Conformational Analysis of GM1 Oligosaccharide in Water Solution with a New Set of Parameters for the Neu5Ac Moiety†

Anna Bernardi* and Laura Raimondi*

Dipartimento di Chimica Organica e Industriale, via Venezian 21, 20133 Milano, Italy

Received July 12, 1994 (Revised Manuscript Received February 28, 1995®)

An unconstrained Monte Carlo/energy minimization (MC/EM) conformational search was performed on ganglioside GM1 oligosaccharide, and the results were compared to the known solution conformation of this molecule. The search was performed using the continuum dielectric water solvation model as implemented in MacroModel/Batchmin and the force field AMBER*. The latter was modified to include new parameters for the Neu5Ac residue contained in GM1. The parameters were developed based on molecular orbital calculations on simple model systems and therefore should have general validity for molecular mechanics calculations on sialyl oligosaccharides. A nice agreement was reached between the computed GM1 structure and the available NMR data.

Introduction

Gangliosides are amphiphylic molecules characterized by the presence of one or more residues of sialic acid. They are located on the outer surface of vertebrate cell membrane (particularly in the nervous system) with the saccharide portion extending from the cell surface and interacting with a number of external biologically active factors.[1] Ganglioside GM1 [Galβ1–3GalNAcβ1–4(Neu5Acα2–3)Galβ1–4Glcβ1–1Cer, **1a**] (Figure 1) is the membrane receptor of some toxin proteins causing diarrhea-related diseases, namely cholera toxin (CT) and the closely related (more than 80% homologous to CT) heat-labile toxin of *E. coli* (LT). Both LT and CT are made of two different polypeptide chains, called A (27K) and B (11.6K), that combine to form an AB_5 complex of $M_r \sim 86$ K. The recognition of GM1 by CT and LT is the first step in the pathogenic action of these toxins. It occurs on the host cell surface *via* the specific interaction of the doughnut-shaped B_5 pentamer with the oligosaccharide portion of GM1. After binding, a fragment of the A subunit (A1) is transferred through the cell membrane and initiates a complex sequence of reactions that ultimately leads to the efflux of ions and fluids from the infected cell.

The binding of GM1 to the B_5 pentamer is one of the best understood cases of those protein–saccharide interactions that appear to play a major role in the recognition processes at cell surface. The mechanism of complexation has been studied in great detail.[2–4] Recently, the crystal structure of unbound LT[5a] and of its complex with lactose (Galβ1–4Glu)[5b] have also been determined. This latter structure seemed to give important indications about the nature of GM1 binding to LT, which were later confirmed by the X-ray structure of CT B-pentamer bound to GM1.[6]

The solution structure of GM1 itself has been investigated with various methods. These studies have combined experimental data, such as NOE-derived distances, and computational techniques to derive a tridimensional model of GM1 in DMSO and water–dodecylphosphocholine solution.[7]

As part of a program aimed to devise a good computational approach to the study of protein–saccharide interactions, we needed to establish a computational protocol that would successfully reproduce the main features of the experimental structure of GM1 in an unbound search starting from a random conformation. The assumption is that such protocol could then be used to perform a conformational analysis of the saccharide in the LT and CT binding pockets and lead to a meaningful model of the GM1/LT and GM1/CT complexes. Since GM1 contains a neuraminic acid residue (Neu5Ac), the first step toward this goal was to determine whether this monosaccharide can be adequately treated with existing molecular mechanics parameters. As it turns out, a new set of AMBER-type parameters had to be developed to account for the known NMR data on sialic acid derivatives. The parameters were derived from molecular orbital calculations on simple model systems, added to the AMBER* force field, and tested on Neu5AcOMe (**2**, R = Me; see Figure 2). Finally, the force field augmented of the new parameters was used for a conformational analysis of GM1 pentasaccharide. The results were then compared with the solution structure of the saccharide as obtained from NMR data.[7]

Computational Section

General Methods. *Ab initio* calculations were performed using the GAUSSIAN 90[8] package on a SGI Iris workstation. All variables were optimized at the RHF/3-21G[9] level, and the stationary points were characterized as minima by vibrational analysis. The RHF/6-31G*[9] single point calculations were performed on the 3-21G geometries. Semiempirical calculations were carried out using standard procedures as imple-

† This work was presented in part at the First European Conference on Computational Chemistry (ECCC1), Nancy, France, May 23–27, 1994.

® Abstract published in *Advance ACS Abstracts,* April 15, 1995.

(1) Fishman, P. H. In *New Trends in Ganglioside Research*; Liviana Press; Padova, 1988.

(2) (a) Cuatrecasas, P. *Biochemistry* **1973**, *12*, 3547. (b) Fishman, P. H.; Moss, J.; Osborne, J. C., Jr. *Biochemistry* **1978**, *17*, 711.

(3) (a) Iida, T.; Tsuji, T.; Honda, T.; Miwatani, T.; Wakabayashi, S.; Wada, K.; Matsubara, H. *J. Biol. Chem.* **1989**, *264*, 14065. (b) De Wolf, M. J. S.; Fridkin, M.; Kohn, L. D. *J. Biol. Chem.* **1981**, *256*, 5489.

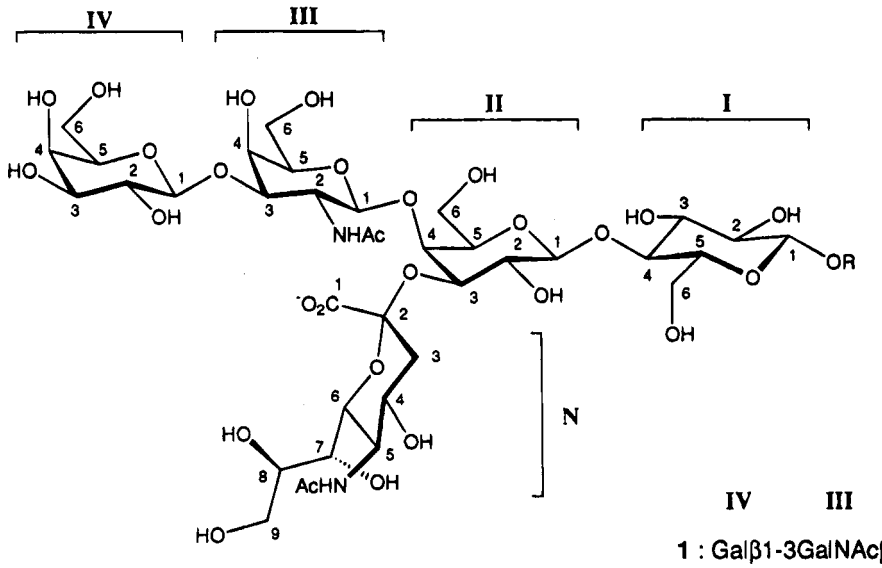
(4) Jobling, M. G.; Holmes, R. K. *Mol. Microbiol.* **1991**, *5*, 1755.

(5) (a) Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; Wartna, E. S.; v. Zanten, B. A. M.; Witholt, B.; Hol, W. G. J. *Nature* **1991**, *351*, 371. (b) Sixma, T. K.; Pronk, S. E.; Kalk, K. H.; Wartna, E. S.; v. Zanten, B. A. M.; Berghuis, A. M.; Hol, W. G. J. *Nature* **1992**, *355*, 561.

(6) Merritt, E. A.; Sarfaty, S.; v. d. Akken, F.; L'Hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.* **1994**, *3*, 166.

(7) (a) Acquotti, D.; Poppe, L.; Dabrowski, J.; v. d. Lieth, C-W; Sonnino, S.; Tettamanti, G. *J. Am. Chem. Soc.* **1990**, *112*, 7772 and references therein. (b) Sabesan, S.; Block, K.; Lemieux, R. U. *Can. J. Chem.* **1984**, *62*, 1034.

0022-3263/95/1960-3370$09.00/0

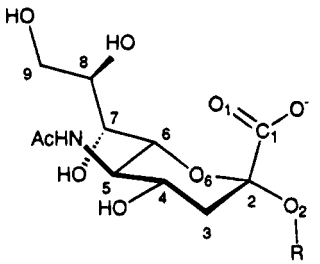


IV III N II I

1 : Galβ1-3GalNAcβ1-4(Neu5Acα2-3)Galβ1-4Glcβ1-R

1a: R = 1-Ceramide
1b: R = Me

Figure 1. Ganglioside GM1 pentasaccharide.



2

φ_C C_1-C_2-O_2-R

χ O_1-C_1-C_2-O_6

θ_1 H_6-C_6-C_7-H_7

θ_2 H_7-C_7-C_8-H_8

θ_3 H_8-C_8-C_9-O_9

Figure 2. Neu5Ac residue of GM1.

mented in MOPAC 6.0.[10] Geometries were optimized with the PRECISE option, and minima were characterized as such by a FORCE calculation. The single and double drive options as implemented in MOPAC 6.0 were used to construct the Ramachandran plot with 30° resolution. Molecular mechanics calculations were performed using MacroModel/Batchmin[11] 3.5 version. The force field used was AMBER* [12] augmented of the parameters described in the text. Calculations in water were performed using the continuum dielectric water solvent model (GB/SA) of Batchmin.[13] Extended cutoff distances were employed (8 Å in van der Waals, 20Å in charge/electrostatics and 10 Å in charge/multipole electrostatics).

Conformational analyses of Neu5AcOMe (**2**, R = Me) and GM1 pentasaccharide (**1b**) were performed in GB/SA water using the pseudosystematic Monte Carlo procedure of Still and Goodman.[14] The search proceeds by altering in a pseudosystematic way the torsional angles of a starting structure (a minimum energy conformation). Thus, a new geometry is generated, which is then energy minimized. This process produces a new minimum energy conformation, which in turn is tested for duplication with previously found conformations. The procedure therefore alternates random changes of coordinates (MC), which allow wide sampling of the potential energy surface, and energy minimizations (EM). This MC/EM procedure has been proven to be among the most effective methods at finding all (or nearly all) low energy conformations of flexible molecules. For **2** the anomeric torsion angle φ_C, all the side chain (θ_1, θ_2, and θ_3, see Figure 2) and C−O torsions, as well as the carbon−nitrogen bond were allowed to vary in the MC steps, for a total of nine explicit variables. For GM1 all the anomeric torsions φ, φ_C, and ψ, the C_5−C_6 ω torsions, all the side chain θ_1, θ_2, and θ_3 torsions of the Neu5Ac residue, and all the carbon−nitrogen bonds (18 torsions total) were used as explicit MC variables. The number of torsion angles allowed to vary simultaneously during each MC step ranged from two to eight for Neu5AcOMe and from two to 17 for GM1. A total of 3500 search steps for Neu5AcOMe and 14 000 for GM1 were performed using the united atom version of AMBER*, starting from a structure derived from NMR data.[7,15] Energy minimization (EM) was performed using the truncated Newton conjugate gradient (TNCG) procedure and was terminated either after 500 iterations or when the energy gradient root mean square (rms) fell below 0.024 kcal/Å mol.

After completion of the MC/EM conformational search, the low energy conformers were subjected to further energy minimization in GB/SA water using the TNCG method to reduce the gradient rms to less than 0.0024 kcal/Å mol. All conformers were saved that differed from the global minimum-energy conformation by no more than 12 kcal/mol. To eliminate duplicate conformations a comparison was performed on the heavy atoms only. Thus, only the global minimum for each family of conformers was retained, independent of the OH conformations. All these minima were reminimized with the all-atom version of AMBER* to a gradient rms of less than 0.0024 kcal/Å mol. The elimination of duplicate conformers was again performed by comparing the heavy atoms only. This procedure was followed because the all-atom AMBER* gives a more accurate reproduction of the experimental data, while the united atom AMBER* force field is more convenient for the initial computational search, since it requires less CPU time for each minimization.

The 10 lowest energy minima of each search were fully

(8) GAUSSIAN 90, Revision H: Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzales, C.; Defrees, D. J.; Fox, D. J.; Seeger, R.; Whiteside, R. A.; Melius, C. F.; Baker, J.; Martin, R. L.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A. Gaussian, Inc., Pittsburgh, PA, 1990.

(9) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley and Sons: New York, 1986.

(10) Stewart, J. J. P. MOPAC 6.0, 1990, QCPE Program no. 455.

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

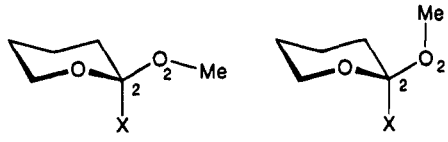
(12) Homans, S. W. *Biochemistry* **1990**, *29*, 9110.

(13) Still, W. C.; Tempzyk, A.; Hawley, R.; Hendrickson, T. F. *J. Am. Chem. Soc.* **1990**, *112*, 6127.

(14) Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110.

(15) Brown, E. B.; Brey, W. S.; Weltner, W., Jr. *Biochim. Biophys. Acta* **1975**, *399*, 124.

Chart 1. 2-Methoxytetrahydropyrans 3 and 4



a, *gauche* **b**, *anti*

3 X = H

4 X = Me

Table 1. Relative Energies (kcal/mol) of *Gauche* and *Anti* Conformations of 2-Methoxytetrahydropyrans 3 and 4 (See Chart 1)

entry	X	**a**, *gauche* (φ_X)[a]	**b**, *anti* (φ_X)[a]
	3 (X = H)		
1	RHF/6-31G*//RHF/3-21G[b]	0.0	2.8
2	MNDO	0.0 (−44)	1.6 (160)
3	AM1	0.1 (−50)	0.0 (−177)
4	PM3	0.5 (−20)	0.0 (174)
	4 (X = Me)		
5	RHF/6-31G*//RHF/3-21G	0.0 (−72)	0.2 (−157)
6	MNDO	0.0 (−76)	0.3 (−154)

[a] φ_X = X−C_2−O_2−Me. [b] Reference 17.

characterized as minimum energy conformations by computing second derivatives (MTST option of Batchmin).

Filtering of the output files for selection of conformations featuring $\theta_2 = 180° \pm 30°$ (see below) was performed using the 3D-search (3DSch) facility of MacroModel.

The computations were performed on a HP 720 or on a HP735 workstation and visualized on a SGI-Iris workstation.

Development of the Parameters. A new set of parameters had to be developed to properly describe the known conformational preferences of sialic acid containing oligosaccharides.[7,15,16] In particular, the two features that appeared to be critical (*vide infra*) were the Neu5Ac anomeric torsion φ_C (C_1−C_2−O_2−R in Figure 2) and the C_7−C_8 torsion in the side chain (θ_2).

Neu5Ac Anomeric Torsion. The NMR data on sialyl oligosaccharides[7,15,16] are consistent with an almost antiperiplanar relationship between the sialyl anomeric substituent R and the carboxy group (Figure 2). This is in contrast with the usual *gauche* conformation adopted by most aldopyranosides. A complete *ab initio* study for the conformational preferences at the anomeric position of 2-methoxytetrahydropyran (**3**) (Chart 1) was performed by Wiberg and Murcko[17] in 1989 (Table 1, entry 1). For the equatorial isomer **3** the *gauche* conformation **3a** (X−C_2−O_2−Me = −60°) was found to be more stable by 2.8 kcal/mol than the *anti* **3b** (X−C_2−O_2−Me = 180°) at the RHF/6-31G*//RHF/3-21G level of theory (Table 1). The latter conformer has a X−C_2−O_2−Me antiperiplanar arrangement, analogous to the one found in the sialic derivatives.[7,15,16] The stabilization of the *anti* rotamer observed for these compounds could well be due to the steric hindrance of the carboxy substituent in the axial position. Therefore, we initially carried out a similar study on 2-methyl-2-methoxytetrahydropyran (**4**) (Chart 1), in order to assess the steric influence of an axial substituent at the anomeric position. The geometry of the two conformers **4a** and **4b** was fully optimized at the RHF/3-21G level of theory,[9] and single point calculations at the RHF/6-31G*[9] level were performed. Indeed, the *gauche*/*anti* energy difference of 2.8 kcal/mol vanishes on passing from **3** to **4** (Table 1, entry 5). The *anti* conformer **4b** is destabilized by only 0.2 kcal/mol with respect to the *gauche* conformation **4a**.

The above calculations gave indications about the nature of the effect. However, what is needed in order to define a reliable parameter set is a complete analysis of the torsional profile for the two anomeric bonds (φ_C = C_1−C_2−O_2−R and χ = O_1=C_1−C_2−O_6; see Figure 2) of the sialic derivatives or of a suitable model system such as 2-carbomethoxy-2-methoxytetrahydropyran (**5**) (Figure 3). In view of the size of the problem, this analysis was performed using semiempirical methods. The performance of the different semiempirical Hamiltonians AM1, PM3, and MNDO[18] in the case at hand was tested by comparing the energy differences calculated with these methods for the rotamers of **3** and **4**. Quite surprisingly, both AM1 and PM3 proved to be rather unreliable on these systems: for 2-methoxytetrahydropyran the two conformers **3a** and **3b** were found to have almost the same energy, in complete disagreement with Wiberg's data (Table 1, entries 3 and 4).[9] On the contrary, MNDO was found to reasonably reproduce the energy differences determined *ab initio*, both for **3** and **4** (Table 1, entries 2 and 6). Therefore, MNDO was chosen to calculate the Ramachandran plot relative to the anomeric torsions φ_C (C_1−C_2−O_2−Me) and χ (O_1=C_1−C_2−O_6) of **5**. The potential energy surface obtained with this Hamiltonian is reported in Figure 3a. It features a large valley for $-150° \le \varphi_C \le -70°$ with a very shallow potential along χ. Barriers to rotation for χ are 1−2 kcal/mol. These features agree both with experimental data and preliminary *ab initio* calculations performed on **5**.[19] Although the use of semiempirical methodologies may be regarded as not completely satisfactory from a theoretical point of view, it can be considered as an "interim solution" whose merits, if any, will be proved by further use.

The MNDO torsional profile for φ_C and χ was reproduced in the molecular mechanics calculations by using the new AMBER-type parameters reported in Chart 2 (Neu5Ac anomeric torsion substructure) in MacroModel substructure format. The AMBER* Ramachandran plot is shown in Figure 3b: the C_1−C_2−O_2−Me torsion shows only a single large minimum centered at −120°. Two minima separated by low barriers (1.2−3.6 kcal/mol) are located along the χ coordinate. Optimization of the two minima yielded two structures with $\varphi_C, \chi = -114°, +130°$ (rel E = 0.0 kcal/mol) and $\varphi_C, \chi = -120°, -16°$ (rel E = 0.7 kcal/mol).

Neu5Ac Side Chain. Houk and co-workers have recently discussed in detail the conformational properties of O−C−C−C fragments, on the basis of *ab initio* calculations.[20] Accurate fit of these data may be quite important for a good description of sialic derivatives' side chain. Therefore, we tested the performance of AMBER* in reproducing the *ab initio* energy differences reported for the four relevant cases collected in Table 2 (1-propanol and 2-butanol; see Chart 3) and Table 3 (2-ethyltetrahydropyran and 3-hydroxy-2-ethyltetrahydropyran; see Chart 4). While the all-atom set of AMBER* parameters (Tables 2 and 3, entries 2 and 6) seems to perform reasonably well, the united atom data (Tables 2 and 3, entries 3 and 7) are consistently overstabilizing the *anti* conformation. We corrected this trend by including in the united atom force field the parameters $V_1 = -0.4$, $V_3 = 2.0$ kcal/mol for the C−C−C−O torsion (Chart 2, C glycoside substructure). The all atom parameters were left unchanged, but they had to be duplicated in the form of a substructure to avoid superseding by the

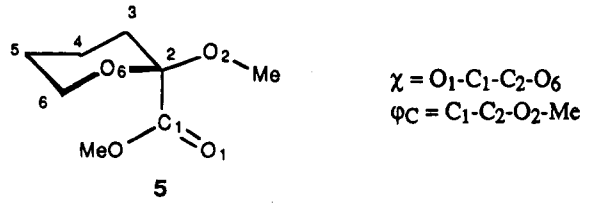
(16) a) Sabesan, S.; Bock, K.; Paulson, J. C. *Carb. Res.* **1991**, *218*, 27. (b) Sabesan, S.; Duus, J. Ø.; Bock, K.; Ludvigsen, S. *J. Am. Chem. Soc.* **1991**, *113*, 3236. (c) Ohrui, H.; Nishida, Y.; Hoh, H.; Meguro, H. *J. Org. Chem.* **1991**, *56*, 1726. (d) Ichikawa, Y.; Lin, Y.-C.; Dumas, D. P.; Shen, G.-J.; Garcia-Junceda, E.; Williams, M. A.; Bayer, R.; Ketcham, K.; Walker, L. E.; Paulson, J. C.; Wong, C.-H. *J. Am. Chem. Soc.* **1992**, *114*, 9283. (e) Acquotti, D.; Fronza, G.; Ragg, E.; Sonnino, S. *Chem. Phys. Lipids* **1991**, *59*, 107. (f) Sabesan, F.; Duus J. Ø.; Neira, S.; Domaille, P.; Kelm, S.; Paulson, J. C.; Bock, K. *J. Am. Chem. Soc.* **1992**, *114*, 8363. (g) Mukhopadhyay, C.; Bush, C. A. *Biopolymers* **1994**, *34*, 11.

(17) (a) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821. (b) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138.

(18) For general reviews on semiempirical methods, see: (a) Stewart, J. J. P. Semiempirical Molecular Orbital Methods. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers, New York, 1990; Vol. 1, Chapter 1. (b) Zerner, M. C. Semiempirical Molecular Orbital Methods. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991; Vol. 2, Chapter 8.

(19) Bernardi, A.; Raimondi, L. Unpublished results.

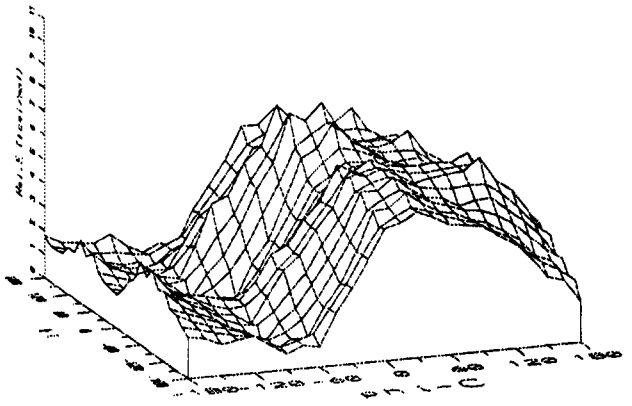
(20) Houk, K. N.; Eksterowicz, J. E.; Wu, Y.-D.; Fuglesang, C. D.; Mitchell, D. B. *J. Am. Chem. Soc.* **1993**, *115*, 4170.



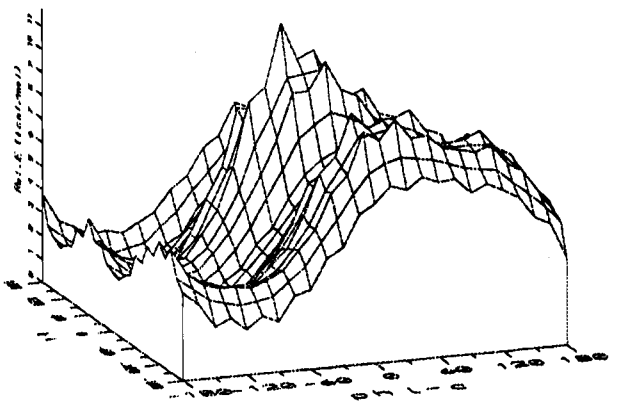
5

$\chi = O_1\text{-}C_1\text{-}C_2\text{-}O_6$

$\varphi_C = C_1\text{-}C_2\text{-}O_2\text{-}Me$



a. MNDO plot



b. Modified AMBER* plot

Figure 3. φ_C, χ Ramachandran plots calculated for 2-carbomethoxy-2-methoxytetrahydropyran (**5**).

Chart 2. New AMBER* Parameters in MacroModel Substructure Format[a]

```
-3
 C  C glycosides (ab initio data from JACS 4170(1993))
 9   CT-CT-CT-O3
-2
 4   1   2   3   4          -0.4000      0.0000      2.0000
-3
 C  C glycosides (ab initio data from JACS 4170(1993))
 9   C3-C3-C3-O3
-2
 4   1   2   3   4           0.0000      0.0000      0.1440
-3
 C  Ethandiol  (ab initio data from JACS 9620 (1993))
 9  O3-CT-CT-O3
-2
 4   1   2   3   4           0.4000      0.0000      2.0000
-3
 C  Ethandiol  (ab initio data from JACS 9620 (1993))
 9  O3-C3-C3-O3
-2
 4   1   2   3   4           0.7000      0.0000      2.0000
-3
 C  Neu5Ac anomeric torsion
 9  O3-CT(-C2(=O2))-O3-CT-CT-CT-CB-2
-2
 4  00   3   2  00           0.0000      0.0000      0.0000
 4   3   2   1  CT           0.6000      0.0000     -0.1000
 4   9   2   1  CT           0.0000      0.0000     -0.2000
 4   5   2   1  CT          -1.5000      0.0000      0.0000
```

[a] The first four substructures were introduced in AMBER* before the pyranose substructure, the last one afterwards.

united atom substructure. With this parameter a better fit was reached with *ab initio* data for 1-propanol, 2-butanol, and the 2-ethyltetrahydropyrans (Tables 2 and 3, entries 4 and 8). The new parameter was correctly read, as expected, for the sialic acid side chain as well as for all the ω torsions of **1b**.

Finally, a major point was addressed, concerning the preferred diol orientation in the sialic acid side chain. The high value of $J_{7,8}$ ($J \approx 9$ Hz; Figure 2) in the ^{1}H-NMR spectra of Neu5Ac itself[15] and of many sialyl oligosaccharides,[16] including GM1,[7] has been interpreted as arising from an antiperiplanar conformation for the $O\text{-}C_7\text{-}C_8\text{-}O$ dihedral angle in the side chain. On the other hand, it is well known that, for 1,2-diols, a *gauche* arrangement (*tG*g^- conformer) is strongly preferred over the corresponding antiperiplanar *tTt* one (Table 4),[21] and this preference is recognized by AMBER*,

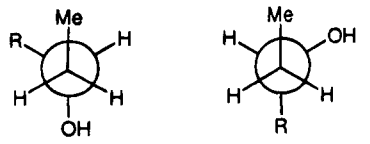
(21) (a) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 3892 and references therein. (b) Van Halsenoy, C.; Van Den Enden, L.; Schafer, L. *THEOCHEM* **1984**, *108*, 121. (c) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620.

Table 2. Relative Energies (kcal/mol) of *Anti* and *Gauche* Conformations of 1-Propanol and 2-Butanol (See Chart 3)

entry	R	*anti*	*gauche*
	R = H		
1	RHF/6-31G*[a]	0.3	0.0
2	AMBER*[b]	0.2	0.0
3	AMBER*[c]	0.0	0.4
4	modif.AMBER*[c,d]	0.2	0.0
	R = Me		
5	RHF/6-31G*[a]	0.9	0.0
6	AMBER*[b]	0.6	0.0
7	AMBER*[c]	0.7	0.0
8	modif.AMBER*[c,d]	1.2	0.0

[a] Reference 20. [b] All atoms. [c] United atoms. [d] Force field modified with inclusion of substructures reported in Chart 2.

Chart 3. *Anti* and *Gauche* Conformations of 1-Propanol and 2-Butanol



anti *gauche*

R = H 1-Propanol

R = Me 2-Butanol

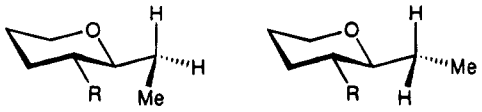
as well as by most other molecular mechanics force fields. In fact, the energy difference between the two conformations of 1,2-ethanediol as calculated by AMBER* (Table 4, entries 4 and 5) was definitely too high when compared with the MP3/6-31+G* value (Table 4, entry 3). A third substructure (Chart 2, 1,2-diol substructure) was thus introduced in AMBER* with parameters designed to reduce this difference according to the *ab initio* data. The values selected for the torsional parameters were $V_1 = 0.4$, $V_3 = 2.0$ kcal/mol for the united atoms $O\text{-}C\text{-}C\text{-}O$ torsion and $V_1 = 0.7$, $V_3 = 2.0$ kcal/mol for the all-atom corresponding torsion. In this way the agreement with the *ab initio* data was improved (Table 4, entries 6 and 7), thus reducing the destabilization of the *tTt* conformation,

Table 3. Relative Energies (kcal/mol) of *Anti* and *Gauche* Conformations of 2-Ethyltetrahydropyrans (See Chart 4)

entry	R	*anti*	*gauche*
	R = H		
1	RHF/6-31G*//RHF/3-21G[a]	0.7	0.0
2	AMBER*[b]	0.7	0.0
3	AMBER*[c]	0.4	0.0
4	modif.AMBER*[c,d]	0.7	0.0
	R = OH		
5	RHF/6-31G*//RHF/3-21G[a]	2.2	0.0
6	AMBER*[b]	2.0	0.0
7	AMBER*[c]	2.0	0.0
8	modif.AMBER*[c,d]	2.2	0.0

[a] Reference 20. [b] All atoms. [c] United atoms. [d] Force field modified with inclusion of substructures reported in Chart 2.

Chart 4. 2-Ethyltetrahydropyrans



R = H, OH

Table 4. Relative Energies (kcal/mol) of *tGg*$^-$ and *tTt* Conformations of Ethane-1,2-diol

entry		*tGg*$^-$	*tTt*
1	RHF/4-21G//RHF/4-21G[a]	0.0	2.6
2	MP2/6-31G*//RHF/6-31G*[a]	0.0	3.4
3	MP3/6-31+G*//RHF/6-31G*[a]	0.0	2.7
4	AMBER*[b]	0.0	4.8
5	AMBER*[c]	0.0	4.0
6	modif.AMBER*[b,d]	0.0	2.7
7	modif.AMBER*[c,d]	0.0	2.8

[a] Reference 21. [b] All atoms. [c] United atoms. [d] Force field modified with inclusion of substructures reported in Chart 2.

which is indicated by NMR experiments as the favored one for the sialic acid side chain.[7,15]

Results and Discussion

The force field AMBER* is known to be very effective for modeling peptides and proteins, and it has recently been augmented of parameters for pyranoses, based on the work of Homans.[12] Since our final goal is modeling the GM1/LT complex, AMBER* appeared to be the force field of choice for our calculations. However, GM1 contains a Neu5Ac residue, and, as far as we know, AMBER* has never been tested in the case of sialic acid derivatives. On the other hand, NMR studies on Neu5Ac itself (**2**, R = H),[15] as well as on a large number of gangliosides[7] and sialyl oligosaccharides,[16] have indicated some typical features of the Neu5Ac moiety that ought to be reproduced by the calculation. In particular, the NMR data show a large presence in solution of conformers in which the sialic acid anomeric substituent is almost antiperiplanar to the carboxy group. This means that the anomeric torsion angle φ_C (C_1–C_2–O_2–R in Figure 2) of sialic acid assumes values between ±120° and ±170°, rather than the −60° value that would normally be found for aldopyranosides. Additionally, the coupling constants observed for the glycerol side chain both in Neu5Ac[15] (Figure 2, R = H) and in GM1[7] have been interpreted as arising from an antiperiplanar disposition of H_7 and H_8 (Figure 2, $J_{7,8}$ = 9–10 Hz, θ_2 = H–C_7–C_8–H = 180°). This is also an unusual feature, since simple 1,2-diols are usually found to adopt a *gauche* conformation by computational and experimental means.[21]

A pseudosystematic Monte Carlo search performed on Neu5AcOMe **2** (Figure 2) using AMBER* as implemented in MacroModel 3.5x revealed that neither of these features were reproduced correctly by the calculation. The calculated minimum energy conformation showed a *gauche* relationship both between the carboxylate and the methoxy group (φ_C = −54°) and between H_7 and H_8 (θ_2 = 48°). The first conformer featuring φ_C = 180° was found to be 4.3 kcal/mol higher in energy than the global minimum, and the first one with an *anti* relationship between H_7 and H_8 was 4.8 kcal/mol higher. Therefore, a set of parameters for the sialic acid moiety was developed as described in detail in the Computational Section (see above). The new parameters were based on molecular orbital calculations on simple model systems and were introduced in AMBER* in the form of the substructures collected in Chart 2 in MacroModel format. The fifth substructure is used to reproduce the sialic acid anomeric torsion, while the remaining ones are meant to improve the description of the side chain.

With the foregoing new parameters a second MC/EM analysis of Neu5AcOMe **2** was undertaken. Ten conformations were found within 2.4 kcal/mol from the global minimum. They all had −97° ≥ φ_C ≥ −124°, in agreement with the experimental data.[15] The side chain appeared to be more flexible. A *gauche* arrangement of the 7,8-diol moiety is still calculated to be favored. However, as expected, the new parameters stabilize the *anti* conformation, which is now only 2.9 kcal/mol above the global minimum, as compared to 4.8 kcal/mol calculated with the default AMBER* parameters.

Although it would obviously be possible to force the side chain parameters in order to impose antiperiplanarity of the C_7–C_8 fragment, the available NMR data give no quantitative information on the population distribution and, therefore, no means to calibrate the parameters. Furthermore, the fact that our current set accurately reproduces relevant model systems in the gas phase suggests that the reason for the discrepancy with the Neu5Ac NMR data may reside in problems with the solvation model.[22,26] The minimum energy conformations which satisfy the NMR requirements (θ_2 = 180° ± 30°) could be calculated by imposing a torsional constraint during the MC/EM search. In the MacroModel framework, this can also be done by filtering the output files of the unbound search (3DSch facility) to screen out all the conformations which do not satisfy the constraints. In this way, 14 conformations (all with −95° ≥ φ_C ≥ −105° and θ_2 = 180° ± 30°) were found within 2.4 kcal/mol from the new global minimum.

Results of the Conformational Analysis of GM1. The Pseudosystematic Monte Carlo search on ganglioside GM1 pentasaccharide **1b** was performed in GB/SA water using the AMBER* united atom force field, augmented of the parameters described above. After 14 000 steps, 3385 different structures were retained within 12 kcal/mol from the global minimum. To these conformers hydrogens were added and the resulting structures were reminimized in GB/SA water using the all-atom AMBER*

(22) A 500 ps mixed-mode molecular dynamics simulation (see ref 23) of Neu5AcOMe **2** run with our force field in MacroModel 4.5 confirmed that the population distribution of θ_2 is mostly gauche (+g: 66%; −g: 32%; ap: 2%).

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

Table 5. Minimum Energy Conformations of GM1 Pentasaccharide within 1.5 kcal/mol from the Global Minimum of the Unbound MC/EM Search

	rel E (kcal/mol)	φ_{IV-III}	ψ_{IV-III}	φ_{III-II}	ψ_{III-II}	φ_{II-I}	ψ_{II-I}	φ_{N-II}	ψ_{N-II}	θ_1	θ_2
CONF 1	0.0	50	−6	23	34	−31	−30	−158	−29	−68	−58
CONF 2	0.5	51	−5	23	34	−30	−30	−158	−30	−68	−58
CONF 3	0.8	49	6	18	29	174	−1	−149	−17	−76	54
CONF 4	0.9	51	−3	21	29	−30	−28	−72	7	−63	−52
CONF 5	1.1	51	−5	24	34	26	−59	−158	−27	−68	−58
CONF 6	1.1	51	1	21	30	−30	−30	−154	−25	−68	−59
CONF 7	1.2	50	−2	21	30	−30	−30	−155	−25	−68	−58
CONF 8	1.2	51	0	21	29	−30	−30	−156	−24	−79	55
CONF 9	1.5	51	2	20	29	174	−1	−155	−28	−56	−174
CONF 10	1.5	54	19	20	27	−30	−30	−161	−21	−65	−57
expl[a]		25−30	flex.[b]	30−25	30−25	flex.[c]	flex.[c]	−165	−18	−55[d]	−155[e]

[a] Reference 7. Values calculated on the basis of the NOE experiments. [b] Observed values = 30 and −40.[7] [c] Observed values (φ, ψ): 55, 0; 35, −50; 5, −30; 30, −170 (or 30, −5).[7] [d] Observed $J_{6,7}$ = 0 Hz.[7] [e] Observed $J_{7,8}$ = 8.5 Hz.[7]

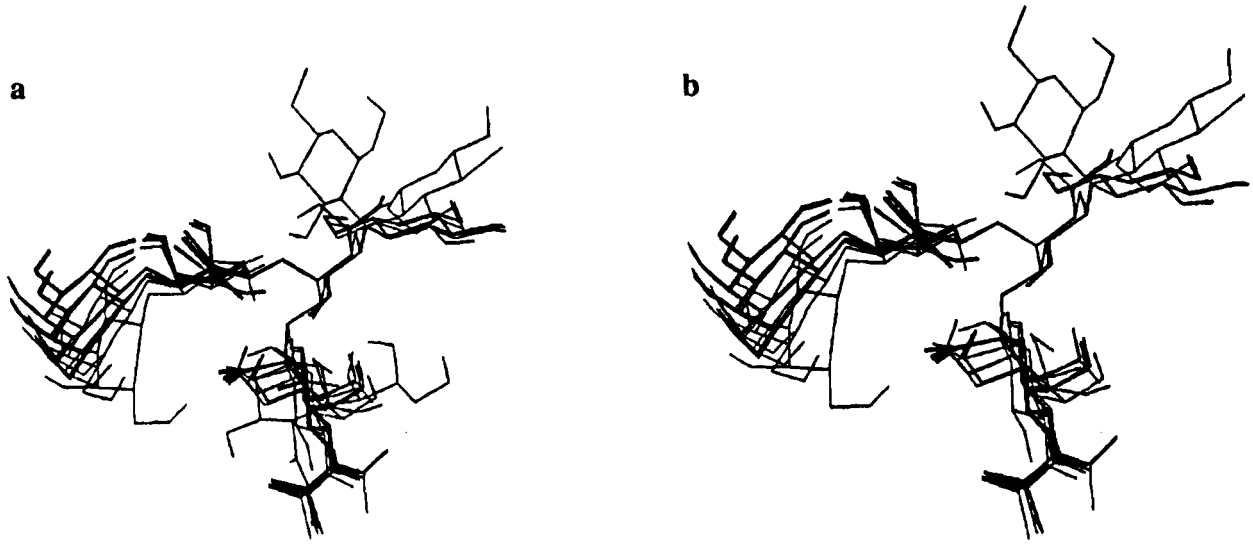


Figure 4. Superimposition plots of GM1 pentasaccharide conformations within 1.5 kcal/mol from global minimum. (a) Superimposition of first 10 conformers. (b) Superimposition of residue II for the first 10 conformers, excluding CONF4 (see text).

force field. Within a range of 1.5 kcal/mol from the global minimum only 10 conformations were found. All 10 structures present a similar arrangement of the pentasaccharide skeleton, as appears clearly by the values of the glycosidic torsion angles reported in Table 5 and by inspection of the superimposition plots of Figure 4. The only exception is the fourth conformer (CONF 4 in Table 5, ΔE = 0.9 kcal/mol), which features a *gauche* conformation for the Neu5Ac anomeric bond ($\varphi_C = -72°$, see Table 5). If this conformer is not considered (Figure 4b), the superimposition plot shows more clearly a common skeleton conformation.

In Figure 5 the global minimum conformation of the MC/EM search (CONF 1 of Table 5, shown in Figure 5a) is compared to the structure proposed for GM1 in water solution on the basis of NMR studies (shown in Figure 5b).[7,24] Superimposition of the two is reported in Figure 5c. As it can be seen, the calculated conformation of the pentasaccharide skeleton agrees very well with the experimentally determined one[7] (rms = 0.964 Å).

The glycosidic torsion angles determined by NOE experiments and those calculated for the ten lowest energy conformations are also compared in Table 5. Clearly, the calculated values are very close to the observed ones, and flexible areas are nicely distinguished from conformationally restricted ones. This is also apparent in Figure 4b, where flexibility is observed mainly at the Galβ1−4Glc linkage (φ_{II-I} and ψ_{II-I} in Table 5). NOE experiments led to the conclusion that at least four main conformations are significantly populated for this moiety.[7] Interestingly, the terminal galactose, which is believed to be involved in binding of LT[5b] and CT,[6] appears to be rather conformationally restricted in our calculations (Table 5, φ_{IV-III} and ψ_{IV-III}).

The most significant difference between calculated and experimental structures concerns the conformation of the sialic acid side chain (Table 5, θ_2). As we have mentioned, the NMR data show a large value for the $J_{7,8}$ vicinal coupling constant ($J_{7,8}$ = 8.5 Hz), which has been interpreted as originating from an antiperiplanar disposition of the two hydroxy groups at C_7(N) and C_8(N).[7,15,16] In our calculations, a *gauche* arrangement appears to be adopted by all conformers within 1.5 kcal/mol but CONF 9. This is not surprising in view of the results described above for Neu5AcOMe (**2**). However, it should be noted that in all 10 low energy conformations the OH_8(N) appears to be hydrogen-bonded to the carboxylate. This H-bond was hypothesized on the basis of NMR data[7] and generally related to the *anti* conformation of the C_7(N)−C_8(N) bond. Moreover, the *gauche* relationship between

(24) This last structure was obtained by graphical input of the conformation proposed in ref 7, followed by minimization in GB/SA water with AMBER*.

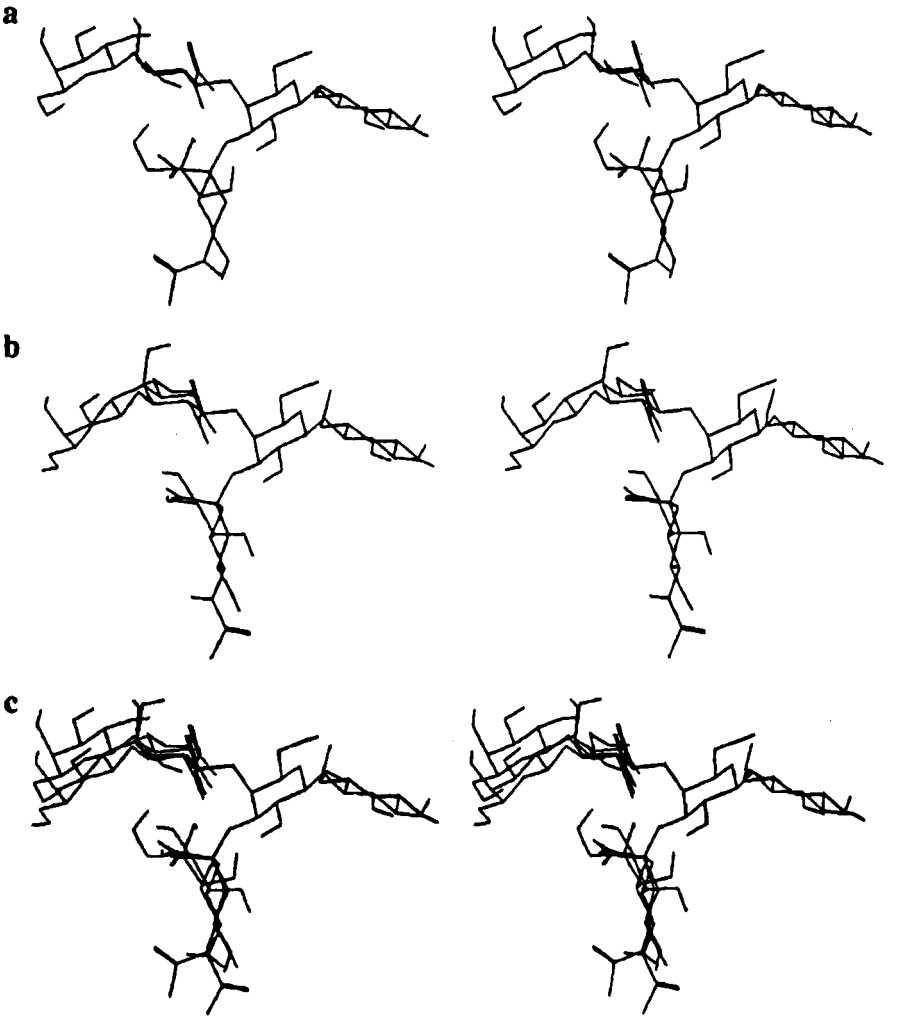


Figure 5. Stereoview of GM1 pentasaccharide. (a) Stereoview of GM1 pentasaccharide minimum energy conformation. (b) Stereoview of GM1 pentasaccharide NMR structure (from ref. 7). (c) Stereoview of superimposed **a** and **b**.

Table 6. Calculated and Experimental[a] Interresidue Distances (Å) for GM1

	H_1(IV)				H_1(III)		H_2(II)	H_{3ax}(N)
	H_2(III)	H_3(III)	H_4(III)	COMe(III)	H_4(II)	H_8(N)	COMe(III)	H_3(II)
calcd[b]	4.1	2.4	4.0	3.8	2.4	4.6 (2.8[c])	3.0	2.5
expl[a]	3.5	2.5	3.5	3.8	2.2	3.1	3.1	2.1

[a] Reference 7. Values calculated on the basis of the NOE experiments. [b] Boltzmann-weighted averages from unbound MC/EM search. [c] From CONF 9 (see text).

H_6(N) and H_7(N), which leads to a $J_{6,7} = 0$ Hz in the ^{1}H-NMR spectrum of GM1,[7] is well reproduced by our calculations (Table 5, θ_1). This means that the position of the chain relative to the Neu5Ac ring, which would be relevant in docking GM1 to the toxins, is also correct in the calculations.

Filtering of the output files to locate the minima which satisfy the NMR C_7–C_8 conformations led to four conformers within 2 kcal/mol from the new global minimum (CONF 9 of the unbound search). Their superimposed structures are reported in Figure 6, which shows the familiar disposition of the pentasaccharide skeleton.

The calculated and experimental (NOE-derived) interresidue distances are reported for comparison in Table 6. The calculated values are in excellent agreement with the experimental results. The only exception is represented by the H_1(III)–H_8(N) average distance, and this is again a product of the poor reproduction of the C_7(N)–

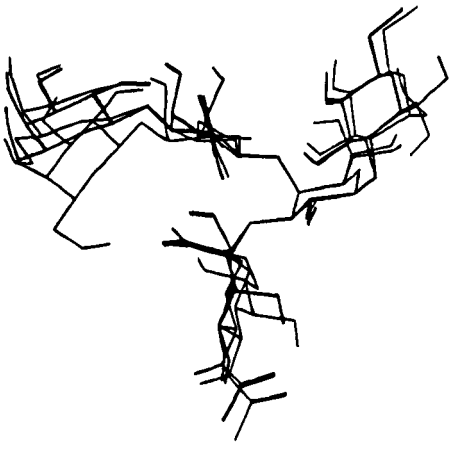


Figure 6. Superimposition plots of GM1 low energy conformations after θ_2 filtering.

$C_8(N)$ potential. In fact, if this distance is calculated from CONF 9 (the only low energy conformation which exhibits an *anti* relationship for the $C_7(N)-C_8(N)$ diol), the agreement with the experimental data is complete.

Conclusions

In this paper we have shown that an MC/EM conformational search of GM1 pentasaccharide **1b** led to results that are fully compatible with the known experimental data on **1b** in solution. The search was performed in GB/SA water, using the continuum solvent model included in Batchmin. The force field was AMBER* augmented of parameters for the Neu5Ac residue that were derived from molecular orbital calculations in ways that are described in the text. Agreement between calculated and experimental structures was very good, the only exception being constituted by the conformation of the $C_7(N)-C_8(N)$ linkage in the sialic acid side chain.

It is now possible to use the computational protocol described in this paper to perform a computational search of GM1 pentasaccharide in the binding site of its natural receptors LT and CT,[25] with some confidence of deriving a meaningful model for the complex.

Acknowledgment. The authors are grateful to Centro Interuniversitario Lombardo di Elaborazione Automatica (CILEA) for a generous gift of computer time.

Supplementary Material Available: RHF/3-21G-optimized geometries of 2-methoxy- and 2-methyl-2-methoxytetrahydropyrans (**3** and **4**); MNDO-optimized geometry of 2-carbomethoxy-2-methoxytetrahydropyran (**5**) (3 pages). This material is contained in libraries on microfiche, immediately follows this article in the microfilm version of the journal, and can be ordered from the ACS; see any current masthead page for ordering informations.

JO941155Q

(25) For an application of MC/EM techniques to conformational searches within an enzyme binding site, see: Guida, W. C.; Bohacek, R. S.; Erion, M. D. *J. Comput. Chem.* **1992**, *13*, 214.

(26) This hypothesis is strongly supported by a detailed study of GB/SA water solvation of 1,2-diols that has just appeared in the literature: Nagy, P. I.; Bitar, J. E.; Smith, D. A. *J. Comput. Chem.* **1994**, *15*, 1228.