

## PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/98891>

Please be advised that this information was generated on 2017-12-06 and may be subject to change.

## Probing the Glycosidic Linkage: UV and IR Ion-Dip Spectroscopy of a Lactoside

Rebecca A. Jockusch,<sup>†</sup> Romano T. Kroemer,<sup>||</sup> Francis O. Talbot,<sup>†</sup> Lavina C. Snoek,<sup>†</sup> Pierre Çarçabal,<sup>†</sup> John P. Simons,<sup>\*,†</sup> Martina Havenith,<sup>§</sup> Joost M. Bakker,<sup>‡</sup> Isabelle Compagnon,<sup>‡</sup> Gerard Meijer,<sup>‡,⊥</sup> and Gert von Helden<sup>⊥</sup>

Contribution from the Department of Chemistry, Physical and Theoretical Chemistry Laboratory, Oxford University, Oxford OX1 3QZ, United Kingdom, and FOM Institute for Plasmaphysics Rijnhuizen, Edisonbaan 14, NL-3439 MN Nieuwegein, The Netherlands

Received December 12, 2003; E-mail: john.simons@chem.ox.ac.uk

**Abstract:** The  $\beta(1\rightarrow4)$  glycosidic linkage found in lactose is a prevalent structural motif in many carbohydrates and glycoconjugates. Using UV and IR ion-dip spectroscopies to probe benzyl lactoside isolated in the gas phase, we find that the disaccharide unit adopts only a single, rigid structure. Its fully resolved infrared ion-dip spectrum is in excellent agreement with that of the global minimum structure computed *ab initio*. This has glycosidic torsion angles of  $\phi_{\text{H}}$  ( $\text{H1}-\text{C1}-\text{O}-\text{C4}'$ )  $\approx 180^\circ$  and  $\psi_{\text{H}}$  ( $\text{C1}-\text{O}-\text{C4}'-\text{H4}'$ )  $\approx 0^\circ$  which correspond to a rotation of  $\sim 150^\circ$  about the glycosidic bond compared to the accepted solution-phase conformation. We discuss the biological implications of this discovery and the generality of the strategies employed in making it.

### Introduction

The biological activities of oligosaccharide chains bound in glycoproteins or glycolipids are intimately linked to their structures and conformations.<sup>1,2</sup> Many selective cell–cell interactions and molecular recognition processes involve glycoconjugates. Specificity is often already present even at the monosaccharide level, and specificity for disaccharides may be 1000-fold higher. Exploration and characterization of the hydrogen-bonded conformational landscapes of oligosaccharides is an essential first step in understanding the origins of their molecular specificity.

The component sugar ring units in oligosaccharide chains are generally taken to be rigid, but the same assumption cannot be made about the glycosidic links between them. Because the types of linkages and torsion angles about these links are primary determinants of the overall 3-D conformation of carbohydrates, their accurate characterization is critical. Unfortunately, the assignment of oligosaccharide structure and conformation is notoriously demanding due to the chemical similarity of component monosaccharides, which all contain multiple hydroxyl groups, capable of extensive intra- and intermolecular hydrogen bonding networks. In some circumstances, the glycosidic linkages may be “stiffened” by direct or solvent-

mediated hydrogen-bonded interactions across them, thereby influencing both their shape and their structural and biological specificity.<sup>3–5</sup> In others, the glycosidic linkages may be highly flexible. This flexibility hinders crystallographic structural determinations; it also complicates or limits NMR (solution-phase) structural assignments based upon time-averaged measurements.<sup>5</sup> Computational studies of (gas-phase) oligosaccharide structure are handicapped by uncertainties in the parameters employed in molecular mechanics calculations, and *ab initio* calculations are restricted by molecular size and complexity constraints,<sup>6–10</sup> though, with the current rate of increase in computer power, these restrictions are diminishing rapidly. Calculations of oligosaccharide conformation in solution are constrained by the additional complexity introduced by interactions with the solvent.<sup>4,6,11</sup> When taken together, these limitations represent a very significant challenge to experiment and to reliable molecular modeling.

A new approach to meeting this challenge has been signaled recently by successful structural studies of two monosaccharides, phenyl- $\beta$ -D-galactopyranoside<sup>12</sup> and phenyl- $\beta$ -D-glucopyranoside,<sup>13</sup> as well as the size-selected hydrated clusters of the

<sup>†</sup> Oxford University.

<sup>‡</sup> FOM Institute for Plasmaphysics Rijnhuizen.

<sup>§</sup> Present address: Department of Physical Chemistry II, Ruhr-University Bochum, D-44780 Bochum, Germany.

<sup>||</sup> Present address: Molecular Modelling and Design, Department of Chemistry, Pharmacia, Viale Pasteur 10, 20014 Nerviano (MI), Italy.

<sup>⊥</sup> Present address: Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-14195 Berlin, Germany.

(1) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.

(2) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637–674.

(3) Chen, J. Y. J.; Naidoo, K. J. *J. Phys. Chem. B* **2003**, *107*, 9558–9566.

(4) Naidoo, K. J.; Chen, J. Y. *J. Mol. Phys.* **2003**, *101*, 2687–2694.

(5) Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371–386.

(6) Imberty, A.; Perez, S. *Chem. Rev.* **2000**, *100*, 4567–4588.

(7) French, A. D.; Johnson, G. P.; Kelterer, A. M.; Dowd, M. K.; Cramer, C. J. *Int. J. Quantum Chem.* **2001**, *84*, 416–425.

(8) Strati, G. L.; Willett, J. L.; Momany, F. A. *Carbohydr. Res.* **2002**, *337*, 1833–1849.

(9) Asensio, J. L.; Jimenez-Barbero, J. *Biopolymers* **1995**, *35*, 55–73.

(10) Engelsen, S. B.; Koca, J.; Braccini, I.; Dupenhoat, C. H.; Perez, S. *Carbohydr. Res.* **1995**, *276*, 1–29.

(11) Naidoo, K. J.; Brady, J. W. *J. Am. Chem. Soc.* **1999**, *121*, 2244–2252.

(12) Jockusch, R. A.; Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1502–1507.

glucopyranoside,<sup>14</sup> all isolated in the gas phase. In each case, a combination of vibrational spectroscopy and ab initio computation provided a clear, atomic-scale perspective of the conformational landscapes of the glycosides, both free and hydrated. The strategy was greatly facilitated by the extraordinary sensitivity of the O–H vibrational frequencies to subtle changes in the hydrogen-bonded environment, encouraging optimism about its application to the more challenging target of oligosaccharide structural determination.

Lactose, a commonly found disaccharide, which joins galactose (Gal) and glucose (Glc) units through a  $\beta(1\rightarrow4)$  linkage, presents an obvious first choice target. In addition to its presence in lactose, this type of  $\beta(1\rightarrow4)$  linkage is a common structural element found in numerous carbohydrate structures including plant cell walls and insect and crustacean exoskeletons, and it is in the common core oligosaccharide in N-linked glycoproteins. The ubiquity of this linkage is an indication of its crucial biological importance.

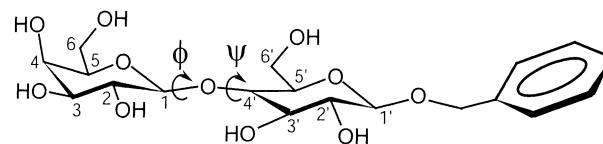
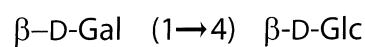
NMR studies of lactose in solution indicate the dominance (>85%) of a single conformation in which the hydroxymethyl groups of the two rings adopt a trans orientation about the glycosidic linkage [with the glycosidic torsion angles,  $\phi_H$  (H1–C1–O–C4')  $\approx 35^\circ$ ,  $\psi_H$  (C1–O–C4'–H4')  $\approx -21^\circ$ ].<sup>9,15,16</sup> X-ray diffraction and NMR studies of lactose loosely bound to proteins also indicate the adoption of a trans conformation.<sup>17</sup> In contrast, empirical force field and ab initio computational studies of free lactose and of its near relative cellobiose (Glc $\beta$ -(1 $\rightarrow$ 4)Glc) indicate at least three regions of torsional stability about the glycosidic linkage.<sup>7,9,10,18,19</sup> The mimic C-lactose, in which the oxygen bridge is replaced with a CH<sub>2</sub> group, appears to be much more flexible; it populates all three regions of torsional stability in solution, and its conformers are differentially selected by different (O-lactose binding) proteins.<sup>20,21</sup>

The results of our gas-phase investigation are presented in the following report, which describes a combined spectroscopic and ab initio computational determination of the structure of benzyl  $\beta$ -lactoside (O-4-benzyl  $\beta$ -D-Gal-(1 $\rightarrow$ 4)- $\beta$ -D-Glc, Scheme 1). The addition of the benzyl ring to lactose provides a chromophore which can be ionized by resonant two-photon ionization (R2PI), allowing for mass spectrometric separation and detection, but which does not disturb the original lactose conformational preferences.

## Methods

**Calculations.** An ensemble of starting structures for benzyl  $\beta$ -lactoside was generated using the random search procedure implemented

**Scheme 1**



glycosidic linkage torsion angles:

$$\phi_H = \text{H1-C1-O-C4}' \quad \psi_H = \text{C1-O-C4}'\text{-H4}'$$

in Sybyl6.5.<sup>22</sup> In the search, all rotatable bonds (i.e., the seven OH groups, the hydroxymethyl groups, the glycosidic linkage, and the linkage to the benzyl moiety) were included as were different ring conformations. 10 000 tries/searches were performed, and 214 unique low-energy conformations were identified. Subsequent ab initio calculations were done using the Gaussian98 package Rev.A11.<sup>23</sup> The 214 conformers were submitted to full geometry optimization at the HF/3-21G level, and for each of the optimized conformers, a single-point calculation at the B3LYP/6-31+G\* level was carried out. The resulting energies indicated the presence of 20 unique conformers within 25 kJ/mol of the minimum energy conformer. These 20 structures were then reoptimized at the B3LYP/6-31+G(d) level, following which 12 conformers were within the 25 kJ/mol energy cutoff. Harmonic frequencies for each of the remaining 12 conformers were calculated at the B3LYP/6-31+G(d) level of theory. Single-point energies for each of these conformers were evaluated at the MP2/6-311++G(d,p) level. Tabulated values of  $E_0$  include the MP2/6-311++G(d,p) energies and vibrational zero-point energy corrections from the B3LYP/6-31+G(d) frequencies. Thermochemistry calculations also include the MP2 electronic energies and thermal corrections from the B3LYP frequencies, which were done using the standard procedures implemented within the Gaussian program. For comparison with measured near-infrared spectra, calculated O–H stretching modes were uniformly scaled by 0.9734, an empirical scaling factor which compensates for systematic errors arising from the level of theory used as well as anharmonic effects. This scaling factor provided an excellent match for the weakly perturbed O–H stretching frequencies measured for the component monosaccharides.<sup>12,13</sup>

**R2PI and Near-IR Spectroscopy.** The instrumentation described here was also used in previous studies of monosaccharide conformation.<sup>12–14</sup> The disaccharide derivative benzyl  $\beta$ -lactoside (Sigma Chemical Co.) was placed in an oven attached to the expansion side of a pulsed nozzle (0.8 mm diameter, 10 Hz pulse rate, 4 bar argon backing pressure). The samples were heated to 505 K to generate a continuous stream of vapor, which was subsequently entrained and cooled in a pulsed expanding argon jet. The skimmed beam was intersected by infrared (IR) and/or ultraviolet (UV) laser beam(s) tuned to the desired wavelength(s) for excitation and ionization. Molecular ions were separated and detected using a time-of-flight mass spectrometer (Jordan); a digital oscilloscope (Tektronix TDS 320) interfaced to a PC, running in-house software, was used to collect the gated, mass-selected ion signal.

Mass-selected resonant two-photon ionization (MS-R2PI) spectra were recorded using tuneable UV laser radiation generated by frequency

- (13) Talbot, F. O.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3562–3565.  
 (14) Jockusch, R. A.; Kroemer, R. T.; Talbot, F. O.; Simons, J. P. *J. Phys. Chem. A* **2003**, *107*, 10725–10732.  
 (15) Hayes, M. L.; Serianni, A. S.; Barker, R. *Carbohydr. Res.* **1982**, *100*, 87–101.  
 (16) Cheetham, N. W. H.; Dasgupta, P.; Ball, G. E. *Carbohydr. Res.* **2003**, *338*, 955–962.  
 (17) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.  
 (18) Cloran, F.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1999**, *121*, 9843–9851.  
 (19) Engelsen, S. B.; Rasmussen, K. *J. Carbohydr. Chem.* **1997**, *16*, 773–788.  
 (20) Espinosa, J. F.; Montero, E.; Vian, A.; Garcia, J. L.; Dietrich, H.; Schmidt, R. R.; Martin-Lomas, M.; Imbert, A.; Canada, F. J.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120*, 1309–1318.  
 (21) Espinosa, J. F.; Canada, F. J.; Asensio, J. L.; Martin-Pastor, M.; Dietrich, H.; Martin-Lomas, M.; Schmidt, R. R.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1996**, *118*, 10862–10871.

- (22) Clark, M.; Cramer, R. D.; van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982–1012.  
 (23) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Salvador, P.; Dannenberg, J. J.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Baboul, A. G.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gonzalez, C.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, revision A.11; Gaussian, Inc.: Pittsburgh, PA, 2001.

doubling the output of an Nd:YAG-pumped pulsed dye laser [GCR-170/Laser Analytical Systems (LAS)] operated with Coumarin 153 dye ( $\sim 0.4\text{--}1$  mJ pulse $^{-1}$ ). The same laser was employed in the UV/UV hole-burn experiments to provide the initial (burn) laser pulse: the “probe” pulse, fired  $\sim 200$  ns after the scanned “burn” laser, was generated by a second, excimer-pumped pulsed dye laser (Lambda Physik FL3002), tuned to successive features of interest in the R2PI spectrum. Resonant IR ion-dip spectra were recorded using tuneable pulsed IR radiation ( $\sim 3000\text{--}4000$  cm $^{-1}$ ,  $0.4$  cm $^{-1}$  bandwidth,  $4\text{--}5$  mJ pulse $^{-1}$ ), generated by difference frequency mixing the output of an Nd:YAG laser (Powerlite Precision 8000) and that of a pumped dye laser (ND6000, Continuum) in a LiNbO $_3$  crystal: the UV probe pulse was provided by the LAS laser system, fired  $\sim 150$  ns after the IR pulse. In all experiments the timing of the pulsed free-jet expansion and laser pulses was controlled via a digital delay pulse generator (DG535 Stanford Research Systems, Inc.).

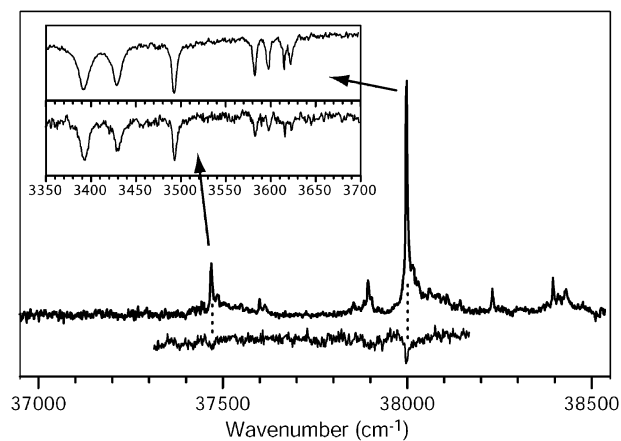
**Mid-IR Spectroscopy.** Benzyl  $\beta$ -lactoside was embedded in a graphite matrix which was placed just below the orifice of a pulsed valve (R. M. Jordan,  $0.5$  mm diameter,  $10$  Hz pulse rate,  $2$  bar argon backing pressure). Benzyl  $\beta$ -lactoside molecules were desorbed from the matrix by a pulsed Nd:YAG laser (Thales Diva II,  $1064$  nm,  $<1$  mJ per pulse,  $10$  ns pulse duration) and subsequently entrained and cooled in the pulsed expanding argon jet. The skimmed beam was intersected by IR and UV laser beams tuned to the desired excitation wavelength. Molecular ions were separated and detected using a time-of-flight mass spectrometer (Jordan); a digital oscilloscope (LeCroy 9430) interfaced to a PC, running in-house software, was used to collect the gated, mass-selected ion signal.

The source of mid-IR radiation was the free-electron laser for infrared experiments (FELIX).<sup>24,25</sup> The FELIX output consisted of a few microseconds long burst (macropulse) of micropulses. The micropulse spacing within the burst was set to  $1$  ns, while the macropulses were repeated at  $5$  Hz. The micropulse duration was set to about  $100$  optical cycles, which resulted in a bandwidth of approximately  $1\%$  (fwhm) of the central frequency. The wavelength range that can be covered by FELIX extends from  $40$  to  $2000$  cm $^{-1}$ , although, in this experiment, only the region between  $300$  cm $^{-1}$  and  $1500$  cm $^{-1}$  was used. In a macropulse, typically, energies of up to  $100$  mJ could be reached.

The UV probe pulse was provided by frequency doubling the output of an Nd:YAG-pumped pulsed dye laser [Spectra Physics GCR-100/PDL3] operated with Coumarin 153 dye ( $\sim 0.4\text{--}1$  mJ per pulse). The timing of the pulsed free-jet expansion and UV laser pulses was controlled via a digital delay pulse generator (SRS, DGS 535) that was synchronized to the FELIX master clock. The FELIX macropulse frequency of  $5$  Hz (i.e., half the repetition rate of the UV laser) allowed for recording alternating IR *on* and IR *off* signals to provide a good normalization of the ion-dip signal. The ion-dip spectrum was then converted into an absorption spectrum following the method outlined previously.<sup>26</sup>

## Results and Discussion

The near ultraviolet spectrum of the lactoside is remarkably simple (Figure 1): the ultraviolet “hole-burning” of the intense band at  $38\,000$  cm $^{-1}$  leads to a “dip” in the intensity of the main features in the R2PI spectrum (accounting for at least  $\sim 80\%$  of the total measured signal), establishing their association with a *single, unique* conformation. This conclusion is confirmed by the result of a similar near-infrared “ion-dip” experiment, also shown in Figure 1; the principal ultraviolet features are each associated with identical infrared spectra.



**Figure 1.** Resonant two-photon ionization and ultraviolet “hole-burn” spectrum of jet-cooled benzyl  $\beta$ -lactoside and (inset) the associated resonant near-infrared ion-dip spectrum.

Despite the potentially large range of conformations that might have been expected a priori in the vapor (different rotamers of the hydroxyl and hydroxymethyl groups<sup>12,13,27</sup> in conjunction with the three sets of glycosidic linkage torsion angles), only one is significantly populated. This could be due to its selective population in the vaporized sample or to the relaxation of an initial population distribution into a single conformation (the global minimum?) as the lactoside is cooled to low temperatures during the free-jet expansion. In similar experiments the component monosaccharides Gal and Glc populate more than one conformer each (two and three, respectively) associated with distinct orientations of the hydroxymethyl group,<sup>12,13</sup> so the first explanation is perhaps the more likely. Regardless, this conformation may have profound implications as an “energetic starting point” structure for lactose. The lack of extensive vibronic progressions suggests that the disaccharide conformation presents a more rigid structure than its isolated component monosaccharides.

So, what is its preferred conformation? To answer this question, we turn to the near-infrared vibrational spectrum of jet-cooled benzyl  $\beta$ -lactoside (which displays all seven of its separately resolved O–H stretch modes) and compare it with the results of a series of *ab initio* calculations (Figure 2). The figure shows the conformations and associated (uniformly scaled) harmonic O–H stretching frequencies of the four lowest lying structures computed using density functional theory (DFT) together with (0 K) relative energies from MP2 calculations.

Inspection of Figure 2 reveals a remarkably close and unique match between the pattern of vibrational bands recorded experimentally and that associated with the calculated global minimum energy conformation. None of the other four low-lying (shown) or eight other higher-energy conformations (Supporting Information: Figure S1 and Tables S1–S3) present so similar a spectral pattern, and there is no agreement at all with the pattern that would be presented by the lowest-lying *trans* conformer (Figure 2D), which has a structure similar to the accepted solution-phase conformation of lactose and is calculated to lie at an energy ca.  $9$  kJ mol $^{-1}$  above the global minimum.

The global minimum conformation, which is associated with a *cis* orientation of the two hydroxymethyl groups, corresponds

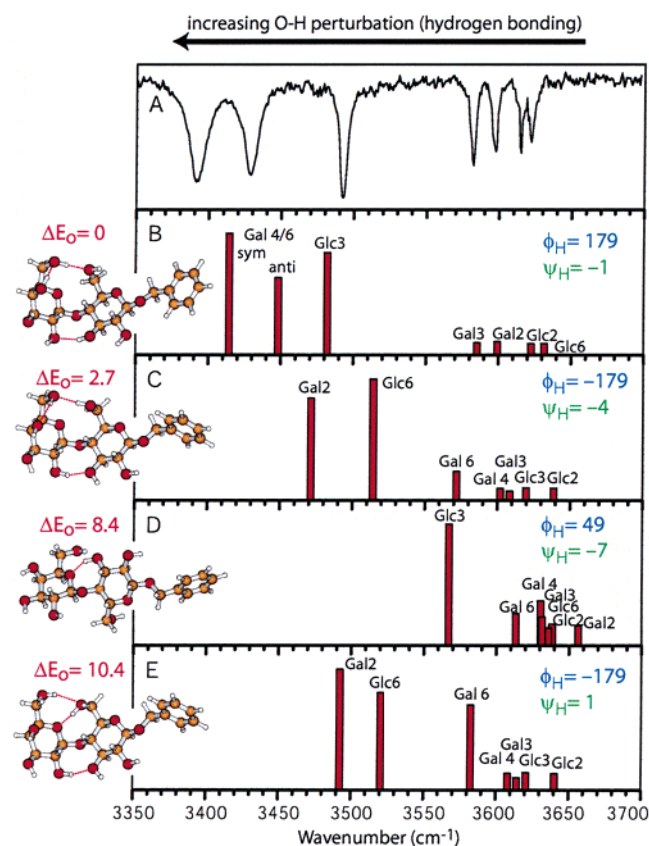
(24) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *Infrared Phys. Technol.* **1995**, *36*, 297–308.

(25) Bakker, J. M.; MacAleese, L.; Meijer, G.; von Helden, G. *Phys. Rev. Lett.* **2003**, *91*, art. no. 203003.

(26) Bakker, J. M.; MacAleese, L.; von Helden, G.; Meijer, G. *J. Chem. Phys.* **2003**, *119*, 11180–11185.

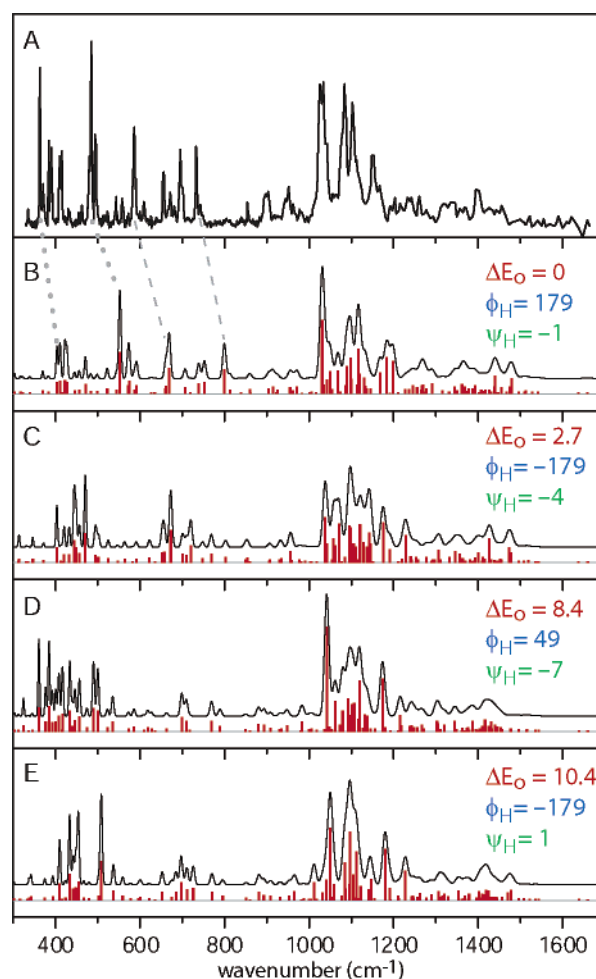
(27) Ro en, A.; Padr on, J. I.; V azquez, J. T. *J. Org. Chem.* **2003**, *68*, 4615–4630.





**Figure 2.** Experimental (A) and computed (B–E) near-infrared ion-dip spectra of benzyl  $\beta$ -lactoside. Structures and frequencies are computed using density functional theory at the B3LYP/6-31+G\* level of theory. Computed bands are labeled with the O–H stretching motion to which they correspond according to the normal coordinate analysis. All frequencies are scaled by 0.9734. Relative energies ( $\text{kJ mol}^{-1}$ ) are computed at the MP2/6-311++G\*\* level and include zero-point energy corrections from unscaled B3LYP/6-31+G(d) level harmonic frequency calculations.

to the dihedral angles of  $\phi_{\text{H}} \approx 179^\circ$ ,  $\psi_{\text{H}} \approx -1^\circ$  (Figure 2B). This structure is similar to one of the observed conformations of the mimic C-lactose,<sup>20</sup> and it parallels the density functional theory predictions for the analogous  $\beta$ -(1 $\rightarrow$ 4) linked disaccharide, cellobiose (Glc $\beta$ (1 $\rightarrow$ 4)Glc).<sup>3,8</sup> Calculations using the semiempirical AMBER force field with a low dielectric constant also favor a cis orientation.<sup>9</sup> On the other hand, calculations using another popular semiempirical force field, MM3, predict a trans conformation about the glycosidic linkage to be the global minimum for methyl  $\beta$ -lactoside.<sup>28</sup> Moreover, the gas-phase results presented here contrast with the trans glycosidic conformation adopted by (room temperature) methyl lactoside in solution,<sup>9</sup> in its crystalline methanol solvate<sup>29</sup> and loosely bound in proteins.<sup>17</sup> Interestingly, the only tightly-binding enzyme for which the enzyme–lactose complex structure has been solved shows lactose in a trans configuration,<sup>30</sup> while the C-lactose analogue bound to the same enzyme shows a cis configuration.<sup>20,31</sup>



**Figure 3.** Experimental (A) infrared spectrum of benzyl  $\beta$ -lactoside recorded in the mid-infrared (“fingerprint” region), using a free-electron laser (FELIX), and computed (B–E) infrared spectra. Computed spectra are unscaled but include a convolution of the stick spectrum with the spectral profile of FELIX. Dotted and dashed lines indicate the correspondence between the experimental bands and those computed for the most stable conformer in the lower energy portion of the spectrum. In this region, the intense bands are predominantly coupled C–O torsions and are more poorly described by the calculations than the more localized modes at higher energy.

The experimental observations are reinforced when the strategy is extended much further into the infrared, by using radiation in the “fingerprint” region between 300 and 1700  $\text{cm}^{-1}$ , supplied by the free-electron laser “FELIX”, recently used to study the structure of structural conformers of gas-phase tryptophan.<sup>25</sup> Figure 3 shows the resonant infrared spectrum of laser desorbed, jet-cooled benzyl  $\beta$ -lactoside recorded in the mid-infrared, together with the calculated (unscaled) harmonic infrared spectra associated with each of the four lowest-lying computed conformations. About 40 separated lines can be identified, with line widths mainly determined by the FELIX spectral profile; therefore, to facilitate the comparison of the calculations to the recorded spectrum in the mid-IR region, the figure includes a convolution of the calculated spectra with the spectral profile of FELIX. The lines seem to fall into two groups, one at low wavenumbers between 350 and 800  $\text{cm}^{-1}$  (intense modes associated with twisting motions of the hydroxyl groups coupled through the hydrogen-bond chains) and one at higher wavenumbers between 900 and 1300  $\text{cm}^{-1}$  (associated principally with C–C and C–O–C stretching modes). Bands in the 900–1300  $\text{cm}^{-1}$  region match the predicted (harmonic) fre-

(28) Rivet, A.; Sabin, C.; Mazeau, K.; Imbert, A.; Perez, S. <http://www.cer-mav.cnrs.fr/>, accessed May 2003.

(29) Stenutz, R.; Shang, M. Y.; Serianni, A. S. *Acta Crystallogr. Sect. C: Cryst. Struct. Commun.* **1999**, *55*, 1719–1721.

(30) Juers, D. H.; Heightman, T. D.; Vasella, A.; McCarter, J. D.; Mackenzie, L.; Withers, S. G.; Mathews, B. W. *Biochemistry* **2001**, *40*, 14781–14794.

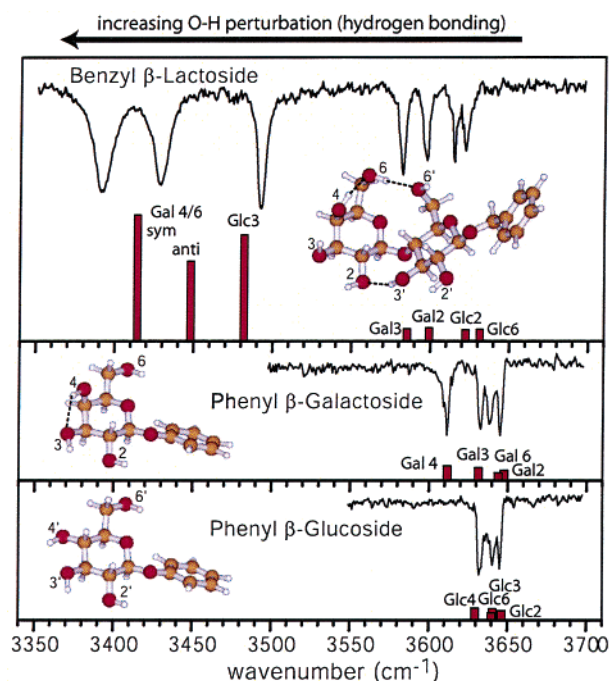
(31) Garcia-Herrero, A.; Montero, E.; Munoz, J. L.; Espinosa, J. F.; Vian, A.; Garcia, J. L.; Asensio, J. L.; Canada, F. J.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **2002**, *124*, 4804–4810.

quencies quite well, but the C–O torsions appear at a slightly lower wavenumber ( $\sim 50\text{ cm}^{-1}$ ) than calculated.

Comparison of the experimental mid-IR spectrum with the computed spectra again favors the assignment to the global minimum structure. The region between  $900\text{ cm}^{-1}$  and  $1300\text{ cm}^{-1}$  matches all the computed spectra except C quite well, although arguably the match to the global minimum, B, is the best. The measured spectrum between  $350$  and  $800\text{ cm}^{-1}$  does not correspond very well to any of the calculated spectra, though, of these, B provides the closest match, especially in the  $550$ – $800\text{ cm}^{-1}$  portion of this region. The poorer appearance of the mid-IR match compared to the near-IR is an indication that the frequencies and intensities of the lower energy modes, particularly the twisting motions of the hydroxyl groups coupled through hydrogen-bond chains, are clearly not as well reproduced by the harmonic calculations used here. The spectrum is much richer in the mid-IR than in the near-IR and thus intrinsically contains more information, but its full assignment is complicated by the spectral congestion and difficulties inherent in calculating the vibrational frequencies and intensities of these lower energy modes.

The experimental (and computed) spectra reveal a remarkable sensitivity of the O–H modes to their hydrogen-bonded environment: (for reference, the “free”, nonbonded O–H stretch of isolated methanol lies around  $3680\text{ cm}^{-1}$ , while in the methanol dimer the hydrogen-bonded O–H stretch is shifted to a lower wavenumber by  $\sim 110\text{ cm}^{-1}$ ).<sup>32</sup> In the three lowest-lying conformers of benzyl  $\beta$ -lactoside, the correspondence between the relative energies of the computed structures and the spectral shifts of the most strongly hydrogen bonded modes is striking. Less stable structures have less strongly shifted frequencies, and both correlate with the decreased stabilization from weaker hydrogen bonding. In the global minimum configuration (Figure 2B), the three O–H bands shifted most strongly to low wavenumbers are all associated with strong hydrogen-bonding across the glycosidic ( $1\rightarrow 4$ ) linkage, which connects the two rings through the interaction between their hydroxymethyl groups, Gal OH6 $\rightarrow$ Glc OH6', and through the Glc OH3' $\rightarrow$ Gal OH2 bond. The vibrational mode associated with the first of these is split into a symmetric and an antisymmetric pair by the cooperative hydrogen-bonded chain Gal OH4 $\rightarrow$ Gal OH6 $\rightarrow$ Glc OH6'. The quartet of bands lying in the higher wavenumber range is associated with the four remaining hydroxyl groups, Gal OH3, Gal OH4, Glc OH2', and Glc OH6'. All are members of the hydrogen-bonded chain, OH2' $\rightarrow$ OH3' $\rightarrow$ OH2 $\rightarrow$ OH3 $\rightarrow$ OH4 $\rightarrow$ OH6 $\rightarrow$ OH6', which circles around the lactose unit. Since the chain terminates at OH6', the mode Glc OH6' is the least displaced from a “free” O–H stretching frequency.

The near-infrared spectrum and global minimum structure of benzyl  $\beta$ -lactoside are compared with those of the two component monosaccharides phenyl  $\beta$ -Gal and phenyl  $\beta$ -Glc in Figure 4. This series of spectra further highlights the sensitivity of an O–H vibrational band structure to the molecular environment: the slightly greater shift toward lower wavenumber in the weakly bonded quartet of the lactoside compared to the monosaccharides results from the enhanced cooperativity associated with the extended hydrogen-bonded chain circling



**Figure 4.** Experimental and computed near-infrared ion dip spectra of benzyl  $\beta$ -lactoside, and of its component monosaccharides, phenyl  $\beta$ -galactoside and phenyl  $\beta$ -glucoside. Computed wavenumbers, calculated at the B3LYP/6-31+G\* level have been scaled by 0.9734.

around the two rings. The change in the chirality at C4 in Gal compared to Glc facilitates an enhanced hydrogen bond linking OH4 $\rightarrow$ OH3 and a correspondingly enhanced spectral shift in the Gal OH4 mode.

The particularly favorable hydrogen-bonding across the glycosidic linkage in the global minimum conformation of the lactoside, facilitated by the rotation of the two hydroxymethyl groups, creates a remarkably rigid structure. The reversal in the orientation of the circling belt of OH groups, which switch from a counterclockwise configuration in the monosaccharides to a clockwise configuration (as oriented in the figures) in the lactoside, also helps stabilize the lactoside (compare Figure 2, frames B and C).

## Conclusions

The rigid, cis oriented structure of the benzyl  $\beta$ -lactoside isolated in a cold ( $\sim 10\text{ K}$ ) gas-phase environment differs from the commonly accepted structures of (room temperature) lactose and methyl lactoside in solution and bound in proteins, where the hydroxymethyl groups adopt a trans configuration about the glycosidic linkage. The difference is not due to the presence of the benzyl group. Calculated structures of methyl  $\beta$ -lactoside show virtually no change in either the relative energies or conformation (of the disaccharide portion of the molecule) compared with benzyl  $\beta$ -lactoside; the global minimum of methyl  $\beta$ -lactoside is also a cis oriented structure. Substitution of a hydroxyl group for the benzyl in the Glc 1 position (to create lactose) would most likely simply lengthen the encircling hydrogen-bond chain through formation of a Glc OH1' $\rightarrow$ Glc OH2' bond. However, both the low temperature and the isolation may well affect the lactoside structure. The cis configuration found at low temperature in the isolated lactoside allows the formation of strong inter-ring hydrogen bonds on two sides of the glycosidic linkage to create a rigid disaccharide structure.

(32) Huisken, F.; Kulcke, A.; Laush, C.; Lisy, J. M. *J. Chem. Phys.* **1991**, *95*, 3924–3929.

Because of this, the enthalpic benefit of the strong hydrogen bonds is somewhat offset by the loss in flexibility. There is an entropic cost, and it is the trans configuration that is favored entropically.<sup>19</sup> Gibbs free energy calculations (room temperature; harmonic frequencies) indeed significantly stabilize the trans conformer over the cis and predict the trans to be the most stable conformer at room temperature by  $\sim 2$  kJ mol<sup>-1</sup> (Supporting Information: Table S1). However, the low-temperature favored cis conformer should still be present, at about half the abundance of the trans conformer given this energy difference. Specific complexing of the disaccharide with water molecules may also alter the preferred conformation about the glycosidic linkage.<sup>4</sup> We plan further gas-phase studies of this lactoside to investigate explicit hydration effects.

The identification of an unusual and very strongly “enthalpically favored” cis conformation, encourages speculation as to its biological implications. The type of  $\beta(1\rightarrow4)$  glycosidic linkage found in the lactoside examined here (in which the O2–C2–C1–O1 dihedral angle is  $\sim -60^\circ$ ) is by far the most prevalent structural disaccharide motif in all of the critical carbohydrate structures found in nature. Key examples include the GlcNAc $\beta(1\rightarrow4)$ GlcNAc chitobiose core of *N*-linked glycoprotein glycan structures and chitin and the Glc $\beta(1\rightarrow4)$ Glc cellobiose repeating unit found in cellulose which, as a vital plant cell wall component, is the most prevalent carbon-based chemical building block on the planet. Given the striking conformational bias and consequential rigidity observed here, it may well be that nature has selected this  $\beta(1\rightarrow4)$  glycosidic motif to be a primary source of substructure in certain circumstances, providing highly specific receptor docking interactions for small ligands (such as the Gal $\beta(1\rightarrow4)$ Glc lactose unit), cell wall rigidification (e.g., through the Glc $\beta(1\rightarrow4)$ Glc linkage in cellulose), and the creation of a spacer unit to create distance from protein surfaces (e.g., in the first two carbohydrate units, GlcNAc $\beta(1\rightarrow4)$ GlcNAc, bound to proteins in *N*-linked glycans). Subtle alteration of the groups, but not the core configuration (e.g., changing OH2 to NHAcetyl to create GlcNAc), allows further fine-tuning of this conformational rigidity.

Finally, we note that a strong interaction of the lactose disaccharide with a biological receptor could have the effect of mitigating the entropic effects that lead to formation of trans orientations in the “free” lactose. In essence, during ligand–receptor docking, the receptor would pay the entropic cost incurred by the ligand in its undocked room temperature state. The isolated gas-phase structure could then represent a snapshot of the true ligand structure in biological interactions, glimpsed in a “naked” state. Examination of the few, available structures in which  $\beta(1\rightarrow4)$  disaccharide ligand units are *tightly* bound to protein structures<sup>33</sup> supports this suggestion. Cis structures are observed for a lactose mimic (*C*-linked Gal $\beta(1\rightarrow4)$ Glc)<sup>20,30,34</sup>

and inferred for a cellobiose (Glc $\beta(1\rightarrow4)$ Glc) unit<sup>35</sup> when bound in enzyme active sites. These tight-binding proteins would provide sufficient binding energy to pay the entropic cost described above.

In conclusion, the great sensitivity of the O–H modes in the disaccharide to their hydrogen-bonded environment, which has the effect of dispersing its associated infrared spectrum over a wide spectral range, has facilitated its full resolution and enabled a full and unambiguous structural assignment to be accomplished. When isolated in the gas phase, benzyl  $\beta$ -lactoside presents only a single, rigid conformational structure, supported by strong inter-ring hydrogen bonds. Its infrared ion-dip spectrum is in excellent agreement with that of the global minimum structure computed ab initio; the glycosidic torsion angles,  $\phi_H$  (H1–C1–O–C4')  $\approx 180^\circ$  and  $\psi_H$  (C1–O–C4'–H4')  $\approx 0^\circ$ , correspond to a rotation of  $\sim 150^\circ$  about the glycosidic bond compared to the accepted solution-phase conformation. Infrared ion-dip spectroscopy provides a general and valuable structural procedure which will be applicable to a broad range of oligosaccharides in the future and may also be extended to include the structural investigation of their hydrated complexes.

**Acknowledgment.** We greatly appreciate the contributions made by Dr. Ben Davis (Dyson Perrins Laboratory), discussions with Professor Raymond Dwek and Dr Mark Wormald of the Oxford Glycobiology Institute, and the support provided by the EPSRC, the Royal Society (R.A.J., Postdoctoral Research Fellowship; L.C.S, University Research Fellowship), the Leverhulme Trust (Grant No. F/08788D), the CLRC Laser Loan Pool, and the Physical and Theoretical Chemistry Laboratory at Oxford. We gratefully acknowledge the support by the *Stichting voor Fundamenteel Onderzoek der Materie* (FOM) in providing the required beam time on FELIX and highly appreciate the skillful assistance by the FELIX staff.

**Supporting Information Available:** Optimized structures of all 12 calculated low-energy conformations of benzyl  $\beta$ -lactoside, both as Cartesian coordinates and figures, are available as Supporting Information. Tables listing the relative energies of these conformations (including thermochemistry calculations), glycosidic bond angles, and the calculated O–H stretching modes are also available. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA031679K

- (33) The majority of lectin and enterotoxin structures found with lactose as a ligand in the RCMB-PDB show that such proteins tend only to bind the galactose (Gal) “tip” of the lactose structure. In addition, such proteins typically operate with low  $K_D$  ( $\leq 10^{-3}$ ) for such monovalent ligands. As such, they are not representative of binding interactions of the entire lactose disaccharide.
- (34) Although NMR and X-ray crystallography structures of (O-linked) lactose in complex with an *inactive* mutant of the same enzyme show that the lactose is bound as a trans structure, as detailed in refs 31 and 32.
- (35) Alzari, P. M.; Souchon, H.; Dominguez, R. *Structure* **1996**, *4*, 265–275.