A Force Field for Monosaccharides and (1 → 4) Linked Polysaccharides

TIMOTHY M. GLENNON, YA-JUN ZHENG,[†] SCOTT M. LE GRAND, BRAD A. SHUTZBERG,[‡] and KENNETH M. MERZ, JR.*

Department of Chemistry, 152 Davey Laboratory, Pennsylvania State University, University Park, Pennsylvania 16802

Received 5 October 1993; accepted 18 March 1994

ABSTRACT

A force field for monosaccharides that can be extended to (1 → 4) linked polysaccharides has been developed for the AMBER potential function. The resulting force field is consistent with the existing AMBER force field for proteins and nucleic acids. Modifications to the standard AMBER OH force constant and to the Lennard-Jones parameters were made. Furthermore, a 10–12 nonbonded term was included between the hydroxyl hydrogen of the saccharide and the water oxygen (TIP3P, SPC/E, etc.) to reproduce better the water–saccharide intermolecular distances. STO-3G electrostatic potential (ESP) charges were used to represent the electrostatic interactions between the saccharide and its surrounding environment. To obtain charges for polysaccharides, a scheme was developed to piece together saccharide residues through 1 → 4 connections while still retaining a net neutral charge on the molecule as a whole. Free energy perturbation (FEP) simulations of D-glucose and D-mannose in water were performed to test the resulting force field. The FEP simulations demonstrate that AMBER overestimates intramolecular interaction energies, suggesting that further improvements are needed in this part of the force field. To test further the reliability of the parameters, a molecular dynamics (MD) simulation of α-D-glucose in water was also performed. The MD simulation was able to produce structural and conformational results that are in accord with experimental evidence and previous theoretical results. Finally, a relaxed conformational map of β-maltose was assembled and it was found that the present force field is consistent with available theoretical and experimental results.

*Author to whom all correspondence should be addressed.

[†]Current address: Department of Chemistry, University of California, Berkeley, Berkeley, CA, 94143.

[‡]Participant in the National Science Foundation-Research Experience for Undergraduates program.

Journal of Computational Chemistry, Vol. 15, No. 9, 1019–1040 (1994)

CCC 0192-8651 / 94 / 091019-22

Introduction

The importance of carbohydrates in numerous biological contexts has stimulated much interest in their structural and conformational properties.[1-7] Over the years, considerable experimental progress into the structural analysis of carbohydrates using X-ray techniques has made possible the complete description of a large number of carbohydrates.[7] Furthermore, modern nuclear magnetic resonance (NMR) techniques are now being used to determine the three-dimensional structure of polysaccharides in solution.[8] These advances, coupled with advances in the synthesis of carbohydrates, have led researchers to realize the importance of molecular conformation in determining the biological functions of these compounds.

Modeling complex molecules has been useful in the study of the structure, function, and dynamics of a wide range of biomolecules.[9] The majority of these studies have been carried out using force-field methods due to the size of biomolecular systems, which precludes the use of more sophisticated quantum mechanical methods. Most of the focus has been on proteins and nucleic acids,[9] with more modest efforts being directed at the carbohydrate area.[10-27] The increasing attention given to carbohydrate chemistry, however, has produced the need for an extendible and accurate force field for free energy simulations as well as structural studies. Moreover, it is important to develop a force field for carbohydrates that can take advantage of existing force fields for proteins and nucleic acids to study mixed carbohydrate/protein systems.

In this article we present the development of a force field for carbohydrates that is compatible with the existing AMBER force field.[28,29] The development of strategies for a carbohydrate force field is difficult because carbohydrate structure can be complex, but it is this complexity that gives carbohydrates the ability to carry substantial biological information.[4] The main difficulty in accomplishing this goal has been the determination of partial charges for carbohydrates. This can be done by using one set of charges that are transferable to all cases, or it can be done on an individual basis. We have chosen the latter strategy because it allows us to be compatible with the existing AMBER force-field representation for nucleic acids and proteins and we feel it is an accurate representation. The former point is germane when one wants to study glycoproteins and other mixed polymer systems. Others have chosen to use the transferable charge strategy in their carbohydrate force field development and have had success with this representation.[19,22]

This article is divided into two main sections. First we describe the force-field development, and then we describe several tests that were carried out to explore the accuracy of our approach in terms of the available structural and thermodynamic information on carbohydrates.

Computational Procedure

THE POTENTIAL FUNCTION

The AMBER potential function was used throughout and has the following form[28,29]:

$$
\begin{aligned}
E_{\text{total}} = & \sum_{\text{bonds}} \frac{K_r}{2}(r - r_{\text{eq}}) + \sum_{\text{angles}} \frac{K_\theta}{2}(\theta - \theta_{\text{eq}}) \\
& + \sum_{\text{dihedrals}} \sum_{n} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] \\
& + \sum_{i<j} \varepsilon_{ij}\left[\left(\frac{R^*_{ij}}{R_{ij}}\right)^{12} - \left(\frac{R^*_{ij}}{R_{ij}}\right)^{6}\right] \\
& + \frac{1}{VDW_{\text{scale}}} \sum_{\substack{i<j \\ \text{1-4 terms}}} \varepsilon_{ij}\left[\left(\frac{R^*_{ij}}{R_{ij}}\right)^{12} - \left(\frac{R^*_{ij}}{R_{ij}}\right)^{6}\right] \\
& + \sum_{h\text{-bonds}} \left[\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}}\right] \\
& + \sum_{i<j} \frac{q_i q_j}{\varepsilon R_{ij}} + \frac{1}{EE_{\text{scale}}} \sum_{\substack{i<j \\ \text{1-4 terms}}} \frac{q_i q_j}{\varepsilon R_{ij}}
\end{aligned}
$$

The first three terms represent the bonded interactions present in a molecule (namely, the bond, angle, and torsion interactions). For the bond stretching and angle bending motions, Hooke's law is used, which is represented by a quadratic potential; the torsional interactions are represented by a truncated Fourier series. K_r and K_θ are the force constants for the bond and angle, respectively, whereas r_{eq} and θ_{eq} are the equilibrium bond length and angle, respectively. V_n, n, ϕ, and γ represent the torsional barrier, the periodicity, the calculated dihedral angle, and, finally, the phase.

The next five terms represent the nonbonded interactions in a molecule and are the Lennard-Jones [the $(R_{ij})^{-12}$, $(R_{ij})^{-6}$, or 6–12 terms], the hydrogen bond (the 10–12 terms), and the electrostatic interaction terms. R_{ij} is the distance between atoms i and j, R^*_{ij} and ε_{ij} are parameters that define the shape of the Lennard-Jones potential for the interaction between atoms i and j, C_{ij} and D_{ij} define the shape of the hydrogen bond potential, q_i and q_j are the atomic point charges for atoms i and j, and ε is the dielectric constant. $1/VDW_{\text{scale}}$ and $1/EE_{\text{scale}}$ are scaling factors for the 1–4 van der Waals and electrostatic terms, respectively. These are set to 1/2 in AMBER.

FORCE FIELD DEVELOPMENT

The initial intramolecular parameters were obtained directly from AMBER with the exception of the OS—CT—OH parameter, which was taken from Ha et al.[19] The nonbonded parameters were also taken initially from AMBER, and the atomic point charges were obtained using STO-3G ESP[30-32] calculations using a modified version of Gaussian 88.[33] This was done to be consistent with the AMBER force field for proteins and nucleic acids, which also uses STO-3G ESP charges.[28,29] The final parameter set is given in Table I, and the final charge sets for two monosaccharides are discussed later. In the next sections we shall discuss how the necessary parameters were determined and tested for both the inter- and intramolecular terms.

Intramolecular Terms

To test the bond, angle, and dihedral parameters, normal mode analysis was carried out on α-D-glucose. These calculations all used STO-3G ESP charges and the final Lennard-Jones parameters described below. Table II summarizes the results. Included for comparison are results from a valence force field and other force-field studies. Only part of the spectrum for α-D-glucose has been resolved experimentally. Furthermore, six O—H stretching frequencies were identified, indicating the presence of coupling between the vibrations. The agreement between the valence force field (which was parameterized to reproduce experimental information) and Ha et al. is good, in the range of about ±10%. This is also true for the T1 and PEF force fields except in the OH stretching region, where the error is large when compared to the valence force field. Clearly, the AM-

TABLE I.
Final Force Field Parameters.[a]

Bond Stretches Bond Type	K_r (kcal/mol · Å)	r_o (Å)
CT—CT	310.0	1.526
CT—HC	331.0	1.090
CT—OH	320.0	1.410
CT—OS	320.0	1.410
HO—OH	460.5	0.960

Bond Angles Angle Type	K_θ	θ_o
OS—CT—OH	92.6	111.55
CT—CT—CT	40.0	109.50
CT—CT—HC	35.0	109.50
CT—CT—OH	50.0	109.50
CT—CT—OS	50.0	109.50
HC—CT—HC	35.0	109.50
HC—CT—OH	35.0	109.50
HC—CT—OS	35.0	109.50
CT—OH—HO	55.0	108.50
CT—OS—CT	60.0	109.50

Torsion Torsion Type	V_n	n	γ
X—CT—CT—X	1.30	3	0.0
X—CT—OH—X	0.50	3	0.0
X—CT—OS—X	1.15	3	0.0
OS—CT—CT—OH	0.14	1	0.0
	0.50	2	0.0
OS—CT—CT—OH	0.14	3	0.0
	0.50	2	0.0
OH—CT—CT—OH	0.14	3	0.0
	0.50	2	0.0
CT—CT—OS—CT	0.38	3	0.0
	0.20	2	180.0

H-bond H-bond Type	C	D
HO OH	7557.0	2385.0
HO OS	7557.0	2385.0

van der Waals Atom Type	R^*	ε
HO	1.000	0.020
HC	1.540	0.010
OH	1.550	0.100
OS	1.650	0.150
CT	1.800	0.060

[a]The atom labels are: CT—tetrahedral carbon; OS—ether oxygen; OH—hydroxyl oxygen; HO—hydroxyl hydrogen; HC—alkyl hydrogen; X—any atom.

TABLE II.
Summary of Normal Mode Analysis Results.

Mode	VF[a]	Ha[b]	AM1[c]	T1[d]	T2[e]	PEF[f]
1	59	54	60	72	72	65
2	79	74	88	101	101	97
3	133	78	94	120	120	113
4	186	107	151	138	138	150
5	139	139	228	193	193	212
6	223	211	247	232	232	243
7	224	226	272	242	242	245
8	229	245	303	248	248	262
9	229	259	309	258	258	270
10	236	270	318	270	270	297
11	240	279	333	292	293	313
12	286	329	341	311	311	322
13	301	341	392	316	316	345
14	328	346	406	321	321	352
15	348	369	417	352	352	377
16	383	393	424	371	372	391
17	424	403	447	436	436	417
18	472	430	500	440	441	438
19	494	470	559	458	458	491
20	553	518	587	480	480	514
21	558	576	619	499	499	537
22	610	580	681	529	529	550
23	725	649	811	606	606	582
24	894	801	935	761	760	830
25	922	842	1024	824	824	856
26	940	903	1123	904	903	925
27	993	917	1155	919	919	953
28	1008	941	1170	929	928	977
29	1034	957	1184	954	954	1008
30	1060	965	1197	966	966	1036
31	1074	1007	1213	970	970	1068
32	1085	1017	1231	978	979	1071
33	1111	1021	1241	1010	1009	1123
34	1124	1061	1249	1019	1919	1141
35	1143	1068	1266	1024	1024	1166
36	1169	1096	1291	1053	1053	1194
37	1181	1200	1302	1093	1093	1207
38	1211	1205	1314	1119	1118	1235
39	1224	1210	1323	1176	1176	1252
40	1243	1235	1336	1179	1178	1264
41	1253	1245	1349	1208	1203	1279
42	1258	1264	1353	1235	1227	1293
43	1271	1270	1375	1252	1246	1305
44	1276	1287	1386	1269	1263	1317
45	1288	1298	1395	1286	1273	1339
46	1306	1329	1413	1296	1291	1369
47	1317	1345	1423	1312	1302	1407
48	1326	1397	1431	1364	1357	1459
49	1331	1405	1447	1384	1382	1486
50	1336	1426	1532	1401	1400	1496
51	1344	1434	1541	1433	1431	1510
52	1357	1449	1544	1470	1466	1524
53	1398	1489	1553	1474	1472	1545
54	1466	1496	1560	1506	1504	1563
55	2884	2897	3026	2870	2870	2904
56	2929	2931	3035	2903	2903	2931
57	2933	2934	3046	2905	2905	2938
58	2937	2937	3055	2907	2907	2945
59	2939	2939	3062	2910	2910	2951
60	2943	2943	3066	2913	1913	2961
61	2983	2969	3106	2940	2940	2972
62	3397	3384	3457	3709	3385	3655
63	3397	3386	3461	3710	3385	3663
64	3397	3387	3465	3711	3386	3663
65	3397	3390	3466	3714	3390	3663
66	3397	3391	3481	3716	3392	3663

TABLE II.
(Continued)

[a]Valance force field results, see P. D. Vasko, J. Blackwell, and J. L. Koenig, *Carbohydr. Res.*, **19**, 297 (1971). P. D. Vasko, J. Blackwell, and J. L. Koenig, *Carbohydr. Res.*, **23**, 407 (1972).
[b]See S. N. Ha, A. Giammona, M. Field, and J. W. Brady, *Carbohydr. Res.*, **180**, 207 (1988).
[c]Semiempirical molecular orbital calculations, see M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).
[d]AMBER without modification using STO-3G ESP charges.
[e]AMBER with modified HO — OH bond stretching force constant (K = 460.5 kcal / mol · Å) using STO-3G ESP charges.
[f]PEF 424 results. See K. Rasmussen, *Acta. Chem. Scand.*, **A36**, 323 (1982).

BER parameter for the OH bond stretch is too large (553 kcal/mol · Å^2) and, therefore, the calculated stretching frequencies are too high. Hydrogen bonding could reduce this value; however, it has been shown experimentally that the formation of an OH ··· O type hydrogen bond only reduces the OH stretching by ~ 200 cm^{-1}.[34] Apparently, the AMBER force constant for OH is too stiff for carbohydrates; we therefore used a smaller force constant (460.5 kcal/mol · Å^2), which resulted in the model labeled T2 and gave better agreement with experiment.

In a previous article we extensively studied the torsion motions of carbohydrates.[35] We found that STO-3G charges as well as the standard AMBER parameters were able to reproduce the observed conformational preferences. Hence, we did not modify the AMBER torsion parameters when we used STO-3G ESP charges. For a further verification of this observation, we analyzed the rotational barrier in methanol using several different methods. These results are shown in Table III. It is clear that AMBER (with STO-3G ESP charges for methanol and the Lennard-Jones parameters de-

TABLE III.
Calculated and Experimental Rotational Barrier for Methanol.

Model	Barrier (kcal / mol)
STO-3G // STO-3G	2.0[a]
3-21G // 3-21G	1.5[a]
3-21G* // 3-21G*	1.5[a]
6-31G* // 6-31G*	1.4[a]
MP2 / 6-31G* // MP2 / 6-31G*	1.5[a]
MP2 / 6-311G** // MP2 / 6-311G**	1.35[b]
AMBER	1.04[b]
Exp.	1.1[a]

[a] See W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley-Interscience, New York, 1986, and the references therein.
[b] This work.

TABLE IV.
The Calculated Results for Methanol – Water ($CH_3OH - OH_2$) Interactions.

Model	ΔE (kcal / mol)	r(OO) (Å)	r(OH) (Å)
STO-3G // STO-3G	6.28	2.708	1.716
3-21G // 3-21G	11.08	2.772	1.800
6-31G* // 6-31G*	5.59	2.961	2.014
6-31G** // 6-31G**	5.52	2.970	2.026
6-311G** // 6-311G**	5.65	2.958	2.015
MP2 / 6-31G** // MP2 / 6-31G**	7.30	2.893	1.933
MP4 / 6-31G** // MP2 / 6-31G**	7.10	2.893	1.933
MP2 / 6-311G** // MP2 / 6-31G**	7.70	2.851	1.888

scribed below) reproduce the methanol profile well. This observation is in line with what we observed previously.[35]

Lennard-Jones Parameters

STO-3G ESP charges perform well with regard to torsion potentials, but they result in energies and geometries for water–methanol and water–dimethyl ether interactions that are not in good agreement with high-level *ab initio* calculations.[36] To alleviate this we have modified the Lennard-Jones (L-J) parameters within AMBER to bring them into better accord with the *ab initio* data given in Tables IV and V. All of the *ab initio* results reported in Tables IV and V were obtained using Gaussian 92[37] and Gaussian 88.[33]

The results of our *ab initio* calculations follow the trend that as we increase the basis set size the interaction energy decreases, and upon inclusion of correlation the interaction energy increases slightly. It has been pointed out that for the water dimer the 6-31G* basis set gives a good representation of the experimental interaction energy, whereas MP2/6-31G* is too negative.[38] It is not certain whether this trend holds up in other cases, so we have focused on fitting our L-J parameters in such a way as to reproduce an interaction energy in the range defined by the 6-31G* and MP2/6-31G* values.

We have found that the standard AMBER L-J parameters for the ether oxygen atom (parameter label OS) can satisfactorily reproduce the data given in Table V, but this is not so for the methanol case (see Tables IV and VI). Furthermore, our tests indicate that the L-J parameters for hydrogen and carbon have little effect on the interaction energy and the interatomic O ··· O distance for the methanol case. In light of this, we have focused on modifying the L-J parameters for only the hydroxyl oxygen atom (parameter label OH). According to our tests, changes in these L-J parameters can reproduce the *ab initio* interaction energies well (e.g., $R^* = 1.65$ Å, $\varepsilon = 0.05$) for water/methanol interactions; however, the interatomic distances are found to be too short (see Table VI). To alleviate this problem, we have found it necessary to add a 10–12 nonbonded interaction between the hydroxyl hydrogen on the saccharide and the water oxygen (parameter labels of HO and OW, respectively). This was also done in the original AMBER parameterization; however, the hydrogen bond term was only used for hydrogen bond interactions between amino acid residues within a protein. For hydrogen bonds between solvent and protein, this term was not applied because the original AMBER was designed primarily to be

TABLE V.
Calculated Results for Dimethyl Ether – Water Complex at Different Levels of Approximation.

Model	ΔE (kcal / mol)	r(OH) (Å)	r(OO) (Å)
STO-3G // STO-3G	4.76	1.794	2.782
3-21G // 3-21G	11.06	1.814	2.722
6-31G* // 6-31G*	5.31	2.020	2.946
MP2 / 6-31G* // MP2 / 6-31G*	7.67	1.908	2.880

used for gas-phase simulations using a distant-dependent dielectric. In the present case, because we are more interested in condensed phase simulations, we find that we need to use this term to acquire a good representation of sugar hydroxyl/water hydrogen bonding interactions. We have found that using the AMBER hydrogen bond, 10–12 parameters for the hydroxyl hydrogen/hydroxyl oxygen (HO—OH atom types[28,29]) interaction worked the best (C = 7557, D = 2385). Table VII gives a sample of the test results obtained by holding the hydrogen bond parameter constant while varying the L-J parameters. Clearly, the hydrogen bond parameter is important in reproducing the intermolecular distance. We feel that the best compromise parameter set is $R^* = 1.55$ and $\varepsilon = 0.10$, with the 10–12 parameters described earlier (compare Tables VI and VII).

Recently, Ferguson and Kollman[39] argued that the Lennard-Jones 6–12 function can replace the 10–12 form in force fields. This can be done easily if the interaction energy is too negative to start with, but we have found it to be impossible when the starting interaction energy is too positive (as is the case when STO-3G ESP charges are used). We have found in tests that when 6-31G* ESP charges are used the interaction energy is too negative using standard AMBER parameters. In this case it is much easier to only use the L-J parameter and, indeed, in our tests we find that with 6-31G* ESP charges we can get good interaction energies and hydrogen bond distances using only the L-J potential (data not shown).

TABLE VI.
Variation of Methanol – Water Interaction Energy and Hydrogen Bond Distances with the Lennard-Jones Parameter for OH.[a]

R^*	ε	ΔE (kcal / mol)	r(OH) (Å)	r(OO) (Å)
1.65	0.01	8.60	1.33	2.32
1.65	0.02	7.50	1.43	2.41
1.65	0.03	6.96	1.49	2.47
1.65	0.05	6.40	1.56	2.54
1.65	0.10	5.70	1.66	2.64
1.65[b]	0.15	5.40	1.72	2.69
1.65	0.20	5.20	1.76	2.73
1.65	0.30	4.95	1.81	2.79
1.55	0.10	6.70	1.55	2.53
1.55	0.13	6.35	1.59	2.56
1.55	0.15	6.20	1.61	2.58
1.55	0.20	6.07	1.64	2.62
1.55	0.30	5.70	1.70	2.67
1.55	0.35	5.60	1.72	2.69

[a]The TIP3P water model was used throughout. The SPC / E water model gave similar results.
[b]The original AMBER L-J values.

TABLE VII.
Variation of Methanol – Water Interaction Energy and Hydrogen Bond Distances, Using the Modified Hydrogen Bond Parameter (C = 7557, D = 2385), with the Lennard-Jones Parameter for OH.

R^*	ε	ΔE (kcal / mol)	r(OH) (Å)	r(OO) (Å)
1.65	0.01	6.10	1.80	2.76
1.65	0.02	6.00	1.81	2.77
1.65	0.03	5.90	1.81	2.78
1.65	0.30	5.35	1.88	2.85
1.55	0.01	6.23	1.79	2.75
1.55	0.10	6.00	1.82	2.79
1.55	0.15	5.90	1.83	2.79
1.55	0.30	5.80	1.85	2.82

Electrostatics of Monosaccharides

The STO-3G ESP charges for several common hexose monosaccharides (allose, altrose, galactose, glucose, gulose, idose, mannose, and talose) have been calculated and are available as supplementary material.* The crystal structure of α-D-glucose[40] was used as a template for each sugar structure. For each saccharide standard AMBER parameters were used to position the hydroxyl groups that differed in conformation from α-D-glucose. The charge sets for the α and β anomers for glucose and mannose are provided as examples in Table VIII. (Scheme 1 indicates the atom labeling system used.) The most interesting trend in these results is the observation that there is a significant difference between the charge distribution of an α anomer and a β anomer.[41] Hence, charge models that use the same point charges for the α and β anomers may not be accurately reproducing the electrostatic properties of these molecules.

Electrostatics of Polysaccharides

The charges described earlier are for monosaccharides only. Extension of these charge models to handle polysaccharides is straightforward. The

*For the charges of the terminal and non-terminal residues for the eight saccharides listed (allose, altrose, galactose, glucose, galactose, idose, mannose, and talose) and the ESP fitting program, please write the author.

quantum mechanical ESPs can be readily fit to a charge model in which several of the charges have been fixed. For example, for terminal residues (e.g., monosaccharides missing a terminal OH) we can fix the charges for one of the OH groups at the STO-3G ESP charges for the OH radical. We then optimize all of the remaining charge centers to obtain a charge set for the terminal residue. The key point in this is retaining a net neutral charge on the terminal saccharide residue. By fixing the OH charges to that of the OH radical (which is net neutral), we ensure that the fit charges are also neutral. Charges arrived at in this manner are given for glucose and mannose in Table IX for the $1 \rightarrow 4$ linkage case only. This same procedure could be used for the cases in which only the hydroxyl hydrogen is fixed, but this is a bit more problematic because there is only one atom whose charge we need to fix. To accomplish this we can set the charge of the terminal H to 0.0 and then optimize all of the remaining charges. Using this procedure, we did not obtain satisfactory results because setting the hydrogen charge to 0.0 when it is usually about +0.3 results in a large perturbation in the overall charge distribution. To resolve this problem, we have fit ESPs for the correspond carbohydrate radical. In this method we remove the hydroxyl hydrogen in question and then do the ESP fit at the UHF/STO-3G level. The charges we obtain in this manner have similar magnitudes to those obtained from the closed-shell ESP calculations. Example charge sets are given in Table X.

TABLE VIII.
The Calculated STO-3G ESP Charges for D-Glucose and D-Mannose.

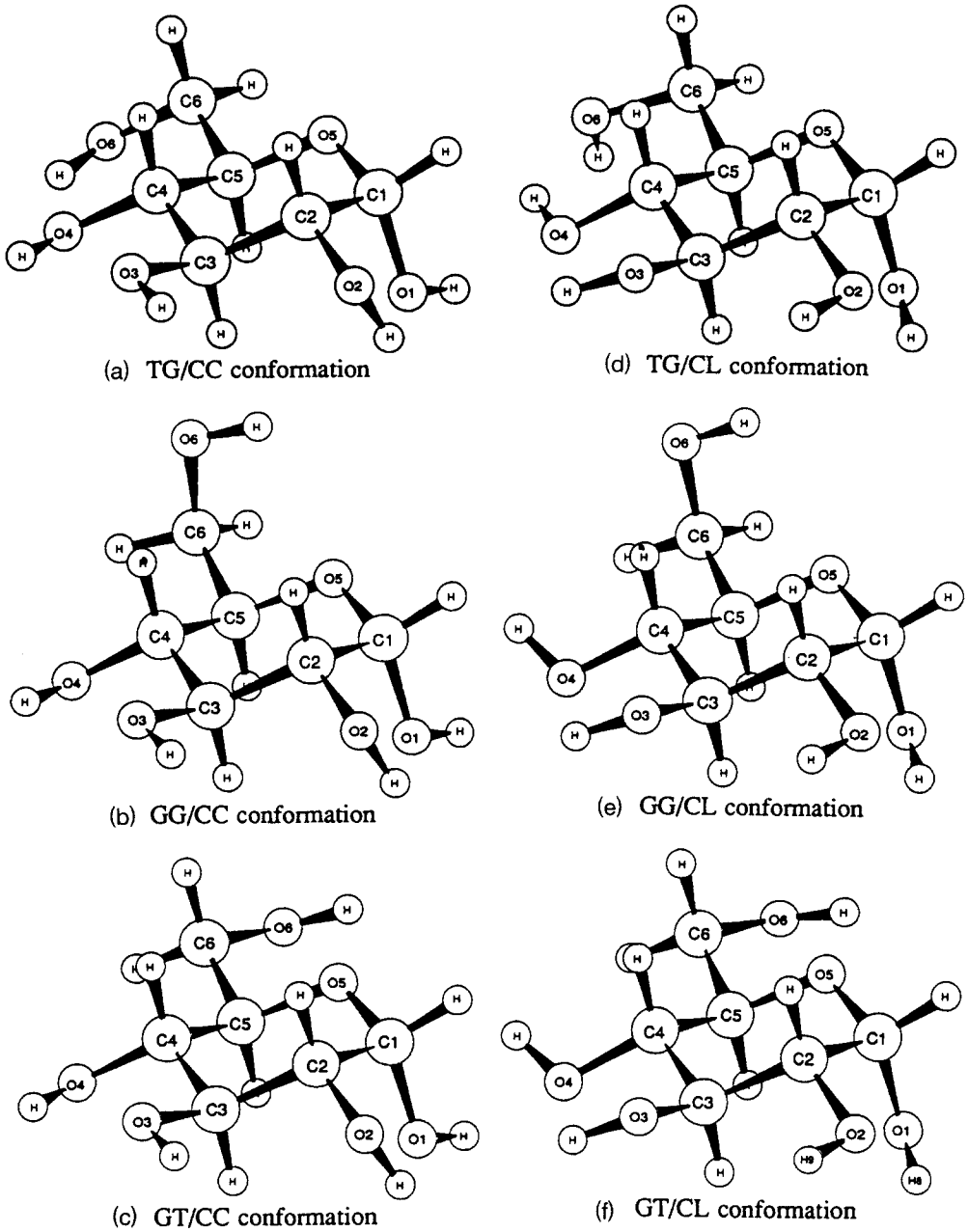
	Glucose		Mannose	
	α	β	α	β
O_5(ring)	−0.465	−0.342	−0.374	−0.307
O_1	−0.546	−0.516	−0.561	−0.515
O_2	−0.496	−0.539	−0.519	−0.489
O_3	−0.539	−0.536	−0.510	−0.499
O_4	−0.497	−0.520	−0.512	−0.500
O_6	−0.518	−0.595	−0.513	−0.512
C_1	0.419	0.355	0.360	0.372
C_2	0.051	0.141	0.004	−0.072
C_3	0.458	0.189	0.287	0.240
C_4	−0.156	0.209	0.110	0.167
C_5	0.287	−0.029	0.128	0.025
C_6	0.190	0.382	0.206	0.212
HC_1	0.023	0.030	0.043	0.021
HC_2	0.052	0.073	0.079	0.101
HC_3	0.004	0.034	0.019	0.009
HC_4	0.092	0.036	0.066	0.057
HC_5	0.045	0.032	0.078	0.078
HC_{6a}	−0.006	−0.016	−0.006	0.001
HC_{6b}	0.012	−0.029	0.009	0.015
HO_1	0.334	0.323	0.343	0.334
HO_2	0.302	0.315	0.338	0.344
HO_3	0.306	0.346	0.308	0.310
HO_4	0.330	0.317	0.304	0.295
HO_5	0.318	0.340	0.313	0.313

To obtain charges for the nonterminal saccharide residues, we adopted the procedure used for the hydroxyl case. To illustrate this we will consider 1–4 linkages again. In this case we need to eliminate the elements of water, so we determine the STO-3G ESP charges for water and use these values to fix the H and OH charges in our fitting procedure. Charges obtained in this way for glucose and mannose are given in Table XI. In all cases we end up with neutral charge sets and a library of charges that allow us to build up any polysaccharide based on $1 \rightarrow 4$ linkages.

These procedures might cause some of the optimized charges to have unusually large or small values. However, we do not find this to be the case when we examined the fit of the optimized charges to those obtained in an unrestrained optimization. We find that there is a good correlation between these two charge sets. Furthermore, this can be confirmed by examining the charges given in Tables IX through XI. Other linkages can be treated in the same manner and, indeed, there are many possible linkage patterns possible. However, we have not attempted to derive charge models for all of these. If anyone is interested in other linkages, we will supply the electrostatic potentials and the fitting program to do so.

Results and discussion

The final parameters are given in Table I. Our carbohydrate force field is compatible with the existing AMBER force field for proteins and nucleic acids and as such is an extension of this force field. To test the efficacy of these parameters, we have carried out several tests. We first calculated the free energy differences between two anomers for D-glucose and D-mannose in water. Next, a molecular dynamics simulation of α-D-glucose in water was performed to study the conformational changes, extent of hydrogen bonding, and the structure of the surrounding solvent for this simple sugar. For convenience, from now on AMBER refers to our modified force field unless otherwise



SCHEME 1. GG = gauche-gauche; GT = gauche-trans; TG = trans-gauche; CL = clockwise; CC = counter-clockwise.

specified. We will also call the force field reported by Ha et al.[19] the CHARMm force field.

FREE ENERGY PERTURBATION

The free energy difference between carbohydrate anomers is thought to be a result of differences in intramolecular (e.g., anomeric effect) and intermolecular interactions (e.g., solvation) between the two anomers. This has been suggested by both experimental[42,43] and theoretical investigations.[20,21] Computational studies of carbohydrates using molecular mechanics have thus far proven to be reasonably capable in reproducing free energy differences between saccharide anomers.[21,44] From simulations of this sort, Ha et al.[21] suggested that "solvation is a dominant factor favoring the β anomer of D-glucose." This conclusion was reached even though there was also a large intramolecular contribution observed that favored the α anomer. Experimentally it has been suggested that solvation effects are important for

TABLE IX.
STO-3G ESP Charges for Terminal Residues of D-Glucose and D-Mannose.[a]

	Glucose				Mannose			
	O_1 Net Neutral		O_4 Net Neutral		O_1 Net Neutral		O_4 Net Neutral	
	α	β	α	β	α	β	α	β
O_5(ring)	−0.291	−0.202	−0.409	−0.316	−0.205	−0.101	−0.316	−0.250
O_1	−0.268	−0.268	−0.533	−0.525	−0.268	−0.268	−0.550	−0.522
O_2	−0.421	−0.467	−0.516	−0.544	−0.444	−0.421	−0.512	−0.483
O_3	−0.521	−0.541	−0.450	−0.464	−0.498	−0.524	−0.440	−0.421
O_4	−0.481	−0.538	−0.268	−0.268	−0.490	−0.506	−0.268	−0.268
O_6	−0.513	−0.597	−0.528	−0.562	−0.506	−0.508	−0.524	−0.525
C_1	−0.214	−0.332	0.426	0.446	−0.300	−0.416	0.368	0.355
C_2	−0.104	0.101	0.148	0.134	−0.119	0.137	0.077	0.050
C_3	0.443	0.249	0.224	0.162	0.281	0.113	0.184	0.081
C_4	−0.148	0.245	−0.617	−0.478	0.121	0.250	−0.557	−0.462
C_5	0.311	0.059	0.175	−0.112	0.152	−0.024	0.037	−0.059
C_6	0.195	0.323	0.266	0.491	0.228	0.194	0.271	0.293
HC_1	0.229	0.264	0.012	0.012	0.260	0.244	0.036	0.030
HC_2	0.153	0.115	0.052	0.087	0.170	0.108	0.066	0.079
HC_3	0.021	0.031	0.105	0.105	0.018	0.047	0.099	0.105
HC_4	0.099	0.027	0.285	0.281	0.072	0.042	0.310	0.286
HC_5	0.019	−0.002	0.112	0.109	0.050	0.096	0.164	0.157
HC_{6a}	−0.002	−0.006	−0.010	−0.041	−0.011	0.010	−0.007	−0.010
HC_{6b}	−0.005	−0.013	0.011	−0.062	−0.010	0.012	0.018	0.018
HO_1	0.268	0.268	0.329	0.314	0.268	0.268	0.336	0.336
HO_2	0.292	0.293	0.312	0.315	0.328	0.309	0.334	0.336
HO_3	0.307	0.336	0.293	0.327	0.310	0.332	0.297	0.298
HO_4	0.320	0.316	0.268	0.268	0.291	0.296	0.268	0.268
HO_6	0.311	0.339	0.313	0.321	0.302	0.310	0.309	0.309

[a] O_1 and O_4 hydroxyls are fixed at the STO-3G ESP charge values of the hydroxide radical.

some sugars (e.g., glucose), whereas for others it is less important (e.g., altrose).[42]

Interestingly, the conclusion that solvation effects are important in determining the anomeric preferences of sugars is based on the observation that force-field models predict that the intramolecular energy difference is small and, therefore, a solvation term is needed.[42,45] Given the inherent difficulty in parameterizing force-field models for the complex intramolecular interactions present in carbohydrates, one has to wonder whether this conclusion is warranted. Indeed, Cramer and Truhlar have recently concluded, based on *ab initio* and semiempirical calculations, that there is "no solvation effect on the β—D → α—D free energy difference."[46] Herein, we have used molecular dynamics free energy perturbation (MD-FEP) simulations to further study this issue using our force-field model.

MD-FEP simulations[9,47-54] were carried out to test the reliability of the force-field parameters in reproducing the experimental free energy difference between the α and β anomers of glucose and mannose. The MD-FEP simulations were carried out using the slow growth method. The SPC/E[55] water model was used, and a timestep of 1.5 fs was employed in conjunction with SHAKE.[56] Periodic boundary conditions were used, and the temperature and pressure were maintained at 298.15 K and 1.0 atm, respectively. The nonbonded pairlist had a cutoff of 9 Å and was updated every 10 timesteps; and a constant dielectric of 1 was used throughout. The molecule was first solvated and was then subjected to 1000 steps of energy minimization, followed by an equilibration period of 45 ps. The simulation time period was 180 ps in each direction, and the simulations were carried out in the forward ($\lambda = 1 \rightarrow 0$) and backward ($\lambda = 0 \rightarrow 1$) directions for a total of 360 ps of simulation time.

Tests were performed on both D-glucose and D-mannose for which experimental values are available.[12,57-59] It has been shown experimentally that the β-anomer is favored for D-glucose (by 0.33 kcal/mol) whereas the α-anomer is favored in the case of D-mannose (by 0.40 kcal/mol at 20°C) in

TABLE X.
STO-3G ESP Charges for the Radicals of D-Glucose and D-Mannose.

	Glucose				Mannose			
	O_1 Radical		O_4 Radical		O_1 Radical		O_4 Radical	
	α	β	α	β	α	β	α	β
O_5(ring)	−0.512	−0.525	−0.459	−0.372	−0.401	−0.353	−0.373	−0.318
O_1	−0.242	−0.230	−0.571	−0.517	−0.240	−0.196	−0.573	−0.512
O_2	−0.476	−0.450	−0.481	−0.483	−0.502	−0.453	−0.521	−0.481
O_3	−0.514	−0.546	−0.532	−0.529	−0.499	−0.508	−0.503	−0.496
O_4	−0.499	−0.536	−0.225	−0.212	−0.510	−0.510	−0.247	−0.242
O_6	−0.523	−0.550	−0.542	−0.500	−0.513	−0.513	−0.506	−0.511
C_1	0.559	0.583	0.416	0.404	0.471	0.478	0.369	−0.384
C_2	−0.126	−0.123	0.051	0.025	−0.218	−0.223	−0.036	−0.153
C_3	0.507	0.517	0.332	0.361	0.418	0.367	0.244	0.233
C_4	−0.182	−0.054	0.137	0.120	0.022	0.122	0.261	0.270
C_5	0.350	0.201	0.075	0.081	0.206	0.078	0.096	0.051
C_6	0.207	0.265	0.267	0.169	0.175	0.197	0.190	0.190
HC_1	0.012	0.003	0.025	0.030	0.043	−0.009	0.045	0.020
HC_2	0.098	0.087	0.057	0.060	0.114	0.116	0.096	0.128
HC_3	0.004	−0.015	0.026	0.005	0.003	−0.015	0.036	0.022
HC_4	0.087	0.072	0.052	0.019	0.083	0.057	0.044	0.038
HC_5	0.010	0.027	0.059	0.045	0.049	0.063	0.064	0.051
HC_{6a}	−0.013	0.006	0.008	0.025	0.004	0.005	−0.003	0.006
HC_{6b}	0.006	−0.015	−0.007	0.024	0.013	0.020	0.015	0.020
HO_1	—	—	0.360	0.322	—	—	0.350	0.335
HO_2	0.301	0.283	0.311	0.302	0.358	0.353	0.345	0.350
HO_3	0.293	0.314	0.306	0.306	0.295	0.306	0.296	0.302
HO_4	0.335	0.338	—	—	0.312	0.304	—	—
HO_6	0.318	0.348	0.335	0.315	0.317	0.314	0.311	0.313

aqueous solution. Two sets of parameters were tested. One has the H-bond terms, whereas the second does not. For comparison, we have also used the Ha et al. force field in some of our simulations.[19]

The calculated free energy difference between α and β-D-glucose using AMBER without (run **1a**) and with (run **2a**) 10–12 terms are calculated to be -4.44 ± 0.08 kcal/mol and -3.03 ± 0.04 kcal/mol, respectively. The results are summarized in Table XII, and all free energy values are reported for the conversion of the β-anomer into the α one. Both of these values are too large and are in the incorrect direction when compared to experiment ($\Delta G_{\beta-\alpha} = 0.33$ kcal/mol). To garner further insight into this problem, we have carried out FEP simulations (runs **1b** and **2b**) using AMBER, in which we include only the intermolecular interactions present (i.e., water–sugar interactions). From these runs we obtain values of 3.30 ± 0.22 kcal/mol (**1b**, without 10–12) and 2.35 ± 0.54 kcal/mol (**2b**, with 10–12), which are too large in magnitude but correctly predict that the β-anomer of glucose is more stable than the α one. Interestingly, Ha et al. obtained a value for the intermolecular contribution to the free energy of solvation difference of 3.03 ± 0.5 kcal/mol, which is close to the values we have obtained.[21] To make further comparisons to the previous work of Ha et al., we have redone their FEP calculations using their parameters within the AMBER suite of programs (run **3**). These runs will also reflect to some extent the parameterization of Homans, because this author used the Ha et al. charge set (or a slight modification to it) for his force field.[22] From this calculation we obtain a value of 0.46 ± 0.14 kcal/mol, which is in fair agreement with the -0.31 ± 0.43 kcal/mol value Ha et al. obtained,[21] but is in excellent agreement with the experimental value of 0.33 kcal/mol. The difference in the computed values is probably due to the use of a different water model in the two simulations, cutoff schemes, and other protocol parameters.

From these data we postulate that AMBER is not giving an accurate representation of the intramolecular contribution to the free energy. Ha et

TABLE XI.
Residue STO-3G ESP Charges for D-Glucose and D-Mannose.[a]

	Glucose		Mannose	
	α	β	α	β
O_5(ring)	−0.470	−0.338	−0.413	−0.301
O_1	−0.502	−0.489	−0.522	−0.477
O_2	−0.475	−0.522	−0.514	−0.484
O_3	−0.563	−0.558	−0.535	−0.544
O_4	−0.616	−0.616	−0.616	−0.616
O_6	−0.487	−0.616	−0.502	−0.503
C_1	0.415	0.307	0.419	0.310
C_2	−0.040	0.108	−0.181	−0.064
C_3	0.461	0.161	0.399	0.276
C_4	0.364	0.627	0.419	0.516
C_5	0.240	−0.076	0.198	0.016
C_6	0.116	0.356	0.151	0.168
HC_1	0.030	0.056	0.043	0.027
HC_2	0.060	0.071	0.117	0.097
HC_3	−0.048	0.001	−0.040	−0.030
HC_4	−0.067	−0.089	−0.045	−0.065
HC_5	0.005	0.006	0.023	0.052
HC_{6a}	0.015	−0.001	0.003	0.012
HC_{6b}	0.019	−0.020	0.009	0.016
HO_1	0.308	0.308	0.308	0.308
HO_2	0.304	0.313	0.352	0.342
HO_3	0.311	0.355	0.306	0.322
HO_4	0.308	0.308	0.308	0.308
HO_6	0.312	0.348	0.313	0.314

[a] O_4 (−0.616), HO_1 (0.308), and HO_4 (0.308) are fixed at the STO-3G ESP charge values for water.

al.[21] obtained a compensation between the inter- and intramolecular interactions, which led to a free energy value that was in good agreement with experiment. In our case the intermolecular contributions (3.30 ± 0.22 kcal/mol and 2.35 ± 0.54 kcal/mol) are much smaller in magnitude than the intramolecular contribution (~ −7.74 kcal/mol and ~ −5.38 kcal/mol). This leads to a free energy value that is not in good agreement with experiment. The Ha et al. simulations used the same charge set for both the α and β anomers of glucose, whereas we use separate charge sets for the two anomers. This may be part of the difference in the results our two simulations obtain; but, clearly, the intramolecular interactions within AMBER appear to be overestimated whereas they are better balanced in the Ha et al. case.

We have also examined D-mannose with the 10–12 terms only. The case without the 10–12 terms has not been considered because in the D-glucose example this approach performed more poorly than the case with the 10–12 terms. The results from this study are given in Table XIII. The experimental value is −0.40 kcal/mol (again β → α direction). From the MD-FEP simulations on this system, we obtain a value of −6.79 ± 0.10 kcal/mol, which is in the correct direction but too large in magnitude (run **4a**). To again determine if the intermolecular component is causing difficulty, we have carried out an MD-FEP simulation in which we only accumulate the intermolecular contributions (run **4b**). This simulation gives a free energy value of 0.02 ± 0.1 kcal/mol, which is in good agreement with the experimental value in that the free energy difference is small. This result again suggests that the AMBER force field is not giving an accurate representation of the intramolecular interactions within a saccharide. Presently, no MD-FEP simulations using the Ha et al. force field have been reported for D-mannose, but by using the D-glucose charges and other CHARMm parameters we have been able to arrive at a force-field representation for D-mannose. This is not to suggest that we have developed a definitive CHARMm representation for D-mannose, but the way we have put together this model is consistent with what is done within CHARMm. From this simulation (run **5**) we calculate a free energy difference of −2.96 ± 0.02 kcal/mol in favor of the α-anomer, which is in the correct direction but again too great in magnitude.

To further our understanding of the performance of these two force fields, we have carried out quantum and classical mechanical calculations in the gas phase. Tables XIV and XV summarize the results for α-D-glucose and β-D-glucose, respectively, and the various conformers examined are given in Scheme 1. Gaussian 92[37] was used in all of the *ab initio* calculations, whereas MOPAC 5.0 (Ref. 60) was used for the semiempirical calculations. The hydroxymethyl position is defined by two dihedrals: O_5—C_5—C_6—O_6 and C_4—C_5—C_6—O_6. Scheme 1 illustrates the various orientations studied: (a) gauche, trans (GT; the O_5—C_5—C_6—O_6 is 60°, the C_4—C_5—C_6—O_6 is 180°); (b) gauche, gauche (GG; −60°, 60°); (c) trans, gauche (TG; 180°, −60°). In all cases full geometry optimizations were carried out from a configuration in which the primary alcohol was correctly positioned to be in the GT, TG, or GG well. No constraints were used, and we found that none were needed because the primary alcohol conformation varied only slightly from the initial positioning.

In the case of α-D-glucose, all methods employed predicted that the trans–gauche/counterclockwise (TG/CC) (counterclockwise indicates the

TABLE XII.
The Calculated Free Energy Difference between α- and β-anomers of D-Glucose.[a]

Run	Elec	14Elec	NB	14NB	BADH	Total (kcal / mol)
1a						
α-β[b]	−2.062	−2.900	0.195	0.233	0.032	−4.50
β-α	−1.801	−3.221	0.439	0.216	0.029	−4.34
1b						
α-β	3.430	—	0.076	—	—	3.51
β-α	3.147	—	−0.067	—	—	3.08
2a						
α-β	1.179	−4.692	0.302	0.215	0.011	−2.99
β-α	1.157	−4.653	0.177	0.218	0.032	−3.07
2b						
α-β	3.780	—	−0.160	—	—	3.62
β-α	2.490	—	0.046	—	—	2.54
3						
α-β	1.599	−1.936	−0.108	1.068	−0.302	0.32
β-α	−0.327	0.406	−0.148	1.010	−0.343	0.60

[a]Elec includes the free energy contributions of the inter- and intramolecular electrostatic interactions. 14Elec includes the free energy contributions of the 1–4 electrostatic interactions. NB includes the free energy contributions of all of the nonbonded interactions of the van der Waals type. 14NB includes the free energy contributions of the 1–4 nonbonded interactions. BADH includes the free energy contributions of the bond, angle, and dihedral interactions. All values are given in kcal / mol. All free energies are given for the $\beta \rightarrow \alpha$ direction.
[b]α-β and β-α indicate the direction in which the free energy simulation was run. All free energies are given for the $\beta \rightarrow \alpha$ direction.

direction in which the secondary hydroxyl groups point around the pyranose ring) conformer was the lowest in energy (see Scheme 1). This conformer was also found to be the low-energy structure in a related study by Polavarapu and Ewig[61] using the 4-31G basis set (single-point calculations were also done at the 6-31G* level). The ordering of the remaining conformers is highly dependent on the method used (see Table XV), but the Ha et al. force field is in the best agreement with the 3-21G results, whereas AMBER is in better agreement with the 6-31G* results. For β-D-glucose all

TABLE XIII.
The Calculated Free Energy Difference between α- and β-anomers of D-Mannose.[a]

Run	Elec	14Elec	NB	14NB	BADH	Total (kcal / mol)
4a						
α-β[b]	−4.415	−1.188	−0.645	−0.147	−0.293	−6.69
β-α	−4.728	−1.058	−0.645	−0.163	−0.300	−6.89
4b						
α-β	0.094	—	0.032	—	—	0.13
β-α	−0.248	—	0.190	—	—	−0.06
5						
α-β	13.731	−17.580	0.538	0.965	−0.636	−2.98
β-α	15.185	−18.376	0.014	0.674	−0.446	−2.95

[a]Elec includes the free energy contributions of the inter- and intramolecular electrostatic interactions. 14Elec includes the free energy contributions of the 1–4 electrostatic interactions. NB includes the free energy contributions of all of the nonbonded interactions of the van der Waals type. 14NB includes the free energy contributions of the 1–4 nonbonded interactions. BADH includes the free energy contributions of the bond, angle, and dihedral interactions. All values are given in kcal / mol. All free energies are given for the $\beta \rightarrow \alpha$ direction.
[b]α-β and β-α indicate the direction in which the free energy simulation was run. All free energies are given for the $\beta \rightarrow \alpha$ direction.

TABLE XIV.
The Calculated Relative Energies of α-D-Glucose and β-D-Glucose at Various Levels.

Structure	AM1	STO-3G	α-D-Glucose 3-21G	6-31G*	AMBER	Ha et al.[a]
(1) TG / CC	0.0	0.0	0.0	0.0	0.0	0.0
(2) TG / CL	2.8	1.3	0.7	0.8	1.8	0.1
(3) GG / CC	0.5	1.4	1.1	0.1	0.4	0.8
(4) GG / CL	3.2	3.3	2.6	0.9	0.5	0.5
(5) GT / CC	0.1	1.3	2.1	0.2	1.0	0.5
(6) GT / CL	3.8	4.0	5.6	2.8	1.2	1.6
Structure	AM1	STO-3G	β-D-Glucose 3-21G	6-31G*	AMBER	Ha et al.
(1) TG / CC	2.7	2.7	3.2	1.0	0.0	4.6
(2) TG / CL	3.4	1.9	7.2	4.4	6.0	1.8
(3) GG / CC	0.0	0.0	0.0	0.0	8.2	0.0
(4) GG / CL	3.5	3.8	9.2	4.4	4.4	3.1
(5) GT / CC	2.5	3.7	6.1	2.0	2.3	2.5
(6) GT / CL	4.0	4.3	11.6	5.7	3.7	4.1

[a]Using the force field given in Ha et al. (see Ref. 19).

methods except for the AMBER charge model predict that the gauche–gauche/counterclockwise (GG/CC) conformer is the most stable structure. The ordering of the remaining conformers is again dependent on the method employed. It is interesting to note that Polavarapu and Ewig[61] predict that the TG/CC conformer is more stable than the GG/CC conformer for β-D-glucose by ~ 0.02 kcal/mol.

The calculated ΔE values between the two low-energy structures for α-D-glucose and β-D-glucose are given in Table XV. We find that all methods predict a small energy difference favoring the α-anomer, but if we take the energy difference using the "correct" conformer ordering using AMBER we obtain a value of −8.2 kcal/mol. This suggests that AMBER is sampling a different part of phase space in MD-FEP simulations relative to the other force-field method. It is interesting to note that the Ha et al. force field gives a conformational energy surface as flat as the *ab initio* computed one while still giving a good ΔE value between the low-energy structures of the α and β anomers. Finally, it is interesting to speculate why the intramolecular free energy calculated by Ha et al.[21] of −3.6 kcal/mol (favoring the α-anomer) is

TABLE XV.
The Relative Energy Ordering and Energy Differences for α-D-Glucose and β-D-Glucose at Various Levels.

	AM1	STO-3G	3-21G	6-31G*	AMBER	Ha et al.
			α-D-glucose			
Order	1 < 5 < 3 < 2 < 4 < 6	1 < 5 < 2 < 3 < 4 < 6	1 < 2 < 3 < 5 < 4 < 6	1 < 5 < 3 < 2 < 4 < 6	1 < 5 < 6 < 4 < 2 < 3	1 < 2 < 4 < 5 < 3 < 6
			β-D-glucose			
Order	3 < 5 < 1 < 2 < 4 < 6	3 < 2 < 1 < 5 < 4 < 6	3 < 1 < 5 < 2 < 4 < 6	3 < 1 < 5 < 2 < 4 < 6	1 < 5 < 6 < 4 < 2 < 3	3 < 2 < 5 < 4 < 1 < 6
			ΔE (α-D-glucose-β-D-glucose) (kcal / mol)[a]			
ΔE	−0.3	−1.4	−1.1	−0.1	−1.1[b]	−0.7

[a]The energy differences are between the lowest energy structures.
[b]AMBER gives the wrong conformer order (1–1 versus 1–3). If we use the quantum mechanical order (i.e., 6-31G*), we get a ΔE of −8.2 kcal / mol.

so large relative to the value predicted by Cramer and Truhlar[46] (~ -0.5 kcal/mol) even though this force field gives a fairly good accounting of the potential energy surface (see Tables XIV and XV). We suspect that the origin of the large intramolecular free energy using this force field is due to insufficient conformational sampling in the MD-FEP simulation and not due to inherent deficiencies in the force-field representation. In other words, during the MD-FEP simulation one (or several) conformers are sampled, but other important points are not. This could be biasing the MD-FEP results. This phenomenon of insufficient sampling can be clearly seen in the MD simulations described later.

In summary, our results suggest that we do not have a clear picture on the role that inter- and intramolecular interactions play in the anomeric preferences for carbohydrates. It is clear that to get a definitive picture, one must be confident that the force field that is being used is capable of reproducing both the inter- and intramolecular interactions present in carbohydrate/aqueous solutions. Furthermore, it is possible that free energy simulation methods are being hampered by insufficient sampling of intramolecular degrees of freedom even in cases in which the force field being used reproduces *ab initio* data on intramolecular conformational preferences. Clearly, this question must be resolved to determine the reliability of force-field methods.

MOLECULAR DYNAMICS SIMULATION

A molecular dynamics simulation of α-D-glucose has been carried out to study the ability of the new force field in predicting the dynamics of carbohydrates. The molecule was set in a solvent box of 436 SPC/E (Ref. 55) water molecules using periodic boundary conditions at constant pressure. The STO-3G charges for α-D-glucose (Table VIII) were used. SHAKE[56] was also employed throughout the simulation and the temperature of the system was set at 298 K. In treating nonbond interactions, a 9.0-Å cutoff was used, with the nonbond pairlist updated every 25 steps. The system was minimized for 500 cycles using the steepest descent method followed by 500 cycles of conjugate gradient minimization. Equilibration was carried out for 52.5 ps. Coordinates and velocities were then collected for 150 ps using a 1.5-fs timestep.

As indicated in Scheme 1, the hydroxymethyl group can adopt three conformations (GG, GT, and TG) defined by the $O_5—C_5—C_6—O_6$ and $C_4—C_5—C_6—O_6$ dihedral angles. In the present simulation, the starting minimized structure was in the GG orientation; however, it quickly started sampling the GT conformation during the sampling phase of the simulation. This was continued until 46 ps into the sampling phase of the simulation, when the GG conformation was formed again. The molecule retained this conformation throughout the sampling period except for one brief time period ($\sim$ 2 ps) in which it again sampled the GT region. This can be seen easily from a history of the $C_4—C_5—C_6—O_6$ dihedral during the sampling phase (see Fig. 1). In the polar plots the 0° position is at 3 o'clock, and by rotating counterclockwise the torsion angle increases; the simulation time increases radiating outwardly. Hence, during the sampling phase of the simulation the GG well was sampled 68% of the time, the GT for 32%, and the TG well was not sampled at all. These results are supported by available experimental evidence that shows the GG and GT conformations to be predominant.[62-64]

Previous simulations studying the orientation of the hydroxymethyl group have produced contradicting results with respect to which conformation of the D-glucose molecule is favored. An *in vacuo* study[27] using the Rasmussen potential function[26] favored the GT (55%) over the TG (38%), with GG occupying only a small percentage (7%) of the sampling period. In an MD simulation of glucose in water, using the CHARMm potential function and the SPC[55] water model, a transition of the exocyclic group from the GT to the TG conformation occurred.[20] In a series of free energy simulations,[21] also using the CHARMm potential function but with the TIP3P[65,66] water model, the hydroxymethyl rotated from the starting TG orientation to either the GG or the GT orientation. These authors have attributed the differences between the observed conformation of the exocyclic hydroxymethyl to the different water models used,

FIGURE 1. Dihedral History of the $C_4—C_5—C_6—O_6$ torsion angle.

insufficient equilibration and sampling, and differences in the truncation method employed.[21]

Two sets of averaged structural parameters over the 150-ps sampling period were calculated according to the orientation of the hydromethyl group (see Scheme 1) to determine its effect on the stability of the ring geometry. As can be seen from the results (Table XVI), a comparison of the two conformations observed in our simulation demonstrates that the position of the hydroxymethyl has little effect on the geometrical parameters of the ring. The largest deviations are 1.9 and 1.4 degrees for the $O_1—C_1—C_2—O_2$ and $O_1—C_1—C_2—C_3$ dihedrals, respectively. Of the geometrical parameters in proximity to the hydroxymethyl, the $C_5—O_5—C_1—C_2$ dihedral has the largest difference between the two averaged structures of 1.2°. The root mean square (rms) values of the various angles and dihedrals are also relatively the same between the two structures, with the $O_1—C_1—C_2—O_2$ dihedral fluctuating the most in both structures. The averaged structural parameters are in excellent agreement with the crystal structure[40,67] and with previous MD simulation studies of glucose in both water[20] and *in vacuo*.[27] The most profound difference, as compared to the MD CHARMm solution study, is 10.9° for the $O_1—C_1—C_2—O_2$ dihedral. This effect could be due to our force field's tendency to overestimate intramolecular interactions, causing the O_2 hydroxyl to be moved toward the O_1 oxygen to interact with it. Overall, though, we see that the various force fields that have been used in the aforementioned simulations and in the present simulation have been able to produce results that are in agreement with the available experimental data for the ring structure but fail to agree on the orientation of the exocyclic hydroxymethyl group. However, this latter effect could be due to insufficient sampling of the conformational space of α-D-glucose.

Another point of interest is the degree to which the sugar hydroxyls are hydrogen bonding to one another and to solvent. Brady[20,27,44] has observed long-lived intramolecular hydrogen bonds in his simulations on α-D-glucose in solution and in vacuum. Recently, we have suggested that the accuracy of this observation is uncertain because of the lack of accurate information regarding torsion potentials for sugar hydroxyls.[35] However, Cramer and Truhlar[46] have suggested on the basis of semiempirical calculations coupled with a continuum solvation model that intramolecular hydrogen bond rings represent minima in solution. This observation can be questioned because of the observation that AM1 in many cases is not able to reproduce correctly the conformational propensi-

TABLE XVI.
Average Ring Structure of α-D-Glucose.

Angle	Average GG	rms	Average GT	rms	CHARMm[a]	rms	Crystal Structure[b]
$C_1—C_2—C_3—C_4$	−54.4	6.0	−54.4	6.1	−52.5	6.0	−51.3
$C_2—C_3—C_4—C_5$	58.8	5.9	57.8	5.8	52.5	6.2	53.3
$C_3—C_4—C_5—C_6$	−177.3	6.6	−178.1	6.6	−176.1	6.7	−176.6
$C_3—C_4—C_5—O_5$	−60.3	6.2	−60.0	6.2	−53.9	6.0	−57.5
$C_4—C_5—O_5—C_1$	60.0	6.5	60.8	6.4	59.7	5.7	62.2
$C_5—O_5—C_1—C_2$	−55.5	7.0	−56.7	6.5	−60.2	5.6	−60.9
$O_5—C_1—C_2—C_3$	51.3	6.7	52.5	6.8	54.8	5.8	54.1
$C_1—C_2—C_3—O_3$	−174.0	6.6	−173.6	6.6	−175.0	6.9	−172.0
$C_2—C_3—C_4—O_4$	−179.3	6.5	180.0	6.4	175.6	6.8	175.3
$O_1—C_1—C_2—O_2$	47.0	8.1	48.9	7.8	58.5	7.8	56.9
$O_1—C_1—C_2—C_3$	−72.6	7.5	−71.2	7.5	−66.2	6.9	−68.7
$C_1—C_2—C_3$	109.4	3.4	109.6	3.4	109.7	3.1	111.1
$C_2—C_3—C_4$	109.5	3.5	109.5	3.4	110.7	3.3	109.8
$C_3—C_4—C_5$	107.3	3.5	107.4	3.4	110.3	3.2	111.1
$C_4—C_5—C_6$	110.7	3.7	111.1	3.8	113.3	3.8	111.5
$C_4—C_5—O_5$	109.9	3.5	110.2	3.6	109.8	2.8	108.7
$C_5—O_5—C_1$	114.6	3.2	114.0	3.2	115.2	2.5	113.7
$O_5—C_1—O_1$	111.2	3.1	111.1	3.0	108.6	2.9	111.5

[a] See J. W. Brady, *J. Am. Chem. Soc.*, **111**, 5155 (1989).
[b] See G. M. Brown and H. A. Levy, *Science*, **147**, 1038 (1965), and G. M. Brown and H. A. Levy, *Acta Cryst.*, **B35**, 656 (1979).

ties of mono and di hydroxylated pyrans.[35] Nonetheless, the available evidence suggests that intramolecular hydrogen bond rings are present in solution and, indeed, we also find this to be the case in our MD simulation of α-D-glucose.

The dihedral histories (Fig. 2) provide good insights into this by illustrating the extent of rotation of the hydroxyls. Examination of the plots reveal that the most rigid hydroxyl is $C_5—C_6—O_6—HO_6$ (Fig. 2e), and to a lesser extent $O_5—C_1—O_1—HO_1$ (Fig. 2a) and $C_3—C_4—O_4—HO_4$ (Fig. 2d), indicating that these hydroxyls are hydrogen bonding more frequently than the flexible $C_1—C_2—O_2—HO_2$ and $C_2—C_3—O_3—HO_3$ (Figs. 2b and 2c).

Further analysis was carried out to determine the percentage of time that hydrogen bonds were actually occurring. Any H ··· O distance less than 2.6 Å, which formed an angle greater than 120° with the oxygen donor, was defined as a hydrogen bond.[68] Table XVII summarizes the results for the various types of intra- and intermolecular hydrogen bonds that could occur. However, the hydrogen bond analysis, using this set of criteria, was unable to produce any evidence that supported the dihedral histories given in Figure 2. For this reason, the same analysis was performed but without employing the angle dependence. It is interesting to note that in the case of carbohydrates "nonstandard" (in the sense of the preferred angle between the hydrogen bond donor and acceptor) hydrogen bonds are formed in the intramolecular "rings." The distance-only criterion results do predict that $C_5—C_6—O_6—HO_6$ is the most rigid due to the presence of hydrogen bonding interactions. In this case the $O_6—HO_6 \cdots O_5$ hydrogen bond exists 86% of the time and the average $O_5—O_6$ distance is 2.68 Å, which is the smallest of all the oxygen-oxygen distances.

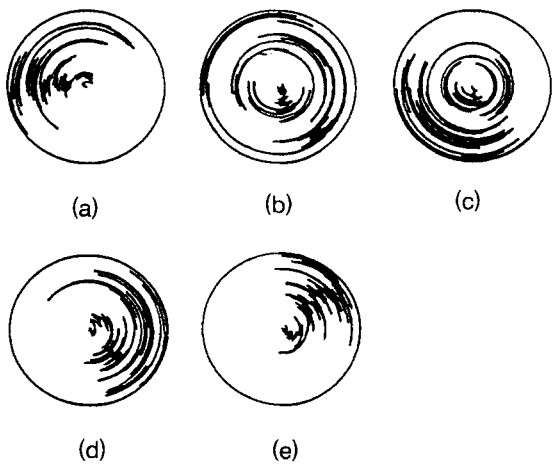


FIGURE 2. Dihedral histories of the five hydroxyls. (a) $O_5—C_1—O_1—HO_1$. (b) $C_1—C_2—O_2—HO_2$. (c) $C_2—C_3—O_3—HO_3$. (d) $C_3—C_4—O_4—HO_4$. (e) $C_5—C_6—O_6—HO_6$.

From these data we can also explain the extreme fluctuations that HO_2 and HO_3 hydroxyls undergo. A comparison of the graph of the $O_2—HO_2 \cdots O_1$ versus the $O_2—HO_2 \cdots O_3$ hydrogen bond (graph not included) demonstrates that when HO_2 is not hydrogen bonding to O_1, it is hydrogen bonding to O_3. O_3 also exhibits this same characteristic, fluctuating between hydrogen bonds with O_2 and O_4. Thus, the possibility of different intramolecular hydrogen bonds leads to the greater fluctuations of the HO_2 and HO_3 hydroxyls. In general, we find that transitions for the hydroxyls occur more rapidly and frequently than those in the MD CHARMm solution study[20] but similar to those in the MD vacuum simulation.[27]

Intermolecular hydrogen bonding was studied as well. When the glucose hydroxyls are the hydrogen bond donor, we see much similarity among the hydrogen bond percentages of the five hydroxyls. HO_6 was hydrated least often (86%) because of its close relationship with the ring oxygen, which decreases its ability to form hydrogen bonds with water. When the glucose hydroxyls are the hydrogen bond acceptors, the percentages range from 59% for O_5 to 99% for O_6. The low percentage for O_5 is due to the proximity of the hydroxymethyl group, which reduces the availability of O_5 for forming hydrogen bonds. The similarity among the intermolecular hydrogen bond data suggests that intramolecular hydrogen bonding may play a more substantial role in determining sugar conformation because its variation is more substantial.

Radial distribution functions (RDF) for water oxygen-glucose oxygen pairs were evaluated to get more insight into water structure around α-D-glucose. Table XVIII summarizes the results from this analysis. The position (2.68 Å–2.73 Å) and height of the first peak (1.5–1.7) are similar among the five hydroxyl oxygen- water oxygen RDFs. The first minimum also occurs at the same relative position (~ 3.4 Å) in all cases. For this reason, the coordination number was calculated out to ~ 3.4 Å for all water oxygen-hydroxyl oxygen pairs. The RDF for the O_1 hydroxyl is given as an example in Figure 3. The solute volume was not excluded in the RDF calculations to make a better comparison to the CHARMm simulation analysis. The RDFs for O_2 and O_3 have the same coordina-

TABLE XVII.
Hydrogen Bond Data from the α-D-Glucose MD Simulation.

Bond	Percent of Time: Angle and Distance Criteria	Percent of Time: Distance Only Criterion
Intramolecular Hydrogen Bonds		
Acceptor-Donor		
$O_1 \cdots HO_2 - O_2$	26.5	49.1
$O_2 \cdots HO_1 - O_1$	19.6	54.6
$O_2 \cdots HO_3 - O_3$	5.5	29.4
$O_3 \cdots HO_2 - O_2$	13.3	21.8
$O_3 \cdots HO_4 - O_4$	0.7	51.6
$O_4 \cdots HO_3 - O_3$	8.8	27.2
$O_5 \cdots HO_6 - O_6$	11.5	86.8
Intermolecular Hydrogen Bonds		
Donor-Acceptor		
$O_1 - HO_1 \cdots OW$	48.8	95.4
$O_2 - HO_2 \cdots OW$	42.9	94.4
$O_3 - HO_3 \cdots OW$	57.7	93.4
$O_4 - HO_4 \cdots OW$	69.0	94.8
$O_6 - HO_6 \cdots OW$	52.0	86.0
Acceptor-Donor		
$O_1 \cdots HW - OW$	62.9	93.4
$O_2 \cdots HW - OW$	81.0	95.1
$O_3 \cdots HW - OW$	68.0	98.0
$O_4 \cdots HW - OW$	50.2	98.0
$O_5 \cdots HW - OW$	37.9	59.3
$O_6 \cdots HW - OW$	60.8	99.4

TABLE XVIII.
Radial Distribution Function Data from the MD Simulation of α-D-Glucose.

Atom	Position of First Peak	Height of first Peak	Coordination Number out to 3.4 Å
O_1	2.73	1.63	2.9
O_2	2.68	1.65	3.2
O_3	2.73	1.61	3.2
O_4	2.73	1.70	2.9
O_5[a]	4.93	1.20	2.4
O_6	2.68	1.50	3.0

[a]The coordination number for the ring oxygen is calculated out to a distance of 3.7 Å.

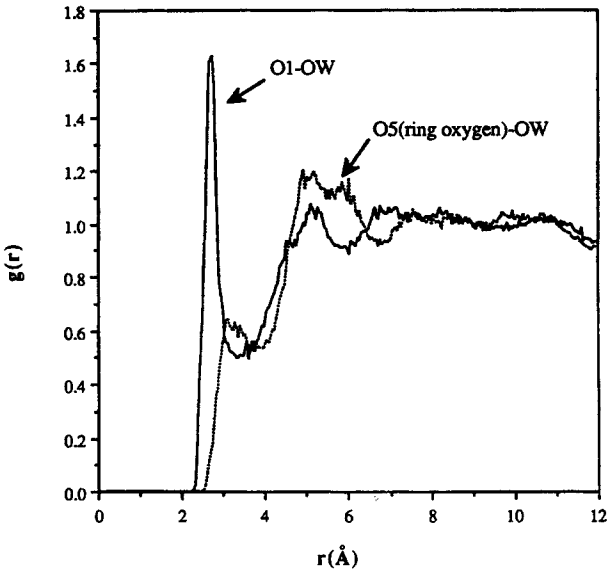


FIGURE 3. Radial distribution functions for glucose oxygen–water oxygen pairs.

tion number of 3.2. O_1 and O_4 also have identical coordination numbers of 2.9, and O_6 has a similar coordination number of 3.0. The slightly higher coordination numbers for the more flexible O_2 and O_3 can be explained by the fact that they have two adjacent hydroxyl group neighbors that bring more water molecules into close proximity, whereas O_1 and O_4 only have one.

The RDF results are similar to those obtained from the MD CHARMm simulation,[20] including similarities between atoms and positions of peaks and minima. There are differences quantitatively, though, with respect to the height of the peaks and the coordination number for those solute atoms that display different flexibility between the two simulations. The coordination numbers for O_2 and O_3 are the ones that disagree most between the two simulations. This is expected if we compare the dihedrals for the hydroxyls, which show that the MD CHARMm simulation does not predict the degree of flexibility that this study has (whereas the flexibility for O_1, O_4, and O_6 is comparable between the two studies). Hence, there is greater similarity of the coordination numbers between the two simulations for these hydroxyls than for the others. These differences can be attributed to the different water models used and to the different force fields employed.

The lower partial charge on the ring oxygen causes the first peak and minimum to be shifted outward radially (see Fig. 3) compared to the water oxygen–hydroxyl oxygen pair RDFs described

earlier. The ring oxygen–water oxygen minimum occurs at ~ 3.7 Å with a coordination number of 2.4 out to this point. A large second solvation shell is also observed at 5.0 Å. The MD CHARMm simulation does not display the large second peak that we observe. This second peak is due to the displacement of water molecules away from the ring oxygen caused by the blocking hydroxymethyl group that is hydrogen bonded to O_5. Hence, this dissimilarity is more a result of the differences observed in the conformation of the hydroxymethyl group than in the differences in the water models employed.

The mean square displacement (MSD) of water molecules from the solute oxygen atoms was calculated to get further insights into the water dynamics surrounding α-D-glucose. The water molecules were divided in the same manner as in the MD CHARMm simulation.[20] All solute oxygens were given a 3.5-Å radius sphere, and any water that did not fall within these regions was defined as bulk water.

Diffusion coefficients were calculated from the linear portion of the MSD plot for each type of water. We observe that in all cases the diffusion constants (ranging from 2.3 to 2.4×10^{-5} cm^2/s) are similar to that observed for bulk SPC/E water (diffusion coefficient of 2.5×10^{-5} cm^2/s).[69] These results are in contrast to the CHARMm results. The CHARMm simulation had a bulk water diffusion coefficient of 3.6×10^{-5} cm^2/s, which agrees with the SPC water model it used.[55] However, the diffusion constants varied substantially depending on the hydroxyl group considered. Also, the CHARMm simulation predicted a diffusion rate faster than that of bulk water for the water molecules surrounding the exocyclic methylene group, whereas the diffusion coefficient of 2.3×10^{-5} cm^2/s for C_6 waters in this simulation is less than that of the bulk water and seems more reasonable. We feel that the major factor giving rise to the differences in the diffusion coefficients is the lengthier equilibration time and sampling period used here, which provides for more adequate statistics.

CONFORMATIONAL ANALYSIS OF β-MALTOSE

Because this force field is designed to allow for extended 1–4 linkages, we need to examine how well this potential function treats conformational maps. A convenient model is maltose [the α(1 → 4) linked dimer of D-glucose; see Fig. 4]. Moreover, this molecule has been studied theoretically,[70–77] and its crystal structure has been solved.[78] Hence, there is substantial information available to allow for comparison.

The crystal structure of β-maltose was used as the starting point for the conformational map. Each conformer was minimized until the rms gradient was 0.1 kcal/mol and a distance-dependent dielectric was used to mimic the presence of solvent. ψ (C_1—O_1—$C_{4'}$—$H_{4'}$) and ϕ (H_1—C_1—O_1—$C_{4'}$) were held fixed by restraining the five atoms involved in the two dihedral angles. The conformational map was determined in 20-degree increments. After an initial scan of the potential energy surface was done, we searched the conformations of the primary hydroxyl groups. We considered GG–GG, GG–GT, GT–GG, GT–GT, TG–GG, and TG–GT. No TG conformers were considered on the reducing glucose residue because this would result in destabilizing interactions.[72] Upon identification of the most stable conformation of the primary hydroxyl groups in each minimum, we assembled the conformational maps (not provided). Although a more intensive study has been undertaken in the past by Dowd et al.[73] using 24 different starting conformers and using the MM3 force field,[79–81] for our purpose here of validating the new force field, we feel that the method we have employed for generating the relaxed PES is sufficient.

Our conformational maps are similar to the ones obtained by Ha et al.[71] (using a CHARMm-based force field) and by Tran et al. (using an MM2-based force field).[72] The GG–GG, GG–GT, GT–GG, GT–GT, TG–GG, and TG–GT plots all show a broad minimum centered at $\psi = -40$ and $\phi = -40$. This has been labelled as the **AB** well by Tran et al. Further minima centered at $\psi = 20$ and $\phi = 20$ (the **C** well) and $\psi = -160$ and $\phi = -40$ (the **E** well) are also present in these figures. Figure 4a shows the structure of β-maltose in the **A** well, which does not have an intraresidue hydrogen bond (as does the structure shown in Fig. 4b for the **C** well). Tran et al. also observed another well labeled the **F** minimum (centered around $\psi = 160$ and $\phi = -50$) which is related to the **E** minimum. However, in our case and in the work of Ha et al., only one continuous minimum was observed (i.e., the **E** well). Tran et al.[72] also identified a **D** well (centered around $\psi = 40$ and $\phi = -40$), but in the present profiles we see no evidence for this minimum. The recent work of Dowd et al.[73] displayed only two regions, the **AB** well and the **E** well, and as suggested by the authors the flatter

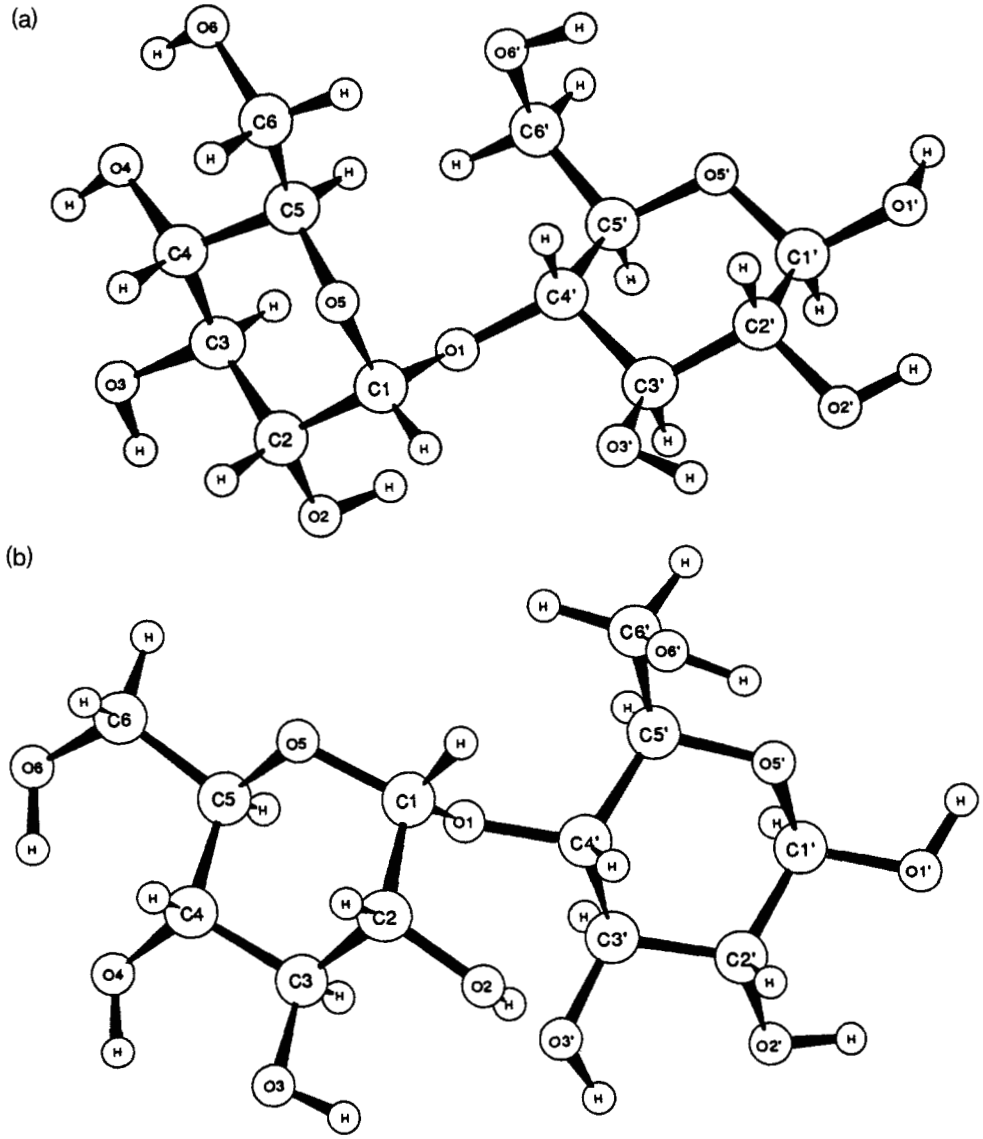


FIGURE 4. Minimum energy structures of β-maltose from the relaxed PES. (a) Structure of β-maltose in the **A** well. (b) Structure of β-maltose in the **C** well.

PE surface could be a result of the large number of starting conformers that were considered. Whether differences between any of the conformational maps are attributable to the force field or to the method employed is not entirely clear, although both surely play a role.

The fully relaxed plot is given in Figure 5. The labeling scheme of the wells is that of Tran et al.,[72] who defined ψ as (C_1—O_1—$C_{4'}$—$C_{5'}$) and ϕ as (O_5—C_1—O_1—$C_{4'}$), which slightly shifts the overall profiles relative to the definition we, Ha et al.,[71] and Dowd et al.[73] have used. In this plot we have taken the lowest energy points contained in the six initial PE surfaces and relaxed the primary hydroxyl conformation constraints. The primary minima do not change location relative to the six starting maps, but the relaxed map's features are not as steep as the latter. This map has the **A**, **B**, **C**, and **E** minima of Tran et al.[72], with the addition of a shallow well which is related to the **D** minimum (centered around $\psi = 50$ and $\phi = -50$) reported by Tran et al. However, our **D** minimum is very small and is not as deep as several of the other minima on the relaxed potential energy surface.

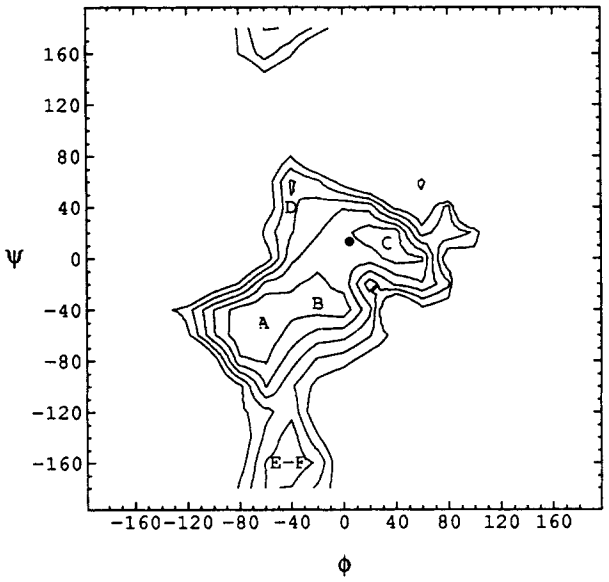


FIGURE 5. The totally relaxed PES of β-maltose. The lowest energy conformation of the starting six PESs at each grid point was chosen. The constraints on the hydroxymethyl groups were relaxed and the structure minimized to within an rms gradient of 0.1 kcal / mol. Contours range from 2.0 to 10.0 kcal / mol above the minimum energy, in levels of 2.0 kcal / mol. The position of the β-maltose crystal structure (●) is also labeled.

Finally, Ha et al.[71] have identified another minimum (the **G** well, centered around $\psi = 140$ and $\phi = 40$), but neither we nor Tran et al. see any hint of this point on the potential energy surface. Also labeled on the PES (Fig. 5) is the crystal structure of β-maltose, the template for the initial conformers. It falls just outside of the 2-kcal/mol contour of the **C** well. Of course, discrepancies always arise when making comparisons between gas-phase calculated structures and crystal structures, which are profoundly affected by packing forces.

Overall, the relaxed potential surface we have presented is similar to those reported by Tran et al.,[72] Ha et al.,[71] and Dowd et al.,[73] with some slight differences. All have the **A** and **B** minimum as being the lowest in energy. However, we and Ha et al. predict that the **C** minimum is similar in energy to the **A** and **B** minima, whereas Tran et al. and Dowd et al. predicted that the **C** minimum is higher in energy than the **A** and **B** minima. The energy ordering of the remaining points on the potential energy surface are ordered differently by the different force fields.

Conclusions

A force field for monosaccharides and 1 → 4 linked polysaccharides has been described. A number of tests (vibrational frequencies, MD simulations, MD-FEP simulations on the solvation free energy differences of the α and β anomers of D-glucose and D-mannose and the conformational map of maltose) have been carried out and indicate that this force field is capable of reproducing experimental data regarding the structure and dynamics of carbohydrates. Furthermore, our force field gives results that are consistent with theoretical results from other labs employing force-field methods. The most glaring weakness of the present force field has to do with the ability to reproduce the experimental solvation energy difference between the α and β anomers of D-glucose and D-mannose. Refinement of the intramolecular interactions within this force field appears necessary based on these results; however, an exploration of the effect MD-FEP timescales have on the convergence of the calculated solvation free energy differences also appears warranted before a definitive conclusion can be reached.

Acknowledgments

We thank the Office of Naval Research for generous support of this research. We also thank the Pittsburgh Supercomputer Center for generous allocations of supercomputer time. The Biomedical Starter Grant (BRSG) program at The Pennsylvania State University is also acknowledged for providing support for this research.

References

1. N. Sharon, *Complex Carbohydrates: Their Chemistry, Biosynthesis and Functions*, Addison-Wesley, Reading, MA, 1975.
2. R. C. Hughs, *Glycoproteins*, Chapman & Hall, London, 1983.
3. N. Sharon, and H. Lis, *Chem. Eng. News*, **59**, 21 (1981).
4. N. Sharon, and H. Lis, *Science*, **246**, 227 (1989).
5. V. T. Marchesi, V. Ginsburg, P. W. Robbins, and C. F. Fox, Eds., *Cell Surface Carbohydrates and Biological Recognition, Proceedings of the ICN-UCLA Symposium*, Vol. 23, Keystone, CO, February, 1977, Liss, New York, 1978.
6. M. Fucuda, Ed., *Cell Surface Carbohydrates and Cell Development*, CRC Press, Boca Raton, FL, 1992.
7. R. U. Lemieux, *Chem. Soc. Rev.*, **7**, 423 (1978).
8. R. Yamasaki, and B. Bacon, *Biochemistry*, **30**, 851 (1991).

9. C. L. Brooks, M. Karplus, and B. M. Pettit, *Proteins: A Theoretical Perspective of Dynamics, Structures, and Thermodynamics*, Vol. LXXI, John Wiley & Sons, New York, 1988.
10. E. Eliel, N. L. Allinger, S. J. Angyal, and G. A. Morrison, *Conformational Analysis*, John Wiley & Sons, New York, 1965.
11. S. J. Angyal, *Angew. Chem. Int. Ed. Engl.*, **8**, 157 (1969).
12. J. F. Stoddart, *Stereochemistry of Carbohydrates*, Wiley-Interscience, New York, 1971.
13. D. A. Brant, *Ann. Rev. Biophys. Bioeng.*, **1**, 369 (1972).
14. D. A. Brant, *Quart. Rev. Biophys.*, **9**, 527 (1976).
15. D. A. Brant, In *The Biochemistry of Plants*, Vol. 3, J. Preiss, Ed., Academic Press, New York, 1980, p. 425.
16. I. Tvaroska, In *Theoretical Chemistry of Biological Systems*, Vol. 41, G. Naray-Szabo, Ed., Elsevier, New York, 1986, p. 283.
17. A. D. French and J. W. Brady, Eds., *Computer Modeling of Carbohydrate Molecules*, American Chemical Society, Washington, DC, 1990.
18. K. Kildeby, S. Melberg, and K. Rasmussen, *Acta Chem. Scand.*, **A31**, 1 (1977).
19. S. N. Ha, A. Giammona, M. Field, and J. W. Brady, *Carbohydr. Res.*, **180**, 207 (1988).
20. J. W. Brady, *J. Am. Chem. Soc.*, **111**, 5155 (1989).
21. S. Ha, J. Gao, B. Tidor, J. W. Brady, and M. Karplus, *J. Am. Chem. Soc.*, **113**, 1553 (1991).
22. S. W. Homans, *Biochemistry*, **29**, 9110 (1990).
23. A. Marsden, B. Robson, and J. S. Thompson, *J. Chem. Soc., Faraday Trans. I*, **84**, 2519 (1988).
24. V. S. R. Rao, M. Biswas, C. Mukhopadhyay, and P. V. Balaji, *J. Mol. Structr.*, **194**, 203 (1989).
25. K. Rasmussen, *Potential Energy Functions in Conformation Analysis*, Vol. 37, Springer-Verlag, New York, 1985.
26. K. Rasmussen, *Acta Chem. Scand.*, **A36**, 323 (1982).
27. J. W. Brady, *J. Am. Chem. Soc.*, **108**, 8153 (1986).
28. S. J. Weiner, *J. Comp. Chem.*, **7**, 230 (1986).
29. S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. S. Alagona, J. Profeta, and P. Weiner, *J. Am. Chem. Soc.*, **106**, 765 (1984).
30. D. E. Williams and J.-M. Yan, *Adv. At. Mol. Phys.*, **23**, 87 (1988).
31. L. E. Chirlian and M. M. Francl, *J. Comp. Chem.*, **8**, 894 (1987).
32. U. C. Singh and P. A. Kollman, *J. Comp. Chem.*, **5**, 129 (1984).
33. M. J. Frisch, M. Head-Gordon, H. B. Schlegel, K. Raghavachari, J. S. Binkley, C. Gonzalez, D. J. Defrees, D. J. Fox, R. A. Whiteside, R. Seeger, C. F. Melius, J. Baker, R. Martin, L. R. Kahn, J. J. P. Stewart, E. M. Fluder, S. Topiol, and J. A. Pople, Gaussian 88, Gaussian, Inc., Pittsburgh, PA, 1988.
34. M. D. Joesten and L. J. Schaad, *Hydrogen Bonding*, Marcel Dekker, New York, 1974.
35. Y.-J. Zheng, S. M. Le Grand, and K. M. Merz, Jr. *J. Comp. Chem.*, **13**, 772 (1992).
36. Y.-J. Zheng and K. M. Merz, Jr. *J. Comp. Chem.*, **13**, 1151 (1992).
37. J. M. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian 92, Gaussian, Inc., Pittsburgh, PA, 1992.
38. K. M. Merz, Jr., *J. Am. Chem. Soc.*, **112**, 7973 (1990).
39. D. M. Fergusen and P. A. Kollman, *J. Comp. Chem.*, **12**, 620 (1991).
40. G. M. Brown and H. A. Levy, *Science*, **147**, 1038 (1965).
41. K. M. Merz, Jr., *J. Comp. Chem.*, **13**, 749 (1992).
42. F. Franks, *Pure Appl. Chem.*, **59**, 1189 (1987).
43. A. Sugget, In *Water: A Comprehensive Treatise*, Vol. 4, F. Franks, Ed., Plenum, New York, 1975, p. 519.
44. J. W. Brady, *Carbohydr. Res.*, **165**, 306 (1987).
45. L. G. Dunfield and S. G. Whittington, *J. Chem. Soc., Perkin II*, 654 (1977).
46. C. J. Cramer and D. G. Truhlar, *J. Am. Chem. Soc.*, **115**, 5745 (1993).
47. J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, London, 1987.
48. M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press, Oxford, UK, 1987.
49. W. F. van Gunsteren and H. J. C. Berendsen, *Angew. Chem. Int. Ed. Engl.*, **29**, 992 (1990).
50. M. Mezei and D. L. Beveridge, *Ann. NY Acad. Sci.*, **482**, 1 (1986).
51. W. F. van Gunsteren, *Protein Eng.*, **2**, 5 (1988).
52. W. L. Jorgenson, *Acc. Chem. Res.*, **22**, 184 (1989).
53. D. L. Beveridge and F. M. DiCapus, *Ann. Biophys. Biophys. Chem.*, **18**, 431 (1989).
54. P. A. Kollman and K. M. Merz, Jr. *Acc. Chem. Res.*, **23**, 246 (1990).
55. H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, and J. Hermans, In *Intermolecular Forces*, B. Pullman, Ed., Dordecht, The Netherlands, 1981; p. 331.
56. W. F. van Gunsteren and H. J. C. Berendsen, *Mol. Phys.*, **34**, 1311 (1977).
57. R. S. Shallenberger, *Advanced Sugar Chemistry, Principles of Sugar Stereochemistry*, AVI Publishing, Westport, CT, 1982.
58. S. J. Angyal, *Aust. J. Chem.*, **21**, 2737 (1968).
59. I. Tvaroska and T. Bleha, *Adv. Carbohydr. Chem. Biochem.*, **47**, 45 (1989).
60. K. M. Merz, Jr. and B. H. Besler, *QCPE Bull.*, **10**, 15 (1990).
61. P. L. Polavarapu and C. S. Ewig, *J. Comp. Chem.*, **13**, 1255 (1992).
62. A. D. Bruyn and M. Anteunis, *Carbohydr. Res.*, **47**, 311 (1976).
63. Y. Nishida, H. Ohuri, and H. Meguro, *Tetrahedron Lett.*, **25**, 1575 (1984).
64. S. J. Perkins, L. N. Johnson, D. C. Philips, and R. A. Dwek, *Carbohydr. Res.*, **59**, 19 (1977).
65. W. L. Jorgenson, *J. Am. Chem. Soc.*, **103**, 335 (1981).
66. W. L. Jorgenson, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, *J. Chem. Phys.*, **79**, 926 (1983).
67. G. M. Brown and H. A. Levy, *Acta Cryst.*, **B35**, 656 (1979).
68. J. Tirado-Rives and W. L. Jorgenson, *J. Am. Chem. Soc.*, **112**, 2773 (1990).
69. H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.*, **91**, 6269 (1987).

70. J. W. Brady and R. K. Schmidt, *J. Phys. Chem.*, **97**, 958 (1993).
71. S. N. Ha, L. J. Madsen, and J. W. Brady, *Biopolymers*, **27**, 1927 (1988).
72. V. Tran and A. Buleson, *Biopolymers*, **28**, 679 (1989).
73. M. K. Dowd, J. Zeng, A. D. French, and P. J. Reilly, *Carbohydr. Res.*, **230**, 223 (1992).
74. C. V. Goebel, W. L. Dimpfl, and D. A. Brant, *Macromolecules*, **3**, 644 (1970).
75. D. A. Rees and P. J. C. Smith, *J. Chem. Soc. Perkin II*, 836 (1975).
76. S. Perez, M. Roux, J. F. Revol, and R. H. Marchessault, *J. Mol. Biol.*, **129**, 113 (1979).
77. S. Perez and C. Vergelati, *Polym. Bull.*, **17**, 141 (1987).
78. M. E. Gress and G. A. Jeffrey, *Acta Cryst.*, **B33**, 2490 (1977).
79. N. L. Allinger, Y. H. Yuh, and J. H. Lii, *J. Am. Chem. Soc.*, **111**, 8551 (1989).
80. J. H. Lii and N. L. Allinger, *J. Am. Chem. Soc.*, **111**, 8566 (1989).
81. N. L. Allinger, M. Rahman, and J. H. Lii, *J. Am. Chem. Soc.*, **112**, 8293 (1990).