RESEARCH PAPER

PCCP
www.rsc.org/pccp

Adding water to sugar: A spectroscopic and computational study of α- and β-phenylxyloside in the gas phase

Isabel Hünig,[a] Alexander J. Painter,[a] Rebecca A. Jockusch,[a] Pierre Çarçabal,[a] Elaine M. Marzluff,[b] Lavina C. Snoek,[a] David P. Gamblin,[c] Benjamin G. Davis[c] and John P. Simons*[a]

[a] *Department of Chemistry, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford, UK OX1 3QZ. E-mail: john.simons@chem.ox.ac.uk*
[b] *Department of Chemistry, Grinnell College, Grinnell, IA, USA*
[c] *Department of Chemistry, Chemistry Research Laboratory, Mansfield Road, Oxford, UK OX1 3TA*

Received 23rd March 2005, Accepted 27th April 2005
First published as an Advance Article on the web 17th May 2005

The gas phase structures of phenyl α- and β-D-xylopyranoside (α- and β-pXyl) and their mono-hydrates have been investigated using a combination of resonant two-photon ionization (R2PI), ultra-violet hole-burning and resonant infrared ion dip spectroscopy, coupled with density functional theory (DFT) and *ab initio* computation. The hole-burning experiments indicate the population of a single conformer only, in each of the two anomers. Their experimental and calculated infrared spectra are both consistent with a conformational assignment corresponding to the computed global minimum configuration. All three OH groups are oriented towards the oxygen atom (O1) on the anomeric carbon atom to form an all *trans* (*ttt*) counter-clockwise chain of hydrogen bonds. The mono-hydrates, α- and β-pXyl(H_2O) each populate two distinct structures in the molecular beam environment, with the water molecule inserted between OH4 and OH3 or between OH3 and OH2 in α-pXyl(H_2O), and between OH2 and O1 in either of two alternative orientations, in β-pXyl(H_2O). In all of the mono-hydrated xyloside complexes, the water molecule inserts into the weakest link of the sugar molecules' hydrogen-bonded chain of hydroxy groups, creating a single extended chain, strengthened by co-operativity. The all-*trans* configuration of the xylose moiety is retained and the mono-hydrate structures correspond to those calculated to lie at the lowest relative energies.

1. Introduction

The biological diversity and functional importance of carbohydrates in biological systems is immense.[1] One of the most complex and interesting processes in which they are involved is the specific recognition of *N*-linked glycans by carbohydrate-binding proteins.[2,3] The role of water molecules in carbohydrate–protein binding, or indeed, in ligand–protein-binding, is not yet well understood. However, the importance of the specifics of such interactions is suggested by the presence of structurally conserved water between carbohydrate ligands and their protein binding partners.[4,5] To elucidate the role of specific water molecules, an effective approach is to first eliminate bound water molecules altogether to provide a structural baseline against which the consequences of their subsequent *controlled* replacement can be explored. This enables direct investigation of water–carbohydrate binding and also of the influence of bound water molecules on the carbohydrate structure and conformation. Their structural preferences can be identified through a combination of mass, and conformationally selective ultra-violet and infrared spectroscopy conducted in the gas phase (in the low temperature environment of a free jet expansion), compared with the results of DFT and *ab initio* calculations. This strategy allows conformational and structural assignments to be made both for the isolated carbohydrate molecules free of any environmental effects and for carbohydrate–water complexes with a defined number of water molecules, in contrast to NMR studies conducted in the condensed phase or X-ray diffraction studies conducted in the crystalline state.[6–9] The data we provide can be used to improve the parameterization of force-fields which are known to be unreliable for the determination of carbohydrate properties.[10]

The strategy, while still taking its first steps,[11–14] is already beginning to generate a set of 'structural working rules' and also contribute to an understanding of the physical processes underlying protein–carbohydrate molecular recognition.[11] The structural working rules have been based upon the identified structural preferences of a series of derivatived mono- and disaccharides and their mono-hydrated complexes, including phenyl-β-D-glucoside,[11,13,15] phenyl-β-D-galactoside,[11,12] phenyl-α-D-mannoside,[11] and the disaccharide benzyl-β-lactoside.[11,14]† Their structures all display a common pattern in which the hydroxy groups form a hydrogen-bonded chain around the molecule, oriented either counter-clockwise, with an all-*trans* '*ttt*' configuration, or clockwise, with an all-*gauche* '*ggg*' configuration (see Fig. 1(a)). When the sugar is hydrated, the extra molecule is inserted into the weakest link(s) in the hydrogen-bonded chain, to generate an extended and co-operatively bonded chain incorporating both the sugar and the bound water molecule. Comparisons with crystallographic data suggest that nature also appears to exploit similar facets of globally co-operative hydrogen bond recognition in its ability to bind simple monosaccharide ligands (see ref. 11 and references therein).

These investigations are widened in the present study through an exploration of the conformations of phenyl

† The phenyl or benzyl moiety, joined through a glycosidic linkage to the carbohydrate provides a chromophore, compatible with the UV excitation scheme, which has little effect on the carbohydrates' conformational preferences in these systems.[15]

DOI: 10.1039/b504230e

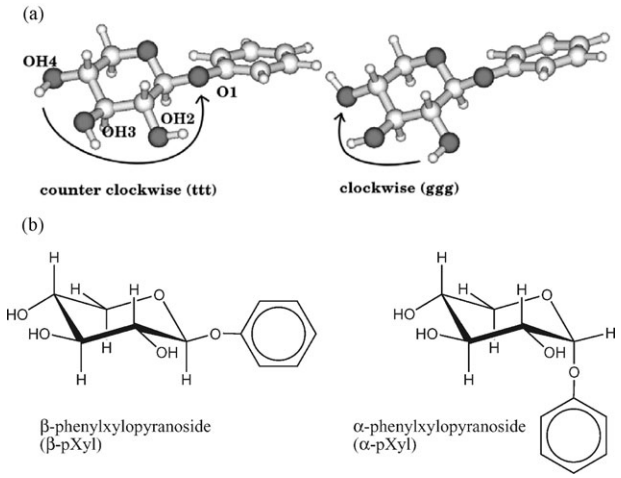

Fig. 1 (a) β-pXyl with counter-clockwise and clockwise orientations of the hydroxy groups. (b) Schematic representation of α-pXyl and β-pXyl in the most stable chair conformation ($^4C_1$).

α- and β-D-xylopyranoside (α/β-pXyl, see Fig. 1(b)) and their singly hydrated complexes (α/β-pXyl(H_2O)). D-Xylose is amongst the approximately ten monosaccharide units that dominate the glycobiology of higher animals. It is a characteristic building block in *N*-linked glycans in plants and invertebrates and is of immunological importance, occurring in binding sites of plant epitopes.[16,17] Guler *et al.* have studied the gas phase structure of D-xylose (along with four other five-carbon monosaccharides) by means of mass spectrometry and DFT (B3LYP/6-311++G(d,p)) calculations.[18] Their study focused on the question of anomeric stability in the gas phase; they found the anomeric form to be preserved following thermal desorption and laser-induced acoustic desorption. The anomeric equilibrium of D-xylose in aqueous solution has also been investigated recently using molecular dynamics (MD) calculations;[19,20] the calculations indicated preferential solvation of the anomeric hydroxy group of the β-anomer over that in the α-anomer, in conformity with the experimentally observed (^{1}NMR) population ratio, $\beta : \alpha \approx 2 : 1$.[20,21]

2. Experimental and theoretical methods

α- and β-pXyl were synthesized through $BF_3 \cdot Et_2O$-catalyzed glycosylation using the corresponding peracetate as a donor.[22] Zemplén deprotection[23] of the two peracetates proceeded to afford α-pXyl and β-pXyl in yields of 59–72% from D-xylose. Both anomers were found pure by NMR, matching characteristic literature values.[24]

A detailed description of the molecular beam apparatus has been published previously.[25] α- and β-pXyl were vaporized in an oven (~130 °C) mounted in front of the nozzle (0.5 mm orifice) of a pulsed valve (General Valve) in the vacuum and seeded into an argon expansion (~3–5 bar Ar). In order to form water complexes, the argon line was saturated with water vapor and the oven temperature was increased by 10–30 °C. The free jet expansion passed through a 1 mm skimmer to form a molecular beam, which was subsequently crossed by one, or two tunable laser beams in the extraction region of a linear time-of-flight mass spectrometer (Jordan). Resonant two photon ionization (R2PI) spectra were recorded using the frequency-doubled output of a pulsed Nd:YAG-pumped dye laser (Spectra-Physics GCR 170 at 355 nm/LAS 205, operating at 10 Hz). Conformer-specific spectra were obtained through 'hole-burning' (UV-HB or IR-HB) experiments, employing UV-UV and IR-UV double resonance spectroscopy. The UV-UV experiments (UV-HB) employed the frequency-doubled output of an excimer-pumped dye laser (Lambda-Physik EMG 201/Lambda-Physik FL3002). IR experiments employed laser radiation in the region between 3000 and 3700 cm^{-1} generated by difference frequency mixing of the fundamental of a Nd:YAG with the output of a dye laser in a $LiNbO_3$ crystal (Continuum Powerlite 8010, ND6000, IRP module). The resolution of the resulting IR laser beam was 0.5 cm^{-1}. The delay between the pump and the probe laser pulses in the IR-HB and UV-HB experiments, both with a pulse duration of ~10 ns, was 150–400 ns. The output energy of the two lasers averaged ~0.5 mJ $pulse^{-1}$ (UV) and ~2 mJ $pulse^{-1}$ (IR), respectively.

Starting structures for *ab initio* calculations of the α- and β-pXyl and their hydrated complexes, were generated by performing molecular mechanics conformational searches using the Monte Carlo multiple minimization (MCMM) method as implemented in the MacroModel software (MacroModel v.8.5, Schrödinger, LLC[26]). 10 000 structures within 50 kJ mol^{-1} were generated for each anomer using the force-field MMFFs; other force-fields (MM3, Amber*, OpalsAA) were also used but they did not generate any additional structures. Eleven of the α-pXyl, twenty-four of the β-pXyl, eleven of the α-pXyl(H_2O) and eighteen of the β-pXyl(H_2O) conformer structures, generated by the MCMM within 35 kJ mol^{-1} of the global minima, were selected for density functional theory geometry optimization (B3LYP/6-31+G(d)) using the Gaussian 03 software package.[27] The selection of conformers for the unhydrated anomers included boat and chair ($^4C_1$ and $^1C_4$) structures. Nine different conformers of α-pXyl within 42 kJ mol^{-1} of the global minimum and eighteen of β-pXyl within 32 kJ mol^{-1}, remained distinguishable after the calculations converged. Within the sugar moiety the lowest energy structures were virtually the same as those reported by Guler *et al.*[18] for the underivatised D-xylose computed at the B3LYP/6-311++G(d,p) level of theory. After DFT geometry optimization, eleven of the α-pXyl(H_2O) structures, all lying within 17 kJ mol^{-1} of the global minimum, and eighteen of the β-pXyl(H_2O) structures, lying within 25 kJ mol^{-1}, remained distinguishable. Harmonic vibrational frequencies were also calculated at the B3LYP/6-31+G(d) level of theory. Because DFT/B3LYP calculations do not take dispersion interactions into account the energies obtained are known to be unreliable.[28,29] Therefore, the relative energies used to find the global minimum structures were obtained from single point MP2/6-311++G(d,p) calculations, corrected for the zero-point energy (ZPE) derived from the B3LYP/6-31+G(d) frequency calculations. The calculated harmonic frequencies were scaled by a factor of 0.9734, a factor known to provide good agreement with experimental data for carbohydrates in the OH stretching region.[12,15]

3. Results and discussion

3.1 α- and β-p-xyloside monomers

Fig. 2 shows the R2PI and UV-HB spectra of (a) α-pXyl and (b) β-pXyl in the range 36 810 to 37 100 cm^{-1}. The two spectra are very similar and their electronic origins (located at 36 888 cm^{-1} (α) and 36 894 cm^{-1} (β)) are separated by only 6 cm^{-1}, suggesting very similar environments for the (phenyl) chromophore. The UV-HB spectra of the two anomers, α- and β-pXyl, displayed below the respective R2PI spectra in Figs. 2a and b reveal the presence of only one significantly populated conformer in each of the xyloside anomers detected in the molecular beam. Their UV spectra display an intense and sharp origin, followed by a much weaker vibronic band at about +30 cm^{-1} and a stronger band at about +45 cm^{-1}. (A very similar pattern has also been found in the individually resolved conformers of phenyl-α-D-mannopyranoside, phenyl-β-D-glucopyranoside and phenyl-β-D-galactopyranoside.[11]) On close examination, despite the anomeric purity indicated in the NMR spectra, the two R2PI spectra also reveal some very

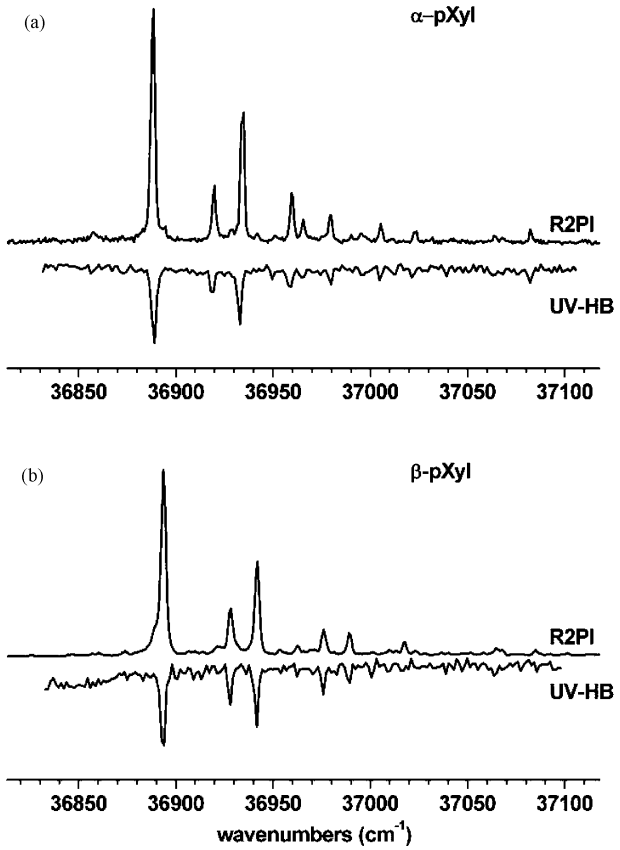

Fig. 2 R2PI and UV-HB spectra of α-pXyl (a) and β-pXyl (b). The UV-HB spectra were recorded with the probe laser tuned to the respective electronic origins: 36 888 cm^{-1} for α-pXyl and 36 894 cm^{-1} for β-pXyl.

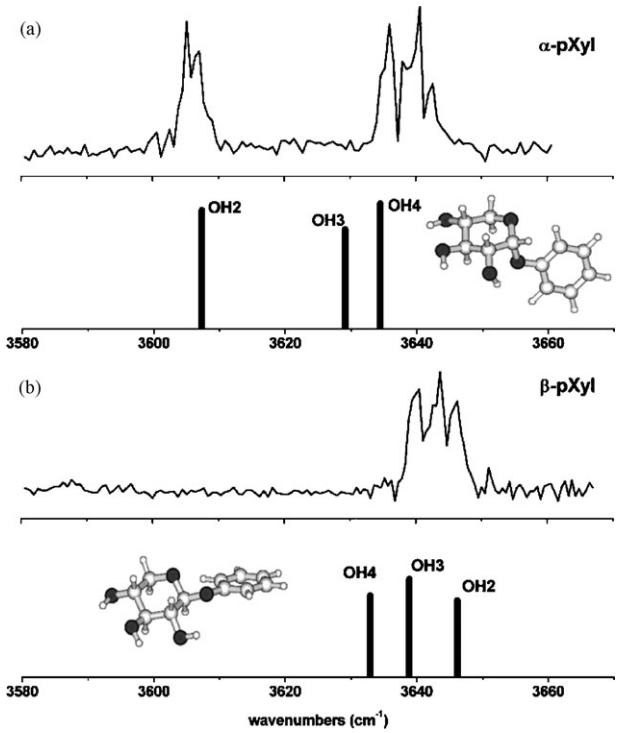

Fig. 3 Experimental IR-HB spectra of α-pXyl (a) and β-pXyl (b). The calculated spectra of the lowest energy conformers to which they are assigned, are shown for comparison. The structures were geometry optimized and frequencies calculated at the B3LYP/6-31+G(d) level of theory. The calculated frequencies are scaled by 0.9734.

weak features indicating cross-contamination of the β-pXyl and α-pXyl anomers which most likely happened after synthesis during the sample preparation. This is most apparent in the origin line of β-pXyl which displays a shoulder on the low-frequency side, located exactly at the origin of the α-anomer.

The IR-HB spectra of the two xyloside anomers, recorded with the UV probe laser centered on their respective electronic origin bands, are shown in Fig. 3, together with the calculated IR spectra of their respective lowest energy conformers and their computed structures. The experimental spectrum of α-pXyl is unexpectedly complex, presenting twice the number of features anticipated, given the presence of three OH groups in the xylose moiety. Their average location is in good agreement with the spectrum predicted for the lowest energy, all-*trans* α-anomer. Both the measured and calculated IR spectra show the same pattern: the separation into two, weakly H-bonded OH bands (labelled OH3 and OH4) located around 3640 cm^{-1} and the more strongly shifted OH2 band, located around 3605 cm^{-1}, and associated with the stronger, OH2–O1 interaction (see Table 1). However, each of these three features appears to be split into doublet, separated by *ca.* 3–4 cm^{-1}. At first sight, the IR spectrum of the β-anomer appears not to show split lines, displaying three closely spaced bands located at 3640–3646 cm^{-1} and in good agreement with the spectrum calculated for the lowest energy, all-*trans* conformer of the β-anomer, where the interactions OH4–OH3, OH3–OH2 and OH2–O1 are all comparable and weak (see Table 1). However, close inspection of the experimental spectrum reveals an underlying, partially resolved doublet splitting (2 cm^{-1}), reflected in the asymmetry of the individual band contours. A number of possible explanations for the unexpected spectral complexity might be entertained. One might be the presence of two near equally populated conformers with structures so similar that their UV band origins overlapped—a very unlikely possibility given the 'text-book' quality of the UV hole-burn spectra shown in Fig. 2. Admittedly, two low-lying conformers are predicted for the β-anomer (differing only in the orientation of the phenyl chromophor) but the two lowest-lying conformers of the α-anomer are separated by *ca.* 12 kJ mol^{-1}. Given the long delay (>150 ns) between the IR and UV laser pulses we can rule out the absorption from molecules in the S(1) state. An alternative possibility might be the presence of an Ar–xyloside van der Waals dimer, but that would require a virtual zero displacement of its band origin compared to the xyloside monomers and the absence of any associated low frequency vibronic progressions—again, not very likely. The splitting is much more likely to have an intramolecular origin, for example a contribution from a combination band (though the very small splittings argue strongly against this), or a tunnelling process *e.g.* of a ring bending motion or an OH group interacting with one or the other lone pair of an adjacent oxygen atom. For the moment, the explanation remains an open question which will not be pursued further. The general features of the IR spectra are in good agreement with the conformational assignments proposed. They both exhibit an all-*trans* chain of hydrogen-bonded OH groups. The notable difference between the spacing of the IR bands in the spectra of α-pXyl and β-pXyl reflects the different geometry of the anomeric group (see Table 1 for OH···O distances and angles). The OH2–O1 separation in β-pXyl is 2.54 Å but in α-pXyl it is significantly shorter, 2.25 Å, and somewhat more linear (110° compared to 100°). This creates a stronger hydrogen bond which shifts the OH-stretching vibration by 41 cm^{-1} towards lower wavenumbers. The distances between the hydrogen atom of the OH3(OH4) and the oxygen atom of the OH2(OH3) on the other hand are considerably greater, varying between 2.46 and 2.49 Å.

3.2 α- and β-pXyl(H_2O) complexes

The R2PI and UV-HB spectra of the water complexes of α-pXyl and β-pXyl are shown in Fig. 4. The R2PI spectrum of α-pXyl(H_2O), in contrast to the unhydrated anomer,

Table 1 OH···O distances and angles in the lowest energy conformers of α- and β-pXyl and their water complexes as revealed by B3LYP/6-31+G(d) calculations. All distances are given in Å, the angles in degrees

Species	Distances/Å (OH···O angles/°)				
	OH4···O3	OH3···O2	OH2···O1	OH4/3/2···OH2	HOH···O3/2/1
α-pXyl	2.45 (106)	2.49 (106)	2.25 (110)	—	—
β-pXyl	2.46 (105)	2.48 (103)	2.54 (100)	—	—
α-pXyl(H_2O)^{I}	2.36 (111)	—	2.13 (113)	1.85 (169)	1.87 (146)
α-pXyl(H_2O)^{II}	—	2.35 (111)	2.26 (110)	1.86 (161)	1.88 (152)
α-pXyl(H_2O)^{III}	2.43 (107)	2.41 (105)	—	1.88 (170)	—
α-pXyl(H_2O)^{IV}	2.46 (106)	2.43 (108)	—	1.88 (163)	—
β-pXyl(H_2O)^{I}	2.45 (106)	2.40 (106)	—	1.86 (160)	1.94 (149)
β-pXyl(H_2O)^{II}	2.45 (106)	2.40 (107)	—	1.85 (161)	1.95 (149)
β-pXyl(H_2O)^{III}	2.37 (108)	—	2.38 (105)	1.85 (161)	1.90 (151)
β-pXyl(H_2O)^{IV}	—	2.33 (108)	2.53 (100)	1.86 (162)	1.88 (152)

presents two, near equally intense peaks in the low wavenumber region of the spectrum. UV-HB spectroscopy reveals these as the origin bands of two different hydrated structures present in the molecular beam. Assuming that both have the same transition dipole moment and absorption cross section (since they both have the same chromophore) their abundances are almost equal. Their UV band origins (at 36 885 cm^{-1}, structure **a** and 36 898 cm^{-1}, structure **b**) are only separated by 13 cm^{-1} and lie very close to the origin of the unhydrated anomer (36 888 cm^{-1}); in these complexes, the phenoxy chromophore is not significantly perturbed by the addition of water.

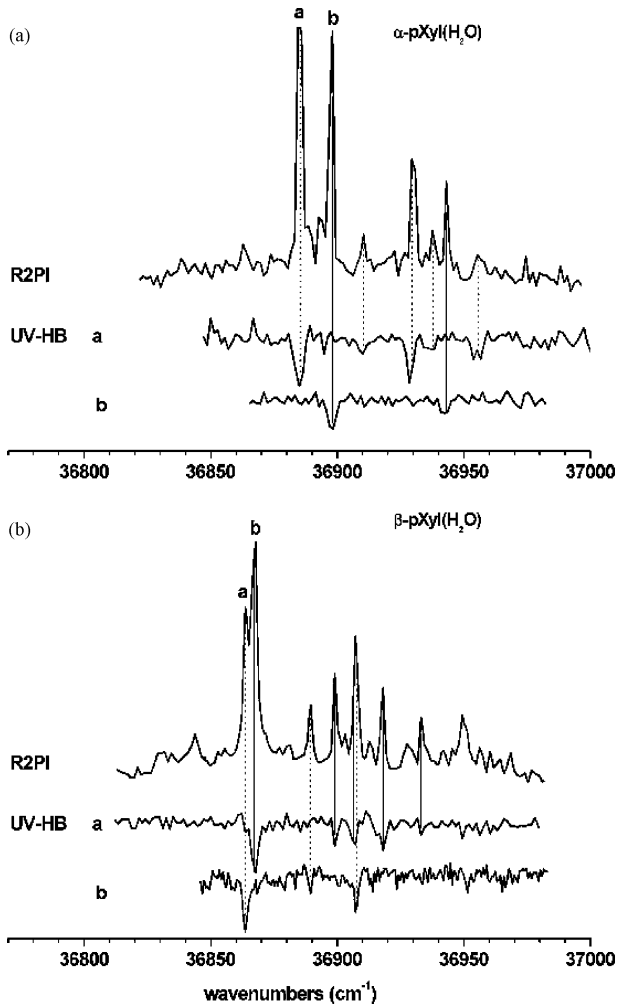

Fig. 4 R2PI and UV-HB spectra of α-pXyl(H_2O) (a) and β-pXyl (H_2O) (b). The UV-HB spectra of structures **a** and **b** were recorded with the probe laser tuned to the transitions labelled as **a** and **b**, respectively, in the R2PI spectra: 36 885 cm^{-1} (36 898 cm^{-1}) for α-pXyl(H_2O) and 36 864 cm^{-1} (36 867 cm^{-1}) for β-pXyl(H_2O).

The R2PI and UV-HB spectra of the mono-hydrated β-anomer, shown in Fig. 4b, also indicate the population of two distinct structures of β-pXyl(H_2O) in the molecular beam. Their UV band origins, located at 36 864 cm^{-1}, **a**, and 36 867 cm^{-1}, **b**, are shifted by only 3 cm^{-1} with respect to each other, but are shifted much more strongly to lower wavenumbers, by *ca.* 30 cm^{-1} with respect to the unhydrated anomer (36 894 cm^{-1}). Binding a water molecule to the β-anomer perturbs the phenoxy chromophore of both structures and both are equally perturbed, suggesting a similar interaction of the water molecule with the O1 atom in each case.

Fig. 5 shows the IR-HB spectra of conformers **a** and **b** of α-pXyl(H_2O). As expected, the IR spectra of the two singly hydrated structures present five transitions in the investigated spectral range, covering the expected five OH stretching vibrations of a singly hydrated xyloside complex. The spectrum of conformer **b** exhibits a small additional peak at 3540 cm^{-1}, which most probably stems from a weak contribution from conformer **a** underlying the UV spectrum associated with conformer **b**. Comparison of the two measured mid-IR spectra with those calculated for the four lowest energy structures, α-pXyl(H_2O) (I, II, III and IV) reveals close agreement with prediction for the two lowest energy 'insertion' structures, I and II, both with respect to the pattern of vibrational wavenumbers and their relative intensities. The experimental and the calculated spectra of I and II both display two strong (and characteristic) transitions in the low wavenumber range associated with strongly hydrogen-bound OH stretching vibrations; two weak transitions at higher wavenumber in the region of weakly hydrogen-bound OH vibrations; and one medium intensity transition at high wavenumber associated with a free OH stretching mode. The calculated spectra associated with the additional structures, III and IV, present a quite different pattern, displaying only one strong transition at low wavenumber, and three transitions in the intermediate range. The monohydrate spectra are assigned to the structures I and II.

The two bands at low wavenumber arise through coupling of the symmetric stretching mode of the bonded water molecule and the OH mode associated with the sugar hydroxy group which acts as the hydrogen bond donor to the water molecule, (H_2O_s ± OH3 in I; H_2O_s ± OH4 in II). These relatively strongly hydrogen-bonded modes are generally not too well reproduced by the DFT method employed here (using the same scaling factor for all bands). Clearly, it is more difficult to calculate accurate geometries and frequencies for molecular complexes and clusters than for covalently bound systems. Nevertheless, frequencies at this level can be used for conformational assignment *via* pattern matching.

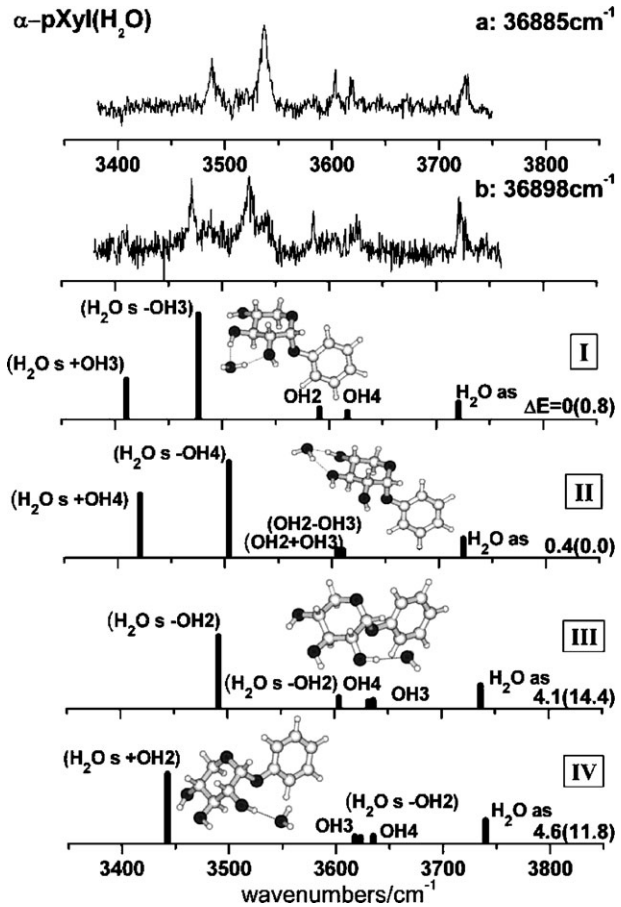

Fig. 5 The IR-HB spectra of the complexes **a** and **b** of the α-pXyl(H_2O), recorded with the analyzing laser at 36 885 cm^{-1} (a) and 36 898 cm^{-1} (b). The spectra of the four lowest energy hydrated structures calculated at the B3LYP/6-31+G(d) level are shown for comparison. Relative energies (in kJ mol^{-1}) from MP2/6-311++G(d,p)//B3LYP/6-31+G(d) calculations are given in the figure. B3LYP/6-31+G(d) energies are shown in parentheses. The spectra associated with **a** (**b**) are assigned to the structures II (I). All frequencies were scaled by 0.9734.

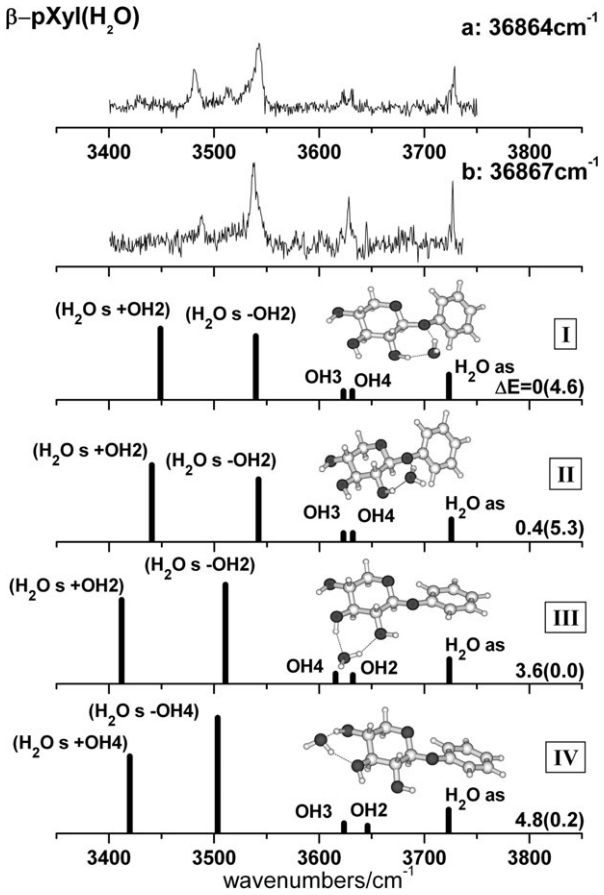

Fig. 6 The IR-HB spectra of the complexes **a** and **b** of the β-pXyl(H_2O), recorded with the analyzing laser at 36 864 cm^{-1} (a) and 36 867 cm^{-1} (b). The spectra of the four lowest energy hydrated structures calculated at the B3LYP/6-31+G(d) level are shown for comparison. Relative energies (in kJ mol^{-1}) from MP2/6-311++G(d,p)//B3LYP/6-31+G(d) calculations are given in the figure. B3LYP/6-31+G(d) energies are shown in parentheses. The spectra associated with **a** and **b** are assigned to one (or other) of the calculated structures, I and II; their close similarity prevents a more specific assignment. All frequencies were scaled by 0.9734.

The experimental IR spectra, **a** and **b**, differ significantly in the intermediate wavenumber range, ($\sim$3600 cm^{-1}) associated with the two, weakly hydrogen-bonded vibrations. The bands are very closely spaced in the spectrum associated with structure **a** (16 cm^{-1}) but are more widely separated in the spectrum associated with structure **b** (40 cm^{-1}). Their calculated spacings follow a similar trend: the scaled wavenumbers of the bands associated with OH2 and OH4 in the structure I are separated by 28 cm^{-1} but those associated with the coupled vibrations, OH2–OH3 and OH2 + OH3 in structure II are separated by only 5 cm^{-1}. Accordingly, the two spectra, **a** and **b** are assigned to the structures II and I, respectively. They are calculated to have almost the same energies (structure II is only 0.4 kJ mol^{-1} higher in energy than the global minimum structure I) and experimentally, they have almost equal populations. The undetected structures III and IV lie *ca.* 4 kJ mol^{-1} higher in energy. Both populated structures retain the all-*trans* conformation of the unhydrated α-pXyl sugar. In the hydrated structure **a**(II), the water molecule is inserted between OH4 and OH3 to form a co-operative chain of hydrogen bonds, $OH4\cdots OH^{water}\cdots OH3\cdots OH2\cdots O1$; in the alternative structure, **b**(I) the water molecule is inserted between OH3 and OH2 to create a similar chain. In unhydrated α-pXyl the hydrogen bonds between OH4 and OH3 and between OH3 and OH2 are both (equally) weak and significantly weaker than the hydrogen bond between OH2 and O1 (see Table 1 for OH$\cdots$O distances and angles). Therefore, in both structures the inserted water molecule creates a similar extension of the hydrogen-bonded chain, by introducing two strong hydrogen bonds bridging its weakest link.

The two measured IR-HB spectra of β-pXyl(H_2O) conformers and the calculated spectra of the four lowest energy computed structures are shown in Fig. 6.‡ All four incorporate the same all-*trans* configuration of unhydrated β-pXyl with the water molecule inserted into the hydrogen-bonded chain of hydroxy groups. The global minimum structure, I, is calculated to lie slightly below structure II, by 0.4 kJ mol^{-1}; the other two structures, III and IV are located at 3.6 and 4.8 kJ mol^{-1}, respectively. The two experimental spectra are very similar to each other and also resemble all four calculated spectra, displaying two strongly hydrogen bound OH stretch vibrations, two weakly hydrogen bound OH modes and the free OH stretch. Despite their close similarity, the experimental observed spectra are assigned to the structures I and II for the following reasons:

(1) The UV-HB spectra of the two β-pXyl(H_2O) complexes are very similar to each other and their electronic origins are

‡ Calculations indicate that two different minima of the structures with the water molecule inserted between OH4(3) and OH3(2) (insert 4 (insert 3)) exist. In one of them, the free hydrogen atom is pointing up, in the other down. The minima have virtually the same energy with a relatively small barrier that could be overcome in the cooling process. For the insert 2 structures, the corresponding barrier was found to be significantly higher.

separated by only 3 cm^{-1}, while both display a relatively large shift, $\sim$30 cm^{-1}, with respect to the electronic origin of the unhydrated β-pXyl. This suggests an equally strong perturbation of the phenoxy chromophore in the two hydrates of β-pXyl(H_2O), in agreement with expectation for the two alternative OH2–O1 insertion structures, I and II. In the two OH2–OH3/OH3–OH4 insertion structures, III and IV, the chromophore would remain unperturbed by the addition of water as *e.g.* in α-pXyl.

(2) The structures I and II are almost isoenergetic and they lie significantly lower in energy than the structures III and IV.

(3) In each of the two experimental IR-spectra the two strongly hydrogen-bound OH stretches occur at nearly the same wavenumbers (3481/3537 cm^{-1} for **a**; 3488/3537 cm^{-1} for (**b**), which is reflected in the calculated spectra in the pairs of structures I and II (3448/3540 cm^{-1} for I; 3441/3541 cm^{-1} for II) and III and IV (3411/3511 cm^{-1} for III; 3418/3504 cm^{-1} for IV).

The anomeric equilibrium of D-xylose in aqueous solution is weighted in favor of the β-anomer (at 20 °C, 65% of D-xylose correspond to β-D-xylose and only 35% to α-D-xylose[19]) while calculations of isolated α- and β-pXyl at zero Kelvin (MP2/6-311++G(d,p)//B3LYP/6-31+G(d)) favor the α-anomer by 8.5 kJ mol^{-1}. The addition of water lessens the stability difference to 3.6 kJ mol^{-1}. This trend, which is conserved but less pronounced when temperature effects are taken into account, is easily understood when considering that the water molecule is always inserted into the weakest hydrogen bond, which in the β-anomer is significantly weaker than in the α-anomer. In β-pXyl(H_2O) the water insertion involves the anomeric group whereas for the α-anomer it does not. This could be the origin of the better solvation of the anomeric hydroxy group in the β-anomer in aqueous solution found by Schmidt *et al.*[19] and Höög and Widmalm[20] with MD calculations which they attribute to the preference for the β-anomer in aqueous solution.

4. Conclusion

The structures found for α-pXyl and β-pXyl and their hydrated complexes in this investigation follow the common pattern of conformational structures observed in all the phenyl-tagged saccharides studied so far. In general, the hydroxy groups located around the pyranoside ring can form a chain of rather weak hydrogen bonds, which can be oriented in a *gauche* or *trans*, clockwise or counter-clockwise configuration, and re-oriented (or not) by hydration.[11–13,15] The two xyloside anomers, and their singly hydrated water complexes all display the same all-*trans*, counter-clockwise orientation of their hydroxy groups. In a recent investigation[11] we noted that the origin bands of the S(1)–S(0) electronic spectra of all the phenyl-tagged saccharides studied to date, lie at wavenumbers $\geq$ 36 800 cm^{-1} when the hydroxy chain is oriented counter-clockwise, and below 36 800 cm^{-1} when the chain is oriented clockwise. In the former case an OH2–O1 (with the phenoxy chromophore) interaction is present while in the latter it is not. The location of the electronic origins of both the isolated and the hydrated anomers of xylose at wavenumbers $>$ 36 800 cm^{-1} lends further support to the rule.

The water molecule in α-pXyl(H_2O) inserts between the OH4 and OH3 groups or between OH3 and OH2. The observed, near equally populated structures correspond to the two calculated to lie at the lowest energy and to be almost isoenergetic. In the two β-pXyl(H_2O) complexes, the water molecule inserts in two distinct orientations between OH2 and O1. Again these two structures correspond to those calculated to lie at the lowest energy. In both anomers, the water molecule is inserted into the weakest link of the hydrogen-bonded chain, to create a stronger, extended chain of hydrogen bonds, in concordance with the rule established through previous observations.[11]

In contrast to the α-anomer, the water molecule inserts into the β-anomer between OH2 and the anomeric group (O1), which could account for its preferential population in aqueous solution.[19,20]

Acknowledgements

This work was funded by the EPSRC, The Royal Society (R. A. J., USA Research Fellowship; L. C. S., University Research Fellowship), the Leverhulme Trust (Grant No. F/08788D), the CLRC Laser Loan Pool and the Physical and the Theoretical Chemistry Laboratory at Oxford.

References

1 R. Dwek, *Chem. Rev.*, 1996, **96**, 683.
2 A. Helenius and M. Aebi, *Annu. Rev. Biochem.*, 2004, **73**, 1019.
3 H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero and H. Rüdiger, *ChemBioChem.*, 2004, **5**, 740.
4 K. Turton, R. Natesh, N. Thiyagarajan, J. A. Chaddock and K. R. Acharya, *Glycobiology*, 2004, **14**, 923.
5 (*a*) R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia and M. Vijayan, *Curr. Sci.*, 1997, **72**, 855; (*b*) R. Ravishankar, M. Ravindran, K. Suguna, A. Surolia and M. Vijayan, *Curr. Sci.*, 1999, **76**, B39.
6 A. Imberty and S. Peréz, *Chem. Rev.*, 2000, **100**, 4567.
7 A.-J. Petrescu, A.-L. Milac, S. M. Petrescu, R. A. Dwek and M. R. Wormald, *Glycobiology*, 2004, **14**(2), 103.
8 M. R. Wormald, A.-J. Petrescu, Y.-L. Pao, A. Glithero and T. Elliot, *Chem. Rev.*, 2002, **102**, 371.
9 H. Kogelberg, D. Solis and J. Jiménez-Barbero, *Curr. Opin. Struct. Biol.*, 2003, **13**, 646.
10 L. Hemmingsen, D. Madsen, A. Esbensen, L. Olsen and S. Engelsen, *Carbohydr. Res.*, 2004, **339**, 937.
11 P. Çarçabal, R. A. Jockusch, I. Hünig, L. Snoek, R. Kroemer, B. Davis, D. Gamblin, J. Simons, I. Compagnon and J. Oomens, submitted.
12 R. A. Jockusch, F. O. Talbot and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2003, **5**, 1502.
13 R. A. Jockusch, R. T. Kroemer, F. O. Talbot and J. P. Simons, *J. Phys. Chem.*, 2003, **107**, 10725.
14 R. A. Jockusch, R. T. Kroemer, F. O. Talbot, L. C. Snoek, P. Çarçabal, J. P. Simons, M. Havenith, J. M. Bakker, I. Compagnon, G. Meijer and G. von Helden, *J. Am. Chem. Soc.*, 2004, **126**, 5709.
15 F. O. Talbot and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2002. **4**, 3562.
16 R. v. Ree, M. Cabanes-Macheteau, J. Akkerdaas, J. Milazzo, C. Loutelier-Bourhis, C. Rayon, M. Villalba, S. Koppelman, R. Aalberse, R. Rodriguez, L. Faye and P. Lerouge, *J. Bio. Chem.*, 2000, **275**(15), 11451.
17 M. Bardor, C. Faveeuw, A. Fischette-Laine, D. Gilbert, L. Galas, F. Trottein, L. Faye and P. Lerouge, *Glycobiology*, 2003, **13**, 427.
18 L. Guler, Y.-Q. Yu and H. Kenttämaa, *J. Phys. Chem. A*, 2002, **106**, 6754.
19 R. Schmidt, M. Karplus and J. W. Brady, *J. Am. Chem. Soc.*, 1996, **118**, 541.
20 C. Höög and G. Widmalm, *J. Phys. Chem.*, 2001, **105**, 6375.
21 S. Angyal, *Adv. Carbohydr. Chem. Biochem.*, 1984, **42**, 15.
22 Y. Lee, E. Rho, Y. Min, B. Kim and K. Kim, *J. Carbohydr. Chem.*, 2001, **6**, 503.
23 G. Zemplen, *Ber. Dtsch. Chem. Ges.*, 1927, **60B**, 1555.
24 E. Montgomery, N. Richtmeyer and C. S. Hudson, *J. Am. Chem. Soc.*, 1942, **64**, 690.
25 E. G. Robertson and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2001, **3**(1), 1.
26 F. Mohamadi, N. G. J. Tichards, R. L. W. C. Guida, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440.
27 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, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli,

J. 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, J. V. Ortiz, Q. Cui, 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. G. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *GAUSSIAN 03 (Revision C.02)*, Gaussian, Inc., Wallingford, CT, 2004.

28 S. Cybulski and C. Seversen, *J. Chem. Phys.*, 2005, **122**, 14117.

29 T. van Mourik and R. Gdanitz, *J. Chem. Phys.*, 2002, **116**, 9620.

