Quantum Mechanical Analysis of 1,2-Ethanediol Conformational Energetics and Hydrogen Bonding

Olgun Guvench and Alexander D. MacKerell, Jr.*

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, 20 Penn Street, HSF II-629, Baltimore, Maryland 21201

Received: April 14, 2006

A proper understanding of the conformational energetics of 1,2-ethanediol (ethylene glycol) is important to the construction of molecular mechanics force fields for the treatment of carbohydrates since these biologically important molecules have a prevalence of vicinal hydroxyl groups. In the present study, quantum mechanical analysis of the 10 unique minimum-energy conformations of ethylene glycol is performed by using 10 model chemistries ranging from HF/6-311++G(d,p) up to a hybrid method that approximates CCSD(T)/cc-pVQZ. In addition, natural bond orbital (NBO) analysis of these conformations with deletion of pairings of CO bond/antibonding and lone pair/antibonding orbitals is used to investigate contributions from the "gauche" effect to ethylene glycol conformational energetics. MP2 with the "correlation consistent" basis sets and DFT/6-311++G(d,p) do the best job of matching the approximate CCSD(T)/cc-pVQZ energies while MP2/6-31G(d) and Hartree−Fock both fare poorly. NBO analysis shows the conformational energies to be independent of the deletion of matrix elements associated with (i) CO bonding and antibonding orbital interactions and (ii) lone pair and antibonding orbital interactions, whereas the energetic ordering correlates with geometric parameters consistent with internal hydrogen bonds. Thus, the present results suggest that standard molecular mechanics potential energy functional forms, which lack explicit terms to account for stereoelectronic effects, are appropriate for carbohydrates.

Introduction

Monosaccharides are now widely recognized as being the letters of a third alphabet of molecular recognition, with the first two alphabets being composed of the more widely studied amino acids and nucleic acids. With this realization have come efforts at developing molecular mechanics force fields optimized specifically for modeling carbohydrates,[1−12] with the ultimate aim of allowing for solvated simulations of biologically relevant polysaccharides, both alone and interacting with other biopolymers. Proper force-field-based treatment of the energetics of polysaccharides in such biological contexts requires a careful balance between intra- and intermolecular interactions. A first, crucial step to the development of such a force field is an accurate understanding of physical phenomena responsible for the intramolecular energetics of carbohydrates.

Unlike amino acids and nucleic acids, monosaccharides pose a particular challenge because of the many possibilities for intramolecular hydrogen bonding. For example, glucose in the six-membered ring form D-glucopyranose, whose polymeric derivatives occur in such biologically important contexts as celluose and starch as well as glycosylated proteins in eukaryotes[13,14] and the cell wall of disease-causing prokaryotes,[15] contains three pairs of vicinal hydroxyl groups. The thermodynamics of D-glucopyranose overwhelmingly favor the chair conformation with the hydroxymethyl moiety in the axial position.[16] This conformation in turn places all of these hydroxyl groups in a gauche conformation relative to each other, thereby allowing the possibility of hydrogen bonding between all three pairs (Figure 1). In the context of force field development, the dependence of the conformational energetics on such intramolecular hydrogen bonding must be well-represented by the molecular mechanics nonbonded terms so that when the molecule is simulated in an aqueous environment with and without other biopolymers, changes in hydrogen bonding as a function of environment are captured.

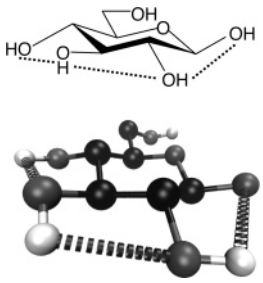

Figure 1. Intramolecular hydrogen bonding in β-D-glucopyranose. β-D-glucopyranose and α-D-glucopyranose (which has the C1 hydroxyl in the axial position) can both form three intramolecular hydrogen bonds among the four annular hydroxyls. CPK representation prepared with VMD.[36]

1,2-Ethanediol (ethylene glycol) holds a special place in the study of carbohydrates as it is the smallest molecular fragment containing a pair of vicinal hydroxyl moieties. With three torsional degrees of freedom, this molecule assumes 10 unique minimum energy conformations after accounting for symmetry. The conformations are commonly labeled using the trans (*t*), gauche (*g*), and gauche minus (*g′*) nomenclature to describe the values of the dihedral angles. Thus, the *tGg′* conformation has values of ∼180°, ∼60°, and ∼−60° for its HOCC, OCCO and CCOH dihedral angles, respectively. The earliest gas-phase quantum mechanical (QM) calculations on the full set of 10

* Address correspondence to this author. Phone: 410-706-7442. Fax: 410-706-5017. E-mail: alex@outerbanks.umaryland.edu.

10.1021/jp0623241 CCC: $33.50
Published on Web 07/20/2006

minimum energy conformations employed the HF/4-21G model chemistry and revealed the most stable two conformations to be *tGg′* and *gGg′*.[17] A subsequent pair of studies on the full set of 10 applied electron correlation methods that included MP4/6-311G(d,p)//MP2/6-31G(d,p)[18] and an approximation to CCSD-(T)/cc-pVTZ.[19] These calculations showed the *tGg′*, *gGg′*, and *g′Gg′* conformations to be within ~1.5 kcal/mol of each other in energy while the fourth-lowest energy conformation was at ~3 kcal/mol. The current study increases the level theory to the approximate CCSD(T)/cc-pVQZ level; full optimizations of all 10 conformations are also performed by using nine model chemistries that range from HF/6-311++G(d,p) up to MP2/cc-pVTZ. Additionally, natural bond orbital (NBO) analysis,[20] with the HF/cc-pVTZ and B3LYP/6-311++G(d,p) model chemistries, involving deletion of matrix elements corresponding to the CO bonding/antibonding orbital as well as lone pair/antibonding orbital interactions expands upon recent work that investigated stereoelectronic contributions to gauche stabilization through analysis of NBO $E^{(2)}$ transfer energies.[21]

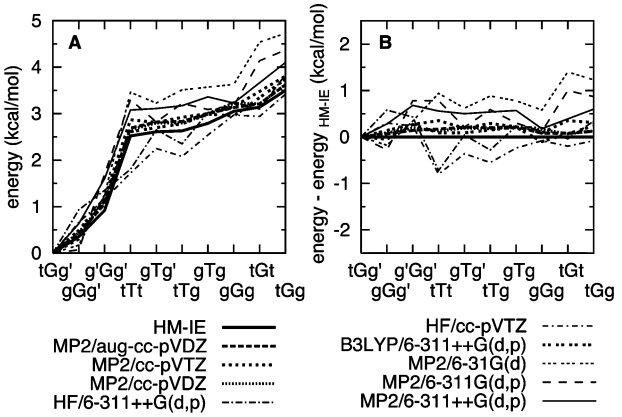

Figure 2. Ethylene glycol (A) energies and (B) energies relative to HM-IE. All values for a particular model chemistry are after geometry optimization with that model chemistry. HM-IE approximate CCSD-(T)/cc-pVQZ energies are calculated from MP2/cc-pVDZ optimized conformations.

Methods

Initial conformations for the 10 ethylene glycol conformations were built and molecular-mechanics force-field minimized with version c32b2 of the program CHARMM[22,23] and force field parameters derived from the CHARMM molecular mechanics force field.[24,25] All quantum chemical calculations were performed with the Gaussian03 package.[26] Unconstrained minimizations starting from the force-field-minimized conformations were performed by using the following model chemistries: HF/6-311++G(d,p), HF/cc-pVTZ, B3LYP/6-311++G(d,p), MP2/6-31G(d), MP2/6-311G(d,p), MP2/6-311++G(d,p), MP2/cc-pVDZ, MP2/aug-cc-pVDZ, and MP2/cc-pVTZ. Additionally, a hybrid method, denoted HM-IE, was applied to approximate energies at the CCSD(T)/cc-pVQZ level, as successfully applied in a recent optimization of alkane dihedral parameters.[27] This method involves first a MP2/cc-pVDZ minimization. The HM-IE energy is then calculated by using single-point energies of the optimized structure and combined with the formula CCSD-(T)/cc-pVDZ + MP2/cc-pVQZ − MP2/cc-pVDZ.[28]

Natural bond orbital (NBO) analysis[20] with the NBO software version 3.1[29] as implemented in Gaussian03 was performed on all 10 conformations with both the HF/cc-pVTZ and B3LYP/6-311++G(d,p) model chemistries. Energies of the optimized conformations were calculated by using all the NBOs and after deletion of the off-diagonal matrix elements of the effective one-electron Hamiltonian (the Fock matrix for HF and the Kohn−Sham matrix for B3LYP) corresponding to pairings of CO bonding and all antibonding orbitals or lone pair and all antibonding orbitals. The NBO deletion energy was calculated as the difference of the total energies before and after matrix-element deletion.

Results and Discussion

The geometry-optimized energies of the 10 unique ethylene glycol conformations using the various model chemistries are shown diagrammatically in Figure 2 and are listed in Table 1 for quantitative comparison. The general trend is the same for all combinations of theory and basis set in this study and also when compared to prior work.[17−19,21] After rank-ordering the conformations by using the HM-IE approximate CCSD(T)/cc-pVQZ energy, the lowest energy conformation is *tGg′*, followed by *gGg′* and *g′Gg′*. Depending on the model chemistry, these three low-energy conformations are within 0.92 (HM-IE) to 1.70 kcal/mol (MP2/6-311G(d,p)) of each other, and all model chemistries rank these three in the same order. Common to these three conformations is the gauche OCCO dihedral angle, which allows for close enough approach of the vicinal hydroxyl moieties to support intramolecular hydrogen bonding, as discussed in detail below. In contrast, the next four lowest energy conformations as rank-ordered with HM-IE have the OCCO dihedral in the trans geometry, thereby precluding such an interaction. With this model chemistry, these four conformations span only a 0.27 kcal/mol range of energy and are well separated from *g′Gg′* by 1.60 kcal/mol. The three highest energy conformations are all greater than 3 kcal/mol higher in energy than the global minimum with use of HM-IE, and though all three have the OCCO dihedral gauche, their HOCC dihedrals are in conformations inconsistent with intramolecular hydrogen bonding, as discussed below.

While all the model chemistries tested show the same general behavior in terms of being able to select the *tGg′*, *gGg′*, and *g′Gg′* as the minimum energy conformations and also rank these three in the same order, there are noticeable energetic differences between the various methodologies for the remaining conformations. Both sets of HF data, one using the 6-311++G(d,p) and the other the cc-pVTZ basis set, systematically underestimate the energies of the *tTt* conformations by 0.82 and 0.90 kcal/mol, respectively. This in turn obscures the clear energetic and conformational separation between the three low-energy conformations and the other seven. Poor results are also seen for MP2 with the small 6-31G(d) basis set: while accounting for the energetics of the three low-energy conformations, its chief shortcoming is the overestimation of the energies of the seven high-energy conformations by an average of 0.92 kcal/mol compared to the HM-IE results, and in particular the *tGt* and *tGg*, by 1.40 and 1.23 kcal/mol, respectively. Increasing the basis to 6-311G(d,p) helps to partially mitigate this overestimation, with the average error ($\sum$(energy$_i$ − energy$_{\mathrm{HM-IE},i}$)/n) for these seven being +0.55 kcal/mol and the *tGt* and *tGg* energies too large by 1.00 and 0.87 kcal/mol, respectively. Employing the larger 6-311++G(d,p) does reduce the maximum error to 0.68 kcal/mol. However, the best MP2 energies relative to the HM-IE approximate CCSD(T)/cc-pVQZ results are obtained with the "correlation-consistent" basis sets[30] cc-pVDZ, aug-cc-pVDZ, and cc-pVTZ and are within 0.35 kcal/mol for all conformations. As a point of comparison, for the representation of ethylene glycol the widely used 6-31G(d) basis set employs 72 basis functions as compared to 108 for 6-311G(d,p), 130 for 6-311G++(d,p), 86 for cc-pVDZ, 146 for aug-cc-pVDZ, 204

TABLE 1: Optimized Energy (in kcal/mol) as a Function of Model Chemistry and Conformation

	HF/6 -311++G(d,p)	HF/ cc-pVTZ	B3LYP/ 6-311++G(d,p)	MP2/ 6-31G(d)	MP2/ 6-311G(d,p)	MP2/ 6-311++G(d,p)	MP2/ cc-pVDZ	MP2/ aug-cc-pVDZ	MP2/ cc-PVTZ	HM-IE
tGg′	0	0	0	0	0	0	0	0	0	0
gGg′	0.93	0.66	0.49	0.17	0.07	0.63	0.44	0.42	0.30	0.35
g′Gg′	1.35	1.16	1.19	1.26	1.70	1.60	1.11	1.06	1.19	0.92
tTt	1.80	1.72	2.61	3.46	3.30	3.08	2.65	2.69	2.87	2.52
gTg′	2.65	2.25	2.70	3.23	2.82	3.11	2.80	2.82	2.81	2.61
tTg	2.35	2.08	2.78	3.51	3.22	3.17	2.80	2.83	2.92	2.63
gTg	3.02	2.54	2.99	3.57	3.09	3.36	2.99	3.00	3.00	2.79
gGg	3.25	2.97	2.95	3.63	3.17	3.23	3.22	3.13	3.15	3.05
tGt	3.10	2.94	*a*	4.54	4.14	*a*	3.21	3.21	3.48	3.14
tGg	3.84	3.42	3.74	4.73	4.37	4.10	3.63	3.62	3.81	3.50

[a] Geometry optimization yields the *g′Gg′* conformation.

TABLE 2: Optimized OCCO Dihedral Angles (in deg) as a Function of Model Chemistry

	HF/ 6-311++G(d,p)	HF/ cc-pVTZ	B3LYP/ 6-311++G(d,p)	MP2/ 6-31G(d)	MP2/ 6-311G(d,p)	MP2/ 6-311++G(d,p)	MP2/ cc-pVDZ	MP2/ aug-cc-pVDZ	MP2/ cc-pVTZ	HM-IE
tGg′	62.0	62.0	62.0	59.9	61.0	61.9	59.5	62.2	60.6	59.5
gGg′	59.9	59.4	58.2	55.4	57.0	59.0	54.9	57.6	56.3	54.9
g′Gg′	60.0	60.2	60.6	55.3	56.8	58.9	55.7	60.6	59.0	55.7
tTt	179.9	180.0	180.0	179.9	−179.9	179.3	180.0	180.0	180.0	180.0
gTg′	−179.9	−180.0	180.0	−180.0	180.0	−180.0	180.0	−180.0	−180.0	180.0
tTg	−180.0	179.8	180.0	179.8	179.1	−179.0	178.8	179.6	179.6	178.8
gTg	177.6	177.4	176.6	176.8	175.7	176.7	175.4	176.6	176.7	175.4
gGg	54.3	54.4	53.7	45.1	51.4	54.1	46.0	54.0	51.8	46.0
tGt	74.1	73.2	*a*	72.7	74.5	*a*	72.5	74.0	74.4	72.5
tGg	65.7	65.4	66.4	62.9	62.9	65.3	62.6	65.3	65.2	62.6

[a] Geometry optimization yields the *g′Gg′* conformation.

for cc-pVTZ, and 400 for cc-pVQZ as used in the HM-IE approximate CCSD(T)/cc-pVQZ calculations. The B3LYP/6-311++G(d,p) DFT results are quite good, with a maximum error of 0.27 kcal/mol. Though DFT is known to suffer inaccuracies in modeling weak dispersive interactions,[25,31−35] the ethylene glycol energetics are most likely dominated by electrostatics in the form of "hydrogen bonding" due to the vicinal hydroxyl groups, hence the favorable results relative to HM-IE. With the 6-311++G(d,p) basis set, both MP2 and B3LYP optimization of the *tGt* conformation leads to conversion to the *g′Gg′* conformation. Thus, the choice of basis set for this system affects not only the relative energetics of the conformations, but also the shape of the potential energy surface such that *tGt* is not always a stable minimum.

The optimized OCCO dihedral values of the various conformations and using each of the various methodologies are listed in Table 2. In general, the results using the larger basis sets are similar regardless of the level of theory. In contrast, the two smallest basis sets, 6-31G(d) and cc-pVDZ, demonstrate similar systematic errors. In particular, for all conformations that contain a gauche OCCO conformation, the MP2 optimizations using either of these two basis sets yield values too small relative to optimizations using the larger basis sets. This systematic error is particularly pronounced for *g′Gg′*, where these two smaller basis sets give values of 55.3° and 55.7°, compared to an average of 59.4° with a standard deviation of 1.4° for the other model chemistries. Similarly, for the *gGg* conformation, the MP2/6-31G(d) and MP2/cc-pVDZ optimized values are 45.1° and 46.0°, respectively, as compared to an average and standard deviation of 53.4° and 1.3° for the other model chemistries. Thus, the MP2/6-31G(d) optimizations yield poor energies relative to the HM-IE approximate CCSD(T)/cc-pVQZ method and poor geometries relative to optimizations using any of the HF, B3LYP, or MP2 levels of theory with the larger basis sets. In contrast, though the MP2/cc-pVDZ calculations suffer from the same type of geometric inaccuracy as MP2/6-31G(d), the MP2/cc-pVDZ energies are very good relative to the HM-IE method, having an average error of +0.13 kcal/mol with a standard deviation of 0.13 kcal/mol among the errors. Such energetic accuracy is as good as that of MP2/aug-cc-pVDZ and B3LYP/6-311++G(d,p) calculations and slightly better than that of the MP2/cc-pVTZ model chemistry. It is important to note that the HM-IE energies are defined by using MP2/cc-pVDZ optimized structures. Given the close agreement between the MP2/cc-pVTZ-optimized energies and the HM-IE energies as calculated after optimization with MP2/cc-pVDZ, it is apparent that the choice of basis set has a large effect on determining the energetics whereas the energetics are robust to variations in the OCCO dihedral for a particular conformation. Indeed, a final piece of evidence for the robustness of the energies relative to variation in the OCCO dihedral is that the RMS energetic error of the MP2/cc-pVTZ energies of the 10 structures optimized at the MP2/6-31G(d) level (MP2/cc-pVTZ//MP2/6-31G(d); individual data not shown) compared to the MP2/cc-pVTZ-optimized structures is 0.04 kcal/mol, with a maximum error of +0.11 kcal/mol (conformation *gGg*).

Both a geometric and an NBO analysis were performed with the goal of explaining the energetic ordering of the 10 unique ethylene glycol minimum energy conformations. The geometric analysis involved investigating geometric parameters consistent with a hydrogen bonding interaction or unfavorable electrostatic interaction between the vicinal hydroxyl moieties. In particular, the parameters were O···H distances and O−H···O angles, as well as O···O and hydroxyl H···hydroxyl H distances. NBO analysis involved calculations of energies to identify stereoelectronic contributions to the energetic ordering. The geometric and NBO analysis used conformations optimized with the HF/cc-pVTZ and B3LYP/6-311++G(d,p) methods. These two model chemistries were used not only due to the large basis sets necessary to get accurate OCCO dihedral geometries but also their one-electron Hamiltonians, which allow for NBO analysis of the energetic effect of deleting off-diagonal matrix elements and which is not possible with MP2. After transformation to the NBO basis, such elements correspond to interactions

TABLE 3: Geometric Parameters, Relative Energies, and NBO Deletion Energies of Optimized Geometries for (A) HF/cc-pVTZ and (B) B3LYP/6-311++G(d,p)[a]

(A) HF/cc-pVTZ

	interaction 1		interaction 2		repulsive interactions		relative energy	NBO deletion energy	
	H···O	O−H···O	H···O	O−H···O	O···O	Ho···Ho		$\sigma_{CO} \rightarrow \sigma^*$	LP $\rightarrow \sigma^*$
tGg′	3.61	26.7	2.42	104.2	2.81	3.20	0.00	2.87	41.33
gGg′	3.31	53.3	2.43	107.2	2.85	2.89	0.66	2.77	43.80
g′Gg′	2.76	87.2	2.76	87.2	2.87	2.57	1.16	2.65	41.12
conformations lacking a hydrogen bond									
tTt	4.27	36.8	4.27	36.8	3.56	5.05	1.72	3.59	38.56
gTg′	3.95	64.5	3.95	64.5	3.65	4.44	2.25	3.42	42.77
tTg	4.32	35.7	3.92	63.3	3.60	4.59	2.08	3.50	40.75
gTg	3.95	64.7	3.95	64.8	3.65	4.12	2.54	3.41	43.45
gGg	2.98	75.4	2.98	75.4	2.89	3.33	2.97	2.30	46.82
tGt	3.64	33.2	3.64	33.2	2.90	4.44	2.94	2.59	38.54
tGg	3.75	23.1	3.23	61.9	2.91	3.93	3.42	2.38	42.76

(B)B3LYP/6-311++G(d,p)

	interaction 1		interaction 2		repulsive interactions		relative energy	NBO deletion energy	
	H···O	O−H···O	H···O	O−H···O	O···O	Ho···Ho		$\sigma_{CO} \rightarrow \sigma^*$	LP $\rightarrow \sigma^*$
tGg′	3.62	29.4	2.39	106.4	2.82	3.16	0.00	9.32	80.28
gGg′	3.33	53.8	2.40	109.5	2.87	2.87	0.49	8.85	82.38
g′Gg′	2.79	87.9	2.79	87.9	2.92	2.59	1.19	8.32	76.19
conformations lacking a hydrogen bond									
tTt	4.31	37.9	4.32	37.9	3.61	5.11	2.61	8.99	75.59
gTg′	3.99	66.2	3.99	66.3	3.71	4.46	2.70	8.80	74.14
tTg	4.38	36.7	3.95	65.0	3.65	4.62	2.78	8.85	75.23
gTg	3.99	66.2	3.99	66.2	3.71	4.15	2.99	8.85	75.26
gGg	3.02	76.1	3.02	76.1	2.94	3.36	2.95	7.69	84.17
tGt	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*
tGg	3.80	24.8	3.26	63.1	2.96	3.96	3.74	8.23	81.24

[a] Distances are in Å, angles in deg, and energies in kcal/mol. [b] Geometry optimization yields the *g′Gg′* conformation.

between "natural" basis functions that represent, for example, lone pairs, bonding orbitals, and antibonding orbitals. Thus, such a deletion analysis in the context of NBO can be used to calculate the energetic contribution from the interaction between natural orbitals.

Table 3 lists the geometric parameters for the (A) HF/cc-pVTZ and (B) B3LYP/6-311++G(d,p)-optimized conformations, along with the relative energies and NBO deletion energies for these respective model chemistries calculated at the optimized conformations. The following discussion of geometries is limited to the HF/cc-pVTZ representation since the geometric parameters are nearly identical in the B3LYP/6-311++G(d,p) representation, with the distances being within hundredths of ångstroms and the angles within a few degrees of the HF/cc-pVTZ values. The three low-energy conformations, *tGg′*, *gGg′*, and *g′Gg′*, all have interaction geometries consistent with intramolecular hydrogen bonding. The global energy minimum *tGg′* has hydroxyl moieties posed such that one of the two H···O distances is 2.42 Å and the corresponding O−H···O angle is 104.2°. Similarly, the *gGg′* conformation, which is ranked second lowest in energy by all 10 of the tested model chemistries, has these values at 2.43 Å and 107.2°, respectively. The geometric parameter that appears to correlate with the energetic ordering of the three low-energy conformations is the repulsive hydroxyl H···hydroxyl H distance. In the *gGg′* conformation, these hydrogen atoms, which carry partial positive charges since they are bonded to the electronegative oxygens, are 0.31 Å closer than for *tGg′* and are thus expected to contribute more unfavorably to the *gGg′* energy than the *tGg′* energy. The third low-energy conformation, which in all but the HF theory is energetically well-separated from the next seven conformations, has the H···O and O−H···O values at 2.76 Å and 87.2°. Though both this distance and angle are not as favorable as the other two low-energy conformations with regard to the formation of a hydrogen bond, due to internal symmetry there are two identical such interactions in this conformation. Nevertheless, the poorer hydrogen bonding geometry and the close H···H distance of 2.57 Å correlate with this being the highest energy conformation among the three low-energy conformations that have geometric parameters consistent with intramolecular hydrogen bonding.

For six of the seven high-energy conformations, namely *tTt*, *gTg′*, *tTg*, *gTg*, *tGt*, and *tGg*, intramolecular hydrogen bonding is precluded because of long H···O distances and/or small O−H···O angles. Among these six, *tGg* has the shortest (best) H···O distance at 3.23 Å but with a poor angle of 61.9°. The largest (best) angle is in *gTg*, but it is still quite poor for hydrogen bond formation with a value of 64.8°. Additionally, the trans OCCO dihedral in this conformation prohibits hydrogen bond formation and the corresponding H···O distance is a large (poor) 3.95 Å. The seventh of the set, *gGg*, has both distances and angles of 2.98 Å and 75.4° due to internal symmetry. Though these values are better than any of the other conformations in this set of seven, they are worse than those for the *g′Gg′* conformation that belongs to the set of three low-energy conformations and has values of 2.76 Å and 87.2°.

While the energetic difference between the *g′Gg′* and *gGg* appears to be the better hydrogen-bond geometries for *g′Gg′*, another possibility may be the energetic consequences of stereoelectronic effects. This possibility was tested with NBO deletion analysis. In particular, the elements of the Fock or Kohn−Sham matrices corresponding to the CO bonding orbital and all antibonding orbital interactions ($\sigma_{CO} \rightarrow \sigma^*$) or to lone pair and all antibonding orbital interactions (LP $\rightarrow \sigma^*$) were deleted and the energies compared before and after deletion. If stereoelectronic effects account for the observed energetic

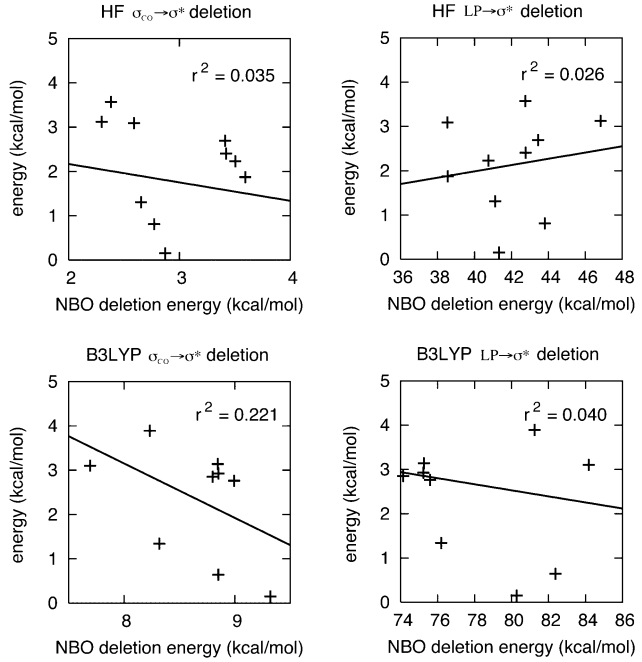

Figure 3. Relative energies as a function of NBO deletion energies.

difference between the two conformations, the deletion energies would be larger for *g′Gg′* than for *gGg*. This is indeed the case for the $\sigma_{CO} \rightarrow \sigma^*$ deletion with deletion energies of 2.65 and 2.30 kcal/mol as compared to relative energies of 1.16 and 2.97 kcal/mol in the case of HF/cc-pVTZ and 8.32 and 7.69 kcal/mol for the deletion energies and 1.19 and 2.95 kcal/mol for the relative energies in the case of B3LYP/6-311++G(d,p). However, the case is weakened by the fact that the trend is opposite for LP $\rightarrow \sigma^*$ deletion, with the deletion energies being larger for the higher energy *gGg* conformation than for *g′Gg′*. In the case of HF/cc-pVTZ these values are 41.1 and 46.8 kcal/mol for *g′Gg′* and *gGg*, respectively, and are 76.2 and 84.2 kcal/mol in the case of B3LYP/6-311++G(d,p). Thus, the small geometric variation between *g′Gg′* and *gGg* conformations, with 0.22 Å and 11.8° differences in the H···O distances and O−H···O angles (0.23 Å and 11.8° for the B3LYP geometries), and hence variation in their intramolecular hydrogen bonding, cannot be ruled out as a significant contributing factor to their relative energetics.

The deletion analysis was also carried out on the other eight conformations to expand the question of the ability of stereoelectronic effects to account for energetic ordering to the entire set of 10 conformations. The results, in conjunction with the relative energies and geometric parameters, are listed in Table 3. Figure 3 shows the relative energies as a function of the deletion energies. Linear regression on these data quite clearly demonstrates that the NBO deletion energies, either in the case of $\sigma_{CO} \rightarrow \sigma^*$ or LP $\rightarrow \sigma^*$ and with either the HF/cc-pVTZ or B3LYP/6-311++G(d,p) model chemistries, have essentially no predictive power with regard to the relative energies. In all cases, the correlation coefficients (r^2) are very small: 0.035 and 0.026 for the HF $\sigma_{CO} \rightarrow \sigma^*$ and LP $\rightarrow \sigma^*$ deletions, respectively, and 0.221 and 0.040 for the B3LYP $\sigma_{CO} \rightarrow \sigma^*$ and LP $\rightarrow \sigma^*$ deletions.

Conclusions

Taken together, the geometric and NBO data point to hydrogen bonding as the main determinant of the ethylene glycol conformational energetics. These data corroborate Trindle et al.'s conclusion that simple notions of electrostatic interaction account for ethylene glycol conformational preferences.[21] In that recent study, energetic analysis of the 10 conformations was performed with the G2(MP2) method. Additionally, the authors discussed the $E^{(2)}$ NBO energies as calculated with the B3LYP/6-311G-(d,p) model chemistry. An $E^{(2)}$ energy is defined as the charge-transfer energy calculated via second-order perturbation analysis.[20] In contrast to deletion energies as applied in the present study, the $E^{(2)}$ energies correspond to a single interaction pair (e.g., a particular lone pair and a particular antibonding orbital) and are nonadditive, which can make it difficult to assess cumulative effects. Thus, the present results have expanded on that prior work by (1) increasing the model chemistry to the HM-IE approximate CCSD(T)/cc-pVQZ method and including comparisons to nine other model chemistries and (2) including the NBO deletion energy analysis and using both HF/cc-pVTZ and B3LYP/6-311++G(d,p).

The clear lack of correlation between the total energies and either the $\sigma_{CO} \rightarrow \sigma^*$ or LP $\rightarrow \sigma^*$ deletion energies ("stereoelectronic effects") is an important point from the perspective of carbohydrate force-field development since molecular mechanics force fields do not typically have explicit terms to represent stereoelectronic effects. Rather, in the molecular mechanics representation, all conformational energetics must be accounted for by (1) bonded terms (including dihedrals) and (2) nonbonded terms (typically Coulomb and Lennard-Jones). As hydrogen bonding geometries do correlate with the ethylene glycol energetics, such molecular mechanics representations should be appropriate for the construction of an optimized carbohydrate molecular mechanics force field.

Acknowledgment. This work was supported by NIH GM070855. The authors wish to thank Dr. Igor Vorobyov for his help with the NBO analysis and to acknowledge a generous grant of computer time from the National Cancer Institute Advanced Biomedical Computing Center.

References and Notes

(1) Glennon, T. M.; Zheng, Y. J.; Legrand, S. M.; Shutzberg, B. A.; Merz, K. M. *J. Comput. Chem.* **1994**, *15*, 1019.
(2) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraserreid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
(3) Ott, K. H.; Meyer, B. *J. Comput. Chem.* **1996**, *17*, 1068.
(4) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 2078.
(5) Durier, V.; Tristram, F.; Vergoten, G. *THEOCHEM* **1997**, *395*, 81.
(6) Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, *326*, 194.
(7) Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, *326*, 210.
(8) Kuttel, M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2002**, *23*, 1236.
(9) Lii, J. H.; Chen, K. H.; Allinger, N. L. *J. Comput. Chem.* **2003**, *24*, 1504.
(10) Damm, W.; Frontera, A.; TiradoRives, J.; Jorgensen, W. L. *J. Comput. Chem.* **1997**, *18*, 1955.
(11) Kony, D.; Damm, W.; Stoll, S.; van Gunsteren, W. F. *J. Comput. Chem.* **2002**, *23*, 1416.
(12) Reiling, S.; Schlenkrich, M.; Brickmann, J. *J. Comput. Chem.* **1996**, *17*, 450.
(13) Helenius, A.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019.
(14) Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477.
(15) Boneca, I. G. *Curr. Opin. Microbiol.* **2005**, *8*, 46.
(16) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219.
(17) van Alsenoy, C.; van den Enden, L.; Schäfer, L. *J. Mol. Struct.* **1984**, *108*, 121.
(18) Oie, T.; Topol, I. A.; Burt, S. K. *J. Phys. Chem.* **1994**, *98*, 1121.
(19) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1994**, *116*, 3892.
(20) Weinhold, F. Natural Bond Orbital Methods. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley and Sons: Chichester, UK, 1998; Vol. 3, p 1792.
(21) Trindle, C.; Crum, P.; Douglass, K. *J. Phys. Chem. A* **2003**, *107*, 6236.

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

(23) MacKerell, A. D., Jr.; Brooks, B.; Brooks, C. L., III; Nilsson, L.; Roux, B.; Won, Y.; Karplus, M. CHARMM: The energy function and its paramerization with an overview of the program. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., III, Schreiner, P. R., Eds.; John Wiley & Sons: Chichester, UK, 1998; Vol. 1, p 271.

(24) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586.

(25) Yin, D. X.; Mackerell, A. D. *J. Comput. Chem.* **1998**, *19*, 334.

(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, K.; Kitao, O.; Nakai, H.; Klene, M.; Li, T. W.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision B.04; Gaussian, Inc.: Pittsburgh, PA, 2003.

(27) Klauda, J. B.; Brooks, B. R.; MacKerell, A. D.; Venable, R. M.; Pastor, R. W. *J. Phys. Chem. B* **2005**, *109*, 5300.

(28) Klauda, J. B.; Garrison, S. L.; Jiang, J. W.; Arora, G.; Sandler, S. I. *J. Phys. Chem. A* **2004**, *108*, 107.

(29) Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F. NBO version 3.1.

(30) Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007.

(31) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175.

(32) Pérez-Jordá, J. M.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134.

(33) Hobza, P.; Sponer, J.; Reschel, T. *J. Comput. Chem.* **1995**, *16*, 1315.

(34) Tsuzuki, S.; Uchimaru, T.; Tanabe, K. *Chem. Phys. Lett.* **1998**, *287*, 202.

(35) Rappé, A. K.; Bernstein, E. R. *J. Phys. Chem. A* **2000**, *104*, 6117.

(36) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33.