DOI: 10.1002/chem.200901830

Hydration of Sugars in the Gas Phase: Regioselectivity and Conformational Choice in *N*-Acetyl Glucosamine and Glucose

Emilio J. Cocinero,[a] E. Cristina Stanca-Kaposta,[a, c] Mark Dethlefsen,[a] Bo Liu,[a] David P. Gamblin,[b] Benjamin G. Davis,*[b] and John P. Simons*[a]

Abstract: The influence of an acetamido group in directing the preferred choice of hydration sites in glucosamine and a consequent extension of the working rules governing regioselective hydration and conformational choice, have been revealed through comparisons between the conformations and structures of "free" and multiply hydrated phenyl *N*-acetyl-β-D-glucosamine (βpGlcNAc) and phenyl β-D-glucopyranoside (βpGlc), isolated in the gas phase at low temperatures. The structures have been assigned through infrared ion depletion spectroscopy conducted in a supersonic jet expansion, coupled with computational methods. The acetamido motif provides a hydration focus that overwhelms the directing role of the hydroxymethyl group; in multiply hydrated βpGlcNAc the water molecules are all located around the acetamido motif, on the "axial" faces of the pyranose ring rather than around its edge, despite the equatorial disposition of all the hydrophilic groups in the ring. The striking and unprecedented role of the C-2 acetamido group in controlling hydration structures may, in part, explain the differing and widespread roles of GlcNAc, and perhaps GalNAc, in nature.

Keywords: carbohydrates · hydrogen bonding · IR spectroscopy · solvation

Introduction

N-Acetyl glucosamine (GlcNAc) and glucose (Glc) are the primary building blocks of many biopolymers. The disaccharide unit, cellobiose (Glc-β1,4-D-Glc), for example, provides the basic, repeat structural unit of cellulose which is a key component of plant cell walls. Similarly, chitobiose (GlcNAc-β1,4-D-GlcNAc) provides the repeat unit in chitin, found in the exo-skeletons of many insects and crustaceans; it also provides the core linkage (to asparagine residues) in all *N*-linked glycoproteins. GlcNAc residues are key building blocks in many proteoglycans, which are involved in a host of biological functions, for example, heparan sulphate, a hetero-glycosaminoglycan which incorporates an alternating co-polymer rich in glucosamine and uronic acid residues.[1,2]

Despite their important role in many living organisms, there have been very few spectroscopic[3] or structural investigations of 2-deoxy amino sugars, such as GlcNAc or GalNAc,[4,5] although the conformational structures and anomeric preferences of the corresponding sugars, Glc and Gal, and many other hexapyranose carbohydrates, have been the focus of a large number of molecular dynamics simulations and NMR measurements (see Kraütler et al.[6] for a very thorough review of the recent literature). There is now a general acceptance that their (dynamical) conformational preferences in aqueous solutions are strongly influenced by hydrogen-bonded interactions, both externally, with the solvent and internally, within the carbohydrate.[4,7,8] Until recently, however,[9–11] there were no benchmark experimental investigations of their intrinsic conformational structures (or those of oligosaccharides containing them) in the absence of a bulk solvent, or of the influence of explicit

[a] Dr. E. J. Cocinero, Dr. E. C. Stanca-Kaposta, M. Dethlefsen, Dr. B. Liu, Prof. J. P. Simons
Chemistry Department, University of Oxford
Physical and Theoretical Chemistry Laboratory
South Parks Road, Oxford OX1 3QZ (UK)
Fax: (+44) 1865-275410
E-mail: John.Simons@chem.ox.ac.uk

[b] Dr. D. P. Gamblin, Prof. B. G. Davis
Chemistry Department, University of Oxford
Chemistry Research Laboratory, 12 Mansfield Road
Oxford OX1 3TA (UK)
Fax: (+44) 1865-285002
E-mail: Ben.Davis@chem.ox.ac.uk

[c] Dr. E. C. Stanca-Kaposta
Institut für Experimentalphysik, Freie Universität Berlin
Arnimallee 14, 14195 Berlin (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.200901830.

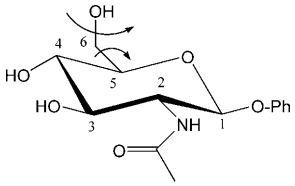
and specific hydration on their conformational preferences in the gas phase. Those that had been available were almost all computational and many were dependent on the quality of the assumed force field.[6,12–14]

In the last few years it has become possible to derive this information experimentally through vibrational spectroscopy conducted under molecular beam conditions, combined with molecular mechanics, density functional and ab initio theoretical calculations.[9–11] Investigations of this kind have led to a set of predictive "working rules", governing binding site selectivity, conformational choice and the influence of spatially localized hydrophilicity (vs hydrophobicity) of individual mono- and oligosaccharides.[9] Such information may be at least as important as that obtained in aqueous solution since it provides a reference point (a ground state) against which the effect of the introduction of environment, including biologically relevant environments, may be gauged. Much of the biological activity of oligosaccharides and glycans, for example, in cell adhesion and signaling, is mediated through membrane bound protein–carbohydrate interactions.[15–17] This can involve direct hydroxyl hydrogen-bonding and purely dispersive "CH–π" binding,[18–20] both of which may, or may not displace the water molecules surrounding the carbohydrate or the protein[21] or binding may alternatively be mediated by bound water molecules.[22,23] In all of these scenarios it is likely that the carbohydrate conformations associated with complexed or explicitly, and sometimes only partially, hydrated structures in the gas phase might provide a more precise model than dynamical structures in aqueous solution.

The present work reports the first experimental investigations of the intrinsic gas-phase conformation and structure of an amino sugar, phenyl *N*-acetyl-β-D-glucosamine[1] (βpGlcNAc) shown in Scheme 1, and its singly and multiply hydrated complexes, isolated at low temperatures in a molecular beam. The influence of the C-2 acetamido group in directing the preferred choice of hydration sites, and a consequent extension of the working rules governing selective hydration and conformational choice, are revealed through comparisons with multiply hydrated phenyl β-D-glucopyranoside (βpGlc) and recent investigations[9,10] of other hydrated monosaccharides.

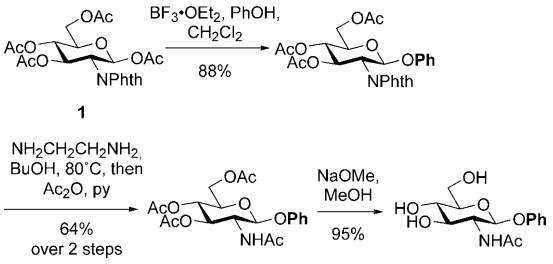
Results

Synthesis: β-Phenyl *N*-acetylglucosamine (βpGlcNAc), bearing an appropriate C-1 phenyl chromophore, was synthesized using the known phthalimide protected peracetyl donor **1** (see Scheme 2). Lewis acid-mediated glycosylation of phenol proved superior to the use of either GlcNAc-derived halo sugar donors or the use of pentaacetyl-D-glucosamine and proceeded in good yield and with excellent stereoselectivity (>98% β) by virtue of participation of the C-2 imide. Use of pentaacetyl-D-glucosamine bearing an acetamide at C-2 resulted in reduced yield in part due to competing oxazoline formation. Removal of the phthalimide and installation of the acetamide at C-2 followed by Zémplen deprotection proceeded cleanly to give βpGlcNAc in 54% overall yield from **1**.



dihedral angle: O5-C5-C6-O6	
-60°	G-
+60°	G+
180°	T
dihedral angle: C5-C6-O6-H6	
-60°	g-
+60°	g+
180°	t

Scheme 1. Schematic representation of phenyl *N*-acetyl-β-D-glucosamine and the notation indicating possible conformations of the *exo*-cyclic hydroxymethyl group.



Scheme 2. Synthesis of βpGlcNAc.

UV and IR spectroscopy of βpGlcNAc: The resonant two-photon ionization (R2PI) spectrum of βpGlcNAc recorded under molecular beam conditions is shown in Figure 1 together with two of its UV "hole-burn" spectra, recorded with the probe laser tuned either to the band origin at 36894 cm^{-1}, labeled "1", or the vibronic band at 36960 cm^{-1}, labeled "2". The hole-burn spectra are rather noisy but while the first clearly displays each of the leading four bands in the vibronic progression, the next band, 2, is lost in the noise; the same result was obtained when the UV probe was tuned onto the other vibronic bands lying between peaks 1 and 2. The situation is reversed in the second hole-burn spectrum where band 2 is noticeably stronger but the first four bands can only just be discerned. These spectra suggest the population of more than one conformer in the cold molecular beam, associated with inter-leaved vibronic spectra.

The infrared ion depletion (IRID) spectra shown in the top half of Figure 2 reinforce this interpretation. The two

[1] The phenyl group provides the UV chromophore required for the 2-photon ionisation process; DFT and ab initio calculations have confirmed that its presence has a negligible influence on the conformational landscape of the carbohydrate.[10]

experimental spectra, labeled 1 and 2, which were recorded with the UV probe laser centred on the corresponding vibronic bands, although having several features in common are not the same. A careful analysis of the two IRID spectra and comparisons with the predictions of the DFT and ab initio calculations shown in the lower half of Figure 2, lead to their association with one or more of the three lowest energy conformers of the glucosamine.

The global minimum conformation, ccG+g−, is predicted to lie at an energy 2.0 kJ mol^{-1} below that of the ccG−g+ conformer and 3.2 kJ mol^{-1} below the ccTg+ conformer.[2] Their free energies calculated at 298 K follow the same rank order. In all three of these conformers, a strong hydrogen bond to the acetamido group, OH3→O=C−NH with an OH3−O distance ≈1.75 Å, intensifies and broadens the OH vibrational band σ3 and displaces it by some 300 cm^{-1} towards lower wavenumber, at ≈3300 cm^{-1}. It also "rotates" the acetamido group towards the pseudo plane of the pyranosyl ring and the dihedral angle, H2-C2-*N*-H is ≈76° in all three conformers. The fourth lowest-lying conformer, with a relative energy of 5.6 kJ mol^{-1} also has a ccG+g− conformation but since there is no OH3→O=C−NH hydrogen bond the acetamido group can adopt an out-of-plane "perpendicular-NH-up" orientation and the H2-C2-*N*-H angle moves to ≈−16°. The computed shift of the NH vibration to higher wavenumber also reflects the loss of the weak NH→O1 hydrogen bonding present in the lower energy conformers.

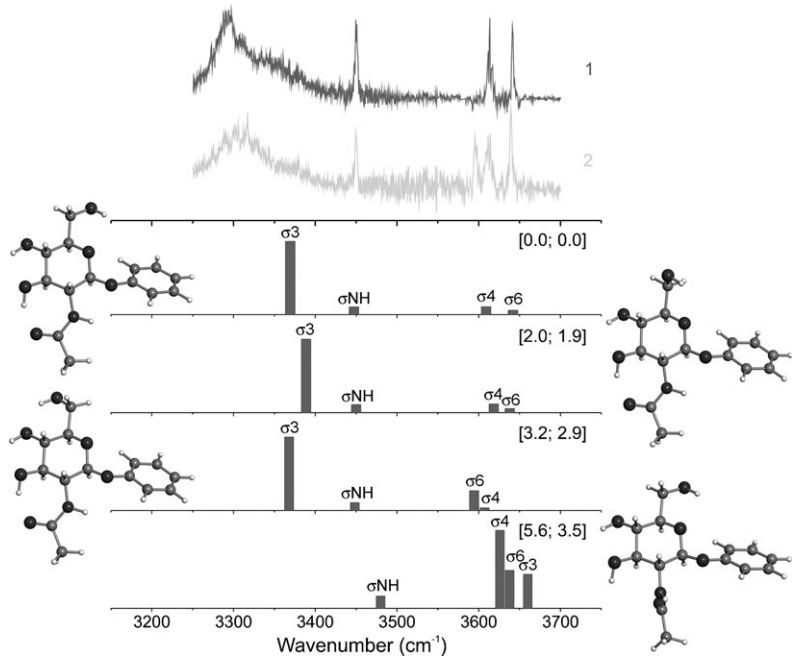


Figure 2. Top: IR ion dip spectra of βpGlcNAc recorded with the UV laser centred on the vibronic bands labeled 1 and 2 in Figure 1. Bottom: computed vibrational spectra and structures of the four lowest energy conformers. Their relative energies in kJ mol^{-1}, at 0 K, and free energies, at 298 K, are shown in brackets.

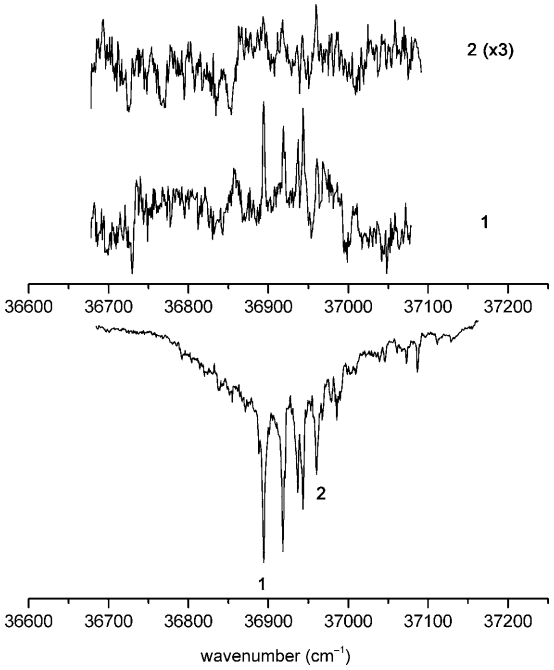


Figure 1. R2PI spectrum of βpGlcNAc (bottom trace) and UV/UV hole-burn spectra recorded with the probe laser centred on the vibronic peaks 1 and 2 (upper traces).

The IRID spectrum 1 is in reasonable accord with the vibrational spectrum predicted for the global minimum conformer but the assignment of spectrum 2 is more subtle. The appearance of a new band at ≈3595 cm^{-1} and the broadening of its nearest neighbour, centred around 3610 cm^{-1}, indicate contributions from more than one conformer. Furthermore, while the band at 3450 cm^{-1}, associated with the "free" NH mode, is unchanged there are small but measurable displacements in the sharp band located at ~3640 cm^{-1} (to lower wavenumber) and the broad band associated with σ3 (to higher wavenumber). Comparisons with the UV hole-burn and the computed vibrational spectra favour its assignment principally to a combination of the ccG−g+ and ccTg+ conformers, together with a minor component associated with the global minimum conformer, ccG+g−. The

[2] The notation is similar to that employed in previous publications.[9,10] "cc" indicates the counter-clockwise orientation of the OH groups on the pyranosyl ring and G+g−, G−g+, Tg+ indicate the orientation of the exocyclic hydroxy (g+, g−) methyl (G+, G−, T) group—see Scheme 1. Alternative orientations of the acetamido group are designated as "parallel" (dihedral angle, H2-C2-*N*-H ~90°), "perpendicular-NH up" (H2-C2-*N*-H≈0°), or "perpendicular-NH down" (H2-C2-*N*-H ≈180°).

"new" band appearing at $\approx 3595\ cm^{-1}$ can then be assigned to the displaced mode σ6, in the ccTg+ conformer, shifted by $\approx 50\ cm^{-1}$ to lower wavenumber through the weak OH6→OH4 interaction, OH6–O4 ≈2.07 Å.

Phenyl β-D-glucopyranoside, βpGlc, displays a similar conformational landscape, presenting the same rank order for its three lowest energy conformers and the same counter clockwise "cc" orientation of the hydroxyl groups, OH4, OH3 (and OH2),[10] promoted by the weak interaction, OH2→O1, OH2-O1≈2.2 Å. In βpGlcNAc, where OH2 is replaced by an acetamido group, the orientation is imposed by the much stronger OH3→O=C hydrogen bond. This is also reflected in the position of the band σ4, which lies at $\sim 3630\ cm^{-1}$ in the three lowest energy conformers of βpGlc[10] but is shifted to $\sim 3610\ cm^{-1}$ in the corresponding conformers of βGlcNAc, reflecting the enhanced co-operativity in the OH4→OH3→O=C–NH chain.

The hydrated complexes, βpGlcNAc·(H$_2$O)$_{n=1-3}$: When the Ar gas stream was seeded with water vapour, the time of flight mass spectrum of photo-ionised βpGlcNAc displayed a series of peaks corresponding to singly, and multiply hydrated complexes. Their R2PI spectra shown in Figure 3 are all broadened, though there is a hint of residual structure, and each one is shifted by $\approx 200\ cm^{-1}$ to lower wavenumber compared to the bare carbohydrate.

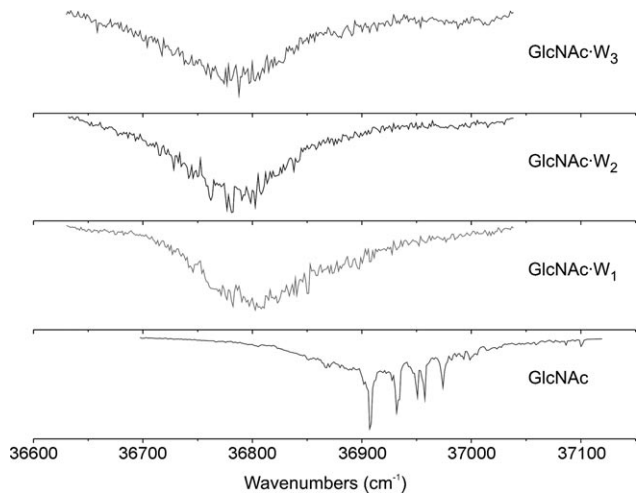


Figure 3. R2PI spectra of singly, doubly and triply hydrated complexes of βpGlcNAc recorded in the corresponding βpGlcNAc·$(H_2O)_{n=1-3}^+$ ion channels; the corresponding spectrum of "bare" βpGlcNAc is also shown for comparison.

The mono-hydrate, βpGlcNAc·(H$_2$O): IRID spectra recorded in the βpGlcNAc·$(H_2O)^+$ ion channel, with the UV probe laser tuned to different positions near ($36809\ cm^{-1}$), and slightly below ($36781\ cm^{-1}$) the peak of the R2PI spectrum, are shown in Figure 4, together with the computed vibrational spectra of the two lowest energy βpGlcNAc·(H_2O) structures. The two experimental IRID spectra, while similar, are not identical which indicates contributions to the R2PI spectrum from more than one structure. Comparisons with the computed spectra are in good qualitative agreement with an assignment in which each of the two IRID spectra includes a contribution, in comparable but slightly different proportions, from the two lowest lying structures. Each one retains the ccG+g– conformation of the monosaccharide; in the global minimum, the pre-existing OH3→O=C bond remains intact with the water molecule bound in an addition complex, to the acetamido group through the hydrogen bond sequence OH3→O=C–NH→OH_W. The acetamido group re-orients into the pseudo plane of the pyranosyl ring and the dihedral angle, H2-C2-*N*-H is ≈120°. In the other structure, which lies very close in energy, the water molecule inserts into the OH3→O=C bond and it is bound through the sequence OH3→OH_W→O=C–NH; the acetamido group now rotates out of the plane into a "perpendicular-NH down" orientation with an angle H2-C2-*N*-H ≈172°. Assignment of both structures to a singly hydrated complex is confirmed by the appearance of only one sharp feature at $\approx 3720\ cm^{-1}$, which is associated in the addition complex, with the antisymmetric OH_W vibration, $\sigma_{W(a)}$, and in the insertion complex with the "free" OH_W vibration, σ_{Wf}.

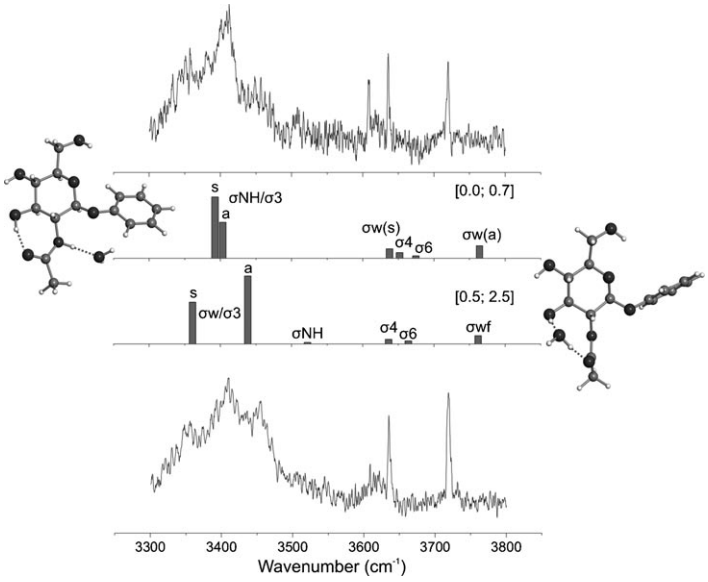


Figure 4. IR ion dip spectra recorded in the βpGlcNAc·$(H_2O)^+$ ion channel with the UV probe laser tuned to $36{,}781\ cm^{-1}$ (top) and to $36{,}809\ cm^{-1}$ (bottom), and computed vibrational spectra, structures and relative energies ($kJ\,mol^{-1}$) of the two lowest energy βpGlcNAc·(H_2O) complexes.

The di-hydrate, βpGlcNAc·(H$_2$O)$_2$: Figure 5 presents the IRID spectrum recorded in the βpGlcNAc·$(H_2O)_2^+$ in channel, with the UV probe laser tuned near to the peak of the R2PI absorption (like the spectra shown in Figure 4, detuning the probe toward lower wavenumber, again led to a slight change in the spectra profile). Comparisons with the

computed vibrational spectra, also shown in Figure 5, indicate its association predominantly, with the global minimum energy structure together with a minor component associated with its closest neighbour, which has the same free energy at 298 K and a relative energy $\approx 1\,kJ\,mol^{-1}$ (at 0 K). Each one retains the preferred ccG+g− conformation of the bare monosaccharide and accommodates a bound water dimer (and not two separate water molecules) inserted between the acetamido and the OH3 groups, breaking the OH3→O=C−NH bond to create extended co-operative loops, OH4→OH3→OH_{W1}→OH_{W2}→O=C−NH. The water dimer lies above the pyranosyl ring in the global minimum structure, where the acetamido plane adopts a "perpendicular-NH down" configuration with a dihedral angle H2-C2-*N*-H≈162° (Figure 5b). In the neighbouring structure the plane rotates through ≈180° into a "perpendicular-NH up" orientation with a dihedral angle ≈1°, and the bound water dimer lies below the pyranosyl ring (see Figure 5c for comparison). In both cases, the location of the bound water molecules is dictated by the location of the carbonyl group. The corresponding "perpendicular-NH up" conformer of bare βpGlcNAc (H2-C2-*N*-H ≈16°) lies well above the global minimum with a calculated relative energy of $5.6\,kJ\,mol^{-1}$, see Figure 2, and it is not populated in the molecular beam.

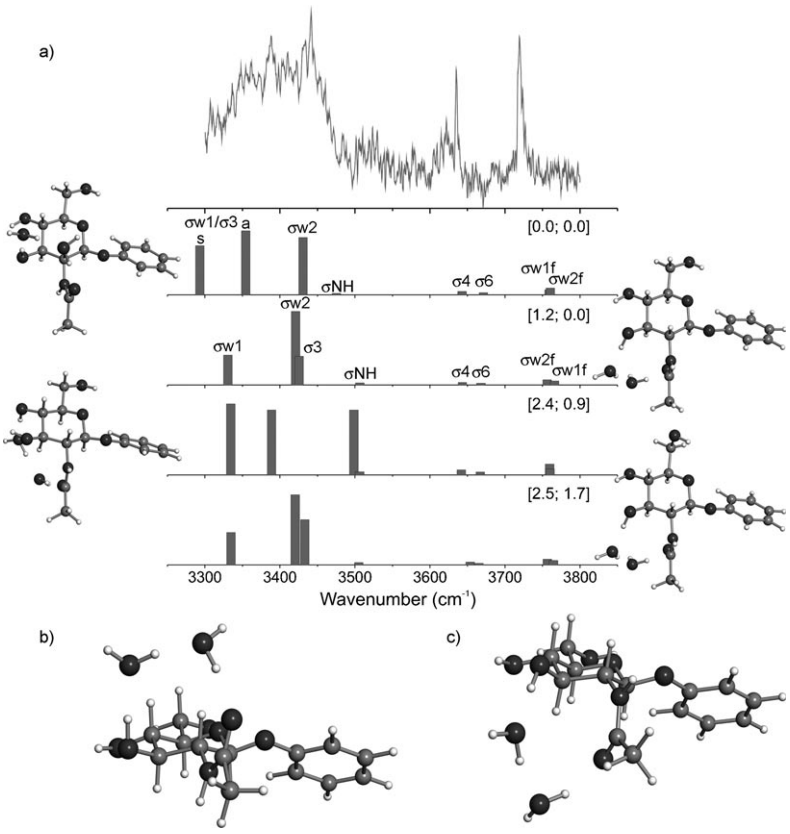
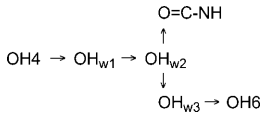


Figure 5. a) IR ion dip spectrum recorded in the βpGlcNAc·$(H_2O)_2^+$ ion channel with the UV probe laser tuned to 36,809 cm^{-1} and the computed vibrational spectra, structures, relative energies and free energies ($kJ\,mol^{-1}$) of the four lowest energy βpGlcNAc·$(H_2O)_2$ complexes. (b,c) "side-on" views of the two, lowest energy populated structures, "NH-down" (b) and "NH-up" (c).

The tri-hydrate, βpGlcNAc·$(H_2O)_3$: The IRID spectra recorded at the two UV probe frequencies in the trihydrated ion channel are shown in Figure 6, together with the computed vibrational spectra of its two lowest-lying structures which have almost identical relative energies but lie $\geq 3.6\,kJ\,mol^{-1}$ below their nearest neighbours. As with the other hydrated complexes, the two experimental IRID spectra differ slightly; note particularly, the change in the relative intensity of the band peaking at $\approx 3440\,cm^{-1}$, associated with (a) strongly hydrogen bonded OH group(s). Comparisons with the two computed spectra reveal a composite spectrum reflecting comparable populations of each of the two low energy structures but there is no correspondence with the vibrational spectra associated with the higher energy structures (shown in the Supporting Information).

In the global minimum structure one water molecule is inserted between OH3 and the acetamido C=O group located above the plane of the pyranosyl ring, while the other two are bound as a dimer inserted between OH3 and the acetamido NH group located below the plane, to create the cyclic hydrogen-bonded sequence: OH3→OH_{W1}→O=C−NH→OH_{W2}→OH_{W3}→OH3 (Figure 6b). The alternative structure depicted in Figure 6c accommodates a branched water trimer located below the pyranosyl plane between OH4 and OH6. It incorporates a doubly hydrogen bonded central water molecule connected to the C=O group, to create the sequence:



Like the dihydrate, the hydroxymethyl group continues to retain its G+g− conformation and the acetamido plane is oriented roughly perpendicular to the plane of the pyranosyl ring, either "NH-down" in the global minimum, or "NH-up" in the alternative structure, with H2-C2-*N*-H dihedral angles of 178 or 11°.

βpGlc·$(H_2O)_{n=2}$: The original investigations of phenyl β-D-glucopyranoside (βpGlc) were limited to its singly hydrated complex.[10] Two structures were detected: in the more strongly populated one, associated with the global minimum structure, the water molecule was bound

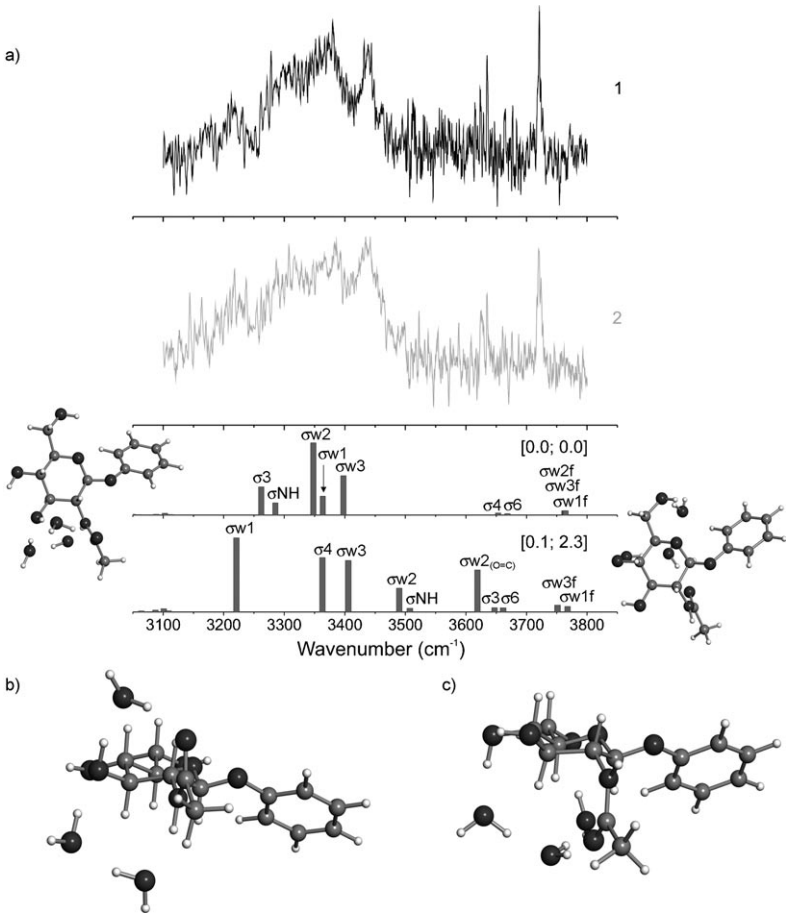


Figure 6. a) IR ion dip spectra recorded in the βpGlcNAc•$(H_2O)_3^+$ ion channel with the UV probe laser tuned, 1) to 36,781 cm^{-1} and 2) to 36,809 cm^{-1} and the computed vibrational spectra, structures, relative energies and free energies ($kJ\,mol^{-1}$) of the two lowest energy βpGlcNAc•$(H_2O)_3$ complexes. b,c) "side-on" views of the two, lowest energy populated structures, "NH-down" (b) and "NH-up" (c).

to the hydroxymethyl group, inserted between OH4 and OH6. This promoted a conformational switch from ccG+ g− to cG−g+, to create an extended, counter-clockwise $(OH\text{-}O)_n$ chain. In the less favoured, alternative structure the closest agreement with experiment indicated retention of the ccG+g− conformation with the bound water molecule attached to the hydroxymethyl group, but this time inserted on the opposite side, between OH6 and the ring oxygen, O5.[10]

The IRID spectrum of the dihydrate, βpGlc•$(H_2O)_2$, is shown in Figure 7a, together with the computed vibrational spectra of its lowest energy structures. The spectrum associated with the global minimum is in good correspondence with observation, but the others are not able to reproduce the experimental spectrum; since the relative energy (and free energy) of the global minimum structure lies well below that of its nearest neighbour, its assignment to the global minimum structure is clear. In order to best accommodate the bound water molecules, now located on each side of the (re-oriented) hydroxymethyl group, the monosaccharide conformation is again reshaped, from ccG+g− to cG−g+, to create the extended sequence, OH2→OH3→OH4→OH_{W1}→OH6→OH_{W2}→O5, in conformity with the general "propensity rules" identified earlier for singly and multiply hydrated monosaccharides. The alternative structures, each of which would accommodate a bound dimer rather than two separate water molecules, are energetically less favoured and are unpopulated.

Although it was possible to detect R2PI signals in the βpGlc•$(H_2O)_3^+$ ion channel they were too weak unfortunately, to enable an IRID spectrum to be recorded at an adequate signal/noise level.

Discussion

Although the intrinsic conformational landscapes of βpGlc and its *N*-acetamido derivative, βpGlcNAc, are quite similar, the enhanced hydrogen bonding to the acetamido group, OH3→O=C−NH, dominates and anchors the orientation of the pyranosyl OH groups, locking them into a counter-clockwise orientation around the ring. More importantly, it also provides a focus for the location of bound water molecules. In hexoses such as glucose, galactose and mannose, hydration regioselectivity is controlled primarily by the flexibility of the exocyclic hydroxymethyl group, and secondarily by the orientation (equatorial or axial) of the pyranosyl hydroxyl groups.[9] Water molecules bound to phenyl α/β-D-glucosides (α/βpGlc) and -mannosides (α/βpMan) inserted preferentially into the attractive "pocket" between OH4 and OH6, which either pre-exists (in αpMan, cG−g+) or is newly created through re-orientation of the hydroxymethyl group and a consequent conformational switch (to cG−g+ in βpMan and α/βpGlc).[9,24] When this site is filled the second water molecule, in βpGlc•$(H_2O)_2$ or βpMan•$(H_2O)_2$ for example, occupies the vacant neighbouring site between OH6 and the ring oxygen, O5. When OH4 is axial rather than equatorial, as in α- or β-galactosides, enhanced hydrogen bonded interactions between OH4 and its nearest neighbour, either OH3 or OH6, prevent insertion of a water molecule into the (4,6) site and instead, the neighbouring (6, O5) site becomes the first port of call for an (in)bound water molecule.[9,24]

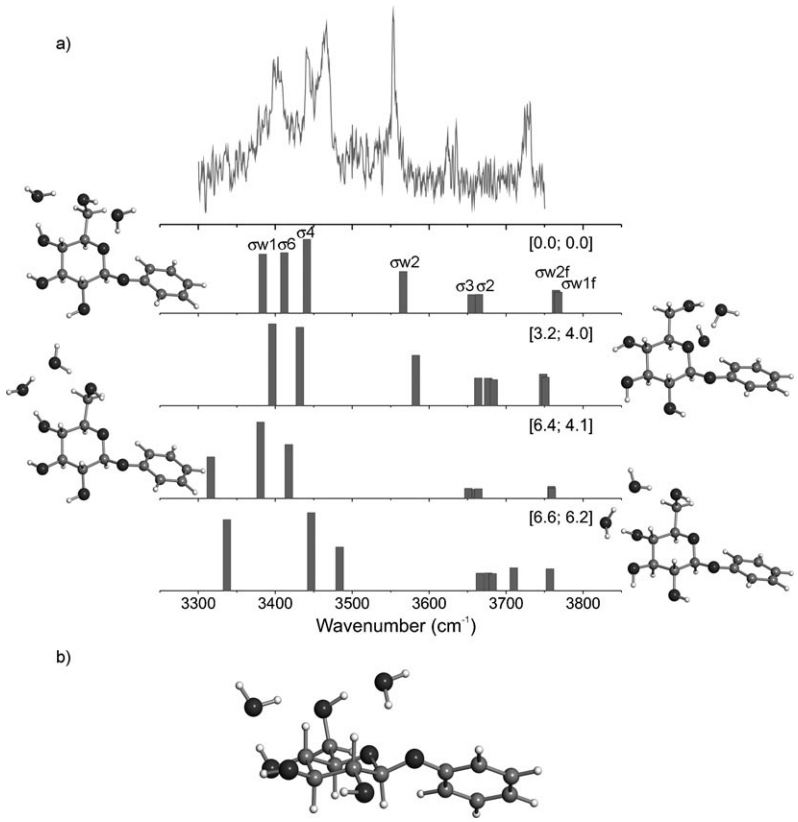


Figure 7. a) IR ion dip spectra of hydrated βpGlc recorded in the βpGlc·$(H_2O)_2^+$ ion channel and the computed vibrational spectra, structures, relative energies and free energies ($kJ\,mol^{-1}$) of the four lowest energy βpGlc·$(H_2O)_2$ complexes. b) Side-on view of the lowest energy, populated structure.

Conclusion

In the 2-deoxyamino sugar, βpGlcNAc, bound water molecules are marshalled into place by the acetamido group which, through the orientation of the C=O and NH, now provides the dominant control. Indeed, so great is its propensity to act as a hydration focus that in βpGlcNAc·$(H_2O)_3$ all three water molecules "loop" the [NHAc-2,OH-3] motif, leaving the remaining motifs "dry". It can act both as a hydrogen bond donor and an acceptor and like the hydroxymethyl group in βpGlc or βpMan, it can fine-tune the preferred choice of binding site by rotating into alternative configurations: "in-plane" (H2-C2-*N*-H ≈90°), "perpendicular-NH up" (H2-C2-*N*-H ≈0°) or "perpendicular-NH down" (H2-C2-*N*-H ≈180°). These properties generate hydrate structures where the bound water molecules are located either above, or below, or above and below the pyranosyl ring, see Figures 5 and 6. Similarly, in doubly hydrated βpGlc, the bound water molecules are also located above the ring rather than around its edge, directed by the "perpendicular" G−g+ orientation of the hydroxymethyl group. Thus, despite the equatorial orientation of all hydrophilic groups in both βpGlc and βpGlcNAc, it is in fact, the "axial" α and β faces of the carbohydrates that are the eventual sites of hydration.

When neither an acetamido nor a hydroxymethyl group is present, for example, in xylose or fucose, regioselectivity can still be provided by the pre-existing counter-clockwise $(OH\text{-}O)_n$ chain. Bound water molecules located further round the pyranose ring select peripheral sites associated with the weakest link in the chain,[9] between OH3 and OH2 in the α-anomers and OH2 and O1 in the β-anomers. Thus, in general the degree of control exerted by the groups around the pyranosyl ring lies in the order acetamido > hydroxymethyl > pyranosyl $(OH\text{-}O)_n$ chain. The striking and unprecedented role of the C-2 acetamido group in controlling the hydrated structures determined here may, in part, explain the differing and widespread roles of GlcNAc and perhaps GalNAc, in nature.

Experimental Section

Spectroscopy: A detailed description of the molecular beam experimental strategy has been published previously.[9] The carbohydrate samples were mixed with graphite powder, vaporized through laser ablation at 1064 nm (pulse energy ≈4 mJ) in front of the nozzle (0.8 mm orifice) of a pulsed valve (Jordan Valve), and seeded into a free jet argon expansion (≈3–5 bar Ar). Hydrated complexes were formed by seeding the argon with water vapor (>0.25%) allowing the investigation of their size-selected structures. The free jet expansion passed through a 2 mm skimmer to form a collimated molecular beam, crossed by one, or two tuneable laser beams in the extraction region of a linear time-of-flight mass spectrometer (Jordan). Mass-selected resonant two photon ionization (R2PI) spectra were recorded using the frequency-doubled output of a pulsed Nd/YAG-pumped dye laser (Continuum Surelite III/Sirah PS-G, 1.5 mJ per pulse UV). Conformer-specific spectra were obtained through UV hole-burning (UV-HB) and IR ion dip experiments, employing UV/UV and IR/UV double resonance spectroscopy. The UV-HB experiments employed the frequency-doubled output of an excimer-pumped dye laser (Lambda-Physik EMG 201/Lambda-Physik FL3002, 0.5 mJ UV). IR experiments employed radiation in the range 3000–3800 cm^{-1}, generated by difference frequency mixing of the fundamental of a Nd/YAG laser with the output of a dye laser in a $LiNbO_3$ crystal (Continuum Powerlite 8010/ND6000/IRP module) or by a tunable OPO/OPA laser system (LaserVision); all laser pulses were ≈10 ns duration. The delay between the pump and the probe laser pulses was ≈150 ns in both the IR ion dip and UV-HB experiments.

Computation and assignment: Conformational and structural assignments of the experimental IRID spectra were performed through comparison with calculations using a combination of molecular mechanics, ab initio and density functional theory (DFT) methods. Initial structures were generated by an extensive molecular mechanics conformational search using the Monte Carlo multiple minimization procedure as implemented in the

MacroModel software (MacroModel v.8.5, Schrödinger, LLC21).[25] Relevant structures, filtered on the basis of their relative energies, information gained from previous studies on carbohydrates and their experimental vibrational signatures, were selected for geometry optimizations using the HF/6-31+G* method, as implemented in the Gaussian 03 program package.[26] The most stable conformers were re-optimized at the B3LYP/6-31+G* level for the unhydrated sugars and 6-311+G* for the hydrated ones. Their relative energies were subsequently calculated more accurately, at the MP2/6-311++G** level of theory, to take proper account of dispersion interactions. Zero-point and free energy corrections were performed using the harmonic frequency calculations performed on the B3LYP structures; for comparison with the experiments, the frequencies predicted for the O-H and N-H modes were scaled by factors of 0.9734 and 0.96, respectively.

Acknowledgements

Financial support provided by the Engineering and Physical Sciences Research Council (EPSRC), the Leverhulme Trust (grant F/08788G) and the Spanish Ministry of Education and Science (EJC), is gratefully acknowledged. We are also grateful for additional technical support provided by the Oxford Supercomputing Centre, the STFC Laser Loan Pool and the Physical and Theoretical Chemistry Laboratory.

[1] R. J. Linhardt, T. Toida, *Acc. Chem. Res.* **2004**, *37*, 431–438.
[2] N. S. Gandhi, R. L. Mancera, *Chem. Biol. Drug Des.* **2008**, *72*, 455–482.
[3] A. Kovács, B. Nyerges, V. Isvekov, *J. Phys. Chem. B* **2008**, *112*, 5728–5735.
[4] M. Mobli, A. Almond, *Org. Biomol. Chem.* **2007**, *5*, 2243–2251.
[5] a) F. Yu, J. H. Prestegard, *Biophys. J.* **2006**, *91*, 1952–1959; b) A. Fernández-Tejada, F. Corzana, J. H. Busto, G. Jiménez-Orsés, J. Jiménez-Barbero, A. Avenoza, J. M. Peregina, *Chem. Eur. J.* **2009**, *15*, 7297–7301.
[6] V. Kräutler, M. Müller, P. H. Hünenberger, *Carbohydr. Res.* **2007**, *342*, 2097–2124.
[7] F. Corzana, J. H. Busto, G. Jimenez-Oses, J. L. Asensio, J. Jimenez-Barbero, J. M. Peregrina, A. Avenoza, *J. Am. Chem. Soc.* **2006**, *128*, 14640–14648.
[8] K. N. Kirschner, R. J. Woods, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 10541–10545.
[9] E. J. Cocinero, E. C. Stanca-Kaposta, E. M. Scanlan, D. P. Gamblin, B. G. Davis, J. P. Simons, *Chem. Eur. J.* **2008**, *14*, 8947–8955; E. J. Cocinero, D. P. Gamblin, B. G. Davis, J. P. Simons, *J. Am. Chem. Soc.* **2009**, *131*, 11117–11123.
[10] J. P. Simons, R. A. Jockusch, P. Çarçabal, I. Hünig, R. T. Kroemer, N. A. Macleod and L. C. Snoek, *Int. Rev. Phys. Chem.* **2005**, *24*, 489–532.
[11] E. C. Stanca-Kaposta, D. P. Gamblin, E. J. Cocinero, J. Frey, R. T. Kroemer, B. G. Davis, A. J. Fairbanks, J. P. Simons, *J. Am. Chem. Soc.* **2008**, *130*, 10691–10696.
[12] O. Coskuner, *J. Chem. Phys.* **2007**, *127*, 015101.
[13] L. Hemmingsen, D. E. Madsen, A. L. Esbensen, L. Olsen, S. B. Engelsen, *Carbohydr. Res.* **2004**, *339*, 937–948.
[14] J. L. Dashnau, K. A. Sharp, J. M. Vanderkooi, *J. Phys. Chem. B* **2005**, *109*, 24152–24159.
[15] R. A. Dwek, *Chem. Rev.* **1996**, *96*, 683–720.
[16] H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, H. Rüdiger, *ChemBioChem* **2004**, *5*, 740–764.
[17] C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357–2364.
[18] A. B. Boraston, D. N. Bolam, H. J. Gilbert, G. J. Davies, *Biochem. J.* **2004**, *382*, 769–781.
[19] Z. Su, E. C. Stanca-Kaposta, E. J. Cocinero, B. G. Davis, J. P. Simons, *Chem. Phys. Lett.* **2009**, *471*, 17–22.
[20] R. K. Raju, A. Ramraj, M. A. Vincent, I. H. Hillier, N. A. Burton, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6500–6508.
[21] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373–380.
[22] S. M. Tschampel, R. J. Woods, *J. Phys. Chem. A* **2003**, *107*, 9175–9181.
[23] C. Clarke, R. J. Woods, J. Gluska, A. Cooper, M. A. Nutley, G.-J. Boons, *J. Am. Chem. Soc.* **2001**, *123*, 12238–12247.
[24] P. Çarçabal, R. A. Jockusch, I. Hünig, L. C. Snoek, R. T. Kroemer, B. G. Davis, D. P. Gamblin, I. Compagnon, J. Oomens, J. P. Simons, *J. Am. Chem. Soc.* **2005**, *127*, 11414–11425.
[25] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrikson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.
[26] Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, Q. C. J. V. Ortiz, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, J. A. Pople, Gaussian, Inc., Wallingford CT, **2004**.

Received: July 2, 2009
Revised: August 10, 2009
Published online: November 9, 2009

