



Coarse-grained molecular dynamics simulations of α -1,3-glucan

Journal:	<i>Soft Matter</i>
Manuscript ID	SM-ART-03-2019-000580.R1
Article Type:	Paper
Date Submitted by the Author:	02-May-2019
Complete List of Authors:	<p>Beltran-Villegas, Daniel; University of Delaware, Chemical and Biomolecular Engineering Intriago, Daniel; University of Delaware, Chemical and Biomolecular Engineering Kim, Kyle; DowDuPont Specialty Products Division, Biomaterials, Industrial Biosciences Behabtu, Natnael; DowDuPont Specialty Products Division, Biomaterials, Industrial Biosciences Londono, David; DowDuPont Specialty Products Division, Science and Innovation, DuPont Specialty Products Jayaraman, Arthi; University of Delaware, Chemical & Biomolecular Engineering</p>

Submitted to Soft Matter

Coarse-grained molecular dynamics simulations of α -1,3-glucan

Daniel J. Beltran-Villegas¹, Daniel Intriago¹, Kyle H. C. Kim², Natnael Behabtu², J. David Londono³, and Arthi Jayaraman^{1,4,}*

1. Dept. of Chemical and Biomolecular Engineering, University of Delaware, Newark DE 19716
2. Biomaterials, DuPont Industrial Biosciences, DowDuPont Specialty Products Division, Wilmington DE 19803
3. Science and Innovation, DuPont Specialty Products, DowDuPont Specialty Products Division, Wilmington DE 19803
4. Dept. of Materials Science and Engineering, University of Delaware, Newark DE, 19716

**corresponding author arthij@udel.edu*

Abstract

In this paper we present a computational study of aggregation in aqueous solutions of α -1,3-glucan captured using a coarse-grained (CG) model that can be extended to other polysaccharides. This CG model captures atomistic geometry (i.e., relative placement of the hydrogen bonding donors and acceptors within the monomer) of α -1,3-glucan monomer, the directional interactions due to the donor-acceptor hydrogen bonds, and their effect on aggregation of multiple α -1,3-glucan chains without the extensive computational resources needed for simulations with atomistic models. Using this CG model, we conduct molecular dynamics simulations to assess the effect of varying α -1,3-glucan chain length and hydrogen bond interaction strengths on the aggregation of multiple chains at finite concentrations in implicit solvent. We quantify the hydrogen bonding strength needed for multiple chains to aggregate, the distribution of inter- and intra- chain hydrogen bonds within the aggregate and in some cases, the shapes of the aggregate. We also explore the effect of substitution/silencing of some randomly selected or specific hydrogen bonding sites in the chain on the aggregation and aggregate structure. In the unmodified α -1,3-glucan solution, the inter-chain hydrogen bonds cause the chains to aggregate into sheets. Random silencing of hydrogen bonding donor sites only increases the hydrogen bond strength needed for aggregation but retains the same aggregate structure as the unmodified chains. Specific silencing of the hydrogen-bonding site on the C6 carbon leads to the chains aggregating into planar sheets that then fold over to form hollow cylinders at intermediate hydrogen bond strength – 4.7 to 5.3 kcal/mol. These cylindrical aggregates assemble end-to-end to form larger aggregates at higher hydrogen bond strengths.

I. Introduction

Polysaccharides are linear polymers (e.g., cellulose, amylose) or branched polymers (e.g., starch) of monosaccharide units bound together by glycosidic linkages. Their unique structure, function (e.g., energy storage), properties (e.g., solubility, toughness, biodegradability), and the need for environmentally friendly biofuels and bioplastics have led to these polymers becoming the focus of extensive research by biologists, materials scientists, chemists, physicists, and engineers.¹⁻³ In this study we focus on one specific type of polysaccharide, α -1,3-glucan, comprised of glucose repeat units linked through α -1,3 glycosidic bonds, which is a known fungal cell wall pathogenesis agent⁴⁻⁶ and is a major component of branched polysaccharides found in dental plaque^{7, 8}. As a synthetic bioplastic, α -1,3-glucan is produced by enzymatic synthesis using sucrose as a base material^{1, 9, 10} to yield highly linear-, water-insoluble polysaccharide chains. The hydroxyl groups in each repeat unit form intra-chain and inter-chain hydrogen bonds⁷. The location and pattern of hydrogen bonds between chains varies in the different crystalline states⁷. Studies of hydroxyl group substitution of synthetic α -1,3-glucan show promising results as thermoplastic polymers –acetates of ester-substituted polymer¹¹ and as viscosity modifiers – ether-substituted polymers.¹² Since the use of α -1,3-glucan in a broad range of technological applications depends on development of α -1,3-glucan-derived polymers with desired solubility and processability, it is important to establish the link between monomer chemistry and macromolecular structure of polysaccharides in both crystalline and dissolved states. Experimental techniques are often limited in their capabilities to characterize these materials at both molecular and macroscopic scales. Therefore, the use of simulations in conjunction with experiments is a powerful tool for connecting molecular level design of polysaccharides to its macroscopic behavior.

Past simulation studies on polysaccharides have mostly used atomistic models which retain the chemical structure (i.e., chirality, chemical composition) at the monomer level but in turn, due to increased computational intensity, can only efficiently sample pre-assembled structures^{2, 13-19} or simulate assembly of a few small molecular weight chains (i.e., oligosaccharides to monosaccharides).^{20, 21} Therefore, there is a need for less-detailed (or coarse-grained) models that can retain details at the monomer level (e.g., placement of hydrogen bonding donors and acceptors) and simultaneously enable computationally fast simulations at experimentally relevant length and time scales. In this paper, we present one such coarse-grained (CG) model for α -1,3-glucan polysaccharides that retains the key chemical features at the monomer level, especially chirality (i.e., distinguishing isomer α -glucan conformation), while enabling simulations of large molecular weight polysaccharides at experimentally relevant solution concentrations.

In past CG simulations of polysaccharides, the CG models were built to determine the mechanical properties of fibers or crystalline structures of cellulose²²⁻²⁴, match the behavior of monosaccharides or oligosaccharides in solution²⁵⁻²⁷, study the interaction between oligosaccharides and proteins²⁸, and study the solvation of polysaccharide chains from crystalline structures by solvent treatment²⁹. These CG models were built to capture the behavior of polysaccharides in different solvents driven by hydrogen bonding but represented those directional hydrogen bonds using effective isotropic interactions by structure mapping from atomistic simulations.^{17, 22-38} We can categorize these polysaccharide CG models into one of three main categories: (1) single bead per monomer models^{23, 24, 33} where an overall attractive strength is imposed on a monomer bead (2) three beads per monomer (M3B)-like models^{27, 31, 36} where three beads represent the sugar ring, and (3) MARTINI-like models^{29, 37, 38} that systematically assign specialty coarse-grained beads to groups of atoms. These models maintain chain stiffness and shape with angle and dihedral potentials. Later CG models have incorporate more detail by using more CG beads per monomer.³² In particular, the work by Markutsya and collaborators³⁰ systematically shows the benefits of increasing complexity in CG models and the importance of decoupling the sugar ring from the oxygen in the glycosidic bond. In these studies, the CG models were specific to one polysaccharide chemistry and the inherently directional hydrogen bonding pattern of polysaccharides was pre-set through *isotropic* interactions. In contrast to all of these past studies, in our α -1,3-glucan CG polysaccharide model presented here the directionality of the hydrogen bonding between the acceptor and donor sites on the CG monomers is retained. Because hydrogen bonding is known to be the driving force for the aggregation of polysaccharides, we chose to develop this CG model by extending a previous DNA CG model developed by Ghobadi and Jayaraman³⁹ where the directionality and specificity of the hydrogen bonding between complementary bases was captured. This CG model of Ghobadi and Jayaraman has been expanded to represent different oligonucleic acid architectures⁴⁰, polymer-conjugated oligonucleic acids^{41, 42}, and collagen-like polypeptides⁴³. In all of these cases this CG model captures melting thermodynamics and some structural features. In this paper we extend this model to represent polysaccharides, and in particular, α -1,3-glucan. The CG model captures the relative placement of acceptor and donor sites on the polysaccharide monomer CG bead as seen in atomistic representation. This CG model also captures the chiral nature of these chains, which to the best of our knowledge no previous CG model does. This CG model allows us to explore the effects of chain length, hydrogen-bonding strength, and concentration on the aggregation and morphology of α -1,3-glucan.

This paper is organized as follows. In section II we describe our simulation and analysis methods, as well as the CG α -1,3-glucan model. In section III we present our results for aggregation and structure of α -1,3-glucan solutions at varying hydrogen bond strengths and chain length, as well as an exploration

of the effect of hydrogen bond donor site substitution/silencing on both aggregation and structure of polymer solutions. We conclude in section IV with a summary of our findings and a vision for future work.

II. Methods

A. Model

The CG model presented in this paper is an extension of our recent work on synthetic and biopolymer chemistries where hydrogen bonding is an important interaction.³⁹⁻⁴³ In the CG model used in this work each monosaccharide ring is represented by a single backbone bead and each hydrogen-bonding site is represented by smaller beads embedded partly in the backbone bead, as shown in Figure 1. Each glucose monomer has three additional small beads for the three OH sites, one additional bead for the one O site and another linker bead for the oxygen in the glycosidic bond. We denote the linker oxygen with L (linker), the OH sites with D (donor) and O site with A (acceptor). We note that the O in the OH site can also serve as an acceptor to a donor OH group, and this is captured in the interactions described later. The dimensions of the simulation are defined in reduced units, with reduced length $\sigma = 0.42\text{nm}$ and reduced energy $\varepsilon = 0.1 \text{ kcal/mol}$. In these reduced units, the diameter of the backbone bead is $\sigma^{\text{BB}} = 1.0\sigma$. The diameter of the hydrogen bonding beads and linker beads are $\sigma^{\text{HB}} = 0.15\sigma$. The size of the backbone bead is modeled to be approximately equal to the size of a glucose ring, as explained in the Electronic Supplementary Information, section S.A. The size of the hydrogen bonding beads is chosen to capture the directionality in the hydrogen bonding interaction, as described later. Our CG model does not have solvent beads, as the solvent is treated implicitly.

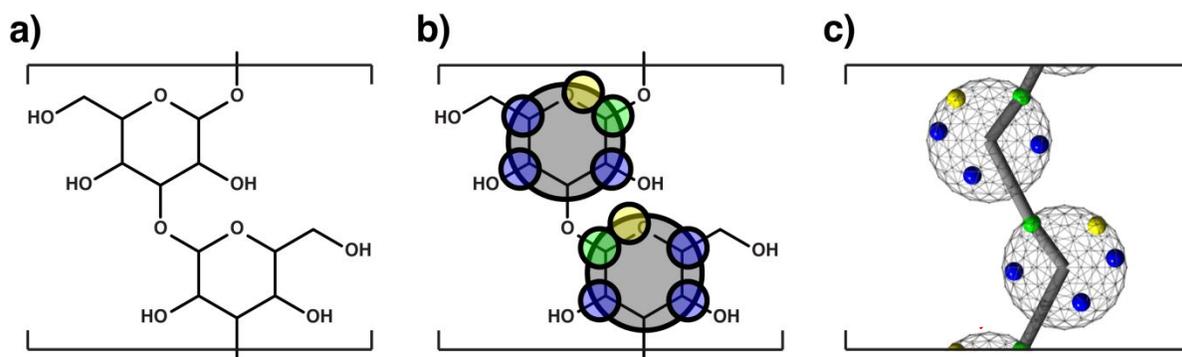


Figure 1. Chemical structure of α -1,3-glucan repeat unit and the CG model developed in this paper. a) Chemical structure, b) 2D representation of backbone (gray), donor (blue), acceptor (yellow), and linker (green) beads in the CG model, and c) 3-D view of the CG model. The reader can view supplementary information movie S.1 which shows the image in panel c) rotate.

A.1 Bonded interactions

The position of hydrogen bonding beads with respect to the center of the backbone bead is chosen to best represent the relative position of the hydrogen bonding sites in the monosaccharide. This can be tuned for different polysaccharide configurations such as cellulose, amylose, etc. For the remainder of the discussion, the OH sites will be denoted by the letters *DC* followed by the number of the carbon to which the OH site is bonded. The oxygen atom in the monosaccharide ring will be denoted by the letters *AR*. The linker oxygen in the glycosidic bond, denoted by *L*, acts as a hinge keeping contiguous monomers together in a similar fashion as in the CG model by Markutsya and collaborators³⁰. A short atomistic simulation of short oligosaccharide chains (see electronic supplementary information, ESI, section S.A. and section II.B.1) is used to determine the relative positions of all atoms in the most stable state. The backbone bead is located at the center of mass of the monomer ring and is used as a reference position to define the positions of other CG beads. The positions of DC2 and DC4 attached directly to the ring are determined by calculating the position vector from the center of mass of the ring to the oxygen in the OH group. The position of DC6 attached to a substituent carbon in α -1,3-glucan is determined by calculating the position vector from the center of mass of the ring to the *carbon* to which that OH group is attached. The position of the AR bead relative to the L bead is determined by finding the average of the center of mass-glycosidic oxygen vector and the center of mass-ring oxygen vector. The L bead position is determined by the vector from the center of mass of the monomer ring to the glycosidic oxygen. All the hydrogen-bonding beads are fixed *rigidly* at 0.445σ relative to the backbone bead, as seen in Figure 2a. The combination of the size of the hydrogen bonding beads and 0.445σ relative placement ensures that the hydrogen bonding beads are mostly *within* the backbone bead with only part of the hydrogen-bonding bead being exposed to interact with another hydrogen-bonding bead. This ensures the directionality of hydrogen bonds formed by said beads. A monomer is thus comprised of a backbone, DC2, DC4, DC6, AR, and L beads.

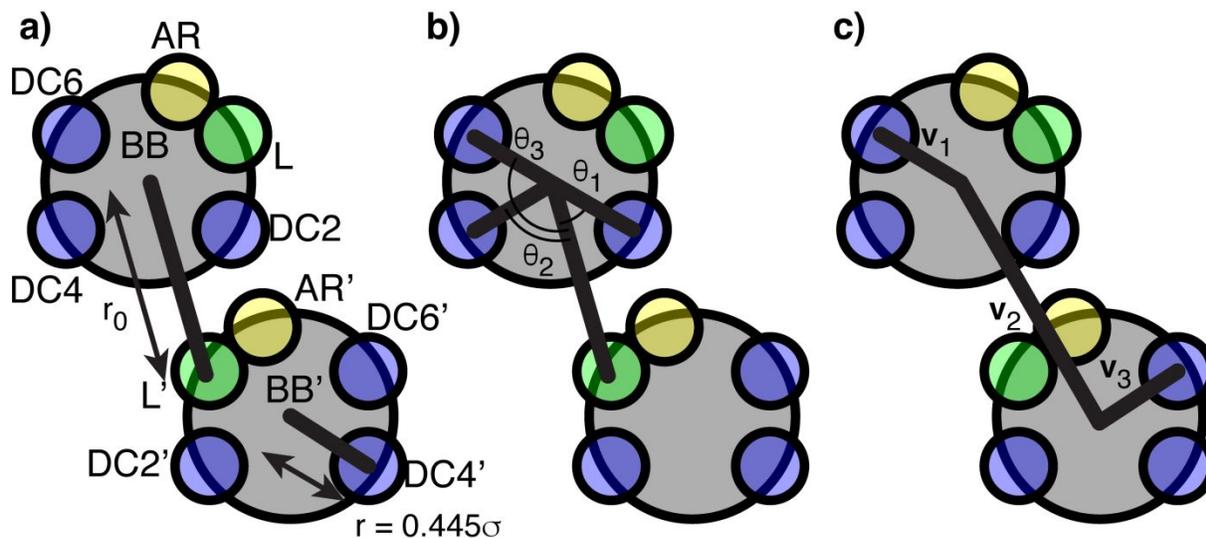


Figure 2. Schematic representation of DC, AR and L bead placement in the CG model of α -1,3-glucan. a) Relative position of DC, AR and L beads with respect to the backbone bead in the same monomer and relative position of the L bead with respect to the backbone bead in the adjacent monomer. b) Relative orientation of a L bead relative to the adjacent monomer shown as anchor angles between linker-adjacent backbone-DC beads. c) Beads representing the improper dihedral potential to maintain restriction on torsional motion between adjacent monomers.

To connect adjacent monomers, the L bead in one rigid body is connected to the *following* backbone bead (the center of the monomer) by a bond as visualized in Figure 2a. These monomer-linker bonds are modeled using the harmonic potential given by

$$U_{bond}(r) = k(r - r_0)^2 \quad (1)$$

where k represents the spring force constant. For backbone-linker bonds, the force constant is set to 2500 ϵ/σ^2 to model the rigidity of glycosidic linkages. The equilibrium bond length is set to $r_0 = 0.75\sigma$, which is determined by the relative position of adjacent backbone beads in atomistic simulation.

To maintain the relative orientation of a linker bead with respect to the adjacent monomer – and ensure hinge-like connectivity between contiguous monomers – we define three angle potentials as shown in Figure 2b : (a) [L]-[BB']-[DC2'], (b) [L]-[BB']-[DC4'], and (c) [L]-[BB']-[DC6'], where X' is the X site on adjacent monomer. These are modeled using a harmonic potential given by

$$U_{angle} = k_\theta(\theta - \theta_0)^2 \quad (2)$$

here the angle force constant k_θ is set to 2500 ϵ/rad^2 to minimize deviation of the linker's position with respect to the adjacent donors. The equilibrium angle θ_0 is set to 63.0°, 62.4°, and 123.1°, for [L]-[BB']-[DC2], [L]-[BB']-[DC4], [L]-[BB']-[DC6], respectively, and are determined from atomistic simulations. The combination of the rigid body representation of these sites on the monomer and these angle potentials ensures that the chirality of each monomer is conserved.

To restrict rotational motion about the glycosidic bond, a dihedral potential is introduced. The dihedral angle is the angle between the planes formed by [DC6']-[BB']-[BB] and [BB']-[BB]-[DC6], as shown in Figure 2c. This dihedral angle potential is based on the deviation of the dihedral from an equilibrium angle of $\psi_0 = 180^\circ$ and is defined by a harmonic improper dihedral potential as

$$U_{dihedral} = K_\psi(\psi - \psi_0)^2. \quad (3)$$

The value of K_ψ is tuned for α -1,3-glucan to mimic structural features seen in atomistic simulations as described in the simulation protocol in section II.B.2. Similarly, an angle potential between three consecutive L beads is introduced as an additional tuning parameter to capture atomistically observed polysaccharide stiffness. The parameters k_θ and θ_0 of this angle potential are tuned for α -1,3-glucan. Table 1 summarizes all the bonded interaction parameters for our CG model of α -1,3-glucan.

Table 1: Bonded interaction parameters for CG model of α -1,3-glucan.

Type of interaction	Participating beads	Parameters	Equation
Bond	L-BB'	$k = 2500 \text{ } \epsilon/\sigma^2, r_0 = 0.75\sigma$	(1)
Angle	L-BB'-DC6'	$k_\theta = 2500 \text{ } \epsilon/\text{rad}^2, \theta_0 = 123.1^\circ$	(2)
	L-BB'-DC4'	$k_\theta = 2500 \text{ } \epsilon/\text{rad}^2, \theta_0 = 62.4^\circ$	
	L-BB'-DC2'	$k_\theta = 2500 \text{ } \epsilon/\text{rad}^2, \theta_0 = 63.0^\circ$	
	L-L'-L''	$k_\theta = 60 \text{ } \epsilon/\text{rad}^2, \theta_0 = 180^\circ$	
Improper dihedral	DC6-BB-BB'-DC6'	$k_\psi = 60 \text{ } \epsilon/\text{rad}^2, \psi_0 = 180^\circ$	(3)

A.2 Non-bonded interactions

To represent all possible hydrogen bonding in the monomer (O-H---O and H-O---HO), the DC beads interact attractively with other DC beads and with AR beads. These attractive interactions are modeled by a Lennard-Jones (LJ) potential given by

$$U^{HB}(r) = 4\epsilon_{AD} \left[\left(\frac{\sigma^{HB}}{r} \right)^{12} - \left(\frac{\sigma^{HB}}{r} \right)^6 \right]. \quad (4)$$

The depth (minimum) of the above potential well is the effective hydrogen bond strength, ϵ_{AD} , which occurs at a separation of $2^{1/6}\sigma^{HB}$. To model the effect of changing temperature and/or changing solvent types we vary the hydrogen bond strength, ϵ_{AD} . Despite the isotropic interaction potential between

hydrogen bonding DC and AR (or DC) beads, our model ensures that the hydrogen bond interaction is effectively directional because of the size and placement of hydrogen bonding beads within the backbone bead, as specified in section II.A.1. Furthermore, specificity is accomplished by a) a smoothing function (GROMACS⁴⁴) used to ramp the potential to zero from $1.9\sigma^{\text{HB}}$ to $2.0\sigma^{\text{HB}}$ to ensure no donor/acceptor trimers are formed, and b) a repulsive acceptor-acceptor Weeks-Chandler-Andersen⁴⁵ (WCA) potential described as

$$U_{\text{repulsive}}(r) = 4\varepsilon_{\text{rep}} \left[\left(\frac{2.3\sigma^{\text{HB}}}{r} \right)^{12} - \left(\frac{2.3\sigma^{\text{HB}}}{r} \right)^6 \right] + \varepsilon_{\text{rep}}. \quad (5)$$

The cutoff for the above repulsive potential is set to $(2^{1/6}) * 2.3\sigma^{\text{HB}}$ to ensure the interactions are purely repulsive. When a hydrogen bonding pair forms, the repulsive shell will prevent additional donors/acceptors to approach. The strength of the repulsive shell is set to $\varepsilon_{\text{rep}} = 1.0\varepsilon$ (=0.1 kcal/mol). In the cases explored in this paper we find no evidence for the formation of trimers of hydrogen bonding sites.

All other non-bonded pair-wise interactions between all other beads are modeled by the WCA potential²⁰, given by

$$U_{ij}(r) = 4\varepsilon_{\text{WCA}} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right] + \varepsilon_{\text{WCA}}, \quad (6)$$

where $\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}$, cutoff set to $2^{1/6}\sigma_{ij}$, and $\varepsilon_{\text{WCA}} = 1.0\varepsilon$ (0.1 kcal/mol). The choice of using repulsive WCA potential for the remaining bead pairs is justified by the fact that main driving force for polysaccharide chain aggregation is hydrogen bonds.

B. Simulations details

B.1 Atomistic simulations

We run atomistic simulations of α -1,3-glucan not only to guide the relative placements of the DC, AR and L beads but also to tune the CG model parameters K_{ψ} , and the k_{θ} and θ_0 of the angle potential between three consecutive L beads. Specifically, the stiffness of the polymer chains (i.e., persistence length) for different degrees of polymerization, DP , are calculated and the values of the K_{ψ} , and the k_{θ} and θ_0 in coarse-grained simulations that best match the atomistic persistence length are chosen. Atomistic simulations are run using the GROMACS⁴⁶ (version 4.6.7) simulation package. The initial structure for 7-mer, 10-mer, and 16-mer α -1,3-glucan chains is generated with Maestro (Schrödinger Release 2018–4: Maestro, Schrödinger, LLC, New York, NY, 2018) and minimized with the OPLS forcefield⁴⁷. Polysaccharide chains are modeled with Antechamber using the Generalized Amber Forcefield (GAFF)⁴⁸.

⁴⁹. Partial charges are assigned using the AM1-BCC method⁵⁰. Water is modeled with the SPC/E forcefield⁵¹. Simulations are run for a total of 10 ns, with a time step of 1 fs at a temperature of $T = 298$ K and a pressure of $P = 1$ bar in the NPT ensemble. The cubic simulation box contains a single 7-mer, 10-mer, or 16-mer polymer chain with 16895, 17062, and 57497 water molecules, respectively. The initial box lateral dimension is 8 nm for 7-mer and 10-mer, and 12 nm for 16-mer. Results for the mapping of L-L-L angle potential in the CG model from the atomistic simulation results are shown in the ESI, section S.B.

B.2 Coarse-grained simulations

The CG model is used in Langevin dynamics (MD) simulations run using the LAMMPS package⁵² (April 2015 version). The simulation is run in a periodic cubic box in an NVT ensemble with a Langevin thermostat. All simulations are run at a temperature of 298K (5.92 in reduced units). In Langevin dynamics simulations the frictional damping coefficient, γ , is set as 100 time steps. The initial configuration of the polysaccharides in the cubic simulation box is created by a random placement of all the chains across the simulation box with the number of chains and box size chosen to mimic a specific concentration of the solution. The number of chains in a simulation box is equal to 100. Initial velocities of the beads are randomized at the temperature of the production run. The time step is set to $\Delta t = 0.001$ (in reduced units). The dihedral potential strength, K_ψ , is initially set to an arbitrarily small value of $0.1 \text{ } \epsilon/\text{rad}^2$ and increased stepwise over 4×10^4 time steps until the desired production value is reached. All other bond and angle potential strength constants are initially set to the value described in Table 1.

The production run of a given set of parameters is done by initially setting the hydrogen bond strength, ϵ_{AD} , to an arbitrarily small value of 30.0ϵ (3 kcal/mol) and stage-wise increasing ϵ_{AD} by intervals of 1.0ϵ (0.1 kcal/mol) every 4×10^6 time steps, till a final value of 75ϵ (7.5 kcal/mol). This range of ϵ_{AD} represents the typical hydrogen bond strengths.^{53, 54} The stage-wise increase in ϵ_{AD} is chosen over a quench (i.e., a “step” function from the minimum to the maximum ϵ_{AD}) to avoid kinetically trapped states and to explore the equilibrium structures that α -1,3-glucan forms at different ϵ_{AD} . The rationale behind the choice of the number of steps per ϵ_{AD} is discussed in detail in section S.C. in the ESI, Briefly, we keep the ϵ_{AD} increment to be 0.1 kcal/mol and increase the number of timesteps at each ϵ_{AD} till the aggregation behavior, discussed in the next section, as a function of ϵ_{AD} does not change with any additional number of time steps. This confirms that we are reporting analysis of equilibrium structures. Trajectories of all beads are saved in each stage so as to have 500 frames of data for analysis.

C. Analyses

C.1 Aggregation behavior and cluster identification

We quantify the aggregation behavior of polymer chains by calculating the number of polymer chains that belong to clusters. Two chains are aggregating if there is at least one hydrogen bond between their hydrogen bonding sites that are within 0.2σ distance and they belong to the same cluster. Additionally, any polymer chain aggregating with *any* member of a cluster belongs to the said cluster. A friends-of-friends algorithm is used to identify all aggregated chains belonging to a cluster⁵⁵. We record the number of chains in each of the clusters, N_{cluster} . N_{cluster} is recorded for the last 250 frames in each stage at constant ε_{AD} for three simulation replicas to determine both an average N_{cluster} corresponding to a given ε_{AD} , $\langle N_{\text{cluster}} \rangle$, and its standard deviation.

C.2 Hydrogen bond distribution

We calculate the number and type of hydrogen bonds – two hydrogen bonding sites within 0.2σ of each other – by looping through all pairs of hydrogen bonding sites. Hydrogen bonds occur within a chain (intra-chain bond) or between different chains (inter-chain bond). We determine the frequency of hydrogen bonds, $f_{\text{H-bond}}$, by counting all bonds of a given type (intra- or inter-chain and between different pairs of hydrogen-bonding sites) and normalizing the count by number of frames (the last 250 frames in a stage at constant ε_{AD}), number of replicas, and total number of monomers in a simulation box.

C.3 Single chain conformation and shape

We determine the conformation of α -1,3-glucan chains by calculating the gyration tensor, \mathbf{S} , of each chain as

$$S_{mn} = \frac{1}{N} \sum_i^N r_m^{(i)} r_n^{(i)}, \quad (7)$$

where N is the total number of monomers in a chain, m and n represent the cartesian coordinates along which the position vector r from the center of mass of the chain is calculated. The gyration tensor \mathbf{S} is a 3x3 matrix in the Cartesian coordinates x, y, z such that its eigenvalues – $\lambda_1, \lambda_2, \lambda_3$ – are positive real values with corresponding eigenvectors v_1, v_2, v_3 that represent the principal directions of mass distribution. The radius of gyration of a chain is calculated from the gyration tensor eigenvalues as

$$R_g^2 = \lambda_1 + \lambda_2 + \lambda_3. \quad (8)$$

From the gyration tensor we also calculate the relative shape anisotropy, κ^2 , as

$$\kappa^2 = \frac{3}{2} \frac{\lambda_1^4 + \lambda_2^4 + \lambda_3^4}{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)^2} - \frac{1}{2}. \quad (9)$$

κ^2 provides a measure of the shape of chains, with $\kappa^2 = 0$ representing a perfect spherical shape, $\kappa^2 = 1$ representing a perfectly linear chain, and $\kappa^2 = 0.25$ representing a planar configuration. We report κ^2

values for the last 250 frames of a given stage at constant ε_{AD} over 3 simulation replicas to provide an average and standard deviation.

C.4. Chain alignment angle

To determine the parallel/antiparallel orientation of neighboring chains within an aggregate, we calculate the alignment angle, Θ , between close-by monomers. This is done by first computing the monomer director vector, \mathbf{u} , taken as the unit vector in the direction between the two linkers bracketing a monomer (i.e. the linker that forms part of the rigid body, and the bonded linker of the following monomer). Θ is calculated as

$$\Theta = \text{acos}(\mathbf{u}_i \cdot \mathbf{u}_j). \quad (10)$$

We calculate Θ for both intra and inter-chain alignment for monomers that have a maximum BB-to-BB distance of 5σ . The chain alignment angle provides information about the relative alignment between monomers within a chain, and between monomers in close-by chains. For instance parallel alignment of chains results in Θ distributions at angles close to 0° , and anti-parallel alignment of chains results in Θ distributions at angles close to 180° .

D. Design space – parameters varied

We consider five values of degree of polymerization, DP , defined as the number of monomers per chain, 5-mer, 7-mer, 10-mer, 14-mer, and 20-mer. We also consider two values of polymer concentration, c_p , at 60 mg/mL and 10 mg/mL. We present our results for $c_p = 60$ mg/mL in the paper and present the $c_p = 10$ mg/mL results in the ESI, sections S.C, S.D, and S.E. The rationale for the range of DP and c_p is to explore the solubility and structure of α -1,3-glucan chains and the structure of aggregates formed for different chain lengths and at dilute to moderate concentrations.

We also explore the effect of substitution/silencing of donor sites. A given donor site is substituted or silenced by “turning off” its attractive potential and replacing it with a purely repulsive WCA interaction with all other beads. Experimentally one could silence a donor site by say, methylation of the OH group to prevent hydrogen bond formation in that particular site. The degree of substitution, DS , is defined as the average number of substituted donors *per* monomer and varies from 0 (no substitution) to 3 (all three DC sites are substituted). We perform two different kinds of substitution: (1) random substitution, where a random set of DC sites (the number of them corresponding to the desired DS value) in the entire simulation box are chosen for substitution, and (2) specific or targeted substitution, where all donors of one type – i.e. DC2, DC4, or DC6 – are substituted. For random substitution we use

DS values of 0.25, 0.5, 0.75, 1, 1.25, and 1.5. For targeted substitution, given all donors of one type are substituted, the equivalent DS value is 1.

III. Results

A. Effect of hydrogen-bonding strength and degree of polymerization on aggregation for unmodified α -1,3-glucan

In this section we present results for unmodified α -1,3-glucan at varying DP (5-mer, 10-mer, 20-mer) at a concentration $c_p = 60$ mg/mL. Results for 7-mer and 14-mer chains, as well as results for $c_p = 10$ mg/mL can be found in the ESI, section S.D, and follow similar trends as the results presented herein.

A.1 Aggregation behavior for unmodified α -1,3-glucan

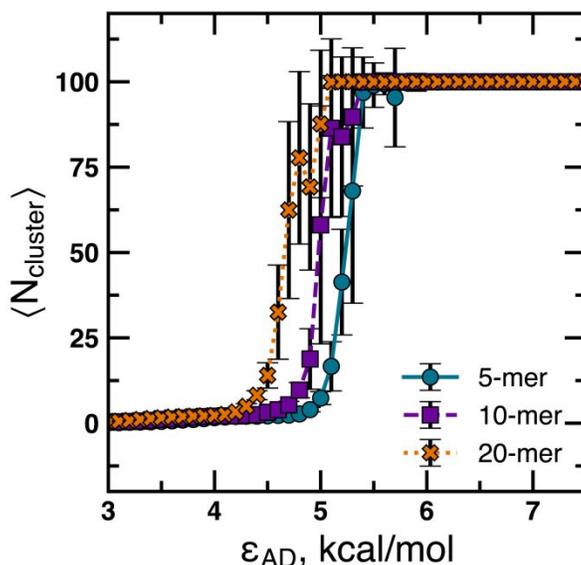


Figure 3. Effect of hydrogen bonding strength, ϵ_{AD} , and degree of polymerization on average number of chains per cluster, N_{cluster} , of α -1,3-glucan at $c_p = 60$ mg/mL. Error bars denote standard deviation over 3 replicas and 250 frames per replica. The total number of chains in a simulation box is 100.

We plot the average number of chains belonging to a cluster, $\langle N_{\text{cluster}} \rangle$, as defined in section II.C.1, as a function of the hydrogen bond strength, ϵ_{AD} , in Figure 3. Our results show that $\langle N_{\text{cluster}} \rangle$ for all DPs increases with increasing ϵ_{AD} from ~ 0 when chains are well dispersed – to ~ 100 when all chains in the simulation box assemble into one cluster. Our results also show that the ϵ_{AD} required to cause aggregation, ϵ_{ADt} , identified as the ϵ_{AD} that results in $\langle N_{\text{cluster}} \rangle$ equal to 10% of the maximum $\langle N_{\text{cluster}} \rangle$ value, decreases with increasing chain length – $\epsilon_{\text{ADt}} \sim 5.0$ kcal/mol for 5-mers, $\epsilon_{\text{ADt}} \sim 4.8$ kcal/mol for 10-mers, and $\epsilon_{\text{ADt}} \sim 4.4$ kcal/mol for 20-mers. The decrease in ϵ_{ADt} with increasing DP is expected, and consistent with observations of molecular weight-dependent solubility of many different polymers.⁵⁶ The above

results show that taking a typical^{53, 54} value for the hydrogen bond strength, $\epsilon_{AD} = 4.7$ kcal/mol, results in shorter chains being well dispersed, while longer chains, 10-mers and up, are insoluble. Recent experimental work on the solubility of synthetic α -1,3-glucan confirms that the polysaccharides are insoluble for $DP > 8$.⁵⁷

The information presented in Figure 3 is a good example of the capabilities of our CG model, where a large number of polymer chains are simulated to predict the solubility, or insolubility, of polysaccharide solutions. In experiments one could vary the hydrogen bonding interaction strength by changing solvent chemistry and/or adding salts. One may need to map/evaluate the effect of solvent chemistry on the hydrogen bond strength to find the exact solvent condition that corresponds to the specific value of ϵ_{AD} . Furthermore, our CG model provides invaluable information that is not as readily available from experimentation, including polysaccharide chain conformations and aggregate structure, which will be described in the following sections.

A.2 Single chain structure

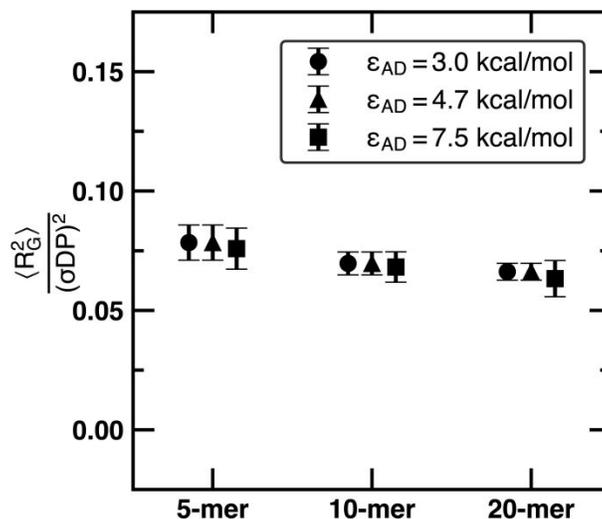


Figure 4. Average squared radius of gyration of each chain normalized by contour length (σDP) at $c_p = 60$ mg/mL. Error bars denote standard deviation over 3 replicas and 250 frames per replica.

We characterize chain conformations by calculating the average squared radius of gyration of each chain, R_g^2 , and its relative shape anisotropy, κ^2 , as defined in section II.C.3, as a function of ϵ_{AD} and DP in Figures 4 and 5. Our results in Figure 4 show that the chain $\langle R_g^2 \rangle$ decreases slightly with increasing ϵ_{AD} for all DP and decreases with increasing DP but both changes are small and differences are within a standard deviation. By normalizing radius of gyration by the contour length, σDP , we see that $\langle R_g^2 \rangle / (\sigma DP)^2 \sim 1/12$, which corresponds to fully extended chains⁵⁸. To confirm that chains are indeed

extended and rod-like, we calculate κ^2 in Figure 5, and find that $\kappa^2 > 0.5$, which corresponds to mainly 1-dimensional shapes, as discussed in section II.C.3. Our results are consistent with the relatively stiff α -1,3-glucan chains we find in atomistic simulations, as described in the ESI, section S.B. Namely the persistence length of α -1,3-glucan chains is ~ 5 nm, and the longest contour length we explore is ~ 8 nm for 20-mer chains. As the contour lengths of the chains we explore are close to the persistence length of the polymer or slightly less, one can expect to see extended rod-like conformations.

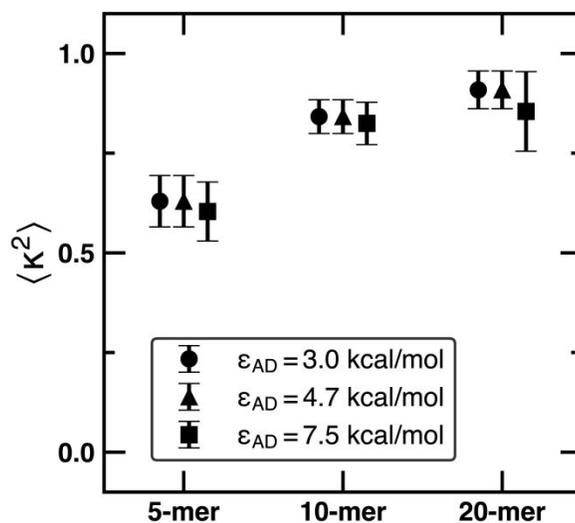


Figure 5. Average relative shape anisotropy, κ^2 , for each chain at $c_p = 60$ mg/mL. Error bars denote standard deviation over 3 replicas and 250 frames per replica.

A.3 Assembled state structure

Next, we quantify the frequency of hydrogen bonds, $f_{\text{H-bond}}$, as defined in section II.C.2; this is shown in Figure 6 as well as in Figure 7 simulation snapshots. Figure 6a shows all the possible intra-chain and inter-chain hydrogen bonds one could form with our model and Figure 6b and 6c show the frequency of each of these hydrogen bonds at $\epsilon_{AD} = 7.5$ and 4.7 kcal/mol, respectively. We see that $f_{\text{H-bond}}$ of all hydrogen bond pairs increases with increasing DP and ϵ_{AD} . Specifically, the 5-mer chains form very few inter-chain hydrogen bonds, while 10-mer and 20-mer chains form primarily DC2-DC4, DC2-DC6, DC6-AR, and DC4-AR hydrogen bond pairs. In our discussion on the aggregation behavior in section III.A.1 we mentioned that short chains, such as 5-mers, at $\epsilon_{AD} = 4.7$ kcal/mol remain well dispersed, and the few inter-chain hydrogen bonds, seen in Figure 6c, confirms the observation. The $f_{\text{H-bond}}$ of all hydrogen bond pairs are higher at $\epsilon_{AD} = 7.5$ kcal/mol than those seen for $\epsilon_{AD} = 4.7$ kcal/mol. The higher $f_{\text{H-bond}}$ with increasing ϵ_{AD} is expected; as the enthalpic driving force for hydrogen bonding increases, more hydrogen bonds are formed and there is a higher $\langle N_{\text{cluster}} \rangle$. The main hydrogen bond pairs observed at $\epsilon_{AD} = 7.5$ kcal/mol are the same as for $\epsilon_{AD} = 4.7$ kcal/mol.

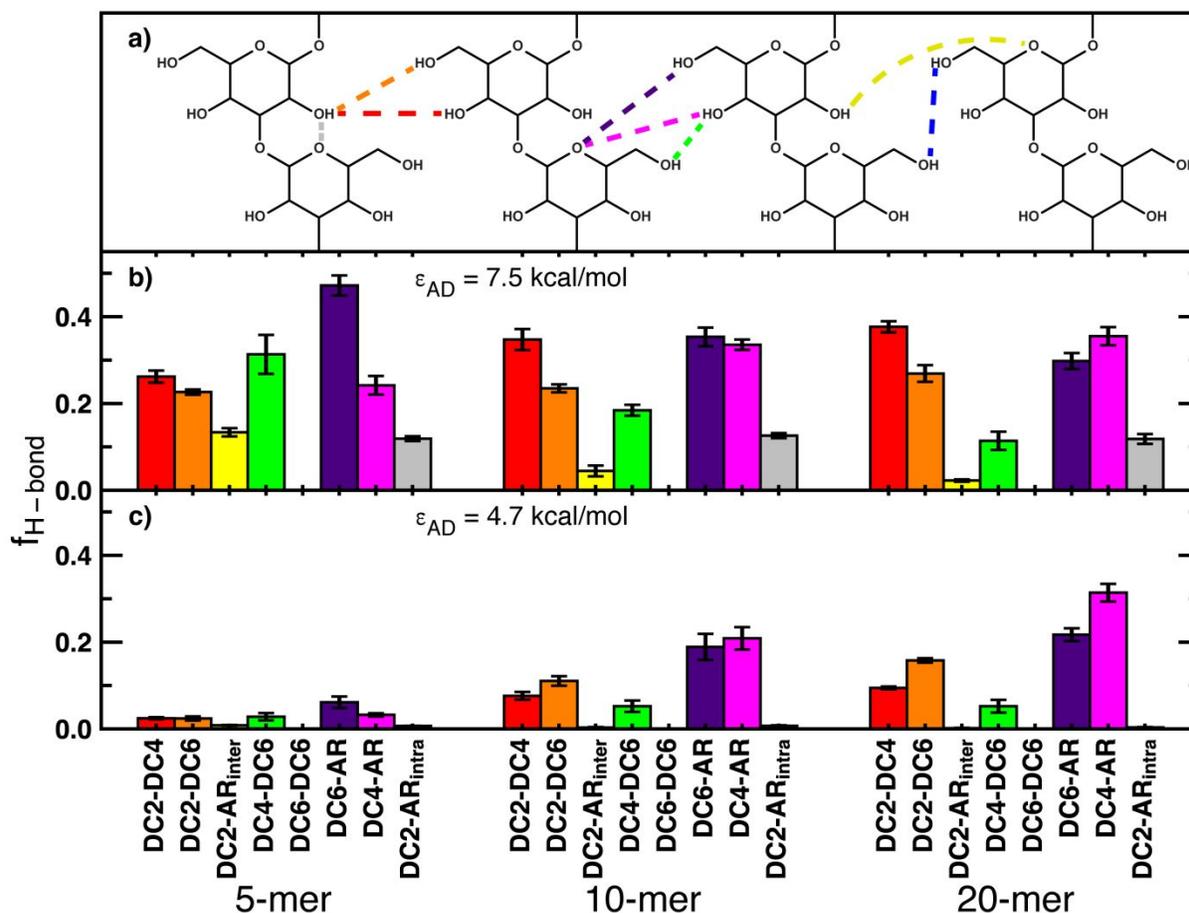


Figure 6. Hydrogen bonding pattern for α -1,3-glucan at $c_p = 60$ mg/mL. a) schematic of different hydrogen bonds colored according to bar colors in panels b and c. b) and c) Frequency of specific hydrogen bonds being formed for 5-mer, 10-mer, and 20-mer chains at hydrogen bonding strength $\epsilon_{AD} = 7.5$ kcal/mol (panel b), and $\epsilon_{AD} = 4.7$ kcal/mol (panel c). Error bars denote standard deviation over 3 replicas.

Figure 7 provides a closer look at the kinds of structures observed for unmodified α -1,3-glucan. We present here results for the 20-mer chains, but structures are similar to those formed by chains with different DP, as seen in the ESI, Figures S12 and S13. Figure 7a shows the sheet-like nature of the aggregates. This sheet-like/layer structure is consistent with experimentally observed chain arrangement in the four known α -1,3-glucan crystal polymorphs^{7, 59-61} which consist of layers of polysaccharide chains held together by hydrogen bonds. It is also consistent with recent TEM imaging of crystalline synthetic α -1,3-glucan assembled from solution that show fibrils or lamellae formed through further assembly of these layers.⁶² In Figure S15 the reader can see a reproduction of the aforementioned TEM images and the schematic representation of α -1,3-glucan chain arrangement⁶². A closer inspection of one such aggregate shows that chains are aligned side-by-side, as seen in Figure 7b and in Figure 7c and held together by the hydrogen bonds shown in Figure 6. The chains in 2-dimensional sheets display some splay (see Supplementary Figure S16) and arrange in both parallel and antiparallel arrangements (Figure 7d). We

hypothesize at this point that hydrogen bonds involving the DC6 site are responsible for this arrangement; the DC6 site is significantly out of the plane that is formed by other hydrogen bonding sites and linker beads in a monomer, as seen in Supplementary Figure S17. Thus, substitution/silencing of the DC6 site, either randomly or specifically, may alter this arrangement. This is investigated in the next two sections.

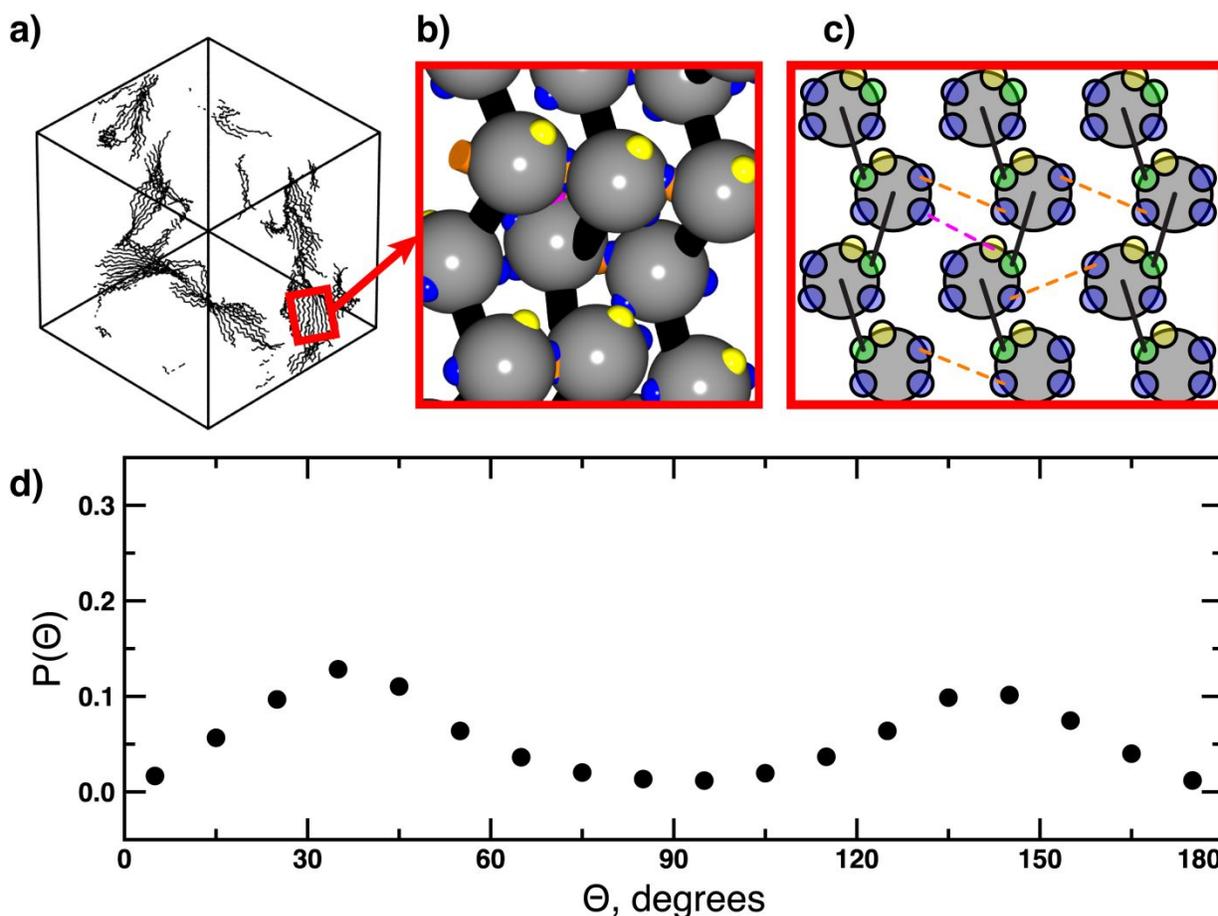


Figure 7. Simulation box snapshot and detail of structure formed by α -1,3-glucan 20-mer chains at $\epsilon_{AD} = 4.7$ kcal/mol and $c_P = 60$ mg/mL. a) Simulation rendering showing only the backbones of polymer chains in the aggregated state. b) Detail of side-by-side arrangement of chains following a similar rendering scheme as figure 1c, as well as example hydrogen bonds depicted using the same color scheme as Figure 6a. c) Schematic representation of panel b to show the hydrogen bonds formed with the same coloring scheme as panel b. d) Inter-chain alignment angle, Θ , distribution between near-by monomers.

B. Effect of random substitution of hydrogen bonding sites in the monomer on the aggregation of α -1,3-glucan chains

Random substitution is accomplished by selecting a random set of hydrogen bonding sites (DC2, DC4, or DC6) from the simulation box and “turning-off” their attractive potential. Specifically, we present here results for $DS = 0.5, 1.0,$ and $1.5,$ for 5-mer, 10-mer, and 20-mer chains and $c_P = 60$ mg/mL

and show the results for the expanded parameter space $DS = 0.25, 0.5, 0.75, 1.0, 1.25, \text{ and } 1.5$, 5-mer, 7-mer, 10-mer, 14-mer, and 20-mer chains in the ESI, section S.E.

B.1 Aggregation behavior for randomly substituted α -1,3-glucan

Figure 8 shows $\langle N_{\text{cluster}} \rangle$ as a function of ϵ_{AD} for 20-mer chains for the unmodified α -1,3-glucan, and for $DS = 0.5, 1.0, \text{ and } 1.5$. As was the case for the unmodified polymer shown in Figure 3, $\langle N_{\text{cluster}} \rangle$ increases with increasing ϵ_{AD} . Here, the ϵ_{ADt} increases with increasing DS – $\epsilon_{\text{ADt}} = 4.4$ kcal/mol for unmodified polymer, $\epsilon_{\text{ADt}} = 4.6$ kcal/mol for $DS = 0.5$, $\epsilon_{\text{ADt}} = 5.0$ kcal/mol for $DS = 1.0$, and $\epsilon_{\text{ADt}} = 5.4$ kcal/mol for $DS = 1.5$. The increase in ϵ_{ADt} with increasing DS is expected because with fewer sites available for hydrogen bonding, additional hydrogen bonding strength is needed to form similar aggregates. Our results suggest that random substitution results in aggregation behavior that is qualitatively similar to that of the unmodified polymer. Characterization of hydrogen bonding pattern of the structures formed by randomly substituted α -1,3-glucan is presented next.

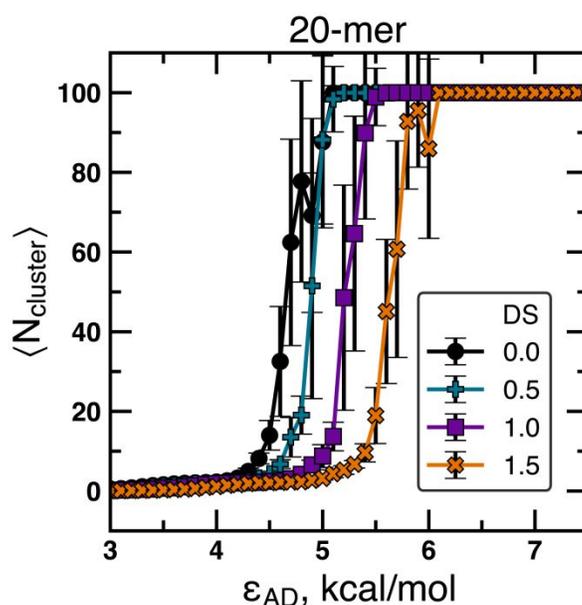


Figure 8. Effect of hydrogen bonding strength, ϵ_{AD} , on number of chains per cluster, $\langle N_{\text{cluster}} \rangle$, for randomly substituted α -1,3-glucan for 20-mer chains at $c_p = 60$ mg/mL. Error bars denote standard deviation over 3 replicas and 250 frames per replica.

B.2 Assembled state structure for randomly substituted α -1,3-glucan

The quantification of $f_{\text{H-bond}}$ for the cases with random substitution is presented in Figures 9 and 10 for $\epsilon_{\text{AD}} = 4.7$ kcal/mol and 7.5 kcal/mol, respectively. The overall $f_{\text{H-bond}}$ increases with increasing DP and with decreasing DS for $\epsilon_{\text{AD}} = 4.7$ kcal/mol, as shown in Figure 9. The increase in $f_{\text{H-bond}}$ with increasing DP is related to the aggregation state of 5-mer, 10-mer, and 20-mer chains at $\epsilon_{\text{AD}} = 4.7$ kcal/mol, seen in Figures S24 and S25, where 5-mer chains are well dispersed, while 20-mer chains are

aggregated; a higher $\langle N_{\text{cluster}} \rangle$ means more hydrogen bonds and a higher $f_{\text{H-bond}}$. It is not surprising that $f_{\text{H-bond}}$ decreases with increasing DS , as fewer sites of all types are available to form hydrogen bonds. The overall $f_{\text{H-bond}}$ decreases with increasing DS and remains fairly constant for all DP for $\epsilon_{\text{AD}} = 7.5$ kcal/mol, as seen in Figure 10. As with results for $\epsilon_{\text{AD}} = 4.7$ kcal/mol, it is not surprising that $f_{\text{H-bond}}$ decreases with increasing DS . The insensitivity of $f_{\text{H-bond}}$ with DP stems from the fact that at $\epsilon_{\text{AD}} = 7.5$ kcal/mol, for all cases explored here, all chains are part of the same cluster, as seen in Figures S18 and S19. The hydrogen bond pattern remains fairly unchanged from the unmodified polymer case, with the main hydrogen bonds formed by the same pairs as the unmodified polymer, shown in Figure 6. Both the aggregation and the hydrogen bond pattern for random substitution is qualitatively similar to those of the unmodified polymer. The structure of the aggregates formed with randomly substituted α -1,3-glucan is qualitatively similar as those formed with unmodified polymer. Simulation renderings of said structures are shown in Figures S28 through S31.

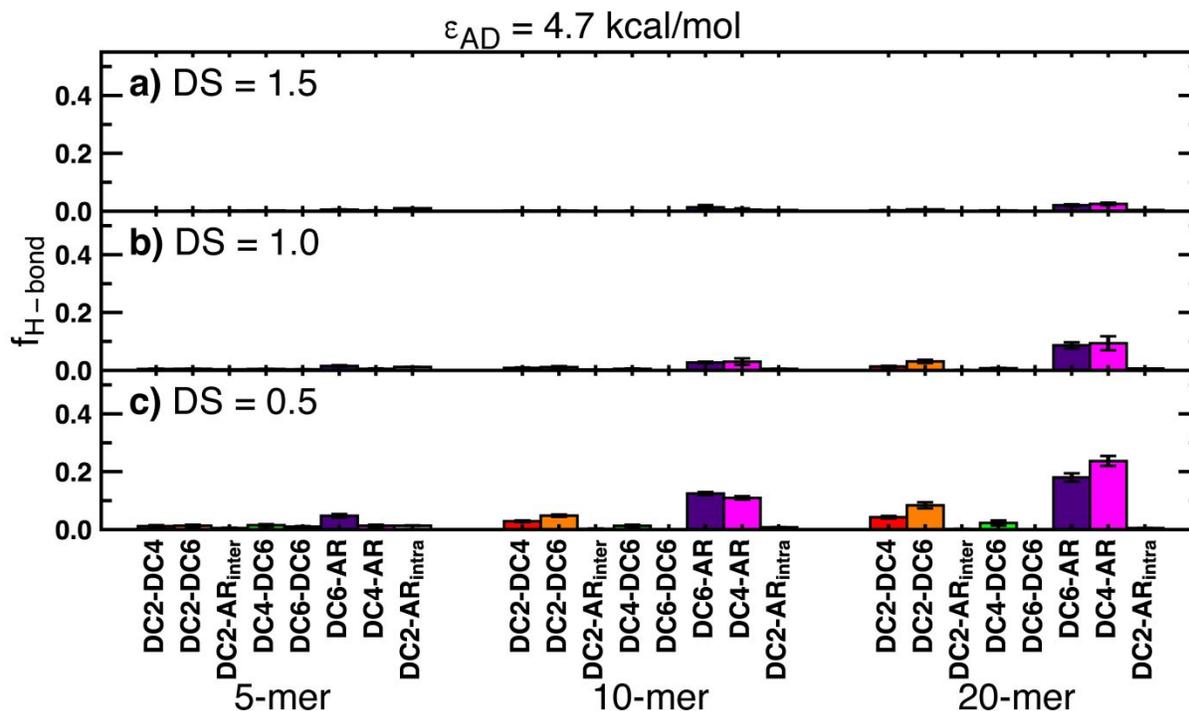


Figure 9. Hydrogen bonding pattern for randomly substituted α -1,3-glucan at $c_p = 60$ mg/mL at hydrogen bonding strength $\epsilon_{\text{AD}} = 4.7$ kcal/mol. Frequency of specific hydrogen bonds being formed for 5-mer, 10-mer, and 20-mer chains at degree of substitution $DS = 1.5$ (panel a), $DS = 1.0$ (panel b), and $DS = 0.5$ (panel c). Error bars denote standard deviation over 3 replicas.

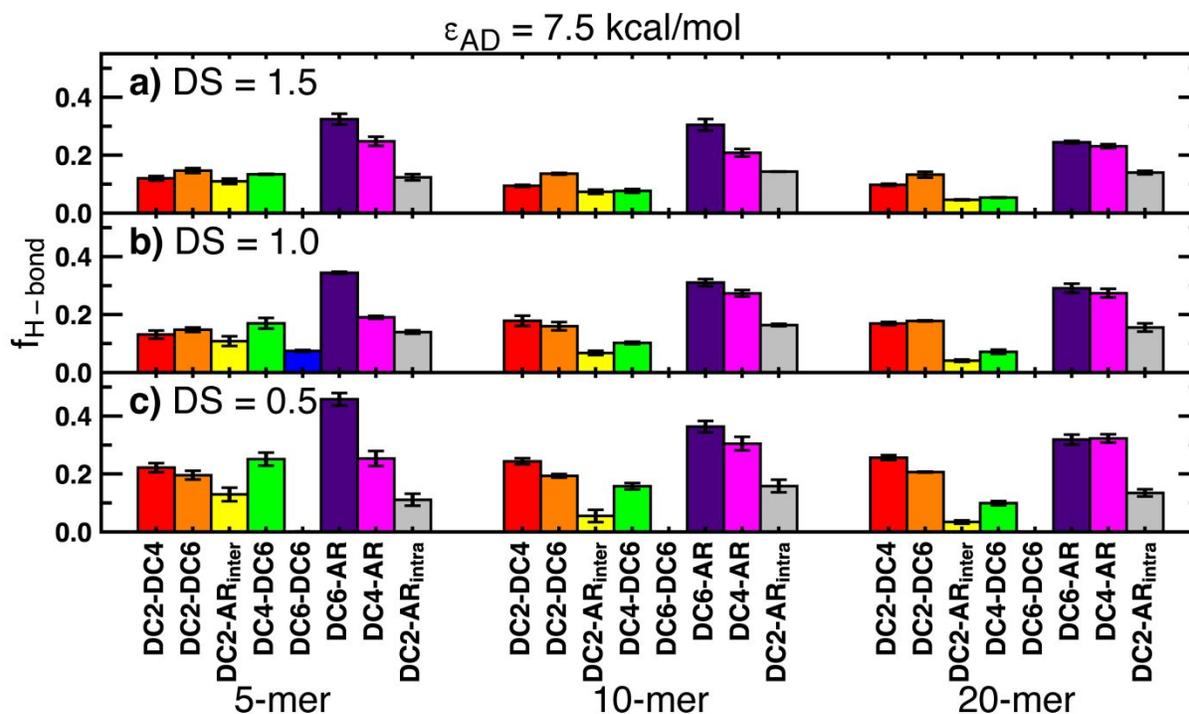


Figure 10. Hydrogen bonding pattern for randomly substituted α -1,3-glucan at $c_p = 60 \text{ mg/mL}$ at hydrogen bonding strength $\epsilon_{AD} = 7.5 \text{ kcal/mol}$. Frequency of specific hydrogen bonds being formed for 5-mer, 10-mer, and 20-mer chains at degree of substitution $DS = 1.5$ (panel a), $DS = 1.0$ (panel b), and $DS = 0.5$ (panel c). Error bars denote standard deviation over 3 replicas.

C. Effect of targeted substitution on the aggregation of α -1,3-glucan

Targeted substitution of all DC6 donor sites in a simulation box is accomplished by “turning-off” their attractive potential and modeling it as a repulsive-only interaction. The rationale for this type of substitution was mentioned in section III.A.3, where targeted substitution of the DC6 site (“silent” DC6) was hypothesized to likely prevent the splay of the sheet-like structures formed by unmodified α -1,3-glucan. We present here results for 5-mer, 10-mer, and 20-mer chains at $c_p = 60 \text{ mg/mL}$, but report in the ESI, section S.F, results for 5-mer, 7-mer, 10-mer, 14-mer, and 20-mer chains at $c_p = 10 \text{ mg/mL}$ and 60 mg/mL and for “silent” DC4 and “silent” DC2.

C.1 Aggregation behavior for targeted substitution of α -1,3-glucan

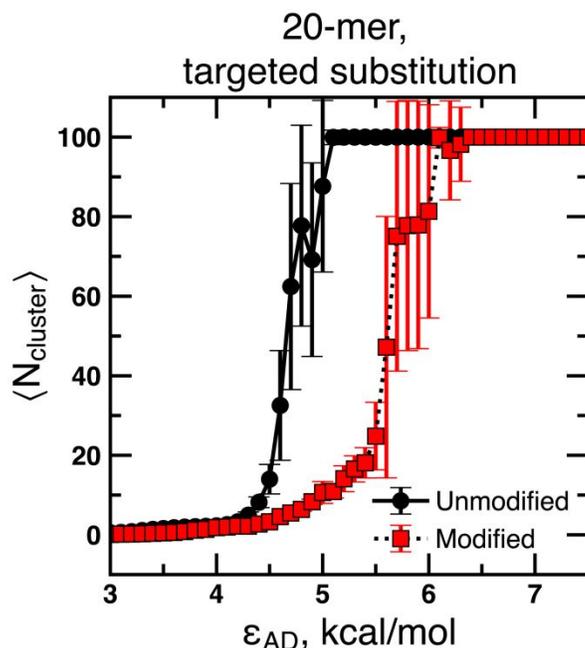


Figure 11. Effect of hydrogen bonding strength, ϵ_{AD} , on number of chains per cluster, $\langle N_{\text{cluster}} \rangle$, for targeted substitution of the DC6 site in α -1,3-glucan for 20-mer chains at $c_p = 60$ mg/mL. Error bars denote standard deviation over 3 replicas and 250 frames per replica.

Figure 11 shows $\langle N_{\text{cluster}} \rangle$ as a function of ϵ_{AD} for both unmodified and “silent” DC6 20-mer chains. Aggregation of “silent” DC6 chains occurs in a qualitatively different fashion as compared with the unmodified polymer and randomly substituted chains. First, a higher ϵ_{AD} is required to cause aggregation, consistent with our observations for random substitution of $DS = 1$, shown in Figure 8. Nonetheless we see aggregation in two regimes as opposed to the one transition from well solvated to fully aggregated seen in unmodified and randomly substituted chains. In the first regime seen in the range $\epsilon_{\text{AD}} \sim 4.7$ to 5.3 kcal/mol the $\langle N_{\text{cluster}} \rangle$ has a value of ~ 10 chains/cluster. In the second regime, at $\epsilon_{\text{AD}} > 5.3$ kcal/mol, $\langle N_{\text{cluster}} \rangle$ reaches its maximum value of 100, signifying that aggregation of all chains occurs. We explore the hydrogen bond pattern and structure of aggregates formed in the next section to help understand these structures.

C.2 Assembled state structure for targeted substitution of α -1,3-glucan

The hydrogen bond pattern for “silent” DC6 structures formed at $\epsilon_{\text{AD}} = 4.7$ kcal/mol for 20-mer chains is shown in Figure 12. The main change compared to unmodified and randomly substituted α -1,3-glucan is that a single inter-chain hydrogen bond, DC4-AR, is formed to stabilize the structure. Silencing of the DC6 site results in all hydrogen bonds involving DC6 becoming unfavorable, leaving the way open for other donor sites to take the place of DC6. Similar results are seen for other DP values, as seen in Figures S38 through S41.

In a similar fashion, targeted substitution of the DC4 results in the DC6-AR hydrogen bond becoming the main inter-chain bond, as seen in Figures S38b, S39b, S40b, and S41b. Targeted substitution of DC2, shown in Figures S38a, S39a, S40a, and S41a, does not result in one type of hydrogen bond taking over, but results in a similar hydrogen bonding pattern as random substitution at $DS = 1$ (see Figure 9b). The relative insensitivity of the hydrogen bonding pattern to the silencing of DC2 is explained by the fact that DC2 is not involved in the most frequent hydrogen bonds (see Figure 6) and DC4 and DC6 are involved in the key hydrogen bonds that keep the chains in the aggregates together.

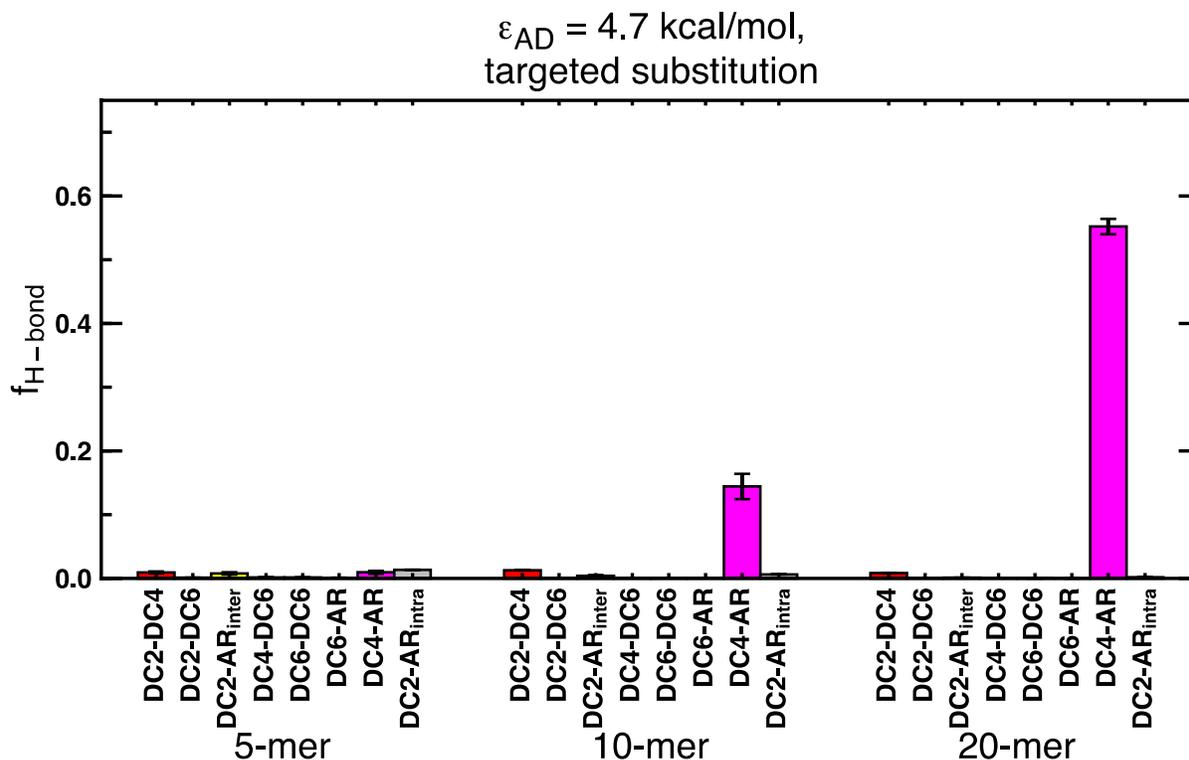


Figure 12. Hydrogen bonding pattern for targeted substitution of the DC6 donor site in α -1,3-glucan at $c_p = 60$ mg/mL at hydrogen bonding strength $\epsilon_{AD} = 7.5$ kcal/mol. Frequency of specific hydrogen bonds being formed for 5-mer, 10-mer, and 20-mer chains. Error bars denote standard deviation over 3 replicas.

Figure 13 provides a closer look at the structure formed at intermediate ϵ_{AD} (the first stage of aggregation) for “silent” DC6 20-mer chains. The main difference between structures formed with “silent” DC6 and unmodified chains (seen in Figure 7) is the formation of cylinders. Supplementary movies S.2 and S.3 show a rotating view of structures for unmodified and “silent” DC6, respectively. A view of a simulation box is shown here in Figure 13a, and a detailed view of one cylinder present in the simulation box is shown in Figure 13b. The hydrogen bonds holding chains together are depicted in Figures 13c and 13d, and correspond to the DC4-AR bond, in accordance to the hydrogen bond pattern shown in Figure 12. Similar to the unmodified polymer, chains are aligned side-by-side, but the side-by-side alignment is different in the two cases: in the unmodified polymer the chains align in both parallel

and anti-parallel fashion as shown in Figure 7d, while in the “silent” DC6 polymer the chains align in primarily parallel fashion, as shown in Figures 13e and S44. We conjecture that the parallel and antiparallel arrangement of the unmodified polymer chains effectively cancels the inherent “helicity” that each chain may have, while the additive “helicity” of each chain in parallel arrangement of the “silent” DC6 polymer leads to the cylinder or nanotube configuration highlighted in Figure 13b.

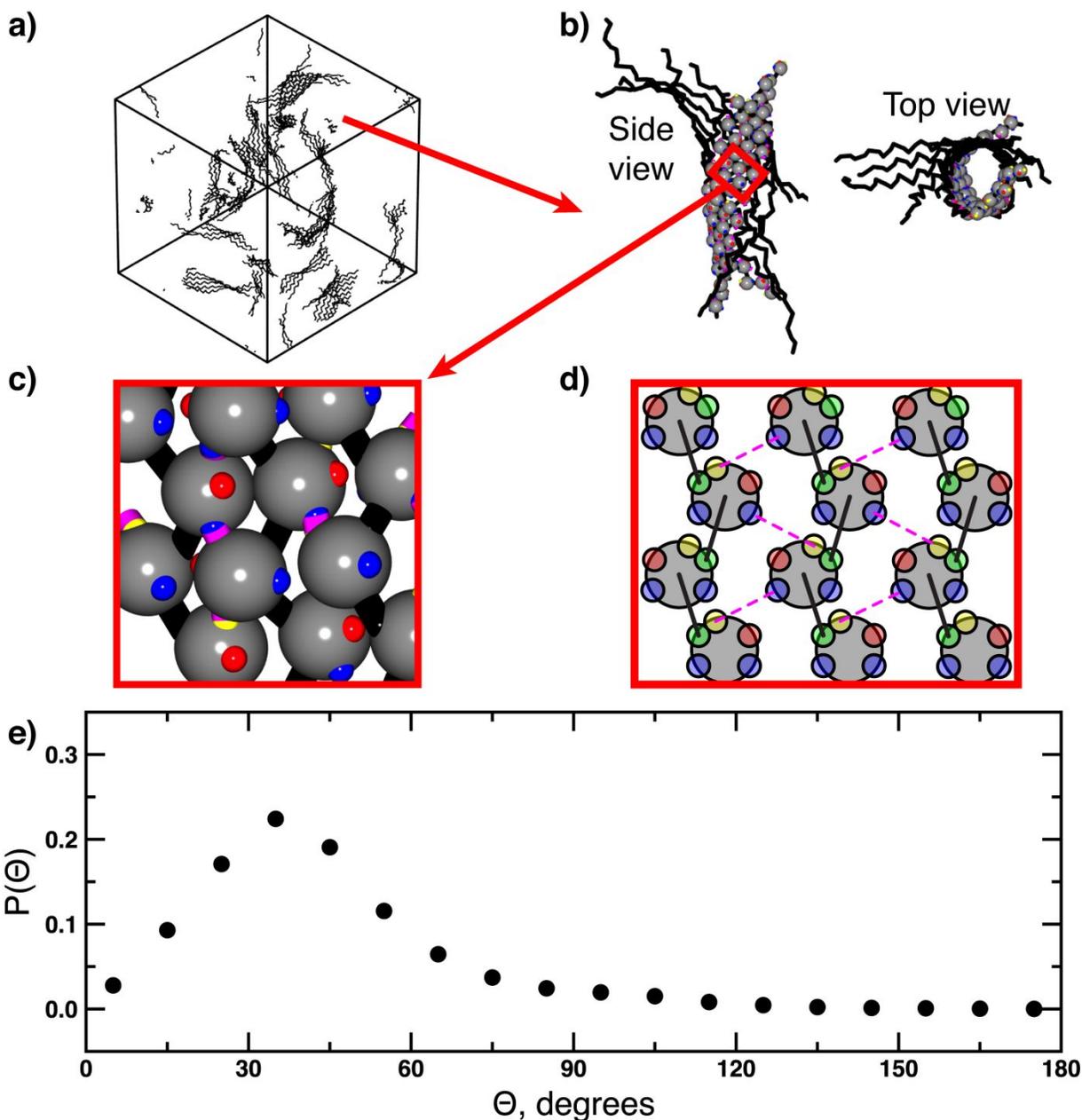


Figure 13. Representative simulation snapshot and structure formed by targeted substitution of the DC6 donor site in α -1,3-glucan 20-mer chains at $\epsilon_{AD} = 4.7$ kcal/mol and $c_P = 60$ mg/mL. a) Simulation snapshot showing the backbones of polymer chains. b) Single cluster side and top views. c) Detail of side-by-side arrangement of chains following a similar rendering scheme as Figure 1c, as well as example hydrogen bonds depicted using the same color scheme as Figure 6a with the “silent” DC6 site shown as red beads. d) Schematic representation of panel b to

show the hydrogen bonds formed with the same coloring scheme as panel c. e) Inter-chain alignment angle, Θ , distribution between near-by monomers.

At higher ϵ_{AD} , shown in Figure S43, the cylinders shown in Figure 13 fuse end-to-end, resulting in overall aggregation consistent with the high values of $\langle N_{\text{cluster}} \rangle$ shown in Figure 11.

IV. Conclusions

In this paper, we presented a coarse-grained (CG) model of α -1,3-glucan based on the hydrogen bond model of Ghobadi and Jayaraman³⁹ that captures the chirality of the α -glucose monomer and directionality of hydrogen bonds formed between monomers in the same chain and between chains. The hydrogen bonding site positions, relative neighboring monomer location and orientation, and effective chain stiffness in the CG model was built from atomistic MD simulation data. The resulting CG model was then used to determine the effect of chain length on the aggregation behavior in solutions of α -1,3-glucan. The model predicted that the hydrogen bond strength required to cause aggregation decreases with increasing chain length. The model also showed that polymer chains remain in a fairly straight configuration, with small changes in conformation upon aggregation. The aggregates that α -1,3-glucan forms are sheet-like structures made of side-by-side stacks of chains (in parallel and antiparallel orientation) held together mainly by inter-chain hydrogen bonds between the donors in carbons C4 and C6 with the acceptor oxygen in the ring.

We tested two kinds of substitution of donor sites: random substitution and targeted substitution. In random substitution we found that the hydrogen bond strength required to cause aggregation increases with increasing degree of substitution. The aggregate shape and the pattern of hydrogen bonds in systems with random substitution are qualitatively similar to those found for the unmodified polymer. For targeted substitution we found that substitution of the donor in carbon C6 results in qualitatively different aggregation behavior: at intermediate hydrogen bond strengths – 4.7 to 5.3 kcal/mol – a population of cylinders formed by folded sheets made of side-by-side stacks of chains in parallel orientation, held together by one main inter-chain hydrogen bond between the donor in carbon C4 and the acceptor in the oxygen in the ring. At the higher hydrogen bond strengths, these cylinders fuse end-to-end.

This paper shows how a coarse-grained model that accounts for directionality in interactions surpasses traditional modeling efforts by (1) requiring less computational power than high-in-detail atomistic simulations, and (2) capturing crucial aspects of aggregated structures not present in traditional CG models where interactions are isotropic. Furthermore, our model allows us to predict aggregation

conditions and aggregate structure to help design novel technologies based on renewable resources, as is the case of our exploration of the effect of random and targeted substitution of hydrogen-bonding sites. Future work in this topic will focus on extending the model to other polysaccharides (e.g. starch, cellulose) and refining the different parameters that make up the model by benchmarking with experimental data.

Conflict of interest

There are no conflicts of interest to declare

Acknowledgements

We acknowledge discussions with scientists in the DuPont Industrial Biosciences, E. I. duPont de Nemours and Co., Experimental Station, Wilmington, Delaware 19803. We acknowledge financial support from U.S. Department of Energy Office of Science grant number DE-SC0017753. The computational work in this paper was supported through the use of information technologies resources at the University of Delaware, specifically the Farber high-performance computing resources. The authors also thank Z. Wu for his feedback and reading of the manuscript.

References

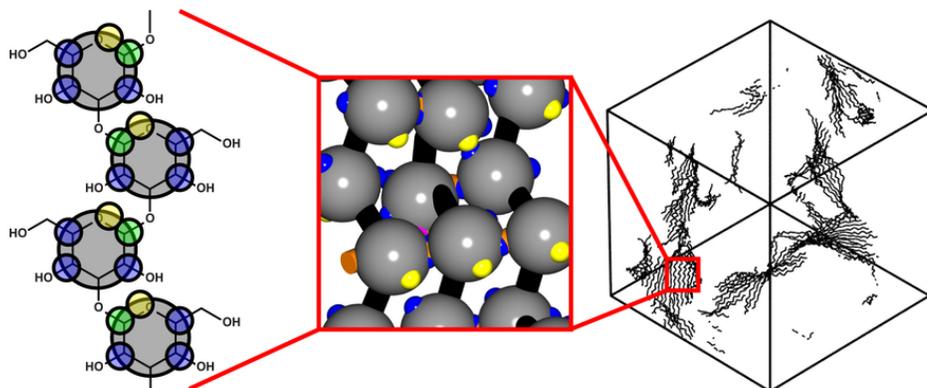
1. T. Iwata, *Angewandte Chemie International Edition*, 2015, **54**, 3210-3215.
2. R. J. Moon, A. Martini, J. Nairn, J. Simonsen and J. Youngblood, *Chemical Society Reviews*, 2011, **40**, 3941-3994.
3. A. Engel, S. Thoms, U. Riebesell, E. Rochelle-Newall and I. Zondervan, *Nature*, 2004, **428**, 929.
4. J. A. Leal, C. Guerrero, B. Gómez-Miranda, A. Prieto and M. Bernabé, *FEMS microbiology letters*, 1992, **90**, 165-168.
5. A. Yoshimi, K. Miyazawa and K. Abe, *Journal of Fungi*, 2017, **3**, 63.
6. A. Choma, A. Wiater, I. Komaniecka, R. Paduch, M. Pleszczyńska and J. Szczodrak, *Carbohydrate polymers*, 2013, **91**, 603-608.
7. K. Ogawa, K. Okamura and A. Sarko, *International Journal of Biological Macromolecules*, 1981, **3**, 31-36.
8. W. J. Loesche, *Microbiological reviews*, 1986, **50**, 353.
9. *US Pat.*, 9,080,195.B2, 2015.
10. *US Pat.*, 7,000,000.A1, 2006.
11. *US Pat.*, 10,000,580.B2, 2018.
12. *US Pat.*, 10,005,850.B2, 2016.
13. S. Neyertz, A. Pizzi, A. Merlin, B. Maigret, D. Brown and X. Deglise, *Journal of applied polymer science*, 2000, **78**, 1939-1946.

14. Y. T. Yao, K. L. Alderson and A. Alderson, *Cellulose*, 2016, **23**, 3429-3448.
15. A. Paajanen, Y. Sonavane, D. Ignasiak, J. A. Ketoja, T. Maloney and S. Paavilainen, *Cellulose*, 2016, **23**, 3449-3462.
16. T. Feng, M. Li, J. Zhou, H. Zhuang, F. Chen, R. Ye, O. Campanella and Z. Fang, *Innovative Food Science & Emerging Technologies*, 2015, **31**, 1-13.
17. T. Feng, X. Zhu and O. Campanella, *Current Opinion in Food Science*, 2016, **9**, 62-69.
18. E. Fadda and R. J. Woods, *Drug discovery today*, 2010, **15**, 596-609.
19. X. Xiong, Z. Chen, B. P. Cossins, Z. Xu, Q. Shao, K. Ding, W. Zhu and J. Shi, *Carbohydrate research*, 2015, **401**, 73-81.
20. J. D. Murillo, M. Moffet, J. J. Biernacki and S. Northrup, *AIChE Journal*, 2015, **61**, 2562-2570.
21. C. S. Pereira, D. Kony, R. Baron, M. Müller, W. F. van Gunsteren and P. H. Hünenberger, *Biophysical journal*, 2006, **90**, 4337-4344.
22. A. B. Poma, M. Chwastyk and M. Cieplak, *Physical Chemistry Chemical Physics*, 2017, **19**, 28195-28206.
23. D. C. Glass, K. Moritsugu, X. Cheng and J. C. Smith, *Biomacromolecules*, 2012, **13**, 2634-2644.
24. G. Srinivas, X. Cheng and J. C. Smith, *Journal of chemical theory and computation*, 2011, **7**, 2539-2548.
25. A.-P. Hynninen, J. F. Matthews, G. T. Beckham, M. F. Crowley and M. R. Nimlos, *Journal of chemical theory and computation*, 2011, **7**, 2137-2150.
26. H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periolo and S. J. Marrink, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2014, **4**, 225-248.
27. J. r. Sauter and A. Grafmüller, *Journal of chemical theory and computation*, 2016, **13**, 223-236.
28. A. B. Poma, M. Chwastyk and M. Cieplak, *The Journal of Physical Chemistry B*, 2015, **119**, 12028-12041.
29. H. M. Cho, A. S. Gross and J.-W. Chu, *Journal of the American Chemical Society*, 2011, **133**, 14033-14041.
30. S. Markutsya, A. Devarajan, J. Y. Baluyut, T. L. Windus, M. S. Gordon and M. H. Lamm, *The Journal of chemical physics*, 2013, **138**, 214108.
31. P. Liu, S. Izvekov and G. A. Voth, *The Journal of Physical Chemistry B*, 2007, **111**, 11566-11575.
32. V. H. Rusu, R. Baron and R. D. Lins, *Journal of chemical theory and computation*, 2014, **10**, 5068-5080.
33. W. Huang, R. Ramesh, P. K. Jha and R. G. Larson, *Macromolecules*, 2016, **49**, 1490-1503.
34. A. B. Poma, M. Chwastyk and M. Cieplak, *Cellulose*, 2016, **23**, 1573-1591.
35. S. Markutsya, Y. A. Kholod, A. Devarajan, T. L. Windus, M. S. Gordon and M. H. Lamm, *Theoretical Chemistry Accounts*, 2012, **131**, 1162.
36. V. Molinero and W. A. Goddard, *The Journal of Physical Chemistry B*, 2004, **108**, 1414-1427.
37. C. A. López, A. J. Rzepiela, A. H. De Vries, L. Dijkhuizen, P. H. Hünenberger and S. J. Marrink, *Journal of Chemical Theory and Computation*, 2009, **5**, 3195-3210.
38. J. Wohllert and L. A. Berglund, *Journal of Chemical Theory and Computation*, 2011, **7**, 753-760.
39. A. F. Ghobadi and A. Jayaraman, *Soft matter*, 2016, **12**, 2276-2287.
40. J. E. Condon and A. Jayaraman, *Soft matter*, 2017, **13**, 6770-6783.
41. A. F. Ghobadi and A. Jayaraman, *The Journal of Physical Chemistry B*, 2016, **120**, 9788-9799.
42. A. Prhashanna and A. Jayaraman, *Journal of Polymer Science Part B: Polymer Physics*, 2019.
43. J. E. Condon and A. Jayaraman, *The Journal of Physical Chemistry B*, 2018, **122**, 1929-1939.
44. E. Lindahl, B. Hess and D. Van Der Spoel, *Molecular modeling annual*, 2001, **7**, 306-317.
45. J. D. Weeks, D. Chandler and H. C. Andersen, *The Journal of chemical physics*, 1971, **54**, 5237-5247.
46. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19-25.
47. J. L. Banks, H. S. Beard, Y. Cao, A. E. Cho, W. Damm, R. Farid, A. K. Felts, T. A. Halgren, D. T. Mainz and J. R. Maple, *Journal of computational chemistry*, 2005, **26**, 1752-1780.

48. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *Journal of computational chemistry*, 2004, **25**, 1157-1174.
49. J. Wang, W. Wang, P. A. Kollman and D. A. Case, *Journal of molecular graphics and modelling*, 2006, **25**, 247-260.
50. A. Jakalian, B. L. Bush, D. B. Jack and C. I. Bayly, *Journal of computational chemistry*, 2000, **21**, 132-146.
51. H. Berendsen, J. Grigera and T. Straatsma, *Journal of Physical Chemistry*, 1987, **91**, 6269-6271.
52. S. Plimpton, *Journal of computational physics*, 1995, **117**, 1-19.
53. G. C. Pimentel and A. McClellan, *Annual Review of Physical Chemistry*, 1971, **22**, 347-385.
54. G. C. Pimentel and A. L. McClellan, *The Hydrogen Bond*, W.H. Freeman, 1960.
55. M. E. Newman and R. M. Ziff, *Physical Review E*, 2001, **64**, 016706.
56. C. S. Brazel and S. L. Rosen, *Fundamental principles of polymeric materials*, John Wiley & Sons, 2012.
57. *US Pat.*, 20180340199.A1, 2018.
58. M. Rubinstein and R. H. Colby, *Polymer physics*, Oxford University Press New York, 2003.
59. J. Jelsma and D. R. Kreger, *Carbohydrate Research*, 1979, **71**, 51-64.
60. K. Ogawa, A. Misaki, S. Oka and K. Okamura, *Carbohydrate Research*, 1979, **75**, C13-C16.
61. K. Ogawa, T. Yui, K. Okamura and A. Misaki, *Bioscience, biotechnology, and biochemistry*, 1994, **58**, 1326-1327.
62. K. Kobayashi, T. Hasegawa, R. Kusumi, S. Kimura, M. Yoshida, J. Sugiyama and M. Wada, *Carbohydrate polymers*, 2017, **177**, 341-346.

TABLE OF CONTENT ENTRY**Journal: Soft Matter****Manuscript ID: SM-ART-03-2019-000580****Title: "Coarse-grained molecular dynamics simulations of α -1,3-glucan"****Author(s): Beltran-Villegas, Daniel J.; Intriago, Daniel; Kim, Kyle H. C.; Behabtu, Natnael; Londono, J. David; Jayaraman, Arthi**

In this work we present a coarse-grained model for α -1,3-glucan that captures hydrogen bonding directionality and polysaccharide monomer structure.



79x39mm (300 x 300 DPI)