suggests that some of the components in the mixture play a role in breaking these hydrophobic interactions. Indeed, it has been suggested [9, 8] that poor breaking down of hydrophobic interactions plays a role in why water is not able to dissolve cellulose.

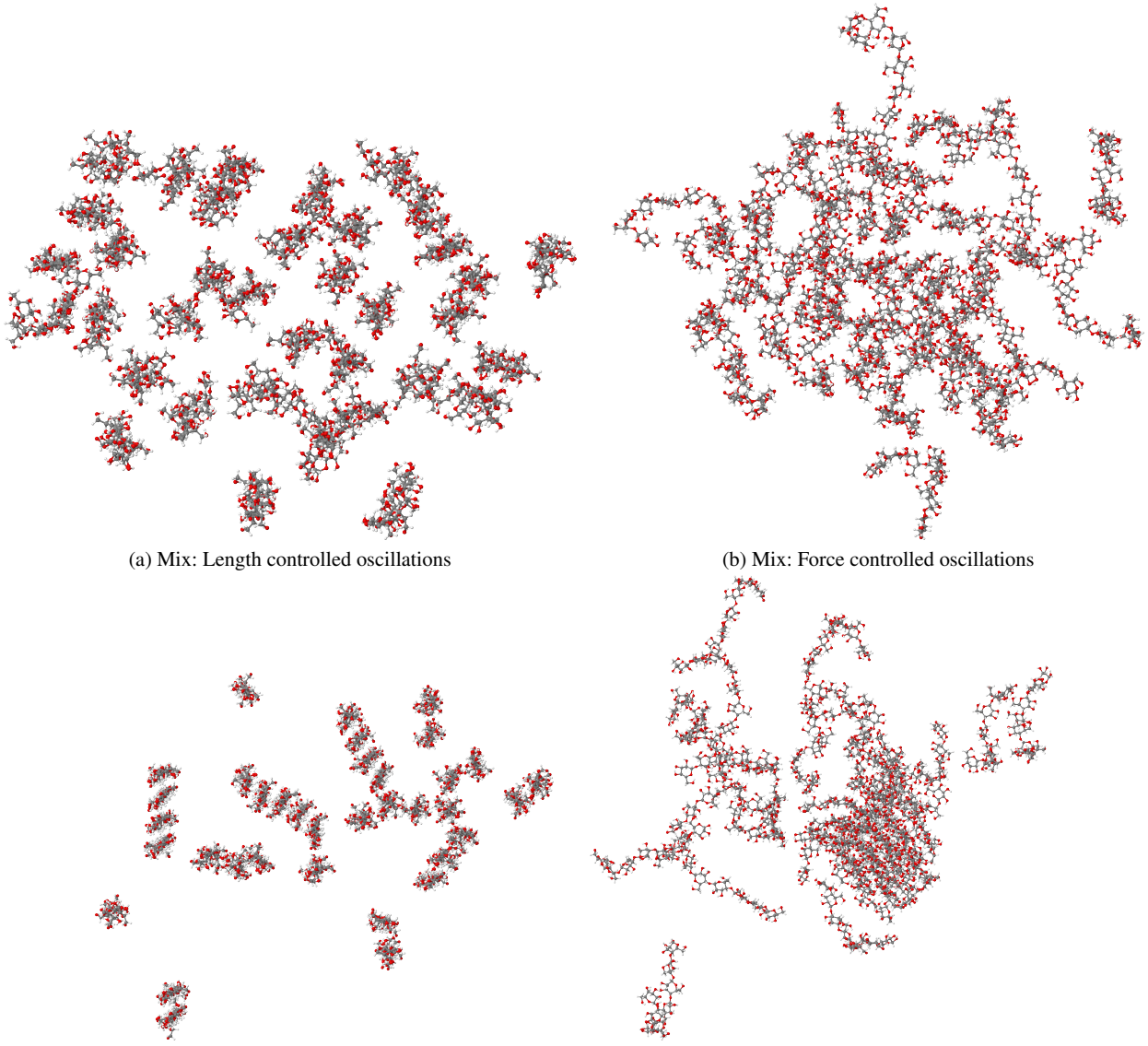
The difference between the length-controlled and force-controlled simulations cannot be explained by more energetic agitation. Conversely, the amplitude for the length controlled simulations is such that the forces on the end points vary by approximately 5 nN/chain. This is an order of magnitude more than in the force-controlled simulations, where the amplitude was 0.35 nN/chain.

The faster dissolution in the force-controlled simulations therefore must be related to the geometry of the two different boundary conditions. In the length controlled simulations the ends are more restricted in space than in the force-controlled ones. This means that the length-controlled boundary conditions are more representative of the edge dynamics of a section in the middle of a larger bundle, while the force-controlled boundary conditions are more representative of an actual end of a bundle. This indicates that there may be some dissolution mechanism that involves chains loosening at the ends of the bundles. This is consistent with the fact that in experiments, shorter bundles are known to dissolve more easily.

With length-controlled oscillations, the bundle is twisting during compression, as the chains are free to move in the axial planes. This induces shear stress on the bundle, and the mechanics are reminiscent of shear stresses induced by an agitated solvent.

In Fig. 11 (a) we show the density as a function of the time since the start of the oscillations. In all four cases, the density immediately starts to decrease as the bundle dissolves. However, the decrease is more rapid for the systems subjected to force-controlled oscillations than those subjected to length-controlled oscillations.

We note that the density here is an indication of the maximum spreading out of the cellulose molecules, which happens by diffusion once they have detached from the bundle. While the density in water decreases at a similar rate



(a) Mix: Length controlled oscillations

(b) Mix: Force controlled oscillations

(c) Water: Length controlled oscillations

(d) Water: Force controlled oscillations

Fig. 9: Cross sections of bundles with 36 chains after 30 ns of length controlled oscillations and 10 ns of force controlled oscillations in mixture and water. The force-controlled oscillations dissolves the bundle fast in both water and the solvent mixture. With length-controlled oscillations, the dissolution is incomplete also after 30 ns and qualitatively different in water and in the solvent mixture.

to the density in the solvent mixture, this does not necessarily mean that the bundles dissolve at a similar rate, merely that the size of the particles diffusing is the same (i.e. single chains). In the case of length-controlled simulations, the decrease in density is slower, due to the fact that the units diffusing away from the bundle are larger and heavier sheet-like structures.

The difference in the dissolution process in water and in the solvent mixture is shown more clear in Fig. 11 (b), displaying the time evolution of the number of clusters in each system. While the two systems with oscillating force both reach full dissolution, the bundle in the mixture solvent is dissolving significantly faster. The difference is largest around 10 ns, which is the time frame of the trajectories displayed in Fig. 9.

4 Conclusions

We have performed a series of molecular dynamics simulations investigating the dissolution of 36-chain cellulose bundles in both water and a mixture solvent consisting of water with Na^+, OH^- and urea. We find that without additional external forces, cellulose bundles of 36 chains, which is what

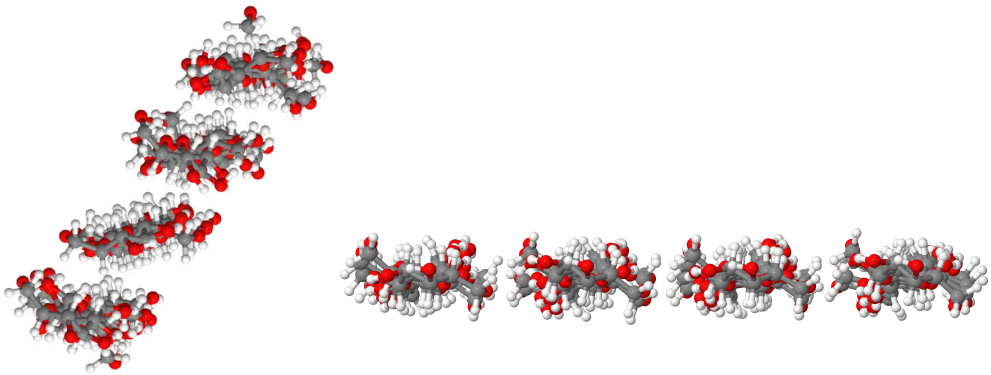
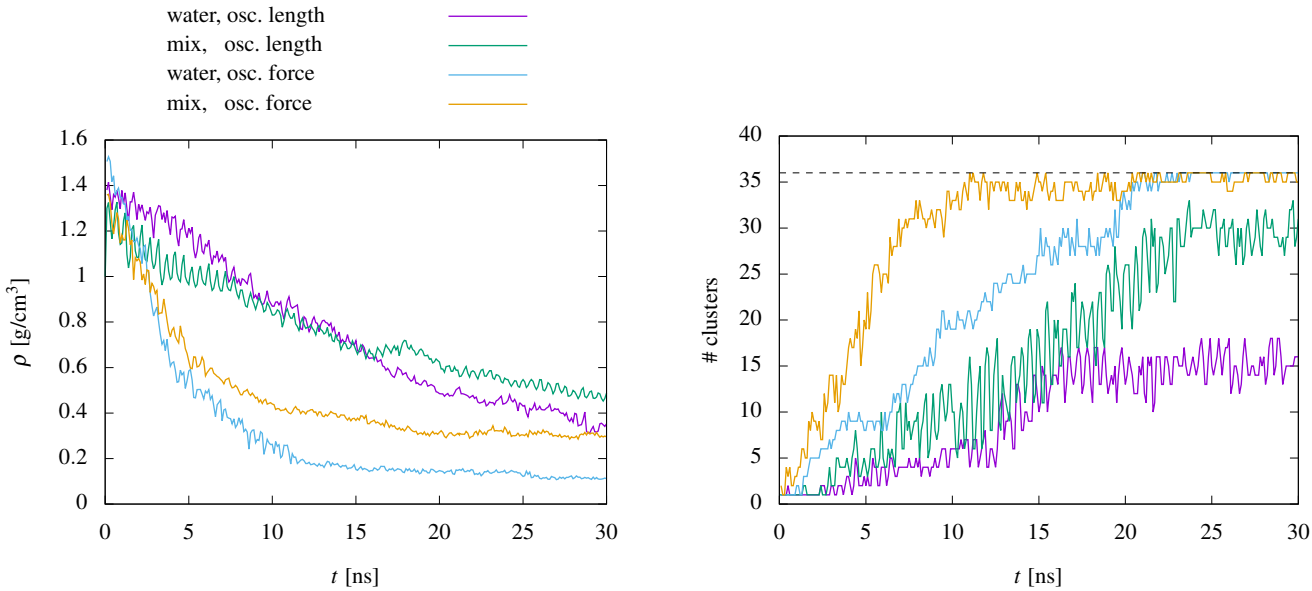


Fig. 10: Comparison of a representative bundle of vertically stacked chains formed in the length-controlled simulations in water (left), and a sheet of 4 chains extracted from the energy minimized maze configuration (right). The chains in sheet configuration interact via in-plane hydrogen bonds, while these interactions are not present between the vertically stacked chains.



(a) Density

(b) Number of clusters

Fig. 11: Time evolution of density (a) and number of clusters (b) for bundles of 36 chains of cellulose at $T = 258$ K with oscillating length and force. The difference in behaviour between the bundle in water and the bundle in the solvent mixture is most clear in the figure for the number of clusters, (b), where the horizontal line represents complete dissolution.

is found in nature, do not dissolve. This is consistent with experimental results. We do however find that in the mixture solvent, the urea intercalates into the bundles, increasing their volume significantly. This is in line with e.g. Liu et al. [26], who found numerically that urea had to be removed from cellulose chains in order for them to agglomerate. For smaller bundles, we find similar results regarding the stability in the two different solvents. We also find that sheets of cellulose are stable in the mixture solvent, but in pure water they immediately collapse into bundles. This appears to be caused by shielding of cellulose hydrophobic interactions by urea in the mixture solvent.

When we add periodic external forces to mimic stirring or agitation of the solvent-cellulose mixture, we see that the bundles are being dissolved. We find that the cellulose dissolves more readily in the mixture solvent than in pure water, which is consistent with experimental results. In the water solvent, the dissolution is incomplete. Under boundary conditions that are likely more representative of longer cellulose bundles, they tend to leave sheets of cellulose that are bound together with hydrophobic interactions. This is

further evidence of the urea stabilizing effect. It is difficult to compare our applied forces quantitatively to specific levels of stirring, and therefore to compare directly to dissolution in experiments. Our agitation may be significantly more energetic than what is achievable by stirring in experiments. Nevertheless, qualitatively, our findings are consistent with the experimental results that the NaOH and urea solvent is able to completely dissolve cellulose under experimental conditions when water only partially dissolves cellulose. Partial dissolution of cellulose in water, should be interpreted together with the fact that NaOH and water is indeed experimentally capable of dissolving cellulose with less than 200 glucose units [13]. It is therefore not surprising that partial dissolution occurs in a small (eight glucose units) modeled system with only water. The different behaviour of water and solvent is likely related to cellulose-cellulose hydrophobic interactions, which appear to be more stable in pure water than in the mixture solvent, where urea is absorbed into the bundles.

Acknowledgements The authors are grateful to the Research Council of Norway for funding through project no. 250158 and project no. 262644 (PoreLab Centre of Excellence). The simulations were performed on resources provided by UNINETT Sigma2 - the National Infrastructure for High Performance Computing and Data Storage in Norway and the Idun cluster provided by the NTNU High Performance Computing Group.

5 Declarations

5.1 Funding

The authors received funding by the Research Council of Norway through project no. 250158 and project no. 262644.

5.2 Conflicts of interest

The authors declare that they have no conflict of interest.

5.3 Availability of data and material and code availability

LAMMPS scripts will be made available upon acceptance.

5.4 Ethics approval

The authors have followed the guidelines of the Committee on Publication Ethics.

5.5 Consent to participate

Not applicable.

5.6 Consent for publication

All authors have read and agreed to publish the latest version of the manuscript.

References

1. A. Dufresne. Nanocellulose From Nature to High Performance Tailored Materials. (2012). DOI 10.1515/9783110254600
2. S. Zhang, C. Chen, C. Duan, H. Hu, H. Li, J. Li, Y. Liu, X. Ma, J. Stavik, Y. Ni, BioResources **13**(2), 1 (2018). DOI 10.15376/biores.13.2.Zhang
3. A.D. French, Cellulose **24**(11), 4605 (2017). DOI 10.1007/s10570-017-1450-3
4. D.P. Delmer, Y. Amor, Plant Cell **7**(7), 987 (1995). DOI 10.1105/tpc.7.7.987
5. Y. Habibi, L.A. Lucia, O.J. Rojas, Chemical Reviews **110**, 3479 (2010)
6. D.J. Cosgrove, Current Opinion in Plant Biology **22**, 122 (2014). DOI 10.1016/j.pbi.2014.11.001
7. M.C. Jarvis, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences **376**(2112) (2018). DOI 10.1098/rsta.2017.0045
8. B. Lindman, G. Karlström, L. Stigsson, Journal of Molecular Liquids **156**(1), 76 (2010). DOI 10.1016/j.molliq.2010.04.016
9. B. Medronho, B. Lindman, Current Opinion in Colloid and Interface Science **19**(1), 32 (2014). DOI 10.1016/j.cocis.2013.12.001
10. C. Yamane, T. Aoyagi, M. Ago, K. Sato, K. Okajima, T. Takahashi, Polymer Journal **38**(8), 819 (2006). DOI 10.1295/polymj.PJ2005187
11. v.H. Staudinger, K.H. In den Birken, M. Staudinger, Die Makromolekulare Chemie **9**(1), 148 (1953). DOI 10.1002/macp.1953.020090109
12. K. Pal, A.K. Banthia, D.K. Majumdar, Biomedical materials (Bristol, England) **1**(2), 85 (2006). DOI 10.1088/1748-6041/1/2/006
13. K. Kamide, K. 0kajima, T. Matsui, K. Kowsaka, Polymer Journal **16**(12), 857 (1984)
14. T. Yamashiki, K. Kamide, K. Okajima, K. Kowsaka, T. Matsui, H. Fukase, Polymer Journal **20**(6), 447 (1988). DOI 10.1295/polymj.20.447
15. J. Cai, L. Zhang, Macromolecular Bioscience **5**(6), 539 (2005). DOI 10.1002/mabi.200400222
16. H. Jin, C. Zha, L. Gu, Carbohydrate Research **342**(6), 851 (2007). DOI 10.1016/j.carres.2006.12.023
17. L. Yan, Z. Gao, Cellulose **15**(6), 789 (2008). DOI 10.1007/s10570-008-9233-5
18. A. Isogai, R.H. Atalla, Cellulose **5**(4), 309 (1998). DOI 10.1023/A:1009272632367
19. S. Fischer, H. Leipner, K. Thümmler, E. Brendler, J. Peters, Cellulose **10**(3), 227 (2003). DOI 10.1023/A:1025128028462
20. Z. Liu, C. Zhang, R. Liu, W. Zhang, H. Kang, P. Li, Y. Huang, Cellulose **23**(1), 295 (2016). DOI 10.1007/s10570-015-0827-4
21. B. Xiong, P. Zhao, K. Hu, L. Zhang, G. Cheng, Cellulose **21**(3), 1183 (2014). DOI 10.1007/s10570-014-0221-7
22. N. Isobe, K. Noguchi, Y. Nishiyama, S. Kimura, M. Wada, S. Kuga, Cellulose **20**(1), 97 (2013). DOI 10.1007/s10570-012-9800-7
23. Z.A. Parray, M.I. Hassan, F. Ahmad, A. Islam. Amphiphilic nature of polyethylene glycols and their role in medical research (2020). DOI 10.1016/j.polymertesting.2019.106316
24. E. Wernersson, B. Stenqvist, M. Lund, Cellulose **22**(2), 991 (2015). DOI 10.1007/s10570-015-0548-8
25. M. Bergenstråhle-Wohlert, L.A. Berglund, J.W. Brady, P.T. Larsson, P.O. Westlund, J. Wohlert, Cellulose **19**(1), 1 (2012). DOI 10.1007/s10570-011-9616-x

26. G. Liu, W. Li, L. Chen, X. Zhang, D. Niu, Y. Chen, S. Yuan, Y. Bei, Q. Zhu, Colloids and Surfaces A: Physicochemical and Engineering Aspects **594**, 124663 (2020). DOI 10.1016/j.colsurfa.2020.124663
27. L. Cai, Y. Liu, H. Liang, Polymer **53**(5), 1124 (2012). DOI 10.1016/j.polymer.2012.01.008
28. V. Agarwal, G.W. Huber, W.C. Conner, S.M. Auerbach, Journal of Chemical Physics **135**(13), 0 (2011). DOI 10.1063/1.3646306
29. Tulus, Mardiningsih, Sawaluddin, O.S. Sitompul, A.K. Ihsan, IOP Conference Series: Materials Science and Engineering **308**(1) (2018). DOI 10.1088/1757-899X/308/1/012048
30. S.Y. Ding, S. Zhao, Y. Zeng, Cellulose **21**(2), 863 (2014). DOI 10.1007/s10570-013-0147-5
31. R. Woods, Co-workers. (2015-2019) GLYCAM Web. Complex Carbohydrate Research Center, University of Georgia, Athens, GA
32. T.C. Gomes, M.S. Skaf, Journal of Computational Chemistry **33**(14), 1338 (2012). DOI 10.1002/jcc.22959
33. Y. Nishiyama, P. Langan, H. Chanzy, Journal of the American Chemical Society **124**(31), 9074 (2002). DOI 10.1021/ja0257319
34. M.R. Shirts, C. Klein, J.M. Swails, J. Yin, M.K. Gilson, D.L. Mobley, D.A. Case, E.D. Zhong, Journal of Computer-Aided Molecular Design **31**(1), 147 (2017). DOI 10.1007/s10822-016-9977-1
35. S. Plimpton, Journal of Computational Physics **117**(1), 1 (1995). DOI 10.1006/jcph.1995.1039
36. L. Martínez, R. Andrade, E.G. Birgin, J.M. Martíne, Journal of Computational Chemistry **30**(13), 2157 (2009). DOI 10.1002/jcc
37. R. Kirschner, K.N., Yongye, A.B., Tschampel, S.M., Daniels, C.R., Foley, B.L, Woods, Journal of Computational Chemistry **29**, 622 (2008)
38. H.J. Berendsen, J.R. Grigera, T.P. Straatsma, Journal of Physical Chemistry **91**(24), 6269 (1987). DOI 10.1021/j100308a038
39. P. Li, L.F. Song, K.M. Merz, Journal of Chemical Theory and Computation **11**(4), 1645 (2015). DOI 10.1021/ct500918t
40. G.A. ÖzpInar, W. Peukert, T. Clark, Journal of Molecular Modeling **16**(9), 1427 (2010). DOI 10.1007/s00894-010-0650-7
41. T. Schneider, E. Stoll, Phys. Rev. B **17**(3), 1302 (1978). DOI 10.1103/PhysRevB.17.1302
42. P. Virtanen, R. Gommers, T.E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S.J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A.R. Nelson, E. Jones, R. Kern, E. Larson, C.J. Carey, lhan Polat, Y. Feng, E.W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. Quintero, C.R. Harris, A.M. Archibald, A.H. Ribeiro, F. Pedregosa, P. van Mulbregt, S... Contributors, Nature Methods (2020)
43. R.T. McGibbon, K.A. Beauchamp, M.P. Harrigan, C. Klein, J.M. Swails, C.X. Hernández, C.R. Schwantes, L.P. Wang, T.J. Lane, V.S. Pande, Biophysical Journal **109**(8), 1528 (2015). DOI 10.1016/j.bpj.2015.08.015
44. A. Shrake, J.A. Rupley, Journal of Molecular Biology **79**(2), 351 (1973). DOI 10.1016/0022-2836(73)90011-9
45. P. Chen, Y. Nishiyama, K. Mazeau, Cellulose **21**(4), 2207 (2014). DOI 10.1007/s10570-014-0279-2