# Hydrated Sugars in the Gas Phase: Spectroscopy and Conformation of Singly Hydrated Phenyl β-D-Glucopyranoside†

**Rebecca A. Jockusch, Romano T. Kroemer,‡ Francis O. Talbot, and John P. Simons***

*Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, U.K.*

*Received: April 30, 2003; In Final Form: September 22, 2003*

The structural investigation of a hydrated monosaccharide, phenyl β-D-glucopyranoside (Pheβ Glc), in the gas phase is presented. It is based upon ab initio computation coupled with the analysis of the resonant 2-photon ionization and IR ion-dip spectra of the singly hydrated clusters stabilized in a free jet expansion and follows an earlier investigation of the structure and conformations of the unhydrated sugar.[1] Despite the potentially large set of possible binding sites and conformations, only two singly hydrated complexes are formed in the free jet expansion. Tentative structural assignments are made on the basis of comparisons with those already established for related systems; comparisons between the observed O−H vibrational frequencies and those computed for structures optimized at the B3LYP/6-31+G(d) level of theory, and in light of the relative energies of these structures, calculated at the single-point MP2/6-311+G(d,p) level are made. A discrepancy between the latest revision of the Gaussian 98 package (revision A.11) and earlier revisions, which can lead to different computed structures when optimizing noncovalently bound molecular complexes, is discussed in an appendix.

## 1. Introduction

Since the first pioneering spectroscopic studies of amino acids and small peptides in the gas phase by Levy and coworkers,[2−4] a rapidly increasing number of molecular spectroscopists have extended their horizons to include structural (and a few dynamical[5,6]) studies of isolated and clustered bioactive molecules, biomolecular building blocks, and larger molecular assemblies involved in the biophysics and biochemistry of living bodies, exploiting methods previously employed only for "simple" molecules.[7−9] A suite of powerful strategies has now evolved that employs a combination of experimental techniques such as laser ablation for transferring molecules into the gas phase, rapid cooling in a free jet expansion or in helium nanodroplets[10] to stabilize their conformers or clusters, a highly selective and sensitive armory of IR and UV laser-based optical spectroscopies, coupled with mass spectrometry, to probe their structural landscapes, and the ready accessibility of powerful ab initio quantum chemical computational codes for interpretation. Fast quantum calculation and optical spectroscopic techniques are now providing access to neurotransmitters and enzyme blockers, amino acids and peptides, nucleic acid bases and nucleosides, and most recently to sugars and glycopeptides, all studied under experimental conditions previously used only for simpler molecules. Considering the ubiquity of water and of hydrogen bonding in biological systems, an essential second step toward an understanding of their structure/function relationships is the determination of hydration effects. Sugars present an obvious and important target.

A major difficulty in the study of hydrated sugars, however, arises from both their conformational flexibility and the high number of very similar hydroxyl groups they accommodate, even in a single monosaccharide; each hydroxyl group is capable of being a donor and an acceptor of hydrogen bonds. Despite the strong interest already shown in bare glucose,[11−13] there have been, to our knowledge, no published ab initio investigations of its specifically hydrated clusters (in the gas phase). Hydration effects in a condensed medium have been modeled by Molteni and Parrinello, who ran a 6-ps quantum dynamic simulation on three conformers of β-D-glucopyranose (βGlc) in water[14] to provide, inter alia, some insight into its anomeric equilibrium. The average number of H bonds formed during the QM simulation of the βGlc suggests that its five hydroxyl groups display similar H-bonding behavior. A gas-phase investigation of selectively hydrated clusters of glucose offers an opportunity to determine which, if any, of the hydroxyl groups is actually the preferred site(s) for hydration. The earlier examination[1] of the unhydrated glycoside, phenyl β-D-glucopyranoside (Pheβ Glc) in the gas phase, provides an excellent platform for such an investigation. As well as identification of the preferred water binding site(s), a structural investigation of the hydrated clusters of Pheβ Glc should also provide insight into the influence of hydration on its conformational preferences.[7−9] These hydration studies benefit greatly from the extraordinary sensitivity of the O−H stretching modes to subtle changes in the H-bonded conformational structures, which has been revealed by the subsequent investigation of phenyl β-D-galactoside (Pheβ Gal).[15] The earlier study of the glycoside Pheβ Glc revealed the presence of three distinct conformers, two with a gauche (G+g− and G−g+) and one with a trans (Tg+) orientation of the hydroxy methyl group, a result in contrast with the NMR solution-phase data for methyl β-D-glucopyranoside that reported the presence of only the two gauche structures.[16−18]

## 2. Experimental and Computational Procedures

The systems used for mass-selected resonant 2-photon ionization (R2PI), UV hole burning, and resonant IR ion-dip

† Part of the special issue "Charles S. Parmenter Festschrift".

* Corresponding author. E-mail: john.simons@chem.ox.ac.uk. Fax: 44 (0) 1865 275410.

‡ Present address: Molecular Modeling & Design, Department of Chemistry, Pharmacia, Viale Pasteur 10, 20014 Nerviano (MI), Italy.

10.1021/jp0351730 CCC: $25.00 
Published on Web 10/29/2003

spectroscopy were identical to those employed in earlier studies of jet-cooled amino acids,[19,20] the glucoside Phe$\beta$Glc,[1] and the galactoside Phe$\beta$Gal.[15] The sugar derivative was heated constantly in a tubular oven held at a temperature of 433 K, which was attached to the outlet of a pulsed nozzle; the plume of vaporized molecules was "picked up" and subsequently cooled by the argon jet (stagnation pressure, 4 bars) expanding into the high-vacuum chamber. UV light was generated by frequency doubling the output of a Nd:YAG-pumped pulsed dye laser (LAS). UV−UV double resonance "hole burn" experiments employed a second (probe) excimer-pumped pulsed dye laser (LambdaPhysik FL3002), fired ∼200 ns after the first (burn) pulse. Tuneable IR laser radiation (ca. 3000−4000 $cm^{-1}$, 0.4-$cm^{-1}$ bandwidth, 4 mJ/pulse) was generated by difference frequency-mixing of a Nd:YAG output (Continuum Powerlite Precision 8000) and a Nd:YAG-pumped (Continuum ND6000) dye laser operating with the LD765 dye in a $LiNbO_3$ crystal. The IR beam was focused onto the jet antiparallel to the UV beam, and the delay between the IR and UV lasers was set to 150 ns. Starting structures for ab initio calculations of Phe$\beta$Glc−($H_2O$) were constructed from the optimized structures of the three low-energy conformers of Phe$\beta$Glc.[1] Singly hydrated complexes were generated by systematically inserting water molecules between an OH group (acting as a donor to the water molecule) and a nearby oxygen atom of the sugar ring (acting as an acceptor) at all possible positions. For each of these structures, the phenyl ring orientation was set to the two different favorable positions (see section 3.2). This procedure led to a total of 24 insertion structures. A number of complexes in which the water molecule forms only one hydrogen bond with the sugar were also generated for comparison.

Ab initio calculations were performed using the Gaussian 98 suite of programs. The conformers were initially optimized using revision A.7[21] of G98 (R7) and subsequently optimized using revision A.11[22] (R11): the reason for, and consequences of, the change to R11 are detailed and discussed in the Appendix. All conformers were fully optimized using the B3LYP hybrid density functional theory method[23,24] in conjunction with the 6-31+G(d) basis set.[25,26] The analytical second derivatives of these stationary points were then computed to ensure that all structures were minima and to obtain the corresponding harmonic frequencies. All frequencies and zero-point energies were uniformly scaled by 0.9734, a scaling factor that provided an excellent match with the experimentally observed, weakly perturbed OH frequencies of the unhydrated Phe$\beta$Glc and Phe$\beta$Gal.[1,15] Single-point MP2[27] energies were then evaluated for all minima, using the 6-311+G(d,p) basis set, to obtain better estimates of the relative energies of the lowest-lying structures. The tabulated values all include scaled zero-point energy corrections from the B3LYP/6-31+G(d)-level calculations.

The notation employed to identify potential binding sites (Figure 1) and molecular conformations follows the scheme used previously.[1,15] Briefly, the first three letters indicate gauche, "g" or trans, "t" orientations of the $H_n-O_n-C_n-C_{n+1}$, $n = 2-4$, dihedral angles; the final letters indicate the orientation of the hydroxy methyl group with respect to the ring oxygen, G+, G−, or T, and its hydroxyl group, g+, g−, or t.

## 3. Results and Discussion

**3.1. Spectroscopy.** *R2PI and UV Hole-Burn Spectra.* The UV R2PI spectrum of the Phe$\beta$Glc−($H_2O$) complex recorded in the $[Phe\beta Glc-(H_2O)]^+$ mass channel in the spectral range of 36 500−37 000 $cm^{-1}$ is shown in Figure 2. It displays a strong band, labeled P, centered at 36 767 $cm^{-1}$ (101 $cm^{-1}$ below the origin of the most abundant unhydrated conformer) followed by a series of weaker transitions at higher wavenumbers. The pair of UV hole-burn spectra associated with the two features labeled P and U and shown in the same Figure include virtually all of the R2PI transitions and indicate the presence of just two different conformers, with band origins at 36 767 (P) and 36 870 (U) $cm^{-1}$. (Two other features, which also appear in the $[Phe\beta Glc-(H_2O)_2]^+$ mass channel, are associated with a doubly hydrated cluster).

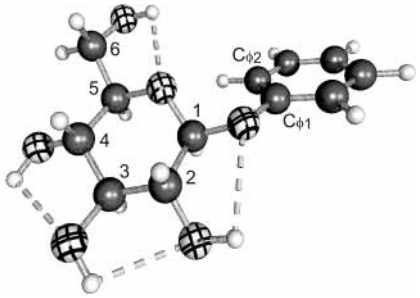


**Figure 1.** Most stable conformer, tttG+g−, of phenyl-$\beta$-D-glucopyranose (Phe$\beta$Glc). The patterned spheres represent the oxygen atoms.

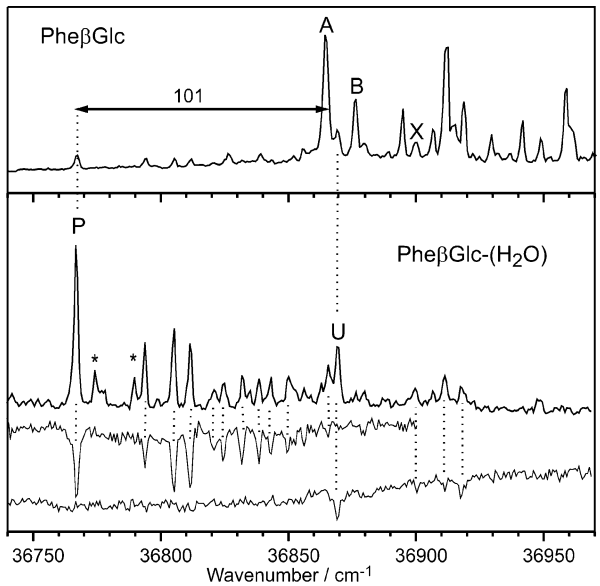


**Figure 2.** Resonant 2-photon ionization (R2PI) and UV/UV ion-dip spectra of the Phe$\beta$Glc−($H_2O$) complex. The two bands marked with asterisks belong to the Phe$\beta$Glc−$(H_2O)_2$ cluster. The spectrum of the bare monomer, Phe$\beta$Glc, is also displayed for reference (top frame).

Several inferences can already be drawn from these data. At least one of the three conformers observed for the unhydrated monomer is no longer populated. The appearance of only two distinct monohydrates indicates a high level of selectivity, and the predominance of one of the two conformers indicates a clear preference between the alternatives. The numerous H-bonding sites on the pyranose ring do not have identical H-bonding properties.

The vibronic progressions associated with the red-shifted band system, based on the origin band P, and the weaker progressions built upon the origin band U are listed in Table 1. Four low-frequency vibronic progressions can be identified in each band system, associated with low-frequency modes at 27, 38, 45, and 58 $cm^{-1}$ for P and at 30, 37, 42, and 48 $cm^{-1}$ for U. The two gauche conformers of unhydrated Phe$\beta$Glc display low-frequency vibrations at 30 ($\nu_1$) and 47 ($\nu_2$) $cm^{-1}$ (G+g− conformer), and 30 ($\nu_1$), 42 ($\nu_2$) and 53 ($\nu_3$) $cm^{-1}$ (G−g+ conformer) associated with (two) inter-ring wagging modes and

**TABLE 1: Observed Vibronic Bands[a] and Tentative Assignments of the Two Singly Hydrated Phe$\beta$Glc Complexes[b] and the Three Unhydrated Conformers of Phe$\beta$Glc**

Singly Hydrated Complexes

| | experimental $\tilde{\nu}$/cm$^{-1}$ | |
|---|---|---|
| assignment | complex P | complex U |
| origin | 36 767 [100] | 36 870 [41] |
| $\tilde{\nu}_1$ | 27 [37] | 30 [16] |
| $\tilde{\nu}_{inter}$ | 38 [48] | 37 [13] |
| $\tilde{\nu}_2$ | 45 [39] | 42 [27] |
| $2\tilde{\nu}_1$ | 54 [15] | |
| $\tilde{\nu}_3$ | 58 [18] | 48 [20] |
| $\tilde{\nu}_1 + \tilde{\nu}_{inter}$ | 65 [20] | |
| $\tilde{\nu}_1 + \tilde{\nu}_2$ | 72 [18] | |
| $2\tilde{\nu}_{inter}$ | 76 [19] | 79 [08] |
| $\tilde{\nu}_{inter} + \tilde{\nu}_2$ | 83 [22] | |
| $\tilde{\nu}_1 + \tilde{\nu}_3$ | 86 [15] | |
| $2\tilde{\nu}_2$ | 89 [13] | |
| $2\tilde{\nu}_1 + \tilde{\nu}_{inter}$ | 92 [09] | |
| $\tilde{\nu}_{inter} + \tilde{\nu}_3$ | 96 [17] | |
| $2\tilde{\nu}_1 + \tilde{\nu}_2$ | 99 [29] | |

Unhydrated Conformers $\tilde{\nu}$/cm$^{-1}$

| assignment | A | B | X |
|---|---|---|---|
| origin | 36 868 | 36 880 | 36 903 |
| $\tilde{\nu}_1$ | 29.9 | 29.9 | |
| $\tilde{\nu}_2$ | 47.0 | 42.2 | 42.7 |
| $\tilde{\nu}_3$ | | 53.1 | |
| $\tilde{\nu}_1 + \tilde{\nu}_2$ | 77.0 | 72.2 | |
| $2\tilde{\nu}_2$ | 94.0 | 83.8 | 84.2 |
| $\tilde{\nu}_2 + \tilde{\nu}_3$ | | 94.7 | |

[a] Experimental shifts are given as $\Delta\tilde{\nu} = \tilde{\nu} - \tilde{\nu}_{origin}$. [b] Values in square brackets are the relative intensities.

(one) torsional mode[1]. The extra low-frequency mode in the two hydrated complexes, most probably the mode at 37−38 cm$^{-1}$ and labeled $\nu_{inter}$ in Table 1, is tentatively assigned to one of the six intermolecular vibrations associated with the bound water molecule.

The two band origins (P and U) associated with the hydrated complexes are separated by 103 cm$^{-1}$, a gap that contrasts with the small separation, 12 cm$^{-1}$, between the two main conformers (G+g− and G−g+) of unhydrated Phe$\beta$Glc. Whereas band P is red-shifted by 101 cm$^{-1}$ with respect to the band origin of conformer Phe$\beta$Glc(G+g−), band U is shifted slightly to higher energy. Whatever the conformation adopted by the sugar ring in conformer U might be, the addition of the water molecule does not seem to affect the electronic transition very much because its origin lies within 32 cm$^{-1}$ of that of any of the three unhydrated conformers. The considerably larger shift associated with the UV spectrum of complex P suggests a more "direct" contact of the bound water molecule and the phenoxy ring in that conformer than in complex U. In the monohydrate of 2-phenoxyethanol (POX), in which the water molecule is inserted between the hydroxyl group and the phenoxy oxygen,[28] hydration shifts the UV band origin 62 cm$^{-1}$ to the red.

*IR Ion-Dip Spectroscopy.* Figure 3 shows the mid-IR IR/UV ion-dip spectra recorded when monitoring origin bands P and U, together with the corresponding spectrum associated with the most strongly populated conformer, G+g−, in unhydrated Phe$\beta$Glc. The spectra associated with complexes P and U are quite similar, with only subtle differences between them. As expected, they include one "free" OH stretching vibration associated with a single bound water molecule, three very weakly perturbed O−H stretching vibrations associated with the sugar (cf. the quartet of bands displayed by the unhydrated

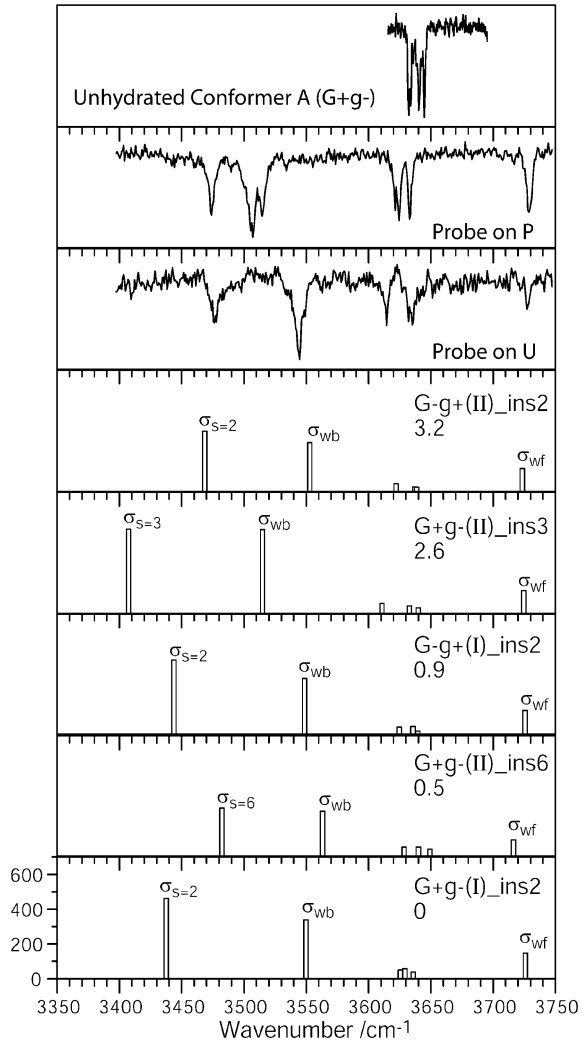


**Figure 3.** IR/UV ion-dip spectra in the O−H stretching region recorded by monitoring bands P and U, alongside the calculated IR spectra of the five lowest-energy conformers of Phe$\beta$Glc−($H_2O$). Calculated wavenumbers are at the B3LYP/6-31+G(d) level and are scaled by 0.9734. MP2/6-311+G(d,p)//B3LYP/6-31+G(d) relative energies are listed in kJ mol$^{-1}$. The experimental IR spectrum of the most abundant unhydrated conformer (G+g−) is also shown in the top frame for comparison. See Figure 6 for an illustration of the conformers.

sugar), and a pair of bands shifted more strongly to lower wavenumber, one associated with the sugar hydroxyl group bonding to the water corresponding to that labeled as $\sigma_{s=2,3,4 \text{ or } 6}$ in the calculated spectra below (as in ref 15) and the other associated with the H-bonded OH group on the water molecule itself; all of the band positions are listed in Table 3. The additional feature, appearing in the wings of the strong band at 3507 cm$^{-1}$ in complex P, is due most likely to a Fermi resonance between the strong O−H stretch and an overtone or combination band of lower frequencies; there is no reasonable singly hydrated structure that could create three strongly H-bonded OH groups.

The most strongly shifted O−H stretching band, which appears at approximately the same position in both complexes P (at 3474 cm$^{-1}$) and U (at 3477 cm$^{-1}$) and is labeled $\sigma_s$ in the computed spectra shown in Figure 3, is associated with the stretching of the sugar hydroxyl group bound (as a proton donor) to the water molecule; it is the displaced band associated with the "missing" member, $\sigma_{2,3,4 \text{ or } 6}$, of the quartet in unhydrated Phe$\beta$Glc. The less strongly shifted O−H stretching vibration, corresponding to the band labeled $\sigma_{wb}$ in the calculated spectra and associated with the bound water molecule, appears 67 cm$^{-1}$

**TABLE 2: Experimentally Observed IR Wavenumbers ($cm^{-1}$) of the OH Stretching Modes in Singly Hydrated Phe$\beta$Glc Complexes[a]**

| singly hydrated Phe$\beta$Glc | | unhydrated Phe$\beta$Glc | | | unhydrated Phe$\beta$Gal | |
|---|---|---|---|---|---|---|
| complex P | complex U | G+g− (A) | G−g+ (B) | Tg+ (X) | G+g− | Tg+ or G−g− |
| 3729 | 3727 | | | | | |
| 3633 | 3635 | 3645 | 3645 | 3642 | 3645 | 3644 |
| 3625 | 3632 | 3641 | 3640 | 3627 | 3638 | 3632 |
| 3622 | 3614 | 3634 | 3636 | 3604 | 3633 | |
| | | 3632 | 3634 | | 3611 | 3606 |
| 3507 | 3544 | | | | | |
| 3474 | 3477 | | | | | |

[a] The wavenumbers of the corresponding modes in the unhydrated sugar, Phe$\beta$Glc, and in its epimer, Phe$\beta$Gal, are also listed for comparison.

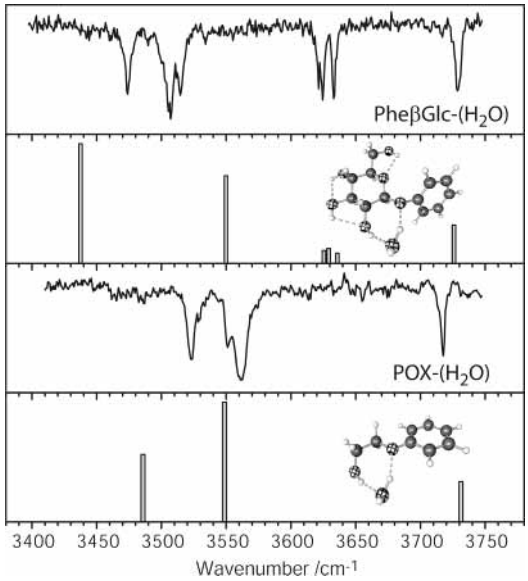


**Figure 4.** Comparison of the IR spectra of the Phe$\beta$Glc and 2-phenoxy ethanol (POX) hydrated complexes.

**TABLE 3: Relative Energies (kJ $mol^{-1}$) of the Lowest-Energy Conformers of the Phe$\beta$Glc−($H_2O$) Complex, Calculated at the Single-Point MP2/6-311+G(d,p)//B3LYP/6-31+G(d) Level and Including Scaled Zero-Point Energy Corrections at the B3LYP/6-31+G(d) Level**

| conformer | ΔE |
|---|---|
| G+g−(I)ins_2 | 0.0 |
| G+g−(II)ins_6 | 0.5 |
| G−g+(I)ins_2 | 0.9 |
| G+g−(II)_ins_3 | 2.6 |
| G−g+(II)ins_2 | 3.2 |
| G−g+(II)ins_6 | 3.3 |
| G+g−(II)ins_2 | 3.4 |
| G−g+(II)ins_6 | 3.5 |
| Tg+(I)ins_2 | 3.6 |
| G−g+(II)ins_3 | 4.2 |
| G−g+(II)ins_4 | 4.3 |
| G+g−(II)ins_4 | 4.6 |
| G+g−(I)ins_3 | 4.6 |
| gggG−g+(I)ins_6 | 5.2 |
| Tg+(II)ins_4 | 5.7 |
| Tg+(II)ins_3 | 5.8 |
| Tg+(II)_ins2 | 6.0 |

above the $\sigma_s$ band in complex U, but in P the separation is only 33 $cm^{-1}$. This is quite close to the corresponding separation, 39 $cm^{-1}$, in the monohydrate of 2-phenoxy ethanol[28] (Figure 4), suggesting a similar water-binding site in the two molecules. An insertion of the water molecule between the (OH)2 hydroxyl group and the (anomeric) phenoxy atom, O1, would "fit the bill". Note that in both cases the scaled frequencies computed at the B3LYP/6-31+G* level fail to reproduce the small gap between the two most strongly shifted experimental bands, although similar calculations for the weakly perturbed bands in unhydrated Phe$\beta$Glc[1] (and Phe$\beta$Gal[15]) conducted at the same level of theory provided excellent agreement with experiment. Calculating the potential energy surfaces of hydrogen-bonded molecular clusters is significantly more challenging than for the bare molecular partner (see the next section).

The three weakly perturbed intramolecular stretching vibrations centered around 3630 $cm^{-1}$, although not directly involved in the water binding, are actually shifted to lower wavenumber by ∼10 $cm^{-1}$ compared to their averaged position in the populated conformers of unhydrated Phe$\beta$Glc[1] and its epimer, Phe$\beta$Gal.[15] These otherwise "uninvolved" vibrations seem therefore to be more sensitive to the addition of a single bound water molecule than to epimerization. The two clusters display slightly different vibrational patterns: complex P presents a doublet at a lower wavenumber followed by a single band to form a "2−1" pattern, but complex U shows the opposite behavior, presenting a "1−2" pattern. In each case, the shift of the lowest-frequency band, appearing at 3622 $cm^{-1}$ (P) and 3614 $cm^{-1}$ (U), is markedly less than that of the moderately H-bonded (OH)6 → (OH)4 hydroxy methyl band, $\sigma_6$, which appears at 3604 $cm^{-1}$ and is associated with the trans Tg+ configuration of the bare molecule. Both measured band patterns suggest instead binding to gauche G+g− and/or G−g+ conformer(s), stripped of one of the bands associated with the OH modes, $\sigma_{2,3,4 \text{ or } 6}$.

**3.2. Ab Initio Calculations.** *Phe$\beta$Glc Revisited*. In a recent investigation of the bare epimer, phenyl-$\beta$-D-galactopyranoside,[15] the calculated OH stretching frequencies were found to vary slightly with a change in the orientation of the phenyl ring (promoting shifts of ca. ±5 $cm^{-1}$), though its orientation had a negligible effect on the conformation of the sugar ring. However, the orientation of the phenyl ring could have a more significant effect on the position of the noncovalently bound water molecule in the hydrated complex Phe$\beta$Glc−($H_2O$), especially if it interacts with ring oxygen O5. Furthermore, because no attempt had been made earlier to investigate the effect of its torsional orientation in the bare molecule[1] (due to its apparently minute influence on the molecular conformation), a relaxed potential energy scan was conducted (at the B3LYP/6-31+G* level) along the torsion angle of the phenyl ring, $\omega$ = C1−O1−C$\phi$1−C$\phi$2 in steps of 30° (Figure 5). The data points indicate the existence of two distinct minima, one associated with a "twisted" orientation located at $\omega \approx -90°$ (position I) and the other with a "flat" orientation located at $\omega \approx 20°$ (position II). For all three rotamers, the B3LYP calculations consistently favor the minimum located around 20° (II), and the barrier for interconversion from I → II appears to be very low (a few kJ $mol^{-1}$). The structures calculated previously[1] all had the phenyl ring located in the less favored "twisted" orientation, I. When the three lowest-energy conformers, G+g−, G−g+, and Tg+, were reoptimized with the phenyl ring in the flat position, II, they

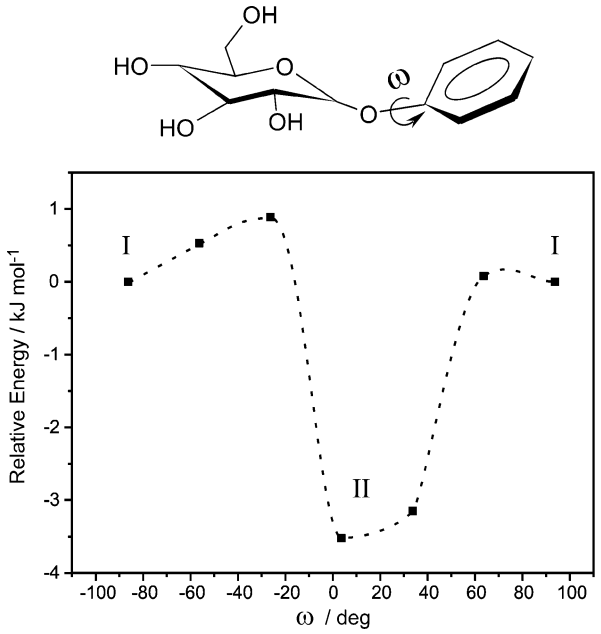


**Figure 5.** Relaxed potential energy scan of the phenyl torsion. The dashed curve presents a guide to the eye and has no physical meaning.

were all stabilized by ca. 4 kJ mol$^{-1}$, providing B3LYP relative energies of 0.0, 1.2, and 2.2 kJ mol$^{-1}$, respectively, and noticeably improving the agreement with the experimentally observed population ratios of 1, 0.37, and 0.1. Single-point MP2/6-311+G(d,p) calculations on the optimized structures further destabilized the trans rotamer with respect to the two gauche structures, yielding relative energies of 0, 1.2, and 4.7 kJ mol$^{-1}$. These MP2 energies also favored the flat orientation of the phenyl ring by ca. 1.5 kJ mol$^{-1}$ for all three rotamers of the bare Phe$\beta$Glc. This trend was, however, not preserved in the Phe$\beta$Glc$-(H_2O)$ clusters; no consistent preference was observed for either one of the two phenyl ring orientations.

*Phe$\beta$Glc$-(H_2O)$ Complexes: Optimized Structures and Relative Energies.* Figure 6 shows the six lowest-energy conformational structures of the singly hydrated Phe$\beta$Glc$-(H_2O)$ complex (optimized at the B3LYP/6-31+G(d) level of theory) together with their relative energies (calculated at the MP2/6-311+G(d,p)//B3LYP/6-31+G(d) level). An extended list of relative energies, including all of the structures lying within 6 kJ mol$^{-1}$ of the global minimum, is presented in Table 3. The low-energy conformers are all associated with insertion structures,[32] labeled X(I or II)ins_n where X indicates the conformation of the sugar, I or II indicates the twisted or flat orientation of the phenyl group, and n describes the proton-donating OH group in the insertion chain, $(OH)_n \rightarrow (OH)_{water} \rightarrow O_{acceptor}$. At both levels of applied theory, the G+g− conformation remains the preferred arrangement of the glucose ring upon single hydration.

Despite their large number—more than 10 structures are calculated to lie within the lowest 5 kJ mol$^{-1}$ of the global minimum, see Table 3—there is a significant gap between the three lowest-lying structures, G+g−ins_2, G+g−ins_6, and G−g+ins_2, and the rest. The next most stable structure, G+g−ins_3, lies 2.6 kJ mol$^{-1}$ above the global minimum, suggesting (on an energetic basis) the strong population of only three clusters in the gas phase. In each of the three lowest-lying structures, the water molecule selects one of the two lowest-energy conformers of the unhydrated sugar (i.e., G+g− or G−g+) and inserts itself either between (OH)2 and phenoxy oxygen atom O1 or between (OH)6 and ring oxygen atom O5. The structural consequences of the insertion at position 2

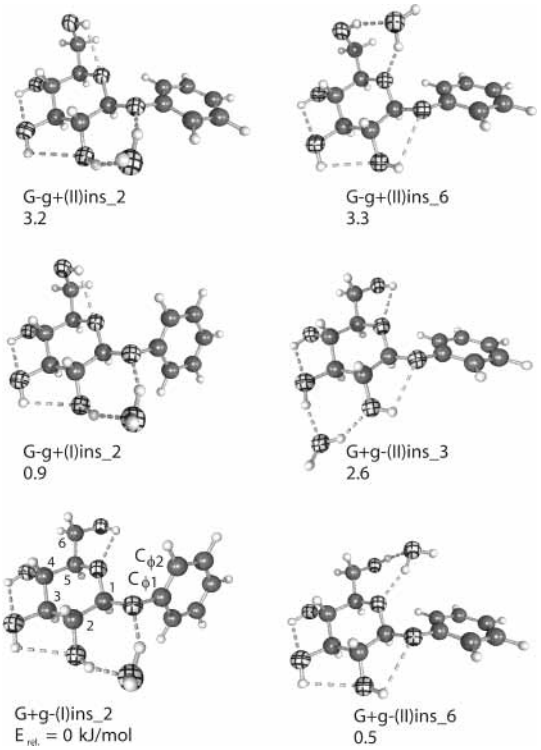


**Figure 6.** Low-energy conformers of Phe$\beta$Glc$-(H_2O)$. Relative energies (in kJ mol$^{-1}$) listed are at the MP2/6-311+G(d,p)//B3LYP/6-31+G(d) level.

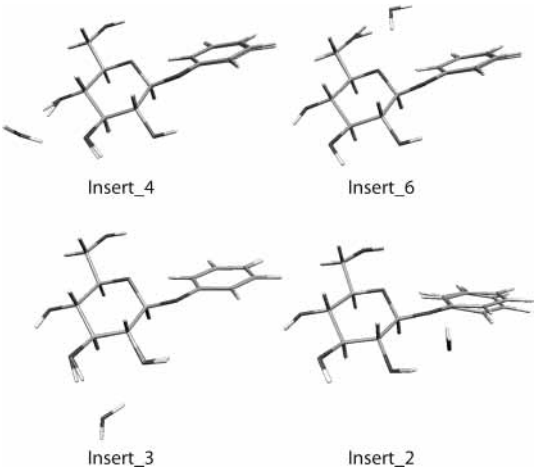


**Figure 7.** G+g−II conformer of Phe$\beta$Glc complexed with water: the unhydrated molecule is shown superimposed with complexes with water inserted at each of the four most favored positions. The phenyl ring is significantly affected by water molecule insertion only at position 2.

compared to that of the other positions are illustrated in Figure 7; only an insertion at position 2 appears to disturb the orientation of the phenyl ring because of the (direct) hydrogen-bonded interaction of the water with phenoxy oxygen O1. The two lowest-energy structures are consistent with the observed UV shifts of the band origins of the two clusters, P and U: assigning the more strongly populated cluster, P, to the global minimum structure, G+g−ins_2, in which the water molecule is H-bonded to phenoxy oxygen O1, which is directly bonded to the phenyl ring, would explain the strong red shift of its origin band (Figure 7). Assigning cluster U to the next most stable structure G+g−ins_6, in which the bound water molecule is relatively far away from the phenyl ring, would account both for its lower intensity and the small UV spectral shift.

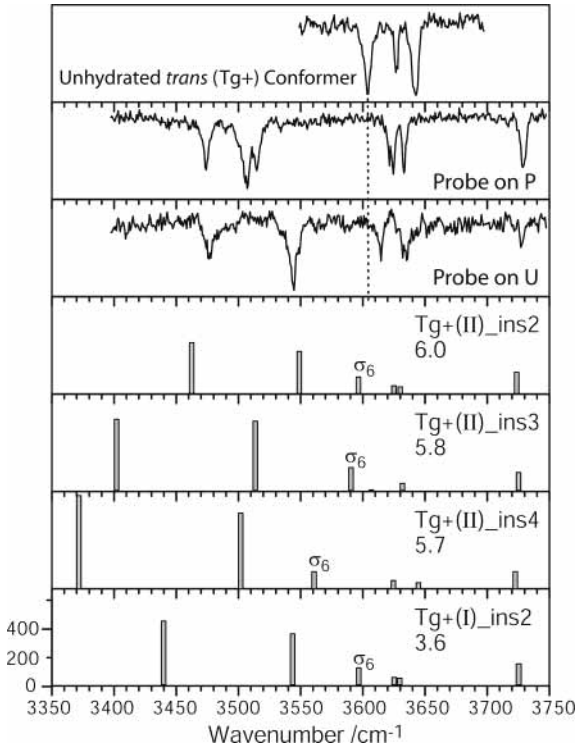


**Figure 8.** Comparison between the computed IR spectra of the lowest-energy trans conformers and the experimental IR spectra of bare and hydrated Phe$\beta$Glc. Relative energies of the conformers are listed in kJ mol$^{-1}$.

*Phe$\beta$Glc−($H_2O$) Complexes: Infrared Spectra.* The computed infrared spectra of almost all of the (insertion) conformers contain (in order of decreasing wavenumbers) one free O−H stretching vibration associated with the bound water molecule ($\sigma_{wf}$), a trio of very weakly perturbed sugar O−H stretching vibrations, and two more strongly shifted O−H modes, associated with the two H-bonded OH groups involved in the insertion of the water molecule into the sugar ring; one of them (labeled $\sigma_{wb}$) is associated with the water, and the other (labeled $\sigma_s$) is associated with the sugar. This broad pattern is in agreement with those of the two complexes, P and U, observed experimentally, confirming their assignment to singly hydrated insertion structures.

A more detailed structural and conformational assignment, based solely on a "match" between the computed and experimental infrared spectra in the O−H stretching region, presents some problems, however. The exclusion of a trans conformation can almost certainly be confirmed because the characteristic pattern displayed by the trio of weakly perturbed OH modes in the calculated spectra in which one band is shifted ∼30 cm$^{-1}$ to lower wavenumber compared to its partners, (Figure 8) does not match the observed experimental IR spectra. The absence of a trans conformation in the singly hydrated sugar cluster reflects the experimental liquid-phase populations of methyl $\beta$Glc derived from NMR measurements in aqueous solution.[16−18] The IR spectra computed for each of the lowest-lying gauche conformers (at the B3LYP/6-31+G(d) level) all display fairly similar spectral patterns; unfortunately, none of them is able to provide a fully quantitative match with the experimental spectra (Figure 3). As noted previously, none of the calculated spectra reproduces the small gap between the two H-bonded OH stretching bands displayed by complex P. However, the structure associated with the global minimum, G+g−(I)ins_2, does (uniquely among the low-energy structures) display the 2−1 pattern of weakly perturbed OH stretches associated with complex P, and as noted earlier, this insertion structure closely resembles that of singly hydrated 2-phenoxy ethanol, which has an IR spectrum that is strikingly similar to that of conformer P (Figure 4). These factors, taken together, all favor the assignment of complex P to the G+g−(I)ins_2 structure. The structural assignment of cluster U is more problematic. Several of the computed IR spectra associated with low-lying gauche cluster structures provide a fairly good match with the O−H vibrational spectrum of cluster U, particularly those labeled G+g−(II)-ins_6 and G−g+(II)ins_2. The first of these lies closest to the global minimum, and its structure locates the bound water molecule relatively far from the phenyl ring, consistent with the low UV shift of conformer U's band origin; cluster U, therefore, is tentatively assigned to the G+g−(II)ins_6 structure.

*Caveat.* The failure of the computational method used here to reproduce quantitatively the positioning or spacing of the strongly H-bonded bands in complex P casts serious doubts on the capacity of the B3LYP density functional theory calculations used to estimate correctly these stronger intermolecular interactions to the level of precision necessary for reliable discrimination between a broad range of possible structures. For example, the computed OH frequencies of the G−g+ins_2 structure (both rotamers I and II) seem to match fairly well those of conformer U. Assigning conformer P to G+g−(I)ins_2 and conformer U to another ins_2 structure, however, does not seem reasonable considering the rather large difference in the shift of the $\sigma_{wb}$ stretching mode between the two observed conformers. Indeed, such an assignment would imply that rotating the hydroxy methyl group through ca. 90° from G+g− to G−g+ is enough to affect the water molecule inserted at position 2, located on the opposite side of the ring, and shift the (OH)2 stretching frequency, $\sigma_{s=2}$, by 34 cm$^{-1}$, which would be quite surprising.

## 4. Conclusions

The resonant 2-photon ionization and IR ion-dip spectra of the singly hydrated glycoside phenyl $\beta$-D-glucopyranoside, recorded in the gas phase in a free jet expansion, have revealed the presence of only two populated structures and have provided their individually resolved UV and IR spectra. The presence of just two distinct complexes indicates a surprising degree of selectivity. The preferred H-bonding sites/structures favored by theory correspond first to an insertion of the water molecule between the (OH)2 group and the phenoxy atom, O1, and second to an insertion between (OH)6 of the hydroxy methyl group and the ring oxygen atom, O5. In both, the gauche hydroxy methyl conformation associated with the global minimum-energy structure of the unhydrated glycoside, G+g−, appears to be retained, but the neighboring G−g+ and trans (Tg+) conformations are no longer sufficiently populated to be observed.

Although the broad pattern in the O−H vibrational structure of the examined complexes can be reproduced by harmonic frequency calculations conducted at the B3LYP/6-31+G(d) level of theory, subtle differences between their infrared spectra, particularly those associated with the binding site, cannot. Higher-level ab initio calculations are required to describe better the noncovalent interactions of water with the numerous hydrogen-bonding sites of sugar rings. Further computational investigations are currently in progress, together with experimental investigations of deuterated derivatives and of a modified glycoside, phenyl $\beta$-D-xylopyranoside, in which the hydroxy methyl group of the sugar is replaced by a hydrogen atom.

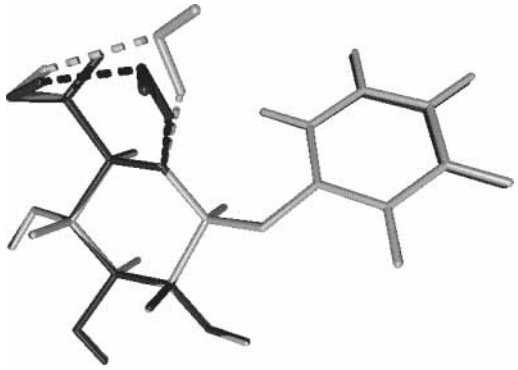

**Figure 9.** Overlay of the two structures of the G−g+ins_6 conformer of Phe$\beta$Glc−($H_2O$) optimized using revision A.7 (both structures) and revision A.11 (dark gray) of the Gaussian 98 package.

**Acknowledgment.** We have benefited from the support provided by the Leverhulme Trust (award F/08788D), the Royal Society (R.A.J., USA Postdoctoral Research Fellowship), the EPSRC, the Laser Support Facility of the CLRC (laser loan), and the Physical and Theoretical Chemistry Laboratory at Oxford. J.P.S. has also benefited greatly over the years from the friendship, acumen, and wisdom of Professor Charles Parmenter.

## Appendix

In the course of the computational investigations, a discrepancy between the default optimization options of the various revisions of the Gaussian 98 (G98) package was encountered. When optimizing the singly hydrated complex, the latest revision, namely, revision A.11 (R11), produced, in some cases, an optimized structure that differed from those generated by the earlier revisions, A.7 (R7) or A.3 (R3). An example is illustrated in Figure 9. Both of the conformers shown are minima (confirmed by analytical frequency calculations) on the B3LYP/6-31+G(d) potential energy surface according to G98 revision A.7, with the dark-gray conformer more stable by (a sizable!) 2 kJ/mol$^{-1}$. However, only the (lower-energy) dark-gray conformer is a minimum according to R11. Not surprisingly, the calculated O−H stretching frequencies associated with the two computed structures are significantly different.[33] The difference in the assessment of stationary points is apparently due to an "improved selection of the redundant internal coordinates" in revision A10/11 for weakly bound complexes (see the release notes of revision A10/11); the redundant internal coordinates are the default coordinate system for optimizations in G98. For the structure shown, a total of 22 additional coordinates are defined in the default R11 optimization compared to the number defined in the R7 procedure. The manual addition of these extra coordinates to an R7 optimization using the ModRedundant keyword yielded results identical to those obtained with R11, confirming the extra redundant coordinates as the source of the difference between optimized structures. This change in the default optimization process complicates the comparison between results obtained with revision A.10 and later revisions and those obtained using revision A.7 and earlier revisions. It is necessary to be aware of this difference when comparing new calculations to old ones. In the present paper, all of the structures were initially optimized with Gaussian 98 revision A.7 and subsequently reoptimized using revision A.11.

## References and Notes

(1) Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562.
(2) Philips, L. A.; Levy, D. H. *J. Phys. Chem.* **1986**, *90*, 4921.
(3) Cable, J. R.; Tubergen, M. J.; Levy, D. H. *J. Am. Chem. Soc.* **1987**, *109*, 6198.
(4) Cable, J. R.; Tubergen, M. J.; Levy, D. H. *Faraday Discuss.* **1988**, 143.
(5) Dian, B. C.; Longarte, A.; Zwier, T. S. *Science* **2002**, *296*, 2369.
(6) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wales, D. J.; Zwier, T. S. *J. Chem. Phys.* **2002**, *117*, 10688.
(7) Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1.
(8) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827.
(9) Simons, J. P. *Compt. Rend. Chim.* **2003**, *6*, 17.
(10) Dong, F.; Miller, R. E. *Science* **2002**, *298*, 1227.
(11) Barrows, S. E.; Dulles, F. J.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *Carbohydr. Res.* **1995**, *276*, 219.
(12) Barrows, S. E.; Storer, J. W.; Cramer, C. J.; French, A. D.; Truhlar, D. G. *J. Comput. Chem.* **1998**, *19*, 1111.
(13) Csonka, G. I.; Elias, K.; Csizmadia, I. G. *Chem. Phys. Lett.* **1996**, *257*, 49.
(14) Molteni, C.; Parrinello, M. *J. Am. Chem. Soc.* **1998**, *120*, 2168.
(15) Jockusch, R. A.; Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1502.
(16) Nishida, Y.; Hori, H.; Ohrui, H.; Meguro, H. *J. Carbohydr. Chem.* **1988**, *7*, 239.
(17) Tvaroška, I.; Taravel, F. R.; Utille, J. P.; Carver, J. P. *Carbohydr. Res.* **2002**, *337*, 353.
(18) Bock, K.; Duus, J. Ø. *J. Carbohydr. Chem.* **1994**, *13*, 513.
(19) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. *Chem. Phys. Lett.* **2000**, *321*, 49.
(20) Snoek, L. C.; Kroemer, R. T.; Hockridge, M. R.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1819.
(21) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.7; Gaussian, Inc.: Pittsburgh, PA, 1998.
(22) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 1998.
(23) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
(24) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(25) Hariharan, P. C.; Pople, R. A. *Theor. Chim. Acta* **1973**, *28*, 213.
(26) Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. *J. Comput. Chem.* **1983**, *4*, 294.
(27) Moller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618.
(28) Macleod, N. A.; Simons, J. P. *Chem. Phys.* **2002**, *283*, 221.
(29) Butz, P.; Kroemer, R. T.; Macleod, N. A.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3566.
(30) Graham, R. J.; Kroemer, R. T.; Mons, M.; Robertson, E. G.; Snoek, L. C.; Simons, J. P. *J. Phys. Chem. A* **1999**, *103*, 9706.
(31) Gerhards, M.; Unterberg, C.; Kleinermanns, K. *Phys. Chem. Chem. Phys.* **2000**, *2*, 5538.
(32) Instead of insertion, the bound water molecule might simply form an addition structure, acting either as a proton donor to one of the oxygen atoms (cf. the singly hydrated neurotransmitters, ephedrine, and pseudoephedrine[29]) or the noradrenaline mimic 2-amino-1-phenylethanol[30] or as a proton acceptor from just one of the OH groups (cf. the singly hydrated catechol cluster[31]). The preexisting network of internal H bonds in the glucopyranose ring limits the potential addition-only sites to O6 and O4 for the two gauche conformers and O6 and O5 for the trans conformer. When these structures were optimized at the B3LYP/6-31+G(d) level, however, they were located at energies ≥17 kJ mol$^{-1}$ above the global minimum. Similarly, if the water were to act only as a proton acceptor, then it would undoubtedly act also as a proton donor to the nearby oxygen

atom (to create an insertion structure) because every hydroxyl group on the sugar ring is close to another oxygen atom, with which it interacts weakly.

(33) When contacted regarding these difficulties, Gaussian technical support confirmed that older versions of Gaussian suffered from an algorithmic error in how coordinates were chosen for some complexes, which could result in false minima. Here, we highlight the dangers of over-reliance on default optimization procedures because we believe this may be relevant to the increasing number of researchers studying noncovalent complexes.