Conformation of Amylose in Aqueous Solution: Small-Angle X-ray Scattering Measurements and Simulations

Jiro Shimada,*,† Hiroki Kaneko,† Toshikazu Takada,† Shinichi Kitamura,‡ and Kanji Kajiwara§

Fundamental Research Laboratories, NEC Corporation, 34 Miyukigaoka, Tsukuba, Ibaraki 305-8501, Japan, Department of Biological Resource Chemistry, Kyoto Prefectural University, Shimogamo, Kyoto 606-8522, Japan, Faculty of Engineering and Design, Kyoto Institute of Technology, Matsugasaki, Kyoto 606-8585, Japan

Received: July 28, 1999; In Final Form: December 20, 1999

Small-angle X-ray scattering profiles for an amylose fragment (maltoheptaose) in aqueous solution were observed and compared with the theoretical profiles calculated for an ensemble of chain conformations generated by molecular dynamics simulations and Monte Carlo simulations. The Monte Carlo results based on the disaccharide conformation energy map obtained without explicitly considering surrounding water molecules were in satisfactory agreement with the experimental results, provided that the effective dielectric constant was set to four. In contrast, the results of the fully solvated molecular dynamics simulations performed using the Cff91, Cff, Gromos, Glycam93, and Glycam99 force-fields were unexpectedly different from each other. Among them, Cff91 gave most satisfactory agreement with experiment, but the other fields yielded conformations that were somewhat or highly extended. It was also shown that recently developed Glycam99 is a significant improvement over Glycam 93. The representative snapshots of the two successful simulations resembled the regular helical structure reported by Goldsmith et al. (*J. Mol. Biol.* **1982**, *156*, 411). The source of the large force-field dependence was investigated by examining the various Ramachandran-like plots for the glycosidic torsion angles. For comparison, similar plots of ab initio energy for maltose (i.e., a fragment with two glucose units) were also calculated at the Hartree−Fock level, although in a simplified manner. These plots suggest that the extended conformation arises from too strong a preference for a certain rotational isomeric state of the glycosidic linkage. A procedure to remedy this over-preference can be devised, although a need of further elaboration of the force-field is indicated. The significance of force-fields is also illustrated in modeling a cyclodextrin composed of 14 glucose units.

I. Introduction

Molecular dynamics (MD) simulations based on force-field models are gaining wide popularity in structural studies of various biomolecules.[1−4] In the studies of amylose chains composed of α(1→4)-linked glucose units, most MD simulations have dealt with their smallest fragment (maltose) composed of only two glucose units,[5−9] or with circular chains including small α- and β-cyclodextrins[10−12] (see ref 13 for a review) and much larger cyclodextrins[14,15] discovered recently by Takaha et al.[16] The single glucose unit and its derivatives have also been the subject of many theoretical works that have paid particular attention to local phenomena such as the equilibrium between α- and β-anomers.[6,17−19] There were, however, only a few MD studies[20] of somewhat longer linear chains, for which helical behavior becomes progressively more pronounced. Although these longer chains can be efficiently mimicked by using a Monte Carlo technique based on the disaccharide conformation energy map, without considering explicit water molecules,[21−23] when we want to investigate the solvation of these chains and their interaction with proteins through water-mediated hydrogen bonds, we need to use models taking explicit water molecules into account.[24] Before investigating such complex problems, we should find out whether the conformation of somewhat longer amylose chains can be reasonably described by MD simulations performed with currently available force-fields.

Empirical force-fields developed for peptides and proteins have been tested in various ways,[25,26] and Pérez et al.[27] have recently made a systematic analysis of 20 types of force-fields by comparing the energies and structures of seven different carbohydrate molecules isolated in a vacuum or in dielectric continuum. Such tests for amylose chains in an aqueous environment, however, have not been reported, and the reasons for this are both experimental and computational. Until recently, the length of typical MD simulations taking into account many surrounding water molecules had been limited to the order of 0.1 ns, a length of time during which conformation transitions occur only rarely. Hence, it was not easy to obtain reliable simulation results within the framework of the empirical force-field employed. Experimental results suitable for comparison with simulation results were also unavailable. However recent computational and experimental developments have markedly improved the situation. New computers have made MD simulations at least 1 ns long commonplace, and accurate information about the structure of amylose in aqueous solution can now be obtained by observing small-angle X-ray scattering (SAXS) when using a synchrotron radiation source. Moreover, the recently resolved 3D structures of three cyclodextrins (CD_{10},

* To whom correspondence should be addressed. e-mail: shimada@frl.cl.nec.co.jp. Fax: 81-298-56-6173.
† NEC Corporation.
‡ Kyoto Prefectural University.
§ Kyoto Institute of Technology.

10.1021/jp992650z CCC: $19.00
Published on Web 02/10/2000

CD_{14}, and CD_{26})[28−31] are available for comparison. Thus we are now in a position to make a detailed comparison of simulation and experimental results.

This paper reports the results of such a comparison for maltoheptaose (G7) in aqueous solution. (In this article G*m* designates an amylose fragment composed of *m* glucose units.) Experimental and computational results are compared with respect to radius of gyration and SAXS profile. Although many force-fields have been and are being developed[5,6,17,32] (see also references cited in ref 27), here we report MD results for only five (Gromos,[10,11] Glycam93/99,[33,34] and Cff91/Cff[35]). These five are suitable for MD simulations of solutions. Gromos was one of the earliest force-fields that dealt with amylose chains in aqueous solution. The Glycam force-field parameters for wide variety of biologically important saccharides are available,[33] and these parameters have already been shown to be successful in simulating various saccharides and in modeling a protein−saccharide complex.[36] We also carry out Monte Carlo calculations of an isolated solute molecule for two force-fields (MM3[37] and a simple force-field[22]) by assigning a dielectric constant of four and regarding the interatomic potential as a mean potential averaged over solvent configuration. There are fundamental differences between the MD models and the Monte Carlo models. The former takes into account all interactions between all the glucose residues of G7. The latter are based on a disaccharide energy map and take into account only the interactions of nearest neighbor residues. Thus we are comparing here more than just the various force-fields. In one case, the longer range interaction is taken into account. In another case they are not. Moreover the MD models include explicit solvent.

We show here that the Monte Carlo results were quite satisfactory but that the MD results were unexpectedly force-field dependent: Cff91 was successful in reproducing both the SAXS profile of G7 and the radius of gyration for very long chains, but other force-fields were not. This may be partly due to the fact that dimensional properties of the chain are very sensitive to force-field parameters, as is well-known in polymer chain statistics based on a rotational isomeric state model.[38] Rather than abandoning unsuccessful force-fields, we attempt to seek a source of the weakness (only of Glycam93) by analyzing the conformation preference exhibited in various Ramachandran-like plots for the glycosidic linkages. Given that the MD results based on force-field models are puzzling, it seems sensible to make ab initio calculations. Thus we proceed to make a similar plot of ab initio energies calculated at the Hartree−Fock (HF) level. However it must be noted that our HF calculations have only limited accuracy, partly because dispersion attraction is nonexistent at the HF level and partly because our HF energies are calculated at molecular structures optimized by using the Cff91 force-field. Our analysis from various standpoints, however, suggests that the weakness detected in this particular case is most likely due to an over-preference of a certain rotational state of the glycosidic linkage, rather than to other effects such as the neglect of the polarization of solvent.[39] On the basis of the HF results, we make a modification of Glycam93 by devising a remedy applicable only to amylose. We also perform MD simulations with Glycam99, a new version which became available after completing our modification of Glycam93. Our MD simulation results indicate that Glycam99 is appreciably better than Glycam93, as far as G7 is concerned. Finally, the significance of force-fields is illustrated in modeling CD_{14}.

It has been widely accepted that all currently available force-fields have weaknesses to some extent, because each of them is tuned only to a limited number of molecular properties. This study should, therefore, serve to to explore the direction of the next-generation force-fields for poly- and oligosaccharides.

II. Experimental Methods

A. Materials. Maltoheptaose (G7) was purchased from Hayashibara Biochemical Laboratories, Inc. The sample was further purified by gel-filtration chromatography on a Biogel P-4 column (extra Fine 1.5 i.d. × 120 cm) with water as the eluent. The purity was checked by high-performance liquid chromatography on a Polyamine II column (4.6 i.d. × 25 cm, YMC Co. Ltd., Japan) and fast-atom bombardment mass spectrometry (JEOL-JMS-SX102A mass spectrometer). α- and *β*-cyclodextrins were purchased from Nakarai tesque and further purified by recrystalization, once from 1-propanol and twice from water. These hydrate crystals were dried at 80 °C in a vacuum under P_2O_5 for 2 days. Solutions measured were prepared by weight using milli-Q water.

B. Small-Angle X-ray Scattering. The synchrotron radiation small-angle X-ray scattering was observed from aqueous solutions of linear amylose (G7). The concentrations of the solutions measured were 20.13 and 40.25 mg/mL. The SAXS experiments were performed with SAXES (small-angle X-ray scattering equipment for solutions) optics installed at BL10C in the Photon Factory, Tsukuba, Japan, according to the method previously described.[15,23] The solutions were placed into 1-mm-thick flat cells (1 × 0.5 × 0.1 cm) with a pair of 20-mm-thick quartz windows, which were subjected to SAXS measurements. The observed range of the magnitude q of the scattering vector **q** was from 2.50×10^{-2} to 0.5 $Å^{-1}$, which is equivalent to the Bragg spacings from 251 to 12.6 Å. The excess scattering intensities were calculated by subtracting the scattering intensity of water from that of the solution. The particle scattering function $P(q)$ is defined by $P(q) = I(q)/I(0)$, where $I(q)$ is the scattered intensity at q, and $I(0)$ is the corresponding intensity extrapolated to $q = 0$.

Measurements for cyclodextrins were carried out in a similar manner. The concentrations of the α-cyclodextrin solutions were 43.08, 21.54, and 11.02 mg/mL, and the concentrations of the *β*-cyclodextrin solutions were 12.70, 6.35, and 4.23 mg/mL. Note that *β*-cyclodextrin is less soluble in water.

III. Computational Procedures

A. Molecular Dynamics Simulation. The initial conformation of G7 and G2 was, unless otherwise noted, assumed to be the regular helical structure built from the atomic coordinates obtained by Goldsmith et al.[40] Their structure is characterized by the torsion angles of $\phi = 105°$ and $\psi = -135°$ associated with the glycosidic linkage connecting two adjacent glucose units. These angles ϕ and ψ are the major factors that determine the backbone conformation of amylose. In this work we adopt the convention $\phi = O(5)_n-C(1)_n-O(4)_{n-1}-C(4)_{n-1}$ and $\psi = C(1)_n-O(4)_{n-1}-C(4)_{n-1}-C(5)_{n-1}$. A relation to other conventions is noted.[41]

The orientation of the hydroxymethyl group relative to the pyranose ring is classified into three rotational states: gg, gt, and tg, where the first letter refers to the torsion angle about O(5)−C(5)−C(6)−O(6), and the second letter to the torsion angle about C(4)−C(5)−C(6)−O(6). If the first (second) torsion angle is designated χ, these three states, respectively, correspond to $\chi = -60°$, 60°, and 180° (60°, 180°, and −60°). These states are reported to occur in the ratio of 60:40:0 for carbohydrates containing gluco configuration,[42] and 3.5:1:0 for various CDs in crystal environments.[43] In many simulations, we took gt states

TABLE 1: MD Simulations of an Amylose Fragment (G7)

force field	run	simulation time (ns)	R_G (Å)[a]	C_∞[b]	start point
Cff91	1	3.58	7.704	5	gt
Cff	1	3.5	8.272	10	gt
Glycam93	1	2.0	7.924		tg
	2	1.4	9.023	12[c]	gt
	A	1.4	9.048	40	gt
Glycam99	1	3.5	8.089	11	gt
	A	3.0	7.981	11	gt
modGlycam93	1	2.2	8.089		*d*
	2	2.2	7.410	2[c]	*e*
GROMOS	1	1.4	8.334	12	gt
	A	1.4	8.560	15	gt
(Obs.)			7.4 ± 0.2[f]	5 ± 1[g]	

[a] Root-mean-square average over simulated values beyond 0.5 ns. Run A was performed under the simulation condition used for Cff91 and Cff. [b] Defined for infinitely long chains (see eq 8). The values of C_∞ are rough estimates based on the distributions of ϕ and ψ (see the Discussion section). [c] Average over runs 1 and 2. [d] Snapshot at 1 ns of Glycam93/run 1. [e] Final snapshot of run 1, followed by 0.2 ns simulation at 500 K. [f] Determined from SAXS measurements. [g] Experimental values.[59]

as the initial states, expecting transitions to other states. Molecular dynamics simulations with the Gromos[10,11] and Amber/Glycam[33,34,44] parameters with the TIP3P water model[45] were carried out in a manner previously described.[14,15] The solute molecule was immersed in a large water sphere with a radius of about 25 Å (1999 water molecules), and simulations were carried out for at least 1 ns with a time step of 2 fs at 300 K. Results before 0.5 ns were excluded from analysis. In some cases, two runs differing in initial structure were carried out (see Table 1). Molecular dynamics calculation employed a house-made program (NEC), which incorporates a fast multipole method to accurately calculate long-range electrostatic interactions.[46,47] All bond lengths were kept constant by applying the SHAKE procedure.[48] Temperature was controlled by using the Nose−Hoover scheme.[49] The center of mass of the solute molecule was restrained near the center of the sphere by applying a harmonic potential with a weak force constant of 2 kcal mol^{-1} Å^{-1}. When a water molecule left the sphere surface and moved more than 1.5 Å from the sphere surface, it was pulled back by applying a boundary force (with a force constant of 20 kcal mol^{-1} Å^{-1}) directed toward the center.

Molecular dynamics calculations with the Cff91 and Cff force-fields were carried out by using the DISCOVER program (MSI Corp.). The calculation protocols for them were slightly different from those described above because not all protocols can be implemented by the built-in options of DISCOVER. To prevent water molecules from escaping from the sphere, a short MD run was performed to equilibrate water positions, and then a list of water molecules whose oxygen atom was more than 21 Å from the sphere center was created. During subsequent simulations, these water oxygens were made to work as obstacles to the escape of molecules inside (i.e., they were restrained to their respective initial positions by using the tether option of DISCOVER). To prevent the solute from floating out toward the sphere surface, the heavy atoms of the pyranose ring of the central (i.e., 4th) glucose residue were also restrained to their initial positions by using the tether option. All bond lengths were kept constant by applying the RATTLE algorithm[48] with a time step of 0.5 fs. The number of water molecules was 2385 for G7 and 2154 for G2.

All calculations in this work were performed on SX4 supercomputers and ONYX2 or EWS workstations.

B. Monte Carlo Simulations. We carried out Monte Carlo calculations as previously described.[22] The Monte Carlo simulation requires a map of probability distribution $P(\phi,\psi)$ of glycosidic torsion angles ϕ and ψ. Under the assumption that one glycosidic linkage has no appreciable influence on the rotation (ϕ,ψ) of another glycosidic linkage (independent rotations), $P(\phi,\psi)$ may be evaluated as the Boltzmann weight

$$P(\phi,\psi) = c \exp[-E(\phi,\psi)/k_B T] \quad (1)$$

where $E(\phi,\psi)$ is the conformational energy of maltose evaluated as a function of ϕ and ψ, k_B is the Boltzmann constant, T is the absolute temperature, and c is a normalization constant. The probability maps used in this work are the ones that have already been calculated.[22,50] The first probability map[50] is the so-called flexible map.[51] In short, energy minimization of each of 24 conformers differing in the orientations of the hydroxymethyl and terminal hydroxyl groups was carried out under the constraint that ϕ and ψ were fixed, and then the lowest energy obtained among the 24 minimization runs was used in the calculation of the probability map. The force-field used was MM3.[37] The second probability map[22] is the so-called rigid map[51] (i.e., a map calculated without energy minimization). The force-field used was a simple one that takes account of only nonbonded (i.e., van der Waals plus electrostatic) potential terms.[22] This simple force-field is designated as K, for later convenience. In both maps the effective dielectric constant was set to four.

The angles ϕ and ψ of each chain were generated according to the probability map, and then the entire chain structure was built on the basis of the Gress−Jeffrey[52] coordinates of the glucose residue, with the hydroxymethyl group in the gg or gt state. For comparison, calculations were also performed by using the Brown−Levi[53] and Arnott−Scott[54] coordinates. Over 1000 chains were generated, and the atomic coordinates of each were stored for later use. The excluded volume effect can safely be ignored for chains as short as G7.

C. Scattering Function. The scattering function $F(q)$ for each chain was calculated according to the Debye formula and then averaged over the ensemble of all generated chains:

$$F(q) = \sum_{i=1}^{n} g_i^2 \phi_i^2(q) + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} g_i g_j \phi_i(q)\phi_j(q) < \sin r_{ij}q/r_{ij}q > \quad (2)$$

where q is the magnitude of the scattering vector $\mathbf{q}$, g_i is an atomic scattering factor (which may be equated to the atomic number), $\phi_i(q)$ is the form factor for a single atom, r_{ij} is the distance between atoms i and j, and $\langle\cdots\rangle$ designates the ensemble average. If the electron density associated with atom i is assumed to be uniform within a sphere centered at the atomic nucleus, $\phi_i(q)$ is given by

$$\phi_i(q) = 3(\sin \sigma_i q - \sigma_i q \cos \sigma_i q)/(\sigma_i q)^3 \quad (3)$$

where σ_i is the radius of the sphere. In refs 15 and 23, σ_i was taken as the van der Waals radius of the ith atom. Very recently, however, Liu et al.[55] have pointed out that the van der Waals radius when scaled by a factor of 0.6 for C and O atoms improves agreement of $\phi_i(q)$ with tabulated atomic structure factor at high q region ($q \geq 0.3$ Å^{-1}). We therefore used $\sigma_i = 1.64 \times 0.6$ Å for the C atoms and 1.5×0.6 Å for the O atoms. We used $\sigma_i = 1$ Å for the H atoms, which are of minor

importance. The intensity plotted in figures is the reduced intensity defined by

$$P(q) = F(q)/F(q=0) \tag{4}$$

The (root-mean-square) radius of gyration, R_G, of the chain molecule can be determined from an analysis of the scattering intensity at low q because $P(q)$ is expanded as

$$\begin{aligned} P(q) &= 1 - \frac{1}{3}q^2{R_G}^2 + ... \\ &= \exp\left(-\frac{1}{3}q^2{R_G}^2 + ...\right) \end{aligned} \tag{5}$$

Strictly speaking, the radius of gyration ${R_G}^2$ appearing in eq 5 is slightly different from the g_i-weighted radius of gyration ${R_G}^{(0)2}$ calculated with $\sigma_i = 0$:

$${R_G}^2 = {R_G}^{(0)2} + \frac{3}{5}\langle\sigma^2\rangle \tag{6}$$

where $\langle\sigma^2\rangle$ is the g_i-weighted mean-square average of the radius σ_i. Here the g_i-weighted average in general is defined by

$$\langle A\rangle = \sum_{i=1}^{N} g_i A_i / \sum_{i=1}^{N} g_i \tag{7}$$

However, contribution of this additional term is generally small, and the difference between ${R_G}^2$ and ${R_G}^{(0)2}$ is less than 0.05 Å in our case. For simplicity, we shall hereafter ignore this additional term, and report ${R_G}^{(0)}$ values as R_G.

IV. Results

In this work, five types of force-fields (Gromos,[10,11] Glycam93/99,[33,34] and Cff91/Cff[35]) were used. Both Gromos and Glycam have a parameter set specifically developed for amylose, while Cff91 is a general-purpose field targeting biomolecules. More recently developed Cff includes parameters derived for carbohydrate compounds with proper account of the anomeric effect on the glycosidic linkages, which effect plays an important role in determining the conformation of amylose. The important difference between Glycam and Gromos is in the treatment of the exoanomeric effect on the glycosidic linkage. Glycam accounts for this effect by adding special torsional terms determined from conformational energy analysis on the basis of ab initio geometry optimizations of axial 2-methoxytetrahydropyran (a model for α-pyranosides) performed at the HF 6-31G* level. Glycam 99 is a new version parametrized with SCEE = 1.2 adopted in the newest version of Amber.[26] The original Gromos field developed by Koehler et al.[10,11] does not incorporate this effect. Although Ott and Meyer[8,9] later proposed a modification based on results for dimethoxymethane ($CH_3-O-CH_2-O-CH_3$) at the 6-31G* level, we used the original field because their modification is expected to lead to too extended conformations (see Discussion section).

A. Radius of Gyration. The (root-mean-square) radius of gyration R_G was evaluated from the initial slope of the Guinier plot (ln $P(q)$ versus q^2) of the experimental data (see eq 5).[15] The observed value of R_G was 7.4 ± 0.2 Å. The simulated value of R_G was the g_i-weighted average computed directly from the coordinates of the atoms.

Figure 1 shows the instantaneous value of R_G plotted against MD simulation time. The correlation of R_G with atomistic molecular structure is shown in Figure 2 by representative MD and Monte Carlo snapshots along with their R_G values given in the caption. Figure 2 also includes, for comparison, a few regular helical structures (a−c) proposed for amylose.[40,56,57] Their geometrical parameters are listed in Table 2.

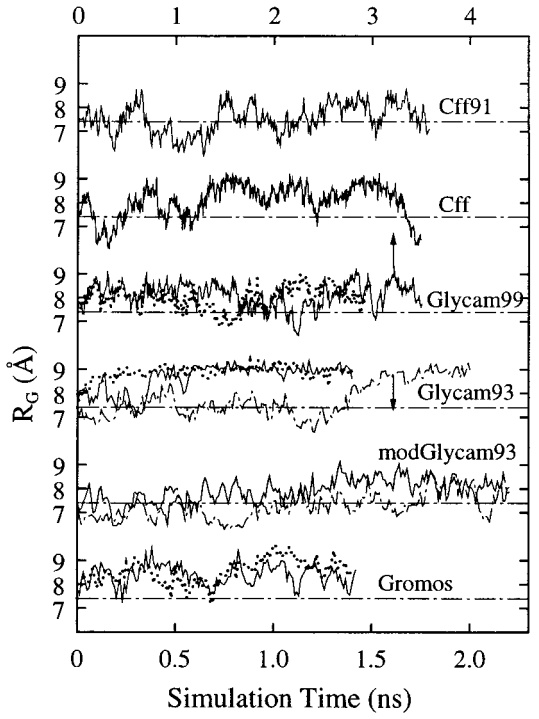

Figure 1. Radius of gyration R_G of an amylose fragment (G7) composed of seven glucose units plotted against simulation time. The solid (run 1) and dashed−dotted (run 2) curves represent the MD values calculated by using the indicated force-field. For comparison, the MD values (run A) for Glycam93, Glycam99, and Gromos performed under the simulation condition used for Cff and Cff91 are shown by the dotted curves. The horizontal lines represent the observed value of 7.4 Å.

When Gromos was used, it is seen from Figure 1 that the computed R_G was consistently larger than the observed value of 7.4 Å, indicating that the simulated chain is more expanded than the actual chain. It is important, however, to note that the chain exhibits rather large conformational fluctuation. The two versions of Glycam are quite different in behavior. The computed R_G for Glycam99 fluctuates around 7.4 Å and the average is 7.9−8.1 Å, which is not far from the experimental value of 7.4 Å, while R_G for Glycam93 increases to a large plateau value of 9.0 Å when the simulations were continued sufficiently long. The simulated chain for Glycam93 is fairly extended and very slender (Figure 2e). For the simulation using Cff91, in contrast, there is good agreement between experimental and simulation results; the Cff91 values fluctuate around 7.4 Å, and the average (7.7 Å) is only slightly large than 7.4 Å experimentally obtained. Thus Cff91 can well mimic the conformation of linear amylose. In summary, the agreement with the observed value of R_G is in the order Cff91 > Glycam99 > Cff ≈ Gromos > Glycam93.

The Gromos/Glycam and Cff91/Cff simulations were performed under different conditions (solvent boundary condition and restraint of central ring atoms). To examine the effect of these conditions, we made additional Gromos/Glycam simulations (run A) under the Cff/Cff91 condition, and the obtained results are indicated by the dotted curves in Figure 1. It is seen that the simulation results are essentially independent of the simulation condition.

Thus the MD results show unexpectedly large force-field dependence. It is surprising to us that Glycam93 is not

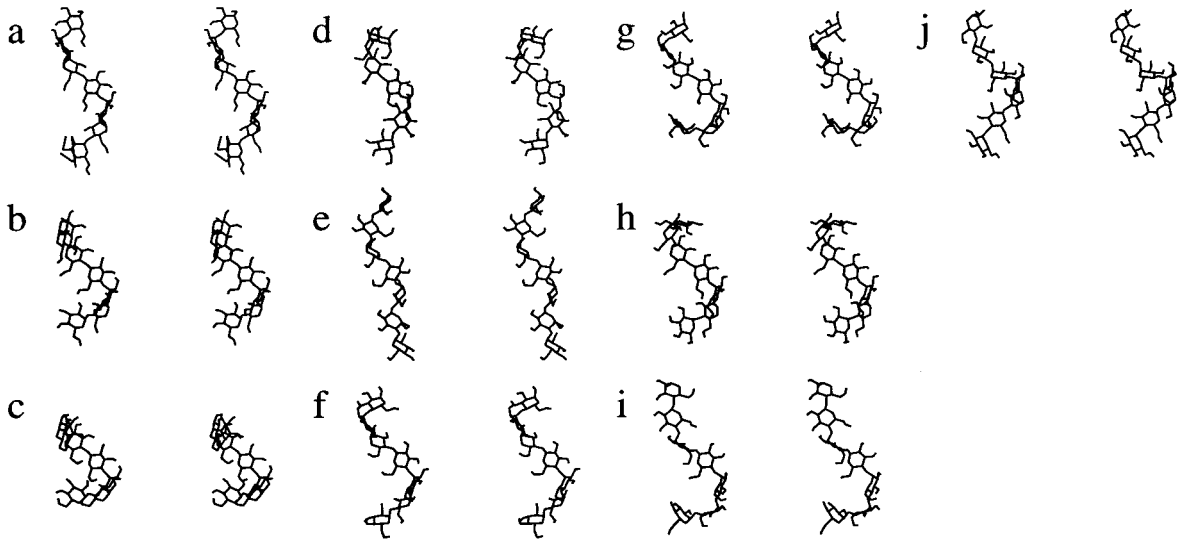

Figure 2. Stereoviews of typical conformations of G7. Regular helical forms: (a) Imberty et al.[56] (8.3), (b) Goldsmith et al.[40] (7.4), (c) Vh form[57] (5.9), whose helix parameters are listed in Table 2. The other conformations are representative snapshots of the simulations. Monte Carlo: (d) MM3 (7.57). Molecular dynamics: (e) Glycam93 (8.85), (f) Glycam99 (8.08), (g) modified Glycam93 (7.72), (h) Cff91 (7.83), (i) Cff (8.30), (j) Gromos (8.32). Each of the values enclosed in parentheses is the radius of gyration (Å). For clarity, hydrogens connected the carbon atoms are omitted. The centers of the right and left views are separated by 31.25 Å.

TABLE 2: Parameters of the Regular Helices Shown in Figure 2

structure	R_G (Å)	radius (Å)	rise (Å)	pitch (Å)	repeat (Å)	ϕ,ψ (°)	notes
a	8.336	2.75	3.55	21.42	6.0	91.8, −153.2	*a*
b	7.387	5.38	2.44	17.60	7.2	105, −135	*b*
c	5.904	4.13	1.36	8.16	6.0	105.6, −127.5	*c*

[a] A strand of the parallel double helical structure determined by Imberty et al.[56] Listed is the radius value of the helical trajectory formed by the glycosidic oxygens. [b] Regular helix model proposed by Goldsmith et al.[40] on the basis of the maltopentaose structure on phosphorylase. [c] The Vh helix structure determined by Rappenecker and Zugenmaier.[57] Its overall geometry is comparable to the V helix structure recently determined by Gessler et al.[31] at atomic resolution.

TABLE 3: Monte Carlo Simulations of an Amylose Fragment (G7)[a]

force field	τ (°)[b]	glucose geometry[c]	R_G (Å)	C_∞[d]
MM3, ϵ=4	117	GJ, gg	7.572	4.4[e]
	117	GJ, gt	7.510	
	117	AS, gg	7.489	
	117	AS, gt	7.414	
	117	BL, gg	7.885	
	117	BL, gt	7.813	
K[f], ϵ=4	118	GJ, gt	7.2	5.0
(Obs.)			7.4 ± 0.2	5 ± 1

[a] Number of generated chains is 1000. [b] τ is the bond angle at the glycosidic oxygen. [c] GJ, AS, and BL represent the glucose geometry used, and gg and gt designate hydroxymethyl orientation (see the text). [d] Defined for infinitely long chains. See eq 8. [e] Yui and Kitamura.[50] [f] Simple force-field with only nonbonded potential terms.[22]

successful, because in our view the derivation of the force-field parameters from ab initio calculations is on the sound theoretical basis. We therefore modified Glycam93 also by using ab initio results but in a different manner (described in detail later in this article). As seen from Figure 1, the obtained result shows improved agreement with the experimental results.

Monte Carlo calculations, on the other hand, were quite satisfactory, giving 7.5 Å (MM3) and 7.2 Å (K) for G7, both of which values are in good agreement with the experimental values. To characterize R_G for very long chains (unperturbed by the excluded volume effect[38,58]), it is convenient to express results in terms of the so-called characteristic ratio defined by

$$C_\infty = \lim_{m\to\infty} 6R_G^2/ml^2 \qquad (8)$$

where m is the number of glucose residues in the chain and l is the virtual bond length ($\approx$ 4.5 Å), i.e., the distance between consecutive glycosidic oxygens. It has already been confirmed that the probability maps give C_∞ = 4.4 (MM3)[50] and C_∞ = 5.0 (K)[22] in close agreement with $C_\infty = 5 \pm 1$ experimentally observed.[59] Note that the flexible map based on MM2 had been abandoned, because that map gave $C_\infty = 3$,[21] which is too low.

Monte Carlo result slightly depends on the glucose geometry used, as summarized in Table 3. As far as G7 is concerned, the Brown−Levi (BL) structure[53] gives somewhat larger (0.4 Å) values of R_G than does the Gress−Jeffrey (GJ) structure,[52] whereas the Arnott−Scott (AS) structure[54] gives slightly smaller (0.1 Å) values. Influence of the conformation gt or gg is generally small.

B. Small-Angle X-ray Scattering Profile. Figure 3 shows a comparison of the experimental values (circles) and the simulated values (curves) with respect to the scattering profile. The curves are for six MD simulations (Glycam93/99, Cff91/Cff, Gromos, modified Glycam93) and two Monte Carlo simulations, one based on MM3 and the other based on the simple potential (K).[22] In the figure, the corresponding average R_G values are also indicated. In view of the scatter of experimental data for high q ($\geq$0.3 Å^{-1}), the Cff91 values may be regarded as in good agreement with experimental data. The Gromos and Glycam93 values are, however, in rather poor agreement even for low q. Both Monte Carlo values show good agreement with experimental values. Although the difference in the glucose geometry became visible for the high q region, general tendency of the profiles was essentially similar (not shown). Modification of Glycam93 (see below) is again encouraging. Thus the agreement and disagreement for the profile are in accord with those for R_G.

Of the regular helix models whose parameters are listed in Table 2, which helix is closest to the conformation found in

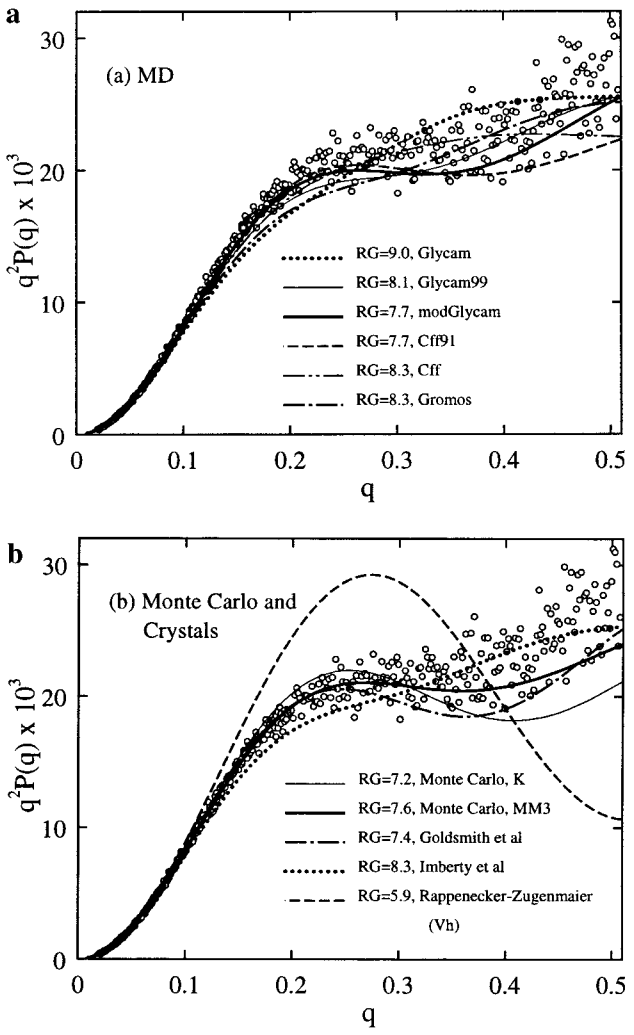

Figure 3. Kratky plots of SAXS experimental data (at 25 °C and at 20.13 mg/mL) for G7 along with simulated scattering profiles. The curves are profiles obtained by averaging over the ensemble of amylose chains generated by the MD or Monte Carlo calculations. Values of the radius of gyration R_G (Å) are also included. (a) MD results, and (b) Monte Carlo results and profiles calculated from crystalline regular helical structure models.

aqueous solution? The helix (b) of Goldsmith et al.[40] is the most satisfactory with respect to both R_G and SAXS profile (Figure 3b). Of particular interest is that the Vh (or V) form[57] found in a crystal environment cannot explain the SAXS data for G7 in aqueous solution, because the Vh form gave too low R_G value of 5.9 Å compared to the experimental value of 7.4 Å and gave too high first peak in the SAXS profile. The V or Vh form has a relatively small pitch, and its stability depends entirely on having successive turns of the helix in favorable contact with each other. This probably the reason crystalline V or Vh is not a good model for the solvated small fragment G7 with an open annulus.

Figure 4 compares the SAXS profiles observed for α- and β-cyclodextrins (CD_6 and CD_7) with the theoretical profiles calculated from the corresponding crystal structures.[60,61] These molecules are suitable for comparing the experimental and theoretical treatments used in the present work, because their conformation in solution should be rather close to the conformation of the crystal structures, since the small ring size results in reduced flexibility. The good agreement between the experimental data and the theoretical values thus confirms the validity of the treatments.

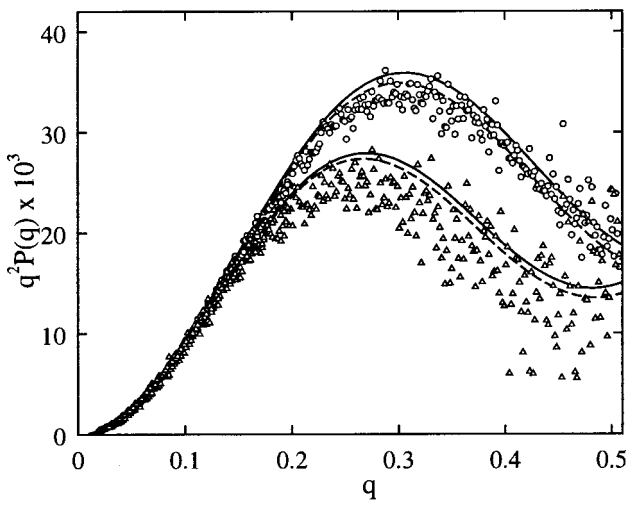

Figure 4. Kratky plots of SAXS experimental data for CD_6 (upper) and CD_7 (lower). The points are data at 25 °C and at 11.02 mg/mL for CD_6 and 12.70 mg/mL for CD_7. The solid and dashed curves represent the profiles calculated from the crystal structure with and without a scaling of the van der Waals radius, respectively. For CD_6, the value of R_G determined from the SAXS profile is 5.686 Å, which is in good agreement with 5.356 Å calculated from the crystal structure. For CD_7, the corresponding values are 6.299 Å (SAXS) and 6.076 Å (crystal).

V. Discussion

A question one may naturally ask is why did the R_G values and the scattering profiles show such unexpectedly large force-field dependence? To answer this question, we made various Ramachandran-like plots of glycosidic torsion angles ϕ and ψ and used these plots to characterize the simulated conformations for the various force-fields.

A. Probability Distribution Obtained from MD Calculation in Aqueous Environment. Figures 5a−e show the distribution of the angles (ϕ,ψ) for various MD calculations carried out for G7 in an aqueous environment. The values calculated every 10 ps intervals are shown by the small circles. It is helpful to classify the entire domain into several rotational isomeric states (A, B, C1, C2, D, E, and G),[62,63] as shown schematically in Figure 6. These states are based on various crystal structures[30,60,61,64−68] and energy calculations. The angles obtained from Gromos are rather broadly distributed, centered at B and C2. The angles from Glycam99 are also broadly distributed, but shifted to higher ψ region. The angles from Glycam93, in contrast, are sharply populated in a different domain centered at A. When every glycosidic linkage is in state A, the chain takes a slender and extended regular helical structure (like Figure 2e), whose helix repeat of 4 is small compared with 6−8 normally found in crystals. The distribution from Cff91 is the most widespread; most of the angles are near G and C2, but they spread into the right-handed helix region (the lower boundary of which is indicated by the dashed curve). Cff and Cff91 are different in behavior; two peaks appears (at A and G) for Cff. An amylose molecule has many hydroxymethyl and hydroxyl groups, and its conformational energy depends on their orientations. Although such details are smeared out in Figure 5, the results show that different force-fields tend to produce different conformational distribution of (ϕ,ψ).

For comparison, simulations were also carried out for α-maltose (G2), which is expected to reach equilibrium more rapidly. It was found that the distribution for G2 is essentially similar to that for G7 (results not shown). It is pertinent here to

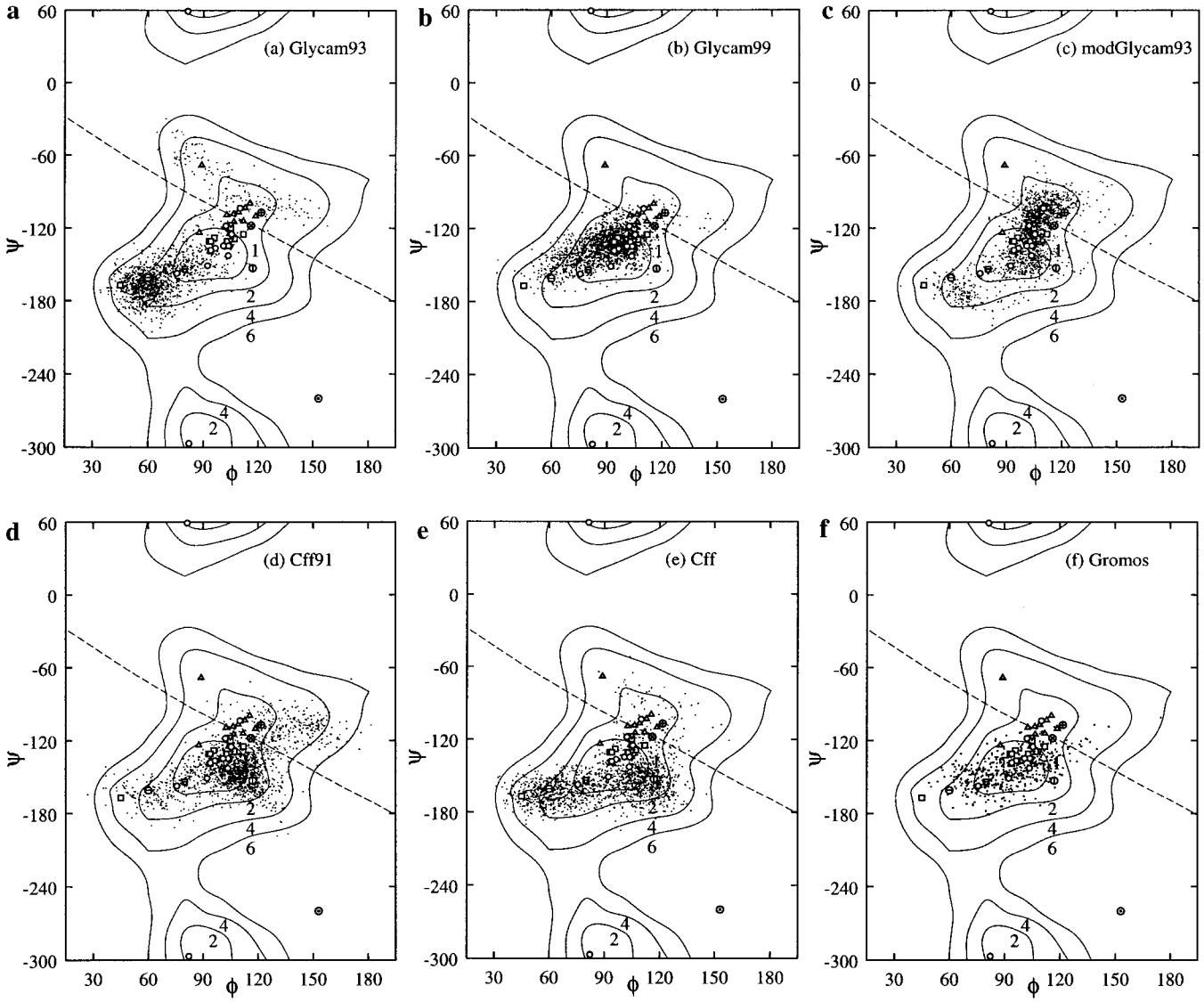

Figure 5. Ramachandran-like plot for G7. Distributions of glycosidic torsional angles ϕ and ψ are shown. The small points represents values calculated from the snapshots (at 10 ps intervals) of MD simulations in an aqueous environment. The force-fields are (a) Glycam93, (b) Glycam99, (c) modGlycam93, (d) Cff91, (e) Cff, and (f) Gromos. Shown by the solid curves are the MM3 iso-energy contours calculated with $\epsilon = 4$. The numbers attached to the curves indicate the energy (kcal/mol). The dashed curve separates the map into two domains corresponding to the left-handed (lower) and right-handed (upper) chilarities. Other symbols represent published values for various crystal structures: (○) CD_{14};[30] (△) α- and β-cyclodextrins;[60,61] (square) antiparallel double helix;[64,65] (▽) maltopentaose on phosphorylase;[40] (circle with a dot) maltotetraose bound to β-amylase;[66] (⊗,⊕) α- and β-maltoses;[52,67] (circle with a vertical or horizontal bar) modeled structure of amylose complexed with KBr or KOH.[68]

mention the result of Ott and Meyer.[8,9] Their simulation result for G2 shows a narrow distribution centered at (71°, −156°),[8] and is very close our Glycam93 distribution for G7. Hence, their modification of Gromos is expected to lead to too extended conformations. For this reason, we have not adopted their modification.

MD simulations of long chains are not feasible at present. However, conformational properties of long chains can be approximated by those of the oligo- or disaccharides, if independent rotations about different glycosidic linkages are assumed. In practice, we made Monte Carlo calculations by regarding the distribution of (ϕ,ψ) just obtained from our MD simulations as $P(\phi,\psi)$. Note that every MD samples of (ϕ,ψ) have an equal weight. In this way, rough estimates of R_G for long chains and of C_∞ were obtained. The obtained results for C_∞ are included in Table 1. It is seen that Cff91 gave $C_\infty = 5$, in satisfactory agreement with $C_\infty = 5 \pm 1$ experimentally obtained.[59] Other force-fields gave too large or too small values. However, it must be noted that these estimates are rough because unusual domain E with the inverted conformation of the glycosidic linkage is not explored in our MD simulations. If domain E was assumed to be populated at 5%, C_∞ was roughly estimated to decrease to 6 for Cff/Glycam99 and by a factor of 0.7 for the other force-fields. (A population at E is 4% for MM3, which has a rather low local minimum of 1.5 kcal/mol.) Although a population at E has negligible influence on short chains such as G7, its effect get amplified for longer chains. This indicates a need of more detailed study of the inverted conformation in future.

Hydrogen bonds between $O(2)_n$ and $O(3)_{n-1}$, bridging consecutive residues, are known to be structurally important. According to the hydrogen bond criteria described in the caption of Figure 6, these hydrogen bonds occurred about 50% on average during our simulations for Cff91, Glycam99, and modified Glycam93, and occurred about 30% for Cff and Gromos, while these bonds scarcely occurred for Glycam93. Thus the occurrence of the hydrogen bonds was fairly force-field dependent. It is important to note that this population is

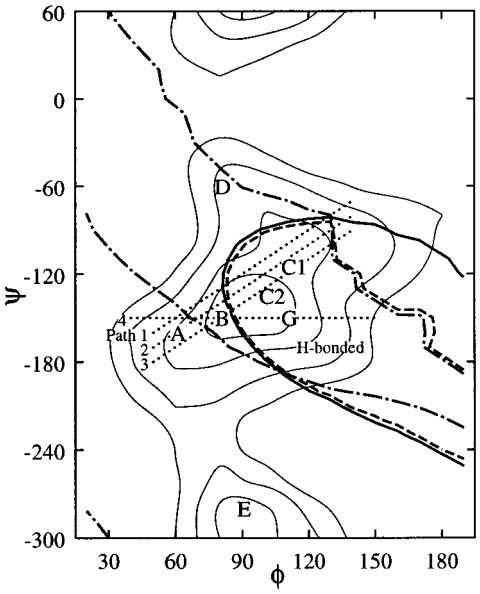

Figure 6. Important domains on the Ramachandran-like plot. Representative rotational isomeric states are indicated by A, B, ..., and G, where their names follow from Tran et al.[62] and Pérez et al.[63] Three criteria for the hydrogen bonding between O(2) and O(3)′ are satisfied in the domain enclosed by the following curves: (a) the solid curve on which O(2)···O(3)′ = 3.4 Å, (b) the dashed curve on which O(2)···H(3)′ = 2.4 Å, and (c) the dashed−dotted curve on which angle O(2)···H(3)′···O(3)′ = 150°. The structures optimized with Cff91 were used in drawing these curves. Also shown are the dotted lines indicating paths 1−4 (see the text).

closely related to the values of angles ϕ and ψ, because these angles considerably limit the allowable range of the distance between $O(2)_n$ and $O(3)_{n-1}$. Specifically for the structures of maltose optimized with Cff91 (see the next subsection), these hydrogen bonds are formed at (ϕ,ψ) in the domain indicated in Figure 6. Rotational isomeric state A is seen to be far apart from that domain. This may explain why Glycam93 that showed a strong preference for state A scarcely gave these hydrogen bonds. Ott and Meyer also observed only a small population (3%) of the hydrogen bonds in their simulation of maltose.[8]

B. Conformational Energies in Vacuum. Here we compare the force-field and ab initio energies of maltose (G2) rather than the more complex maltoheptaose (G7). Under the constraint that ϕ and ψ are fixed at their grid points in the range between 20° $\leq \phi \leq$ 190° and −300° $\leq \phi \leq$ 60° (with a spacing of 10°), energy minimization of the structure of α-maltose was carried out using the DISCOVER program, for Cff91 with $\epsilon = 1$ (vacuum). For each optimized structure thus obtained, the corresponding Glycam and ab initio energies were evaluated. Ab initio calculations were carried out at the HF level by the AMOSS program.[69] The basis set is MIDI4 plus polarization. Because many HF calculations are still computationally intensive, the HF energies for high-energy region were computed only at a fourth of the total grid points, and values at the remaining grid points were interpolated from them. In this process of minimization, the initial orientation of hydroxymethyl group was assumed to be in the gg state. This restriction should not be a bad one, because gg is the state most frequently found in a crystal environment.

Figures 7a−d show such energy maps for Glycam93/99, Cff91, and HF. It is seen that the low-energy area of Glycam93 is rather narrow and shifted to the left; compare the solid contours drawn at 4 kcal/mol level. Before proceeding to the next section, it pertinent to make three remarks on computational details. (i) To compare the HF and force-field energies, molecular structure must ideally be optimized by using the respective methods. However, the minimization of many (≈200) structures with HF is still computationally too intensive. We have therefore used, for simplicity, the structure optimized with Cff91 throughout. (ii) For simple molecules, it is rather straightforward to compare energies for the structure optimized by using different force-fields. However, for such molecules as maltose having many flexible hydroxyl groups, it is not. Indeed when we optimized structures by using Glycam93, we occasionally encountered a structure with the hydroxyl group orientation flipped from the Cff91-minimized one. Energy comparison for essentially the same structure may be accomplished by keeping the orientation of all hydroxyl groups fixed. We found that the Glycam93 energy profile calculated in this manner is not much different from those in Figure 9a shown below. (iii) As a result of the flipping, a slight complication arises: the HF and Glycam energies are not always continuous with respect to ϕ and ψ. This explains some irregularities seen in Figure 7.

C. Conformational Energies in Aqueous Environment. To obtain a simple insight into the solvent effect from a different viewpoint, we calculate energy profiles by using a continuum dielectric approach,[70] in which the surrounding water is approximated by a continuum solvent having an effective dielectric constant of 80.

In this approach, total conformational energy is expressed as the sum of the molecular energy evaluated in a vacuum and the solvation energy, the latter of which is in turn a sum of electrostatic and nonpolar contributions. The electrostatic contribution caused by the polarization of solvent around the polar solute can be estimated by solving the linearized Poisson−Boltzmann equation, while the nonpolar contribution can be estimated as a surface-area dependent term. Calculations were carried out using the DelPhi module.[71] Dielectric constants of the solute and solvent domains were assumed to be 2 and 80, respectively. Grid spacing was 0.3/Å, solute extent was 30%, and the molecular exclusion radius was based on Cff91. The charge on each atom was set to the value specified by the force-field used. The surface energy was found to be insensitive to ϕ and ψ, as far as this molecular system is concerned.

Figures 8a−c show conformation energy maps (force-field plus Delphi) thus constructed for maltose in an aqueous environment. Comparison with the corresponding maps in a vacuum (Figure 7) reveals that solvation modifies the potential surface somewhat, but not to the extent that conformation preference of (ϕ,ψ) is decisively different.

To clarify this point more quantitatively, we compared the energy profiles along four paths in the (ϕ,ψ) map: path 1, $\psi = \phi - 210$; path 2, $\psi = \phi - 220$; path 3, $\psi = \phi - 230$; path 4, $\psi = -150$. These paths are illustrated by the dotted lines in Figure 6. Figure 9 shows that the energy generated by Glycam is generally lower with or without solvation than that from the other force-fields for $\phi < 90°$ and generally higher for $\phi > 90°$. This explains why region A is favored by Glycam93, and therefore why Glycam93 overestimates the dimensions of the molecule. It is also seen that Cff91 shows smaller deviation from the HF values than does Glycam. But noticeably large deviation of Cff91 for $\phi > 100°$ on path 4 suggests that even the Cff91 energies are underestimated in this region (state G).

It is interesting to compare the correlation between these energy maps and the MD population maps, that is, to compare the energy domains enclosed by the solid curves (drawn at 4 kcal/mol) in Figures 8a−c with the corresponding populations

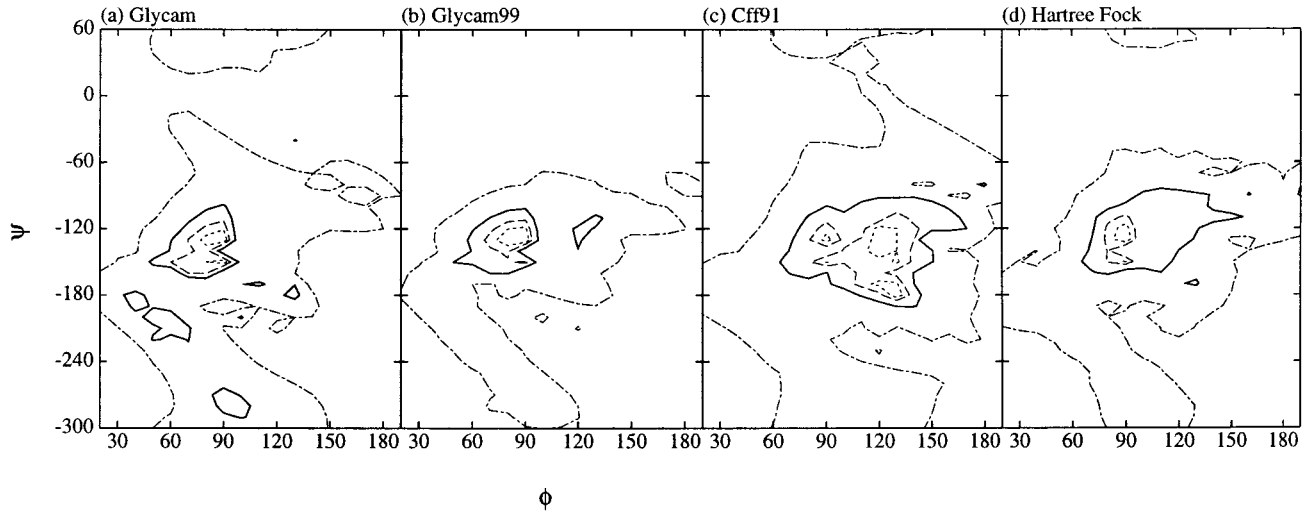

Figure 7. Ramachandran-like plot for α-maltose (G2) in a vacuum. Iso-energy contours for 1, 2, 4, and 10 kcal/mol are, respectively, shown by the short-dashed, dashed, solid, and dashed−dotted curves.

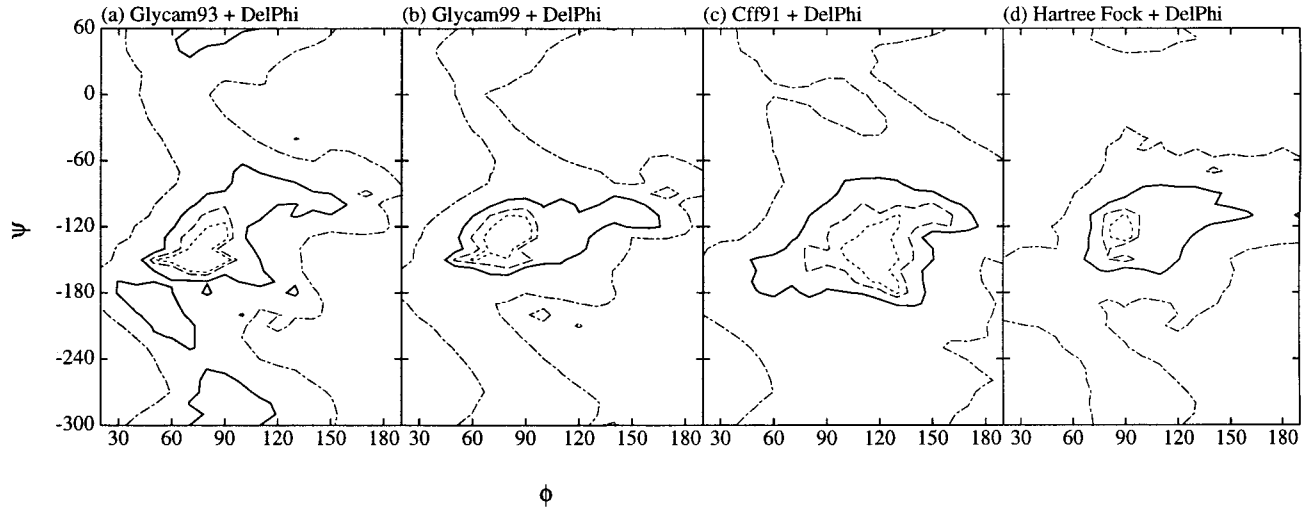

Figure 8. Ramachandran-like plot for α-maltose (G2) in an aqueous environment. Conformation energy is the sum of force-field (or HF) energies and solvation energies. Solvation energies were calculated by using the Delphi program based on the continuum dielectric approach. For a key to the various curves, see the caption of Figure 7.

shown in Figure 5. There is seen a good correlation for Cff91, with which the maltose structure was optimized, but only a modest correlation for Glycam. However, the correlation should be considered as only qualitative in nature, because the conformational (free) energy thus calculated ignores the contribution from the conformational entropy of maltose. Indeed Schmidt et al.[7] have already shown the importance of this contribution in their simulation of maltose in a vacuum. They found that conformational entropy considerably flattens the overall potential shape between $\phi = 50°$ and $\phi = 100°$, lowering the potential barrier by as much as 2 kcal/mol.

Similar conformation energy map and profiles were constructed for the HF results and are included in Figures 8d and 9b. When these profiles and maps were calculated, the electrostatic charges used in the Delphi calculation were taken as the Mulliken charges computed for each conformation. As seen by comparing Figure 9a with 9b, solvation does not appreciably alter the position of the potential minimum, even when the conformation dependence of charges is taken into account in this manner. Surely the conformation dependence of charges may be important in a vacuum, but the present result suggests that its effect on solvation is of rather minor importance. A more complete analysis based on more suitable ESP (electrostatic potential fitted) charges instead of the Mulliken charges would be better, but is beyond the scope of the present article.

D. Modification of the Torsional Potential. In attempting to modify the Glycam93 field, there are at least two approaches worthy of consideration. One approach is to use additional torsional potentials associated with the glycosidic linkage. We modified Glycam93 in such a way that the modified force-field can reproduce the HF potential energy of maltose evaluated as a function of ϕ and ψ. Although there is a limitation that the attractive van der Waals interaction cannot be taken into account by the HF level calculations, we expected that the HF result can give us some hint for improving Glycam in a systematic manner. The details of the modification are given in the Appendix. As already stated, the results included in Figures 1 and 3 are rather encouraging despite this limitation.

The modified Glycam93 was, however, not successful in reproducing experimentally observed value of C_∞ (Table 1). Thus the present modification is not adequate for modeling long chains. This problem is probably caused by the neglected contribution of the attractive van der Waals interaction. Surely further investigation and elaboration are needed in future.

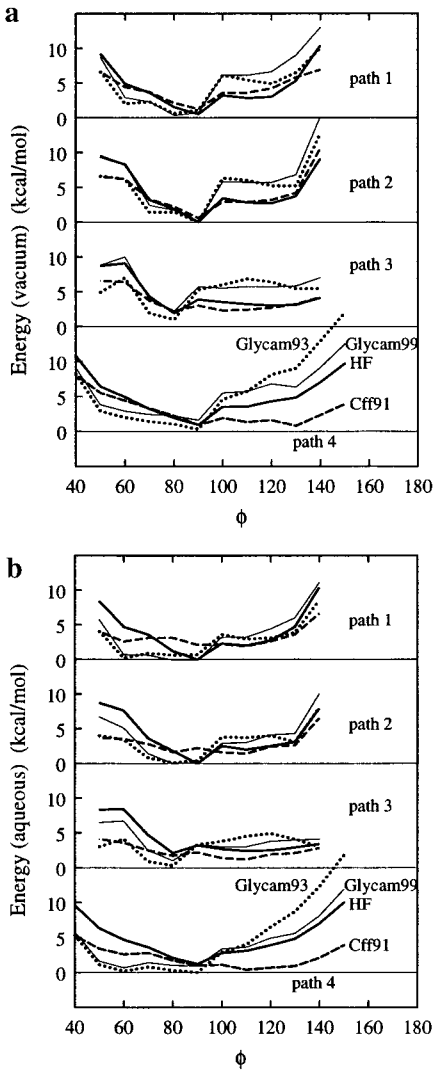

Figure 9. Detailed analysis of the energy profiles for α-maltose (G2) (a) in a vacuum and (b) in an aqueous environment. These profiles are along paths 1−4 illustrated in Figure 6. The solid, dashed, dotted, and thin solid curves represent the HF, Cff91, Glycam93, and Glycam99 values, respectively.

Another approach is the reparametrization of the force-field parameters in terms of the multiple-conformation RESP charge fitting method.[72] For glucose, the RESP charges to be used in the Amber force-field has already been determined by Simmerling et al.[17] Thus, we made MD simulations of G7 by using their parameter set. Unfortunately, however, we obtained too large value of $R_G = 8.8$ Å. This indicate that the simple use of the RESP charges is not sufficient for the correct description of amylose conformation, as far as their force-field is concerned. It seems interesting to examine whether the quality of Glycam can be improved by the reparametrization with the RESP charges, especially when the charges are tuned for maltose instead of glucose.

VI. Concluding Remarks

We have compared various force-fields on their ability to reproduce amylose conformation. It was found that Cff91 was most satisfactory in all respect but that others were not. In particular, Glycam93 showed too strong a preference for rotational isomeric state A (Figure 6). However, the performance of Glycam93 for many other saccharides should not be judged from the present result for amylose. Indeed, Glycam93 has already been successfully applied to various saccharides and has also been used in modeling a protein−saccharide complex.[36] Reproducing amylose conformation might be one of the most difficult tasks assigned to Glycam93. It was also found that this weak point of Glycam was appreciably improved by its new version Glycam99. We also have attempted to improve Glycam93 by adding an energy term dependent on a pair of glycosidic torsion angles in such a way that the resultant total potential energy fit to the HF result for maltose. Our modification is a tentative remedy, because it lacks transferability to other saccharide molecules. Both Glycam99 and our modGlycam93 were moderately successful in reproducing conformation of G7, but both failed to reproduce C_∞; the former gave $C_\infty = 10$, which is too large, and the latter gave $C_\infty = 2$, which is too low, compared with the experimental value of $C_\infty = 5$. Thus, development of a force-field that can overcome this shortcoming is quite beneficial.

TABLE 4: Root-Mean-Square Deviation (rmsd) of the MD Snapshots at 1 ns from the Crystal Structure of CD_{14}

force field	rmsd (Å)	force field	rmsd (Å)
Glycam93	4.958	Cff91	4.131
Glycam99	3.802	Cff	1.991
modified Glycam93	1.404	Gromos	3.431

To illustrate the significance of force-fields, we used it in a preliminary simulations of CD_{14}, whose crystallographic structure is characterized by two unusually inverted glycosidic linkages,[28−30] (ϕ,ψ) of which is in unusual domain E (Figure 6). Figure 10 shows the MD snapshots at 1 ns along with the crystal structure, the rms deviations from which are summarized in Table 4. A striking deviation from the initial saddlelike structure[28,30] is seen for some of the force-fields. If the quality of the force-field can be assessed in terms of those deviations, as is often done for proteins and nucleic acids,[26] modified Glycam93 is the best. In the case of amylose, however, we think the situation is not so simple that this conclusion can be drawn unconditionally. This is because amylose is much more flexible than proteins and can make conformational transitions easiliy. Indeed conformational flexibility must be taken into account to explain the fact that all glucoses of CD_{14} to CD_{26} in solution are found to be equivalent on the ^{13}C NMR time scale, whereas they are not equivalent in a crystal environment because of the occurrence of inverted glycosidic linkages.[30] Thus there remains a fundamental question of whether the deviation is a real one caused by conformational fluctuation or an artifact due to an inadequacy of the force-field. SAXS studies now in progress for CD_{10}, CD_{14}, and CD_{26}—whose crystal structures have already been resolved—should provide information to help us answer this question.

VII. Appendix

Modification of the Glycam93 Potential. As described in the Discussion section, the Hartree−Fock (HF) energies of α-maltose were evaluated as a function of ϕ and ψ at the grid points (with a spacing of 10°) between $40° \le \phi \le 160°$ and $-300° \ \psi \le 60°$, the two hydroxymethyl configuration being fixed to gg. To express the obtained potential in the B-spline form, we first calculate the difference

$$\Delta U(p,q) = U_{\text{glycam}} - U_{\text{HF}} \qquad \text{(A1)}$$

between the HF and Glycam energies at the grid points (p,q) corresponding to $\phi = 10p$ and $\psi = 10q$, p and q being integers. Energy values were set to zero at (90°, −130°), for convenience. Suppose that ϕ is between grid points p and $p + 1$, and that ψ

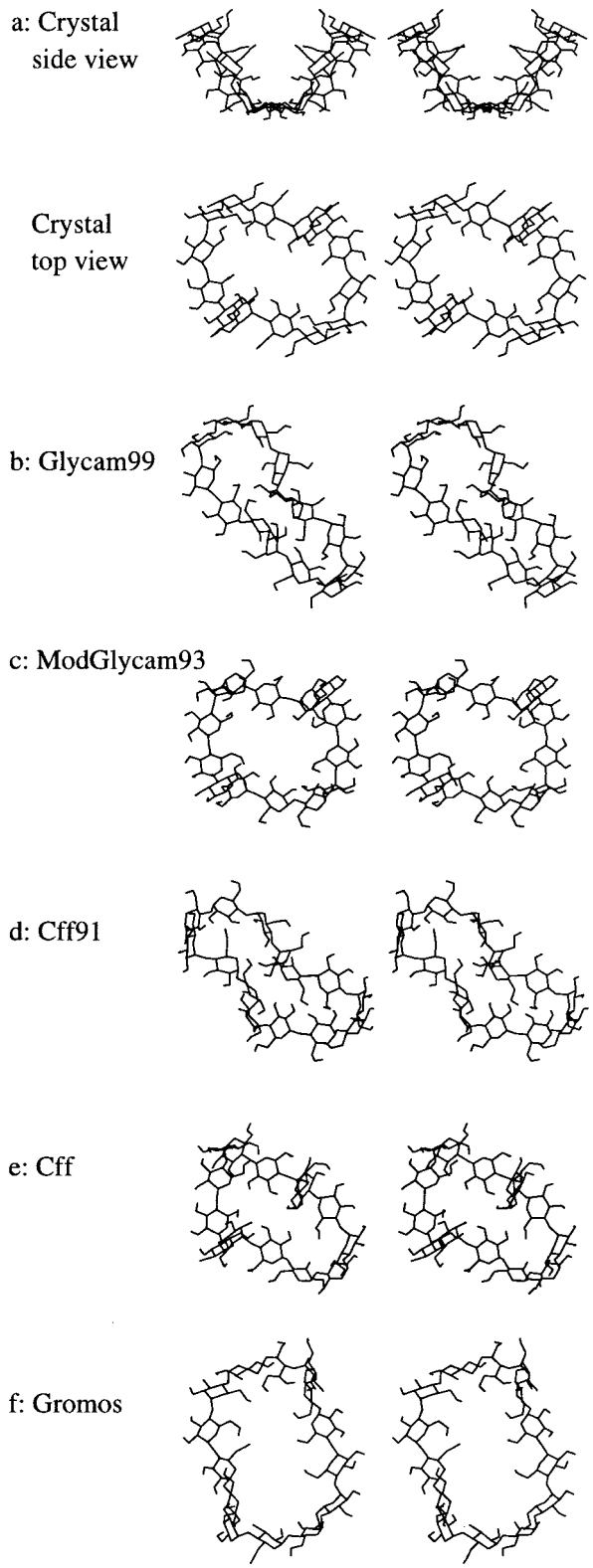

Figure 10. Stereoviews of the MD snapshots of CD_{14} at 1 ns.

is between q and $q + 1$. The modified potential at (ϕ,ψ) may be represented in terms of these differences as

$$U(\phi,\psi) = U_{\text{glycam}} - f \sum_{i=-1}^{2} \sum_{j=-1}^{2} F_i(u)F_j(v)\Delta U(p + i, q + j) \tag{A2}$$

where f is an adjustable parameter normally taken as unity, and

$$F_1(u) = (-u^3 + 3u^2 - 3u + 1)/6$$

$$F_2(u) = (3u^3 - 6u^2 + 4)/6$$

$$F_3(u) = (-3u^3 + 3u^2 + 3u + 1)/6 \tag{A3}$$

$$F_4(u) = u^3/6$$

Here u and v are parameters ranging from 0 to 1. In our case,

$$u = (\phi - 10p)/10$$

$$v = (\psi - 10q)/10 \tag{A4}$$

To smoothly join the modified and unmodified regions of ϕ, the ΔU values were damped to zero in the following way: $\Delta U = V_1/4$ and 0 for ϕ= 30° and 20°, respectively, where V_1 is the ΔU value at $\phi = 40°$; $\Delta U = V_2/2$, $V_2/4$, and 0 for $\phi = 170°$, 180°, and 190°, respectively, where V_2 is the ΔU value at $\phi = 160°$. Actual values of $\Delta U(p,q)$ are given in Table 5S.

Acknowledgment. The authors thank Dr. M. Mimura for performing the SAXS measurement and thank Dr. K. Gessler at the Free University Berlin for the coordinates of CD_{14}. The authors also thank Prof. R. J. Woods at the University of Georgia for drawing their attention to the new force-field Glycam99, and the two anonymous referees for many critical comments. This work was supported in part by a grant from the project "Development of bioreactor system for food industries (1997−2001)" of the Society for Techno-Innovation of Agriculture, Forestry, and Fisheries (STAFF), which is in turn financially supported by MAFF, Japan. Additional support was provided by the Ministry of Education, Science, and Culture, Japan, to S.K. (11896004). The synchrotron SAXS measurements were performed with authorization from the Photon Factory Advisory Committee (Proposal No. 94G293).

Supporting Information Available: Table 5S containing values used for modified Glycam93. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.
(2) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, UK, 1987.
(3) Brooks, C. L., III; Karplus, M.; Petitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; Wiley: New York, 1988.
(4) (a) van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947. (b) Hünenberger, P. H.; van Gunsteren, W. F. In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997; Volume 3, Chapter 1.
(5) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927.
(6) Glennon, T. M.; Zheng, Y.-J.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M., Jr. *J. Comput. Chem.* **1994**, *15*, 1019.
(7) Schmidt, R. K.; Teo, B.; Brady, J. W. *J. Phys. Chem.* **1995**, *99*, 11339.
(8) Ott, K.-H.; Meyer, B. *Carbohydr. Res.* **1996**, *281*, 11.
(9) Ott, K.-H.; Meyer, B. *J. Comput. Chem.* **1996**, *17*, 1068.
(10) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 197.
(11) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *J. Mol. Biol.* **1988**, *203*, 241.
(12) Lichtentahler, F. W.; Immel, S. *Tetrahedron: Asymmetry* **1994**, *5*, 2045.
(13) Lipkowitz, K. B. *Chem. Rev.* **1998**, *98*, 1829.
(14) Shimada, J.; Handa, S.; Kaneko, H.; Takada, T. *Macromolecules* **1996**, *29*, 6408.
(15) Kitamura, S.; Isuda, H.; Shimada, J.; Takada, T.; Takaha, T.; Okada, S.; Mimura, M.; Kajiwara, K. *Carbohydr. Res.* **1997**, *304*, 303.

(16) Takaha, T.; Yanase, M.; Takata, H.; Okada, S.; Smith, S. M. *J. Biol. Chem*. **1996**, *271*, 2902.

(17) Simmerling, C.; Fox, T.; Kollman, P. A. *J. Am. Chem. Soc*. **1998**, *120*, 5771.

(18) Gregurick, S. K.; Liu, J. H.-Y.; Brant, D. A.; Gerber, R. B. *J. Phys. Chem. B* **1999**, *103*, 3476.

(19) See for example: Cramer, C. J.; Truhlar, D. G.; French, A. D. *Carbohydr. Res.* **1997**, *298*, 1.

(20) Trommsdorff, U.; Tomka, I. *Macromolecules* **1995**, *28*, 6128.

(21) (a) Jordan, R. C.; Brant, D. A.; Cesàro, A. *Biopolymers* **1978**, *17*, 2617. (b) Brant, D. A.; Christ, M. D. In *Computer Modeling of Carbohydrate Molecules*; French, A. D. Brady, J. W., Eds.; American Chemical Society: Washington, DC, 1990; Chapter 4.

(22) (a) Kitamura, S.; Okamoto, T.; Nakata, Y.; Hayashi, T.; Kuge, T. *Biopolymers* **1987**, *26*, 537. (b) Nakata, Y.; Kitamura, S.; Takeo, K.; Norisuye, T. *Polym. J.* **1994**, *26*, 1085.

(23) Mimura, M.; Urakawa, H.; Kajiwara, K.; Kitamura, S.; Takeo, K. *Macromol. Symp.* **1995**, *99*, 43.

(24) Woods, R. J. *Curr. Opin. Struct. Biol.* **1995**, *5*, 591.

(25) Beachy, M. D.; Chasman, D.; Murphy, R. B.; Halgren, T. A.; Friesner, R. A. *J. Am. Chem. Soc*. **1997**, *119*, 5908.

(26) (a) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., Jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179. (b) Kollman, P.; Dixon, R.; Cornell, W.; Fox, T.; Chipot, C.; Pohorille, A. In *Computer Simulation of Biomolecular Systems, Theoretical and Experimental Applications*; van Gunsteren, W. F., Weiner, P. K., Wilkinson, A. J., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1997; Volume 3, pp 83−94.

(27) Pérez, S.; Imberty, A.; Engelsen, S. B.; Gruza, J.; Mazeau, K.; Jimenez-Barbero, J.; Poveda, A.; Espinosa, J.-F.; van Eyck, B. P.; Johnson, G.; French, A. D.; Kouwijzer, M. L. C. E.; Grootenuis, P. D. J.; Bernardi, A.; Raimondi, L.; Senderowitz, H.; Durier, V.; Vergoten, G.; Rasmussen, K. *Carbohydr. Res.* **1998**, *314*, 141.

(28) Saenger, W.; Jacob, J.; Gessler, K.; Steiner, T.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T. *Chem. Rev.* **1998**, *98*, 1787.

(29) Harata, K. *Chem. Rev.* **1998**, *98*, 1803.

(30) Jacob, J.; Gessler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha T.; Saenger, W. *Angew. Chem., Int. Ed.* **1998**, *37*, 606.

(31) Gessler, K.; Usón, I.; Takaha T.; Krauss, N.; Smith, S. M.; Okada, S.; Sheldrick, G. M.; Saenger, W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4246.

(32) (a) Senderowitz, H.; Still, W. C. *J. Org. Chem.* **1997**, *62*, 1427. (b) Kouwijzer, M. L. C. E.; van Eijck, B. P.; Kooijman, H.; Kroon, J. *Acta Crystallogr*. **1995**, *B51*, 209. (c) Reiling, S.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem.* **1996**, *17*, 450. (d) Grootenhuis, P. D. J.; Hassnoot, C. A. G. *Mol. Simul.* **1993**, *10*, 75. (e) Damm, W.; Frontera, A.; Tirado-Rives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955.

(33) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.

(34) Woods, R. J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 9, Chapter 3.

(35) (a) Maple, J. R.; Dinur, U.; Hagler, A. T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5350. (b) Hwang, M.-J.; Ni, X.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *Biopolymers* **1998**, *45*, 435.

(36) (a) Brisson, J.-R.; Uhrinova, S.; Woods, R. J.; van der Zwan, M.; Jarrell, H. C.; Paoletti, L. C.; Kasper, D. L.; Jennings, H. J. *Biochemistry* **1997**, *36*, 3278. (b) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Eur. J. Biochem*. **1998**, *258*, 372. (c) Peräkylä, M.; Kollman, P. A. *J. Am. Chem. Soc.* **1997**, *119*, 1189.

(37) (a) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982. (b) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

(38) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Wiley: New York, 1969.

(39) Berne, B. J.; Straub, E. *Curr. Opin. Struct. Biol.* **1997**, *7*, 181.

(40) Goldsmith, E.; Sprang, S.; Fletterick, R. *J. Mol. Biol.* **1982**, *156*, 411.

(41) There are some other conventions for measuring glycosidic torsion angles: If the IUPAC 1983 rule is obeyed, $\phi_I = \phi$, and $\psi_I = C(1)_n-O(4)_n-C(4)_{n-1}-C(3)_{n-1}$. In another rule, H atoms connected to C(1) and C(4) are used in measuring angles. $\phi_H = H(1)_n-C(1)_n-O(4)_{n-1}-C(4)_{n-1}$ and $\psi_H = C(1)_n-O(4)_{n-1}-C(4)_{n-1}-H(4)_{n-1}$. These angles are approximately related each other by $\phi_I = \phi$, $\psi_I = \psi - 240$, $\phi_H = \phi - 120$, $\psi_H = \psi + 120$. Also note that glucose units are serially numbered from the reducing end; i.e., in the direction from C(1) to C(4) of each unit.

(42) Marchessault, R. H.; Pérez, S. *Biopolymers* **1979**, *18*, 2369.

(43) Lichtenthaler, F. W.; Immel, S. *Liebig Ann.* **1996**, *1996*, 27.

(44) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S., Jr.; Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765.

(45) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(46) Greengard, L. *Science* **1994**, *265*, 909.

(47) Shimada, J.; Kaneko, H.; Takada, T. *J. Comput. Chem.* **1994**, *15*, 28.

(48) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, 1987.

(49) (a) Nose, S. *J. Chem. Phys.* **1984**, *81*, 511. (b) Hoover, W. G. *Phys. Rev. A.* **1985**, *31*, 1695.

(50) Yui, T.; Kitamura, S., unpublished work.

(51) French, A. D.; Rowland, R. S.; Allinger, N. L. In *Computer Modeling of Carbohydrate Molecules*; French, A. D.; Brady, J. W., Eds.; American Chemical Society: Washington, DC, 1990; Chapter 7.

(52) Gress, M. E.; Jeffrey, G. A. *Acta Crystallogr.* **1977**, *B33*, 2490.

(53) Brown, G. M.; Levi, H. A. *Acta Crystallogr.* **1978**, *B35*, 656.

(54) Arnott, S.; Scott, W. E. *J. Chem. Soc., Perkin Trans. 2* **1972**, *2*, 324.

(55) Liu, J. H.-Y.; Brant, D. A.; Kitamura, S.; Kajiwara, K.; Mimura, M. *Macromolecules* **1999**, *32*, 8611.

(56) Imberty, A.; Chanzy, H.; Pérez, S.; Buléon, A.; Tran, V. *J. Mol. Biol.* **1988**, *201*, 365.

(57) Rappenecker, G.; Zugenmaier, P. *Carbohydr. Res*. **1981**, *89*, 11.

(58) Yamakawa, H. *Modern Theory of Polymer Solutions*; Harper & Row: New York, 1971.

(59) (a) Nakanishi, Y.; Norisuye, T.; Teramoto, A.; Kitamura, S. *Macromolecules* **1993**, *26*, 4220. (b) Norisuye, T. *Macromol. Symp.* **1995**, *99*, 31. (c) Brant, D. A.; Dimpful, W. L. *Macromolecules* **1970**, *3*, 655. (d) Banks, W.; Greenwood, C. T. *Carbohydr. Res.* **1968**, *7*, 349.

(60) Lindner, K.; Saenger, W. *Acta Crystallogr*. **1982**, *B38*, 203.

(61) Betzel, C.; Saenger, W.; Hingerty, B. E.; Brown, G. M. *J. Am. Chem. Soc.* **1984**, *106*, 7545.

(62) Tran, V.; Buléon, A.; Imberty, A.; Pérez, S. *Biopolymers* **1989**, *28*, 679.

(63) Pérez, S.; Taravel, F.; Vergelati, C. *Nouv. J. Chim.* **1985**, *9*, 561.

(64) Hinrichs, W.; Büttner, G.; Steifa, M.; Betzel, Ch.; Zabel, V.; Pfannemüller, B.; Saenger, W. *Science* **1987**, *238*, 205.

(65) Hinrichs, W.; Saenger, W. *J. Am. Chem. Soc.* **1990**, *112*, 2789.

(66) Mikami, B.; Degano, M.; Hehre, E. J.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 7779.

(67) Takusagawa, F.; Jacobson, R. A. *Acta Crystallogr*. **1978**, *B34*, 213.

(68) Goebel, C. V.; Dimpfl, W. L.; Brant, D. A. *Macromolecules* **1970**, *3*, 644.

(69) *AMOSS*, Release 3.0; NEC Corp.; Tokyo, Japan, 1995.

(70) (a) Sharp, K. A.; Nicholls, A.; Fine, R. M.; Honig, B. *Science* **1991**, *252*, 106. (b) Sitkoff, D.; Sharp, K. A.; Honig, B. *J. Phys. Chem.* **1994**, *98*, 1978.

(71) Nicholls, A.; Sharp, K. A.; Honig, B. *Delphi*; Department of Biochemistry and Molecular Biophysics; Columbia University: New York, 1990.

(72) Bayly, C. I.; Cieplak, P.; Cornell, W. D.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269.