CONDENSED
MATTER
PHYSICS

Review

Multiscale theory, modelling, and simulation of hemicellulose and lignin in solution*

A. Kovalenko [ID] †

Software for Multiscale Modelling, Inc., Edmonton, Alberta, Canada T6E 5J5

This review examines multiscale modelling approaches for cellulose nanocrystals (CNCs) and lignocellulosic plant cell walls, with a focus on hemicellulose and lignin interactions in aqueous environments. The three-dimensional reference interaction site model with the Kovalenko–Hirata closure (3D-RISM-KH) is highlighted as a powerful molecular solvation theory applied in nanochemistry and biomolecular simulations. The method has been successfully employed to investigate hemicellulose hydrogels, the influence of hemicellulose composition on nanoscale forces in primary cell walls, and lignin-lignin and lignin-hemicellulose interactions. Findings indicate that these interactions are predominantly hydrophobic and entropy-driven, arising from water exclusion effects. Insights gained through this modeling framework deepen the understanding of molecular-scale forces in plant cell walls and inform strategies for biomass valorization, including genetic engineering and pretreatment technologies aimed at enhancing cellulose extraction and utilization.

Key words: *cellulose nanocrystals, hemicellulose, lignin, molecular solvation theory, solvation thermodynamics, molecular simulations*

1. Introduction

Molecular interactions in nanochemical and biomolecular systems differ considerably from those in bulk phases due to the effects of size, shape, and chemical composition. Understanding these interactions is essential for designing advanced materials and improving bioprocessing strategies. Various theoretical models have been developed to describe solvation and interaction processes in liquid media, ranging from continuum solvation frameworks to quasichemical approaches. However, these models often face limitations in system size and accuracy for complex biological and polymeric structures [1–9].

In the context of sustainable technologies, a major challenge lies in overcoming the intrinsic recalcitrance of plant biomass [10, 11]. Lignocellulosic materials possess a highly ordered and resistant architecture, where cellulose is closely associated with hemicellulose and lignin. These noncellulosic polymers act as protective matrices that restrict enzymatic access to cellulose, thereby limiting efficient deconstruction [12, 13]. Environmentally friendly pretreatment methods seek to fractionate lignocellulose into its major components while minimizing the energy input and chemical waste [14]. Emerging catalytic processes such as "lignin-first" biorefining have shown promise in selectively removing lignin while preserving carbohydrate fractions, enabling the generation of high-value bioproducts [15].

Multiscale frameworks combining molecular dynamics (MD), kinetic modelling, and machine learning are emerging to overcome the significant challenge of linking atomistic MD mechanisms to larger-scale process predictions [16]. Much attention has been paid to using density functional theory (DFT) and MD to understand the solubilization and conversion processes of the cellulose, hemicellulose, and lignin components of plant biomass to produce high value-added chemicals and fuels [17]. DFT is often used

---

*This paper is dedicated to the memory of Stefan Sokołowski, Professor at Maria Curie-Skłododowska University in Lublin, Poland, an extraordinary scientist in microscopic theory and computer simulations of fluids and solutions.

†Email: akovalenko@smmodeling.com

---

with model compounds (e.g., xylose for hemicellulose, guaiacyl dimers for lignin) because the full polymers are too complex for direct quantum treatment [17]. Simulations of model lignocellulosic mixtures of cellulose, hemicellulose, and lignin using MD and DFT showed that intercomponent interactions influence the decomposition dynamics, that is hydrogen bonding networks affect accessibility and reaction initiation [18]. DFT shows how solvent molecules or catalysts change the activation barriers, in particular, how energy barriers for certain lignin decomposition pathways are lowered by water/methanol [19]. Studies of base-catalyzed β-O-4 cleavage provide insight into selective depolymerization mechanisms relevant to chemical valorization [20]. DFT work on lignin dissolution in ionic liquids reveals hydrogen bonding interactions and structural effects relevant to pretreatment [21]. DFT analysis of xylose, xylobiose, xylan decomposition reveals preferred pathways to furans and small carbonyl compounds [22]. For hemicellulose, DFT elucidates initial bond cleavages and subsequent fragmentation sequences which are valuable for kinetic modelling and understanding the product formation [22]. MD studies of deep eutectic solvent (DES) interactions with hybrid systems of lignin and hemicellulose helped explain dissolution/pretreatment efficiency, especially via hydrogen bond analyses and interaction energies [23].

The reference interaction site model (RISM) [24–26] complemented by a variety of closure relations [3–9, 24–39] provides an alternative capable of delivering molecular-level insights into solvation phenomena with computational efficiency. Its three-dimensional extension (3D-RISM) [24, 25] calculates solvent site distributions around solute molecules, enabling a spatially resolved description of solvation structure. Among the closure relations, the Kovalenko–Hirata (KH) approximation [28] has proven especially effective across diverse molecular systems, ranging from small molecules [40] and biomacromolecules [41] to nanomaterials [42–45], drug binding [46, 47], ionic liquids [48], protein folding [49], and molecular recognition processes [50, 51]. Moreover, integration of 3D-RISM-KH with Density Functional Theory (DFT) [52] has extended its applicability to electronic structure calculations.

By applying 3D-RISM-KH modelling, significant advances have been made in clarifying intermolecular interactions within plant cell walls that were studied earlier in [53, 54]. These previous studies included enzymatic deconstruction pathways, supramolecular polymer self-assembly [55], and cooperative interactions within polysaccharides [56]. Such previous modelling has been instrumental in elucidating molecular recognition, carbohydrate–π, and π–π stacking interactions, which underlie biological signaling and structural stability [57, 58]. Of particular interest, glucuronoarabinoxylan hemicellulose, which is a major constituent of biofuel feedstocks such as corn and sugarcane has been studied for its role in modulating cell wall mechanics. Genetic modifications altering hemicellulose composition have demonstrated a potential for improving cellulose accessibility without impairing plant growth [13].

Recent 3D-RISM-KH investigations have provided detailed insights into CNC-hemicellulose interactions, showing that carboxylate groups significantly contribute to the reinforcement of primary cell walls [59]. Furthermore, studies have confirmed that lignin-lignin and lignin-hemicellulose associations are largely hydrophobic and entropy-driven [60]. These interactions are further stabilized by lignin methoxy groups, which strengthen the supramolecular matrix and thereby contribute to the persistence of biomass recalcitrance.

Collectively, multiscale modelling approaches such as 3D-RISM-KH offer a comprehensive understanding of the molecular determinants of cell wall integrity. This knowledge supports the development of improved pretreatment processes, genetic engineering strategies, and the design of bio-based materials with tailored mechanical properties.

2. Molecular theory of solvation for nanochemical and biomolecular systems in electrolyte solutions

2.1. Integral equations

The 3D-RISM-KH molecular theory of solvation [61–67] is based on the probability density function $g_\gamma(\mathbf{r})$, which describes the spatial likelihood of locating solvent site γ around a solute of arbitrary geometry. In solution, $g_\gamma(\mathbf{r}) > 1$ indicates enhanced density regions, while $g_\gamma(\mathbf{r}) < 1$ denotes depletion,

converging to unity in the bulk. This distribution is determined by solving the 3D-RISM integral equation

$$h_\gamma(\mathbf{r}) = \sum_\alpha \int d\mathbf{r}' g_\alpha(\mathbf{r} - \mathbf{r}') \chi_{\alpha\gamma}(r'), \tag{1}$$

complemented with the 3D-KH closure approximation

$$g_\gamma(\mathbf{r}) = \begin{cases} \exp\left(d_\gamma(\mathbf{r})\right) & \text{for } g_{\alpha\gamma}(\mathbf{r}) \leqslant 1, \\ 1 + d_\gamma(\mathbf{r}) & \text{for } g_\gamma(\mathbf{r}) > 1, \end{cases} \tag{2}$$

$$d_\gamma(\mathbf{r}) = -u_\gamma(\mathbf{r})/(k_{\mathrm{B}}T) + h_\gamma(\mathbf{r}) - c_\gamma(\mathbf{r}),$$

which are solved for the 3D site total correlation function $h_\gamma(\mathbf{r})$ and the 3D site direct correlation function $c_\gamma(\mathbf{r})$ that follows the asymptotic behavior $c_\gamma(\mathbf{r}) \sim -u_\gamma(\mathbf{r})/(k_{\mathrm{B}}T)$ related to the 3D solute-solvent interaction potential $u_\gamma(\mathbf{r})$. The 3D-KH closure effectively combines the Mean Spherical Approximation (MSA) and Hypernetted Chain (HNC) approximation, applying MSA in density-enhanced regions and HNC in density-depleted areas. Unlike the 3D-RISM-HNC model, which may overestimate associative interactions and introduce numerical instability, 3D-RISM-KH demonstrates robustness in electrolyte and strongly interacting systems.

The solvent site-site susceptibility function

$$\chi_{\alpha\gamma}(r) = \omega_{\alpha\gamma}(r) + \rho_\alpha h_{\alpha\gamma}(r) \tag{3}$$

comprises the intramolecular matrix describing the geometry of solvent molecules and the site-site total correlation function $h_{\alpha\gamma}(r)$ of bulk solvent at number site density ρ_α. These terms account for solvent molecular geometry and intermolecular correlations. The site-site total correlation functions are derived from the dielectrically consistent RISM (DRISM) theory of Perkyns and Pettitt [68] incorporating dielectric bridge corrections to ensure consistency with macroscopic dielectric properties, which is complemented with the site-site KH closure [25]. Such refinements make the 3D-RISM-KH theory broadly applicable to strongly associating species across varying temperatures, pressures, and solution compositions.

2.2. Solvation free energy and potential of mean force

Within the 3D-RISM-KH framework, solvation free energy of solute "u" in solvent "v" is expressed as an exact differential

$$\mu_{\mathrm{u}} = \epsilon_{\mathrm{u}} - TS_{\mathrm{V}} \tag{4}$$

derived via Kirkwood's thermodynamic integration as

$$\mu_{\mathrm{u}} = \sum_\gamma \int d\mathbf{r}\, \Phi_\gamma(\mathbf{r}), \tag{5a}$$

$$\Phi_\gamma(\mathbf{r}) = \rho_\gamma k_{\mathrm{B}}T \left(\frac{1}{2} h_\gamma^2(\mathbf{r}) \Theta\left(-h_\gamma(\mathbf{r})\right) - c_\gamma(\mathbf{r}) - \frac{1}{2} h_\gamma(\mathbf{r}) c_\gamma(\mathbf{r}) \right), \tag{5b}$$

where $\Phi_\gamma(\mathbf{r})$ denotes the 3D solvation free energy density, and $\Theta(x)$ is the Heaviside step function.

The solvation free energy is decomposed into the solute-solvent and solvent-solvent enthalpic contributions ϵ^{uv} and ϵ^{vv}, and the entropic term S_V,

$$\mu_{\mathrm{u}} = \epsilon^{\mathrm{uv}} + \epsilon^{\mathrm{vv}} - TS_{\mathrm{V}}, \tag{6a}$$

$$S_V = -\frac{1}{T} \left(\frac{\partial \mu}{\partial T} \right)_V, \tag{6b}$$

while the partial molar volume is obtained using Kirkwood-Buff theory [69] adapted for the 3D-RISM formalism as [70, 71]

$$\bar{V} = k_{\mathrm{B}}T\chi_T \left(1 - \sum_\gamma \rho_\gamma \int d\mathbf{r}\, c_\gamma(\mathbf{r}) \right), \tag{7}$$

and the isothermal compressibility of bulk solvent χ_T is obtained from DRISM-KH theory.

To mitigate systematic overestimation of hydration free energies, a universal correction is applied under ambient conditions [72]:

$$\mu_{\text{hydrated}}^{\text{3D−RISM−UC}} = \mu_{\text{hydrated}}^{\text{3D−RISM−KH}} + a\left(\rho\bar{V}\right) + b, \tag{8}$$

where the regression coefficients are $a = -3.312$ kcal/mol and $b = 1.152$ kcal/mol.

Furthermore, the potential of mean force (PMF) between nanoparticles can be computed by combining Lennard-Jones and electrostatic interactions with solvation free energy differences, offering a molecular-scale perspective on colloidal interactions in solution.

2.3. Efficient numerical solution with MDIIS

Solving the 3D-RISM-KH integral equations requires iterative numerical methods due to the complexity of three-dimensional correlation functions. The Modified Direct Inversion in the Iterative Subspace (MDIIS) algorithm [73] provides a computationally efficient and stable solver with quadratic convergence properties. It reduces residuals between the integral and closure equations while minimizing memory requirements. Comparable iterative solvers include Pulay's DIIS for Hartree-Fock equations [74], the Generalized Minimal Residual (GMRes) method [75], and Newton-Raphson approaches [76, 77]. The MDIIS solver has proven to provide highly effective quasi-quadratic convergence for large-scale nanochemical and biomolecular systems, enabling the accurate exploration of solvation structures and intermolecular forces.

3. Effect of hemicellulose composition on nanoscale forces governing cell wall strength in decomposing resilent cell wall plant structures

The efficient utilization of lignocellulose for producing second-generation biofuels and high-value chemicals requires the deconstruction of the highly resilient plant cell wall. The degree of recalcitrance differs significantly across plant species and phenotypes, largely due to variations in the chemical composition of the non-cellulosic matrix. Structural modifications in the composition and branching pattern of hemicellulose backbones can substantially alter the cell wall strength and microarchitecture.

Glucuronoarabinoxylan is the dominant hemicellulose in grasses widely employed for biofuel production, including sugarcane and corn. Its structure consists of a xylan backbone decorated with glucuronic acid and arabinose side chains, the distribution and frequency of which vary across plant genotypes [78]. These arabinose, glucuronic acid, and glucuronate substituents play a critical role in stabilizing primary cell walls by forming hydrogen-bonding interactions with cellulose surfaces. A molecular-level understanding of such interactions provides essential insights for genetic modification strategies and pretreatment methods aimed at enhancing biomass conversion efficiency.

To explore the role of hemicellulose in primary cell wall assembly, the three-dimensional reference interaction site model with the Kovalenko-Hirata closure (3D-RISM-KH) was employed [24, 25, 36, 61]. This statistical mechanical framework permits explicit sampling of molecular interactions, capturing both electrostatic and nonpolar contributions, including hydrogen bonding, solvophobic effects, solvent structuring, salt bridge formation, and steric constraints [59]. Semi-quantitative analyses using this method confirmed the substantial contribution of hemicellulose to cell wall stabilization, highlighting how specific xylan substitutions impact the nanoscale strength. The findings further indicated that stabilization primarily originates from functional group chemistry, rather than stereochemical configurations. For example, in galactoglucomannans, monosaccharide residues such as galactose, glucose, and mannose contribute only marginally to stability, whereas carboxylate groups of randomly substituted glucuronate residues are the major stabilizing factor.

Emerging approaches to mitigate biomass recalcitrance include engineering crops to modulate lignin content and xylan substitution patterns, thereby enabling less severe pretreatment and more effective enzymatic hydrolysis [13, 79]. Additionally, mild lignin removal from genetically modified plants has been shown to improve cellulose accessibility without disrupting the overall wall integrity [80]. Such strategies
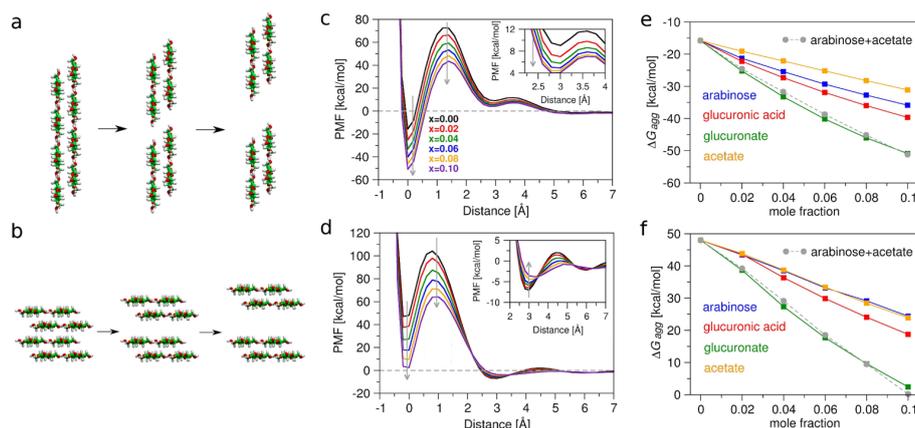
Figure 1. (Colour online) Potential of mean force $W_{\text{PMF}}(d)$ and aggregation free energy ΔG_{agg} [59]. Disaggregation of CNs in hemicellulose hydrogels through the disruption of hydrophilic (**a**) and hydrophobic (**b**) interactions. PMF profiles along these disaggregation pathways (**c**) and (**d**) are shown for glucuronate molar fractions ranging from $x = 0.0$ to 0.1 (legend in (**c**)). PMF variations with glucuronate concentration are indicated by grey arrows. G_{agg} for hydrophilic and hydrophobic contacts in hemicellulose hydrogels is displayed in (**e**) and (**f**), respectively. The sum of arabinose and acetate curves is represented by a grey dotted line.

facilitate integrated biomass valorization, where hemicellulose and lignin fractions are converted into low-molecular-weight compounds, while cellulose microfibrils are retained for nanocellulose applications [81]. These approaches offer a viable alternative to complete depolymerization of lignocellulose into biofuels, which remains limited by the crystalline nature of cellulose and its resistance to enzymatic breakdown [82].

3.1. Models of hemicellulose particles

A primary cell wall fragment model was constructed, consisting of a cellulose nanocrystallite (CN) composed of two four-chain, eight-glucose-length cellulose fragments immersed in aqueous solutions containing varying concentrations of arabinose, glucuronic acid, and glucuronate [59]. The CN models were generated using Cellulose-Builder [83].

The 3D-RISM-KH solvation theory was applied to examine hydrophilic and hydrophobic aggregation pathways of CNs (figures 1a and 1b). By solving the integral equations with the CN interaction potentials, the structural and solvation free energy properties of hemicellulose around CNs were determined through three-dimensional solvent site density maps.

The potential of mean force (PMF) between CNs in glucuronate solutions was computed as

$$W_{\text{PMF}}(d) = u_{12}(d) + \mu_{12}(d) + \mu_{\text{LJ}}(d) - \mu_1 - \mu_2, \tag{1}$$

where $u_{12}(d)$ is the CNs interaction potential, $\mu_{12}(d)$ is the solvation free energy of the CN aggregate at separation d, and μ_1 and μ_2 are the solvation free energies of the individual CNs. The PMF curves corresponding to hydrophilic (figure 1a) and hydrophobic (figure 1b) contact modes are shown in figures 1c and 1d.

During disaggregation, both the configurations exhibited two local minima: a direct face-to-face contact ($d_{\text{fc}} \approx 0$ Å) and a solvent-separated state ($d_{\text{ss}} \approx 3$ Å). In the hydrophilic case, the global minimum occurs at a direct contact due to extensive hydrogen bonding, whereas in the hydrophobic case the solvent-separated configuration is more stable (≈ -7 kcal/mol in pure water). These results suggest that CN aggregation in primary walls is predominantly mediated by hydrophilic interactions [80]. Further, the outer layers of CNs are packed less densely than the core ones, and significant energy barriers must be overcome to detach glucan chains from the CN or to assemble dispersed chains into CN structures.

With a hydronium counterion modelled using [84], glucuronate increases CN aggregation by a factor of 2 compared to arabinose or glucuronic acid. The stabilization in the hydrophobic face contact

(figure 1b) at 0.1 mole fraction of glucuronate results in the aggregation free energy $\Delta G_{\text{agg}} \approx 0$, making aggregation spontaneous. A similar trend is observed for the hydrophilic face contact (figure 1a), where the aggregation well deepens from $\Delta G_{\text{agg}} \approx -15$ kcal/mol in pure water to -30 kcal/mol in arabinose and -55 kcal/mol for glucuronate at a mole fraction of 0.1.

Increasing glucuronate concentration reduced aggregation free energy barriers while enhancing stabilization effects (figures 1c–1f). Glucuronate proved to be twice as effective as arabinose or glucuronic acid in stabilizing CNs, with stabilization effects arising from both the sugar ring and the carboxylate group. Notably, aggregation free energy decreased nearly linearly with an increasing hemicellulose concentration, consistent with the displacement of water molecules from cellulose surfaces. These results also explain why alkaline pretreatments, such as sodium hydroxide, efficiently remove hemicellulose: OH^- ions disrupt glucuronate-cellulose interactions, weakening the structural integrity [77].

The significant role of hemicellulose carboxylate groups in primary cell wall integrity poses two key questions: (i) Is the stabilization glucuronate effect due to the basic carboxylate group or the acidic hydronium ion? (ii) What is the contribution from the sugar-like cyclic part of glucuronate? The 3D-RISM-KH results with an acetate anion as a basic hemicellulose branch model and a hydronium counterion show that the effect of acetate is minor and similar to that of arabinose (figures 1e and 1f). About half of the effect is attributed to the sugar ring and the other half to the carboxylate group. Moreover, the sum of ΔG_{agg} for acetate and arabinose is close to ΔG_{agg} for glucuronate. This means that both the sugar moiety and the functional group contribute to CN stabilization (grey line in figures 1e and 1f).

Furthermore, ΔG_{agg} decreases almost linearly with increasing hemicellulose concentration due the ability of hemicellulose branches to displace the water molecules from the cellulose surface. This agrees with previous findings [13] that inhibiting the incorporation of glucuronic acid and glucuronate into xylan branches degrades the plant cell walls. The 3D-RISM-KH calculations explained why sodium hydroxide effectively removes hemicellulose from biomass [59], in agreement with the experimental findings [85]. The strong base OH^- disrupts glucuronate-cellulose interactions and destabilizes the cell wall microstructure, which can be used in the design of advanced biomass pretreatment strategies [59].

3.2. Hemicellulose binding to the aggregated cellulose nanocrystallites

A molecular representation of hemicellulose binding to aggregated cellulose nanostructures (CNs) is illustrated through the 3D density distributions of all carbon atoms within the corresponding hemicellulose branches, as shown in figure 2. The hydrogen (H) and oxygen (O) atoms on the nanocellulose surface that participate in hydrogen bonding are identified in figure 2a, distinguishing hydrogen bond donors from acceptors. Figures 2b–2e depict the regions where hemicellulose exhibits an increased local density at the nanocellulose surface relative to its bulk phase, allowing for the reconstruction of the preferred binding modes of hemicellulose branches around cellulose.

The uncharged hemicellulose branches of arabinose and glucuronic acid primarily interact with the hydrophilic faces of nanocellulose as hydrogen bond donors by coordinating with available oxygen atoms on the nanocellulose surface. Conversely, glucuronate anions act as hydrogen bond acceptors due to their strong interactions with the primary hydroxyl groups on the nanocellulose surface.

Furthermore, the distribution of acetate anions in figure 2e closely resembles that of glucuronate anions, underscoring the significance of basic functional groups in coordinating with hydroxyl groups. These findings indicate that hemicellulose branches within the plant cell walls interact with the nanocellulose surface in a complementary manner, contributing to the stabilization of the cell wall microstructure. The polar interactions between the exocyclic groups of hemicellulose and the polar groups on nanocellulose dictate the preferred orientation of hemicellulose branches over the hydrophobic surface. Additionally, interactions involving endocyclic sites with the nanocellulose surface exhibit a less specific, hydrophobic character, which does not produce distinct features in the 3D site distributions.

3.3. Solvation free energy density at the nanocellulose surface

Figure 3 illustrates the three-dimensional spatial maps of the solvation free energy density (3D-SFED) contributed by glucuronate at the nanocellulose surface. The solvation free energy denoted as $\mu_{\text{solv}} = \int_s d\mathbf{r} \Phi_s(\mathbf{r})$ represents the total solvation free energy density (3D-SFED) contributions $\Phi_s(\mathbf{r})$ from

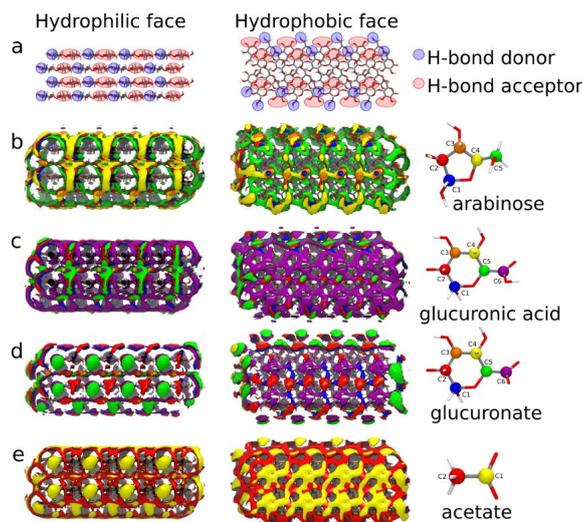**Figure 2.** (Colour online) 3D site density distributions $g(\mathbf{r})$ of aggregated CNs [51]. (**a**) highlights the hydrogen bond donor and acceptor sites within the CN structure. $g(\mathbf{r})$ of hemicellulose carbon atoms from arabinose (**b**), glucuronic acid (**c**), glucuronate (**d**), and acetate (**e**) monomers in proximity to the CNs. The isosurfaces of $g(\mathbf{r}) = 1.4$ are displayed in the same colors as the corresponding atoms, except for the glucuronate C6 atom which is represented at $g(\mathbf{r}) = 2.0$.

all hydrogel species s integrated over the system space for CNs immersed in a hemicellulose hydrogel. These 3D-SFED maps provide a comprehensive representation of the solvation structure while encoding convolved information regarding interaction forces. They explicitly characterize the ensemble-averaged effective interactions of hemicellulose components spatially distributed over the nanocellulose surface, thereby elucidating their influence on CN aggregation.

The 3D-SFED of the glucuronate component varies from highly negative values, corresponding to its most thermodynamically favorable arrangements at the nanocellulose surface, to less negative values, indicating less favorable configurations. As shown in figure 3a, the isosurface with larger negative values is highly localized around the polar sites on the nanocellulose surface, signifying hydrogen bonding interactions between hemicellulose and cellulose. By contrast, the isosurface with smaller negative values, depicted in figure 3b, represents a more diffuse second layer of hemicellulose monomers stacking over the nanocellulose surface. Within this secondary layer, stacking interactions are weaker and less specific compared to hydrogen bonding, primarily arising from hydrophobic interactions and enhanced intermolecular C-H...O interactions, as observed in the cellulose X-ray structure [86].

Although stacking interactions contribute significantly, hemicellulose-cellulose binding is predominantly driven by site-specific hydrogen bonds. The 3D-SFED maps of arabinose and glucuronic acid exhibit trends similar to those observed in the potential of the mean force W_{PMF} and aggregation free energy ΔG_{agg} presented in figure 1.

4. Effect of lignin chemical composition on lignin-lignin and lignin–hemicellulose supramolecular interactions in secondary plant cell walls

The inherent recalcitrance of lignocellulosic biomass represents a major obstacle to the advancement of plant-based green technologies. This resistance arises from the hierarchical arrangement of cellulose, hemicellulose, and lignin within the plant cell walls, forming highly ordered and robust supramolecular structures that limit the accessibility to enzymatic hydrolysis [10, 11]. In particular, the three-dimensional architecture of secondary cell walls comprising cellulose microfibrils embedded within a matrix of hemicellulose and lignin operates as a protective barrier that significantly contributes to recalcitrance [10, 87].
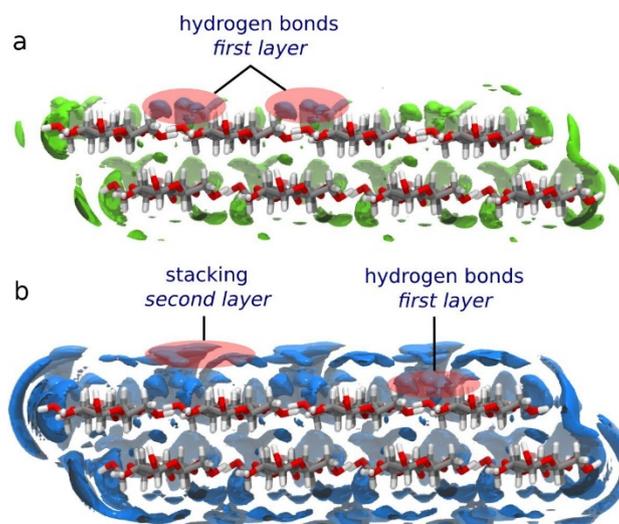
Figure 3. (Colour online) Isosurfaces of the 3D solvation free energy density of glucuronate around cellulose nanocrystals [59]. (**a**) Localized regions of strong hydrogen bonding are shown in green (larger negative isovalue). (**b**) Diffuse second-layer distributions with weaker interactions, shown in blue (smaller negative isovalue), represent hemicellulose–cellulose stacking.

Lignin, a heterogeneous aromatic polymer built from phenylpropanoid monomers, is a key determinant of this resistance and is considered to be the only large-scale renewable source of aromatic carbon [88, 89]. Together with hemicellulose, an amorphous heteropolysaccharide [90], lignin occupies the interstitial spaces between cellulose fibrils, thereby restricting cellulose accessibility and reducing its enzymatic digestibility.

Given the complexity of lignocellulosic biomass, research has increasingly emphasized multiscale investigations into its structural and dynamic properties [89]. Small-angle neutron scattering and MD simulations have demonstrated that lignin aggregates form fractal-like, multiscale structures with pores sufficiently large to accommodate cellulolytic enzymes [91]. Furthermore, MD simulations indicate that lignin undergoes entropy-driven collapse under ambient conditions [92].

Studies further reveal that lignin and hemicellulose interpenetrate at their interface, underscoring a strong molecular affinity between these two polymers [93]. To investigate these interactions more rigorously, the 3D-RISM-KH molecular theory of solvation has been applied, enabling the study of lignin-lignin and lignin-hemicellulose interactions as a function of lignin chemical composition [60]. Unlike conventional MD simulations, this statistical mechanics based approach accounts for long-timescale solvation processes in complex aqueous environments. Results suggest that supramolecular interactions between lignin and hemicellulose play a critical role in the assembly and cohesion of secondary cell walls. These interactions are largely entropy-driven, with hydrophobic forces expelling water from polymer surfaces. Importantly, chemical composition of lignin particularly its degree of methoxy substitution-modulates both the mechanical stability and digestibility of plant biomass. Similarly, hemicellulose composition influences how effectively it integrates into the plant cell wall matrix.

4.1. Lignin-hemicellulose and lignin-lignin interactions

To probe lignin-hemicellulose and lignin-lignin interactions at lignin concentrations between 0 and 10^{-2} molar fraction, hemicellulose and lignin oligomers were modelled in aqueous solutions of the three primary lignin monomers: p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) (figure 4a) [60]. In lignin polymers, the dominant linkage is the β-O-4$'$ ether bond [94, 95]. To approximate polymeric residues more accurately (figure 4b), ether linkages were substituted with methoxy groups (figure 4c). Optimized monomer geometries were generated using molecular mechanics and integrated into the DRISM-KH framework [60] to compute the solvation structures and thermodynamic properties.
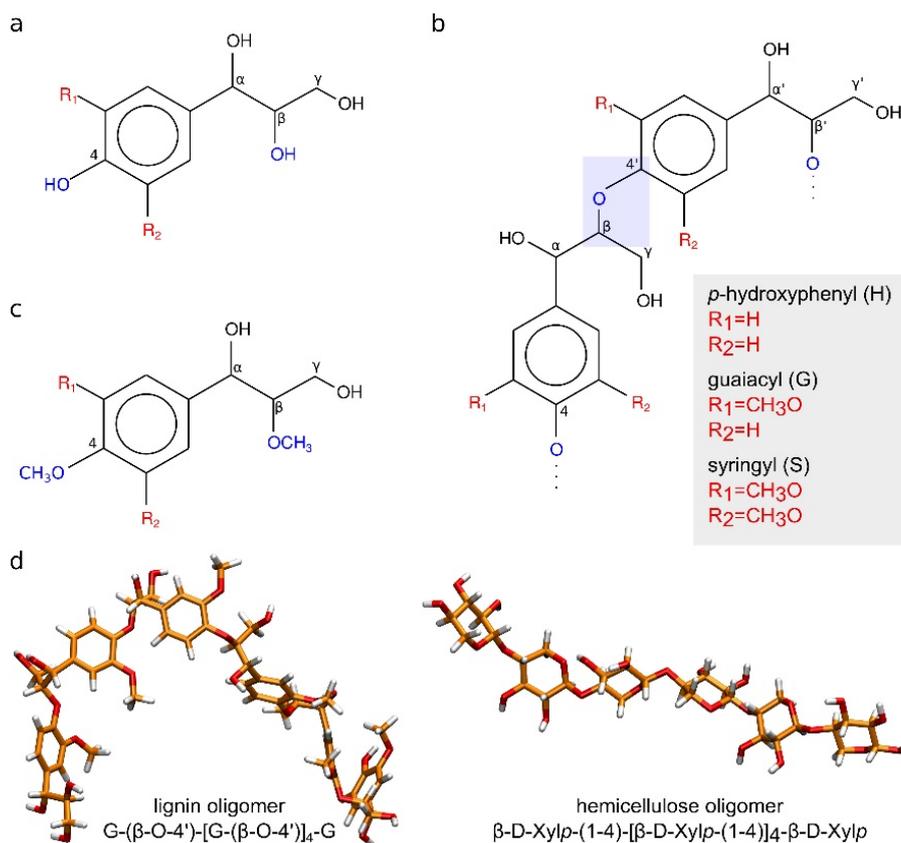
Figure 4. (Colour online) Hemicellulose and lignin oligomers [60]. (**a**) Lignin residues, with substituents R_1 and R_2 highlighted in red. (**b**) Lignin dimer containing a β-O-4′ bond (light-blue background). (**c**) Lignin monomer with hydroxyl groups replaced by methoxy groups. (**d**) Stick model of lignin and hemicellulose oligomers, with β-O-4′ bonding in blue.

MD simulations of 10 ns duration were performed to generate conformers of lignin and hemicellulose oligomers. Xylohexaose, composed of six β $(1 \rightarrow 4)$-linked xylose units, represented hemicellulose, while hexaguaiacyl, containing six guaiacyl residues linked by β-O-4′ bonds, represented lignin (figure 4d). Hemicellulose was parameterized using the CHARMM force field [96], while lignin was modelled with CHARMM parameters developed by Petridis and Smith [97]. Water was represented using a modified TIP3P model [98, 99]. Thermodynamic properties, including solvent site-site susceptibility and solvation free energy, were evaluated using 3D-RISM-KH theory (1)–(2) at 300 K, density $\rho = 1.0$ g/cm^3, and dielectric constant $\epsilon = 78.5$. Entropy (6b) was calculated at temperatures 295 K and 305 K.

4.2. Solvation free energy of the hemicellulose and lignin oligomers

The solvation free energies (SFEs) of hemicellulose and lignin oligomers revealed that the presence of lignin monomers stabilizes both types of oligomers, with stabilization increasing as lignin concentration rises. The strength of these interactions follows the order H < G < S, highlighting the impact of methoxy substitution on supramolecular affinity (figure 5). At a molar fraction of 0.01, SFEs for hemicellulose ranged from −10 to −15 kcal/mol, while lignin values reached −17 to −27 kcal/mol depending on the monomer type. These results confirm that lignin exhibits both strong self-affinity and a pronounced tendency for aggregation.

Experimental evidence aligns with these findings: lignin composition modulates biomass digestibility, with syringyl-rich lignin imparting greater recalcitrance than guaiacyl-rich lignin, whereas p-hydroxyphenyl units are associated with reduced resistance [79, 100–104]. As the interactions medi-
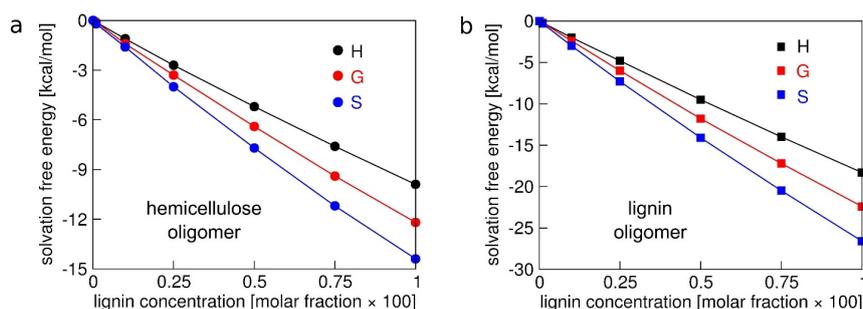
Figure 5. (Colour online) Solvation free energies of (**a**) hemicellulose and (**b**) lignin oligomers in aqueous solutions containing H, G, or S monomers at molar fractions $x = 0.0 - 0.1$, relative to pure water [60]. Interaction strength increases in the order H < G < S.

ated by S lignin are about 25% stronger than those mediated by G lignin, a decrease in the S/G ratio reduces the secondary cell wall recalcitrance, as was experimentally observed in sugarcane and switchgrass [101, 102]. Further, lignin enriched with H units was related to lower recalcitrance compared to lignin enriched with G or S units [103, 104]. These experimental findings confirm the theoretical predictions [60] that methoxy substituents in lignin enhance cell wall stability. Nevertheless, contradictory reports suggested that a higher S/G ratio may also enhance digestibility, underscoring the influence of genetic and environmental variability [105, 106].

Importantly, lignin composition is not the sole determinant of biomass recalcitrance. Other structural factors, such as lignin–carbohydrate complexes, porosity, cellulose crystallinity, and enzyme inhibition, also contribute [10, 11, 95]. The present solvation studies isolate supramolecular contributions, thereby providing insight into molecular-scale interactions free from larger-scale variability. The effect of lignin composition on its interaction with hemicellulose strongly affects the biomass processing, in particular, when hemicellulose removal is essential for efficient digestion, such as in switchgrass [107]. The lignin composition may affect the severity of pretreatment necessary for hemicellulose extraction.

With biomass recalcitrance determined by multiple interdependent factors beyond lignin content and composition, such as biomass porosity, lignin-carbohydrate complexes, lignin-induced enzyme inhibition, and cellulose crystallinity and accessibility [10, 11, 96], the calculations [60] illustrated the effects of supramolecular interactions in the plant cell walls while excluding other uncontrolled factors at larger scales.

At the molecular level, hemicellulose-lignin and lignin-lignin interactions display similar characteristics, since changes in lignin concentration or chemical structure yield comparable effects on the surface free energy (SFE) of both oligomers. Figure 6 shows the respective water and lignin contributions to SFE, derived by partitioning the solvation free energy (5) into interaction components.

For hemicellulose (figure 6a), the favorable negative contribution from water increases with lignin concentration, while the lignin contribution increases unfavorably. Entropy loss from restricted lignin monomer motion is only partly offset by lignin interactions, making the water contribution the dominant stabilizing factor. The effect of lignin composition is stronger in the water contribution, following H < G < S, consistent with the overall SFE trend. A similar mechanism drives lignin-lignin interactions, with methoxy substitution further enhancing water contributions (figure 6b).

SFE and solvation entropy values at 300 K (table 1) confirm that entropy increases with methoxy content (H < G < S), explaining the observed stabilization of both hemicellulose-lignin and lignin-lignin interactions. These results indicate predominantly hydrophobic interactions, where dehydration of interfaces facilitates favorable water contributions. Increased methoxy substitution strengthens the lignin hydrophobicity, enhancing solvation entropy and water exclusion.

Relative hydrophobicity is also supported by DRISM-KH predictions of bulk compressibility, which increases with lignin concentration and composition (H < G < S; figure 7). These results align with MD simulations showing the entropy-driven lignin collapse and solvent-mediated lignin-cellulose association at ambient conditions. Overall, lignin probably contributes to biomass recalcitrance through its hydrophobic interactions with cellulose, emphasizing its role in the plant cell wall structure and stability.
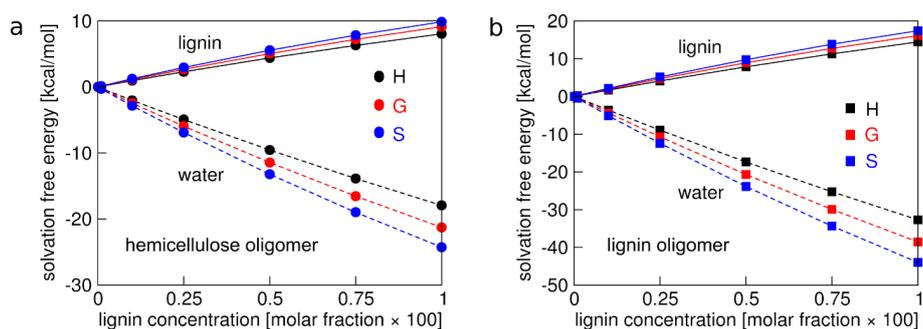
Figure 6. (Colour online) Contributions of lignin and water to solvation free energy [60] for hemicellulose (**a**) and lignin (**b**) oligomers as a function of lignin concentration for p-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin monomers. Solid lines denote water contributions, dashed lines represent lignin contributions. Total SFE values correspond to the sum of these contributions (figure 5).
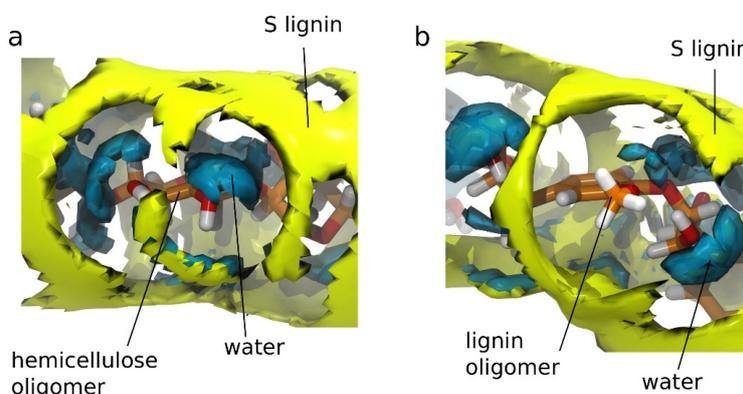

Figure 7. (Colour online) 3D-SFED isosurfaces of a water molecule (cyan) and a syringyl lignin monomer (yellow) around hemicellulose (**a**) and lignin (**b**) oligomers, shown in stick form [60]. Water isosurfaces highlight hydrogen bonding with polar groups, while syringyl distributions reflect non-specific hydrophobic stacking in hemicellulose-lignin and lignin-lignin interactions.

Table 1. Solvation free energy $\Delta\mu$ and solvation entropy $T\Delta S_V$ of hemicellulose and lignin in an aqueous solution containing lignin monomers at a molar fraction of $x = 0.01$ and a temperature of 300 K, expressed relative to H lignin [60].

z	Hc oligomer		Lignin oligomer	
Monomer	$\Delta\mu$ [kcal/mol]	$T\Delta S_V$ [kcal/mol]	$\Delta\mu$ [kcal/mol]	$T\Delta S_V$ [kcal/mol]
H	0	0	0	0
G	−2.2	2.2	−4.1	5.0
S	−4.5	3.8	8.3	6.8

4.3. Solvation free energy density of hemicellulose and lignin oigomers in solution

The three-dimensional solvation free energy density (3D-SFED), calculated using equation (5b), is illustrated in figure 7. These isosurfaces provide a molecular-scale view of the supramolecular interactions underlying the thermodynamic properties described in earlier sections. The isovalues corresponding to the complete first solvation shell of water remain unchanged, while larger isovalues capture the second hydration layer. By partitioning the 3D-SFED contributions from water and different lignin monomers, a molecular-level representation of the effective interactions was achieved; namely, those between hemicellulose and lignin oligomers (solute) and between water and lignin monomers (solvent).

This analysis highlights the supramolecular interactions between lignin and hemicellulose, the two predominant heteropolymers of the non-cellulosic matrix in the secondary plant cell walls, while also assessing how variations in lignin composition affect their interaction strength. Both hemicellulose-lignin and lignin-lignin interactions display comparable features, dominated by hydrophobic and non-specific associations. Such interactions are probably critical to the organization of the amorphous secondary wall, potentially driven by entropic forces that limit water accessibility to hydrophobic domains. Furthermore, the chemical structure of lignin, particularly the degree of methoxy substitution, significantly influences the cohesive interactions within the wall, thereby shaping its mechanical resilience. Studies employing model compounds rather than intact plant cell walls offer an effective means of probing the molecular mechanisms underpinning lignocellulosic biomass recalcitrance.

5. Conclusion

Cellulose, the most abundant biopolymer on Earth, serves as a vital renewable resource for sustainable biofuel production. The susceptibility of lignocellulosic biomass to depolymerization can be enhanced through enzymatic or chemical pretreatments that disrupt cellulose microfibrils, often within thermochemical processes. Solvents play a central role in these treatments, and elucidating the molecular mechanisms underlying biomass modification is critical for improving their efficiency.

In this context, the three-dimensional reference interaction site model with the Kovalenko–Hirata closure (3D-RISM-KH) molecular solvation theory was applied to investigate the interactions of cellulose nanocrystals embedded in a hemicellulose hydrogel. The analysis predicted how variations in hemicellulose chemical composition influence the nanoscale forces that regulate the assembly of primary cell walls. Furthermore, the study demonstrated that lignin-lignin and lignin-hemicellulose associations are predominantly hydrophobic and entropy-driven, arising from the displacement of water molecules at their contact surfaces.

These findings provide valuable insight into the thermodynamic principles governing the biomass restructuring, particularly during high-temperature treatments that increase the average diameter of cellulose microfibrils. By advancing the molecular-level understanding of biomass pretreatment, the study contributes to the development of more effective strategies for plant biomass conversion. Additionally, the results support integrated valorization approaches, in which non-cellulosic fractions are depolymerized into low-molecular-weight chemicals, while cellulose microfibrils are utilized for the production of nanocrystalline and nanofibrillar cellulose.

6. Funding

References

1. Gerez G., di Remigio Eikås R., Rune Jensen S., Bjørgve M., Frediani L., J. Chem. Theory Comput., 2023, **19**, 1986–1997, doi:10.1021/acs.jctc.2c01098.
2. Skyner R. E., McDonagh J. L., Groom C. R., van Mourik T., Mitchell J. B. O., Phys. Chem. Chem. Phys., 2015, **17**, 6174–6191, doi:10.1039/C5CP00288E.
3. Vyboishchikov S. F., Voityuk A. A., J. Chem. Inf. Model., 2021, **61**, 4544–4553, doi:10.1021/acs.jcim.1c00885.
4. Klamt A., Diedenhofen M., J. Phys. Chem. A, 2015, **119**, 5439–5445, doi:10.1021/jp511158y.
5. Chaudhari M. I., Vanegas J. M., Pratt L. R., Muralidharan A., Rempe S. B., Annu. Rev. Phys. Chem., 2020, **71**, 461, doi:10.1146/annurev-physchem-012320-015457.
6. Pratt L. R., Rempe S. B., AIP Conf. Proc., 1999, **492**, 172–201, doi:10.1063/1.1301528.
7. Asthagiri D., Dixit P., Merchant S., Paulaitis M., Pratt L., Rempe S. B., Varma S., Chem. Phys. Lett., 2010, **485**, 1–7, doi:10.1016/j.cplett.2009.12.013.
8. Rogers D. M., Rempe S. B., Phys. Chem. B, 2011, **115**, 911—9129, doi:10.1021/jp2012864.
9. Rogers D. M., Jiao D., Pratt L. R., Rempe S. B., Annu. Rep. Comput. Chem., 2012, **8**, 71–127, doi:10.1016/B978-0-444-59440-2.00004-1.
10. Himmel M. E., Ding S., Johnson D. K., Adney W. S., Nimlos M. R., Brady J. W., Foust T. D., Science, 2007, **315**, 804–807, doi:10.1126/science.1137.
11. Chundawat S. P. S., Beckham G. T., Himmel M. E., Dale B. E., Annu. Rev. Chem. Biomol. Eng., 2011, **2**, 121–145, doi:10.1146/annurev-chembioeng-061010-114205.
12. Abolore R. S., Jaiswal S., Jaiswal A. K., Carbohydr. Polym. Technol. Appl., 2024, **7**, 100396, doi:10.1016/j.carpta.2023.100396.
13. Mortimer J. C., Miles G. P., Brown D. M., Zhang Z., Segura M. P., Weimar T., Yu X., Seffen K. A., Stephens E., Turner S. R., Dupree P., Proc. Natl. Acad. Sci. U.S.A., 2010, **107**, 17409–17414, doi:10.1073/pnas.1005456107.
14. Clark J. H., Green Chem., 2019, **21**, 1168–1170, doi:10.1039/C9GC90021G.
15. Abolore R. S., Jaiswal S., Jaiswal A. K., Carbohydr. Polym. Technol. Appl., 2024, **7**, 100396, doi:10.1016/j.carpta.2023.100396.
16. Mazumder S., Golbabaei M. H., Zhang N., Biomimetics, 2025, **10**, 802, doi:10.3390/biomimetics10120802.
17. Kong Y., Fu S., BioResources, 2022, **17**, 7135–7166, doi:10.15376/biores.17.4.Kong.
18. Mazumder S., Zhang N., Biomimetics, 2023, **8**, 188, doi:10.3390/biomimetics8020188.
19. Luo X., Li Q., Fuel Process. Technol., 2024, 258, 108093, doi:10.1016/j.fuproc.2024.108093.
20. Mensah M., Tia R., Adei E., Nora H. de Leeuw N. H., Front. Chem., 2022, **10**, 793759, doi:10.3389/fchem.2022.793759.
21. Zhang Y., He H., Dong K., Fan M., Zhang S., RSC Adv., 2017, **7**, 12670–12681, doi:10.1039/c6ra27059j.
22. Kostetskyy P., Broadbelt L.J., Energy Fuels, 2020, **34**, 15195–15216, doi:10.1021/acs.energyfuels.0c02295.
23. Wang W., Xu Y., Zhu B., Ge H., Wang S., Li B., Xu H., Bioresour. Technol., 2023, **385**, 129401, doi:10.1016/j.biortech.2023.129401.
24. Hirata F. (Ed.), Molecular Theory of Solvation, Understanding Chemical Reactivity, Kluwer, Dordrech, 2003, 358, doi:10.1007/1-4020-2590-4.
25. Roy D., Kovalenko A., Int. J. Mol. Sci., 2021, **22**, 5061, doi:10.3390/ijms22105061.
26. Hansen J.-P., McDonald I. R., In: Theory of Simple Liquids: With Applications to Soft Matter, Academic Press, Oxford, fourth edn., 2013, 455–510, doi:10.1016/B978-0-12-387032-2.00011-8.
27. Verlet L., Mol. Phys., 1980, **41**, 183–190, doi:10.1080/00268978000102671.
28. Verlet L., Levesque D., Mol. Phys., 1982, **46**, 969–980, doi:10.1080/00268978200101711.
29. Martynov G. A., Sarkisov G. N., Mol. Phys., 1983, **49**, 1495–1504, doi:10.1080/00268978300102111.
30. Rogers F. J., Young D. A., Phys. Rev. A, 1984, **30**, 999–1007, doi:10.1103/PhysRevA.30.999.
31. Zerah G., Hansen J., J. Chem. Phys., 1986, **84**, 2336–2343, doi:10.1063/1.450397.
32. Ballone P., Pastore G., Galli G., Gazzillo D., Mol. Phys., 1986, **59**, 275–290, doi:10.1080/00268978600102071.
33. Martynov G. A., Vompe A. G., Phys. Rev. E, 1993, **47**, 1012, doi:10.1103/PhysRevE.47.1012.
34. Duh D., Haymet A. D. J., J. Chem. Phys., 1995, **103**, 2625–2633, doi:10.1063/1.470724.
35. Lee L. L., J. Chem. Phys., 1995, **103**, 9388–9396, doi:10.1063/1.469998.
36. Kovalenko A., Hirata F., J. Chem. Phys., 1999, **110**, 10095–10112, doi:10.1063/1.478883.
37. Kast S. M., Phys. Rev. E, 2003, **67**, 041203, doi:10.1103/PhysRevE.67.041203.
38. Kast S. M., Kloss T., J. Chem. Phys., 2008, **129**, 236101, doi:10.1063/1.3041709.
39. Tsednee T., Luchko T., Phys. Rev. E, 2019, **99**, 032130, doi:10.1103/PhysRevE.99.032130.
40. Johnson J., Case D.A., Yamazaki T., Gusarov S., Kovalenko A., Luchko T., J. Phys.: Condens. Matter, 2016, **28**, 344002, doi:10.1088/0953-8984/28/34/344002.
41. Imai T., Hiraoka R., Kovalenko A., Hirata F., J. Am. Chem. Soc., 2005, **127**, 15334–15335,

doi:10.1021/ja054434b.

42. Kondratenko M., Stoyanov S. R., Gusarov S., Kovalenko A., McCreery R. L., J. Phys. Chem. C, 2015, **119**, 11286–11295, doi:10.1021/jp5128332.

43. Stoyanov S. R., Gusarov S., Kuznicki S. M., Kovalenko A., J. Phys. Chem. C, 2008, **112**, 6794–6810, doi:10.1021/jp075688h.

44. da Costa L. M., Stoyanov S. R., Gusarov S., Seidl P. R., de M. Carneiro J. W., Kovalenko A., J. Phys. Chem. A, 2014, **118**, 896–908, doi:10.1021/jp408005h.

45. Fafard J., Lyubimova O., Stoyanov S. R., Dedzo G. K., Gusarov S., Kovalenko A., Detellier C., J. Phys. Chem. C, 2013, **117**, 18556–18566, doi:10.1021/jp4064142.

46. Genheden S., Luchko T., Gusarov S., Kovalenko A., Ryde U., J. Phys. Chem. B, 2010, **114**, 8505–8516, doi:10.1021/jp101461s.

47. Imai T., Miyashita N., Sugita Y., Kovalenko A., Hirata F., Kidera A., J. Phys. Chem. B, 2011, **115**, 8288–8295, doi:10.1021/jp2015758.

48. Malvaldi M., Bruzzone S., Chiappe C., Gusarov S., Kovalenko A., J. Phys. Chem. B, 2009, **113**, 3536–3542, doi:10.1021/jp810887z.

49. Omelyan I., Kovalenko A., J. Chem. Theory Comput., 2015, **11**, 1875–1895, doi:10.1021/ct5010438.

50. Kovalenko A., Kobryn A. E., Gusarov S., Lyubimova O., Liu X., Blinov N., Yoshida M., Soft Matter, 2012, **8**, 1508–1520, doi:10.1039/C1SM06542D.

51. Blinov N., Dorosh L., Wishart D., Kovalenko A., Biophys. J., 2010, **98**, 282–296, doi:10.1016/j.bpj.2009.09.062.

52. Casanova D., Gusarov S., Kovalenko A., Ziegler T., J. Chem. Theory Comput., 2007, **3**, 458–476, doi:10.1021/ct6001785.

53. Carpita N. C., McCann M. C., J. Biol. Chem., 2020, **295**, 15144–15157, doi:10.1074/jbc.REV120.014561.

54. Roman M., Cannizzo C., Pinault T., Isare B., Andrioletti B., van der Schoot P., Boutellier L., J. Am. Chem. Soc., 2010, **132**, 16818–16824, doi:10.1021/ja105717u.

55. Aida T., Meijer E. W., Stupp S. I., Science, 2012, **335**, 813–817, doi:10.1126/science.1205962.

56. Varki A., Cummings R. D., Esko J. D., Stanley P., Hart G. W., Aebi M., Darvill A. G., Kinoshita T., Packer N. H., Prestegard J. H., Schnaar R. L., Seeberger P. H. (Eds.), Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor (NY), 3rd edn., 2015, URL https://pubmed.ncbi.nlm.nih.gov/27010055.

57. Dong C., Beis K., Nasper J., Brunkan-LaMontagne A. L., Clarke B. R., Whitfield C., Naismith J. H., Nature, 2006, **444**, 226–229, doi:10.1038/nature05267.

58. Gama C. I., Tully S. E., Sotogaku N., Clark P. M., Pawat M., Vaidehi N., Goddard W. A. III, Nishi A., Hsieh-Wilson L. C., Nature Chem. Biol., 2006, **2**, 467–473, doi:10.1038/nchembio810.

59. Silveira R. L., Stoyanov S. R., Gusarov S., Skaf M. S., Kovalenko A., J. Am. Chem. Soc., 2013, **135**, 19048–19051, doi:10.1021/ja405634k.

60. Silveira R. L., Stoyanov S. R., Gusarov S., Skaf M. S., Kovalenko A., J. Phys. Chem. Lett., 2015, **6**, 206–211, doi:10.1021/jz502298q.

61. Kovalenko A., In: Springer Handbook of Electrochemical Energy, Breitkopf C., Swider-Lyons K. (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 2017, 95–139, doi:10.1007/978-3-662-46657-5_5.

62. Kovalenko A., Gusarov S., Phys. Chem. Chem. Phys., 2018, **20**, 2947–2969, doi:10.1039/C7CP05585D.

63. Gusarov S., Pujari B. S., Kovalenko A., J. Comput. Chem., 2012, **33**, 1478–1494, doi:10.1002/jcc.22974.

64. Kovalenko A., In: Molecular Theory of Solvation, Understanding Chemical Reactivity, Vol. 24, Hirata F. (Ed.), Kluwer Academic Publishers, Dordrecht, 2004, 169–275, doi:10.1007/1-4020-2590-4_4.

65. Kovalenko A., Hirata F., J. Chem. Phys., 2000, **112**, 10391–10402, doi:10.1063/1.481676.

66. Kovalenko A., Hirata F., J. Chem. Phys., 2000, **112**, 10403–10417, doi:10.1063/1.481677.

67. Kovalenko A., Hirata F., Chem. Phys. Lett., 1998, **290**, 237–244, doi:10.1016/S0009-2614(98)00471-0.

68. Perkyns J. S., Pettitt B. M., J. Chem. Phys., 1992, **97**, 7656–7666, doi:10.1063/1.463485.

69. Kirkwood J. G., Buff F. P., J. Chem. Phys., 1951, **19**, 774–777, doi:10.1063/1.1748352.

70. Harano Y., Imai T., Kovalenko A., Kinoshita M., Hirata F., J. Chem. Phys., 2001, **114**, 9506–9511, doi:10.1063/1.1369138.

71. Imai T., Harano Y., Kovalenko A., Hirata F., Biopolymers, 2001, **59**, 512–519, doi:10.1002/1097-0282(200112)59:7<512::AID-BIP1056>3.0.CO,2-C.

72. Palmer D. S., Frolov A. I., Ratkova E. L., Fedorov M. V., J. Phys.: Condens. Matter, 2010, **22**, 492101, doi:10.1088/0953-8984/22/49/492101.

73. Kovalenko A., Ten-no S., Hirata F., J. Comput. Chem., 1999, **20**, 928–936, doi:10.1002/(SICI)1096-987X(19990715)20:9<928::AID-JCC4>3.0.CO,2-X.

74. Pulay P., Chem. Phys. Lett., 1980, **73**, 393–398, doi:10.1016/0009-2614(80)80396-4.

75. Saad Y., Schultz M. H., SIAM J. Sci. Stat. Comput., 1986, **7**, 856–869, doi:10.1137/0907058.

76. Howard J. J., Perkyns J. S., Choudhury N., Pettitt B. M., J. Chem. Theory Comput., 2008, **4**, 1928–1939, doi:10.1021/ct8002817.
77. Minezawa N., Kato S., J. Chem. Phys., 2007, **126**, 054511, doi:10.1063/1.2431809.
78. Pauly M., Keegstra M., Plant J., 2008, **54**, 559–568, doi:10.1111/j.1365-313X.2008.03463.
79. Chen F., Dixon R. A., Nat. Biotechnol., 2007, **25**, 759–761, doi:10.1038/nbt1316.
80. Ding S.-Y., Liu Y.-S., Zeng Y., Himmel M. E., Baker J. O., Bayer E. A., Science, 2012, **338**, 1055–1060, doi:10.1126/science.1227491.
81. Habibi Y., Lucia L. A., Rojas O. J., Chem. Rev., 2010, **110**, 3479–3500, doi:10.1021/cr900339w.
82. Shishir P. S., Chundawat S. P. S., Bellesia G., Uppugundla N., da Costa Sousa L., Gao D., Cheh A. M., Agarwal U. P., Bianchetti C. M., Phillips Jr. G. N., Langan P., Balan. V., Gnanakaran S., Dale B. E., J. Am. Chem. Soc., 2011, **133**, 11163–11174, doi:10.1021/ja2011115.
83. Gomes T. C. F., Skaf M. S., J. Comput. Chem., 2012, **33**, 1338–1346, doi:10.1002/jcc.22959.
84. Yoshida N., Phongphanphanee S., Maruyama Y., Imai T., Hirata F., J. Am. Chem. Soc., 2006, **128**, 12042–12043, doi:10.1021/ja0633262.
85. Agbor V. B., Cicek N., Sparling R., Berlin A., Levin D. B., Biotechnol. Adv., 2011, **29**, 675–685, doi:10.1016/j.biotechadv.2011.05.005.
86. Nishiyama Y., Sugiyama J., Chanzy H., Langan P., J. Am. Chem. Soc., 2003, **125**, 14300–14306, doi:10.1021/ja037055w.
87. Achyuthan K. E., Achyuthan A. M., Adams P. D., Dirk S. M., Harper J. C., Simmons B. A., Singh A. K., Molecules, 2010, **15**, 8641–8688, doi:10.3390/molecules15118641.
88. Ragauskas A. J., Beckham G. T., Biddy M. J., Chandra R., Chen F., Davis M. F., Davison B. H., Dixon R. A., Gilna P., Keller M., Langan P., Naskar A. K., Saddler J. N., Tschaplinski T. J., Tuskan G. A., Wyman C. E., Science, 2014, **344**, No. 6185, 1246843, doi:10.1126/science.1246843.
89. Tuck C. O., Pérez E., Horávth I. T., Sheldon R. A., Poliakoff M., Science, 2012, **337**, 695–699, doi:10.1126/science.121893.
90. Scheller H. V., Ulvskov P., Annu. Rev. Plant Biol., 2010, **61**, 263–289, doi:10.1146/annurev-arplant-042809-112315.
91. Petridis L., Pingali S. V., Urban V., Heller W. T., O'Neill H. M., Foston M., Ragauskas A., Smith J. C., Phys. Rev. E, 2011, **83**, 061911, doi:10.1103/PhysRevE.83.061911.
92. Petridis L., Schulz R., Smith J. C., J. Am. Chem. Soc., 2011, **133**, 20277–20287, doi:10.1021/ja206839u.
93. Charlier L., Mazeau K., J. Phys. Chem. B, 2012, **116**, 4163–4174, doi:10.1021/jp300395k.
94. Boerjan W., Ralph J., Baucher M., Annu. Rev. Plant Biol., 2003, **54**, 519–546, doi:10.1146/annurev.arplant.54.031902.134938.
95. Pu Y., Hu F., Huang F., Davison B. H., Ragauskas A. J., Biotechnol. Biofuels, 2013, **6**, 15–13, doi:10.1186/1754-6834-6-15.
96. Guvench O., Greene S. N., Kamath G., Brady J. W., Venable R. M., Pastor R. W., MacKerell A. D. Jr., J. Comput. Chem., 2008, **29**, 2543–2564, doi:10.1002/jcc.21004.
97. Petridis L., Smith J. C., J. Comput. Chem., 2009, **30**, 457–467, doi:10.1002/jcc.21075.
98. Jorgensen W. L., Chandrasekhar J., Madura J. D., Impey R. W., Klein M. L., J. Chem. Phys., 1983, **79**, 926–935, doi:10.1063/1.445869.
99. MacKerell A. D. Jr., Bashford D., Bellott R. L., Dunbrack R. L. Jr., Evanseck J. D., Field M. J., Fischer S., Gao J., Guo H., Ha S., et al., J. Phys. Chem. B, 1998, **102**, 3586–3616, doi:10.1021/jp973084f.
100. Donohoe B. S., Decker S. R., Tucker M. P., Himmel M. E., Vinzant T. B., Biotechnol. Bioeng., 2008, **101**, 913–925, doi:10.1002/bit.21959.
101. Jung J. H., Fouad W. M, Vermerris W., Gallo M., Altpeter F., Plant Biotechnol. J., 2012, **10**, 1067–1076, doi:10.1111/j.1467-7652.2012.00734.x.
102. Fu C., Mielenz J. R., Xiao X., Ge Y., Hamilton C. Y., Rodriguez M. Jr., Chen F., Foston M., Ragauskas A., Bouton J., Dixon R.A., Wang Z.-Y., Proc. Natl. Acad. Sci. U.S.A., 2011, **108**, 3803–3808, doi:10.1073/pnas.1100310108.
103. Bonawitz N. D., Kim J. I., Tobimatsu Y., Ciesielski P. N., Anderson N. A., Ximenes E., Maeda J., Ralph J., Donohoe B. S., Ladisch M., Chapple C., Nature, 2014, **509**, 376–380, doi:10.1038/nature13084.
104. Ziebell A., Gracom K., Katahira R., Chen F., Pu Y., Ragauskas A., Dixon R. A., Davis M., J. Biol. Chem., 2010, **285**, 38961–38968, doi:10.1074/jbc.M110.137315.
105. Li X., Weng J.-K., Chapple C., Plant J., 2008, **54**, 569–581, doi:10.1111/j.1365-313X.2008.03457.x.
106. Mansfield S. D., Curr. Opin. Biotechnol., 2009, **20**, 286–294, doi:10.1016/j.copbio.2009.05.001.
107. DeMartini J. D., Pattathil S., Miller J. S., Li H., Hahn M. G., Wyman C. E., Energy Environ. Sci., 2013, **6**, 898–909, doi:10.1039/C3EE23801F.

Багатомасштабна теорія та моделювання геміцелюлози і лігніну в розчині

А. Коваленко

Центр програмного забезпечення для багатомасштабного моделювання, Едмонтон, Альберта, Канада, T6E 5J5

У цьому огляді розглядаються підходи до багатомасштабного моделювання целюлозних нанокристалів (ЦНК) та лігноцелюлозних клітинних стінок рослин з акцентом на взаємодії геміцелюлози та лігніну у водних середовищах. Тривимірна модель опорного сайту взаємодії із замиканням Коваленко-Хірати (3D-RISM-KH) лежить в основі перспективної теорії молекулярної сольватації, що застосовується в нанохімії та біомолекулярному моделюванні. Запропонований метод був успішно застосований для дослідження геміцелюлозних гідрогелів, впливу складу геміцелюлози на нанорозмірні сили в первинних клітинних стінках, а також взаємодій лігнін-лігнін та лігнін-геміцелюлоза. Результати дослідження показують, що ці взаємодії є переважно гідрофобними та зумовленими ентропією і виникають внаслідок ефектів виключення води. Результати моделювання, отримані за допомогою цієї моделі, поглиблюють розуміння молекулярних сил у клітинних стінках рослин та допомагають розробити стратегії утилізації біомаси, включаючи генну інженерію та технології попередньої обробки, спрямовані на покращення екстракції та використання целюлози.

Ключові слова: *нанокристали целюлози, геміцелюлоза, лігнін, теорія молекулярної сольватації, термодинаміка сольватації, молекулярне моделювання*