

Molecular-level driving forces in lignocellulosic biomass deconstruction for bioenergy

Loukas Petridis and Jeremy C. Smith

University of Tennessee/Oak Ridge National Laboratory, Center for Molecular Biophysics, P.O. Box 2008 Oak Ridge TN 37831-6309, USA.

Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville TN 37996.

email: smithjc@ornl.gov

Abstract | The plant cell wall biopolymers lignin, cellulose and hemicellulose are potential renewable sources of clean biofuels and high-value chemicals. However, the complex 3D structure of lignocellulosic biomass is recalcitrant to deconstruction. Major efforts to overcome this recalcitrance have involved pretreating biomass before catalytic processing. This Perspective describes recent work aimed at elucidating the molecular-level physical phenomena that drive biomass assembly and are at play in commonly-employed aqueous-based and thermochemical pretreatments. Several key processes have been found to be driven by biomass solvation thermodynamics, an understanding of which therefore facilitates the rational improvement of methods aimed at the complete solubilization and fractionation of the major biomass components.

[H1] Introduction

The cell walls of plants are mostly comprised of lignocellulosic biomass, which itself consists mainly of aligned bundles of partly crystalline and partly amorphous cellulose fibrils embedded in a disordered matrix of hemicellulose and lignin (FIG. 1) (REF 0-1, 0-2). The inherent resistance of this assembly to deconstruction is the major obstacle to the cost-effective transformation of biomass into biofuels, such as ethanol, and high-value bioproducts (REF 0-3). Biomass recalcitrance arises from many factors, many of which are related to the inaccessibility of cleavable bonds (REF 0-4). Such inaccessibility is a result of the self-association and crystallinity of cellulose and the compaction and aggregation of lignin and its binding to cellulose and hydrolytic enzymes. To overcome biomass recalcitrance, chemists have devised and explored¹ various pretreatment methods involving physical, biological, chemical (catalytic)² and solvent-based processes. Pretreatments are performed in order to alter the 3D structure, interactions and composition of biomass, with the goal of enabling its efficient dissolution and fractionation into components that are then deconstructable by enzymatic or artificial catalysts^{3,4}.

Among the simplest physical pretreatment methods are comminutory processes such as ball-milling⁵, which are performed to convert macroscopic particles into smaller and more reactive fragments. In contrast, relatively harsh industrial pretreatment methods such as pyrolysis and gasification have been examined but do not preserve the monomer structures of the biomass components.⁶ Thus, there is intense research interest in developing milder thermochemical pretreatment methods. Procedures using dilute acid or ammonia⁷ (REF 7-1) can be effective. Similarly, ionic liquids^{8,9}, deep eutectic solvents¹⁰ and mixtures of H₂O with organic co-solvents can also help break down lignocellulosic biomass. Pretreating biomass leads to coupled physical and chemical changes, and the litany of possible chemical reactions has been summarized in more

specialist review articles^{6,11,12}. In this Perspective, we instead focus on broadly applicable, fundamental, molecular-level physical phenomena that have recently been revealed to drive the formation of the structures and associations of the biopolymers. Understanding what physicochemical forces keep these structures together allows us to better understand how to break them down. The concepts described are relevant to most pretreatment processes in which biomass is in aqueous solution, with or without other solvents.

The configurational behavior of biomass polymers can be understood within the general framework of the ‘quality’ of a solvent¹⁵. Three classes of solvent can be considered. In a bad solvent, polymer–polymer interactions are favored and the polymer assumes collapsed conformations in which chains are tightly packed. In a θ solvent, polymer–polymer and polymer–solvent interactions balance exactly, leading to the polymer adopting Gaussian ‘random-coil’ conformations, similar to an ‘ideal’ chain without excluded volume interactions. Finally, in a good solvent, polymer–solvent interactions are energetically favorable, and the polymer adopts more extended, self-avoiding conformations. For example, H₂O is a good solvent for hemicellulose^{16,17}. In contrast, at room temperature H₂O is a bad solvent for both lignin and cellulose. Finding a good or θ solvent for biomass polymers is important for developing effective chemical pretreatment.

[H1] Biomass components

We now describe the major components of biomass in more detail, paying particular attention to the intra- and intermolecular interactions responsible for the structures that they adopt *in vivo*. Underlying the thermodynamics discussed are solvation effects and non-covalent interactions, particularly hydrophobic interactions and hydrogen bonds.

[H2] Cellulose

Cellulose is synthesized^{19,20} (REF 19-1) *in vivo* by cellulose synthase enzymes embedded in the plasma membrane. The catalytic subunit within cellulose synthase acts on the substrate uridine diphosphate glucose to form a cellulose polysaccharide, which is then translocated across the membrane through a pore subunit. Several such enzymes are usually nearby, forming complexes known as “rosettes”, and the crystallization of polysaccharide chains into microfibrils (FIG. 1B) occurs away from the enzyme complex (REF 18-1). Crystalline regions in native cellulose consist of a mixture of two distinct ‘cellulose I’ forms, which differ slightly in their unit cells^{26,27}. Microfibrils have often been represented as hexagonal arrangements of 36 chains^{21,22}, but more recent estimates are in the range of 18–24 chains.²³⁻²⁵ The microfibrils assemble into fibres in the cell wall.¹⁸

Cellulose is completely insoluble in H₂O below ~300 °C (REF. 28). The molecular-level interactions and thermodynamics behind this insolubility are of interest. In general, dissolution is favored by the increase in entropy associated with removing the impermeable partition between the solid and liquid phases. However, the contribution of this entropy of mixing is greater for small molecules than it is for polymers, because a greater number of individual molecules is released in the former case. Another aspect favoring dissolution is the significantly larger conformational entropy of single polymer molecules in solution compared to the solid state. Indeed, the cores of cellulose fibers are highly ordered (FIG. 1B), such that cellulose has impressive mechanical properties — its axial elastic modulus is greater than that of Kevlar²⁹ and the persistence length for cellulose nanofibers in H₂O has been estimated to be ~2.5 μm ³⁰. However, individual short

cello-oligomer chains are also fairly rigid,³¹ with the persistence length for a single cellulose chain in H₂O being relatively high (~10 nm),^{32,33} reflecting a solution phase rigidity that means there is relatively little entropic gain associated with dissolution. A third effect favoring dissolution is hydrogen bonding — the number of hydrogen bonds per cellulosic OH group is lower in the crystalline state than in aqueous solution³¹.

It turns out that microfibril assembly is driven by the intrinsic structural anisotropy of single cellulose chains — the hydrophilic CH₂OH and OH groups of each monomer are located at the equatorial positions of the glucopyranose ring, whereas the top and bottom ring surfaces are more hydrophobic²⁷. The cellulose chains are thus amphiphilic. Equatorial inter-chain and intra-chain OH···O hydrogen bonding leads to several chains organizing into a sheet. The hydrophobic surfaces of the sheets then stack leading to the microfibrils retaining corresponding hydrophilic and hydrophobic surfaces. The microfibrils are twisted on account of the hydrogen bonding between the chiral monomers (REF 34-36). Quantum chemical calculations and molecular dynamics (MD) simulations have shown the interactions between cellulose sheets to be enthalpically strong (REF 37). Indeed, the interaction energies per residue are greater between sheets than they are within them³⁷. A variety of cooperative interactions contribute to the stacking enthalpy^{31,34-3} involving dispersion, charge-transfer, exchange and electrostatic interactions, which result in numerous, relatively weak C–H···O hydrogen bonds and van der Waals interactions. The stacking of the cellulose sheets also gives rise to a substantial hydrophobic stabilization free energy change. H₂O molecules lining the hydrophobic surfaces form stronger hydrogen bonds with each other than they would in bulk H₂O, such that their spatial density correlation is higher when hydrating cellulose^{38,39}. Further, the interaction of the glucan OH groups with these H₂O molecules further restricts their configurational freedom³¹. Consequently, the solvent entropy contribution to cellulose crystalline fibre formation is substantial³⁸. Separating individual chains stacked in solution has been estimated to require ~2 kcal mol monomer⁻¹(REF. 31).

There is also evidence that surface hydration rigidifies cellulose fibrils, with the calculated persistence length of the fibril increasing on hydration⁴⁰. This increased rigidity is, perhaps counterintuitively, linked to increased surface disorder, which is manifested in both experiments and calculations as a progressive loss of order from the centre of the fibril outwards^{21,41}. The surface disorder leads to a significant increase in the number of hydrogen bonds ‘bridging’ monomers, and these rigidify the fibril.

[H2] Hemicelluloses

Whereas cellulose is assembled from only a single monomer (D-glucose) and is unbranched, hemicelluloses are built from multiple types of monomer and are branched. Therefore, of the two polymers, it is only cellulose that crystallizes. Polysaccharide hydrolysis rates are influenced by the accessibility of β-1,4-glycosidic bonds to solvent, but H₂O is absent from crystalline regions of cellulose. Therefore, whereas amorphous cellulose is quickly hydrolyzed, the dissolution of crystalline cellulose is limited⁴² and H₂O is a good solvent for hemicelluloses but a bad one for crystalline cellulose microfibrils. As shown by small-angle neutron scattering experiments, hemicelluloses adopt structures that are penetrated by H₂O (REF. 17), and this accessibility helps us to remove hemicellulose from biomass. Indeed, acid hydrolysis of Avicel (a microcrystalline form of cellulose) is much slower than that of xylan hemicelluloses on account of the crystallinity of the former⁴².

[H2] Lignin

Monolignols such as **1–3** (FIG. 1Ac) polymerize to form lignin, whose respective G, S and H units differ in their degree of methoxy substitution. Certain bonds between these units are easier to cleave than others. Lignins rich in G feature more recalcitrant β -5', β - β' and 5-5' linkages and stronger π - π interactions than lignins rich in S and H, making G-rich lignins difficult to solubilize⁴³. In contrast, lignins rich in S are less cross-linked and feature a higher proportion of β -O-4 linkages — bonds that are the most easily cleaved — making them more easily depolymerized and extracted into solution⁴⁴. Lignin polymers with a high relative abundance of H usually have lower molecular weights, which also contributes to them being more easily deconstructed^{45,46}.

Overall, lignin is hydrophobic on account of its aromatic rings. Thus, H₂O is a poor solvent for lignin, which adopts compact, collapsed, glassy, ellipsoidal forms in aqueous solution near room temperature (FIG. 2a)^{43,47-49}. Lignin is always compact in H₂O at temperatures below ~210 °C (FIG. 2b) and at higher temperatures is chemically degraded. Therefore, the random coil state is never observed in pure H₂O although it can be observed when a co-solvent mixture is used (FIG. 2c). At room temperature, aqueous lignin can be represented by a 'blob' model⁵⁰ consisting of ~15-monomer 'blobs' — segments in which monomers proximal along the chain are proximal in space⁴⁷. These blobs interpenetrate to form a relatively spherical 'equilibrium' globule (FIG. 2a).

We now consider the molecular driving forces that stabilize the lignin equilibrium globule at room temperature. The enthalpy change associated with compacting lignin into a globule is negative, because lignin-H₂O interactions are stronger than lignin-lignin interactions⁴⁷. The lignin conformational entropy also favours extended states. Rather, lignin undergoes compaction at room temperature in aqueous solution because H₂O molecules on the surface of lignin have lower translational freedom and lower density fluctuations (lower compressibility) than they do in the bulk. The collapse of lignin to compact equilibrium globules is thus driven by the entropy of the H₂O molecules increasing as they are displaced into the bulk.^{47,51} This lignin collapse is distinct from the collapse of a purely hydrophobic polymer because the latter process is enthalpically driven, with the change in hydration entropy being unfavorable⁵².

BOX 1 | The hydrophobic effect

The hydrophobic effect is different on 'small' (<1 nm) and 'large' (>1 nm) length scales¹⁰⁶. On the small length scale (for example, in the hydration of small nonpolar solutes), the hydrogen-bonding network between solvating H₂O molecules remains intact, with the hydrogen bonds 'reaching around' the solute. However, the required specific spatial organization of hydrogen-bonding patterns has an entropic cost, and correspondingly the main contribution to the hydrophobic interaction of small molecules is entropic: strong correlations in solute-solvent centre-of-mass translational motion¹⁰⁷. In contrast, large length-scale hydrophobicity is enthalpically driven and involves surface dewetting. At room temperature, the hydrophobic surfaces of cellulose and lignin, although physically extended, are heterogeneous and have thus been observed in simulations to show the hydration signature of the small-scale hydrophobic effect^{37,47} in which the H₂O hydrogen bonding network is not substantially perturbed by the solute. For example, in aqueous lignin, a H₂O molecule in the hydration shell participates in only 3% fewer hydrogen bonds on average than does a bulk H₂O molecule⁴⁷.

[H1] Biomass pretreatment

We now discuss the molecular driving forces leading to morphological changes during biomass pretreatment. These insights can suggest ways to rationally improve the process.

[H2] Effects of temperature on biomass

The above discussion has already made clear how the entropy of H₂O solvent molecules can be the deciding molecular-scale factor in the behavior of biomass polymers in aqueous solution. Recently, the water entropy effect has also been seen to drive pretreatment effects. In the case of poplar biomass, a combination of multiple experimental and computational probes allowed the identification of two fundamental H₂O-driven processes responsible for molecular-scale morphological changes during steam explosion pretreatment (SEP, FIG. 3)⁵⁴⁻⁵⁶.

The first process is a growth in the crystalline regions of cellulose, as evidenced by sharper X-ray diffraction peaks. This coincides with the coalescence of the microfibrils, indicated by the small-angle scattering features moving to lower scattering angles. Core H₂O molecules are inserted between separate microfibrils in biomass prior to SEP⁵⁴. These H₂O molecules form strong hydrogen bonds to cellulose, and the insertion of these molecules into the cellulose core is highly favorable at lower temperatures. However, at the temperatures used in SEP (~160–200 °C), the term $-T\Delta S$ associated with the confinement of the core H₂O becomes more unfavourable⁵⁴, while the density of the core H₂O decreases, leading to a weakening of the cellulose–H₂O interactions that disfavor fibril coalescence⁵⁸. Thus, at high temperatures H₂O molecules are irreversibly released from the core, such that the microfibrils coalesce and form larger crystalline domains. An increase in crystallinity can also be seen upon the removal of hemicellulose⁵⁷.

The second process at play in SEP is a temperature-induced attenuation of lignin–hemicellulose entanglement, allowing a separation of lignin and hemicellulose phases that causes the plant cell wall to become more porous⁵⁴. This weaker lignin–hemicellulose association coincides with an increase in the entropy of the H₂O molecules hydrating lignin such that, at temperatures above ~147–207 °C, the density fluctuations of H₂O molecules solvating lignin become similar to bulk H₂O. Thus, the entropic penalty for H₂O confinement to the lignin surface is reduced and softwood lignin undergoes a transition between collapsed states: from a globular equilibrium globule state (FIG. 2a) to a less spherical, ‘crumpled’ globule (FIG. 2b). The crumpled globule has a higher solvent-exposed surface area and the lignin blobs remain intact but become separated from each other. In turn, the larger surface area of the crumpled globule weakens the binding of hemicellulose to lignin. We again stress that lignin is different to a purely hydrophobic polymer,⁵⁹ for which such a shape change is not usually observed.

The two processes we have described — cellulose aggregation and lignin:hemicellulose phase separation — proceed not only in pure H₂O but also other common aqueous thermochemical pretreatments, which might involve using dilute acid or NH₃-induced fiber expansion⁵⁴. An additional temperature effect is the lignin ‘glass-to-liquid’ transition, which is common in biopolymers and amorphous polymers. In the case of dry lignin, this glass transition temperature, T_g falls in the range 50–150 °C, with the value depending on the plant source material, the processing conditions, and the method used to measure T_g (REF. 60,61). At temperatures below

T_g , lignin is glassy — hard and stiff, with its monomer units trapped in cages formed by other parts of the polymer such that lignin is structurally arrested and cannot undergo substantial rearrangements on timescales shorter than ~ 100 s (REF. 62-64). At temperatures above T_g , lignin assumes a liquid-like phase, which is softer with subunits that experience redistribution and relocalization, facilitating downstream processing⁶⁵.

During thermochemical pretreatment, phase-separated lignin self-aggregates⁶⁶ as a result of driving forces that are analogous to those that cause single lignin molecules to collapse in H_2O . Further, lignin- H_2O interactions are less favorable than lignin-cellulose interactions because the latter can involve hydrophobic regions in both polymers. Thus, the relative weakness of lignin- H_2O interactions means that, in aqueous solution, aggregates of lignin strongly associate with cellulose^{65,67} and render it inaccessible to enzymes. Therefore, although purely aqueous high-temperature pretreatments cause lignin-hemicellulose phase separation and hemicellulose removal, such pretreatments are of limited use because lignin is difficult to completely remove on account of its strong interactions with cellulose⁶⁸. Notably, MD simulations show that lignin interacts more weakly with amorphous cellulosic regions than it does with crystalline regions. The origin of this is yet another result of hydration — the amorphous cellulosic regions interact more strongly with H_2O and their desolvation barrier hinders their interaction with lignin⁶⁹. Therefore, promoting cellulose decrystallization may have an added benefit in that it reduces the degree of cellulose-lignin association.

A further undesirable process is the interaction of phase-separated lignin with cellulolytic enzymes (cellulases)^{70,71}. The adsorption of enzymes on lignin surfaces has been found to correlate with the degree of clustering of nonpolar residues on the enzyme surface⁷². Other MD studies implicated two competitive binding processes to be at play. The first of these is the preferential binding of lignin to the hydrophobic faces of cellulose, which also happen to be the preferred binding sites of cellulases⁷³. The second is the specific binding of lignin to the three tyrosine residues of the cellulose-binding modules of cellulases⁷⁴. Lignin thus binds exactly where, for industrial purposes, it is least desired. Knowledge of these effects gives us a simple, plausible explanation as to why lignin is so effective at stopping cellulases from hydrolyzing cellulose^{70,71}.

The solubilization of hemicellulose during pretreatment is hindered by covalent and non-covalent interactions with other cell-wall components^{75,76}. Hemicellulose associates non-covalently with cellulose through hydrogen bonds⁷⁷ and these interactions are strongest when the hemicellulose substitution follows an even pattern⁷⁸. Again, hydration is critical — MD results in the case of the xylan (a hemicellulose derived from xylose) show that the loss of the cellulose hydration shell ordering at high temperatures leads to stronger xylan-cellulose interactions, leading to lower cellulose conversion⁷⁹.

[H2] Organic solvents

We now turn our attention to the desirable properties for a solvent (or mixture) to be considered for the pretreatment of lignocellulosic biomass. A key requirement is to increase the quality of the solvent for both cellulose and lignin - H_2O is a bad solvent for these biopolymers whereas we want a liquid that is a θ or good solvent. Finding such a solvent will need us to recognize the amphiphilic character of both biopolymers. Whereas hydration entropy plays a critical role in aqueous solution, for efficient cellulose and lignin dissolution direct polar and nonpolar solvent interactions are also

required. We now present as a case study an effective method that employs a mixture of H₂O with tetrahydrofuran (THF), a relatively nonpolar co-solvent⁸⁰⁻⁸². The principles discussed here are also applicable to other co-solvents such as γ -valerolactone (GVL), 1,4-dioxane, Me₂SO and Me₂CO. These polar aprotic solvents are useful because they are miscible with H₂O, can accept hydrogen bonds, and their application leads to substantial biomass delignification⁸³.

For THF/H₂O mixtures act as θ solvents, because in these mixtures solvent–lignin and lignin–lignin interactions are approximately equal in strength. THF preferentially solvates the aromatic rings of lignin, and in doing so shifts the equilibrium configurational distribution of the biopolymer from a compact globule to a random coil⁵³ (FIG. 2c). Further, lignin does not self-aggregate in THF/H₂O solution. Also, when in dilute acid solution above ~130 °C, broken lignin linkages may recombine with other groups leading to troublesome lignin ‘recondensation’ reactions⁸⁴; but the separation of the individual lignin chains in THF/H₂O solution precludes such unwanted reactions. Moreover, the H₂O molecules in the mixed solvent hydrogen bond to the labile ether linkages of lignin, the hydrolysis of which is thus not impeded by the presence of THF⁵³.

In the case of cellulose, an effective pretreatment solvent needs to dissolve chains by disrupting their hydrogen bonds to other chains. Also, aside from competing for hydrogen bonding, the solvent must also interfere with the hydrophobic stacking interactions between cellulose sheets. Evidence from computer simulations suggests that solvent–H₂O mixtures have rather variable local phase separation behaviours at the cosolvent–cellulose interface⁸⁵. THF–H₂O, GVL–H₂O, EtOH–H₂O and Me₂CO–H₂O mixtures all undergo demixing on the cellulose surface, with the extent of this demixing being predictable from their degree of deviation from Raoult’s law, *i.e.* the difference between the strengths of interaction of the solvent components. For example, THF and H₂O spontaneously phase separate at the local surfaces of a cellulose microfibril, with the H₂O molecules hydrogen bonding to the hydrophilic cellulose faces while the THF molecules aggregate on the hydrophobic faces (FIG. 4)⁸⁶. The binding of THF to the hydrophobic faces also blocks lignin aggregation on those faces⁷⁴.

[H2] Ionic liquids

High yields of biomass deconstruction can be achieved using ionic liquids (ILs)⁸⁷. Early work led to the suggestion that cellulose dissolution in ILs arises from hydrogen bonding between the polysaccharide OH and the IL anions⁸⁸ — interactions that break up the hydrogen bonding network within cellulose microfibrils⁸⁹. Indeed, cellulose dissolution in ILs proceeds best when the anion has a high basicity towards hydrogen bonding⁹⁰. However, both IL anions and cations are now known to perform this hydrogen bonding function^{88,91,92}. Furthermore, just as amphiphilic effects are important in how H₂O–cosolvent mixtures interact with cellulose, these effects are also important for ILs because they also enable disruption of the non-polar faces of cellulose.^{93,94} MD simulations using 1-*n*-butyl-3-methylimidazolium chloride ([bmim]Cl) indicate that while Cl[−] anions disrupt accessible intramolecular cellulose hydrogen bonds, [bmim]⁺ cations stack on and intercalate between the hydrophobic surfaces of cellulose⁹⁵. A synergistic mechanism has been proposed in which anions initially insert in the cellulose strands and then encourage the subsequent insertion of the cations⁹⁶.

Other, less direct, effects may also play a role in the efficacy of a given IL. For example, single cellulose chains explore greater conformational variability in some ILs than they do in H₂O (REF.

97), and this entropic difference can see the dissolution of cellulose in these ILs be more favorable. However, not all ILs are created equal — while MD simulations and X-ray scattering show that cellulose has greater conformational variability in [bmim]Cl than in H₂O (REF. 99), when in other ILs cellulose molecules can be more rigid and exist as rod-like polymers^{98, 100, 101}.

[H2] Length scale considerations.

Scaling up from molecular driving forces all the way to modeling cell-wall ultrastructure is challenging. However, in some cases multiscale rationalization is possible⁴³. For example, the behavior of large polymer chains can sometimes be predicted by considering the properties of oligomeric forms. There is controversy surrounding the degree of polymerization (N) and branching of lignins. N values depend on the plant source and the values reported range from ~ 10 to at least ~ 60 ,¹⁰² in addition to lignin being found in both branched¹⁰³ and linear¹⁰⁴ forms. MD calculations have suggested that at room temperature in H₂O, a qualitative transition in the dependence of the shape of lignin on length occurs at $N \approx 15$ (REF. 59), below which the molecule is less spherical. Chain-length dependent thermodynamic competition determines this change. For small N , the favorable lignin–H₂O interaction dominates and causes lignin to adopt elongated shapes. For larger N , the lignin self-interaction and the increase in entropy of liberated H₂O — both favored by surface area minimization — dominate, causing lignin to assume more spherical shapes. Scaling up further brings us to a result we described earlier: the bad nature of H₂O as a solvent sees lignin aggregate (FIG. 2). Neutron scattering experiments and MD simulations⁴⁹ have revealed these aggregates to have highly folded surfaces⁴⁹, described by a surface fractal dimension d_s invariant under change of scale from 1–1000 Å. d_s is a measure of the roughness or irregularity of a surface, taking values between 2 for a smooth surface (*e.g.* $d_s = 2.07$ for graphite) and 3 for a rough surface (*e.g.* $d_s = 2.90$ for carbonate rock): lignin was found to have $d_s = 2.62 \pm 0.02$. Non-spherical aggregate shapes are adopted because lignin's polar OH groups interact favorably with H₂O, thus reducing the need to minimize the lignin–H₂O interfacial area. Thus, a detailed multiscale picture now exists of the $\sim \mu\text{m}$ lignin aggregates that have been observed in electron micrographs to coalesce on cellulose surfaces during various pretreatments¹⁰⁵.

[H1] Conclusion and outlook

Plant cell walls are biosynthesized in recalcitrant metastable states. Solvation thermodynamics, in particularly H₂O entropy, is a major driving force behind stabilizing these states. During pretreatment biomass can escape these metastable states and undergo changes in structure. Counteracting the hydration entropy effect by modifying the solvent quality can enable us to develop an effective pretreatment strategy. An ideal pretreatment solvent would be good or θ for all biomass components, would decrystallize cellulose fibers, dissociate lignin aggregates, expand lignin molecules and fractionate the disassembled biopolymers. Ensuring that the amphiphilicities of co-solvents and biomass are complementary is an important principle guiding us in this direction, and allows one to disrupt both intrapolymeric hydrogen bonds and associations between hydrophobic surfaces.

Our discussion here has been limited to the physical effects at play in biomass dissolution, but we must also have an understanding of chemical reactivity and how the two are interdependent. Gaining such an understanding will require the relevant 3D structures to be determined in order to tell us what reactive groups are exposed to solvent. The solvation of these groups must be known at an atomistic level, such that we can then consider which chemical reactions can take place.

Several examples of relevant differential solvent effects exist. For example, the rate of xylose dehydration to furfural is 40 times faster in GVL–H₂O than it is in pure H₂O (REF. 108). Also, the different acid-catalyzed biomass reaction rates in different solvent environments have been attributed to the formation and properties of H₂O-rich or H₂O-deficient local domains near the reactants¹⁰⁹. Rates of acid-catalysed reactions in the liquid phase can be enhanced by altering the extents of solvation of the initial and transition states (REF 109-1). Useful in this regard are studies on biomass pyrolysis using coupled MD and quantum chemical calculations¹¹⁰ or coupled chemical kinetics and transport model approaches¹¹¹. A further example of coupling of physical and chemical changes is cleavage of lignin–hemicellulose covalent bonds that might lead to the phase separation of lignin–hemicellulose occurring at lower temperatures.

Mild pretreatment methods that limit chemical changes are desirable¹¹² because they can deliver¹¹³ both high-purity sugar and native-type lignin streams in biorefining applications. The optimal pretreatment design will combine the physical changes we have described with specific chemical reactions that give a certain small-molecule products¹¹⁴. It may also be possible to judiciously choose solvents that optimize the timing of different biomass deconstruction processes so as to synchronize product formation in a ‘one-pot’ process¹⁶. The challenge for pretreatment is thus further shaped and the common molecular-level driving forces discussed here provide a basis for rationalizing the design of ever more efficient and economic lignocellulose deconstruction procedures.

[H1] References

- 0-1. Cosgrove, D.J. & Jarvis, M.C. Comparative structure and biomechanics of plant primary and secondary cell walls. *Front. Plant Sci.* **3**, 204 (2012).
- 0-2. Burton, R.A., Gidley, M.J. & Fincher, G.B. Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat. Chem. Biol.* **6**, 724-732 (2010).
- 0-3. Himmel, M.E. et al. Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* **315**, 804-807 (2007).
- 0-4. Meng, X.Z. & Ragauskas, A.J. Recent advances in understanding the role of cellulose accessibility in enzymatic hydrolysis of lignocellulosic substrates. *Curr. Opin. Biotechnol.* **27**, 150-158 (2014).
1. Meng, X. et al. Physicochemical structural changes of poplar and switchgrass during biomass pretreatment and enzymatic hydrolysis. *ACS Sustainable Chem. Eng.* **4**, 4563–4572 (2016).
2. Sun, Z. et al. Complete lignocellulose conversion with integrated catalyst recycling yielding valuable aromatics and fuels. *Nat. Catal.* **1**, 82–92 (2018).
3. Ragauskas, A. J. et al. Lignin valorization: improving lignin processing in the biorefinery. *Science* **344**, 1246843 (2014).
4. Chen, H. et al. A review on the pretreatment of lignocellulose for high-value chemicals. *Fuel Process. Technol.* **160**, 196–206 (2017).
5. Balch, M. L. et al. Lignocellulose fermentation and residual solids characterization for senescent switchgrass fermentation by *Clostridium thermocellum* in the presence and absence of continuous *in situ* ball-milling. *Energy Environ. Sci.* **10**, 1252–1261 (2017).

6. Kan, T., Strezov, V. & Evans, T. J. Lignocellulosic biomass pyrolysis: A review of product properties and effects of pyrolysis parameters. *Renewable Sustainable Energy Rev.* **57**, 1126–1140 (2016).
7. da Costa Sousa, L. et al. Next-generation ammonia pretreatment enhances cellulosic biofuel production. *Energy Environ. Sci.* **9**, 1215–1223 (2016).
- 7-1. Chundawat, S.P.S. et al. Multi-scale visualization and characterization of lignocellulosic plant cell wall deconstruction during thermochemical pretreatment. *Energy Environ. Sci.* **4**, 973–984 (2011).
8. George, A. et al. Design of low-cost ionic liquids for lignocellulosic biomass pretreatment. *Green Chem.* **17**, 1728–1734 (2015).
9. Socha, A. M. et al. Efficient biomass pretreatment using ionic liquids derived from lignin and hemicellulose. *Proc. Natl Acad. Sci. USA* **111**, E3587–E3595 (2014).
10. Kumar, A. K., Parikh, B. S. & Pravakar, M. Natural deep eutectic solvent mediated pretreatment of rice straw: bioanalytical characterization of lignin extract and enzymatic hydrolysis of pretreated biomass residue. *Environ. Sci. Pollut. Res.* **23**, 9265–9275 (2016).
11. Zhang, K., Pei, Z. & Wang, D. Organic solvent pretreatment of lignocellulosic biomass for biofuels and biochemicals: a review. *Bioresour. Technol.* **199**, 21–33 (2016).
12. Agbor, V. B., Cicek, N., Sparling, R., Berlin, A. & Levin, D. B. Biomass pretreatment: fundamentals toward application. *Biotechnol. Adv.* **29**, 675–685 (2011).
13. Vanholme, R., Demedts, B., Morreel, K., Ralph, J. & Boerjan, W. Lignin biosynthesis and structure. *Plant Physiol.* **153**, 895–905 (2010).
14. Park, Y. B. & Cosgrove, D. J. Xyloglucan and its interactions with other components of the growing cell wall. *Plant Cell Physiol.* **56**, 180–194 (2015).
15. *Polymer Physics* (Eds Rubinstein, M. & Colby, R. H.) Vol. 23 (Oxford University Press, Oxford, 2003).
16. Smith, M. D., Cai, C. M., Cheng, X., Petridis, L. & Smith, J. C. Temperature-dependent phase behaviour of tetrahydrofuran–water alters solubilization of xylan to improve co-production of furfurals from lignocellulosic biomass. *Green Chem.* **20**, 1612–1620 (2018).
17. Muller, F. et al. SANS measurements of semiflexible xyloglucan polysaccharide chains in water reveal their self-avoiding statistics. *Biomacromolecules* **12**, 3330–3336 (2011).
18. Zhang, T., Vavylonis, D., Durachko, D. M. & Cosgrove, D. J. Nanoscale movements of cellulose microfibrils in primary cell walls. *Nat. Plants* **3**, 17056 (2017).
- 18-1. Guerriero, G., Fugelstad, J. & Bulone, V. What Do We Really Know about Cellulose Biosynthesis in Higher Plants? *J. Integr. Plant Biol.* **52**, 161–175 (2010).
19. Somerville, C. Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* **22**, 53–78 (2006).
- 19-1. Sethaphong, L. et al. Tertiary model of a plant cellulose synthase. *P. Natl. Acad. Sci. USA* **110**, 7512–7517 (2013).
20. McNamara, J. T., Morgan, J. L. W. & Zimmer, J. A molecular description of cellulose biosynthesis. *Annu. Rev. Biochem.* **84**, 895–921 (2015).
21. Cosgrove, D. J. Re-constructing our models of cellulose and primary cell wall assembly. *Curr. Opin. Plant Biol.* **22**, 122–131 (2014).
22. Ding, S.-Y. & Himmel, M. E. The maize primary cell wall microfibril: a new model derived from direct visualization. *J. Agric. Food Chem.* **54**, 597–606 (2006).

23. Vandavasi, V. G. et al. A structural study of CESA1 catalytic domain of arabidopsis cellulose synthesis complex: evidence for CESA trimers. *Plant Physiol.* **170**, 123–135 (2016).
24. Newman, R. H., Hill, S. J. & Harris, P. J. Wide-angle X-Ray scattering and solid-state nuclear magnetic resonance data combined to test models for cellulose microfibrils in mung bean cell walls. *Plant Physiol.* **163**, 1558–1567 (2013).
25. Wang, T. & Hong, M. Solid-state NMR investigations of cellulose structure and interactions with matrix polysaccharides in plant primary cell walls. *J. Exp. Bot.* **67**, 503–514 (2016).
26. Nishiyama, Y., Sugiyama, J., Chanzy, H. & Langan, P. Crystal structure and hydrogen bonding system in cellulose I_α from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **125**, 14300–14306 (2003).
27. Nishiyama, Y., Langan, P. & Chanzy, H. Crystal structure and hydrogen-bonding system in cellulose I_α from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **124**, 9074–9082 (2002).
28. Medronho, B., Romano, A., Miguel, M. G., Stigsson, L. & Lindman, B. Rationalizing cellulose (in)solubility: reviewing basic physicochemical aspects and role of hydrophobic interactions. *Cellulose* **19**, 581–587 (2012).
29. Moon, R. J., Martini, A., Nairn, J., Simonsen, J. & Youngblood, J. Cellulose nanomaterials review: structure, properties and nanocomposites. *Chem. Soc. Rev.* **40**, 3941–3994 (2011).
30. Usov, I. et al. Understanding nanocellulose chirality and structure–properties relationship at the single fibril level. *Nat. Commun.* **6**, 7564 (2015).
31. Bergenstråhle, M., Wohler, J., Himmel, M. E. & Brady, J. W. Simulation studies of the insolubility of cellulose. *Carbohydr. Res.* **345**, 2060–2066 (2010).
32. Kamide, K., Saito, M. & Suzuki, H. Persistence length of cellulose and cellulose derivatives in solution. *Makromol. Chem., Rapid Commun.* **4**, 33–39 (1983).
33. Kroon-Batenburg, L. M. J., Kruiskamp, P. H., Vliegthart, J. F. G. & Kroon, J. Estimation of the persistence length of polymers by MD simulations on small fragments in solution. Application to cellulose. *J. Phys. Chem. B* **101**, 8454–8459 (1997).
34. Zhao, Z. et al. Cellulose microfibril twist, mechanics, and implication for cellulose biosynthesis. *J. Phys. Chem. A* **117**, 2580–2589 (2013).
35. Bu, L., Himmel, M. E. & Crowley, M. F. The molecular origins of twist in cellulose I-beta. *Carbohydr. Polym.* **125**, 146–152 (2015).
36. Hadden, J. A., French, A. D. & Woods, R. J. Unraveling cellulose microfibrils: a twisted tale. *Biopolymers* **99**, 746–756 (2013).
37. Gross, A. S. & Chu, J.-W. On the molecular origins of biomass recalcitrance: the interaction network and solvation structures of cellulose microfibrils. *J. Phys. Chem. B* **114**, 13333–13341 (2010).
38. Gross, A. S., Bell, A. T. & Chu, J. W. Entropy of cellulose dissolution in water and in the ionic liquid 1-butyl-3-methylimidazolium chloride. *Phys. Chem. Chem. Phys.* **14**, 8425–8430 (2012).
39. Miyamoto, H., Schnupf, U. & Brady, J. W. Water structuring over the hydrophobic surface of cellulose. *J. Agric. Food Chem.* **62**, 11017–11023 (2014).
40. Petridis, L. et al. Hydration control of the mechanical and dynamical properties of cellulose. *Biomacromolecules* **15**, 4152–4159 (2014).

41. Phyto, P., Wang, T., Yang, Y., O'Neill, H. & Hong, M. Direct determination of hydroxymethyl conformations of plant cell wall cellulose using ^1H polarization transfer solid-state NMR. *Biomacromolecules* **19**, 1485–1497 (2018).
42. Kanchanalai, P., Temani, G., Kawajiri, Y. & Realff, M. J. Reaction kinetics of concentrated-acid hydrolysis for cellulose and hemicellulose and effect of crystallinity. *Bioresources* **11**, 1672–1689 (2016).
43. Zhao, W. et al. From lignin subunits to aggregates: insights into lignin solubilization. *Green Chem.* **19**, 3272–3281 (2017).
44. Perras, F. A. et al. Atomic-level structure characterization of biomass pre-and post-lignin treatment by dynamic nuclear polarization-enhanced solid-state NMR. *J. Phys. Chem. A* **121**, 623–630 (2017).
45. Sangha, A. K. et al. Chemical factors that control lignin polymerization. *J. Phys. Chem. B* **118**, 164–170 (2014).
46. Ziebell, A. et al. Increase in 4-coumaryl alcohol Units during lignification in alfalfa (*Medicago sativa*) alters the extractability and molecular weight of lignin. *J. Biol. Chem.* **285**, 38961–38968 (2010).
47. Petridis, L., Schulz, R. & Smith, J. C. Simulation analysis of the temperature dependence of lignin structure and dynamics. *J. Am. Chem. Soc.* **133**, 20277–20287 (2011).
48. Ratnaweera, D. R. et al. The impact of lignin source on its self-assembly in solution. *RSC Adv.* **5**, 67258–67266 (2015).
49. Petridis, L. et al. Self-similar multiscale structure of lignin revealed by neutron scattering and molecular dynamics simulation. *Phys. Rev. E* **83**, 061911 (2011).
50. Grosberg, A. Y., Nechaev, S. K. & Shakhnovich, E. I. The role of topological constraints in the kinetics of collapse of macromolecules. *J. Phys.* **49**, 2095–2100 (1988).
51. Silveira, R. L., Stoyanov, S. R., Gusarov, S., Skaf, M. S. & Kovalenko, A. supramolecular interactions in secondary Plant cell walls: effect of lignin chemical composition revealed with the molecular theory of solvation. *J. Phys. Chem. Lett.* **6**, 206–211 (2015).
52. Athawale, M. V., Goel, G., Ghosh, T., Truskett, T. M. & Garde, S. Effects of lengthscales and attractions on the collapse of hydrophobic polymers in water. *Proc. Natl Acad. Sci. USA* **104**, 733–738 (2007).
53. Smith, M. D. et al. Cosolvent pretreatment in cellulosic biofuel production: effect of tetrahydrofuran–water on lignin structure and dynamics. *Green Chem.* **18**, 1268–1277 (2016).
54. Langan, P. et al. Common processes drive the thermochemical pretreatment of lignocellulosic biomass. *Green Chem.* **16**, 63–68 (2014).
55. Pingali, S. V. et al. Morphological changes in the cellulose and lignin components of biomass occur at different stages during steam pretreatment. *Cellulose* **21**, 873–878 (2014).
56. Nishiyama, Y., Langan, P., O'Neill, H., Pingali, S. V. & Harton, S. Structural coarsening of aspen wood by hydrothermal pretreatment monitored by small- and wide-angle scattering of X-rays and neutrons on oriented specimens. *Cellulose* **21**, 1015–1024 (2014).
57. Driemeier, C., Mendes, F. M., Santucci, B. S. & Pimenta, M. T. B. Cellulose co-crystallization and related phenomena occurring in hydrothermal treatment of sugarcane bagasse. *Cellulose* **22**, 2183–2195 (2015).
58. Silveira, R. L., Stoyanov, S. R., Kovalenko, A. & Skaf, M. S. Cellulose Aggregation under Hydrothermal Pretreatment Conditions. *Biomacromolecules* **17**, 2582–2590 (2016).

59. Petridis, L. & Smith, J. C. Conformations of Low-Molecular-Weight Lignin Polymers in Water. *ChemSusChem* **9**, 289–295 (2016).
60. Li, W. et al. Rapid dissolution of lignocellulosic biomass in ionic liquids using temperatures above the glass transition of lignin. *Green Chem.* **13**, 2038–2047 (2011).
61. Hatakeyama, H. & Hatakeyama, T. Lignin structure, properties, and applications. *Adv. Polym. Sci.* **232**, 1–63 (2009).
62. Khodadadi, S. & Sokolov, A. P. Protein dynamics: from rattling in a cage to structural relaxation. *Soft Matter* **11**, 4984–4998 (2015).
63. Vural, D. et al. Impact of hydration and temperature history on the structure and dynamics of lignin. *Green Chem.* **20**, 1602–1611 (2018).
64. Vural, D., Smith, J. C. & Petridis, L. Dynamics of the lignin glass transition. *Phys. Chem. Chem. Phys.* **20**, 20504–20512 (2018).
65. Li, H., Pu, Y., Kumar, R., Ragauskas, A. J. & Wyman, C. E. Investigation of lignin deposition on cellulose during hydrothermal pretreatment, its effect on cellulose hydrolysis, and underlying mechanisms. *Biotechnol. Bioeng.* **111**, 485–492 (2014).
66. Pingali, S. V. et al. Breakdown of cell wall nanostructure in dilute acid pretreated biomass. *Biomacromolecules* **11**, 2329–2335 (2010).
67. Selig, M. J. et al. Deposition of lignin droplets produced during dilute acid pretreatment of maize stems retards enzymatic hydrolysis of cellulose. *Biotechnol. Prog.* **23**, 1333–1339 (2007).
68. Gao, X. et al. Comparison of enzymatic reactivity of corn stover solids prepared by dilute acid, AFEXTM, and ionic liquid pretreatments. *Biotechnol. Biofuels* **7**, 71 (2014).
69. Lindner, B., Petridis, L., Schulz, R. & Smith, J. C. Solvent-driven preferential association of lignin with regions of crystalline cellulose in molecular dynamics Simulation. *Biomacromolecules* **14**, 3390–3398 (2013).
70. Nakagame, S., Chandra, R. P., Kadla, J. F. & Saddler, J. N. Enhancing the enzymatic hydrolysis of lignocellulosic biomass by increasing the carboxylic acid content of the associated lignin. *Biotechnol. Bioeng.* **108**, 538–548 (2011).
71. Rahikainen, J. et al. Inhibition of enzymatic hydrolysis by residual lignins from softwood—study of enzyme binding and inactivation on lignin-rich surface. *Biotechnol. Bioeng.* **108**, 2823–2834 (2011).
72. Sammond, D. W. et al. Predicting enzyme adsorption to lignin films by calculating enzyme surface hydrophobicity. *J. Biol. Chem.* **289**, 20960–20969 (2014).
73. Liu, Y. S. et al. Cellobiohydrolase hydrolyzes crystalline cellulose on hydrophobic faces. *J. Biol. Chem.* **286**, 11195–11201 (2011).
74. Vermaas, J. V. et al. Mechanism of lignin inhibition of enzymatic biomass deconstruction. *Biotechnol. Biofuels* **8**, 217 (2015).
75. Luo, Y. et al. The production of furfural directly from hemicellulose in lignocellulosic biomass: a review. *Catal. Today* <https://doi.org/10.1016/j.cattod.2018.06.042> (2018).
76. Jacobsen, S. E. & Wyman, C. E. Cellulose and hemicellulose hydrolysis models for application to current and novel pretreatment processes. *Appl. Biochem. Biotechnol.* **84**, 81–96 (2000).
77. Pereira, C. S., Silveira, R. L., Dupree, P. & Skaf, M. S. Effects of xylan side-chain substitutions on xylan–cellulose interactions and implications for thermal pretreatment of cellulosic biomass. *Biomacromolecules* **18**, 1311–1321 (2017).

78. Grantham, N. J. et al. An even pattern of xylan substitution is critical for interaction with cellulose in plant cell walls. *Nat. Plants* **3**, 859–865 (2017).
79. Kumar, R. et al. Cellulose–hemicellulose interactions at elevated temperatures increase cellulose recalcitrance to biological conversion. *Green Chem.* **20**, 921–934 (2018).
80. Nguyen, T. Y., Cai, C. M., Kumar, R. & Wyman, C. E. Co-solvent pretreatment reduces costly enzyme requirements for high sugar and ethanol yields from lignocellulosic biomass. *ChemSusChem* **8**, 1716–1725 (2015).
81. Cai, C. M., Nagane, N., Kumar, R. & Wyman, C. E. Coupling metal halides with a co-solvent to produce furfural and 5-HMF at high yields directly from lignocellulosic biomass as an integrated biofuels strategy. *Green Chem.* **16**, 3819–3829 (2014).
82. Cai, C. M., Zhang, T., Kumar, R. & Wyman, C. THF co-solvent enhances hydrocarbon fuel precursor yields from lignocellulosic biomass. *Green Chem.* **15**, 3140–3145 (2013).
83. Shuai, L. & Luterbacher, J. Organic solvent effects in biomass conversion reactions. *ChemSusChem* **9**, 133–155 (2016).
84. Sannigrahi, P., Kim, D. H., Jung, S. & Ragauskas, A. Pseudo-lignin and pretreatment chemistry. *Energy Environ. Sci.* **4**, 1306–1310 (2011).
85. Smith, M. D., Cheng, X., Petridis, L., Mostofian, B. & Smith, J. C. Organosolv–water cosolvent phase separation on cellulose and its influence on the physical deconstruction of cellulose: a molecular dynamics Analysis. *Sci. Rep.* **7**, 14494 (2017).
86. Mostofian, B. et al. Local phase separation of co-solvents enhances pretreatment of biomass for bioenergy applications. *J. Am. Chem. Soc.* **138**, 10869–10878 (2016).
87. Sun, J. et al. One-pot integrated biofuel production using low-cost biocompatible protic ionic liquids. *Green Chem.* **19**, 3152–3163 (2017).
88. Remsing, R. C., Swatloski, R. P., Rogers, R. D. & Moyna, G. Mechanism of cellulose dissolution in the ionic liquid 1-*n*-butyl-3-methylimidazolium chloride: a ¹³C and ^{35/37}Cl NMR relaxation study on model systems. *Chem. Commun.* 1271–1273 (2006).
89. Swatloski, R. P., Spear, S. K., Holbrey, J. D. & Rogers, R. D. Dissolution of cellulose with ionic liquids. *J. Am. Chem. Soc.* **124**, 4974–4975 (2002).
90. Anderson, J. L., Ding, J., Welton, T. & Armstrong, D. W. Characterizing ionic liquids on the basis of multiple solvation interactions. *J. Am. Chem. Soc.* **124**, 14247–14254 (2002).
91. Pinkert, A., Marsh, K. N., Pang, S. S. & Staiger, M. P. Ionic liquids and their interaction with cellulose. *Chem. Rev.* **109**, 6712–6728 (2009).
92. Feng, L. & Chen, Z.-I. Research progress on dissolution and functional modification of cellulose in ionic liquids. *J. Mol. Liq.* **142**, 1–5 (2008).
93. Medronho, B. & Lindman, B. Brief overview on cellulose dissolution/regeneration interactions and mechanisms. *Adv. Colloid Interfac. Sci.* **222**, 502–508 (2015).
94. Lindman, B., Karlström, G. & Stigsson, L. On the mechanism of dissolution of cellulose. *J. Mol. Liq.* **156**, 76–81 (2010).
95. Mostofian, B., Smith, J. C. & Cheng, X. Simulation of a cellulose fiber in ionic liquid suggests a synergistic approach to dissolution. *Cellulose* **21**, 983–997 (2014).
96. Li, Y. et al. Dissolving process of a cellulose bunch in ionic liquids: a molecular dynamics study. *Phys. Chem. Chem. Phys.* **17**, 17894–17905 (2015).
97. Liu, H., Sale, K. L., Holmes, B. M., Simmons, B. A. & Singh, S. Understanding the interactions of cellulose with ionic liquids: a molecular dynamics study. *J. Phys. Chem. B* **114**, 4293–4301 (2010).

98. Jiang, X., Kitamura, S., Sato, T. & Terao, K. Chain dimensions and stiffness of cellulosic and amylosic chains in an ionic liquid: cellulose, amylose, and an amylose carbamate in BmimCl. *Macromolecules* **50**, 3979–3984 (2017).
99. Mostofian, B., Cheng, X. & Smith, J. C. Replica-exchange molecular dynamics simulations of cellulose solvated in water and in the ionic liquid 1-butyl-3-methylimidazolium chloride. *J. Phys. Chem. B* **118**, 11037–11049 (2014).
100. Nakamura, Y. & Norisuye, T. Brush-Like Polymers. in *Soft Matter Characterization* (eds. Borsali, R. & Pecora, R.) 235–286 (Springer, Dordrecht, 2008).
101. Hirosawa, K., Fujii, K., Hashimoto, K. & Shibayama, M. Solvated structure of cellulose in a phosphonate-based ionic liquid. *Macromolecules* **50**, 6509–6517 (2017).
102. Tolbert, A., Akinoshio, H., Khunsupat, R., Naskar, A. K. & Ragauskas, A. J. Characterization and analysis of the molecular weight of lignin for biorefining studies. *Biofuels, Bioprod. Bioref.* **8**, 836–856 (2014).
103. Ralph, J. et al. Lignins: natural polymers from oxidative coupling of 4-hydroxyphenylpropanoids. *Phytochem. Rev.* **3**, 29–60 (2004).
104. Crestini, C., Melone, F., Sette, M. & Saladino, R. Milled wood lignin: a linear oligomer. *Biomacromolecules* **12**, 3928–3935 (2011).
105. Donohoe, B. S., Decker, S. R., Tucker, M. P., Himmel, M. E. & Vinzant, T. B. Visualizing Lignin Coalescence and Migration Through Maize Cell Walls Following Thermochemical Pretreatment. *Biotechnol. Bioeng.* **101**, 913–925 (2008).
106. Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature* **437**, 640–647 (2005).
107. Sumi, T. & Sekino, H. Integral equation study of hydrophobic interaction: a comparison between the simple point charge model for water and a Lennard-Jones model for solvent. *J. Chem. Phys.* **126**, 144508 (2007).
108. Mellmer, M. A. et al. Solvent effects in acid-catalyzed biomass conversion reactions. *Angew. Chem. Int. Ed.* **53**, 11872–11875 (2014).
109. Walker, T. W. et al. Universal kinetic solvent effects in acid-catalyzed reactions of biomass-derived oxygenates. *Energy Environ. Sci.* **11**, 617–628 (2018).
- 109-1. Mellmer, M.A. et al. Solvent-enabled control of reactivity for liquid-phase reactions of biomass-derived compounds. *Nat. Catal.* **1**, 199–207 (2018).
110. Zheng, M. et al. Initial reaction mechanisms of cellulose pyrolysis revealed by ReaxFF molecular dynamics. *Fuel* **177**, 130–141 (2016).
111. Di Blasi, C. Modeling chemical and physical processes of wood and biomass pyrolysis. *Prog. Energy Combust. Sci.* **34**, 47–90 (2008).
112. Rahimi, A., Ulbrich, A., Coon, J. J. & Stahl, S. S. Formic-acid-induced depolymerization of oxidized lignin to aromatics. *Nature* **515**, 249–252 (2014).
113. Das, A. et al. Lignin conversion to low-molecular-weight aromatics via an aerobic oxidation-hydrolysis sequence: comparison of different lignin sources. *ACS Sustainable Chem. Eng.* **6**, 3367–3374 (2018).
114. Thomas, V. A. et al. Adding tetrahydrofuran to dilute acid pretreatment provides new insights into substrate changes that greatly enhance biomass deconstruction by *Clostridium thermocellum* and fungal enzymes. *Biotechnol. Biofuels* **10**, 252 (2017).

[H1] Acknowledgements

This research was supported by the Genomic Science Program, Office of Biological and Environmental Research, U. S. Department of Energy (DOE), under Contract FWP ERKP752. This research used resources of three User Facilities supported by DOE: NERSC (contract No. DE-AC02-05CH11231), HFIR/SNS (DE-AC02-05CH11231) and OLCF (contract No. DE-AC05-00OR22725).

[H1] Author contributions

All authors contributed equally to the preparation of this manuscript.

[H1] Competing interests

The authors declare no competing interests.

[H1] Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[H1] How to cite this article

Petridis, L. & Smith, J. C. Molecular-level driving forces in lignocellulosic biomass deconstruction. *Nat. Rev. Chem.* **2**, xxxx (2018).

[H1] 45 word summary

Physical molecular driving forces are described that stabilize native lignocellulosic plant biomass structures and govern thermochemical biomass pretreatments. This understanding can be used in the design of methods for efficiently deconstructing biomass for biofuels and other bioproducts.

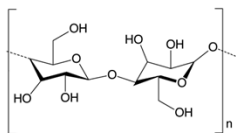
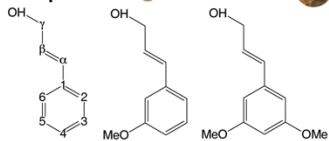
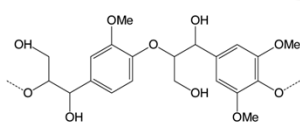
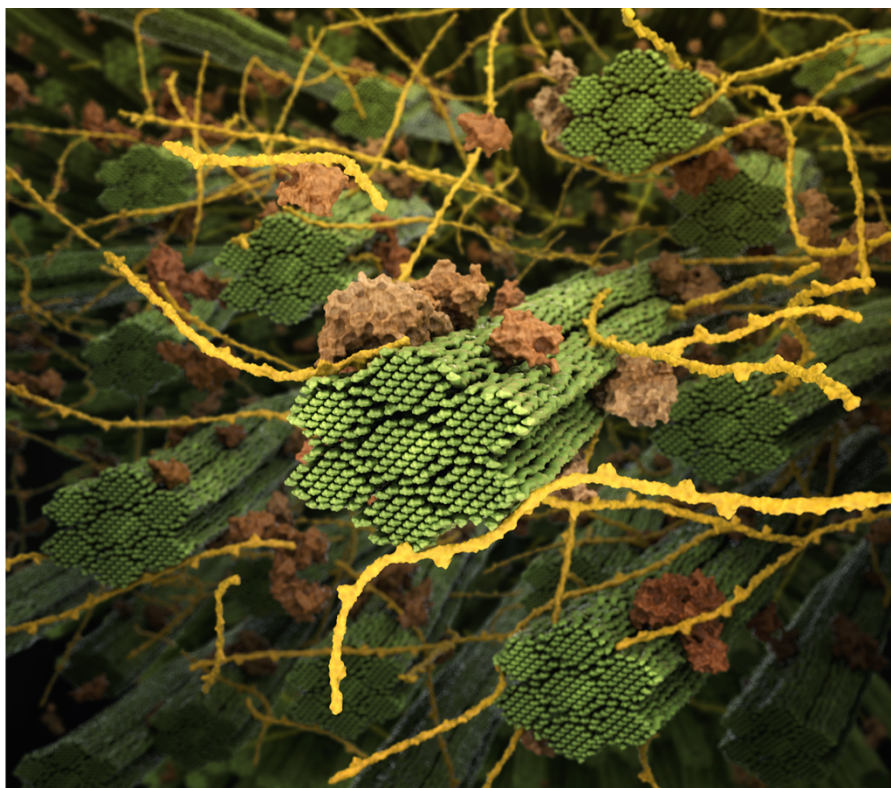
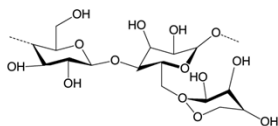
Aa | cellulose**Ab | monolignols****Ac | lignin****Ad | hemicellulose**

Figure 1 | The structure of lignocellulosic biomass and its components. **Aa** | Cellulose consists of unbranched polysaccharide chains, with β -(1 $\rightarrow$ 4) linkages between D-glucose units affording a polymer referred to as β -1,4-glucan. **Ab** | In contrast, hemicelluloses are often branched and contain more than one type of residue. The example here is a xyloglucan, in which the monomers in β -1,4-glucan often feature α -(1 $\rightarrow$ 6) linkages to xylose. In turn, these sidechains can be further decorated with β -(1 $\rightarrow$ 2) linkages to galactose¹⁴. **Ac** | *p*-Coumaryl (**1**), coniferyl (**2**) and sinapyl (**3**) alcohols differ in their degree of methoxylation. These three predominant monolignols give rise to the hydroxyphenyl (H), guaiacyl (G), and syringyl (S) units that make up lignin. **Ad** | Lignin is a cross-linked amorphous polymer¹³. In this example, G units are linked through β -O-4' and 5-5' linkages. **B** | Native biomass is a complex material comprising the components in **A**. The simulation-inspired structure (REF. 54) here shows the colour-coded cellulose (green, with seven microfibrils depicted as forming a fibre), hemicellulose (yellow) and lignin (brown) domains.

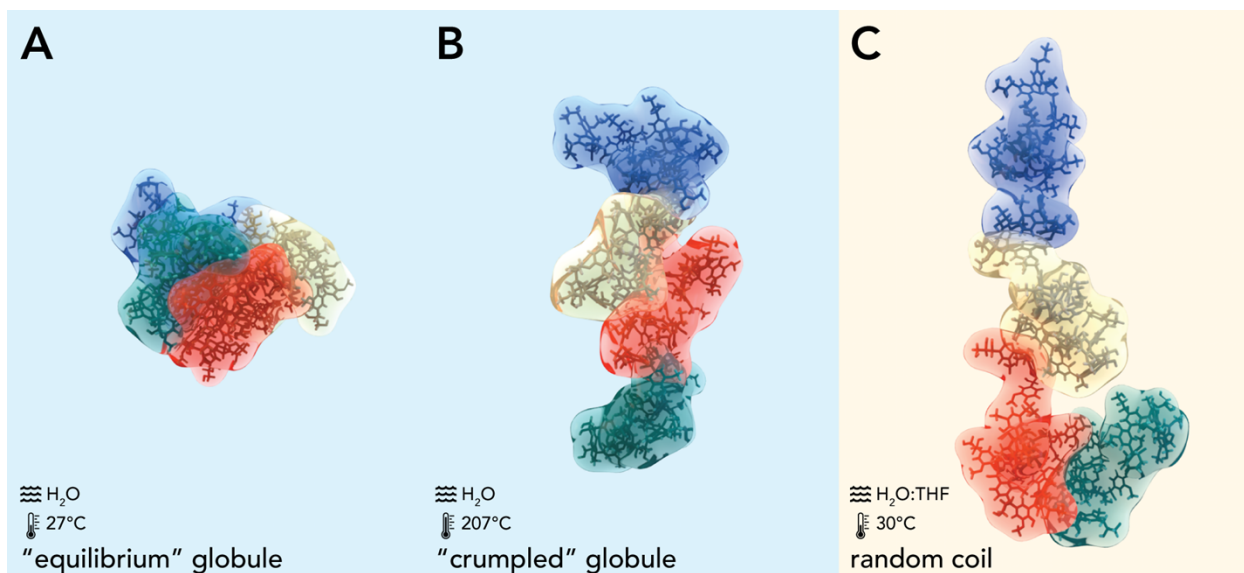


Figure 2 | **Molecular dynamics models of three states of lignin**^{47,53}. Softwood lignin can be modeled as multiple 'blobs', each of which is a ~15 monolignol oligomer. The blobs can interpenetrate and are coloured differently. **a** | In H_2O at 27 °C lignin exists as compact, spherical 'equilibrium' globules comprising several blobs. **b** | In H_2O above ~147-207 °C (temperatures typical for biomass pretreatment) the solvent remains poor but lignin assumes more extended, aspherical 'crumpled' globular forms with more separated blobs. **C** | THF/ H_2O (in 1:1 and 1:2 THF: H_2O v/v ratios) is a θ solvent in which lignin adopts random coil configurations, with the blobs are more extended, such that the lignin chain is exposed to the solvent along its entire contour. Figure drawn using data from REFs 16 & 47.

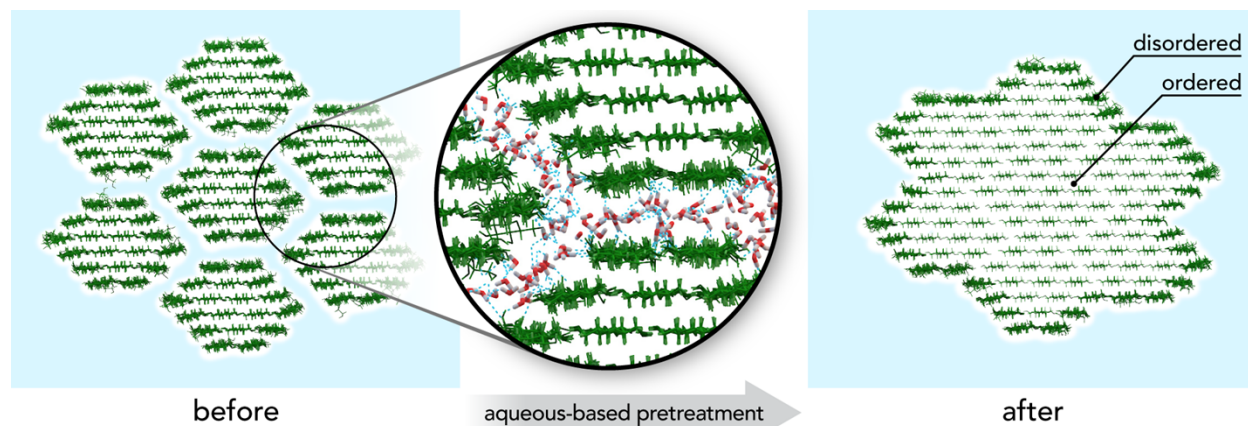


Figure 3 | **Steam explosion pretreatment causes cellulose fibrils to coalesce**. Prior to pretreatment, cellulose fibrils (green) are separated by layers of H_2O (blue background). These hydration layers are ordered relative to bulk H_2O but their presence induces the cellulose surface to become disordered (see inset; hydrogen bonds are shown as blue dashed lines). At the high pretreatment temperatures, H_2O is released and the cellulose fibrils coalesce. This model was based on experimental data and simulations.⁵⁴

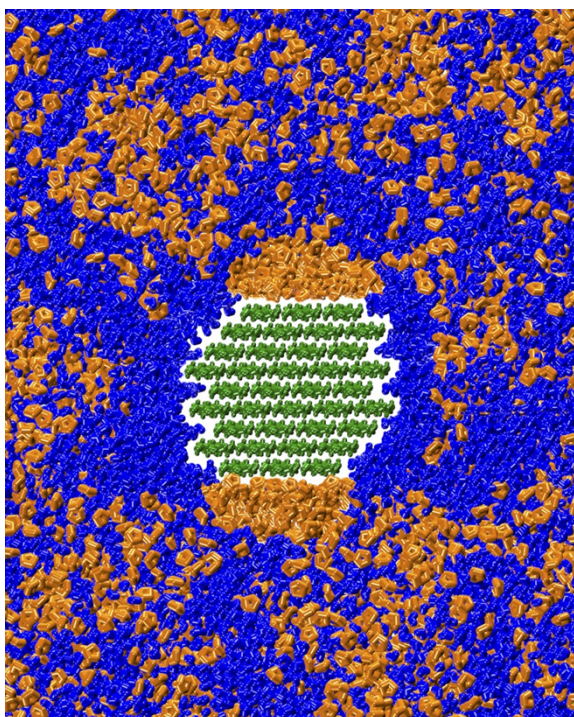


Figure 4 | **A molecular dynamics snapshot of cellulose in THF–H₂O.** THF (orange) associates with the hydrophobic surfaces of cellulose (green), and H₂O (blue) with the hydrophilic surfaces, while the bulk co-solvent is mixed. Figure adapted with permission from Ref. 86, American Chemical Society.