ELSEVIER

Carbohydrate Research 281 (1996) 11–34

CARBOHYDRATE RESEARCH

Molecular dynamics simulations of maltose in water [1]

Karl-Heinz Ott *, Bernd Meyer [2]

Complex Carbohydrate Research Center and Department of Biochemistry and Chemistry, The University of Georgia, 220 Riverbend Road, Athens, GA 30602, USA

Received 21 February 1995; accepted 26 August 1995

Abstract

The conformational preferences and flexibility of α-maltose were calculated using a modified GROMOS potential energy function. Six molecular dynamics (MD) simulations of α-maltose with explicit inclusion of water were run for up to 1000 ps each, starting from different conformations of the glycosidic linkage, the hydroxymethyl groups, and the hydroxyl groups. Comparison of calculated ensemble-averaged optical rotations with experimental data demonstrated an excellent agreement. Reasonable agreement was also found between experimental and calculated time-averaged NMR parameters. The global minimum centered at $\boldsymbol{\varphi/\psi}$ [3] = −49°/−36° was populated to more than 90% of the time. The second minimum with ≈ 11 kJ/mol higher potential energy at $\boldsymbol{\varphi/\psi}$ = −29°/−173° represented the inverted conformation. Transitions between both minimum regions were only found from the local into the global minimum within the total simulation time of 5400 ps. The overall ratio between the staggered populations of the hydroxymethyl groups is *gg*:*gt*:*tg* = 70:23:6 for the reducing glucose, and *gg*:*gt*:*tg* = 55:38:6 for the nonreducing residue. Average lifetimes of the rotamers of the hydroxymethyl groups range between 5 ps for *tg* and 700 ps for *gg* conformers. The lowest energy barriers between the rotameric populations were estimated to be between 12 kJ/mol and 16 kJ/mol. Cluster analysis techniques employed to analyze the large amount of data in multidimensional space revealed correlations between the glycosidic linkage and the staggered forms of the hydroxymethyl groups. Interresidue hydrogen bonds that were found in less than 5% of all conformations occurred when the glycosidic linkage adopted conformations with relatively high φ and ψ values. Several additional minima that had been found from in vacuo studies are not stable once water is included since the importance of

* Corresponding author. Present address: American Cyanamid Co., P.O. Box 400, Princeton, NJ 08543, USA.

[1] This work was supported in part by a grant from the scientific committee of the NATO via the DAAD.

[2] Present address: Institut für Organische Chemie, M.L. King Platz 6, 20146 Hamburg, Germany.

[3] Dihedral angles φ = H-1′–C-1′–O-1′–C-4, ψ = C-1′–O-1′–C-4–H-4. Bold font indicates mean values.

SSDI 0008-6215(95)00320-7

intramolecular hydrogen bonds is drastically reduced. Methodological aspects and the efficiency of MD simulations for reaching conformational equilibria are discussed.

Keywords: Conformation; Cluster analysis; Dynamics; Energy minima

1. Introduction

Maltose, one of the most abundant disaccharides in nature, is a cleavage product of the starch components amylose and amylopectin, carbohydrate storage forms of plants. Investigation of the conformational preferences of maltose may explain amylose structure and structural irregularities that influence the physical and chemical properties of amylose. Furthermore, it is an ideal model for the α-(1 $\rightarrow$ 4) glycosidic linkage. From previous experimental and theoretical analyses of maltose [1–13], the strength of the intramolecular hydrogen bonding and the number, location, and relative energy of local energy minima in the conformational space of maltose remain controversial.

In our attempt to establish and validate an empirical force field for carbohydrates and to analyze the requirements and possibilities of MD simulations for the detailed conformational analysis of solvated carbohydrates, we performed a series of different extended MD simulations. The incorporation of water in the simulations is mandatory for a reasonable description of conformations and entropy as we demonstrated in a previous publication [14] that also describes the force field parametrization. A high temperature MD simulation at 600 K showed that only two local minima are present in the φ/ψ space of maltose in water as described by the modified GROMOS force field. However, at 600 K, artifacts from imbalances in the force field became unacceptable. In this paper we report, as a consequence, MD simulations performed at 300 K that extended to up to 1000 ps. Starting from different conformations we expected a comprehensive sampling of the conformational space. Locations and entropy of local minima, their energy differences, resulting population distributions, transition probabilities, and associated activation energies are derived. The efficiency of simulation protocols and parameters, as well as biases from start conformations are evaluated. Cluster analysis is employed and its use for analyzing large amounts of higher-dimensional conformational data is discussed.

2. Experimental procedures

Details of the computational procedures and programs used are described in Ott and Meyer [14]. Six different starting conformations were generated by setting the *exo-* and *endo*-cyclic torsion angles to values obtained from previous simulations [6,7]. After a few additional energy minimization (EM) steps, the maltose conformers were 'solvated' by fitting them into a truncated octahedron box with 160 water molecules. The resulting ensemble was energy minimized using steepest descent and conjugate gradient EM with position restraining for the solute atoms using a harmonic potential with a force constant of 6250 kJ/mol^1 nm^2. This position restraining potential was also applied during each

of the first 50 ps of the following MD simulations, which were therefore excluded from the analysis.

For the MD simulations, initial velocities were assigned randomly according to a Maxwell distribution at 300 K. This virtual temperature was held constant by coupling the velocities of the atoms to an external bath with a relaxation constant of 0.1 ps [15]. Solvent and solute atoms were scaled separately. The pressure of the system was kept at 1.013 bar by isotropically coupling the atom positions and the size of the computational box to a pressure bath with a relaxation constant of 0.5 ps [15]. The width of the integration steps of the MD simulations was 2 fs at a 64 bit precision of the computer word. Fast atomic motions due to bond length vibrations were removed and bond lengths were held constant with a tolerance of 0.001 nm using the SHAKE algorithm [16]. Nonbonded interactions were calculated on the basis of an atom pair list that was updated every 25 integration cycles with a cut-off radius of 0.8 nm under periodic boundary conditions. The concept of charge groups assured electrical neutrality while calculating the electrostatic interactions with a dielectric constant of one. The energy values of the different force field terms and the coordinates were stored every 100 fs for further analysis.

The cluster analysis [14,17,18] grouped the elements of the trajectories into families of related conformations [19]. Each conformation is represented as a vector in the multidimensional space spanned by cosine and sine values of φ, ψ, and both ω dihedral angles. The relation between conformations is expressed by the correlation coefficient between all pairs of conformations ('complete clustering'). A hierarchical cluster tree is built by starting out with single conformations as vertices and successively defining new vertices by unifying the most related vertex pairs. In a first step, a subset of 2500 snapshots was selected by using every 20th record of the pooled dataset of all simulations and clustered [19]. The resulting cluster tree was analyzed to select a suitable cutting level for the cluster tree such that, for each of the selected dihedral angles, a single region in the conformational space was assigned. In the second step, the remaining conformations were assigned into the most appropriate cluster. These grouped datasets were analyzed statistically and graphically.

3. Results

Six MD simulations of maltose, enumerated 1–6, in boxes containing 160 water molecules were performed. Each was started in a different conformation. The selection of the starting conformations and the naming conventions of different maltose conformations are based on earlier reports of Ha et al. [6] and Tran et al. [7] that describe empirical force field calculations of maltose in vacuo (Table 1). Using position restraining for the atoms of the solute for the first 50 ps of each simulation allowed the water configuration to equilibrate in a conformation that may stabilize the preset conformation of the maltose. The first two simulations, runs 1 and 2, started near the global minimum found by Ha et al. [6] with run 2 having a different setting for the hydroxyl and hydroxymethyl dihedral angles that was taken from minimum F of Ha et al. [6] where the glycosidic linkage is in an inverted conformation (cf. Table 1). Run 3

Table 1
Overview of the setup of the MD simulations, runs 1–6 [a]

MD run	1	2	3	4	5	6
φ (°)	−56	−50	−24	−42	22	24.1
ψ (°)	−28	−55	−276	−180	−143	−222.8
Conf	A	A	D	E/F	G	–
$\omega_1-\omega_2$	*gg–tg*	*tg–tg*	*tg–tg*	*tg–tg*	*tg–tg*	*gt–tg*
Time (ps)	1000	1000	1000	400	1000	1000
Temp. (K)	300	300	300	300	300	300
O-5–C-1–O-1–OH-1	−159.2	−15.6	−18.6	62.0	−147.8	−18.1
C-1–C-2–O-2–OH-2	169.7	49.1	48.3	16.0	75.1	51.7
C-2–C-3–O-3–OH-3	177.6	−41.7	−38.4	−46.0	−53.3	−48.4
C-5–C-6–O-6–OH-6	63.7	81.8	50.6	55.0	49.6	60.7
C-1–C-2–O-2–OH-2	−141.3	−163.5	36.5	−169.0	−159.3	−68.4
C-2–C-3–O-3–OH-3	166.5	168.5	−43.2	161.0	165.2	−40.6
C-5–C-6–O-6–OH-6	−61.3	−71.8	37.5	−78.0	−174.5	41.7
C-3–C-4–O-4–OH-4	173.8	−164.8	41.1	−147.0	−150.4	42.4

[a] Starting values for the dihedral angles φ (= H-1′–C-1′–O-1′–C-4), ψ (= C-1′–O-1′–O-4–H-4), ω_1 (reducing glucose), ω_2 (nonreducing glucose), and of all the hydroxyl groups are given. The letters in row 'Conf' indicate the location of the start conformation in the φ/ψ space according to the nomenclature of Ha et al. [6] and Tran et al. [7]. The total simulation times of the runs are listed in row 'Time'. In all runs, the first 50 ps were run under position restraining for the atoms of the maltose, while the 160 water molecules were allowed to equilibrate.

started in a local minimum that was only found by Tran et al. [7] and is of interest because it could serve as an intermediate for transitions between the global and the side minimum. The inverted conformations of the glycosidic linkage were investigated in runs 4 and 5. Run 6 served as a control with its initial conformation far removed from the local minima.

As a summary of the results, the trajectories for the φ, ψ, and both ω angles are depicted in Figs. 1 and 2. Corresponding averages, standard deviations and ranges are listed in Tables 2 and 3. Energy statistics are summarized in Table 4. Typically, the total energy of the maltose/water system decreased rapidly during the first few picoseconds and was stable after 50 ps where data analysis started. Additional evidence that the systems are in thermal equilibrium comes from the facts that there was no heat to be moved into the 'temperature bath' during the free MD runs (except run 2, see below) and that the size of the calculation boxes remained constant.

During the equilibrium phase of the first MD simulation, from 50 ps to 1000 ps, the maltose did not leave the area of the global minimum for the glycosidic linkage conformation. The hexose ring conformations were fluctuating around $^4C_1$ conformations with RMS-deviations of less than 9° for the endocyclic dihedral angles. The hydroxymethyl group of the reducing monosaccharide remained in the *gg* conformation for 95% of the simulation. The hydroxymethyl group at the nonreducing end showed a considerably higher flexibility with 18 transitions between all three rotamers (Fig. 2).

Intramolecular hydrogen bonds of significant lifetime of several hundred femtosec-

onds were only found in 6% of the time between the two OH-6 groups. Both OH-6 act as a donor and as an acceptor for hydrogen bonds (Fig. 3). The other hydroxyl groups of the maltose were surrounded by water molecules that are the partners for hydrogen bond interactions. The total portion of the time that a hydroxyl group is hydrogen bonded to an individual water molecule did not exceed 3% of the total MD simulation. Similar observations regarding the pyranose ring conformations and hydroxyl orientations are found for all other runs. Hydrogen bond patterns varied between the runs. However, in general, hydrogen bonds were found only a small percentage of the time, mostly between O-3 and O-2′ and between the primary hydroxyl groups. The hydrogen bond lifetimes were typically very short and rarely exceed 1 ps.

The second MD simulation, run 2, was tailored to examine the influence of the initial setting of the hydroxyl and hydroxymethyl group orientations on the progress of the MD run. Although the same conformational space of the glycosidic linkage is occupied in runs 1 and 2, and the hydroxyl groups rearrange very rapidly and are similarly distributed over the different rotamers in both runs, these simulations differ significantly in terms of their hydroxymethyl group trajectories, and the average total energy, which is ≈ 9 kJ/mol lower in run 1.

To test the stability of the local minimum D that was described by Tran et al. [7] but was not apparent in the modified AMBER [6] or modified GROMOS [14] force field studies, a third MD simulation was started at $\varphi = -5°$, $\psi = 84°$ (Fig. 4). The starting conformation immediately changes after the release of the position restraints and occupies for 20 ps a local minimum with $\varphi/\psi \sim 24°/18°$, close to minimum C found in the in vacuo simulations [6,7]. The next 10 ps are characterized by a transition to the region of the global minimum that is reached at 85 ps and not left any more. This conformational change was accompanied by decreases of ≈ 20 kJ/mol in the energy values of each, the dihedral, the van der Waals, and the bond angle terms. The

Table 2
Characteristic values for the φ/ψ angles for each MD run [a]

Run	φ				ψ			
	Avg	Std	Min	Max	Avg	Std	Min	Max
1	−48	14	−112	2	−35	12	−90	24
2	−50	15	−99	−1	−37	14	−109	15
3	−47	19	−105	65	−34	16	−88	79
4	−16	24	−103	8	−13	60	−180	180
5	−42	16	−101	25	−49	91	−180	180
6	−47	16	−100	17	−39	31	−180	180
Run	φ (150–240)				ψ (−120–150)			
	Avg	Std	Pop	%	Avg	Std	Pop	%
4	183	8	1097	12	−33	9	8403	88
5	188	9	3256	34	−36	13	6244	66

[a] For each MD simulation (Run), the minimum (Min), the maximum (Max), the average values (Avg) and the standard deviations (Std) are tabulated. The averages, standard deviation, population (Pop), and the percentage population (%) of the ψ-values for runs 4 and 5 are shown for the whole MD simulation as well as being split into the main minimum, from $\psi = -120°$ to 150°, and local minimum, from $\psi = 150°$ to 240°.

electrostatic interaction energy between the two glycosyl residues was temporarily decreased by ≈ 40 kJ/mol. At the same time, the volume of the box increased by 4%. Around this time, and in a 14 ps time interval around 600 ps, HO-3 forms a hydrogen

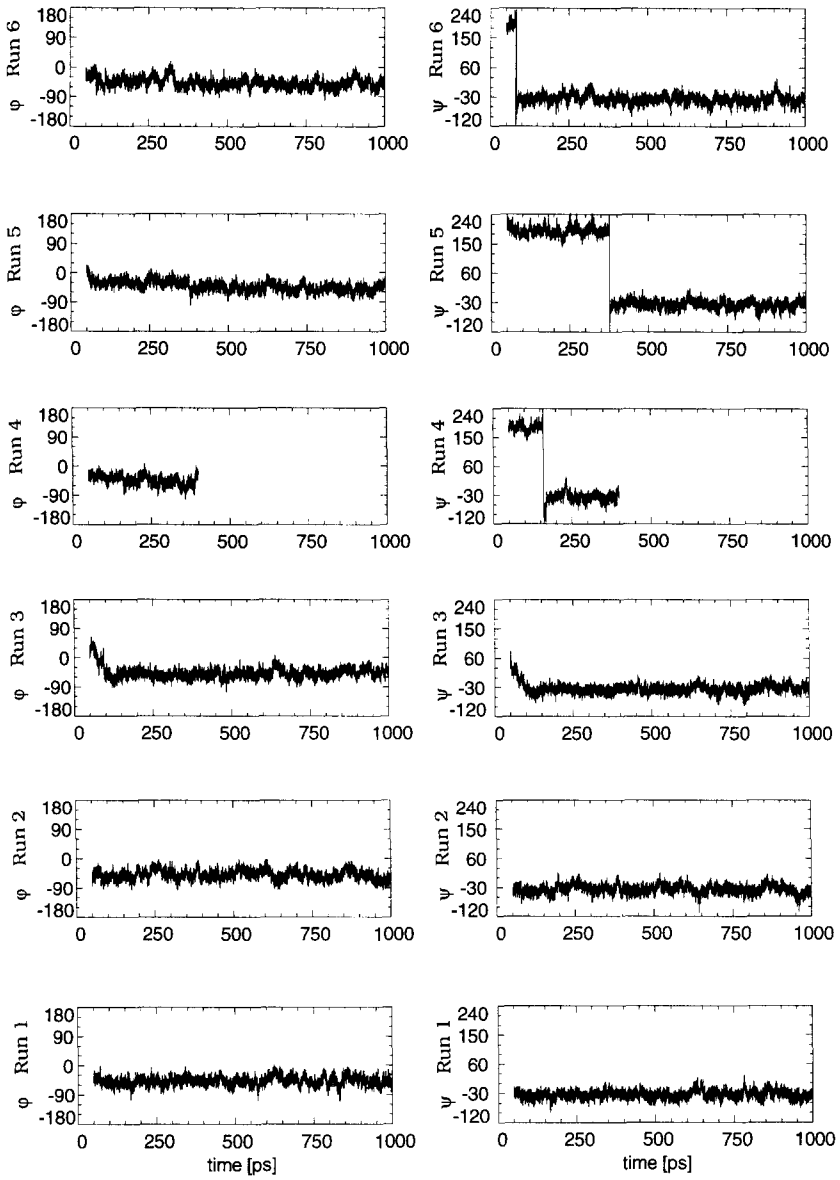

Fig. 1. Trajectories of the φ/ψ angles of MD runs 1–6 (bottom to top). Left: φ trajectories; right: corresponding ψ trajectories. The ψ angles are scaled between $-120°$ and 240°, the φ angles are scaled from $-180°$ to 180°. MD runs 1, 2, 3, 5, and 6 were run for 1000 ps and MD run 4 was run for 400 ps. The first 50 ps that were run under conformational restraints of all maltose atoms to relax the water box are not shown.

bond with O-2′ when the glycosidic linkage adopted a conformation of $\varphi/\psi \approx -20°/-15°$. The higher proportion of this hydrogen bond is also reflected in a significantly higher population of the C-2–C-3–O-3–HO-3 dihedral angle in the interval between −180° to −60° from which OH-3 can potentially act as a hydrogen bond donor.

Run 4 was started from the 'inverted' conformation of the glycosidic linkage where one glucose ring is turned upside-down with respect to the normal orientation found in the global minimum (Figs. 4 and 5). This conformation was stabilized in vacuo by hydrogen bonds between O-2′ and OH-6 that is also present during the first 100 ps free

Table 3
Statistics of the hydroxymethyl group's conformation [a]

Run	Rot	ω_1				ω_2			
		Pop	% Pop	Avg	Std	Pop	% Pop	Avg	Std
1	*gt*	237	2	66	15	5533	55	65	15
	tg	169	2	170	15	1024	10	147	19
	gg	9594	96	294	12	3443	34	295	12
2	*gt*	3166	32	69	18	8109	81	65	15
	tg	1890	19	159	15	1610	16	164	15
	gg	4944	49	294	12	281	3	292	16
3	*gt*	6	0	109	6	1265	13	63	15
	tg	243	3	173	17	690	7	166	13
	gg	9251	97	294	12	8045	80	295	13
4	*gt*	3156	90	69	16	869	25	67	14
	tg	344	10	154	17	176	5	167	19
	gg	0	0			2455	70	294	12
5	*gt*	5038	50	67	16	2322	23	65	15
	tg	1219	12	156	16	1314	13	169	17
	gg	3743	37	294	13	6364	64	295	12
6	*gt*	274	3	72	18	1479	16	66	15
	tg	904	9	167	16	395	4	163	15
	gg	8822	88	294	12	7626	80	295	12
Global mini-mum	*gt*	10581	24	67	17	15875	36	65	15
	tg	2251	5	157	17	2440	5	162	17
	gg	31676	71	294	12	26193	59	295	12
Local mini-mum	*gt*	979	23	70	15	2439	56	65	15
	tg	441	10	159	18	414	10	163	18
	gg	2907	67	294	12	1474	34	294	14
Pooled data set	*gt*	11877	23	68	17	19551	38	65	15
	tg	3267	6	159	18	3233	6	162	17
	gg	35854	70	294	12	28214	55	295	12

[a] For each of the six MD simulations, the population (Pop), the percentage population (% Pop), the average dihedral angle (Avg) and its standard deviation (Std) of the hydroxymethyl orientations of the reducing glucose (ω_1) and the nonreducing glucose (ω_2) are listed. The statistical analysis from data of all six runs are also tabulated separately for conformations that are in the region of the global and of the local minimum of the glycosidic linkage. Additionally, the statistical analysis of the whole pooled data set of all six simulations is added. For the analysis, the bins for the three staggered conformations are defined by the regions from 0°–120° (*gt*), 120°–240° (*tg*), and 240°–360° (*gg*).

Table 4
Average values of the individual energy contributions for the six different MD simulations 1, 2, 3, 4, 5, and 6 in kJ/mol [a]

Run	Time (ps)	E_{bond}	$E_{el\text{-}m}$	$E_{LJ\text{-}m}$	$E_{el\text{-}m\text{-}s}$	$E_{LJ\text{-}m\text{-}s}$	$E_{sum\text{-}m}$	$E_{LJ\text{-}s}$	$E_{el\text{-}s}$	$E_{p\text{-}tot}$	$E_{k\text{-}tot}$	$Std_{E\text{-}tot}$
1	50–1000	123.3	578.9	−23.5	−461.3	−101.1	116.2	1157.1	−7662.5	−6389.2	1265.3	65.2
2	50–1000	123.5	577.3	−21.6	−455.9	−101.2	122.0	1155.9	−7656.6	−6380.4	1265.7	65.9
3	50–1000	124.7	575.0	−22.3	−460.9	−100.3	116.2	1156.1	−7660.8	−6388.4	1266.0	65.0
4	50– 400	122.9	580.4	−19.8	−463.0	−100.9	119.6	1158.1	−7668.5	−6390.9	1266.4	69.9
5	50–1000	124.8	579.0	−22.6	−463.2	−100.2	117.8	1157.3	−7662.7	−6387.5	1265.6	64.3
6	50–1000	122.6	575.5	−23.2	−461.1	−99.2	114.6	1154.5	−7657.7	−6388.5	1265.8	67.8
1–6	Avg	123.8	578.1	−22.0	−460.9	−100.8	118.4	1156.9	−7662.2	−6387.3	1265.8	
4	50– 150	128.1	573.2	−18.7	−440.8	−105.7	136.1	1162.6	−7687.3	−6388.6	1264.9	74.0
4	200– 400	121.2	582.6	−20.2	−471.9	−99.0	112.7	1154.7	−7657.6	−6390.2	1267.1	70.5
5	100– 350	129.1	582.0	−21.8	−461.3	−101.6	126.4	1155.1	−7663.1	−6381.5	1264.8	63.3
5	400–1000	122.0	578.9	−23.3	−465.8	−99.2	112.6	1158.8	−7663.7	−6392.2	1265.9	n.d.

[a] The sum of all bonded interactions are labeled E_{bond}. Shown are also the sum of all electrostatic interactions within the maltose ($E_{el\text{-}m}$), between the maltose and water ($E_{el\text{-}m\text{-}s}$), and within the water ($E_{el\text{-}s}$). The sum of all van der Waals interactions within the maltose are labeled $E_{LJ\text{-}m}$, between the maltose and water $E_{LJ\text{-}m\text{-}s}$, and within the water $E_{LJ\text{-}s}$. The sums of the interaction energies E_{bond}, $E_{el\text{-}m}$, $E_{el\text{-}m\text{-}s}$, $E_{LJ\text{-}m}$, and $E_{LJ\text{-}m\text{-}s}$ are labeled $E_{sum\text{-}m}$, the total potential energies are under $E_{p\text{-}tot}$, and the kinetic energies are in column E_{kin}. $S_{E\text{-}tot}$: standard deviation for total energy. First seven rows: averages over the whole unrestrained part of the MD simulations, i.e., 50 ps to the end. Last four rows: subaverages for the simulations 4 and 5 that correspond to the local minimum with the inverted conformation of the glycosidic linkage (50–150 ps in run 4 and 100–350 ps in run 5) and for the parts of simulations 4 and 5 during which the global minimum was occupied.

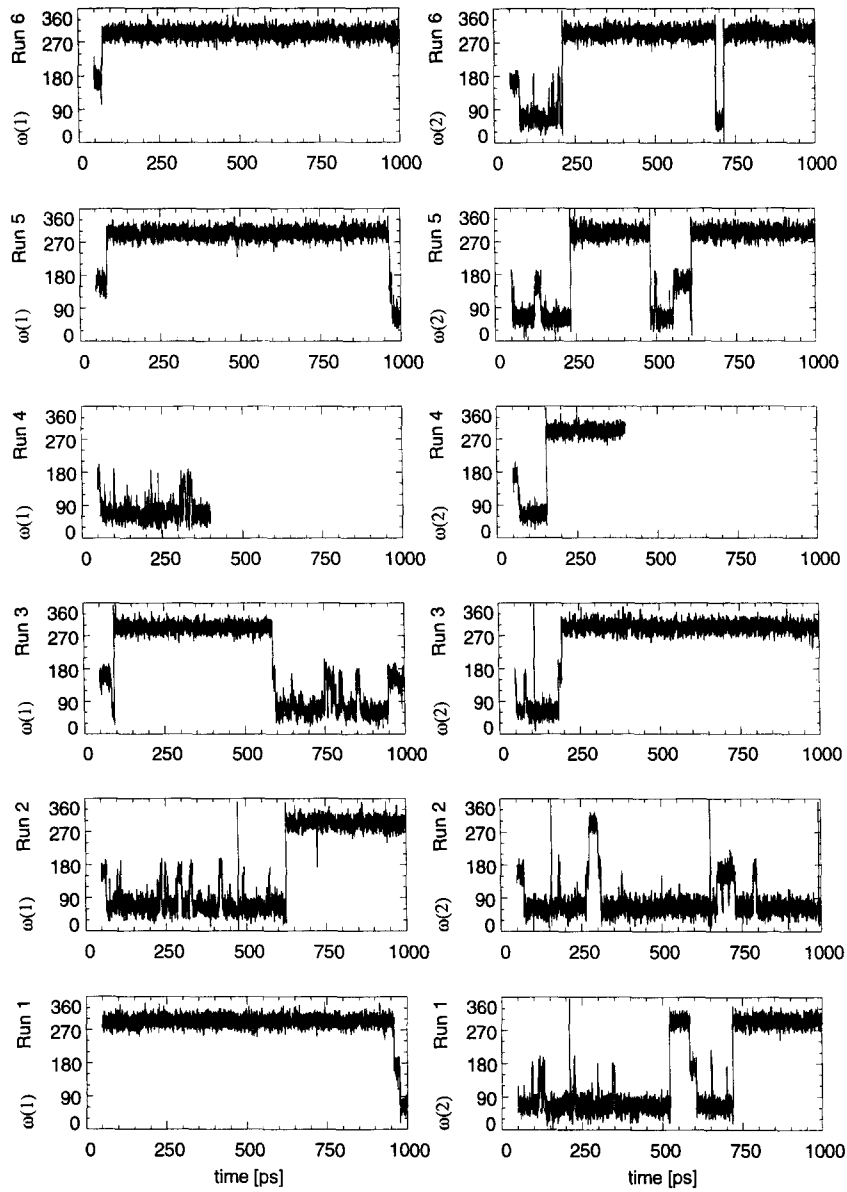

Fig. 2. Trajectories of the ω_1/ω_2 angles of MD runs 1–6 (bottom to top). Left: ω_1 trajectories of the reducing glucose; right: ω_2 trajectories of the nonreducing glucose. The ω angles are scaled between 0° and 360°. MD runs 1, 2, 3, 5, and 6 were run for 1000 ps and MD run 4 was run for 400 ps. The first 50 ps that were run under conformational restraints of all maltose atoms to relax the water box are not shown.

MD simulation where the local minimum with $\varphi/\psi = -33°/-178°$ is sampled. The molecule underwent a slow conformational change between the 159th ps and the 170th ps (Fig. 6). In a first fast transition, an intermediate conformation at $\varphi/\psi = 120°/-77°$

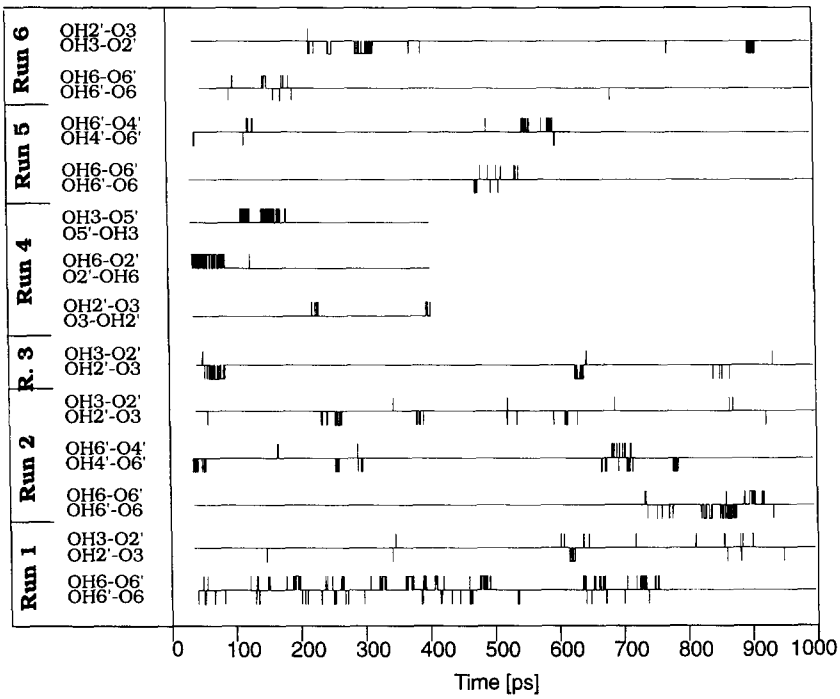

Fig. 3. Hydrogen bond trajectories obtained from the six MD simulations 1–6 (bottom to top). OH-2′/OH-3 hydrogen bonds are shown for all runs. Trajectories for the other H-bonds are shown if the H-bonds occurred more than 1% of the time. The hydrogen bonds are listed for pairs of hydroxyl groups, where the donor is marked as OH-x and the acceptor as O-x. The central horizontal lines represent the conformations without hydrogen bonding. Deviations from this line indicate the occurrence of hydrogen bonds. Hydrogen bonds were selected by a combined criterion using the donor–acceptor distance and the donor–hydrogen–acceptor angle. It is obvious that hydrogen bonds are relatively instable. A correlation between these hydrogen bond 'trajectories' and the φ/ψ trajectories indicates that the OH-2′–O-3 hydrogen bonds occur preferably at higher φ/ψ values (Fig. 1).

was reached that was stable for 10 ps. A second fast conformational transition led to the final conformation with $\varphi/\psi = -49°/-35°$. This transition was accompanied by a change in the total potential energy of only 1.6 kJ/mol, and was mainly driven by the decrease in the electrostatic interaction between solute and solvent by 31 kJ/mol. At the same time, the electrostatic energy within the water and within the solute increased by 30 and 9 kJ/mol, respectively.

The start conformation G was left within the first few picoseconds of the unrestrained part of MD run 5. The local minimum with the inverted conformation discussed in run 4 was accessed during the next 300 ps. A rapid transition of the glycosidic linkage conformation led to the region of the global minimum which was stable for the rest of the MD run (cf., Fig. 4). The total potential energy of the ensemble averaged over 400 to 1000 ps was 11 kJ/mol lower than the corresponding average for the the time from 75 to 375 ps. The transition at about 400 ps was accompanied by a 7 kJ/mol reduction of

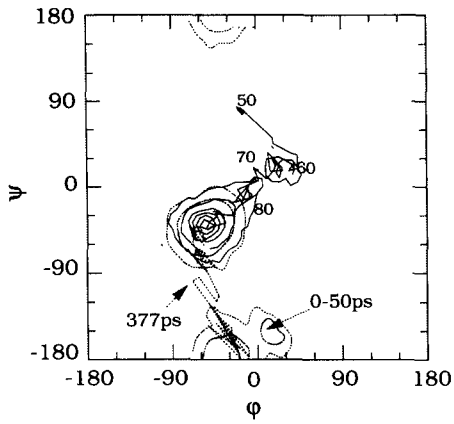

Fig. 4. Contour plots of the φ and ψ angles for MD runs 3 (solid lines) and 5 (dotted lines) from 50–1000 ps. The conformations are pooled into 6° bins. The contour lines represent populations (outside to inside) of 1%, 10%, 30%, 50%, 70%, and 90% relative to that of the highest populated bin, which was normalized to 100%. For run 5, only the 1% and 10% contours are shown. The solid line connects points of the averages over 1 ps intervals of run 3 between 45 ps and 80 ps. The transition from the local minimum of the inverted conformation at $\psi \approx 180°$ into the global minimum is depicted with a dotted line connecting conformations observed in 100 fs steps between 375 and 380 ps.

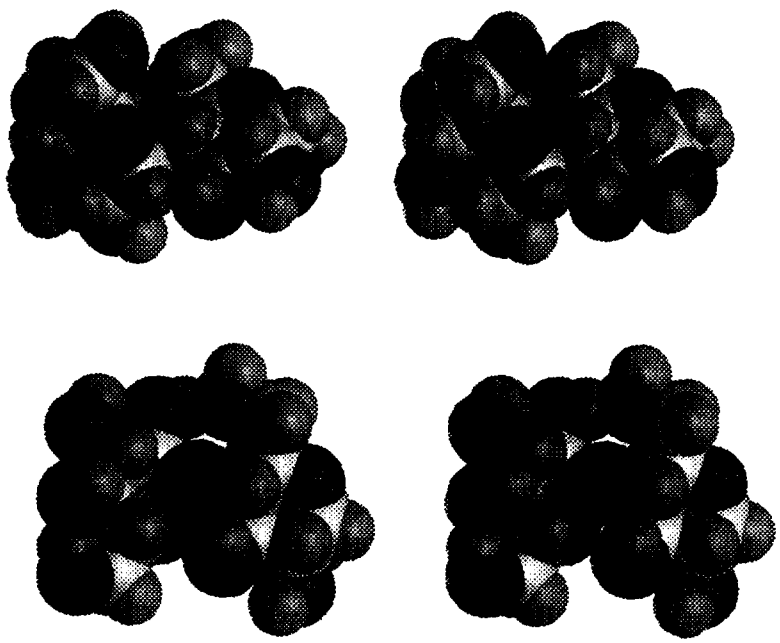

Fig. 5. Stereo-plot of CPK models of maltose in normal (top) and inverted conformation (bottom). The reducing end is on the right.

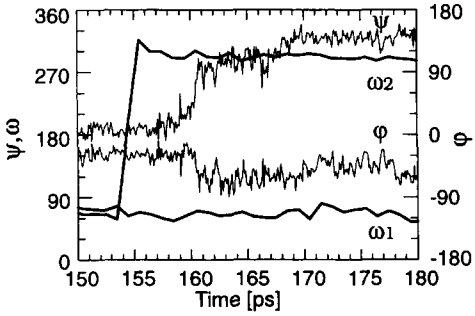

Fig. 6. Trajectory of the φ, ψ, and ω angles between 150 and 180 ps in MD run 4. For clarity, 1 ps averages are depicted for the ω trajectories, while for the φ and ψ trajectory, the data points are plotted in 100 fs intervals.

the bonded interaction energy within the maltose. The electrostatic interaction energies were reduced by 8 kJ/mol, while van der Waals interactions increased by ≈ 4 kJ/mol.

In run 6, which was started from a high energy conformation, the global minimum was reached in two conformational transitions through a metastable intermediate conformation of the inverted type (Fig. 7). The first transition was characterized by a fast change of the φ angle from the initial value of 13° to $-24°$ and a slow change of the ψ angle from 148° to $\approx 190°$ that took 10 ps. After only 16 ps, a second transition from 76 to 86 ps led to the global minimum. Nine picoseconds before the glycosidic linkage changes the first time, the hydroxymethyl group of the reducing glucose showed a transition from *tg* to *gg*. At about the same time, the ω angle of the nonreducing unit jumped from its *tg* conformation into a *gt* conformation.

Clustering of the conformations in the four dimensional space of the dihedral angles φ, ψ, ω_1, and ω_2 was performed to evaluate more detailed possible correlations between the conformation of the glycosidic linkage and the orientations of the hydroxy-

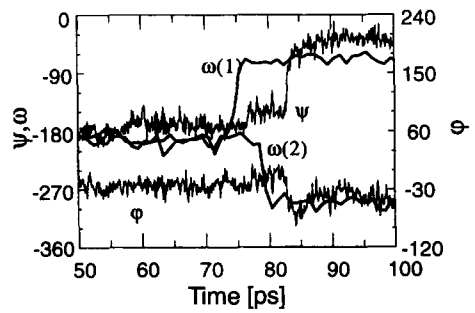

Fig. 7. Trajectories between 50 and 100 ps of run 6. The left y-axis scaled from 0° to 360° is for the ψ, ω_1, and ω_2 angles. The axis for the φ angle is shown at the right side and is scaled from $-120°$ to 240°. The trajectories for the ω angles are shown as 1 ps averages. It is clear that no direct correlation exists between the transition of the ω angles and the φ/ψ transitions. However, an indirect correlation mediated through water rearrangement might exist.

methyl groups. In addition, it served as a test case for higher dimensional clustering, because the results can easily be visualized and compared with sorting procedures. The analysis of the cluster tree resulted in 28 clusters, which have standard deviations of the dihedral angle distributions below 20°. This cut-off ensures a complete separation of local minima for all relevant dihedral angles.

The clusters can be grouped according to the conformation of the glycosidic linkage (Table 5). The major group (M) contains 18 clusters with the conformation of the glycosidic linkage near the global minimum, i.e., $\varphi/\psi \approx -49°/-36°$. A second group (I) contains seven clusters that represent conformers with an inverted conformation of the glycosidic linkage at $\varphi/\psi \approx -29°/173°$. Two additional clusters (#26 and #27) with a cumulative population of 0.65% form the third group (P) that is characterized by conformations of the glycosidic linkage with positive φ/ψ values. The fourth group (T) formed by cluster #28 holds only 0.2% of the total population and represents the transition between the main minimum (M) and the side minimum (I) with a ψ population from $-75°$ to $-190°$.

Within the main group (M), the clusters that have the same conformations of the hydroxymethyl groups can be combined, resulting in nine subgroups representing all the possible permutations of *gg*, *gt*, and *tg* of both hydroxymethyl groups (Table 5). The most populated cluster #1 with 19,199 members belongs to the first group (M). Conformers in cluster #1 have typically a long average lifetime of hundreds of nanoseconds. Most of the transitions out of this cluster or into this cluster are *gg–gt* transitions of the nonreducing hydroxymethyl group (Figs. 8 and 9).

Cluster #2, with 6358 members and cluster #3 with 5247 members together occupy a nearly identical dihedral angle space as cluster #1, but they differ from cluster #1 by their *gt* orientation in the nonreducing hydroxymethyl group. A positive correlation between the ψ and the $\omega(1)$ angles is indicated from contour plots (not shown) of the two clusters in $\psi/\omega(1)$ space. In this projection, the two clusters are separated by a

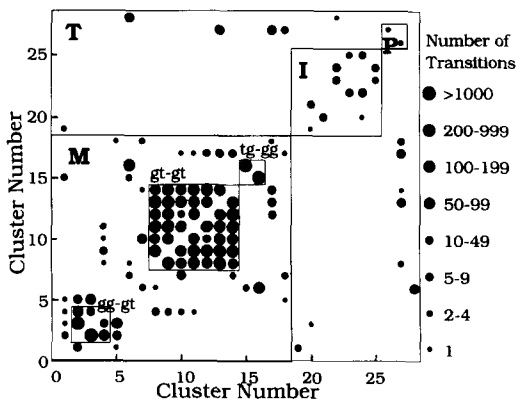

Fig. 8. Representation of the number of transitions between the clusters listed in Table 5 in a bubble plot. The number of transitions is indicated by the radius of the bubbles. The boxes are drawn according to the grouping of the clusters as explained in the text.

Table 5
Results of the cluster analysis of the $\varphi/\psi/\omega_1/\omega_2$ dihedral angle space for the pooled data set from all six MD simulations [a]

Cluster #	Pop	Sub %	Tot %	φ		ψ		ω_1		ω_2		Rot	
				Avg	Std	Avg	Std	Avg	Std	Avg	Std	$\omega_1-\omega_2$	
1	19199	41.8	38	−49	14	−36	13	−66	12	−65	12	*gg* – *gg*	Group M
2	6358	13.8	12	−49	14	−30	10	−61	11	61	14	*gg* – *gt*	
3	5247	11.4	10	−51	14	−46	11	−72	11	70	14	*gg* – *gt*	
4	67	0.1	0	−35	15	−32	12	−14	21	58	14	*gg* – *gt*	
5	1992	4.3	4	−51	13	−38	11	−66	12	160	18	*gg* – *tg*	
6	5666	12.3	11	−47	15	−35	14	67	15	−65	12	*gt* – *gg*	
7	294	0.6	1	−46	14	−38	15	64	15	168	17	*gt* – *tg*	
8	564	1.2	1	−40	13	−29	12	50	9	51	11	*gt* – *gt*	
9	828	1.8	2	−54	9	−32	10	55	9	69	8	*gt* – *gt*	
10	339	0.7	1	−39	14	−31	14	58	11	95	13	*gt* – *gt*	
11	599	1.3	1	−40	12	−51	10	60	10	65	9	*gt* – *gt*	
12	562	1.2	1	−64	9	−38	11	73	11	51	10	*gt* – *gt*	
13	1047	2.3	2	−35	12	−25	13	79	12	62	11	*gt* – *gt*	
14	705	1.5	1	−57	9	−43	11	79	12	75	10	*gt* – *gt*	
15	1094	2.4	2	−51	13	−38	12	165	13	−65	12	*tg* – *gg*	
16	750	1.6	1	−40	13	−30	13	133	16	−69	12	*tg* – *gg*	
17	401	0.9	1	−52	13	−34	12	157	18	69	15	*tg* – *gt*	
18	219	0.5	0	−40	20	−29	15	164	17	156	14	*tg* – *tg*	
19	1423	30.7	3	−27	12	−169	13	−66	11	−66	13	*gg* – *gg*	Group I
20	1348	29.0	3	−31	12	−176	14	−67	12	63	15	*gg* – *gt*	
21	241	5.2	0	−26	11	−169	12	−70	12	159	17	*gg* – *tg*	
22	823	17.7	2	−34	11	−178	13	68	13	67	14	*gt* – *gt*	
23	109	2.3	0	−29	10	−176	9	82	16	166	22	*gt* – *tg*	
24	327	7.0	1	−26	15	−168	13	154	20	67	13	*tg* – *gt*	
25	371	8.0	1	−22	14	−171	13	170	16	163	16	*tg* – *tg*	
26	24		0	17	15	55	38	164	24	160	13	*tg* – *tg*	Group P
27	317		1	13	25	11	19	154	20	67	15	*tg* – *gt*	
28	82		0	−33	12	−145	33	64	12	−60	16	*gt* – *gg*	Group T
Avg (2–4)	11672	25.4	23	−50	14	−37	10	−66	11	65	14	*gg* – *gt*	
Avg (8–14)	4644	10.1	9	−47	11	−35	12	66	11	65	10	*gt* – *gt*	
Avg (M)	45931		90	−49	14	−36	12						
Avg (I)	4642		9	−29	12	−173	13						
Total:	51000			−49		−39							

[a] The clusters are sorted according to the conformation of the glycosidic linkage into four groups resulting in the major group (M) representing the global minimum, the minor group (I) with conformers having the inverted conformation of the glycosidic linkage, the group (P) with positive φ/ψ angle values, and a group (T) representing conformational transitions. The population (Pop) of each cluster (10 conformations are equivalent to 1 ps MD simulation), the percentage of the population respective to the total population (% Tot), the percentage of the population within its group (% Sub), the averages (Avg) and standard deviations (Std) for the φ, ψ, ω_1, and ω_2 angles are listed.

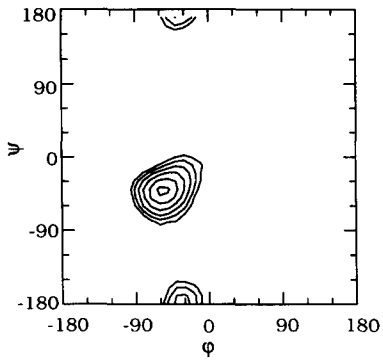

Fig. 9. Contour plot of the free energy of maltose as obtained from the cumulative results of all MD runs. The free energy was obtained by converting the populations using the Boltzmann statistics. Contour levels are shown at 10.5, 8.4, 6.3, 4.2, 2.1, and 0.4 kJ/mol (outside to inside). The energy of the inverted conformation is probably too low because of the repeated starting of the MD simulations in the inverted conformation and the inability to reach an equilibrium state between the two minima (see text). The enthalpy differences obtained directly from the MD simulations were larger than 10 kJ/mol, i.e., the local minimum at $\psi \approx 180°$ would not be visible in this map.

sharp boundary that runs perpendicular to the line connecting their centers. The high number of transitions between them results in an average lifetime of only ≈ 0.3 ps and proves that no significant energy barrier exists between these clusters. Cluster #5 has more transitions with cluster #3 than with cluster #2 indicating that transitions between *gt*(2) and *tg*(2) are favored, probably by electrostatic interactions between both hydroxymethyl groups when lower ψ angles are adopted.

The glycosidic linkage conformations represented in cluster #6 are very similar to that of cluster #1. For the reducing glucose, around 100 transitions occurred between the *gt* conformer in cluster #6 and *tg* conformer in cluster #16 that covers the ω_1 range between 100° and 160°. The ω_1 range from 150° to 190° is covered by cluster #15 that is separated by the clustering algorithm primarily because *tg* rotamers have a very wide distribution. Cluster #16 has its φ population centered around $-40°$, while cluster #15 has an average φ angle value of $-51°$. This difference correlates with changes of the ψ angles that are $-30°$ and $-38°$ for cluster #16 and #15, respectively. This correlation is likely to be caused by the reducing hydroxymethyl group that interacts in its *tg* conformation with the nonreducing ring near H-5. A change of the conformation around the C-1′–O-1′ bond could relieve this interaction.

The standard deviations within a cluster had to be smaller than about 20° to ensure a clear separation of local energy minimum regions. This stringent requirement resulted in the creation of seven different clusters that have both hydroxymethyl groups in a *gt* conformation and that represent conformers within the range of the global energy minimum of the φ/ψ space (subgroup gtgt). The average φ and ψ angles differ by up to 30° between these seven clusters but they can be treated as a single cluster for most statistical analysis because no correlation exists between their preferred dihedral angles.

and the high number of transitions (Fig. 8) between them proves the lack of a significant energy barrier separating them. Separating these clusters is of importance when transitions into other clusters are being analysed (compare, e.g., transition probabilities from clusters #4 and #17 into the gtgt group in Fig. 8). As expected, transitions are mostly found for those pairs of clusters that have the greatest overlap in their populations.

The group P is represented by only two clusters, both having $tg(1)$ conformations. The population of conformers with similar glycosidic linkage conformations is too small such that ω transitions mostly lead to the assignment into another group, e.g., ≈ 15 transitions are found with cluster #13 which is part of the gtgt subgroup. Thereby, the population of the group P as a representative of conformers with positive φ/ψ dihedral angles might be underestimated due to the current definition of the clusters. The change in the glycosidic linkage between the group P minimum and the global minimum is represented by transitions between clusters #27 and #17.

The second major group (I) consists of seven clusters with 4642 conformations that have the inverted conformation of the glycosidic linkage. These clusters contain conformations in a φ angle range from $-70°$ to 30°. The range for the ψ angles extends from 150° to $-120°$. A correlation between increased ψ and decreased φ values is indicated for the clusters within group I (cf., Table 5). Three of them, clusters #20, #22, and #23, have average ψ values of about 183° and φ values of less than $-29°$. The other four clusters have ψ angles of $\approx 190°$ and φ angles of $\approx -25°$. These correlations can be attributed partly to interactions between the hydroxymethyl groups. There is, however, some bias in the analysis due to incomplete sampling of the conformational space for some of these less populated clusters.

4. Discussion

The description of molecular properties on the basis of molecular dynamics simulations requires the accessible conformational space of the molecule to be sampled representatively (Fig. 9). Simulations in the sub-nanosecond range are often too short to represent all conformational states properly [13]. By running six MD simulations with a total simulation time of 5.4 ns and starting each simulation from a different conformation, an effective sampling of the conformational space of maltose was achieved. Moreover, we are dealing with a relatively restricted conformational space, at least for the solute. This special case allows for a more in-depth analysis of the applied methods from a practical point of view.

Global minimum.—The global minimum conformation with mean values for $\varphi/\psi = -49°/-36°$ was found in all MD simulations of maltose in water independent of the start conformation of the glycosidic linkage and the initial orientations of hydroxyl and hydroxymethyl groups. No transition out of this minimum region was found in any MD run. While the glycosidic linkage was in the global minimum region, the hydroxymethyl groups occupied all three staggered conformations, with a ratio of $gg:gt:tg = 71:23:4$ and 58:35:5 for the reducing and the nonreducing residue, respectively.

Only a few intramolecular hydrogen bonds were observed in the MD simulations.

The most frequently occurring intramolecular hydrogen bonds were between the OH-3 of the reducing residue acting as a hydrogen bond donor and O-2′ of the nonreducing glucose acting as an acceptor. This hydrogen bond can only be formed if the φ and ψ angles adopt values of about $-10°$ that are larger than the average. Typical lifetimes of this hydrogen bond are a few picoseconds. However, conformations of the glycosidic linkage that would allow the formation of this hydrogen bond have a significantly longer lifetime of up to 30 ps. Thus, the HO-3–O-2′ hydrogen bond contributes little to the stabilization of those glycosidic linkage conformations. In water, the hydroxyl groups preferentially form intermolecular hydrogen bonds with the solvent molecules. The trajectories of the internal electrostatic energy terms of maltose between the residues indicate that the stabilization energy due to the intramolecular hydrogen bond is around 20 kJ/mol. However, this gain of internal energy is offset by an increase of similar size in the interaction energy between the solvent and the maltose. This equilibrium is very sensitive to force field parameter settings such that the actual hydrogen bonding statistic and the stabilization of the conformations by hydrogen bonding is strongly influenced by the choice of nonbonded interaction parameters (see comparison with ref. [13] below).

Local minima.—One energetically higher local minimum with mean values of $\varphi/\psi = -29°/-177°$ was stable long enough to derive reliable estimates of its relative energy and lifetime. The accessible area within this local minimum is of similar size as that of the global minimum. Conformations in this minimum were observed in three MD simulations, run 4 that was started within this minimum and runs 5 and 6 that were started from nearby positions in φ/ψ space (Table 1) that were not stable. This local minimum conformation was sampled for about 20 ps in run 6, for 100 ps in run 4, and for 320 ps in run 5. The global minimum was reached in all three MD simulations via a relatively low energy barrier that is located at ψ values between $-90°$ and $-150°$. In the region of the local minimum, hydrogen bonds between OH-6 and OH-2′ are formed. By analogy to the hydrogen bonds observed in the global minimum, these intramolecular hydrogen bonds only occur transiently and contribute little to the stabilization of the glycosidic linkage conformations.

A third region with metastable conformations had positive mean values for φ and ψ of 17° and 55°, respectively. These conformations, observed during run 6 for less than 20 ps, are in the region of minimum C that was found to be the global minimum in the in vacuo molecular mechanics studies of Ha et al. [6]. Other minima than these are not expected for the modified GROMOS force field in water as evident from an extensive search through the conformational space of maltose using a high temperature MD simulation [14].

Energy differences.—The energy difference between the two major minima with $\psi \approx -50°$ and $\psi \approx 180°$ can be used to assess their relative populations. The energy values of each minimum were obtained by averaging the energy terms of the MD simulations in the corresponding time intervals. The internal energy of the water is stable throughout run 5, which allows a comparison of the energies of the two minima, i.e., energy averages are determined for the time period from 80 ps to 350 ps where the local minimum is accessed, and from 400 ps to 1000 ps which represents the global minimum (Table 4). The energy differences of 11 kJ/mol of the total potential energy and 14 kJ/mol of the maltose energy (Table 4, $E_{\text{sum-m}}$) are estimates of the energy

difference between the maltose with normal and inverted conformations of the glycosidic linkage. The error estimates for these energies are ± 1.3 kJ/mol. An energy difference of 11 kJ/mol is equivalent to a population of the inverted conformation at 300 K of 2.5%. Runs 4 and 6 also contain the inverted conformation but were not used for this type of energy comparison because in run 6 the inverted conformation was not stable long enough, and in run 4 the average of the internal solvent interaction energies drifted by -38 kJ/mol over the course of the MD simulation.

The energy difference between both minima can be analyzed in more detail by comparing the averages of the individual energy terms for the time intervals during which the local and the global minima are accessed (Table 4). Most energy terms favor the population of the global minimum over the local minimum. Only the van der Waals interaction energy between maltose and water and the electrostatic interactions within the water favored the local minimum by 1.5 kJ/mol and 3.7 kJ/mol, respectively, reflecting the larger hydrophilic surface of the inverted conformation.

Dynamics and transitions.—The average φ/ψ angles and their population distributions within the global minimum region were very similar in each of the six MD simulations. Thus, a single simulation is sufficient to represent the conformation of the glycosidic linkage of this minimum with only little bias (Fig. 4). Nevertheless, the local minimum is not covered representatively in all simulations, even not in those simulations that start from the local minimum. Furthermore, a single MD simulation, even if it extends to 1 ns, does not represent the equilibrium of a disaccharide like maltose over the whole conformational space as indicated by the lack of transitions from the global minimum into the local minimum. Lifetime estimates for the global minimum of greater than 1000 ps and for the local minimum of about 150 ps can be made from the six MD simulations. These estimates carry a large margin of uncertainty because of the lack of enough transitions. As can be seen in Table 3, even the hydroxymethyl groups are normally not in equilibrium in any single run. The population distributions for the different runs vary strongly. Typically, the ratio of *gt*:*tg* conformations reaches an equilibrium within a single simulation. The energy barrier between these two conformers is relatively low and about 8 transitions can be observed within 1 ns, equivalent to a 50–100 ps lifetime for *gt* and a 5–10 ps lifetime for *tg* conformers. On the other hand, the *gg* conformer has the lowest energy and is separated by higher energy barriers from the *gt* and *tg* conformers, reflected by average lifetimes of the *gg* conformer of 200 ps in the nonreducing and 700 ps in the reducing glucose residue. Summarized over all MD runs, 21 transitions of the hydroxymethyl groups from and into the *gg* conformations are found for both residues, and hence, all simulations taken together give a population distribution for the ω populations that is based on enough transitions such that the equilibrium is approached. The final, cumulative statistics of the hydroxymethyl population reads *gg*:*gt*:*tg* = 70:23:6 for the reducing glucose and *gg*:*gt*:*tg* = 55:38:6 for the nonreducing residue. Nearly equal ratios are found for the glycosidic linkage in the normal and in the inverted conformation.

Using the Arrhenius equation with a preexponential factor of 10^{12}, the energy barriers between *gt* and *tg* conformers are estimated from these transition frequencies of 8×10^9 s^{-1} to be about 12 kJ/mol. Using the cumulative transition frequencies of the hydroxymethyl groups in the *gg* conformation from all MD runs, which is equal to

1.5×10^9 s^{-1}, the lowest barrier from *gg* is estimated to be about 16 kJ/mol. These barriers are basically in agreement with those found for similar molecules.

Energies.—Average energy values that were derived from the energy trajectories can be compared directly between and within the calculations since the molecular ensemble (and the potential energy function) was the same in all runs. Independent of the maltose conformation and the start conditions of the runs, the average potential and kinetic energy terms are very similar in all six MD calculations (Table 4). The averages of the total potential energy in the six calculations vary between −6380 kJ/mol and −6891 kJ/mol. Surprisingly, the maximal value of this energy term is found in run 2 where the global minimum region of the glycosidic linkage was occupied throughout the run. Compared with run 1 that was started from very similar φ/ψ angles but other hydroxyl and hydroxymethyl group orientations, the water–water and water–maltose interactions in run 2 have energies that are higher by about 6 and 5 kJ/mol, respectively. The RMS fluctuations of the total potential energy term within the runs are between 60 and 72 kJ/mol, which results in standard deviations of the average values of ≈ 0.4 kJ/mol, so that a difference of more than 2 kJ/mol of the total potential energy can be statistically significant. In run 2, the gradient of the total potential energy is only 5 kJ/mol over 950 ps, but the averages over 100 ps indicate that this energy value goes through a maximum of about −6370 kJ/mol at 600 ps and finally reaches an average energy of −6390 kJ/mol. The latter value is comparable with the total potential energy values in the other runs. Therefore, over 800 ps of simulation time was necessary to achieve a complete relaxation of the solvent box. A similar relaxation of the water occurred in run 4. The largest variation between the runs was found for the averages of the electrostatic energy within the water where the minimum and the maximum of the average were about 12 kJ/mol apart (which is equivalent to about 0.15% of the average energies). Hence, the initial position of the water atoms and the relaxation of the solvent box is a critical factor in the simulations, and without the ability to compare different runs of the same ensemble, it would not have become evident that a single run is not completely equilibrated. Furthermore, shorter MD runs of only a few hundred picoseconds might also not reach an equilibrium for the solute. However, for simulations 1, 3, 5, and 6 proper equilibration has been proven. The energies of the portion of run 5 during which the molecule stayed in the global minimum are very similar to those of run 1 which remains in the global minimum all the time, indicating that the molecular ensemble, including the solvent, reaches an equilibrium very shortly after the major change of the conformation of the glycosidic linkage occurred in run 5. This is not obvious because the conformational change requires also a major reorientation of the water molecules. Moreover, the system seems to equilibrate faster when the starting conformation had a higher total energy, which allows for a faster rearrangement of the molecular ensemble.

Influence of start conditions.—Starting from different conformations provided a substitute for the lack of sufficient transitions between the local and the global minimum. Local minima regions and the transition pathways can only be explored, and rough estimates or boundary values for the lifetimes can be obtained based on the few transitions that were found. Although the population differences may not correctly reflect the equilibrium values, energy differences of minima can be determined if statistically significant periods had been sampled for each minimum. The energy

differences can, in turn, be used to weight the population of the minima. In practice, these calculations can be run simultaneously on different computer systems, allowing for a fast turn around.

The hydroxyl and hydroxymethyl conformations changed shortly after position restraints were removed and were of minor influence for the solute conformations observed during the simulations. However, their initial setting influenced the initial positions of the water molecules. As seen from the comparison between run 1 and 2, this may have dramatic influence on the equilibration of the ensemble, with strong effects on the energy of the whole system. Thereby, the use of different starting conditions can be helpful in arriving at a meaningful interpretation of the MD data.

Cluster analysis.—The major problem in the cluster analysis is the determination of the level where the cluster tree is being cut. Since the regions of the minima are sometimes rather extended compared to their separation, it becomes necessary to cut the cluster tree at fairly low levels to separate the regions of the different minima and recombine some of the clusters identified in this way to avoid an artificial splitting of conformations that belong to a single energy minimum. The transition frequencies between cluster pairs are a good indicator of clusters that are closely related. However, even though a transition analysis of some cluster pairs might show that they should be pooled into one group of clusters, a detailed analysis of the individual clusters can reveal correlations between dihedral angles that otherwise might remain undetected. A simple classification into a local and a global minimum and further splitting into the nine different ω populations would not have revealed correlations between those dihedral angles. It is remarkable that clustering even in such a low dimensional space allowed a much better analysis than sorting. Sorting procedures become far more difficult in higher dimensional conformational space.

Comparison with other calculations.—Molecular mechanics and MD studies of maltose in vacuo by Ha et al. [6] and Tran et al. [7] showed several minima for the glycosidic linkage. The minimum with the lowest energy using MM2CARB [7] was found at $\varphi/\psi = -60°/-40°$ (minimum A). The MD calculation and molecular mechanics study using the CHARMM potential [6] showed the global minimum at $\varphi/\psi = 38°/16°$ (minimum C) and a second local minimum with a relative energy of about 0.42 kJ/mol at $\varphi/\psi = -61°/-40°$ (minimum A). In both studies, several additional local minima were found. The most prominent of which are located at $\varphi/\psi \approx -40°/-30°$ (B), $-20°/20°$ (C) and $-4°/60°$ (D) in the study of Tran et al. [7] with relative energies of 0.8 kJ/mol, 9.0 kJ/mol, and 10.9 kJ/mol, respectively. Ha et al. [6] found additional minima at $\varphi/\psi = -3°/-40°$ (B) and $49°/-142°$ (G) that have relative energies of 3.8 kJ/mol and 11.7 kJ/mol, respectively. In addition, the minimum of maltose with the inverted glycosidic linkage conformation was found at $\varphi/\psi = -30°/-160°$ (E) and $-20°/170°$ (F) [7] with relative energies of 14.7 kJ/mol and 17.2 kJ/mol, respectively. Ha et al. [6] found only one minimum (F) for the inverted conformation that had $\varphi/\psi = 48°/180°$ with a relative energy of 3.3 kJ/mol.

Only two minima remain once water is considered in the MD simulations with GROMOS and the local minimum is populated at about 2.5%. Minima A of both studies, MM2CARB and CHARM, have about the same size and shape as our global minimum. The inverted conformations represented by minima E and F of Tran et al. [7],

and F of Ha et al. [6] also overlap almost exactly with the region of the inverted conformation described here. The local minimum would only be populated at about 0.3% in the MM2CARB force field [7]. The energy difference reported by Ha et al. [6] of 3.3 kJ/mol between their global minimum C and their minimum F with the inverted conformation results in a population of the inverted conformation of about 21%.

Minimum C reported in the above studies was only occupied transiently for 20 ps in MD run 3 that was started from minimum D. Several very short-lived transitions into the outskirts of region C were also observed but the maltose never remained in this region for more than a few picoseconds. This clearly indicated that minimum C could be reached during the simulations but is not stable when water is present and the *exo*-anomeric effect is taken into account. Ha et al.'s minimum B [6] lies largely outside the global minimum of maltose in water described here.

The local minima D and G were not stable for even a short time in the water simulations. This is not obvious because simulations 3 and 5 were started in these minima and our set-up procedure was designed to allow for a relaxation and stabilization of the starting conformation.

The hydrogen bonds and the orientations of the hydroxymethyl groups strongly determine the location and the shape of the minimum regions of the glycosidic linkage in the in vacuo studies [6,7]. In contrast, the MD simulations with explicit inclusion of the solvent clearly indicate that the conformation of the glycosidic linkage determines whether intramolecular hydrogen bonds may occur or not. MM2CARB [7] seems to represent the global minimum and the inverted conformation very well. However, the study showed additional minima that are artificially stabilized by hydrogen bonds. The modified CHARMM force field [6] on the other hand does not reproduce the global minimum found in water and shows many differences compared to the data obtained in this study. It seems likely that strong hydrogen bonds were largely responsible for the additional minima found and for the shape and size of those minima.

Recently, Brady and Schmidt [13] have published a study of maltose in water. They used a derivative of the CHARMM force field that had been used before in the in vacuo study of Ha et al. [6]. This force field differs from the modified GROMOS force field used in this study primarily by much higher partial charges. It also does not account for the *exo*-anomeric effect. Brady and Schmidt [13] reported the results of several short MD simulations that were started from the previously found minima A and C. The trajectories were well within the regions of low energy found in the in vacuo studies, indicating that the solvent affects the strength of the interresidue hydrogen bonds only marginally. They were able to reduce the effect of the strong intramolecular hydrogen bonds by arbitrarily reducing the partial charges on the oxygen and carbon atoms by 0.1 e.s.u. A 60 ps trajectory indicates that the hydrogen bonds were in fact weakened by this approach. They suggest that hydrogen bonds might be weakened in water. Nevertheless, the resulting trajectory stays within the energy barriers already found in the in vacuo simulation. In particular, they find a high population at positive φ/ψ angles of $\approx 10°/40°$. Brady and Schmidt [13] have pointed out in their paper that the simulation times they used were too short to produce meaningful statistics.

Comparison with experimental data.—The results of the MD simulations in aqueous solution may be compared with experimental data. X-ray [20,21] and neutron diffraction

[22] studies of maltose have been performed. Normally, in crystal structure analysis, inter- and intra-molecular hydrogen bonds are found. In particular, a hydrogen bond between OH-2′ and O-3 was found in most of the studies. Correlated with this finding, the dihedral angles of the glycosidic linkage adopt values around $0° \pm 10°$ for φ and ψ. A single conformation like the one found in the crystal structure analyses is not in agreement with NMR experiments [3–5,23] or optical rotation studies [8,11] in aqueous solution. Rees and Thom [24] found a strong temperature and solvent dependence of the experimental optical rotation of maltose, concluding that in water, the hydrogen bond between O-2′ and O-3 is weakened compared to less polar solvents. Even moderately polar solvents, like Me_2SO, show significant changes of the observed optical rotations compared to water [24]. Also, temperature and alkali have a significant influence on the conformation of the glycosidic linkage as measured by optical rotations [8,11,25,26].

Recently, Stevens [11] published that conformations with $\varphi/\psi = -70°/-35°$ should contain the highest population. Molar rotations for maltose as a function of the conformation of the glycosidic linkage and the hydroxymethyl groups have been calculated [11]. The values for the molar rotation were given in 10° bins on the φ/ψ map. The estimated error [11] of these values is $\pm 20°$ cm^2 dmol^{-1}. Multiplication of these molar optical rotations with the corresponding populations obtained from our MD simulations can be used to arrive at an estimate for the molar rotation of maltose. This ensemble averaged value of the molar rotation was calculated to $387° \pm 20°$ cm^2 dmol^{-1} using the complete set of conformations of the MD simulations described here. The experimental value for the molar rotation in maltose [27] is $[M] = 402°$ cm^2 dmol^{-1} under standard conditions showing an excellent agreement between experiment and our theoretical result. The ensemble average value for the global minimum obtained here is 402° cm^2 dmol^{-1} and for the inverted conformation is 238° cm^2 dmol^{-1}. Considering only the centers of the minima B, C, D, E, and G reported earlier [6], one obtains molar rotations [cm^2 dmol^{-1}] of 184°, 140°, 286°, 259°, and −120°, respectively. This indicates that the minima B, C, and G must be extremely low in population and that minima D and E cannot have a significant contribution to account for the experimentally found molar rotation. This holds, since the calculated molar rotation of minimum A [7] is 426° cm^2 dmol^{-1}, close to the maximum calculated value of 516° cm^2 dmol^{-1} [11].

Pérez et al. [4] conclude from their NMR study of the solvent dependence of the interresidue $^3J_{CH}$ coupling constants in maltose that the solution conformation/conformational distribution is different in solution than in the crystal. They found that the hydrogen bond between OH-3 and O-2′ is weakened in aqueous solution. The $^3J_{C\text{-}1',H\text{-}4}$ coupling constants reported for α,β-maltose and methyl β-maltoside are 2.5 Hz [28], 4.5 Hz [4] (at 353 K), and 3.9 Hz ± 0.3 Hz [5], respectively. For the $^3J_{C\text{-}4,H\text{-}1'}$ coupling constant, values of 3 Hz [28], 4.3 Hz [4] (at 353 K), and 3.5 Hz ± 0.3 Hz [5] have been reported. From our MD study, $^3J_{C\text{-}1',H4}$ and $^3J_{C\text{-}4,H\text{-}1'}$ calculate to 3.9 Hz and 2.7 Hz, respectively, using the parametrization of Tvaroska et al. [29] or Mulloy et al. [30]. Therefore, the calculated values agree reasonable well with the experimentally derived data, considering that Karplus type relationships for these coupling constants normally carry a margin of uncertainty of at least ±1 Hz. The 0.8 Hz difference between the calculated coupling constant for the φ angle and the data of Shashkov et al. [5] could theoretically be explained if the inverted conformation had a population of ≈ 30%.

Unfortunately, the high error margin for the Karplus equation and the error of the experimental data do not allow for such a detailed analysis of the data in terms of distribution of inverted versus global minimum or shapes and location of the minima.

NOE Experiments [5] on methyl β-maltoside showed only a weak enhancement from H-1′ to H-3 as shown by the ratio of $NOE_{H\text{-}1'-H\text{-}3}$:$NOE_{H\text{-}1'-H\text{-}4}$ = 1:6. Using a full relaxation matrix approach for the MD results shown here, the calculated time averaged NOE values of H-1′–H-3 and H-1′–H-4 are 5.7:1 and 1:3 for the local and the global minimum, respectively. This indicates that the inverted conformation is only populated for a short portion of the time because the inverted conformation should create a strong NOE between the H-1′ and H-3 protons.

The experimentally determined coupling constants between the H-5 and both C-6 protons of the two glucose rings are $^3J_{H\text{-}5,H\text{-}6S}$ = 2 Hz and $^3J_{H\text{-}5,H\text{-}6R}$ = 5 Hz [5]. They agree very well with the average coupling constants derived from the MD simulations presented here, which are $^3J_{H\text{-}5,H\text{-}6S}$ = 2.4 Hz and $^3J_{H\text{-}5,H\text{-}6R}$ = 4.7 Hz for the reducing residue and $^3J_{H\text{-}5,H\text{-}6S}$ = 2.5 Hz and $^3J_{H\text{-}5,H\text{-}6R}$ = 5.8 Hz for the nonreducing residue, respectively.

5. Conclusion

MD simulations in water using the GROMOS force field with additional energy terms for the *exo*-anomeric effect are capable of assessing conformational preferences of disaccharides in aqueous solutions. Furthermore, relative energies of local minima, estimates of the corresponding energy barriers, and lifetimes of conformational states can be determined when the conformational space is extensively sampled. All MD simulations presented in this study have reached the global minimum at $\varphi/\psi \approx -49°/-36°$ but no transition occurred from the global into the local minimum at $\varphi/\psi = -29°/-173°$. An equilibrium between different local minima was difficult to achieve because the lifetimes of the conformational states normally exceeded many hundred picoseconds. Even for transitions between rotamers of the hydroxymethyl groups for which energy barriers of only ≈ 12 kJ/mol have to be overcome, a single 1 ns MD simulation is too short to reach equilibrium.

Using the Arrhenius equation, an estimate of the lifetime in the global minimum can be derived. The average lifetime of the inverted conformation of 150 ps translates into a height of 13 kJ/mol for the barrier from the inverted conformation into the global minimum. Together with the energy difference between the global and the local minimum of more than 10 kJ/mol, the barrier to reach the inverted conformation from the global minimum is more than 23 kJ/mol. Therefore, the average lifetime of the global minimum exceeds 8 ns. To reach an equilibrium of the glycosidic linkage conformation, it would be necessary to run simulations for about 150 ns or longer, i.e., 1000 times longer than a typical simulation is run currently. Starting from different conformations that are outside the global minimum strongly improves the efficacy in producing a comprehensive description of local minima. The relative energy determined thereby can be used to weight the relative populations of different minima that are not in equilibrium.

Furthermore, the time needed to equilibrate the water box can occasionally take more than several hundred picoseconds. The question of whether a water box has equilibrated can be answered from a comparison between different MD simulations of the same molecular ensemble. One way to address the problems of reaching an equilibrium is the use of several distinctly different starting conformations for MD simulations.

References

[1] S. Melberg and K. Rasmussen, *Carbohydr. Res.*, 76 (1979) 23–27.
[2] S. Melberg and K. Rasmussen, *Carbohydr. Res.*, 76 (1979) 27–38.
[3] G.M. Lipkind, V.E. Verovsky, and N.K. Kochetkov, *Carbohydr. Res.*, 133 (1984) 1–13.
[4] S. Perez, F. Taravel, and C. Vergelati, *Nouv. J. Chim.*, 4 (1985) 561–564.
[5] A.S. Shashkov, G.M. Lipkind, and N.K. Kochetkov, *Carbohydr. Res.*, 147 (1986) 175–182.
[6] S.N. Ha, L.J. Madsen, and J.W. Brady, *Biopolymers*, 27 (1988) 1927–1952.
[7] V. Tran, A. Buleon, A. Imberty, and S. Perez, *Biopolymers*, 28 (1989) 679–690.
[8] E.S. Stevens and B.K. Sathyanarayana, *J. Am. Chem. Soc.*, 111 (1989) 4149–4154.
[9] A.D. French, *Carbohydr. Res.*, 188 (1989) 206–211.
[10] H. Lim, C.S. Setser, and J.V. Paukstelis, *Cereal Chem.*, 69 (1992) 387–390.
[11] E.S. Stevens, *Biopolymers*, 32 (1992) 1571–1579.
[12] M.K. Dowd, J. Zeng, A.D. French, and P.J. Reilly, *Carbohydr. Res.*, 230 (1992) 223–244.
[13] J.W. Brady and R.K. Schmidt, *J. Phys. Chem.*, 97 (1993) 958–966.
[14] K.-H. Ott and B. Meyer, *J. Comp. Chem.*, in press.
[15] H.J.C. Berendsen, P.J.M. Postma, W.F. Van Gunsteren, A. di Nola, and J.R. Haak, *J. Chem. Phys.*, 81 (1984) 3684–3690.
[16] W.F. Van Gunsteren and H.J.C. Berendsen, *Mol. Phys.*, 34 (1977) 1311–1327.
[17] J.A. Hartigan, *Clustering Algorithms*, Wiley, New York, 1975.
[18] B. Everitt, *Cluster Analysis*, Halsted, New York, 1980.
[19] D.M. Abraham and F. Rizzardy, *The Berkley Interactive Statistical System*, Norton, New York, 1988.
[20] S.S.C. Chou and G.A. Jeffrey, *Acta Crystallogr.*, 23 (1967) 1038–1040.
[21] F. Takusagawa and G.A. Jacobson, *Acta Crystallogr., Sect. B*, 34 (1978) 213–218.
[22] M.E. Gress and G.A. Jeffrey, *Acta Crystallogr., Sect. B*, 33 (1977) 2490–2495.
[23] I. Christiansen-Brams, M. Meldal, and K. Bock, *J. Carbohydr. Chem.*, 11 (1992) 813–835.
[24] D.A. Rees and D. Thom, *J. Chem. Soc. Perkin Trans. II*, (1977) 191–201.
[25] E.S. Stevens and K. Bangalore, *Biopolymers*, 27 (1988) 415–421.
[26] E.S. Stevens, B.K. Sathyanarayana, and E.R. Morris, *J. Phys. Chem.*, 93 (1989) 3434–3436.
[27] F.J. Bates, *NBS Circular C440*, USG, NBS, Washington, DC, 1942.
[28] A. Parfondry and N. Cyr, *Carbohydr. Res.*, 59 (1977) 299–309.
[29] I. Tvaroska, *Carbohydr. Res.*, 206 (1990) 55–64.
[30] B. Mulloy, T.A. Frenkiel, and D.B. Davies, *Carbohydr. Res.*, 184 (1988) 39–46.