J|A|C|S ARTICLES

Published on Web 07/20/2009

The Building Blocks of Cellulose: The Intrinsic Conformational Structures of Cellobiose, Its Epimer, Lactose, and Their Singly Hydrated Complexes

Emilio J. Cocinero,[†] David P. Gamblin,[‡] Benjamin G. Davis,*[,‡] and John P. Simons*[,†]

Contribution from the Department of Chemistry, University of Oxford, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, United Kingdom and the Department of Chemistry, University of Oxford, Chemical Research Laboratory, Mansfield Road, Oxford OX1 3TA, United Kingdom

Received April 24, 2009; E-mail: john.simons@chem.ox.ac.uk; Ben.Davis@chem.ox.ac.uk

Abstract: A combination of vibrational spectroscopy conducted under molecular beam conditions and quantum chemical calculation has established the intrinsic three-dimensional structures of the cellulose disaccharide and, focusing on the critical β1,4-linkage at the nonreducing end of the growing cellulose polymer, its C-4′ epimer. Left to their own devices they both adopt a *cis* (*anti-*ϕ*/syn-*ψ) glycosidic configuration, supported in the epimer by strong, cooperative inter-ring hydrogen bonding. In the cellulose disaccharide, however, where the OH-4′(Glc) group is equatorial, the cooperativity is reduced and the corresponding inter-ring hydrogen bonding is relatively weak. The *cis* conformational preference is still retained in their singly hydrated complexes. In the cellulose disaccharide insertion of the water molecule at the favored binding site between OH-4′ and the neighboring hydroxyl group OH-6′ promotes a structural reorganization to create a configuration that parallels that of its *un*hydrated epimer and greatly strengthens the inter-ring hydrogen bonding. In the C-4′ epimer, the axial orientation of OH-4′ blocks this binding site and the bound water molecule simply adds on at the end of the $(OH\text{-}O)_n$ chain, which has a negligible effect on the (already strong) inter-ring bonding. The implications of these results are discussed with respect to the structure and insolubility of native cellulose polymers.

Introduction

Cellobiose, one of nature's fundamental building blocks, unites two glucose (Glc) residues through a β1,4 glycosidic linkage to provide the basic, repeat structural unit of cellulose, Glc*p*-β1,4-D-Glc*p*, a key component of plant cell walls. Cellulose and its subunit, D-glucose, are the most abundant organic molecules on earth. The native extended chains of β1,4-linked cellobiose units in cellulose laterally aggregate to form fibrils with crystalline regions from which analyses suggest that the polymer typically displays an alternating "*trans*" structure with successive residues rotated by ~180° about the glycosidic linkages, facilitating inter-ring hydrogen bonding between OH3 and O5′,[1] see Figure 1. It remains a matter of debate as to whether this *trans* structure is created directly at each enzyme-catalyzed polymerization step, now generally accepted to occur at the nonreducing end of the growing β1,4-linked chain, or whether it is accessed subsequently by relaxation from an initial "*cis*" orientation.[2]

The free cellulose disaccharide unit, cellobiose, also adopts a *trans* conformation, both in the solid state[3] and in aqueous solution;[4] so too, does its C-4′ epimer, lactose Gal*p*-β1,4-D-Glc*p* - where OH4′ (Gal) is oriented axially rather than equatorially. NMR measurements of residual dipolar coupling in lactose in liquid crystal environments,[5] or of inter-residue nuclear Overhauser effects (NOEs) and scalar coupling of the glycosides, methyl-α-cellobioside[3] and methyl-α-lactoside[6] in solution, together with molecular dynamics simulations, lead to average glycosidic dihedral angles similar to those found in crystalline cellulose.

Access to the *intrinsic* conformational landscapes of carbohydrates in the absence of solvent or other environmental interactions, a necessary prerequisite to understanding how they can be altered through interactions in condensed phase environments, has until recently, been possible only through computational modeling. Thus Ramachandran diagrams of glycosides '*in vacuo*', have been modeled through molecular mechanics,[7] molecular dynamics (MD),[8] and quantum mechanical electronic

[†] Physical and Theoretical Chemistry Laboratory, University of Oxford.
[‡] Chemistry Research Laboratory, University of Oxford.

(1) (a) Zugenaier, P. *Prog. Polym. Sci.* **2001**, *26*, 1341–1417. (b) Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.
(2) Saxema, I. M.; Brown, R. M., Jr. *Ann. Botany* **2005**, *96*, 9–21.
(3) Chu, S. S. C.; Jeffrey, G. A. *Acta Crystallogr.* **1968**, *B24*, 830–838.
(4) (a) Olsson, U.; Serianni, A. S.; Stenutz, R. *J. Phys. Chem. B* **2008**, *112*, 4447–4453. (b) Larsson, E. A.; Staaf, M.; Söderman, P.; Höög, C.; Widmalm, G. *J. Phys. Chem. A* **2004**, *108*, 3932–3937. (c) Cheetham, N. W. H.; Dasgupta, P.; Ball, G. E. *Carbohydr. Res.* **2003**, *338*, 955–962.
(5) Martin-Pastor, M.; Cañales, A.; Corzana, F.; Asensio, J. L.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **2005**, *127*, 3589–3595.
(6) Asensio, J. L.; Jiménez-Barbero, J. *Biopolymers* **1995**, *35*, 55–73.
(7) Stortz, C. A.; French, A. D. *Mol. Simul.* **2008**, *34*, 373–389.
(8) Campen, R. K.; Verde, A. V.; Kubicki, J. D. *J. Phys. Chem. B* **2007**, *111*, 13775–13785.

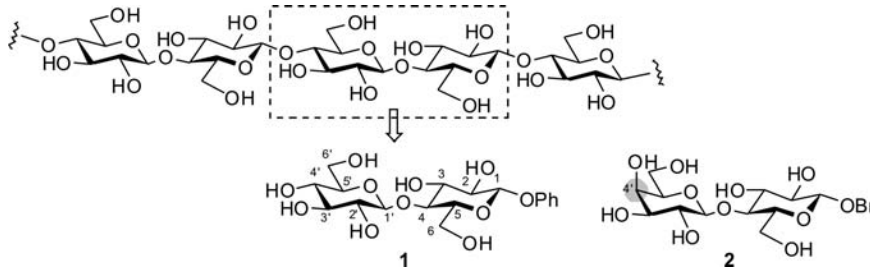


Figure 1. Structural representations of cellulose, the cellulose disaccharide unit (**1**), and its C-4′epimer, the lactoside (**2**).

structure calculations.[9] Interestingly, density functional theoretical (DFT) calculations of cellobiose *in vacuo* have predicted a global minimum energy structure corresponding to an *anti-*ϕ/*syn-*ψ conformation with dihedral angles (ϕ_H and ψ_H) [defined as ϕ_H: H1′−C1′−O′−C4, ψ_H: C1′−O′−C4−H4] ≈ 180°, 0°, respectively, and the two hydroxymethyl groups adopting a *cis* arrangement, rather than the *trans* (*syn*/*syn*) conformation adopted by the disaccharide in condensed phase environments.[10] The range of possible conformations that might in principle, be accessed by the disaccharide units, however, is very large,[7] and their calculated relative energies often lie within the limited accuracy of the computational methods. Without experimental confirmation, identification of their true intrinsic conformational structures remains open to doubt. The same critique also applies to molecular modeling of explicitly hydrated (microsolvated) carbohydrates *in vacuo* although DFT calculations (at 0 K) of microsolvated cellobiose, have predicted relative stabilization of the *trans* configuration with increasing hydration.[11]

Fortunately, it is now possible to determine the intrinsic structures of carbohydrates *and* their microhydrated complexes in the gas phase experimentally, through analysis of their mass and conformer selective vibrational spectra recorded under molecular beam conditions coupled with a combination of molecular mechanics, DFT and *ab initio* calculations.[12−14] The vibrational signatures associated with OH and NH stretching modes are particularly informative since they are extremely sensitive to the local hydrogen bonded conformational environment.[15,16] This strategy has already been used to identify the intrinsic, gas phase structure of benzyl β-lactoside,[12] which turns out to be *cis* rather than *trans*, and it has also exposed the specificity and the consequences of explicit hydration in a series of monosaccharides commonly found in biological systems, including glucose, galactose, mannose, xylose, and fucose.[14] The precise new strategy is ideally suited therefore to address, for the first time, the influence of hydration (and hence, microsolvation) on the glycosidic conformation of disaccharides (and oligosaccharides[17]).

It is applied here to the cellulose disaccharide and, focusing on the critical β1,4-linkage at the nonreducing end of the growing cellulose polymer, its C-4′ epimer, to identify and compare (1) the intrinsic structures of phenyl β-cellobioside (phenyl β1,4-D-glucopyranosyl-β-D-glucopyranoside) **1** and its singly hydrated complex and (2) the corresponding structures of its isolated and singly hydrated epimer, benzyl β-lactoside **2**, see Figure 1. [The phenyl and benzyl groups in **1** and **2** provide the ultraviolet chromophore required for mass selective detection via the resonant two-photon ionization step in the infrared ion-depletion (IRID) technique. Their presence has a negligible influence on the conformational landscapes of the two disaccharides.[12,15]] The changed hydroxyl configuration at C-4′ in **2**, from equatorial to axial, turns out to be highly significant, influencing the selected hydration site, the structural consequences of hydration and the rigidity of the β1,4-glycosidic linkage.

Methods

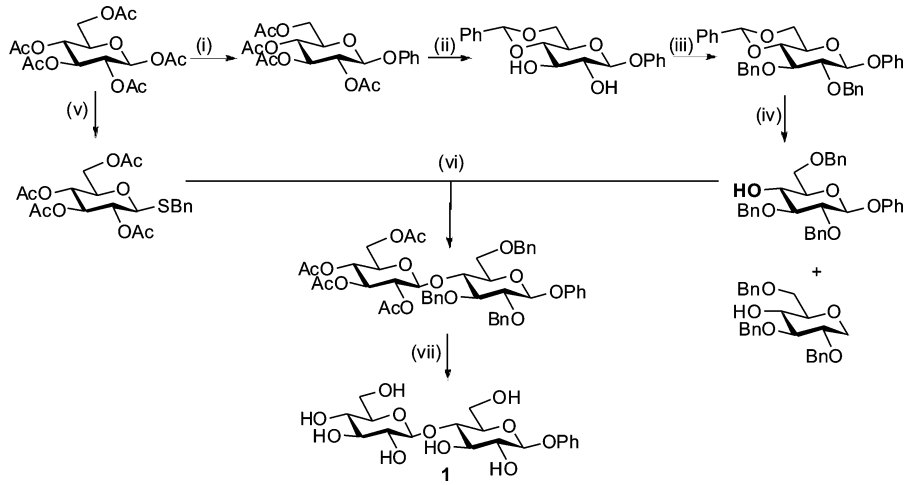
Synthesis. The phenyl glycoside of the cellulose disaccharide unit (**1**) was readily prepared from parent carbohydrate D-glucose (Scheme 1) in >30% overall yield. Thus, the chromophoric phenyl 'tag' was installed into the peracetate of D-glucose using $BF_3 \cdot Et_2O$-catalyzed glycosylation with phenol. Protecting group manipulation allowed selective access to OH-4 in a suitable tribenzylated acceptor monosaccharide. MeOTf-activated glycosylation of OH-4 using the peracetylated benzyl thioglucoside donor allowed access to the cellobioside framework with excellent (>98%) β-stereoselectivity consistent with the presence of a participatory group at C-2 of the donor sugar. Global deprotection (hydrogenolysis and methanolysis) yielded **1**. This route of seven steps (longest linear route), although requiring the stereoselective formation of the interglycosidic β1,4 bond, compared favorably with previous syntheses from less readily available starting block cellobiose[18] (e.g., 16% overall yield)[19] as well as allowing greater potential flexibility in the design of analogues.

The lactose disaccharide (**2**), 'tagged' with a benzyl group was commercially available.

Computation. Initial structures were generated through an extensive conformational search using a combination of the Large-scale Low Mode (which uses low frequency modes to construct conformational changes) and Monte Carlo Multiple Minimization

(9) (a) French, A. D.; Johnson, G. P. *Cellulose* **2004**, *11*, 449–462. (b) French, A. D.; Johnson, G. P. *Mol. Simul.* **2008**, *34*, 365–372.
(10) Strati, G. L.; Willett, J. L.; Momany, F. A. *Carbohydr. Res.* **2002**, *337*, 1833–1849.
(11) (a) Bosma, W. B.; Appell, M.; Willett, J. L.; Momany, F. A. *J. Mol. Struct. (THEOCHEM)* **2006**, *776*, 1–19. (b) Bosma, W. B.; Appell, M.; Willett, J. L.; Momany, F. A. *J. Mol. Struct. (THEOCHEM)* **2006**, *776*, 21–31.
(12) Jockusch, R. A.; Kroemer, R. T.; Talbot, F. O.; Snoek, L. C.; Çarçabal, P.; Simons, J. P.; Havenith, M.; Bakker, J. M.; Compagnon, I.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2004**, *126*, 5709–5714.
(13) Çarçabal, P.; Jockusch, R. A.; Hünig, I.; Snoek, L. C.; Kroemer, R. T.; Davis, B. G.; Gamblin, D. P.; Compagnon, I.; Oomens, J.; Simons, J. P. *J. Am. Chem. Soc.* **2005**, *127*, 11414–11425.
(14) Cocinero, E. J.; Stanca-Kaposta, E. C.; Scanlan, E. M.; Gamblin, D. P.; Davis, B. G.; Simons, J. P. *Chem. Eur. J.* **2008**, *14*, 8947–8955.
(15) Simons, J. P.; Çaŗ; cabal, P.; Davis, B. G.; Gamblin, D. P.; Hünig, I.; Jockusch, R. A.; Kroemer, R. T.; Marzluff, E. M.; Snoek, L. C. *Int. Rev.Phys. Chem.* **2005**, *24*, 489–532.
(16) Cocinero, E. J.; Stanca-Kaposta, E. C.; Gamblin, D. P.; Davis, B. G.; Simons, J. P. *J. Am. Chem. Soc.* **2009**, *131*, 1282–1287.
(17) Stanca-Kaposta, E. C.; Gamblin, D. P.; Cocinero, E. J.; Frey, J.; Kroemer, R. T.; Fairbanks, J. A.; Davis, B. G.; Simons, J. P. *J. Am. Chem. Soc.* **2008**, *130*, 10691–10696.
(18) Montgomery, E. M.; Richtmyer, N. K.; Hudson, C. S. *J. Am. Chem. Soc.* **1943**, *65*, 1848–1854.
(19) Stanek, J.; Kocourek, J. *Chem. Listy Vedu Prum.* **1953**, *47*, 607–702.

Scheme 1. Synthesis of the Cellulose Disaccharide Unit (**1**)[a]



[a] Reagents and conditions: (i) $BF_3 \cdot OEt_2$, PhOH, DCM, 63%; (ii) NaOMe, MeOH (96%), then $PhCH(OMe)_2$, MeCN, pTSA, 82%; (iii) NaH, BnBr, DMF, 88%; (iv) $NaCNBH_3$, HCl, THF, 86% (+7% 1-deoxy impurity); (v) $BF_3 \cdot OEt_2$, BnSH, DCM, 93%; (vi) MeOTf, DCM, 85%; (vii) NaOMe, MeOH then Pd/C, H_2, 95%.

procedures,[20] as implemented in the MacroModel software (MacroModel v. 8.5, Schrödinger, LLC21). Relevant structures were selected on the basis of their calculated relative energies and information gained from previous studies on carbohydrates and their experimental vibrational signatures. These were submitted for further *ab initio* and DFT (B3LYP functional) geometry optimizations using the Gaussian 03 program package[21] and 6-311+G* basis sets for both the isolated and hydrated structures. The lowest energy conformers (cut off ~25 kJ mol^{-1}) were submitted for further higher level calculations, MP2/6-311++G**, to determine their relative energies, which were corrected for zero point and free energy using the frequency calculations performed at the B3LYP level with 6-311+G* basis sets. The frequencies of the OH stretch modes, expressed in wavenumbers, were corrected for anharmonicity using the multiplier, 0.9734.

Spectroscopy. Detailed descriptions of the experimental strategy have been published previously.[12] The carbohydrate samples were mixed with graphite powder and vaporized into a supersonic jet of argon using a focused, Q-switched Nd/YAG laser desorption system operating at 1064 nm; their hydrated complexes were formed by seeding the carrier gas with water vapor (~0.25%). Passage through a 2 mm skimmer created a collimated molecular beam which was crossed by one or two tunable laser beams in the extraction region of a linear time-of-flight mass spectrometer (Jordan). Mass-selected R2PI spectra were recorded using a pulsed Nd/YAG-pumped dye laser (Continuum Powerlite II/Sirah PS-G, 1−3 mJ/pulse UV) operating at 10 Hz. Conformer-specific UV and IR spectra were recorded through IR ion dip (IRID) double resonance spectroscopy. The IRID experiments employed radiation in the range 3100−3800 cm^{-1}, ~3−5 mJ/pulse, generated by difference frequency mixing of the fundamental of a pulsed Nd/YAG laser with the output of a dye laser in a $LiNbO_3$ crystal (Continuum Powerlite 8010/ND6000/IRP module). The delay time between opening the pulsed valve and triggering the UV laser was adjusted to optimize the generation of either 'naked' or hydrated molecules in the supersonic expansion and the delay between the pump and the probe laser pulses was ~150 ns in the IRID hole burning experiments.

Results and Discussion

Spectroscopy and Structure. The UV R2PI excitation spectra of the cellobioside **1** and its C-4′ epimer **2** are shown in Figure 2 and the IR ion dip (IRID) spectrum of **1**, recorded in the OH stretch region, is shown in Figure 3 where it can be compared with that of **2**.[12] The IRID spectrum remained the same regardless of the location of the UV probe laser within its sharply structured R2PI spectrum and both indicated predominant populations of single conformers only, a remarkable result, given the number of potential conformations that might conceivably have been stabilized in the cold molecular beam. Figure 3 also includes the computed vibrational spectra of the four lowest lying conformations of the cellulose disaccharide predicted theoretically. The experimental IRID spectrum is in excellent correspondence with the predicted spectrum associated with its global minimum energy conformation, calculated to lie well below the energies, and free energies of its nearest neighbors. In this structure the disaccharide **1** adopts a *cis* conformation, with (ϕ_H, ψ_H) ≈ 180°, 0°. The calculated relative energies of the lowest lying *trans* conformer are 17.8 kJ mol^{-1} and 11.2 kJ mol^{-1}, respectively.

The appearance of a cluster of partially overlapping bands located between 3600 and 3650 cm^{-1} separated by some 100 cm^{-1} from the intense blended feature centered at ~3545 cm^{-1}, reflects two distinct OH environments associated, respectively, with weak and moderately strong hydrogen bonding. This signature is in quantitative agreement with the predicted vibrational assignments. The global minimum conformation is supported by two inter-ring hydrogen bonds, OH6′→OH6 and OH2′→OH3 (a link in the 'counterclockwise' cooperative $(OH \rightarrow O)_n$ chain, OH4′→OH3′→OH2′→OH3→OH2). The three least displaced (and overlapping) bands correspond to the OH-2 and OH-6 'spectrator' modes, $\sigma 2$ and $\sigma 6$, and the weakly coupled OH-4′ mode, $\sigma 4'$; the two bands at slightly lower wavenumber can be assigned to the modes $\sigma 3$ and $\sigma 3'$; and the (blended) pair of bands at ~3545 cm^{-1} is associated with the two coupled modes, $\sigma 6'$ and $\sigma 2'$.

The IRID spectrum of the C-4′ epimer, the lactoside **2**, and the vibrational spectrum associated with its intrinsic global minimum conformational structure were reported earlier,[12] and are shown for comparison, in the lower part of Figure 3. Like the cellulose disaccharide **1**, the epimer **2** also adopts a *cis* configuration, with (ϕ_H, ψ_H) ≈ 179°, −1°, but epimerization at C-4′ has a striking consequence. The OH4′ group, now oriented axially rather equatorially, promotes a strong OH4′→OH6′

(20) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrikson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440–467.
(21) Frisch, M. J.; et al. *Gaussian 03*, Revision B.03; Gaussian, Inc.: Pittsburgh, PA, 2003.

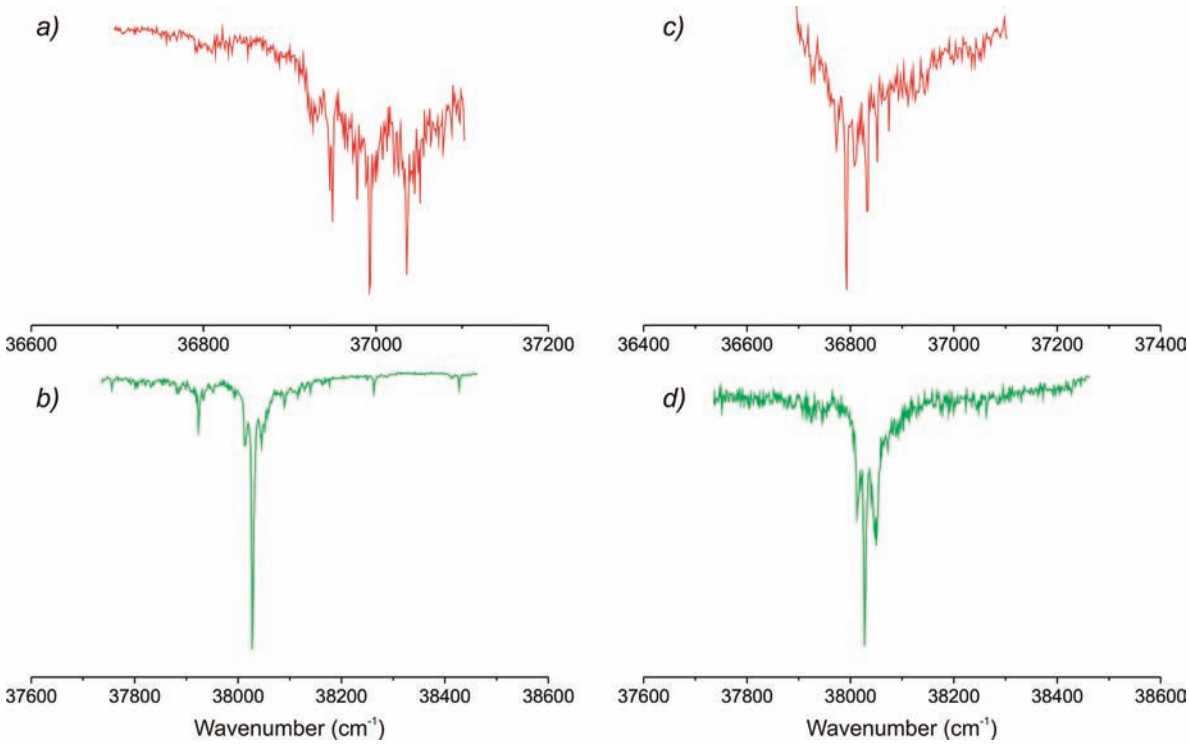


Figure 2. R2PI spectra of (a) the cellobioside **1**, (c) the lactoside **2**, and (b) and (d), their singly hydrated complexes all recorded near their respective UV band origins.

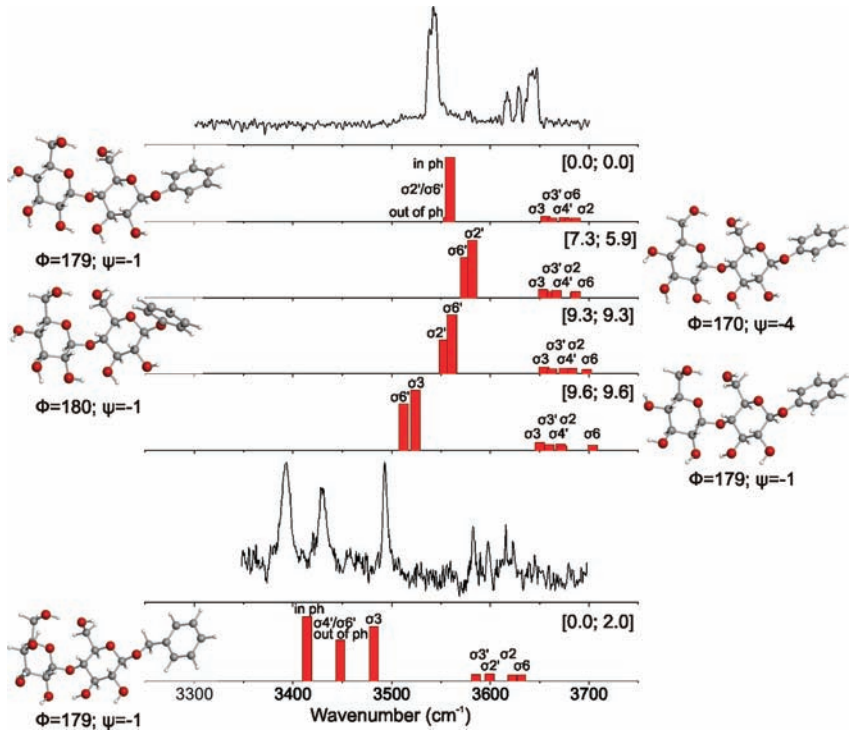


Figure 3. (Top) Experimental IR ion depletion spectrum of the cellulose disaccharide **1** and the computed structures and vibrational spectra of each of its four lowest lying conformers, presented as stick diagrams. (Bottom): Experimental IR and computed vibrational spectrum and structure of the lactoside **2** reported earlier.[12] Relative energies (kJ mol^{-1}) at 0 K and free energies at 298 K are shown in brackets.

interaction which dramatically alters the conformational landscape, reversing the orientation of the cooperative $(OH\rightarrow O)_n$ chain from counterclockwise to clockwise and extending it. This creates much stronger inter-ring binding through the OH6′→OH6 and OH3→OH2′ hydrogen bonds and is reflected in its vibrational signature which displays three intense bands shifted to substantially lower wavenumbers in comparison with the cellobioside; they are associated with the coupled modes, σ4′ and σ6′, connected through strong hydrogen bonding from OH4′→OH6′→OH6, and the σ3 mode.[12]

Hydration. The UV R2PI excitation spectra of the singly hydrated complexes of the cellobioside **1** and its epimer **2** are

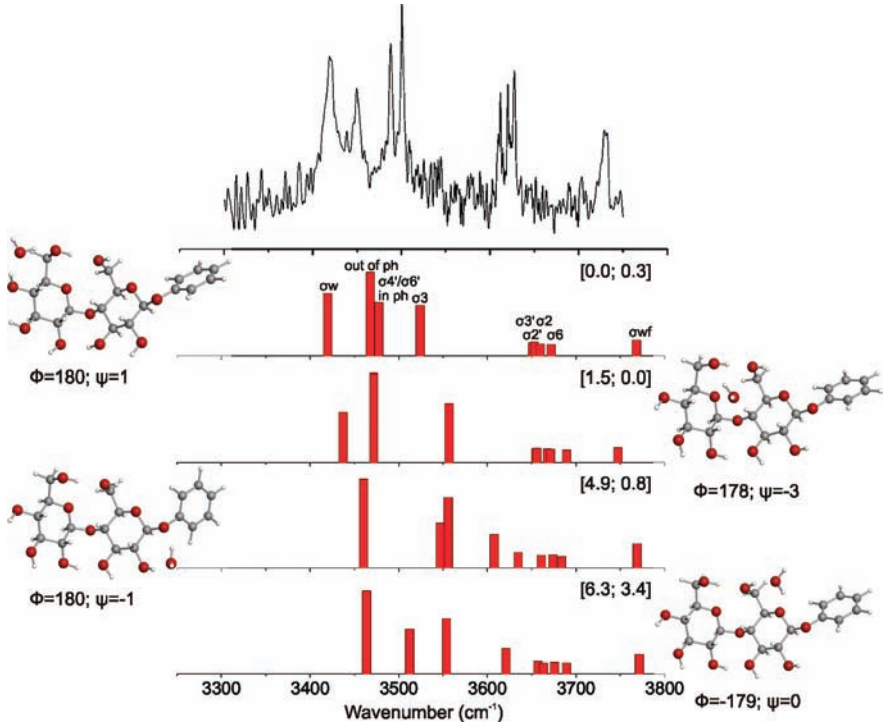


Figure 4. Experimental IR ion depletion and computed vibrational spectra and structures of the lowest lying conformers of the singly hydrated cellulose disaccharide **1**. Relative energies (0 K) and free energies (298 K), in kJ mol^{-1}, are shown in brackets.

shown in the lower half of Figure 2; hydration shifts the spectrum of the cellobioside ~160 cm^{-1} to the red. There was no significant displacement for the epimer, and this was also true for the spectrum recorded in the doubly hydrated ion channel. If the hydrated ions were fragmenting, because of the spectral overlap the signals recorded in the monohydrate ion channel could conceivably have included contributions from doubly, or even multiply hydrated lactose conformers. Fortunately, the recorded IRID spectra, which are shown in Figures 4 and 5 together with the computed vibrational spectra, structures, and relative energies of their four lowest lying singly hydrated conformers, allow this possibility to be excluded, see later discussion.

The principal features in the experimental IRID spectrum of the hydrated cellulose disaccharide **1** are in good agreement with the vibrational spectrum predicted for its global minimum energy structure, particularly the prediction of four bands strongly displaced to lower wavenumbers, but the subsequent predicted spectra are in increasingly poor correspondence with experiment. The global minimum structure, uniquely, incorporates a water molecule inserted between OH4′ and OH6′. This is also the favored binding site in the monosaccharides, phenyl β-D-glucopyranoside and phenyl β-D-mannopyranoside, both of which have an equatorial OH4 group, but not phenyl β-D-galactopyranoside, where OH4 is axial.[14] Hydration at the 4′−6′ site of the cellulose disaccharide **1** reverses the orientation of OH4′ to create a greatly extended, cooperatively hydrogen bonded, clockwise chain, OH2→OH3→OH2′→OH3′→OH4′→H_2O→OH6′→OH6, now incorporating eight rather than five consecutive hydrogen-bonds. This strengthens the inter-ring bond, OH6′→OH6 and also strengthens (and reverses) the bonding between OH3 and OH2′. The dihedral angles, (ϕ_H, ψ_H) ≈ 180°, 1° about the glycosidic linkage do not change, however; the *cis* structure of the hydrated cellulose disaccharide is retained but its rigidity is very much strengthened by the incorporation of a bound water molecule in the extended clockwise $(OH\rightarrow O)_n$ chain. (DFT calculations of *multiply* hydrated cellobiose appear to suggest that a conformational switch from *cis* to *trans* could be promoted by the formation of explicitly hydrated and 'water bridged' structures.[22] The results presented here indicate the switch certainly must require the presence of more than one water molecule). The stabilization in aqueous solution cannot be attributed simply to the polarity of the surrounding solvent since the calculated dipole moment of the lowest lying *trans* conformer of *O*-phenyl cellobiose, 5.2 D, is a little less than that of the *cis* conformer, 7.0 D.

Strikingly, the conformational structure of the monohydrated cellulose disaccharide **1** is now identical to that of the *bare* lactoside **2**, cf. Figure 3, where the OH4′→OH6′ bonding, and the consequent extended clockwise chain, is promoted by the axial orientation of OH4′. In the hydrated cellulose disaccharide **1**, where OH4 is equatorial, the link between OH4′ and OH6′ is mediated by the insertion of a water molecule at the favored (4′−6′) binding site. Its experimental IR spectrum presents a single band at high wavenumber, ~3730 cm^{-1}, readily assigned to the 'free' OH of a bound water molecule; a cluster of bands around 3610 to 3630 cm^{-1}, associated with weakly hydrogen-bonded (intra-ring), peripheral OH groups, and assigned to σ2′, σ3′, σ2, and σ6; and a quartet of strongly displaced bands lying between 3400 and 3500 cm^{-1} associated with OH groups in much more strongly hydrogen-bonded inter-ring and intermolecular environments. They are assigned to σ3, and the coupled vibrations, σ4′ and σ6′, together with the OH mode, σ_W, associated with the bound water molecule. Eight of the nine OH bands are individually resolved. The lowest lying *trans* conformer of the singly hydrated disaccharide **1** has a calculated relative energy of 13.3 kJ mol^{-1} and free energy of 7.9 kJ mol^{-1}, a little less than those of the bare *trans*-disaccharide, 17.8 kJ mol^{-1} and 11.2 kJ mol^{-1}, respectively, but well above those of the corresponding *cis* conformer.

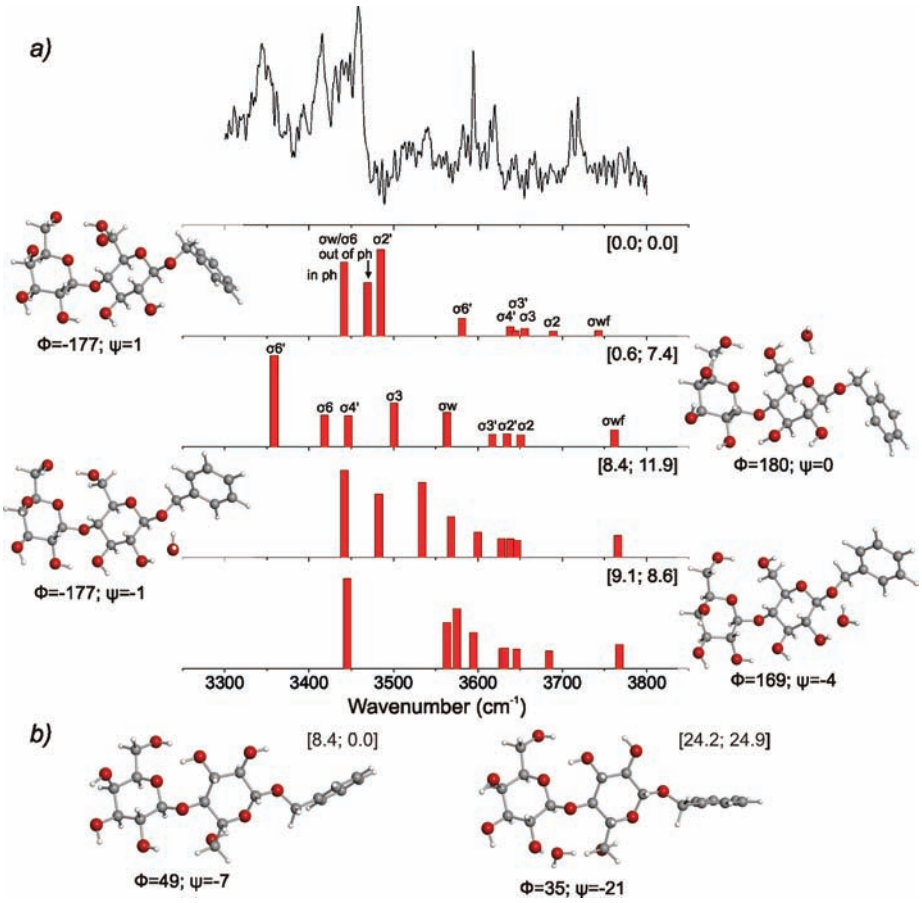


Figure 5. Experimental IR ion depletion and computed vibrational spectra and structures of (a) the lowest-energy conformers of the singly hydrated lactoside **2** and (b) the lowest-energy *trans* conformers of the bare lactoside and its singly hydrated complex. Relative energies (kJ mol^{-1}) at 0 K and free energies at 298 K, are shown in brackets.

The experimental IRID spectrum of the hydrated lactoside **2**, monitored in the singly hydrated ion channel, is shown in Figure 5. It corresponds best to a composite of the two lowest lying computed structures, predicted to be separated by only 0.6 kJ mol^{-1} but lying at energies >8 kJ mol^{-1} below those of their nearest neighbors. If, as noted earlier, there were very extensive ion fragmentation, the doublet at ~3710 cm^{-1} and 3720 cm^{-1} might indicate the binding of two water molecules. Unless there were near 100% fragmentation, however (an improbable scenario given the absence of fragmentation in hydrated cellobiose), three (or more) OH bands would appear in the region around 3720 cm^{-1}. A qualitative comparison between the full experimental spectrum and the computed spectra of the two lowest lying structures strongly favors the alternative interpretation: the experimental spectrum is a composite. The strong band centered at ~3350 cm^{-1} is associated with the structure lying at a relative energy of 0.6 kJ mol^{-1}, and the bands at 3710 and 3720 cm^{-1} correspond to the vibration of the 'free' OH group of the bound water molecule in the two lowest lying, isomeric monohydrate structures.

The strong, pre-existing $(OH4')_{axial}\rightarrow OH6'$ interaction in the galactose ring of lactose prevents insertion of the bound water molecule at the (4′−6′) site, just as it did in the hydrated monosaccharide, phenyl β-D-galactopyranose·(H_2O).[14] In lactose, the water molecule simply "adds on", either to the pre-existing, clockwise chain through an $OH6\rightarrow OH_W$ bond at its terminus, in the structure with relative energy 0.6 kJ mol^{-1}, or in the global minimum energy structure to a reversed, counterclockwise $(OH\text{-}O)_n$ chain through an $OH_W\rightarrow OH6$ bond at its new starting point. In this structure, the three bands at low wave numbers are associated with the strongly hydrogen-bonded, OH2′→OH3, inter-ring mode, $\sigma 2'$, and the coupled $OH_W\rightarrow OH6\rightarrow OH6'$ vibrations, σ_W and $\sigma 6$, involving the bound water molecule and the other inter-ring mode. The vibrational signature of the higher-energy structure reflects the clockwise orientation of the extended $(OH\rightarrow O)_n$ chain. The most strongly displaced band is now $\sigma 6'$, reflecting the strong OH6′→OH6 hydrogen bond, while the next three, $\sigma 6$, $\sigma 4'$, and $\sigma 3$, reflect the $OH6\rightarrow OH_W$ interaction, the inter-ring bond, OH3→OH2′, and the strong $(OH4')_{axial}\rightarrow OH6'$ interaction.

Figure 5 also presents the computed structures of the lowest-energy *trans* conformer of benzyl β-D-lactoside and its singly hydrated complex, (the conformation identified in methyl β-D-lactoside through NMR measurements of its average structure in aqueous solution).[6] Explicit hydration of the lactoside **2** by a single water molecule bridging across the Gal and Glc rings between OH2 and OH6′, actually *de*-stabilizes the *trans* conformation. At 0 and 298 K its relative energy and free energy are calculated to be 24.2 and 24.9 kJ mol^{-1}; the corresponding figures for the unhydrated *trans* conformer are 8.4 and 2.0 kJ mol^{-1}.[12]

Concluding Remarks

The combination of vibrational spectroscopy conducted under molecular beam conditions and quantum chemical calculation

(22) Bosma, W. B.; Appell, M.; Willett, J. L.; Momany, F. A. *J. Mol. Struct. (THEOCHEM)* **2006**, *776*, 21–31.
(23) French, A. D. *Cellulose* **2004**, *11*, 39–42.

has established the intrinsic three-dimensional structures of the cellulose disaccharide **1**, and its C-4′ epimer **2**, in the gas phase at low temperature. Left to their own devices they both adopt a *cis* (*anti/syn*) glycosidic configuration, supported in the lactoside **2** by strong, cooperative inter-ring hydrogen bonding, promoted by the axial orientation of OH4′(Gal). In the cellulose disaccharide **1**, however, where the OH-4′(Glc) group is equatorial, the cooperativity is reduced, and the corresponding inter-ring hydrogen bonding is relatively weak.

The *cis* conformational preference is still retained in their singly hydrated complexes. In the cellulose disaccharide **1** insertion of the water molecule at the favored binding site between OH-4′ and the neighboring hydroxyl group OH-6′ promotes a structural reorganization to create a configuration that parallels that of its *un*hydrated epimer, greatly strengthening the inter-ring hydrogen bonding. In the C-4′ epimer **2** the axial orientation of OH-4′ blocks this binding site, and the bound water molecule simply adds on at the end of the $(OH\text{-}O)_n$ chain; this has a negligible effect on the (already strong) inter-ring bonding. When the disaccharide unit **1** is incorporated into a cellulose polymer, the OH-4′ group is capped in a glycosidic bond, breaking the cooperative $(OH\text{-}O)_n$ chain in the free disaccharide, and removing the favored (4′−6′) water binding site. This might well contribute to the intrinsic resistance of cellulose polymer to solvation,[23] and it may also facilitate a change in conformation about the glycosidic linkages from *cis* to *trans* in the growing polymer, allowing formation of the sequence of OH3→O5′ inter-ring hydrogen-bonds *along* the extended polymer chain. The typical *trans* conformation in cellulose may not be a result of inherent inter-residue interaction but, instead, arises through a combination of its extended structure and microhydration.

Acknowledgment. We have appreciated discussions with Drs. Cristina Stanca-Kaposta and Timothy Vaden and are grateful for the support provided by EPSRC (B.G.D., J.P.S.), the Leverhulme Trust (J.P.S., Grant F/08788G), the Spanish Ministry of Education and Science (E.J.C.), the Oxford Supercomputing Centre, the STFC Laser Loan Pool, and the Physical and Theoretical Chemistry Laboratory.

Supporting Information Available: Detailed description of the synthesis and characterization of phenyl β-D-cellobioside, **1**, Cartesian coordinates and total energies of the conformations and structures shown in Figures 3, 4, and 5, and the complete ref 21 This material is available free of charge via the Internet at http://pubs.acs.org.

JA903322W