Anomeric Free Energy of D-Mannose in Water: A Comparison of Free Energy Perturbation, Potential of Mean Force, and MC(JBW)/SD Simulations

Hanoch Senderowitz* and W. Clark Still

Department of Chemistry, Columbia University, New York, New York 10027

Received: August 13, 1996[⊗]

The anomeric free energy difference between the α (axial) and β (equatorial) anomers of D-mannose was calculated using standard free energy methods (free energy perturbation (FEP) and potential of mean force (PMF)) and with the recently described MC(JBW)/SD simulation technique. The final results obtained with our new united atom AMBER* force field and GB/SA continuum water are FEP 0.28 ± 0.10, PMF 0.20 ± 0.10, MC(JBW)/SD, and 0.25 ± 0.02 kcal/mol in favor of the α anomer, all in good agreement with the experimental value (0.34−0.45 kcal/mol). Convergence, as measured by the stability of the results and the standard deviations, is reached at least an order of magnitude faster with MC(JBW)/SD than with the other two methods.

Introduction

Only a few examples are present in the literature where molecular modeling methods have been used to successfully determine anomeric free energy differences of simple monosaccharides in aqueous solution. This is not surprising since several difficulties are involved in such a computation. First, the ring-opening mechanism involved in the anomeric equilibrium is outside the scope of a standard molecular mechanic treatment. Moreover, within each configurational family (*i.e.*, α or β) the rotational freedom of the primary and secondary hydroxyl groups gives rise to numerous conformations that must be adequately sampled. Finally, the conformational behavior of monosaccharides is governed by a combination of steric and stereoelectronic (*i.e.*, anomeric, *exo*-anomeric, and *gauche*) effects. A careful force field parametrization is therefore needed. Indeed, several carbohydrate force fields have been described[1] but few have been used to calculate anomeric free energies of monosaccharides.[2] These studies typically used traditional, nondirect free energy techniques such as free energy perturbation (FEP),[2a−c] thermodynamic integration (TI),[2b,c] and potential of mean force (PMF) calculations.[2d] These require a series of simulations along a reactant → product mutation pathway (FEP, TI) or a reaction coordinate (PMF). Accurate results are obtained only when each intermediate simulation is converged.[3]

Alternatively, stereoisomeric (*e.g.*, conformational) free energies can be obtained directly from a single simulation provided it can rapidly interconvert stereoisomers to generate a converged equilibrium ensemble. Recently, we have developed such a simulation technique, MC(JBW)/SD,[4] and showed that when coupled with a carefully parametrized force field and an efficient solvent model, it provides converged anomeric free energies in good agreement with experiment, in simulation times totaling a few nanoseconds.[5] Here we compare the performance of MC-(JBW)/SD with FEP and PMF in the calculations of the anomeric free energy difference of D-mannose in water.

Theory

In this section, we briefly outline the basic theory of each of the computational methods used in this work.

Free Energy Perturbation (FEP). The free energy difference between two states, A and B, with Hamiltonians H_A and H_B is given by[6]

$$G_B - G_A = -kT \ln\langle \exp[-(H_B - H_A)/kT]\rangle_A \qquad (1)$$

where $\langle\ \rangle_A$ indicates an ensemble average over state A and T is the simulation temperature. Equation 1 holds only when the two states are similar enough so that all the significant contributions to the average come from regions of the potential surface that are well sampled in both. In practice, this limits successful applications of eq 1 to cases where $\Delta G = \sim kT$. Since the free energy difference in many chemical processes is larger than that (which implies insufficient sampling of important configurations of state B while running the simulation on a potential energy surface defined by state A), such perturbations are usually broken into a series of intermediate states (windows) whose energy differences do not exceed $\sim kT$. The total free energy is then obtained by summing the free energy differences of the intermediate states:[7]

$$\Delta G = \sum_{i=0}^{n} \Delta G(\lambda_{i+1} - \lambda_i) \qquad (2)$$

$$\Delta G(\lambda_{i+1} - \lambda_i) = -kT \ln\langle \exp[-(H_{\lambda i+1} - H_{\lambda i})/kT]\rangle_{\lambda i} \qquad (3)$$

where λ is a continuous coupling parameter such that $\lambda = 0$ describes the initial state and $\lambda = 1$, the final state.

Potential of Mean Force (PMF)**.** When the transformation between two states can be defined by a single reaction coordinate, it is useful to introduce the concept of potential of mean force, $W(\phi)$. $W(\phi)$ is the reversible work needed to take ϕ from some reference value, ϕ_1, to other values along the reaction coordinate:[8]

$$W(\phi) = -kT \ln P(\phi) \qquad (4)$$

where $P(\phi)$ is the probability distribution function of the reaction coordinate. If we assign certain ranges to states A and B, the free energy difference between them can be found by integrating over the probability distribution function:

$$K = \frac{[B]}{[A]} = \frac{\int_{\phi \epsilon B} \exp[-W(\phi)/kT]\, d\phi}{\int_{\Phi \epsilon A} \exp[-W(\phi)/kT]\, d\phi} \qquad (5)$$

[⊗] Abstract published in *Advance ACS Abstracts,* January 1, 1997.

S1089-5647(96)02437-6 CCC: $14.00

It is often difficult to determine the PMF from simple molecular simulations since the reaction coordinate frequently moves through high-energy regions of the potential energy surface and so adequate sampling is hard to obtain. This difficulty may be overcome by introducing an "umbrella" sampling potential, $U(\phi)$, to obtain a uniform sampling along the reaction coordinate. In practice, a series of simulations is run, each constrained, using $U(\phi)$, to successive regions along the reaction coordinate. These biased windows samples are then corrected to remove the effect of $U(\phi)$ and the complete PMF is generated by joining the results of the individual simulations.

MC(JBW)/SD. The free energy difference between two states, A and B, is given by

$$\Delta G = -kT \ln K \tag{6}$$

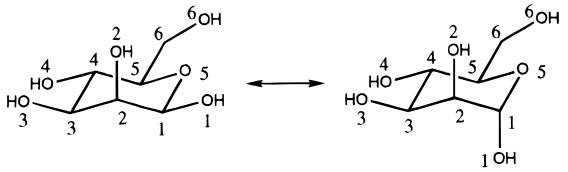
$$K = [\mathrm{B}]/[\mathrm{A}] \tag{7}$$

where K is the population ratio of the two states at equilibrium. K can be obtained directly from a single simulation provided that (1) the simulation frequently interconverts the states available to the molecular system to give a Boltzmann-weighted ensemble and (2) it is possible at each point in the simulation to determine which state the currently simulated structure corresponds to. A short while ago we developed a new simulation technique to achieve this goal.[4] The method, MC(JBW)/SD, generates the canonical ensemble by alternating between stochastic dynamics (SD) and jumping between wells (JBW) steps. JBW is a smart Monte Carlo algorithm which operates by locating the various conformations of a molecular system (*e.g.*, via a conformational search) and subsequently driving a Metropolis Monte Carlo-like simulation to jump repeatedly between them. The hybrid algorithm combines the strengths of SD and Monte Carlo and allows for rapid convergence even for multiconformational or stereoisomeric (*e.g.*, anomeric) systems.[4,5]

Results and Discussion

Molecular simulations were carried out using our recently developed united atom AMBER* force field[5a] and the GB/SA continuum solvent model for water,[9] as implemented in MacroModel 5.5.[10] All simulations were run at 300 K with a 1 fs time step, 8 Å cutoff for VdW interactions, and 20 Å cutoff for electrostatic interactions. In all cases, standard deviations were calculated by the method of block averages. Additional computational details for each particular simulation are given below.

FEP. Free energy perturbations proceeded from β-D-mannose to α-D-mannose (see **1**) using double-wide sampling.



1

Dummy atoms were used as place holders for the OH group and were assigned stretch and bend force constants of 10 kcal/(mol Å) and 1 kcal(mol rad), respectively. Simulations of individual windows were carried out using either SD or the recently described MC/SD[11] simulation technique. Several simulation protocols were attempted with different number of windows and simulation times. These are shown in Table 1 together with the energetic results and relevant timing information. Figure 1 shows a representative example of the free energy profile for the interconversion of β-D-mannose to α-D-mannose.

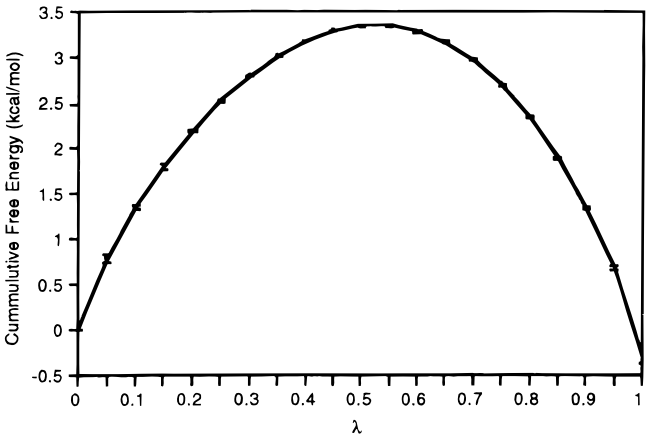


Figure 1. Cumulative free energy profile for the β-D-mannose → α-D-mannose interconversion as obtained from a 20, 500 ps windows, FEP simulation using MC/SD. Error bars are shown at 1σ.

TABLE 1: Anomeric Free Energies (kcal/mol) and Timing Data (CPU Minutes on a R4400 200 MHz SGI INDY Workstation) for the $\beta \rightarrow \alpha$ Interconversion of D-Mannose As Calculated by Free Energy Perturbation

		SD		MC/SD	
windows	ps/window	free energy	time	free energy	time
10	10	0.50 ± 0.56	15	−0.01 ± 0.28	24
	50	0.45 ± 0.40	46	0.32 ± 0.20	73
	100	0.83 ± 0.68	92	0.52 ± 0.38	145
	200	0.45 ± 0.40	169	0.19 ± 0.19	266
	500	0.39 ± 0.28	423	0.31 ± 0.15	665
20	10	0.38 ± 0.90	29	0.12 ± 0.48	46
	50	0.10 ± 0.31	88	0.33 ± 0.34	139
	100	0.07 ± 0.19	176	0.25 ± 0.26	277
	200	0.13 ± 0.19	323	0.23 ± 0.16	508
	500	0.25 ± 0.16	808	0.28 ± 0.10	1270
expt		0.34−0.45		0.34−0.45	

TABLE 2: θ_0 (deg) and Force Constants (kJ/mol) for the 33 PMF Windows along the $\beta \rightarrow \alpha$ Reaction Coordinate of D-Mannose

ϕ	K	ϕ	K	ϕ	K
−45	3000	−10	6000	12	6000
−40	3200	−8	6000	14	6000
−35	3400	−6	6000	16	6000
−30	5600	−4	6000	18	6000
−25	5800	−2	6000	20	6000
−22.5	6000	0	5800	22.5	6000
−20	6000	2	6000	25	5800
−18	6000	4	6000	30	5600
−16	6000	6	6000	35	3400
−14	6000	8	6000	40	3200
−12	6000	10	6000	45	3000

PMF. Following the PMF calculations of van Eijck *et al.* for glucose,[2d] we defined the reaction coordinate for the $\beta \rightarrow \alpha$ interconversion of D-mannose as the improper torsion ϕ = C1−O5−O1−C2 (see **1**). Thus, ϕ was driven in the range of −45° to +45° (ϕ values for the minima are β −37.3°, α 35.3°) according to the protocol presented in Table 2, and its distribution in each simulation was recorded at 1° bins. The explicit form of the "umbrella" potential is given by

$$E = -K(\cos(\theta - \theta_0) - 1) \tag{8}$$

where θ and θ_0 are the actual and constrained torsions, respectively, and K is the force constant. In order to maintain a chair conformation of the pyranose ring along the entire reaction coordinate (in particular near the transition state region),

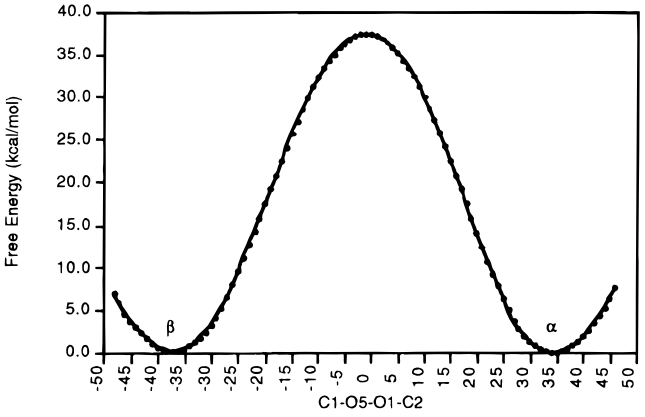


Figure 2. Potential of mean force profile for the β-D-mannose → α-D-mannose interconversion as obtained from a 33, 500 ps windows, PMF simulation using MC/SD.

TABLE 3: Anomeric Free Energies (kcal/mol) and Timing Data (CPU Minutes on a R4400 200 MHz SGI INDY Workstation) for the β → α Interconversion of D-Mannose by Potential of Mean Force Calculations

ps/window	free energy	time
50	0.09 ± 0.30	139
100	0.28 ± 0.15	270
500	0.20 ± 0.10	1370
expt	0.34−0.45	

the C3−C2−C1−O5 torsion was constrained to 60 ± 30° using a flat bottom potential and a force constant of 1000 kJ/mol. This addition was not expected to perturb the equilibrium significantly. In order to obtain good overlap between subsequent simulations and consequently sufficient sampling along the entire reaction coordinate, we found it necessary to use different constraint values for each simulation. These data are gathered in Table 2. The results of the individual simulations were processed using the weighted histogram analysis method (WHAM) of Kollman *et al.*[12] The convergence criterion for iterating the WHAM equations was set to 5×10^{-7}. Three different simulation protocols were attempted with 50, 100, and 500 ps per window. In all cases, individual simulations were run using MC/SD. The free energy difference between the two anomers was calculated from eq 5 where $\phi < 0$ and $\phi > 0$ were taken as the β and α regions, respectively, and the integration was replaced by a summation over the function values at 1° intervals along the reaction coordinate. The results, together with relevant timing information are collected in Table 3. Figure 2 presents an example of the PMF profile for the interconversion of β-D-mannose to α-D-mannose along the defined reaction coordinate.

MC(JBW)/SD. The strategy for calculating anomeric free energies using MC(JBW)/SD has beend described in detail elsewhere and is discussed only briefly here.[5] Our simulation of D-mannose began with a conformational search of the α and β monomers. The lowest 100 structures spanned an energy window of 2.5 kcal/mol and were used as input for a 1 ns MC-(JBW)/SD simulation. The α/β ratios were obtained by monitoring the C2−C1−O5−O1 improper torsion (α −120 ± 30°; β 120 ± 30°; see **1**). Interconversion between different conformers or anomers occurred once every 0.03 ps on average, assuring good convergence of the anomeric ratios. The results of our simulation, together with relevant timing information are summarized in Figure 3.

As the results of Tables 1 and 3 and of Figure 3 indicate, all methods converged to similar (*i.e.*, within 1 standard deviation)

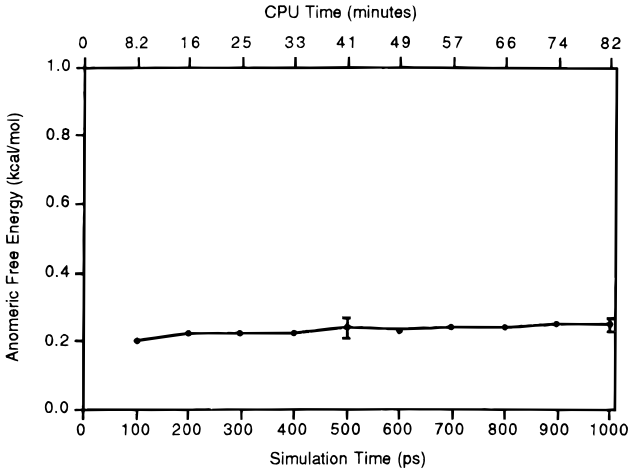


Figure 3. Anomeric free energies (kcal/mol) and timing data (CPU minutes on a R4400 200 MHz SGI Indy workstation) for D-mannose as calculated by MC(JBW)/SD. Error bars are shown at 1σ and were calculated from five blocks by the method of block averages.

anomeric free energies of D-mannose in good agreement with the experimental value of 0.34−0.45 kcal/mol in favor of the α anomer.[13]

Table 1 summarizes the free energy perturbation results for different combinations of number of windows and simulation times. As expected, longer simulations (either with more windows or longer simulation time per window) generally gave results which were closer to the final ones and with smaller standard deviations. Both SD and MC/SD converged to almost identical values (0.25 and 0.28 kcal/mol, respectively). Two general trends are observed. (1) MC/SD simulations converged faster than SD ones. In fact, the last 4, 20 windows MC/SD simulations (50, 100, 200, and 500 ps per window) differ from the final result (0.28 kcal/mol) by not more than 0.05 kcal/mol. The differences between the results of the corresponding SD simulations are much larger. (2) Better results were obtained when the perturbation was broken into more windows rather than when longer simulations per window were attempted. 1 and 2 ns, for example, were better spent on 20, 50 ps (0.33 kcal/mol) and 20, 100 ps (0.25 kcal/mol) windows rather than on 10, 100 ps (0.52 kcal/mol) or 10, 200 ps (0.19 kcal/mol) ones. Thus, for this system convergence within each window occurred quite rapidly (after ∼50 ps). However, the requirements of eq 1 were strictly met only when the perturbation was broken into a rather large number of intermediate steps (*i.e.*, 20). This is in contrast with the results obtained by Kollman *et al.* in deriving the PMF of association for several nonpolar dimers.[3] Both simulation methods required several hundreds of CPU minutes for reasonable convergence, MC/SD being approximately 1.6 computationally more costly (but with smaller standard deviations).

Our PMF results are summarized in Table 3. As expected, the standard deviations decrease with an increase in the simulation time per window. PMF calculations depend on the sampling along the reaction coordinate which is controlled by the umbrella potential (*i.e.*, number of windows and values of constraints). Preliminary attempts with a small number of windows did not lead to a satisfactory overlap between successive simulations along the reaction coordinate and were therefore abandoned. Next, we devised a simulation protocol consisting of 29 windows with varying constraints which lead to an anomeric free energy difference of 0.08 kcal/mol in favor of the α anomer. Sampling along the reaction coordinate was improved by increasing the number of windows (to 33) and modifying the values of the constraints (see Table 2), resulting

in a final value of 0.20 ± 0.10 kcal/mol in favor of the α anomer. Overall, the computational efficiency, as measured by the final results, convergence, and CPU time, is similar to that obtained with free energy perturbation.

The results of our MC(JBW)/SD simulation are presented in Figure 1. The final value, 0.25 ± 0.02 kcal/mol favoring α-mannose, is statistically indistinguishable (*i.e.*, within 1 standard deviation) from the PMF and FEP results. The standard deviation, however, is much smaller than those obtained with the other two methods, reflecting the superior convergence properties of MC(JBW)/SD. In fact, the data in Figure 1 suggest that only 100 ps of simulation time (8 CPU minutes on a R4400 200 MHz SGI INDY workstation) were sufficient to obtain anomeric free energy for D-mannose which differed from the final result by only 0.05 kcal/mol. MC(JBW)/SD is therefore at least an order of magnitude faster to converge than either PMF or free energy perturbation in this system.

Conclusions

In this work we compared three different simulation techniques by calculating the anomeric free energy of D-mannose in water. Though no special effort was made to optimize the performance of any of them, the results clearly suggest that MC(JBW)/SD converges at least an order of magnitude faster than the other two methods. In contrast with PMF calculations or free energy perturbation, MC(JBW)/SD is not limited to cases where a suitable reaction coordinate can be defined or by the computational difficulties arising from large, complicated mutations, and can be equally applied to any multiconfigurational molecular system. Indeed, the method requires an initial set of input conformations. However, with available conformational search methods and for simple molecular systems, these may be obtained at a relatively small computational cost. These results suggest that the MC(JBW)/SD simulation technique should find wide applications in free energy calculations.

Acknowledgment. The authors thank Dr. Quentin McDonald for writing the computer program for the iterative solution of the WHAM equations and for helpful discussions. This work was supported by grants from the National Science Foundation (CHE95 44253) and the Kanagawa Academy of Science and Technology.

References and Notes

(1) (a) Melberg, S.; Rasmussen, K. *J. Mol. Struct.* **1979**, *57*, 215. (b) Melberg, S.; Rasmussen, K. *Carbohyd. Res.* **1979**, *76*, 23. (c) Melberg, S.; Rasmussen, K. *Carbohyd. Res.* **1980**, *78*, 215. (d) Rasmussen, K. *Acta. Chem. Scand.* **1982**, *A36*, 323. (e) Jeffrey, G. A.; Taylor, R. *J. Comput. Chem.* **1980**, *1*, 99. (f) Ha, S. N.; Giammona, A.; Field, M.; Brady, J. W. *Carbohyd. Res.* **1988**, *180*, 207. (g) Homans, S. W. *Biochemistry* **1990**, *29*, 9110. (h) Mardsen, A.; Robson, B.; Thompson, J. S. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84*, 2519. (i) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832. (j) Glennon, T. M.; Zheng, Y. J.; Le Grand, S. M.; Shutzberg, B. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15*, 1019. (k) Grootenhuis, P. D. J.; Haasnoot, C. A. G. *Mol. Simul.* **1993**, *10*, 75. (l) Reiling, S.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem.* **1996**, *17*, 450.

(2) (a) Zheng, Y.-J.; Le Grand, S. M.; Merz, K. M. *J. Comput. Chem.* **1992**, *13*, 772. (b) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553. (c) Schmidt, R. K.; Karplus, M.; Brady, J. W. *J. Am. Chem. Soc.* **1996**, *118*, 541. (d) van Eijck, B. P.; Hooft, R. W. W.; Kroon, J. *J. Phys. Chem.* **1993**, *97*, 12093.

(3) Chipot, C.; Kollman, P. A.; Pearlman, D. A. *J. Comput. Chem.* **1996**, *17*, 1112.

(4) Senderowitz, H.; Guarnieri, F.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *117*, 8211.

(5) (a) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078. (b) Senderowitz, H.; Still, W. C. *J. Org. Chem.*, in press.

(6) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420.

(7) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431.

(8) (a) Paci, E.; Ciccotti, G.; Ferrario, M.; Kapral, R. *Chem. Phys. Lett.* **1991**, *176*, 581. (b) Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. *J. Chem. Phys.* **1992**, *97*, 3639.

(9) (a) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, *112*, 6127. See also: Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745. (b) Qui, D.; Shenkin, P. S.; Hollinger, F.; Still, W. C. *J. Phys. Chem.*, in press.

(10) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(11) Guarnieri, F.; Still, W. C. *J. Comput. Chem.* **1994**, *15*, 1302.

(12) Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M. *J. Comput. Chem.* **1992**, *13*, 1011.

(13) (a) Rudrum, M.; Shaw, D. F. *J. Chem. Soc.* **1965**, 52. (b) Kirby, A. J. *The Anomeric Effect and Related Stereoelectronic Effects on Oxygen*; Springer Verlag: Berlin, 1983; p 7. (c) Stoddart, J. F. *Stereochemistry of Carbohydrates*; Wiley-Interscience: New York, 1971; p 92.