Ab Initio and Experimental Studies on the Protonation of Glucose in the Gas Phase

Kimberly A. Jebber,[†] Kui Zhang, Carolyn J. Cassady,* and Alice Chung-Phillips*

Contribution from the Department of Chemistry, Miami University, Oxford, Ohio 45056

Received February 9, 1996[⊗]

Abstract: Protonations of α- and β-D-glucopyranose in the gas phase were investigated using the ab initio molecular orbital approach at the HF/6-31G* level with full geometry optimization. Minimum-energy structures of three neutral and six protonated species for each anomer were calculated. Geometries, energies, and intramolecular hydrogen bonding in these structures are discussed. For the neutral species at 298 K the order of stability for the hydroxymethyl conformers is calculated to be *GT* > *GG* > *TG* for the α anomer and *GG* > *GT* > *TG* for the β anomer. Protonated species that least disrupt the internal hydrogen bonding network in the neutral species are considered; these include protonations on the oxygen sites labeled as the hydroxymethyl *O6*, the ring *O5*, and the exocylic hydroxyl *O4*. The *O6* protonation in the *TG* conformation is electronically most favored. Energy corrections for basis-set deficiency and electron-correlation omission in the adopted theoretical procedure were estimated from high-level calculations on ethanol, 2-propanol, and dimethyl ether. In addition, the gas-phase basicity (GB) of glucose was measured by proton transfer reactions in a Fourier transform ion cyclotron resonance mass spectrometer. The experimental GB values for both anomers were determined to be 188 ± 3 kcal/mol. The experimental values are compared with the ab initio estimates of 178−190 and 177−189 kcal/mol for the respective α and β anomers. Theoretical structures for the lowest-electronic-energy protonated species in the three hydroxymethyl conformations of each anomer are also presented to serve as reference data for postulating various kinetic pathways.

Introduction

Carbohydrates constitute one of the most important classes of biomolecules because they provide the major source of energy required by every living organism. The importance of carbohydrate research was demonstrated by the large number of papers presented in the recent symposium titled, Conformation of Carbohydrates: Experiment and Calculation,[1] at the 209th National Meeting of the American Chemical Society. Carbohydrates are one of the most actively investigated groups of biomolecules in the chemical community today.

In the recent past we undertook several mass spectral and ab initio studies on the protonations of small amino acids and peptides in the gas phase.[2−5] These studies led to accurate thermodynamic data for gas-phase basicity (GB) and proton affinity (PA) and a quantitative understanding of the structural properties of the relevant molecular species. Our experience from the peptide work provided the impetus to continue protonation studies on carbohydrates. We report here our first in-depth study on glucose, a monosaccharide.

Earlier experiments on glucose include neutron and X-ray diffraction determinations of crystal structures,[6] as well as optical rotation and nuclear magnetic resonance (NMR) analyses of anomeric[7] and rotamer[8] distributions in aqueous solution. Preliminary results on the GBs of glucose, maltose, cellobiose, sucrose, and lactose were first obtained from our laboratory using a Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer.[9a] For these compounds ionization by fast atom bombardment (FAB) gave low yields of protonated ions because these species readily fragmented by H_2O loss; this resulted in relatively large error ranges for the GB values. Therefore, in the present study additional experiments were carried out to refine the preliminary data. Because of the difficulties in performing the mass spectral measurements for carbohydrates, theoretical calculations are particularly important to verify experiments.[9b]

Computational research on carbohydrates has been conducted mainly in the force-field and semiempirical MO arenas. There have been activities in computer modeling,[10] MM2/MM3[11] and MNDO/AM1/PM3[12,13] calculations on individual molecules, and exploratory studies of solvent effects on glucose using molecular dynamics[14] and quantum statistical[13] methods. To our knowledge neutral glucose and its simple derivatives are the only

[†] Current address: Lilly Laboratory for Clinical Research, Indianapolis, IN 46202.

[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1996.

(1) Sponsored jointly by the Division of Computers in Chemistry and the Division of Carbohydrate Chemistry. See Final Program for 209th ACS National Meeting in Anaheim, CA, April 2−6, 1995: *Chem. Eng. News* **1995**, *73*, 41.

(2) Zhang, K.; Zimmerman, D. M.; Chung-Phillips, A.; Cassady, C. J. *J. Am. Chem. Soc.* **1993**, *115*, 10812.

(3) McKiernan, J. W.; Beltrame, C. E. A.; Cassady, C. J. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 718.

(4) Zhang, K.; Cassady, C. J.; Chung-Phillips, A. *J. Am. Chem. Soc.* **1994**, *116*, 11512.

(5) Cassady, C. J.; Carr, S. R.; Zhang, K.; Chung-Phillips, A. *J. Org. Chem.* **1995**, *60*, 1704.

(6) (a) Brown, G. M.; Levy, H. A. *Science* **1965**, *147*, 1038. (b) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1968,** *B24*, 830.

(7) Angyal, S. J. *Aust. J. Chem.* **1968**, *21*, 2737; *Angew. Chem., Int. Ed. Engl.* **1969**, *8*, 157.

(8) Nishida, Y.; Ohrui, H.; Meguro, H. *Tetrahedron Lett.* **1984**, *25*, 1575.

(9) (a) Cassady, C. J. *Gas-Phase Basicities of Small Carbohydrates*; The 41st ASMS Conference on Mass Spectrometry, San Francisco, CA, May 31, 1993. (b) Jebber, K. A.; Chung-Phillips, A. *Computational Studies of Neutral and Protonated* α- *and* β-*D-Glucose*; The 208th ACS National Meeting, Washington, DC, August 1994.

(10) See, e.g.: (a) *Computer Modeling of Carbohydrate Molecules, ACS Symposium Series, No. 430;* French, A. D., Brady, J. W., Eds.; American Chemical Society; Washington, DC, 1990. (b) Senderowitz, H.; Parish, C.; Still, W. C. *J. Am. Chem. Soc.* **1995**, *118*, 2078.

(11) (a) Dowd, M. K.; Reilly, P. J.; French, A. D. *J. Comput. Chem.* **1992**, *13*, 102. (b) Koca, J.; Perez, S.; Imberty, A. *J. Comput. Chem.* **1995**, *16*, 296.

(12) Khalil, M.; Woods, R. J.; Weaver, D. F.; Smith, V. H., Jr. *J. Comput. Chem.* **1991**, *12*, 584.

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

(14) (a) Ha, S.; Gao, J.; Tidor, B.; Brady, J. W.; Karplus, M. *J. Am. Chem. Soc.* **1991**, *113*, 1553. (b) Eijck, B. P. v.; Hooft, R. W. W.; Koon, J. *J. Phys. Chem.* **1993**, *97*, 12093.

S0002-7863(96)00427-1 CCC: $12.00

carbohydrate species that have been studied with the ab initio approach.[15−17] Clearly, the large and complex carbohydrate structures are the major obstacles to ab initio calculations.[18]

We now recap here aspects of previous ab initio studies of relevance to our study. Polavarapu and Ewig[15] initiated a comprehensive investigation on the α and β anomers of glucose at the Hartree−Fock level using the split-valence 4-31G basis set with geometry optimizations (i.e., at the HF/4-31G level[18]). They discussed the geometries, energies, dipole moments, and thermodynamic properties of five anomeric pairs of low-energy hydroxymethyl conformers plus several higher-energy conformers. Subsequently Salzner and Schleyer[16] added polarization functions to the basis set (at the HF/6-31G* level) in optimizing the geometries of three anomeric pairs of conformers in a detailed examination of the anomeric effect. Recently, Barrows, Dulles, Cramer, French, and Truhlar[17] obtained optimized structures of four conformers of the β anomer using a very large basis set and the second-order Møller−Plesset perturbation approach for electron correlation[18] (at the MP2/cc-pVDZ level) to study the relative stability of selected conformations in the $^4C_1$ and $^1C_4$ chair forms of glucose.

Previous ab initio studies on glucose[15−17] help us define a reasonable and practical scope for this investigation: to study protonations of the three lowest-energy hydroxymethyl conformers of the α and β anomers in the $^4C_1$ chair form at the HF/6-31G* level of theory with geometry optimizations. The completed theoretical work for comparison with the experimental GBs is presented here, while additional theoretical studies on glucose structures are reported elsewhere.[19] This work is the first ab initio study to date on the protonation of glucose. In view of the potential applications of these ab initio results to molecular modeling on polysaccharides, the calculated energies and geometries are carefully examined to gain an understanding of the relationship between intramolecular hydrogen bonding and the relative stability of different glucose species.

Computational Methods

Ab initio MO calculations were carried out employing the Gaussian 92 program[20] on the IBM ES/9121/480 and DEC 4000−710 AXP computers at Miami University and the CRAY Y-MP8/864 computer at the Ohio Supercomputer Center. The computational procedure is one evolved from our previous investigations on peptides.[5]

For the glucose molecules minimum-energy structures of neutral and selected protonated species ($C_6H_{12}O_6$ and $C_6H_{13}O_6^+$) are determined at the HF/3-21G and HF/6-31G* levels with full geometry optimizations. Vibrational frequencies are obtained at the HF/6-31G* level for the neutral species and at the HF/3-21G level for both neutral and protonated species. The reason for employing the 3-21G basis set is to spare the costly vibrational frequency computations associated with the larger 6-31G* set for the numerous protonated species involved in this protonation study. Consequently the HF/6-31G* electronic energies and HF/3-21G thermodynamic properties, at their respective optimized geometries, are used for the GB and PA calculations: this we call the medium level (Level M) calculation.

For the smaller model compounds (ethanol, 2-propanol, and dimethyl ether) a higher level (Level H) is introduced where geometry optimizations are carried out at the MP2/6-31+G** level. The GBs are determined from the MP4/6-31+G(2d,2p) electronic energies and MP2/6-31+G** thermodynamic properties, both at the MP2/6-31+G** optimized geometries. Here MP2 = FULL.[18] The difference in the GBs calculated from Level H and Level M for each model compound is used as a "correction" value for the comparable type of oxygen involved in the protonation of glucose.

Experimental Methods

All experiments were performed using a Bruker CMS-47X FT-ICR mass spectrometer.[21] Protonated glucose ions were produced in an external source by FAB with a 6−10-kV beam of xenon or argon atoms/ions from a Phrasor Scientific FAB gun.[22] The α and β anomers of glucose were analyzed separately and each were dissolved in a glycerol matrix prior to analysis.

The ions were transferred from the external source into the FT-ICR cell by electrostatic focusing. Protonated glucose ions (MH^+) were mass-selected by resonant frequency ejection techniques[23] and allowed to react with reference compounds (B) present at static pressures of $(2-20) \times 10^{-8}$ Torr. The major process observed was deprotonation, reaction 1:

$$MH^+ + B \rightarrow M + BH^+ \quad (1)$$

Glucose MH^+ were produced in low abundance by FAB. Although broadband measurements were made to observe product formation, high-resolution measurements were employed to obtain MH^+ intensity measurements for use in kinetics calculations. (In our mass spectrometer, high resolution is a more sensitive mode than broadband.) In addition, ions that had the same nominal mass as MH^+ were sometimes produced by reactions during the ion accumulation period between the B neutrals and FAB matrix ions. Narrow band measurements eliminated this problem because the resolving power was adequate to separate the interfering ions from MH^+.

Rate constants for reaction 1 were determined by monitoring the pseudo-first-order change in MH^+ intensity as a function of reaction time at a constant pressure. Pressures were measured with a calibrated ionization gauge.[4,24] Reported reaction efficiencies are the ratio of the experimental rate constant to the collision rate constant that was obtained from the average dipole orientation model.[25] All experiments were performed at room temperature (*ca.* 298 K).

Results

The numbering scheme for the atoms of a neutral D-glucose molecule in the pyranose form is illustrated in Figure 1. The two anomers of D-glucopyranose are distinguished by the position of the C1−O1 bond relative to the ring: axial for α and equatorial for β. For either anomer three classes of low-energy conformers arise from internal rotation about the C5−C6 bond. The conformers are named according to the orientations of the hydroxymethyl C6−O6 bond relative to the C5−O5 and C4−C5 bonds in the ring: trans gauche (*TG*), gauche gauche (*GG*), and gauche trans (*GT*). Using these notations the structure in Figure 1 is designated as *βGG*, which corresponds to the *GG* conformer of β-D-glucopyranose. The same naming scheme is followed for the other neutral species (see Figure 2). The name of a protonated species is derived from the name of its parent neutral species appended by the numeral *n* to denote *On* as the protonation site, e.g., an *O6*-protonated

(15) Polavarapu, P. J.; Ewig, C. S. *J. Comput. Chem.* **1992**, *13*, 1255.
(16) Salzner, U.; Schleyer, P. v. R. *J. Org. Chem.* **1994**, *59*, 2138.
(17) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219 .
(18) Hehre, W. J.; Radom, L.; Schleyer, P. v. R.; Pople, J. A. *Ab initio Molecular Orbital Theory*; Wiley: New York, 1986.
(19) Chung-Phillips, A.; Zhang, K.; Jebber, K. A. Manuscript in preparation.
(20) 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. *Gaussian 92*, Gaussian: Pittsburgh, 1992.
(21) Kofel, P.; Allemann, M.; Kellerhals, H. P.; Wanczek, K. P. *Int. J. Mass Spectrom. Ion Proc.* **1985**, *65*, 97.
(22) Perel, J.; Faull, K.; Mahoney, J. F.; Tyler, A. N.; Barchas, J. D. *Am. Lab.* **1984**, *16*, 94.
(23) Comisarow, M. B.; Grassi, V.; Parisod, G. *Chem. Phys. Lett.* **1978**, *57*, 413.
(24) Bartmess, J. E. In *Structure/Reactivity and Thermochemistry of Ions*; Ausloss, P., Lias, S. G., Eds.; Reidel: Dordrecht, The Netherlands, 1987; pp 367−371.
(25) Su, T.; Bowers, M. T. *Int. J. Mass Spectrom. Ion Phys.* **1973**, *12*, 347.

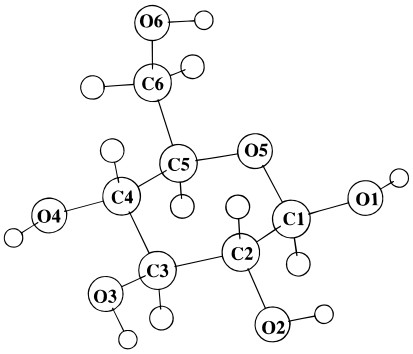

Figure 1. Neutral β-D-glucopyranose in the *GG* conformation (βGG). Carbon and oxygen atoms are identified with atom labels. Hydrogen atoms are numbered in a clockwise direction first for those bonded to carbons and next for those bonded to oxygens. Numbering begins at C1−C6 for H1−H7 and at O1−O6 for H8−H12. In a protonated species the additional hydrogen is specified as H13.

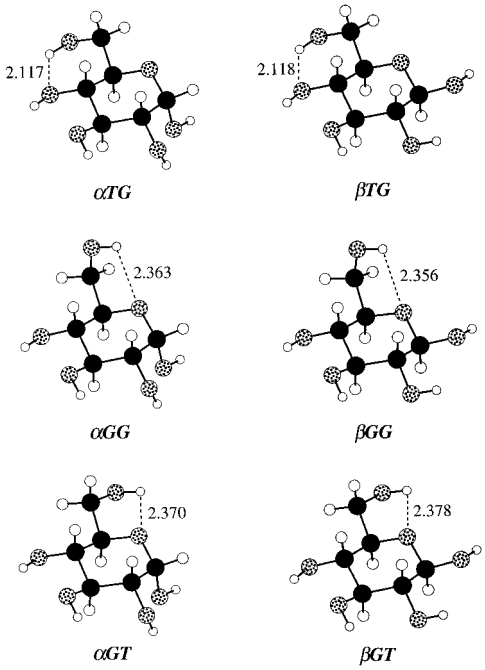

Figure 2. HF/6-31G* structures for the neutral α- and β-D-glucopyranose. The H-bond distance of O6−H···O4 or O6−H···O5 is shown in Å.

BGG is called $\beta GG6$ (see Figure 3). For the ring oxygen *O5*, the modifier *a* or *e* is used to show whether proton addition is axial or equatorial, e.g., $\alpha TG5a$ (see Figure 4).

The most stable structures for the three hydroxymethyl conformations of the neutral α and β anomers, obtained at the HF/6-31G* level with full geometry optimizations, are shown in Figure 2. Optimized internal coordinates (bond lengths, bond angles, and dihedral angles) for the lowest-electronic-energy conformers of the two anomers, αTG and βGG, and atomic Cartesian coordinates for all six neutral species are provided as Supporting Information in Tables S-1 and S-2 to facilitate future computational work on polysaccharides. The most stable protonated species that conserve the conformations of the neutral species (Figure 2) are given in Figures 3 and 4; these are employed in the mechanistic study of the simplest elementary step in a protonation reaction. Average bond lengths and average absolute bond deviations are given separately for the neutral and protonated species in Table 1. These data are used for the discussion on geometries.

The structures in Figures 2 exhibit a network of intramolecular hydrogen interactions in the form O*i*−H···O*j*: counterclockwise

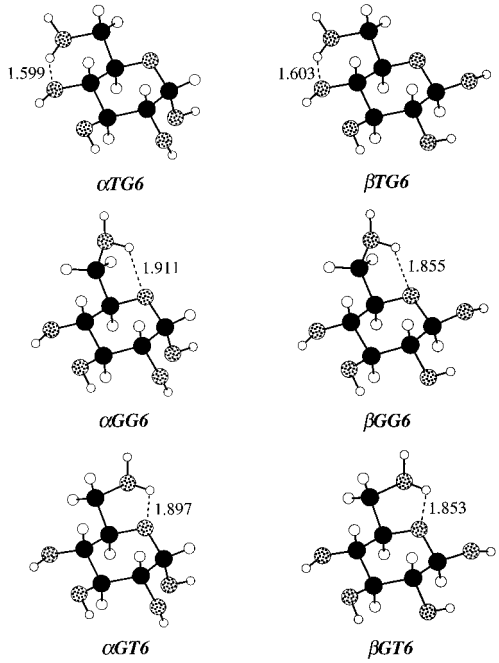

Figure 3. HF/6-31G* structures for the *O6*-protonated α- and β-D-glucopyranose. The H-bond distance of O6−H···O4 or O6−H···O5 is shown in Å.

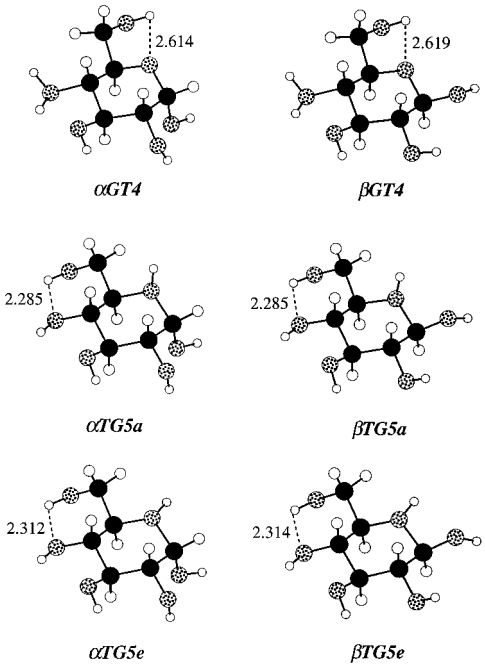

Figure 4. HF/6-31G* structures for selected *O4*- and *O5*-protonated α- and β-D-glucopyranose. The H-bond distance of O6−H···O5 or O6−H···O4 is shown in Å.

O1−H···O5, O2−H···O1, O3−H···O2, and O4−H···O3 and clockwise O5−H···O1 for the four exocyclic hydroxyl groups; counterclockwise O6−H···O4 and clockwise O6−H···O5 for the hydroxymethyl chain. Because intramolecular hydrogen bonding involving the hydroxymethyl O6 is particularly important in the interpretation of gas-phase mass spectral data, the distances for H···O in O6−H···O4 and O6−H···O5 are explicitly shown in the figures. A complete listing of H···O distances relevant to subsequent discussions on the relative stability of glucose conformers is given in Table 2.

Ab initio energy quantities for the individual species obtained at different theoretical levels are presented in Tables 3 and 4 for glucose and as Supporting Information in Table S-3 for the

Table 1. HF/6-31G* Average Bond Lengths and Deviations (Å) for α- and *β*-D-Glucopyranose

bonds[a]		total no.[b]	α		*β*	
			length	dev	length	dev
		Neutral Species[c]				
C−C		15	1.522	0.003	1.521	0.003
C−H		21	1.085	0.003	1.087	0.002
C−O	C5−O5	3	1.421	0.003	1.414	0.002
	C1−O5	3	1.386	0.000	1.394	0.000
	C1−O1	3	1.393	0.000	1.378	0.000
	C*i*−O*i*, *i* = 2, 3, 4, 6	12	1.400	0.002	1.399	0.002
O−H		15	0.949	0.001	0.949	0.000
		O6-, *O5*-, and *O4*-Protonated Species				
C−C		30	1.524	0.005	1.523	0.003
C−H		42	1.083	0.004	1.084	0.004
C−O	C5−O5	4	1.406	0.007	1.399	0.006
	C1−O5	4	1.415	0.008	1.426	0.007
	C1−O1	4	1.378	0.002	1.362	0.000
	C*i*−O*i*, *i* = 2, 3, 4, 6	20	1.390	0.005	1.390	0.004
O−H		26	0.951	0.001	0.950	0.001
C−O*	C5−O5*	2	1.514	0.001	1.499	0.002
	C1−O5*	2	1.528	0.007	1.537	0.002
	C1−O1 (O5*)[d]	2	1.340	0.004	1.336	0.001
	C*i*−O*i**, *i* = 4, 6	4	1.513	0.008	1.511	0.008
O*−H		10	0.967	0.010	0.968	0.010

[a] See Figures 1−4 and text. A neutral species has seven C−H, five O−H, five C−C, and seven C−O bonds. A protonated species has one O−H bond added. [b] Total number of bonds in the three neutral species or the six protonated species used in the averaging. [c] Experimental solid-state average bond lengths in Å are as follows: C−C = 1.523, C−H = 1.098, C5−O5 = 1.428, C1−O5 = 1.428, C1−O1 = 1.389, C*i*−O*i* = 1.418, OH = 0.968 for α (ref 6a); and C−C = 1.539, C−H = 1.060, C5−O5 = 1.437, C1−O5 = 1.437, C1−O1 = 1.383, C*i*−O*i* = 1.425, OH = 0.967 for *β* (ref 6b). [d] The C1−O1 bond in a O5 protonated species.

Table 2. HF/6-31G* Intramolecular Hydrogen Interactions in α- and *β*-D-Glucopyranose[a]

structure	O1−H···O5 (O5−H···O1)	O2−H···O1 (O1−H···O2)	O3−H···O2 (O2−H···O3)	O4−H···O3 (O3−H···O4)	O6−H···O4 (O4−H···O6)	O5−H···O6 (O6−H···O5)
			Neutral Species			
α*TG*	2.529	2.238	2.512	2.391	2.117	
α*GG*	2.533	2.254	2.494	2.448		(2.363)
α*GT*	2.529	2.248	2.501	2.413		(2.370)
*β*TG	2.437	2.531	2.493	2.398	2.118	
*β*GG	2.461	2.525	2.472	2.460		(2.356)
*β*GT	2.475	2.525	2.480	2.424		(2.378)
			O6-Protonated Species			
α*TG6*	2.652	2.295	2.563	2.431	1.599*	
α*GG6*	2.662	2.335	2.570	2.463		(1.911*)
α*GT*6	2.712	2.336	2.563	2.480		(1.897*)
*β*TG6	2.559	2.624	2.535	2.425	1.603*	
*β*GG6	2.639	2.666	2.531	2.455		(1.855*)
*β*GT6	2.692	2.660	2.532	2.375		(1.853*)
			O4- and *O5*-Protonated Species			
α*GT4*	2.637	2.271	2.599	2.052*		(2.614)
α*TG5a*	2.688	2.465	2.617	2.397	2.285	
α*TG5e*	2.831 (2.676*)	2.470	2.608	2.422	2.312	
*β*GT4	2.562	2.643	2.560	2.050*		(2.619)
*β*TG5a	2.719 (2.375*)	2.843	2.587	2.383	2.285	
*β*TG5e	2.849 (2.429*)	2.864	2.581	2.412	2.314	

[a] See Figures 1−4. Distance H···O*j* in the interacting group O*i*−H···O*j* is given in Å. Double headings are used for each column to include two interactions in opposite orientations from O*i* to H to O*j*: top for counterclockwise (without parentheses); bottom for clockwise (with parentheses). The asterisk indicates an interaction involving a protonated O*i* in O*i*−H.

model compounds. Representing these quantities are the following: (a) electronic energy (E_{elec}) that determines the optimized geometrical structure at 0 K; and (b) thermodynamic properties including zero point energy (E_{ZP}), energy change from 0 to 298.15 K ($E - E_0$), and entropy at 298.15 K (S). In view of a general interest in the relative energies and distributions of different conformers of neutral glucose at room temperature, the thermodynamic contribution (E_{therm}),

$$E_{therm} = E_{ZP} + (E - E_0) - TS \qquad (2)$$

and Gibbs free energy (G),

$$G = E_{elec} + E_{therm} + RT \qquad (3)$$

where RT = 0.59249 kcal/mol for T = 298.15 K, are also calculated and shown in Table 3.

For the protonation study, $M + H^+ \rightarrow MH^+$, the gas-phase basicity and proton affinity correspond to the changes in Gibbs free energy and enthalpy of the reaction. These quantities, in kcal/mol, may be expressed as follows:[5]

Table 3. Ab Initio Calculations for Neutral α- and β-D-Glucopyranose[a]

A. HF/6-31G* Electronic and Thermodynamic Properties

structure	E_{elec} (hartree)	E_{ZP} (kcal/mol)	$E - E_0$ (kcal/mol)	S [cal/(K·mol)]
α*TG*	−683.334 048 0	135.259	7.050	102.206
α*GG*	−683.333 865 0	135.101	7.129	102.810
α*GT*	−683.333 736 5	135.048	7.153	102.991
β*TG*	−683.332 190 4	134.835	7.172	103.087
β*GG*	−683.332 273 9	134.711	7.235	103.629
β*GT*	−683.331 952 2	134.658	7.264	103.833

B. HF/6-31G* Relative Energies (kcal/mol) and Distributions

structure	E_{elec}	E_{therm}[b]	G	exp(−G/RT)	structure	E_{elec}	E_{therm}	G	exp(−G/RT)
α*TG*	0.000	0.000	0.000	1.000	β*TG*	0.000	0.000	0.000	1.000
α*GG*	0.115	−0.260	−0.145	1.277	β*GG*	−0.052	−0.222	−0.274	1.588
α*GT*	0.195	−0.342	−0.147	1.282	β*GT*	0.149	−0.308	−0.159	1.308

C. Anomeric Distributions in a Gaseous Mixture

	HF/4-31G					HF/6-31G*			
structure	E_{elec}[c]	E_{therm}[c]	G	exp(−G/RT)	structure	E_{elec}	E_{therm}	G	exp(−G/RT)
α*TG*	0.00	0.00	0.00	1.00	α*TG*	0.000	0.000	0.000	1.000
α*GG*	0.34	−0.36	−0.02	1.03	α*GG*	0.115	−0.260	−0.145	1.277
α*GT*	0.67	−0.46	0.21	0.70	α*GT*	0.195	−0.342	−0.147	1.282
β*TG*	2.49	−0.68	1.81	0.05	β*TG*	1.166	−0.565	0.601	0.363
β*GG*	2.51	−0.93	1.58	0.07	β*GG*	1.113	−0.787	0.326	0.577
β*GT*	3.15	−1.10	2.05	0.03	β*GT*	1.315	−0.872	0.443	0.473

[a] See Figure 2 and text. [b] In calculating E_{therm}, E_{ZP} is not scaled for more direct comparisons with values given in refs 13 and 15. [c] Deduced from values listed in Tables I and II of ref 15.

Table 4. Ab Initio Calculations for the Protonations of α- and β-D-Glucopyranose[a]

A. HF/6-31G* Electronic Energies[b]

structure	E_{elec} (hartree)	ΔE_{elec} (kcal/mol)	structure	E_{elec} (hartree)	ΔE_{elec} (kcal/mol)
α*TG6*	−683.664 877 2	−207.60	β*TG6*	−683.661 819 9	−206.85
α*GG6*	−683.655 327 1	−201.72	β*GG6*	−683.654 755 3	−202.36
α*GT6*	−683.654 873 6	−201.52	β*GT6*	−683.653 362 4	−201.69
α*GT4*	−683.646 077 9	−196.00	β*GT4*	−684.642 449 5	−194.84
α*TG5a*	−683.641 889 7	−193.17	β*TG5a*	−683.647 191 5	−197.67
α*TG5e*	−683.643 451 8	−194.15	β*TG5e*	−683.645 742 0	−196.76

B. HF/3-21G Thermodynamic Properties

structure	E_{ZP} (kcal/mol)	$E - E_0$ (kcal/mol)	S [cal/(K·mol)]	ΔE_{therm}[c] (kcal/mol)	$-T\Delta S$[d] (kcal/mol)
α*GG*	132.568	7.132	102.710		
α*GG6*	139.851	7.201	102.977	6.62	7.68
α*GT*	132.450	7.193	103.312		
α*GT4*	140.049	7.269	103.706	6.87	7.64
α*TG*	132.826	7.027	102.060		
α*TG5*	139.661	7.444	104.561	5.89	7.01
β*GG*	132.046	7.321	104.153		
β*GG6*	138.982	7.345	104.326	6.28	7.71
β*GT*	131.878	7.406	104.923		
β*GT4*	139.474	7.476	105.329	6.86	7.64
β*TG*	132.210	7.259	103.764		
β*TG5*	139.247	7.576	105.662	6.15	7.19

[a] See Figures 2−4 and text. This table provides data for Level M calculations: HF/6-31G*//HF/6-31G* for ΔE_{elec} and HF/3-21G//HF/3-21G for ΔE_{therm}. [b] E_{elec} for the respective neutral species are given in Table 3. [c] In calculating E_{therm} using the HF/3-21G values, E_{ZP} is scaled by a factor of 0.91. See ref 5. [d] See eq 6.

$$GB = -(\Delta E_{elec} + \Delta E_{therm} + 6.28) \quad (4)$$

$$PA = GB - T\Delta S \quad (5)$$

The ΔE_{elec} and ΔE_{therm} terms refer to the change $M \rightarrow MH^+$ while the $-T\Delta S$ term is for protonation

$$-T\Delta S = -298.15[S(MH^+) - S(M)] + 7.76 \quad (6)$$

Results for the protonation calculations are summarized in Tables 4 and 5 for glucose and Table 6 for the model compounds.

For glucose protonation the protonated species MH^+ is taken to have the same conformation (*TG*, *GG*, or *GT*) as the neutral species M, e.g., α*TG* → α*TG6*; the protonation data in Table 4 (ΔE_{elec}, ΔE_{therm}, and $-T\Delta S$) follow this formal definition. However, only one calculation for each selected protonation site of a given anomer is performed in order to reduce the number of vibrational frequency calculations; the resulting thermodynamic terms are then shared by all the protonated conformers of the given anomer at this protonation site. Hence the thermodynamic properties are identified by the protonation site (*O6*, *O5*, or *O4*) of an anomer regardless of the specific conformation. For the *O6*-protonation on α, e.g., the ΔE_{therm} and $-T\Delta S$ values listed in Table 4B for the α*GG* → α*GG6* protonation are used for the α*TG* → α*TG6* and α*GT* → α*GT6*

Table 5. Ab Initio Gas-Phase Basicities and Proton Affinities (in kcal/mol) for α- and β-D-Glucopyranose: Comparisons with Experiments[a]

protonation site	structures: neutral → protonated	GB HF/6-31G*	GB adjusted	PA HF/6-31G*	PA adjusted
	α-D-Glucopyranose				
O6	α*TG* → α*TG6*	194.70	190	202.38	198
	α*GG* → α*GG6*	188.82	184	196.50	193
	α*GT* → α*GT6*	188.62	184	196.30	192
O4	α*GT* → α*GT4*	182.85	178	190.49	187
O5	α*TG* → α*TG5a*	181.00	177	188.01	185
	α*TG* → α*TG5e*	181.98	178	188.99	186
	β-D-Glucopyranose				
O6	β*TG* → β*TG6*	194.29	189	202.00	198
	β*GG* → β*GG6*	189.80	185	197.51	194
	β*GT* → β*GT6*	189.13	184	196.84	193
O4	β*GT* → β*GT4*	181.70	177	189.34	186
O5	β*TG* → β*TG5a*	185.24	181	192.43	189
	β*TG* → β*TG5e*	184.33	180	191.52	188
	Experimental GB for α or β anomer:	188 ± 3			

[a] The "HF/6-31G*" value refers to the Level M calculation employing data provided in Table 4. The "adjusted" value is the "HF/6-31G*" value plus the "correction"; the latter is deduced in Table 6 for each oxygen type.

Table 6. Ab Initio Results (in kcal/mol) for the Protonations of Ethanol, Dimethyl Ether, and 2-Propanol: Comparisons with Experiments[a]

A. Ab Initio Results

	O type:	CH_3CH_2OH *n*-alcohol	$(CH_3)_2CHOH$ *sec*-alcohol	$(CH_3)_2O$ ether
Level M:[b]	ΔE_{elec}	−195.42	−200.05	−199.28
	ΔE_{therm}	5.87	6.01	6.37
	GB	183.27	187.76	186.63
Level H:[c]	ΔE_{elec}	−192.18	−196.81	−196.52
	ΔE_{therm}	7.58	7.26	7.82
	GB	178.32	183.27	182.42
experimental GB:[d]		180.2	183.4	184.3

B. Corrections on GB for Glucose

	protonation site:	O6	*O4*	*O5*
for Level M → Level H, ΔGB:		−4.95	−4.49	−4.25

[a] Deduced from the ab initio data provided in Table S-3. [b] Same as the glucose calculations (Tables 4 and 5). [c] Data for Level H calculations: MP4/6-31+G(2d,2p)//MP2(FU)/6-31+G** for ΔE_{elec} and MP2(FU)/6-31+G**//MP2(FU)/6-31+G** for ΔE_{therm}. In calculating E_{therm} using the MP2(FU)/6-31+G** value, E_{ZP} is scaled by a factor of 0.96. See ref 5. [d] Reference 26.

protonations as well. This approximation may be taken as reasonable considering that variations in thermodynamic properties among different conformers are usually small compared with the variation in electronic energy; the former therefore have a smaller impact on the calculated GB and PA.

In the interest of exploring different protonation mechanisms, HF/6-31G* optimized structures for the protonated species with the lowest electronic energies in the three hydroxymethyl conformations of each anomer are presented in Figure 5. Additional information on these structures is provided in a separate publication.[19]

The experimental reaction efficiencies for the deprotonation of the α and β anomers of D-glucopyranose are given in Table 7. For both anomers, very little reaction is seen with 2-methylpropene (GB = 187.3 kcal/mol[26]) but a greater than 10-fold

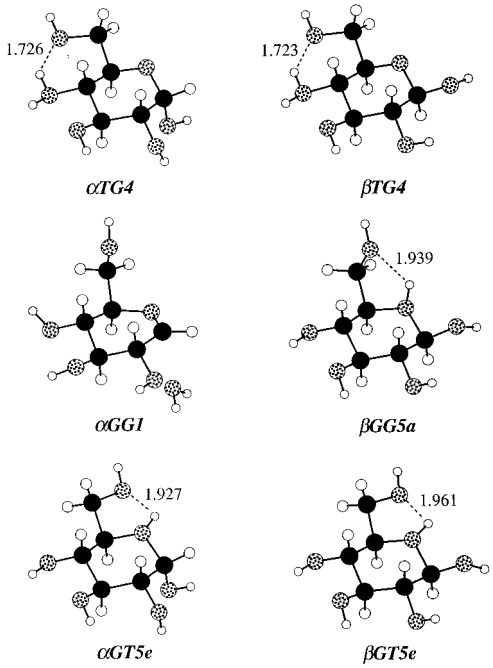

Figure 5. HF/6-31G* structures for the lowest-electronic-energy protonated species in each hydroxymethyl conformation of α- and β-D-glucopyranose. The shorter H-bond distance involving the protonated *O4* or *O5* is shown in Å.

Table 7. Experimental Efficiencies for the Reactions of α- and β-D-Glucopyranose with Reference Compounds

ref compd	GB[a] (kcal/mol)	reaction efficiency α	reaction efficiency β
2-propanol	183.4	0.09	0.12
2-methyl-2-propanol	185.9	0.12	0.10
1,4-dioxane	186.0	0.14	0.12
2-methylpropene	187.3	0.02	0.03
acetone	188.9	0.40	0.44
methyl acetate	190.0	0.35	0.29
tetrahydrofuran	191.4	0.46	0.51
ethyl acetate	192.0	1.10	0.83
ammonia	195.6	0.77	0.82

[a] Reference 26.

increase in the rate of reaction occurs with acetone (GB = 188.9 kcal/mol[26]). Therefore, the experimental GBs of the glucose anomers are assigned as 188 kcal/mol, which is intermediate to the GBs to these reference compounds. The reported uncertainty of ±3 kcal/mol is the difference in GB between glucose and its bracketing reference compounds plus an additional 2.0 kcal/mol to account for any experimental error in the reference GB values. The slow, but observable reactions of protonated glucose ions with the less basic 2-propanol, 2-methyl-2-propanol, and 1,4-dioxane primarily involve proton-bound dimer formation (MHB^+). This process is observed for near thermoneutral reactions of protonated peptides[2−5] and disaccharides[9a] with oxygen or nitrogen containing compounds, but does not occur appreciably with a hydrocarbon such as 2-methylpropene.

The α- and β-D-glucopyranose were analyzed separately by mass spectrometry. As the results in Table 7 indicate, no significant difference was seen in the rates of deprotonation of these two species. However, some mutarotation may occur when a pure anomer is dissolved in glycerol, making it possible that the solutions analyzed were mixtures of the two anomers.

(26) Lias, S. G.; Liebman, J. F.; Levin, R. D. *J. Phys. Chem. Ref. Data* **1984**, *13*, 695.

Discussion

Neutral Species. We first focus on the HF/6-31G* geometries and electronic energies of the hydroxymethyl conformers *TG*, *GG*, and *GT* of α- and *β*-D-glucopyranose in Figure 2. Statistical data presented in Table 1 show remarkably constant C−C, C−H, C−O, and O−H bond lengths, with average absolute deviations no greater than 0.003 Å for each anomer. Note that the C−O bonds are listed separately for those bonded to the ring oxygen (O5), the anomeric carbon (C1), and the hydroxyl and hydroxymethyl oxygens (O*i*, *i* = 2, 3, 4, and 6).

Even though the calculated average bond lengths refer to isolated (or gas-phase) molecules, a comparison with those deduced from experimental solid-state values[6] is still beneficial. (The experimental values pertinent to this discussion are provided in footnote *c* of Table 1.) For the α anomer the HF/6-31G* data show average absolute differences of 0.001 Å for C−C, 0.01 Å for C−H, 0.01−0.02 Å for C−O (excluding C5−O5, C1−O5, and C1−O1), and 0.02 Å for O−H from the neutron diffraction data;[6a] in particular, the calculated C−O and O−H bond lengths are about 0.02 Å shorter than experiments. These discrepancies are consistent with those found for small compounds containing highly electronegative elements (such as oxygen) at the HF/6-31G* level of theory.[18] Similar conclusions are reached for the *β* anomer[6b] except that the X-ray values for the C−H bond appear to be unusually small. (The calculated bond lengths usually compare better with data from neutron diffraction than X-ray because the former focusses on nuclear positions whereas the latter depends on electron densities. The solid-state C−H and O−H bonds in the *β* anomer are understandably shorter because of the X-ray experiments.[6b])

The HF/6-31G* geometries show a difference between the two anomers in the relative lengths of the two bonds associated with the anomeric carbon: C1−O5 is shorter than C1−O1 in α while the reverse is true in *β* (Table 1). This difference, though small, is in agreement with the anomeric effect in dihydroxymethane, dimethoxymethane, 2-hydroxytetrahydropyran, 2-methoxytetrahydropyran, and D-glucopyranose discussed in the literature.[15−17,27,28] Yet, the solid-state data indicate that both α and *β* anomers have a longer C1−O5 bond than the C1−O1 bond.

Bond angles and dihedral angles of α*TG* and *βGG* in Table S-1 provide typical values for the two anomers. Generally the calculated values for the same angle among the three conformers of each anomer stay nearly the same except those that differentiate the specific conformation (*TG*, *GG*, or *GT*). Comparisons with the experimental bond angles pertaining to the central ring show average absolute difference of *ca.* 1° for C2−C1−O5, C3−C2−C1, C4−C3−C2, and C5−C4−C3. Larger differences are found for the C5−O5−C1 angle: the calculated average values are 117.2° (α) and 114.9° (*β*) vs. the experimental[6] 113.8° (α) and 112.7° (*β*). Again, the calculated C5−O5−C1 angle for the α anomer conforms with the anomeric effect. Owing to the existence of extensive intermolecular hydrogen bonding in the crystalline states of glucose, comparisons with the measured values for the endocyclic bond and torsional angles are not made. (Ideally, the way to access the accuracy of the calculated gas-phase geometry is to optimize the geometry of the molecule in the crystal environment[17] and compare the resulting structure with the experimental solid-state structure.)

The hydrogen···oxygen distances in Table 2 for intramolecular hydrogen interactions in neutral glucose species of Figure 2 fall in the range of 2.12−2.53 Å. The H···O*j* distance in O*i*−H···O*j* roughly increases with decreasing number of intervening carbon atoms between oxygens O*i* and O*j*. Accordingly the hydrogen interactions are subdivided into three classes: (a) O6···O4 in 1,5-positions; (b) O6···O5, O4···O3, O3···O2, and O2···O1 in 1,4-positions; and (c) O1···O5 in 1,3-positions. Applying the widely accepted criterion for a H-bond[29] that the O*i*−H···O*j* angle be greater than 90°, we find that the relevant angles in glucose allow class (a) to be a H-bond,[17] class (b) a borderline H-bond, and class (c) not a H-bond. The relative strengths of the H···O attraction are expected to be largest in class (a) and smallest in class (c), where the nature of interaction is presumably electronic in class (a) and electrostatic in class (c). For ease of discussion we simply address classes (a) and (b) as H-bonds and class (c) as a hydrogen interaction.

A shorter hydrogen bonding distance usually implies a stronger H-bond and thus greater stability. The shortest H-bonds attained by O6−H in the *TG* conformation (O6−H···O4 = 2.117 Å in α*TG* and 2.116 Å in *βTG*), synchronized by a relatively short H-bond of its neighbor O4−H (O4−H···O3 = 2.391 in α*TG* and 2.398 Å in *βTG*), seem to be responsible for the greater electronic stability of the *TG* conformer over the *GG* and *GT* conformers. On the other hand, the attraction of O6−H for O5 in *βGG* appears to be stronger than expected judging from its O6−H···O5 bond distance (2.356 Å) relative to those of the same H-bond in the α*GG*, α*GT*, and *βGT* conformers. This may in part explain the slightly greater stability of *βGG* over *βTG*, despite the shorter O6−H···O4 bond in *βTG*. To conclude, the relative H-bond distances in Table 2 seem to support the relative electronic stabilities of α*TG* > α*GG* > α*GT* and *βGG* > *βTG* > *βGT* shown by E_{elec} in Table 3B. The very small difference in E_{elec} between *βGG* and *βTG* (0.052 kcal/mol) should be taken as insignificant.

The HF/6-31G* thermodynamic contributions at room temperature (E_{therm} in Table 3B) clearly indicate that nuclear motion favors the *GG* and *GT* conformers over the *TG* conformer. The driving force for the relative nuclear stability of *GT* > *GG* > *TG* again seems to come from a weaker O6−H···O5 bond in *GT* and *GG* as compared with the O6−H···O4 bond in *TG*. The weaker O5 interaction with O6−H results in a smaller torsional frequency involving the hydroxymethyl chain and leads to lower E_{therm} as a result of smaller zero-point energy and higher entropy.

Adding E_{therm} to E_{elec} yields the relative free energies (G in Table 3B) among the three hydroxymethyl conformers of each anomer. The order of stability is *GT* > *GG* > *TG* for α and *GG* > *GT* > *TG* for *β*. Overall, the differences in free energy among the different conformers are very small (0.1−0.3 kcal/mol) because of nearly equal electronic energies but slightly decreasing thermodynamic contributions from *GT* to *GG* to *TG*. Here, the difference in G between *aGT* and α*GG* (0.002 kcal/mol) is negligible. Additional calculations of exp(−*G*/RT) provide relative populations of 36% *GT*, 36% *GG*, and 29% *TG* for α and 40% *GG*, 33% *GT*, and 26% *TG* for *β*.

The gas-phase basicities reported here were measured separately for each glucose anomer; however, mutarotation could occur to some extent when the pure anomer was dissolved in

(27) (a) Jeffrey, G. A.; Pople, J. A.; Radom, L. *Carbohydr. Res.* **1972**, *25*, 117; **1974**, *38*, 81. (b) Jeffrey, G. A.; Pople, J. A.; Binkley, J. S.; Vishveshwara *J. Am. Chem. Soc.* **1978**, *100*, 373. (c) Wiberg, K. B.; Murcko, M. A. *J. Am. Chem. Soc.* **1989**, *111*, 4821. (d) Tvaroska, I.; Carver, J. P. *J. Phys. Chem.* **1994**, *98*, 9477.

(28) (a) Woods, R. J.; Szarek, W. A.; Smith, V. H., Jr. *J. Chem. Soc., Chem. Commun.* **1991**, 334. (b) Cramer, C. J. *J. Org. Chem.* **1992**, *57*, 7034.

(29) (a) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, 1991. (b) Steiner, T.; Saenger, W. *Acta Crystallogr.* **1992**, *B48*, 819.

(30) Del Bene, J. E.; Shavitt, I. *J. Phys. Chem.* **1990**, *94*, 5514.

glycerol. In view of this experimental condition and the past interest on anomeric equilibrium[13−16] the following topics are discussed.

Consider first a gaseous mixture of α and β anomers at 298 K. To estimate the relative anomeric stability, energies of the three α and three β conformers need be expressed on the same scale. Using α*TG* as the common reference conformer a second set of β values are obtained and shown in the right half of Table 3C under "HF/6-31G*". Comparison of each β value to the α value of the same conformation indicates that β is 1.0−1.2 kcal/mol higher in E_{elec} and 0.5−0.6 kcal/mol lower in E_{therm} than α. The greater electronic stability of the α anomer relative to the β anomer has been mainly attributed to the anomeric effect.[16] Using the exp(−*G*/RT) values, the overall free-energy stability is deduced as α*GT* > α*GG* > α*TG* > β*GG* > β*GT* > β*TG* with respective populations of 26%, 26%, 20%, 12%, 9%, 7%. These populations lead to an anomeric distribution of 72% α and 28% β and a hydroxymethyl rotamer distribution of 38% *GG*, 35% *GT*, and 27% *TG*.

Next we consider an aqueous mixture of glucose anomers at ambient temperatures. Experiments showed that the equilibrium distributions are about 36% α and 64% β for the anomers[7] and 54% *GG*, 44% *GT*, and 2% *TG* for the hydroxymethyl conformers.[8] These experimental aqueous distributions are clearly very different from those calculated for the gas phase. While the aqueous distributions obtained from different theoretical models have yet to reach a quantitative agreement with experiments, explanations have been advanced that hydration stabilizes the β anomer more and that solvation disfavors the *TG* conformation.[13,14]

In estimating the change in free energy from the β anomer to the α anomer of glucose in aqueous solution, $\beta \rightarrow \alpha$, Cramer and Truhlar (CT)[13] used an expression for the aqueous free energy (G°_{aq}) as a sum of gas-phase free energy (G°_g) and the free energy of solvation (G°_S). They obtained $\Delta G^\circ_g(\beta \rightarrow \alpha) = -0.5$ kcal/mol based on HF/6-31G*//HF/4-31G electronic energies and HF/4-31G//HF/4-31G thermodynamic properties,[15] and $\Delta G^\circ_{aq}(\beta \rightarrow \alpha) = -0.5$ kcal/mol after adding the AM1-SM2 G°_S value to each conformer. The calculated ΔG°_{aq}, though close in magnitude to the experimental value of 0.3 kcal/mol,[7] is wrong in sign.

We now upgrade the gaseous part of CT's calculation to the HF/6-31G* level for both electronic energies and thermodynamic properties. Using the exp(−*G*/RT) values of Table 3C and by averaging over the three conformations for each anomer,[13] we obtain $\Delta G^\circ_g(\beta \rightarrow \alpha) = -0.547$ kcal/mol. Employing CT's AM1-SM2 G°_S values,[13] we arrive at $\Delta G^\circ_{aq}(\beta \rightarrow \alpha) = -0.5$ kcal/mol, which is no different from the CT result shown above. To conclude, theory at this general level contradicts experiment in predicting a predominance of the α anomer in water.

Finally, we discuss topics of importance to ab initio calculations on carbohydrates. Specifically, recent results obtained at different theoretical levels for molecular systems related to D-glucopyranose are reviewed. In so doing we hope to gain better perspectives for future applications.

As ab initio computations on structures larger than glucose will most likely involve basis sets smaller than 6-31G*, it is of value to examine the differences between Polavarapu and Ewig's[15] HF/4-31G and our HF/6-31G* results for the neutral glucose in the gas phase. The relevant HF/4-31G data for direct comparisons are presented in the left half of Table 3C. A casual inspection between the corresponding columns in the left and right halves of Table 3C reveals that the orders of stability for the six species resulting from the two different basis sets are essentially the same in both E_{elec} and E_{therm}, except for the relative E_{elec} between β*TG* and β*GG*. But differences in the E_{elec} or E_{therm} values among conformers of the same anomer are notably larger with 4-31G than 6-31G*. The more striking discrepancy is the anomeric electronic energy difference shown by HF/4-31G: the E_{elec} of β are 2.2−2.5 kcal/mol higher than those of α, doubling the differences given by HF/6-31G*. (Our HF/6-31G* results in Table 3C are in agreement with the E_{elec} values obtained by Salzner and Schleyer[16] for α*TG*, α*GT*, β*TG*, and β*GT*.) As a result, HF/4-31G yields an anomeric distribution of 95% α and 5% β.

Given sufficient computation resources our HF/6-31G* results may be improved in principle by enlarging the basis set and inclusion of electron correlation. But in practice, there is the difficulty of not knowing to what extent the level of theory needs to be upgraded before the calculated energies show convergence. A good example is provided by Barrows *et al.*,[17] who calculated the electronic energies for the β*TG* and β*GG* conformers at 13 different ab initio levels, including geometry optimizations from HF/STO-3G to MP2/cc-pVDZ plus higher-level single-point approximations. They finally adopted a composite E_{elec} value of 0.30 kcal/mol for β*TG* relative to β*GG* (as compared with the HF/6-31G* value of 0.052 kcal/mol in Table 3B). This example demonstrates the expense involved when a thorough search is conducted to find a reliable estimate for the small energy difference between two hydroxymethyl conformers.

Fortunately it is more straightforward to determine the effect of the basis set on the anomeric electronic energy difference at the Hartree−Fock level. From our analysis on neutral glucose we find an upgrade from the HF/4-31G to HF/6-31G* narrows this anomeric gap by 1.2−1.4 kcal/mol (cf. Table 3C). Recent calculations on 2-hydroxytetrahydropyran and 2-methoxytetrahydropyran also show a reduction in the energy difference between the equatorial and axial conformers upon expanding the basis set.[16,27d,10b] The reduction is *ca.* 0.5−0.6 kcal/mol from HF/6-31G* to HF/O; Huz, C,H:6-31G//HF/6-31G* or HF/6-311++G**//HF/6-31G*, and from HF/6-31G** to HF/6-311++G**//HF/6-31G**. Even greater reduction is achieved for 2-methoxytetrahydropyran when electron correlation is introduced at the MP2 level for geometry optimizations (MP2/6-31G*) followed by a basis set expansion (HF/6-311++G**).[27d] Based on these findings we anticipate a smaller anomeric energy difference and a larger β population from those shown in Table 3C when a theoretical level higher than HF/6-31G* is applied to D-glucopyranose in the gas phase.

Gas-Phase Protonation. Consider transferring a proton from a proton-donor molecule to a neutral α- or β-D-glucopyranose molecule M. There are six oxygen sites for protonation: one primary (*n*) hydroxyl oxygens (*O6*), four secondary (*sec*) hydroxyl oxygen (*O1*, *O2*, *O3*, and *O4*), and one ether oxygen (*O5*).

In a gas-phase protonation reaction, certain oxygen sites are more relevant than others. First, consider the energetic factor. The H-bond network as depicted in Figure 2 for the lowest-electronic-energy state of each conformation of the neutral species is best left unchanged in protonation. Otherwise, energy must be spent to break an intramolecular H-bond in order to rotate an O−H bond away from the existing equilibrium position. The energy expenditure for internal rotation of any O−H bond from a conformation in Figure 2 would require 3 kcal/mol or more.[19] Next, consider the spatial factor. A casual inspection at the structures in Figure 2 reveals that *O1*, *O2*, and *O3* is each involved in one C−O bond, one O−H bond, and one intramolecular H-bond; consequently there is only one "slot" open on the oxygen to receive the proton originating from

the proton-donor molecule. On the other hand, the *O6* in the *TG*, *GG*, and *GT* conformers, *O4* in the *GG* and *GT* conformers, and *O5* in the *TG* conformers each has two "slots" available. These three oxygen sites (*O6*, *O4*, and *O5*) in their appropriate conformations are therefore sterically favored for protonation: They provide ample space for the approach of a proton-donor. Besides, protonation at each of these sites has the best chance of leaving the original H-bond network least disturbed and thereby achieving the maximum energy benefit. Under these considerations 12 stable protonated species are produced (Figures 3 and 4).

The kinetic reasons given above apply equally well to the mass spectral measurement of the GB of M from a deprotonation reaction by gas-phase collisions, reaction 1, where B is a reference compound. In this situation a proton in MH^+ is to be transferred via the intermediate $MH^+\cdots B$ to B; this proton is more likely to be the "added" proton H13 on the *O6*, *O5*, or *O4* site of a species in Figure 3 or 4 with reference to its parent neutral species in Figure 2 (see definition for H13 in Figure 1 caption). Note that H13 in MH^+ is not engaged in intramolecular hydrogen bonding and therefore (a) it is physically more accessible for complexing with B (especially for the *O6*-protonated variety) and (b) it is energetically less demanding to be removed from MH^+. Based on this reasoning we proceed to study this set of 12 protonated species for comparisons with the measured GBs.

Protonated Species. In regard to geometry the most pronounced "local" change accompanying protonation is the lengthening of bonds attached to the protonated oxygen O*i**. (The asterisk designates O*i* as the protonated oxygen.) The C−O* and O*−H bond lengths in the protonated species listed in Table 1 indicate increases of 0.1 Å in C−O and 0.02 Å in O−H over the corresponding values in the neutral species. When O*i** is the proton donor in a H-bond, the increases in the bond length of C−O*i** and O*i**−H, coupled with an increase of positive charge on H due to protonation, brings a significant reduction to the H···O*j* distance from the neutral H-bond O*i*−H···O*j* to the ionic H-bond O*i**−H···O*j*.[29b] (For example, compare the H···O4 distance in the neutral α*TG* of Figure 2, 2.117 Å, to the corresponding distance in the protonated α*TG6* of Figure 3, 1.599 Å.) This reduction is *ca.* 0.5 Å in O6*−H···O4 and O6*−H···O5 due to *O6*-protonation and *ca.* 0.4 Å in O4*−H···O3 due to *O4*-protonation (see Table 2). The decrease in the particular H-bond distance enhances significantly the strength of the particular H-bond. Thus *O6*-protonation is expected to give rise to a stronger ionic H-bond than *O4*-protonation mainly because the CH_2OH chain is more flexible and has a longer reach than the exocylic OH group. In the case of *O5*-protonation, no such enhancement is observed because *O5* is not the proton donor of any H-bond in the neutral species. Nonetheless, there are inconsequential effects coming from *O5*-protonation: the original neutral O1−H···O5 interaction in α*TG5e*, *βTG5a*, or *βTG5e* is now replaced by a comparable ionic O5*−H···O1 interaction.

Protonation induces a small overall increase in the hydrogen bonding distances for all except for the ionic H-bond that underwent the more pronounced "local" change already described (Table 2). This may be called the "distant" change: Protonation destabilizes the H-bonds away from the protonation site. (For example, compare the H···O5 distance in the neutral α*GT* of Figure 2, 2.370 Å, to the corresponding distance in the protonated α*GT4* of Figure 4, 2.614 Å.)

The analysis of hydrogen bonding aids in the understanding of the relative stability of the six selected protonated species in each anomer. Based on the electronic energies of Table 4, protonation at *O6* is clearly preferred. The strong attraction arising from the remarkably short H-bond distances between the O6*−H donor and the O4 or O5 acceptor places the *O6*-protonated species 5−15 kcal/mol lower in energy than the *O4*- or *O5*-protonated species. Among the *O6*-protonated species, the *TG* conformer is some 5−6 kcal/mol more stable than the *GG* and *GT* conformers, again, because of a shorter H-bond distance for O6*−H···O4 (1.6 Å) than O6*−H···O5 (1.9 Å). The α*TG6* and *βTG6* species therefore become the most stable protonated species for the respective α and *β* anomers in the elementary step of protonation as defined earlier.

At this point it seems meaningful to compare the electronic protonation energy (PE = $-\Delta E_{elec}$) calculated for the model compounds (Table 5) where H-bonds are absent with those of D-glucopyranose where extensive H-bond networks exist. At the HF/6-31G* level the PE of ethanol (*n*-alcohol) is 195 kcal/mol, which is 7−13 kcal/mol smaller than the PEs resulting from *O6*-protonations. This large energy difference may be attributed to the significant enhancement of the particular H-bond due to protonation (O6*−H···O4 or O6*−H···O5). Surprisingly the PEs of 2-propanol (*sec*-alcohol) and ether, 200 and 199 kcal/mol, are 2−6 kcal/mol greater than those of the respective *O4*- and *O5*-protonations. One explanation could be that protonation destabilizes the original H-bond network and this adverse impact on stability is greater than the gain from the enhancement of a particular H-bond (O4*−H···O3) or the emergence of a new hydrogen interaction (O5*−H···O1). There is also the anomeric effect which operates in glucose but not in the model compounds: *O5*-protonation on an α anomer destroys in part the hyperconjugative stabilization present in the parent neutral species. The notably lower PEs of the *O5*-protonated α anomers vs those of the *β* anomers are a manifestion of the anomeric effect (e.g., 4.5 kcal/mol difference between α*TG5a* vs *βTG5a*). (The anomeric effect as it influences basicity has been extensively studied for dihydroxymethane.[28])

Gas-Phase Basicity. The theoretical GBs for D-glucopyranose, listed under "HF/6-31G*" in Table 5, are derived from the HF/6-31G* electronic energies and HF/3-21G zero-point and thermodynamic energies in Table 4. To obtain a more accurate theoretical GB, the HF/6-31G* electronic protonation energies need be corrected for deficiencies in the size and composition of the 6-31G* basis set and the lack of electron correlation in the HF procedure. Our previous studies on amino acids and peptides[4,5] indicate that the MP4/6-31+G(2d,2p)//MP2/6-31+G** level suggested by Del Bene and Shavitt[30] for protonation calculations yields reasonably accurate GBs. We therefore apply this higher level to study the *O*-protonations of model compounds ethanol, 2-propanol, and dimethyl ether, which emulate the *O6*-, *O4*-, and *O5*-protonations at the respective hydroxymethyl, hydroxyl, and ring ether oxygen atoms in glucose. The differences in the GBs calculated from the MP4/6-31+G(2d,2p) and HF/6-31G* levels (or Level H and Level M) for the model compounds are used as "correction constants" on the GBs of glucose calculated from the HF/6-31G* level. We adopt this approach based on our experience that the "correction constant" derived from a model compound for a particular type of basic atom (e.g., amino N, hydroxyl O, and amide carbonyl O in peptides) is nearly transferrable from molecule to molecule.[5] (Note that the protonation of dihydroxymethane, a model compound for *O1*-protonation associated with the anomeric effect, has been examined at the MP2/6-311++G**//HF/6-31G* level.[28a] *O1*-protonation is not explicitly considered here.)

After making the corrections, the theoretical GBs for D-glucopyranose are listed under "adjusted" in Table 5, which

show estimates of 178−190 and 177−189 kcal/mol for the respective α and *β* anomers. The upper values 190 and 189 kcal/mol corresponding to *O6*−protonations in the *TG* conformation, α*TG* → α*TG6* and *βTG* → *βTG6*, compare especially well with the experimental 188 ± 3 kcal/mol. The theoretical proton affinities are also listed in Table 5. Using the theoretical $-T\Delta S$ value of 7.7 kcal/mol (average of the α*GG6* and *βGG6* values), the "experimental" PA is estimated to be 196 kcal/mol.

The kinetic pathway for the protonation reaction as postulated above represents the simplest elementary step in a gas-phase mechanism. In this step the conformation of the neutral reactant is essentially conserved during protonation or that the protonated product retains the reactant conformation. This elementary step may be expressed, e.g., by the reaction α(*TG*) + H^+ → αH^+-(*TG*), where each species is in its lowest-energy state. The particular example is depicted explicitly by the α*TG* and α*TG6* structures in Figures 2 and 3 and by the change specified as α*TG* → α*TG6* in Table 5.

But is this simple kinetic pathway a realistic description of the protonation reaction involved in the nonequilibrium GB measurement conducted in real life? The answer is probably negative. What we have provided here are sample calculations of typical low-to-high GB values corresponding to protonations that range from destabilizing to enhancing the intramolecular hydrogen bonding network in the lowest-energy states of isolated neutral species.

At any rate it is of great interest to know what the most stable protonated species of α- and *β*-D-glucopyranose are in the gas phase, regardless of the circumstances under which they are formed. After extensive searches, we identified the six structures in Figure 5.[19] Relative to the neutral *TG*, *GG*, and *GT* conformers of α and *β* anomers in Table 3, the respective relative electronic energies (ΔE_{elec}) for the new protonated structures in kcal/mol are the following: −210.95 (α*TG4*), −206.30 (α*GG1*), and −205.14 (α*GT5e*); −209.96 (*βTG4*), −208.40 (*βGG5a*), and −207.11 (*βGT5e*).

The very short H-bond distance involving the protonated O4 or O5 (O4* or O5*) is given in Figure 5 to show one reason for their exceptional stability. These five new *O4*- and *O5*−protonated species seem to be related to the *O6*-protonated species in Figure 3. For example, α*TG6* may be changed to α*TG4* (and vice versa) via intramolecular proton transfer from O6* to O4 (and from O4* to O6). Gas-phase collisions may equilibrate these pairs of related structures. The surprise is the *O1*-protonated species, α*GG1*, which has a geometry drastically different from any of those encountered in this study. The protonated *O1* in α*GG1* is nearly dissociated from the ring as a H_2O molecule. (The "C1−O1" bond in α*GG1* is 2.357 Å long.) The high stability of α*GG1* is an anomeric phenomenon[28b] as the counter species *βGG1* (not shown here) is 11 kcal/mol higher in energy. The clockwise orientations of exocylic hydroxyl H-bonds in α*GG1* are common to species protonated at the *O1*, *O2*, or *O3* site.

Among the new protonated species one α (α*TG4*) and all three *β* (*βTG4*, *βGG5a*, and *βGT5e*) species are more stable than the respective most stable α (α*TG6*) and *β* (*BTG6*) species of the original set (Table 4A). One reason could be that *sec*-alcohol and ether type oxygens invoke greater protonation energies ($-\Delta E_{elec}$) than *n*-alcohol type oxygen. The adjusted GB values deduced from the most stable species, α*TG4* and *βTG4*, are 193 and 192 kcal/mol which represent the largest theoretical GB values obtainable for the respective α and *β* anomers. On the other hand, the most probable protonated species may be α*GT5e* in view of the fact that α*GT* is calculated to be the most populated neutral species at room temperature (Table 3C and ref 13). The adjusted GB deduced for the change, α*GT* → α*GT5e*, is 189 kcal/mol which is in excellent agreement with the experimental GB of 188 ± 3 kcal/mol.

Concluding Remarks

The gas-phase basicities of D-glucopyranose from our mass spectral measurements supply the first experimental data without the interference of crystal packing or solvation effects. The measured GBs are shown to be in general agreement with the theoretical interpretation based on our ab initio calculations. The minimum-energy structures calculated at the HF/6-31G* level for the numerous D-glucopyranose species provide valuable structural data for postulating kinetic pathways in a protonation reaction and for the development of force-field parameters in computer modeling of polysaccharides.

Acknowledgment. The computational work was supported by Miami University through the Miami Computing and Information Services (MCIS) and by a grant from the Ohio Supercomputer Center. Grants from the National Institutes of Health to support both the theoretical (R15-GM52670−01) and experimental (R15-GM47657−01) research are gratefully acknowledged.

Supporting Information Available: The HF/6-31G* internal coordinates for the lowest-electronic-energy conformers of neutral α- and *β*-D-glucopyranose, α*TG* and *βGG* of Figure 2, are shown in Table S-1, the HF/6-31G* atomic Cartesian coordinates (in bohrs) are provided for all six neutral α- and *β*-D-glucopyranose species of Figure 2 in Table S-2, and the ab initio data for the ethanol, 2-propanol, and dimethyl ether used to deduce the protonation results of Table 6 are provided in Table S-3 (6 pages). See any current masthead page for ordering and Internet access instructions.

JA960427Z