Conformation of Novel Cycloamylose: Topological Aspects and Simulations

Jiro Shimada,* Susumu Handa,† Hiroki Kaneko, and Toshikazu Takada

Fundamental Research Laboratories, NEC Corporation, Miyukigaoka 34, Tsukuba, Ibaraki 305, Japan

Received March 1, 1996; Revised Manuscript Received May 31, 1996[⊗]

ABSTRACT: It has recently been discovered that cycloamylose, cyclic α-1,4-glucan with a degree of polymerization ranging from 17 to several hundred, can be produced by the action of potato D-enzyme (EC 2.4.1.25) on amylose. To obtain structural insights into this new series of cycloamylose, the topological aspects of its circular structure were discussed in terms of a ribbon model properly introduced. A simple analysis of the chain-length dependence of the product population was performed on the basis of an elastic wire model. If the amylose chain is rather stiff, a double-helical structure with foldbacks is more reasonable than a circularized single-helical structure. This is because theoretical calculations for the latter predict a periodic population oscillation, which is caused by the need for the helix ends to meet, whereas such an oscillation was not experimentally observed. However, due to the unknown chain stiffness, it is difficult to draw a definite conclusion. To answer some questions that arose in the above analysis, atomistic models are necessary. Therefore, crude atomistic models were created in close analogy with the modeling of supercoiled DNA. The models were then refined by molecular dynamics simulation performed with a fast multipole method. Both structures formed a hollow tube with internal diameter of 5−6 Å. As far as the enthalpic contribution is concerned, the simulation results support the double-helical structure. The simulation results also indicated appreciable chain flexibility, which might play an important role in accommodating guest molecules.

1. Introduction

Recently, Takaha et al.[1] discovered that a novel series of amylose comprising 17 to several hundred glucose residues is produced by the action of potato D-enzyme on amylose. Although their structural information has not been obtained, there is strong biochemical evidence showing that these molecules are circular. This new series of cycloamylose represents a new family other than the α-, *β*-, and *γ*-cyclodextrins, which comprise only a small number (6−8) of glucose residues. Cycloamylose may have a cavity geometry different from that of these small cyclodextrins known for forming various inclusion complexes. Furthermore, cycloamylose is highly soluble in cold water, unlike linear amylose, which is insoluble.[1] For these reasons, cycloamylose may find novel applications in chemical, pharmaceutical, and food industries. It is anticipated that resolution of its 3-D structure will open up a new avenue for these interesting molecules that can accommodate a variety of guest molecules in its cavity. However, at present and probably in the near future, determining its structure is a difficult problem, as may be foreseen from the rather scarce structural information concerning carbohydrates compared with that concerning proteins. Therefore, we attempt to theoretically predict its plausible 3-D structure. Our prediction will help to explain its various properties and also to make practical applications.

However, despite recent progress in computational methods and computers, it is still difficult to construct molecular models entirely on the basis of an atomic model. This is partly due to the difficulty of making simulations long enough to allow many conformational transitions, and partly due to the lack of precision of the empirical force field used. Furthermore, because of the circularity condition imposed on the circular chain, its backbone torsion angles cannot freely rotate. As a consequence, it is far more difficult to model than linear chains,[2,3] and the conformation search and modeling methods, which are efficiently used for linear carbohydrates,[4] are no longer useful for circular chains. To overcome these difficulties to some extent, our strategy here involves crude modeling on the basis of an elastic wire model,[5−9] followed by its refinement by atomistic simulations on the basis of the so-called molecular mechanical model.[10−13] This is similar to the strategy adopted by Hao and Olson in developing an all-atom model of supercoiled DNA.[9]

Before proceeding, it is helpful to make a brief historical survey of the conformational studies on *linear* amylose. Early X-ray fiber diffraction studies, combined with model building, revealed that amylose complexed with small ions or molecules adopts a variety of left-handed helices; see, for example, the work of Goebel et al.[14] Among these, the most popular V-form helix is a tightly coiled helix characterized by a radius of $\rho = 4$ Å and a rise of $D_z = 1.4$ Å per residue.[15] In this paper, ρ refers to the helix formed by the glycosidic oxygens. Although it is difficult to determine a conformation of a single amylose chain, crystallographic analyses of maltoheptaose bound to phosphorylase suggested a loosely coiled left-handed helical structure.[16] Compared with the V-form helix, $D_z = 2.3$ Å is approximately doubled and $\rho = 5$ Å is somewhat larger. Most of the conformational energy calculations for maltose comprising two glucose residues revealed that the chain prefers a left-handed helix at a principal potential minimum.[17−29]

As for starch, glycogen, and various dextrins, Kainuma and French[30] pointed out the possibility of taking parallel and anti-parallel double helices from model building studies with a space-filling model. Subsequently, after some controversy on the helix chirality, an analysis by Imberty et al.[20,31] of various X-ray data for A- and B-type starches allowed structural determination of such parallel helices, which were found to be

* Fax: 81−298−56−6173. E-mail: shimada@cs.nis.nec.co.jp.
† Previously at NEC Informatec Systems, Ltd. Present address: Kyushu Institute of Technology, Iizuka 820, Japan.
⊗ Abstract published in *Advance ACS Abstracts,* August 15, 1996.

S0024-9297(96)00338-5 CCC: $12.00

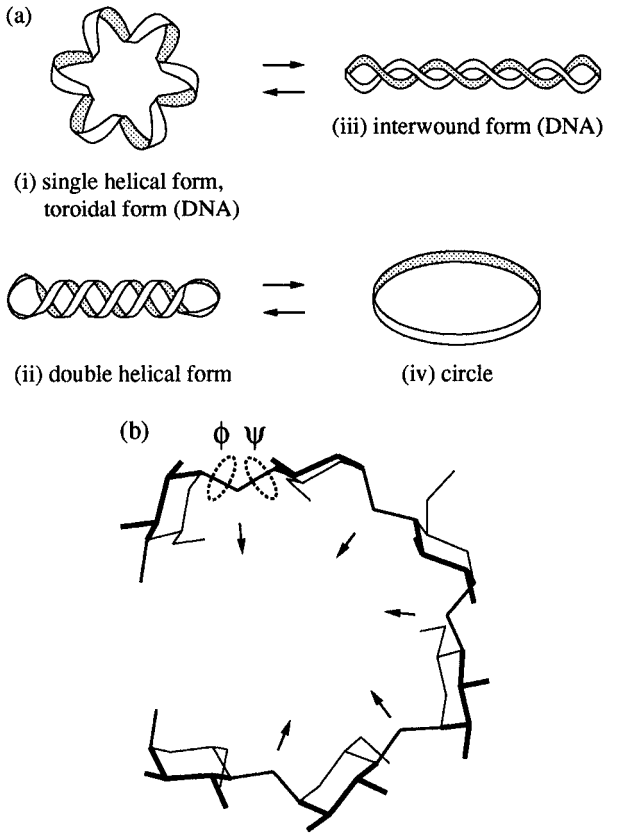

Figure 1. Ribbon representations of cycloamylose. The amylose chain may be modeled as a ribbon, provided that the V-shape of the glycosidic linkage is aligned with the normal of a ribbon plane. (a) When a ribbon end is twisted several times and then closed, a single helical form i of cycloamylose is created. By contrast, when a circular ribbon iv simply closed is wrapped around a cylinder surface, an anti-parallel double-helical form ii of cycloamylose is created, with foldbacks at each end. As is well known in DNA supercoiling, it is possible to convert ribbon i (toroidal form) to ribbon iii (interwound form). However, note that ribbons iii and iv are quite unlikely for cycloamylose. (b) Ribbon-like structure of amylose chains. The curvature vector (arrow) is aligned with the V-shape of the glycosidic linkage. Actually, the projection shown is a part of the crystal structure of β-cyclodextrin.

rather slender left-handed helices with $\rho \simeq 3$ Å and $D_z \simeq 3.5$ Å. Hinrichs et al.[32,33] observed an anti-parallel helix for *p*-nitrophenyl-α-maltohexaose$_2$ complexed with $Ba(I_3)_2$. Its detailed structure shows that it has a large helix radius of $\rho = 5$ Å, which results in a large central cavity with a 5 Å diameter, large enough to accommodate iodines. In addition to these experimentally derived structures, conformation energy calculations by Schulz et al.[34] predicted different forms of double helices.

From these studies, it is reasonable to assume that linear amylose adopts a single- or double-helical form in solution and in the solid state. Therefore, when the linear molecule is forced to cyclize, there are two possible outcomes: an anti-parallel double helix with foldbacks at each end, and a circularized single helix, as suggested by Takaha et al.[1]

Among the ribbons shown in Figure 1a, ribbons i and ii schematically represent the single- and double-helical forms, respectively. Ribbons i and ii are topologically different, because ribbon ii can be smoothly deformed to circular ribbon iv, which is topologically different from ribbon i. This means that the single-helical form i cannot converts into the double-helical form ii without breaking and rejoining the ribbon, which just corresponds to a 360° rotation of the chain axis on either side of the glycosidic linkage. (Actually, an $N \times 360°$ rotation is necessary, where $N = 6$ is the number of times the two strands are intertwined.) This 360° rotation is unlikely for cycloamylose molecules, at least for ones that are not very large.

It is important to note the following to avoid confusion with the situation encountered in DNA supercoiling. In the case of macroscopic rubber tubing, the cyclic single helix i easily converts into another anti-parallel double helix iii of a different type. This is well known as an interconversion from the toroidal form to the interwound form of closed circular DNA (see, for example, the textbook by Cantor and Schimmel[35]). However, this interwound form should clearly be distinguished from the double-helical form of cycloamylose under consideration, because they are topologically different. Indeed, there are substantial structural differences: ribbon ii is left-handed, while ribbon iii is right-handed; the ribbon plane of ii is smoothly wrapped around the enclosing cylinder surface, whereas the plane of iii is not, as can be easily confirmed by a paper ribbon experiment.

The ribbon of amylose, as opposed to that of DNA, may be characterized by its intrinsic tendency to coil, specifically in such a way that the V-shape of the glycosidic linkage C−O−C aligns with the curvature (Figure 1b). Hence, the introduced ribbon tends to curve considerably toward the shaded side. However, the circular form iv requires a small curvature, while the interwound form iii requires curving toward the unshaded as well as shaded sides. For these reasons, the circular and interwound forms are considered to be unlikely for cycloamylose above DP = 17, where DP designates the degree of polymerization.

Thus, the primary object of this paper is to theoretically determine whether structure i or ii is more likely. To do this, our theoretical development proceeds in two parts.

In the first part, the molecular weight distribution of the reaction product is theoretically calculated for the single- and double-helical structures on the basis of simple models: an elastic wire model called the helical wormlike model[5,6] is used for the former, and a two-state model is used for the latter. Comparison with experimental data supports the double-helical structure. However, this conclusion depends on the assumed stiffness parameter of the amylose chain. Moreover, a few questions arise on the underlying assumptions about the (simplified) use of the elastic wire model. Specifically, it is not certain whether such models are structurally possible, even down to DP = 17.

Therefore, in the second part, we attempt to build models in atomic detail.

First, crude models are created. For the double-helical form we assume an ideal anti-parallel double helix, except for the several terminal residues. To help atomic modeling of foldbacks, we again employ an elastic wire model based on a regular arrangement of glycosidic oxygens along a smoothly varying elastic string. In a manner similar to that suggested for closed circular DNA by Hao and Olson,[9] we represent the string as a B-spline curve and determine the conformation of the foldbacks by minimizing the elastic energy. The local elastic energy is assumed to be a minimum in the regular helical form. For the single-helical form, we make a single-helical pathway of several turns deformed around a large circle. Once the positions of

the oxygens are established, the rigid glucose units are then added.

Second, the obtained crude structures are refined by atomic energy minimization in a vacuum. Subsequently, molecular dynamics simulations are carried out in a water sphere or ellipsoid. Structures i and ii are compared with respect to their molecular mechanical energies.

2. Simple Analysis of the Product Population

Here we attempt to analyze the product population as a function of DP. If we assume that the product is formed simply by cyclization equilibria of linear amylose, its molecular weight distribution can be calculated by the standard theories of Jacobson and Stockmayer[36] and Flory and Semlyen.[37] For linear species, the weight concentration w_n of the species comprising n (=DP) units is described by

$$w_n = \text{const} \times np^{n-1} \qquad \text{(linear)} \qquad (1)$$

where p ($\simeq$1) is the fraction of reacted functional groups of the linear chains in equilibria. For circular species, this relation is modified to

$$w_n = \text{const} \times np^{n-1}G_n(0) \qquad \text{(circular)} \qquad (2)$$

where $G_n(0)$ is the ring-closure probability.

For the single-helical form, we can evaluate the ring-closure probability, taking account of both enthalpic and entropic contributions in the framework of the elastic wire model.[5,6,38] Because theoretical results shown later can easily be understood without going into mathematical detail, the evaluation is deferred to Appendix A.

For the double-helical form, we can make a semiquantitative analysis based on a simple two-state model. Suppose that $n - n_c$ glucose units of cycloamylose are in the helix (H) state and the remaining n_c units are in the loop (L) state. Then the total energy E may be expressed as

$$E = E_H(n - n_c) + E_L n_c \qquad (n \geq n_c) \qquad (3)$$

where E_H and E_L are the energies of the respective states relative to that of linear amylose. When the chain length is decreased below n_c, the chain comprises only energetically unstable loops, which are progressively strained. As a rough estimate, we may assume that the strain energy per glucose unit is proportional to the square of the loop curvature, which is inversely proportional to the loop size n. Its energy may then be expressed as

$$E = nE_L(n_c/n)^2 \qquad (n \leq n_c) \qquad (4)$$

This means that the energy increases sharply as n decreases, resulting in a low yield of the product below n_c. To simplify the final expression, we may expand the right of eq 4 in terms of $n_c - n$ as

$$\begin{aligned} E &= E_L n_c + E_L(n_c - n) + \cdots \\ &= [E_H(n - n_c) + E_L n_c] - (E_L + E_H)(n - n_c) + \cdots \end{aligned} \qquad (5)$$

Now, the ring-closure probability should be proportional to $\exp(-E/RT)$, where R is the gas constant and T is the absolute temperature, and we have assumed that the linear molecule already takes a double-helical form in a physiological environment. After some rearrangements, the weight fraction is calculated from eq 2 with the factor $G_n(0)$ redefined as

$$\begin{aligned} G_n(0) &= \text{const} && \text{for } n > n_c \\ &= \text{const} \times \exp\left[(n - n_c)\frac{\Delta E}{RT}\right] && \text{otherwise (double helical)} \end{aligned} \qquad (6)$$

where the factor $\exp(-nE_H/RT)$ showing chain-length dependence similar to p^n has been included in the latter. Assuming that the loop state corresponds to the planar or right-handed state, as found in various energy calculations so far carried out for maltose,[17-28] we may expect that $\Delta E = E_L + E_H$ is in the range of 0.5−2 kcal/mol (1 cal = 4.184 J).

Figure 2a shows the population of cycloamylose expected for these two forms, where the theoretical values for the single-helical form are calculated for the two estimated values 43 (more stiff) and 24 Å (more flexible) of the stiffness parameter λ^{-1} of the wormlike chain model[5,6,37] as described in Appendix A, and those for the double-helical form are calculated with parameters $n_c = 21$ and $\Delta E/RT = 1$, the same value of $p = 0.9$

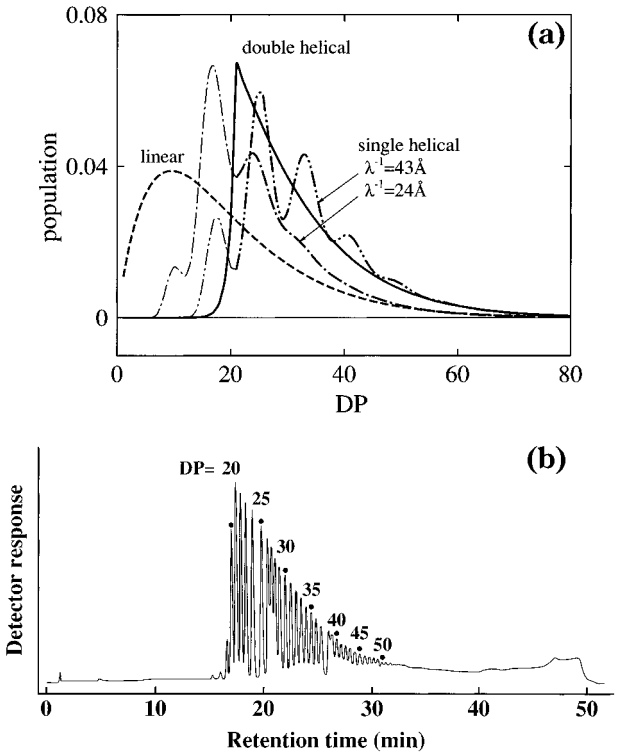

Figure 2. (a) Calculated weight fraction of cycloamylose products as a function of DP. The full curves represent the theoretical values for the double-helical form calculated from eq 2 with 3, while the oscillatory curves represent those for the single-helical forms calculated from eq 2 with Appendix A, for the indicated stiffness parameter $\lambda^{-1} = 43$ (more stiff) and 24 Å (more flexible) of the wormlike chain. See also the last two paragraphs of Appendix A. The values indicated by the thin curves might be diminished by an incorporation of the excluded volume effect (no calculation). The dashed curve represents the population for linear species. (b) Observed population of cycloamylose by experiment (courtesy of T. Takaha). Cycloamylose was prepared as described by Takaha et al.,[1] except that synthetic amylose AS-10 (average molecular weight, 10 000) was used as a substrate. Cycloamylose was separated by high-performance anion-exchange chromatography with a pulsed amperometric detector.[1] The numbers above the peaks indicate the DP of the products.

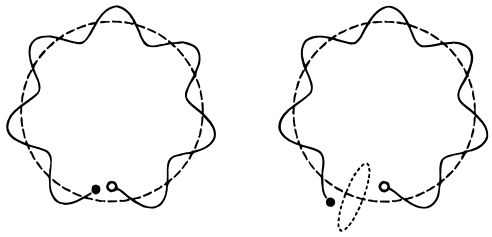

Figure 3. Cyclization of linear amylose with (a) DP = $7n_0$ and (b) DP = $6.5n_0$, where n_0 is the helix repeat (number of glucose units in the helix turn). If the chain length is an integral multiple of n_0, as in (a), the initial and final ends can easily meet. By contrast, if the chain length is smaller (or longer) by half a helix turn, either end must be twisted around its helix axis.

being used throughout. For comparison, the product population expected for linear species is also included. These theoretical values should be compared with the experimental data shown in Figure 2b (courtesy of T. Takaha). Both theoretical values can reproduce a low yield for small DP. For the respective form, this may be explained in terms of the large deformation energy necessary to bend a molecule into compact toroidal form, or the inability to create two terminal foldbacks progressively strained when DP is decreased below 20.

The most important feature expected for the single-helical form is the marked oscillation of the population with a period of the helix repeat n_0 (i.e., the number of residues per turn) assumed to be 8 in the present calculations. This oscillation is caused by the need for the twisted chain ends to meet at ring-closure.[38−51] The reason for this oscillation is schematically explained in Figure 3. Actually, such an oscillation was observed for DNA with a period of about 10.5 base pairs.[3] However, it has not been observed in the present fully equilibrated sample of cycloamylose. In contrast, the experimental data agree rather well with the semiquantitative values calculated for the double-helical form. Moreover, (over- and underwound) isomers differing only in the topological linking number were not observed for any DP. Thus, we consider that the population profile supports the double-helical form as a conformation of cycloamylose.

However, there remain certain points which may favor the single-helical form. First, the observed occasional gaps of the elution times in Figure 2b might be related to the transition from the over- to underwound conformation as DP is increased. Second, incorporation of the large excluded volume effect (steric repulsion as shown below in Figure 6 for DP = 18), which is not fully taken into account in the present treatment, might significantly diminish the theoretical values below DP ≃ 20. If so, experimental data can also be explained by assuming the single-helical form with a more flexible estimate of $\lambda^{-1} = 24$ Å, for which the oscillation is fairly damped above DP ≃ 20. Thus, further information on the chain stiffness in aqueous solution is required to draw a definite conclusion.

3. Crude Modeling of Cycloamylose

Necessity of an Atomistic Model. Although we have considered the problem on the basis of the coarse-grained models, a few questions naturally arise. First, is it possible to build realistic atomistic models, even down to DP = 17, for which cycloamylose is observed? Are they reasonable from an energetical point of view? Second, in everyday life, we have experienced that an isotropic toroidal rubber tube spontaneously changes its shape to the interwound form. Indeed, as far as closed circular DNA is concerned, extensive studies support the interwound form (with branching in certain cases), as recently remarked by Vologodskii and Cozzarelli in their review.[42] Although the interwound form itself is unlikely for cycloamylose, does that molecule retain its initial toroidal shape? If it undergoes some large structural deformation, the simple picture illustrated in Figure 3 and the reasoning based on the absence of the periodic oscillation might not be justified. To answer these questions, it is necessary to build an atomistic model.

The amylose chain trajectory has often been described in terms of virtual bonds connecting consecutive glycosidic oxygens, because its repeating unit is rather complex in structure when compared with, for example, that of the polyethylene chain. Indeed, snapshots of Monte Carlo calculations for linear amylose indicate that these oxygens form a smooth space curve.[17,43,44] Specifically, the chains in a crystal environment adopt smooth left-handed helices or circular shapes (β-cyclodextrin, etc.), see Figure 1b. These observations naturally lead us to describe the chain in terms of a continuous space curve. Thus, we proceed to theoretically develop a crude modeling procedure, representing amylose again as an elastic wire or string. This is followed by affixing glucose units to the wire.

Modeling the Double-Helical Form. We represent the amylose molecule as a regular arrangement of glycosidic oxygens along a smoothly varying elastic wire. These oxygens are serially numbered 1, 2, ..., n (n = DP), each oxygen being denoted by i. We assume that the double-helical form is composed of two parts: a regular double helix at the center and variable terminal loops.

For concreteness, suppose that the regular antiparallel structure at the center is maintained in the range

$$i_1 \le i \le i_2 \qquad \text{(upward strand)}$$

$$i_3 \le i \le i_4 \qquad \text{(downward strand)} \tag{7}$$

while a hairpin loop of unknown structure is localized to n_t units at the top and n_b units at the bottom ($n_t = i_3 - i_2$; $n_b = i_1 + n - i_4$), as shown in Figure 4 for DP = 24. Exact values of n_t and n_b are not known. However, considering the fact that the smallest cycloamylose produced in the reaction has DP = 17, we may assume that n_t, $n_b \simeq 8$.

B-Spline Fitting. To build loop structures, we make use of curve fitting methods as proposed by Hao and Olson[9] for closed circular DNA. Following their method, the space curve is approximated by fourth-order cyclic B-spline polynomials, which smoothly connect a sequence of controlling points, P_1, P_2, ..., P_n. The ith segment of the space curve is calculated from

$$\mathbf{r}_i(u) = F_1(u)\mathbf{P}_{i-1} + F_2(u)\mathbf{P}_i + F_3(u)\mathbf{P}_{i+1} + F_4(u)\mathbf{P}_{i+2} \tag{8}$$

where u is the chain contour parameter ranging from 0 to 1, $\mathbf{P}_j$ represents the position of P_j, and $F_l(u)$ are cubic polynomials defined by

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

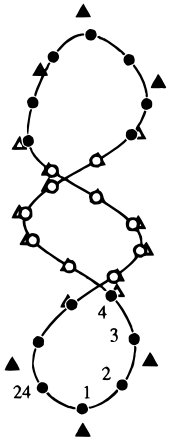

Figure 4. Projection of the double-helical cycloamylose with DP = 24, obtained by crude modeling. The full curve represents the B-spline curve generated from the controlling points indicated by the triangles, and the circles affixed to the curve represent the glycosidic oxygens. Here, the open symbols correspond to the fixed anti-parallel regular helix, while the filled symbols correspond to the variable loops. Each terminal loop consists of eight residues; the top loop is between glycosidic oxygens $i_2 = 9$ to $i_3 = 17$, while the bottom loop is between $i_4 = 21$ to $i_1 = 5$.

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

$$F_3(u) = (-3u^3 + 3u^2 + 3u + 1)/6$$

$$F_4(u) = u^3/6 \tag{9}$$

If a sufficiently large number of controlling points are used, as in standard numerical interpolation, $\boldsymbol{r}_i(u)$ smoothly interpolates from P_i to P_{i+1}. However, because there are only a few controlling points in the present case, we can only state that $\boldsymbol{r}_i(u)$ is located in the vicinity of P_i and P_{i+1}.

For the anti-parallel regular helical part fixed during crude modeling, each controlling point (x_j,y_j,z_j) may be identified with the glycosidic oxygen itself:

$$x_j = \tilde{\rho}\cos(j\Omega - \phi_0)$$

$$y_j = \tilde{\rho}\sin(j\Omega - \phi_0)$$

$$z_j = z_0 + jD_z \tag{10}$$

where Ω and D_z are the rotation angle and axial rise, respectively, per residue. The coordinates ϕ_0 and z_0 may be arbitrary chosen to specify the initial position (x_0, y_0, z_0) at $j = 0$. If $\tilde{\rho}$ is related to the helix radius ρ by the relation

$$\tilde{\rho} = \frac{3}{2 + \cos\Omega}\rho \tag{11}$$

then all controlling points (except those at the boundary between the loop and helix) can be shown to be located on the helix of radius ρ and pitch $h = D_z n_0$.

It is worthwhile to note that Zhang et al.[45] recently developed a more systematic method for creating constrained DNA loops (protein-bound supercoiled DNA) by using Bézier functions combined with Fourier series augmentation.

Elastic Energy. For the loop parts of unknown configuration, we may determine their controlling points from a minimization condition of the elastic energy for a continuous chain of total contour length L. In this paper, considering numerical tractability within the framework of B-spline fitting, we use for its elastic energy a simplified form,

$$U = \frac{k_B T}{4\lambda}\int_0^L [(\kappa(s) - \kappa_0)^2 + (1 + \sigma)^{-1}(\tau(s) - \tau_0)^2]\,ds \tag{12}$$

Here, $\kappa(s)$ and $\tau(s)$ are the differential geometrical curvature and torsion at contour point s of the space curve, respectively, λ^{-1} is a stiffness parameter, σ is the Poisson ratio, and $k_B T$ has its usual meaning. At the minimum of energy, the chain takes a regular helix form with constant curvature κ_0 and torsion τ_0, which may be determined from the helix radius ρ and pitch h of the assumed helix by solving the equations

$$\rho = \kappa_0/(\kappa_0^2 + \tau_0^2) \tag{13}$$

$$h = 2\pi\tau_0/(\kappa_0^2 + \tau_0^2) \tag{14}$$

In this way, the intrinsic tendency of amylose to preferentially coil left-handed is taken into account.

It is important to note that a more appropriate expression for the elastic energy is

$$U = \frac{k_B T}{4\lambda}\int_0^L [\omega_\xi(s)^2 + (\omega_\eta(s) - \kappa_0)^2 + (1 + \sigma)^{-1}(\omega_\zeta(s) - \tau_0)^2]\,ds \quad \text{(helical wormlike model)} \tag{15}$$

In eq 15, $\omega(s) = (\omega_\xi,\omega_\eta,\omega_\zeta)$ is an infinitesimal rotation vector describing the deformation of the isotropic wire at contour point s. This is the energy expression assumed in the helical wormlike model,[5,7] whose statistical and dynamical properties have already been extensively studied.[5,6] As discussed by Yamakawa and Shimada,[7] eq 12 corresponds to a sheetlike chain with $\omega_\xi(s) = 0$ (rather than an isotropic wire), for which certain deformations are not allowed. In this sense, eq 12 is physically unnatural. However, this defect will not be significant, as long as we confine ourselves to a crude search of the loop conformation of cycloamylose (but not of DNA).

In principle, loop configuration at the minimum of *rigorous* elastic energy may be determined by solving the Euler−Lagrange equation expressed in terms of Euler angles with certain boundary conditions. It is remarkable that such a problem was numerically solved by Westcott et al.,[46] who recently determined the loop conformation of DNA composed of naturally curved and straight segments.

Although contour length is a constant in the elastic wire model, it fluctuates in B-spline fitting. To adjust the contour length L_b of the loop to a given value L_b^0, we added to the potential energy a penalty function of the form

$$U_{penal} = {}^1/_2 k_B T\gamma(L_b - L_b^0)^2 \tag{16}$$

The factor γ is a (dimensionless) force constant to control the magnitude of the fluctuation of the contour length. There is no unique way to estimate L_b^0. However, for the ideal anti-parallel double helix, contour length L can be expressed in terms of the number n of the glucose units and virtual bond length $l_u = 4.5$ Å as

$$L = 1.02nl_u \quad \text{(ideal anti−parallel double helix)} \tag{17}$$

We have assumed, for convenience, that this relation also holds for a loop composed of n_b units:

$$L_b^0 = 1.02n_b l_u \tag{18}$$

The excluded volume effect between glucose units is taken into account by assuming that nonadjacent glycosidic oxygens interact through repulsive potential of the form $\epsilon(\sigma/r)^{12}$, with $\epsilon = 0.6$ kcal/mol and $\sigma = 4$ Å.

Computational Details of Modeling the Double-Helical Form. In our actual calculations, we assumed an ideal anti-parallel double helix, except for several terminal residues. We built each loop from five controlling points, the two points P_1 and P_5 being on the regular helix given by eq 10, and the remaining points P_2-P_4 being free.

We carried out 1000−3000 steps of Metropolis Monte Carlo calculations for this model by assigning $\lambda^{-1} = 140$ Å, $\sigma = 0$, and $\gamma^{-1} = 0.02n$. We note that the estimate of λ^{-1} is from Fujii et al.,[43] and the value of γ has been chosen so that rms fluctuation of each virtual bond length is suppressed to about 0.1 Å. Differential geometrical properties, such as L, κ, and τ, were numerically calculated by dividing the continuous chain into 45 segments per glucose unit. Figure 5 shows the elastic energy plotted against Monte Carlo cycles for cycloamylose with DP = 24, created with the geometrical parameters specified by code b in Table 1. The energy rapidly decreases within the first 50 steps and may be regarded as reaching equilibrium beyond 600 steps. Figure 4 shows a projection of a snapshot of the simulation for DP = 24. The final snapshot has been used as an initial guess for later calculations.

Modeling the Single-Helical Form. The toroidal pathway of a closed circular superhelix of radius ρ and k turns, deformed around a large circle of radius R, is given by[9]

$$\mathbf{r} = R[\cos(2\pi u)\mathbf{e}_x + \sin(2\pi u)\mathbf{e}_y] + \rho\{\cos(2\pi ku)\mathbf{e}_z + \sin(2\pi ku)[\cos(2\pi u)\mathbf{e}_x + \sin(2\pi u)\mathbf{e}_y]\} \tag{19}$$

As u varies from 0 to 1, the chain makes k toroidal revolutions. Actually, the number k may be chosen as an integer nearest the ratio of DP to the helix repeat n_0, which we assume as 6 or 8. Note that the following relation holds:

$$R = nD_z/2\pi \tag{20}$$

where R is the radius of the central axis of the formed doughnut.

Affixing Glucose Units to the Wire. After thus establishing the position of the glycosidic oxygens, our next task is to affix the glucose units to the continuous chain. By inspecting the 3-D structures of the left-handed helix of linear amylose and α- and β-cyclodextrins, we found that the V-shaped glycosidic linkage almost aligned with the curvature of the chain contour, as already noted in Figure 1b. Therefore, we added the rigid glucose units in such a way that the V-shape coincides with the curvature vector. Each glucose unit was assumed to have $C_4-C_5-C_6-O_6$ in the gauche conformation ($\phi = 60°$). Slight mismatches in the chain ends were equally partitioned into each glucose unit by adjusting the virtual bond length around 4.5 Å.

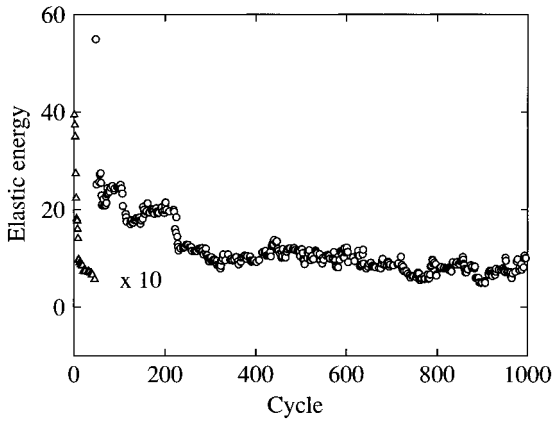

Figure 5. Elastic energy of double-helical cycloamylose with DP = 24, plotted against Monte Carlo cycles. The initial trajectory is generated with parameter code b given in Table 1. The values indicated by the triangles immediately after the start of the simulation should be understood to be multiplied by a factor of 10.

Table 1. Geometrical Parameters Used for Generating Various Initial Configurations

code	form	radius (Å)	rise (Å/res)	repeat (res)	loop size (res)
a	double	5.185	2.33	8	6
b	double	5.185	2.33	8	8
c	double	5.185	2.33	8	10
d	double	2.989	3.43	6.2	6
e	double	2.989	3.43	6.2	8
f	double	2.989	3.43	6.2	10
g	single	5.185	2.33	8	
h	single	2.989	3.43	6	

4. Molecular Dynamics Simulation Procedure

Energy Calculations in a Vacuum. To examine the effect of the force field used, energy calculations were carried out in a vacuum, with two types of interatomic potentials: a GROMOS field,[47,48] with a simple distance-dependent dielectric constant, and an all-atom MM2 (77) field,[49] with effective dielectric constants of 4, 80, and so on. A brief interface between the molecular dynamics and MM2 programs was written by the author. In all molecular dynamics simulations, equations of motion have been integrated with a central difference algorithm, and the temperature has been controlled by the Nosé−Hoover method,[50] with a characteristic time of 0.1 ps. When the GROMOS field was employed, all bond lengths were constrained by applying the SHAKE procedure,[50] with a tolerance of 10^{-10}. For the GROMOS calculations, total simulation time was 100 ps, with a time step of 1 fs. For the MM2 calculations (without SHAKE), the simulation time was 25 ps, with a time step of 1 fs, and the mass of all the hydrogen atoms was increased to 3, which serves to make the time step sufficiently small compared with the fastest motion involving those atoms, without affecting the equilibrium distribution of conformations. The second half of the simulation was used for sampling.

Molecular Dynamics Simulations in Water. An ellipsoid sufficiently large to enclose the cycloamylose molecule was determined in such a way that the surrounding water shell was at least 7 Å in width. The cycloamylose molecule was then immersed into a bulk of waters of ellipsoidal shape, and the overlapping waters were removed. A few water molecules were found to remain within the cavity formed by cycloamylose. The system we adopted for DP = 48, 24, and 18 contained 1975, 1086, and 1513 water molecules, respectively, for the double-helical form and 2870, 1526, and 1513 for the single-helical form.

After minimization to eliminate bad contacts between waters and cycloamylose, simulations of the whole system were

carried out at 300 K. During the first 5 ps, each glycosidic oxygen was constrained to its initial position. When a water molecule left the ellipsoid by more than 1.5 Å, it was pulled back by applying a weak boundary force directed toward the ellipsoid center. This boundary force, which prevents waters from escaping, was calculated as described in Appendix B, with an increased water width $w = 7 + 1.5$ Å. Total simulation times (except the first 5 ps) for DP = 48, 24, and 18 were 125, 200, and 150 ps, respectively, for the double-helical form; they were 75, 200, and 50 ps for the single-helical form.

For DP = 24, after equilibration for 200 ps in the ellipsoid, we added waters until cycloamylose was contained in a large sphere of radius 28.2 Å (2960 waters) and continued the simulation for 175 ps, and then we used the results from the last 100 ps for analysis.

Our calculations employed new programs that incorporate a new fast multipole method[51,52] to accurately calculate long-range electrostatic interactions. Specifically, this method can suppress the rms error in the electric potential felt by each atom below 0.1 kcal/mol. The GROMOS parameters[47,48] and the TIP3P water model of the OPLS field[53] were used, their van der Waals parameters being merged by using a combination rule adopted by the latter field. A time step of 2 fs was employed. Simulations of linear amylose in a water cube (953 waters) were carried out in a similar manner but under the periodic boundary condition. The temperature and pressure were controlled to 300 K and 1 bar, respectively, by the use of the Nosé−Hoover−Andersen method,[50] with characteristic times of 0.1 (temperature) and 2 ps (pressure).

5. Results and Discussion

In this section, we consider cycloamylose with DP = 18, 24, and 48. This set has been selected from the following considerations: cycloamylose is only marginally detected when DP = 18; it is highly populated and stained with iodine when DP = 48; and when DP = 24, it is also highly populated but cannot form a complex with iodine.

Table 1 summarizes the geometrical parameters used for generating various initial starting models. The double-helical forms created with the parameters specified by codes a−c have plump structures with a larger helix radius of 5.2 Å, while those specified by d−f have slimmer structures with a smaller radius of 3 Å, each structure differing in the size of the terminal loop. Single-helical forms are also built with a larger (g) and smaller (h) helix radius.

Among the parameter codes a−h, we anticipated that codes b and g would be most probable. This is based on the facts that their helix parameters correspond to the structures found in the crystals, and that the loop size of eight residues is nearly half the minimum DP of the 17 experimentally observed. We did not attempt any model building with the V-form helix[15] because it is too tightly coiled to take a toroidal form for small DP.

Simulation in a Vacuum. We have generated various initial configurations and then minimized them to their nearest local minimum, the minimized structures being graphically almost indistinguishable from the starting structures. Figure 6a shows the minimized configurations generated with the most probable parameter codes b and g. It is seen that both structures form a hollow tube with internal diameter of 5−6 Å, which may accommodate ions and small guest molecules. The opening at each end of the double-helical form is wide enough to allow passage of iodines whose van der Waals radius is 2.35 Å. Although the central part of the single-helical form is rather crowded for DP < 24, steric hindrance is not so severe as to inhibit its formation. The minimized conformational energy is appreciably lower for the double-helical form than for the corresponding single-helical one: 10.8 *vs* 15.1 for DP = 48; 13.1 *vs* 13.7 for DP = 24; and 15.3 *vs* 15.3 for DP = 18, where the energy per glucose unit is expressed in kcal/mol.

However, such minimization is highly probable to fall into the nearest local minimum, not to a true minimum. Hence, we made additional MD calculations at 300 K, starting from several initial configurations, although only for DP = 24. To examine the influence of the force fields used, we also made calculations with the MM2 field,[49] with effective dielectric constants of 4 (crystal environment) and 80 (water).

Figure 7 shows the initial configurations, and Figure 8 shows the corresponding stereoviews of the snapshots of the simulation at 25 (MM2) and 100 ps (GROMOS), respectively. In the latter figure, the progression of only the glycosidic oxygens is traced, for the sake of clarity. Note that each configuration (except g and h) is a best-fit trajectory to the corresponding one depicted in Figure 7. The mean potential energy calculated over the second half of the simulation time is summarized in Table 2.

Because configurations a−f belong to the same topological class in the sense of ribbon modeling, it is expected that sufficiently long simulations would lead to an essentially similar configuration. However, the obtained final snapshots are fairly dependent on the initial configuration, indicating that our simulation time is still insufficient. The situation was not appreciably improved, even when some of the simulations were continued further.

However, it is possible to draw a general tendency from the results. First, slender initial configurations (d−f) have higher energies than the corresponding plumper ones (a−c), and the slender initial configurations themselves gradually change their shapes to more plumper ones. Second, in all force fields, the initial single-helical forms are generally not stable and spontaneously take rather different configurations. Specifically, the MM2 field with $\epsilon = 4$ brings the configuration, within about 10 ps, to a rather collapsed form, which may be regarded as a compromise between the instability of the toroidal form and the inaccessibility to the interwound form. At any rate, the configuration reached from the single-helical form is energetically disfavored. Calculated conformational energy shows that configuration b is the most stable irrespective of the force fields used, in accord with our expectations. Hence, we consider that code b is the most probable candidate for the cycloamylose configuration.

Distribution of Glycosidic Torsional Angles in a Vacuum. The major factor that determines the amylose chain conformation is the distribution of torsional angles ϕ and ψ associated with the glycosidic linkage connecting two adjacent glucose units (Figure 1b). More specifically, ϕ and ψ are the angles for $O_{5'}-C_{1'}-O_4-C_4$ and $C_{1'}-O_4-C_4-C_5$, respectively. In most of the experimental and theoretical studies carried out for amylose chains, (ϕ,ψ) falls in a broad low-energy domain centered at (90°,−150°), where all torsional angles are measured in degrees with the cis rule (Figures 9 and 10).

However, during the simulations, we observed that some of the pairs (ϕ,ψ) were trapped into an unusual range around (90°,60°), at which the energy is higher than the minimum by 3−4 kcal/mol. The number of such unusual pairs of torsion angles is also included in Table 2. The table shows that configurations with such

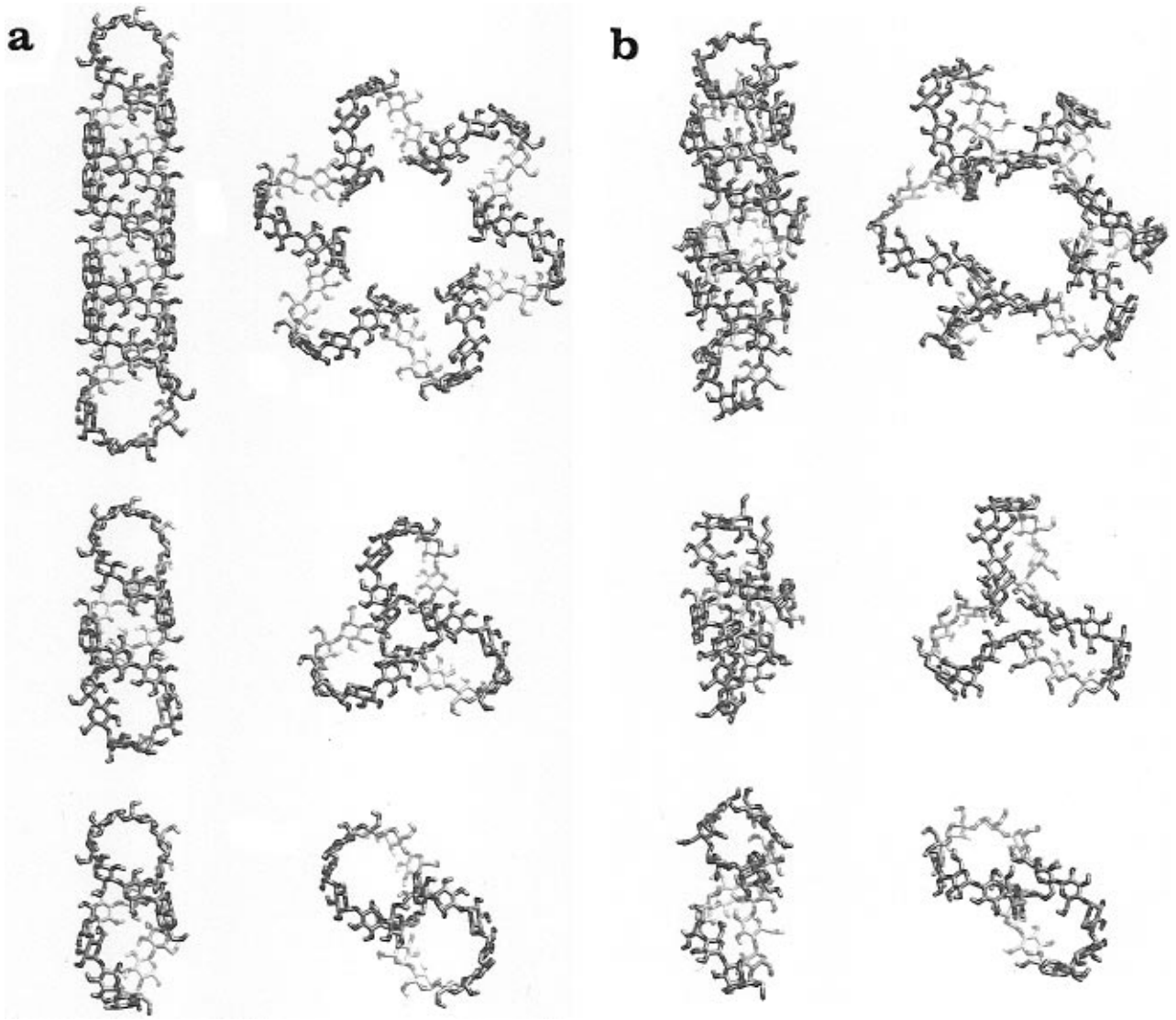

Figure 6. (a) Conformation of cycloamylose after local energy minimization in a vacuum. Shown on the left are the double-helical forms with DP = 48 (top), 24 (middle), and 18 (bottom), while the corresponding single-helical forms are on the right. All conformations were generated with parameter codes b and g, as given in Table 1. Actual molecular size may be measured by noting that the center-to-center separation between the conformations with the same DP is 40 Å along the horizontal direction. C, gray; O, red; H, blue. The figure was drawn using VISCAP (NEC Corp). (b) Snapshots of cycloamylose at the end of the molecular dynamics simulations in water at 300 K. Each snapshot should be compared with the corresponding one depicted in (a). Total simulation times are 125, 375, and 150 ps (from top to bottom) for the double-helical form and 75, 375, and 50 ps for the single-helical form.

(ϕ,ψ) generally lead to higher energy. Simulated annealings from 500 K were also carried out, but they were unsuccessful in reducing either the energy or number of unusual pairs. The failure indicates that annealing in the present situation requires additional structural constraints (atomic distances, etc.) to augment the precision of the empirical force fields used. To our knowledge, only one reported example of an unusual (ϕ, ψ) is the pair (153°,100°) observed by Mikami et al.[54] for maltotetraose (DP = 4) bound to β-amylase. The occurrence of such an unusual pair cannot be ruled out, but it is probably related to the activated state during the scission of the amylose chain.

Simulation in Water. It is now widely accepted that simulations in a vacuum should be used only in an exploratory manner to obtain qualitative results.[13] Although we consider that the coincidence of the conformational stabilities derived from three different force fields represents something of a reality, simulations in water are clearly more realistic. We thus proceeded to make simulations of cycloamylose in water, although only for some selected codes (b and g), because of the still large computational burden. Figure 6b shows the snapshots obtained at the end of the 50−375 ps simulations carried out in water at 300 K. Note that each configuration is a best-fit to the corresponding one depicted in Figure 6a. It is seen that the essential features of the double- and single-helical structures are maintained during the simulation. The stability of the single-helical form sharply contrasts with the results in a vacuum, indicating the important role of water molecules, as already noted by Eisenhaber and Schulz,[55] in interpreting conformational stability of various types of double-helical amylose.

It is known that a certain cutoff treatment of the electrostatic interaction causes artifacts in various properties, as demonstrated by Saito.[56] In particular, artificial instability or stability of the α-helix of polypeptides was observed by Schreiber and Steinhauser.[57] However, we have accurately calculated the electrostatic interaction by the use of the fast multipole method,[51] implemented as described in our previous papers,[52] so

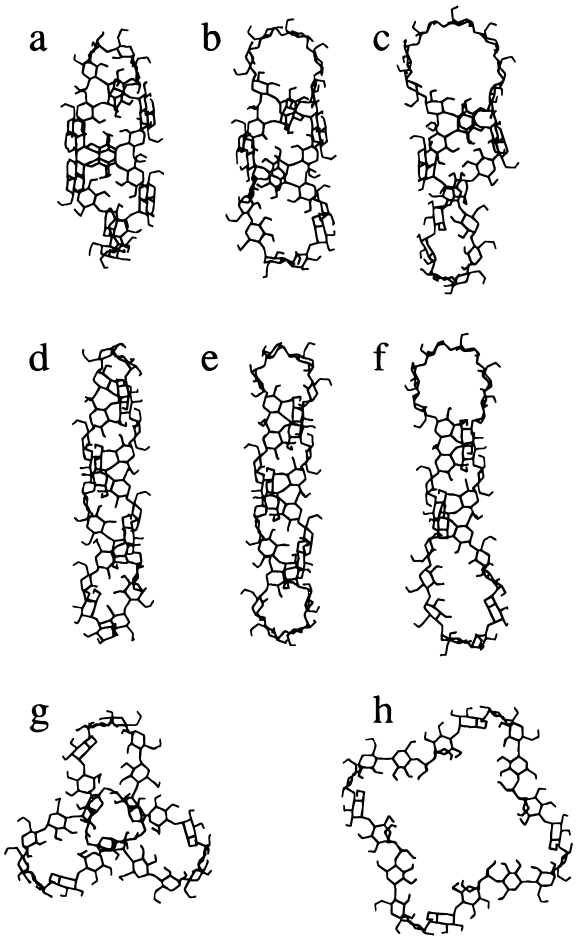

Figure 7. Various initial conformations of cycloamylose with DP = 24. Conformations a−h are generated with the corresponding parameter codes given in Table 1. Conformations a and b are separated by 25 Å along the horizontal direction.

Table 2. Mean Potential Energy of Cycloamylose with DP = 24 at 300 K, Obtained by Molecular Dynamics Simulations Starting from Various Initial Configurations a−h

	mean energy (kcal/mol)			no. of unusual torsions		
code	GROMOS	MM2 ($\epsilon = 4$)	MM2 ($\epsilon = 80$)	GROMOS	MM2 ($\epsilon = 4$)	MM2 ($\epsilon = 80$)
			Cycloamylose in a Vacuum			
a	329.7	709.2	710.0	0	1	1
b	328.7	697.7	710.0	0	0	0
c	359.5	717.4	714.5	0	1	0
d	370.7	710.8	720.0	1	0	0
e	360.9	718.7	730.3	0	1	3
f	371.4	715.1	718.7	2	0	1
g	363.8	726.9	724.9	1	1	2
h	337.3	727.5	737.6	2	0	1
			Cycloamylose + 2960 Waters			
b	−25 822			0		
g	−25 812			0		

that our results are considered to be free from such artifacts.

There are some other salient points to be noted. First, the double-helical form shrinks along the helix axis compared with the energy-minimized configuration, narrowing the opening at each end. This feature is most clearly seen for cycloamylose with DP = 24, whose snapshot shows a somewhat deformed double-helical state, which may be related to the experimental observation that cycloamylose with DP ≃ 24 is not appreciably stained with iodine, in contrast to linear amylose with the same DP (Takaha, private communication). Second, the conformational flexibility of both forms is large compared with that of α-cyclodextrin. Indeed, the rms fluctuation of mass-weighted atomic position for DP = 24 is 1.2 Å (double) and 1.3 Å (single), which are both large compared to the corresponding simulation value of 0.58 Å (80 ps average) calculated for the heavy atoms of α-cyclodextrin in an aqueous solution.[48] The exhibited flexibility might be important for accommodating a variety of small guest molecules.

Energetic comparison was also carried out for the two forms with DP = 24 contained in a large water sphere, the number of water molecules 2960 being the same. As shown in Table 2, the energy difference again supports the formation of the double-helical form by about 10 kcal/mol, which has been obtained by averaging over the last 100 ps of the 175 ps simulation, preceded by 200 ps equilibration in a smaller water ellipsoid. However, there are unknown entropic contributions to the free energy, and fluctuation of the potential energy is somewhat large compared to that in vacuum. In this sense, the observed energy difference alone should not be regarded as providing decisive evidence.

Distribution of Glycosidic Torsional Angles in Water. Here we examine distribution of torsion angles ϕ and ψ in light of energy calculations and crystal structures. Many conformational energy calculations have been carried out for maltose, i.e., linear amylose with DP = 2, by using various force fields, such as CHARMm-based,[22,26] AMBER-based,[25] MM2,[21,23] MM3,[24,27] PCFF91,[28] GROMOS,[29] and others.[17−20] Some studies, especially the early ones, employed rigid glucose conformation,[17−20,28] while most of the recent studies take account of the relaxation of glucose units.[21−25,27] Recently, calculated results[22−24] have been compared in detail by Merz and co-workers.[25] Finally of note is the recent potential map of mean force obtained by Brady and co-workers.[26] In this connection, it is worthwhile to recall that certain relaxed maps characterized by a broader potential well failed to reproduce the observed chain dimension and its temperature coefficient, whereas the rigid maps could be refined to reproduce them, as shown by Brant and his co-workers[17] and by Kitamura et al.[18]

The dotted curves in Figures 9 and 10 represent isoenergy contours of the MM2 prediction that carefully takes account of the relaxation of glucose units, where the potential values have been taken from the work of French and Brady.[21] All other calculations with various force fields predict such a broad minimum approximately centered at $(\phi,\psi) = (90°,-150°)$, in the left-handed helix domain (L). The low-energy domain extends to the right-handed helix domain (R), in which a second potential minimum was observed for some of the force fields. The depth of the second minimum depends on the force field used. Some calculations predict a second minimum higher than the first by 1.5−2 kcal/mol,[21,23,24] while other fields predict a minimum of comparable depth[25] or a deeper minimum.[22]

In Figures 9 and 10, the points other than the open circles indicate the experimental values obtained for various amyloses such as α- and β-cyclodextrins,[58,59] double-helical amylose, and so on.[60,61] It is seen that experimental points are distributed in a low-energy region of the (ϕ, ψ) map. Indeed, they are located within the 4 kcal/mol contour of the MM2 prediction. The point at (90°,−60°), far apart from the other points, corre-

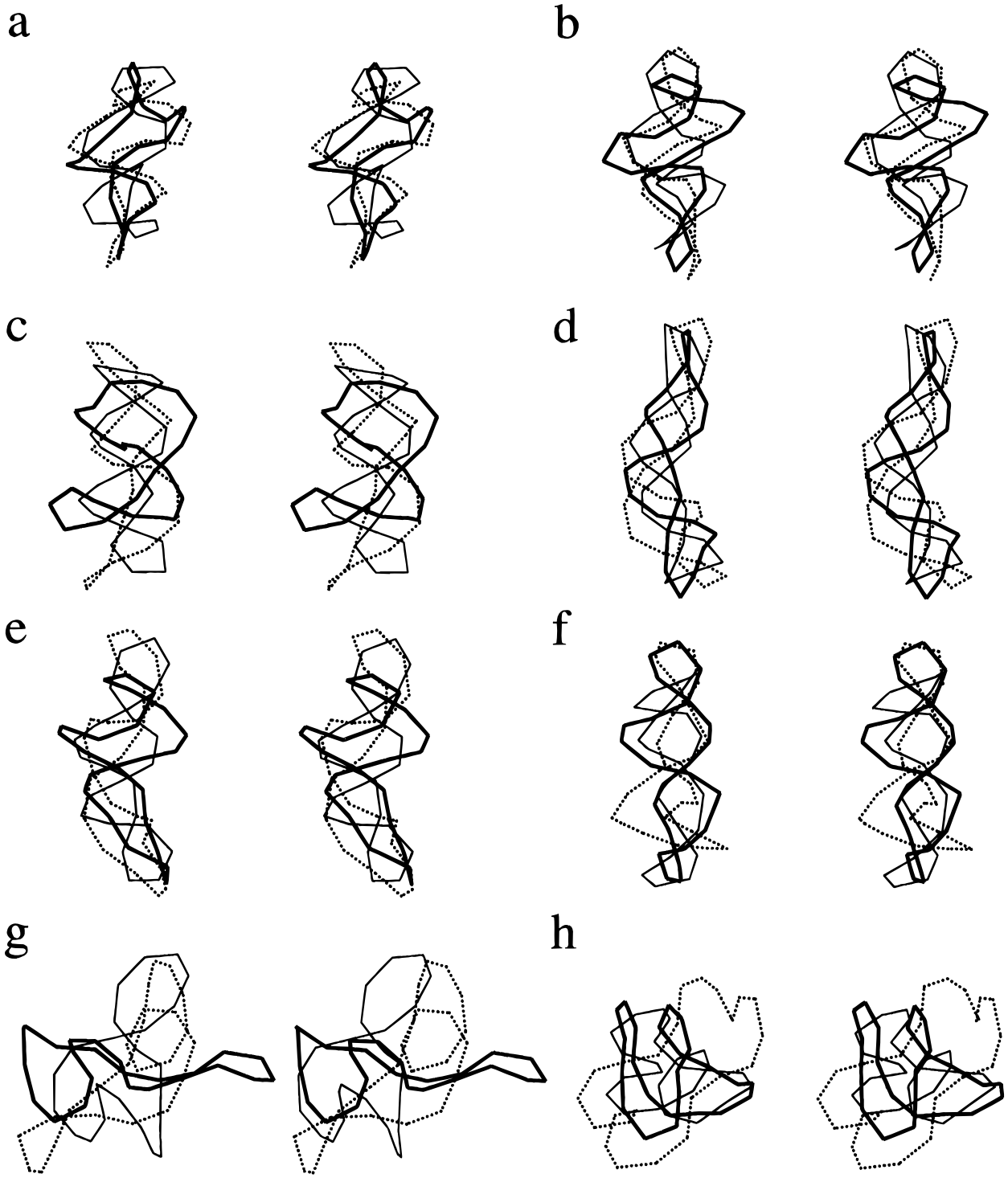

Figure 8. Stereoview of the snapshots of cycloamylose with DP = 24, obtained by molecular dynamics simulations in a vacuum at 300 K. Each of the simulations for a−h was initiated with the corresponding conformation in Figure 7. For clarity, only the positions of the glycosidic oxygens are traced. The thick-line segments (ϵ = 4) and dotted segments (ϵ = 80) represent the conformations after 25 ps of molecular dynamics simulation calculated with the MM2 field, and the thin-line segments represent those after 100 ps of simulation with the GROMOS field. Right and left views are separated by 28.6 Å along the horizontal direction.

sponds to the glycosidic linkage adjacent to the bent glucose unit of α-cyclodextrin. This large deviation is caused by an inclusion of water molecules in the central cavity, illustrating the large effect of water molecules on amylose conformation. Therefore, it is not surprising that the mean potential for (ϕ,ψ) obtained in the presence of waters and other glucose units somewhat deviates from that calculated for maltose isolated in a vacuum.

The open circles in Figures 9 and 10 represent the last 100 ps averages of (ϕ,ψ) calculated for each glycosidic linkage of the two forms of cycloamylose with DP = 24. The distribution of the simulation points in the L region may be regarded as being consistent with various experimental data. This distribution is also consistent with the 200 ps simulation results for a linear amylose fragment with DP = 6 in water (results not shown). In that R region in which the simulation points

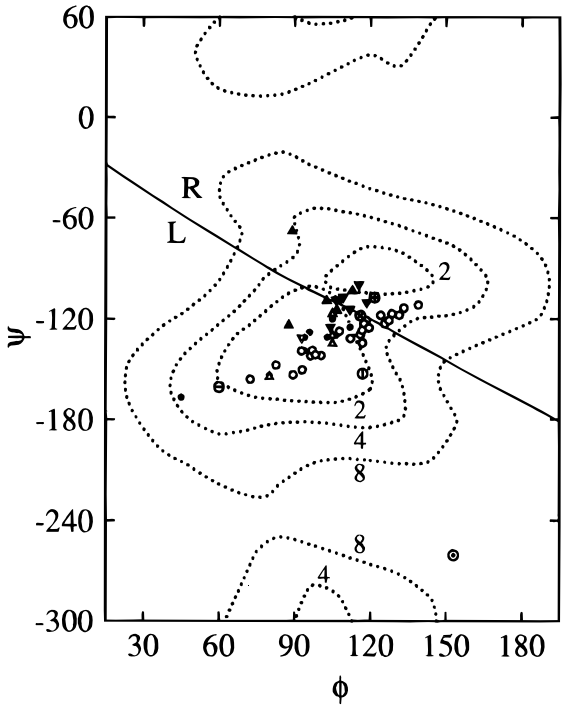

Figure 9. Distribution of glycosidic torsional angles ϕ and ψ for the double-helical cycloamylose with DP = 24. The open circles represent the average over 100 ps of simulation in water at 300 K. The dotted curves represent the MM2 isoenergy contours carefully calculated by French and Brady,[21] allowing relaxation of other degrees of freedom. The numbers attached to the curves indicate the energy expressed in kcal/mol. Other symbols represent the values found in published crystal structures: (▲, ▼) α- and β-cyclodextrins,[58,59] (●) anti-parallel double helix,[32,33] (▽) α-panose,[19] (△) maltopentaose on phosphorylase,[16] (⊙) maltotetraose bound to β-amylase,[54] (⊗, ⊕) α- and β-maltoses,[60,61] (⊖, ⏀) modeled structure of amylose complexed with KBR or KOH.[14]

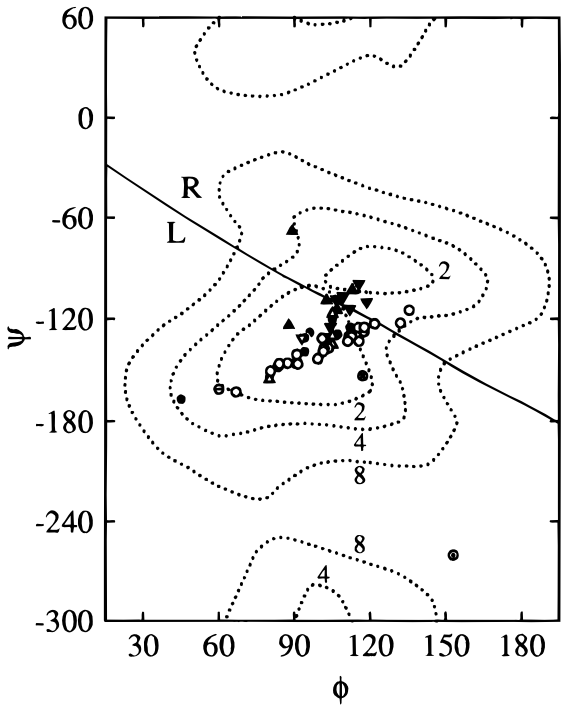

Figure 10. Same as Figure 9 except that simulation results for the single-helical cycloamylose are shown.

are located, there are no corresponding experimental data. However, the simulation points are still within the 4 kcal/mol contour, and therefore they may also be regarded as reasonable. Unusual pairs of (ϕ,ψ), as found in some of the simulations in a vacuum, were never observed during the simulations; even if the temperature was elevated to 500 K for 50 ps, unusual pairs did not appear.

6. Concluding Remarks

There are two plausible conformations of the newly discovered cycloamylose: the first is a circularized single-helical structure, and the second is a double-helical structure with foldbacks at each end. It is interesting to note their structural and topological similarities and dissimilarities to the supercoiled forms of closed circular DNA. We have made a simple analysis of the product population as a function of chain length (DP) on the basis of an elastic wire model. Assumption of rather high stiffness suggested that the double-helical structure is more reasonable, the single-helical structure being unlikely because of the absence of periodic oscillation of populations with a period of 6−8 residues. However, the conclusion drawn from the analysis depends on whether the amylose chain is rather stiff or flexible in a physiological environment.

Subsequently, we created atomistic models for these structures and carried out molecular dynamics simulations. From the viewpoint of single- *versus* double-helical structures, our simulation results may be summarized as follows.

When simulations were carried out in a vacuum, it was found that the single-helical form cannot maintain its initial toroidal shape, irrespective of the force fields used in this study. This instability may indicate that our analysis of the product population is not completely justified. However, the mean energy of the single-helical form is appreciably higher than that of the double-helical form. Therefore, we consider that the single-helical form is unlikely.

By contrast, when more realistic simulations were carried out in water, the toroidal single-helical form was maintained at 300 K (Figure 6b). Even when the temperature was raised to 500 K for a short time (50 ps for DP = 24), the initial shape still persisted, although two of the three leaves somewhat approached each other (figure not shown). If the form is stable, periodic oscillation of the product population should appear, as mentioned in the previous section. The absence of such an oscillation in the experimental data implies that the single-helical form is unlikely. From an energetic point of view, the double-helical form may be regarded as more stable, because its mean potential energy is lower than that of the single-helical form. Of course, this requires comments. There remains ambiguity arising from the conformational and solvation entropies, which were ignored in the above analysis because of their extreme difficulty in theoretical evaluation. Intuitively considered, however, they will work in the opposite direction, canceling each other to some extent; the conformational entropy favors the single-helical form having larger flexibility, while the solvation entropy favors the other having a smaller exposed area.[62]

At any rate, the prediction of the two molecular dynamics calculations are different with respect to the stability of the single-helical form. However, as far as the enthalpic contribution is concerned, both calculations agree in indicating that the double-helical form is more probable. In this sense, our calculation supports that cycloamylose has an anti-parallel double helix

structure, although final confirmation requires further experimentation. The two proposed atomistic structures should serve as useful starting points for further theoretical and experimental studies.

Acknowledgment. The authors are greatly indebted to Drs. T. Takaha, S. Okada (Ezaki Glico Co., Ltd.), and Dr. S. M. Smith (University of Edinburgh) for their many valuable suggestions and for guiding the authors into the study of cycloamylose. The authors also thank Dr. L. Sawyer (University of Edinburgh) for valuable, Professor W. K. Olson (Rutgers, the State University of New Jersey) for valuable comments, and Dr. S. Kitamura (Kyoto Prefectural University) for many discussions on linear and cyclic amylose conformations. Thanks are also extended to Professor Emeritus H. Yamakawa (Kyoto University) for his interest in this work. This work has been supported by the project "Improvement of biologically active compounds in foods (1992−1997)", organized by the Ministry of Agriculture, Forestry, and Fishery, Japan.

Appendix A: Ring-Closure Probability of Single Helical Amylose

In this appendix, we derive an approximate expression for the ring-closure probability of linear single-helical amylose on the basis of the helical wormlike chain model[5,6] without excluded volume. Static properties of this model of total contour length L are described in terms of four parameters: constant differential geometrical curvature κ_0 and torsion τ_0 of the characteristic helix, the static stiffness parameter λ^{-1}, and the shift factor M_L, defined as the molecular weight per unit contour length.

For amylose, the helical wormlike model was successfully applied to an analysis of the Monte Carlo results of Jordan et al.[17] It is remarkable that the progression of the virtual bond connecting glycosidic oxygens has been faithfully mimicked.[43] Evaluation of the ring-closure probability for atomistic models is fairly difficult. However, it is rather easy for the helical wormlike model, although only for some limited but important cases. Here, we show that the ring-closure probability for amylose can be *approximately* calculated from theoretical expressions previously derived for an twisted wormlike model, which is one of the extreme cases of the helical wormlike model.

The amylose chain of Jordan et al. was described in terms of rather large $\kappa_0/\lambda = 28$ and $\tau_0/\lambda = -12$. Although the atomistic models used by Jordan et al. and by us are different, their qualitative features are expected to be similar. For such cases of model parameters with large magnitude, the chain takes a fairly tightly coiled helical form, and its helix axis may behave as a uniformly bending rod. This means that, if we take the helix axis as the contour on a coarse-grained length scale, the chain may well be described by the ordinary Kratky−Porod wormlike chain[5,6,37] or the twisted wormlike chain.[5,38] It should be noted that contour length depends on the length scale to be adopted.

To show the coarse graining more explicitly, consider the mean-square end-to-end distance $\langle R^2\rangle$, whose expression is given, for example, by eq 14 of the paper of Fujii et al.[43] For the above cases of model parameters, it is well approximated by

$$\langle R^2\rangle = \frac{\lambda^{-2}\tau_0^{\,2}}{\kappa_0^{\,2}+\tau_0^{\,2}}\left[\lambda L - \frac{1}{2} + \frac{1}{2}\exp(-2\lambda L)\right] \quad \text{(A1)}$$

as long as contour length L is not very small. This expression resembles that for the wormlike chain of total contour length $\tilde{L}$ with stiffness parameter $\tilde{\lambda}^{-1}$:

$$\langle R^2\rangle_{KP} = \tilde{\lambda}^{-2}\left[\tilde{\lambda}\tilde{L} - \frac{1}{2} + \frac{1}{2}\exp(-2\tilde{\lambda}\tilde{L})\right] \quad \text{(A2)}$$

In fact, the two expressions become identical if their parameters and contour length are related by

$$\tilde{\lambda}^{-1} = \lambda^{-1}\left(\frac{\tau_0^{\,2}}{\kappa_0^{\,2}+\tau_0^{\,2}}\right)^{1/2} \quad \text{(A3)}$$

$$\tilde{\lambda}\tilde{L} = \lambda L \quad \text{(A3)}$$

This means that the chain trajectory of the amylose chain can be approximately described by the (auxiliary) wormlike chain, as expected.

Thus ring-closure probability $G_n(0)$ of the amylose chain composed of n glucose units can be calculated from that of the wormlike chain by

$$G_n(0) = \tilde{\lambda}^3 G_{KP}(0;\tilde{\lambda}\tilde{L}) \quad \text{(A4)}$$

Ring-closure probability G_{KP} for small L were evaluated directly from the configuration integral, taking proper account of fluctuations in the configuration of the closed ring around its most probable one (circle), while certain asymptotic expansion methods were used for large L. Its explicit expression as a function of $\tilde{\lambda}\tilde{L}$, τ_0, and the Poisson ratio σ are given by eq 73 of the article of Shimada and Yamakawa,[38] where the rate $\tilde{\tau}_0$ of the rotation around the helix axis is expressed as

$$\tilde{\tau}_0 = (\kappa_0^{\,2}+\tau_0^{\,2})/\tau_0 \quad \text{(A5)}$$

To actually calculate $G_{KP}(0)$, we must know the value of $\tilde{\lambda}^{-1}$. Because of the lack of experimental data, we are obliged to make two simplifying assumptions: (i) κ_0 and τ_0 are calculated by solving eqs 13 and 14 with the helix radius and pitch determined in a crystal environment, and (ii) L is equated to nl_u, where $l_u = 4.5$ Å is the virtual bond length.

The characteristic ratio C_∞, which is experimentally defined as

$$C_\infty = \lim_{n\to\infty}\langle R^2\rangle/nl_u^{\,2} \quad \text{(A6)}$$

can be expressed as

$$C_\infty = \tau_0^{\,2}/\lambda l_u(\kappa_0^{\,2}+\tau_0^{\,2}) \quad \text{(A7)}$$

for the auxiliary wormlike chain. Therefore, from eqs A3 and A7, its stiffness parameter $\tilde{\lambda}^{-1}$ may be determined as

$$\tilde{\lambda}^{-1} = l_u C_\infty(\kappa_0^{\,2}+\tau_0^{\,2})^{1/2}/|\tau_0| \quad \text{(A8)}$$

From the experimentally determined value[63] of $C_\infty = 4.8$, with radius $\rho = 5.2$ Å and pitch $h = 18.6$ Å, we have $\lambda^{-1} = 86$ Å and $\tilde{\lambda}^{-1} = 43$ Å. The oscillatory curve in Figure 2a was calculated with these parameters, assuming $\sigma = 0$.

There are two remarks regarding Figure 2a. The first is concerned with the Poisson ration σ. The amplitude

of the oscillation is proportional to the factor $(1 + \sigma)^{-1}$, so that the use of the more popular value of $\sigma = 0.5$ leads to somewhat reduced but still appreciable oscillations. The second remark is concerned with the magnitude of the stiffness parameter λ_{KP}^{-1}. On the large length scale adopted in an analysis of intrinsic viscosity, Nakanishi et al.[64] found that the amylose chain in dimethyl sulfoxide can be better described by the helical wormlike chain with $\kappa_0/\lambda = 3.5$ and $\tau_0/\lambda = 4$ rather than the KP chain. When they analyzed the data by the KP chain, they found the closest fit at $\lambda_{KP}^{-1} =$ 24 Å, which is half the above value of 43 Å. If the λ_{KP}^{-1} takes such a low value even in a physiological condition, as claimed by Norisuye,[65] then the oscillation becomes fairly damped, although it is still observed.

Appendix B: Boundary Forces for a System of Ellipsoidal Shape

In molecular dynamics simulations of biomolecules without periodic boundary, target molecules are often placed in a sphere of waters. It is then necessary to prevent an escape of water molecules from the sphere. For this purpose, several maneuvers have already been devised. In the simplest method, the escaped water is pulled back by applying harmonic boundary force directed toward the center of the sphere.

However, we often encounter a system that is ellipsoidal in shape rather than spherical. In such a case, an ellipsoid of water should be preferable from the viewpoint of consuming time, especially at an early stage of the simulation. Therefore, in this appendix, we make simple extensions required for treating the ellipsoidal systems. The calculation proceeds in four steps:

(1) We determine a minimum rectangular enclosing the target molecule. We designate its center by $\mathbf{r}_0 = (x_0,y_0,z_0)$ and its three sides by $2a_0$, $2b_0$, and $2c_0$.

(2) We then choose an ellipsoid large enough to enclose any of the target molecule. If we assume that its center is also located at $\mathbf{r}_0$, and that its three radii are $a' = ta_0$, $b' = tb_0$, and $c' = tc_0$, then t may be determined by

$$t = \max[(x - x_0)^2/a_0^2 + (y - y_0)^2/b_0^2 + (z - z_0)^2/c_0^2]^{1/2} \quad \text{(B1)}$$

where (x,y,z) extends over all the (initial) atomic positions of the target molecule.

(3) We add surrounding water layers of width w. The three radii then increase to $a = a' + w$, $b = b' + w$, and $c = c' + w$.

(4) We assume that the boundary force acts toward the center, with its magnitude being proportional to the distance (measured along the force) between the particle and the surface of the ellipsoid. To give expressions for the boundary force, we introduce for each particle a parameter λ defined by

$$\lambda = [(x - x_0)^2/a^2,\ \ (y - y_0)^2/b^2,\ \ (z - z_0)^2/c^2]^{-1/2} \quad \text{(B2)}$$

where $\mathbf{r} = (x,y,z)$ is the position of the particle. If $\lambda \leq 1$, the particle is outside of the ellipsoid, and it is inside otherwise. When the particle is outside, the boundary potential U_{BF} and force $\mathbf{F}_{BF}$ acting on that particle are given by

$$U_{BF} = {}^1/_2\kappa(1 - \lambda)^2(\mathbf{r} - \mathbf{r}_0)^2 \quad \text{(B3)}$$

$$\mathbf{F}_{BF} = -\kappa(1 - \lambda)^2(\mathbf{r} - \mathbf{r}_0) - \kappa(1 - \lambda)\lambda^3(\mathbf{r} - \mathbf{r}_0)^2[(x - x_0)/a^2,\ \ (y - y_0)/b^2,\ \ (z - z_0)/c^2] \quad \text{(B4)}$$

where κ is a force constant. When the particle is inside the ellipsoid, $U_{BF} = \mathbf{F} = 0$, as is natural.

References and Notes

(1) Takaha, T.; Yanase, M.; Takata, H.; Okada, S.; Smith, S. M. *J. Biol. Chem.* **1996**, *271*, 2902.
(2) Go, N.; Scheraga, H. A. *Macromolecules* **1970**, *3*, 178.
(3) Bruccoleri, R. E.; Karplus, M. *Biopolymers* **1987**, *26*, 137.
(4) Koča, J.; Pérez, S.; Imberty, A. *J. Comput. Chem.* **1995**, *16*, 296.
(5) Yamakawa, H. *Annu. Rev. Phys. Chem.* **1984**, *35*, 23.
(6) Yamakawa, H. In *Molecular Conformation and Dynamics of Macromolecules in Condensed Systems*; Nagasawa, M., Ed.; Elsevier: Amsterdam, 1988; p 21.
(7) Yamakawa, H.; Shimada, J. *J. Chem. Phys.* **1978**, *68*, 4722.
(8) Olson, W. K.; Babcock, M. S.; Gorin, A.; Liu, G.; Marky, N. L.; Martino, J. A.; Pedersen, S. C.; Srinivasan, A. R.; Tobias, I.; Westcott, T. P.; Zhang, P. *Biophys. Chem.* **1995**, *55*, 7.
(9) Hao, M.-H.; Olson, W. K. *J. Biomol. Struct. Dyn.* **1989**, *7*, 661.
(10) 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.
(11) McCammon, J. A.; Harvey, S. C. *Dynamics of proteins and nucleic acids*; Cambridge Univ. Press: Cambridge, UK, 1987.
(12) Brooks, C. L., III; Karplus, M.; Petitt, B. M. *Proteins: A theoretical Perspective of Dynamics, Structure and Thermodynamics*; Wiley: New York, NY, 1988.
(13) van Gunsteren, W. F.; Mark, A. E. *Eur. J. Biochem.* **1992**, *204*, 947.
(14) Goebel, C. V.; Dimpfl, W. L.; Brant, D. A. *Macromolecules* **1970**, *3*, 644.
(15) Winter, W. T.; Sarko, A. *Biopolymers* **1974**, *13*, 1447.
(16) Goldsmith, E.; Sprang, S.; Fletterick, R. *J. Mol. Biol.* **1982**, *156*, 411.
(17) (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.; ACM: Washington, DC, 1990; Chapter 4.
(18) (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.
(19) Imberty, A.; Pérez, S. *Carbohydr. Res.* **1988**, *181*, 41.
(20) Imberty, A.; Chanzy, H.; Pérez, S.; Buléon, A.; Tran, V. *J. Mol. Biol.* **1988**, *201*, 365.
(21) French, A. D.; Brady, J. W. In *Computer Modeling of Carbohydrate Molecules*; French, A. D., Brady, J. W., Eds.; ACM: Washington, DC, 1990; Chapter 1.
(22) Ha, S. N.; Madsen, L. J.; Brady, J. W. *Biopolymers* **1988**, *27*, 1927.
(23) Tran, V.; Buleon, A. *Biopolymers* **1989**, *28*, 679.
(24) Dowd, M. K.; Zeng, J.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *230*, 223.
(25) Glennon, T. M.; Zheng, Y.-J.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M., Jr. *J. Comput. Chem.* **1994**, *15*, 1019.
(26) Schmidt, R. K.; Teo, B.; Brady, J. W. *J. Phys. Chem.* **1995**, *99*, 11339.
(27) Fringant, C.; Tvaroska, I.; Mazeau, K; Rinaudo, M.; Desbrieres, J. *Carbohydr. Res.* **1995**, *278*, 27.
(28) Trommsdorff, U.; Tomka, I. *Macromolecules* **1995**, *28*, 6128.
(29) Ott, K.-H.; Meyer, B. *Carbohydr. Res.* **1996**, *281*, 11.
(30) Kainuma, K.; French, D. *Biopolymers* **1972**, *11*, 2241.
(31) Imberty, A.; Pérez, S. *Biopolymers* **1988**, *27*, 1205.
(32) Hinrichs, W.; Büttner G.; Steifa M.; Betzel, Ch.; Zabel, V.; Pfannemüller, B.; Saenger, W. *Science* **1987**, *238*, 205.
(33) Hinrichs, W.; Saenger, W. *J. Am. Chem. Soc.* **1990**, *112*, 2789.
(34) Schulz, W.; Sklenar, H.; Hinrichs, W.; Saenger, W. *Biopolymers* **1993**, *33*, 363.
(35) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry*; Freeman: San Francisco, CA, 1980; Chapter 24.
(36) Jacobson, H.; Stockmayer, W. H. *J. Chem. Phys.* **1950**, *18*, 1600.
(37) (a) Flory, P. J. *Statistical Mechanics of Chain Molecules*; Wiley: New York, NY, 1969; Appendix D. (b) Flory, P. J.; Semlyen, J. A. *J. Am. Chem. Soc.* **1966**, *88*, 3209.
(38) Shimada, J.; Yamakawa, H. *Macromolecules* **1984**, *17*, 689.
(39) Shore, D.; Baldwin, R. L. *J. Mol. Biol.* **1983**, *170*, 957.
(40) Levene, S. D.; Crothers, D. M. *J. Mol. Biol.* **1986**, *189*, 73.

(41) Hagerman, P. J.; Ramadevi, V. A. *J. Mol. Biol.* **1990**, *212*, 351.
(42) Vologodskii, A. V.; Cozzarelli, N. R. *Annu. Rev. Biophys. Biomol. Struct.* **1994**, *23*, 609.
(43) Fujii, M.; Nagasaka, K.; Shimada, J.; Yamakawa, H. *Macromolecules* **1983**, *16*, 1613.
(44) Mimura, M.; Urakawa, H.; Kajiwara, K.; Kitamura, S.; Takeo, K. *Macromol. Symp.* **1995**, *99*, 43.
(45) Zhang, P.; Tobias, I.; Olson, W. K. *J. Mol. Biol.* **1994**, *242*, 271; **1995**, *251*, 721.
(46) Westcott, T. P.; Tobias, I.; Olson, W. K. *J. Phys. Chem.* **1995**, *99*, 17926.
(47) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 197.
(48) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *J. Mol. Biol.* **1988**, *203*, 241.
(49) Burkert, U.; Allinger, N. L. *Molecular Mechanics*; American Chemical Society: Washington, DC, 1982.
(50) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press: Oxford, UK, 1987.
(51) Greengard, L. *Science* **1994**, *265*, 909.
(52) Shimada, J.; Kaneko, H.; Takada, T. *J. Comput. Chem.* **1993**, *14*, 867; **1994**, *15*, 28.
(53) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
(54) Mikami, B.; Degano, M.; Hehre, E. J.; Sacchettini, J. C. *Biochemistry* **1994**, *33*, 7779.
(55) Eisenhaber, F.; Schulz, W. *Biopolymers* **1992**, *32*, 1643.
(56) Saito, M. *J. Chem. Phys.* **1994**, *101*, 4055.
(57) Schreiber, H.; Steinhauser, O. *Biochemistry* **1992**, *31*, 5856.
(58) Lindner, K.; Saenger, W. *Acta Crystallogr.* **1982**, *B38*, 203.
(59) Betzel, C.; Saenger, W.; Hingerty, B. E.; Brown, G. M. *J. Am. Chem. Soc.* **1984**, *106*, 7545.
(60) Takusagawa, F.; Jacobson, R. A. *Acta Crystallogr.* **1978**, *B34*, 213.
(61) Gress, M. E.; Jeffrey, G. A. *Acta Crystallogr.* **1977**, *B33*, 2490.
(62) Ooi, T.; Oobatake, M.; Némethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086.
(63) Banks, W.; Greenwood, C. T. *Carbohydr. Res.* **1968**, *7*, 349.
(64) Nakanishi, Y.; Norisuye, T.; Teramoto, A.; Kitamura, S. *Macromolecules* **1993**, *26*, 4220.
(65) Norisuye, T. *Macromol. Symp.* **1995**, *99*, 31.

MA9603387