JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Perspective

pubs.acs.org/JACS

Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from *ab Initio* Quantum Mechanics/Molecular Mechanics Dynamic Simulations

Albert Ardèvol[†,#] and Carme Rovira*[,†,‡]

[†]Departament de Química Orgànica and Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain

[‡]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, 23, 08018 Barcelona, Spain

ABSTRACT: Carbohydrate-active enzymes such as glycoside hydrolases (GHs) and glycosyltransferases (GTs) are of growing importance as drug targets. The development of efficient competitive inhibitors and chaperones to treat diseases related to these enzymes requires a detailed knowledge of their mechanisms of action. In recent years, sophisticated first-principles modeling approaches have significantly advanced in our understanding of the catalytic mechanisms of GHs and GTs, not only the molecular details of chemical reactions but also the significant implications that just the conformational dynamics of a sugar ring can have on these mechanisms. Here we provide an overview of the progress that has been made in the past decade, combining molecular dynamics simulations with density functional theory to solve these sweet mysteries of nature.

1. INTRODUCTION

Carbohydrates are the most abundant biomolecules on Earth. They play a huge diversity of roles in living organisms, ranging from structural elements in cell walls (e.g., cellulose, chitin, starch, or glycogen) to cell−cell recognition processes of non-photosynthetic cells. It is widely recognized that carbohydrates attached to a cell surface and cell proteins (glycoforms or glycoconjugates), such as the acetylated sugar derivatives *N*-acetyl-glucosamine, *N*-acetyl-galactosamine, and *N*-acetyl-neuraminic acid (sialic acid), control how cells communicate with each other and recognize other cells and cell invaders, such as bacteria and viruses. It has been recognized that N-linked glycosylation controls viral virulence and immune evasion from several prominent human pathogens.[1]

The vast amount and diversity of carbohydrate-based structures present in nature requires a large group of enzymes responsible for their synthesis, degradation, and modification.[2] This is the function of carbohydrate-active enzymes (CAZymes) such as glycoside hydrolases (GHs), polysaccharide lyases (PLs), or glycosyltransferases (GTs), which constitute approximately 1−2% of the genome of any organism.[3] These enzymes have a myriad of industrial and biotechnological applications, ranging from biofuel production (search for improved cellulase enzymes for biomass degradation) to drug design (search for new and more potent inhibitors for CAZymes involved in diseases).[4] Lack or malfunction of CAZymes, which in most cases is genetically inherited but can also be acquired, such as lactase misproduction, breaks down cell−cell communication. This, in turn, can lead to serious health problems, from allergies and autoimmune diseases to severe recessive lysosomal storage diseases (e.g., Tay-Sachs, Hunter, Fabry, Gaucher, and Krabbs diseases, caused by lack of hexosaminidase A, iduronate-2-sulfatase, α-galactosidase A, and glucocerebrosidase, respectively), among others. CAZymes, CAZyme inhibitors, and chaperones are used to treat these diseases, but there is still a long way to go before they can be cured.

GTs and GHs constitute the major catalytic machinery for the synthesis and breakage of glycosidic bonds, respectively. Deciphering their reaction mechanisms at a molecular level is crucial to develop new specific enzyme inhibitors. Quantitative modeling of glycosylation reactions, hampered for many years by the lack of structures of CAZyme complexes with their corresponding carbohydrate substrates, is experiencing a renaissance, enhanced by the development of robust theoretical tools such as quantum mechanics/molecular mechanics (QM/MM) techniques, most of them being able to take into account protein dynamics. This has made possible the study of chemical reactions in CAZymes at an unprecedented level of detail, providing insight into the detailed motion of the atoms involved in the chemical reaction.

It is currently recognized that proteins should not be pictured as static objects, but rather as dynamic systems.[5] CAZymes constitute a special case, as not only the enzymes themselves but also the substrates they bind are particularly flexible. In fact, carbohydrates have been compared to "wiggling snakes",[6] as they do not have a single 3D structure and their sugar components might exhibit a myriad of ring conformations. As will be discussed later in this Perspective, the precise conformation of carbohydrate sugar units might have a significant effect in catalysis (e.g., GHs achieve their high catalytic efficiency by changing the conformation of one specific sugar, among other factors). On the other hand, the enzyme scaffold is also highly flexible and dynamic, as exemplified by the significant conformational changes that GTs undergo upon substrate binding.[7] Proper description of the conformations of the carbohydrate ligand, together with the dynamics of the enzyme, represent a challenge for theoretical approaches. Traditionally, static calculations based on geometry optimizations along a pre-defined reaction coordinate (usually monodimensional) and using small models of the enzyme active site were performed to study glycosylation

Received: February 2, 2015
Published: May 13, 2015

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

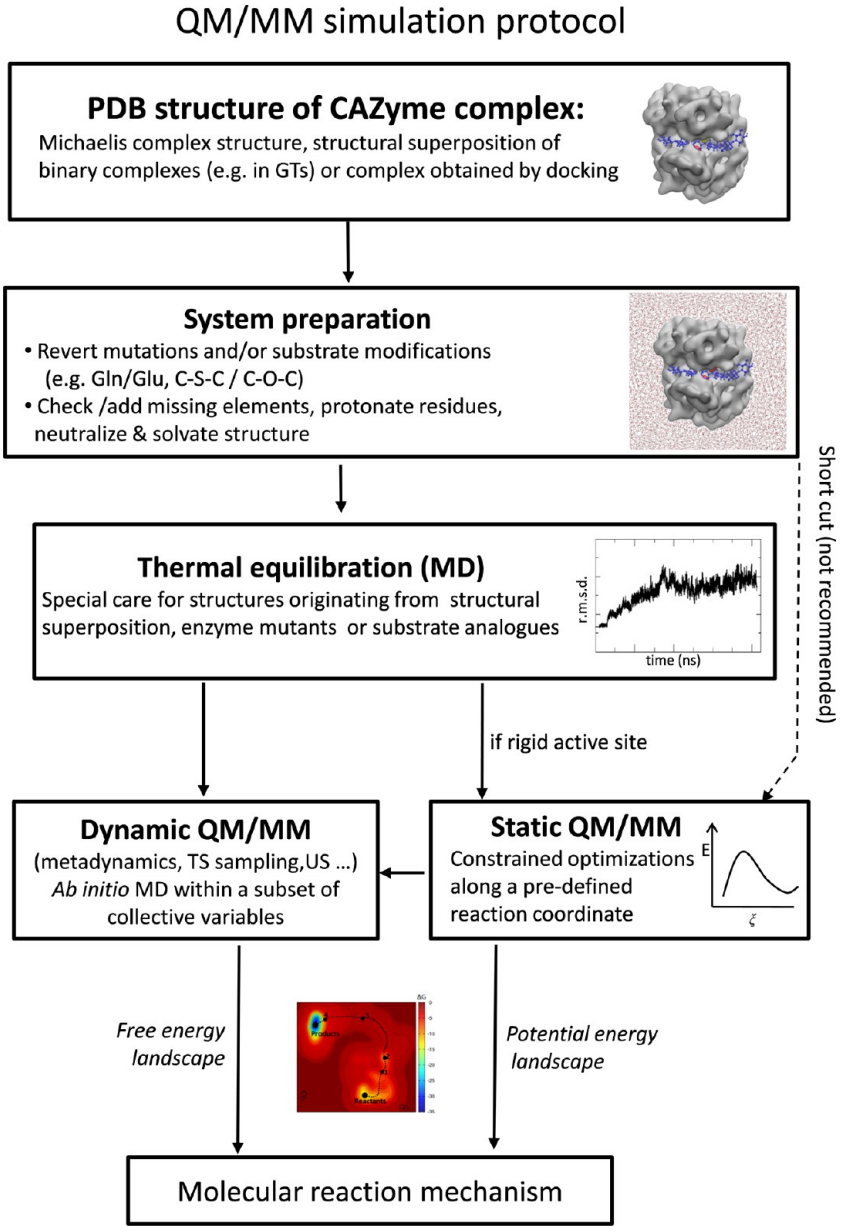


Figure 1. Standard protocol to perform a QM/MM simulation starting from the pristine crystallographic coordinates.

reactions. The first QM study of the lysozyme reaction mechanism confirmed the covalent glycosyl−enzyme intermediate for retaining the GHs used this approach.[8] Similarly, the first Density Functional Theory (DFT) study of the reaction mechanism of galactosyltransferase LgtC,[9] which provided indications that that the unusual front-face mechanism for glycosyl transfer[10] might be feasible in GTs, was performed using a gas-phase model. With the coming-of-age of QM/MM techniques, these mechanisms were reassessed, and new mechanistic details became accessible to computation.[11] QM/MM treatment of enzymatic reactions is now the method of choice,[12] as one can model the chemical reaction in the full enzyme, thus minimizing size effects. These methods can be used within the static approach, but one can also use QM/MM molecular dynamics (MD) simulations,[13] based on DFT (i.e., first-principles QM/MM MD) or a semiempirical model, to solve the electronic structure on the fly. The static approach provides one of the most stable conformations of the system at 0 K, whereas the dynamic approach gives also the different protein conformations at given temperatures. This Perspective gives our point of view on the work that is currently being done on the modeling of catalytic mechanisms in GHs and GTs, with special emphasis on studies performed by first-principles QM/MM dynamic simulations.

2. INITIAL STRUCTURE FOR THE CALCULATIONS: IS THE CRYSTAL STRUCTURE GOOD ENOUGH?

One of the crucial aspects in the modeling of reactive processes in enzymes (CAZymes are far from being an exception) is the initial structure chosen for the calculations. Ideally, one should find a starting structure that corresponds to the portion of the protein's landscape in which one is interested. A high-resolution crystal structure of the reactants' configurations is a good starting point, assuming all reactive species are present (e.g., the glycoside substrate in the case of GHs and the donor/acceptor substrates in the case of GTs). However, subtle motions of active-site residues,

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

essential for catalysis, may not or may only partially be evidenced by analysis of the crystal structures.[14] Atomic positions obtained by X-ray crystallography are time and ensemble averages over many molecules in the crystal,[15] and in some cases they might not fully capture the flexibility of the protein/substrate. As discussed by several authors, in the refinement procedure, heterogeneity (i.e., differences among the different molecules in the crystal) is approximated by a single, average structure with isotropic variance,[16] and correlated motions are neglected. In other words, average structural parameters (distances, angles, etc.) obtained by time averaging of given atomic positions do not correspond to the time average of the given parameter. Classical MD studies have quantified these differences. For instance, it has been found that, for distances shorter than 5 Å, corrections greater than 0.5 Å may apply.[15] Similarly, a first-principles study of the dynamics of a metal−ligand bond (the Fe−O_2 bond in myoglobin) concluded that distances and angles are strongly correlated.[17] As a result, structural parameters taken from the X-ray structure can be affected by errors that, even though not relevant for a description of protein structure, are significant for the purpose of QM/MM modeling. For this reason, the crystal structure needs to be carefully analyzed and thermally equilibrated prior to starting a QM/MM calculation.

Figure 1 illustrates the standard protocol of a QM/MM calculation on CAZymes, which does not differ qualitatively from the protocol used to model any enzyme complex in general. Because the crystal structure does not correspond to the enzyme at a given instant of time but rather a time/ensemble average, the X-ray protein structure needs to be equilibrated thoroughly to bring the model as close as possible to a "true snapshot" of the protein dynamics, or a realistic microscopic state. How much time should the structure be pre-equilibrated? There is no single recipe for all systems, but a consensus seems to be that at least the protein root-mean-square deviation (rmsd) should be converged, which for many systems is normally achieved in a few nanoseconds. Some authors recommend that the final rmsd value of the computed structure must converge to a value within the resolution of the crystal structure,[18] even though the computed rmsd is not directly related with its resolution. It should be also remarked that looking at the rmsd of the simulated system and comparing it to the experimental structure is not the only possible option in assessing the equilibration, and hence the reliability, of the simulated system.

A common strategy when the enzyme−substrate complex structure is not available is to build the complex from structural alignment (e.g., the apoenzyme vs enzyme complexes with different substrates or the structures at different reaction stages[19]). A particular example is the building of the initial structure, a ternary complex (enzyme + donor + acceptor, E:D:A), of a GT enzyme (Figure 2)[20] from structural alignment of two binary complex structures (E:D and E:A). This can be considered as a type of docking, as a new ligand molecule is introduced in a previously empty cavity. In these cases, the initial structure deviates significantly from that of the true ternary complex but can be recovered after MD simulation. However, a long-time scale molecular dynamics might be needed to accommodate the substrates in their binding sites and account for protein conformational changes occurring upon binding (it is not guaranteed that the protein secondary structure, after energy minimization or a short MD, corresponds to a representative state along the reaction pathway). Skipping the MD equilibration step (e.g., just minimizing the energy of the initial structure) is a common shortcut that makes the calculation faster (Figure 1), at the risk of not capturing the reaction mechanism in full detail or obtaining an alternative mechanism.

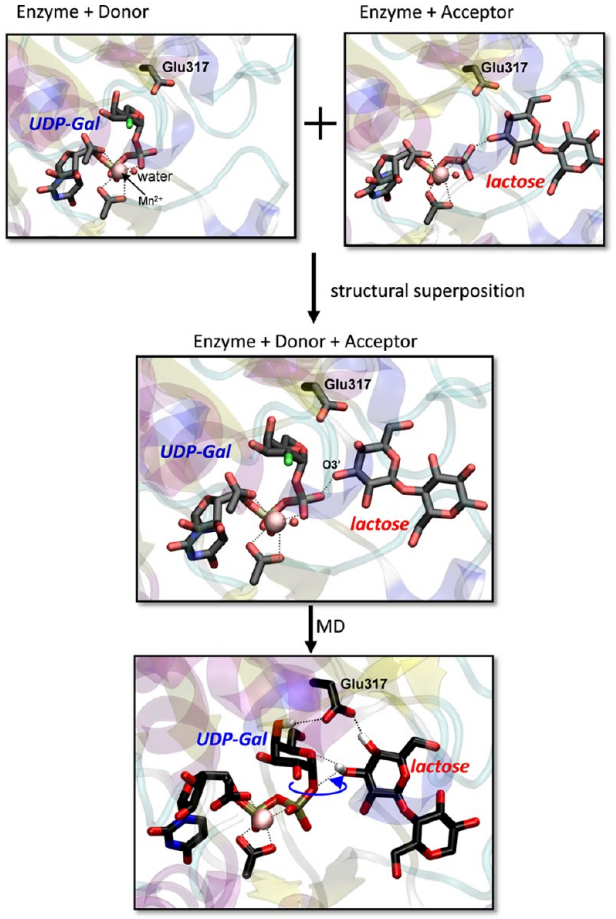


Figure 2. Standard procedure to generate a ternary complex structure of a GT with its two substrates (donor and acceptor) when only structures of binary complexes are available. In the case of α3-galactosyltransferase (α3GalT), the donor is UDP-Gal and the acceptor is lactose. The two binary complexes correspond to the PDB entries 2VFZ (α3GalT + UDP-2F-Gal) and 1GWF (α3GalT + UDP + lactose). The 2-OH of the Gal donor is restored prior to starting MD simulation.

Protein flexibility is increasingly recognized as being essential for protein function.[21] In fact, protein motions that are apparently not involved in the reaction coordinate have been found, in some cases, to promote catalytic turnover.[22] Several authors have proposed that slow vibrational modes of the protein affect the chemical reaction so much as to be part of the reaction coordinate, named the *collective reaction coordinate*[23] (an instructive discussion is provided in ref 24). Therefore, proper description of protein dynamics is an important issue in realistic QM/MM calculations not only at the initial (reactants) state but also along the chemical reaction. Introducing the enzyme environment in the calculation can introduce severe errors if the "environment" is not a realistic one. Needless to say, critical compensation of errors can hide the real chemistry behind the reaction under investigation, even if the most sophisticated methods are used *a posteriori* to describe the electron density.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

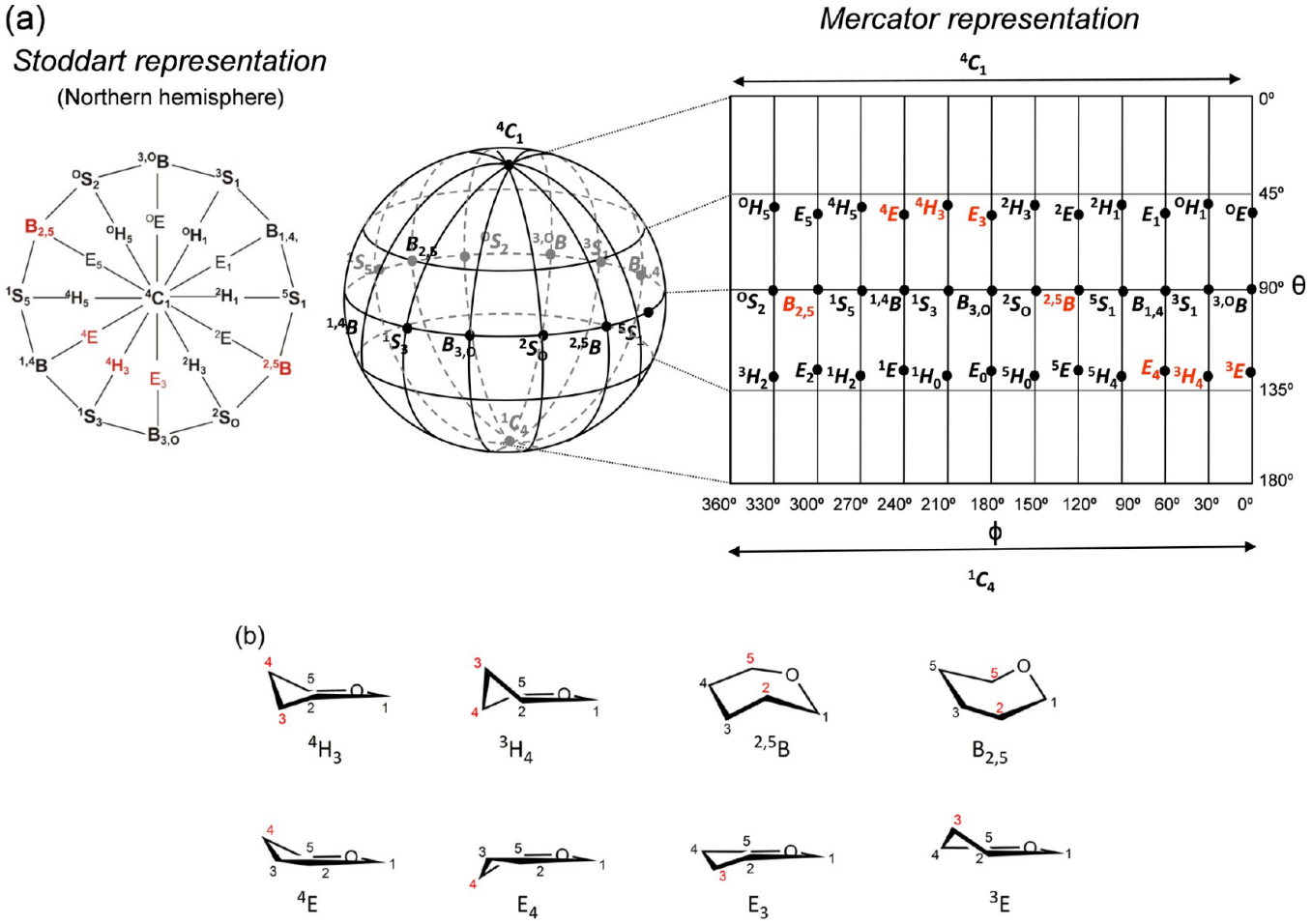


Figure 3. (a) Two of the most used representations of the conformations of a pyranose ring: Stoddart diagram or polar projection (only the Northern projection is represented) and Mercator representation (cylindrical projection). Adapted with permission from ref 29a. Copyright 2012 American Chemical Society. (b) Favored conformations for a pyranose oxocarbenium ion (TS-like conformations).

3. CATALYTIC MECHANISMS OF GLYCOSIDE HYDROLASE: ROLE OF SUBSTRATE CONFORMATIONS

3.1. Conformations of Pyranose Rings. Following IUPAC nomenclature,[25] pyranoses exhibit 38 canonical conformations that can be classified into chair, half-chair, boat, skew-boat, and envelope conformations (represented by the letters C, H, B, S, and E, respectively). Each conformation type exhibits four atoms on the same plane and two out-of-plane atoms (except E conformers, which have only one).

Apart from the letter describing the type of conformation, specific conformations are identified by two indexes indicating the atoms that are out-of-plane (superscript and subscript indicate whether they are up or down, respectively, with respect to the plane).

i. Chair conformers exhibit the two out-of-plane atoms on opposite sides of the reference ring plane, one up and one down. The lowest-numbered carbon atom in the ring must be exoplanar. There are two chairs: $^{4}C_{1}$ and $^{1}C_{4}$.

ii. Boat conformers display the two out-of-plane atoms on opposite sides of the ring plane, either both up or both down. There are six boats: $^{3,O}B$, $B_{1,4}$, $^{2,5}B$, $B_{3,O}$, $^{1,4}B$, and $B_{2,5}$.

iii. Half-chair conformers exhibit four atoms on the plane that are consecutive. There are 12 half-chair conformers: $^{O}H_{5}$, $^{O}H_{1}$, $^{2}H_{1}$, $^{2}H_{3}$, $^{4}H_{3}$, $^{4}H_{5}$, $^{3}H_{2}$, $^{1}H_{O}$, $^{5}H_{4}$, $^{5}H_{O}$, $^{3}H_{4}$, and $^{1}H_{2}$.

iv. Envelope conformers have only one out-of-plane atom.

v. Skew-boat conformers display three atoms on the plane that are consecutive. The lowest-numbered carbon atom in the ring, or the atom numbered next above it, is exoplanar, in that order of preference. There are six skews: $^{3}S_{1}$, $^{5}S_{1}$, $^{2}S_{O}$, $^{1}S_{3}$, $^{1}S_{5}$, and $^{O}S_{2}$.

In 1971, Stoddart introduced a diagram[26] (Figure 3) integrating all possible conformations of a pyranose ring, together with their inter-relations, that has become very popular for representing conformational itineraries of carbohydrates during catalysis.[27] Of course, Stoddart's diagram does not inform on the relative energy of each conformation, but theoretical simulations can be used to obtain this information. It has been shown that canonical conformations on the Stoddart diagram do not necessarily correspond to stable structures. Even for the simplest six-membered ring, cyclohexane, boat conformations are not local minima but transition states between skew-boat conformations.[28]

Stoddart's diagram, along with the so-called plate carrée or Mercator representation (Figure 3, right) of the Cremer–Pople sphere (see next section), constitute simple ways to relate all conformations of a pyranose ring and have proven to be very

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528–7547

operative to discuss substrate distortions in GHs (see refs 27 and 29 for research monographs on the topic).

3.2. Cremer–Pople Puckering Coordinates. Cremer and Pople introduced in 1975 a general definition of ring puckering coordinates.[30] For the particular case of a six-atom sugar ring, any of the possible conformations in Figure 3 can be unequivocally assigned using three polar coordinates: Q, φ, and θ (Figure 4).

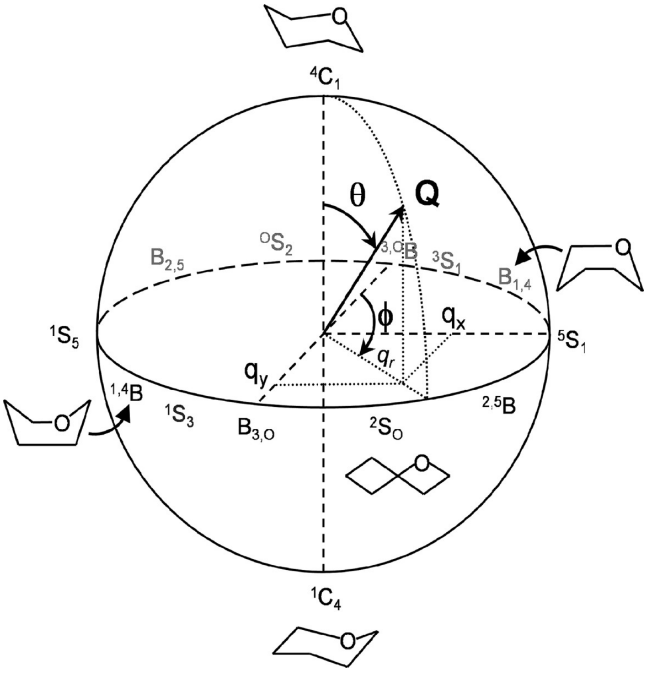


Figure 4. Cremer and Pople puckering coordinates of a pyranose ring (Q, ϕ, and θ) and their projection in the xy plane (q_x and q_y). Adapted with permission from ref 64. Copyright 2010 American Chemical Society.

The Q coordinate is the sum of the perpendicular distance of each ring atom (j) to the ring average plane, $Q = \sum_1^6 Z_j$. The φ, θ, and Q coordinates are obtained by solving the following system of equations:

$$\begin{cases} Q \sin\theta \cos\phi = \sqrt{\frac{1}{3}} \sum_{j=1}^{6} z_j \cos\left[\frac{2\pi}{6} 2(j-1)\right] \\ Q \sin\theta \sin\phi = \sqrt{\frac{1}{3}} \sum_{j=1}^{6} z_j \sin\left[\frac{2\pi}{6} 2(j-1)\right] \\ Q \cos\theta = \sqrt{\frac{1}{6}} \sum_{j=1}^{6} (-1)^{j-1} z_j \end{cases}$$

Any ring conformation would fall within the puckering sphere-like volume of Figure 4. The two chair conformers ($^4C_1$ and $^1C_4$) are located on the poles ($\theta = 0$ or π), whereas the six boat and six skew-boat structures are sequentially placed in steps of $\phi = \pi/6$ on the equator ($\theta = \pi/2$). Thus, the θ and ϕ coordinates are sufficient to differentiate among all conformers.

Two representations (polar or rectangular, Figure 3a) have been historically used to map the Cremer and Pople 3D plot into a simpler two-dimensional plot. Stoddart's diagram[26] (Figure 3a, left) corresponds to the projection of the polar coordinates onto the equatorial plane and can be defined by the Cartesian coordinates q_x and q_y (Figure 4).[31] On the other hand, the Mercator representation (Figure 3a, right) is an equidistant cylindrical projection that results in a rectangular map with respect to θ and ϕ.[32] Which representation to use is a matter of choice. Catalytic GH itineraries are easy to interpret when they are drawn on Stoddart's representation, as there are no discontinuities between conformations (they are all interconnected). Moreover, since GHs normally follow straight itineraries on either the Northern or Southern hemisphere, but not on both (i.e., they do not follow transitions from $^4C_1$ to $^1C_4$ or vice versa), only one diagram is sufficient to study the experimentally relevant conformations, which further simplifies the problem. In contrast, the Mercator representation comprises all conformations on a single diagram, thus mapping the complete Cremer–Pople sphere. This can be most convenient in a more general context, if one is interested in sampling all conformations of the sugar ring.

3.3. Catalytic Mechanisms of Glycoside Hydrolases. The reaction mechanism of GHs, a classical S_N2 reaction, has been the subject of several QM/MM studies. Despite the large number of GH families known (to date 133, classified according to sequence similarities),[3c] they share a common catalytic mechanism: acid/base catalysis with retention or inversion of the anomeric configuration, with the exception of the family GH4.[33] The acid/base reaction is assisted by two essential residues: a proton donor and a nucleophile or general base residue.[34] The former is usually glutamate or aspartate, whereas the latter is glutamic or aspartic acid (remarkable exceptions are sialidases, in which the nucleophile is an activated tyrosine[35]).

Inverting GHs operate by a single nucleophilic substitution (Figure 5a), while retaining GHs follow a double-displacement mechanism via formation and hydrolysis of a covalent intermediate (Figure 5b). Regardless of the type of mechanism, each reaction step involves an oxocarbenium ion-like transition state (TS), as evidenced by kinetic measurements of the isotope effect.[36] Figure 6 illustrates the TS of the first step of the reaction mechanism in retaining GHs. The oxocarbenium ion-like TS is characterized by sp^2 hybridization and development of positive charge at the anomeric carbon, which is partially stabilized by electron donation from the ring oxygen. In this structure, the sugar ring is distorted from the relaxed $^4C_1$ conformation of a pyranose in solution into a conformation in which the C2, C1, O5, and C5 atoms are as coplanar as possible, as required for a stable oxocarbenium ion. Only eight pyranose conformations (hereafter named as TS-like conformations, Figure 3b) conform to this criteria: two boat conformations ($B_{2,5}$ and $^{2,5}B$), two half-chair conformations ($^4H_3$ and $^3H_4$), and four envelope conformations (^3E, E_3, ^4E, and E_4). Therefore, the TS of GH-catalyzed enzymatic reaction is characterized by the sugar ring adopting one of these eight conformations.

Even though the main steps of the GH mechanisms are well established (Figure 5), the specific TS conformations used for several GHs remains uncertain. Typically, the TS conformation is conserved within a given GH family (i.e., similar protein structures accommodate similar sugar conformations). It is also conserved for enzymes acting on a given substrate and stereochemical outcome (e.g., retaining β-glucosidases and retaining β-mannosidases, Table 1), with some exceptions (e.g., β-xylanases use a $B_{2,5}$ or $^1S_3$ TS, depending on the GH family).[28b] Identifying the particular TS conformation for each GH, GH family, or type of substrate is a topic of ongoing interest, as it is essential for the development of TS mimics.[4] Drug molecules such as Relenza and Tamiflu, for instance, which adopt a boat conformation, are potent inhibitors for viral neuraminidases.[4]

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528–7547

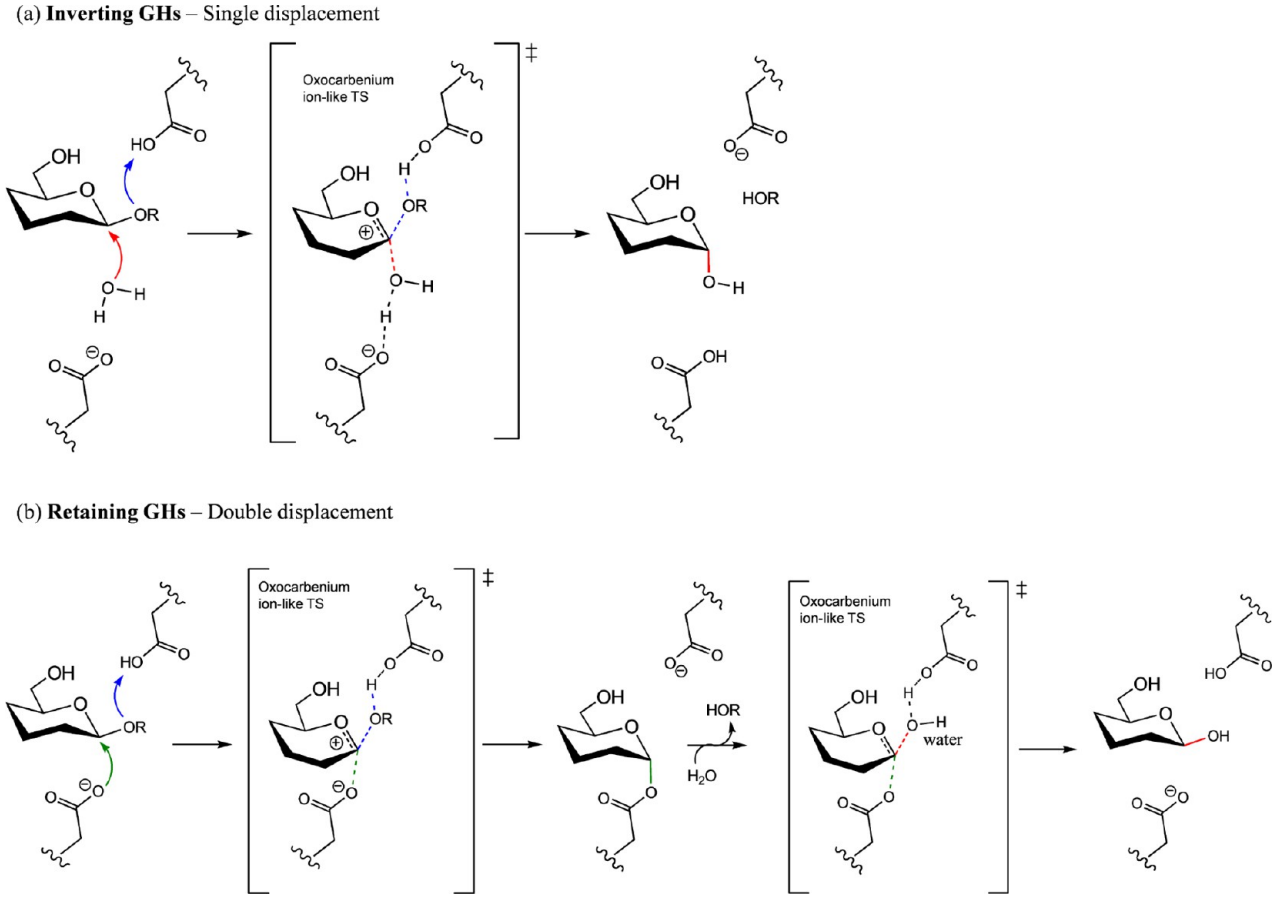


Figure 5. Generalized mechanisms for enzymatic glycoside hydrolysis: (a) inverting GHs and (b) retaining GHs.

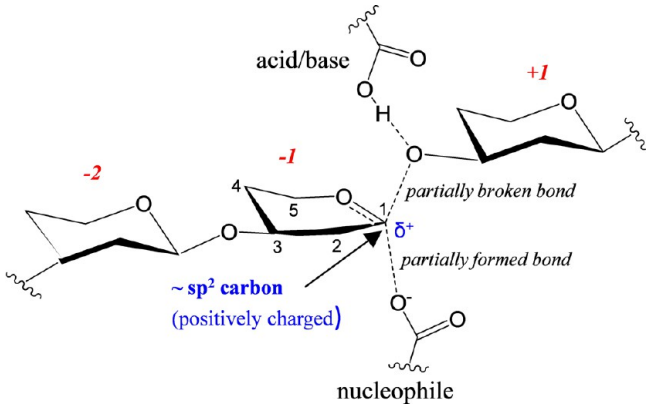


Figure 6. TS of the glycosylation step of the retaining glycoside hydrolases' catalytic mechanism for the particular case of a 1,3 glycosidic bond.

Large efforts are being made to incorporate features of the TS (not only conformation but also a charge at the anomeric center and/or the endocyclic oxygen of the corresponding substrate) into the development of TS inhibitors,[4] surmounting the considerable amount of energy necessary to selectively bind the TS.

3.4. Sugar Distortion in Glycoside Hydrolase Michaelis Complexes. It was found by X-ray[37] and confirmed by NMR[38] studies in the 1990s that the substrate binds to GHs in a distorted conformation. In particular, the saccharide unit binding at subsite −1 (the binding site accommodating the saccharide ring with the scissile glycosidic bond, as in Figure 7a) often adopts a boat- or skew-boat-type conformation (unlike the relaxed $^4C_1$ chair conformation observed for the remaining substrate saccharide rings).

The distortion of the −1 sugar unit in GH complexes was early associated with an enhancement of glycoside hydrolysis, as the distortion places the substrate on the pathway to reach the TS of the reaction. This somehow resembles the so-called "near attack configuration" (NAC) concept introduced by Bruice and Lightstone to explain catalysis in chorismate mutase.[39] In GHs, the distortion of the substrate reduces the steric interaction between the hydrogen and the anomeric carbon (Figure 7b) and places the aglycon (i.e., the leaving group) in a pseudoaxial position. This causes the glycosidic oxygen to approach the acid/base residue and facilitates nucleophilic attack on the anomeric carbon.[40] These features have been confirmed by QM/MM studies,[41] which also demonstrated that the distortion introduces subtle electronic and structural changes that pre-activate the substrate for the hydrolysis reaction.[41a]

Substrate distortion (boat or skew-boat) has recently been observed in a number of Michaelis complexes of GHs.[29] With a few exceptions that remain unexplained,[29a] β-GHs and several α-GHs bind the substrate in a distorted conformation (the distortion is not so advantageous for α-GHs, as the leaving group is already axial for a $^4C_1$ α-sugar). It is important to highlight, however, that structural characterization of GH Michaelis complexes requires crystallizing the enzyme with non-hydrolyzable substrate mimics (see, e.g., ref 37b), using enzyme mutants (e.g., mutating one of the two catalytic residues as in ref 42), or working under conditions in which the enzyme is

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

Table 1. Catalytic Conformational Itineraries Obtained from Structural Analyses of Some Glycoside Hydrolases and Theoretical Confirmation

enzyme group	stereochemical outcome at C1	GH families	general itinerary	evidence
α-mannosidases	retention	GH38, GH76	$^{O}S_2 \rightarrow B_{2,5}^{\ddagger} \rightarrow {}^{1}S_5$	X-ray[112] α-mannose FEL[47] *Michaelis complex simulation:* • QM/MM metadynamics of α-mannoside FEL (GH76)[113] *Reaction simulation:* • QM/MM metadynamics (GH38)[49]
	inversion	GH92 GH47	$^{O}S_2 \rightarrow B_{2,5}^{\ddagger} \rightarrow {}^{1}S_5$ $^{3}S_1 \rightarrow {}^{3}H_4^{\ddagger} \rightarrow {}^{1}C_4$	X-ray[47,114] α-mannose FEL[47] *Michaelis complex simulation:* • QM/MM metadynamics of α-mannoside FEL (GH47)[47]
β-mannosidases	retention	GH2,GH26, GH113	$^{1}S_5 \rightarrow B_{2,5}^{\ddagger} \rightarrow {}^{O}S_2$	X-ray[67b,115] β-mannose FEL[64] mannoimidazole FEL[80]
β-glucosidases	retention	GH5, GH16, GH7	$^{1}S_3 \rightarrow {}^{4}H_3^{\ddagger} \rightarrow {}^{4}C_1$[a]	X-ray[116] β-glucose FEL[31] *Reaction simulations:* • QM/MM metadynamics (GH16)[46] • QM/MM transition path sampling (GH7)[50] • QM/MM[b] adaptive reaction coordinate (GH7)[75]
	inversion	GH8, GH6	$^{2}S_O \rightarrow {}^{2,5}B^{\ddagger} \rightarrow {}^{5}S_1$	X-ray[42,117] β-glucose FEL[31] *Reaction simulation:* • QM/MM metadynamics (GH8)[48]
β-xylosidases	retention	GH10, GH39 GH11, GH120	$^{1}S_3 \rightarrow {}^{4}H_3^{\ddagger} \rightarrow {}^{4}C_1$ $^{2}S_O \rightarrow {}^{2,5}B^{\ddagger} \rightarrow {}^{5}S_1$	X-ray[118] β-xylose FEL[28b] *Michaelis complex simulation:* • QM/MM MD (GH10, GH11)[28b,41b] • Classical MD (GH10, GH11)[119]
	inversion	GH43	$^{2}S_O \rightarrow {}^{2,5}B^{\ddagger} \rightarrow {}^{5}S_1$	X-ray[120] *Reaction simulation:* • QM/MM metadynamics (GH43)[72]
α-L-fucosidases	retention	GH29, GH95	$^{1}C_4 \rightarrow {}^{3}H_4^{\ddagger} \rightarrow {}^{3}S_1$	X-ray[65,121] α-L-fucose FEL[65]
α-galactosidases	retention	GH27	$^{4}C_1 \rightarrow {}^{4}H_3^{\ddagger} \rightarrow {}^{1}S_3$	X-ray[122] *Reaction simulation:* • QM/MM[b] metadynamics[123]
β-galactosidases	retention	GH2	$^{4}H_3 \rightarrow {}^{4}E^{\ddagger} \rightarrow {}^{4}C_1$	*Reaction simulation:* • QM/MM potential energy profile[73]
sialidases	retention	GH33	$^{4}B_{2,5} \rightarrow {}^{4}E^{\ddagger} \rightarrow {}^{4}C_1$	X-ray[35a] *Reaction simulation:* • QM/MM[b] umbrella sampling[35b]

[a]Slightly different itineraries around the $^{4}H_3$ TS conformation were obtained for different enzymes: $^{1,4}B/{}^{1}S_3 \rightarrow [{}^{4}E/{}^{4}H_3]^{\ddagger} \rightarrow {}^{4}C_1$ for GH16 endoglucanase;[46] $^{4}E \rightarrow {}^{4}H_3^{\ddagger} \rightarrow {}^{4}C_1$ and $^{4}H_5 \rightarrow {}^{4}H_3^{\ddagger} \rightarrow {}^{4}C_1$ for GH7 cellobiohydrolase (refs 75 and 50, respectively). [b]A semiempirical model was used to describe the QM atoms. The study in ref 123 was apparently inspired by previous DFT-based work.[48,49]

inactive (e.g., lowering the pH[43]). This naturally raises the question of whether the distortion is also present in catalytically competent crystals of the wild-type (WT) enzyme in complex with the natural substrate—in other words, whether the distortion of the substrate could be a consequence of the structural modifications introduced to avoid the catalytic reaction taking place during the time scale of the experiment (i.e., a sort of "observer effect").[44] First-principles QM/MM simulations have confirmed that the Michaelis complex of several GHs indeed features a distorted substrate.

3.5. First-Principles Quantum Mechanics/Molecular Mechanics Simulations of Glycoside Hydrolase Michaelis Complexes. The first study that quantitatively addressed the problem of substrate distortion in GHs was focused on 1,3−1,4-β-glucanase,[45] a member of the family GH16, in complex with a 4-methylumbelliferyl tetrasaccharide.[41a] QM/MM MD simulations, performed without any restriction on the motion of protein atoms, demonstrated that a distorted conformation of the −1 sugar ring (an intermediate $^{1}S_3/{}^{1,4}B$ conformation) is required in order to bind to the enzyme. Because both distorted and nondistorted (i.e., $^{4}C_1$) conformations correspond to local minima, the properties of the two forms could be compared, revealing that the distorted substrate exhibits peculiar electronic and structural properties that favor the cleavage of the glycosidic bond.[41a] In particular, the glycosidic bond elongates (by 0.06 Å), the intra-ring C1−O5 distance shrinks (by 0.03−0.05 Å), and the anomeric charge increases (from 0.06 to 0.11 electron) in the distorted substrate compared to the nondistorted one.[41a,46] It thus became clearer that the changes induced by substrate distortion (Figure 7b) are in the direction of the TS of the reaction, in which the glycosidic bond is partially broken and the C1−O5 bond acquires partial double bond character (Figure 6).

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

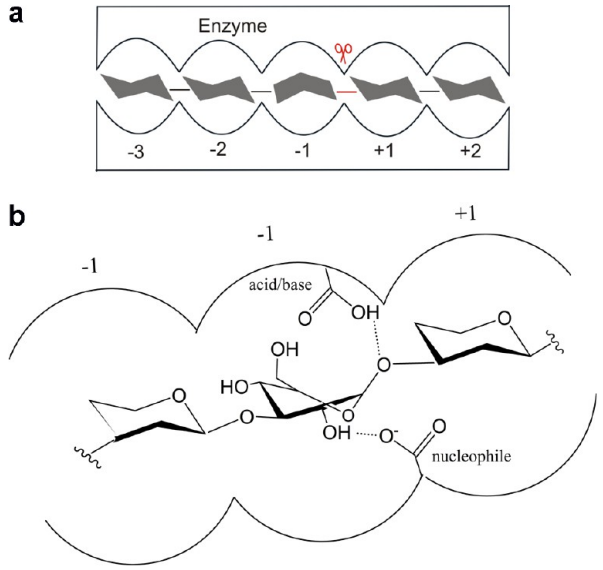


Figure 7. (a) Sugar-binding subsites in glycoside hydrolases. The bond to be cleaved by the enzyme is located between subsites −1 and +1 (indicated with the scissors).[110] In subsites −3, −1, +1, and +2, the sugar adopts the relaxed $^4C_1$ (chair) conformation, whereas in the "catalytic" −1 subsite, the sugar moiety is distorted to one of the eight TS-like conformations (Figure 4b) (boat type in the picture). (b) Sugar distortion at the −1 subsite.

Therefore, pre-organization of the substrate in a distorted conformation ($^1S_3/^{1,4}B$ in 1,3−1,4-β-glucanase) prepares the substrate for the enzymatic reaction, from both structural and electronic points of view.

How is distortion of the substrate achieved by the enzyme? Biarnés et al.[41a,44a] hypothesized that the shape of the binding cavity favors an axial orientation of the sugar leaving group and governs substrate distortion, but interactions with the sugar hydroxyl substituents (especially the strong bond between the 2-OH and the negatively charged catalytic nucleophile) also play a role. It would be extremely interesting to find out whether the change from an all-chair configuration of the substrate in solution (*out-of-enzyme*) to the distorted conformation takes place as the substrate enters the enzyme active site and begins interacting with the catalytic residues or, alternatively, if the change takes place once the substrate is in the active site. QM/MM simulations of conformational changes on-enzyme (section 3.3) show that the energetic cost for a conformational change is significant (>10 kcal/mol) once the substrate is in the active site. Therefore, the first of the two hypotheses seems more plausible.

Several QM/MM MD studies of GH Michaelis complexes have been reported since 2006, including GH47 retaining α-mannosidase,[47] GH8 inverting β-endoglucanase,[48] GH11 retaining β-endoxylanase,[41b] GH38 retaining α-mannosidase,[49] and GH7 cellobiohydrolase Cel7A.[50] These studies have confirmed that the substrate is distorted in the Michaelis complex of these GHs. To our knowledge, no clear evidence of a nondistorted substrate in a β-GH has been reported so far, which seems to indicate that substrate distortion is a general feature of Michaelis complexes of not only β-GHs but also several α-GHs. The modeling studies also confirmed that most X-ray structures of modified enzymes or complexes with non-hydrolyzable substrates are good mimics of the "real" (i.e., distorted) Michaelis complexes. For instance, QM/MM MD simulations revealed that mutation of the acid/base residue in GH8 endoglucanase results in the same substrate distortion ($^2S_O/^{2,5}B$) as in the WT enzyme.[48] Likewise, the $^{3,O}B/^3S_1$ distortion observed in GH47 α-mannosidase in complex with a S-linked disaccharide is the same as that of the natural substrate.[47] However, there are cases in which the opposite is found. For instance, the X-ray structure of the complex of the D204A mutant (D204 being the nucleophile residue) of the GH38 Golgi α-mannosidase II shows a $^4C_1$ conformation for the −1 saccharide, whereas a $B_{2,5}$ conformation was obtained in QM/MM calculations of the WT enzyme complex.[49] Similarly, a complex of GH11 β-xylanase E177Q mutant (E177 being the acid/base residue) was found to exhibit a different conformation than the WT enzyme.[28b] Therefore, one can conclude that enzyme mutation might affect the conformation of the substrate. Recent structures showing undistorted substrate conformations obtained under noncatalytic conditions[43,51] might require further analysis. As pointed out by Speciale et al., structural modifications made to allow kinetic trapping can perturb substrate interactions that are important for defining the conformational catalytic itinerary.[29c]

3.6. Performance of Classical Force-Fields in Predicting Substrate Distortion. Simulations on 1,3−1,4-β-glucanase[45] showed that classical MD (i.e., force-field-based MD) does not reproduce the distortion of the substrate, as only the relaxed $^4C_1$ conformation survived the classical MD simulations. The reason for the apparent lack of substrate distortion is related to the previously discussed changes in anomeric charge and bond lengths that occur upon distortion. Because the anomeric charge changes with ring conformation (it is higher for the distorted $^1S_3/^{1,4}B$ conformation compared with the $^4C_1$ one), it is hard for standard force-fields, which rely on fixed atomic charges, to capture the relative stability of the two conformations of the sugar ring in the enzyme. In fact, when the charge of the anomeric carbon was manually increased (to +0.6 or more), the distorted conformation turned out to be stable under MD. This situation, also observed for GH47 α-mannosidase, might not be general for all GHs and carbohydrate force-fields.[52] In fact, classical MD simulations on Golgi α-mannosidase II reproduced the distortion of the −1 saccharide into a $B_{2,5}$ conformation (the $^4C_1$ conformation observed for the D204A mutant changed spontaneously to $B_{2,5}$).[49] However, it casts a word of caution on the use of classical MD to describe substrate distortion in Michaelis complexes of GHs. In case the −1 sugar ring evolves spontaneously to a $^4C_1$ conformation during energy minimization or MD equilibration, one should wonder whether this is a real change or a consequence of overestimating the relative stability of the chair conformation by the given force-field.

4. SUGAR PUCKERING AND CONFORMATIONAL FREE ENERGY LANDSCAPES

4.1. Conformational Free Energy Landscape of β-Glucose. As discussed in section 3.2, distortion of the substrate in the Michaelis complex of GHs leads to small structural (elongation of the glycosidic bond and shortening of the intra-ring C1−O5 distance) and electronic (increase of the anomeric charge) changes that place the substrate on the pathway to reach the reaction TS. It would be reasonable to assume that the enzyme environment (i.e., both the electrostatic field and the interactions with active-site residues) is responsible for these changes. However, the above-mentioned study on 1,3−1,4-β-glucanase showed that this is not the case, as similar changes were found in calculations of the isolated substrate (i.e., in the absence

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

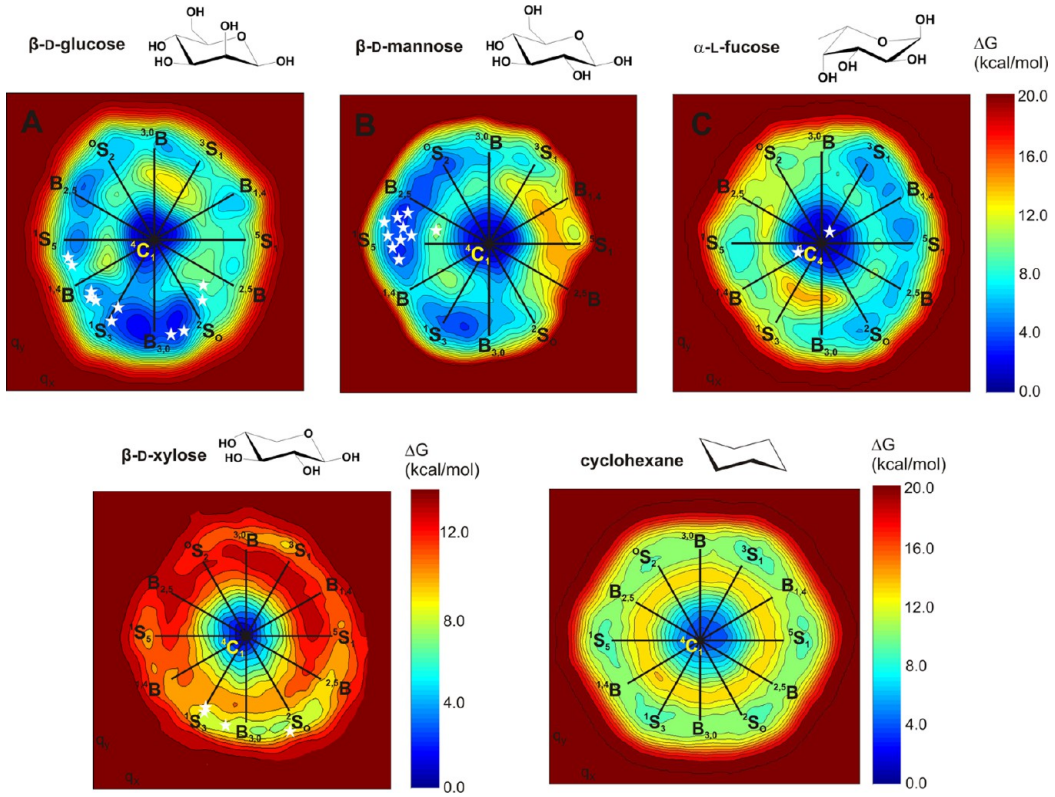


Figure 8. Conformational free energy landscapes (Stoddart representation of the puckering sphere Northern hemisphere) obtained for β-D-glucose,[31] β-D-mannose,[64] α-L-fucose,[65] β-D-xylose,[28b] and cyclohexane.[28b] Each contour line of the diagram corresponds to 0.5 kcal/mol (β-D-glucose, β-D-mannose, and α-L-fucose) and 1 kcal/mol (β-D-xylose and cyclohexane). Star symbols indicate the observed conformations for Michaelis complexes of β-D-glucosidases, β-D-mannosidases, α-L-fucosidases, and β-D-xylosidases. Adapted with permission from ref 64. Copyright 2010 American Chemical Society.

of the protein environment). In particular, the glycosidic bond distance is 0.06 Å longer for a $^{1}S_{3}$ distorted substrate than for a relaxed $^{4}C_{1}$ substrate. Likewise, the C1−O5 distance shrinks by 0.04 Å. Therefore, even though the shape and enzyme−substrate interactions at the −1 and +1 subsites enforce the conformation of the substrate, the small structural and electronic changes observed upon distortion are an intrinsic property of the substrate, and the enzyme has probably evolved to use these properties (e.g., adapting the shape of subsites −1 and +1 to accommodate a distorted substrate) for a more efficient catalysis.

The fact that the "TS-like properties" of the substrate are already observed for an isolated substrate suggests that one could identify which conformations are the best pre-activated for catalysis by quantifying these properties for all conformations of isolated sugars. In other words, one could predict which conformations will be present in GH active sites by analyzing the conformational properties of the carbohydrates they bind. Using first-principles metadynamics (briefly summarized below), Biarnés et al. computed the conformational free energy landscape (FEL) of β-glucose,[31] adapting the Cremer and Pople puckering coordinates as collective variables, for the catalytically relevant Northern hemisphere. Interestingly, the most stable distorted conformations were located on one side of the diagram (south and west in Figure 8A). These conformations feature a small but significant lengthening of the C1−O1 bond, a shortening of the C1−O5 bond, and a development of charge at the anomeric center. Thus, the conformations corresponding to these minima are pre-activated for catalysis. Recent calculations using standard quantum chemistry approaches have confirmed these features.[53] It was compelling to see that the pre-activated conformations correlate with the ones found in X-ray structures of glucoside-active β-GHs. Therefore, the FEL computed by metadynamics can be regarded as a *fingerprint* of the sugar ring that can be used to predict of the occurrence of distorted sugar conformations in β-glucosidases.

The FEL of β-glucose also evidenced that very few of the nine free energy minima correspond to "canonical" conformations in Stoddart's diagram (Figure 8A). This is not surprising for a nonsymmetric molecule that, in addition, exhibits many hydrogen-bonding and repulsive interactions affecting the free energy surface. Similar considerations on the prevalence of noncanonical conformations have also been discussed in a recent study using potential energy searches.[54]

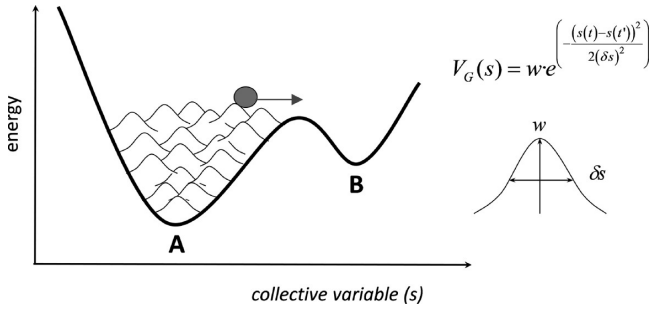
It should be noted that free energy differences obtained for isolated sugars by first-principles methods might depend on the QM method used (or the functional employed within DFT). Nevertheless, their relative order remains usually unchanged. For instance, energy differences for β-glucose were found to vary by ±0.6 kcal/mol using several DFT functionals, within a PW basis set with 70 Ry cutoff. Results obtained using atom-localized basis sets or correlated methods will most likely differ slightly. However, the essence of the analysis, which should be independent of the method used, is that low-energy conformations on a given region of the Stoddart diagram are the best pre-activated ones for catalysis, and they generally correspond to the experimentally observed ones.

4.2. Metadynamics. Metadynamics[55] is a MD technique aimed at enhancing the sampling of the configuration space during the simulation and at estimating the FEL. The method is based on a dimensionality reduction: A set of collective variables,

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

which enclose the essential modes that are associated with the transitions in the analyzed process, are defined (e.g., puckering coordinates in the case of ring distortion or distances/coordination numbers in the case of covalent bond cleavage/formation). Small repulsive potential terms (Gaussian potentials, V_G) are added in the regions of the space that have already been explored (Scheme 1). These repulsive potentials make the

Scheme 1



system escape from already visited regions in phase space to others, as soon as biasing potential counterbalances the underlying free energy.[55,56] The method can be exploited not only for accelerating rare events but also for mapping the FEL, which can be estimated, after a sufficient time, as the negative of the sum of the added Gaussians. This method has recently been applied to a variety of problems in the areas of biophysics, chemistry, and material science.[57]

Suitable collective variables in the metadynamics procedure should fulfill the following requirements:[58] (i) they are explicit functions of the atomic positions, and (ii) they are able to distinguish the different states of the system (e.g., all conformers in Stoddart or Mercator diagrams). The accuracy of free energies computed by metadynamics has been estimated as $\propto(\delta s w/\tau_G)^{1/2}$ (where δs and w are the width and height of the repulsive Gaussian potentials and τ_G is the Gaussian deposition time) when the system reaches the diffusion regime.[59] For chemical reactions, Nair et al.[60] proposed a protocol to obtain the free energy barrier after just one crossing of the TS, by restarting the simulation sometime before this event using Gaussian heights $\leq 1 k_B T$. The results obtained with metadynamics have also been shown to be quantitatively similar to the ones obtained with the umbrella sampling approach (of course, a fair comparison requires that, in both cases, the same number and composition of collective variables are considered), but at a much lower computational cost.[61] Several variations of the original method have been developed that result in higher efficiency. The metadynamics method, which was mathematically proved recently,[62] has been applied with success to decipher enzyme reaction mechanisms since 2006.[63]

4.3. Conformational Free Energy Landscape of Other Monosaccharides: The Pre-activation Index. The analysis of the β-glucose FEL has been extended to other saccharides such as α- and β-mannose,[47,64] α-L-fucose,[65] and, most recently, β-xylose[28b] (all monosaccharides mentioned in this Perspective refer to the D enantiomer, unless specified). Qualitative changes on the FEL were observed for different sugars (Figure 8), which can be traced back to differences in the molecular structures. As shown for β-mannose,[64] relative free energy values alone (ΔG_{rel}) are normally not enough to identify the most pre-activated conformations for catalysis, and relevant changes in structural and electronic parameters should also be considered. A simple

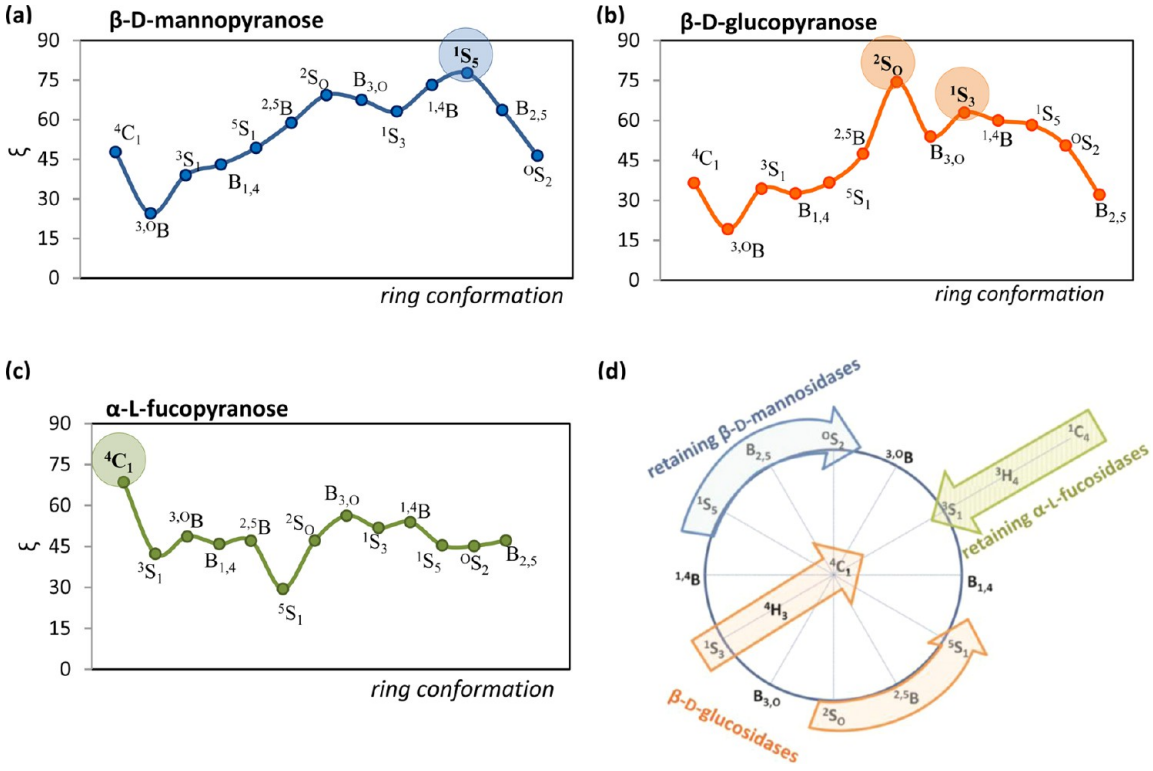


Figure 9. (a–c) Variation of the values of the ξ pre-activation index (as defined in the text) as a function of ring conformation obtained for β-D-glucose, β-D-mannose, and α-L-fucose. (d) Experimentally predicted catalytic conformational itineraries for β-D-glucosidases (orange), β-D-mannosidases (blue), and α-L-fucosidases (green).[64] Reprinted with permission from ref 64. Copyright 2010 American Chemical Society.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528–7547

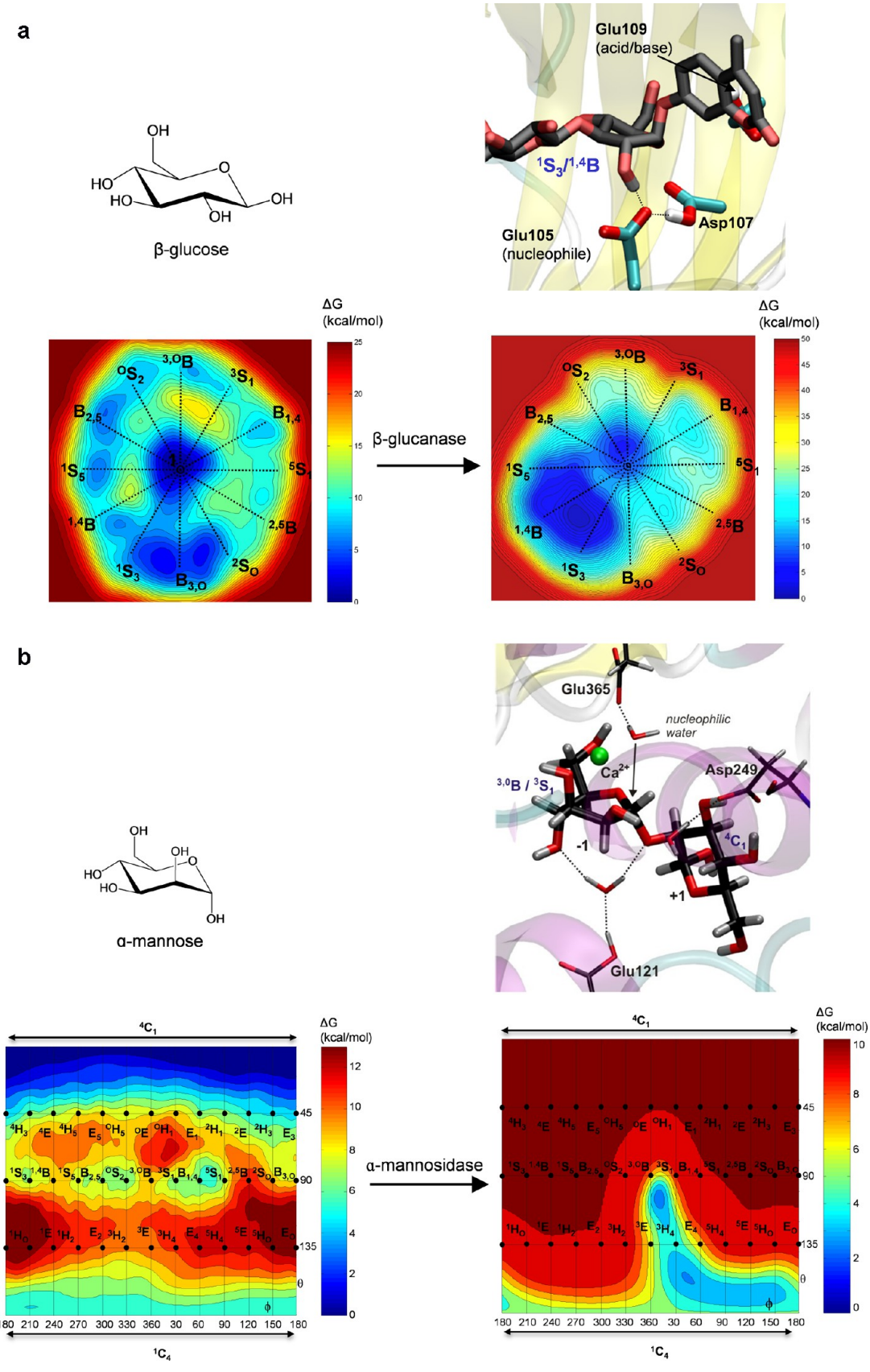


Figure 10. (a) Conformational free energy landscapes (FELs) of isolated β-D-glucose (Stoddart representation of the Northern hemisphere) and the β-glucoside residue at the −1 enzyme subsite of *Bacillus* 1,3−1,4-β-glucanase, contoured at 0.5 kcal/mol. (b) Conformational FELs of isolated α-D-mannopyranose (Mercator projection including both Northern and Southern hemispheres) and the α-mannosyl residue at the −1 enzyme subsite of *Caulobacter* α-1,2-mannosidase, contoured at 1 kcal/mol.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

index integrating several structural, electronic, and energetic properties can be defined to rank all conformations of a given sugar.

The conformational pre-activation index (ξ), introduced by Ardèvol et al.,[64] integrates changes in electronic, structural, and energetic parameters of saccharide rings with pucker conformation.[64] The parameters considered are those that relate the properties of the substrate to the oxocarbenium ion-like TS of the glycosidase reaction.[41a] These comprise the C1−O1 bond orientation (measured as the angle Ω between the C1−O1 bond and the average plane of the saccharide ring), the C1−O1 and the C1−O5 bond lengths; the partial charge at the C1, O1, and O5 atoms, and the conformational relative free energy (Ω, d_{C1-O1}, d_{C1-O5}, q_{C1}, q_{O1}, q_{O5}, and ΔG_{rel}, respectively). For a given j conformation, the ξ score is defined as

$$\xi_j = \frac{1}{n}\sum_{i=1}^{n} \text{score}(x_{i,j})$$

where i sums all the parameters mentioned above, normalized over all conformations:

$$\text{score}(x_{i.j}) = \frac{x_{i,j} - x_{i,j}^{\min}}{x_{i,j}^{\max} - x_{i,j}^{\min}} \times 100 \qquad \text{for}$$

$$x_i = (d_{C1-O1},\ q_{C1},\ q_{O5},\ \Omega)$$

$$\text{score}(x_{i.j}) = \frac{x_{i,j}^{\max} - x_{i,j}}{x_{i,j}^{\max} - x_{i,j}^{\min}} \times 100 \qquad \text{for}$$

$$x_i = (d_{C1-O5},\ q_{O1},\ \Delta G_{rel})$$

Typically, $x_{i,j}$ is calculated as the average value of several optimized structures of a given conformation j. In this formulation, the conformations that display the highest values of ξ are the most likely candidates to be observed in the Michaelis complex of a given GH. Subsequent studies on other saccharide FELs have improved the definition of the index, making it more general.[47] Different weights can be applied to each parameter (e.g., C1−O1 bond length might be more important than the C1−O1 bond orientation, or vice versa), or other parameters could be included when more data are being collected. The analysis of sugar conformations in terms of the ξ index, in its current simple definition, showed that this index can be used to predict the conformation of the substrate in Michaelis complexes of GHs (e.g., $^{2}S_{O}$ and $^{1}S_{3}$ for β-glucosidases,[31] $^{1}S_{5}$ for β-mannosidases,[64] $^{O}S_{2}$ and $^{3}S_{1}$ for α-mannosidases,[47] $^{1}C_{4}$ for α-L-fucosidases,[65] and $^{1}S_{3}$ and $^{2}S_{O}$ for β-xylosidases,[28b] Figure 9). This again demonstrates that sugar distortions present in GH Michaelis complexes are influenced by the conformational preferences of single sugar units.

4.4. Enzyme Reshaping of the Sugar Conformational Free Energy Landscape. To quantify the effect of the enzyme on the conformations of the substrate, one can map the conformational free energy of the substrate in the Michaelis complex, i.e., computing the "on-enzyme" FEL. QM/MM metadynamics simulations on 1,3−1,4-β-endoglucanase showed that the flexibility of the −1 sugar ring is notably reduced in comparison to that of a free glucopyranoside unit (Figure 10a). In addition, whereas the $^{4}C_{1}$ chair is the global minimum of isolated glucopyranose,[31] the sugar ring preferentially adopts a distorted conformation ($^{1,4}B/^{1}S_{3}$) in the E·S complex,[46] with $^{4}C_{1}$ being a higher energy local minimum (Figure 10a). It is thus clear that the enzyme environment restricts the conformational space available for the glucosyl unit, such that only two main conformations are stabilized ($^{4}C_{1}$ and $^{1,4}B/^{1}S_{3}$), and inverts the relative stability of the chair conformer with respect to the most stable distorted structure. A similar analysis on GH47 α-mannosidase showed how the enzyme restricts the conformational space of the mannose unit upon binding (Figure 10b).[47] In spite of the FEL differences between the E·S complex and the isolated sugar, one can easily recognize the on-enzyme conformation among the conformations of the isolated sugar that are pre-activated for catalysis. This highlights the importance of analyzing the FEL of isolated sugars to predict on-enzyme conformations of carbohydrates based on these sugars.

5. SUBSTRATE CONFORMATIONAL CHANGES DURING CATALYSIS: THE CATALYTIC ITINERARY

A fascinating line of research on GHs, and one with a major impact on the design of enzyme inhibitors, is the conformational analysis of reaction pathways within the diverse enzyme families.[29a] The catalytic itinerary of a given GH can be experimentally predicted from the structure of the Michaelis complex and the glycosyl enzyme intermediate (for retaining GHs) or the product complex (for inverting GHs). For instance, structural studies on retaining β-mannosidases from families 2 and 26 have revealed that the β-mannosyl substrate adopts a $^{1}S_{5}$ conformation in the Michaelis complex and an $^{O}S_{2}$ conformation in the covalent intermediate of the hydrolysis reaction.[27b,66] In addition, TS-like inhibitors were found to adopt a $B_{2,5}$ type of distortion when bound to a GH2 β-mannosidase.[67] On the basis of this evidence, it was proposed that retaining β-mannosidases follow a $^{1}S_{5} \rightarrow B_{2,5}^{\ddagger} \rightarrow {}^{O}S_{2}$ catalytic itinerary. Even in cases for which no TS-like inhibitor structures are available, the TS can be guessed from the structures of the Michaelis complex and covalent intermediate structures (or the reaction products in the case of inverting GHs). In these cases, the catalytic itinerary is often inferred, invoking the principle of less nuclear motion,[68] from the shortest path that connects both conformations in Stoddart or Mercator diagrams.

First-principles QM/MM studies have contributed to ascertain some of these itineraries. Moreover, the simulations provide structural, energetic, and electronic details of all the species along the reaction pathway, including the reaction TS. It is experimentally very difficult to obtain such information due to the extremely short lifetime of an oxocarbenium ion-like species.

The double-displacement reaction of glycosidases was the topic of the seminal work by Warshel and Levitt in 1976, in which they set the basis of the QM/MM approach.[69] At that time, there was a debate on whether the oxocarbenium ion is a TS along the pathway toward the covalent intermediate or a reaction intermediate itself. It was later demonstrated by measurements of the kinetic isotope effect that an oxocarbenium-ion species is the TS of the glycosylation reaction,[70] which proceeds via a covalent intermediate.[71] These results were later confirmed by QM and QM/MM calculations.[8,11a]

In recent years, QM/MM simulations have unveiled catalytic itineraries of several families of GHs (Table 1), not only lysozyme[11a] but also GH8 inverting β-glucosidase,[48] GH16 retaining endo-β-glucosidase,[46] GH38 retaining α-mannosidase,[49] GH43 inverting β-xylosidase,[72] GH2 retaining β-galactosidase,[73] GH18 retaining chitinase,[74] and GH7 retaining cellobiohydrolase (Cel7A).[50,75] Most of these studies were performed using first-principles methods (e.g., DFT) for the QM region, but semiempirical methods (mainly AM1, PM3, and scc-

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

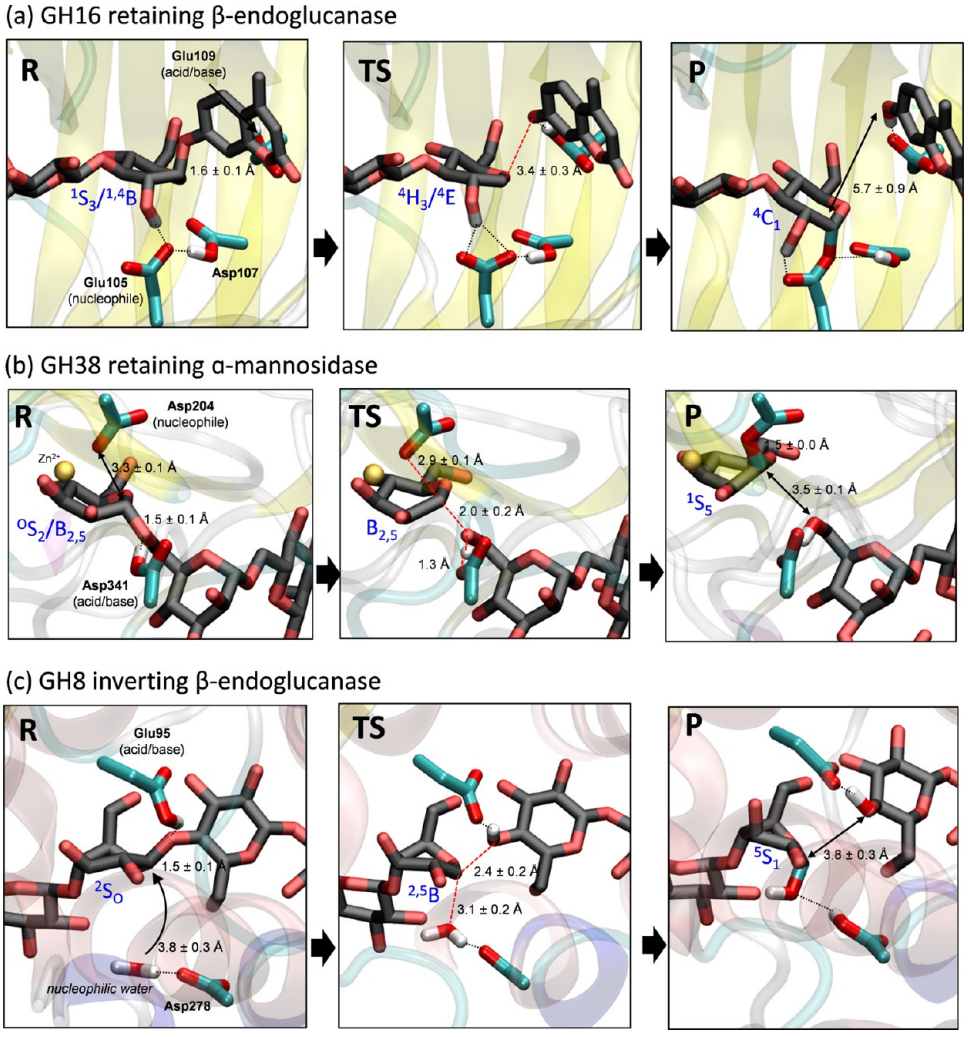


Figure 11. Representative structures along the reaction pathway, corresponding to the stationary points of the reaction free energy surface. (a) For *Bacillus* 1,3−1,4-β-glucanase (family 16 retaining GH). Adapted with permission from ref 46. Copyright 2011 American Chemical Society. (b) For Golgi α-mannosidase II (family 38 retaining GH). Adapted with permission from ref 49. Copyright 2010 American Chemical Society. (c) For *Clostridium thermocellum* endo-1,4-glucanase A (family 8 inverting GH). Adapted with permission from ref 48. Copyright 2009 American Chemical Society.

DFTB) have also been used.[75,76] Jitonnom et al.[74] noted drastic changes in reaction energy barriers (~20 kcal/mol) when DFT corrections were used on top of the semiempirical values, but a recent work seems to contradict this point.[77] In the field of first-principles simulations (e.g., DFT-based) of enzymatic reactions, it is usual to find that computed reaction energy barriers change with the QM method used (even for the same molecular mechanism).[78] Thus, it is important to check properties other than the energy barrier to ensure that the correct molecular mechanism has been captured. For GHs, the catalytic itinerary of the enzymatic reaction should involve one of the TS-like conformations.

QM/MM studies have shown that both the glycosylation step of the hydrolysis mechanism of retaining GHs and the hydrolysis mechanism of inverting GHs feature dissociative oxocarbenium-ion-like TSs, with the glycosidic bond cleavage occurring before nucleophilic attack on the anomeric carbon.[46,72] Figure 11 shows the detailed molecular structure of the main states along the reaction pathway for two retaining GHs and one inverting GH, computed by QM/MM metadynamics using as collective variables the covalent bonds that are formed/broken during the reaction. Although the anomeric carbon acquires positive charge at the TS, it was found by Biarnés et al.[46] that the maximum oxocarbenium ion character of the −1 sugar ring does not coincide with the TS, but takes place afterward. In other words, the TS of the glycosylation reaction corresponds to an early TS with respect to charge development on the anomeric carbon.

An interesting feature of the GH hydrolysis reaction is the shape of the FEL in the products region (i.e., after the oxocarbenium-ion-like TS). For GH16 β-endoglucanase, an intermediate state (P′) was found in which the leaving group is not yet protonated. The presence of this intermediate can be interpreted in terms of the chemical properties of the leaving group. For a relatively good leaving group such as the methylumbelliferyl aglycon (Figure 11a) ($pK_a \approx 7$), the reaction can take place without its protonation, even though protonation is required to complete the reaction until the final products (P). In the case of a poorer leaving group (e.g., a sugar alcohol), one would expect this intermediate to vanish. This was the case, for instance, for the cleavage of two mannosyl residues by Golgi α-mannosidase II (GMII), a family 38 retaining GH (Figure 11b).[49] In contrast, a good leaving group such as *p*-dinitrophenyl is expected to stabilize further the P′ intermediate at the expense of

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

the protonated leaving group. Therefore, the final shape of the reaction FEL in the products region reflects the nature of the leaving group. This facilitates the interpretation of Hammet/Brønsted plots of the GH-catalyzed hydrolysis for distinct glycoside substrates.[79]

The above QM/MM studies, which took into account the dynamics of the enzyme during the chemical reaction, allowed drawing the detailed conformational itinerary that the substrate follows during catalysis. In the particular case of GH16 β-endoglucanase, as well as other GHs, it was found that the main states along the reaction pathway do not correspond to "canonical conformations" in Stoddart's diagram. Moreover, the precise catalytic itinerary ($^{1,4}B/^{1}S_3 \rightarrow [^{4}E/^{4}H_3]^{\ddagger} \rightarrow {}^{4}C_1$, Figure 11a) is not a radial straight line on Stoddart's diagram, as usually assumed, but a warped one. Other retaining β-glucanase enzymes might feature similar itineraries around the characteristic $^{1}S_3 \rightarrow {}^{4}H_3{}^{\ddagger} \rightarrow {}^{4}C_1$ one.

It is interesting that the computed catalytic itinerary for GH16 β-endoglucanase ($^{1,4}B/^{1}S_3 \rightarrow [^{4}E/^{4}H_3]^{\ddagger} \rightarrow {}^{4}C_1$) is similar to the minimum free energy pathway that connects the two main minima of the FEL of the Michaelis complex ($^{1,4}B/^{1}S_3 \rightarrow [E_5/^{4}H_5]^{\ddagger} \rightarrow {}^{4}C_1$), which in turn can be predicted from the FEL of the isolated substrate (Figure 10a). This reassesses the valuable "hidden" information contained in the FEL of the isolated substrate, which informs not only about the conformation of the Michaelis complex but also about the TS-like conformation, and thus the conformational itinerary used for catalysis. This correspondence has been used to assess/exclude conformational itineraries in α-L-fucosidases,[65] β-xylosidases,[28b] and α-mannosidases.[47] The analysis of conformational FELs has also been extended to assess the TS-like properties of GH inhibitors.[80]

6. CATALYTIC MECHANISMS OF GLYCOSYLTRANSFERASES

GTs catalyze the formation of glycosidic linkages by the transfer of a saccharide, typically a monosaccharide, from a sugar nucleotide donor to an acceptor substrate. Compared to GHs, GTs exhibit a much smaller diversity of structural folds. Most GTs fall into one of the two distinct structural folds (GT-A and GT-B, Figure 12),[81] although about 73 families have been found on the basis of sequence similarities.

Approximately 65% of all GTs known use nucleotide-activated sugars (Leloir GTs) and acceptors varying from sugars to lipids and protein residues (O- and N-glycosylation). GTs follow a sequential ordered bi-bi catalytic mechanism, whereby non-covalent binding of the donor substrate to the active site is followed by the binding of the acceptor substrate, yielding the enzyme−substrates ternary complex.[82] To prevent hydrolysis of the activated monosaccharide, the enzyme stays in an inactivated state until the acceptor substrate is bound.[82b] When the ternary complex is formed, the enzyme undergoes a conformational change that activates it for catalysis. In the case of fold-type A GTs, crystal structures of the apo form,[83] as well as complexes with the donor or acceptor substrates,[82b,84] show a substrate-dependent folding of a very flexible N-terminal loop (Figure 12a). Typically, binding of the acceptor substrate folds the N-terminal loop in a partial α helix secondary structure, while in the absence of the substrate the N-terminal loop is disordered and has not been solved by X-ray crystallography. In the case of fold-type B GTs, there is a transition from an "open" to a "closed" state that involves an approach of two large Rossmann fold domains (Figure 12b).[85]

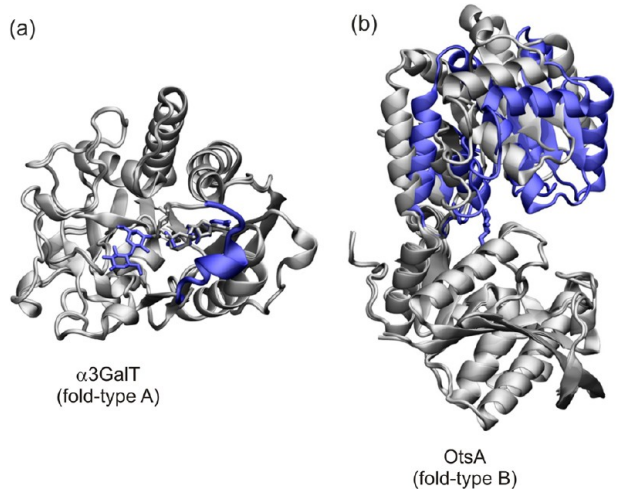


Figure 12. (a) Superposition of the crystal structures of fold-type A α-3-galactosyltransferase in complex with UDP-2F-Gal (PDB 2VFZ)[111] and UDP/lactose (PDB 1GWV).[84] The N-terminal loop is folded when the acceptor substrate is bound in the active site. The N-terminal loop and the lactose molecule are shown in blue. (b) Superposition of the "open" (gray) and the "closed" (blue) structures of fold-type B glycogen synthase, taken from PDB 3FRO.

GTs catalyze the synthesis of the glycosidic bond with two possible stereochemical outcomes: retention or inversion of the anomeric configuration. Inverting GTs operate via an S_N2 reaction in a single displacement step with a general base catalyst that increases the nucleophility of the attacking group (Figure 13a). This is analogous to the mechanism of inverting GHs (Figure 3a), except that the nucleophile is one of the acceptor sugar hydroxyl groups and the leaving group is the phosphate group of the nucleotide. Most inverting GTs exhibit an aspartic or glutamic general acid residue whose side chain serves to partially deprotonate the incoming acceptor hydroxyl group, rendering it a better nucleophile. Measurements of the kinetic isotope effect for β-1,4-GalT[86] and α-1,3-FucT,[87] as well as inhibition experiments,[88] are consistent with the presence of an oxocarbenium-ion-like TS, hence confirming a dissociative S_N2 type of reaction. Theoretical QM/MM studies on inverting β-4Gal-T1[89] and α-GT13[90] have given support to an S_N2 type of mechanism. There is controversy, however, concerning the identity of the catalytic base in the reaction mechanism of O-linked *N*-acetylglucosamine transferase.[91]

In contrast to inverting GTs, the mechanism of retaining GTs is not well understood, with two main mechanisms being currently discussed. By analogy with retaining GHs, a double-displacement mechanism involving a covalent glycosyl−enzyme intermediate was proposed early on (Figure 13b). In the glycosylation step, an aspartate or glutamate side chain in the active site plays the role of a nucleophile, reacting with the anomeric carbon of the donor sugar and forming a glycosyl−enzyme covalent intermediate. In a second step, the acceptor molecule attacks the anomeric carbon, breaking the glycosyl−enzyme covalent bond and forming a new glycosidic bond with overall retention of stereochemistry. Evidence of the formation of covalent glycosyl−enzyme adducts has been reported in mutants of fold type A GTs (family 6 mammalian α3-galactosyltransferase, α3GalT, and blood group GTs) by means of chemical rescue,[92] mass spectrometry,[93] and theoretical calculations on the WT enzyme.[20,94] However, most GT families do not exhibit an appropriately positioned nucleophile within the

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

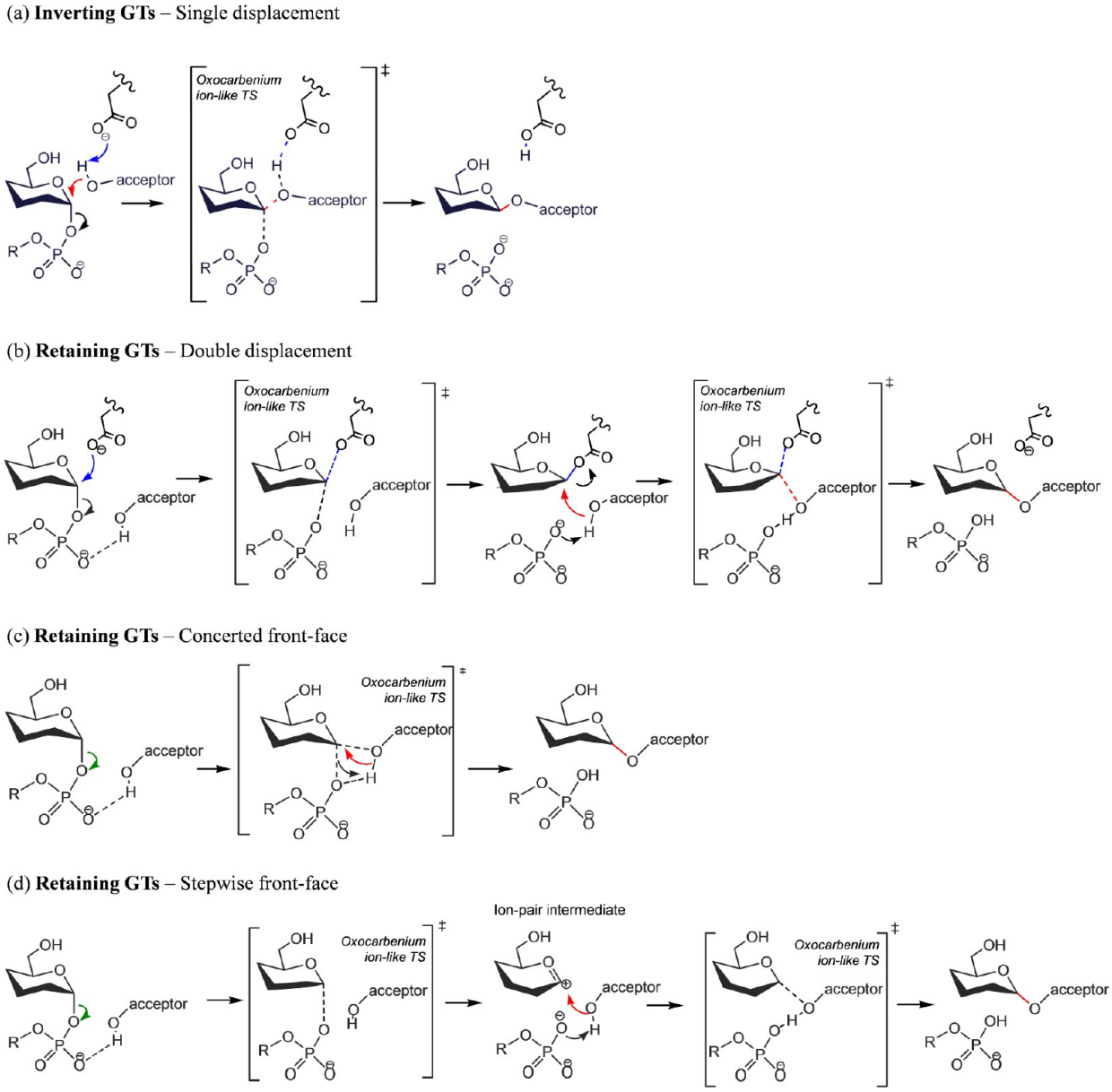


Figure 13. Reaction mechanisms of glycosyltransferases (GTs). (a) S_N2 mechanism for inverting GTs. (b) Double-displacement mechanism for retaining GTs. (c) Front-face mechanism for retaining GTs, considering a fully concerted reaction. (d) Front-face mechanism for retaining GTs, in which a short-lived ion-pair intermediate is formed.

active-site residues. Examples of this large group of GTs are glycogen phosphorylase, a retaining GT that does not use nucleotide sugars,[70] lipopolysaccharyl α-galactosyltransferase (LgtC), and trehalose-6-phosphate synthase (OtsA). This prompted some authors to suggest an unusual mechanism (Figure 13c), in which the reaction proceeds by a front-side single displacement,[95] similar to the solvolysis reaction of glycosyl fluoride that was described as a "type of internal return" by Sinnott and Jencks.[10] In this mechanism, the nucleophilic hydroxyl group of the acceptor attacks the anomeric carbon atom from the same side from which the leaving group departs, thus explaining the retention of stereochemistry. The key aspect of the front-face reaction is that the leaving group and nucleophile need to interact, as the leaving group acts as the catalytic base to deprotonate the incoming nucleophile.

Although the mechanism is often drawn as totally concerted (as in Figure 13c), this is an oversimplification, and most authors agree that the reaction takes place via a short-lived intermediate (as in Figure 13d).[81c]

7. QUANTUM MECHANICS/MOLECULAR MECHANICS CALCULATIONS OF GLYCOSYLTRANSFERASE REACTION MECHANISMS

QM/MM is the preferred method to study GT reaction mechanisms[96] and has contributed to clarify the apparent controversy about the mechanisms among GTs. Since the first QM/MM investigation on a retaining GT,[11b] all theoretical studies have agreed that those enzymes that do not include a carboxylic acid residue in the active site that can serve as a nucleophile follow a front-face type of mechanism (Figure 13c,d). However, essential details of these mechanisms remain unclear.

7.1. Glycosyltransferase Enzymes Exhibiting a Putative Active-Site Nucleophile. To date, only two family 6 GTs, α-

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

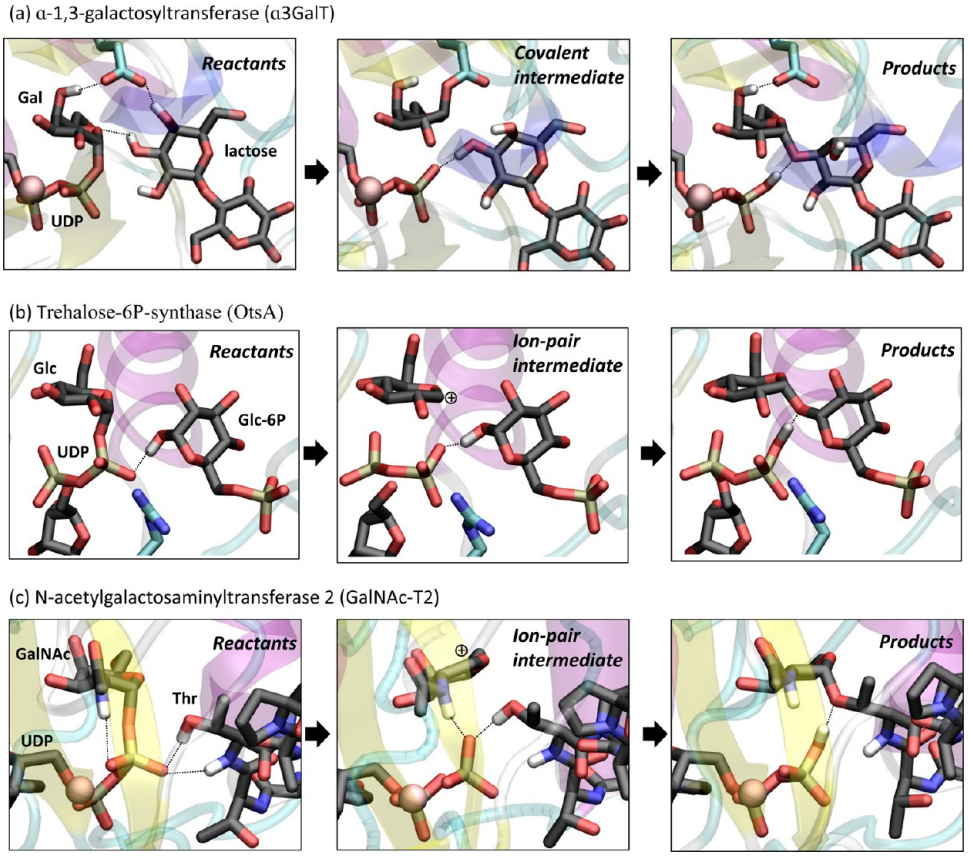


Figure 14. Reaction mechanism for selected retaining glycosyltransferases obtained by QM/MM metadynamics simulations. (a) Double-displacement mechanism of glycosyl transfer catalyzed by mammalian α3GalT. Adapted with permission from ref 20. Copyright 2013 John Wiley & Sons, Inc. (b) Front-face mechanism of glycosyl transfer catalyzed by OtsA. Adapted with permission from ref 11b. Copyright 2011 John Wiley & Sons, Inc. (c) Front-face mechanism of O-glycosylation catalyzed by GalNAc-T2. Adapted with permission from ref 82b. Copyright 2014 John Wiley & Sons, Inc.

1,3-galactosyltransferase (a3GalT) and blood-group A and B α-1,3-glycosyltransferase (GTA/GTB), both enzymes adopting fold type A, have been found to exhibit a nucleophile within the active-site residues (aspartate or glutamate, as in retaining GHs). There is indirect experimental evidence supporting the double-displacement mechanism in both cases. First, mutation of the nucleophilic residue (E317 in α3GalT and E303 in GTA/B) to alanine inactivates both enzymes.[92,93] Second, a chemical rescue experiment with azide on α3GalT E317A mutant yielded the β-galactosyl azide adduct.[92] Finally, a covalent glycosyl−enzyme adduct was observed by mass spectrometry in the E303C mutant of GTA.[93] It was suggested not only that the residue is essential for catalysis but also that a covalent intermediate is formed during the reaction.

Static QM/MM studies on the α3GalT reaction[97] showed that Glu317 was essential for catalysis, either by stabilizing the TS in a front-face mechanism or by forming a glycosyl−enzyme covalent intermediate in a double-displacement mechanism. However, the potential energy barriers obtained were very similar; thus, the authors concluded that both mechanisms compete with similar rates. A subsequent, more accurate analysis[94b] confirmed the competing mechanisms and showed that the highest energy state along the front-face reaction coordinate (whether it is a true TS or an ion-pair intermediate could not be discerned) is stabilized by Glu317. In contrast, dynamic QM/MM simulations[20] on the same enzyme showed that the nucleophile collapses with the anomeric carbon of the sugar donor, which indicates that the double-displacement mechanism is the only feasible pathway for α3GalT (Figure 14a). The GTA/B enzymes have been less studied due to the lack of a good crystal structure of the ternary complex, but a recent study[94c] proposed that the double-displacement mechanism in GTB does not proceed via the typical two-fold S_N2 reactions (Figure 13b) but via two S_N1 reactions, each of them with a stable oxocarbenium ion intermediate. In that mechanism, the covalent glycosyl enzyme adduct of the double-displacement mechanism would dissociate to produce a more stable oxocarbenium ion intermediate. The covalent glycosyl−enzyme intermediate with the E303C mutant was found to be much more stable than the WT enzyme, in agreement with the experimentally trapped adduct.

7.2. Glycosyltransferase Enzymes Lacking an Active-Site Nucleophile. For those enzymes that do not include a suitably positioned carboxylic acid residue in the active site, various residues other than glutamate or aspartate have been proposed to act as a nucleophile,[81c,85c] but such a role is not fully supported by any experimental observation or theoretical calculation. In fact, QM/MM studies on LgtC[98] showed that the double displacement would not be possible. In general, it is assumed that enzymes without a putative nucleophilic residue follow a front-face mechanism, and the focus is on characterizing the reaction path and the features of the chemical species involved.

Early studies on a reduced model of the active site of LgtC[9] concluded that the front-face reaction takes place via a single TS, as depicted in Figure 13b, albeit with a high energy barrier. More recently, a QM/MM study including the whole LgtC environ-

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

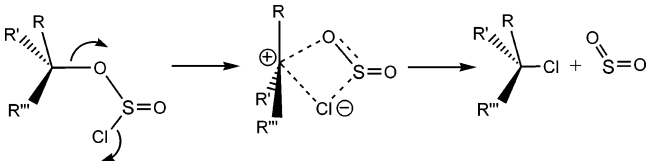
ment and the solvent confirmed the single TS with a more reasonable energy barrier (ca. 12 kcal/mol).[98] The TS was of dissociative character and took place very late in the reaction coordinate, after the cleavage of the UDP-glycosyl bond and concomitant with the nucleophilic attack by the lactose acceptor.

A different energy profile was found for another fold-type B GT, trehalose-6-phosphate synthase (OtsA),[99] by means of dynamic QM/MM simulations, using the metadynamics approach.[100] In contrast to a single TS, a short-lived ion-pair intermediate (Figure 14b) was identified as a minimum in the FEL. Experimental and theoretical calculations for glucosyl transfer in solution, as well as static QM/MM calculations and thorough TS searches on the potential energy of α-mannosyltransferase,[101] were also interpreted in terms of an ion-pair species. Analogously, QM/MM metadynamics simulations identified the ion-pair intermediate in the free energy surface (Figure 14c) for GalNAc-T2, a fold-type A human GT responsible for the post-translational modification of many cell-surface proteins.[102] However, no ion-pair intermediate was found by static QM/MM calculations on the same enzyme.[82b] It remains to be investigated why different QM/MM models obtain different results. It could be argued that the fine details of the mechanism in the region of the ion-pair free energy surface are just a matter of nuance. However, in our opinion it is of fundamental importance to establish whether the reaction is one-step (Figure 13c) (i.e., a concerted mechanism) or stepwise (Figure 13d).

QM/MM calculations have also helped to analyze other important aspects of the glycosyl-transfer reaction, such as the effect of certain mutations in the active-site residues,[102] the implications of substrate−substrate interactions in the reaction mechanism and substrate selectivity,[94b] or the role of metal ions.[101]

7.3. S_Ni, S_Ni-like, or S_N1-like? The terms used to name the front-face mechanism demand particular attention. As mentioned previously, the chemical antecedents of this reaction go back to the internal return mechanism described by Lewis and Boozer for the decomposition of alkylchlorosulfites.[103] The reaction was termed S_Ni (nucleophilic substitution with internal return), as the "returning" nucleophile results from the decomposition of the leaving group (Scheme 2). Later on,

Scheme 2



Sinnott and Jencks applied the same concept to the retention of configuration observed for the solvolysis reaction of glycosyl fluoride[104] in a mixture of ethanol and trifluoroethanol. They described the reaction, with retention of configuration, as "a type of internal return",[103] reflecting the observation that the nucleophile and the leaving group interacted to aid nucleophilic attack (see also discussion on page 24 of ref 81c).

Concerning GTs, several authors have adopted the term S_Ni to mean "front-face reaction",[70,95,105] even though the reaction is obviously different from that of alkyl chlorosulfites (the main one being that the nucleophile is not internal but external). However, other authors use the names S_Ni-like or S_Ni-type.[99] The terminology is somewhat confusing, as some authors use these terms indistinctly (S_Ni, S_Ni-like, or S_Ni-type) to refer to a stepwise front-face reaction (Figure 13d),[99,105], while others, especially the theoreticians,[98] use them to distinguish a concerted reaction (commonly named S_Ni) from a stepwise reaction (S_Ni-like) (Figure 13c,d, respectively). Since the main difference between the S_Ni and the S_N1 mechanisms is the absence or presence of a *separated* ion-pair intermediate, the reaction could also be termed (and this would avoid comparison with the alkylchlorosufites) as S_N1-like, as a carbocation (although short-lived) is formed. For the sake of clarity, in this Perspective we have avoided the labels S_Ni, S_Ni-like, and S_N1-like but refer to a concerted or stepwise front-face reaction.

8. SUMMARY AND CONCLUSIONS

The enzymatic hydrolysis and synthesis of the glycosidic bond are catalyzed by diverse enzymes generically termed glycoside hydrolases or glycosidases (GHs) and glycosyltransferases (GTs), respectively. First-principles QM/MM simulations contribute to clarifying the mechanisms underlying catalysis in these enzymes, as exemplified by the works discussed in this Perspective.

One of the greatest challenges of theoretical approaches to describe CAZyme mechanisms is describing enzyme flexibility, and different studies have addressed the problem of disentangling the molecular mechanism of catalysis by taking into account the dynamics of the enzyme to different degrees, from not describing at all the dynamics of the enzyme before or during the chemical reaction, i.e., the full static approach, to partially describing enzyme dynamics before and during the chemical reaction, i.e., the dynamic approach. Because CAZymes are flexible enzymes binding particularly flexible substrates, we believe that a fair description of the dynamics of the system is necessary to capture the reaction mechanism in full.

A current issue in the research on GHs, and one that has a major impact on the design of enzyme inhibitors, is the conformational analysis of reaction pathways within the diverse families.[29a] Enzymatic hydrolysis of the glycosidic bond in a carbohydrate may occur with either inversion or retention of the configuration of the anomeric carbon. Inversion involves a single step and TS, whereas retention is a two-step process with two TSs (a double-displacement reaction). The TSs of both inverting and retention reactions have been shown to have oxocarbenium ion character, which implies that the sugar ring is distorted in one of the possible TS-like conformations (Figure 3b). Such distortion is already evidenced in the Michaelis complex, where the −1 saccharide is also distorted away from its lowest energy solution chair conformation and is *pre-activated* for catalysis. Calculations of the FELs of isolated saccharides have demonstrated that the conformations that the substrate attains on-enzyme are related to the intrinsic conformational preferences of the glycoside ligand, and the enzyme has adapted the active site to recognize the best pre-activated conformation for catalysis. Several theoretical studies have characterized the TS (conformation and energetics) for several GH families, as well as the complete itinerary that the substrate follows during catalysis. A promising future direction in this field is to use conformational insight in the design of enzyme inhibitors, and inhibitor conformational FELs[80] can be a useful tool to achieve this goal.

Whereas the double-displacement mechanism for retaining GHs is well established, the mechanism of retaining GTs has been surrounded by controversy, as two distinct reaction mechanisms—a double displacement (two S_N2 reactions) and a front-face mechanism—have been proposed. Different enzyme

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

families seem to have evolved to follow either mechanism, depending on the presence or absence of a putative nucleophile residue near the anomeric carbon of the donor sugar. This might not be the only factor that determines the mechanism, as evidenced by experiments on LgtC and GalNAc-T2 inactivation by N189D and N335D mutation,[106] respectively, as well as the remarkable increase of the reaction energy barrier when the front-face mechanism was enforced in α3GalT by QM/MM modeling.[20] Nevertheless, different types of QM/MM approaches (static and dynamic) have led to different conclusions about the details of the front-face mechanism in GTs without a putative nucleophile, as well as the competition between front-face and double-displacement mechanisms in nucleophile-containing GTs. In fact, a concerted mechanism (Figure 13c) has only been obtained for LgtC using the full static QM/MM approach,[98] whereas calculations performed after protein equilibration obtained either a more planar energy profile or an ion-pair intermediate.[82b,94c,100,101] In the case of nucleophile-containing GT, static QM/MM calculations[94b,c,97] showed a competition between front-face and double-displacement mechanisms, whereas dynamic QM/MM simulations[20] predict that only the double-displacement mechanism is possible. Further work is necessary to solve these issues.

The two GT mechanisms, double-displacement and front-face, can be seen as two sides of the same coin, and one prevails over the other, depending on the electrostatic environment and nucleophilicity of the active site.[20] To reach a compromise between catalytic activity and protection against hydrolysis, the nucleophilicity of the active site must be properly tuned. It has been suggested,[81c] and proved by QM/MM calculations,[11b,82b,94b,96] that the leaving diphosphate group plays the role of a base catalyst, activating the incoming acceptor hydroxyl group for nucleophilic attack, and that the acceptor substrate assists the departure of the leaving phosphate. Therefore, hydrolysis is precluded in the absence of the right donor and acceptor substrates.

The underlying reason why different families have evolved to use a specific type of mechanism remains unknown. As in the case of GHs, the solution of this fundamental problem could be hidden in the intrinsic properties of the donor/acceptor substrates. It is also likely that family 6 GTs need the presence of a nearby glutamate residue to better orient the 3-OH of the acceptor in the active site. We can foresee that, in the next years, MD and hybrid QM/MM calculations will continue deciphering these complex and fascinating mechanisms, providing hints for enzyme engineering. On a larger scale, constant developments on classical simulations and enhanced sampling methods will enable simulations of large protein conformational motions that are essential to understand the regulatory mechanisms in GTs, substrate binding, and product release. The development of novel approaches (GPU accelerated electronic structure theory[107] and novel linear-scaling DFT[108]) able to treat full proteins at a QM level is also particularly promising. Although these methods cannot account for protein flexibility (sampling) yet, recent calculations of T4 lysozyme mutants are encouraging.[109]

AUTHOR INFORMATION

Corresponding Author

*c.rovira@ub.edu

Present Address

[#]A.A.: Max-Planck Institut für Biophysik, Max-von-Laue-Straße 3, 60438 Frankfurt am Main, Germany

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We are indebted to our former and current collaborators, X. Biarnés, J. Iglesias-Fernández, V. Rojas-Cervellera, L. Raich, L. Petersen, A. Laio, M. Parrinello, P. Reilly, A. Planas, R. Hurtado-Guerrero, G. J. Davies, and S. J. Williams, for previous work that is reviewed here. Our research is supported by Generalitat de Catalunya (grant 2014SGR-987) and Ministerio de Economía y Competitividad (MINECO) (grant CTQ2014-55174-P). We acknowledge the computer support, technical expertise, and assistance provided by the Barcelona Supercomputing Center–Centro Nacional de Supercomputación (BSC–CNS).

REFERENCES

(1) Vigerust, D. J.; Shepherd, V. L. *Trends Microbiol.* **2007**, *15*, 211–218.

(2) Larsbrink, J.; Rogers, T. E.; Hemsworth, G. R.; McKee, L. S.; Tauzin, A. S.; Spadiut, O.; Klinter, S.; Pudlo, N. A.; Urs, K.; Koropatkin, N. M.; Creagh, A. L.; Haynes, C. A.; Kelly, A. G.; Cederholm, S. N.; Davies, G. J.; Martens, E. C.; Brumer, H. *Nature* **2014**, *506*, 498–502.

(3) (a) Davies, G. J.; Gloster, T. M.; Henrissat, B. *Curr. Opin. Struct. Biol.* **2005**, *15*, 637–645. (b) Hart, G. W.; Copeland, R. J. *Cell* **2010**, *143*, 672–676. (c) Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. *Nucleic Acids Res.* **2014**, *42*, D490–495.

(4) Gloster, T. M.; Vocadlo, D. J. *Nat. Chem. Biol.* **2012**, *8*, 683–694.

(5) (a) Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964–972. (b) McCammon, J. A.; Karplus, M. *Acc. Chem. Res.* **1983**, *16*, 187–193. (c) Hammes-Schiffer, S.; Benkovic, S. J. *Annu. Rev. Biochem.* **2006**, *75*, 519–541.

(6) Woods, R. J.; Tessier, M. B. *Curr. Opin. Struct. Biol.* **2010**, *20*, 575–583.

(7) (a) Urresti, S.; Albesa-Jove, D.; Schaeffer, F.; Pham, H. T.; Kaur, D.; Gest, P.; van der Woerd, M. J.; Carreras-Gonzalez, A.; Lopez-Fernandez, S.; Alzari, P. M.; Brennan, P. J.; Jackson, M.; Guerin, M. E. *J. Biol. Chem.* **2012**, *287*, 24649–24661. (b) Albesa-Jove, D.; Giganti, D.; Jackson, M.; Alzari, P. M.; Guerin, M. E. *Glycobiology* **2014**, *24*, 108–124. (c) Lira-Navarrete, E.; de las Rivas, M.; Compañon, I.; Pallarés, M. C.; Kong, Y.; Iglesias-Fernández, J.; Bernardes, G. J. L.; Peregrina, J. M.; Rovira, C.; Bernadó, P.; Bruscolini, P.; Clausen, H.; Lostao, A.; Corzana, F.; Hurtado-Guerrero, R. *Nat. Commun.* **2015**, In press.

(8) Bottoni, A.; Miscione, G. P.; De Vivo, M. *Proteins* **2005**, *59*, 118–130.

(9) Tvaroska, I. *Carbohydr. Res.* **2004**, *339*, 1007–1014.

(10) Sinnott, M. L.; Jencks, W. P. *J. Am. Chem. Soc.* **1980**, *102*, 2026–2032.

(11) (a) Bowman, A. L.; Grant, I. M.; Mulholland, A. J. *Chem. Commun.* **2008**, 4425–4427. (b) Ardèvol, A.; Rovira, C. *Angew. Chem., Int. Ed.* **2011**, *50*, 10897–10901.

(12) (a) Gao, J.; Truhlar, D. G. *Annu. Rev. Phys. Chem.* **2002**, *53*, 467–505. (b) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229. (c) Lonsdale, R.; Ranaghan, K. E.; Mulholland, A. J. *Chem. Commun.* **2010**, *46*, 2354–2372.

(13) Brunk, E.; Rothlisberger, U. *Chem. Rev.* **2015**, DOI: 10.1021/cr500628b.

(14) Garcia-Viloca, M.; Poulsen, T. D.; Truhlar, D. G.; Gao, J. *Protein Sci.* **2004**, *13*, 2341–2354.

(15) Kuzmanic, A.; Kruschel, D.; van Gunsteren, W. F.; Pannu, N. S.; Zagrovic, B. *J. Mol. Biol.* **2011**, *411*, 286–297.

(16) (a) DePristo, M. A.; de Bakker, P. I.; Blundell, T. L. *Structure* **2004**, *12*, 831–838. (b) Kuzmanic, A.; Pannu, N. S.; Zagrovic, B. *Nat. Commun.* **2014**, *5*, 3220.

(17) Rovira, C.; Parrinello, M. *Biophys. J.* **2000**, *78*, 93–100.

(18) Boero, M. *J. Phys. Chem. B* **2011**, *115*, 12276–12286.

(19) Sanchez-Martinez, M.; Marcos, E.; Tauler, R.; Field, M.; Crehuet, R. *J. Phys. Chem. B* **2013**, *117*, 14261–14272.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528–7547

(20) Rojas-Cervellera, V.; Ardèvol, A.; Boero, M.; Planas, A.; Rovira, C. *Chemistry* **2013**, *19*, 14018−14023.
(21) (a) McCammon, J. A. In *Simplicity and complexity in proteins and nucleic acids*; Frauenfelder, H., Deisenhofer, J., Wolynes, P. G., Eds.; Dahlem University Press: Berlin, 1999; Vol. *83*, pp 193−198. (b) Hammes-Schiffer, S. *Biochemistry* **2013**, *52*, 2012−2020.
(22) (a) Ivanov, I.; Tainer, J. A.; McCammon, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1465−1470. (b) Hammes-Schiffer, S. *Biochemistry* **2002**, *41*, 13335−13343. (c) Agarwal, P. K.; Billeter, S. R.; Rajagopalan, P. T.; Benkovic, S. J.; Hammes-Schiffer, S. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2794−2799.
(23) (a) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. *Nature* **2007**, *450*, 913−916. (b) Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D. *Nature* **2007**, *450*, 838−844. (c) Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E. *Science* **2011**, *332*, 234−238. (d) Pu, J.; Gao, J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 3140−3169.
(24) Moliner, V. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 15013−15014.
(25) IUPAC. *Eur. J. Biochem.* **1980**, *111*, 295−298.
(26) Stoddart, J. F. *Stereochemistry of carbohydrates*; Wiley-Interscience: New York, 1971.
(27) (a) Davies, G. J.; Ducros, V. M.; Varrot, A.; Zechel, D. L. *Biochem. Soc. Trans.* **2003**, *31*, 523−527. (b) Vocadlo, D. J.; Davies, G. J. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539−555.
(28) (a) Stortz, C. A. *J. Phys. Org. Chem.* **2010**, *23*, 1173−1186. (b) Iglesias-Fernandez, J.; Raich, L.; Ardèvol, A.; Rovira, C. *Chem. Sci.* **2015**, *6*, 1167−1177.
(29) (a) Davies, G. J.; Planas, A.; Rovira, C. *Acc. Chem. Res.* **2012**, *45*, 308−316. (b) Satoh, H.; Manabe, S. *Chem. Soc. Rev.* **2013**, *42*, 4297−4309. (c) Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J. *Curr. Opin. Struct. Biol.* **2014**, *28C*, 1−13.
(30) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(31) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686−10693.
(32) Dowd, M. K.; French, A. D.; Reilly, P. J. *Carbohydr. Res.* **1994**, *264*, 1−19.
(33) Jongkees, S. A.; Withers, S. G. *Acc. Chem. Res.* **2014**, *47*, 226−235.
(34) White, A.; Rose, D. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 645−651.
(35) (a) Amaya, M. F.; Watts, A. G.; Damager, I.; Wehenkel, A.; Nguyen, T.; Buschiazzo, A.; Paris, G.; Frasch, A. C.; Withers, S. G.; Alzari, P. M. *Structure* **2004**, *12*, 775−784. (b) Pierdominici-Sottile, G.; Horenstein, N. A.; Roitberg, A. E. *Biochemistry* **2011**, *50*, 10150−10158.
(36) Zechel, D. L.; Withers, S. G. *Acc. Chem. Res.* **2000**, *33*, 11−18.
(37) (a) Tews, I.; Perrakis, A.; Oppenheim, A.; Dauter, Z.; Wilson, K. S.; Vorgias, C. E. *Nat. Struct. Biol.* **1996**, *3*, 638−648. (b) Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schulein, M.; Davies, G. J. *Biochemistry* **1996**, *35*, 15280−15287.
(38) Espinosa, J. F.; Montero, E.; Vian, A.; Garcia, J. L.; Dietrich, H.; Schmidt, R. R.; Martin-Lomas, M.; Imberty, A.; Canada, F. J.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120*, 1309−1318.
(39) (a) Bruice, T. C.; Lightstone, F. C. *Acc. Chem. Res.* **1999**, *32*, 127−136. (b) Schowen, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11931−11932.
(40) Kirby, A. J. *Acc. Chem. Res.* **1984**, *17*, 305−311.
(41) (a) Biarnés, X.; Nieto, J.; Planas, A.; Rovira, C. *J. Biol. Chem.* **2006**, *281*, 1432−1441. (b) Soliman, M. E.; Ruggiero, G. D.; Pernia, J. J.; Greig, I. R.; Williams, I. H. *Org. Biomol. Chem.* **2009**, *7*, 460−468.
(42) Guerin, D. M.; Lascombe, M. B.; Costabel, M.; Souchon, H.; Lamzin, V.; Beguin, P.; Alzari, P. M. *J. Mol. Biol.* **2002**, *316*, 1061−1069.
(43) Hill, C. H.; Graham, S. C.; Read, R. J.; Deane, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 20479−20484.
(44) (a) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C. *Biocatal. Biotransform.* **2010**, *28*, 33−40. (b) Ardèvol, A., *Study of molecular mechanisms in glycoside hydrocases and transferases by ab initio molecular dynamics*, Thesis, Autonomous University of Barcelona, Spain, 2012.
(45) Viladot, J. L.; de Ramon, E.; Durany, O.; Planas, A. *Biochemistry* **1998**, *37*, 11332−11342.
(46) Biarnés, X.; Ardèvol, A.; Iglesias-Fernández, J.; Planas, A.; Rovira, C. *J. Am. Chem. Soc.* **2011**, *133*, 20301−20309.
(47) 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.
(48) Petersen, L.; Ardèvol, A.; Rovira, C.; Reilly, P. J. *J. Phys. Chem. B* **2009**, *113*, 7331−7339.
(49) Petersen, L.; Ardèvol, A.; Rovira, C.; Reilly, P. J. *J. Am. Chem. Soc.* **2010**, *132*, 8291−8300.
(50) Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Gotz, A. W.; Sandgren, M.; Withers, S. G.; Stahlberg, J.; Beckham, G. T. *J. Am. Chem. Soc.* **2014**, *136*, 321−329.
(51) Wan, Q.; Zhang, Q.; Hamilton-Brehm, S.; Weiss, K.; Mustyakimov, M.; Coates, L.; Langan, P.; Graham, D.; Kovalevsky, A. *Acta Crystallogr. D: Biol. Crystallogr.* **2014**, *70*, 11−23.
(52) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; Gonzalez-Outeirino, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622−655.
(53) Mayes, H. B.; Broadbelt, L. J.; Beckham, G. T. *J. Am. Chem. Soc.* **2014**, *136*, 1008−1022.
(54) Kurihara, Y.; Ueda, K. *Carbohydr. Res.* **2014**, *344*, 2266−2269.
(55) Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(56) (a) Ensing, B.; De Vivo, M.; Liu, Z. W.; Moore, P.; Klein, M. L. *Acc. Chem. Res.* **2006**, *39*, 73−81. (b) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(57) (a) Barducci, A.; Bonomi, M.; Parrinello, M. *WIREs Comput. Mol. Sci.* **2011**, *1*, 826−843. (b) Biarnés, X.; Bongarzone, S.; Vargiu, A. V.; Carloni, P.; Ruggerone, P. *J. Comput. Aided. Mol. Des.* **2011**, *25*, 395−402. (c) Leone, V.; Marinelli, F.; Carloni, P.; Parrinello, M. *Curr. Opin. Struct. Biol.* **2010**, *20*, 148−154.
(58) Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. *J. Phys. Chem. B* **2005**, *109*, 6676−6687.
(59) Laio, A.; Rodriguez-Fortea, A.; Gervasio, F. L.; Ceccarelli, M.; Parrinello, M. *J. Phys. Chem. B* **2005**, *109*, 6714−6721.
(60) Nair, N. N.; Schreiner, E.; Marx, D. *J. Am. Chem. Soc.* **2008**, *130*, 14148−14160.
(61) Palmer, J. C.; Car, R.; Debenedetti, P. G. *Faraday Discuss.* **2013**, *167*, 77−94.
(62) Dama, J. F.; Parrinello, M.; Voth, G. A. *Phys. Rev. Lett.* **2014**, *112*, No. 240602.
(63) Boero, M.; Ikeda, T.; Ito, E.; Terakura, K. *J. Am. Chem. Soc.* **2006**, *128*, 16798−16807.
(64) Ardèvol, A.; Biarnés, X.; Planas, A.; Rovira, C. *J. Am. Chem. Soc.* **2010**, *132*, 16058−16065.
(65) Lammerts van Bueren, A.; Ardèvol, A.; Fayers-Kerr, J.; Luo, B.; Zhang, Y.; Sollogoub, M.; Bleriot, Y.; Rovira, C.; Davies, G. J. *J. Am. Chem. Soc.* **2010**, *132*, 1804−1806.
(66) Offen, W. A.; Zechel, D. L.; Withers, S. G.; Gilbert, H. J.; Davies, G. J. *Chem. Commun. (Camb.)* **2009**, 2484−2486.
(67) (a) Vincent, F.; Gloster, T. M.; Macdonald, J.; Morland, C.; Stick, R. V.; Dias, F. M.; Prates, J. A.; Fontes, C. M.; Gilbert, H. J.; Davies, G. J. *Chembiochem* **2004**, *5*, 1596−1599. (b) Tailford, L. E.; Offen, W. A.; Smith, N. L.; Dumon, C.; Morland, C.; Gratien, J.; Heck, M. P.; Stick, R. V.; Bleriot, Y.; Vasella, A.; Gilbert, H. J.; Davies, G. J. *Nat. Chem. Biol.* **2008**, *4*, 306−312.
(68) (a) Hine, J. In *Advances in Physical Organic Chemistry*; Gold, V., Bethel, D., Eds.; Academic Press: San Diego, 1978; Vol. *15*, pp 1−61. (b) Sinnott, M. L. In *Advances in Physical Organic Chemistry*; Bethell, D., Ed.; Academic Press: San Diego1988; Vol. *24*, pp 113−204.
(69) Warshel, A.; Levitt, M. *J. Mol. Biol.* **1976**, *103*, 227−249.
(70) Sinnott, M. L. *Chem. Rev.* **1990**, *90*, 1171−1202.
(71) Vocadlo, D. J.; Davies, G. J.; Laine, R.; Withers, S. G. *Nature* **2001**, *412*, 835−838.
(72) Barker, I. J.; Petersen, L.; Reilly, P. J. *J. Phys. Chem. B* **2010**, *114*, 15389−15393.
(73) Brás, N. F.; Fernandes, P. A.; Ramos, M. J. *J. Chem. Theory Comput.* **2010**, *6*, 421−433.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547

(74) Jitonnom, J.; Lee, V. S.; Nimmanpipug, P.; Rowlands, H. A.; Mulholland, A. J. *Biochemistry* **2011**, *50*, 4697−4711.
(75) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2011**, *133*, 19474−19482.
(76) Barnett, C. B.; Wilkinson, K. A.; Naidoo, K. J. *J. Am. Chem. Soc.* **2010**, *132*, 12800−12803.
(77) Govender, K.; Naidoo, K. J. *J. Chem. Theory Comput.* **2014**, *10*, 4708−4717.
(78) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. *Org. Biomol. Chem.* **2004**, *2*, 968−980.
(79) Planas, A.; Abel, M.; Millet, O.; Palasi, J.; Pallares, C.; Viladot, J. L. *Carbohydr. Res.* **1998**, *310*, 53−64.
(80) Williams, R. J.; Iglesias-Fernández, J.; Stepper, J.; Jackson, A.; Thompson, A. J.; Lowe, E. C.; White, J. M.; Gilbert, H. J.; Rovira, C.; Davies, G. J.; Williams, S. J. *Angew. Chem., Int. Ed.* **2014**, *53*, 1087−1091.
(81) (a) Malik, V.; Black, G. W. In *Advances in Protein Chemistry and Structural Biology*; Christo, C., Tatyana, K.-C., Eds.; Academic Press: New York, 2012; Vol. *87*, pp 87−115. (b) Breton, C.; Fournel-Gigleux, S.; Palcic, M. M. *Curr. Opin. Struct. Biol.* **2012**, *22*, 540−549. (c) Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008**, *77*, 521−555.
(82) (a) Zuccotti, S.; Zanardi, D.; Rosano, C.; Sturla, L.; Tonetti, M.; Bolognesi, M. *J. Mol. Biol.* **2001**, *313*, 831−843. (b) Lira-Navarrete, E.; Iglesias-Fernández, J.; Zandberg, W. F.; Companon, I.; Kong, Y.; Corzana, F.; Pinto, B. M.; Clausen, H.; Peregrina, J. M.; Vocadlo, D. J.; Rovira, C.; Hurtado-Guerrero, R. *Angew. Chem., Int. Ed.* **2014**, *53*, 8206−8210.
(83) Gastinel, L. N.; Bignon, C.; Misra, A. K.; Hindsgaul, O.; Shaper, J. H.; Joziasse, D. H. *EMBO J.* **2001**, *20*, 638−649.
(84) Boix, E.; Zhang, Y.; Swaminathan, G. J.; Brew, K.; Acharya, K. R. *J. Biol. Chem.* **2002**, *277*, 28310−28318.
(85) (a) Baskaran, S.; Roach, P. J.; DePaoli-Roach, A. A.; Hurley, T. D. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 17563−17568. (b) Giganti, D.; Alegre-Cebollada, J.; Urresti, S.; Albesa-Jové, D.; Rodrigo-Unzueta, A.; Comino, N.; Kachala, M.; López-Fernández, S.; Svergun, D. I.; Fernández, J. M.; Guerin, M. E. *J. Biol. Chem.* **2013**, *288*, 29797−29808. (c) Diaz, A.; Diaz-Lobo, M.; Grados, E.; Guinovart, J. J.; Fita, I.; Ferrer, J. C. *IUBMB Life* **2012**, *64*, 649−658.
(86) Kim, S. C.; Singh, A. N.; Raushel, F. M. *Arch. Biochem. Biophys.* **1988**, *267*, 54−59.
(87) Murray, B. W.; Wittmann, V.; Burkart, M. D.; Hung, S. C.; Wong, C. H. *Biochemistry (Mosc)* **1997**, *36*, 823−831.
(88) Burkart, M. D.; Vincent, S. P.; Duffels, A.; Murray, B. W.; Ley, S. V.; Wong, C. H. *Bioorg. Med. Chem.* **2000**, *8*, 1937−1946.
(89) Krupicka, M.; Tvaroska, I. *J. Phys. Chem. B* **2009**, *113*, 11314−11319.
(90) Kozmon, S.; Tvaroska, I. *J. Am. Chem. Soc.* **2006**, *128*, 16921−16927.
(91) Withers, S. G.; Davies, G. J. *Nat. Chem. Biol.* **2012**, *8*, 952−953.
(92) Monegal, A.; Planas, A. *J. Am. Chem. Soc.* **2006**, *128*, 16030−16031.
(93) Soya, N.; Fang, Y.; Palcic, M. M.; Klassen, J. S. *Glycobiology* **2011**, *21*, 547−552.
(94) (a) Andre, I.; Tvaroska, I.; Carver, J. P. *Carbohydr. Res.* **2003**, *338*, 865−877. (b) Gomez, H.; Lluch, J. M.; Masgrau, L. *J. Am. Chem. Soc.* **2013**, *135*, 7053−7063. (c) Bobovska, A.; Tvaroska, I.; Kona, J. *Glycobiology* **2015**, *25*, 3−7.
(95) Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. *Nat. Struct. Biol.* **2001**, *8*, 166−175.
(96) Tvaroska, I. *Carbohydr. Res.* **2015**, *403*, 38−47.
(97) Gomez, H.; Lluch, J. M.; Masgrau, L. *Carbohydr. Res.* **2012**, *356*, 204−208.
(98) Gomez, H.; Polyak, I.; Thiel, W.; Lluch, J. M.; Masgrau, L. *J. Am. Chem. Soc.* **2012**, *134*, 4743−4752.
(99) Lee, S. S.; Hong, S. Y.; Errey, J. C.; Izumi, A.; Davies, G. J.; Davis, B. G. *Nat. Chem. Biol.* **2011**, *7*, 631−638.
(100) Ardevol, A.; Rovira, C. *Angew. Chem., Int. Ed.* **2011**, *50*, 10897−10901.
(101) Bobovska, A.; Tvaroska, I.; Kona, J. *Org. Biomol. Chem.* **2014**, *12*, 4201−4210.
(102) Gomez, H.; Rojas, R.; Patel, D.; Tabak, L. A.; Lluch, J. M.; Masgrau, L. *Org. Biomol. Chem.* **2014**, *12*, 2645−2655.
(103) Lewis, E. S.; Boozer, C. E. *J. Am. Chem. Soc.* **1952**, *74*, 308−311.
(104) Sinnott, M. L.; Jencks, W. P. *J. Am. Chem. Soc.* **1980**, *102*, 2026−2032.
(105) Errey, J. C.; Lee, S. S.; Gibson, R. P.; Martinez Fleites, C.; Barry, C. S.; Jung, P. M.; O'Sullivan, A. C.; Davis, B. G.; Davies, G. J. *Angew. Chem., Int. Ed.* **2010**, *49*, 1234−1237.
(106) Lairson, L. L.; Chiu, C. P.; Ly, H. D.; He, S.; Wakarchuk, W. W.; Strynadka, N. C.; Withers, S. G. *J. Biol. Chem.* **2004**, *279*, 28339−28344.
(107) Luehr, N.; Ufimtsev, I. S.; Martínez, T. J. *J. Chem. Theory Comput.* **2011**, *7*, 949−954.
(108) Bowler, D. R.; Miyazaki, T. *Rep. Prog. Phys.* **2012**, *75*, 036503.
(109) Fox, S. J.; Dziedzic, J.; Fox, T.; Tautermann, C. S.; Skylaris, C. K. *Proteins* **2014**, *82*, 3335−3346.
(110) Davies, G. J.; Wilson, K. S.; Henrissat, B. *Biochem. J.* **1997**, *321* (Pt2), 557−559.
(111) Jamaluddin, H.; Tumbale, P.; Withers, S. G.; Acharya, K. R.; Brew, K. *J. Mol. Biol.* **2007**, *369*, 1270−1281.
(112) Zhong, W.; Kuntz, D. A.; Ember, B.; Singh, H.; Moremen, K. W.; Rose, D. R.; Boons, G. J. *J. Am. Chem. Soc.* **2008**, *130*, 8975−8983.
(113) Thompson, A. J.; Speciale, G.; Iglesias-Fernandez, J.; Hakki, Z.; Belz, T.; Cartmell, A.; Spears, R. J.; Chandler, E.; Temple, M. J.; Stepper, J.; Gilbert, H. J.; Rovira, C.; Williams, S. J.; Davies, G. J. *Angew. Chem., Int. Ed.* **2015**, *54*, 5378−5382.
(114) Zhu, Y.; Suits, M. D.; Thompson, A. J.; Chavan, S.; Dinev, Z.; Dumon, C.; Smith, N.; Moremen, K. W.; Xiang, Y.; Siriwardena, A.; Williams, S. J.; Gilbert, H. J.; Davies, G. J. *Nat. Chem. Biol.* **2010**, *6*, 125−132.
(115) (a) Ducros, V. M.; Zechel, D. L.; Murshudov, G. N.; Gilbert, H. J.; Szabo, L.; Stoll, D.; Withers, S. G.; Davies, G. J. *Angew. Chem., Int. Ed.* **2002**, *41*, 2824−2827. (b) Cartmell, A.; Topakas, E.; Ducros, V. M.; Suits, M. D.; Davies, G. J.; Gilbert, H. J. *J. Biol. Chem.* **2008**, *283*, 34403−34413.
(116) Davies, G. J.; Mackenzie, L.; Varrot, A.; Dauter, M.; Brzozowski, A. M.; Schulein, M.; Withers, S. G. *Biochemistry* **1998**, *37*, 11707−11713.
(117) Zou, J.; Kleywegt, G. J.; Stahlberg, J.; Driguez, H.; Nerinckx, W.; Claeyssens, M.; Koivula, A.; Teeri, T. T.; Jones, T. A. *Structure* **1999**, *7*, 1035−1045.
(118) (a) Suzuki, R.; Fujimoto, Z.; Ito, S.; Kawahara, S.; Kaneko, S.; Taira, K.; Hasegawa, T.; Kuno, A. *J. Biochem.* **2009**, *146*, 61−70. (b) Czjzek, M.; Ben David, A.; Bravman, T.; Shoham, G.; Henrissat, B.; Shoham, Y. *J. Mol. Biol.* **2005**, *353*, 838−846. (c) Vandermarliere, E.; Bourgois, T. M.; Rombouts, S.; Van Campenhout, S.; Volckaert, G.; Strelkov, S. V.; Delcour, J. A.; Rabijns, A.; Courtin, C. M. *Biochem. J.* **2008**, *410*, 71−79. (d) 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.
(119) Kankainen, M.; Laitinen, T.; Peräkylä, M. *Phys. Chem. Chem. Phys.* **2004**, *6*, 5074−5080.
(120) Brux, C.; Ben-David, A.; Shallom-Shezifi, D.; Leon, M.; Niefind, K.; Shoham, G.; Shoham, Y.; Schomburg, D. *J. Mol. Biol.* **2006**, *359*, 97−109.
(121) Sulzenbacher, G.; Bignon, C.; Nishimura, T.; Tarling, C. A.; Withers, S. G.; Henrissat, B.; Bourne, Y. *J. Biol. Chem.* **2004**, *279*, 13119−13128.
(122) Guce, A. I.; Clark, N. E.; Salgado, E. N.; Ivanen, D. R.; Kulminskaya, A. A.; Brumer, H., 3rd; Garman, S. C. *J. Biol. Chem.* **2010**, *285*, 3625−3632.
(123) Pan, X. L.; Liu, W.; Liu, J. Y. *J. Phys. Chem. B* **2013**, *117*, 484−489.

NOTE ADDED AFTER ASAP PUBLICATION

This paper was published ASAP on June 4, 2015, with a nomenclature error. The corrected version was posted on June 5, 2015.

DOI: 10.1021/jacs.5b01156
J. Am. Chem. Soc. 2015, 137, 7528−7547