1
2
3 **Comparison of the Simulations of Cellulosic Crystals**
4 **with Three Carbohydrate Force Fields**
5
6
7

8
9 Hitomi Miyamoto^{#,†}, Udo Schnupf^{f,†}, Michael Crowley[§]
10 and John W. Brady^{*,†}
11
12
13
14

15 [†]Department of Food Science
16 Cornell University
17 Ithaca, NY 14853
18

19
20 [§]Biosciences Center and the National Bioenergy Center
21 National Renewable Energy Laboratory
22 Golden, CO 80401
23

24
25 [#]Department of Chemical Engineering
26 Technion – Israel Institute of Technology
27 Technion City, Haifa 32000, Israel
28

29
30 ^f Department of Chemistry and Biochemistry
31 Bradley University
32 Peoria, IL 61625
33

34
35 * Author to whom correspondence should be addressed; +1 (607) 255-2897;
36 jwb7@cornell.edu

Abstract

Three independently developed molecular mechanics force fields for carbohydrates have been used to simulate a suite of small molecule analogs of cellulose for which crystal structures have been reported, as a test to determine which might be best for simulations of cellulose itself. Such evaluation is necessary since the reported cellulose crystal structure is not stable in molecular dynamics simulations with any available force field. The present simulations found that all three resulted in small deviations from the reported crystal structures, but that all were reasonably accurate and none was clearly superior to the others for the entire suite of structures examined.

Introduction

Molecular mechanics simulations can only give a useful description of the physical behavior of a system if the force fields that are employed are a reasonably accurate approximation of the actual variation in system energy with atomic coordinates. Within that broad statement, some physical properties are more sensitive to force field details than others. The first molecular dynamics (MD) simulation, of a two-dimensional rare gas fluid at low temperature, used a simple hard disk potential function, but nonetheless exhibited a melting/freezing transition¹ even though there was no attractive component to the force field at all. Similarly, the initial studies by Ramachandran and co-workers of peptide conformations using a similar hard-wall force field were able to delineate many of the most important broad features governing protein conformations.² On the other hand, studies in recent decades have found many cases where important physical properties are very sensitive to force field details or approximations, such as the treatment of long range interactions and truncations, charge distributions, rotational barriers, etc. Fortunately, for many biologically important types of molecules, force fields have been developed that perform adequately for most important properties of interest, which is usually demonstrated by the successful calculation of known molecular properties that were not used in the force field development. For proteins, nucleic acids, lipids, and carbohydrates, generally several force fields have been independently developed, which often give similar qualitative results, both validating the general approach and indicating that all have captured the important broad features. However, it is also generally true that one force field parameterization or another might give the best fit to a suite of physical properties, and

it has been accepted in the past to conduct head-to-head comparisons to determine which force field might be most accurate for a particular purpose.³

Carbohydrates are one of the more important categories of biological molecules, given their many roles in energy storage, molecular recognition, as structural components in the nucleic acids, and as the most abundant single biomolecule in the biosphere, cellulose. Many carbohydrate force fields have been developed in the past, and recent parameterizations have reached a very high level of sophistication.⁴ These force fields play a crucial role in the efforts to understand cellulose structure and its enzymatic deconstruction into free sugars for use as a fermentation feedstock in the production of bioethanol, as well as the behavior many other biologically and medically important carbohydrates containing glucose monomers.

Unfortunately, cellulose may be one of those difficult cases where force field details play a crucial role in determining important calculated properties, including its crystalline structure. In the reported crystal structure for cellulose I β ,⁵ the chains exhibit parallel packing in sheets with individual chains in a two-fold screw conformation, with all of the exocyclic C6 primary alcohol groups in the so-called *tg* conformation,⁶ making hydrogen bonds along the same chain and between chains in the same crystal layer, but with no hydrogen bonds between layers, much like the structure of the peptide that makes up silk. This situation is unusual, since few other glucose-containing molecules exhibit this exocyclic conformation preference. It is not seen in either the crystal structures of glucose nor observed in MD simulations of glucose in aqueous solution, and is also not found in the crystal structures of cellobiose, cellobioside,

cellotriose, or cellotetraose,⁷⁻¹³ the most important small-molecule component analogs of cellulose, **although there are rare examples of this conformation.**¹⁴ In this context it is interesting that recent very long-time MD simulations of small diameter fibrils^{15,16} have found that this putative crystal structure of cellulose I β is not stable with any of the three major force fields recently developed to model carbohydrates: CHARMM (C35),^{17,18} GROMOS 45a4,¹⁹ or the AMBER-related GLYCAM06.²⁰ It is possible **the tg O-6 orientation** of the reported crystal structure of cellulose I β is in error, but it is also possible that the crystal packing of cellulose produces a collective field that alters the electronic structure of the polymer in a way that cannot be adequately captured by the static charge models used in these standard force fields. If so, the new generation of polarizable force fields for carbohydrates currently under development^{4,21} may be required to realistically model cellulose. However, it is still useful to conduct the same sort of direct comparisons between available force fields as have been used in the past to determine which might be best for conventional MD simulations of cellulose. Here we report such a comparison of three force fields using the known crystal structures of a suite of small molecule cellulose analogs as standards.

Procedures

MD simulations were carried out with the general molecular mechanics program CHARMM, version 36.^{22,23} The sugar atoms were modeled using three force fields specifically developed for carbohydrates, CHARMM35 (C35),^{17,18} GLYCAM06,²⁰ and an earlier CHARMM-style parameterization, designated here as CSFF.²⁴⁻²⁶ Water

120 molecules were represented using the CHARMM-modified TIP3P force field.^{27,28} The
 121 force fields previously developed for ethanol and methanol molecules were used for the
 122 crystals containing alcohols.²⁹⁻³¹

Figure 1. Packing arrangements of seven different types of cellulosic crystals: a, cellobiose; b, cellobioside; c, cellobioside; d, cellobioside; e, cellobioside+methanol; f, cellotetraose+water; g, cellotriose+water+ethanol.

129 The initial simulation models were constructed using the unit cells of seven
 130 different crystals as shown in Figure 1: a, cellobiose;⁷ b, c, and d, three types of
 131 cellobioside;⁸⁻¹⁰ e, cellobioside+methanol;¹¹ f, cellotetraose+water;¹² g,
 132 cellotriose+water+ethanol.¹³ The unit-cell parameters of the crystals are summarized
 133 in Table 1. The cellobiose crystal has the space group $P2_1$ with two molecules per
 134 unit-cell (Figure 1a). One of the cellobioside crystals is also in the $P2_1$ space group,

with two molecules in the unit cell (Figure 1b). Another cellobioside crystal is in the $P1$ space group, with one cellobioside per unit cell (Figure 1c), and a third is again in the $P2_1$ space group with two molecules per unit cell (Figure 1d). The cellobioside+methanol crystal is in the $P2_1$ space group with two cellobioside and two methanol molecules for per unit cell (Figure 1e). The cellotetraose hydrate crystal is in the $P1$ space group with two cellotetraose molecules and one water molecule per unit cell (Figure 1f), while the cellotriose+water+ethanol crystal is in the $P2_1$ space group with eight cellotriose molecules, eight water molecules, and two ethanol molecules per unit cell (Figure 1g).

	Unit-cell parameters						space group
	a / Å	b / Å	c / Å	α / °	β / °	γ / °	
Cellobiose	10.972	13.048	5.091	90.0	90.83	90.0	$P2_1$
Cellobioside (Figure1b)	6.606	14.074	9.318	90.0	108.95	90.0	$P2_1$
Cellobioside (Figure1c)	4.6726	7.5120	12.4270	84.7	86.0	79.7	$P1$
Cellobioside (Figure1d)	6.617	14.117	9.344	90.0	108.81	90.0	$P2_1$
Cellobioside+methanol	7.652	25.532	4.496	90.0	101.84	90.0	$P2_1$
Cellotetraose+water	8.045	9.003	22.510	89.66	94.83	115.80	$P1$
Cellotriose+water+ethanol	7.9978	76.3800	8.9908	90.0	116.40	90.0	$P2_1$

Table 1. The unit cell parameters of the crystal structures used here.

In setting up the computational models of these crystals, each unit-cell was extended to be three times as large for each axis, and the resulting crystals were used as initial simulation models.

Chemical bonds to hydrogen atoms were kept fixed using the general constraint algorithm SHAKE.^{32,33} van der Waals interactions were smoothly truncated on an atom-by-atom basis using switching functions from 12.0 to 14.0 Å.²² Electrostatic interactions were treated using the particle-mesh Ewald method³⁴ with a real space cutoff of 12 Å and a κ value of 0.3 Å⁻¹. The system was run at a constant temperature of 300 K and at a constant pressure of 1 bar using a Hoover constant pressure/temperature algorithm, maintained with a Hoover piston with a mass of 2000

amu.³⁵⁻³⁷ Initial velocities were assigned from a Boltzmann distribution at 300 K, and the trajectories were run for 70 ns with a step size of 1 fs using full periodic boundary conditions. The last 20 ns of the simulation time of each simulation were used for the analysis. Molecular graphics were generated with Visual Molecular Dynamics (VMD1.9)³⁸ and Cambridge Crystallographic Data Centre (CCDC) Mercury 3.0.^{39,40}

Results and Discussion

The dimensional changes for the unit cells of the seven different cellulosic crystals are shown in Figure 2. This figure illustrates the difference between the crystal structure values and the average during those 20 ns. Therefore, the crystal structure values are zero in the vertical axis in Figure 2. The dimensional changes in simulations with the three force fields for carbohydrates that were tested (C35, GLYCAM06, CSFF) remained less than ± 14.5 % from the crystal structure values. The change in unit-cell dimensions was the highest for the β -angle of the cellotriose+water+ethanol crystal simulated with the C35 force field (Figure 2g) at 14.5 %, followed by the β -angle of the cellobioside crystal using C35 (Figure 2c) at 13.7 %, the a-axis of the cellobioside crystal using CSFF (Figure 2d) at 12.0 %, and the a-axis of the cellobioside crystal with CSFF (Figure 2b) at 11.3 %. From these results, the dimensions of the crystals simulated with the C35 and CSFF force fields changed significantly in several cases,

whereas few major changes in dimensions were observed in the simulations with GLYCAM06. However, the magnitudes of the dimensional changes in crystals simulated with the GLYCAM06 force field were high for cellobiose, cellobioside (Figure 2(b) and (d)), cellotetraose+water (f), and cellotriose+water+ethanol (g) crystals. Thus, the models simulated with the GLYCAM06 force field did not exhibit major changes in single dimensions as illustrated above, but they changed the most with respect to the sum of all dimension changes.

Figure 2. The changes in the unit-cell dimensions of seven different types of cellulosic crystals: a, cellobiose; b, cellobioside; c, cellobioside; d, cellobioside; e, cellobioside+methanol; f, cellotetraose+water; g, cellotriose+water+ethanol. The simulations with C35, GLYCAM06, and CSFF were colored by blue, yellow, and pink, respectively.

The glycosidic linkage dihedral angles ϕ (O5–C1–O1–C4) and ψ (C1–O1–C4–C5) determine the backbone conformation in cellulosic polymers (Figure 3). These angles are indicated by points colored gray in Figure 3 for the conformations observed in the various crystal structures used here.^{7-9,11-13} The conformations of the seven different crystals with each of the force fields used here showed results similar to those of the unit cell parameters. Thus, it can be deduced that all of these carbohydrate force fields are adequate to simulate cellulosic models from the viewpoint of the main chain conformations. A close inspection indicated that slight changes in the main chain conformations were observed in the cellobioside crystal (pink in Figure 3 (b) and (d)) simulated with CSFF and the cellobiose crystal (blue in Figure 3 (a)) modeled with C35. The cellobioside glycosidic bond angles (ϕ , ψ) in these crystals are reported in degrees as follows: Figure 3 (a), (-76.3, -132.3); (b), (-89.1, -152.0); (d), (-88.8, -151.3). The averaged glycosidic bond angles during the last 20 ns of the simulation times were (-81.9, -144.7) simulated with C35 (Figure 3 (a)), (-79.3,-132.2) with CSFF (Figure 3 (b)), and (-79.0, -131.8) with CSFF (Figure 3 (d)).

<i>O3-H-05</i>	<i>C35</i>	<i>GLYCAM06</i>	<i>CSFF</i>
Cellobiose	0.98	0.97	0.96
Cellobioside (Figure1b)	0.91	0.98	0.98
Cellobioside (Figure1c)	0.98	0.99	1.00
Cellobioside (Figure1d)	0.90	0.96	0.98
Cellobioside+methanol	0.98	1.00	1.00
Cellotetraose+water	0.70	0.82	0.96
Cellotriose+water+ethanol	0.82	0.91	0.93

<i>O3-H-06</i>	<i>C35</i>	<i>GLYCAM06</i>	<i>CSFF</i>
Cellobiose	0.27	0.00	0.02
Cellobioside (Figure1b)	0.15	0.00	0.02
Cellobioside (Figure1c)	0.42	0.27	0.34
Cellobioside (Figure1d)	0.16	0.00	0.02
Cellobioside+methanol	0.54	0.17	0.26
Cellotetraose+water	0.06	0.02	0.14
Cellotriose+water+ethanol	0.11	0.06	0.13

Table 2. Intramolecular hydrogen bonds¹ calculated from the simulations.

¹The criteria used for identification of hydrogen bonds were that the A–H distance should be less than 2.8 Å and the D–H–A angle should be greater than 110°, where A and D represent acceptor and donor atoms, respectively. The number listed is the occurrence, that is, the ratio of number of times the H-bond meets the criteria to the total number of samples.

The glycosidic bond angles of the crystals containing dimers such as cellobiose and cellobioside crystals shifted slightly during the simulations, whereas the angles of the crystals which contain cellulosic molecules with degrees of polymerization (DP) greater than two (cellotetraose+water and cellotriose+water+ethanol) with the three force fields were similar to the crystal data. It might be concluded that these three force fields specifically developed for carbohydrates are approximately equally suitable for use in

simulations of cellulose from the results obtained by the analysis of main chain conformations.

Figure 3. ϕ/ψ plot of glycosidic linkage conformations: experimental crystal structures colored by black (circle), C35, GLYCAM06, and CSFF colored by blue (square), yellow (rhombus), and pink (triangle), respectively. Seven different types of cellulosic crystals were simulated: a, cellobiose; b, cellobioside; c, cellobioside; d, cellobioside; e, cellobioside+methanol; f, cellotetraose+water; g, cellotriose+water+ethanol.

The occupancy ratios of the intramolecular hydrogen bonds observed in the simulations are listed in Table 2. The criteria used to identify hydrogen bonds included that the A–H distance should be less than 2.8 Å and the D–H–A angle should be greater than 110°, where A and D represent acceptor and donor atoms, respectively. The occupancy is the ratio of number of times the H-bond meets the criteria to the

245 total number of samples; the measurement covers all glucose residues. For each crystal,
246 it has been reported that the major intramolecular hydrogen bond in these seven
247 crystals is O3–H...O5, in addition to a minor one of O3–H...O6.⁷⁻¹³ This means that if
248 the crystal structure is totally maintained, the occupancy ratio of the hydrogen bonds
249 in the crystal is 1.00. In the cellulose crystal I β , the intramolecular hydrogen bonds
250 such as O2-H-O6 and O6-H-O2 are prevalently formed; however, these hydrogen
251 bonds could not be observed in the seven simulated models, and the occupancy ratios
252 of these hydrogen bonds were less than 0.02 in the simulations with the three force
253 fields. The occupancy ratios of the intra-chain O3–H...O5 hydrogen bond were
254 relatively high, from 0.70 to 1.00, except for a slightly lower degree of occupancy
255 ratio in the cellotetraose+water model simulated with C35 and GLYCAM06, and the
256 cellotriose+water+ethanol model simulated with C35. A close investigation reveals
257 that the occupancy ratio of intramolecular hydrogen bonds in the crystal models
258 simulated with CSFF was highest, followed by GLYCAM06, with C35 the lowest.
259 The occupancy ratios of the O3–H...O6 intra-chain hydrogen bond were lower (with
260 the range being 0.00-0.54) than those of O3-H...O5. In the case of the minor
261 O3-H...O6 intramolecular hydrogen bond, the hydrogen bonds in crystal models

simulated with C35 were highly maintained, followed by CSFF, with GLYCAM06 the lowest.

Figure 4. Color scheme to indicate primary alcohol group (O5–C5–C6–O6): blue, gt conformation; yellow, gg conformation; pink, tg conformation. The gt conformation is reported in the crystals. In the case of the cellobiose crystal, when residues form the gg conformation of more than 50 % during the last 20 ns in the simulation time, those residues are colored by yellow. In the case of the cellotetraose+water crystal, when residues form the gg (tg) conformation of more than 50 % during the last 20 ns in the simulation time, those residues are colored by yellow (pink).

Figure 4 illustrates changes of the primary alcohol group, O5–C5–C6–O6 observed in the simulations. The three low-energy conformations for this group are usually designated as gt, gg, and tg, where the first letter specifies the orientation of the C6–O6 bond as *trans* or *gauche* with respect to the C5–O5 bond, and the second

specifies the orientation with respect to the C4-C5 bond.⁶ It has been reported that this conformation is gt in cellobiose and the cellotetraose+water crystals. In the case of the cellobiose crystal simulated with GLYCAM06, the gt conformation clearly remained present during the last 20 ns in the simulation time, as compared with other force fields as can be seen in Figure 4. That is, GLYCAM06 force field is the best to simulate the cellobiose crystal from the viewpoint of side-chain conformation. In the case of the C35 and CSFF force fields, almost half of the hydroxyl groups changed from the gt conformation to the gg conformation. Interestingly, most of the changed hydroxyl groups were the residues of a reducing end. On the other hand, the cellotetraose+water crystal with C35 maintained the gt conformations for the hydroxyl groups compared with the other force fields, and the hydroxyl groups of a few residues in the cellotetraose+water model changed to the gg or tg conformations. In the case of the simulation with the CSFF force field, about half of the hydroxyl groups in cellotetraose+water model transformed to the gg conformation in the same way as the cellobiose model; however, the tg conformation was not observed during the last 20 ns of the simulation. The hydroxyl groups in the cellobiose crystal simulated with GLYCAM06 maintained the gt conformation, but the hydroxyl groups in the cellotetraose+water crystal were not able to keep the gt conformation, and more than

half of the hydroxyl groups transformed to gg or tg conformations. From the viewpoint of the primary alcohol groups, the best force field differs depending on the crystal studied. As mentioned above, there is a tendency that the reproducibility of the main chain simulated with the C35 and CSFF force fields depends on the length of the polymer. However, it seems that the primary alcohol groups of the simulated crystals with CSFF do not depend on the DP. In the case of the models simulated with C35, it was observed that the conformations of the hydroxyl groups in the cellotetraose+water model were better maintained than those in the cellobiose model. That is, from the results of the primary alcohol group and the glycosidic linkage conformations, the conformation under the C35 force field may depend on the DP of the cellulosic polymer.

The intermolecular hydrogen bonding systems of these crystals and the occupancy ratio of these hydrogen bonds during the last 20 ns in the simulations are shown in Figure 5. The heading “A1-A2” shown on the top of the column chart in Figure 5(a) refers to intermolecular hydrogen bonds between the A1 residue and the A2 residue. This means that the A1 glucose in one molecule formed a hydrogen bond to the A2 glucose in another molecule, where A1 and A2 are the glucose rings as shown in the figures of the molecules to the left. In addition, “HO1-O2” shown at the bottom of the

column chart in Figure 5(a) means the occupancy ratios of intermolecular hydrogen bonds between HO1 of A1 and O2 of A2. The definition of the hydrogen bonds has already been described above. For Figure 5(g), the unit-cell is composed of eight isolated cellotriose chains: two chains for each A, B, C, and D. The maximum occupancy ratio of intermolecular hydrogen bonds in the left column chart in Figure 5(g) is 1 because only one hydrogen bond can be formed per chain (a cellotriose chain has only one HO4). The maximum occupancy ratio of the right column chart in Figure 5(g) is 3 because each of three residues of a cellotriose can potentially form a hydrogen bond.

It can be seen that due to thermal fluctuation, the hydrogen exchanges positions between donor and acceptor in the hydrogen bonds during the dynamics run; for example, HO3-O6 and O3-HO6 of A1-A2 in Figure 5(b), HO3-O6 and O3-HO6 of A1-A2 in Figure 5(d), and HO2-O6 and O2-HO6 of A2-A2 in Figure 5(d). In the cellobioside crystal (Figure 5(b)), the intermolecular hydrogen bond of HO3-O6 between A1 and A2 is reported by Mackie et al.^{8,9} Namely, the intermolecular hydrogen bond in the model simulated with the GLYCAM06 force field was highly maintained in the same way as in the crystal. On the other hand, in Figure 5(d), the intermolecular hydrogen bond of O2-HO6 between A2 residues is known to be

331 observed in the crystal, and this hydrogen bond remained only in the model simulated
332 with the C35 force field. For the HO3-O6 and O3-HO6 hydrogen bonds between A1
333 and A2 in Figure 5(d), both of the intermolecular bonds exist in the cellobioside
334 crystal; that is, the hydrogen bonds in the model with the GLYCAM06 force field
335 were maintained more than with the other force fields.
336

Figure 5. Schematic representations and occupancy ratios of the intermolecular hydrogen bonds in the simulations of the seven different types of cellulosic crystals: a, cellobiose; b, cellobioside; c, cellobioside; d, cellobioside; e, cellobioside+methanol; f, cellotetraose+water; g, cellotriose+water+ethanol. Hydrogen bonds are represented by pink dotted lines. For simplicity, each chain in the unit-cell is shown in the right figure of schematic representation of (g). The criteria used for identification of hydrogen bonds were that the A–H distance should be less than 2.8 Å and the D–H–A angle should be greater than 110°, where A and D represent acceptor and donor atoms, respectively. The vertical axis is the occurrence, that is, the ratio of number of times the H-bond meets the criteria to the total number of samples. Only hydrogen bonds with an occupancy ratio above 0.10 are reported. The color scheme for the three force fields is the same as in the previous figures.

Figure 5. (Cont.).

Figure 5. (Cont.).

360

361 Various molecules such as methanol, ethanol, and water are also present in the
362 cellobioside+methanol, cellotetraose+water, and cellotriose+water+ethanol crystals.
363 These species make hydrogen bonds with the carbohydrate molecules in the crystals,
364 and as can be seen in Figure 5, in the models simulated with the GLYCAM06 force
365 field, these hydrogen bonds were more highly maintained. Thus, it may be concluded
366 that the GLYCAM06 force field is somewhat better at simulating cellulose
367 intermolecular hydrogen bonding than the other force fields.

368

369 **Conclusions**

370

371 Simulating crystal structures is not generally the most demanding test for force
372 fields since such simulations do not necessarily provide opportunities for major
373 transitions over energy barriers such as ring conformational changes or torsional
374 rotations. The advantage of such calculations, however, is that small molecule crystal
375 structures are generally known to high precision, and thus at least constitute a test of
376 minimal acceptability, while generally few other direct, highly accurate experimental
377 tests of force fields, such as solution conformation, are available.⁴¹ Such simulations are
378 more demanding when unit cell dimensions and angles are allowed to vary as in the
379 present simulations.

380

381 From the results presented here, it can be seen that none of the three force
382 fields tested is clearly superior to the others in reproducing all aspects of the crystal
383 structures of the selected set of small molecules. All appeared to model the selected
384 crystals acceptably in a qualitative sense, from which it might be hoped that they would
385 be useful for simulating cellulose itself. Of course, all these force fields use static charge
386 assignments and fixed torsional terms, and thus might not be able to capture any effects
387 that could arise from bond polarization and charge distribution perturbations that may
388 result from the multiple hydrogen bonds formed in cellulose crystals, for example. The
389 observation that the primary alcohol conformation distribution is sensitive to the chain
390 DP using the C35 force field is particularly interesting in this context. This suggests the
391 possibility that the collective field resulting from cellulose crystal packing can
392 significantly affect the orientational preference for this group. This is important since
393 the principal reason that the reported fiber structure of cellulose Ib is not stable using
394 static point charge force fields is due to the tendency of the primary alcohol groups to
395 rotate away from the *tg* conformation. If the fiber structure is correct, it may be because
396 the crystal packing creates an electrostatic environment favoring *tg* that cannot be, or at
397 least is not, adequately mimicked with the present fixed point charge models. In this

context, it would be most interesting to repeat the present simulations using the new generation of polarizable force fields when they become available.^{4,21}

Acknowledgments

The U.S. Department of Energy, Office of Biological and Environmental Research, and the Office of Advanced Scientific Computing Research supported this work under the SciDAC Program. H.M. thanks the JSPS Postdoctoral Fellowships for Research Abroad.

References

- (1) Alder, B. J.; Wainwright, T. E. *J. Chem. Phys.* **1957**, *27*, 1208-1209.
- (2) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95-99.
- (3) Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinoza, J.-F.; van Eyck, B. P.; Johnson, G., et al. *Carbohydrate Research* **1998**, *314*, 141-155.
- (4) Patel, D. S.; He, X. B.; MacKerell, A. D. *J. Phys. Chem. B* **2015**, *119*, 637-652.
- (5) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074-9082.
- (6) Marchessault, R. H.; Pérez, S. *Biopolymers* **1979**, *18*, 2369-2374.

- 422 (7) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallographica* **1968**, B24, 830-838.
- 423 (8) Mackie, I. D.; Röhring, J.; Gould, R. O.; Pauli, J.; Jäger, C.; Walkinshaw, M.;
424 Potthast, A.; Rosenau, T.; Kosma, P. *Carbohydrate Research* **2002**, 337,
425 161-166.
- 426 (9) Mackie, I. D.; Röhring, J.; Gould, R. O.; Pauli, J.; Jäger, C.; Walkinshaw, M.;
427 Potthast, A.; Rosenau, T.; Kosma, P. *Carbohydrate Research* **2002**, 337, 1065.
- 428 (10) Rencurosi, A.; Röhring, J.; Pauli, J.; Potthast, A.; Jäger, C.; Pérez, S.; Kosma,
429 P.; Imberty, A. *Angewandte Chemie-International Edition* **2002**, 41, 4277-4281.
- 430 (11) Ham, J. T.; Williams, D. G. *Acta Crystallographica B* **1970**, 26, 1373-1383.
- 431 (12) Raymond, S.; Heyraud, A.; Qui, D. T.; Kvik, A.; Chanzy, H. *Macromolecules*
432 **1995**, 28, 2096-2100.
- 433 (13) Raymond, S.; Henrissat, B.; Qui, D. T.; Kvik, A.; Chanzy, H. *Carbohydrate*
434 *Research* **1995**, 277, 209-229.
- 435 (14) Yoneda, Y.; Mereiter, K.; Jaeger, C.; Brecker, L.; Kosma, P.; Rosenau, T.;
436 French, A. *J. Am. Chem. Soc.* **2008**, 130, 16678-16690.
- 437 (15) Matthews, J. F.; Beckham, G. T.; Bergensträhle-Wohlert, M.; Brady, J. W.;
438 Himmel, M. E.; Crowley, M. *Journal of Chemical Theory and Computation*
439 **2012**, 8, 735-748.
- 440 (16) Matthews, J. F.; Bergensträhle, M.; Beckham, G. T.; Himmel, M. E.; Nimlos, M.
441 R.; Brady, J. W.; Crowley, M. F. *J. Phys. Chem. B* **2011**, 115, 2155-2166.
- 442 (17) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor,
443 R. W.; Mackerell, A. D. *Journal of Computational Chemistry* **2008**, 29,
444 2543-2564.
- 445 (18) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell, A. D.
446 *Journal of Chemical Theory and Computation* **2009**, 5, 2353-2370.
- 447 (19) Lins, R. D.; Hünenberger, P. H. *Journal of Computational Chemistry* **2005**, 26,
448 1400-1412.
- 449 (20) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzáles-Outeiriño, J.;
450 Daniels, C. R.; Foley, B. L.; Woods, R. J. *Journal of Computational Chemistry*
451 **2008**, 29, 622.
- 452 (21) He, X.; Lopes, P. E. M.; MacKerell, A. D. *Biopolymers* **2013**, 99, 724-738.
- 453 (22) Brooks, B. R.; Brucoleri, R. E.; Olafson, B. D.; Swaminathan, S.; Karplus, M.
454 *Journal of Computational Chemistry* **1983**, 4, 187-217.
- 455 (23) Brooks, B. R.; Brooks, C. L.; A.D. MacKerell, J.; Nilsson, L.; Petrella, R. J.;
456 Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S., et al. *Journal of*
457 *Computational Chemistry* **2009**, 30, 1545-1614.

- 458 (24) Palma, R.; Himmel, M. E.; Brady, J. W. *Journal of Physical Chemistry B* **2000**,
459 *104*, 7228-7234.
- 460 (25) Kuttel, M.; Brady, J. W.; Naidoo, K. J. *Journal of Computational Chemistry*
461 **2002**, *23*, 1236-1243.
- 462 (26) Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.;
463 Sugiyama, J.; Himmel, M. E.; Brady, J. W. *Carbohydrate Research* **2006**, *341*,
464 138-152.
- 465 (27) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L.
466 *J. Chem. Phys.* **1983**, *79*, 926-935.
- 467 (28) Durell, S. R.; Brooks, B. R.; Ben-Naim, A. *Journal of Physical Chemistry* **1994**,
468 *98*, 2198-2202.
- 469 (29) Cézard, C.; Rice, C. A.; Suhm, M. A. *J. Phys. Chem. A* **2006**, *110*, 9839-9848.
- 470 (30) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.;
471 Lelong, D.; Rosanski, W.; Cieplak, P. *Physical Chemistry Chemical Physics*
472 **2010**, *12*, 7821-7839.
- 473 (31) Caldwell, J. W.; Kollman, P. A. *The Journal of Physical Chemistry* **1995**, *99*,
474 6208-6219.
- 475 (32) van Gunsteren, W. F.; Berendsen, H. J. C. *Molecular Physics* **1977**, *34*,
476 1311-1327.
- 477 (33) Miyamoto, S.; Kollman, P. A. *Journal of Computational Chemistry* **1992**, *13*,
478 952-962.
- 479 (34) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089-10092.
- 480 (35) Hoover, W. G. *Physical Review A* **1985**, *31*, 1695-1697.
- 481 (36) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177-4189.
- 482 (37) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. *J. Chem. Phys.* **1995**, *103*,
483 4613-4621.
- 484 (38) Humphrey, W.; Dalke, A.; Schulten, K. *Journal of Molecular Graphics* **1996**,
485 *14*, 33-38.
- 486 (39) Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor,
487 R.; Towler, M.; van De Streek, J. *Journal of Applied Crystallography* **2006**, *39*,
488 453-457.
- 489 (40) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.;
490 Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van De Streek, J.; Wood, P. A.
491 *Journal of Applied Crystallography* **2008**, *41*, 466-470.
- 492 (41) Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Sabounji, M.-L.; Brady, J. W. *J.*
493 *Am. Chem. Soc.* **2005**, *127*, 10991-10998.