

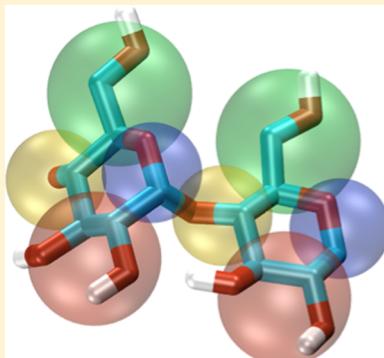
# PITOMBA: Parameter Interface for Oligosaccharide Molecules Based on Atoms

Victor H. Rusu,<sup>†,‡</sup> Riccardo Baron,<sup>†</sup> and Roberto D. Lins\*,<sup>‡</sup>

<sup>†</sup>Department of Medicinal Chemistry, College of Pharmacy, and The Henry Eyring Center for Theoretical Chemistry, The University of Utah, Salt Lake City, Utah 84112-5820, United States

<sup>‡</sup>Departamento de Química Fundamental, Universidade Federal de Pernambuco, Cidade Universitária, Recife, PE 50740-560, Brazil

**ABSTRACT:** A novel four-bead coarse-grained (CG) model for carbohydrates denoted PITOMBA was devised using a bottom-up approach based on the atomistic GROMOS 53A6<sub>GLYC</sub> force field and on experimental thermodynamical data. The model was developed to be used in conjunction with the SPC CG water model (*J. Chem. Phys.* 2011, 134, 084110) and the GROMOS force field functional form. Explicit electrostatic interactions are considered by assigning point charges to each CG bead. Validation of the model is presented to a variety of structural and thermodynamic properties for mono- and oligosaccharides in solution. In addition, the model development philosophy allows for prompt extensions to include hexopyranose chains with diverse glycosidic linkages and branches.



## 1. INTRODUCTION

Carbohydrates are one of the most abundant biomolecules in nature and take part in many biological and chemical processes such as cell signaling and fuel combustion.<sup>1–5</sup> These molecules often comprise of highly flexible large polymeric moieties that can be associated with other types of biomolecules such as lipids and proteins.<sup>6–9</sup> The size and complexity of these macromolecular assemblies may limit practical applications when using an atomistic representation. An alternative to overcome such limitation is to reduce the number of degrees of freedom of the molecular system using a simplified spatiotemporal resolution. This can be achieved by using coarse-grained (CG) models, where groups of atoms are mapped into single beads. Sampling of the configurational phase space using such representation results in a gain of computational efficiency in two ways, (i) the evident reduction of the number of particles, and (ii) smoother potentials that allows longer simulation time steps to be used. However, it is important to keep in mind that such approach represents a compromise between time and scale efficiency and loss of chemical information (e.g., structural details, specific interactions, simplified thermochemical surfaces).

CG models have been developed based on a distinct philosophy used for atomistic force fields,<sup>10–13</sup> including the existing representations for carbohydrates in the past few years.<sup>13–23</sup> Molinero and Goddard III pioneered the development of a CG force field for carbohydrates with the release of the M3B parameter set.<sup>14</sup> The M3B is a three-bead model that maps the position of the CG beads to match C1, C4, and C6 atoms in a  $\alpha$ -D-glucose. The compatible solvent is one-bead CG water that represents a single water molecule. The nonbonded parameters were empirically defined to mimic the behavior of

the atomistic  $\alpha$ -D-glucose as described by the UFF DREIDING<sup>24</sup> force field. The model is based on harmonic bond and harmonic bond-bending angle potentials associated with cosine based torsional equations to express the bonded terms and on the Morse function to compute intermolecular interactions. The parametrization process allowed the M3B to reproduce several properties such as experimental glass transition temperature and density, as well as the atomistic obtained cohesive energy of  $\alpha$ -D-glucose in water. However, the description of the nonbonded interactions could be more efficiently described using other potentials than a Morse function. In fact, this has been subsequently addressed by Liu et al.<sup>15</sup> The authors used the least-squares force matching method<sup>25,26</sup> to derive a tabulated intermolecular interactions from atomistic OPLS/AA<sup>27</sup> simulations using the SPC/E water model.<sup>28</sup> They showed that the model could be used on different temperatures and pressures besides the ones to which it was originally parametrized.

The same level of coarsening was also used by Bu et al. to describe the cellulose I $\beta$  polymorph.<sup>16</sup> However, they used the Boltzmann inversion method to fit the bond, angles and dihedral parameters from atomistic simulations. The nonbonded terms were defined from rescaled Lennard-Jones potentials. The Generalized Born model with switch was used to represent the solvent implicitly.<sup>29</sup> This CG model described reasonably well experimental values of the lattice parameters of the cellulose I $\beta$  polymorph. A more complex CG carbohydrate model was developed by López et al. as an extension to the MARTINI force field. It included mono-, di-, and oligosac-

Received: May 26, 2014

Published: October 9, 2014

charides.<sup>17</sup> This model was later combined with lipids to form the first set of glycolipids parameters for the MARTINI force field family.<sup>30</sup> As the standard MARTINI force field, it is based on the GROMOS force field functional form.<sup>6–9,31,32</sup> The bead types were taken from the list of beads from MARTINI and commonly to all previous models, electrostatic interactions are not explicitly represented. This extension was further appended by including parameters for  $\beta$ -D-glucose to simulate a crystalline cellulose.<sup>10–13,18</sup> Hynninen et al.<sup>13–23</sup> developed another CG force field for  $\beta$ -D-glucose using the same approach as Liu et al. (bonded interactions where mapped using Boltzmann inversion and nonbonded interaction using force matching). The model describes the radial distribution functions and cellulose fibril structures as sampled by the atomistic CHARMM force field.<sup>14,33</sup>

Models using other than three beads to represent a monosaccharide have also been developed. Srinivas et al. introduced a single bead and solvent free model to simulate the crystalline to amorphous transition of cellulose fibril.<sup>19,24</sup> The model based on the CHARMM force field for carbohydrates is<sup>15,33</sup> however, restricted to the applications intended.<sup>19,23,25,26</sup>

Glass et al.<sup>22,27</sup> used the REACH<sup>12,28,34,35</sup> (Realistic Extension Algorithm via Covariance Hessian) method to reproduce the collective motion of the all-atoms simulations using a single bead and solvent free CG model. The REACH method does not use any experimental data, it uses a direct mapping from all-atom MD simulations. The model is applied to describe elastic and mechanical properties of I $\beta$  cellulose. The first five-bead model was presented by Bellesia et al. to study the phase transition of I $\beta$  cellulose to III<sub>I</sub> cellulose.<sup>16,21</sup> The solvent free model was parametrized using experimental crystal structures as reference to set the bonded and the scaled Lennard-Jones parameters. This model is capable of describing structural and thermomechanical properties of cellulose. In a systematic manner, Markutsya et al. applied the force matching approach for different mapping schemes, from one- to four-bead models.<sup>23,29</sup> The authors suggest that one- and two-bead schemes are best suited to model chain-scale properties, such as transitions from crystalline to amorphous structures. The increase in number of beads allowed for a better description of the glycosidic dihedral angles. According to their findings, three- and four-bead models have sufficient degrees of freedom to describe conformational changes of glucose units in polysaccharides. However, their four-bead model was verified to sample unphysical orientations of the glucose ring in cellulose I $\alpha$ . The same research team have also mapped a  $\beta$ -D-glucose molecule into a six-bead CG description.<sup>17,20</sup> This model was parametrized using the force-matching method to a CHARMM-style force field and TIP3P<sup>30,36</sup> water model. The CG model was paired to a solvent described by a single bead representing one water molecule. This model allowed for an even easier dihedral mapping of  $\beta$ -1,4-linked D-glucopyranose  $\phi$  and  $\psi$  dihedral angles, as well as the possibility to compute the ring puckering distribution of the glucose ring. However, the authors report an unexpected puckering transition to boat conformation in both, atomistic (ca. 20%) and CG simulations (>40%).<sup>20</sup> This finding warrants caution in the choice of fine grain models used to parametrize higher order parameters from a bottom-up approach, since NMR data shows that  $\beta$ -D-glucose adopts a  $^4C_1$  conformation in water.<sup>37–39</sup> In fact, among the plethora of carbohydrate atomistic force fields, GROMOS 45A3,<sup>32</sup> GLYCAM06,<sup>40</sup> and MM3<sup>41</sup> have been reported as best choices to describe the experimental conformations of 54

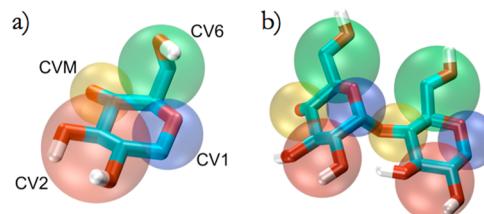
conformers with differing hydroxymethyl, hydroxyl, and glycosidic linkage orientations.<sup>42</sup> It is worth noting that since this report, the GROMOS carbohydrate force field has been further improved to better describe the puckering preferences of saccharides in solution.<sup>43</sup>

Aiming to reproduce properties of carbohydrate chains and moieties in solution, we have devised a novel four-bead CG model using a bottom-up approach. Our parametrization is based on the atomistic GROMOS 53A6<sub>GLYC</sub><sup>43</sup> and on experimental thermodynamical data. Explicit electrostatic interactions are considered by assigning point charges to each CG bead. The model was developed to be used in conjunction with the SPC CG water model<sup>44</sup> and the GROMOS force field functional form. Validation of the model is presented to a variety of structural and thermodynamic properties for carbohydrate systems in solution. In addition, the model development philosophy allows for prompt extensions to include carbohydrate chains with different glycosidic linkages and branches.

## 2. METHODOLOGY

### 2.1. Force Field Parametrization Strategy for $\alpha(1 \rightarrow 4)$

**Glucans.** The PITOMBA force field employs a four-bead model that combines chemical and physical information with the required level of resolution to map the glucose chain torsions; each glucose residue is defined by four charged beads (CV1, CV2, CVM, and CV6) placed at the center of mass of the corresponding group of atoms that constitute the bead, as shown in Figure 1a and Table 1. The di- and oligosaccharides are then built by connecting the CG beads that map the atoms of the glycosidic linkage, as in the case of maltose (Figure 1b).



**Figure 1.** Representation of glucose (a) and maltose (b) residues in the PITOMBA force field and the corresponding atomistic representation in the GROMOS (united atoms) format.

**Table 1.** Mass, in Atomic Units, and the Atomistic Glucose Residue Atoms That Constitute Each Bead in the PITOMBA Model

bead type	mass (au)	constituent atoms
CV1	29	O5–C1–H1
CV2	60	C2–H2–O2–HO2–C3–H3–O3–HO3
CVM	29	C4–H4–O4
CV6	44	C5–H5–C6–H61–H62–O6–HO6

The PITOMBA four-bead model allows the development of a CG model with the following properties: (i) explicit definition of a dihedral angle along the glycosidic bond; (ii) bead charges derived mathematically from the dipole moment and conservation of charge equations; (iii) proper geometrical representation of 1  $\rightarrow$  1, 1  $\rightarrow$  4, and 1  $\rightarrow$  6 axial and equatorial linkages; and (iv) direct mapping between atomistic and CG representations.

The PITOMBA parametrization follows four key steps:

- (1) Using the GROMOS functional form the bond stretching, bond-angle bending, improper dihedral deformation are defined to fit the corresponding average values from a 100 ns atomistic simulation with the GROMOS 53A6<sub>GLYC</sub> force field for the glucose and maltose molecules in SPC water. All beads within the glucose molecule are explicitly connected to avoid the need of bond bending angle force constants within a residue;
- (2) The charge distribution is derived to fit the dipole moment of the glucose residue, in explicit water, for each trajectory frame in the entire 100 ns simulation;
- (3) The van der Waals parameters of each bead is defined in order to reproduce the experimental solvation free energy of their reference molecule (Table 1). The reference molecule is defined as the corresponding hydrogen-capped molecule associated with the CG bead. Solvation free energies for the individual beads were obtained in the CG SPC water model;
- (4) The CG torsional angle for the maltose molecule is fitted to reproduce the corresponding HF/6-31G(d)<sup>45,46</sup> torsional profile, as parametrized in the GROMOS 53A6<sub>GLYC</sub> atomistic force field.

**2.2. Developed Tools.** A set of tools has been devised to aid the PITOMBA force field development. These tools were made generally enough so that they can be used to either further extend the current force field or to assist other CG force fields. The tools are written in C using GROMACS libraries. Installation usage and help follows the GROMACS utility standards. They are publicly available at <http://biomat.dqfnet.ufpe.br>. Among those, we highlight *g\_map* and *g\_cgdipolefitter* that map the center of mass of group of atoms into CG beads and derive bead charges analytically from the dipole moment of atomistic trajectories, respectively. In addition, the Rotational Profiler program (<http://biomat.dqfnet.ufpe.br>) was developed and used to adjust the atomistic CG molecular mechanical profile into the quantum mechanical one, as previously used to derive the atomistic GROMOS 53A6<sub>GLYC</sub> force field.<sup>43</sup>

**2.3. Molecular Dynamics Simulations.** All atomistic GROMOS 53A6<sub>GLYC</sub> D-glucopyranose simulations were performed for 100 ns using the periodic boundary conditions in the *x*, *y*, and *z* directions for the entire N,*p*,T ensemble with the SPC water model.<sup>47</sup> The LINCS<sup>48</sup> method was used to constrain the D-glucopyranose O–H bonds, allowing a 2 fs integration time step. The van der Waals interactions were smoothed by a switch function with 1.2 and 1.4 nm for the inner and outer cutoffs, respectively. The electrostatic interactions were treated using the reaction field<sup>49</sup> method with  $\epsilon = 66.6$  and with cutoff of 1.4 nm. The temperature was kept constant to 298 K by the velocity rescale scheme<sup>50</sup> with relaxation time of 0.2 ps. The isotropic 1 bar pressure was maintained by the Berendsen barostat<sup>51</sup> with a coupled time of 0.5 ps and compressibility ( $\kappa_T$ ) of  $4.51 \times 10^{-5}$  bar<sup>-1</sup>. The simulations were performed using the GROMACS 4.6.1 simulation package.<sup>52</sup>

The PITOMBA simulations were performed for 100 ns in the N,*p*,T ensembles using the GROMOS CG water model. No constrains were applied in the CG simulations. The nonbonded interactions were treated in the same way required by the SPC CG water model, that is, cutoff of 2.0 nm and reaction field with  $\epsilon_{CS} = 2.5$  ( $\epsilon$  within the cutoff sphere) and  $\epsilon = 78.5$  ( $\epsilon$  beyond the cutoff sphere). Except for the  $\kappa_T$  of  $7.51 \times 10^{-4}$

bar<sup>-1</sup>, the temperature and pressure were kept constant to the same values of the atomistic simulation using the same algorithms and conditions. The CG simulations were performed using our modified version of the GROMACS 4.6.1 package, which includes the quartic bond term interaction<sup>44</sup> necessary to describe the SPC CG water. The set of atomistic and CG simulated systems is reported Table 2.

**Table 2. Simulation Time, in Nanoseconds, and Molecular Representation of the Different Simulated Systems**

system	time (ns)	molecular representation
glucose	100	atomistic and coarse-grained
maltose	100	atomistic and coarse-grained
cellobiose	100	atomistic and coarse-grained
SPC CG water	100	coarse-grained
10 and 18 monomers long amylose V	100	atomistic and coarse-grained
18 monomers long amylose A	100	coarse-grained
18 monomers long amylose B	100	coarse-grained
10 monomers long cellulose	100	atomistic and coarse-grained
$\alpha$ -cyclodextrin	100	atomistic and coarse-grained
$\beta$ -cyclodextrin	100	atomistic and coarse-grained
$\gamma$ -cyclodextrin	100	atomistic and coarse-grained

The hydration free energies of the CG beads in GROMOS CG SPC water were estimated using the Bennett Acceptance Ratio method,<sup>53</sup> using a two-step approach in which the electrostatic interactions were vanished first, followed by the van der Waals energy term. Each free energy contribution was calculated using 41 equally spaced  $\lambda$ -values and soft-core potentials.<sup>54</sup> The simulation at each  $\lambda$ -value was 1 ns long with 500 ps of equilibration. The soft-core parameters were  $\lambda^P = 1.0$ ,  $\sigma = 0.3$  nm and  $\alpha = 0.5$  nm. The results were computed with the *g\_bar* tool from GROMACS.

### 3. RESULTS

**3.1. Bonded Interactions.** The GROMOS bond stretching has the functional form

$$V_i = \frac{1}{4} K_i (b_i^2 - b_{0,i}^2)^2 \quad (1)$$

where  $K_i$  is the bond stretching force constant,  $b_i$  and  $b_{0,i}$  are the distance and equilibrium distance of the *i*th CG bond, respectively. The mapped CG bonds, over the 100 ns simulation, behaved as normal distributions, except for CV2–CV6 and CVM–CV6, as expected. The bimodal distribution for these bonds is associated with variations in the hydroxymethyl dihedral angle (O5–C5–C6–O6). It is worth noting that the hydroxymethyl gg conformation is nearly negligible for glucose in water. Therefore, we have opted to take the distance between the mapped beads for the most observed conformation. The hydroxymethyl gg conformation is slightly favored in both GROMOS 45A4, 53A6<sub>GLYC</sub> force fields. The CV2–CV6 and CVM–CV6 bead equilibrium bond distances were taken from the average distances of the mapped atomistic glucose molecule in gg conformation over the current

100 ns run. The CG glycosidic bond during the entire 100 ns atomistic maltose simulation behaves as a normal distribution, and therefore, the equilibrium bond distance was taken as the average of the entire maltose simulation. Force constants were empirically fitted to allow simulations using timesteps up to 7 ps without constraints. Force constants and associated equilibrium bond distances are shown in Table 3. As expected,

**Table 3.** Bond Type, Force Constant, in  $\text{kJ mol}^{-1} \text{nm}^{-4}$ , Equilibrium and Sampled Distance Values, in Nanometers, for All Bonds in the PITOMBA Model and Mapped from the Atomistic GROMOS 53A6<sub>GLYC</sub> Force Field for the Glucose Residue

bond	$K$ ( $\text{kJ mol}^{-1} \text{nm}^{-4}$ )	$b_0$ (nm)	CG distance (nm)	distance 53A6 <sub>GLYC</sub> (nm)
CV1–CV2	$1.110 \times 10^6$	0.280	$0.278 \pm 0.004$	$0.280 \pm 0.005$
CV1–CVM	$5.850 \times 10^5$	0.332	$0.326 \pm 0.003$	$0.332 \pm 0.005$
CV1–CV6	$9.900 \times 10^5$	0.279	$0.266 \pm 0.004$	$0.279 \pm 0.008$
CV2–CVM	$9.800 \times 10^5$	0.298	$0.292 \pm 0.004$	$0.298 \pm 0.006$
CV2–CV6	$5.730 \times 10^5$	0.441	$0.427 \pm 0.003$	$0.441 \pm 0.010$
CVM–CV6	$1.155 \times 10^6$	0.258	$0.249 \pm 0.004$	$0.258 \pm 0.012$
CV1– +CVM <sup>a</sup>	$1.250 \times 10^6$	0.227	$0.227 \pm 0.003$	$0.227 \pm 0.007$

<sup>a</sup>The “+” sign indicates the following residue. Distance measured for the maltose.

the CG bond distributions fall within the average values. However, the CG bonds distributions show smaller deviations compared to the atomistic ones (Table 3). As a consequence the CG glucose ring displays a higher hindrance compared to the atomistic representation by the GROMOS 53A6<sub>GLYC</sub> force field (Table 3).

**Table 4.** Angle Type, Force Constant, in  $\text{kJ mol}^{-1}$ , Equilibrium and Sampled Angle Values, in Degrees, and Relative Positions of All Angles in the PITOMBA Model and Mapped from the Atomistic GROMOS 53A6<sub>GLYC</sub> Force Field

angle	$K$ ( $\text{kJ mol}^{-1}$ )	$\theta_0$ (deg)	position	CG angle (deg)	angle 53A6 <sub>GLYC</sub> (deg)
CV1–CV2–CVM	0	0	ring	$70.0 \pm 0.4$	$69.9 \pm 1$
CV1–CV2–CV6	0	0	ring	$37.8 \pm 0.4$	$37.7 \pm 2$
CV1–CVM–CV2	0	0	ring	$52.4 \pm 0.3$	$52.5 \pm 1$
CV1–CVM–CV6	0	0	ring	$54.7 \pm 0.4$	$54.7 \pm 2$
CV1–CV6–CV2	0	0	ring	$38.0 \pm 0.4$	$38.0 \pm 2$
CV1–CV6–CVM	0	0	ring	$76.3 \pm 0.5$	$76.2 \pm 3$
CV2–CV1–CVM	0	0	ring	$57.5 \pm 0.4$	$57.5 \pm 1$
CV2–CV1–CV6	0	0	ring	$104.2 \pm 0.8$	$104.2 \pm 3$
CV2–CVM–CV6	0	0	ring	$104.8 \pm 0.8$	$104.7 \pm 3$
CV2–CV6–CVM	0	0	ring	$40.8 \pm 0.4$	$40.8 \pm 2$
CVM–CV1–CV6	0	0	ring	$49.0 \pm 0.4$	$49.1 \pm 3$
CVM–CV2–CV6	0	0	ring	$34.4 \pm 0.4$	$34.4 \pm 2$
CV2–CV1–+CVM	850	95.7	$1\rightarrow 4 \alpha$	$96.8 \pm 2$	$95.7 \pm 4$
CV6–CV1–+CVM	850	128.0	$1\rightarrow 4 \alpha$	$131.2 \pm 3$	$128. \pm 10$
CV1–+CVM–+CV1	850	147.4	$1\rightarrow 4 \alpha$	$150.4 \pm 4$	$147.4 \pm 6$
CV1–+CVM–+CV2	850	128.7	$1\rightarrow 4 \alpha$	$127.8 \pm 3$	$128.7 \pm 10$
CV1–+CVM–+CV6	850	106.7	$1\rightarrow 4 \alpha$	$107.1 \pm 2$	$106.7 \pm 10$
CVM–CV1–+CVM	850	121.4	$1\rightarrow 4 \alpha$	$120.3 \pm 2$	$121.4 \pm 5$
CV2–CV1–+CVM	850	108.9	$1\rightarrow 4 \beta$	$110.6 \pm 2$	$108.9 \pm 4$
CV6–CV1–+CVM	850	142.1	$1\rightarrow 4 \beta$	$144.5 \pm 3$	$142.1. \pm 6$
CV1–+CVM–+CV1	850	150.4	$1\rightarrow 4 \beta$	$152.0 \pm 5$	$150.4 \pm 7$
CV1–+CVM–+CV2	850	107.7	$1\rightarrow 4 \beta$	$108.0 \pm 3$	$107.7 \pm 6$
CV1–+CVM–+CV6	850	130.8	$1\rightarrow 4 \beta$	$132.3 \pm 4$	$130.8 \pm 10$
CVM–CV1–+CVM	850	165.5	$1\rightarrow 4 \beta$	$165.2 \pm 3$	$164.5 \pm 4$

The GROMOS functional form of the bond angle bending for the atoms  $i$ ,  $j$ , and  $k$  is given by

$$V_i = \frac{1}{2}K_i(\cos\theta_i - \cos\theta_{0,i})^2 \quad (2)$$

where  $K_i$  is the bond angle bending force constant,  $\theta_i$  and  $\theta_{0,i}$  are the angle and equilibrium angle of the  $i$ th CG angle, respectively. As mentioned above, all CG beads are linked by a bonded term, making it unnecessary to define bending parameters within a CG residue. Therefore, all the force constants of the inner ring angles were set to zero and are only listed in the parameter file to work with the pdb2gmx tool in GROMACS, which would complain about the absence of these angles within the beads. The inter-residue equilibrium values were taken as the average value of the mapped beads into the 100 ns atomistic simulation of a maltose molecule described by the 53A6<sub>GLYC</sub> force field. Force constants were empirically defined to best reproduce the measured average values. CG angles, associated force constant, and their variation are listed in Table 4.

The improper dihedral angle in the GROMOS force field is given by

$$V_i = \frac{1}{2}K_i(\xi_i - \xi_{0,i})^2 \quad (3)$$

where is the improper dihedral force constant,  $\xi_i$  and  $\xi_{0,i}$  are the improper dihedral angle and equilibrium improper dihedral angle of the  $i$ th CG improper angle, respectively. There are two types of improper dihedrals in the proposed force field. The first describes the glucose ring (puckering and hydroxymethyl conformation), while the second is used to describe the anomer (see Table 5). A dependence of the improper dihedral angle with the bead distances results from the adopted four-bead

**Table 5.** Improper Dihedral Type, Force Constant, in  $\text{kJ mol}^{-1} \text{deg}^{-2}$ , Equilibrium Angle, in Degrees, Relative Position in the PITOMBA Model and Sampled Distance Values for All Improper Angles in the PITOMBA Model and Mapped from the Atomistic GROMOS 53A6<sub>GLYC</sub> Force Field

improper angle	$K$ ( $\text{kJ mol}^{-1} \text{deg}^{-2}$ )	$\theta_0$ (deg)	position	CG improper (deg)	53A6 <sub>GLYC</sub> improper (deg)
CV1–CVM–CV2–CV6	600	16.4	ring	$16.4 \pm 3$	$16.4 \pm 5$
CV2–CV6–+CVM–CV1	600	-47.2	$\alpha$ -anomer	$-47.9 \pm 2$	$-47.2 \pm 5$
CV2–CV6–+CVM–CV1	600	-7.0	$\beta$ -anomer	$7.2 \pm 2$	$7.0 \pm 12$

mapping scheme. Therefore, improper equilibrium angles were taken from the average values for the hydroxymethyl gg ensemble. The CG improper force constants were empirically set to have a similar fluctuation of the atomistic improper angles (Table 5). It is worth noting that the GROMOS 53A6<sub>GLYC</sub> force field is a reparameterization that properly describes the distribution (100%) of the  $^4\text{C}_1$  chair conformation of glucose in water at 298 K. Therefore, the actual CG improper angle values shown in Table 5 are associated with the  $^4\text{C}_1$  ring puckering in the CG representation.

The dihedral term for the GROMOS force field is given by

$$V_i = K_i(1 + \cos \delta_i \cos n_i \phi_i) \quad (4)$$

where  $K$  is the dihedral force constant,  $\delta_i$  is the phase shift,  $n_i$  is the multiplicity and  $\phi_i$  is the dihedral angle of the  $i$ th CG dihedral angle, respectively. There are two torsional angles  $\phi$  and  $\psi$  along the glycosidic bond for  $\alpha$ -1,4-linked D-glucopyranoses, defined in the GROMOS atomistic force field as O5–C1–O4–C4 and C1–O4–C4–C3, respectively. For the CG representation a single dihedral angle ( $\chi$ ) has been defined to map the linkage between two adjacent monosaccharides.  $\chi$  is defined along the CV2–CV1–+CVM–+CV2 beads, where the “+” sign indicates the following residue. It is worth noting that  $\chi$  contains geometric information on both  $\phi$  and  $\psi$  dihedral angles as well as all bending angles along the glycosidic linkage. Thus,  $\chi$  was fitted by mapping the CG beads to the corresponding points from a QM-generated rotational profile around  $\phi$  and  $\psi$  dihedral angles. Torsional profiles of maltose were calculated at the HF/6-31G(d) level with the GAUSSIAN09 program.<sup>55</sup> We found that nearly all-lowest energy points of the rotational profile were associated with the value of ca. 96° (maltose) and ca. 94° (cellobiose) for the  $\psi$  dihedral angle. Therefore, fitting of the  $\chi$  dihedral angle to the atomistic representation was performed for these fixed values of  $\psi$ . The resulting dihedral parameters are listed in Table 6. Figure 2 shows the comparison between the CG-mapped and the atomistic profiles ( $x$ -axis corresponds to the angular values of the CG representation).

**Table 6.** Proper Dihedral Type, Force Constant, in  $\text{kJ mol}^{-1}$ , Phase Shift, and Multiplicity of the Glycosidic Dihedral Angle in the PITOMBA Model

dihedral angle <sup>a</sup>	$K$ ( $\text{kJ mol}^{-1}$ )	phase shift (deg)	multiplicity	glycosidic linkage
CV2–CV1–+CVM–+CV2	32.4	80	1	1→4 $\alpha$
CV2–CV1–+CVM–+CV2	6.8	0	0	1→4 $\beta$
CV2–CV1–+CVM–+CV2	-9.8	0	2	1→4 $\beta$
CV2–CV1–+CVM–+CV2	12.9	120	2	1→4 $\beta$

<sup>a</sup>The “+” sign indicates the following residue.

The fitted CG rotational profile of maltose is not able to reproduce all the energy points of the QM-generated curve. However, the curve minimum points have energy difference below 5  $\text{kJ mol}^{-1}$  and differences above 10  $\text{kJ mol}^{-1}$  are only located on the energy barrier, which is above 60  $\text{kJ mol}^{-1}$ .

Despite of the reduced spatial representation of the CG model, the single cosine torsional term allows for similar inter-residue dihedral fluctuations in the CG maltose molecule as in the atomistic counterpart, during 100 ns simulation (Figure 3).

The general trends of the HF/6-31G(d) profile for cellobiose are reproduced in the fitted PITOMBA CG profile. The main differences are on the values of the energy barriers and on the position of central local minimum. The QM-generated barriers are 51 and 32  $\text{kJ mol}^{-1}$  and the PITOMBA ones are 38 and 37  $\text{kJ mol}^{-1}$ , respectively. The second minimum energy point (-30°) is shifted by +60° in the PITOMBA model (+30°) in vacuum. However, we note that such difference is expected to dampen in solvated systems.

The simulated dihedral angle for cellobiose in the atomistic and PITOMBA models show that the 3 cosine based torsional terms of PITOMBA samples an average equilibrium angle of 175°, which is shifted by 10° compared to the atomistic GROMOS 53A6<sub>GLYC</sub> (-175°) (Figure 4).

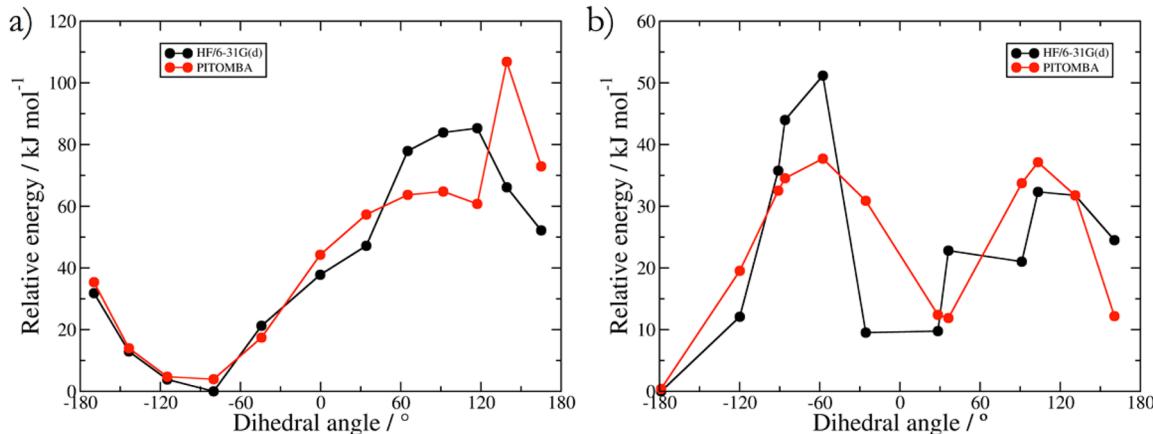
**3.2. Nonbonded Interactions.** Electrostatic interactions in the CG model are included by assigning point charges to the beads in order to reproduce the average GROMOS 53A6<sub>GLYC</sub> dipole moment of the D-glucose residue (Table 7). The bead charges were derived mathematically from the dipole moment and conservation of charge equations. Figure 4 shows GROMOS 53A6<sub>GLYC</sub> and the CG D-glucose dipole moment along 100 ns trajectories.

Even though the charges were defined from a sample of the most observed conformation, the CG dipole moment intensity matches the average value of the dipole moment intensity of the atomistic glucose for the entire 100 ns simulation (Figure 5a). The large fluctuation of the dipole moment in the atomistic model arises from rotation of the multiple glucose hydroxyl groups. The reduced granularity of the CG model, therefore, results in a smaller intensity and direction variation of this quantity (Figure 5b). The average difference between the dipole moment direction in the CG and atomistic model is ca. 21°.

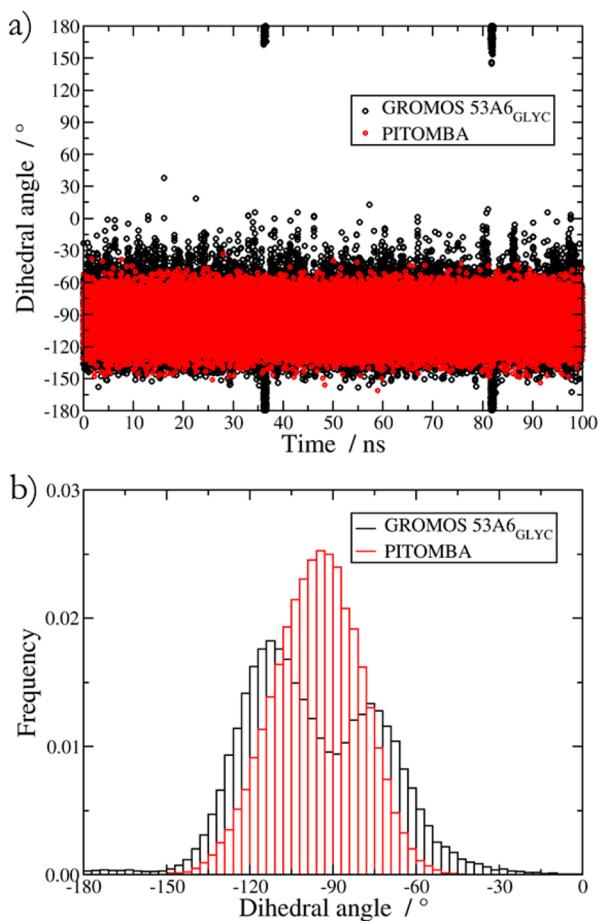
The van der Waals nonbonded interactions for the GROMOS force field uses the Lennard-Jones function in the form

$$V_i = \frac{C_{i,12}}{r_i^{12}} - \frac{C_{i,6}}{r_i^6} \quad (5)$$

where  $r_i$  is the distance in the  $i$ th interaction. Following the GROMOS philosophy for parameter development, van der Waals parameters were derived based on hydration free energies of the individual beads in CG SPC water model. Because one CG SPC bead represents five water molecules, we have assumed a correlation between the van der Waals terms in

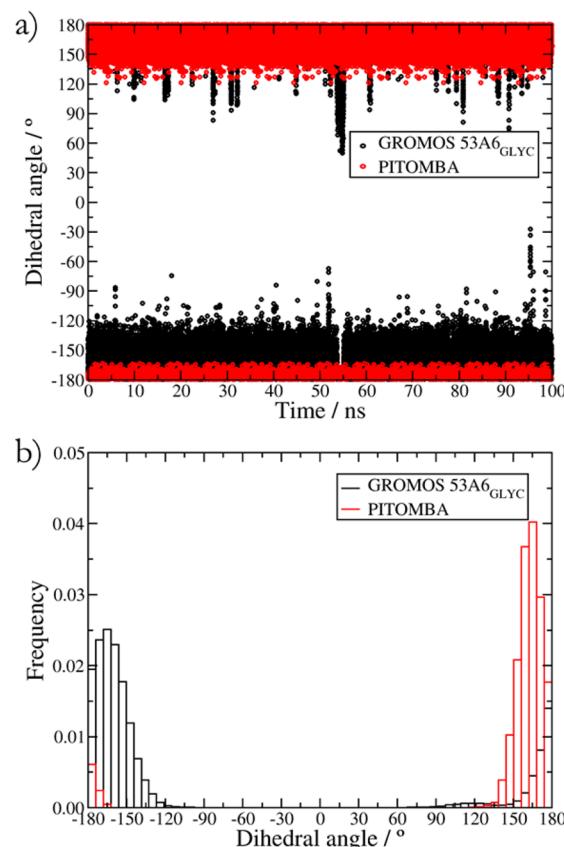


**Figure 2.** Rotational profiles for  $\phi$  (atomistic) and  $\chi$  (CG) dihedral angles in (a) maltose and (b) cellobiose. Atomistic profile is shown as calculated by HF/6-31G(d).  $x$ -axis is shown in CG angular values.



**Figure 3.** Distribution of  $\chi$  dihedral angle for maltose in CG and atomistic simulations (mapped onto the CG beads) over 100 ns simulation period for maltose.

both representations. We have chosen to express this correlation by defining the C6 term and fitting C12. Our choice is based on the fact that dispersion is mainly additive. We have noted that the sum of the C6 of five water molecules is 6.72 times smaller than this term in one CG water bead. Therefore, our parametrization scheme comprised of the sum of the atomistic C6 terms times 6.72, as a constant. Model molecules were chosen to represent the individual beads based on the chemical composition the latter encompasses. The

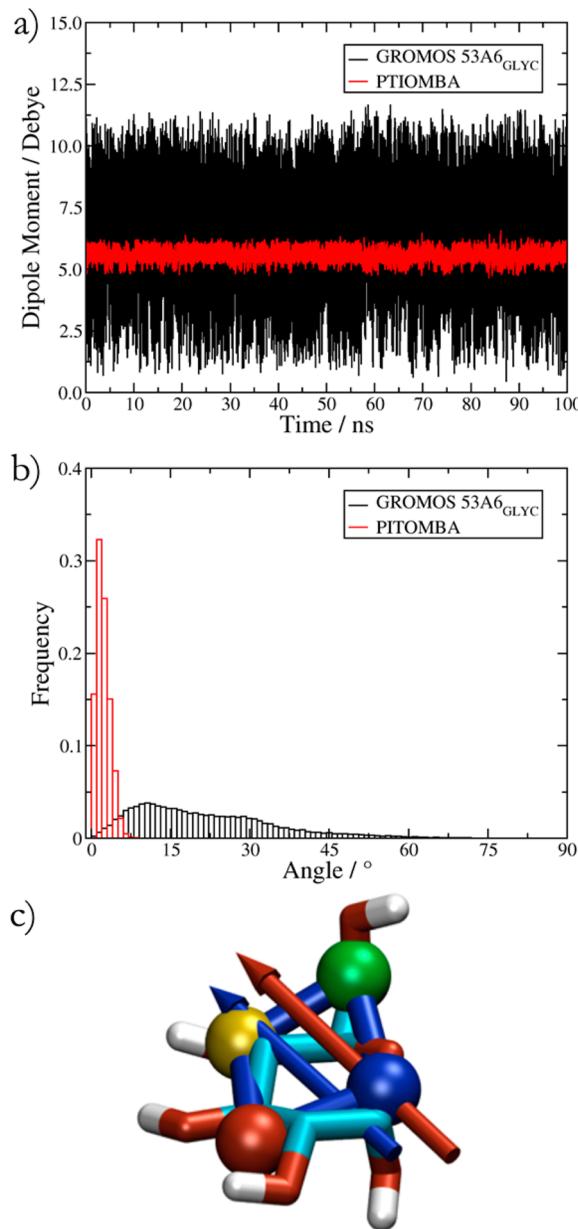


**Figure 4.** Distribution of  $\chi$  dihedral angle for cellobiose in CG and atomistic simulations (mapped onto the CG beads) over 100 ns simulation period for cellobiose.

**Table 7. Charge Values, in Electrons, for All Bead Types in the PITOMBA Model**

bead type	charge (e)
CV1	-0.531
CV2	0.109
CVM	0.106
CV6	0.316

experimental and the fitted standard hydration free energy of each associated molecule are shown in Table 8. The



**Figure 5.** (a) Dipole moment of the glucose in PITOMBA and GROMOS 53A6<sub>GLYC</sub> representations as a function of time over 100 ns simulations. (b) Histogram of the dipole moment angle of glucose in PITOMBA and GROMOS 53A6<sub>GLYC</sub> over 100 ns simulations. (c) Snapshot of one simulation frame showing the CG dipole moment (blue) and the atomistic dipole moment (red). The atomistic dipole moment was shifted in order to better visualize both dipole representations since they overlay on top of each other.

corresponding CG Lennard-Jones parameters are shown in Table 9.

**Table 9. Lennard-Jones C<sub>6</sub> and C<sub>12</sub> Parameters for the Different Beads in the PITOMBA Model**

bead type	C <sub>6</sub> (kJ mol <sup>-1</sup> nm <sup>6</sup> )	C <sub>12</sub> (kJ mol <sup>-1</sup> nm <sup>12</sup> )
CV1	5.59800434 × 10 <sup>-02</sup>	0.98528029 × 10 <sup>-03</sup>
CV2	1.11960087 × 10 <sup>-01</sup>	0.19705606 × 10 <sup>-02</sup>
CVM	5.59800434 × 10 <sup>-02</sup>	0.98528029 × 10 <sup>-03</sup>
CV6	1.06167802 × 10 <sup>-01</sup>	0.13249361 × 10 <sup>-02</sup>

It is worth noting that the hydration free energy difference between a CG water molecule and five atomistic ones is 68.9 kJ mol<sup>-1</sup>,<sup>44</sup> therefore, an equivalence has to be made. We assumed that this difference is associated with the remaining 12 nonconnected hydrogen bonds of the five atomistic molecules that comprised the CG water bead. Therefore, the difference in the hydration free energy of each CG bead compared to the experimental one was rescaled by the number of missing hydrogen bonds between the CG beads and water. This contribution was taken to be 68.9/12 = 5.74 kJ mol<sup>-1</sup> per missing hydrogen bond. The number of missing hydrogen bonds between each sugar CG bead and water was calculated by the radial distribution function on an atomistic simulation of each model compound and rounded to the closest integer (Table 8). The CG force field has a higher average hydration free energy error compared to the atomistic GROMOS 53A6<sup>56</sup> (ca. 6% vs 0.8%, respectively). However, we note that such loss in accuracy can likely be afforded by the smoother CG potential surface.

#### 4. FORCE FIELD VALIDATION

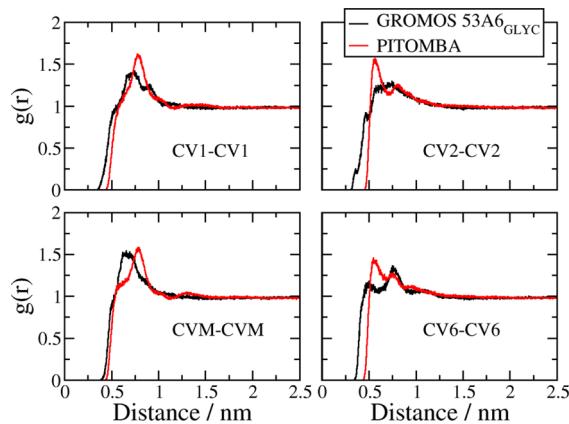
Structural and solvation properties were assessed for selected  $\alpha$  and  $\beta$  ( $1 \rightarrow 4$ ) glucose oligosaccharide-based systems. It should be noted that limited availability of experimental data for carbohydrates in solution often precludes us to obtain a direct comparison between experimental and simulated data.

**4.1. Glucose Molecule.** To access the solubility of glucose in water, the free energy of solvation was calculated using the same protocol for each individual CG bead. We found the value of  $-28.8$  kJ mol<sup>-1</sup>, which rescaled represents  $-109.2$  kJ mol<sup>-1</sup> for the PITOMBA glucose residue. This value is in good agreement with the experimental value of  $-106.5$  kJ mol<sup>-1</sup> for the glucose molecule.<sup>59</sup> To further extend the understanding of the solubility of glucose in water a simulation box containing 64 glucose residues and 2320 CG waters using the PITOMBA force field was compared to 64 glucose molecules and 10820 SPC water molecules using the GROMOS 53A6<sub>GLYC</sub> force field. The radial distribution functions (RDF) between the CG beads in both models (atomistic representation mapped into CG) are shown in Figure 6. Due to a larger volume/spherical

**Table 8. Calculated and Rescaled Hydration Free Energies of the CG Beads in the PITOMBA Model<sup>a</sup>**

bead type	associated molecule	coordination no. (water)	$\Delta G^0_{\text{Hydration}}$ in the CG model (kJ mol <sup>-1</sup> )	rescaled $\Delta G^0_{\text{Hydration}}$ (kJ mol <sup>-1</sup> )	expt. <sup>57,58</sup> value (kJ mol <sup>-1</sup> )
CV1	methanol	2	-32.1	-20.6	-21.3
CV2	1,2-ethanediol	6	-2.4	-36.4	-32.2
CVM	methanol	4	1.9	-21.1	-21.3
CV6	ethanol	2	-10.0	-22.4	-20.9

<sup>a</sup>Experimental energies are associated to the hydrogen-capped molecules comprising each bead.

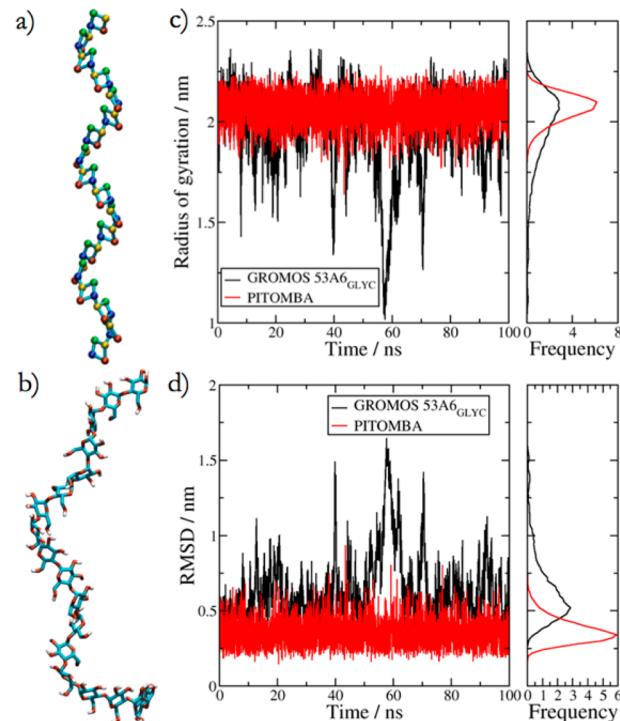


**Figure 6.** Radial distribution functions between identical CG beads in PITOMBA and GROMOS 53A6<sub>GLYC</sub> force fields. The atomistic curves were obtained by mapping the atomistic trajectory into CG representation.

shape of the CG beads, compared to the atomistic molecular groups that comprise them, the atomistic molecules can come close to each other, as it can be seen on the RDF curves of Figure 6. Despite the absence of hydrogen bonding, larger bead volume and associated spherical shape of the beads, the PITOMBA model is capable to describe satisfactory bead–bead interactions, compared to the GROMOS 53A6<sub>GLYC</sub> force field.

**4.1. Single and Double  $\alpha(1 \rightarrow 4)$  Glucose Helices.** The single  $\alpha(1 \rightarrow 4)$  glucose strand polymer is called amylose V, which is highly flexible in water.<sup>60</sup> It cocrystallized with different compounds such as alcohols, DMSO and iodine reveals to be a left-handed helix with six to eight residues per turn.<sup>61</sup> The same number of residues per turn was observed for three different CG amylose V systems (10, 18, and 36 monomers long) in water. The pitch in our simulation in solution is around 20 Å, which is not directly comparable to the available experimental pitch of ca. 8 Å in antiparallel pairs of parallel 6<sub>S</sub>-symmetry amylose helices crystal in nonpolar solvents. Single chains in solution and free of contacts with other polysaccharide chains are expected to exhibit higher flexibility and therefore to present themselves in an *unfolded* state. As a support to this hypothesis, the experimental radius of gyration in the gel phase for a 10-mer amylose V, as measured by SAXS, has been reported as 10.1 Å<sup>62,63</sup> and 10.5 Å.<sup>64</sup> The calculated average radius of gyration for a 10-mer CG amylose V for the PITOMBA model reveals a value of 11.9 Å, whereas the atomistic GROMOS 53A6<sub>GLYC</sub> amylose V 10-mer a value of 9.8 Å. For a 18-mer long amylose V strand the average gyration radii are 19.3 and 20.6 Å for the atomistic and CG models, respectively. However, the atomistic system shows a higher fluctuation compared to the atomistic one (Figure 7). The root mean square deviation (RMSD) of both simulations shows a larger fluctuation of the atomistic model compared to CG one. This is due to the higher flexibility of the atomistic model compared to the PITOMBA model.

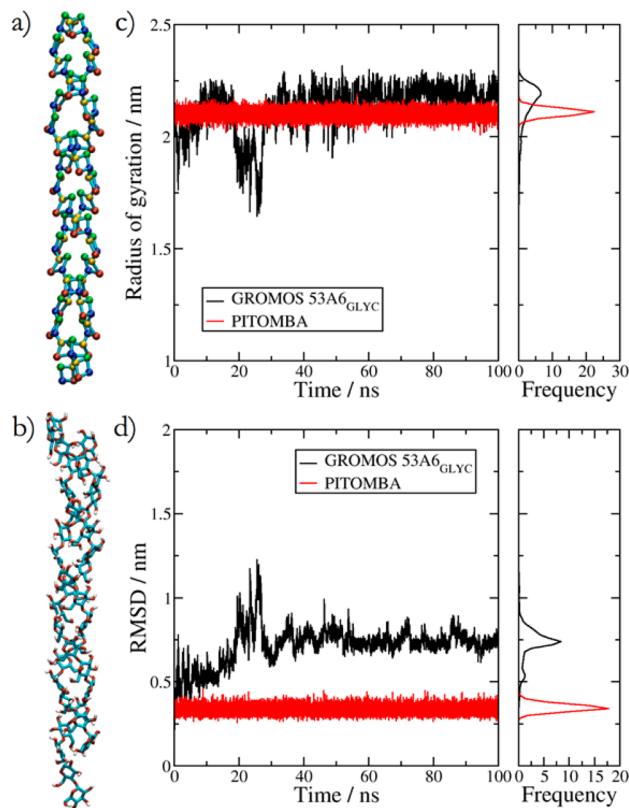
The parallel double helix of  $\alpha$ -1,4-linked D-glucose polymer is called amylose A, and its antiparallel counterpart is called amylose B.<sup>65–67</sup> While both have six residues per turn in their pitch, there seems to be no consensus in the literature about their helical chirality (left or right-handed; e.g., amylose A and B structures in ref 65–69). For validation purposes, we simulated a claimed left-handed amylose A in which the crystallographic X-ray data show a pitch of 21.4 Å.<sup>68,69</sup> For this



**Figure 7.** 18-mer long amylose V in the (a) PITOMBA and (b) atomistic GROMOS 53A6<sub>GLYC</sub> representation and their corresponding (c) radius of gyration and (d) root mean square deviation as a function of time over 100 ns simulations in solution. The histogram of each graph is represented on the right of each graph.

system, our simulations in water predicted a pitch of 24.1 Å. Compared to amylose V (single strand), a much closer agreement between the experimentally estimated and calculated pitch value is observed for the double helix. This is attributed to the fact that in amyloses A the carbohydrate-carbohydrate contacts maintain the oligosaccharide folded state, that is, its helical structure. The PITOMBA model, predicts a single double helical structure as proposed by the experimental crystal data (Figure 8a) and by the corresponding GROMOS 53A6<sub>GLYC</sub> atomistic simulation (Figure 8b). The average radius of gyration of the CG model is 21.1 Å and the last 50 ns of atomistic simulation provides a average of 22.1 Å (Figure 8c). The atomistic model, however, shows a greater fluctuation compared to the CG one (Figure 8d). This difference in RMSD is due to the slide of one chain relative to the other in the atomistic simulation of amylose A (as it can be seen in Figure 8b). This phenomenon, where hydrogen bonds are broken and reformed, is not observed in the PITOMBA model. However, the same structural helical pattern is kept in both cases.

**4.2. Single  $\beta(1 \rightarrow 4)$  Glucose Chain.** Radius of gyration of  $\beta(1 \rightarrow 4)$  oligosaccharide chains in water are not presented since experimental data is restricted to crystalline phases, therefore the PITOMBA validation was performed comparing a 10-mer cellulose strand in water to the atomistic GROMOS 53A6<sub>GLYC</sub> model. The gyration radii in both representations are similar, with the PITOMBA being  $1.49 \pm 0.06$  nm and the atomistic counterpart GROMOS 53A6<sub>GLYC</sub>  $1.51 \pm 0.02$  nm (Figure 9). The RMSD of cellulose in both representations is comparable due to the fact that the simulated cellulose is mainly gg and therefore has the same center of mass distances in both models.

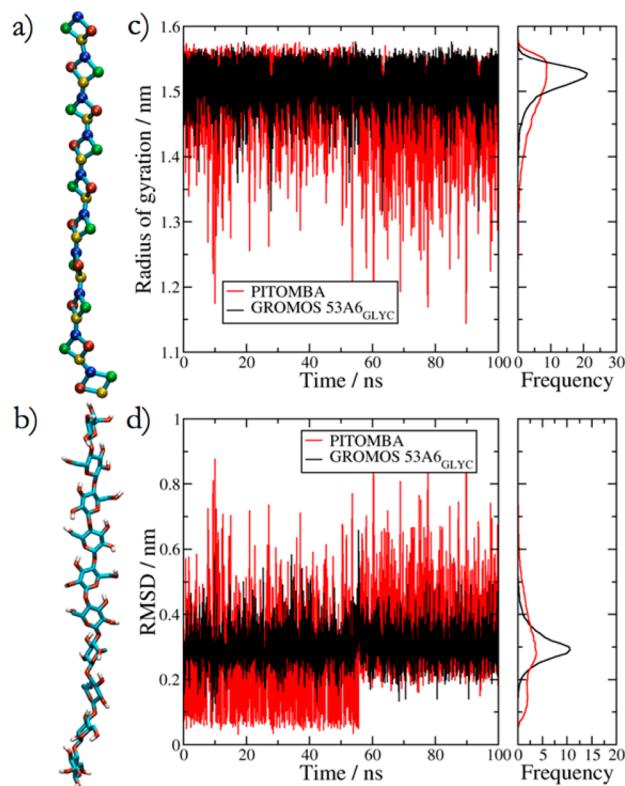


**Figure 8.** 18-mer amylose A representation in the (a) PITOMBA and (b) GROMOS 53A6<sub>GLYC</sub> models and their corresponding (c) radius of gyration and (d) RMSD as a function of time over 100 ns simulations in solution. The histogram of each graph is represented on the right of each graph.

**4.3.  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Cyclodextrins.** Cyclodextrins (CD) are homochiral oligosaccharides composed of 6 to 13  $\alpha$ -1,4-linked D-glucopyranose monomers.<sup>70</sup> Crystallographic X-ray studies show that each glucose monomer possesses a clear  $^4\text{C}_1$  chair ring puckering. Those CDs composed of 6, 7, and 8 monomers are referred as  $\alpha$ -,  $\beta$ -, and  $\gamma$ -cyclodextrins, respectively.<sup>71</sup> A wide range of chemical species can be partially or completely included in their cavities. Generally, the size variation of CDs provide discrimination about the inclusion complexes, with hydrophobic guests being better binders than hydrophilic ones.<sup>72–74</sup> In the absence of guests, CDs are found hydrated in crystal structures. The water content in a CD cavity is highly variable and dynamical; for example, the average number of water molecules in  $\beta$ -cyclodextrin is 6.5; however, up to 12 water molecules can be found in its cavity.<sup>75–77</sup> It is believed that desolvation of both host and guest and conformational changes in the guest are the dominant driving effects of binding for the  $\beta$ - and  $\gamma$ -cyclodextrins complexes.<sup>78–81</sup> In contrast, in  $\alpha$ -cyclodextrins, it has been proposed that polar interactions between the host and guest are the main driving binding forces since they are weakly solvated.<sup>76</sup>

To assess the water–CD interaction balance, the number of water molecules inside the  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs cavities were monitored over 100 ns simulations of both CG and atomistic GROMOS 53A6<sub>GLYC</sub> simulations. Table 10 and Figure 10 show the average number of SPC water molecules inside the CDs cavities.

The number of water molecules inside each CD cavity in both representations was measured in the same way. We



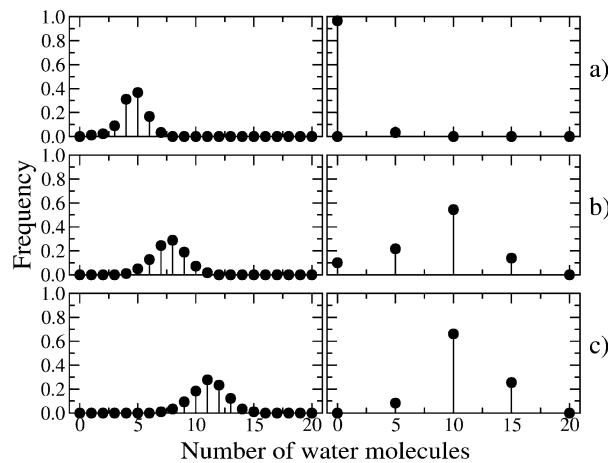
**Figure 9.** Representation of a 10-mer cellulose chain in the (a) PITOMBA and (b) GROMOS 53A6<sub>GLYC</sub> representations and their corresponding (c) radius of gyration and d RMSD as a function of time over the 100 ns simulation in water. The histograms of each graph are represented on the right of each graph.

**Table 10. Average Number of Water Molecules Inside the Cavities of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Cyclodextrins Simulated with the Atomistic GROMOS 53A6<sub>GLYC</sub> and the CG PITOMBA Model**

system	no. of water molecules	
	GROMOS 53A6 <sub>GLYC</sub> <sup>a</sup>	PITOMBA model <sup>b</sup>
$\alpha$ -cyclodextrins	5	0
$\beta$ -cyclodextrins	8	1.8 (9)
$\gamma$ -cyclodextrins	11	2.2 (11)

<sup>a</sup>SPC water. <sup>b</sup>GROMOS CG water. In parentheses the corresponding number of SPC water molecules.

counted the number of water molecules inside of a cylinder comprised of a 4-Å height and with a fixed radius depending on the CD. The cylinder radii for  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs were defined based on the cavity volume of the atomistic representation of these molecules as 4.47, 5.00, and 5.47 Å, respectively. The absence of CG waters inside the CG  $\alpha$ -CD can be explained by the fact that the volume associated with five water molecules in the CG representation (sphere modeled) is larger than the cylindrical-like shape adopted by the five waters inside the atomistic  $\alpha$ -CD. However, the average number of waters inside the  $\beta$ - and  $\gamma$ -CDs is in accordance in both representations, atomistic and CG (Table 10 and Figure 10), even though the observed shift in the distribution of the atomistic water molecules inside of the CD cavities are not reproduced by the CG force field. The number of SPC water molecules inside the atomistic  $\beta$ - and  $\gamma$ -CDs cavities are 8 and 11, respectively. In the CG representation the average number of CG water



**Figure 10.** Number of water molecules inside of the cavities of  $\alpha$ - (a),  $\beta$ - (b), and  $\gamma$ -cyclodextrins (c) in GROMOS 53A6<sub>GLYC</sub> atomistic (left) and PITOMBA CG (right) representations.

molecules inside the  $\beta$ - and  $\gamma$ -CDs cavities equals to 1.8 and 2.2, which correspond to 9 and 11 atomistic water molecules, respectively.

The geometrical differences between the atomistic GROMOS 53A6<sub>GLYC</sub> and CG PITOMBA models for  $\gamma$ -CD were assessed by their RMSD over 100 ns simulations. The atomistic model has a smaller RMSD fluctuation, but the difference between the average RMSD fluctuation of both models is 0.06 nm (Figure 11a). The initial structure for the  $\gamma$ -CD in each run was the same (Figure 11b) but the final structure (Figure 11c) for the PITOMBA model is slightly more compact, which explains the difference in the RMSD. Similar fluctuations are observed for the  $\alpha$ - and  $\beta$ -CDs (data not shown for conciseness).

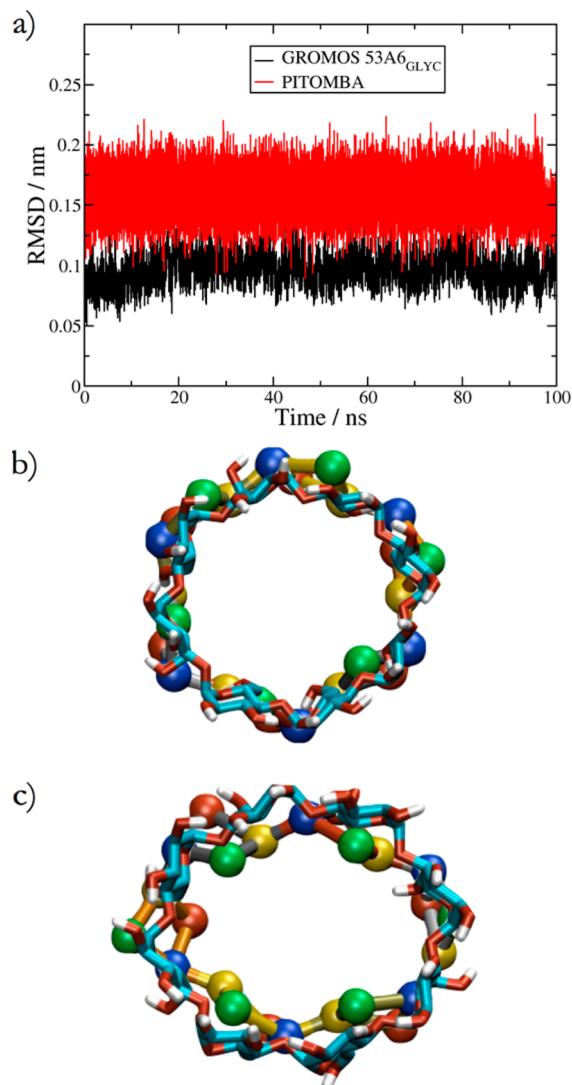
To further access the geometrical properties of  $\gamma$ -CDs in the PITOMBA representation, the cone-like shape of both atomistic and CG models were measured through the ratio between the radius of gyration of the CG beads that compose each face of  $\gamma$ -CD in the PITOMBA representation (Figure 12). Due to the larger flexibility of the hydroxymethyl group in the atomistic representation, the distribution of the GROMOS 53A6<sub>GLYC</sub> model is slightly larger than the CG counterpart. However, the difference between the average values of both representations is 0.01. Similar fluctuations are observed for the  $\alpha$ - and  $\beta$ -CDs (data not shown for conciseness).

## 5. COMPUTATIONAL SPEED-UP

The PITOMBA force field is 1–2 orders of magnitude computationally faster than its atomistic counterparts. The maximum integration step that can be used without constraining bond stretching is 7 fs (Figure 13).

## 6. FINAL CONSIDERATIONS

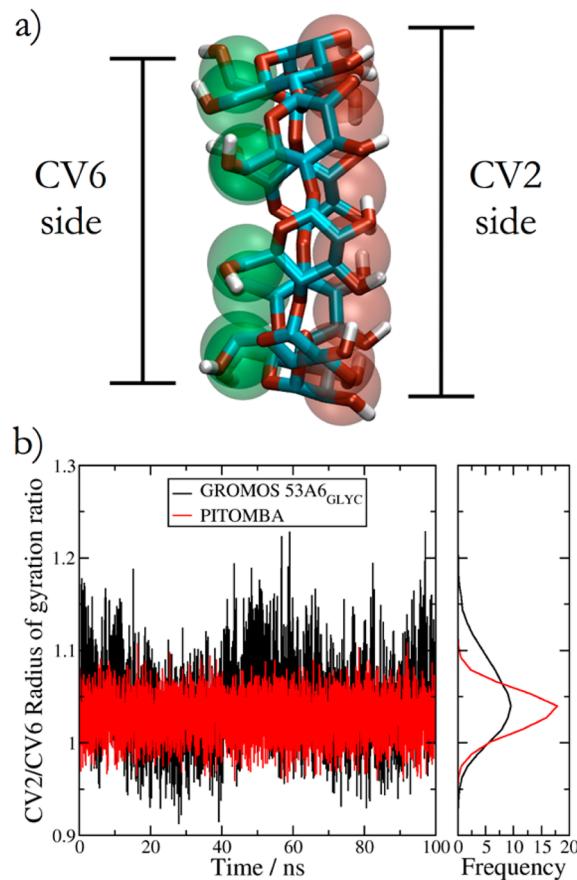
We present our newly created force field for carbohydrates called PITOMBA. The model was developed for condensed phase simulations and it is compatible with the GROMOS CG water model. The glucose molecule is represented by a four-bead model and a single dihedral defining the  $\alpha$  and  $\beta(1 \rightarrow 4)$  glycosidic linkage. The parametrization approach used the GROMOS 53A6<sub>GLYC</sub> force field in order to derive the bonded parameters and the HF/6-31G(d) rotational profile to derive the dihedral terms. Electrostatic interactions are considered by adding charges to the beads that can reproduce the average



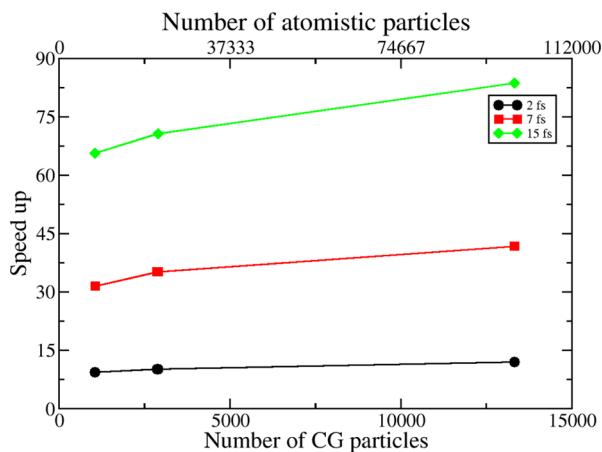
**Figure 11.** (a) Root mean square deviation of  $\gamma$ -cyclodextrin for 100 ns simulations in GROMOS 53A6<sub>GLYC</sub> atomistic and PITOMBA CG models. (b) Initial and (c) final structures of  $\gamma$ -cyclodextrin in GROMOS 53A6<sub>GLYC</sub> atomistic and PITOMBA CG models, respectively.

atomistic dipole moment of glucose. The Lennard-Jones terms are rescaled from the atomistic GROMOS 53A6<sub>GLYC</sub> in order to fit experimental free energies. All the D-glucopyranose monomers are considered to be in the  $^4C_1$  ring conformation, which is the only accessible ring conformation of glucose in water.

The PITOMBA model is able to predict with good accuracy structural features of oligo- and polysaccharides in solution as well as their flexibility in water, when compared to experimental data and atomistic GROMOS 53A6<sub>GLYC</sub> force field simulations. In addition, it also properly describes interparticle (carbohydrate–carbohydrate and carbohydrate–water) interactions, as represented by the number of water molecules inside the  $\alpha$ ,  $\beta$ , and  $\gamma$ -cyclodextrins cavities and the radial distribution functions of a glucose solution. These results demonstrate the compatibility of the PITOMBA with the CG water model and is therefore expected to work well with the remaining GROMOS CG solvents. Moreover, the model is easily



**Figure 12.** (a) Representation of both faces of  $\gamma$ -cyclodextrin and (b) their ratio in the PITOMBA and GROMOS 53A6<sub>GLYC</sub> atomistic models. The histogram of the ratio is represented on the right side of b. The CV2 side encompasses atoms C2, C3, O2, O3, HO2, and HO3 and the CV6 side encompasses the hydroxymethyl group in the atomistic representation.



**Figure 13.** Speed up of the PITOMBA force field for different time steps compared to the 2 fs time step used in GROMOS 53A6<sub>GLYC</sub> for the molecules of glucose,  $\gamma$ -cyclodextrin, and 18-mer of amylose A. CG bond stretching was constrained in the 15 fs runs.

extendable to include carbohydrate chains with different glycosidic linkages and branches.

The main limitation of the current model is the narrower distribution of bond distances and angles compared to GROMOS 53A6<sub>GLYC</sub> force field. Therefore, the PITOMBA

model is likely to hinder intrachain dynamics for long oligo- and polysaccharides. However, it is not expected to affect packing.

This model lays the foundation for large-scale simulations of carbohydrate moieties, such as cellulose slabs, free-energy calculations of structural carbohydrate remodeling and to further development of a CG lipid model to investigate the dynamics of glycolipid membranes and aggregates. Due to its compatibility to the GROMOS SPC CG water model, it can potentially be used for multiscale simulation with atomistic SPC water.

## AUTHOR INFORMATION

### Corresponding Author

\*Email: roberto.lins@ufpe.br.

### Author Contributions

V.H.R. proposed the model, implemented the computational tools, executed the simulations, analyzed the results and wrote the first draft of the manuscript. R.B. contributed with comments and suggestions, and helped write and revise the manuscript. R.D.L. contributed with comments and suggestions, coordinated the project, and helped write and revise the manuscript. All authors read and agreed with the final version of this article.

### Funding

CNPq, FACEPE, INCT-INAMI, CAPES, NanoBiotec-BR, and nBioNet.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Sereina Riniker for valuable comments and for providing the input files for the GROMOS CG water model and Mauricio Coutinho-Neto e Ricardo Longo for thorough discussions. This research was supported by FACEPE, CNPq, CAPES, and STINT. V.H.R. acknowledges support from CAPES through a Nanobiotec-BR/CAPES scholarship Proc. 9614-12-5. Partial computational resources were provided by the Argonne National Laboratory, a facility sponsored by the U.S. Department of Energy and the Extreme Science and Engineering Discovery Environment (XSEDE) supercomputers. XSEDE is supported by National Science Foundation grant number OCI-1053575. R.B. thanks the University of Utah and the Department of Medicinal Chemistry for startup funding.

## REFERENCES

- (1) Marradi, M.; Martín-Lomas, M.; Penadés, S. Glyconanoparticles: Polyvalent Tools to Study Carbohydrate-Based Interactions. In *Advances in Carbohydrate Chemistry and Biochemistry*; Horton, D., Ed.; Academic Press: New York, 2010; Vol. 64, pp 211–290.
- (2) Hollingsworth, R. I.; Wang, G. Toward a Carbohydrate-Based Chemistry: Progress in the Development of General-Purpose Chiral Synthons From Carbohydrates. *Chem. Rev.* **2000**, *100*, 4267–4282.
- (3) Dwek, R. A.; Butters, T. D. Introduction: Glycobiology Understanding the Language and Meaning of Carbohydrates. *Chem. Rev.* **2002**, *102*, 283–284.
- (4) McReynolds, K. D.; Gervay-Hague, J. Chemotherapeutic Interventions Targeting HIV Interactions with Host-Associated Carbohydrates. *Chem. Rev.* **2007**, *107*, 1533–1552.
- (5) Murray, H. E.; Hsieh-Wilson, L. C. The Chemical Neurobiology of Carbohydrates. *Chem. Rev.* **2008**, *108*, 1708–1731.

- (6) López, C. A.; de Vries, A. H.; Marrink, S. J. Amylose Folding Under the Influence of Lipids. *Carbohydr. Res.* **2012**, *364*, 1–7.
- (7) Straathof, A. J. J. Transformation of Biomass Into Commodity Chemicals Using Enzymes or Cells. *Chem. Rev.* **2013**, *114*, 1871–1908.
- (8) Sapsford, K. E.; Algar, W. R.; Berti, L.; Gemmill, K. B.; Casey, B. J.; Oh, E.; Stewart, M. H.; Medintz, I. L. Functionalizing Nanoparticles with Biological Molecules: Developing Chemistries That Facilitate Nanotechnology. *Chem. Rev.* **2013**, *113*, 1904–2074.
- (9) Rauch, J.; Kolch, W.; Laurent, S.; Mahmoudi, M. Big Signals From Small Particles: Regulation of Cell Signaling Pathways by Nanoparticles. *Chem. Rev.* **2013**, *113*, 3391–3406.
- (10) Reith, D.; Pütz, M.; Müller-Plathe, F. Deriving Effective Mesoscale Potentials From Atomistic Simulations. *J. Comput. Chem.* **2013**, *24*, 1624–1636.
- (11) Lyubartsev, A. P.; Laaksonen, A. Calculation of Effective Interaction Potentials From Radial Distribution Functions: A Reverse Monte Carlo Approach. *Phys. Rev. E* **1995**, *52*, 3730–3737.
- (12) Moritsugu, K.; Smith, J. C. Coarse-Grained Biomolecular Simulation with REACH: Realistic Extension Algorithm via Covariance Hessian. *Biophys. J.* **2007**, *93*, 3460–3469.
- (13) Hynninen, A.-P.; Matthews, J. F.; Beckham, G. T.; Crowley, M. F.; Nimlos, M. R. Coarse-Grain Model for Glucose, Cellobiose, and Cellotetraose in Water. *J. Chem. Theory Comput.* **2011**, *7*, 2137–2150.
- (14) Molinero, V.; Goddard, W. A. M3B: A Coarse Grain Force Field for Molecular Simulations of Malto-Oligosaccharides and Their Water Mixtures. *J. Phys. Chem. B* **2004**, *108*, 1414–1427.
- (15) Liu, P.; Izvekov, S.; Voth, G. A. Multiscale Coarse-Graining of Monosaccharides. *J. Phys. Chem. B* **2007**, *111*, 11566–11575.
- (16) Bu, L.; Beckham, G. T.; Crowley, M. F.; Chang, C. H.; Matthews, J. F.; Bomble, Y. J.; Adney, W. S.; Himmel, M. E.; Nimlos, M. R. The Energy Landscape for the Interaction of the Family 1 Carbohydrate-Binding Module and the Cellulose Surface Is Altered by Hydrolyzed Glycosidic Bonds. *J. Phys. Chem. B* **2009**, *113*, 10994–11002.
- (17) López, C. A.; Rzepiela, A. J.; de Vries, A. H.; Dijkhuizen, L.; Hünenberger, P. H.; Marrink, S. J. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.* **2009**, *5*, 3195–3210.
- (18) Wohlert, J.; Berglund, L. A. A Coarse-Grained Model for Molecular Dynamics Simulations of Native Cellulose. *J. Chem. Theory Comput.* **2011**, *7*, 753–760.
- (19) Srinivas, G.; Cheng, X.; Smith, J. C. A Solvent-Free Coarse Grain Model for Crystalline and Amorphous Cellulose Fibrils. *J. Chem. Theory Comput.* **2011**, *7*, 2539–2548.
- (20) Markutsya, S.; Kholod, Y. A.; Devarajan, A.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. A Coarse-Grained Model for  $\beta$ -D-Glucose Based on Force Matching. *Theor. Chem. Acc.* **2012**, *131*, 1162–1176.
- (21) Bellesia, G.; Chundawat, S. P. S.; Langan, P.; Redondo, A.; Dale, B. E.; Gnanakan, S. Coarse-Grained Model for the Interconversion Between Native and Liquid Ammonia-Treated Crystalline Cellulose. *J. Phys. Chem. B* **2012**, *116*, 8031–8037.
- (22) Glass, D. C.; Moritsugu, K.; Cheng, X.; Smith, J. C. REACH Coarse-Grained Simulation of a Cellulose Fiber. *Biomacromolecules* **2012**, *13*, 2634–2644.
- (23) Markutsya, S.; Devarajan, A.; Baluyut, J. Y.; Windus, T. L.; Gordon, M. S.; Lamm, M. H. Evaluation of Coarse-Grained Mapping Schemes for Polysaccharide Chains in Cellulose. *J. Chem. Phys.* **2013**, *138*, 214108.
- (24) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A Generic Force Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.
- (25) Ercolessi, F.; Adams, J. B. Interatomic Potentials From First-Principles Calculations: The Force-Matching Method. *EPL (Europhysics Letters)* **1994**, *26*, 583–588.
- (26) Izvekov, S.; Parrinello, M.; Burnham, C. J.; Voth, G. A. Effective Force Fields for Condensed Phase Systems From Ab Initio Molecular Dynamics Simulation: A New Method for Force-Matching. *J. Chem. Phys.* **2004**, *120*, 10896–10912.
- (27) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
- (28) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. The Missing Term in Effective Pair Potentials. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
- (29) Im, W.; Lee, M. S.; Brooks, C. L. Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.* **2003**, *24*, 1691–1702.
- (30) López, C. A.; Sovova, Z.; van Eerden, F. J.; de Vries, A. H.; Marrink, S. J. Martini Force Field Parameters for Glycolipids. *J. Chem. Theory Comput.* **2013**, *9*, 1694–1708.
- (31) Oostenbrink, C.; Villa, A.; Mark, A. E.; van Gunsteren, W. F. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation: The GROMOS Force-Field Parameter Sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25*, 1656–1676.
- (32) Lins, R. D.; Hünenberger, P. H. A New GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Comput. Chem.* **2005**, *26*, 1400–1412.
- (33) MacKerell, A. D.; Bashford, D.; Bellott, Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
- (34) Moritsugu, K.; Smith, J. C. REACH Coarse-Grained Biomolecular Simulation: Transferability Between Different Protein Structural Classes. *Biophys. J.* **2008**, *95*, 1639–1648.
- (35) Moritsugu, K.; Kurkal-Siebert, V.; Smith, J. C. REACH Coarse-Grained Normal Mode Analysis of Protein Dimer Interaction Dynamics. *Biophys. J.* **2009**, *97*, 1158–1167.
- (36) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (37) Autieri, E.; Sega, M.; Pederiva, F.; Guella, G. Puckering Free Energy of Pyranoses: A NMR and Metadynamics-Umbrella Sampling Investigation. *J. Chem. Phys.* **2010**, *133*, 095104.
- (38) Spiwok, V.; Králová, B.; Tvaroška, I. Modelling of B-D-Glucopyranose Ring Distortion in Different Force Fields: A Metadynamics Study. *Carbohydr. Res.* **2010**, *345*, 530–537.
- (39) Barnett, C. B.; Naidoo, K. J. Ring Puckering: A Metric for Evaluating the Accuracy of AM1, PM3, PM3CARB-1, and SCC-DFTB Carbohydrate QM/MM Simulations. *J. Phys. Chem. B* **2010**, *114*, 17142–17154.
- (40) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. *Carbohydrates. J. Comput. Chem.* **2008**, *29*, 622–655.
- (41) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. Molecular Mechanics. The MM3 Force Field for Hydrocarbons. *I. J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.
- (42) Stortz, C. A.; Johnson, G. P.; French, A. D.; Csonka, G. I. Comparison of Different Force Fields for the Study of Disaccharides. *Carbohydr. Res.* **2009**, *344*, 2217–2228.
- (43) Pol-Fachin, L.; Rusu, V. H.; Verli, H.; Lins, R. D. GROMOS 53A6GLYC, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates. *J. Chem. Theory Comput.* **2012**, *8*, 4681–4690.
- (44) Riniker, S.; van Gunsteren, W. F. A Simple, Efficient Polarizable Coarse-Grained Water Model for Molecular Dynamics Simulations. *J. Chem. Phys.* **2011**, *134*, 084110.
- (45) Roothaan, C. C. J. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.* **1951**, *23*, 69–89.
- (46) Hehre, W. J. Self—Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian—Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261.

- (47) Berendsen, H.; Postma, J.; van Gunsteren, W. F.; Hermans, J. *Intermolecular Forces*; Pullman, B., Ed.; Reidel Publishing Company: Dordrecht, 1981; pp 331–342.
- (48) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. LINCS: A Linear Constraint Solver for Molecular Simulations. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (49) Tironi, I. G.; Sperb, R.; Smith, P. E.; van Gunsteren, W. F. A Generalized Reaction Field Method for Molecular Dynamics Simulations. *J. Chem. Phys.* **1995**, *102*, 5451–5459.
- (50) Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling Through Velocity Rescaling. *J. Chem. Phys.* **2007**, *126*, 014101.
- (51) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (52) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447.
- (53) Bennett, C. H. Efficient Estimation of Free Energy Differences From Monte Carlo Data. *J. Comput. Phys.* **1976**, *22*, 245–268.
- (54) Beutler, T. C.; Mark, A. E.; van Schaik, R. C.; Gerber, P. R.; van Gunsteren, W. F. Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem. Phys. Lett.* **1994**, *222*, 529–539.
- (55) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision D.01; Gaussian, Inc.: Wallingford, CT, 2009.
- (56) Schuler, L. D.; Daura, X.; van Gunsteren, W. F. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *J. Comput. Chem.* **2001**, *22*, 1205–1218.
- (57) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Universal Quantum Mechanical Model for Solvation Free Energies Based on Gas-Phase Geometries. *J. Phys. Chem. B* **1998**, *102*, 3257–3271.
- (58) Gonçalvez, P. F. B.; Stassen, H. Calculation of the Free Energy of Solvation From Molecular Dynamics Simulations. *Pure Appl. Chem.* **2004**, *76*, 231–240.
- (59) Jämbeck, J. P. M.; Mocci, F.; Lyubartsev, A. P.; Laaksonen, A. Partial Atomic Charges and Their Impact on the Free Energy of Solvation. *J. Comput. Chem.* **2013**, *34*, 187–197.
- (60) Mark, H. X-ray Investigation of Carbohydrates. *Chem. Rev.* **2013**, *26*, 169–186.
- (61) Le Bail, P.; Bizot, H.; Pontoire, B.; Buléon, A. Polymorphic Transitions of Amylose–Ethanol Crystalline Complexes Induced by Moisture Exchanges. *Starch/Stärke* **1995**, *47*, 229–232.
- (62) Müller, J. J.; Gernat, C.; Schulz, W.; Müller, E.-C.; Vorwerg, W.; Damaschun, G. Computer Simulations of X-ray Scattering Curves: Gelation and Crystallization Process in Amylose Solutions. *Biopolymers* **1995**, *35*, 271–288.
- (63) Leloup, V. M.; Colonna, P.; Ring, S. G.; Roberts, K.; Wells, B. Microstructure of Amylose Gels. *Carbohydr. Polym.* **1992**, *18*, 189–197.
- (64) Shimada, J.; Kaneko, H.; Takada, T.; Kitamura, S.; Kajiwara, K. Conformation of Amylose in Aqueous Solution: Small-Angle X-ray Scattering Measurements and Simulations. *J. Phys. Chem. B* **2000**, *104*, 2136–2147.
- (65) Imbert, A.; Pérez, S. A Revisit to the Three-Dimensional Structure of B-Type Starch. *Biopolymers* **1988**, *27*, 1205–1221.
- (66) Hinrichs, W.; Saenger, W. Crystal and Molecular Structure of the Hexasaccharide Complex (P-Nitrophenyl  $\alpha$ -Maltohexaose)2·Ba(I<sub>3</sub>)<sub>2</sub>·27H<sub>2</sub>O. *J. Am. Chem. Soc.* **2013**, *112*, 2789–2796.
- (67) Hsein-Chih, H. W.; Sarko, A. The Double-Helical Molecular Structure of Crystalline B-Amylose. *Carbohydr. Res.* **1978**, *61*, 7–25.
- (68) Imbert, A.; Chanzy, H.; Pérez, S.; Buléon, A.; Tran, V. The Double-Helical Nature of the Crystalline Part of  $\alpha$ -Starch. *J. Mol. Biol.* **1988**, *201*, 365–378.
- (69) Hsieh-Chih, H. W.; Sarko, A. The Double-Helical Molecular Structure of Crystalline  $\alpha$ -Amylose. *Carbohydr. Res.* **1978**, *61*, 27–40.
- (70) Kurkov, S. V.; Loftsson, T. Cyclodextrins. *Poorly Soluble Drugs* **2013**, *453*, 167–180.
- (71) Venkatesh, G.; Thulasidhasan, J.; Rajendiran, N. A Spectroscopic and Molecular Modeling Studies of the Inclusion Complexes of Orciprenaline and Terbutaline Drugs with Native and Modified Cyclodextrins. *J. Inclusion Phenom. Macrocyclic Chem.* **2014**, *78*, 225–237.
- (72) Zhang, J.; Ma, P. X. Cyclodextrin-Based Supramolecular Systems for Drug Delivery: Recent Progress and Future Perspective. *Adv. Drug Delivery Rev.* **2013**, *65*, 1215–1233.
- (73) Kahle, C.; Holzgrabe, U. Determination of Binding Constants of Cyclodextrin Inclusion Complexes with Amino Acids and Dipeptides by Potentiometric Titration. *Chirality* **2004**, *16*, 509–515.
- (74) Kolhapurkar, R.; Patil, K. Studies of Volumetric and Activity Behaviors of Binary and Ternary Aqueous Solutions Containing B-Cyclodextrin and Glucose. *J. Mol. Liq.* **2013**, *178*, 185–191.
- (75) Lindner, K.; Saenger, W. B-Cyclodextrin Dodecahydrate: Crowding of Water Molecules Within a Hydrophobic Cavity. *Angew. Chem., Int. Ed. Engl.* **1978**, *17*, 694–695.
- (76) Ponce Cevallos, P. A.; Buera, M. P.; Elizalde, B. E. Encapsulation of Cinnamon and Thyme Essential Oils Components (Cinnamaldehyde and Thymol) in B-Cyclodextrin: Effect of Interactions with Water on Complex Stability. *J. Food Eng.* **2010**, *99*, 70–75.
- (77) Song, L. X.; Bai, L.; Xu, X. M.; He, J.; Pan, S. Z. Inclusion Complexation, Encapsulation Interaction and Inclusion Number in Cyclodextrin Chemistry. *Coord. Chem. Rev.* **2009**, *293*, 1276–1284.
- (78) Gelb, R. I.; Schwartz, L. M.; Cardelino, B.; Fuhrman, H. S.; Johnson, R. F.; Laufer, D. A. Binding Mechanisms in Cyclohexaamyllose Complexes. *J. Am. Chem. Soc.* **2013**, *103*, 1750–1757.
- (79) Inoue, Y.; Hakushi, T.; Liu, Y.; Tong, L.; Shen, B.; Jin, D. Thermodynamics of Molecular Recognition by Cyclodextrins. 1. Calorimetric Titration of Inclusion Complexation of Naphthalenesulfonates with  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Cyclodextrins: Enthalpy–Entropy Compensation. *J. Am. Chem. Soc.* **2013**, *115*, 475–481.
- (80) Rekharsky, M. V.; Goldberg, R. N.; Schwarz, F. P.; Tewari, Y. B.; Ross, P. D.; Yamashoji, Y.; Inoue, Y. Thermodynamic and Nuclear Magnetic Resonance Study of the Interactions of  $\alpha$ - and  $\beta$ -Cyclodextrin with Model Substances: Phenethylamine, Ephedrines, and Related Substances. *J. Am. Chem. Soc.* **2013**, *117*, 8830–8840.
- (81) Liu, Y.; Cao, R.; Chen, Y.; He, J.-Y. Effect of B-Cyclodextrin Charge Type on the Molecular Recognition Thermodynamics of Reactions with (Ferrocenylmethyl)dimethylaminium Derivatives. *J. Phys. Chem. B* **2013**, *112*, 1445–1450.