Interaction between Cellooligosaccharides in Aqueous Solution from Molecular Dynamics Simulation: Comparison of Cellotetraose, Cellopentaose, and Cellohexaose

Myco Umemura,*,† Yoshiaki Yuguchi, and Takahiro Hirotsu

National Institute of Advanced Industrial Science & Technology, AIST, 2217-14 Hayashi-cho, Takamatsu-shi, Kagawa 761-0395, Japan

Received: March 3, 2004; In Final Form: June 16, 2004

A molecular dynamics (MD) simulation was carried out for aqueous solutions of cellooligosaccharides to investigate their interactive behavior. Single and double strands of cellotetraose, cellopentaose, and cellohexaose were simulated in systems with TIP3P water. The cellotetraose double strand separated into two within 1 ns calculation time. An aggregation state of a cellohexaose double strand is firmer than that of cellopentaose in the simulation, supporting experimental results. The self-diffusional motion of solutes is approximately in inverse proportion to their molecular surface areas without special interactive effects, while their atomic positional fluctuations manifest the effect of sugar−sugar interaction. Sugar−sugar interaction restrains the fluctuation inversely proportional to the degree of polymerization, as opposed to proportional promotion by sugar−water interaction. The degree of polymerization is at a critical point in cellopentaose, as viewed from the balance between sugar−sugar and sugar−water interaction. The occupancy of intramolecular hydrogen bonds between the particular oxygen atoms, O3 and O5′, in adjacent residues is remarkably high in the double strand of cellohexaose among the three cellooligosaccharides, which abundance means a stable state of a twisted ribbon structure of the sugar chain. That leads to a higher occurrence of intermolecular hydrogen bonds between sugar chains using the oxygen atom O6 in cellohexaose than in cellotetraose or cellopentaose.

1. Introduction

Cellulose, the main component of wood, is an attractive, environmentally friendly, but hard-to-treat material for precise processing due to its insolubility and lack of thermal plasticity.[1−3] It is composed of bundles of sugar chains that aggregate tightly with each other having both hydrophilic groups and relatively hydrophobic flat faces.[1] Figure 1a depicts the natural crystal made up from metastable Cellulose I with all strands parallel.[4] The crystal is constructed of sheets composed of sugar chains that are hydrogen bonded with each other.[1] The aggregation can be considered to occur by hydrophobic interaction between the sheets on which hydrocarbon groups lie, but its mechanism is still not clearly understood from the molecular point of view. Knowledge of interaction between the sugar chains is crucial to understand the nature of cellulose and other saccharides.

Cellooligosaccharide is the oligosaccharide of cellulose that barely dissolves when its degree of polymerization (DP) is greater than five. This phenomenon is confirmed by the small angle X-ray scattering (SAXS) method that can estimate the state of aggregation on a nanoscale. We have obtained SAXS profiles of cellopentaose, cellohexaose, maltopentaose, and maltohexaose in aqueous solution, in which the former two are penta- and hexaoligosaccharides of cellulose and the latter two those of amylose, respectively, with a concentration of 10 mg/mL at 298 K.[5] Cellulose and amylose, both made of D-glucopyranoses, are different in the configuration of their glycosidic linkages (α and β, respectively). The molecular weights of penta- and hexaoligosaccharides are not so greatly different as to have different scattering curves if their aggregation states are similar. However, we have observed that the profile of cellohexaose has remarkably higher values with steeper slope than those of the other three that have almost the same gentle slope in the small angle area. This means that cellohexaose aggregates are unlike the other three that are in molecular dispersion. Some changes occur from cellopentaose to cellohexaose in the molecular motion of sugar, the hydration structure around it, sugar flexibility, or a balance between hydrophilic and hydrophobic interactions. Analysis on an atomic order is needed because the solubility of cellohexaose is distinctly different from that of cellopentaose, even though the DP increases only one from cellopentaose to cellohexaose.

When we consider the interaction between cellooligosaccharides, one important aspect of cellulose is its flat ribbon surface formed with particular intramolecular hydrogen bonds, which surface has a hydrophobic characteristic.[1] Tashiro et al. theoretically estimate that a particular sort of intramolecular hydrogen bond plays the most essential part in the elasticity of cellulose among all intra- and intermolecular hydrogen bonds; the elastic constant of cellulose decreases at the highest rate when the intramolecular hydrogen bond between O3−H3 and O5′ in adjacent residues (symbols will be defined later) diminishes.[6] Some experimental data confirm that the stability of crystalline cellulose depends on the rigidity of the ribbon structure in strands.[7,8] In addition, Mattinen et al. report that the flat rigid surface of crystalline cellulose makes a facile attachment possible for aromatic residues in cellulases, based on their NMR experiments.[9] These results suggest that the ribbon structure of cellulose or cellooligosaccharides is intrinsic to their aggregation mechanism. Contrary to this, Tanaka et al. report a result from MD simulation that cellulose chains in the aqueous environment do not keep the folded-chain structure but make loose tangles

* Address correspondence to this author. E-mail: maiko-umemura@aist.go.jp.

† Present address: AIST, 2266−98 Anagahora, Shimosidanmi, Moriyama-ku, Nagoya-shi, Aichi, 463-8560 Japan.

10.1021/jp049044a CCC: $27.50
Published on Web 07/31/2004

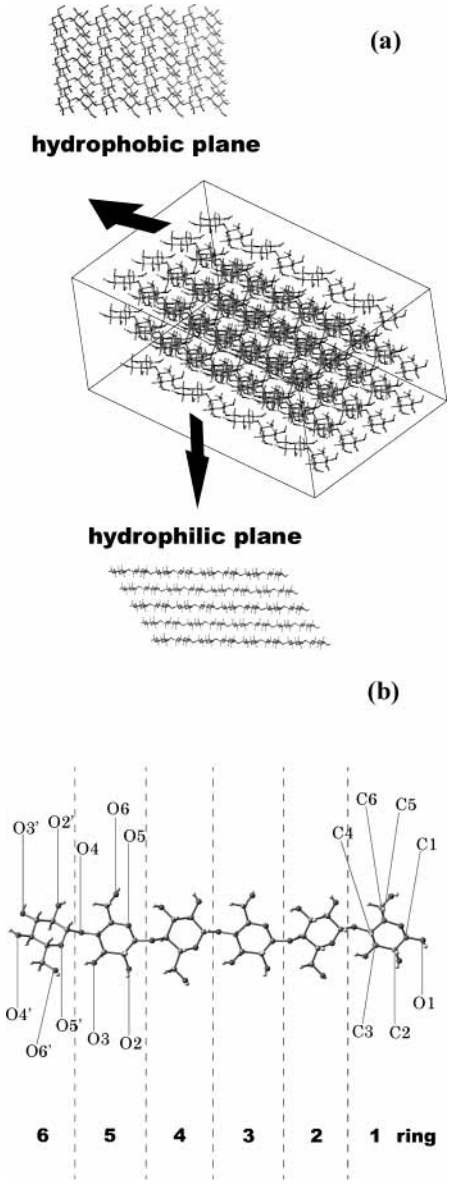

Figure 1. (a) The crystal structure of cellulose I. It has both hydrophobic and hydrophilic faces. Hydrocarbon and hydroxyl groups expose themselves on the hydrophobic and hydrophilic planes, respectively. (b) Symbols for atomic and ring numbering in cellooligosaccharides.

that contain many water molecules among them.[10] Although their result has already stirred controversy with some doubts of the calculation condition,[11] the present study can be another test of their result.

In this paper, we investigate from an atomic point of view the interactive behavior of cellooligosaccharides in aqueous solution particularly in order to understand different dissolving properties between tetra- or penta- and hexasaccharides. We focus on the relation between the ribbon structure and the aggregation state of cellooligosaccharides. For this purpose, we use the MD simulation method that allows us to decide the initial coordinate between cellooligosaccharides and to obtain the atomic information about the position and momentum of each molecule. MD simulation applied to oligosaccharides can result in good agreement with experimental data.[12−15] MD simulations offer the advantage of providing insights into the influence that interaction between oligosaccharides has on the aggregation state on an atomic order, supplementing statistical average information obtained from experiment. We arrange two strands of cellotetraose, cellopentaose, and cellohexaose so that one relatively flat ribbon surface is parallel to the other to see the interaction between the surfaces, and analyze their aggregation states in water through 1 ns calculation time. We also calculate single strand systems of the three cellooligosaccharides in water for comparison to see the effects of sugar−sugar interaction on dynamics and hydrogen bondings of the solutes. There are two kinds of interaction concerning the solutes, sugar−sugar and sugar−water interactions, and it has not been clear which is the main factor in determining the aggregation state of cellulose in water.[10,11] We analyze the dynamics of solutes from diffusion and fluctuation from which we can obtain information about the relationship between the two kinds of interactions. We then examine intra- and intermolecular hydrogen bonds at each hydroxyl group in solutes, which reflect the conformation and arrangements of solutes in an adequate way. We also pay attention to the relationship between the intramolecular hydrogen bond, O3−O5′, and the interaction between cellooligosaccharides.

2. Simulation Procedure

Six kinds of systems with cellotetraose, cellopentaose, or cellohexaose and explicit water were simulated with particle mesh Ewald[16] for this study. Single strands of the three cellooligosaccharides were simulated as solutes for comparison in addition to their double strands. We performed all MD calculations using the SANDER module in the AMBER 7 program.[17] The Parm99 force-field parameters,[18] augmented with the GLYCAM parameters[19] (version 2000a) for oligosaccharides, were used throughout the MD calculations. The conformation and atomic partial charges of the saccharide were retrieved from the work by Woods et al. in which the charge is ensemble averaged with the RESP algorithm.[20,21] We set the initial arrangement of double strands in the crystal structure of cellulose I, in which each ring face is 6.506 Å apart in parallel.[22] A cutoff of 16 Å was used for nonbonded interactions. MD simulations were performed under isothermal−isobaric periodic boundary conditions, in which the saccharides were immersed in a theoretical box of 2000 and 5000 TIP3P[23] waters for single and double strands, respectively. The SHAKE[24] algorithm was applied to all bonds containing hydrogen atoms, consistent with the use of TIP3P waters.

Initial conjugate-gradient energy minimization was performed on all systems, using a 0.01 kcal·mol^{-1}·Å convergence criterion in the energy gradient. The energy minimizations were followed by a period of annealing for water (30 ps), during which the temperature of the system was increased and stabilized at 298 K. During the minimization and annealing, the solute was fixed with a force constant of 200 kcal·mol^{-1}·Å^{-2} and the initial unfavorable contacts made by the solvent were removed. After these preparations, the solute was released and production dynamics were performed for 1 ns with 2 fs time steps at 298 K and a pressure of 1 atm. The calculation proceeded on the PCI card MD Engine II (Fuji Xerox, Tokyo) designed exclusively for the MD calculation.

Symbols for atomic numbering and direction of numbering are in accordance with the recommendations of the IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN).[25] The chain is numbered from the reducing glucose residue to the nonreducing glycosyl group. Note that O4 involving a glycosidic linkage is included in the residue close to the reducing end. We herein call each residue a "ring" for simplicity (Figure 1b). For example, the reducing glucose residue is called "Ring

1", and the residue at the nonreducing end in cellohexaose is called "Ring 6".

3. Results and Discussion

3.1. Aggregation State. Figure 2 denotes the average structure of the three double strands in the initial and last 100 ps, with that of the three single strands in the last 100 ps, in 1 ns calculation time. The conformations of the single strand of cellotetraose, cellopentaose, and cellohexaose are the same in their loose helix characteristics except for periods. We describe their helical properties using the well-established parameters *n*, the number of monomers per helix turn, and *h*, the advance (in Å) along the helix axis per residue.[26] A (+) or no sign and a (−) sign of *n* mean right-handed and left-handed conformations, respectively (see details in ref 26). The chain becomes closer to a flat, 2-fold screw axis conformation in cellulose as the value of *n* is closer to 2. The average values of the solutes in the last 100 ps are $n = -2.61$, -2.62, -2.65 and $h = 4.60$, 4.89, 5.00, in the single strands of cellotetraose, cellopentaose, and cellohexaose, respectively. The flat ribbon structure in cellulose is twisted in cellooligosaccharides. The loose helical conformation of cellooligosaccharides also has been found in two-dimensional NMR spectra.[27] The double strand of cellohexaose steadily aggregates, forming a loose helix with $n = -2.50$ and $h = 5.09$ (Figure 2i), while that of cellopentaose has a separation at the side of Ring 1 ($n = -2.43$ and $h = 5.01$, Figure 2f). That of cellotetraose separates into two single strands being apart approximately 20 Å at the shortest distance ($n = -2.60$ and $h = 4.72$, Figure 2c). It is interesting that the conformation of chains in double strands is closer to that in cellulose than in single strands. The double strands of cellotetraose, cellopentaose, and cellohexaose are in separation, half-aggregation, and aggregation states, respectively. The ribbon surfaces of the double strands are parallel to each other for Ring 3 to 5 of the cellopentaose and for Ring 1 to 6 of the cellohexaose. This indicates the stability of the face-to-face interactions in cellooligosaccharides.

These aggregation states continue through 1 ns except for the initial 100 ps in double strand systems of cellopentaose and cellohexaose. Figure 3 shows the time evolution of aggregation states in each double strand system as the distance between the center of mass of the atoms C1−5 in each corresponding ring. The abbreviations **c4d**, **c5d**, and **c6d** mean the double strand systems of cellotetraose, cellopentaose, and cellohexaose, respectively. The character "**c**", the numbers from 4 to 6, and the character "**d**" denote "cellooligosaccharide", the DP, and "double strand", respectively. Distances between rings show a phased increase and finally have values of more than 20 Å in **c4d**. In **c5d**, the values are unstable especially at Ring 1, and the baselines also fluctuate. In contrast to these two systems, distances are small and stable in **c6d** after about 80 ps with stable baselines at ca. 4.6 Å, except for those at Ring 6 that sometimes have slightly larger values. The distances are averaged over 100 to 1000 ps (Figure 4). The value of each ring pair is around 4.6 Å in cellohexaose, while it deviates from 4.8 to 10 Å in cellopentaose. The distances for cellotetraose are more than 11.8 Å at Ring 3. The reducing end of Ring 1 in cellopentaose has the greatest value of 10 Å, which is approximately twice as great as that in cellohexaose. The aggregation state is more stable and compact in cellohexaose than in cellopentaose. Ends in every double strand have greater values than the middle rings, denoting that water can attack and perturb the former more than the latter part in strands. We cannot find the cause of the specific separation from Ring 1 in cellopentaose at present.

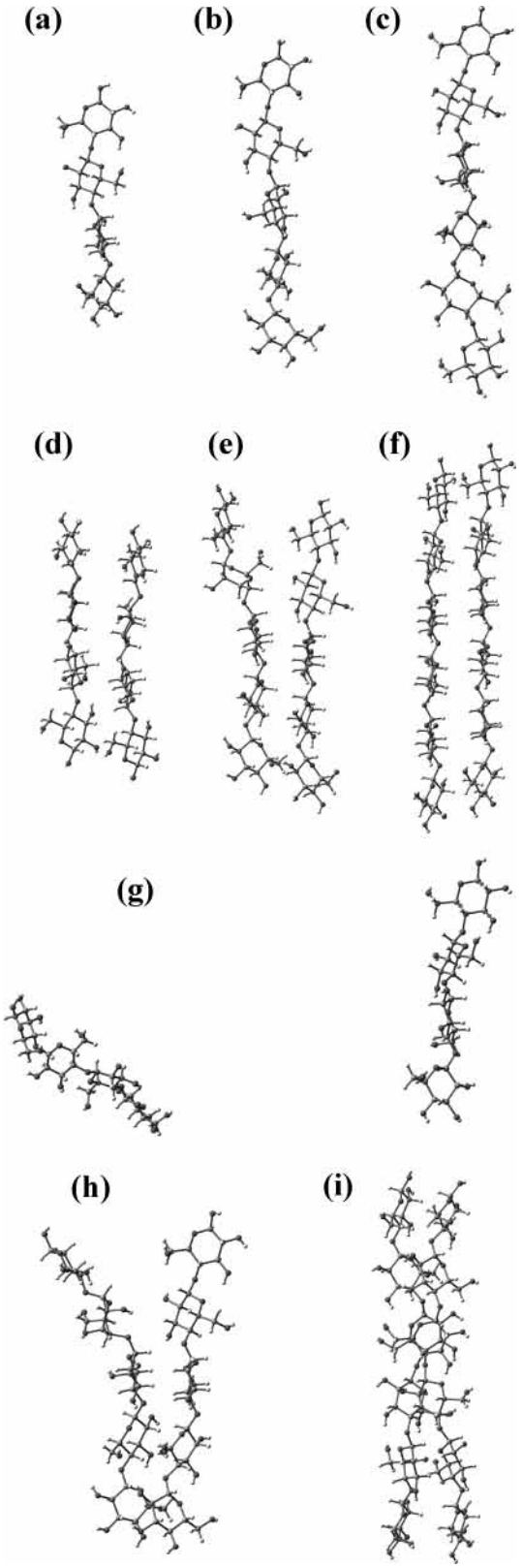

Figure 2. The averaged structure of solutes through the initial or last 100 ps in 1 ns calculation time; single strands of (a) cellotetraose, (b) cellopentaose, and (c) cellohexaose in the last 100 ps, double strands of (d) cellotetraose, (e) cellopentaose, and (f) cellohexaose in the initial 100 ps, and double strands of (g) cellotetraose, (h) cellopentaose, and (i) cellohexaose in the last 100 ps.

The double strand of cellopentaose does not entirely separate through the 1 ns calculation time, although it is in molecular dispersion in experiments.[5] Separation can occur entirely in

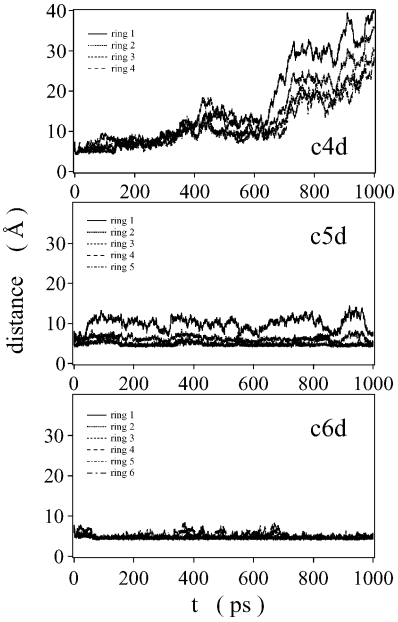

Figure 3. Time evolution of distances between two corresponding rings in the double strand system of cellotetraose (**c4d**), cellopentaose (**c5d**), and cellohexaose (**c6d**).

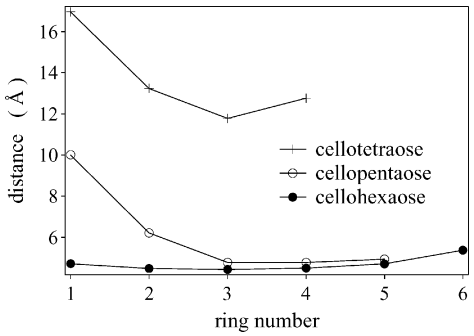

Figure 4. The average distance between two corresponding rings in the double strands of cellotetraose, cellopentaose, and cellohexaose from 100 to 1000 ps.

cellopentaose like cellotetraose with changes of calculation conditions such as temperature, pressure, calculation time, or the initial arrangement between each strand. The choice of an initial arrangement is particularly difficult because it is almost infinite. One stable structure of cellulose, Cellulose I, was used as the initial arrangement in this case, so the results for cellotetraose, cellopentaose, and cellohexaose can be reasonably compared. In this sense, the partly separated state of the cellopentaose double strand in the calculation corresponds well to the result from SAXS profiles.[5]

3.2. Diffusion and Fluctuation. Translational motions of solutes are estimated by mean-square displacement of the center of mass of all atoms in each sugar chain from 0 to 5 ps (Figure 5). We averaged the value at each MD calculation step from 100 to 1000 ps. In each double strand system, the result is averaged over two solutes. All converge well to form straight lines after about 0.5 ps. The steepest slope is that of the single cellotetraose (**c4s**), and the gentlest slopes are those of the double cellohexaose (**c6d**). The slopes of the single cellopentaose (**c5s**) and cellohexaose (**c6s**) are approximately identical with those of the double cellotetraose (**c4d**) and cellopentaose (**c5d**), respectively. Diffusivity is quantitatively estimated by the self-

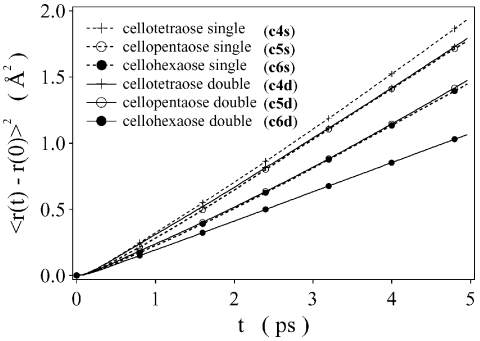

Figure 5. Mean-square displacement of the center of mass of sugar chains from 0 to 5 ps. The values of the double strand systems were the average over each strand.

TABLE 1: Simulated Self-Diffusion Coefficients of Cellotetraose, Cellopentaose, and Cellohexaose[a]

	D (10^{-10} m^2 s^{-1})	
	single	double[b]
cellotetraose (**c4**)	6.87	6.29
cellopentaose (**c5**)	6.35	5.31
cellohexaose (**c6**)	5.26	3.68

[a] The values are evaluated from 100 to 1000 ps in each calculation. [b] The values are averaged over two solutes in each system.

diffusion coefficient D calculated by using the Einstein relation

$$D = \lim_{t\to\infty} \frac{1}{6t} \langle |\boldsymbol{r}_i(t) - \boldsymbol{r}_i(0)|^2 \rangle \qquad (1)$$

where t is time and $\boldsymbol{r}_i$ is the position of the center of mass in each sugar chain. We can obtain the coefficient from the inclination in Figure 5. As can be seen in Table 1, the values of **c5s** and **c6s** are 1.1 and 1.2 times those of **c4s** and **c5s**, respectively. Those of **c5d** and **c6d** are 1.2 and 1.4 times those of **c4d** and **c5d**, respectively. Those of **c4s**, **c5s**, and **c6s** are 1.1, 1.2, and 1.4 times those of **c4d**, **c5d**, and **c6d**, respectively. The values of **c4d** and **c5d** approximately equal those of **c5s** and **c6s**, respectively.

The scale and order of the values are inversely proportional to the size of solute surface areas under the following assumption. We assume a cellooligosaccharide to be a long flat rectangular parallelepiped with sides 4, 8, and 6 multipled by DP plus 3 Å (Figure 6). The length of the end residue is assumed as 1.5 Å longer than the internal one for a hydroxyl group. Assuming that two rectangles are stacked in double strands, the surface areas are estimated as 712, 856, 1000, 928, 1120, and 1312 Å^2 for **c4s**, **c5s**, **c6s**, **c4d**, **c5d**, and **c6d**, respectively. The surface area increases approximately 1.2 times as the DP augments from 4 to 5 or from 5 to 6 in the single or double strands. The area of the double strand is 1.3 times that of the single in cellotetraose, cellopentaose, and cellohexaose. The rate becomes closer to that of D when we consider the difference in the aggregation states of tetra-, penta-, and hexacellooligosaccharides, i.e., almost completely separated, loose, and firm, respectively. The area becomes larger in the calculated **c5d** than in our assumption. In the double strand system of cellotetraose, each strand can move like a single in its separate state, resulting in a greater value of D as average. Therefore, the self-diffusive characteristics of the sugar chains can be considered to roughly follow the extent of their surface areas. This is consistent with

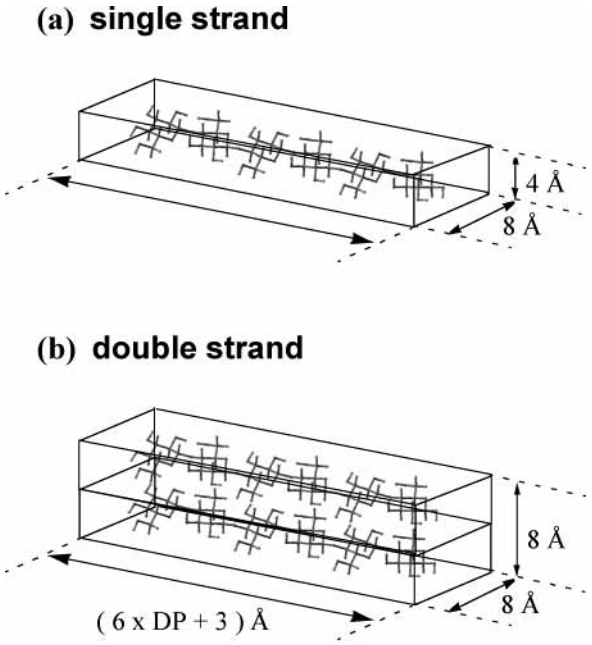

Figure 6. Schematic description explaining the diffusion characteristics with the extent of solute surface areas: (a) single strand and (b) double strand.

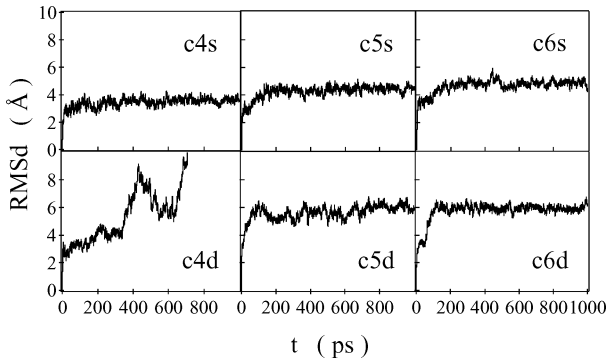

Figure 7. The root-mean-square distance (rmsd) per atom in solutes after rms fittings in each system.

the Einstein relation

$$D = \frac{k_B T}{\xi} \tag{2}$$

where k_B is the Boltzmann constant, T is temperature, and ξ is a coefficient of friction.[28] It is natural to consider that friction increases in proportion to surface area. We leave a detailed discussion for another opportunity, but conclude that the aggregation state of the sugar chain influences its self-diffusivity primarily according to the change of surface areas, without any special effect due to interaction between sugar chains or sugar and water.

Atomic positional fluctuation is defined as the root-mean-square displacement of an atom from its average position during an interval. Advancing to the evaluation of fluctuations, we fit every center of gravity of all atoms in solutes to the previous one for removal of the translational and rotational motions of solutes themselves. We used the PTRAJ module in AMBER 7 for this operation called the rms fitting. Figure 7 describes the root-mean-square distance (rmsd) per sugar atom in each system after the rms fitting. Deviation from the initial coordinate reaches a plateau in the opening 100 ps in each system except for **c4d**, meaning that ca. 100 ps is needed to stabilize the systems sufficiently. The values after ca. 700 ps in **c4d** are truncated in the figure because of their large size. Unremarkable small slopes seen in the plateaus are due to the accumulation of errors while fitting coordinates at each step to the previous, which errors are removed when we fit each coordinate to the initial one. The stable conformation is changed by sugar−sugar interactions because values at the plateaus are larger in the double than in the single strand systems of cellopentaose or cellohexaose, being ca. 6 and 4 Å, respectively. Fluctuation within a 30−40-ps interval is observed in each system with other fine fluctuations. Among these short and long intervals, the double strand system of cellopentaose has a unique ca. 100-ps interval of fluctuation. This fluctuation is caused by a closing−separating movement at the reducing end in the cellopentaose double strand.

TABLE 2: Averaged Atomic Positional Fluctuations per Atom in Cellotetraose, Cellopentaose, and Cellohexaose[a]

	Å/5 ps		Å/40 ps	
	single	double[b]	single	double[b]
cellotetraose (**c4**)	0.65	0.81	0.93	1.59
cellopentaose (**c5**)	0.69	0.71	1.04	1.14
cellohexaose (**c6**)	0.74	0.67	1.26	0.92

[a] The values are evaluated from 100 to 1000 ps in each calculation. [b] The values are averaged over two solutes in each system.

On the basis of the rms fitted coordinates, we evaluated atomic positional fluctuations per atom at two intervals, 5 and 40 ps, corresponding to the short and long ones seen in rmsd's. Table 2 denotes the average values in each system after the opening 100 ps in calculation time during which the results are meaningless. The values are inversely proportional to DP in the single strand systems, while the opposite is true in the double strand. The values at both short and long intervals are the highest and lowest in the single and double strand systems of cellohexaose, respectively. The fluctuations in **c5s** and **c6s** are 1.06 and 1.07 times those in **c4s** and **c5s**, respectively, while those in **c4d** and **c5d** are 1.14 and 1.06 times those in **c5d** and **c6d**, respectively, at the 5-ps interval. The rate is greater than these at the 40-ps interval in the same order. In the single strand systems, the order and rate of fluctuations inversely correspond to those of the self-diffusion coefficient D (Table 1), suggesting that the fluctuation increases as the solute becomes hard to move. In the double strand systems, however, this suggestion is overturned with the same order in the value of fluctuations and diffusion.

The above result can be explained as following. Let us assume that the ratio of the number of interaction sites in cellooligosaccharides is in proportion to DP. Note that the number of interaction sites increases at almost the same rate with the surface area in the single strands of cellooligosaccharide. In the assumption, a solute fluctuation increases in proportion to the number of interaction sites when a solute is in an isolated state, while it decreases in inverse proportion to the number of sites when it is in an aggregating state. In other words, sugar−water interaction increases the atomic positional fluctuation, while sugar−sugar interaction decreases it. A balance between these two interactions determines the aggregation state of saccharides. We call this scheme the "site balance model". This model may be too simple to apply to any saccharides in which there are indeed various kinds of isomers and manner of linkage, but it is roughly true at least in the case of the double strands of tetra-, penta-, and hexacellooligosaccharides. More interaction sites lead to greater fluctuation in an isolated state but conversely confer an advantage over denser interaction between sugar chains in an aggregating state with smaller fluctuation. At DP

TABLE 3: Averaged Occupancy of the Intramolecular Hydrogen Bonds between Adjacent Residues in Cellotetraose, Cellopentaose, and Cellohexaose over Calculation Time[a]

	atom pair (%) single			atom pair (%) double		
	O2−O6′	O3−O6′	O3−O5′	O2−O6′	O3−O6′	O3−O5′
cellotetraose (**c4**)	1.45	4.92	78.17	1.47	3.73	81.12
cellopentaose (**c5**)	1.04	1.41	79.71	1.29	10.24	65.69
cellohexaose (**c6**)	1.10	3.08	77.79	2.01	7.71	88.72

[a] The range of the hydrogen bond angle O−H···O is from 90° to 180°, and the distance O···O is within 3.5 Å. The values are averaged over all pairs in each type from 100 to 1000 ps.

= 5 in cellooligosaccharide, there is considered to be a critical point at which an increasing power of fluctuation from a sugar−water interaction balances with a decreasing one from the sugar−sugar interaction. It is possible that the shape of cellooligosaccharide contributes to the validity of this model, being rigid and stable (see the next section) and thereby offering a constant number of interaction sites to others. More investigation will be needed to verify this model with regard to shape and flexibility in other kinds of saccharides.

3.3. Intramolecular and Intermolecular Hydrogen Bonds. Hydrogen bonding is one of the strongest factors determining the interaction or the static and dynamic states of saccharide chains. We have investigated intra- and intermolecular hydrogen bonds in the saccharide according to the following definition. A pair consisting of an oxygen atom and a hydroxyl group is regarded as making a hydrogen bond when (1) the distance between the two oxygen atoms is within 3.5 Å and also (2) the angle between O−H···O is from 90 to 180°. The distance and angle can well depict the characteristics of hydrogen bonding in aqueous saccharide solutions.[29−31] We herein describe a hydrogen bond O*x*−H*x*···O*y* or O*x*···H*y*−O*y* as O*x*−O*y*′, not distinguishing the former from the latter.

Table 3 shows the averaged occupancy of intramolecular hydrogen bonds in each sugar chain over from 100 to 1000 ps of calculation time. Intraresidue and interactions involving O4 were not included in the evaluation. There are three kinds of atom pairs, O2−O6′, O3−O6′, and O3−O5′, making hydrogen bonds between adjacent residues in a molecule. A chain in **c5d** has four other kinds of atom pairs (not shown) in addition to the above three because it bends between Ring 2 and 3 (the left chain in Figure 2h) unlike the other chains in each system. The atom pair O3−O5′ has a high value around 80% in every system except **c5d**, denoting that a flat ribbon surface in cellulose already appears in its oligosaccharides although it is twisted. As mentioned in Section 1, this sort of intramolecular hydrogen bond is the main contributor among all intra- and intermolecular hydrogen bonds to the firm structure of cellulose.[6] A twisted ribbon surface in cellohexaose is described schematically in Figure 8 with the intramolecular hydrogen bond between O3−O5′. The O3−O5′ pair has the highest value of 88.72% in **c6d** among all six systems. The firm flat ribbon surface in cellulose develops most in the double strand of cellohexaose although it is twisted, as compared with the other five systems. The double strand system of cellopentaose conversely has the lowest value of 65.69% for the O3−O5′ atom pair because of the bend mentioned above. The value of 81.12% in **c4d** is close to that of 78.17% in **c4s**, denoting the separation state of the double strand. It is hence possible that the balance between sugar−water and sugar−sugar interactions is at a critical point leading to instability in the ribbon surface in the cellopentaose double strand.

The values of the O3−O6′ pair increase as the strand becomes double from single in cellopentaose and cellohexaose, i.e., 1.41% to 10.24% and 3.08% to 7.71%, respectively. The values of the O2−O6′ pair also slightly increase as the strand changes from single to double: 1.04% to 1.29% and 1.10% to 2.01% in the cellopentaose and the cellohexaose, respectively. The O6 hydroxyl group can be attacked more by water molecules than other oxygen atoms because it is located outside a ring. We consider that the saccharide decreases its fluctuation by the uptake of more O6 in intramolecular hydrogen bonds in its aggregation states. From the molecular structure of cellooligosaccharides (Figure 1b), the O3−O6′ pair is always accompanied by an O3−O5′ pair with the *gt* position of O6 while the O2−O6′ pair depends on the *tg* conformation of O6. However, the values of the O3−O6′ pair do not coincide with those of the O3−O5′ in cellotetraose and cellopentaose. This result supports the above idea that more O6 atoms are taken into intramolecular hydrogen bonds in the aggregation states. The values of the O2−O6′ pair are smaller than those of the O3−O6′ in each system, although the *tg* conformation of O6 is dominant in native cellulose.[32,33] There is a possibility that the sugar−water interaction disturbs the O2−O6′ hydrogen bonds of cellooligosaccharides in an aqueous environment unlike in the crystalline form, resulting in a helical conformation.

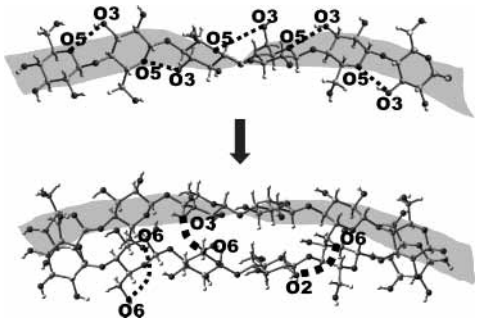

Figure 8. Above: A twisted ribbon structure through the intrahydrogen bond, O3−O5′, described as the gray-scaled twisted rectangle. Below: Intermolecular hydrogen bonds that frequently occur in the cellohexaose double strand. The tickness of the broken lines corresponds to the occupancy of atom pairs.

In Table 4, we show properties related to the intermolecular hydrogen bond between sugar chains in double strand systems. Herein intermolecular hydrogen bonds between sugar and water are not included. There are 77, 59, and 68 interaction sites of intermolecular hydrogen bonds in **c4d**, **c5d**, and **c6d**, respectively, in which a site means a position of the hydrogen bond pair composed of O*x* and O*y*. We distinguish each oxygen atom in sugar chains, for example, the pair O1−O6′ between Ring 1 and 2 is counted as a different site from the pair O1−O6′ between Ring 3 and 4. The ratio of 68 to 59 (1.15) is in good agreement with that of DP, 6 to 5 (1.2), supporting our conjecture in Section 3.2 that the number of interaction sites is approximately proportional to DP. The highest value of 77 in **c4d** is due to the unstable relative arrangement between the two sugar chains. The average occupancies per site are 0.64%,

TABLE 4: Properties Related to the Intermolecular Hydrogen Bond between Sugar Chains in Cellotetraose, Cellopentaose, and Cellohexaose[a]

	cellotetraose (**c4d**)	cellopentaose (**c5d**)	cellohexaose (**c6d**)
no. of interaction sites	77	59	68
av. occ. (%)[b]	0.64	2.93	4.37
dev. occ. (%)[c]	0.73	3.83	6.29
no./step[d]	0.49	1.73	2.97
O6 (%)[e]	59.0	84.1	97.2

[a] The range of the hydrogen bond angle O−H···O is from 90° to 180°, and the distance O···O is within 3.5 Å. The values are evaluated from 100 to 1000 ps. [b] Average occupancy over all intermolecular hydrogen bond pairs. [c] Standard deviation of occupancy over all intermolecular hydrogen bond pairs. [d] Number of intermolecular hydrogen bonds per MD calculation step. [e] The percentage of intermolecular hydrogen bonds involving O6.

TABLE 5: Percentages of Each Kind of Intermolecular Hydrogen Bonds between Sugar Chains in Cellotetraose, Cellopentaose, and Cellohexaose[a]

	atom pair (%)			
	O2−O6′	O3−O6′	O6−O6′	others
cellotetraose (**c4d**)	27.5	10.7	9.9	51.8
cellopentaose (**c5d**)	37.4	24.2	19.1	19.3
cellohexaose (**c6d**)	40.7	39.0	14.5	5.9

[a] The range of the hydrogen bond angle O−H···O is from 90° to 180°, and the distance O···O is within 3.5 Å. The values are evaluated from 100 to 1000 ps.

2.93%, and 4.37%, and the standard deviations of occupancy are 0.73%, 3.83%, and 6.29% in **c4d**, **c5d**, and **c6d**, respectively. The total number of intermolecular hydrogen bonds per MD calculation step in **c6d** is ca. 1.7 times that in **c5d**. These results mean that there are more intermolecular hydrogen bonds with more frequent occurrence per site in cellohexaose than in cellopentaose or cellotetraose. In addition, we have noticed a peculiarity of the oxygen atom, O6, in the interactive behavior of the oligosaccharides. As seen in the bottom line in the table, the percentage of the intermolecular hydrogen bond involving O6 is 97.2% in **c6d** versus 84.1% or 59.0% in **c5d** or **c4d**, respectively. Almost all intermolecular hydrogen bonds are made utilizing O6 in cellohexaose in addition to the abundance of intermolecular hydrogen bonds. This probably relates to the lower occupancy of intramolecular hydrogen bond O3−O6′ in **c6d** (7.71%) than **c5d** (10.24%, Table 3). As seen in Table 5, the percentage of the O3−O6′ among all kinds of intermolecular hydrogen bonds is ca. 15% higher in **c6d** than in **c5d**. Each strand of the cellohexaose double strand releases O6 from its intramolecular O3−O6′ into intermolecular O3−O6′ hydrogen bonds. We can summarize the above results that the stable twisted ribbon surface leads to a stable interaction site using O6 in the cellohexaose double strand (Figure 8).

Dynamics of the intermolecular hydrogen bonds between sugar chains is denoted in Figure 9 as their total number. Although the dynamics of each hydrogen bond shows no tendency (not shown), in total they manifest vibrational characteristics. In **c4d**, the number is zero after ca. 700 ps. In **c5d** and **c6d**, the period of vibration is approximately from 100 to 250 ps, which is quite long compared to the time range known for pure water dynamics. A hydrogen bond diminishes in ca. 1 ps at 283 K in water.[34] Intermittent collective molecular motions occur in water by the rearrangement of the hydrogen bond network on the order of 0.1−10 ps.[35−37] Considering these facts, the long-range vibration in intermolecular hydrogen bonding belongs not only to the dynamics of a hydrogen bond but also

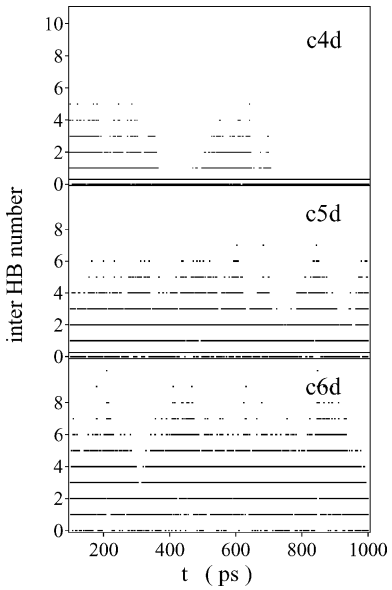

Figure 9. Time evolution in the total number of intermolecular hydrogen bonds between sugar chains in each double strand system from 100 to 1000 ps of calculation time.

to the cooperative motion of sugar−sugar interaction. Moreover, the vibration occurs not only in cellopentaose but also in cellohexaose where rmsd shows no such long-range fluctuation (Figure 7). The range is a little longer in cellohexaose than in cellopentaose, showing the stable aggregation state of the former. It is possible that this long-range vibration denotes some intrinsic properties of sugar−sugar interaction that depend on the number of hydroxyl groups, conformations, and relative arrangements of the sugar chain. Elucidating the property will also broaden the horizon of study in carbohydrate recognition in which sugar−sugar interaction plays an essential role.

4. Conclusions

The difference in aggregation states among cellotetraose, cellopentaose, and cellohexaose depends on the balance of sugar−water and sugar−sugar interactions. Sugar−water interaction promotes fluctuation of the saccharide chain in proportion to the number of interaction sites, while sugar−sugar interaction works in the opposite way. Cellopentaose is considered to be at a critical balance point between the two interactions. The "site balance model" used in the analysis can be adopted not only for cellooligosaccharide, but also for other various kinds of sugar−sugar interaction, although we need more precise investigation in other degrees of polymerization of cellooligosaccharide and other kinds of saccharides. The stability of a twisted ribbon surface composed of O3 and O5′ contributes to the stable aggregation state of the cellohexaose double strand, increasing the interaction with the other chain utilizing O6. This result is in contrast with another simulation study by Tanaka et al.; their cellulose chains do not have ribbon surfaces but loose tangles involving water molecules and do not directly hydrogen bond with each other in aqueous environments.[10] Considering that even cellohexaose is hardly soluble in water, it is considered that their result is a peculiar case with some peculiar conditions of calculation such as the high temperature (500 K) and the force field[38] that are different from ours.

Acknowledgment. We thank T. Endo and colleagues in our laboratory for useful discussions and criticism. This work is

supported by the Ministry of Education, Science and Culture, Japan (support of young researchers with a term to Y.Y.).

References and Notes

(1) Klemm, D.; Philipp, B.; Heinze, T. *Comprehensive Cellulose Chemistry*; Wiley-VCH: Weinheim, Germany, 1998.
(2) Saito, Y.; Okano, T.; Gaill, F.; Chanzy, H.; Putaux, J.-L. *Int. J. Biol. Macromol.* **2000**, *28*, 81.
(3) Endo, T.; Zhang, F.; Shinohara, Y. *Cellul. Commun.* **2002**, *9*, 86.
(4) Sarko, A.; Muggli, R. *Macromolecules* **1974**, *7*, 486.
(5) Yuguchi, Y. Unpublished data.
(6) Tashiro, K.; Kobayashi, M. *Polymer* **1991**, *32*, 1516.
(7) Okajima, K.; et al. *Polym. J.* **1992**, *24*, 71.
(8) Okajima, K.; et al. *Polym. Int.* **1992**, *29*, 47.
(9) Mattinen, M.-L.; Linder, M.; Teleman, A.; Annila, A. *FEBS Lett.* **1997**, *407*, 291.
(10) Tanaka, F.; Fukui, N. *Cellulose* **2004**, *11*, 33.
(11) French, A. D. *Cellulose* **2004**, *11*, 39.
(12) Bernardi, A.; Galgano, M.; Belvisi, L.; Colombo, G. *J. Comput. Aided Mol. Des.* **2001**, *15*, 117.
(13) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Eur. J. Biochem.* **1998**, *258*, 372.
(14) Cheong, Y.; Shim, G.; Kang, D.; Kim, Y. *J. Mol. Struct.* **1999**, *475*, 219.
(15) Shimada, J.; Kaneko, H.; Takada, T.; Kitamura, S.; Kajiwara, K. *J. Phys. Chem. B* **2000**, *104*, 2136.
(16) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
(17) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham, T. E. I.; Wang, J.; Ross, W. S.; Simmerling, C.; Darden, T.; Merz, K. M.; Stanton, R. V.; Cheng, A.; et al. *AMBER 7*; University of California, San Francisco: San Francisco, CA, 2002.
(18) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
(19) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
(20) Basma, M.; Sundara, S.; Calgan, D.; Vernali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125.
(21) Woods, R. J.; Chappelle, R. *J. Mol. Struct.* **2000**, *506*, 149.
(22) Bernet, B.; Xu, J.; Vasella, A. *Helv. Chim. Acta* **2000**, *83*, 2072.
(23) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(24) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(25) International Union of Biochemistry and Molecular Biology. *Biochemical Nomenclature and Related Documents*, 2nd ed.; Portland Press: London, UK, 1992.
(26) French, A. D.; Johnson, G. P. *Cellulose* **2004**, *11*, 5.
(27) Sugiyama, H.; Hisamichi, K.; Usui, T.; Sakai, K.; Ishiyama, J. *J. Mol. Struct.* **2000**, *556*, 173.
(28) Toda, M.; Matsuda, H.; Hiwatari, Y.; Wadachi, M. *Structure and properties of Liquid* (in Japanese); Iwanami shoten: Tokyo, Japan, 1976.
(29) Umemura, M.; Hayashi, S.; Nakagawa, T.; Yamanaka, S.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct.* (*THEOCHEM*) **2003**, *624*, 129.
(30) Umemura, M.; Hayashi, S.; Nakagawa, T.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct.* (*THEOCHEM*) **2003**, *636*, 215.
(31) Umemura, M.; Hayashi, S.; Nakagawa, T.; Urakawa, H.; Kajiwara, K. *J. Mol. Struct.* (*THEOCHEM*) **2003**, *639*, 69.
(32) Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002**, *124*, 9074.
(33) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300.
(34) Stillinger, F. H.; Rahman, A. *J. Chem. Phys.* **1974**, *60*, 1545.
(35) Ohmine, I.; Saito, S. *Acc. Chem. Res.* **1999**, *32*, 741.
(36) Matsumoto, M.; Ohmine, I. *J. Chem. Phys.* **1996**, *104*, 2705.
(37) Sasai, M.; Ohmine, I.; Ramaswamy, R. *J. Chem. Phys.* **1992**, *96*, 3045.
(38) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.O6.