Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supporting Information

for

Selectivity Reactivity of Glycosyl Cation Stereoisomers: The Role of Intramolecular Hydrogen Bonding

M. P. Dvores,[a] P. Çarçabal[b] and R. B. Gerber[a,c]

a. Fritz Haber Centre for Molecular Dynamics, The Institute of Chemistry, The Hebrew University, Jerusalem 91904, Israel.

b. Insitut des Sciences Moléculaires d'Orsay, ISMO, Univ Paris-Sud; CNRS, bat 210, Univ Paris-Sud 91405 Orsay Cedex, France.

c. Department of Chemistry, University of California Irvine, CA, 92697, USA.

Table of Contents

Supplementary Figures

S1……..2

S2……..3

S3……..3

S4…….……4

S5……..5

S6……..6

S7……..7

S8……..8

S9……..9

The Structural Origins of the Vibrational Transitions in the Galactosyl and Talosyl Oxocarbenium Ions……..10

Computational and Experimental Protocols

Conformer search protocol……14

AIMD protocol……..15

VSCF protocol……...…16

Experimental protocol……17

Notes on conformational nomenclature…….18

References……. 21

Supplementary Figures

Fig. S1 shows the IRMPD spectra of the *m/z* 195 species from αMeGal-H^+ and βMeGal-H^+. The experimental spectra from the methyl-galactosyl cation anomers resemble one another closely.

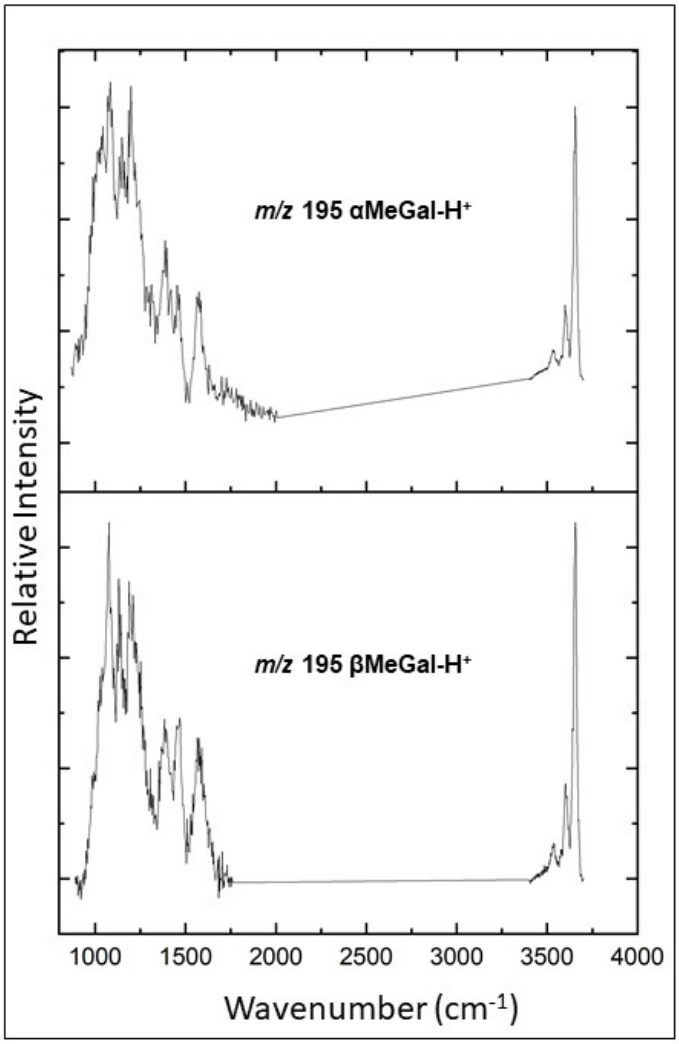
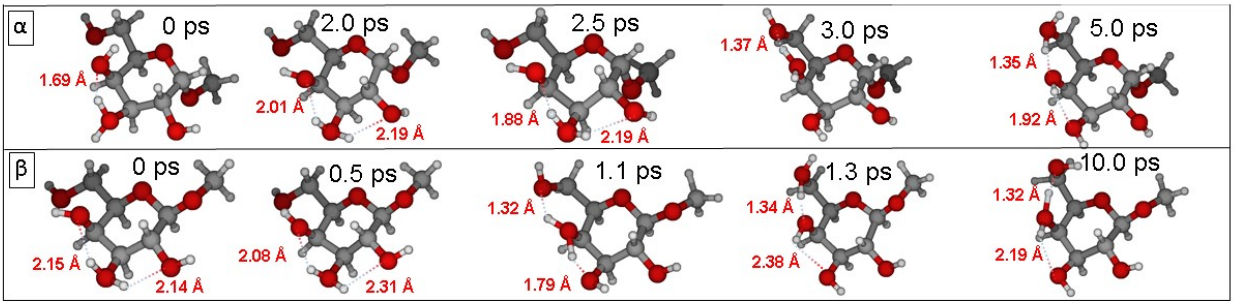


Fig. S2 shows snapshots from representative trajectories initiated from αMeGal-H^+ and βMeGal-H^+ protonated at O3. In both anomers, protonation at O3 led to rapid proton transfer between O3 and O4, to form the lowest energy conformer: the intramolecular proton-bound dimer between O4 and O6. Reverse proton transfer back to O3 is energetically unfavorable and was not observed. No significant differences were observed in the dynamics of the anomers. The lengths



of key hydrogen bonds are noted in the figure.

Fig. S3 shows representative snapshots along reactive trajectories of the lowest energy conformers of αMeGal-H^+ and βMeGal-H^+ with a proton shared between O4 and O6. In both anomers an intramolecular proton-bound dimer was predominant. Proton transfer from O6 to O5 was possible in both anomers and also reversible (not shown). Note: *NVE* ensemble simulations were extended beyond 10 ps to ascertain whether proton transfer was related to anomericity.

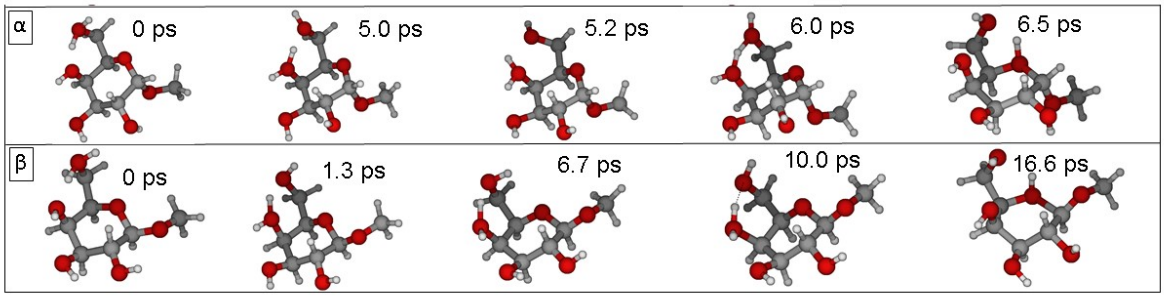


Fig. S4 shows snapshots from representative trajectories initiated from αMeGal-H^+ and βMeGal-H^+ protonated at O2. At the O2 position, which is in a *cis*-disposition to O1 in the α-anomer but has two adjacent *trans*-substituents in the β-anomer, anomericity determines the reactitivy toward proton transfer. In αMeGal-H^+, the proton is transferred to the adjacent cis-substituent at O1. This leads to formation of an oxocarbenium ion-methanol complex as when the initial protonation is at O1. However, in βMeGal-H^+ with O2 uniquely bounded by *trans*-substituents, the excess proton was not transferred in any trajectory. βMeGal-H^+ protonated at O2, therefore, underwent conformer transitions but was unreactive, making it unique among the substituents MeGal-H^+. The lengths of key hydrogen bonds are noted (red) in the figure as well as the C2O2 bond length (black).

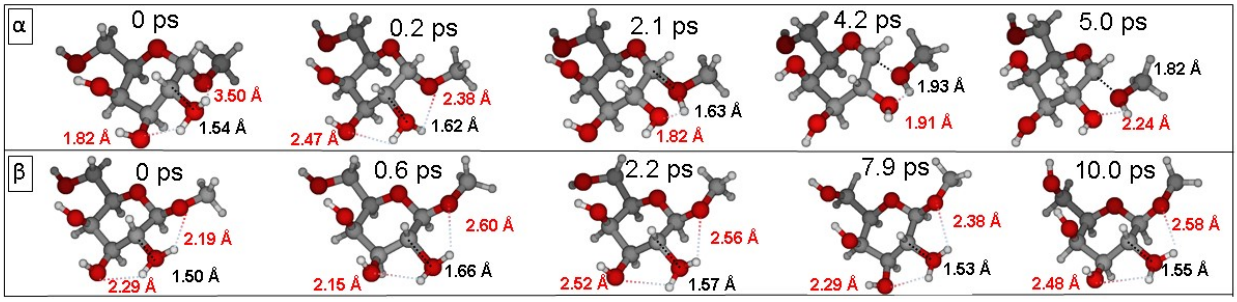


Fig. S5 shows snapshots from representative trajectories initiated from αMeGal-H^+ and βMeGal-H^+ protonated at O5. In both anomers protonation at O5 led to one of two outcomes, slight intermittent ring opening via rupture of the C1O5 bond or proton transfer from O5 to O4, based on the initial conformation. Starting from an extended hydroxymethyl group with a *tg* orientation about the C5C6 bond (see p.6 for an explanation of the conformational nomenclature used in this paper) the distance between C1 and O5 shortened and a *ggTMG* orientation was adopted during the course of the simulation. However, starting the simulation from a *ggTMG* conformation leads to proton transfer from O5 to O6 in some trajectories for both anomers. This demonstrated that in MeGal-H^+ the excess proton is reversibly transferred between O4, O6 and O5 without any observed impact of the anomericity. The lengths of key hydrogen bonds are noted (red) in the figure as well as the C1O5 bond length (black).

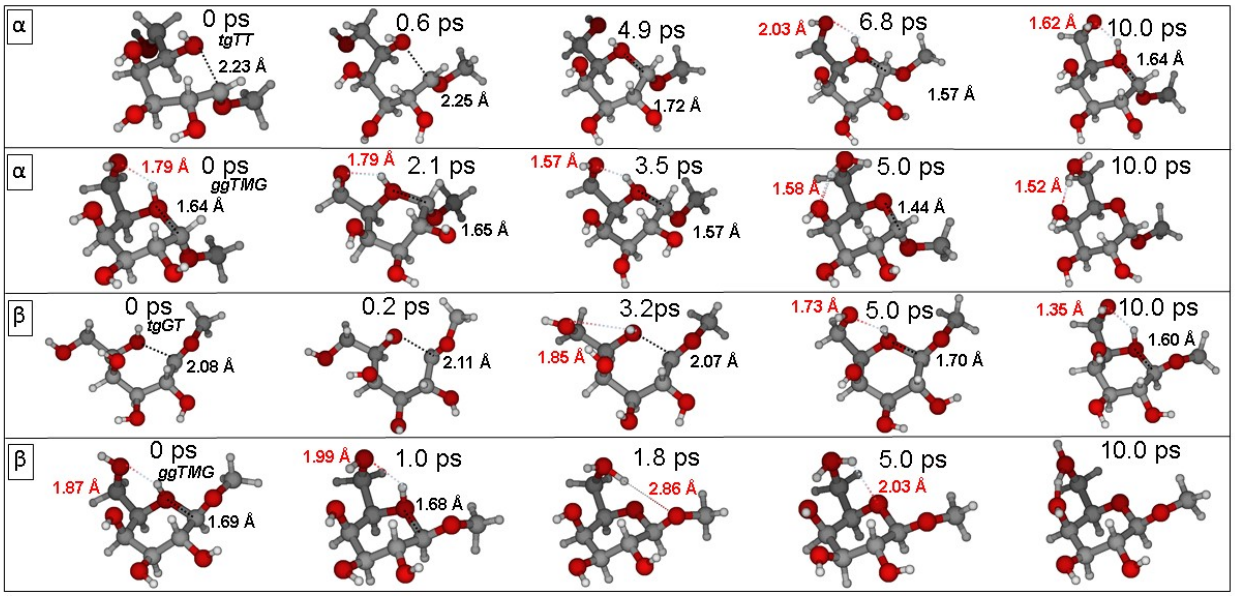


Fig. S6 Representative snapshots along the trajectory initiated from βMeGal-H^+ protonated at O1 showing formation of the intramolecular hydrogen bonded ring between H1 and O4. Here, the detached methanol moiety turns to form a hydrogen bond with O4. This interaction results in reestablishing of the C1O1 bond, followed by reversible proton transfer between O1 and O4. Key intramolecular hydrogen bond lengths are noted in the figure.

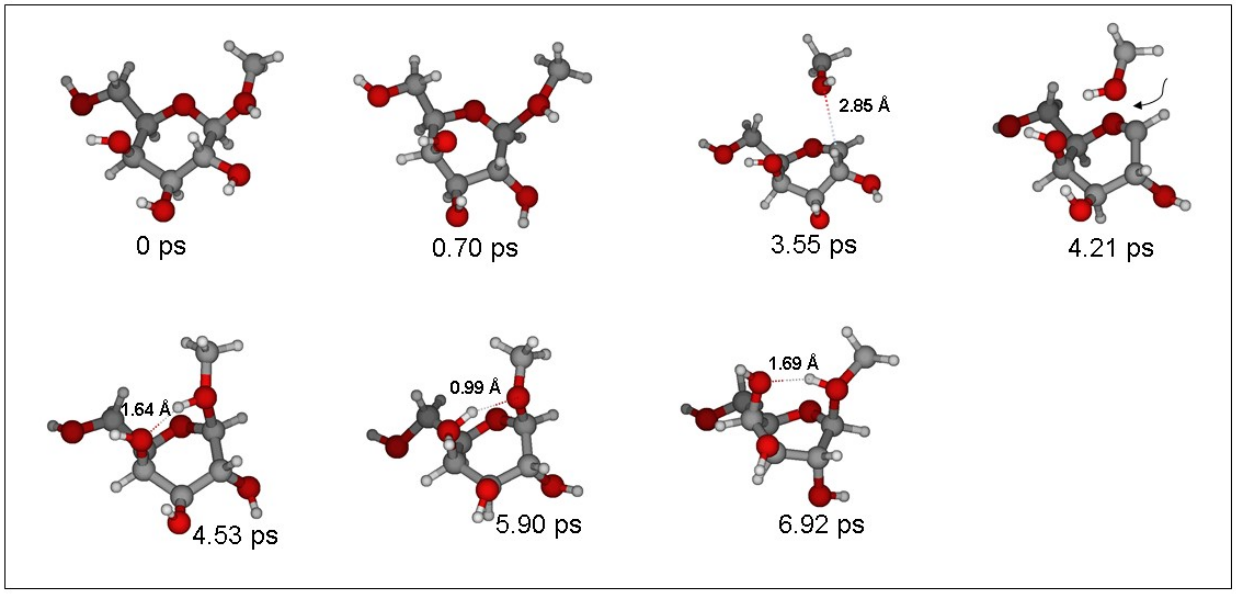


Fig. S7 Representative snapshots along the trajectory initiated from βMeGal-H^+ protonated at O1 showing formation of the intramolecular hydrogen bonded ring between H1 and O6. Here, the detached methanol moiety concertedly reforms the C1O1 bond and forms stabilizing hydrogen bonding interactions with O6. Reversible proton transfer between O1 and O4 was also seen. Key intramolecular hydrogen bond lengths are noted in the figure.

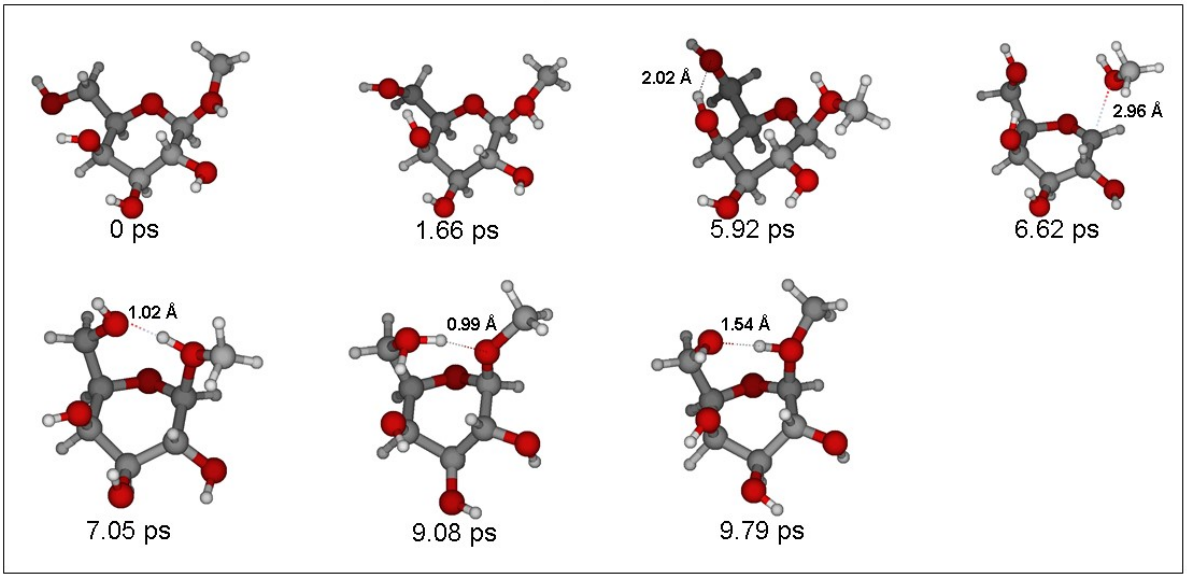


Fig. S8 Optimized structures of βMeGal-H^+ protonated at O1 identified in NVE ensemble simulations. The energies associated with the conformers are relative to the 0.000 eV conformer of βMeGal-H^+. Formation of a seven-membered hydrogen bonded ring had a very stabilizing effect on βMeGal-H^+ protonated at O1, bringing it very close in energy to the minimum energy conformer of βMeGal-H^+ which was an intramolecular proton-bound dimer between O4 and O6. Most trajectories (8/10) initiated from βMeGal-H^+ protonated at O1 resulted in proton transfer to O4 or O6, thereby reversibly removing the excess proton from O1, something unseen in simulations of the α-anomer which is in a trans-disposition to the O4 and O6 substituents.

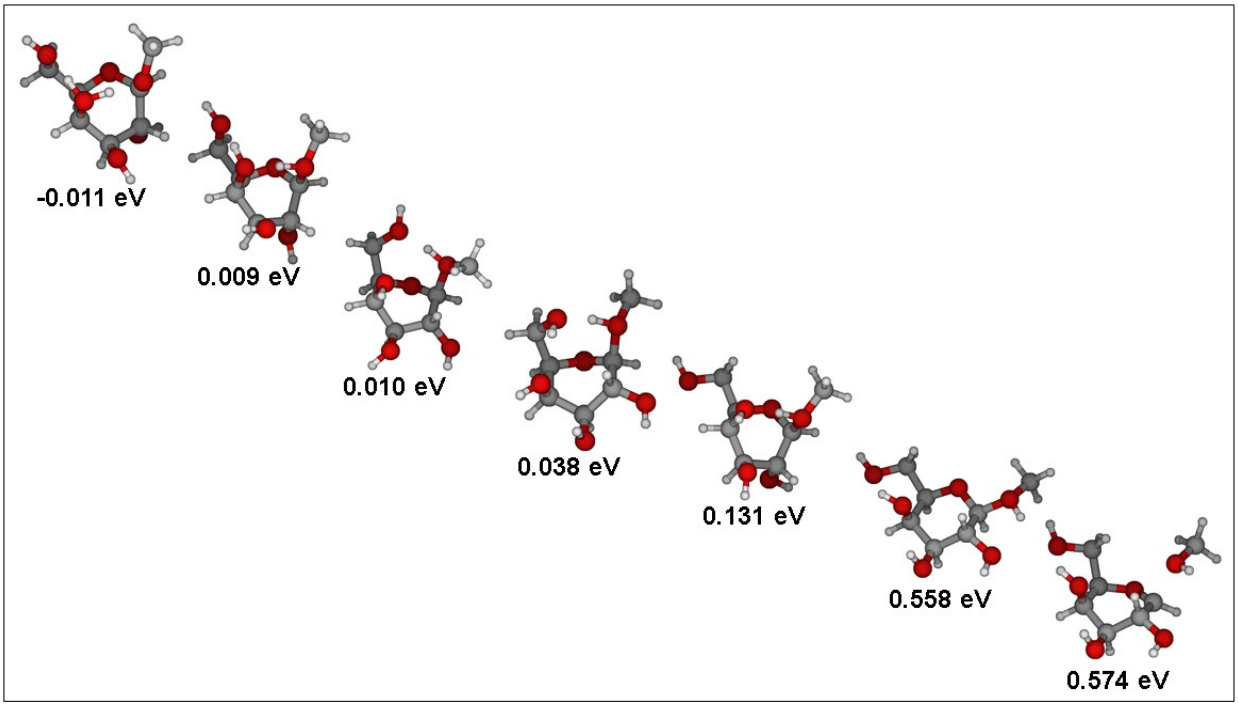
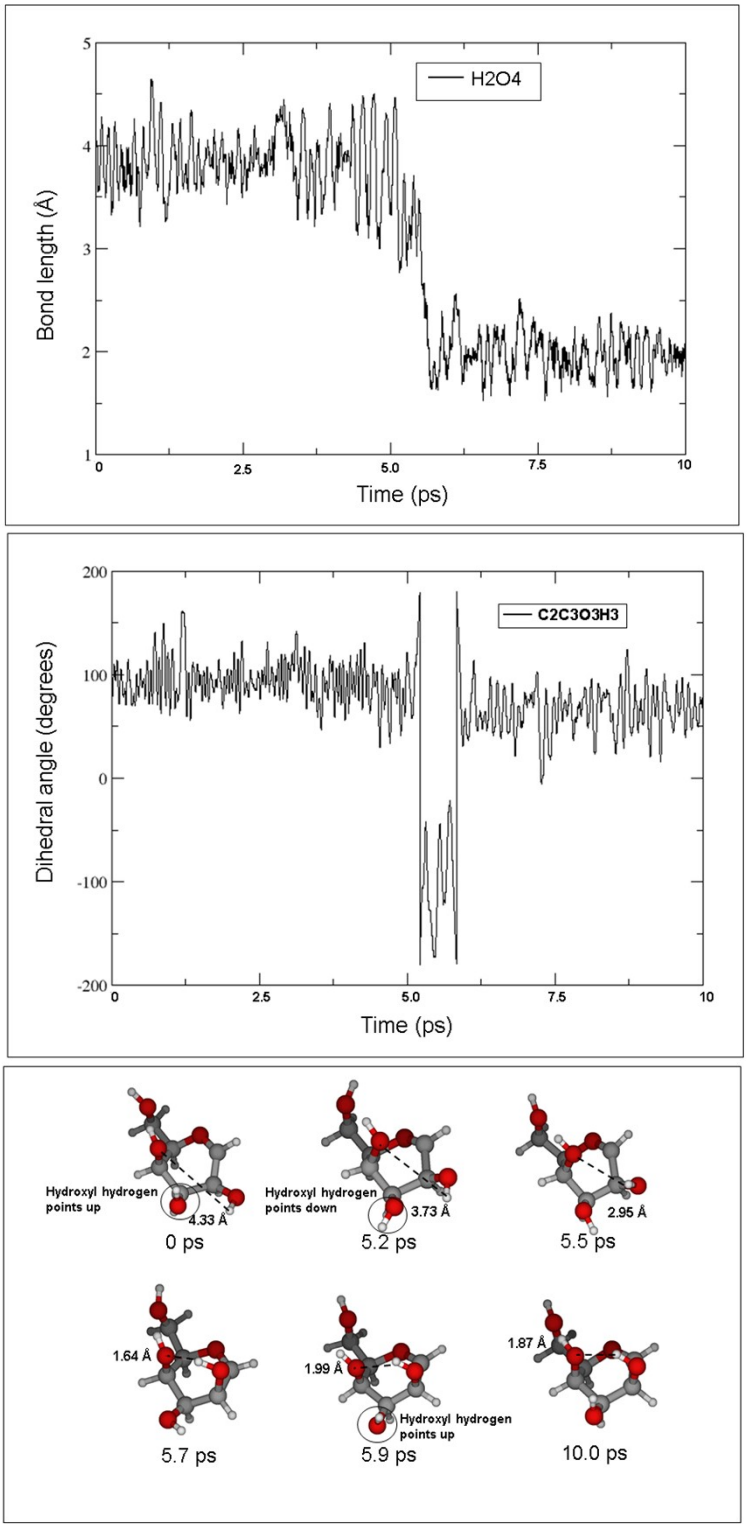


Fig. S9 (top panel) Evolution of the H2O4 hydrogen bond in the talosyl oxocarbenium ion (middle panel) Evolution of the dihedral angle C4C3O3H3, which indicates the orientation of the O3 hydroxyl hydrogen. (bottom panel) Representative snapshots along the trajectory. Changes in the H2O4 hydrogen bond distance and the position of hydrogen in the O3 hydroxyl group are noted. This was the only instance observed of conformer change in the talosyl oxocarbenium ion.



The Structural Origins of the Vibrational Transitions in the Galactosyl and Talosyl Oxocarbenium Ions

The galactosyl oxocarbenium ion has three vibrational transitions (i.e. σ-6, σ-2 and σ-3) in the high wavenumber region which fall reasonable close together to produce the highest intensity transion in this region. The σ-4 transition is red-shifted by ~70 cm^{-1}, as this transition involves the strongest hydrogen bond in the ion (see Fig. S10). Another group of high intensity transtions stands out in the low wavenumber region at 1080-1200 cm^{-1}. Unsuprisingly, these features are all present in the IRMPD spectra, as could be expected in the case of galactose starting material. In the calculated spectrum of the talosyl oxocarbenium ion, however, the high wavenumber transitions are reversed, while high intensity transitions are absent in the low wavenumber region. The former feature is directly related to the stronger hydrogen bonding network in the talosyl epimer. Here, all of transitions except σ-6 undergo strong red-shifts, such that the most intense vibrational transition of the high wavenumber region is red-shifted by ~140 cm^{-1} with repect to the σ-6 transition (see Fig. S11). The structural origin of the high intensity transitions of the low wavenumber region is indeterminate, but as in the spectra of the αMeGal-H^+ and βMeGal-H^+ conformers (see Figs. S12 and S13) this region serves as a highly sensitive fingerprint region in finite-temperature spectra.

Fig. S10 Length of hydrogen bonds in the lowest energy conformers of galactosyl and talosyl oxocarbenium ions.

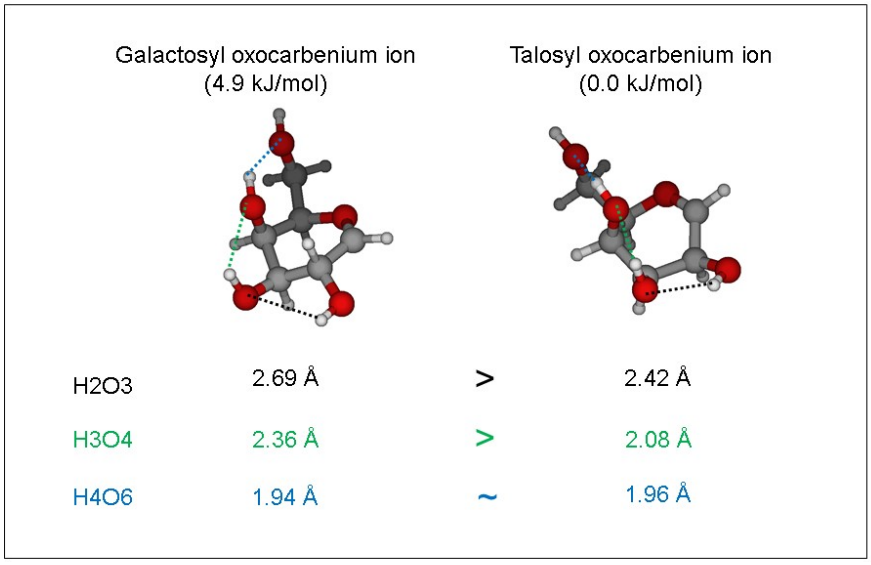


Fig. S11 (top panel) IRMPD from methyl α-(in red) and β-(in blue) D-galactopyranoside compared with the calculated AIMD spectra of the talosyl and galactosyl oxocarbenium ions. (bottom panel) The individual vibrational mode assignments of the AIMD calculated vibrational spectra of the a) talosyl and b) galactosyl oxocarbenium ions. The vibrational modes are distinguished by color.

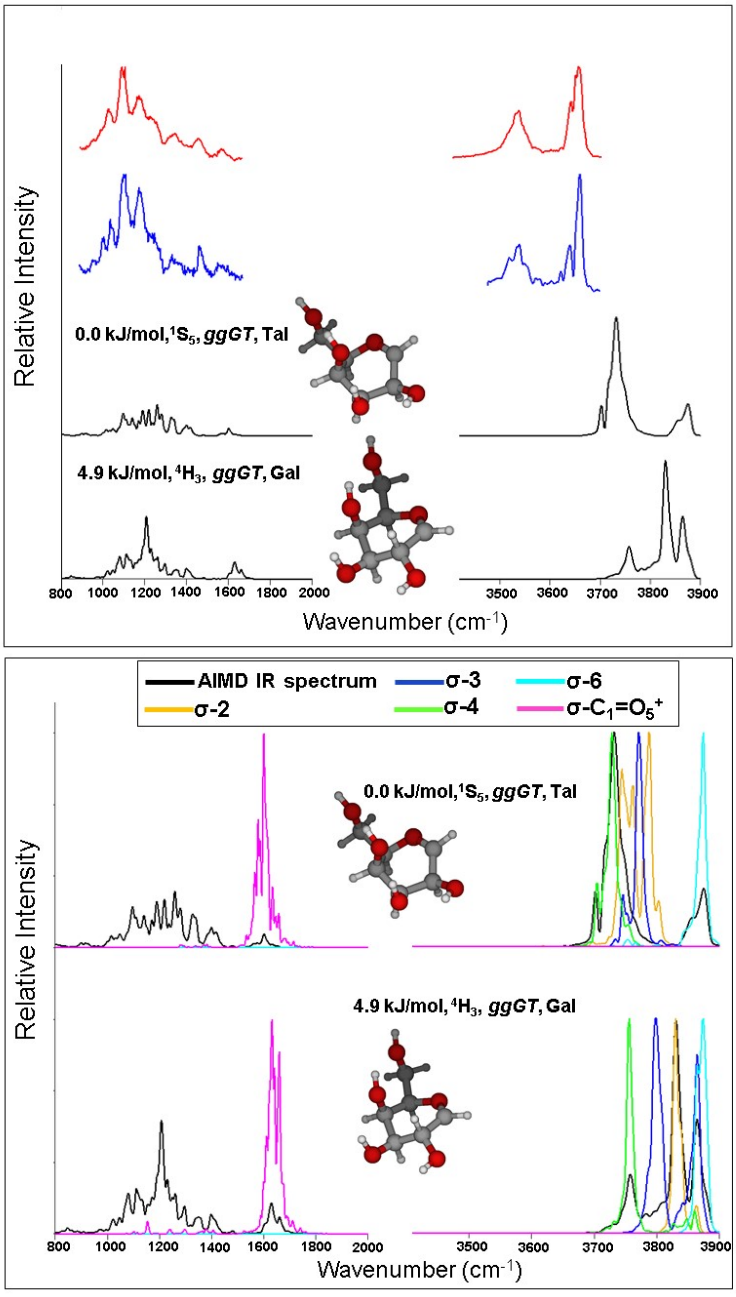


Fig. S12 IRMPD spectra from protonated α-(in red) and β-(in blue) methyl D-galactopyranoside compared with the AIMD calculated spectra of the conformers of αMeGal-H^+ (in order of their relative energies). The protonated oxygen in the conformer is noted.

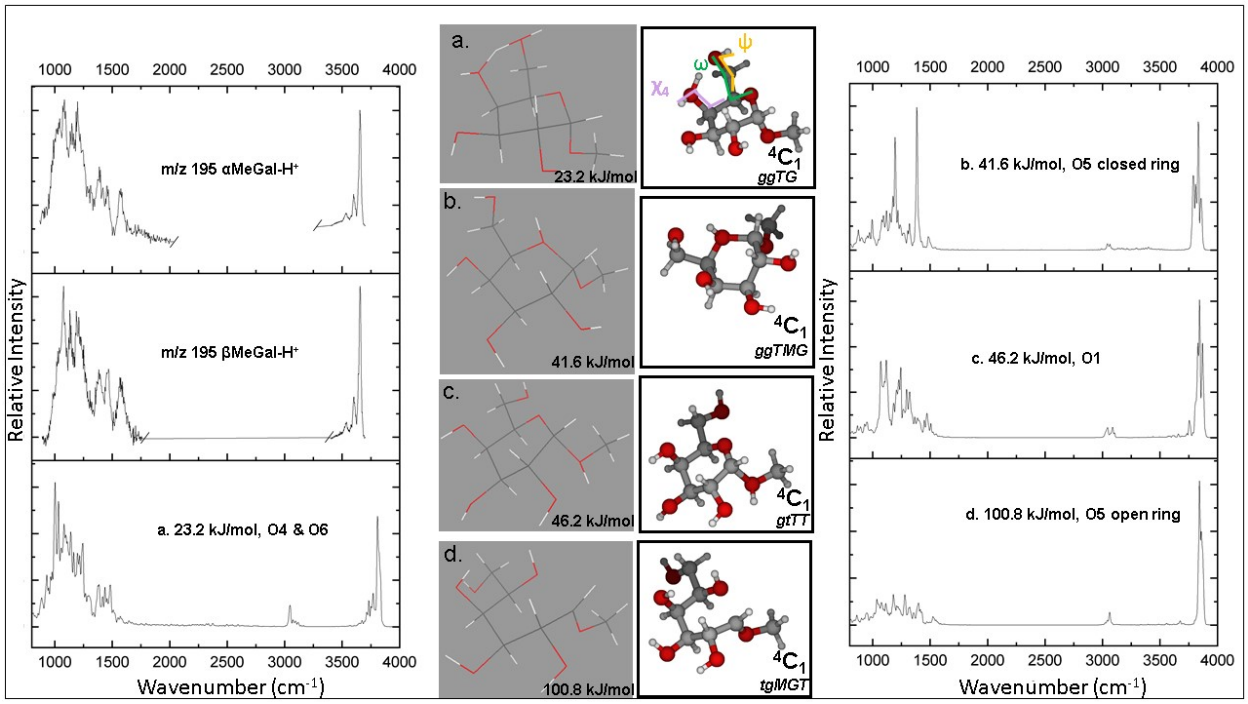
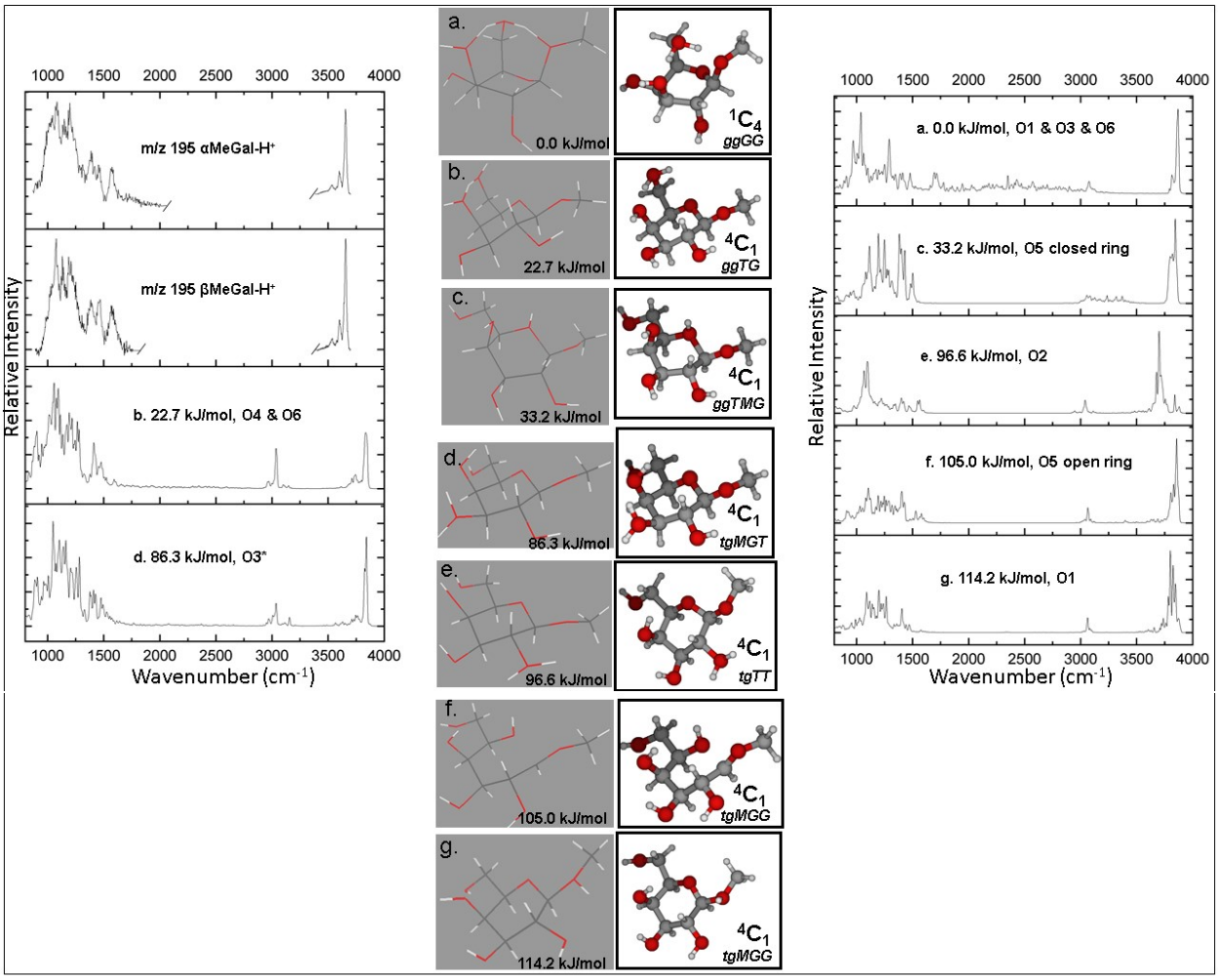


Fig. S13 IRMPD spectra from protonated α-(in red) and β-(in blue) methyl D-galactopyranoside compared with the AIMD calculated spectra of the conformers of βMeGal-H^+ (in order of their relative energies). The protonated oxygen in the conformer is noted.



Computational and Experimental Protocols

Conformer Search Protocol

Low energy conformers of the investigated glycosyl cations were sought by optimizing the neutral glycosides in the proximity of a proton. The result of the optimization would be a minimized glycosyl cation. In order to consider the different possibilities for protonation of the glycoside, optimization was repeated so as to find the optimized protonated structures associated with each of its oxygen atoms. To accomplish this, a proton was placed near (1.9 Å) or on (in the case for which the former failed to yield an optimized glycosyl cation with a novel protonation) an oxygen atom as the starting structure for optimization. Once all optimized structures with all possible protonations were identified, an attempt was made to see if other low energy conformers with this particular protonation could be found. Towards this end, changes in the hydroxymethyl group orientation as well as the orientation of the hydroxyl groups (clockwise or counterclockwise with respect to C1) were made and optimization was begun anew. Optimizations were performed the CP2K suite of programs[1] at the B3LYP-D/6-31+ (d,p) level of theory. The zero-damping empirical dispersion correction of Grimme[2] was used to correct for dispersion interactions in the system. Higher energy conformers quickly decayed into lower energy conformers at the AIMD simulation temperature. Thus, instead of emphasizing an energetic hierarchy of low energy conformers obtained by single-point optimizations, the critical structures for the spectroscopic properties of the glycosyl cations were those identified with an enduring presence in the course of the simulations. These structures were in turn optimized with CP2K at the B3LYP-D/6-31+ (d,p) level of theory and ranked by energy to determine the energetic hierarchy of conformers populated at the simulation temperature.

AIMD Protocol

All AIMD simulations were performed using the CP2K suite of programs. The electronic structure was calculated 'on-the-fly' using the Quickstep module that employs a hybrid Gaussian and plane wave approach to Kohn-Sham density functional theory for *ab initio* Born-Oppenheimer molecular dynamics.[3, 4] A 6-31+G (d,p) basis set was used with an auxiliary plane wave basis set defined by a 350 Ry electron density grid. Goedecker-Teter-Hutter atomic pseudopotentials optimized for the BLYP functional were used to represent the interaction between valence and core electrons.[5] Matrix elements smaller than 10^{-12} Hartree/atom were neglected and the convergence criterion of the SCF wavefunction was set at 10^{-6} Hartree/atom. The default values for the Martyna-Tuckerman Poisson solver[6] were used to establish non-periodic boundary conditions. The motion of the nuclei was evaluated classically using the velocity Verlet algorithm with a 0.5 fs time step.

The simulations commenced with a 5 ps equilibration at ~350 K in an *NVT* ensemble using the Nosé-Hoover Chain thermostat (chain length = 3) to maintain the system temperature. Conformers identified in the equilibration phase were optimized and subject to equilibration under the same conditions so as to form a group of conformers that persist at the experimental temperature of ~350 K. Following equilibration, 10 ps simulations were run for constant energy ensembles at an average temperature of ~350 K for each of these conformers. Only trajectories with an average temperature between 300-385 K were used for spectra calculation so as to maintain a relation of at most ten percent difference between the upper bound of the experimental temperature range (i.e. 350 K) and the conformer population in the simulations. Experience has shown that at high temperatures, T > 300 K, ten picoseconds is an adequate period for exploring the conformational space in system of this size.[7-9] For better sampling of structures in time, ten trajectories or more were run for each conformer.

VSCF Protocol

VSCF calculations of the anharmonic frequencies and intensities of the vibrational transitions of optimized conformer structures at 0 K were performed with the GAMESS program suite.[10] All conformers were optimized with GAMESS prior to VSCF calculations. The gradient convergence tolerance (keyword OPTTOL) was set to 1.0E-7 Hartree/Bohr for optimization and was lowered to 1.0E-5 for the VSCF calculation. The SCF convergence limit was set at 1.0E-6 for optimization and was lowered to 1.0E-5 for the VSCF calculation. The number of grid points used in the VSCF electronic structure calculation was 8. Only couplings between pairs of modes were calculated (keyword ncoup=2). The number of mode-mode anharmonic couplings was adjusted for by examination of the calculated harmonic spectrum of the minimum energy structures. When a prominent higher wavenumber mid-IR transition attributable was present, the number of mode-mode couplings in the calculation was increased so as to include all the modes above ~1500 cm^{-1}. The IMODE keyword was used to select only the higher wavenumber transitions for the VSCF calculation. The B3LYP-D functional was employed along with the 6-31+G (d,p) basis set so as to provide a comparison between the calculated spectra from the AIMD simulation that included contributions from conformer dynamics at ~350 K and line-spectra derived from VSCF calculations of individual conformer structures at 0 K.

Experimental Protocol

Methyl-α-D-galactopyranoside (α-Met-O-Gal) and methyl-β-D-galactopyranoside (β-Met-O-Gal) are available commercially (Sigma-Aldrich) and were used without further purification. The analytes were dissolved in $CH_3OH:H_2O$=1:1 with 0.1% of acetic acid at a concentration of 5.10^{-5} $mol.l^{-1}$ and were sprayed with conventional ESI conditions.

Infrared spectroscopy integrated to mass spectrometry was performed using a 7 Tesla Fourier transform ion cyclotron resonance (FT-ICR) tandem mass spectrometer (Bruker Apex Qe) coupled with IR lasers.[11] The *m/z* 163 and 195 ions of interest were mass selected in the quadrupole and then accumulated in a hexapole ion trap. Ions were then pulse-driven into the ICR cell where they were subjected to IR irradiation.

Infrared spectra in the 900-1800 cm^{-1} and 3200-3700 cm^{-1} spectral ranges were recorded using the free electron laser (IR FEL) CLIO (Orsay, France)[12] and an optical parametric oscillator/amplifier (OPO/OPA from LaserVision pumped by an Innolas Spitlight 600 running at 25Hz), respectively. Details about these experimental setups can be found elsewhere.[11, 13] The IR FEL delivers trains of ~500 picosecond pulses at 25 Hz, and the laser mean power was roughly linearly decreasing from 1000 mW (900 cm^{-1}) to 450 mW (2000 cm^{-1}). One or two attenuators (-3dB each) were necessary to avoid saturation when the IR FEL laser was tuned on resonance with the sugar C-O stretch, in particular, and the irradiation time was set to 500 ms. The typical output energy of the OPO/OPA is ~12-13 mJ/pulse at 3600 cm^{-1}, but rapidly decreases at lower wavenumbers. A significant enhancement of the fragmentation signal could be observed using an auxiliary CO_2 (10watt cw, BFi OPTiLAS, France). As described previously,[14] ions are irradiated with a CO_2 pulse following each OPO/OPA pulse delivered at 25 Hz, the delay being on the order of ~1 μs. Experimental IR spectra correspond to the plot of the fragmentation yield, *i.e.* – ln[P/(P+F)] where P and F are the abundance of the parent and photofragments, as a function of the laser wavelength.

Notes on conformational nomenclature

This work examines the conformational and reactive dynamics of glycosyl cations. As such, it makes frequent reference to saccharide conformational nomenclature. These notes present the conventions adopted by the present work which are used to distinguish between different conformers of glycosyl cation. In this work, the term "conformational" is used in one of its broadest senses to include the saccharide stereochemistry as well as torsional conformation which does not involve the inversion of a stereocenter.

In a cyclic molecule such as a saccharide in its pyranose or furanose forms, a substituent may be located at a small angle to the plane of the ring or the substituent may be found well above or below the plane of the ring. The former, which remain near the plane of the ring, are termed 'equatorial substituents' while the latter which are roughly parallel to the plane of the ring are termed 'axial substituents'. Distortion of the ring is one of the possible conformational changes that can take place in a cyclic system. These distortions introduce changes in the angles at which the substituents meet the plane of the ring, thus altering the axial and equatorial substituents to different degrees. Cremer and Pople[15] introduced a quantitative measure of conformational change to an N-membered ring system known as the 'puckering coordinates'. Cremer-Pople puckering coordinates define a mean ring plane using the cartesian coordinates of all ring atoms from which the atomic elevations Z_j can be computed for a six-membered ring as:

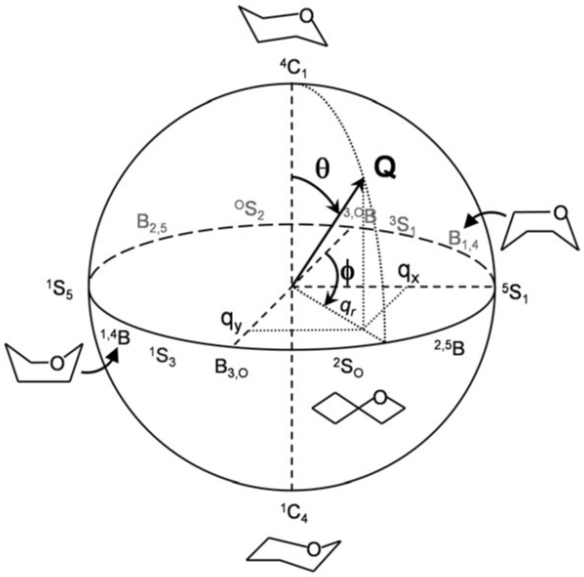
$$Z_j = \frac{1}{\sqrt{3}} q_2 \cos\left[\phi_2 + \frac{2\pi(j-i)}{3}\right] + \frac{1}{\sqrt{6}} q_3 (-1)^{j-1}, \qquad \text{(S.1)}$$

where q_2, ϕ_2, and q_3 are the puckering parameters. These parameters are computed by solving the following system of equations in which they have been transformed into the spherical

$$\begin{cases} q_2 \cos\phi_2 = Q \sin\theta \cos\phi \\ q_2 \sin\phi_2 = Q \sin\theta \sin\phi \\ q_3 = Q \cos\theta \end{cases}, \qquad \text{(S.2)}$$

coordinates (Q, θ, ϕ):

where Q is the radial coordinate known as the total puckering amplitude, is the sum of the perpendicular elevations Z_j from the mean ring plane. Thus, the Cremer-Pople puckering



coordinates of a six-membered ring are calculated as a point on a sphere illustrated in Fig. S.10, where there are 38 canonical ring puckering conformations for a six-membered ring.

Fig. S10 Cremer-Pople puckering coordinates of a six-membered ring (Q, θ, ϕ) taken from ref. [16]

A Mercator representation, shown in Fig. A.3, is an equidistant cylindrical projection of the polar coordinates that is useful for looking at the relationships between the canonical puckering conformations.[17] Intermediate puckering conformations that fall in between two canonical conformations such as $B_{3,o}/^2S_o$are also possible.

Thus far, the conformation of the ring form of the saccharide and its influence on stereochemistry have been discussed. However, many important monosaccharides such as the aldohexoses including glucose and galactose also have an exocyclic C5-C6 linkage, known as the hydroxymethyl group. The hydroxymethyl group of a saccharide will usually be in one of three staggered conformations: gauche–gauche, *gg*; trans–gauche, *tg*; and gauche–trans, *gt*.[18] The conformation of the hydroxylmethyl group has been classified by its orientation about the bonds C4-O4, C5-C6, and C6-O6 in the saccharide. Three torsion angles describe this orientation: χ_4 = (C5-C4-O4-H4), ω = (O5-C5-C6-O6), and ψ = (C5-C6-O6-H6).[18] The conformational designations are based on the approximate values of these three torsion angles as specified in Table S.1.

Table S.1 Conformational designations of the hydroxymethyl group based on approximate torsion angle value.

	~60°	~180°	~-60°
χ_4	*G*[a]	*T*[b]	*MG*[c]
ψ	*G*	*T*	*MG*
ω	*gt*	*tg*	*gg*

[a] synclinal, gauche, 60°

[b] antiperiplanar, trans, 180°

[c] –synclinal, -gauche, -60°

Therefore, a hydroxymethyl conformation will be indicated by designations such as *ggGT* for example which means that ω is approximately in a gauche-gauche conformation while χ_4 is in an approximately gauche conformation and ψ is approximately trans.

[

References

1] T.D. Kühne, M. Iannuzzi, M. Del Ben, V.V. Rybkin, P. Seewald, F. Stein, T. Laino, R.Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M.H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A.S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G.K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C.J. Mundy, C. Plessl, M. Watkins, J. VandeVondele, M. Krack, J. Hutter, CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations, The Journal of Chemical Physics 152(19) (2020) 194103.

[2] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J Chem Phys 132(15) (2010) 154104.

[3] P.Hohenberg, W. Kohn, Inhomogeneous Electron Gas, Phys Rev 136 (1964) B864.

[4] W. Kohn, L.J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, Phys Rev 140(4A) (1965) 1133-&.

[5] S. Goedecker, M. Teter, J. Hutter, Separable dual-space Gaussian pseudopotentials, Phys Rev B 54(3) (1996) 1703-1710.

[6] G.J. Martyna, M.E. Tuckerman, A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters, J Chem Phys 110(6) (1999) 2810-2821.

[7] S. Rudic, H.B. Xie, R.B. Gerber, J.P. Simons, Protonated sugars: vibrational spectroscopy and conformational structure of protonated O-methyl alpha-D-galactopyranoside, Mol Phys 110(15-16) (2012) 1609-1615.

[8] H.B. Xie, R.B. Gerber, Interaction and reaction of the hydroxyl ion with beta-D-galactose and its hydrated complex: an ab initio molecular dynamics study, Phys Chem Chem Phys 14(35) (2012) 12086-12089.

[9] H.B. Xie, L. Jin, S. Rudic, J.P. Simons, R.B. Gerber, Computational Studies of Protonated beta-D-Galactose and Its Hydrated Complex: Structures, Interactions, Proton Transfer Dynamics, and Spectroscopy, J Phys Chem B 116(16) (2012) 4851-4859.

[10] M.W. Schmidt, K.K. Baldridge, J.A. Boatz, S.T. Elbert, M.S. Gordon, J.H. Jensen, S. Koseki, N. Matsunaga, K.A. Nguyen, S.J. Su, T.L. Windus, M. Dupuis, J.A. Montgomery, General Atomic and Molecular Electronic-Structure System, J Comput Chem 14(11) (1993) 1347-1363.

[11] J.M. Bakker, T. Besson, J. Lemaire, D. Scuderi, P. Maitre, Gas-phase structure of a π-allyl-palladium complex: Efficient infrared Spectroscopy in a 7 T Fourier transform mass spectrometer, J. Phys. Chem. A 111(51) (2007) 13415-13424.

[12] R. Prazeres, F. Glotin, C. Insa, D.A. Jaroszynski, J.M. Ortega, Two-colour operation of a Free-Electron Laser and applications in the mid-infrared, Eur. Phys. J. D 3(1) (1998) 87-93.

[13] J.M. Bakker, R.K. Sinha, T. Besson, M. Brugnara, P. Tosi, J.Y. Salpin, P. Maitre, Tautomerism of Uracil Probed via Infrared Spectroscopy of Singly Hydrated Protonated Uracil, J. Phys. Chem. A 112(48) (2008) 12393-12400.

[14] R.K. Sinha, E. Nicol, V. Steinmetz, P. Maitre, Gas Phase Structure of Micro-Hydrated $[Mn(ClO_4)]^+$ and $[Mn_2(ClO_4)_3]^+$ Ions Probed by Infrared Spectroscopy, J. Am. Soc. Mass Spectrom. 21(5) (2010) 758-772.

[15] D. Cremer, J.A. Pople, General Definition of Ring Puckering Coordinates, J Am Chem Soc 97(6) (1975) 1354-1358.

[16] J. Iglesias-Fernandez, L. Raich, A. Ardevol, C. Rovira, The complete conformational free energy landscape of beta-xylose reveals a two-fold catalytic itinerary for beta-xylanases, Chem Sci 6(2) (2015) 1167-1177.
[17] M.K. Dowd, A.D. French, P.J. Reilly, Modeling of aldopyranosyl ring puckering with MM3 (92), Carbohydrate Research 264(1) (1994) 1-19.
[18] I. Tvaroska, F.R. Taravel, J.P. Utille, J.P. Carver, Quantum mechanical and NMR spectroscopy studies on the conformations of the hydroxymethyl and methoxymethyl groups in aldohexosides, Carbohydrate Research 337(4) (2002) 353-367.