JCTC Journal of Chemical Theory and Computation

Sugar Folding: A Novel Structural Prediction Tool for Oligosaccharides and Polysaccharides 2

Junchao Xia,[†,§] Ryan P. Daly,[†,§] Feng-Chuan Chuang,[†] Laura Parker,[†] Jan H. Jensen,[‡] and Claudio J. Margulis*[,†]

Department of Chemistry, University of Iowa, Iowa City, Iowa 52242, and Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

Received February 8, 2007

Abstract: This is the second in a set of two articles where we describe our newly developed scheme to predict conformations of complex oligosaccharides in solution. We apply our fast sugar conformation prediction tool to the case of two complex human milk oligosaccharides LNF-1 and LND-1. As described in detail in the first paper, our protocol initially delivers a set of "unique structures" corresponding to important minima on the potential-energy landscape of a complex sugar using an implicit solvent model. The nuclear Overhauser effect ranking of individual conformations provides a suitable way for comparison with available experiments. The structures obtained agree well with earlier computational predictions but are obtained at a significantly lower computational cost. Sugar conformations corresponding to stable energy minima not found by earlier molecular dynamics studies were also detected using our methodology. In order to evaluate the effects of explicit solvation and thermal fluctuations on several different predicted conformers, we also performed short-time molecular dynamics simulations in an explicit solvent.

1. Introduction

Oligosaccharides; polysaccharides; and their glycoconjugates, glycoproteins, and glycolipids play a very important role in biological phenomena such as cell−cell interaction, inflammatory processes, immunity, and fertilization.[1,2] Conformational studies are crucial to understand biological function.[1] In most cases, the determination of an oligosaccharide's conformation involves the characterizing of the $\phi-\psi$ glycosidic linkages between monosaccharide residues. Glycosidic linkages of oligosaccharides can usually be quite flexible.[3] Such flexibility creates difficulty on crystallization and results in considerable limitations[4,5] when applying standard experimental techniques such as crystallography. Martin-Pastor and Bush[6,7] pointed out that the internal motions of oligosaccharides might be classified into two kinds: the fluctuation around a single energy minimum (conformation) on the order of 15−50° and the interconversion between distinct energy minima. Depending on the free-energy difference between conformers, experimental nuclear Overhauser effect (NOE) data may be dominated by a single conformation corresponding with a structure that is thermally fluctuating around a single minimum, or with an ensemble average over several different conformations when these are close in free energy.[3,8] Even though conformations may be close in free energy, they are not necesarily structurally close. Due to experimental challenges in working with sugars, computational methods have become important tools to understand or predict the conformations of oligosaccharides and glycoconjugates in various environments.[5,9−18]

Different computational methods to perform a conformation analysis of oligosaccharides in vacuo[5,9,10,12,13] have been proposed. Some of these are based on relaxed or adiabatic maps in potential-energy surfaces for disaccharides;[19,20] others are based on the CICADA method combined with simulated annealing to travel through conformational space.[21−23] Monte Carlo methods[24] and genetic algorithms[25,26] have also been used. Our methodology described in the first paper is fully automatic; it works on almost any oligosaccharide; it is based on a ring perception algorithm that automatically detects rotable dihedrals, a systematic coupled dihedral space search for the whole oligosaccharide, and the use of a substructure matching algorithm that recognizes a branch within a

* Corresponding author e-mail: claudio-margulis@uiowa.edu.
[†] University of Iowa.
[‡] University of Copenhagen.
[§] Equal contributions.

10.1021/ct700034q CCC: $37.00
Published on Web 06/14/2007

complex sugar when that branch has already been studied and stored in a database. The methodology is first applied to find regions within the complex oligosaccharide dihedral space that are sterically allowed. Subsequently, minimizations are performed and structures are pooled into what we have defined in the first paper as "unique structures". These unique structures can be sorted on the basis of different criteria such as their root-mean-square deviation against experimental NOEs, their energies in implicit solvent, or any other desired criteria.

Even though the idea of a systematic search over dihedral space for a complex oligosaccharide may appear to be an exponentially untractable problem for which Monte Carlo or other techniques could be more applicable, our experience is that, for biologically interesting sugars which are not linear, crowding severely limits the number of available conformations as the sugar becomes larger. At the same time, since dihedral angle motion becomes coupled, free-energy barriers to rotation appear to be larger and transitions between certain conformations become rare events on the typical length of a molecular dynamics (MD) or Monte Carlo simulation. In this work, we will show that the number of "unique structures" obtained for our larger sugar is smaller than that for the one with less monosaccharide units. It turns out that our systematic sampling of the whole fully coupled dihedral phase space for a complex oligosaccharide with size on the order of seven units can be easily performed within a few hours. In our approach, if further information is desired about particular structures, our "unique structures" can be used as sensible starting conformations for MD in solution. Furthermore, if part of the sugar in question has previously been studied and stored in our database, only those saved conformations and not the whole dihedral space need to be searched when further complexity is added to the molecule. The situation is different with other techniques such as molecular dynamics and Monte Carlo simulations in explicit solvent. Our experience has been that even the longest simulations currently available for complex oligosaccharides (on the order of 50 ns) only visit basins that are close to the initial MD conformation of the sugar. This is mainly because, in the case of branched sugars, torsions are strongly coupled, particularly when the branching occurs on adjacent linkages.

In this paper, we test our tool by predicting the conformations of complex oligosaccharides present in human milk. Our choice is based on the fact that these have been extensively studied via experimental NOEs and MD.

Hundreds of lactose-derived oligosaccarides exist in human milk. These oligosaccharides are the third largest component in milk and are thought to provide mechanisms of breast-feeding protection for infants against enteric pathogens.[27] Although conformational studies have been performed for several human milk oligosaccharides on the basis of NOEs, *J* coupling, residual dipolar couplings, and molecular dynamics simulations,[7,15,28,29] more work remains to be done to fully determine the conformations that these oligosaccharides can take in solution and which of these are relevant to their biological function.

Our results are for oligosaccharides LNF-1 [α-L-Fucp-(1→2)-β-D-Galp-(1→3)-β-D-GlcpNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glcp] and LND-1 [α-L-Fucp-(1→2)-β-D-Galp-(1→3)-[α-L-Fuc-(1→4)]-β-D-GlcpNAc-(1→3)-β-D-Gal-(1→4)-β-D-Glcp]. In a recent paper,[15] Almond et al. investigated these two oligosaccharides by NMR and long (50 ns) molecular dynamics simulations in explicit water. The conclusions from their very interesting work were that these oligosaccharides possess relatively ordered structures (i.e., they fold). The authors also showed that the oligosaccharides can be easily trapped in the "wrong" free-energy minima for times as long as 50 ns, if initial structures were "incorrectly" selected. These wrong initial conditions produced trajectories that yielded incorrect NOEs. This fundamentally important result emphasizes the need to have a fast systematic way of generating all relevant sugar conformers a priori without the need to rely on MD for sampling. In this paper, we will show that our method can accomplish this in a very efficient manner. These conformers can be used to quickly predict which structure is closest to the correct experimental NOE and also to generate a family of initial structures that can be further tested via MD or other methods of choice.

2. Simulation Methods

2.1. Coarse-Graining Systematic Search. We performed conformational searches for the LNF-1 and LND-1 oligosaccharides. The scanning increment for each linkage was 10°. In order to reduce the number of conformations studied, structures have been pooled so that four adjacent points on ϕ and four adjacent points on ψ are converted into a single geometry-averaged point. This was done for each glycosidic linkage. The first dihedral angle of the longest side chain (NAc group) was also rotated with increments of 60°.

As described in the first paper, an energy minimization for each allowed conformation was carried out using the software TINKER[30,31] with the MM3 force field[32] and the generalized Born suface area (GBSA) implicit solvent model.[33,34] For a comparison test, we also performed energy minimization using GROMACS[35,36] with the OPLS-AA force field[37] in the gas phase. CPU times for a full conformational search and energy minimizations were less than a day on a single-processor (Intel Pentium 4 CPU, 2.80 GHz) computer.

2.2. Molecular Dynamics Simulation. After scoring the structures obtained on the basis of a comparison between experimental and computationally obtained NOEs (see section 2.3), short MD simulations on the order of 5 ns were carried out for selected structures. This was done in order to test the stability of these structures in the presence of an explicit solvent and in order to get better solvent-averaged NOEs. MD simulations were carried out using the software GROMACS[35] with the OPLS-AA force field.[37] In each case, the simulation box was 5 nm × 5 nm × 5 nm and the simple point charge (SPC)[38] water model was used to model water explicitely. Simulations were carried out under periodic boundary conditions. Constant pressure, temperature, and number of particles (NPT) simulations were carried out at $T = 300$ K and $P = 1$ atm. The Nose−Hoover thermostat[39,40] and the Berendsen pressure coupling scheme[41] were used for this purpose. A time step of 0.001 ps was used for integration.

2.3. Nuclear Overhauser Enhancement. In the first paper, we showed that sorting structures according to their

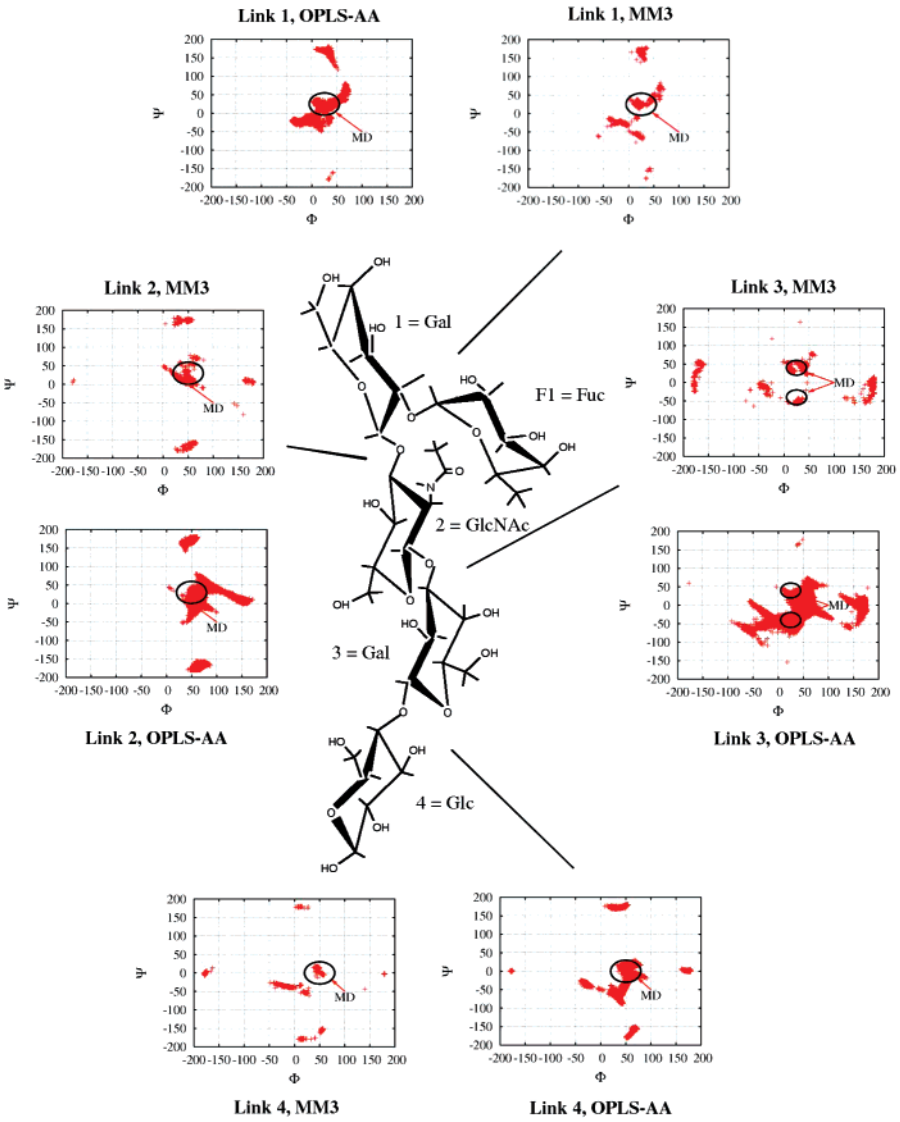

Figure 1. Distributions of all $\phi-\psi$ pairs from 1108 unique conformations of LNF-1 with the MM3 force field and 7030 with the OPLS-AA force field. Clearly, the distribution for each linkage is clustered around several important regions. The circled regions are those found in ref 15 by extensive 50 ns molecular dynamics simulations using the CHARMm[44] force field and the explicit TIP3P[45] water model. As it is clear from this picture, our method captures many more allowed regions than the MD simulations. MD is not able to visit most of these structures because transitions between these are rare events on a nanosecond time scale. In our database, we have linked information about these regions, a vector in $2n$ dimensional space (here, n is the number of dihedral linkages and rotable side chains).

energies in implicit solvent or free energies computed using a harmonic approximation does not necesarily yield the most likely conformations in solution. This is because of the effect of entropy, anharmonicity, and explicit solvation on the free-energy landscape. On the other hand, a good estimator appears to be

$$\mathrm{RMSD} = \sqrt{\frac{\sum_i^N \left(\frac{\sigma_i - \sigma_{0i}}{\sigma_{0i}}\right)^2}{N}} \qquad (1)$$

where σ_i is our calculated NOE value for the ith proton pair and σ_{0i} is the corresponding experiment value, and the summation is over all available N experimental NOEs. This estimator could fail in the case in which several local free-energy minima are within a small fraction of KT from each other. This would correspond to the case in which the sugar does not have a well-defined fold and its structure is more consistent with a random configuration. This does not appear to be the case for the sugars studied in this article.[15] Such a case would be of little interest from a sugar-folding prediction perspective. As we will demonstrate later in this manuscript, our prediction method using this structure-sorting scheme agrees well with very expensive MD simulations previously published[15] when these are started from the proper initial conditions. In order to obtain better thermal averages, once relevant conformations are identified, short MD simulations in explicit solvent can be used to refine the NOE results obtained from individual configurations.

Extensive literature is available on the nuclear Overhauser enhancements of oligosaccharides or glycoconjugates in solution; see, for example, refs 4 and 5. In order to create

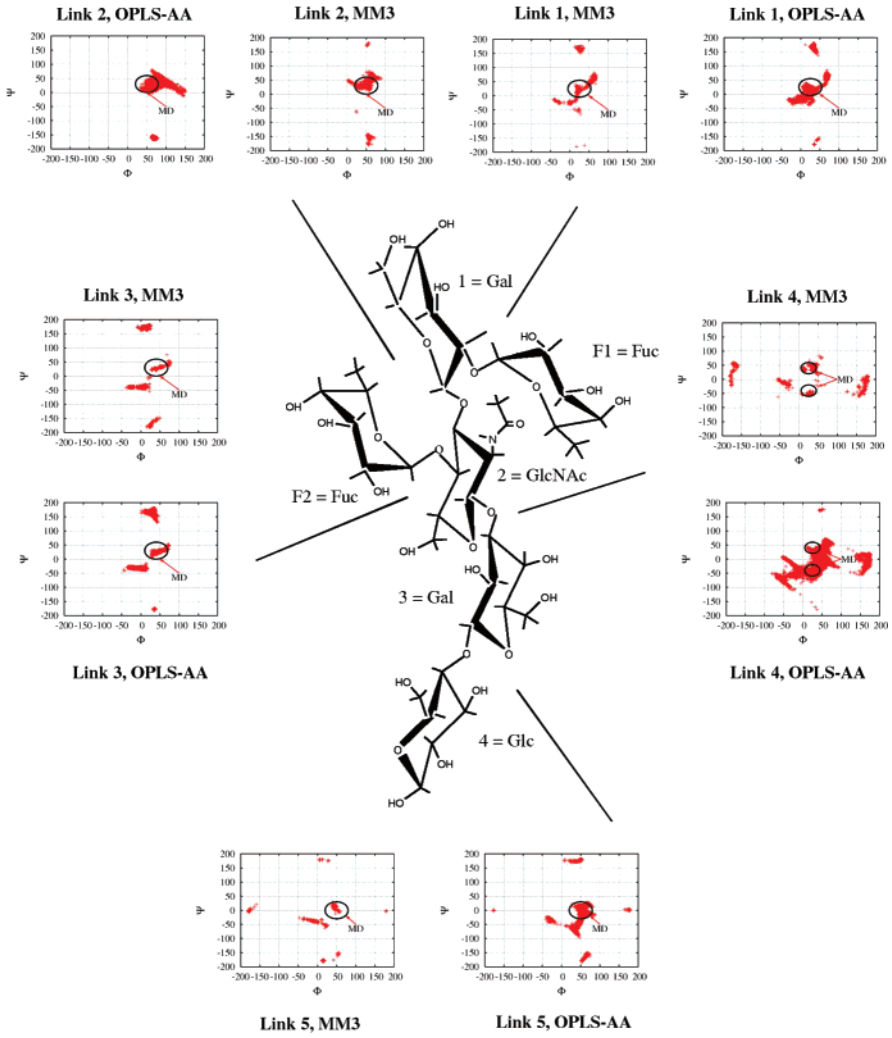

Figure 2. Distributions of all $\phi-\psi$ pairs from 989 "unique conformations" of LND-1 using the MM3 force field and 5220 conformations obtained from minimization using the OPLS-AA force field. The distribution of important regions is similar to that previously found for LNF-1 shown in Figure 1. The circled regions are those sampled by the 50 ns MD simulations in ref 15. Link 2 has a smaller allowed dihedral space than in the case of LNF-1 because of hindrance due to the presence of link 3.

our root-mean-square deviation sorting scheme, we have coded the model-free approach[42,43] into our tool. NOEs from selected initial conditions obtained from MD time averages were also computed according to the scheme of Cumming and Carver.[3,8]

3. Results and Discussions

3.1. Coarse-Graining Grid Search. We applied our method to determine the possible conformations of LNF-1 and LND-1 human milk sugars in solution. Detailed chemical structures of these molecules are shown in Figures 1 and 2. These two fucosylated oligosaccharides have similar structures. LND-1 has an additional α-L-Fuc connected to GlcNAc.

3.1.1. LNF-1 Milk Sugar. Dihedral space search and structure pooling resulted in 24 041 allowed conformations for LNF-1. After energy minimization, only 1108 of these were defined by the program as "unique conformations". We have chosen the criteria $\Delta E < 5.0$ kcal/mol, $\Delta\psi < 10°$, and $\Delta\phi < 10°$ to define a "unique conformation". According to our experience, the number of unique conformations could have been reduced ever further to less than 100 if we would have chosen $\Delta\psi < 50°$ and $\Delta\phi < 50°$. Even though the grid may seem too coarse in this case, 50° is a reasonable number since it is compatible with the size of our energy basins at thermal conditions. We know this from the time evolution of $\phi-\psi$ for each linkage in our molecular dynamics simulations. Nonetheless, we have used the finer grid since the algorithm was fast enough that all minimizations could be carried out on a single PC in less than a day.

The insert graphs in Figure 1 show the distributions of all $\phi-\psi$ pairs of unique conformations for LNF-1 using the MM3 force field and the GBSA implicit solvent model as well as the OPLS-AA force field in the gas phase. As mentioned before, these structures are sterically allowed and energy-minimized. It is obvious from this figure that the $\phi-\psi$ distribution for each linkage is clustered into several different important regions. In particular, the circled regions are those previously reported by Almond et al.[15] from one of their two very long 50 ns molecular dynamics simulations which matched the correct experimental NOE values. Clearly, our exhaustive search generated a much larger pool of allowed conformational regions. Another interesting feature of these plots is that both MM3 in implicit solvent and OPLS-AA in

Table 1. Potential Energy Differences (kcal/mol) of Four Selected Unique Structures of LNF-1 from Figure 1[a]

conformation	link 1	link 2	link 3	link 4	ΔE (GR)	ΔE (TK)
conf. 1	(11.0, 26.4)	(165.7, 11.3)	(23.9, 56.2)	(−177.9, −3.1)	0.0	0.0
conf. 2	(24.1, 22.7)	(45.5, 5.5)	(28.3, −53.7)	(41.8, 4.6)	5.96	4.33
conf. 3	(−4.1, −27.6)	(46.6, 169.5)	(25.0, −55.6)	(−13.3, −37.4)	6.51	4.864
conf. 4	(65.9, 73.1)	(54.7, 12.5)	(15.5, 34.2)	(17.4, −178.4)	11.15	8.563

[a] TK corresponds to energies calculated from TINKER using the MM3 force field and the implicit GBSA solvent model, and GR corresponds to calculations using GROMACS with OPLS-AA in the gas phase. The underlined pairs of angles are not within the circled regions corresponding to best NOE values shown in Figure 1. Only conf. 2 has all linkages within the circled regions.

Table 2. Comparison of NOEs Computed for Different Proton Pairs in Each of Our Selected Conformers of LNF-1 against Experimental and MD Values

				NOE calculated							
				conf. 1		conf. 2		conf. 3		conf. 4	
proton pairs		exp.[15]	MD[15]	ind.	MD	ind.	MD	ind.	MD	ind.	MD
F1 H2	F1 H1	5.5	6.9	10.5	13.6	10.9	14.0	10.6	14.0	13.7	14.4
F1 H5	F1 H1	0.6	0.4	0.5	0.5	0.5	0.5	0.6	0.5	0.4	0.5
F1 H3	F1 H5	4.9	6.5	5.4	4.0	4.8	3.9	5.4	4.2	5.6	4.0
F1 H4	F1 H5	5.8	7.5	7.9	6.7	7.4	6.8	7.7	7.0	8.4	6.9
1 H5	1 H1	6.2	6.8	6.1	9.7	7.7	9.8	6.3	9.5	9.4	9.9
1 H2	1 H1	2.5	1.8	1.5	2.6	2.3	2.7	1.8	2.5	2.1	2.7
1 H3	1 H1	5.7	5.0	4.4	5.1	6.6	5.1	4.8	5.6	5.1	5.2
1 H4	1 H5	5.8	7.2	7.9	7.6	8.6	7.5	8.2	7.3	7.9	7.4
2 H2	2 H1	2.2	1.8	4.2	2.6	2.6	2.5	2.1	2.0	3.9	2.6
2 H3	2 H1	7.0	2.1	5.4	4.0	3.8	3.9	8.0	6.0	5.7	4.1
2 H5	2 H1	7.7	9.9	11.3	10.4	9.1	10.5	9.3	9.8	11.6	11.3
4 H3	4 H1	3.0	4.0	7.3	4.3	6.9	4.1	5.9	4.3	4.1	3.7
4 H5	4 H1	7.8	6.4	9.7	11.0	10.6	11.1	10.8	11.0	7.2	10.5
1 H2	F1 H1	6.8	5.8	11.2	9.0	19.9	9.0	19.2	10.8	−0.4	9.9
1 H3	F1 H1	0.5	0.7	−1.0	0.9	−1.4	0.9	−0.2	0.6	6.0	0.6
1 H2	F1 H5	1.3	2.4	0.1	2.2	0.6	2.3	0.1	1.7	11.4	2.1
2 H2	F1 H5	6.8	4.7	0.1	9.8	7.0	10.0	0.0	0.0	0.2	8.8
2 H4	F1 H5	0.6	2.5	0.1	0.0	−0.3	0.0	0.2	0.1	0.3	0.2
2 H3	1 H1	5.7	5.6	0.5	7.8	9.2	7.9	0.3	0.5	7.9	8.1
3 H1	2 H1	0.4	0.3	−0.2	−0.3	−0.4	−0.2	−0.5	−0.3	−0.3	−0.3
3 H3	2 H1	11.5	9.8	6.1	7.2	10.5	7.2	11.5	8.5	13.5	7.5
3 H4	2 H1	0.8	1.1	−0.4	−0.2	3.6	0.0	5.2	0.3	−0.9	−0.3
3 H3	2 H5	0.5	0.9	−0.2	−0.3	−0.8	−0.4	−0.9	−0.6	−0.7	−0.3
4 H4	3 H1	11.0	11.3	0.4	0.6	13.7	11.7	11.9	11.6	0.4	8.6
	RMSD		0.74	1.41	0.78	1.01	0.74	1.50	0.77	2.89	0.73
	RMSD rank			2	4	1	2	3	3	4	1

[a] The columns labeled "MD[15]" and "exp.[15]" correspond to NOEs from the MD simulations and experimental measurements of Almond and co-workers in ref 15. The subcolumn labeled "ind." represents NOEs calculated from a single individual conformer, and that labeled "MD" corresponds to our time-averaged NOEs computed from short 5 ns MD trajectories using as starting conditions conformer 1, 2, 3, or 4.

the gas phase appear to produce similar sugar conformations. We suspect this is generally true for all available force fields even though relative energies in each case may be different. These energetic differences which we have observed with different force fields and ab initio calculations do not significantly affect our results since, as we have shown in the first paper, the energy ranking in implicit solvent does not generally coincide with the ranking of free-energy minima in solution, which is what determines the corresponding NOE values. Almond and co-workers' important study sheds light on the fact that even during very long molecular dynamics simulations the full configuration space is not readily visited. This is clear from the fact that their two trajectories produced significantly different NOE values. Only one of them being close to the correct experimental value. The reason for this is that, in the case of complex branched oligosaccharides, typical molecular dynamics time scales are not long enough to fully sample this space. Hence, our inexpensive a priori identification of conformational regions together with our ranking of structures based on their RMS deviation with respect to the corresponding experimental NOE values provides not only a good way to identify correct configurations in solution but also a way to generate initial conditions for further sampling with molecular dynamics in explicit solvent without having to rely on the trajectory to sample configuration space.

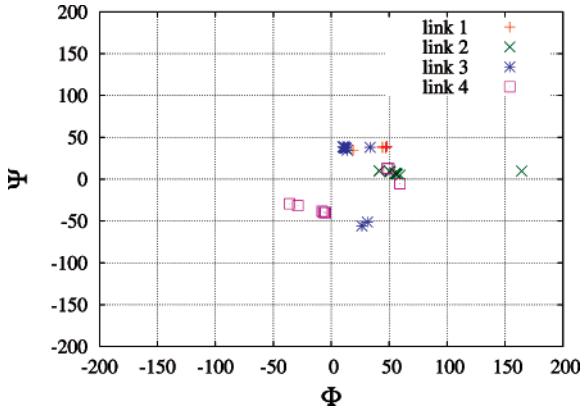

Figure 3. Distributions of all $\phi-\psi$ pairs from the 20 conformations of LNF-1 ranked with smallest RMSD from experimental NOE values. Most conformations are located within the circled regions in Figure 1.

Table 1 exhibits the potential energies and $\phi-\psi$ values for four selected conformations of LNF-1 from Figure 1. Configuration 1 is the lowest-energy minimum found by the

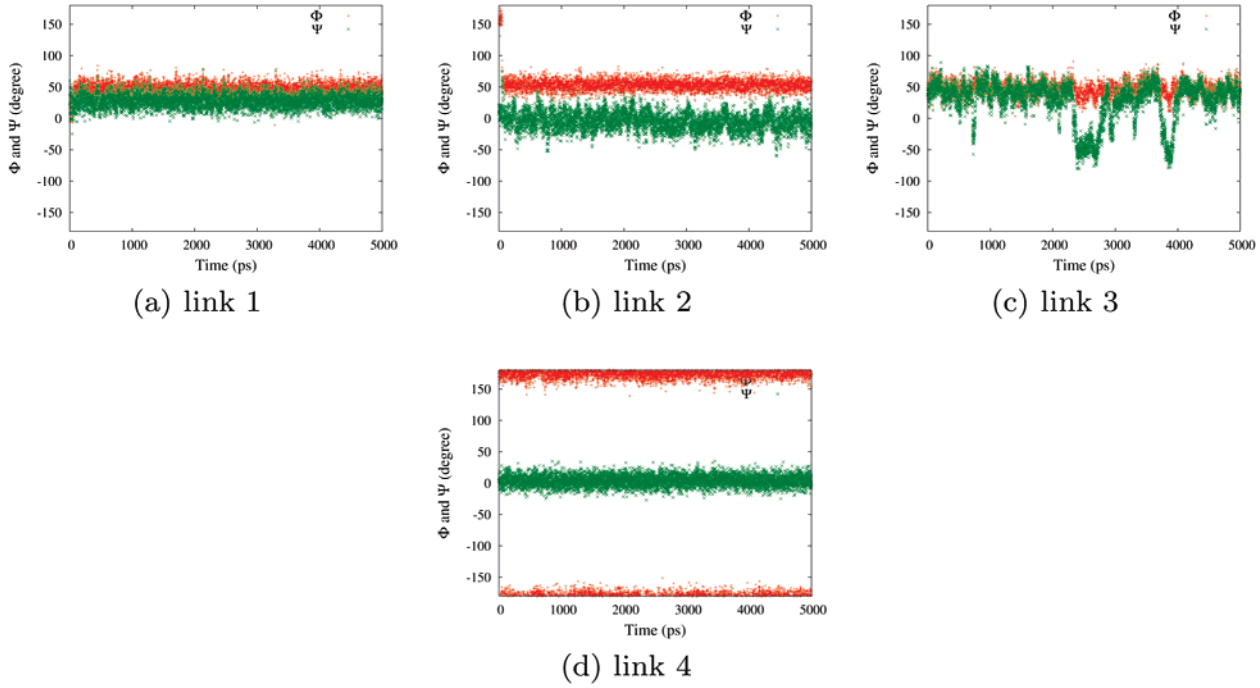

Figure 4. Time evolution of conformation 1 (global energy minimum) of LNF-1 as shown in Table 1 simulated using GROMACS with the OPLS-AA force field and the SPC explicit water model. Link 2 transfers to the circled region in Figure 1 relatively quickly (50 ps). In contrast, link 4 stays outside the circled region during our 5 ns run.

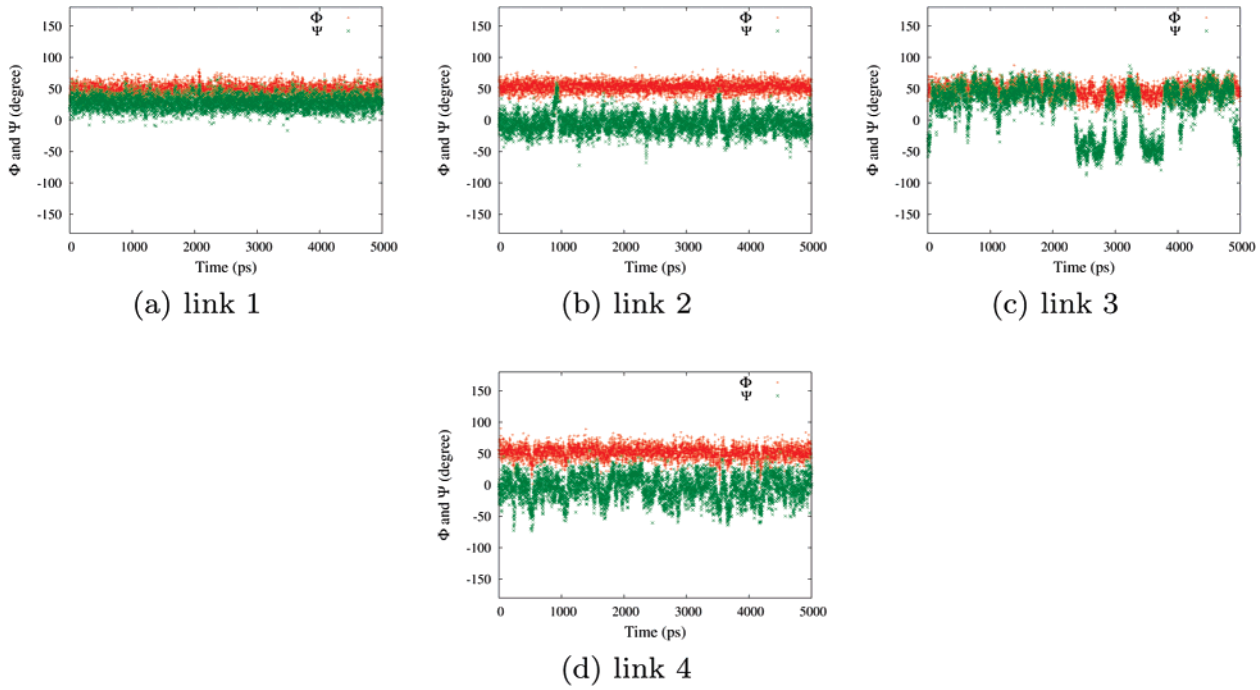

Figure 5. Time evolution of conformation 2 of LNF-1 as shown in Table 1 in explicit water. All linkages fluctuate around the initial values in the circled regions as shown in Figure 1. Time-averaged NOEs in Table 2 show that this final conformation is a good candidate for the most likely structure in solution.

algorithm in implicit solvent. Conformation 2 is our candidate for best structure in solution. For conformer 2, all values of $\phi-\psi$ pairs are located within the circled regions in Figure 1. Conformers in this region have minimal RMS deviation from the experimental NOE values. Conformations 1, 3, and 4 have at least one linkage outside of this region. In particular, conformation 1 (our global-energy minimum in implicit solvent) has two linkages outside the circled regions. The RMSD of its NOE values with respect to experiments is quite large as shown in Table 2. In contrast, conformation 2 is selected from Figure 3 and has the best NOE values compared to experiments,[15] but its energy in implicit solvent is much higher ($\Delta E \approx 5$ kcal/mol). This result is consistent with our findings in the first paper. The effects of explicit solvent and entropy must be taken into account to obtain a good approximation of the free energy of these systems.

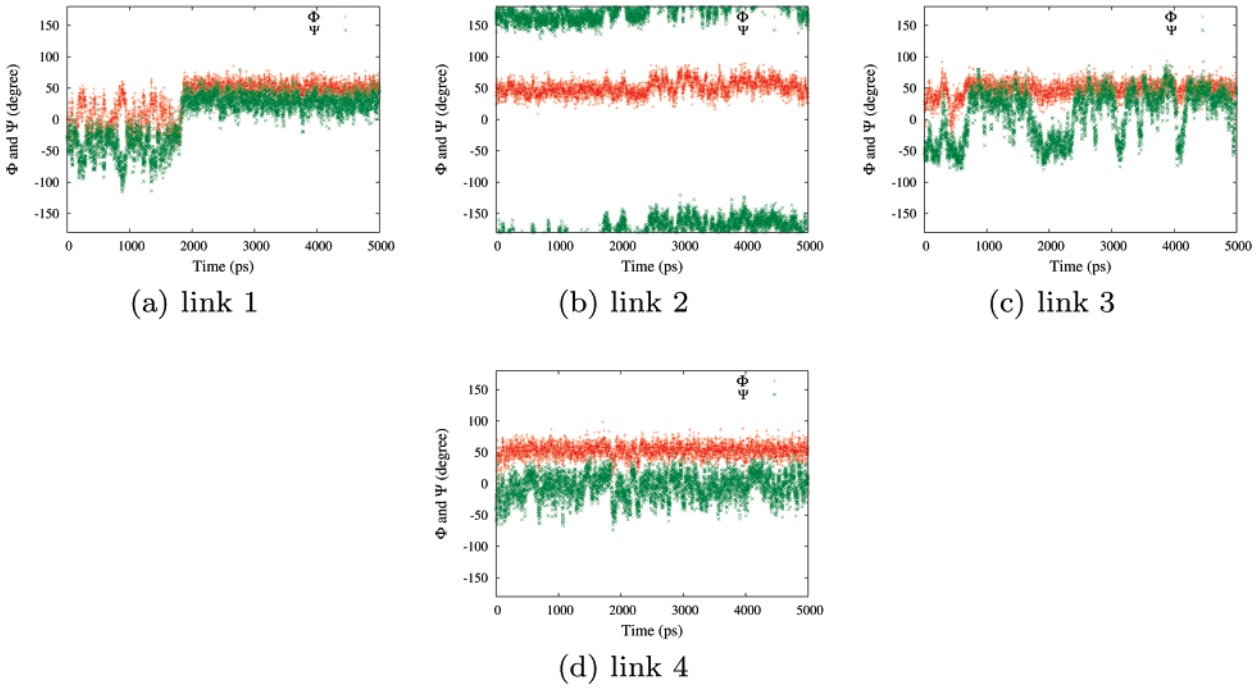

Figure 6. Time evolution of conformation 3 of LNF-1 as shown in Table 1 in explicit water. Link 1 shifts to the circled region in Figure 1 within 2 ns. Link 2 fluctuates outside the circled region.

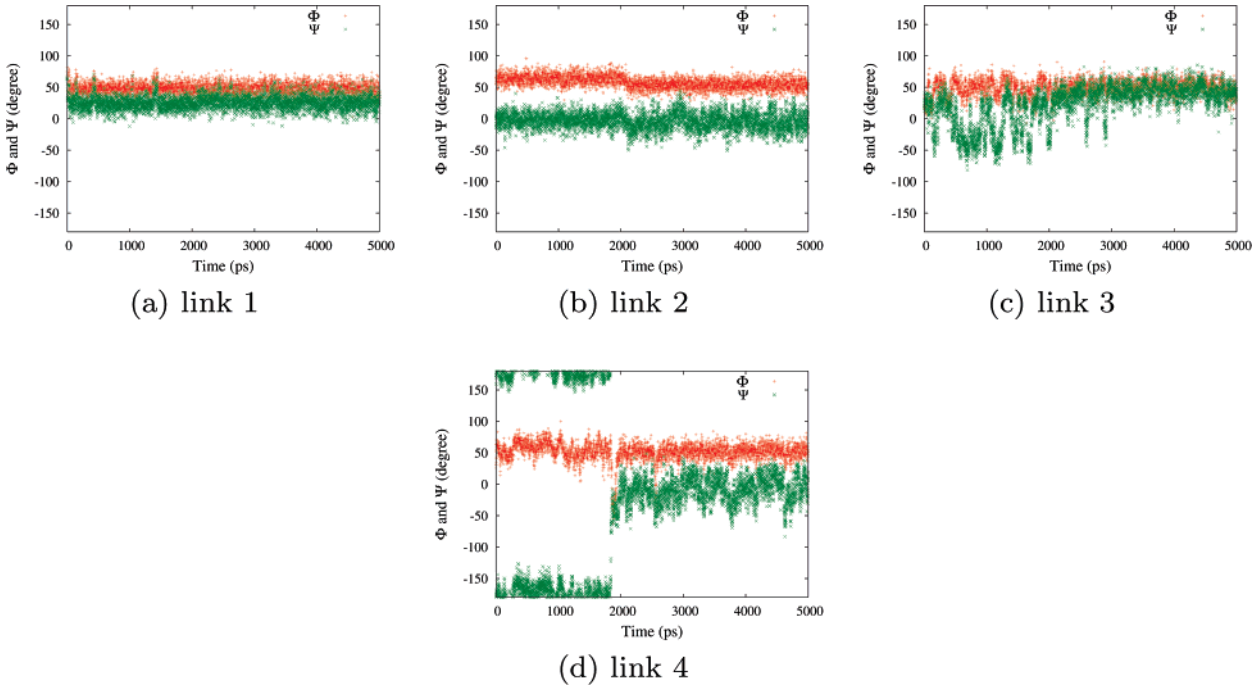

Figure 7. Time evolution of conformation 4 of LNF-1 as shown in Table 1 in explicit water. After the transition of link 4 around 2000 ps, the final conformation is the same as that of conformation 2 in Figure 5.

Conformations 3 and 4 have NOE RMSDs larger than that of conformation 2 and were chosen for comparison as initial conditions for molecular dynamics simulations.

It is clear from Table 2 that conformer 2 is the best candidate on the basis of NOEs. It is also clear from the same table that, for all initial conformers, solvent-averaged NOEs are closer to the experimental values than those resulting from single initial configurations. This is due to significant changes in conformation during MD that bring one or more glycosidic angles closer to the values of conformer 2.

We performed relatively short 5 ns MD simulations using the software GROMACS[35] with the OPLS-AA force field[37] in SPC water[38] for the four selected conformations in Table 1. We are particularly interested in understanding whether the regions that our algorithm singled out as most likely in solution on the basis of NOE RMSDs are stable during explicit solvent simulations or if structures in these regions

Table 3. Final Structures from 5 ns MD Simulations with Explicit Solvent for the Four Unique Conformations in Tables 1 and 4

	LNF-1			
	link 1	link 2	link 3	link 4
conf. 1	(50, 25)	(50, 0)	(50, −50/50)	(180, 0)
conf. 2	(50, 25)	(50, 0)	(50, −50/50)	(50, 0)
conf. 3	(50, 25)	(50, 180)	(50, −50/50)	(50, 0)
conf. 4	(50, 25)	(50, 0)	(50, −50/50)	(50, 0)

	LND-1				
	link 1	link 2	link 3	link 4	link 5
conf. 1	(50, 25)	(50, 0)	(50, 25)	(50, −50/50)	(50, 0)
conf. 2	(50, 25)	(50, 0)	(50, 25)	(50, −50/50)	(50, 0)
conf. 3	(50, 25)	(50, 180)	(50, 180)	(50, −50/50)	(50, 180)
conf. 4	(50, 25)	(50, 0)	(50, 25)	(50, −50/50)	(180, 0)

[a] The $\phi-\psi$ values are in degrees. For LNF-1, initial conf. 2 and initial conf. 4 result in identical final conformations; the final conformations for initial conf. 1 and initial conf. 3 have only one linkage (link 4 or link 2) that is different from conf. 2. For LND-1, confs. 1 and 2 share identical final conformations after MD. Initial confs. 3 and 4 have different final conformations.

undergo significant configurational modifications. Other structures that are also local energy minima but possess several linkages outside this selected configuration space region were studied in order to gauge whether barriers to interconversion were readily crossed.

Figures 4−7 show the time evolution of the dihedral angles for the four selected conformations in explicit water. Final conformations for these four runs are listed in Table 3. Initial conformation 2 (Figure 5) has all (ϕ,ψ) values within the dihedral regions being the best NOE values as compared with those of experiments. Throughout our 5 ns simulation, the trajectory corresponding to initial conformation 2 does not depart from the angular areas circled in Figure 1. It is interesting to notice that these areas correspond to two clearly different conformations that share identical link angles 1, 2, and 4, but link 3 transitions between $\psi = -50°$ and $\psi = +50°$. Both of these structures have good NOEs for the protons considered on each side of linkage 3 since at $\psi = -50°$ and $\psi = +50°$ the proton distances involved are very similar as can be appreciated from Figure 8. This is a clear example that shows how experimental NOEs may correspond to a linear combination of structures in different local basins instead of an average over structures in a single free-energy minimum. One should therefore be careful when experimentally assigning a structure simply on the basis of NOE constraints since these may not correspond to a single structure, but instead to a combination of several different structures.

Results from our simulations with configuration 1 (the global-energy minimum in implicit solvent) as the initial condition are shown in Figure 4. In this case, link 2 transfers to dihedral angles similar to those of configuration 2 within 50 ps. Just as in the case of starting structure 2, link 3 fluctuates between $\psi = +50°$ and $\psi = -50°$. The angles corresponding to link 4 remain almost constant throughout our 5 ns simulations and are different from those in configuration 2. The resulting time-averaged NOEs for trajectories with configuration 1 as the initial condition are tabulated in Table 2. Since most linkages undergo rotations

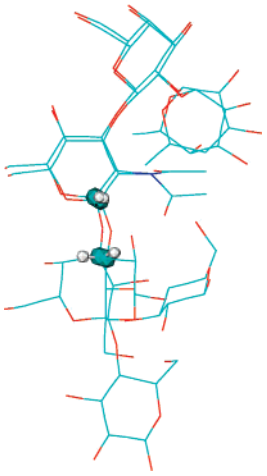

Figure 8. Two conformations of LNF-1 that share identical link 1, 2, and 4 angles but have different link 3 angles ($\psi = -50°$ and $\psi = +50°$). Both of these two structures have NOEs close to the experimental values for the proton pair (3H3−2H1) considered on each side of linkage 3. This is because the distance in each case between the protons considered is very similar, 2.36 and 2.51 Å, respectively.

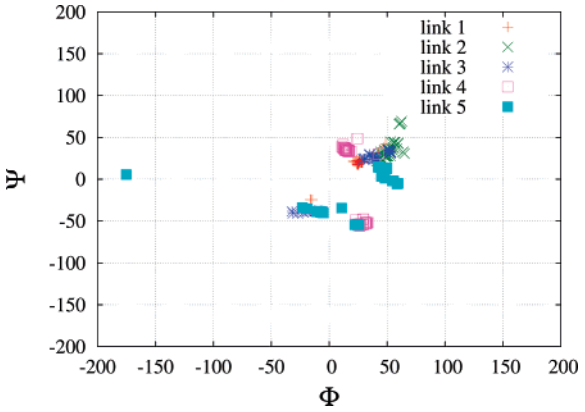

Figure 9. Distribution of $\phi-\psi$ angle pairs corresponding to the 20 conformations of LND-1 with smallest NOE RMSD. Most conformations fall inside the circled regions in Figure 2.

Table 4. The Potential Energy Differences of Four Sterically Allowed Minimized Structures of LND-1 Selected from Figure 2[a]

conformation	link 1	link 2	link 3	link 4	link 5	Δ*E* (GR)	Δ*E* (TK)
conf. 1	(27.6, 20.2)	(62.7, 22.1)	(72.3, 51.6)	(14.9, 38.1)	(−175.1, 5.4)	0.0	0.0
conf. 2	(28.4, 20.7)	(48.6, 27.2)	(51.5, 33.9)	(11.7, 37.8)	(−5.5, −40.0)	1.62	1.14
conf. 3	(65.9, 64.4)	(61.9, −155.8)	(18.0, 176.5)	(12.4, 39.1)	(−175.1, 4.8)	5.14	7.55
conf. 4	(27.9, 20.2)	(47.2, 28.0)	(52.9, 34.3)	(159.2, −37.7)	(−178.3, 0.5)	14.46	8.40

[a] Conf. 1 is the global energy minimum in implicit solvent. Conf. 2 has all linkages within the circled regions in Figure 2 and corresponds to the one with the closest NOE values to experimental values. TK corresponds to energies calculated from Tinker using the MM3 force field and the implicit GBSA solvent model, and GR corresponds to calculations using GROMACS with OPLS-AA in the gas phase. The underlined pairs of angles are not within the circled regions corresponding to the best NOE values shown in Figure 2. Only conf. 2 has all linkages within the circled regions.

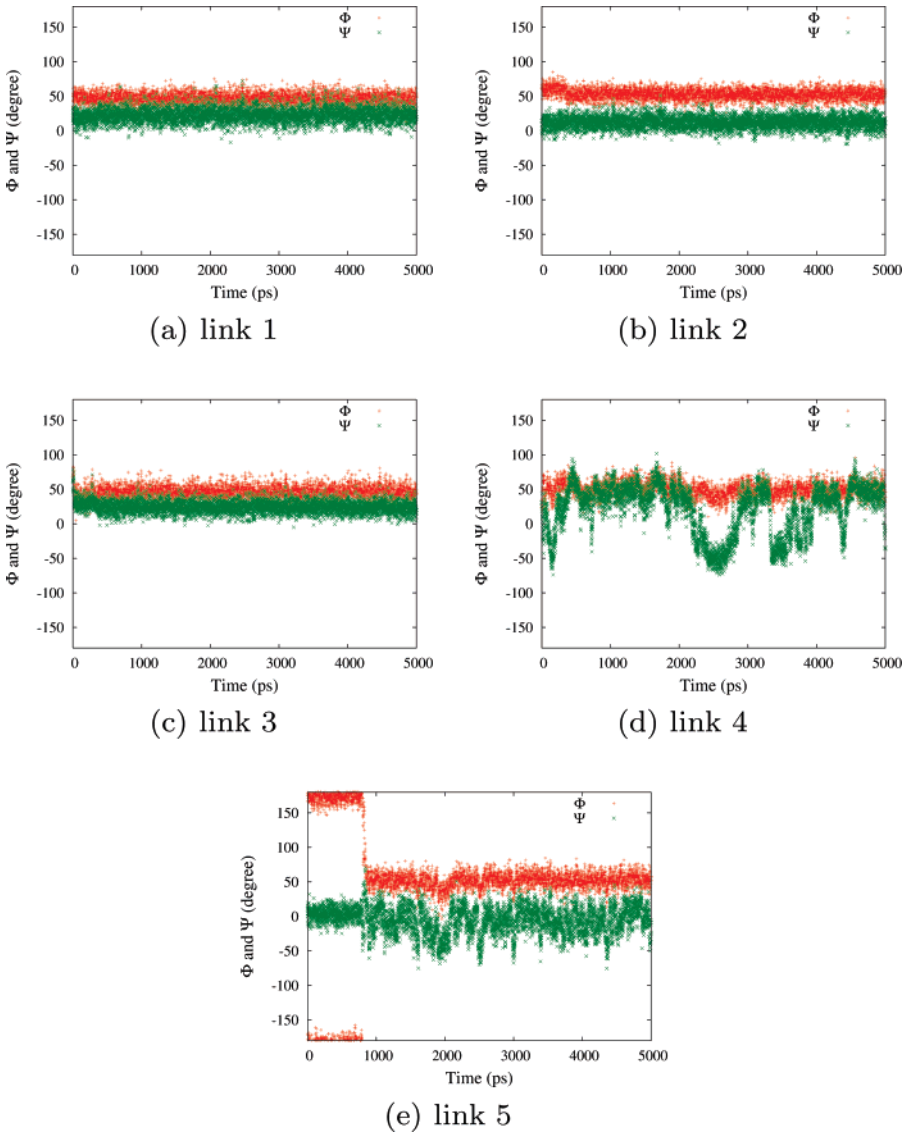

Figure 10. Time evolution of conformation 1 (global energy minimum in implicit solvent) of LND-1 as shown in Table 4 simulated using GROMACS with the OPLS-AA force field and the explicit SPC water model. Initially, link 5 is outside the circled region in Figure 2, but it switches into this region after 1000 ps. The small RMSD (Table 5) indicates that this final conformation is a good candidate for the most likely structure in solution. This structure is the one previously identified by long 50 ns MD simulations.[15]

to final configurations analogous to that of structure 2, it is not surprising that the value of the NOE RMSD is much smaller than that from initial structure 1. Since link 4 is a terminal residue and is far from the crowded linkage, it is likely that on a longer time scale it will undergo conformational changes. The case of conformation 3 is quite different from the previous two. After 2 ns, link 1 (Figure 6) shifts to the same angles as in conformer 2, but link 2 does not. Therefore, the NOEs corresponding to 2 H3 to 1 H1 in Table 2 are quite different from those experimentally observed. In the case of conformer 4, after link 4 of Figure 7 undergoes a transformation at around 2 ns, the conformation of the molecule is identical to that of conformer 2.

3.1.2. LND-1 Milk Sugar. In order to study LND-1, we follow two different approaches. The first one is analogous to our procedure in the case of LNF-1, and it involves the search of the whole dihedral space. The second approach, which we use for comparison and benchmarking, makes use of our substructure recognition algorithm and database. This approach generates "unique conformations" for LND-1 on the basis of a database entry previously stored for the conformations of its subfragment LNF-1.

In the first case, after the coarse-grained grid search, we obtained only 9071 sterically allowed conformations as opposed to the case of LNF-1 (the smaller oligosaccharide) in which our algorithm found 24 041 structures. This is interesting since adding degrees of freedom to the system appears to reduce instead of increase the number of accessible regions in dihedral space. The additional branch α-L-Fuc in LND-1 is the cause for this reduction in number of allowed conformations. We expect this to be a general trend in sugars that are branched, particularly those with adjacent linkages. An extreme case in which torsional degrees of freedom are reduced to a minimum is that of cyclodextrins. When

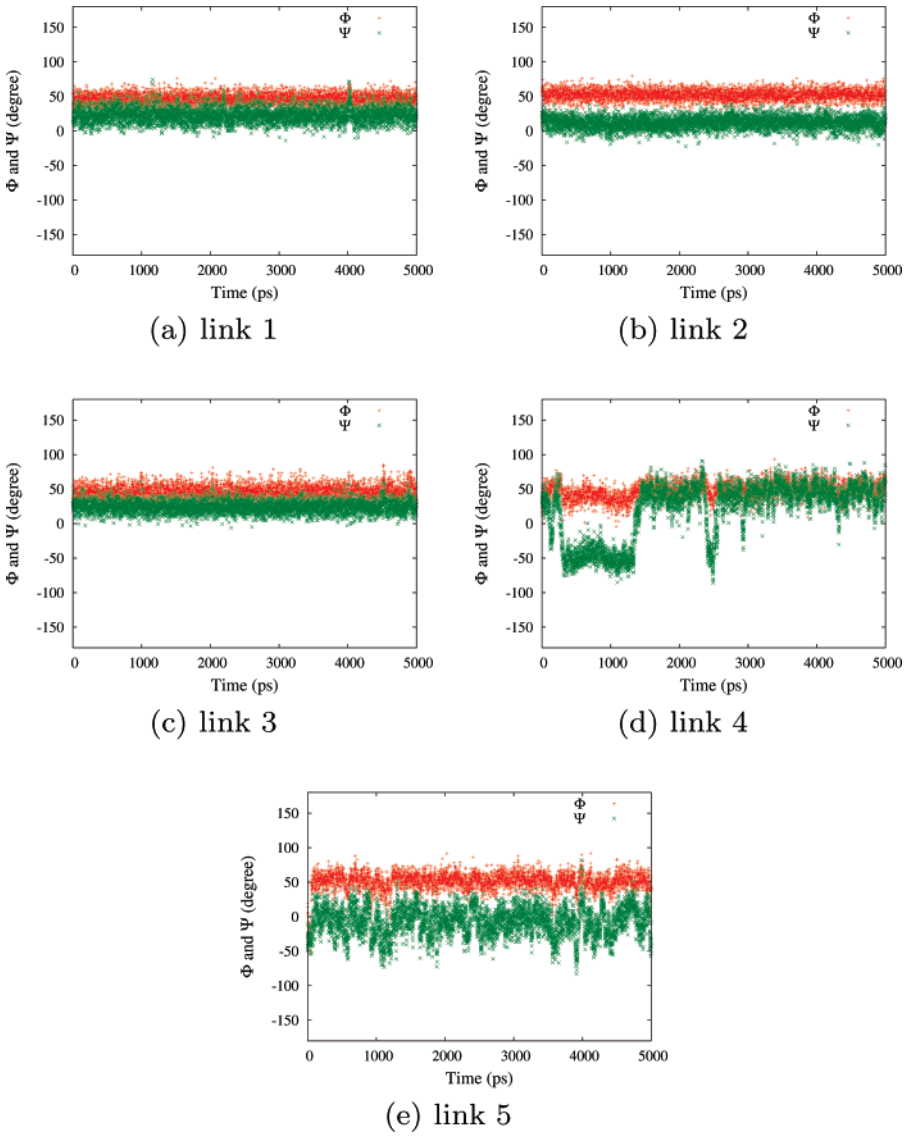

Figure 11. Time evolution of conformation 2 of LND-1 as shown in Table 4 in explicit water. All linkages fluctuate in the circled regions depicted in Figure 2. The final conformation is the same as that resulting from the MD simulation of conformation 1 (Figure 10).

identical angular and energetic criteria are used as previously described in the case of LNF-1, energy minimizations using the MM3 force field and the GBSA implicit solvent model produced 989 "unique conformations".

The set of insert graphs in Figure 2 show the distributions of all $\phi-\psi$ pairs corresponding to the 989 unique conformations of LND-1. For comparison, we also show the 5220 minima obtained using OPLS-AA in the gas phase. It is clear that several energy-minimized regions in $\phi-\psi$ space are present in addition to those previously found during MD simulations[15] (circled regions in Figure 2). Results using OPLS-AA and MM3 are qualitatively similar. When we compare LND-1 with LNF-1, we notice that in the case of LND-1 the width of certain allowed regions is narrowed due to the presence of the additional α-L-Fuc branch. Furthermore, some regions in dihedral space completely dissapear in the case of LND-1. As an example, the region around (180°, 0°) for link 2 (β-D-Galp-(1→3)-β-D-GlcpNAc) is absent in the case of LND-1.

We have ranked the LND-1 conformations on the basis of their energy in implicit solvent and also on the basis of their RMSD with respect to experimental NOEs. Figure 9 exhibits the distributions of all $\phi-\psi$ pairs of the best 20 LND-1 conformations on the basis of RMSD. Most conformations appear to be located in the regions circled in Figure 2.

In Table 4, we compare four conformations from Figure 2. Just as in the case of LNF-1, we have chosen these four structures because one is the global-energy minimum in implicit solvent; the second one is the structure ranked best on the RMSD scale, while the third and fourth are allowed minimum-energy structures that have not been previously reported computationally but are not in the correct regions according to our NOE calculations. Table 4 displays potential-energy values and corresponding $\phi-\psi$ values.

Just as in the case of LNF-1, without the need of expensive MD simulations in explicit solvent, a simple NOE ranking based on our exhaustive search algorithm was able to

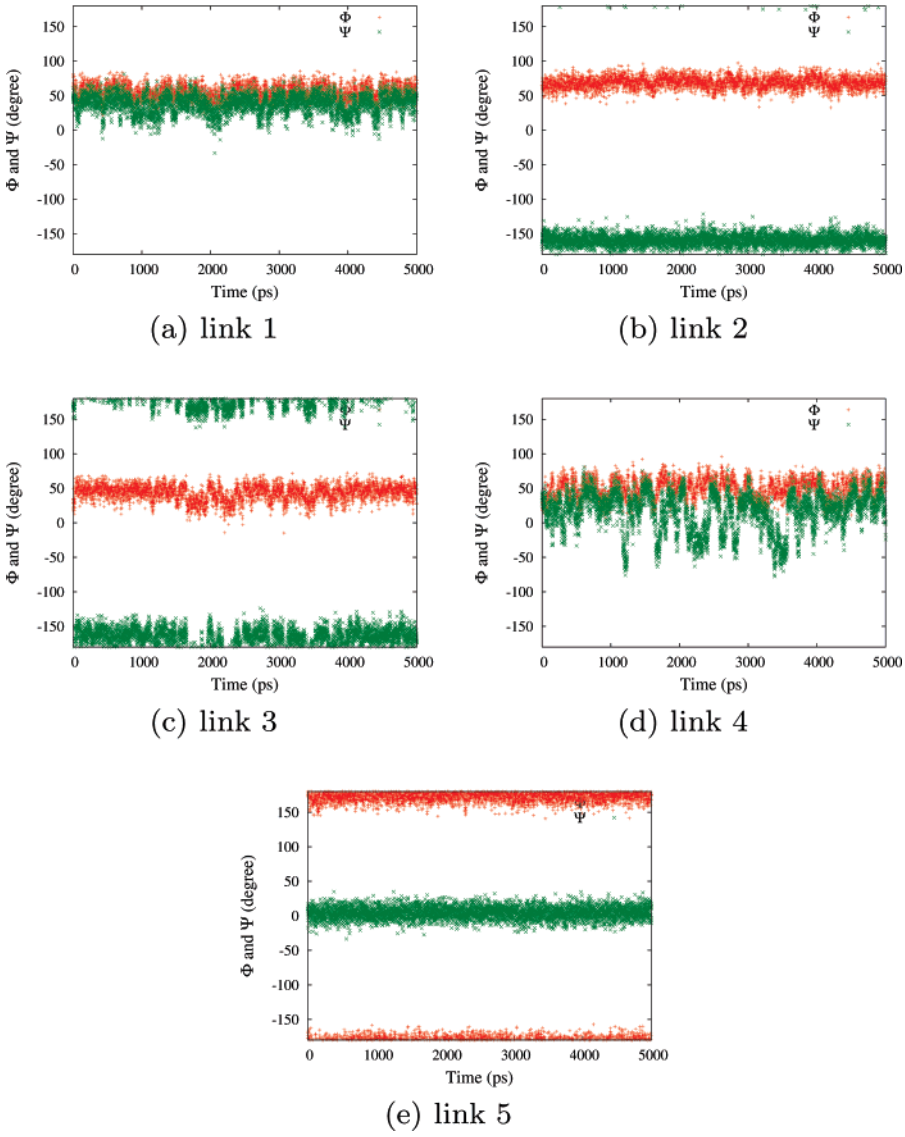

Figure 12. Time evolution of LND-1 conformation 3 as shown in Table 4 in explicit water. Links 2, 3, and 5 fluctuate outside the circled regions in Figure 2. The RMSD with respect to experimental values is large.

efficiently identify the correct regions of configuration space in which conformation 2 is located. This region coincides with that proposed by Almond and co-workers[15] from one of their trajectories that was initiated from an appropriate initial conformation. Our calculation was carried out in less than 1 day on a single-processor desktop PC.

In order to quantify the advantage of using a rotameric substructure database and a substructure matching algorithm, we also analyzed LND-1 using LNF-1 as a database entry. When a new sugar is added to the database, our program adds an entry with the following information: the residue names and topology (residue connectivity, chirality of atoms, etc.) and a unique name for a file in which vectors of allowed dihedral conformations are stored representing points in dihedal phase space from which the whole oligosaccharide can be reconstructed. When in search mode, the program checks all entries in the database and calls our substructure mapping algorithm in order to determine whether there is any molecule in the database that could potentially be a substructure of the new molecule. If several substructures are available, only the largest one is used in order to build a model for the new molecule. These models take as starting points the vectors of dihedral angles stored for the subfragment and only do full searches on the parts of the molecule not originally stored as a substructure in the database. All vectors in the database for that particular substructure are used as starting points to obtain the full dihedral phase space for the new molecule. Every time a new vector from the subfragment is retrieved, the new molecule is reasembled by adding the remaining residues and side-chain linkages. In the case of LND-1, after the conformational search and database storage for LNF-1 was performed, the search for LND-1 was carried out by simply adding one residue, which provides a branching point. The CPU time for a full search of the dihedral space of LND-1 previously described in this paper was 2665 s. In contrast, it only took 1114 s to search using the database.

For the LND-1 milk sugar, Figures 10−13 exhibit the dynamics in explicit solvent of the four initial conformations shown in Table 4. All final conformations are displayed in

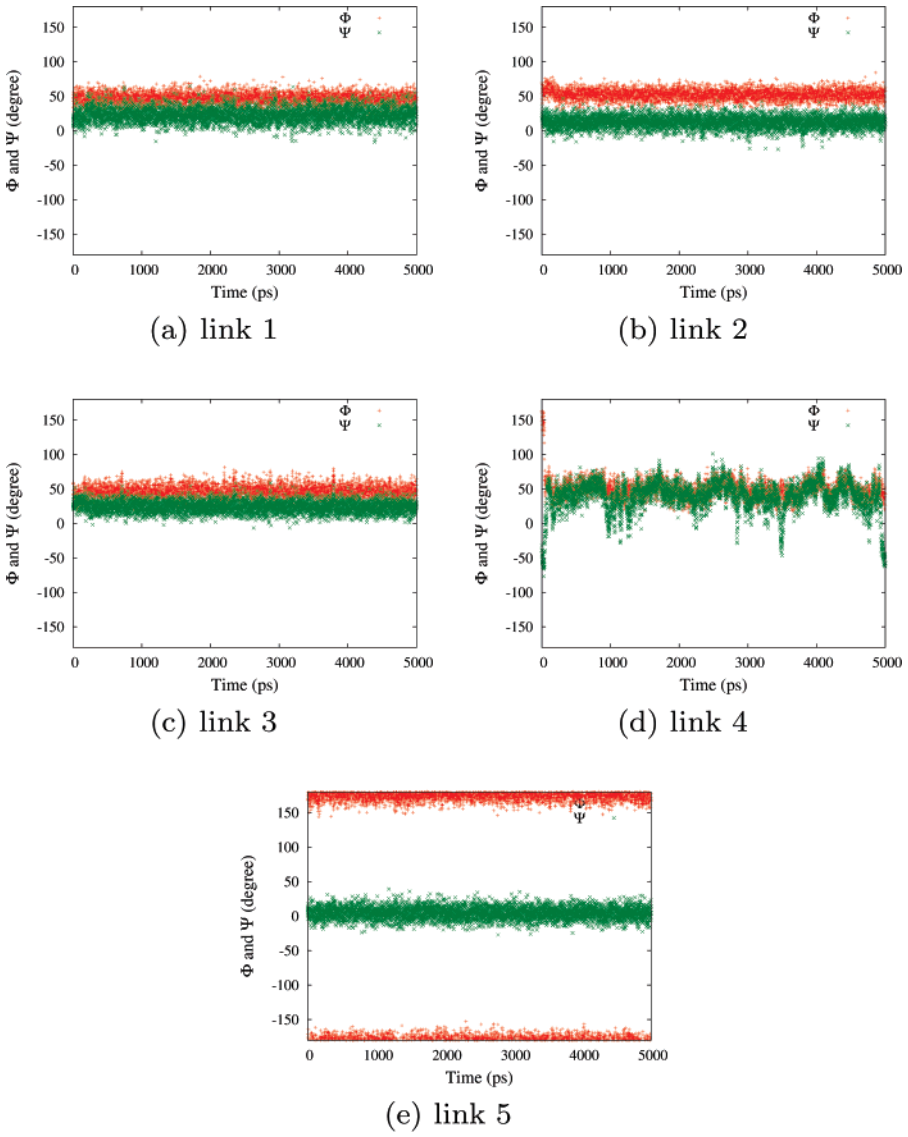

Figure 13. Time evolution of dihedral angles in LND-1 conformation 4 (Table 4) in explicit water. Initially, links 4 and 5 are not within the circled regions in Figure 2. Link 4 transfers into the circled region within a few picoseconds.

Table 3. Similar results to those obtained in the case of LNF-1 are observed. Only conformation 2 (see Figure 11) has all initial dihedral angle values within the circled NOE regions shown in Figure 2. Figure 10 shows that link 5 of conformation 1 (the global-energy minimum in implicit solvent) evolves toward the circled region in Figure 2 within 1000 ps. This final conformation is the one identified by Almond and co-workers[15] and by our prediction algorithm as being the most likely in solution. As is to be expected, the time-averaged RMSD of NOE values is small, as can be appreciated in Table 5. Our time-averaged NOEs appear to be slightly worse than those in ref 15. This is reasonable since their simulations are much longer (50 ns) and their methodology involves obtaining a time correlation function, while in our case, for economy of time in our prediction procedure, we simply average over NOE values of individual snapshots along a 5 ns trajectory.

Figure 11 shows the results from our MD simulations with initial conditions corresponding to conformation 2 (the structure a priori predicted to have the best NOEs). In all cases, dihedral angles fluctuate within the circled regions in Figure 2. The small time-averaged RMSD with respect to experiments (Table 5) indicates that this conformation is indeed a stable structure and the best candidate for solution conformation.

In the case of conformation 3 in Figure 12, links 2, 3, and 5 fluctuate outside the circled regions, and this results in a large RMSD with respect to experimental values. Figure 13 shows that, in the case of conformation 4, link 4 transfers to the corresponding circled region while link 5 remains outside the corresponding circled areas. Because of the small overall deviation of the time-averaged RMSD values (Table 5) with respect to experiments, it is possible that the final conformations from this trajectory correspond to actual structures often visited in solution. This is likely since the only difference between these structures and those in the free-energy basin corresponding to conformer 2 is the terminal unit far from the set of crowded linkages. This situation was also observed in the case of the MD simulation of conformer 1 of LNF-1.

Table 5. Table NOE Values for Different Proton Pairs for the Four Selected Structures of LND-1 from Table 4[a]

proton pairs		exp.[15]	MD[15]	NOE calculated							
				conf. 1		conf. 2		conf. 3		conf. 4	
				ind.	MD	ind.	MD	ind.	MD	ind.	MD
1 H5	1 H1	6.1	8.9	6.6	8.4	7.3	8.5	7.0	9.0	7.2	8.5
1 H3	1 H1	4.1	6.8	5.7	3.8	7.0	3.7	5.3	4.9	6.9	3.9
2 H3	2 H1	9.9	5.6	7.4	4.2	6.6	4.2	5.7	5.4	5.2	4.3
2 H5	2 H1	13.2	10.3	10.4	8.6	10.6	8.9	7.3	8.0	10.9	9.1
F1 H2	F1 H1	8.8	9.9	9.8	12.6	9.9	12.6	12.9	13.2	9.9	12.4
F1 H3	F1 H5	7.9	9.3	4.7	3.8	4.5	3.8	4.9	3.2	4.6	3.7
F1 H4	F1 H5	10.7	11.0	7.2	6.3	7.1	6.1	7.3	6.2	5.4	6.3
F2 H2	F2 H1	13.3	12.0	9.0	12.4	9.6	11.9	11.3	11.0	9.6	12.6
F2 H3	F2 H5	8.4	9.3	3.0	3.2	4.0	3.3	5.0	4.5	4.1	3.2
F2 H4	F2 H5	14.5	11.2	6.8	6.0	6.5	6.0	7.6	6.7	6.5	5.9
2 H3	1 H1	4.4	8.9	5.2	6.6	6.8	6.5	0.2	0.4	6.3	6.5
3 H3	2 H1	16.5	14.7	11.5	6.8	12.4	6.5	11.5	7.9	1.0	5.6
4 H4	3 H1	12.7	14.9	0.3	8.9	10.9	10.8	0.3	0.6	0.4	0.6
1 H2	F1 H1	8.5	9.7	11.3	6.7	12.0	6.8	1.8	5.1	12.1	6.7
2 H2	F1 H5	9.3	10.0	5.5	10.2	9.2	10.3	0.0	0.1	5.2	10.3
2 H4	F1 H5	2.0	1.5	0.2	0.1	0.2	0.0	0.0	0.0	0.5	0.1
2 H3	F2 H1	0.6	0.4	0.3	0.0	0.1	0.0	15.9	14.1	0.1	0.0
2 H4	F2 H1	9.8	8.5	0.6	8.4	5.0	8.0	0.3	0.4	4.6	8.1
2 H5	F2 H1	1.7	1.4	1.0	0.2	0.3	0.1	8.1	2.1	0.2	0.2
1 H2	F2 H5	8.4	10.3	2.1	7.1	7.8	6.9	0.0	0.0	7.8	6.9
2 H3	F2 H5	1.1	0.9	0.2	0.5	0.6	0.5	0.0	0.0	0.6	0.5
2 H4	F2 H5	2.4	1.9	9.5	1.4	3.3	1.5	0.5	0.3	3.3	1.4
	RMSD		0.33	0.83	0.52	0.48	0.52	5.53	4.83	0.56	0.56
	RMSD rank			3	1	1	1	4	4	2	3

[a] Similar to the case of LNF-1 in Table 2, the global energy minimum in an implicit solvent (conf. 1) does not have the best NOE values. Conf. 2 has the best NOEs but is ranked higher on an implicit solvent potential-energy scale. Except in the case of conf. 2, MD-averaged NOEs appear to be closer to experimental data.

By comparing each linkage of the best solution conformation for LND-1 with corresponding linkage for LNF-1, it is easy to see that $\phi-\psi$ values are similar, as displayed in Table 3. The extra linkage prevents certain configurations but otherwise preserves the oligosaccharide fold.

4. Conclusions

We have developed a sugar structure prediction tool based on a ring perception algorithm, automatic recognition of rotable dihedrals, Euler rotations, implicit solvent minimizations, NOE calculations, and molecular dynamics in explicit solvent. We have also implemented a subtree recognition algorithm for finding an oligosaccharide fragment within a more complex molecule and a database for storing structural and rotameric information. Oligosaccharides are complex topological molecules with multiple possible branching points. Since dihedral rotations are strongly coupled, particularly in the case of adjacent linkages or when branching is present, the use of a simple rotameric library to study conformations of these systems is many times not feasible. Our database and subtree recognition algorithm overcomes this problem by storing all coupled $\psi-\phi$ regions of oligosaccharide fragments as vectorial quantities that can be queried when a larger sugar is presented to the program.

This automatic tool for sugar structure prediction was applied to the case of LNF-1 and LND-1, two related oligosaccharides present in milk. Our tool identified and pooled all important "unique conformations" for these oligosaccharides. The distribution of these unique conformations is much wider than previously reported from MD simulations. Structures appear to be clustered around distinct important regions. Previous MD studies show that even very long (50 ns) molecular dynamics studies do not reproduce the correct experimental NOEs unless initial conditions are carefully chosen. This is because in a complex oligosaccharide the sampling of angular space is slow on a molecular dynamics time scale and the molecule remains trapped for very long times in local minima that do not necesarily correspond to the solution structure. Our algorithm overcomes this problem by brute-force sampling of the whole dihedral angular space. Since sugars are bulky and can be branched, this search does not exponentially explode and in fact is much faster than sampling allowed conformations with molecular dynamics or Monte Carlo techniques. Once a full dihedral space search is accomplished, structures are pooled by an implicit solvent minimization procedure. Simple NOE calculations and ranking against experimental data reveals in a very short time which of all allowed minimized conformations are most likely to exist in solution. Short-time MD simulations (5 ns) for different initial structures sampled according to our algorithm allow us to test whether these are stable in explicit solvent and provide a good strategy to obtain "local basin averaged" NOE values. In this article and in the first paper,[46] we have shown that other ranking criteria such as implicit solvent energies are poor estimators of the free-energy difference of oligosaccharides in solution. In general, the lowest-energy structure in implicit solvent is not the overall free-energy minimum in solution and does not correspond to a structure that has correct NOE values as compared to experimental data.

Our algorithm was successful in finding the best possible candidate structures in solution for LNF-1 and LND-1 in a very short time and without the need to resort to MD simulations. Our MD simulations confirm the fact that these structures are indeed stable in solution because when initial conditions were given in these regions of dihedral angular space we did not observe departure from the corresponding basin thoughout our simulations. This is not the case for other

studied initial structures with linkages that although allowed did not match the experimental NOEs. An exception to this is terminal residues far apart from crowded linkages for which rotations are most likely decoupled from other glycosidic torsions.

An interesting question that arises in the case of oligosaccharides is whether as in the case of small proteins a clear fold exists. We find that these two sugars have particularly ordered structures. The additional branch of α-L-Fuc in LND-1 has some influence on the conformations of other linkages, but most allowed conformations for LNF-1 are also allowed in the case of LND-1. It is interesting that we have identified fewer allowed conformations in the case of the larger sugar than in that of the smaller one, pointing to the fact that branching and crowded linkages can indeed shrink the conformational space of larger sugars. The question of whether small sugars have well-defined folded structures in general is interesting and should be the focus of future studies.

Acknowledgment. This research was funded by Grant #05-2182 from the Roy J. Carver Charitable Trust awarded to C.J.M. and by the Skou Fellowship from the Danish Natural Sciences Research Council awarded to J.H.J.

References

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683−720.

(2) Rudd, P. M.; Dwek, R. A. *Crit. Rev. Biochem. Mol. Biol.* **1997**, *32*, 1−100.

(3) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664−6676.

(4) Duus, J. Ø.; Gotfredsen, C. H.; Bock, K. *Chem. Rev.* **2000**, *100*, 4589−4614.

(5) Wormald, M. R.; Petrescu, A. J.; Pao, Y.-L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371−386.

(6) Martin-Pastor, M.; Bush, C. A. *Biochemistry* **1999**, *38*, 8045−8055.

(7) Martin-Pastor, M.; Bush, C. A. *Carbohydr. Res.* **2000**, *323*, 147−155.

(8) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6676−6683.

(9) Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567−4588.

(10) Bush, C. A.; Martin-Pastor, M.; Imbery, A. *Annu. Rev. Biophys. Biomolec. Struct.* **1999**, *28*, 269−293.

(11) Woods, R. *Glycoconjugate J.* **1998**, *15*, 209−216.

(12) Woods, R. J. *Curr. Opin. Struct. Biol.* **1995**, *5*, 591−598.

(13) Woods, R. J. The Application of Molecular Modeling Techniques to the Determination of Oligosaccharide Solution Conformations. In *Reviews of Computational Chemistry*; Lipkowitz, K., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 9, pp 129−165.

(14) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10541−10545.

(15) Almond, A.; Petersen, B. O.; Duus, J. *Biochemisty* **2004**, *43*, 5853−5863.

(16) Almond, A.; Sheehan, J. K. *Glycobiology* **2003**, *13*, 255−264.

(17) Almond, A.; Bunkenborg, J.; Franch, T.; Gotfredsen, C. H.; Duus, J. Ø. *J. Am. Chem. Soc.* **2001**, *123*, 4792−4802.

(18) Woods, R. J.; Pathiaseril, A.; Wormald, M. R.; Edge, C. J.; Dwek, R. A. *Eur. J. Biochem.* **1998**, *258*, 372−386.

(19) Imberty, A.; Tran, V.; Pérez, S. *J. Comput. Chem.* **1989**, *11*, 205−216.

(20) Imberty, A.; Gerber, S.; Tran, V.; Pérez, S. *Glycoconjugate J.* **1990**, *7*, 27−54.

(21) Kocă, J. *THEOCHEM* **1994**, *308*, 13−24.

(22) Kocă, J. *Prog. Biophys. Mol. Biol.* **1998**, *70*, 137−173.

(23) Engelsena, S. B.; Kocab, J.; Braccinic, I.; Penhoatc, C. H.; Pérez, S. *Carbohydr. Res.* **1995**, *271*, 1−29.

(24) Peters, T.; Meyer, B.; Stuike-Prill, R.; Somorjai, R.; Brisson, J.-R. *Carbohydr. Res.* **1993**, *238*, 49−73.

(25) Nahmany, A.; Strino, F.; Rosen, J.; Kemp, G. J.; Nyholm, P.-G. *Carbohydr. Res.* **2005**, *340*, 1059−1064.

(26) Strino, F.; Nahmany, A.; Rosen, J.; Kemp, G. J.; Sá-correia, I.; Nyholm, P.-G. *Carbohydr. Res.* **2005**, *340*, 1019−1024.

(27) Newburg, D. S.; Ruiz-Palacios, G. M.; Morrow, A. L. *Annu. Rev. Nutr.* **2005**, *25*, 37−58.

(28) Martin-Pastor, M.; Bush, C. A. *Biochemistry* **2000**, *39*, 4674−4683.

(29) Landerjö, C.; Jansson, J. L. M.; Maliniak, A.; Widmalm, G. *J. Phys. Chem. B* **2005**, *109*, 17320−17326.

(30) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933−5947.

(31) Ponder, J. W. F. M. R. *J. Comput. Chem.* **1987**, *8*, 1016−1024.

(32) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.

(33) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, *100*, 19824−19839.

(34) Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1995**, *246*, 122−129.

(35) van der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. *J. Comput. Chem.* **2005**, *26*, 1701−1718.

(36) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Biol.* **2001**, *7*, 306−317.

(37) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *117*, 11225−11236.

(38) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981.

(39) Nosé, S. *Mol. Phys.* **1984**, *52*, 255−268.

(40) Hoover, W. G. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31*, 1695−1697.

(41) Berendsen, H. J. C.; Postma, J. P. M.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(42) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.

(43) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4559−4570.

(44) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(46) Xia, J.; Daly, R. P.; Chuang, F.-C.; Parker, L.; Jensen, J. H.; Margulis, C. J. update: **2007**, *4*, 1620−1628.

CT700034Q