Quantum Chemical Conformational Analysis of 1,2-Ethanediol: Correlation and Solvation Effects on the Tendency To Form Internal Hydrogen Bonds in the Gas Phase and in Aqueous Soluti...

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · MAY 1994

Impact Factor: 12.11 · DOI: 10.1021/ja00088a027

2 AUTHORS:

Christopher J Cramer
University of Minnesota Twin Cities
531 PUBLICATIONS **23,142** CITATIONS
SEE PROFILE

Donald Truhlar
University of Minnesota Twin Cities
1,342 PUBLICATIONS **79,615** CITATIONS
SEE PROFILE

Quantum Chemical Conformational Analysis of 1,2-Ethanediol: Correlation and Solvation Effects on the Tendency To Form Internal Hydrogen Bonds in the Gas Phase and in Aqueous Solution

Christopher J. Cramer* and Donald G. Truhlar*

Contribution from the Department of Chemistry and Supercomputer Institute, University of Minnesota, 207 Pleasant Street SE, Minneapolis, Minnesota 55455-0431

*Received November 3, 1993**

Abstract: Correlated *ab initio* calculations with very large (correlation-consistent polarized valence triple-ζ) basis sets predict that 1,2-ethanediol adopts a gas-phase population of conformers at 298 K comprised of rotamers 98% gauche and 2% trans about the C–C bond. Gauche conformers that have internal hydrogen bonds make up 83% of the total population. Changes in relative energy of up to 0.6 kcal/mol are observed upon decreasing the size of the basis set to correlation-consistent polarized double-ζ (which is still larger than commonly used polarized double-ζ basis sets), illustrating the difficulty of even gas-phase conformational analysis in this seemingly simple molecule; the extra variational freedom and more complete polarization space in the larger basis stabilizes trans hydroxyl conformations and increases by a factor of 2 both the predicted fractional population of trans C–C rotamers and the predicted population of conformers with no internal hydrogen bond. Solvation effects were studied using the SM*x* series of quantum statistical aqueous solvation models. By adding calculated free energies of solvation to gas-phase free energies, it is found that the trans population increases from 2 to 12%, and the portion of conformers having no internal hydrogen bond increases from 17 to 25%. The calculated results are in reasonable agreement with experimental results both in the gas phase and in aqueous solution. The results provide a consistent picture of the competition between the various effects (electronic energies, zero point effects, thermal vibrational–rotational free energy components, and electric polarization and first hydration shell contributions to solvation free energies) that, when combined with the proper statistics, contribute to determining the populations of all possible isomers in aqueous solution. Calculated relative solvation free energies for gauche vs trans C–C torsion are also in good agreement with classical Monte Carlo and molecular dynamics simulation results.

1. Introduction

Polyhydroxylated molecules are ubiquitous in biological chemistry, with the most notable examples being sugars.[1,2] Such molecules are especially hydrophilic as a result of the hydroxyl groups serving as both donors and acceptors in hydrogen bonds with surrounding water molecules. An interesting question which arises is that of the extent to which the *intramolecular* hydrogen bonds found between vicinal hydroxyl groups in the gas phase are disrupted in aqueous solution in order to permit additional *intermolecular* hydrogen bonding with water molecules.[3,4] Since molecular recognition of sugar molecules may play a role in cellular uptake of glycosylated substrates,[2] this conformational issue is additionally relevant to the extent to which binding to a specific receptor site also involves formation of new intermolecular hydrogen bonds at the expense of intramolecular or solute–solvent ones.

1,2-Ethanediol is the simplest example of a saturated carbon compound with vicinal hydroxyl groups. Conformational isomers differ by the specification of the three possible rotameric dihedral angles, one about the C–C bond and one about each of the two C–O bonds. Assuming typical 3-fold torsion potentials with minima for angles near 60°, 180°, and 300°, this gives rise to 27 (i.e., 3^3) conformational minima. Symmetry, however, makes many of these degenerate; the 10 unique possible isomers are displayed in Figure 1 with their degeneracies indicated. Following conventional notation, the dihedral angles noted above are designated g^+, t, and g^-, respectively, for gauche clockwise, trans, and gauche counterclockwise C–O torsions, and similarly G^+, T, and G^- for the C–C torsion. Thus, the all-transoid conformer is designated tTt. Symmetry in the form of an external mirror plane parallel to the C–C bond may be thought of as reflecting + into – for all three torsions and vice versa, while t and T are unaffected. In addition, symmetry in the form of an internal plane bisecting the C–C bond relates isomers via the same reflection process, with concomitant interchange of the first and the third descriptors. Thus, by following these two procedures, one finds $g^+G^+g^-$ to be identical to $g^-G^-g^+$ (external reflection), $g^+G^-g^-$ (internal reflection), and $g^-G^+g^+$ (both reflections taken in either order).

1,2-Ethanediol has been studied in the gas phase by a variety of methods including microwave spectroscopy,[5] electron diffrac-

* Abstract published in *Advance ACS Abstracts,* April 1, 1994.

(1) (a) Rademacher, T. W.; Parekh, R. B.; Dwek, R. A. *Annu. Rev. Biochem.* **1988**, *57*, 785. (b) Koehler, J. E. H.; Saenger, W.; Van Gunsteren, W. F. *J. Mol. Biol.* **1988**, *203*, 241. (c) Srivastava, O. M.; Hindsgaul, O.; Shoreibah, M.; Pierce, M. *Carbohydr.* **1988**, *179*, 137. (d) Lemieux, R. U. *Explorations with Sugars: How Sweet it Was*; American Chemical Society: Washington, DC, 1990. (e) Kroon-Batenburg, L. M. J.; Kroon, J. *Biopolymers* **1990**, *29*, 1243. (f) Paulsen, H. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 823. (g) Ragazzi, M.; Ferro, D. R.; Perly, B.; Petitou, M.; Choay, J. *Carbohydr. Res.* **1990**, *195*, 169. (h) Sabesan, S.; Bock, K.; Paulson, J. C. *Carbohydr. Res.* **1991**, *218*, 27. (i) Zhbankov, R. G. *J. Mol. Struct.* **1992**, *275*, 65. (j) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1992**, *233*, 15.

(2) (a) Avery, O. T.; Heidelberg, M. *J. Exp. Med.* **1923**, *38*, 81. (b) Aminoff, D.; Morgan, W. T. J. *Nature* **1948**, *162*, 579. (c) Walker, S.; Valentine, K. G.; Kahne, D. *J. Am. Chem. Soc.* **1990**, *112*, 6428. (d) Lee, M. D.; Ellestad, G. A.; Borders, D. B. *Acc. Chem. Res.* **1991**, *24*, 235. (e) Rao, V. S. R. In *Molecular Conformation and Biological Interactions*; Balaram, P., Ramaseshan, S., Eds.; Indian Academy of Sciences: Bangalore, 1991; p 411. (f) Nicolaou, K. C.; Dai, W.-M.; Tsay, S.-C.; Estevez, V. A.; Wrasidlo, W. *Science* **1992**, *256*, 1172.

(3) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 5745.

(4) For general treatments of intermolecular hydrogen bonding, see: (a) Kollman, P. A. In *Modern Theoretical Chemistry*; Schaefer, H. F., III, Ed.; Plenum: New York, 1977; Vol. 4, Chapter 3. (b) Hobza, P.; Zahradnik, R. *Chem. Rev.* **1988**, *88*, 871. (c) Kumpf, R. A.; Damewood, J. R., Jr. *J. Phys. Chem.* **1989**, *93*, 4478.

(5) (a) Kuhn, L. P. *J. Am. Chem. Soc.* **1952**, *74*, 2492. (b) Walder, E.; Bander, A.; Günthard, H. H. *Chem. Phys.* **1980**, *51*, 223. (c) Caminati, W.; Corbelli, G. *J. Mol. Spectrosc.* **1981**, *90*, 572.

0002-7863/94/1516-3892$04.50/0

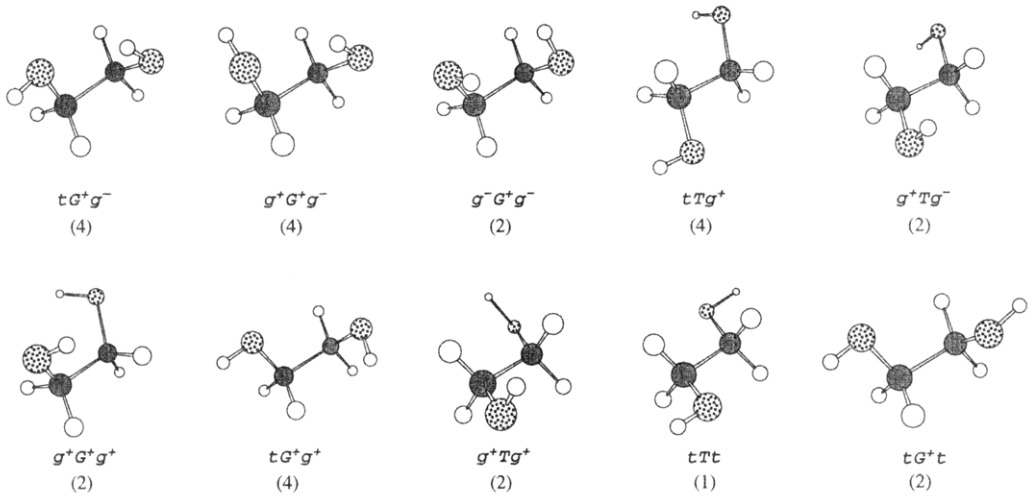

Figure 1. The ten stereochemically unique conformers of 1,2-ethanediol. Structures were fully optimized at the *ab initio* MP2/cc-pVDZ level. The degeneracy of the individual conformers is indicated by the number in parentheses below its designation. The sum of the degeneracies is equal to the total number of possible torsional combinations, i.e., 27.

tion,[6] and infrared (IR) spectroscopy.[7] The latter technique has also been used to study the diol in mixtures with inert gases quickly cooled to form low-temperature matrices.[8] These studies all agree that tG^+g^- is the predominant conformer. While the $g^+G^+g^-$ conformer was not identified in either the microwave or the electron diffraction studies, the IR results suggest that it constitutes an appreciable fraction of the equilibrium mixture,[7,8] and, in addition, Tasumi *et al.* assigned a number of very weak signals in their IR studies to the $g^-G^+g^-$ conformer.[8d] No other isomers have been proposed to contribute to the room temperature equilibrium population to within the detection limits of the various methodologies (probably about 5%).

Ab initio calculations are in general accord with these findings. Early studies at the Hartree–Fock (HF) level employing small split-valence basis sets and only partially optimized geometries were carried out for some or all isomers by a number of workers.[8a,b,9] Van Alsenoy *et al.*,[10] employing the HF/4-21G level,[11] were the first to perform complete geometry optimizations for all 10 conformers. They found valence angle differences of up to 8° among the isomers and bond distance differences of up to 0.02 Å. Subsequently, seven groups published higher level studies[8d,12–17] employing the 6-31G*, 6-31G**, and larger basis sets and taking account of electron correlation with Møller–Plesset perturbation theory up to fourth order (MP4).[11] Also, studies using quadratic configuration interaction including all single, double, and pertubative triple excitations (QCSID(T))[18] and, most recently, density functional theory within the local and nonlocal spin density approximations[19] have been reported. One of these studies[15] also tested molecular mechanics force fields[20] and discussed problems related to parameterizing such force fields using a combination of gas-phase and liquid-phase solution data. All but one of these studies[8d,13–17] included only 4–6 of the conformers, and in four cases[13–16] the only conformer considered with a trans C–C bond was *tTt*. The very recent study of Oie *et al.*,[17] on the other hand, explored all 10 possible isomers at levels up to MP4/6-311G**//MP2/6-31G** (we follow throughout the standard notational convention[11] that $x/y//z/w$ denotes a calculation with method x and basis y at a geometry optimized with method z and basis w) and NLSD/TZVP//NLSD/DZVP2 (where NLSD is the local spin density approximation[19] with nonlocal exchange and correlation corrections included using the methods of Becke[21] and Perdew,[22] respectively, and the TZVP and DZVP2 basis sets[23] are polarized triple-ζ and double-ζ, respectively). These authors found reasonable agreement between relative isomeric energies predicted at the MP4 and NLSD levels of theory, demonstrated the range of energy spanned by all 10 isomers to be less than 4.5 kcal/mol, and observed relative energies for the various isomers to be converged to within about 0.5 kcal/mol with respect to their final increase in basis set size. All of these investigations agree that tG^+g^- and $g^+G^+g^-$ are the lowest and next lowest energy isomers, respectively.

Liquid-phase solution data are more scarce. Based on nuclear magnetic resonance (NMR) studies, Pachler and Wessels assigned the concentration of isomers gauche about the C–C bond to be 88 ± 3% in deuterium oxide.[24] This may be compared to very dilute IR studies in carbon tetrachloride where, if one assumes identical extinction coefficients for the O–H stretching bands from both free and hydrogen-bonded hydroxyl groups, a ratio of 90:10 is inferred for isomers having an internal hydrogen bond

(6) Bastiansen, O. *Acta Chem. Scand.* **1949**, *3*, 415.
(7) Buckley, P.; Giguère, P. A. *Can. J. Chem.* **1967**, *45*, 397.
(8) (a) Ha, T. K.; Frei, H.; Meyer, R.; Günthard, H. H. *Theor. Chim. Acta* **1974**, *34*, 277. (b) Frei, H.; Ha, T. K.; Meyer, R.; Günthard, H. H. *Chem. Phys.* **1977**, *25*, 271. (c) Takeuchi, H.; Tasumi, M. *Chem. Phys.* **1983**, *77*, 21. (d) Park, C. G.; Tasumi, M. *J. Phys. Chem.* **1991**, *95*, 2757.
(9) (a) Radom, L.; Lathan, W. A.; Hehre, W. J.; Pople, J. A. *J. Am. Chem. Soc.* **1973**, *95*, 693. (b) Almlöf, J.; Stymne, H. *Chem. Phys. Lett.* **1975**, *33*, 118.
(10) Van Alsenoy, C.; Van Den Enden, L.; Schafer, L. *J. Mol. Struct. (Theochem)* **1984**, *108*, 121.
(11) For references for the various split-valence basis sets, explanations of basis set notation, the MP2 method of treating electron correlation in electronic structure calculations, and statistical mechanical formulae for deriving $\Delta G_{vib-rot}$, see: Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab Initio Molecular Orbital Theory*; Wiley: New York, 1986.
(12) Cabral, B. J. C.; Albuquerque, L. M. P. C.; Fernandes, F. M. S. S. *Theor. Chim. Acta* **1991**, *78*, 271.
(13) Nagy, P. I.; Dunn, W. J., III; Alagona, G.; Ghio, C. *J. Am. Chem. Soc.* **1991**, *113*, 6719.
(14) Alagona, G.; Ghio, C. *J. Mol. Struct. (Theochem)* **1992**, *254*, 287.
(15) Murcko, M. A.; DiPaola, R. A. *J. Am. Chem. Soc.* **1992**, *114*, 10010.
(16) Carmichael, I.; Chipman, D. M.; Padlasek, C. A.; Serianni, A. S. *J. Chem. Soc.* **1993**, *115*, 10863.
(17) Oie, T.; Topol, I. A.; Burt, S. K. *J. Phys. Chem.* **1994**, *98*, 1121.
(18) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968.
(19) For an overview of density functional techniques, see: (a) Parr, R. G.; Wang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989. (b) Ziegler, T. *Chem. Rev.* **1991**, *91*, 651. (c) Labanowski, J.; Andzelm, J., Eds. *Density Functional Methods in Chemistry*; Springer-Verlag: New York, 1991.
(20) (a) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127. (b) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.
(21) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(22) Perdew, J. P. *Phys. Rev. B* **1986**, *38*, 7406.
(23) Godbout, N.; Salahub, D. A.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.
(24) Pachler, K. G. R.; Wessels, P. L. *J. Mol. Struct.* **1970**, *6*, 471.

compared to those which do not.[25] Note, however, that while it is true that no trans isomer can have an internal hydrogen bond, it is also true that of the gauche isomers only tG^+g^- and $g^+G^+g^-$ have such a bond. It is thus not possible to make a direct comparison between these two experiments, although it is clear that in neither solvent is there a particularly large percentage of trans conformers.

Aqueous solvation effects on the diol have been explored theoretically with explicit solvent simulations by Nagy *et al.*[13] and Hooft *et al.*[26] The former group performed Monte Carlo calculations using the optimized potentials for liquid simulations (OPLS)[27] force field and the 4-point transferable intermolecular potential (TIP4P)[28] water model. Changes in free energies of solvation were monitored while individual dihedral angles were driven from one conformer to another, and four conformers were considered throughout. The latter group employed a molecular dynamics approach using the Gröningen molecular simulation (GROMOS) force field[29] and the simple point charge (SPC)[30] water model. By freezing only bond distances, the molecular dynamics approach was allowed to sample all conformers (assuming that the simulation time was long enough for this to occur in practice). Both studies found that the tTt isomer is better solvated than the tG^+g^-, with the former group finding a 1.2 ± 0.5 kcal/mol difference and the latter 1.1 ± 0.2 kcal/mol. Considering only the tTt trans conformer, Nagy *et al.*[13] predicted a ≥99:1 gauche/trans ratio about the C–C bond in aqueous solution. This ratio employs the required corrections pointed out by Hooft *et al.*[26] On the other hand, Hooft *et al.*, considering all possible conformers, predicted about a 67:33 ratio for the same equilibrium.[26] Most of the difference is due to the gas-phase energy surface to which the differential solvation free energies were added; Nagy *et al.*[13] used gas-phase free energies calculated at the MP2/6-31G*//HF/6-31G* level, while Hooft *et al.*[26] took their gas-phase energies from the GROMOS[29] force field. Hooft *et al.* found that the gas-phase preference for internally hydrogen-bonded solute molecules disappeared completely in aqueous solution. In light of the great overestimate of the trans form population by their force field, one must question whether this dramatic predicted solvation effect is valid.

Alagona and Ghio[14] analyzed the effects of aqueous solvation on 1,2-ethanediol by an alternative method, in which the solvent is represented by a continuum dielectric. Using the polarized continuum model (PCM) of Tomasi and co-workers,[31] they considered differences in only the electrostatic portions of the solvation free energy at the HF/4-31G and HF/6-31G* levels for four conformers. At both of these levels, using geometries optimized *in vacuo*, they found the tTt isomer to be better solvated than the tG^+g^- by only 0.2 kcal/mol. Other relative differences were more sensitive to basis set, but they found $g^+G^+g^-$ and $g^-G^+g^-$ to be better solvated than tG^+g^- by 0.6–1.1 kcal/mol.

We present here a third approach for analyzing the conformational equilibrium in aqueous solution, employing quantum statistical solvation models[32-35] which take account of both the mutual polarization of solute and solvent (like a continuum model) and also local effects specific to the first solvation shell (like an explicit solvent simulation). In particular, the Austin Model 1-Solvation Models 1a and 2 (AM1-SM1a[32] and AM1-SM2[33]) and the Parameterized Model 3-Solvation Model 3 (PM3-SM3[34]), which will be discussed in more detail below, were employed to calculate absolute free energies of solvation; these were added to gas-phase free energies calculated from significantly higher levels of *ab initio* theory than have been used to date. The parameterization of these models is consistent with respect to separating gas-phase and solution components of molecular energies. All 10 possible conformers have been considered. Section 2 describes the computational methodology and the solvation models, section 3 provides the results of the calculations for gaseous and solvation free energies, and section 4 discusses how our data compare with experiment and earlier calculations.

2. Theoretical Methods

Gas-Phase Calculations. Complete geometry optimizations were performed for all 10 isomers at both the HF and the MP2 levels using the correlation-consistent polarized valence double-ζ (cc-pVDZ) basis set of Dunning[36] (which yields 86 contracted basis functions for the present molecule). The MP2/cc-pVDZ geometries were used for additional single point calculations taking account of correlation using the coupled cluster formalism, considering all single and double excitations to infinite order and using a perturbative approach for triple excitations,[37] i.e., CCSD(T)/cc-pVDZ//MP2/cc-pVDZ. HF and MP2 single point calculations at these geometries were also carried out with the much larger cc-pVTZ basis set,[36] which is a valence triple-ζ set including 2p and 1d polarization functions on hydrogen and 2d and 1f polarization functions on the heavy atoms (a total of 204 contracted basis functions for the diol). Relative energies, E_{rel}, were then calculated as [MP2/cc-pVTZ + (CCSD(T)/cc-pVDZ − MP2/cc-pVDZ)], with all calculations at the MP2/cc-pVDZ geometry, i.e., higher order correlation effects calculated with the smaller basis set were added to the largest basis set MP2 calculation. These energies were converted to relative gas-phase free energies employing standard statistical mechanical formulae,[11,38] using vibrational frequencies and moments of inertia calculated at the HF/cc-pVDZ level. The latter level was also used to verify all structures as local minima. Finally, each isomer has a contribution to its gas-phase free energy of $-RT \ln \omega$, where ω is the structural degeneracy of the isomer as listed in Figure 1. All *ab initio* calculations were carried out with the GAUSSIAN92 program suite.[39]

Semiempirical calculations at the Neglect of Diatomic Differential Overlap (NDDO)[40] level were also carried out with the Austin Model 1 (AM1)[41] and Parametrized Model 3 (PM3)[42] Hamiltonians. Again, complete geometry optimizations were performed; the tG^+t isomer was not stationary with either semiempirical Hamiltonian, and other differences in geometry between the semiempirical and *ab initio* levels are discussed in the following sections. We regard these calculations as delivering electronic and nuclear potential energies, E, and convert these to G°_{298} by employing the same thermodynamic formulae[11] as in the *ab initio* case.

(25) (a) Krueger, P. J.; Mettee, H. D. *J. Mol. Spectrosc.* **1965**, *18*, 131. (b) Busfield, W. K.; Ennis, M. P.; McEwen, I. J. *Spectrochim. Acta A* **1973**, *29*, 1259.

(26) Hooft, R. W. W.; van Eijck, B. P.; Kroon, J. *J. Chem. Phys.* **1992**, *97*, 3639.

(27) Jorgensen, W. L. *J. Phys. Chem.* **1986**, *90*, 1276.

(28) Jorgensen, W. L.; Madura, J. D. *Mol. Phys.* **1985**, *56*, 1381.

(29) van Gunsteren, W. F. *Gröningen Molecular Simulation Package (GROMOS)*; University of Gröningen: The Netherlands, 1987.

(30) Berendsen, H. J. C.; Posta, J. P. M.; van Gunsteren, W. F.; Hermans, J. In *Intermolecular Forces*; Pullmann, B., Ed.; Reidel: Dordrecht, 1981; pp 331–342.

(31) (a) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117. (b) Bonaccorsi, R.; Cimiraglia, R.; Tomasi, J. *J. Comput. Chem.* **1983**, *4*, 567.

(32) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1991**, *113*, 8305, 9901(E).

(33) Cramer, C. J.; Truhlar, D. G. *Science* **1992**, *256*, 213.

(34) Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.* **1992**, *13*, 1089.

(35) For a comprehensive overview, see: Cramer, C. J.; Truhlar, D. G. *J. Comput. Aided Mol. Des.* **1992**, *6*, 629.

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

(37) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

(38) Nagy *et al.*[13] considered replacement of some torsional modes in the diol vibrational partition functions with classical hindered (or free) rotors. After correction for an error pointed out by Hooft, van Eijck, and Kroon,[26] the effect of this substitution in the hindered rotor approximation for the tTt isomer is to decrease its relative free energy by about 0.5 kcal/mol (the assumption further being made that no corrections are required for the internally hydrogen-bonded isomers). Sufficient approximations are implicit in this approach that it is unclear whether it is any more accurate than simply using the *ab initio* partition function—especially insofar as *all* isomers have at least four modes with characteristic frequencies below 500 cm^{-1}. We have therefore elected not to attempt any such corrections for hindered rotors.

(39) Frisch, M. J.; Trucks, G. W.; Head-Gordon, M.; Gill, P. M. W.; Wong, M. W.; Foresman, J. B.; Johnson, B. G.; Schlegel, H. B.; Robb, M. A.; Replogle, E. S.; Gomperts, R.; Andres, J. L.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Martin, R. L.; Fox, D. J.; Defrees, D. J.; Baker, J.; Stewart, J. J. P.; Pople, J. A. *GAUSSIAN92*; Gaussian, Inc.: Pittsburgh, PA, 1992.

(40) Pople, J. A.; Beveridge, D. L. *Approximate Molecular Orbital Theory*; McGraw-Hill: New York, 1970.

(41) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.

(42) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209, 221.

Table 1. Calculated Relative Energies of 1,2-Ethanediol Conformers in the Gas Phase[a,b]

	internal H-bond		no internal H-bond							
level	tG^+g^- [c]	$g^+G^+g^-$	$g^-G^+g^-$	tTg^+	g^+Tg^-	$g^+G^+g^+$	tG^+g^+	g^+Tg^+	tTt	tG^+t
AM1	0.00	–0.88	–0.95	1.55	0.22	–0.98	2.54	0.39	2.97	*d*
PM3	0.00	–1.39	–0.78	1.06	–0.68	–0.82	1.76	–0.43	2.55	*d*
HF/4-21G[e]	0.00	0.88	2.28	3.19	3.35	4.30	6.44	3.99	2.55	5.39
HF/6-31G**[f]	0.00	0.64	1.28	2.36	2.44				2.02	
HF/cc-pVDZ	0.00	0.56	1.53	2.37	2.35	3.22	3.94	2.69	2.08	3.61
HF/cc-pVTZ//MP2/cc-pVDZ	0.00	0.80	1.13	1.77	2.10	2.87	3.14	2.29	1.35	2.62
MP2/6-31G*//HF/6-31G*[g]	0.00	0.24	1.25						3.36	
MP2/6-31G**//HF/6-31G**[f]	0.00	0.20	1.22	3.25	2.99				3.23	
MP2/cc-pVDZ	0.00	–0.04	1.58	3.31	2.79	3.27	4.33	3.05	3.46	4.34
CCSD(T)/cc-pVDZ//MP2/cc-pVDZ	0.00	–0.12	1.39	3.14	2.60	3.11	4.20	2.85	3.34	4.27
MP4/6-311G**//MP2/6-31G*[h]	0.00	–0.09	1.49	3.14	2.67	3.03	4.28	2.93	3.31	4.24
NLSD/TZVP//NLSD/DZVP2[h]	0.00	–0.01	1.02	3.05	2.59	2.70	3.68	2.70	3.10	3.60
MP2/cc-pVTZ//MP2/cc-pVDZ	0.00	0.31	1.20	2.91	2.80	3.22	3.81	2.99	2.85	3.48
composite E_{rel}[i]	0.00	0.23	1.01	2.74	2.61	3.06	3.68	2.79	2.73	3.41
ZPVE[j]	0.00	0.05	–0.37	–0.34	–0.26	–0.53	–0.43	–0.29	–0.35	–0.51
$\Delta G_{vib-rot}(T)$[j,k]	0.00	0.19	–0.58	–0.52	–0.38	–0.60	–0.65	0.01	–0.19	–0.39
$-RT \ln \omega$	0.00	0.00	0.41	0.00	0.41	0.41	0.00	0.41	0.82	0.41
composite G°_{298}[l]	0.00	0.42	0.84	2.22	2.64	2.87	3.03	3.21	3.36	3.43

[a] In kcal/mol. [b] Unless otherwise indicated, geometries were fully optimized at the reported level. [c] Absolute energies (E_h) for this column are –36.033 35; –33.707 05; –228.394 43; –228.945 01; –228.953 39; –229.024 30; –229.540 61; –229.596 48; –229.613 37; –229.668 85; not reported; not reported; –229.860 44; –229.915 92; 0.091 88; 0.063 91; –0.001 30; –229.853 31. [d] Not stationary at this level of theory. [e] Reference 10. [f] Reference 8d. [g] References 13 and 14. [h] Reference 17. [i] See Section 2. [j] Calculated from unscaled HF/cc-pVDZ frequencies. [k] ZPVE + vibrational–rotational heat capacity and entropy terms, again at HF/cc-pVDZ level. [l] Calculated as composite $E_{rel} + \Delta G_{vib-rot}(T) - RT \ln \omega$.

Aqueous Calculations. Solvation free energies were calculated using the AM1-SM1a,[32] AM1-SM2,[33] and PM3-SM3[34] solvation models. The standard state for these calculations is 298 K and one mol L^{-1} solute concentration both in the gas phase and in solution. These models divide the free energy of solvation into two parts. The first, known as ΔG_{ENP}, accounts for the favorable effects of mutual polarization of the solute and the solvent and the associated costs of electronic and geometric distortion, at least to the extent that these effects can be simulated using AM1 or PM3 partial charges at the nuclear centers. The second term, G°_{CDS}, includes the effects of cavitation, dispersion, and other interactions specific to the first solvation shell, such as the nonelectrostatic components of solute–solvent and solvent–solvent hydrogen bonding and the entropic effects of perturbing the structure of the first shell of solvent; it also includes electric polarization effects due to local atomic dipoles that are beyond the ability of the atom-centered partial charges to represent and makes up empirically for some of the deficiencies of the AM1 and PM3 partial charges. While all three SM*x* models calculate ΔG_{ENP} in the same manner (but using different optimized parameters), they differ in the way in which G°_{CDS} is treated.

In the AM1-SM1a model, every atom is assigned a surface tension and contributes to G°_{CDS} based on its solvent-accessible surface area.[43] For the diol, four types of atoms are explicitly assigned: carbon, carbon-bound hydrogen, sp^3 oxygen, and oxygen-bound hydrogen. The SM2 and SM3 models, on the other hand, are more general in that they do not require an assignment of atomic "type". Instead, all atoms have a single atom-specific surface tension, which is further adjusted depending on the number of attached hydrogen atoms as evaluated from the self-consistently determined bond order matrix.[35] This prescription allows hydrogen atoms themselves, which have drastically different interactions with the first solvation shell depending on the heavy atom to which they are attached, to be treated much more accurately than would be possible using a single hydrogen-atom surface tension. The attached hydrogen atoms modify other atoms' surface tensions, and their own solvent-accessible surface area is set to zero. Moreover, they do not explicitly block solvent from the underlying heavy atoms, which thereby take on the role of representing fragments (e.g., CH_2 and OH in 1,2-ethanediol) in the G°_{CDS} calculation (although H atoms remain explicit for ΔG_{ENP}). As a result of these differences, the SM1a model can be somewhat more discriminating between distinct rotamers involving $-XH_n$ groups than the SM2 and SM3 models. This is discussed in more detail in the following sections.

For each SM*x* model, the calculation was performed in two ways. First, free energies of solvation were computed for the "frozen" MP2/cc-pVDZ geometry; this approach does not include contributions from nuclear structural relaxation, but it may be more trustworthy in cases where the semiempirical AM1 or PM3 Hamiltonian fails to deliver a reasonable gas-phase (or liquid-phase solution) structure. In addition, calculations were performed for fully "relaxed" structures, e.g., ΔG°_{solv} = PM3-SM3//PM3-SM3 – PM3//PM3. Finally, for the AM1-SM2 model, additional calculations were performed where only the three critical dihedral angles of 1,2-ethanediol were "locked" at the MP2/cc-pVDZ values, all other degrees of freedom being permitted to optimize. All semiempirical calculations, both in the gas phase and in solution, were performed with version 3.5 of the AMSOL computer program.[44]

Equilibrium Populations and Absolute Free Energies of Solvation. Equilibrium populations of conformers were calculated using a standard Boltzmann formalism, i.e., the fraction of conformer A is

$$F(\mathrm{A}) = \frac{e^{-G^\circ_A/RT}}{\sum_i e^{-G^\circ_i/RT}} \qquad (1)$$

where G may be either a gas-phase or a solution value, and i runs over all conformers. In tables, $\%(\mathrm{A}) \equiv 100F(\mathrm{A})$. The absolute free energy of solvation, ΔG°_{solv}, is calculated as[45]

$$\Delta G^\circ_{solv} = -RT \ln \left\{ \sum_A \left[\frac{e^{-G^\circ_{gas}(A)/RT}}{\sum_i e^{-G^\circ_{gas}(i)/RT}} \right] e^{-\Delta G^\circ_{solv}(A)/RT} \right\} \qquad (2)$$

where both i and A run over all conformers.

3. Results and Discussion

Gas Phase. Figure 1 illustrates the optimized structures of all 10 conformers at the MP2/cc-pVDZ level; additional geometric and energetic details are provided as supplementary material. Table 1 provides a comprehensive listing of our own and several earlier gas-phase calculations for these conformers. It is evident that the cc-pVDZ basis set is the best polarized double-ζ basis employed: the Hartree–Fock electronic energy of optimized tG^+g^- is 5.3 kcal/mol lower with the cc-pVDZ basis (86 contracted

(43) (a) Lee, B.; Richards, F. M. *J. Mol. Biol.* **1971**, *55*, 379. (b) Hermann, R. B. *J. Phys. Chem.* **1972**, *76*, 2754. (c) Eisenberg, D.; McLachlan, A. D. *Nature* **1986**, *319*, 199. (d) Ooi, T.; Oobatake, M.; Nemethy, G.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 3086.

(44) Cramer, C. J.; Lynch, G. C.; Hawkins, G. D.; Truhlar, D. G. AMSOL, version 3.5, Quantum Chemistry Program Exchange, program no. 606. *QCPE Bull.* **1993**, *13*, 55.

(45) Ben-Naim, A. *Statistical Thermodynamics for Chemists and Biochemists*; Plenum: New York, 1992.

Table 2. Equilibrium Gas-Phase Population of Conformers at 298 K at Various Levels of Theory, with Higher Levels of Theory to the Right[a]

isomer	AM1[b]	PM3[c]	HF/4-21G[d]	CCSD(T)/cc-pVDZ[e]	MP4/6-311G**[f]	NLSD/TZVP[g]	composite (C)[h]
tG^+g^-	5.2	4.2	82.6	48.7	50.3	49.3	55.9
$g^+G^+g^-$	16.7	32.2	13.5	43.2	42.5	36.3	27.4
$g^-G^+g^-$	34.6	21.1	2.3	6.2	5.4	11.7	13.4
tTg^+	0.9	1.7	0.9	0.6	0.6	0.7	1.3
g^+Tg^-	3.5	12.7	0.3	0.6	0.5	0.6	0.6
$g^+G^+g^+$	37.6	23.3	0.1	0.4	0.4	0.7	0.4
tG^+g^+	0.2	0.6	0.0	0.1	0.1	0.3	0.3
g^+Tg^+	1.3	4.1	0.1	0.2	0.2	0.3	0.2
tTt	0.0	0.0	0.4	0.1	0.1	0.1	0.2
tG^+t	*i*	*i*	0.0	0.0	0.0	0.1	0.2
total C–C gauche	94.3	81.5	98.4	98.6	98.7	98.4	97.6
total C–C trans	5.7	18.5	1.6	1.4	1.4	1.6	2.4
internal H-bond[j]	21.9	36.4	96.0	92.0	92.7	85.6	83.4
no internal H-bond	78.1	63.6	4.0	8.0	7.3	14.4	16.7

[a] % (A) as defined below eq 1. [b] From Table 1, AM1 + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [c] From Table 1, PM3 + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [d] From Table 1, HF/4-21G + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [e] From Table 1, CCSD(T)/cc-pVDZ + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [f] From Table 1, MP4/6-311G** + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [g] From Table 1, NLSD/TZVP + $\Delta G_{vib-rot}(T) - RT \ln \omega$. [h] From composite final row of Table 1, based on MP2/cc-pVTZ and CCSD(T)/cc-pVDZ calculations. [i] Not stationary. [j] Sum of isomers tG^+g^- and $g^+G^+g^-$.

basis functions) than with the nominally comparable 6-31G** basis (90 contracted basis functions). At the frozen core MP2 level, the difference increases to 9.0 kcal/mol (using the HF optimized geometries), illustrating the superior balance of the cc-pVDZ basis for correlated calculations. Finally, although absolute energies were not reported by Oie *et al.*,[17] we note that our largest basis, cc-pVTZ, is considerably more complete (204 contracted basis functions) than was theirs, 6-311G** (108 contracted basis functions); very large basis sets are required in order to accurately calculate correlation energies.[46] Table 2 presents the results of Table 1 in the form of equilibrium populations in the gas phase at 298 K.

It is particularly interesting to examine this system with the correlation-consistent basis sets[36] because some previous studies[47] have pointed out systematic deficiencies in the 6-31G and 6-311G basis sets, in particular, insufficiently diffuse valence orbitals; the correlation-consistent basis sets do not suffer from these problems.

A number of interesting trends are apparent. First, isomeric relative energies differ considerably with improving basis set—even on going from the cc-pVDZ basis, which is quite large by most standards, to the cc-pVTZ basis, energies relative to the global minimum change by up to 0.80 kcal/mol at the HF level (note that slightly different geometries are involved) and by up to 0.86 kcal/mol at the MP2 level (identical geometries). In particular, the larger basis appears to stabilize conformers with *t* hydroxyl orientations relative to the global minimum, which itself has a single *t* hydroxyl, but is unique in that the hydroxyl in question is serving as a hydrogen bond acceptor. Thus, the largest stabilizations are observed for the tTt and tG^+t isomers, where both hydroxyl torsions are trans. In addition, the only other conformer with an internal hydrogen bond, $g^+G^+g^-$, goes up 0.35 kcal/mol in energy, suggesting that this intramolecular interaction is overestimated even with a polarized cc-pVDZ basis. The net result on going from the cc-pVDZ to the cc-pVTZ composite free energies is to about double the equilibrium percentage both of conformers having no internal hydrogen bond and of conformers which are trans about the C–C bond (Table 2). This is in marked contrast to the MP4 results of Oie *et al.*,[17] which most closely resemble our CCSD(T)/cc-pVDZ results and do not capture the changes observed on going to the composite cc-pVTZ level. The NLSD/TZVP results of Oie *et al.*,[17] on the other hand, agree quite well with our best composite triple-ζ calculations. The only deviation of note is in the relative proportions of the two internally hydrogen-bonded isomers, with the NLSD level predicting the tG^+g^- isomer to be only slightly more prevalent than the $g^+G^+g^-$. However, of the *ab initio* calculations, only these two models predict the total amount of these two isomers to be in the 85% range, as compared to the lower level *ab initio* predictions of 92–96%. As density functional methods remain relatively untested for systems with internal hydrogen bonds, this comparison is particularly interesting. Finally, it is noteworthy that in general the larger basis sets decrease the energetic separation between the two internally hydrogen-bonded isomers and the eight others and that this effect is reduced at correlated levels compared to the HF level.

Figure 2. Structure of the $g^-G^+g^-$ conformer optimized at the AM1 level. The perspective shown on the right illustrates the nearly eclipsed nature of the structure.

It is also clear that the semiempirical levels are totally unsuited to gas-phase conformational analysis in this instance. This is a typical result for conformational problems, and it appears to arise in the present case from two key anomalies. The first, unique to AM1, is the tendency for isomer $g^-G^+g^-$ to optimize from a staggered geometry about the C–C bond to one in which the two C–O bonds are nearly eclipsed (Figure 2). This appears to be driven by a strongly favorable interaction between the protons and the oxygens on opposite hydroxyl groups. This is reminiscent of the bifurcated hydrogen bonds which are typically produced by AM1,[48] and the effect is to dramatically increase the equilibrium population of this conformer in the gas phase. The PM3 structures, conversely, are all reasonably similar to their *ab initio* counterparts.

The other anomaly in the semiempirical gas-phase methods is the apparent preference in both methods for the O–H bond to eclipse some other bond in a 1,3-fashion, such that gauche hydroxyl torsions are dramatically stabilized over trans. This effect is sufficient to render the tG^+t conformer nonstationary. Moreover, since one of the two intramolecularly hydrogen-bonded conformers which dominate the *ab initio* population has a trans hydroxyl, its destabilization causes the ratio of hydrogen-bonded to non-hydrogen-bonded conformers to effectively reverse on going from the best *ab initio* level of theory to either of the semiempirical

(46) (a) Almlöf, J.; Taylor, P. R. *J. Chem. Phys.* **1987**, *86*, 4070. (b) Gordon, M. S.; Truhlar, D. G. *Int. J. Quantum Chem.* **1987**, *31*, 81. (c) Bauschlicher, C. W., Jr.; Langhoff, S. R.; Taylor, P. R. In *Supercomputer Algorithms for Reactivity, Dynamics, and Kinetics of Small Molecules*; Lagana, A., Ed.; Kluwer: Dordrecht, 1989; p 1.

(47) (a) Grev, R. S.; Schaefer, H. F., III. *J. Chem. Phys.* **1989**, *91*, 7305. (b) Del Bene, J. E.; Aue, D. H.; Shavitt, I. *J. Am. Chem. Soc.* **1992**, *114*, 1631.

(48) (a) Khalil, M.; Woods, R. J.; Weaver, D. F.; Smith, V. H., Jr. *J. Comput. Chem.* **1991**, *12*, 584. (b) Zheng, Y.-J.; Merz, K. M., Jr. *J. Comput. Chem.* **1992**, *13*, 1151.61.

Table 3. Free Energies of Solvation for 1,2-Ethanediol Conformers[a,b]

	SM1a		SM2			SM3		
isomer	frozen	relaxed	frozen	relaxed	locked	1SCF	frozen	relaxed
tG^+g^-	−8.76	−8.80	−9.45	−9.62	−9.61	−9.19	−8.96	−9.25
$g^+G^+g^-$	−9.20	−9.11	−9.52	−9.70	−9.59	−9.39	−9.05	−9.44
$g^-G^+g^-$	−9.25	−8.77[c]	−9.54	−9.08[c]	−9.52	−9.36	−9.02	−9.39
tTg^+	−9.18	−9.18	−10.41	−10.40	−10.36	−10.02	−10.16	−10.06
g^+Tg^-	−9.54	−9.64	−10.51	−10.48	−10.48	−10.15	−10.28	−10.17
$g^+G^+g^+$	−9.45	−8.68	−9.67	−8.99	−9.55	−9.30	−9.18	−9.33
tG^+g^+	−9.48	−9.62	−10.23	−10.48	−10.23	−9.69	−9.67	−9.74
g^+Tg^+	−9.61	−9.55	−10.52	−10.48	−10.46	−10.12	−10.29	−10.14
tTt	−8.87	−9.05	−10.42	−10.47	−10.46	−10.03	−10.16	−10.05
tG^+t	−9.32	*d*	−10.27	*d*	−10.28	*d*	−9.72	*d*
ΔG°_{solv}[e]	−9.01	−8.92	−9.55	−9.64	−9.64	−9.32	−9.09	−9.37

[a] Solvation free energies refer to a standard state of 298 K and 1 mol L^{-1} solute concentration in both the gas phase and solution. [b] The SM3 1SCF calculations are carried out with the geometry frozen at the gas-phase structures optimized by PM3; see Section 2 for details of how the geometries were chosen in other columns. [c] Structure is quite different from *ab initio* geometry; see Figure 2. [d] Structure is not stationary at this level of theory. [e] Using eq 2 from text.

levels! This is apparently not related to the small basis implicit in the semiempirical techniques, since at the *ab initio* level the tendency is for smaller bases to overemphasize the percentage of the conformers with an internal hydrogen bond. It has been suggested that semiempirical H–H core repulsion functions overstabilize the eclipsing of vicinal X–H bonds, leading to incorrect flattening of 5-membered alicyclic rings at such levels;[49] however, this explanation has been disputed.[50] It is not clear what is responsible for the apparent favorability of the more distant 1,3-interaction in this instance. It is, however, perhaps worth noting that in general only primary aliphatic alcohols can adopt a transoid hydroxyl torsion which fails to 1,3-eclipse another bond. As such, 1,2-ethanediol is a worst-case scenario for this problem, while more complicated molecules, like sugars with a majority of secondary or tertiary alcohols, may be less adversely affected.

The poor accuracy of the AM1 and PM3 models for gas-phase free energies is mentioned for its intrinsic interest, but it will not affect our final results in aqueous solution since, as in previous studies,[3,51] we will combine high-level *ab initio* gas-phase free energies with SM*x* solvation energies.

Several previous studies of 1,2-ethanediol included only one conformer, *tTt*, that is trans about the C–C bond. Table 1 shows that at all HF levels of theory the electronic energy of *tTt* is indeed lower than any other C–C transoid conformer, but it enjoys much less correlation energy; moreover, it is the only isomer which has no geometrical degeneracy (a cost of 0.82 kcal/mol relative to 4-fold-degenerate conformers at 298 K). Considering all of these effects together with a temperature correction which is about average, the *tTt* conformer is found to constitute less than 10% of the total gas-phase trans population at the highest level of theory in Table 2. Conformers tTg^+ and tTg^- are found to be the dominant trans isomers at the highest level. The question of whether aqueous solvation may favor *tTt* sufficiently to make it the dominant C–C trans conformer in solution will be discussed further below on the basis of the present complete conformational analysis of *all* C–C trans conformers.

An interesting trend observed when the zero point vibrational energies and $\Delta G_{vib-rot}(T)$ values are compared for all 10 isomers in Table 1 is that these terms favor the isomers without internal hydrogen bonds by an average of 0.5 kcal/mol compared to those with internal hydrogen bonds. This is reasonable since the five-membered ring structure of the internal hydrogen bonding tightens the torsional potentials and bending frequencies. Attempts to explain the conformational equilibria without including such effects cannot be quantitative.

Two points worthy of comment are the difference in potential energy between the lowest energy internally hydrogen-bonded isomer and the lowest energy isomer without such a hydrogen bond and also the difference in energy between the lowest energy gauche C–C isomers and the lowest energy trans C–C isomer. The former difference is 1.0 kcal/mol, and the latter is 2.6 kcal/mol. These values become 0.8 and 2.2 kcal/mol, respectively, if we consider free energies instead of rigid molecule potential energies at optimized structures.

Aqueous Solution. Solvation free energies calculated from the three solvation models for various constrained or relaxed geometries are presented in Table 3. The final column of Table 3 is our most reliable estimate.

Comparison of the last and seventh (SM3 ISCF) columns of Table 3 provides our most reliable estimate of how much the solvation energy changes when the solute geometry is allowed to relax in solution. It is interesting to have an estimate of this effect since it is typically not addressed in explicit water simulations; we find that it is a small effect, ranging from 0.02 to 0.06 kcal/mol for all conformers.

Adding the solvation energies in Table 3 to the composite gas-phase free energies of Table 1 provides relative free energies in solution, which are presented in Table 4 in the form of the equilibrium population of conformers at 298 K. By using eq 2, it is possible to calculate the conformationally averaged absolute free energy of solvation. Our best value of this quantity (last column of Table 3) is −9.4 kcal/mol. Considering all of the different methods, the theoretical precision is about 0.5 kcal/mol.

4. Comparison to Experiment and Previous Simulations

Although an experimental measurement[52] from 1935 yielded a free energy of solvation of −7.7 kcal/mol, the experimental technique employed suggests considerable error in this value.[53] A much more recent measurement, employing a dew point technique,[54] yielded a free energy of solvation of −9.6 kcal/mol, in excellent agreement with our calculation, which preceded the new measurement.

The more recent value seems quite reasonable from a group contribution perspective since on going from ethane to ethanol, the free energy of solvation changes[55] from 1.8 to −5.0 kcal/mol

(49) Ferguson, D. M.; Gould, I. R.; Glauser, W. A.; Schroeder, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 525.
(50) Csonka, G. I. *J. Comput. Chem.* **1993**, *14*, 895.
(51) Cramer, C. J.; Truhlar, D. G. *J. Am. Chem. Soc.* **1993**, *115*, 8810.
(52) Butler, J. A. V.; Ramchandi, C. N. *J. Chem. Soc.* **1935**, 952.
(53) The experiment bubbled a stream of dry air through a solution of 1,2-ethanediol (0.57 M solution in water) and measured a 34 ppm concentration of the diol in the condensed stream, which was assumed to represent the ratio of vapor pressures over the solution. The potential difficulties arise from nonideality, nonequilibrium of the gas bubbles with the solution, and solute entrainment in small hydrogen-bonded gas-phase water clusters.
(54) Suleiman, D.; Eckert, C. A. *J. Chem. Eng. Data*, submitted for publication.
(55) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563.

Table 4. Equilibrium Gas-Phase and Aqueous Solution Populations of Conformers at 298 K at Various Levels of Theory[a,b]

		C + SM1a		C + SM2			C + SM3	
isomer	gas-phase C[c]	frozen	relaxed	frozen	relaxed	locked	frozen	relaxed
tG^+g^-	55.9	36.3	45.4	47.4	53.9	53.3	44.8	45.8
$g^+G^+g^-$	27.4	37.6	37.7	26.2	30.3	25.3	25.6	31.0
$g^-G^+g^-$	13.4	20.1	10.4	13.3	5.2	11.0	11.9	14.0
tTg^+	1.3	1.7	2.0	5.6	4.7	4.4	8.0	4.2
g^+Tg^-	0.6	1.6	2.2	3.3	2.6	2.5	4.8	2.5
$g^+G^+g^+$	0.4	0.9	0.3	0.5	0.1	0.4	0.5	0.4
tG^+g^+	0.3	0.7	1.1	1.0	1.4	0.9	0.9	0.6
g^+Tg^+	0.2	0.7	0.7	1.3	1.0	1.0	1.9	0.9
tTt	0.2	0.1	0.2	0.8	0.8	0.8	1.2	0.6
tG^+t	0.2	0.3	*d*	0.6	*d*	0.5	0.5	*d*
total C–C gauche	97.6	95.9	94.9	89.1	90.9	91.4	84.3	91.8
total C–C trans	2.4	4.1	5.1	10.9	9.1	8.6	15.7	8.2
internal H-bond[e]	83.3	73.9	83.1	73.6	84.2	78.6	70.4	76.8
no internal H-bond	16.7	26.1	16.9	26.4	15.8	21.4	29.6	23.2

[a] %, using eq 1 from text. Solution free energies are taken as the sum of free energies of solvation from Table 3 and gas-phase composite free energies from the last row of Table 1. [b] See Section 2 for details on how geometries were chosen. [c] C denotes composite gas-phase results in Table 1. [d] Unconstrained geometry is not a stationary point at this level of theory. [e] Sum of isomers tG^+g^- and $g^+G^+g^-$.

(AM1-SM1a, AM1-SM2, and PM3-SM3 predict[35] 0.8, 1.2, 1.2 and −4.4, −4.9, −4.6 for these two molecules, respectively). While the vicinal addition of another hydroxyl group will clearly have less than a fully additive effect, it seems unlikely, even given the intramolecular hydrogen bonding possibility, that only another −2.7 kcal/mol of additional favorable solvation would be gained after the initial change of −6.8; moreover, the SM*x* solvation models slightly underestimate the initial gain, suggesting that the −9.4 kcal/mol prediction is probably not too negative.

We now turn to a discussion of how our results compare to other experimental data, explicit water simulations, and other results obtained with continuum models.

We consider first the relationship between the solvation model results and the experimental measurement in D_2O of an 88 ± 3% concentration of conformers which are gauche about the C–C bond.[24] It is important to note that the quoted error refers only to the precision of the measurement and does not include potential errors in the estimated coupling constants for the "pure" gauche and trans isomers. The latter are derived from a reaction field methodology which does not explicitly account for either intra- or intermolecular hydrogen bonds.[56]

The SM*x* results, however, appear to support this ratio quite closely. Thus, if we eliminate the relaxed AM1-SM*x* models because of their very poor geometries for the $g^-G^+g^-$ conformer, we find a general agreement that the gauche fraction is 88 ± 4% for all of the remaining models except the frozen AM1-SM1a results. The agreement between the frozen SM2, locked SM2, frozen SM3, and relaxed SM3 results is remarkably close for the percentage population of every conformer, affording us some confidence in the performance of the models. A corollary is that the four models are in fairly close agreement on the equilibrium ratio of conformers with and without an internal hydrogen bond, roughly 75:25. These results, together with our best gas-phase calculations, suggest that on passing from the gas phase into aqueous solution there is a relative increase in the population of trans conformers by a factor of 3–5, in absolute terms an increase of about 6–10 percentage points. From inspection of Tables 3 and 4, most of this appears to result from an increase in the population of trans conformers tTg^+ and g^+Tg^- at the expense of conformer $g^-G^+g^-$.

It is unfortunately difficult to compare these results to the explicit water simulations of Nagy *et al.*[13] and Hooft, van Eijck, and Kroon.[26] In the former case, where an estimated 99:1 *G*:*T* ratio in aqueous solution was suggested, only the tTt isomer was considered. However, the SM*x* models suggest that this isomer is a minor player in the equilibrium compared to tTg^+ and g^+Tg^-.

(56) Abraham, R. J.; Cavalli, L.; Pachler, K. G. R. *Mol. Phys.* **1966**, *11*, 49.

Table 5. Relative Populations for Conformers of the T and G Rotamers from Different Methods[a,b]

		C + SM2		C + SM3	
isomer	Hooft, van Eijck, and Kroon[c]	frozen	locked	frozen	relaxed
tG^+g^-	24	53	58	53	50
$g^+G^+g^-$	22	29	28	30	34
$g^-G^+g^-$	10	15	12	14	15
$g^+G^+g^+$	15	1	0	1	0
tG^+g^+	20	1	1	1	1
tG^+t	9	1	1	1	*d*
tTg^+	47	51	51	51	51
g^+Tg^-	24	30	29	30	30
g^+Tg^+	12	12	11	12	11
tTt	17	8	9	7	7

[a] %. [b] SM*x* data from Table 4. [c] Reference 26. [d] Not stationary at this level.

While Hooft, van Eijck, and Kroon calculated the aqueous population for all conformers, they did so with the gas-phase portion of the free energy calculated using the GROMOS force field, leading to an incorrect overall *G*:*T* ratio of 67:33. (This ratio is the average for two separate sampling techniques, one of which predicts [trans]/[gauche] to be 0.54 ± 0.08, and other predicts 0.46 ± 0.04.) However, we may compare the relative populations of the *T* and *G* rotamers separately, thereby eliminating any effect from errors in their internal force field with respect to describing the torsion about the central C–C bond. Such an analysis is provided in Table 5, where it is apparent that the agreement for the *T* isomers is excellent. The SM*x* models predict slightly higher populations of conformers tTg^+ and g^+Tg^- at the expense of conformer tTt, but that is the only significant difference. It is not clear whether this latter discrepancy arises from a poor torsional term in the force field or an inadequacy in either of the solvation models, but whatever the cause, the effect would be quite small if expressed in kcal/mol.

For the dominant *G* rotamers, on the other hand, the agreement is rather poor. We find about a 40–50% relative increase in populations of conformers without an internal hydrogen bond, i.e., water does compete with internal hydrogen bonds but is by no means strong enough to completely eliminate internal hydrogen bonding. In contrast, the results of Hooft, van Eijck, and Kroon predict much higher concentrations of *G* conformers that lack an internal hydrogen bond. This leads to the suspicion that the GROMOS force field fails to accurately estimate the energy of the internal hydrogen bond relative to external ones.

While this last assumption is somewhat speculative, it is supported to some extent by an analysis of another calculation common to all three solvation studies, the solvation free energy of tTt relative to tG^+g^-. Nagy *et al.*[13] and Hooft, van Eijck, and

Kroon[26] predict the latter isomer to be less well solvated by 1.2 ± 0.5 and 1.1 ± 0.2 kcal/mol, respectively. The SM2 frozen, SM2 locked, SM3 frozen, and SM3 relaxed results are 1.0, 0.8, 1.2, and 0.8 kcal/mol, respectively; the simulations employ either unrelaxed or partially relaxed geometries, and these SM*x* results agree very closely. Thus, since the simulation results and the SM*x* results are in harmony for the relative free energies of solvation for these isomers, one with and one without an internal hydrogen bond, it seems likely that population differences among the *G* conformers with and without internal hydrogen bonds must be due to a gas-phase effect, with the high-level *ab initio* results almost certainly being preferred over the molecular mechanics ones in this instance.

We note in addition, with regard to the solvation free energy of *tTt* relative to tG^+g^-, that the frozen AM1-SM1a model performs surprisingly poorly, predicting the latter to be less well solvated by a mere 0.1 kcal/mol. This is particularly surprising insofar as this model had the smallest root-mean-square error for the parametrization set employed in the development of the SM*x* models.[35] However, analysis of the data in Table 3 suggests that the AM1-SM1a model, unlike the others, predicts trans hydroxyl groups to be quite a bit less well solvated than gauche ones. The cause of this effect lies in the CDS component of the free energy of solvation. In the SM2 and SM3 models, the hydroxyl proton does not have a surface tension of its own, nor does it block the accessible surface of the hydroxyl oxygen. Rather, it modifies the underlying oxygen surface tension, making it more negative. This approach makes the CDS term insensitive to hydroxyl rotation, other than to the extent that it changes other aspects of molecular structure. The SM1a model, on the other hand, has a unique surface tension for both the oxygen atom and the hydrogen atom, with the latter being smaller than the former by about a factor of 5. As a result, there is a fairly strong tendency for the hydroxyl proton to prefer a sterically congested (eclipsed) location relative to what might be thought of as an oxygen lone pair. We have not observed this nonphysical anomaly before, and it does not appear to have an adverse impact on the absolute free energy of solvation. However, it does appear to make the SM1a model unsuitable for conformational analysis where such rotations are an issue.

The continuum approach employed by Alagona and Ghio also predicts a very small difference in solvation free energies between *tTt* and tG^+g^-, with the former being favored over the latter by only 0.2 kcal/mol.[14] However, this model fails to take account of directional components of hydrogen bonding (i.e., components not considered in an electrostatic only model), the cost of reorganization of the solvent shells, cavitation, or dispersion. The latter two will not differ much for the various conformers, but the former two may well differ. As such, it is instructive to decompose the SM*x* solvation free energies into their individual ENP and CDS components. The former term is most closely related to the portion of the solvation free energy addressed in the calculations of Alagona and Ghio, while the latter provides some insight into the importance of local effects. We emphasize that the scheme used to parametrize the SM*x* model allows the parameters governing the CDS term to relax in order to account for systematic errors in the ENP term (e.g., poor atomic partial charges from the AM1 or PM3 wave functions or the breakdown of the atom-centered generalized Born charge polarization model itself), and as such this decomposition should be taken to be qualitative in nature. Nevertheless, the separate ENP and CDS terms retain semiquantitative significance, and they provide qualitative insight into the origin of solvation effects.

Table 6 provides the separated energetic components for all three models, using the frozen MP2/cc-pVDZ geometries for consistency. The entries labeled "NOPOL" correspond to solvation free energies in the absence of any solute electronic relaxation as a result of solvent-induced polarization, i.e., this term is a static G_P polarization term only.[35] The entries labeled "1SCF" include the effect of electronic relaxation (which proceeds until the internal energy cost relative to the gas-phase optimum begins to exceed the system's energy gain from interaction with the solvent) and are thus ΔG_{EP} energies. Since frozen geometries are being used, neither term includes the nuclear repulsion contribution ΔG_N; however, this effect is very small, as discussed in section 3.

Table 6. Components of the Aqueous Free Energy of Solvation (kcal/mol) for 1,2-Ethanediol Conformers

		AM1-SM1a		AM1-SM2		PM3-SM3	
isomer		ENP	CDS	ENP	CDS	ENP	CDS
tG^+g^-	NOPOL[a]	−1.45	−7.11	−1.42	−7.85	−1.53	−7.11
	1SCF[b]	−1.65	−7.11	−1.61	−7.84	−1.86	−7.10
$g^+G^+g^-$	NOPOL	−1.51	−7.47	−1.47	−7.84	−1.60	−7.12
	1SCF	−1.73	−7.47	−1.67	−7.84	−1.95	−7.10
$g^-G^+g^-$	NOPOL	−1.51	−7.58	−1.49	−7.90	−1.60	−7.17
	1SCF	−1.67	−7.58	−1.65	−7.89	−1.86	−7.16
tTg^+	NOPOL	−1.25	−7.83	−1.24	−9.08	−1.56	−8.26
	1SCF	−1.35	−7.83	−1.35	−9.07	−1.91	−8.25
g^+Tg^-	NOPOL	−1.33	−8.11	−1.33	−9.10	−1.67	−8.30
	1SCF	−1.43	−8.11	−1.43	−9.08	−2.00	−8.28
$g^+G^+g^+$	NOPOL	−1.55	−7.69	−1.53	−7.95	−1.68	−7.21
	1SCF	−1.75	−7.69	−1.73	−7.94	−1.98	−7.20
tG^+g^+	NOPOL	−1.85	−7.42	−1.83	−8.11	−1.84	−7.35
	1SCF	−2.06	−7.42	−2.15	−8.09	−2.34	−7.33
g^+Tg^+	NOPOL	−1.30	−8.17	−1.29	−9.13	−1.60	−8.31
	1SCF	−1.44	−8.17	−1.42	−9.11	−1.99	−8.30
tTt	NOPOL	−1.29	−7.47	−1.28	−9.03	−1.63	−8.21
	1SCF	−1.39	−7.47	−1.39	−9.02	−1.96	−8.20
tG^+t	NOPOL	−1.91	−7.14	−1.87	−8.14	−1.91	−7.37
	1SCF	−2.18	−7.14	−2.15	−8.12	−2.36	−7.36

[a] Gas-phase semiempirical electronic wave function at gas-phase MP2/cc-pVDZ geometry. [b] Frozen calculation: polarized semiempirical electronic wave function at gas-phase MP2/cc-pVDZ geometry.

Table 6 shows clearly that the majority of the solvation free energy comes from the CDS term. Moreover, it is this term which appears to discriminate most effectively between *G* and *T* rotamers—in general the latter are favored by about 1.2 kcal/mol with the SM2 and SM3 models. The SM1a model exhibits anomalous behavior because of the strong dependence on hydroxyl group torsions noted earlier. Thus, almost all of the difference in the solvation free energies for the *tTt* and tG^+g^- is found in this term, and it is not surprising that the continuum approach of Alagona and Ghio failed to observe much difference between the two in an electrostatic sense. Their value of 0.24 kcal/mol for the difference at the HF/6-31G* level is in good agreement with our frozen SM2 and frozen SM3 1SCF values for the ENP term of −0.22 and 0.10 kcal/mol, respectively. There is less agreement among the various models for the difference in solvation free energies between conformers tG^+g^- and $g^+G^+g^-$. Nagy *et al.*[13] predict the latter to be better solvated by 0.95 ± 0.20 kcal/mol, Hooft, van Eijck, and Kroon[26] predict 0.69 ± 0.16, and Alagona and Ghio[14] predict 1.13 and 0.65 at the HF/4-31G and HF/6-31G* levels, respectively. Our SM3 relaxed result predicts this difference to be quite a bit smaller, 0.19 kcal/mol (the frozen models predict about half this). The results from the continuum models are probably slightly less trustworth in this regard. The predictions of Alagona and Ghio show a large basis set effect and do not take account of several other discriminating components of solvation, as mentioned earlier. The SM2 and SM3 models have difficulty in distinguishing changes in the CDS components of solvation arising from conformational changes associated exclusively with proton locations, since they are explicitly recognized only in the ENP portion of the calculation, not the CDS. Of course, the large variation even between the two simulations for this last energy difference suggests that more work needs to be done to determine which models are most accurate in this respect.

We note in closing that the various SM*x* models predict very little difference between the ENP components for any of the four

most important gas-phase conformers. Dissolution in nonpolar solvents, where the ENP effects will be *roughly* halved with respect to water while the CDS terms will be effectively constant for all isomers (although probably different from aqueous values), is thus predicted to have very little effect on the gas-phase equilibrium population. This appears to be entirely consistent with the IR results in inert gas matrices and carbon tetrachloride solution.[8,25] Additionally, the small variation in the ENP energy across all 10 conformers stands in contrast to the behavior of the molecular dipole moment, which ranges from zero for g^+Tg^- and tTt (by symmetry) to 3.05 D for tG^+g^+ (using the MP2/cc-pVDZ density matrix; see supplementary material for complete list of dipole moments). It is thus evident that the widely used Kirkwood–Onsager approach,[57] which models the electrostatic component of the free energy of solvation as being proportional to the square of the molecular dipole moment, would be entirely unsuitable for 1,2-ethanediol.

5. Conclusions

Conformational analysis of 1,2-ethanediol requires high-level treatments, in terms of both basis sets and correlation, in order to accurately predict relative energies in the gas phase. The combination of correlated, extended basis set, gas-phase free energies with the quantum statistical SMx models allows an exploration of the effect of aqueous solvation on all of the possible conformers of the solute at a more reliable level than previously employed methods. The SMx models are validated by noting that they predict an aqueous equilibrium population in close agreement with available experimental results, an absolute free energy of solvation for the solute which is also consistent with experimental trends, and relative free energies of solvation between conformers, with and without internal hydrogen bonds between the vicinal hydroxyl groups, which agree nicely with and extend available results from explicit water simulations.

(57) (a) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351. (b) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486. (c) Kirkwood, J. G. *J. Chem. Phys.* **1939**, *7*, 911.

The calculations show that the population of trans C–C conformations is increased by a factor of 3–5 in aqueous solution as compared to the gas phase, due in large part to first hydration shell effects. The identity of the hydroxyl rotamers that make the dominant contribution for each C–C conformation is not changed by aqueous solvation. The percentage of molecules with no internal hydrogen bond is increased from 17% to only about 23% by aqueous sovlation; thus the solvent water molecules do not compete so strongly for the solute hydroxyls as to disrupt the qualitative conformational picture that results from internal energetic interactions present in the gas-phase molecule—a very important conclusion for the qualitative understanding of diols and polyols.

The ability of the quantum statistical models to accurately treat both solute polarization and first hydration shell effects like hydrogen bonding suggests that these models will prove particularly useful for analyzing the influence of aqueous solvation on similar conformation issues in more complex polyols, like sugars.[3] The advantage of being able to explore many more conformers than would be typical for much more time consuming simulation approaches is especially significant. It appears that continuum solvation models including both volume electric polarization effects and surface hydration shell effects may be very useful for aqueous conformational problems.

Acknowledgment. The authors are grateful to Professors Peter Carr and Lionel Carreira for helpful discussions and to Dr. Igor Topol and Professor Chuck Eckert for providing preprints of ref 17 and 54, respectively. This work was supported in part by the National Science Foundation.

Supplementary Material Available: ASCII-formatted Gaussian92 archive files including Z-matrices and dipole moments for all structures optimized at the MP2/cc-pVDZ level and vibrational frequencies for all structures at the HF/cc-pVDZ level (6 pages). This material is contained in many 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 information.