Molecular Dynamics Simulations of Solvated Crystal Models of Cellulose I_α and III_I

Toshifumi Yui* and Sachio Hayashi

Department of Applied Chemistry, Faculty of Engineering, University of Miyazaki, Nishi 1-1, Gakuen-kibanadai, Miyazaki 889-2192, Japan

Received September 12, 2006; Revised Manuscript Received December 5, 2006

Swelling behaviors of cellulose I_α and III_I crystals have been studied using molecular dynamics simulations of the solvated finite-crystal models. The typical crystal models consisted of 48 × 10-mer chains. For the cellulose I_α crystal, models consisting of different numbers of chains and chain lengths were also studied. The structural features of the swollen crystal models, including the cellulose I_β crystal model reported previously, were compared. A distinct right-handed twist was observed for models of the native cellulose crystals (cellulose I_α and I_β), with a greater amount of twisting observed for the I_α crystal model. Although the amount of twist decreased with increasing dimensions of the cellulose I_α crystal model, the relative changes in twist angle suggest that considerable twist would arise in a crystal model of the actual dimensions. In contrast to the swelling behavior of crystal models of the native cellulose, the cellulose III_I crystal model exhibited local, gradual disordering at the corner of the reducing end. Comparison of the lattice energies indicated that the cellulose chains of the I_β crystal were packed in the most stable fashion, whereas those of the I_α and III_I crystals were in a metastable state, which is consistent with the crystallization behaviors observed. Upon heating of the native cellulose crystal models, the chain sheets of the I_α model showed a continuous increase in twist angle, suggesting weaker intersheet interactions in this model. The swollen crystal models of cellulose I_α and III_I reproduce well the representative structural features observed in the corresponding crystal structures. The crystal model twist thus characterizes the swelling behavior of the native cellulose crystal models, which seems to be related to the insolubility of the crystals.

Introduction

Cellulose is a linear polysaccharide of β-(1 → 4)-linked D-glucan units and is the most abundant bioresource on Earth. Cellulose is biosynthesized in a highly crystalline form, consisting of crystalline microfibrils. Early X-ray diffraction studies suggested that cellulose chains conform to a typical 2-fold helical structure in the crystalline state.[1] The native cellulose crystal was originally defined as a pure crystalline phase and designated the now redundant name, cellulose I. ^{13}C cross-polarization magic-angle-spinning NMR[2,3] and electron diffraction[4,5] studies revealed that the native form consisted of two allomorphs, I_α and I_β. Whereas the chains in the I_α phase crystallize with a single-chain triclinic unit cell of $P1$ symmetry, the I_β crystal comprises two independent chains in a monoclinic unit of $P2_1$ symmetry.[5] The I_α crystal is known to be the least stable of the two allomorphs, and it can be converted to the I_β phase simply by heating.[6] Probably the most representative feature of the native cellulose crystals is that the chains are aligned along the fiber axis with the same direction, or parallel polarity. Since the parallel-chain structures correspond to a typical metastable form, the diverse crystalline allomorphs, including a series of cellulose II crystals of an antiparallel-chain structure, have been derived from native cellulose crystals via appropriate chemical and physical treatments. Cellulose III_I is also derived from cellulose I samples by treatment with liquid ammonia or a number of suitable amines.[7,8] Cellulose III_I is essentially a single-chain structure and therefore parallel-chain by default, based on the monoclinic unit cell and $P2_1$ space group.[9] Earlier molecular and crystal studies of cellulose I_β[10,11] (previously regarded as cellulose I) and cellulose III_I[12] were carried out using a combination of conventional X-ray diffraction and stereochemical model refinement techniques to reveal the atomic positions. While these crystal structure analyses successfully proposed the fundamental structural features, such as chain-packing arrangements and the orientation of hydroxymethyl groups, the positions of the hydroxyl hydrogen atoms required the definition of a precise hydrogen-bonding scheme (including the donor−acceptor direction) that remained unclear due to weak scattering of the X-rays and limited diffraction data, arising from the physical properties of the hydrogen atoms and the polycrystalline nature of the fiber samples, respectively. Recently, a series of landmark studies on cellulose crystal structures have been reported, in which highly oriented samples of the pure crystalline phase were subjected to synchrotron X-ray diffraction measurement, producing approximately 100−300 reflection spots.[13−18] Neutron diffraction data were also collected for deuterated samples to reveal reliable hydrogen atom positions. These results indicated that for the two native crystals, cellulose I_α[13] and I_β[14], the chains were interconnected by O−H···O-type hydrogen bonds to form flat chain sheets with weak C−H···O hydrogen bonds between them. As had been suggested by the results of the previous electron diffraction studies,[5] the chain-packing structures of the two native cellulose crystals differed in the relative displacement of sheets along the chain axis. As a result, it was considered that the two crystal structures shared a nearly identical appearance in the *ab* base-plane projection. However, in the cellulose III_I structure, its *ab* projections revealed that the pyranose planes were slightly staggered in relation to each other about the helix axis, which resulted in corrugated chain sheet formation along both the [0,1,0] and [1,−1,0] crystal planes, both involving interchain

* Author to whom correspondence should be addressed. Phone: +81-985-58-7319. Fax: +81-985-58-7323. E-mail: tyui@cc.miyazaki-u.ac.jp.

10.1021/bm060867a CCC: $37.00
Published on Web 02/08/2007

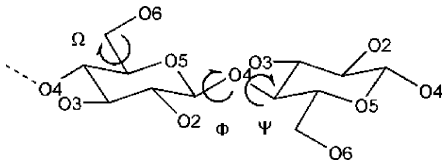

Figure 1. Torsion-angle parameters of cellulose. The torsion angles are defined by Φ = O5−C1−O4−C4, Ψ = C1−O4−C4−C5, and Ω = O5−C5−C6−O6.

hydrogen bonds.[18] It should be noted that the cellulose II structure[15,16] exhibited a similar *ab* projection to that of cellulose III_I,[18] although the former displayed one chain oriented in the opposite direction to the other. Another distinguishing difference between the native cellulose crystals and cellulose III_I concerns the orientation of the hydroxymethyl groups; while the native forms involve a rare tg conformation,[13,14] the cellulose III_I[18] and II[15,16] crystals, together with the related cellotetraose crystals,[19,20] all adopt a gt conformation.

Crystal modeling studies based on theoretical calculations proved useful in examining various carbohydrate crystals. French and co-workers modeled cellulose I[21] and cellulose III_I[22] crystals using a combination of MM3 force field calculations for construction of their small finite model crystals, designated as miniature crystals, or "minicrystals" for short. We adopted the minicrystal modeling method to supplement the polysaccharide crystal structures previously determined using the X-ray fiber diffraction method.[23−25] Recently, Matthews et al.[26] and our own group[27] independently reported crystal modeling studies on finite but fairly large cellulose crystal models, in which the solvated cellulose I_β systems were examined using molecular dynamics (MD) calculations. The most remarkable result observed in these calculations was an overall deformation of the solvated crystal models; these models were quickly and significantly twisted into the right-handed form as soon as the MD calculations were implemented, and these twisted structures persisted during the rest of the simulation time. Since the overall twist was able to decompose into a twisted shape for individual chain sheets, the chain sheets of cellulose I_β crystal models appeared to involve a particular internal stress. These calculations seemed to reproduce the morphology of real microfibrils showing similar right-handed twists,[28,29] prompting an evaluation of the mechanical state of the crystals of interest. The present paper reports on results obtained from similar MD calculations for solvated crystal models of cellulose I_α and III_I. Some of the previous results on the cellulose I_β crystal model[27] are also presented to compare swelling behaviors among the three parallel-chain crystals.

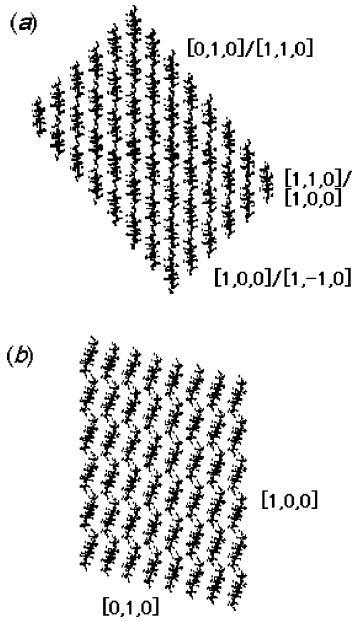

Figure 2. Projections of the *ab* base plane of the cellulose (a) I_α and I_β and (b) III_I crystal models. The crystal planes making up the base-plane edges are labeled with the Miller indices. The complex labeling in part a is presented as [index of I_α]/[index of I_β].

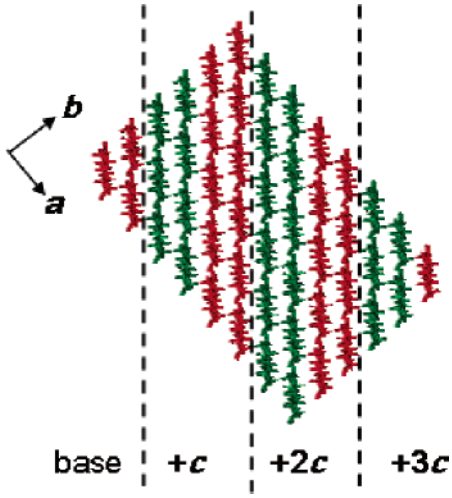

Figure 3. Relative chain staggering positions of a crystal model of cellulose I_α. Chain sheets in green advance to those in red along the *c*-axis by a half-helical pitch. Additional translational operations along the *c*-axis are also shown.

Computational Methods

Crystal Model Building. Figure 1 shows the torsion-angle parameters of primary interest in defining the cellulose conformation. The torsion angle, Ω, defines the orientation of the hydroxymethyl group of which representative conformers are designated as gg, gt, or tg with typical Ω angles of −60°, 60°, or 180°, respectively.

Figure 2 depicts the *ab* base-plane projections of the representative finite-crystal models examined in the present study. Here, 48 cellulose chains are aligned according to the published atomic coordinates and lattice parameters of each corresponding crystalline form.[13,14,18] The shapes of the base plane have been designed such that the hydrophilic faces are mostly exposed to the surrounding solvent environment, with the chain sheets running diagonally along the base plane. Constituent chains were primarily generated by combined translations of an independent chain from the origin along the *a*- and/or *b*-axes. While the cellulose I_β crystal model required two independent chains, i.e., the origin and center chains,[27] the cellulose III_I crystal model was generated using a single chain. Such a process of simple *ab* translations of an independent chain was not applicable to modeling of the cellulose I_α crystal because the resulting crystal model had wedge-shaped termini, reflecting the triclinic symmetry of the unit cell. The wedge-shaped termini, which are more exposed to solvent than the monoclinic- or orthorhombic-symmetry-based termini, emphasize the terminal effect in a finite-crystal model. Thus, additional operations were introduced for construction of the cellulose I_α model (Figure 3). Two types of cellulose chains were defined, which are crystallographically identical but differ in the *c*-axis by a half-helical pitch stagger. Each chain sheet consisted of either of the two chain types, and pairs of these sheets were placed alternately along the [1,1,0] direction. An additional operation of one unit translation along the fiber axis was performed for every other pair of sheets. These operations afforded crystal models with the corresponding base planes at nearly right angles to the fiber axis. The crystal models were then placed in a rectangular periodic box filled with TIP3P water models.[30] The buffer size between the edges

Table 1. Dimensions and Number of Molecules in the Crystal Model Systems

label	base-plane dimensions[a] no. of chains [lattice planes]	DP	no. of molecules residue	water
cellulose I_α				
A1-20	4[1,0,0] × 1[1,1,0] × 6[0,1,0]	20	480	34 448
A2-20	6[1,0,0] × 1[1,1,0] × 4[0,1,0]	20	480	33 781
B1-20	6[1,0,0] × 1[1,1,0] × 8[0,1,0]	20	960	43 748
B2-20	8[1,0,0] × 1[1,1,0] × 6[0,1,0]	20	960	43 658
C-10	6[1,0,0] × 1[1,1,0] × 6[0,1,0]	10	360	16 195
C-20	6[1,0,0] × 1[1,1,0] × 6[0,1,0]	20	720	39 199
C-30	6[1,0,0] × 1[1,1,0] × 6[0,1,0]	30	1080	77 545
cellulose I_β (**D**)[27]	6[1,−1,0] × 1[1,0,0] × 8[1,1,0]	20	960	24 196
cellulose III_I	6[1,0,0] × 8[0,1,0]	20	960	13 476

[a] The sides in the crystal model are given. The sides listed are the minimum set required to describe the base-plane dimensions of each crystal.

of the crystal model and the periodic box was set to 12 Å. Table 1 describes the components of the crystal models examined in the present study, including the dimensions and the number of molecules and the labels assigned to the models. Among the cellulose I_α crystal models, **A1-20**, **A2-20**, **B1-20**, and **B2-20** involve a rectangular chain arrangement on the base plane. The **A1-20** and **A2-20** models exchange the crystal phases assigned to the model sides with each other, as do the **B1-20** and **B2-20** models. A series of **C** models shares a square 6 × 6 chain arrangement, each differing in chain length. As shown in Figure 2, the **B2** model possesses similar projection of the base plane to that of the cellulose I_β crystal model that was designated **D** in our previous report.[27] The cellulose III_I model also consists of 48 × 20-mer chains and is similar to the cellulose I_α (**B**-type) and I_β models with respect to crystal model dimensions.

The deformation or twisting of crystal models during dynamics simulations are evaluated by the degree of deviation from the planarity of the central chain sheet passing diagonally toward the base plane. As shown in Figure 4, the virtual bonds connecting the gravity centers of the residues define the torsion angle, Θ, representing the twisting angle of the chain sheet along the fiber axis. Obviously, a proportion of Θ depends on the residue position along the fiber axis; the largest Θ value would be expected at either the terminal ends ($\Theta_{\pm 10}$ or $\Theta_{\pm 9}$), where the angles decrease to nearly 0° on approaching the middle position, or $\Theta_{\pm 1}$ if an overall uniform twist occurs. When a chain sheet twists in the clockwise direction to give a right-handed twist, the Θ value is defined as being positive.

MD Simulations. Similar procedures to those described in our previous report[27] were adopted to perform the minimization and dynamics calculations for the crystal models. Briefly, the water configurations were first optimized by constraining the atomic positions of the crystal models with a force of 500 kcal/(mol Å^2), followed by full optimization of the whole system. Constant volume dynamics was carried out for 100 ps while gradually increasing the temperature from 20 to 300 K. This was immediately followed by the 500 ps NTP dynamics with 1 bar and 300 K for equilibration. Throughout the initial 600 ps dynamics, the motions of the crystal models were fixed by imposing positional constraints on the heavy atoms with a constrained force of 10 kcal/(mol Å^2). The production runs for NTP dynamics were then implemented for 1 ns without any constraint. A 2 fs time step for integration was combined with the SHAKE option[31] to constrain the motions of hydrogen atoms, while the particle mesh Ewald method[32] was adopted for long-range, nonbonded interactions. Dynamics at a controlled heating rate were also performed for the **B1** crystal model, for which the procedure for the previous cellulose I_β model study[27] was again adopted. The heating run started at the 600 ps trajectory of the preceding constant-temperature run and was continued for a further 500 ps. The heating rate was adjusted to 0.25 K/ps, resulting in a final temperature of 425 K.

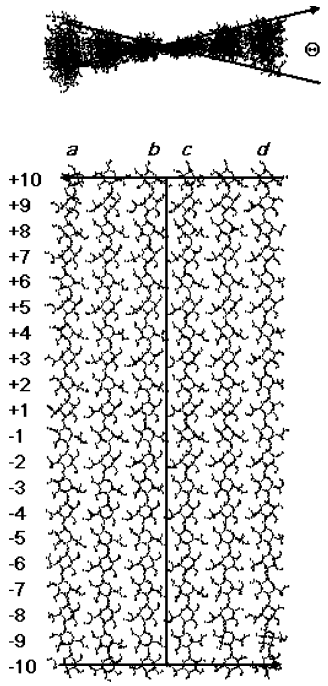

Figure 4. Scheme highlighting the twisting angle (Θ) in the central chain sheet. The Θ angle is defined by the torsion angle consisting of the virtual bonds connecting the centers of gravity of the residues (*G*), as follows: $\Theta_{+10} = G(a;+10) - G(b,c;+10) - G(b,c;-10) - G(d;-10)$, $\Theta_{+9} = G(a;+9) - G(b,c;+9) - G(b,c;-9) - G(d;-9)$, ...$\Theta_{-10} = G(a;-10) - G(b,c;-10) - G(b,c;+10) - G(d;+10)$.

The crystal lattice energy, $\Delta E_{\text{lattice}}$, in the gas phase was calculated based on the optimized geometries and the MD trajectories using the following definition

$$\Delta E_{\text{lattice}} = E_{\text{total}} - (E_{\text{inner}} + E_{\text{outer}}) \quad (1)$$

where E_{total}, E_{inner}, and E_{outer} represent the AMBER[33] steric energies of the whole crystal model, the central chain, and the remaining chains, respectively. A combination of the generalized Born approach[34,35] and the surface area[36] option was adopted to introduce a solvent effect with a default surface tension of 0.005 kcal/(mol Å^2), giving the free energy $\Delta G_{\text{lattice}}$ in the solution state

$$\Delta G_{\text{lattice}} = \Delta E_{\text{lattice}} - T\Delta S + \Delta G_{\text{GB/SA,total}} - (\Delta G_{\text{GB/SA,inner}} + \Delta G_{\text{GB/SA,outer}}) \quad (2)$$

where T is the temperature, ΔS is the entropy change upon formation of the whole model, and $\Delta G_{\text{GB/SA}}$ is the solvation free energy for each component. Assuming a constant ΔS contribution for all time steps, we evaluate $\Delta G^*_{\text{lattice}} = \Delta G_{\text{lattice}} + T\Delta S$ to assess the lattice free energy of a crystal model in the solution state.

The present MD calculations were performed with the SANDER module of the AMBER 8 package[33] using the GLYCAM 04 carbohydrate parameters.[37−39] VMD 1.8.3 software[40] was used for molecular visualization and animation of the trajectory data.

Results and Discussion

Crystal Model Twist and Thermodynamics Features As the constraints were released in the MD run, the cellulose I_α crystal models immediately developed an overall twist within the initial 20 ps, and these twisted structures persisted throughout the rest of the simulation time. This was similar behavior to that observed for the cellulose I_β model studies.[26,27] Figure 5 compares the root-mean-square deviations (rmsd's) for atomic

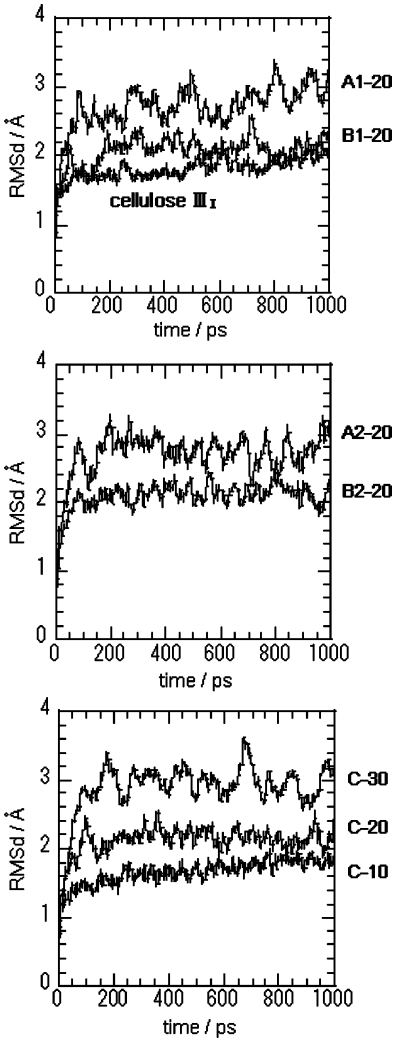

Figure 5. Root-mean-square deviation trajectories of the atomic coordinates calculated for the crystal models.

coordinates calculated from the trajectories of the cellulose I_α and III_I crystal models. While most of the crystal models appeared to reach equilibrium values within ~100−300 ps, the trajectories of the **C-10** model—the smallest crystal model—and the cellulose III_I model continued to increase gradually. As had been already observed in the previous cellulose I_β calculations,[27] crystal models with a larger base plane gave smaller rmsd values at equilibrium. The **A**-type models, therefore, exhibited larger rmsd values than those of the **B**-type models. When comparing models with the same base-plane dimensions, i.e., **A1-20** and **A2-20**, or **B1-20** and **B2-20**, the rmsd trajectories were similar to each other. When comparing the trajectories of the **C**-type models, those consisting of longer chains resulted in larger rmsd values along with larger variations. While the average rmsd value for the cellulose I_β model was reported to be approximately 1.6 Å,[27] its cellulose I_α counterpart, the **B1-20** model, showed increasing deformation with swelling, giving an average rmsd of approximately 2 Å.

Figures 6 and 7 show variations in the sheet twisting angle, Θ, calculated from the final 500 ps trajectories with respect to the residue positions. All of the Θ values are averages of those observed between the positive and negative residue positions. It should be noted that positive positions tended to give larger Θ values than negative positions and that the differences became more obvious for chain sheets with smaller degrees of polymerization (DP) and as the positions reached the termini. Figure

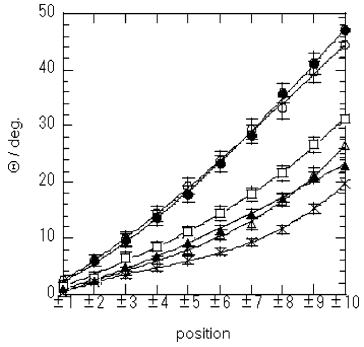

Figure 6. Variations in sheet twist angle, Θ, of the central chain sheet with respect to the chain position for cellulose I_α crystal models (○) **A1-20**, (●) **A2-20**, (□) **C-20**, (△) **B1-20**, and (▲) **B2-20** and (×) the I_β[27] crystal model. The Θ values are averages for the positive and negative positions (see Figure 4 for definition) and the final 500 ps MD trajectories. Error bars indicate the rmsd values calculated for each trajectory.

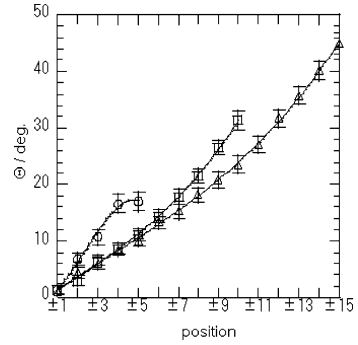

Figure 7. Variations in sheet twist angle, Θ, of the central chain sheet with respect to the chain position for cellulose I_α crystal models (○) **C-10**, (□) **C-20**, and (△) **C-30**.

6 compares Θ plots for cellulose I_α crystal models with various base-plane sizes, along with the previous results for the cellulose I_β model.[27] The Θ curves increase more steeply for models consisting of fewer chains, suggesting that lateral expansion of the base-plane dimensions, namely, extension of the central chain sheet, suppressed twist in the crystal model. This tendency was also observed for the twisting behavior of cellulose I_β crystal models of various sizes.[27] The terminal Θ value at the ±10 position of the 36-chain **C-20** model was 31.3°, which is closer to those for the 48-chain **B**-type models (26.4° for **B1-20** and 22.9° for **B2-20**) than those for the 24-chain **A**-type models (44.7° for **A1-20** and 47.1° for **A2-20**). This implies that a comparable amount of twist would occur even in a larger crystal model. However, the similarity of Θ variations between the two **A**-type models indicates that the length of the central chain sheet is a primary factor that determines the Θ value. The same results were also observed for the twisting behavior of the two **B**-type models. Comparison of Θ curves for the **B1-20** model of cellulose I_α and the cellulose I_β model, both of which essentially shared the same base-plane structure, shows that Θ for the **B1-20** model increases more steeply than that for the cellulose I_β crystal model, resulting in a smaller terminal angle of $\Theta_{\pm 10} =$ 19.5° for the latter. As discussed in our previous report,[27] these twisting motions of the chain sheets occurred spontaneously, as if the crystal models were located in relatively shallow energy-minimum wells. In comparison with the twisting angles for the cellulose I_β model, the I_α model may involve either greater internal stress or weaker intersheet interactions between the chain sheets. The latter seems to be more likely, considering the similarity of the chain sheet structure of the two native

crystals. In fact, on the basis of inspection of the hydrogen bond geometries, Nishiyama et al. pointed out that more enhanced C−H···O interactions existed between the neighboring sheets of cellulose I_β compared to cellulose I_α.[13]

Figure 7 describes chain-length dependence of crystal model twist observed for the **C**-type models. Elongation of cellulose chains suppressed the amount of twist. Only the Θ curve for the **C**-**10** model displays damping behavior at the terminal values, which was reflected by extra disorder or nonuniform deformation occurring at the edges of the chain sheets. As a rough approximation, an increment of Θ per 2-fold helical pitch was estimated to be 2.99°/pitch based on the $\Theta_{\pm 15}$ value for the **C**-**30** model of 44.9°. (Note that advancement by one chain position in the plots implies moving along the chain in both the positive and negative directions, i.e., one helical pitch. See Figure 4.) Similar estimations gave 3.13°/pitch for the **C**-**20** model and 4.15°/pitch for the **C**-**10** model, where the latter value refers to the $\Theta_{\pm 4}$ value of 16.6° instead of the $\Theta_{\pm 5}$ value. These Θ plots suggest that the slope is likely to converge to a certain value for a crystal model with infinite fiber length. Needless to say, the present models were based on a very limited length from the viewpoint of a real microfibril, since it is clearly unrealistic to handle a model with an actual fiber length using currently available computer resources. Although a periodic box system with respect to the fiber axis can define an infinite microcrystal model, it is still unrealistic to construct a periodic box with the side lengths of approximately 1 μm to allow the crystal model twist. Therefore, it is encouraging that the results in Figure 7 seem to confirm the suitability of the present strategy based on crystal models of finite length.

In contrast to the swelling behaviors of the native cellulose crystal models, no distinct overall deformation was observed in the cellulose III_I crystal model in the swollen state. This was not unexpected in light of the relatively isotropic nature of the interchain hydrogen bond network. The crystal model developed local disorder such that the [1,0,0] side was peeled off from a corner of the reducing end. Meanwhile, the central chain sheet maintained its original flat shape, except at the disordered corner. Visual inspection of the MD trajectory animation revealed that this disordered region developed very slowly as the simulation proceeded. A much longer simulation, therefore, would finally result in a crystal model in a noncrystalline state. This should be a fairly reasonable event because, as evident from Figure 2, the chains making up the [1,0,0] lattice plane were more exposed to surrounding water molecules than chains making up the [0,1,0] plane and than any plane of the native cellulose crystal models. Matthews et al. studied a cellulose I_β crystal model that involved the [1,0,0] and [2,0,0] lattice planes as a primary surface, designated the rectangular model.[26] Although the chains on these planes were even more exposed to solvent than the present cellulose III_I chains on the [1,0,0] plane, the rectangular model developed a twisted structure similar to those of the other native cellulose crystal models. This again emphasizes the distinct intermolecular interactions involved in forming the chain sheets in native cellulose crystals, which may be enhanced in an aqueous evironenment.

Table 2 lists the GB/SA lattice energies, $\Delta G^*_{\text{lattice}}$, for the minimized crystal models and those calculated from the final 500 ps MD trajectories. The cellulose I_α crystal model was represented by the **B1**-**20** model, which corresponds to the 48 × 20-mer chain system and is similar to the two other crystal models. Assuming that the steric energy of the chain is relatively constant, the cellulose I_β crystal is expected to exist in a more stable state than the cellulose I_α crystal, whereas the origin chain

Table 2. GB/SA Lattice Energies (kcal/mol per residue) of the Minimized Crystal Models and Average Energies Calculated from the MD Trajectories

crystal and chain types		minimized structure	MD trajectory[a]
I_α (**B1**-**20**)		−31.3	-27.3 ± 2.0[b]
I_β[c]	origin	−35.1	-29.8 ± 2.2
	center	−34.3	-29.7 ± 2.3
III_I		−31.3	-27.6 ± 2.1

[a] Average values from the final 500 ps trajectories. [b] Root-mean-square deviation value. [c] These values were derived from the trajectories of the previous calculations.[27]

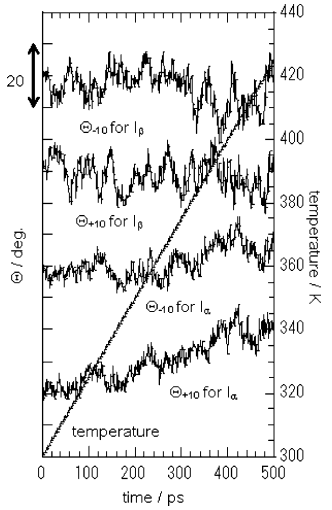

Figure 8. Terminal Θ trajectories at the +10 and −10 positions obtained from heating dynamics for the cellulose I_α (**B1**-**20**) and I_β models. The latter values were calculated from the heating trajectories of the previous study.[27]

may reside within a slightly more stable environment in the cellulose I_β crystal than the center chain. In addition, the lattice energies calculated for the cellulose III_I and I_β crystal models were very similar. Although the average energies derived from the MD trajectories of the three crystal models increased with respect to those for the minimized structures, the same order of increase as for the minimized energies was observed. Such variations in the lattice energy appear to be in reasonable agreement with the crystallizing relationship of the corresponding cellulose allomorphs; irreversible transformation from the I_α to I_β form could be explained by an exothermic process, as indicated by lowering of the lattice energy. Previous modeling studies also support the present results.[21,22] In addition, an a priori prediction of the native cellulose crystals proposed two candidate structures, i.e., one-chain triclinic and two-chain monoclinic crystals, with the former proposed as being less stable.[41] It is known that the I_β form is reversibly transformed to cellulose III_I through a noncrystalline state. Wada et al. suggested that the cellulose III_I crystal is metastable based on comparison of the calculated crystal densities (1.57 for cellulose III_I, 1.61 for cellulose I_α, and 1.63 for cellulose I_β) and inspection of their detailed chain-packing features.[18] Although the crystal density indicated a less effective packing of chains for the cellulose III_I crystal, more extensive interchain hydrogen bonds should contribute to three-dimensional stabilization of the crystal, resulting in a comparable lattice energy to that of cellulose I_α.

Thermodynamics features of the two native cellulose crystals were also examined by heating the crystal models. Figure 8 compares variations of the terminal Θ values for the two native cellulose crystal models upon heating. The Θ values for the cellulose I_β model were derived from the previous MD calcula-

(*a*)

167(16)	41(46)	60(13)	159(11)	167(14)	60(11)	59(14)	55(12)
165(13)	164(13)	166(17)	164(15)	165(13)	164(13)	170(28)	174(48)
22(50)	165(13)	163(13)	164(12)	164(9)	164(12)	164(13)	168(23)
46(25)	163(11)	164(12)	164(14)	164(10)	163(9)	164(13)	56(12)
78(11)	163(9)	163(16)	162(13)	163(15)	164(10)	164(19)	57(12)
-33(47)	56(13)	169(18)	165(11)	-47(8)	58(22)	171(13)	-87(53)

(*b*)

56(10)	52(8)	50(9)	52(11)	51(8)	50(9)	51(9)	42(10)
55(10)	50(8)	50(8)	50(9)	49(8)	51(9)	49(8)	44(9)
55(10)	50(9)	50(8)	50(8)	50(9)	50(8)	48(9)	45(9)
56(9)	50(8)	50(8)	50(9)	49(9)	50(8)	50(8)	44(10)
55(10)	50(9)	50(9)	50(9)	50(8)	49(8)	49(8)	44(9)
32(46)	19(52)	55(12)	-12(53)	-25(47)	-23(52)	53(11)	55(16)

b ↑ → *a*

Figure 9. Average Ω (rmsd) values for the base-plane residues of the +1 position layer of cellulose (a) I_α (**B1**-**20**) and (b) III_I models. The values (given in degrees) were calculated from the final 500 ps trajectories for the respective crystal models.

tions.[27] Crystal twist gradually developed in the cellulose I_α model (**B1**-**20**) as the temperature increased from 300 to 400 K. The Θ values for cellulose I_β showed fluctuations of considerable amplitude around constant values. This difference in Θ variation may be interpreted as the difference in the relative contributions of intersheet interactions between the two native crystals. The intersheet interactions involved in the cellulose I_α crystal, which are initially weaker, are more quickly dissipated as a result of increases in thermal motion of the sheets, allowing intrasheet stress to predominate at elevated temperatures. This is not the case for cellulose I_β crystals, in which intersheet interactions continue to make a significant contribution throughout the heating. It should be noted that the force field parameters have generally been developed to reproduce molecular behavior at room temperature and that the AMBER force field functions, similar to other molecular mechanics functions, involve no temperature term. The results presented for the intermolecular interactions from heating dynamics are thus of a more qualitative nature.

Structural Features of the Swollen Crystal Models The present section compares structural features between the cellulose I_α (**B1**-**20**) and III_I models in the swollen state as represented by the average parameters calculated from the final 500 ps trajectories. Figure 9 depicts the average and rmsd values for the hydroxymethyl conformations, Ω, of the 48 residues making up the base plane at the +1 position. None of the hydroxymethyl orientations deviated significantly from the original staggered positions of tg ($\Omega \approx 180°$) for cellulose I_α and gt ($\Omega \approx 60°$) for cellulose III_I, except those present on the surface of the model. The tg conformations, featuring the native cellulose crystal structures, also prevailed in the swollen crystal models of cellulose I_β.[27] In the cellulose I_α model, some of the surface hydroxymethyl groups were rotated into approximate gt conformations, and the five Ω values involved additional conformational transitions, as indicated by the relatively large rmsd values. Since the gt conformation is most favored for glucopyranose residues,[42−45] the cellulose III_I crystal model apparently exhibited crystal conformations for most of the surface residues as well as the interior residues. Conformation changes took place at the five residues located on the [0,1,0] crystal plane. Variations in Ω for the cellulose I_α crystal model displayed slightly larger rmsd values than those for the cellulose III_I model, which is possibly correlated to twisting motions occurring in the former. These rotational patterns of the hydroxymethyl groups occurred in the remaining base-plane layers. Similarly, the hydroxyl groups located inside the models were mostly oriented into crystal conformations. As a consequence of retaining these crystal conformations, the substitutent groups participated in the original hydrogen bond networks. Table 3 well describes the reproduction of the hydrogen bond geometries in the swollen state. Here, we have defined "the core unit" of the crystal models as a representative part of the crystal structure, which involves tetrasaccharide units, consisting of the +2 to −2 position residues, within either the c (I_α model) or b (III_I model) chain belonging to the central chain sheet (Figure 4). The core unit comprised a few neighboring chains if necessary. For both crystal models, all of the important hydrogen bonds, with original length ranging from 2.5 to 2.9 Å, were reproduced as a stable interaction with nearly 100% occupation time. Weak hydrogen bonds of >3.0 Å in length were also detected with an occupation time of 30−60%. Nishiyama et al. proposed the second type of hydrogen bond pattern based on the minor deuterium positions O2−D and O6−D in their cellulose I_α crystal structure.[13] Our present MD calculations,

Table 3. Average Length and Occupation Time for Hydrogen Bonds in the Core Unit of the Crystal Models[a]

bond type[b]		observed[13,18]	calculated[c]	
atom	atom	length (Å)	length (rmsd) (Å)	occupation time (%)
cellulose I_α (**B1**-**20**) model				
Intrachain Bonds				
O6	H−O2	2.480	2.863 (0.15)	100.00
O6	H−O2	2.465	2.773 (0.11)	100.00
O6	H−O2	2.480	2.858 (0.15)	99.40
O3−H	O5	2.918	2.783 (0.10)	99.20
O3−H	O5	2.868	2.759 (0.10)	100.0
O3−H	O5	2.918	2.786 (0.10)	99.60
Interchain Bonds along the [1,1,0] Plane				
O3	H−O6	2.821	2.812 (0.11)	99.60
O6−H	O3	2.770	2.817 (0.11)	98.40
O3	H−O6	2.821	2.807 (0.11)	99.60
O6−H	O3	2.770	2.824 (0.11)	98.40
O2	H−O6	3.606	3.308 (0.15)	33.00
O6−H	O2	3.641	3.288 (0.14)	53.80
O2	H−O6	3.606	3.320 (0.15)	32.20
O6−H	O2	3.641	3.288 (0.14)	58.20
cellulose III_I model				
Intrachain Bonds				
O3−H	O5	2.866	2.911 (0.12)	99.00
O5	H−O3	2.866	2.909 (0.12)	99.00
O3−H	O5	2.866	2.914 (0.12)	97.60
O3−H	O6	3.021	3.288 (0.13)	65.20
O6	H−O3	3.021	3.297 (0.12)	59.60
O3−H	O6	3.021	3.295 (0.13)	61.00
Interchain along the [1,−1,0] Plane				
O6	H−O2	2.619	2.698 (0.10)	100.00
O2−H	O6	2.619	2.715 (0.10)	100.00
O6	H−O2	2.619	2.706 (0.10)	100.00
O2−H	O6	2.619	2.710 (0.10)	100.00
Interchain along the [0,1,0] Plane				
O6−H	O2	2.642	2.747 (0.12)	98.80
O2	H−O6	2.642	2.743 (0.11)	98.60
O6−H	O2	2.642	2.756 (0.12)	99.80
O2	H−O6	2.642	2.755 (0.12)	99.60

[a] Hydrogen bonds with significant occupation times (>10%) are listed. [b] Atoms are listed in order of position (+2, +1, −1, and −2) of the residues to which the atoms belong. For the atom pairs of interchain bonds, the atoms of the center chain are given in the left-hand column. [c] Values are averages for the final 500 ps trajectories.

however, only predicted the major schemes. In other words, crystal twist may have been driven by further stabilizing the major hydrogen-bonding scheme. It should be noted that the lengths predicted for all bonds, including those having nearly 100% occupation time, in the cellulose III_I model were longer than those observed. When the unit cell parameters of the core unit region were averaged using the final 500 ps trajectories, it was found that the γ angle decreased to 95° from the observed value of 105.1°,[18] accompanied by slight shortening along the *b*-axis from 7.85 to 7.4 Å. This probably resulted from chain sheet sliding along the [1,1,0] crystal plane, accompanied by very slight stretching of the hydrogen bonds. However, the average unit cell parameters of the cellulose I_α crystal model are fairly consistent with the values observed, indicating that the local chain arrangement was not significantly affected by crystal twist in the model. When inspecting the glycosidic linkage conformations, $\Phi-\Psi$, in the core unit, the largest deviation was detected for Ψ angles of the cellulose III_I chain, which slightly rotated to between −101° and −102° from the observed value of −92°.[18] The remaining $\Phi-\Psi$ angles only deviated by 2−5° from those of the crystal structures and the rmsd values were approximately 6° for all the angles. Clearly, the original 2-fold helical structures were conserved in the twisted models.

Conclusions

The present study assessed the swelling behaviors of crystal models of cellulose I_α and III_I allomorphs and compared these with results obtained from a similar crystal model study of cellulose I_β, in particular in terms of the crystal model twist and lattice energy. Both of the native cellulose crystals, I_α and I_β, are characterized by the presence of considerable stress within the chain sheets, as inferred from spontaneous crystal twist. The well-known metastable character of the cellulose I_α crystal is described by a larger amount of twist and lower lattice energy. The former feature is likely to arise from weaker intersheet interactions, which were supported by results for heating MD calculations. Whereas the lattice energy of the cellulose III_I crystal model was similar to that of the cellulose I_α model, their swelling behaviors were essentially different. In the swollen state, instead of showing distinct crystal twist, the cellulose III_I model displayed a slight rearrangement of the chain packing, so the γ angle decreased by approximately 10°, accompanied by appreciable stretching of interchain hydrogen bonds. In contrast to this isotropic deformation, the crystal twist observed in the native cellulose crystal models may reflect their development of resistance to dissolution.

Although the present study examined several types of cellulose I_α crystal models differing in base-plane dimensions, the appropriate size and shape of the base plane remains an issue. The diameter of cellulose microfibrils ranges from 2 to 20 nm, depending on the origins. Cellulose microfibrils of the cell wall of the freshwater alga *Glaucocystis* consist of almost pure I_α phase,[46] for which the diameter was reported to be approximately 10 nm[46] or no more than 20 nm.[13] Electron diffraction data suggested that the microfibrils comprised [1,0,0] and [0,1,0] lattice planes.[46] All of the present cellulose I_α crystal models were constructed with the same lattice planes according to these *Glaucocystis* microfibrils. As for the base-plane dimensions, the **B**-type models, for example, measure 5 nm in diameter, which is no more than half of the size of the *Glaucocystis* microfibrils. As discussed in the preceding section, the amount of twist depends on the base-plane size. The size effect should be taken into account in discussions. From the plots shown in Figure 7, the Θ value per 2-fold helical pitch was estimated to be 2.64°/pitch for the **B1-20** model and 1.97°/pitch for the cellulose I_β model. Elongation of these models using the Θ/pitch values and the helical pitch length of cellulose chains generates microfibrils with complete twist that are 142 nm long for the I_α (**B1-20**) model and 190 nm long for the I_β model. The complete twist length required for the **C**-type model is estimated to be approximately 125 nm based on $\Theta = 2.99°$/pitch. Since the twisting pitch observed for real cellulose microfibrils was reported to be 600 nm to 1 μm for bacterial cellulose[29] and 700 nm for algal cellulose,[28] the present values of the crystal twist angles are clearly overestimated. Considering the size dependence of the twist angle, therefore, additional expansion of the base-plane size for the crystal models and an increase in chain length may improve the twisting pitch to values comparable to those observed. In fact, algal microfibrils giving a twisted shape exhibited base-plane dimensions of 5 × 30 nm.[28]

In addition to the size factor discussed above, it seems worthwhile discussing other possible interpretations for the excessive twist of the crystal models. It should be noted that the variations in Θ shown in Figures 6 and 7 appear to behave nonlinearly; for most of the plots, the slope increases on approaching the terminal positions. The present estimation of Θ/pitch based on the terminal Θ values may not have provided an appropriate approximation. A theoretical treatment based on an appropriate mechanical model would allow us to reasonably extrapolate Θ/pitch to an infinite model. As a tentative evaluation, however, Θ/pitch using the $\Theta_{\pm1} - \Theta_{\pm5}$ values for the **C-30** plot only results in a trivial improvement, giving 2.25°/pitch. Another very undesirable possible cause of the excessive twist could be force field inaccuracies. Although the GLYCAM parameters adopted in the present study have been carefully developed and tested against experimental data, application to a condensed crystalline phase may be outside the scope of this approach. In this respect, it should be pointed out that Matthews et al. proposed values of 2.8−3.2° corresponding to the present Θ/pitch for a cellulose I_β crystal model with dimensions of 36 × 14-mer.[26] Their MD calculations were carried out using the CHARMM force field, and the resulting crystal model twist was in fairly good agreement with our predictions. This consistency among different force fields confirms the present results. Finally, it is known that cellulose microfibrils consist of both I_α and I_β phases in variable proportions, depending on the source. Such microfibrils with composite phases should exhibit different mechanical properties from our present models of pure crystalline phases, such as suppression of the twist angle. This is a very interesting topic, and I_α/I_β crystal modeling will be our next challenge.

Acknowledgment. This work was supported by Strategic Research Fund 2006 from the University of Miyazaki.

References and Notes

(1) Meyer, K. H.; Misch, L. *Helv. Chim. Acta* **1937**, *20*, 232−244.
(2) Atala, R.; VanderHart, D. L. *Science* **1984**, *223*, 282−285.
(3) VanderHart, D. L.; Attala, R. *Macromolecules* **1984**, *17*, 1465−1472.
(4) Sugiyama, J.; Okano, T.; Yamamoto, H.; Horii, F. *Macromolecules* **1990**, *23*, 3196−3198.
(5) Sugiyama, J.; Vuong, R.; Chanzy, H. *Macromolecules* **1991**, *24*, 4168−4175.
(6) Horii, F.; Yamamoto, H.; Kitamaru, R.; Tanahashi, M.; Higuchi, T. *Macromolecules* **1987**, *20*, 2946−2949.
(7) Hess, K.; Trogus, C. *Ber. Dtsch. Chem. Ges. B* **1935**, *68*, 1986−1988.
(8) Barry, A. J.; Petterson, F. C.; King, A. J. *J. Am. Chem. Soc.* **1936**, *58*, 333−337.

(9) Wada, M.; Heux, I.; Isogai, A.; Nishiyama, Y.; Chanzy, H.; Sugiyama, J. *Macromolecules* **2001**, *34*, 1237−1243.
(10) Sarko, A.; Muggli, R. *Macromolecules* **1974**, *7*, 486−494.
(11) Gardner, K. H.; Blackewll, J. *Biopolymers* **1974**, *13*, 1975−2001.
(12) Sarko, A.; Southwick, J.; Hayashi, J. *Macromolecules* **1976**, *9*, 857−863.
(13) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300−14306.
(14) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074−9082.
(15) Langan, P.; Nishiyama, Y.; Chanzy, H. *J. Am. Chem. Soc.* **1999**, *121*, 9940−9946.
(16) Langan, P.; Nishiyama, Y.; Chanzy, H. *Biomacromolecules* **2001**, *2*, 410−416.
(17) Langan, P.; Sukumar, N.; Nishiyama, Y.; Chanzy, H. *Cellulose* **2005**, *12*, 551−562.
(18) Wada, M.; Chanzy, H.; Nishiyama, Y.; Langan, P. *Macromolecules* **2004**, *37*, 8548−8555.
(19) Raymond, S.; Heyraud, A.; Tran, Q. D.; Kvick, Å.; Chanzy, H. *Macromolecules* **1995**, *28*, 2096−2100.
(20) Gessler, K.; Krauss, N.; Steiner, T.; Betzel, C.; Sarko, A.; Saenger, W. *J. Am. Chem. Soc.* **1995**, *117*, 11397−11406.
(21) French, A. D.; Miller, D. P.; Aabloo, A. *Int. J. Biol. Macromol.* **1993**, *15*, 30−36.
(22) Ford, Z. M.; Stevens, E. D.; Johnson, G. P.; French, A. D. *Carbohydr. Res.* **2005**, *340*, 827−833.
(23) Yui, T.; Ogasawara, T.; Ogawa, K. *Macromolecules* **1995**, *28*, 7957−7958.
(24) Yui, T.; Miyawaki, K.; Yada, M.; Ogawa, K. *Int. J. Biol. Macromol.* **1997**, *21*, 243−250.
(25) Yui, T.; Taki, N.; Sugiyama, J.; Hayashi, S. *Int. J. Biol. Macromol.* [Online early access] DOI: 10.1016/j.ijbiomac.2006.08.017. Published Online: Sept 28, 2006.
(26) Matthews, J. F.; Skopec, C. E.; Mason, P. E.; Zuccato, P.; Torget, R. W.; Sugiyama, J.; Himmel, M.; Brady, J. W. *Carbohydr. Res.* **2006**, *341*, 138−152.
(27) Yui, T.; Nishimura, S.; Akiba, S.; Hayashi, S. *Carbohydr. Res.* **2006**, *341*, 2521−2530.
(28) Hanley, S. J.; Revol, J.-F.; Godbout, L.; Gray, D. G. *Cellulose* **1997**, *4*, 209−220.
(29) Hirai, A.; Tsuji, M.; Horii, F. *Sen-i Gakkaishi* **1998**, *54*, 506−510.
(30) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Phys. Chem.* **1983**, *79*, 926−935.
(31) Ryckaert, J. P.; Cicotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(32) Essmann, U.; Pereta, L.; Berkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *14*, 33−38.
(33) Case, D. A.; Darden, T. A.; Cheatham, T. E., III.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Homak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. *AMBER 8*; University of California: San Francisco, 2004.
(34) Hawkins, G. D.; Cramer, C. J.; Truhlar. D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.
(35) Hawkins, G. D.; Cramer, C. J.; Truhlar. D. G. *J. Phys. Chem.* **1996**, *100*, 19824−19839.
(36) Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem.* **1999**, *20*, 217−230.
(37) Kirchner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541−10545.
(38) Basma, M.; Sundara, S.; Calgan, D.; Venali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125−1137.
(39) Kirchner, K. N.; Woods, R. J. *J. Phys. Chem. A* **2001**, *105*, 4150−4155.
(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(41) Viëtor, R. J.; Mazeau, K.; Lakin, M.; Pérez, S. *Biopolymers* **2000**, *54*, 342−354.
(42) Bock, K.; Duus, J. Ø. *J. Carbohydr. Chem.* **1994**, *13*, 513−543.
(43) Nishida, Y.; Ohrui, H.; Meguro, H. *Tetrahedron Lett.* **1984**, *24*, 1575−1578.
(44) Nishida, Y.; Ohrui, H.; Meguro, H. *J. Carbohydr. Chem.* **1988**, *7*, 239.
(45) Marchessault, R. H.; Pérez, S. *Biopolymers* **1979**, *28*, 2369−2374.
(46) Imai, T.; Sugiyama, J.; Itoh, T.; Horii, F. *J. Struct. Biol.* **1999**, *127*, 248−257.

BM060867A