JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

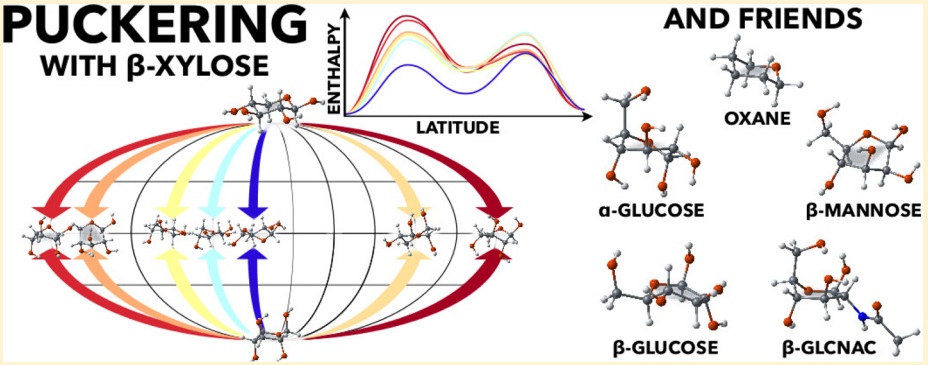
pubs.acs.org/JACS

How Sugars Pucker: Electronic Structure Calculations Map the Kinetic Landscape of Five Biologically Paramount Monosaccharides and Their Implications for Enzymatic Catalysis

Heather B. Mayes,[†,‡] Linda J. Broadbelt,[†] and Gregg T. Beckham*[,‡]

[†]Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, United States
[‡]National Bioenergy Center, National Renewable Energy Laboratory, Golden, Colorado 80401, United States

Supporting Information



ABSTRACT: Glycoside hydrolases (GHs) distort carbohydrate ring geometry along particular "catalytic itineraries" during the cleavage of glycosidic bonds, illustrating the relationship between substrate conformation and reactivity. Previous theoretical studies of thermodynamics of isolated monosaccharides offer insights into the catalytic itineraries of particular sugars. However, kinetic accessibility of carbohydrate puckering conformations and the role of exocyclic groups have not yet been thoroughly addressed. Here we present the first complete library of low-energy local minima and puckering interconversion transition states for five biologically relevant pyranose sugars: *β*-xylose, *β*-mannose, *α*-glucose, *β*-glucose, and *β*-*N*-acetylglucosamine. These were obtained by a thorough theoretical investigation each of the 38 IUPAC designated puckering geometries and all possible conformations of the exocyclic groups. These calculations demonstrate that exocyclic groups must be explicitly considered when examining these interconversion pathways. Furthermore, these data enable evaluation of previous hypotheses of why enzymes perturb ring geometries from the low-energy equatorial chair ($^4C_1$) conformation. They show that the relative thermodynamics alone do not universally correlate with GH catalytic itineraries. For some sugars, particular puckers offer both catalytically favorable electronic structure properties, such as anomeric carbon partial charge, and low kinetic barriers to achieve a given puckering conformation. However, different factors correlate with catalytic itineraries for other sugars; for *β*-*N*-acetylglucosamine, the key N-acetyl arm confounds the puckering landscape and appears to be the crucial factor. Overall, this study reveals a more comprehensive understanding of why particular puckering geometries are favored in carbohydrate catalysis concomitant with the complexity of glycobiology.

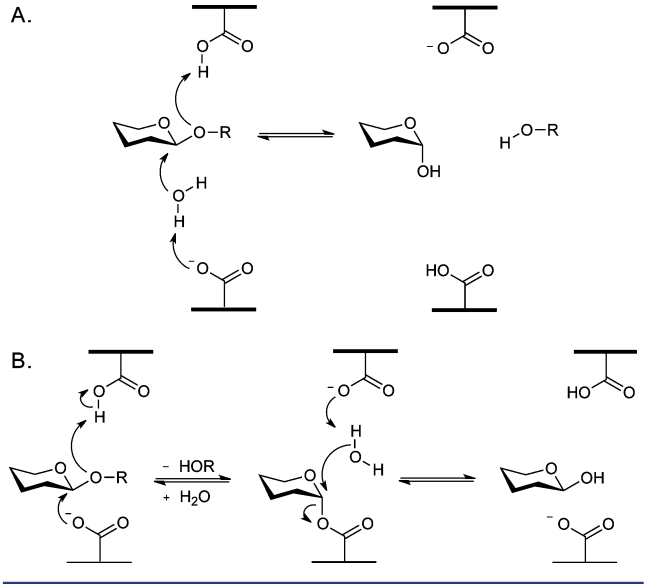
INTRODUCTION

Carbohydrates are the most abundant and diverse set of biological molecules on Earth and are responsible for a vast array of vital biological functions including structure, energy storage, and signaling.[1−5] Glycoside hydrolase (GH)[6,7] enzymes, which depolymerize carbohydrates typically via inverting or retaining hydrolytic mechanisms (Scheme 1), represent 132 known enzyme families with diversity in protein fold concomitant with the diversity present in their substrates.[8] As GHs are important in cell biology, glycobiology, human health, carbon and nitrogen cycling on Earth, and renewable energy, understanding the molecular-level mechanisms employed by GH enzymes is of paramount importance. Extensive crystallographic studies with transition state (TS) analogues have revealed that GH enzymes distort their carbohydrate substrates in the −1 position for catalysis. This distortion of the pyranose ring in "activated" complexes of enzymes is a general feature of GH action[7] and has been documented since the first X-ray studies of a GH on hen egg-white lysozyme.[9] Puckered ring conformations are classified according to nomenclature described by Schwartz[10] and adopted by the International Union of Pure and Applied Chemistry (IUPAC).[11] The 38 canonical pyranose puckering conformations represent all unique ways in which a pyranose ring can be in a chair (C), envelope (E, previously "sofa"), half-chair (H, sometimes called

Received: October 6, 2013
Published: December 24, 2013

   dx.doi.org/10.1021/ja410264d | *J. Am. Chem. Soc.* 2014, 136, 1008−1022

Scheme 1. Koshland's Proposed Inverting (A) and Retaining (B) Mechanisms for GH Enzymes[6]



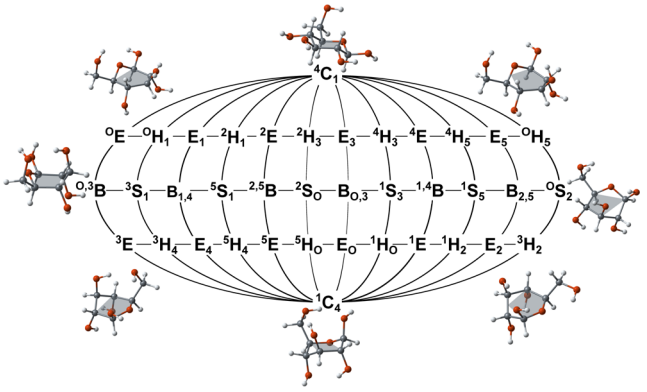


Figure 1. The two-dimensional projection of the CP sphere shows the 38 canonical puckering designations. The letter designates the type of pucker (chair, half-chair, envelope, skew, or boat) and the conformations on the outer ring are illustrated using β-glucose.

half-boat or twist), skew (S, sometimes called twist-boat), and boat (B), as illustrated in Figure 1. Cremer and Pople[12] defined a spherical coordinate system that uniquely defines the puckering conformation for general monocyclic rings, allowing the conformations to be mapped on a sphere (Figure 1). The spherical coordinates of a puckering conformation are referred to as Cremer–Pople (CP) puckering parameters and allow for a precise description of the geometry of pyranose rings.[12] This precise method to describe puckering geometry is especially important because the stable puckering geometries are often intermediate between the IUPAC designated puckers.[13,14] We refer to specific carbon and oxygen atoms by their IUPAC numbering designation.[15]

Why carbohydrate-active enzymes distort substrate ring geometries is the subject of active investigation, and a particular example of the broadly applicable phenomena of how conformation affects reactivity.[16−18] One proposal is that certain puckered geometries align the anomeric substituent leaving group in an axial position posed for the nucleophilic attack,[7,19] rather than the equatorial alignment afforded by the stable chair conformation shown in Scheme 1. It has also been proposed that enzymes distort carbohydrate rings to impart an oxocarbenium-like character to the anomeric carbon (C1) to make it a better target for nucleophilic attack.[7,20−23] Computational analysis of the conformations of polysaccharide monomers in the different puckering conformations can provide evidence of whether perturbing the structure and thus the electron density distribution is sufficient to increase the positive partial charge on the anomeric carbon. Some notable studies include Biarnés et al.'s metadynamics study wherein they mapped the thermodynamic landscape of the northern hemisphere of the CP sphere for β-glucose[24] and similar studies for β-mannose[25] and α-L-fucose.[26] Their chosen collective variables (CVs) sample of the CP coordinates for the northern hemisphere. However, as noted in their work,[24] selection of the CP parameters as the CVs over which to sample does not ensure sampling of the full potential energy surface of the molecules due to the presence of exocyclic groups which, as we show in the present work, play a pivotal role in determining barriers between the different puckering geometries and significantly impact the kinetic landscapes. Furthermore, the metadynamics method does not provide accurate kinetic data if the chosen CVs do not reflect the appropriate reaction coordinates for describing the kinetic landscape.[27−30] Sega et al. have shown that the Cartesian-based CVs chosen in these metadynamics studies lead to systematic overestimation of barriers and sampling of conformations with unphysical total puckering amplitude, with the errors greatest at the most relevant part of the CP sphere for puckering studies, the equator.[31] Nonetheless, for the monosaccharides considered in these studies,[24−26] the authors propose a correlation between experimentally observed puckered orientations and computationally determined lower-free-energy puckered orientations of isolated monosaccharides that also exhibit some oxocarbenium-like character.

Other studies which provide thermodynamic information about monosaccharide puckering geometries include: a study by Autieri et al.[32] using CVs recommended by Sega et al. with the GROMOS 45A4 molecular mechanics force field; work by Sattelle et al. using GLYCAM force fields to explore a range of monosaccharides;[33−36] and by Barnett and Naidoo,[13,14] who employed the flat histogram free energy method (FEARCF)[37,38] to ensure adequate sampling of all parts of the CP sphere with the PM3 semiempirical method. Each of these studies reveals interesting features of monosaccharide puckering landscapes. However, as Barnett and Naidoo reported,[14] important features of these landscapes cannot be accurately resolved by molecular mechanics force fields or semi empirical methods, resulting in incorrect identification of structures and/or inaccurate relative energies. Some kinetic pathways for puckering interconversion have been revealed with DFT and ab initio methods.[39−41] These studies confirm the importance of ring flexibility in describing monosaccharides and offer significant insight into several key ring interconversion reaction paths, but a full kinetic characterization of interconversion between all stable puckering conformations for the sugars in the present work does not yet exist.

Here, we present a comprehensive puckering study of five monosaccharides, including all parts of the CP sphere as well as the degrees of freedom due to rotation of exocyclic groups. Additionally, we map the interconversion pathways between all the different puckering geometries by fully exploring the kinetic landscape to provide insights about monosacchiride ring puckering pathways. The five pyranose sugars studied are: β-

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

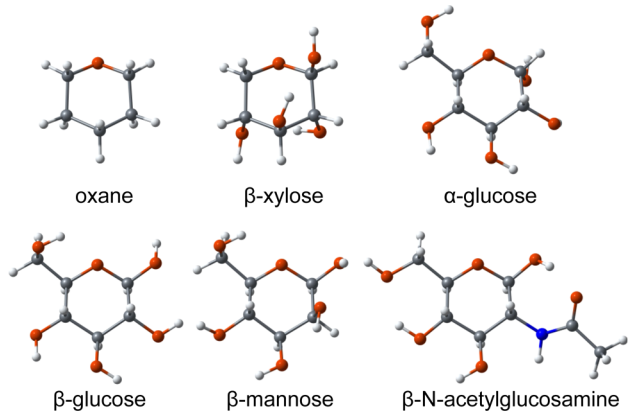


Figure 2. The six molecules included in our study of pyranose ring puckering.

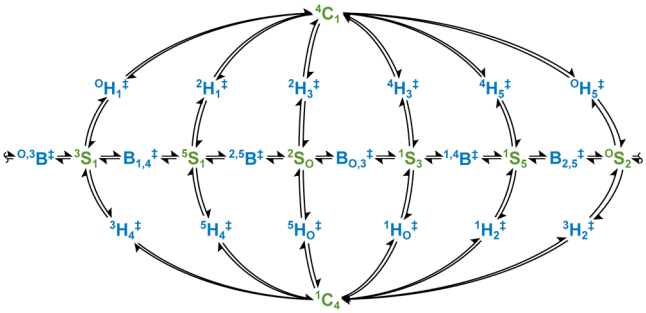


Figure 3. Pyranose puckering interconversion pathways for cyclohexane puckering geometries proposed by Stoddart.[51]

xylose, α-glucose, β-glucose, β-mannose, and β-*N*-acetylglucosamine (GlcNAc) (Figures 2 and 3). Glucose, mannose, and GlcNAc are all important in glycosylation.[4,42,43] β-Glucose is the monomer of cellulose, the main component of plant cell walls.[44] β-Xylose is the main component of hemicellulose, another significant plant component.[45] Note that β-xylose is identical to β-glucose, except for the absence of a hydroxymethyl group at C5. β-Mannose subunits can be another important component of hemicelluloses,[45] and β-mannose is also an epimer of β-glucose, with only a difference in stereochemistry at C2, providing an interesting comparison between the molecules. α-Glucose is another epimer of β-glucose and is important as α-glucosidases play roles in human diseases.[46−48] GlcNAc is the monomer of chitin, an important polymer in the exoskeleton of arthropods and cell walls of fungi, algae, and yeast.[49] Its decomposition is key to carbon and nitrogen recycling.[50] Furthermore, *O*-GlcNAc is one of the most important post-translational modifications in cell biology, with functional implications on par with protein phosphorylation.[4]

Comparing the results for the five sugars to each other and to oxane (tetrahydropyran), a minimal pyranose ring, reveals the crucial role of exocyclic groups in stabilizing puckering conformations; a different arrangement at just one chiral center greatly changes the puckering interconversion landscape, displaying large deviations from the ring geometry interconversion mechanisms for cyclohexane proposed by Stoddart.[51] The high-level electronic structure information for this series of molecules offers additional insight into why enzymes follow certain "catalytic itineraries",[7] and provides a comprehensive foundation for future computational and structural studies with enzymes and chemical catalysts present.

COMPUTATIONAL METHODS

Starting geometries for ring-puckering conformations were generated using the 6ring program of Bérces et al.[52] The exocyclic groups were added using CHARMM.[53] For each molecule in each of the 38 IUPAC puckering configurations, we generated a set of rotamers by rotating each exocyclic group in 120° increments around each C−C−O−H or C−C−C−O dihedral angle and in 180° increments around the C−N−C═O and C−C−N−C dihedral angles in GlcNAc. There are no exocylic groups on oxane, so there is only one conformation per ring pucker. For xylose, 81 rotamers were optimized per puckering conformation for a total of 3078 conformations. For α-glucose, β-glucose, and β-mannose, 729 rotamers were generated for each pucker, resulting in 27 702 conformations for each of these six-carbon monosaccharides. For GlcNAc, 972 conformations per ring pucker were generated for a total of 36 936 geometries. Thus, the initial set of conformations considered for all molecules in this study comprised 123 201 geometries.

Given the large number of structures studied, a hierarchical approach was used to identify a set of low-energy structures for each sugar, which were then analyzed using a higher level of theory. First, the initial structures were optimized using Gaussian 09[54] with the M06-2X/6-31G(d) level of theory keeping the ring puckering dihedral angles fixed. The M06-2X[55] functional is a newer iteration of the M05-2X[56] functional that has been shown to provide accurate energies in carbohydrate QM studies.[57,58] Pople-style basis sets[59,60] are used, including polarization functions.[61,62] We performed all calculations in vacuum, as the system of interest is the protein interior, which has a low dielectric constant on the order of 4.[63]

The next stage of screening involved increasing the size of the basis set and performing fully relaxed optimizations and frequency calculations at the M06-2X/6-31+(d,p) level of theory on unique conformations within 5 kcal/mol of the lowest electronic energy structure for a given pucker, both to find local minima and TS structures. We included diffuse functions[64] for the geometry optimization and frequency calculations as they can be particularly important in systems with hydrogen bonding.[65] Local minima were verified to have no imaginary frequencies and each TS exactly one imaginary frequency. We assigned the resulting structures to the closest IUPAC configuration on the CP sphere based on arc length. For the CP coordinates of the canonical conformations, we used those listed by Hill and Reilly[66] for a pyranose ring. After this step, we discarded any conformations with Gibbs free energies (at 298 K) greater than 5 kcal/mol of the lowest-energy conformation with the same pucker and same number of imaginary frequencies.

The final screen involved reoptimizing this set of low-energy structures using a method shown to obtain highly accurate geometries and frequencies[67−69] and calculating the energy at a high level of theory. Specifically, geometries and vibrational frequencies were obtained at the B3LYP/6-311+G(2df,p) level of theory using an ultrafine integration grid and tight convergence. Electronic energies for these low-energy conformations were calculated at the CCSD(T)[70−73]/6-311+G(d,p) level of theory. Boltzmann-weighted averaging was employed to obtain aggregate properties of puckered conformations from individual conformations with exocyclic groups in different orientations. To calculate entropies, enthalpies, and free energies for individual conformations, standard statistical mechanics formulas were applied.[74] Corrections to the partition function for internal rotation, such as discussed by Pfaendtner et al.,[75] were not included since accurate treatment of internal modes of rotation is impeded by coupled motion and by large contributions to the low frequencies from ring puckering. Thus, the frequencies were calculated based on the harmonic oscillator (HO) assumption. The effect of low frequencies in this study is mediated by the need only to compare different conformations of the same molecule to each other, leading to a large degree of cancellation of contributions from low frequencies. Furthermore, the low frequencies affect the entropy contribution to

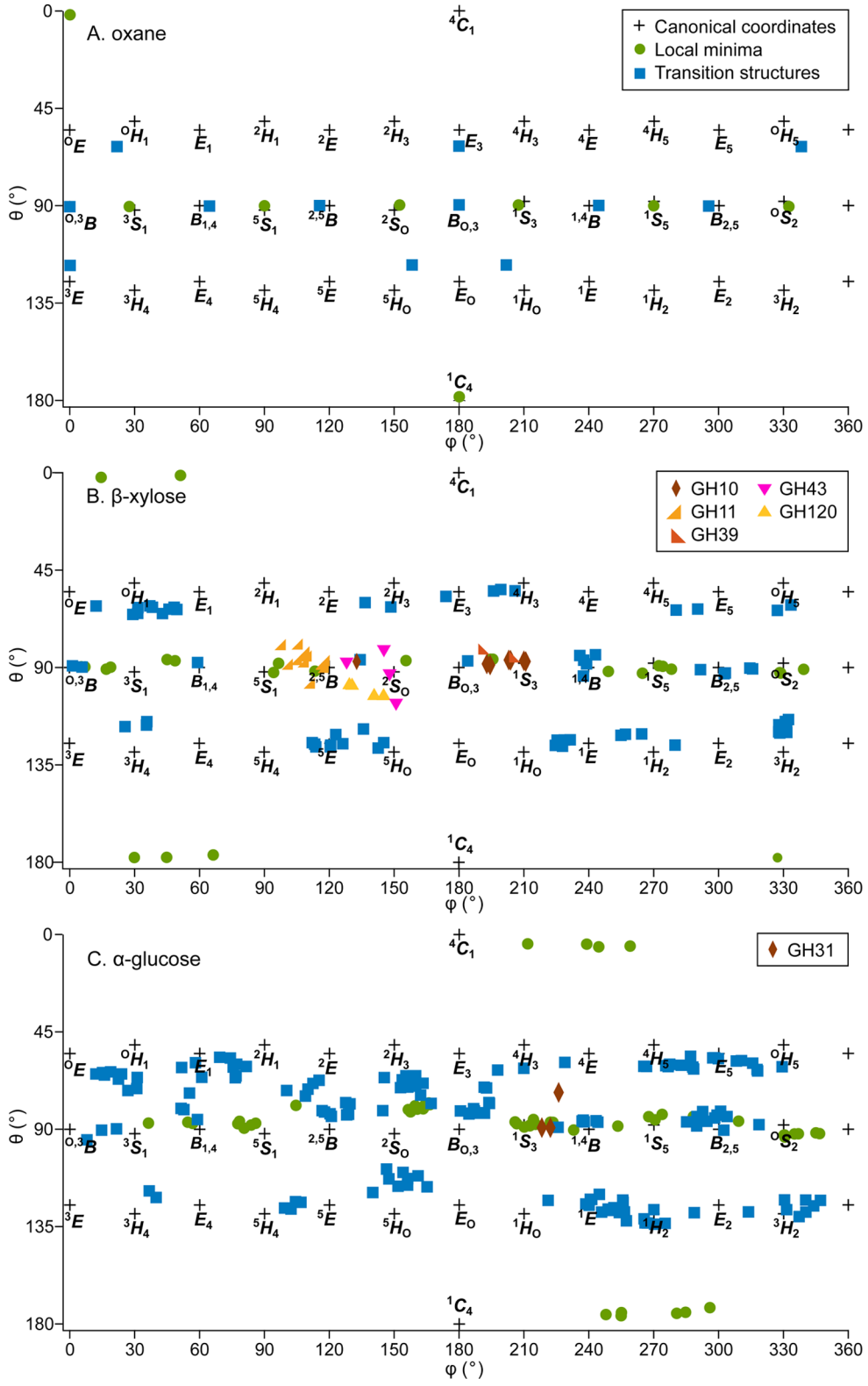


Figure 4. continued

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

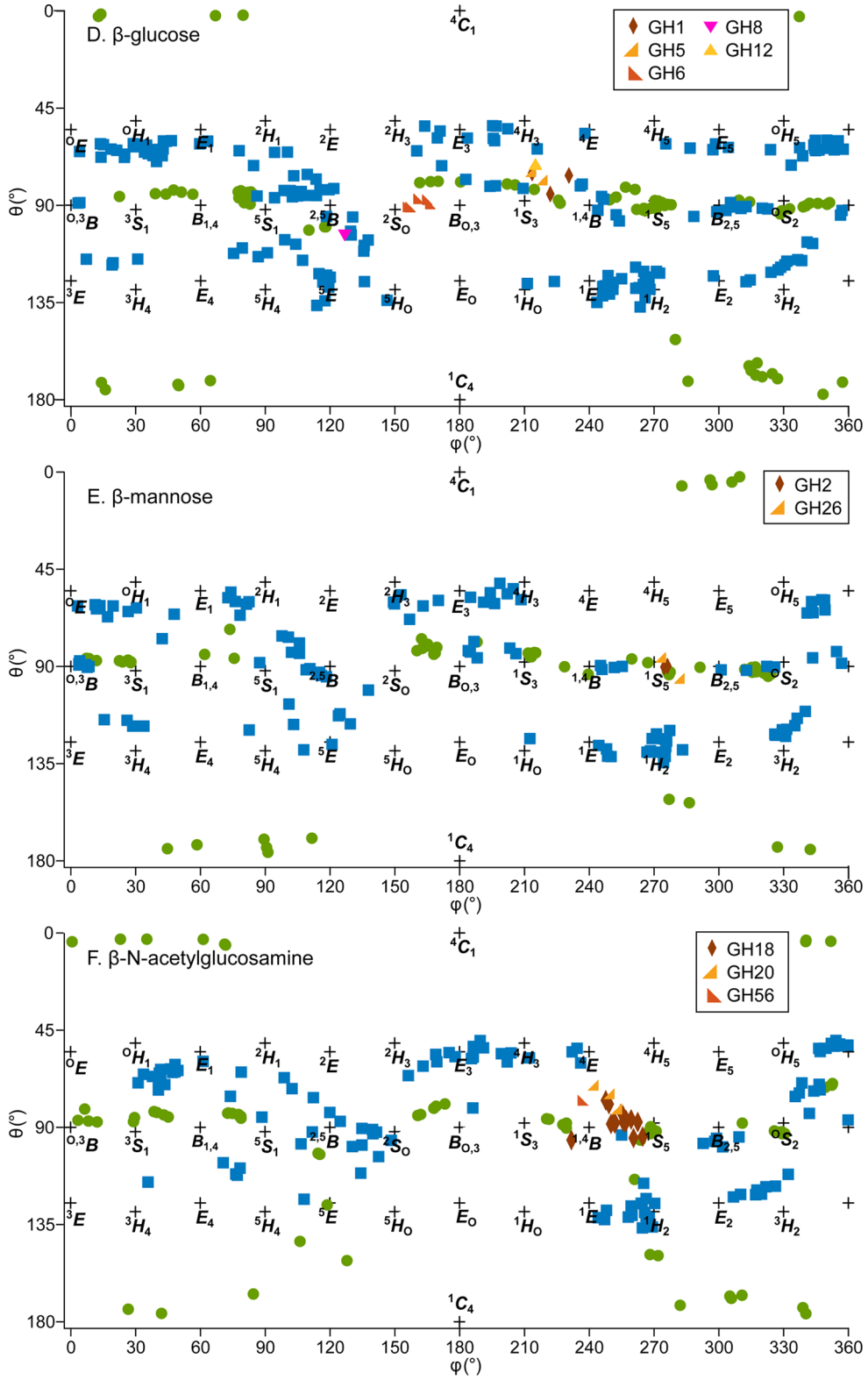


Figure 4. Mercator projections of the CP spheres for each of the molecules in this study, showing the puckering parameters of the following: canonical puckers (labeled black crosses), local minima (green circles), TSs (blue squares), and activated substrates from enzyme crystal structures (remaining points, labeled by enzyme family; further details for each structure such as resolution, substrate, and citation are provided in the SI).

free energy more than the enthalpy. Thus, we do not expect the HO assumption to significantly affect the relative enthalpy data presented here.

Intrinsic reaction coordinates (IRCs) using the Hessian-based Predictor-Corrector integration method[76−79] were followed from each TS to connect it with the correct local minima. In very flat regions of the potential energy surface, the gradient-based stopping criteria was decreased or even eliminated and the Hessian recalculated at the first two steps. TSs corresponding to exocyclic rotation were discarded as they are not the focus of this study.

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

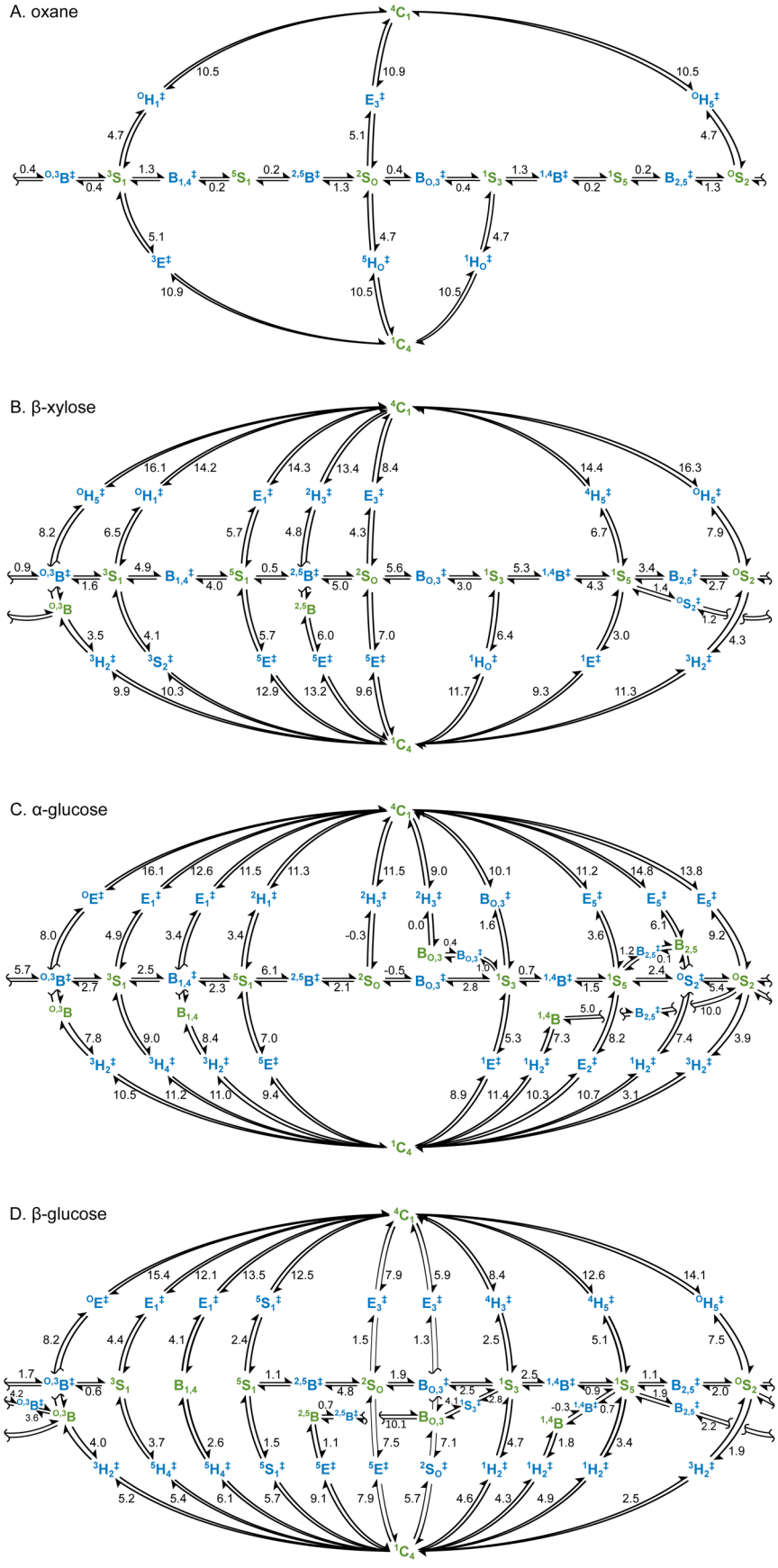


Figure 5. continued

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

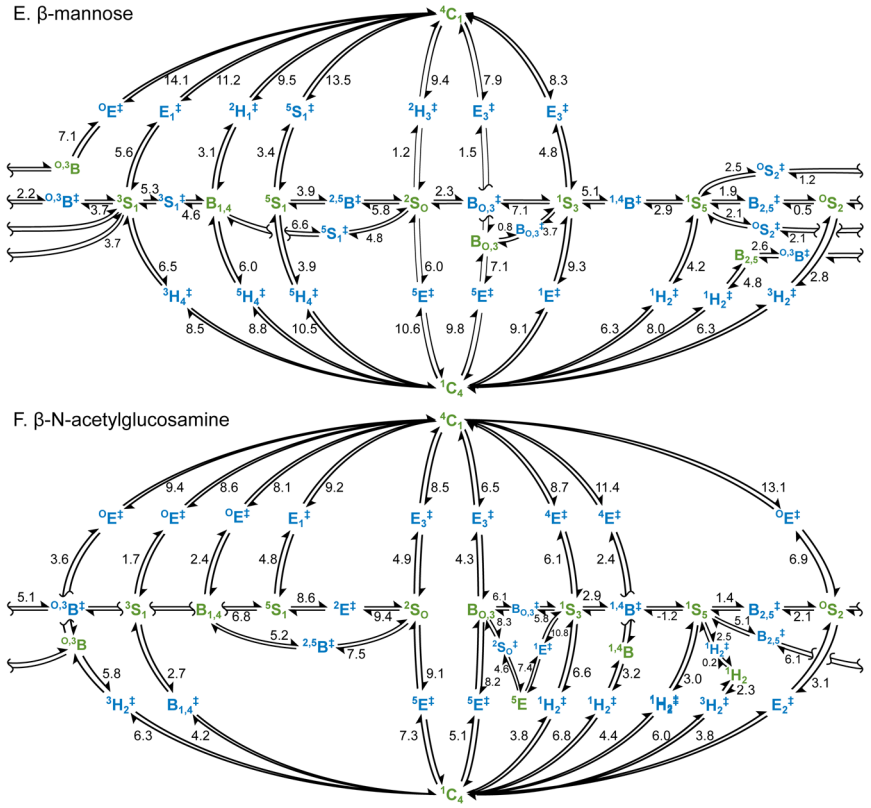


Figure 5. Main puckering interconversion pathways for the molecules in this study. Values shown are the enthalpy change from a local minimum (green) to a TS (blue with ‡), at 298 K in kcal/mol calculated with CCSD(T)6-311+G(d,p)//B3LYP/6-311+G(2df,p).

The results shown in the following sections are for "low-energy" local minima and saddle points, where we define "low-energy" as having a Gibbs free energy within 1.25 kcal/mol (approximately 2 k_BT at 298 K) for a given pucker local minimum or pathway. Electronic densities were calculated at the CCSD/6-311+G(d,p) level of theory. Partial atomic charges were evaluated with natural population analysis[80] using NBO 5[81] and with the ChelpG scheme.[82]

RESULTS

The range of results for puckering conformations across different sugars provides valuable context for considering features of individual sugars. Thus, we first present our results holistically before examining each sugar. In the overview, we also introduce why certain properties are of interest for catalysis. The key results from this work include which puckering conformations are local minima and which are TSs for puckering interconversion. CP parameters allow us to precisely identify the puckering geometry, as shown in Figure 4 for all six molecules. This figure contains Mercator projections of the three-dimensional CP sphere mapped onto two dimensions. As such, there is significant horizontal distortion at the top and bottom of the projection, but little at the equator, which is rich in puckering conformations. The CP parameters of the canonical puckering geometries are marked with a black cross and labeled with their IUPAC name. Green circles identify the CP coordinates of stable local minima, and blue squares identify CP coordinates of puckering TS. The Supporting Information (SI) contains additional Mercator projections on which arrows point from each TS to the local minima obtained by following the IRCs of each TS. The CP parameters of experimentally observed substrates in GHs are also shown on these Mercator projections, identified by the GH family. The SI contains a list of the structures and associated key data, such as the structural resolution and substrate used.

As evident in Figure 4, stable puckering geometries do not conform to the discrete IUPAC puckering designations. While many stable points are intermediate between canonical puckering conformations, the IUPAC designations are nonetheless useful constructs in describing monosaccharide rings. To assign an IUPAC name to each stable point to aid in discussion, we classify each calculated stable point by its closest canonical puckering designation on the CP sphere according to arc length. For oxane, there is only one way to be in a given puckering geometry. The exocyclic groups on the other molecules allow for a diversity of stable puckering geometries for a given pucker, which is reflected in clusters of local minima and TSs on the Mercator projections. We used Boltzmann weighting to obtain aggregate properties for a monosaccharide's puckering conformation. The local minima were grouped by their puckering designation alone, while the TSs were grouped by pathway, defined by the puckering conformations of the TS and the local minima the TS connects.

Figure 5 shows the puckering interconversion map for each molecule, including enthalpic barriers in kcal/mol at 298 K, to move from a local minimum to a TS. For clarity, only the lowest energy pathway between a pair of local minima is shown, and some pathways connecting more distant pairs of local minima were omitted. Unabridged lists of interconversion pathways are included in the SI. To preactivate carbohydrates for catalysis, GHs perturb the monosaccharide ring from $^4C_1$ to a pucker along the CP equator, as shown by the experimentally observed substrate puckering in enzyme active sites in Figure 4. Thus, the enthalpy change from $^4C_1$ is of particular interest. The barriers to interconvert between different conformations

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

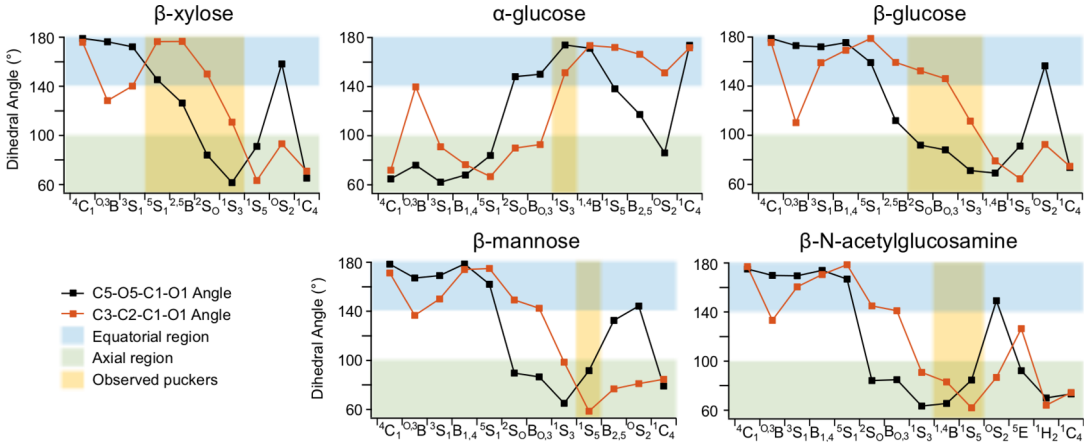


Figure 6. Values of the dihedral angles C5−O5−C1−O1 and C3−C2−C1−O1, in degrees, for each sugar's puckering local minima, to indicate whether the C1−O1 bond displays equatorial character (top, light-blue shaded region) or axial character (bottom, light-green shaded region). Puckers which have been observed in enzyme−substrate activated complexes are highlighted by vertical yellow bands.

on the equator tend to be lower than those to interconvert from the poles. In some cases, the electronic energy surface along the equator is so flat that the enthalpy change from a local minimum to a TS is negative. Although the enthalpies of those TSs are lower than the local minimum, the electronic energy is higher for the TS. Their identification as a ring puckering interconversion TS was, as with all TSs, confirmed by IRC calculations.

Figure 6 provides data to evaluate whether puckering is employed primarily to position the anomeric oxygen bond in an axial position poised for cleavage by S_N2 attack. Specifically, it shows the C5−O5−C1−O1 and C3−C2−C1−O1 dihedral angles for each sugar in each puckering conformation found to be a local minimum. When both of these dihedral angles are near 180°, the C1−O1 bond is unambiguously equatorial, as in the $^4C_1$ conformation of *β*-xylose. To aid in identifying which conformations display equatorial character, the range near 180° is shaded in light blue. Likewise, when both dihedral angles are near 90°, the range shaded with light green, the C1−O1 bond is unambiguously axial, as in the $^1C_4$ conformation of *β*-xylose. The puckered conformations which have been experimentally observed in activated enzyme−substrate complexes are highlighted with a yellow vertical bar. As *α*-glucose clearly demonstrates, the sugar ring is not distorted from $^4C_1$ simply to align the C1−O1 bond axially, as the anomeric oxygen is already axial in the $^4C_1$ conformation. The observed pucker in *α*-glucose activated complexes is not one of the puckered conformations that maintains the axial anomeric oxygen, such as $^5S_1$, but the $^1S_3$ pucker in which the anomeric oxygen is equatorial. The observed puckers for both *β*-mannose and GlcNAc align the anomeric oxygen axially, while the range of observed puckered geometries for *β*-xylose include both primarily axial and primarily equatorial C1−O1 alignments. The *β*-glucose $^{1,4}B$ and $^1S_5$ conformations offer a highly axial anomeric oxygen, but are not the observed puckers; the observed *β*-glucose puckers are more intermediate. Thus, while an axial anomeric bond alignment is likely favorable for glycosidic bond cleavage, as has been proposed,[7,19] it alone is clearly not a necessary or universal feature of activated sugar substrate conformations.

Figure 7 provides a compilation of several other properties to consider as metrics for catalytic susceptibility. Each column of the graphs provides data for a different sugar, displaying properties of the puckering conformations found to be stable local minima. As in Figure 6, the yellow vertical bands highlight conformations observed in crystal structures of activated enzyme−substrate complexes. The first row of Figure 7 shows the one-step enthalpy barrier to move from $^4C_1$ to a different pucker local minimum. An "N/A" indicates that no TS was found for a one-step conversion from $^4C_1$. Thermodynamic stabilities of minima are shown in the second row, with relative Gibbs free energy compared to the $^4C_1$ conformation, which is the lowest-energy conformation in all cases. The remaining three rows highlight properties believed to be indicators of preactivation for catalytic action.[7,20−23] Since GHs employ nucleophilic attack at the C1 carbon, a high partial charge on C1 would make it more susceptible for catalysis. Similarly, a shortened distance between the C1 carbon and the ring oxygen (O5) is consistent with oxocarbenium-like character. The last row quantifies the C1−O1 distance. In a GH, this is the bond that would be broken. A longer C1−O1 distance corresponds to a weakened bond, and thus is another indication that the pucker is preferable for catalysis.

In the following subsections, we discuss each molecule in turn. The difference between the molecules is only due to differences in exocyclic groups, which result in the rich conformational landscapes that affect the ring pucker interconversion pathways (kinetic landscapes) of the pyranose rings (Figure 5). The final subsection reviews the importance of the thorough conformational search included in this study.

Oxane. Oxane, a minimal pyranose ring, contains no exocyclic groups and thus provides an ideal basis to compare the puckering behavior of the monosaccharides included in this study. The number of pathways from oxane's chair conformations (Figure 5A) is significantly reduced from the number in Stoddart's proposed interconversion pathways for cyclohexane (Figure 3). Stoddart described pathways through each of the six envelope and six half-chair conformations, showing only the half-chair conformations.[51] In doing so, he cited Hendrickson's semiempirical strain calculations that show the half-chair TSs have slightly lower energies than the envelopes. Specifically, Hendrickson calculated that a cyclohexane envelope TS is 11.3 kcal/mol higher in energy than the chair, while the half-chair TS is 11.0 kcal/mol higher in energy.[83] Our results for oxane also show a slightly higher barrier for the envelope TS than the half-chair. The smaller

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

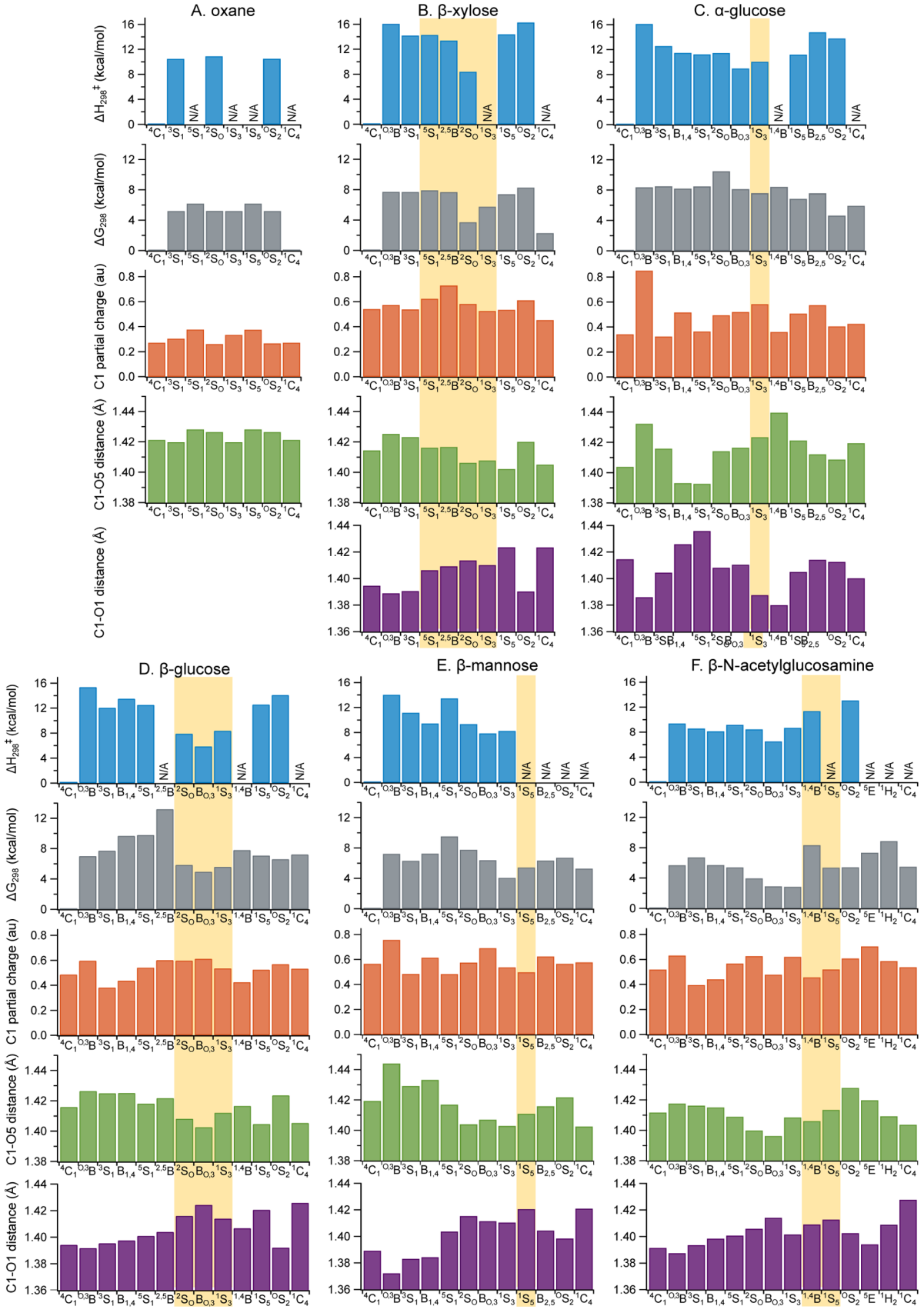


Figure 7. Select Boltzmann-weighted properties of puckering local minima: barrier heights ($\Delta H^{\ddagger}$) in kcal/mol for conversion from the $^4C_1$ geometry to the designated puckering local minima in one step ("N/A" if no TS was obtained for a one-step conversion); relative Gibbs free energy in kcal/mol compared to the $^4C_1$ conformation; ChelpG partial charge on the C1 in atomic units; distance between the C1 and ring oxygen (O5) in Å; and distance between the C1 and O1 in Å. Puckering conformations observed in activated enzyme−substrate structural studies are highlighted in yellow.

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

number of interconversion pathways is consistent with the reduction in symmetry from cyclohexane to oxane, and agrees with early semiempirical[84] and more recent DFT studies.[85] The stable envelope TSs contain the C3 carbon, which is opposite the ring oxygen, above or below the plane (due to remaining symmetry, ^{3}E and E_3 are equivalent). The stable half-chair TSs are those with the ring oxygen and a neighboring carbon participating. As shown in the Mercator projection of the CP sphere in Figure 4A, the CP parameters for the calculated stable points do not always coincide with the parameters for the canonical conformations.

Along the equator, the skew conformations are local minima, and the boat conformations are TSs that allow interconversion along this "pseudorotational itinerary." The barriers for interconversion along the equator are 1.3 kcal/mol or less, describing a very flat portion of the enthalpic surface. Figure 7A conveys some property differences among the local minima. For example, the equivalent $^{5}S_1$ and $^{1}S_5$ conformations have slightly higher Gibbs free energy, slightly higher partial charge on the C1 carbon, and slightly longer C1−O bond lengths than all of the other minima. Since oxane does not have a hydroxyl group bonded to C1, no graph is shown in the final row of graphs in Figure 7. Oxane is not a GH substrate, so the pucker properties are not presented to understand properties of activated substrates, but instead are presented as a point of comparison for the monosaccharides discussed below.

β-Xylose. The potential energy surface for β-xylose is more complicated than that for oxane due to the additional degrees of freedom from the four exocyclic hydroxyl groups. Each group has multiple low-energy orientations, resulting in multiple possible rotamers of each puckering conformation. As previously noted, this results in clusters of local minima (green) and TSs (blue) on the Mercator projection of the CP sphere in Figure 4B. Additionally, there is more variety in heights of the enthalpic barriers. The enthalpic surface along the equator is not as flat as for oxane, while these barriers remain smaller than those to move from the $^{4}C_1$ pole to the equator.

Figure 4B also shows the CP parameters for "activated" substrates obtained from structural studies. Enzyme structures from family GH11 (retaining), GH120 (retaining), and GH43 (inverting) indicate that $^{2,5}B$ and $^{2}S_O$ are preferred ring conformations for catalysis.[21,86,87] Crystal structures from the GH10[88] and GH39[89] (both retaining) adopt the $^{1}S_3$ conformation.

As shown in Figures 5B and 7B, the most kinetically accessible puckered conformation is $^{2}S_O$, which is also more stable (lower Gibbs free energy) than the other puckered conformations along the equator, as confirmed in a recent study by Rönnols et al.[90] Their NMR studies in solvents of varying dielectric constant show significant populations of $^{4}C_1$, $^{1}C_4$, and $^{2}S_O$ geometries. Their studies in low dielectric constant are most relevant to this gas-phase study and to the low dielectric environment found in enzymes.[63] They found that at low dielectric constant, the populations in the $^{1}C_4$ and $^{2}S_O$ geometries are similar. This is consistent with the similar Gibbs free energy we found for these two geometries.

Several puckered conformations display properties favorable to catalysis. $^{2,5}B$ has the highest positive partial charge, making it more susceptible to nucleophilic attack, while $^{2}S_O$, $^{1}S_3$, and $^{1}S_5$ have shorter O5−C1 bond lengths consistent with oxocarbenium character as well as an elongated C1−O1 bond consistent with a weakened leaving-group bond. Of these four conformations, only $^{1}S_5$ is not experimentally observed. It is important to note that free energy does not explain why it is not observed; it has a lower Gibbs free energy than the observed $^{2,5}B$ pucker. $^{1}S_5$ also provides a more axial alignment of the anomeric oxygen (Figure 6). One possible candidate revealed in this study is that it is relatively more difficult kinetically to obtain the $^{1}S_5$ conformation than the other three. As previously noted, $^{2}S_O$ has the lowest barrier from $^{4}C_1$. Unlike $^{1}S_5$, both $^{2,5}B$ and $^{1}S_3$ neighbor $^{2}S_O$ on the relatively flat enthalpic surface of the CP sphere's equator. Thus, the observed structures represent a combination of some catalytically favorable properties and kinetic accessibility, which may be significant.

Of these observed puckering geometries, the differences between $^{2,5}B$, $^{2}S_O$, and $^{1}S_3$ merit further discussion. $^{2}S_O$ and $^{1}S_3$ have both low free energy and shorter C1−O5 bond lengths, while $^{2,5}B$ has relatively high free energy and C1 partial charge. It may be that the higher C1 partial charge offers more catalytic benefit, making it worth overcoming the higher free energy. Amorim et al. tested the hypothesis that the $^{2,5}B$ conformation is especially amenable to catalysis.[91] They compared hydrolysis rates of a xyloside and a bicyclic xyloside analogue locked into a $^{2,5}B$ geometry by a C−C bridge between C2 and C5. They found hydrolysis rates for the locked $^{2,5}B$ xyloside to be at least 2 orders of magnitude faster than the flexible xyloside. The difference between the rates for locked and flexible xylosides indicate there is a barrier for the flexible ring to overcome to adopt the catalytically favorable pucker but that enzymes that overcome this barrier obtain additional catalytic benefit. This finding is consistent with our study—if there were no catalytic advantage to puckering, there would be no reason to leave the $^{4}C_1$ energy well. At the same time, our findings are consistent with enzymes following an efficient path; the β-xylose catalytic itineraries avoid the highest barriers in the path toward activated puckered conformations.

α-Glucose. Just as β-xylose is more complex than oxane and has more interconversion pathways, α-glucose has more ring puckering interconversion pathways than β-xylose (Figure 5C). This α-sugar provides an interesting contrast to the rest of the sugars in this study, which all have β-anomeric oxygens. As previously discussed, the solution-stable $^{4}C_1$ conformation has an axial anomeric oxygen, which is moved to the equatorial conformation in the observed activated pucker geometry. As for all of the pyranose sugars studied, the $^{4}C_1$ conformation is in a deep energy well compared to the puckered conformations along the equator of the CP sphere. The energy landscape around $^{2}S_O$ is so flat that the enthalpy barrier to moving from $^{2}S_O$ to $^{1}S_3$ or $^{4}C_1$ is negative. In this case, although the enthalpies of the relevant TSs are lower than the $^{2}S_O$ local minimum, both the electronic energy and Gibbs free energy were higher for the TS.

The lowest-enthalpy barrier from the $^{4}C_1$ conformation is to $B_{O,3}$. The enthalpy barrier to convert from $B_{O,3}$ to $^{1}S_3$ is only 0.4 kcal/mol, and $^{1}S_3$ is downhill in free energy from $B_{O,3}$. $^{1}S_3$ has the second-lowest enthalpy barrier from $^{4}C_1$. It is also part of the proposed catalytic itinerary for α-glucosidases,[7] based on identification of such distorted rings in the catalytic sites of GH structures (Figure 4C). As shown in Figure 7C, $^{1}S_3$ has a high partial charge on the anomeric carbon, second only to $^{O,3}B$. Both have similar Gibbs free energies, but the kinetic barrier to $^{O,3}B$ is the highest of the one-step conversions from $^{4}C_1$.

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

While the $^{1}S_{3}$ higher anomeric carbon partial charge is catalytically favorable, the C1−O1 and C1−O5 bond distances for $^{1}S_{3}$ are not consistent with what is expected to be catalytically favorable. For this α-anomeric carbon, no puckering conformation yields both a higher C1 partial charge simultaneously with a shorter C1−O5 distance and longer C1−O1 distance. On the basis of the properties that we quantified, the experimental structures in the $^{1}S_{3}$ conformation suggest that following a low-energy path to an electrophilic anomeric carbon may be more catalytically advantageous than a catalytic itinerary that provides an elongated leaving group bond.

The observed enzymatic preference for the $^{1}S_{3}$ conformation also indicates free energy alone is not a reliable predictor of catalytic itineraries. The $^{1}S_{3}$ conformation is approximately 7.6 kcal/mol higher in energy than $^{4}C_{1}$, while $^{O}S_{2}$ is approximately 4.7 kcal/mol higher in energy than $^{4}C_{1}$. However, to our knowledge, no α-glucose monomer has been observed in the $^{O}S_{2}$ conformation in a GH enzyme.

β-Glucose. As shown in Figure 5D, the β-glucose kinetic landscape contains even more puckering interconversion pathways than α-glucose. It also has some lower enthalpy barriers to leave the $^{4}C_{1}$ low-energy well. For α-glucose, the lowest enthalpy barrier to leave $^{4}C_{1}$ is 9 kcal/mol. For β-glucose, three pathways require less than 9 kcal/mol to escape the $^{4}C_{1}$ conformation: 5.9 kcal/mol to reach $B_{O,3}$, 7.9 kcal/mol to reach $^{2}S_{O}$, and 8.4 kcal/mol to reach $^{1}S_{3}$. Intriguingly, these low-energy-barrier puckering conformations are found in crystal structures of substrates in catalytic sites (Figure 4D). The low-energy path for interconversion from $^{4}C_{1}$ to $B_{O,3}$ through E_{3} was also found by Barnett and Naidoo[13] with a barrier of approximately 5 kcal/mol evaluated with the semiempirical method PM3, which is in excellent agreement given the difference in methods. In contrast, the barrier for this path given by Biarnés et al.[24] is more than 8 kcal/mol evaluated with the DFT functional PBE. Their overestimation is not surprising given the analysis by Sega et al.[31]

Figure 7D shows that in addition to having low kinetic barriers, $^{2}S_{O}$, $B_{O,3}$, and $^{1}S_{3}$ are among the lowest Gibbs free energy conformations and have the highest C1 partial charges along the equator. They also have catalytically favorable shorter C1−O5 and longer C1−O1 bonds than does $^{4}C_{1}$. However, these three conformations are not the only ones with low Gibbs free energy, high C1 partial charge, and short C1−O5 bond distances. $^{1}S_{5}$, for example, has similar properties, except for its higher kinetic barrier. $^{1}S_{5}$ is not one of the experimentally observed puckering geometries, suggesting that kinetics may have an important role in the catalytic itineraries of enzymes acting on β-glucose polymers, an aspect only peripherally considered in previous computational studies of β-glucose puckering.[24]

β-Mannose. β-mannose differs from β-glucose only in the stereochemistry of the O2 attachment to the ring; the β-mannose O1−O2−O3 arrangement is cis−cis, whereas the β-glucose O1−O2−O3 arrangement is trans−trans. Different puckering geometries can allow more or less space between the adjacent oxygens, which can be a determining factor in determining conformation stability and energy. A striking feature of the β-mannose puckering landscape is the lack of stable points in the ^{4}E−$^{4}H_{5}$−E_{5} region, as shown in Figures 4E and 5E. This is a region of high cis−cis clash between O1, O2, and O3. Optimizations with the ring pucker held fixed and the rest of the molecule relaxed showed that these three conformations have among the shortest O1−O2 distances, at 2.5 Å (the fully relaxed $^{4}C_{1}$ O1−O2 distance is 2.7 Å and the oxygen van der Waals radius is 1.5 Å[92]). Given the steric hindrance of these puckers, the lack of stable points in this region is logical. It is also remarkable because this region lies between the solution-stable $^{4}C_{1}$ conformation and the $^{1}S_{5}$ conformation, which is the experimentally observed pucker conformation in activated substrate-enzyme structures. We observed no one-step pucker-interconversion paths between $^{4}C_{1}$ and $^{1}S_{5}$. Although $^{1}S_{5}$ is directly below this barren region of the CP sphere, it does not suffer from cis−cis clash. In fact, it offers the greatest distance (2.9 Å) between O1 and O2 of all the β-mannose puckering geometries.

Ardèvol et al. stated, "We have found that $^{1}S_{5}$ is among the most stable conformers and simultaneously is the most preactivated conformation in terms of elongation/shortening of the C1−O1/C1−O5 bonds, C1−O1 orientation, and charge development at the anomeric carbon."[25] Our findings support their assertion that $^{1}S_{5}$ displays a favorable axial C1−O1 orientation (Figure 6) and elongated C1−O1 bond length (Figure 7E). They also note that a shorter C1−O5 bond and higher C1 partial charge is considered catalytically advantageous. However, contrary to their assertion that $^{1}S_{5}$ is preactivated in terms of C1−O5 bond shortening and charge development at the anomeric carbon, both their study and the present study found that $^{1}S_{5}$ has a longer C1−O5 bond length and a lower C1 partial charge than other puckering conformations, including being less preactivated than the lowest-energy $^{4}C_{1}$ conformation. Our results for free energy differ slightly from Ardèvol et al. They report that the local minima near the $^{1}S_{3}$ conformation are the lowest free-energy conformations on the CP sphere's equator, slightly below the minimum near $^{1}S_{3}$. We found that the local minimum near $^{1}S_{3}$ is lower than the local minima nearest to the canonical $^{1}S_{5}$ conformation. We believe our finding is more accurate because we ensured that our calculations were tightly converged and included all orientations of the exocyclic groups.

Our analysis of the pathways to convert from the $^{4}C_{1}$ conformation to a puckering conformation along the equator reveals that $B_{O,3}$ has the lowest enthalpy barrier, followed by $^{1}S_{3}$, with an enthalpy barrier of less than 1 kcal/mol to convert from $B_{O,3}$ to $^{1}S_{3}$. While $^{1}S_{3}$ has the lowest Gibbs free energy of the puckered conformations, $^{O,3}B$ has the highest partial charge on C1. Yet, to date, neither of these puckering geometries has been observed in an activated enzyme−substrate complex. The observed structure, $^{1}S_{5}$, is adjacent to $^{1}S_{3}$ and has the second lowest Gibbs free energy of the puckers on the CP equator, as shown in Figure 7E. However, contrary to the result implied by Ardèvol et al., it has among the lowest C1 partial charges, although it does have the longest C1−O1 bond length. Like β-xylose and α-glucose, β-mannose has no single puckering geometry with simultaneously favorable kinetics, C1 charge, C1−O5 distance, and C1−O1 distance. The longer distance between the O1 and O2 oxygens may play a role in why $^{1}S_{5}$ is observed over the $^{1}S_{3}$ or $B_{O,3}$ conformations, which have O1−O2 distances of 2.6 Å. The longer O1−O2 distance offered by $^{1}S_{5}$ may allow enzyme residues to more easily selectively interact with O2, which is a highly conserved enzyme−substrate interaction,[93] without clashing with O1. Simulations of enzymes active on β-mannose substrates could test this hypothesis but are outside the scope of this study. This study does show that a consideration of free energy, C1−O1 and

C1−O5 bond lengths, and C1 partial charge does not allow simple identification of a preactivated puckering conformation.

***β*-*N*-Acetylglucosamine.** GlcNAc, with its long N-acetyl arm attached to the C2 carbon, has a distinct feature on its puckering landscape compared to the other sugars examined in this study: GlcNAc features stable local minima in the region between the $^{1}C_{4}$ pole and the equator, as shown in Figure 4F. Specifically, there are stable local minima in the ^{5}E conformation and near the $^{1}H_{2}$ canonical puckering conformation. Figure 8 illustrates how the N-acetyl arm stabilizes

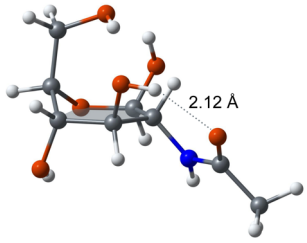


Figure 8. GlcNAc in a ^{5}E geometry that is a stable local minimum.

the ^{5}E conformation. The position of C5 above the plane formed by the other ring atoms allows for intramolecular hydrogen bonding between the hydroxyl groups attached to C1, C3, and C6, as can also occur in the $^{1}C_{4}$ conformation. However, unlike in $^{1}C_{4}$, ^{5}E contains C2 in the same plane as O5, C1, C3, and C4. This position of C2 allows the acetyl oxygen to be in proximity to the O3 hydrogen to allow for hydrogen bonding which stabilizes the structure. Similarly, structures near $^{1}H_{2}$ are stabilized by interactions of the N-acetyl group with the rest of the molecule.

Figure 4F also shows that GlcNAc-active enzymes distort the substrate to a pucker in the vicinity of $^{1,4}B$, to a highly conserved degree. Surprisingly, this preferred pucker conformation has among the highest Gibbs free energy and enthalpy barriers to convert from $^{4}C_{1}$ in one step, as shown in Figures 5F and 7F. Additionally, its C1 partial charge is lower than for $^{4}C_{1}$. Unique patterns appear to be at play with GlcNAc.

van Aalten et al. noted that chitinases take advantage of the N-acetyl arm for nucleophilic attack, instead of a carboxylate protein side chain as is common in other carbohydrate-active enzymes.[94] Thus, the ability of the N-acetyl group to interact with the rest of the molecule has been recognized as an important feature of GlcNAc-modifying enzymes, as well as an important feature affecting the puckering interconversion landscape. This could also be a factor in why the thermodynamically less stable $^{1,4}B$ pucker is adopted for catalysis.

The top row of Figure 9 shows two conformations of the GlcNAc substrate in the −1 position from crystal structures of enzyme activated substrate complexes and, in the bottom row, the lowest energy conformations of two monomer pucker geometries, $B_{O,3}$ and $^{1,4}B$. In order for the N-acetyl group to perform nucleophilic attack at C1, the acetyl group must be below C1, which in turn requires an axial orientation of the N-acetyl group at C2. One measure of whether the N-acetyl group is oriented below C1 is the value of the O5−C1−C2−N angle, which is noted on the conformations in Figure 9. The experimentally observed puckered geometries have an O5−C1−C2−N angle of about 90°, with an even smaller angle in some of the crystal structures listed in the SI. $B_{O,3}$ has among

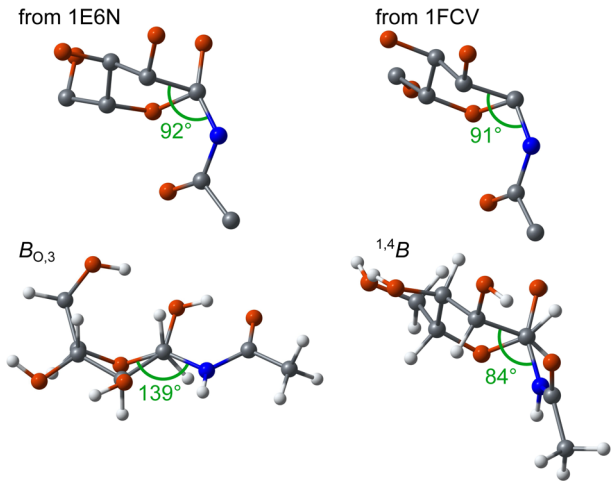


Figure 9. Top row: conformation of the GlcNAc subunit in the −1 position of the active site of *Serratia marcescens* ChiB from PDB entry 1E6N[94] (left) and *Apis mellifera* Hva from PDB entry 1FCV[95] (right). Bottom row: lowest-energy local minima for a GlcNAc monomer in the $B_{O,3}$ (left) and $^{1,4}B$ (right) conformations. The images are annotated with the value for the O5−C1−C2−N dihedral angle, with that angle highlighted in green.

the lowest relative free energies of the GlcNAc puckering conformations and the lowest enthalpy barrier from $^{4}C_{1}$, but the N-acetyl group is in an equatorial position, preventing orientation of the N-acetyl group below C1. In contrast, the observed $^{1,4}B$ conformation has the smallest O5−C1−C2−N dihedral angle of the puckering conformations on the equator of the CP sphere. No other monomer in this study has such a "handle" for the enzyme to manipulate. It appears that for this monomer, the pucker conformation which best allows for the N-acetyl arm to participate in catalysis is employed, trumping other considerations that we examined.

Importance of Conformational Search. Monosaccharides exhibit a vast array of available conformations due to the orientations of each exocyclic group and the multiple puckering conformations. Different rotamers of the same sugar in the same puckering geometry can differ in free energy by several kcal/mol, due to different orientations allowing more or less intramolecular hydrogen bonding. Errors of that magnitude could lead to misidentification of low-energy puckering conformations as well as misestimating rate coefficients by orders of magnitude. This warrants explicit treatment of exocyclic group rotation in theoretical treatments. The explicit consideration of the exocyclic groups in the present investigation for each monosaccharide provides confidence that we have identified the lowest-energy substrate puckering itineraries.

With these data in hand, we can review the results to determine whether such a thorough search was required. For example, if the hydroxymethyl groups of the six-carbon sugars consistently take on the same orientation in the low-energy conformations, we would know that wide conformation sampling was not needed. The orientation of hydroxymethyl groups is commonly classified as gauche−gauche (gg), trans−gauche (tg), or gauche−trans (gt), based on the O5−C5−C6−O6 and C4−C5−C6−O6 dihedral angles, respectively.[96,97] For *β*-glucose, we found that the hydroxymethyl orientation was most commonly in the gg orientation, as previously reported.[24] However, about one-third of the puckering geometries for

dx.doi.org/10.1021/ja410264d | J. Am. Chem. Soc. 2014, 136, 1008−1022

which we isolated local minima have the lowest-energy conformation with a non-gg orientation. The hydroxymethyl orientations were even more diverse for the TSs; around fifty percent of the lowest energy TS conformations contain non-gg orientations. Had we not thoroughly searched the conformation space so thoroughly to find the lowest energy conformations, our calculations of relative free energy and kinetic barriers would likely be incorrect. This could result in either overestimation of barriers (for example, if we compared a higher-energy TS with a low-energy local minima) or underestimation (if we compared a lower-energy TS with a higher-energy local minima). Our exhaustive search for both local minima and TSs for each puckering conformation results in the most comprehensive data set to date for ring interconversion for each of the sugars in this study. Furthermore, our method did not presuppose a reaction coordinate. Internal mode analysis of imaginary frequencies shows that TS reaction coordinates rarely comprise purely ring puckering. While the primary component is reliably rotational movement of dihedral angles in the ring, rotational movement of exocyclic groups can contribute more than 20% of the displacement in the imaginary frequency.

The resulting library of low-energy local minima and TSs for the six pyranose molecules in their range of puckering conformations provides an excellent resource for future computational studies of carbohydrate-modifying enzymes, including both GHs and glycosyltransferases. The SI contains atomic Cartesian coordinates for local minima and TSs which have Gibbs free energy within 1.25 kcal/mol (approximately 2 k_BT at 298 K) for a given pucker local minimum or TS connecting the same two local minima.

DISCUSSION AND CONCLUSIONS

In 1971, Stoddart published a map of the puckering interconversion pathways of cyclohexane.[51] Our present study now provides the detailed interconversion pathways of five biologically important pyranose sugars: β-xylose, α-glucose, β-glucose, β-mannose, and β-*N*-acetylglucosamine. For each of the 38 IUPAC recognized puckering conformations, all degrees of freedom due to rotatable bonds were explored and puckering interconversion TSs were obtained and verified. A library of the low-energy local minima and TSs obtained is provided in the SI.

Comparing the kinetic landscapes of the pyranose sugars with each other and a minimal pyranose molecule (oxane) reveals the large effect exocyclic groups have on the location of stable geometries, the puckering pathways, and the kinetic barriers of interconversion. Figure 7 conveys the effect of exocyclic groups on barrier heights for ring interconversion, as well as other properties of interest. For example, oxane's barrier heights to leave the $^4C_1$ energy well are approximately 10 kcal/mol, with a range of only tenths of kcal/mol to travel from $^4C_1$ to different puckering conformations on the equator of the CP sphere. In contrast, the barrier heights for β-glucose to move from $^4C_1$ to other puckering conformations range from 5 to 13 kcal/mol for different puckering conformation destinations. The presence of exocyclic groups could either lower the energy of TSs or raise them due to steric hindrance. Thus, theoretical studies must consider the role of exocyclic groups when constructing thermodynamic and kinetic landscapes.

A more severe result of the same phenomena is differences in the presence or absence of stable puckering geometries among these molecules. Exocyclic groups can stabilize otherwise unstable geometries, as GlcNAc's N-acetyl arm allows for a local minima at ^5E, which is not a feature of any other molecule in this study. They can also interfere with stability of puckering geometries, as found for β-mannose. Steric clash from β-mannose's cis orientation of C2 with respect to both C1 and C2 results in the $^4E-{}^4H_5-E_5$ region of the CP sphere being devoid of stable points of either type (local minima or TSs). This is in contrast with its epimer β-glucose, where the $^4E-{}^4H_5-E_5$ region is populated with multiple TSs.

In addition to mapping the kinetic landscape of these pyranose sugars, the high-level QM calculations provide rich data about the electronic structure of each conformation. This allowed us to test the hypothesis that different puckering conformations make carbohydrate substrates more amenable to catalysis. As previously reported by another group,[24,25] some molecules show a confluence of catalytically attractive features in the experimentally observed puckering conformations of activated substrates, such as an enriched partial charge for nucleophilic attack, an elongated leaving group bond, and low relative Gibbs free energy. However, such a formula does not hold for all pyranose sugars. While β-glucose has an experimentally observed puckering conformation with simultaneously higher C1 partial charge, longer C1−O1 bond length, and low Gibbs free energy, neither β-xylose nor α-glucose have such a serendipitous structure. The observed GlcNAc activated pucker is not the one with enriched C1 partial charge, an elongated leaving group bond, and low relative Gibbs free energy. Instead, the preferred pucker is the one that allows the enzyme to use the N-acetyl arm at the nucleophile, as van Aalten et al. observed.[94] While electrostatic, thermodynamic, and stereochemical considerations are important to understanding why enzymes pucker carbohydrates in enzyme-active sites, these properties are not sufficient to understand which puckers are favored. Simulations and structural biology studies of enzyme action on substrates will continue to reveal additional important considerations related to enzyme−substrate interactions,[98−100] such as perturbations to glycosidic bond torsional angles.[101] As clearly demonstrated by α-glucose, neither free energy of stable states nor axial alignment of the C1−O1 bond are reliable predictors of conformations on enzyme catalytic itineraries. The picture emerging from this comprehensive study of sugar puckering conformations is fuller, more nuanced, and more complex than previously proposed. Such complexity is to be expected from sugars; sugar chemistry and glycobiology are notoriously complex.[3] The information presented in this work on how puckering perturbs monosaccharide properties provides a valuable foundation for future studies with a substrate interacting with catalysts, allowing disambiguation of the effect of interactions from the effect of substrate distortion alone.

ASSOCIATED CONTENT

Supporting Information

Complete refs 53 and 54; a summary of key information about the crystal structures used in this study, such as resolution and substrate; for each molecule in this study, Mercator projection diagrams showing which TS connects to which local minima; full listings of puckering interconversion pathways for each molecule and their enthalpy barriers; for each of the geometries used to generate the data in main manuscript, we include absolute energies, atomic Cartesian coordinates, partial charges on non-hydrogen atoms assigned using the NBO and ChelpG

schemes, and C−C and C−O bond lengths. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

gregg.beckham@nrel.gov

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the DOE Computational Science Graduate Fellowship (CSGF) which is provided under Grant No. DEFG02-97ER25308 and the Department of Energy (DOE) Office of Energy Efficiency and Renewable Energy (EERE) through the BioEnergy Technologies Office, Grant No. DE-EE0003044. G.T.B. also acknowledges the NREL Laboratory Directed Research and Development Program for funding. This research used resources of the following: the NREL Computational Sciences Center supported by the DOE Office of EERE under Contract No. DE-AC36-08GO28308; the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231; and the San Diego Supercomputing Center (SDSC) SDSC under the NSF XSEDE Grant MCB090159. H.B.M. is also supported by an award from the ARCS Foundation Inc., Chicago Chapter. The authors would like to thank Fernando R. Clemente, George R. Schatz, Jared M. Clark, Christina M. Payne, Joshua V. Vermaas, and Michael F. Crowley for helpful discussions. The authors also thank Chris Mayes for helpful scripts, available to the public at https://github.com/cmayes.

REFERENCES

(1) McNeil, M.; Darvill, A. G.; Fry, S. C.; Albersheim, P. *Annu. Rev. Biochem.* **1984**, *53*, 625−663.

(2) Orpin, C. G. In *Biochemistry and Genetics of Cellulose Degradation*; Aubert, J., Beguin, P., Millet, J., Eds.; Academic Press Inc.: Orlando, FL, 1988; pp 171−179.

(3) Varki, A. *Cell* **2006**, *126*, 841−845.

(4) Hart, G. W.; Copeland, R. J. *Cell* **2010**, *143*, 672−676.

(5) Imperiali, B. *J. Am. Chem. Soc.* **2012**, *134*, 17835−17839.

(6) Koshland, D. E., Jr. *Biol. Rev.* **1953**, *28*, 413−436.

(7) Davies, G. J.; Planas, A.; Rovira, C. *Acc. Chem. Res.* **2012**, *45*, 308−316.

(8) Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. *Nucleic Acids Res.* **2009**, *37*, D233−238.

(9) Blake, C. C. F.; Mair, G. A.; North, A. C. T.; Phillips, D. C.; Sarma, V. R. *Proc. R. Soc. London, Ser. B* **1967**, *167*, 365−377.

(10) Schwarz, J. C. P. *J. Chem. Soc., Chem. Commun.* **1973**, *14*, 505−508.

(11) Joint Commission on Biochemical Nomenclature, I. *Eur. J. Biochem.* **1980**, *111*, 295−299.

(12) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.

(13) Barnett, C. B.; Naidoo, K. J. *Mol. Phys.* **2009**, *107*, 1243−1250.

(14) Barnett, C. B.; Naidoo, K. J. *J. Phys. Chem. B* **2010**, *114*, 17142−17154.

(15) McNaught, A. D. *Pure Appl. Chem.* **2008**, *68*, 1919−2008.

(16) Satoh, H.; Manabe, S. *Chem. Soc. Rev.* **2013**, *42*, 4297−4309.

(17) Heaven, M. C. *Science* **2013**, *342*, 46−47.

(18) Chang, Y.-P.; Długołęcki, K.; Küpper, J.; Rösch, D.; Wild, D.; Willitsch, S. *Science* **2013**, *342*, 98−101.

(19) Walvoort, M. T. C.; van der Marel, G. A.; Overkleeft, H. S.; Codée, J. D. C. *Chem. Sci.* **2013**, *4*, 897−906.

(20) Sinnott, M. L. *Chem. Rev.* **1990**, *90*, 1171−1202.

(21) Rye, C. S.; Withers, S. G. *Curr. Opin. Chem. Biol.* **2000**, *4*, 573−580.

(22) Withers, S. G. *Pure Appl. Chem.* **1995**, *67*, 1673−1682.

(23) Vocadlo, D. J.; Davies, G. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539−555.

(24) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686−10693.

(25) Ardèvol, A.; Biarnés, X.; Planas, A.; Rovira, C. *J. Am. Chem. Soc.* **2010**, *132*, 16058−16065.

(26) Lammerts van Bueren, A. L.; Ardèvol, A.; Fayers-Kerr, J.; Luo, B.; Zhang, Y.; Sollogoub, M.; Blériot, Y.; Rovira, C.; Davies, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 1804−1806.

(27) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(28) Peters, B.; Trout, B. L. *J. Chem. Phys.* **2006**, *125*, 054108.

(29) Peters, B.; Beckham, G. T.; Trout, B. L. *J. Chem. Phys.* **2007**, *127*, 034109.

(30) Peters, B. *Mol. Simul.* **2010**, *36*, 1265−1281.

(31) Sega, M.; Autieri, E.; Pederiva, F. *J. Chem. Phys.* **2009**, *130*, 225102.

(32) Autieri, E.; Sega, M.; Pederiva, F.; Guella, G. *J. Chem. Phys.* **2010**, *133*, 095104.

(33) Sattelle, B. M.; Hansen, S. U.; Gardiner, J.; Almond, A. J. *Am. Chem. Soc.* **2010**, *132*, 13132−13134.

(34) Sattelle, B. M.; Almond, A. *Glycobiology* **2011**, *21*, 1651−1662.

(35) Sattelle, B. M.; Bose-Basu, B.; Tessier, M.; Woods, R. J.; Serianni, A. S.; Almond, A. *J. Phys. Chem. B* **2012**, *116*, 6380−6386.

(36) Sattelle, B. M.; Almond, A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 5843−5848.

(37) Naidoo, K. J. *Sci. China Chem.* **2011**, *54*, 1962−1973.

(38) Naidoo, K. J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 9026−9036.

(39) O'Donoghue, P.; Luthey-Schulten, Z. A. *J. Phys. Chem. B* **2000**, *104*, 10398−10405.

(40) Ionescu, A. R.; Bérces, A.; Zgierski, M. Z.; Whitfield, D. M.; Nukada, T. *J. Phys. Chem. A* **2005**, *109*, 8096−8105.

(41) Seshadri, V.; Westmoreland, P. R. *J. Phys. Chem. A* **2012**, *116*, 11997−12013.

(42) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77*, 521−555.

(43) Henrissat, B.; Sulzenbacher, G.; Bourne, Y. *Curr. Opin. Struct. Biol.* **2008**, *18*, 527−533.

(44) Zugenmaier, P. *Crystalline Cellulose and Derivatives: Characterization and Structures*; Springer-Verlag: Berlin Heidelberg, 2008; Chapter 1.

(45) Scheller, H. V.; Ulvskov, P. *Annu. Rev. Plant Biol.* **2010**, *61*, 263−289.

(46) Hoefsloot, L. H.; Hoogeveen-Westerveld, M.; Reuser, A. J. J.; Oostra, B. A. *Biochem. J.* **1990**, *272*, 493−497.

(47) Kato, A.; Hayashi, E.; Miyauchi, S.; Adachi, I.; Imahori, T.; Natori, Y.; Yoshimura, Y.; Nash, R. J.; Shimaoka, H.; Nakagome, I.; Koseki, J.; Hirono, S.; Takahata, H. *J. Med. Chem.* **2012**, *55*, 10347−10362.

(48) Fischer, P. B.; Collin, M.; Karlsson, G. B.; James, W.; Butters, T. D.; Davis, S. J.; Gordon, S.; Dwek, R. A.; Platt, F. M. *J. Virol.* **1995**, *69*, 5791−5797.

(49) Rinaudo, M. *Prog. Polym. Sci.* **2006**, *31*, 603−632.

(50) Souza, C. P.; Almeida, B. C.; Colwell, R. R.; Rivera, I. N. G. *Mar. Biotechnol.* **2011**, *13*, 823−830.

(51) Stoddart, J. F. *Stereochemistry of Carbohydrates*; Wiley-Interscience: New York, 1971; p 58.

(52) Bérces, A.; Whitfield, D. M.; Nukada, T. *Tetrahedron* **2001**, *57*, 477−491.

(53) Brooks, B. R.; et al. *J. Comput. Chem.* **2009**, *30*, 1545−1614.

(54) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.;

Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision C.01; Gaussian, Inc.: Wallingford, CT, 2009.

(55) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241.

(56) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364−382.

(57) Wodrich, M. D.; Corminboeuf, C.; Schreiner, P. R.; Fokin, A. A.; von Ragué Schleyer, P. *Org. Lett.* **2007**, *9*, 1851−1854.

(58) Csonka, G. I.; French, A. D.; Johnson, G. P.; Stortz, C. A. *J. Chem. Theory Comput.* **2009**, *5*, 679−692.

(59) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257−2261.

(60) Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, *72*, 650−654.

(61) Hariharan, P. C.; Pople, J. A. *Theor. Chim. Acta* **1973**, *28*, 213−222.

(62) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, *77*, 3654−3665.

(63) Gilson, M. K.; Honig, B. H. *Biopolymers* **1986**, *25*, 2097−2119.

(64) Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, *80*, 3265−3269.

(65) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. *J. Chem. Phys.* **2007**, *126*, 084108.

(66) Hill, A. D.; Reilly, P. J. *J. Chem. Inf. Model.* **2007**, *47*, 1031−1035.

(67) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(68) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **2001**, *114*, 108.

(69) Jiménez-Hoyos, C. A.; Janesko, B. G.; Scuseria, G. E. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6621−6629.

(70) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968−5975.

(71) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(72) Raghavachari, K.; Pople, J. A.; Replogle, E. S.; Head-Gordon, M. *J. Phys. Chem.* **1990**, *94*, 5579−5586.

(73) Dunning, T. H., Jr. *J. Phys. Chem. A* **2000**, *104*, 9062−9080.

(74) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000.

(75) Pfaendtner, J.; Yu, X.; Broadbelt, L. J. *Theor. Chem. Acc.* **2007**, *118*, 881−898.

(76) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Phys.* **2004**, *120*, 9918−9924.

(77) Hratchian, H. P.; Schlegel, H. B. In *Theory and Applications of Computational Chemistry: The First Forty Years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G., Eds.; Elsevier Science: Amsterdam, 2005; pp 195−249.

(78) Hratchian, H. P.; Schlegel, H. B. *J. Chem. Theory Comput.* **2005**, *1*, 61−69.

(79) Collins, M. A. *Theor. Chem. Acc.* **2002**, *108*, 313−324.

(80) Reed, A. E.; Weinstock, R. B.; Weinhold, F. *J. Chem. Phys.* **1985**, *83*, 735−746.

(81) Glendening, E.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann,; J. A. Morales, C. M.;Weinhold, F. *NBO 5*; Theoretical Chemistry Institute: Madison, WI, 2011.

(82) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.

(83) Hendrickson, J. B. *J. Am. Chem. Soc.* **1967**, *89*, 7047−7061.

(84) Pickett, H. M.; Strauss, H. L. *J. Am. Chem. Soc.* **1970**, *92*, 7281−7290.

(85) Fernández-Alonso, M. d. C.; Asensio, J. L.; Cañada, F. J.; Asensio, J. L.; Jiménez-Barbero, J.; Cuevas, G. *ChemPhysChem* **2003**, *4*, 754−757.

(86) Huang, C.-H.; Sun, Y.; Ko, T.-P.; Chen, C.-C.; Zheng, Y.; Chan, H.-C.; Pang, X.; Wiegel, J.; Shao, W.; Guo, R.-T. *Biochem. J.* **2012**, *448*, 401−407.

(87) Brüx, C.; Ben-David, A.; Shallom-Shezifi, D.; Leon, M.; Niefind, K.; Shoham, G.; Shoham, Y.; Schomburg, D. *J. Mol. Biol.* **2006**, *359*, 97−109.

(88) Suzuki, R.; Fujimoto, Z.; Ito, S.; Kawahara, S.-I.; Kaneko, S.; Taira, K.; Hasegawa, T.; Kuno, A. *J. Biol. Chem.* **2009**, *146*, 61−70.

(89) Czjzek, M.; Ben David, A.; Bravman, T.; Shoham, G.; Henrissat, B.; Shoham, Y. *J. Mol. Biol.* **2005**, *353*, 838−846.

(90) Rönnols, J.; Manner, S.; Siegbahn, A.; Ellervik, U.; Widmalm, G. *Org. Biomol. Chem.* **2013**, *11*, 5465−5472.

(91) Amorim, L.; Marcelo, F.; Rousseau, C.; Nieto, L.; Jiménez-Barbero, J.; Marrot, J.; Rauter, A. P.; Sollogoub, M.; Bols, M.; Blériot, Y. *Chem.—Eur. J.* **2011**, *17*, 7345−7356.

(92) Alvarez, S. *Dalton Trans.* **2013**, *42*, 8617−8636.

(93) Cartmell, A.; Topakas, E.; Ducros, V. M.-A.; Suits, M. D. L.; Davies, G. J.; Gilbert, H. J. *J. Biol. Chem.* **2008**, *283*, 34403−34413.

(94) van Aalten, D. M. F.; Komander, D.; Synstad, B.; Gåseidnes, S.; Peter, M. G.; Eijsink, V. G. H. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8979−8984.

(95) Marković-Housley, Z.; Miglierini, G.; Soldatova, L.; Rizkallah, P. J.; Müller, U.; Schirmer, T. *Structure* **2000**, *8*, 1025−1035.

(96) Matthews, J. F.; Bergenstråhle, M.; Beckham, G. T.; Himmel, M. E.; Nimlos, M. R.; Brady, J. W.; Crowley, M. F. *J. Phys. Chem. B* **2011**, *115*, 2155−2166.

(97) Shen, T.; Langan, P.; French, A. D.; Johnson, G. P.; Gnanakaran, S. *J. Am. Chem. Soc.* **2009**, *131*, 14786−14794.

(98) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2010**, *132*, 12800−12803.

(99) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2011**, *133*, 19474−19482.

(100) Thompson, A. J.; Dabin, J.; Iglesias-Fernández, J.; Ardèvol, A.; Dinev, Z.; Williams, S. J.; Bande, O.; Siriwardena, A.; Moreland, C.; Hu, T.-C.; Smith, D. K.; Gilbert, H. J.; Rovira, C.; Davies, G. J. *Angew. Chem., Int. Ed.* **2012**, *51*, 10997−11001.

(101) Miyake, H.; Kurisu, G.; Kusunoki, M.; Nishimura, S.; Kitamura, S.; Nitta, Y. *Biochemistry* **2003**, *42*, 5574−5581.