JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

A Trapped Covalent Intermediate of a Glycoside Hydrolase on the Pathway to Transglycosylation. Insights from Experiments and Quantum Mechanics/Molecular Mechanics Simulations

Lluís Raich,[§] Vladimir Borodkin,[‡] Wenxia Fang,[‡] Jorge Castro-López,[#] Daan M. F. van Aalten,[*,‡,▽] Ramón Hurtado-Guerrero,[*,⊥,#] and Carme Rovira[*,∥,§]

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

[‡]Division of Molecular Microbiology and [▽]MRC Protein Phosphorylation and Ubiquitylation Unit, School of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland

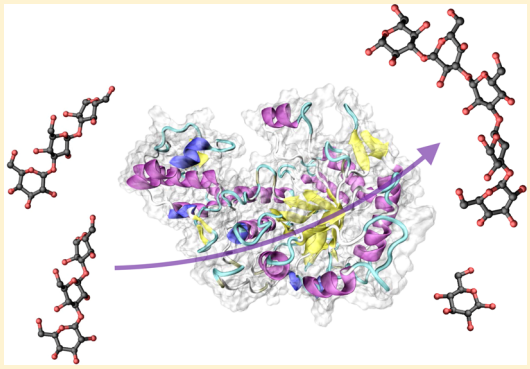
[⊥]Fundación ARAID, Edificio CEEI Aragón, 50018 Zaragoza, Spain

[#]Institute of Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, BIFI-IQFR (CSIC) Joint Unit, Mariano Esquillor s/n, Campus Rio Ebro, Edificio I+D, 50018 Zaragoza, Spain

[∥]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys, 23, 08020 Barcelona, Spain

Supporting Information

ABSTRACT: The conversion of glycoside hydrolases (GHs) into transglycosylases (TGs), i.e., from enzymes that hydrolyze carbohydrates to enzymes that synthesize them, represents a promising solution for the large-scale synthesis of complex carbohydrates for biotechnological purposes. However, the lack of knowledge about the molecular details of transglycosylation hampers the rational design of TGs. Here we present the first crystallographic structure of a natural glycosyl−enzyme intermediate (GEI) of *Saccharomyces cerevisiae* Gas2 in complex with an acceptor substrate and demonstrate, by means of quantum mechanics/molecular mechanics metadynamics simulations, that it is tuned for transglycosylation ($\Delta G^{\ddagger}$ = 12 kcal/mol). The 2-OH···nucleophile interaction is found to be essential for catalysis: its removal raises the free energy barrier significantly (11 and 16 kcal/mol for glycosylation and transglycosylation, respectively) and alters the conformational itinerary of the substrate (from $^{4}C_{1} \rightarrow [^{4}E]^{\ddagger} \rightarrow {}^{1,4}B/{}^{4}E$ to $^{4}C_{1} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{4}C_{1}$). Our results suggest that changes in the interactions involving the 2-position could have an impact on the transglycosylation activity of several GHs.



1. INTRODUCTION

The growth of glycomics and the development of diagnostic tests, vaccines, and new therapeutics based on carbohydrates are hampered by the lack of effective tools for their production.[1,2] Enzymatic synthesis, characterized by its high stereo- and regioselective products obtained under mild conditions, is a promising approach to solve this problem. In this regard, retaining glycoside hydrolases (GHs) have a high potential to be engineered to synthesize carbohydrates by transglycosylation (Figure 1). Along this line, several approaches such as directed evolution,[3−5] site-directed mutagenesis,[6,7] and the use of endo/exo glycosynthases[8−10] have appeared in the past few years.

The reaction mechanism of retaining GHs consists of two steps. In the first one (glycosylation, Figure 1), a glycosyl−enzyme intermediate (GEI) forms[11] and the leaving group is released. In the second step (deglycosylation), an acceptor molecule reacts with the GEI, generating two different products depending on the identity of the acceptor molecule, either a water molecule (R = H, hydrolysis product) or a sugar acceptor (R = sugar, transglycosylation product).[12] Although hydrolysis is thermodynamically favorable (e.g., almost 3 kcal/mol for cellobiose[13]), a few GHs known as transglycosylases (TGs), such as xyloglucan *endo*-TGs,[14] sucrase-type enzymes,[15] or *trans*-sialidases,[16] display significant transglycosylation activities and lead to high yields on reasonable time scales. The overall mechanism of transglycosylation is well known (Figure 1), but it is still unclear how TGs can favor transglycosylation over hydrolysis in 55 M water. Experiments on natural and engineered TGs show that they usually have lower catalytic efficiencies (i.e., higher reaction free energy barriers) in comparison to their hydrolytic relatives,[12,17] leading to long-lived species before the thermodynamic equilibrium is reached.

Received: September 29, 2015
Published: February 9, 2016

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

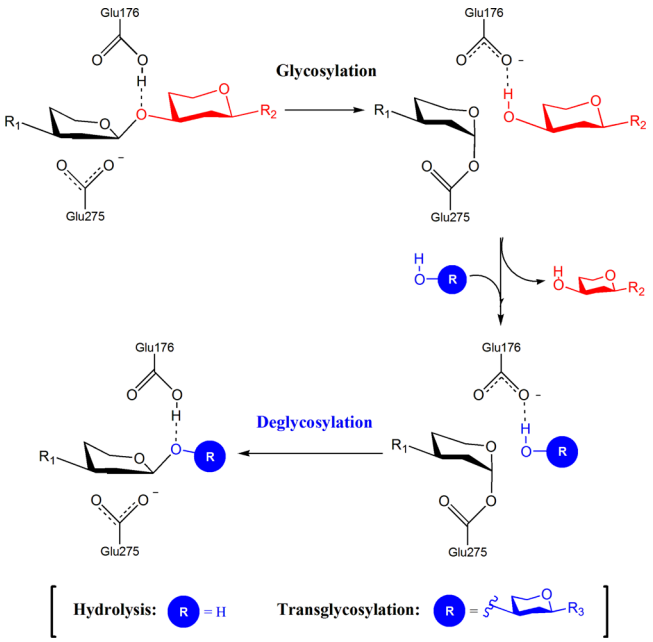


Figure 1. Reaction mechanisms of retaining GHs. The nucleophile (Glu275) and acid/base (Glu176) residues are those of *Sc*Gas2.

In fact, GEI lifetimes as high as 30 min have been reported for wild-type TGs,[18] whereas the intermediate breaks down quickly in purely hydrolytic GHs.

Factors such as substrate acceptor binding, water migration into the active site, and transition state (TS) interactions are known to influence the activation energy; thus, enzyme mutations affecting these factors can modify the transglycosylation/hydrolysis ratio.[12] However, there is not a straightforward, easy, and rational approach for generating such efficient enzyme variants. Moreover, the limited knowledge of the molecular basis of transglycosylation—particularly TS interactions—is hindering research in this field. Of utmost importance are the interactions of the sugar hydroxyl groups with enzyme residues, especially the ones involving the 2-OH group, which are predicted to play a major role in the glycosylation step, according to kinetic studies.[19,20]

The aforementioned issues, encompassing the transglycosylation mechanism and the role of the 2-OH in catalysis, are addressed in this work using the prototypical TG *Saccharomyces cerevisiae* Gas2 (*Sc*Gas2) and applying a multidisciplinary approach. *Sc*Gas2 is a GPI-anchored plasma membrane glycoside hydrolase (GPI = glycophosphatidylinositol) of family 72 (GH72) whose catalytic domain is found in the cell wall of fungi.[21,22] Members of this family, whose function is to regulate the assembly and rearrangement of the β-1,3-glucan that forms part of the fungal cell wall, are known to exhibit high transglycosylation activities.[23] Previous structures of *Sc*Gas2 in complex with different β-glucan substrate/product oligosaccharides suggested that the acceptor substrate protects the GEI.[24] Nevertheless, no structure of the GEI, which would be the best point to understand the diverging hydrolysis and transglycosylation reactions (Figure 1), has been reported to date. To our knowledge, the only example of a GEI structure with a bound acceptor substrate is the recent work on *Hypocrea jecorina* GH7 cellobiohydrolase Cel7A.[25] However, the structure contains non-natural substrates (2-deoxy-2-fluoroglycosides) bound to the enzyme, which are known to disrupt the network of hydrogen bond interactions in the active site.[26,27] Specifically, the crucial hydrogen bond between 2-OH and the nucleophile is missing.

Herein, we report a natural GEI structure of *Sc*Gas2 in complex with an acceptor substrate and demonstrate, using *ab initio* quantum mechanics/molecular mechanics (QM/MM) metadynamics, that it is catalytically competent. Our study identifies crucial interactions between the GEI and the acceptor substrate, as well as the fundamental interactions involving 2-OH, whose contribution to the reaction energy barrier is quantified for the first time using *ab initio* methods. Moreover, we elucidate the substrate conformational itinerary and active-site residues that play a role in catalysis.

2. METHODS

2.1. Crystallization, Data Collection, Structure Determination, and Refinement. The Glu176Gln mutant of *Sc*Gas2 was purified as described previously.[24] The mutated protein was spin-concentrated to 21 mg/mL. Crystals were grown by sitting drop experiments at 20 °C through mixing 1 μL of protein with an equal volume of a reservoir solution (0.2 M ammonium sulfate, 20% PEG3350, 0.1 M sodium BIS-TRIS, pH 4.5). Under these conditions, crystals appeared within 3−7 days. The crystals were soaked with 300 mM laminaribiose β-fluoride for 10 min, followed by a second soaking with 100 mM laminaripentaose. They were cryo-protected with 0.2 M ammonium sulfate, 20% PEG3350, 20% butanediol, 0.1 M sodium BIS-TRIS, pH 4.5, and flash-cooled prior to data collection at 100 K. The structure was solved by molecular replacement using Glu176Gln mutant, PDB 2W61, as a search model. The model was improved through cycles of manual building in Coot[28] and refinement with REFMAC5.[29] Topologies for the oligosaccharide ligands were generated with PRODRG.[30] The final models were validated with PROCHECK;[31] model statistics are given in Table 1.

2.2. Synthesis of Laminaribiose β-fluoride. Synthesis of laminaribiose β-fluoride (Scheme S1) was initiated with regioselective glycosylation of the corresponfing ethyl thioglucoside derivative with perbenzoylated glucosyl trichloroacetimidate, according to the published procedure to give the 1,3-β-linked disaccharide **3** in

Table 1. Data Collection and Refinement Statistics for *Sc*Gas2 GEI with Laminaritetraose and in Complex with a Laminaripentaose Acceptor[a]

wavelength (Å)	1.54
resolution (Å)	20.00−1.8 (1.86−1.80)
cell dimensions (Å)	a = 050.11
	b = 070.25
	c = 149.92
unique reflections	49626
completeness	98.3 (97.3)
R_{merge}	0.073 (0.429)
$I/\sigma(I)$	30 (3.1)
redundancy	4.1 (3.9)
R_{work}/R_{free}	0.216/0.244
RMSD from ideal geometry, bonds (Å)	0.011
RMSD from ideal geometry, angles (deg)	1.33
$\langle B \rangle$ protein (Å^2)	25
$\langle B \rangle$ ligand (Å^2)	43.1
$\langle B \rangle$ solvent (Å^2)	31.6
Ramachandran plot:	
preferred regions (%)	97.4
allowed regions (%)	2.1
outliers (%)	0.4
PDB ID	5FIH

[a]Values in parentheses refer to the highest resolution shell. Ramachandran plot statistics were determined with PROCHECK.

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

87%.[32] The details of the synthesis are provided in the Supporting Information (SI).

2.3. Cloning and Purification of the C-Terminal Histidine-Tagged Glu176Gln Mutant. To facilitate the purification of the Glu176Gln mutant, we previously incorporated a C-terminal histidine tag in the wild-type enzyme. The cloning and purification of the mutant is described in detail in the SI.

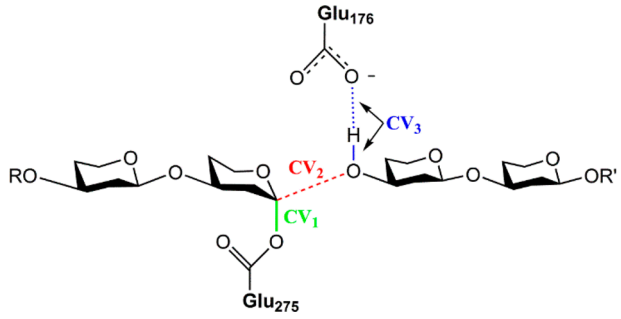
2.4. Mass Spectrometry Detection of Sugar Covalent Binding to the Glu176Gln Mutant. First, 200 mM G2-F was mixed with 2.8 μM Glu176Gln and incubated at room temperature for 1 h. After running on SDS-PAGE, the gel band was cut for in-gel trypsin digestion. Sample was analyzed by LC-MS/MS (Q-exactive-HF mass spectrometer coupled with a Dionex Ultimate 3000 RS, Thermo Scientific). The raw MS data were searched against the protein sequence by the Mascot search engine (Matrix Science, Version 2.2) through proteome discovery software (version 1.4). The MS spectrum is provided in the SI.

2.5. Classical and QM/MM Molecular Dynamics Simulations. The initial structure for the simulations was taken from the present reported structure of *S. cerevisiae* Gas2 covalently bound with laminaritetraose and in complex with an incoming laminaripentaose acceptor. Classical MD simulations using the Amber11 software,[33] together with FF99SB (protein residues),[34] GLYCAM06 (carbohydrates),[35] gaff (glycosylated glutamate),[36] and TIP3P (water)[37] force fields, were performed to equilibrate the ternary complex for about 18 ns. Analysis of the trajectory was carried out using standard tools of AMBER and VMD.[38] Further details of the classical simulations are provided in the SI.

QM/MM MD simulations were performed using the method developed by Laio et al.,[39] which combines Car–Parrinello MD,[40] based on Density Functional Theory (DFT), with force-field MD methodology, as described in the SI. The QM region included the glucose rings at the −1 and +1 subsites, half rings of the saccharides at the −2 and +2 subsites, and the catalytic residues (Glu176 and Glu275), leading a total number of 88 QM atoms (including capping hydrogens; Figure S3) and 91.779 MM atoms for the system. The QM region was enclosed in an isolated supercell of size 18.5 × 17.9 × 21.6 Å^3. Kohn–Sham orbitals were expanded in a planewave basis set with a kinetic energy cutoff of 70 Ry. Norm-conserving Troullier–Martins *ab initio* pseudopotentials[41] were used for all elements. The calculations were performed using the Perdew, Burke, and Ernzerhoff (PBE) generalized gradient-corrected approximation.[42] This functional form has been proven to give a good performance in the description of hydrogen bonds[43] and was already used with success in previous works on GHs and transferases.[44] A fictitious electronic mass of 700 au and a time step of 5 au were used to ensure an adiabaticity of 4.73 × 10^{-5} au·ps^{-1}·atom^{-1} for the fictitious kinetic energy. Further details are given in the SI.

2.6. Metadynamics Simulations. The reaction free enery landscape (FEL) of the transglycosylation reaction was explored using the metadynamics approach with three collective variables (CVs). Two of them, CV_1 and CV_2, were taken as the C1–O_{Glu275} and C1–$O_{Acceptor}$ distances, respectively. The third, CV_3, was taken as the difference between the O_{Glu176}–$H_{Acceptor}$ and $O_{Acceptor}$–$H_{Acceptor}$ distances (see Scheme 1). The hill height was 0.6 kcal/mol, and the deposition time was 24 fs (200 MD steps). A fictitious harmonic coupling was used to diminish the perturbation of the time-dependent potential, according to Lagrangian metadynamics,[45] using mass of 100 amu and constants of 1 au for cv1 and cv2, and mass of 100 amu and a constant of 3.5 au for cv3, following literature recommendations.[46] Walls at 4 Å for each distance and +2 Å and −1.5 Å for the difference of distances were used to reduce the FEL space to the chemical event. First crossing criterion, as recommended for chemical reactions,[46] was taken to determine the simulation end (convergence tests are provided in the SI, section 6.4). The three-dimensional free energy landscape (3D FEL) was completed after 759 deposited Gaussians (Figure S5). The projection of 3D to 2D was done using a Jacobian procedure (see SI of ref 47). An additional 3D metadynamics simulation was performed to model the reaction mechanism for the *off* configuration, using the same CVs as in the previous case (Scheme S3). The metadynamics parameters were taken as those of the previous reaction, with the exception of the hill height, which was set to 1 kcal/mol, as a higher barrier was expected.

Scheme 1



3. RESULTS AND DISCUSSION

3.1. Structure of the Glycosyl–Enzyme Intermediate. To trap the GEI, we synthesized the disaccharide laminaribiose β-fluoride (Scheme S1), which was soaked into crystals of the inactive acid/base Glu176Gln mutant. The combination of fluoride being a good leaving group and the Glu176Gln mutation rendering an inactive enzyme and preventing the second reaction step (deglycosylation via hydrolysis or transglycosylation, see Figure 1) led to the generation of a stable GEI.[48] We also soaked laminaripentaose (G5, Figure S1) into the previous crystals in order to see whether it was accommodated by the positive subsites. As a result, we were able to trap a GEI featuring an enzyme-bound laminaritetraose (G4), with glucose at −1 displaying an unequivocal α-bond with the Glu275 nucleophile (Figure 2 and Table 1).

The observation of an enzyme-bound G4 was an unexpected result since the only source of activated glucose came from laminaribiose β-fluoride (G2F). Given the high regio- and stereospecificity of the reaction, indicating that it occurred on-enzyme, there are two plausible hypothesis to explain the formation of the G4·*Sc*Gas2 GEI: One possible explanation is that, once a G2-F unit reacts with the enzyme (forming a G2·*Sc*Gas2 GEI), another G2-F reacts from +1 and +2 through transglycosylation, leading to G4-F that further reacts with the enzyme, leading to the observed G4·*Sc*Gas2 GEI. Another explanation is that the initial G2·*Sc*Gas2 GEI reacts with a G2-F molecule in positions −3 and −4 through an S_Ni mechanism.[49,50] While the first hypothesis seems unlikely due to the absence of a base to deprotonate the acceptor hydroxyl, the second is unprecedented in GHs and needs further verification. Independently of the mechanism of G4·*Sc*Gas2 GEI formation, this intermediate was further identified by mass spectroscopy (see section 2.4 and SI). The presence of G2·*Sc*Gas2 GEI was also identified, suggesting that in solution both covalent intermediates are present.

Strikingly, we also found G5 spanning all the previously reported positive subsites (+1 to +5), representing the first case in which a GEI has been trapped with natural substrates. The 3-OH of the +1 acceptor sugar is 3.3 Å away from the anomeric carbon of the saccharide at −1 (Figure 3), correctly placed to act as nucleophile for the transglycosylation reaction. The pyranose covalently bound to the Glu275 nucleophile adopts a $^4C_1$ relaxed conformation, consistent with the alpha enzyme–substrate glycosidic bond. The superposition of the present structure with a nonproductive complex of *Sc*Gas2 with laminaripentaose (PDB 2W62, Figure S2) displays a very

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325–3332

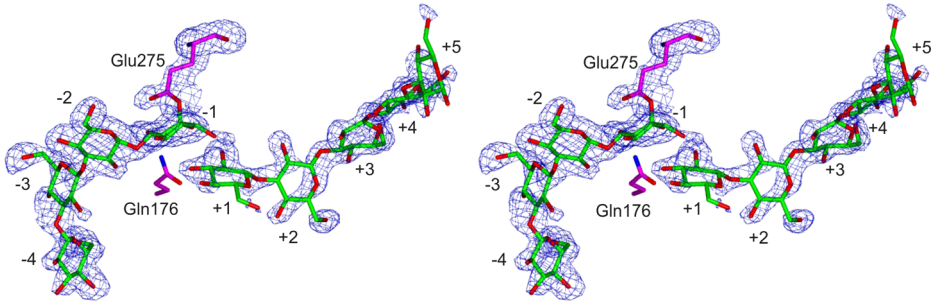


Figure 2. Stereo $F_o - F_c$ electron density map of the G4 GEI of Glu176Gln *Sc*Gas2 in complex with a G5 acceptor. The unbiased (i.e., before inclusion of any ligand model) $|F_o| - |F_c|$, f_{calc} electron density maps are shown at 2.2σ.

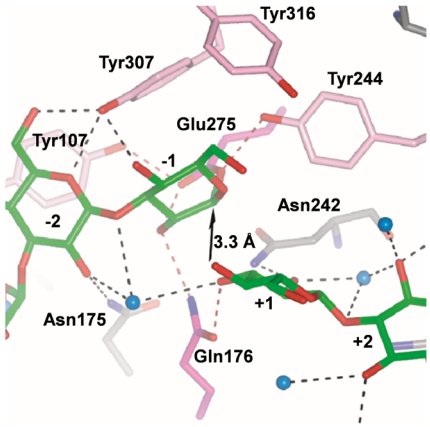


Figure 3. Close view of the catalytic region of the GEI of Glu176Gln *Sc*Gas2 with G4 in complex with a G5 acceptor. The amino acids placed in the negative and positive subsites are shown as sticks with gray carbons, except for Glu275/Gln176 and Tyr107/Tyr244/Tyr307, which are shown as magenta and pink carbon atoms, respectively. Protein−ligand and water−ligand hydrogen bonds are shown as dotted black lines. Relevant hydrogen bonds are shown as dotted orange lines. Water molecules involved in hydrogen bonds with the ligands are shown as cyan spheres. For clarity purposes protein−water hydrogen bonds are not shown. A black arrow indicates the distance among −1 C1 with +1 O3.

small RMSD between the acceptor sugars (0.5 Å), indicating that the sugar units are constrained within the positive subsites. This suggests that the molecular determinants for the transglycosylation activity of the enzyme lie at the negative subsites, as there are few differences in the positive subsites of the nonproductive complex and the present GEI.

As shown in Figures S1 and S2, the negative subsites display different sugar−protein interactions in the new structure compared to PDB entry 2W62, which is not surprising since the latter structure did not contain a sugar at the −1 subsite. A stacking interaction between the −1 sugar and Tyr307 (Figure 3) is formed, as well as a hydrogen bond between the 4-OH hydroxyl group of the −1 saccharide and Tyr307 (O···O distance of 2.7 Å). However, the relevant interaction of the 2-OH hydroxyl group is not fully defined by the electron density alone: it could be interacting either with the nucleophile through a very short hydrogen bond (O···O distance of 2.4 Å, similar as in previous studies[26]) or with the side chain of the mutated Gln176 at a typical hydrogen bond distance (2.8 Å).

3.2. Interactions Involving the Donor 2-OH. To unravel the network of hydrogen bond interactions around the 2-OH, we performed molecular dynamics (MD) simulations on the wild-type enzyme by reverting the mutation of the acid/base residue. A detailed analysis of all active center interactions reveals two features that could not be observed in the crystallographic structure, presumably due to the Glu176Gln mutation. First of all, the 2-OH substituent often changes hydrogen bond partner, from the nucleophile (2-OH··· nucleophile interaction, hereafter named as *on* configuration) to a solvent water molecule (2-OH···H_2O interaction, named as *off* configuration; see Figure 4a and Figure S4, bottom panel), with populations 45.4% and 54.6%, respectively. Second, the 2-OH accepts a hydrogen bond from the amino group of Asn175 (average distance of 2.07 Å; see Figure 4b and middle panel of Figure S4).

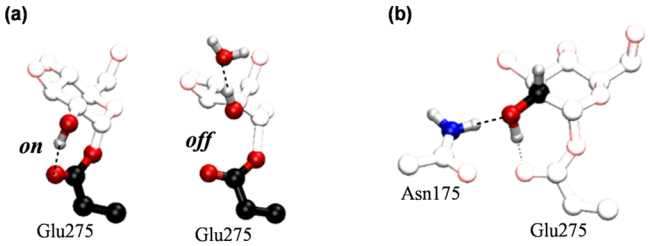


Figure 4. *On* and *off* configurations of the 2-OH···nucleophile interaction (a) and interaction of Asn175 with the 2-OH group (b) observed in the simulations.

To further characterize the dynamics of the C2-OH bond, we performed a QM/MM metadynamics simulation using the H2−C2−O2−H dihedral angle as CV. Consistent with the above results, the free energy profile (Figure S6) displays two minima that correspond to the *on* and *off* configurations, but the *on* configuration is more stable by 3 kcal/mol. Therefore, a snapshot of the *on* configuration was taken to initiate the modeling of the chemical reaction.

3.3. QM/MM Modeling of the Transglycosylation Reaction. The transglycosylation reaction was modeled using metadynamics. This approach has proved to be very useful to study enzymatic reactions, including carbohydrate−active enzymes.[44] A set of three CVs, including all bonds that are formed or cleaved during the reaction, was used to drive the reactants (GEI complex with a G5 acceptor) toward the transglycosylation product: the *Sc*Gas2 complex with laminarinonaose (G9; note that this complex can be considered as a Michaelis complex (MC) of *Sc*Gas2 with G9). CV_1 measures the cleavage of the GEI bond, CV_2 quantifies the degree of formation of the new donor−acceptor glycosidic bond, and CV_3 takes into account the proton transfer between the acceptor and the acid/base residue (Scheme 1).

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

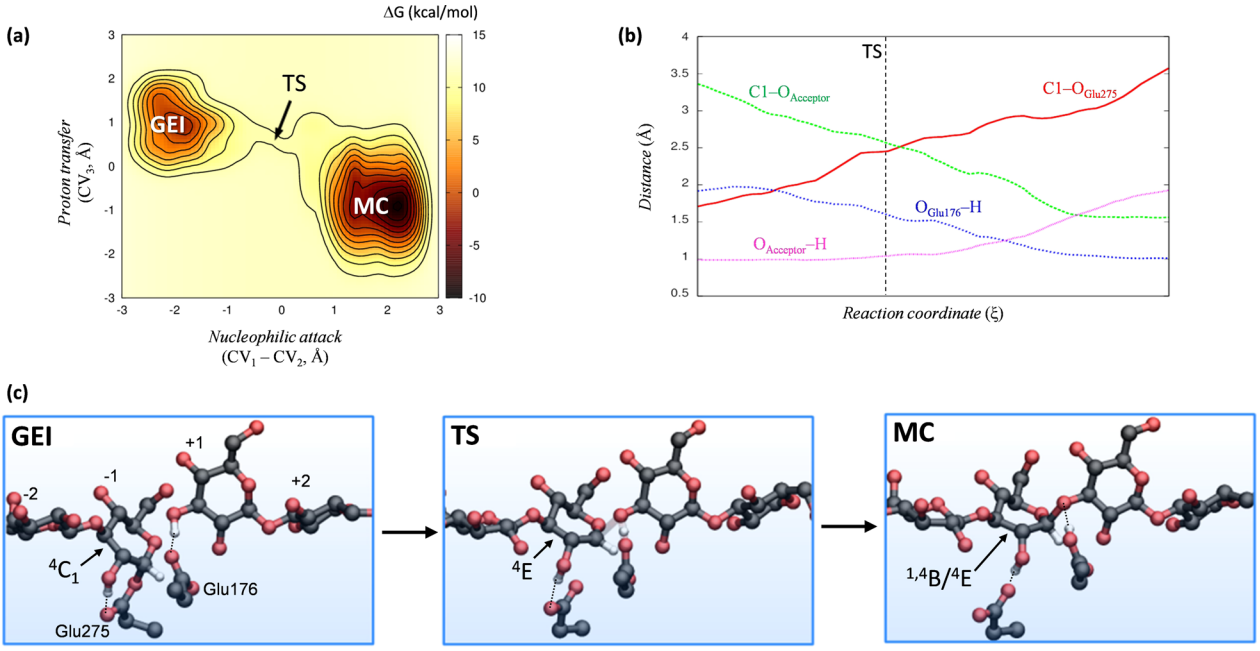


Figure 5. (a) Computed free energy landscape of the transglycosylation reaction (projection on two collective variables) for the *on* configuration. Contour lines are at 2 kcal/mol. (b) Evolution of relevant distances along the reaction coordinate of the *on* configuration: glycosyl−enzyme bond (red), acceptor−donor glycosidic bond (green), $O_{Acceptor}$−H (pink), and O_{Glu176}−H (blue). A running average over five data points has been taken. (c) Mechanism of the transglycosylation reaction obtained from *ab initio* QM/MM metadynamics simulations. Representative states along the lowest free energy pathway.

The 3D FEL obtained from the QM/MM metadynamics simulation is shown in Figure S7. To facilitate the analysis, Figure 5a shows a 2D projection over the nucleophilic attack (CV_1−CV_2) and the proton transfer (CV_3) space. The shape of the FEL clearly indicates that transglycosylation in *Sc*Gas2 is a concerted reaction (there is no local minimum between reactants and products as the one observed, e.g., in some glycosyltransferases[44]), with two minima (GEI and MC) and one TS.

The reaction starts with the elongation of the GEI bond at the same time as the hydroxyl of the incoming sugar approaches toward the anomeric carbon (Figure 5b). At the TS, which is 12 kcal/mol above the GEI, the −1 glucose changes its pyranose ring puckering from $^4C_1$ to ^4E (Figure 5c), a conformation compatible with the requirement of a stable oxocarbenium ion.[44,51] At this stage the new glycosidic bond is nearly formed and the $O_{Acceptor}$-H bond starts to lengthen, conserving the main enzyme···substrate interactions present at the GEI (the −1 saccharide···Tyr307 stacking interaction and the 2-OH···nucleophile, Asn175···2-OH, and 4-OH···Tyr307 hydrogen bonds). The interaction with Asn175 is particularly interesting since mutation of this residue to alanine inactivates *Sc*Gas2,[24] suggesting that it may be an important TS interaction. The change in the anomeric center charge (from +0.1 e^- at the GEI to +0.3 e^- at the TS) reflects the oxocarbenium ion-like character of the TS. From this point, a downhill pathway of 21 kcal/mol leads to the MC of *Sc*Gas2 with G9. Once the MC is formed, the 2-OH changes its hydrogen bond partner from one to another oxygen atom of the nucleophile, and the −1 pyranose ring adopts a distorted $^{1,4}B/^4E$ conformation. It is worth noting that the 2-OH ends up interacting with the same nucleophile oxygen that was initially involved in the covalent bond with the substrate, something that we have previously observed in the study of the GEI formation in a β-endoglucanase.[52] This switch, which in *Sc*Gas2 takes place once the glycosidic bond is formed (Figure S8), breaks the 2-OH···Asn175 interaction. Therefore, this interaction is likely to affect the glycosylation step of the reaction.

The evolution of the distances along the reaction coordinate (Figure 5b) reveals that the transglycosylation reaction is asynchronous and dissociative, as the nucleophile moves away from the anomeric carbon before the acceptor proton is transferred, supporting a D_NA_N type of mechanism.[53] The value of the free energy barrier when going from the MC to the TS (glycosylation of G9), 21 kcal/mol, is compatible with experimental rate constants measured for retaining β-glucosidases.[19] The computed FEL also shows that the glycosylation reaction is endothermic (the MC is ca. 9 kcal/mol more stable than the GEI). Interestingly, the endothermic nature of the glycosylation reaction was recently observed by Piens et al. for a xyloglucan *endo*-TG.[18] It was argued that the GEI retains the energy corresponding to the broken glycosidic bond, contributing to the formation of the new bond of the transglycosylation product. From these results, it is tempting to suggest that an endothermic glycosylation is a prerequisite for efficient transglycosylation.

In summary, the transglycosylation reaction in *Sc*Gas2 consist of a concerted D_NA_N mechanism with a free energy barrier of 12 kcal/mol, characterized by a $^4C_1 \rightarrow [^4E]^{\ddagger} \rightarrow {}^{1,4}B/^4E$ conformational itinerary that agrees with the one expected for retaining β-glucosidases.[54] Moreover, the simulations reveal the formation of key interactions during the reaction, such as Asn175···2-OH and the 2-OH···nucleophile hydrogen bonds, that could play an important role in catalysis.

3.4. Effect of the 2-OH···Nucleophile Interaction on the Catalytic Mechanism. As mentioned in the Introduction, interactions involving the 2-OH group have been predicted to play a major role in catalysis in retaining GHs.[19,20] In fact,

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

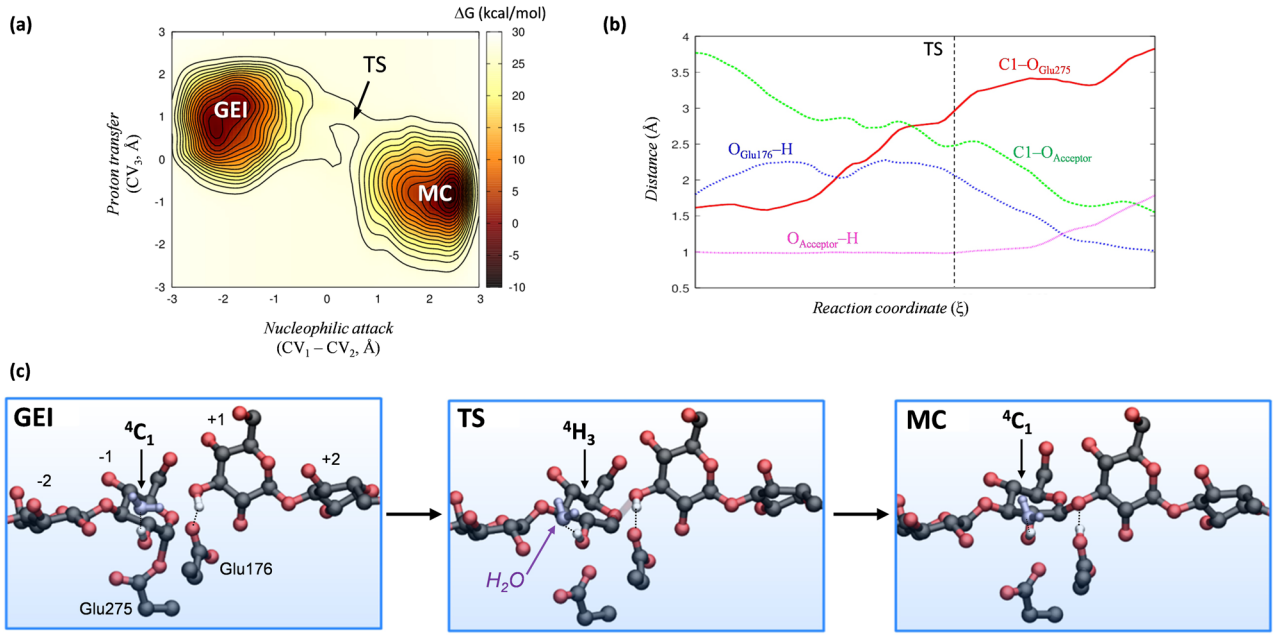


Figure 6. (a) Computed free energy landscape of the transglycosylation reaction (projection on two collective variables) for the *off* configuration. Contour lines are at 2 kcal/mol. (b) Evolution of relevant distances along the reaction coordinate for the *off* configuration: glycosyl−enzyme bond (red), acceptor−donor glycosidic bond (green), $O_{Acceptor}$−H (pink), and OGlu176−H (blue). A running average over five data points has been taken. (c) Mechanism of the transglycosylation reaction obtained from *ab initio* QM/MM metadynamics simulations.

kinetic analyses by Namchuk and Withers[19] of the hydrolysis of a series of monosubstituted aryl glycosides, in which the glycone OH groups were substituted by hydrogen, showed an increase of the activation free energy ($\Delta\Delta G^{\ddagger}$) of the glycosylation step (Figure 1) in the range 4−11 kcal/mol. As discussed by White et al.,[20] removal of the 2-OH has a dual effect: it not only removes the hydrogen bond between the 2-OH and the nucleophile, which decreases the reaction rate, but also inductively stabilizes the oxocarbenium ion-like TS, which inherently increases the reaction rate. Therefore, the observed activation energy increases must be taken as a minimum estimate of the contribution of the interactions involving the 2-OH to the reaction free energy barrier.

To quantify the "net" effect of the 2-OH···nucleophile interaction in the reaction mechanism, we repeated the metadynamics simulation of the transglycosylation reaction starting from the *off* configuration (Figure 4a), in which the 2-OH does not interact with the nucleophile. Like in the previous case (*on* configuration), the FEL reveals a concerted process (Figure 6a). The reaction starts with the approach of the hydroxyl group of the incoming sugar to the anomeric carbon. Once the $C1-O_{Acceptor}$ distance reaches 3 Å, the glycosyl−enzyme bond ($C1-O_{Glu275}$ in Figure 6b) progressively elongates to reach the TS. The −1 saccharide ring changes conformation from $^4C_1$ [GEI] to $^4H_3$ [TS] (Figure 6c), and the charge of the anomeric center increases up to 0.2 e^-, reflecting the formation of an oxocarbenium ion-like species. The reaction turns out to be highly dissociative and more asynchronous than the one for the *on* configuration.

The transglycosylation free energy barrier is 28 kcal/mol, thus being 16 kcal/mol higher than the one obtained when the 2-OH···nucleophile interaction is not formed (Table 2). The reverse reaction, the glycosylation, involves a free energy barrier of 32 kcal/mol, i.e., 11 kcal/mol higher than the corresponding reaction for the *on* configuration. One could guess from the GEI and TS conformations that a similar $^4C_1 \rightarrow [^4E]^{\ddagger} \rightarrow {}^{1,4}B_{/4}E$ catalytic itinerary for transglycosylation, as found for the *on* configuration, is operative here. However, the donor pyranose ring exhibits a $^4C_1$ conformation at the products, leading to an unusual $^4C_1 \rightarrow [^4H_3]^{\ddagger} \rightarrow {}^4C_1$ conformational itinerary (Figure 6c).

Table 2. Computed Values (kcal/mol) of the Reaction Free Energy Barriers for *On* and *Off* Configurations

	$\Delta G^{\ddagger}_{on}$	$\Delta G^{\ddagger}_{off}$	$\Delta\Delta G^{\ddagger}_{on-off}$
transglycosylation	12	28	16
glycosylation	21	32	11

Therefore, net removal of the 2-OH···nucleophile interaction increases the glycosylation and transglycosylation free energy barriers by 11 and 16 kcal/mol, respectively ($\Delta\Delta G^{\ddagger}_{on-off}$: Table 2). These values, quantified for the first time by *ab initio* methods in the native enzyme, reinforce previous experimental estimations that concluded that this interaction contributes >10 kcal/mol to the TS stabilization in retaining GHs.[26,27] In addition, the simulations show that the 2-OH···nucleophile interaction contributes significantly to the conformational itinerary of the substrate during catalysis.

3.5. Effect of the Asn175···2-OH Interaction on the Reaction Energy Barrier. The above results reveal that the 2-OH···nucleophile interaction is crucial for catalysis; thus, any interaction affecting it is expected to have an impact on the enzymatic activity. According to our simulations, the only residue that can directly influence the 2-OH···nucleophile interaction is Asn175, which forms a hydrogen bond with 2-OH. Interestingly, experiments show that replacing Asn175 by alanine, which would suppress the Asn175···2-OH interaction, abolishes enzymatic activity.[24]

To test how Asn175Ala mutation affects the energy barriers, we performed additional QM/MM calculations in the GEI, TS, and MC for both wild-type and mutant systems (see details in the SI), in the spirit of the work of Bueren-Calabuig et al.[55] The

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

transglycosylation energy barrier turned out to be unaffected. However, the glycosylation barrier increased by 6.5 kcal/mol, consistent with the fact that Asn175Ala mutation inactivates *Sc*Gas2.[24]

We can conclude that the TS stabilizing role of Asn175 affects notoriously the glycosylation step (in which the 2-OH··· nucleophile interaction is more important due to the negatively charged nucleophile) but leaves the transglycosylation step equally efficient.

4. FINAL REMARKS AND CONCLUSIONS

Trapping a GH glycosyl−enzyme intermediate (GEI) structure provides direct evidence of the double displacement mechanism. The observation of a GEI together with a bound substrate is particularly difficult, as usually acceptors diffuse out of the active site when the GEI is formed to leave space for amply available water molecules for deglycosylation,[20] precluding the characterization of the corresponding enzyme complex. Here we have trapped a natural GEI structure of *Sc*Gas2 in complex with an acceptor substrate by co-crystallyzation, using a mutant of the acid/base residue and activated substrates. Subsequent enhanced sampling QM/MM approaches show that the reported structure represents a catalytically competent form of the wild-type enzyme. The system with the 2-OH··· nucleophile interaction formed—which is the most favored—is able to react easily by transglycosylation ($\Delta G^{\ddagger}$ = 12 kcal/mol). The QM/MM simulations reveal that the 2-OH··· nucleophile interaction not only affects the energy barriers of the catalytic reaction but also changes the conformational itinerary (from $^{4}C_{1} \rightarrow [^{4}E]^{\ddagger} \rightarrow {}^{1,4}B/{}^{4}E$ for the *on* configuration to $^{4}C_{1} \rightarrow [^{4}H_{3}]^{\ddagger} \rightarrow {}^{4}C_{1}$ for the *off* configuration). The lack of this interaction increases the transglycosylation barrier by 16 kcal/mol and leads to a more asynchronous dissociative mechanism. This highlights the two-fold role of the 2-OH··· nucleophile interaction in the GH mechanism: (i) it stabilizes the reaction TS, lowering the free energy barrier; and (ii) it keeps the substrate at the MC in a conformation that is preactivated for catalysis.

The simulations have solved the network of hydrogen bond interactions around the −1 saccharide, revealing that the 2-OH interacts with the nucleophile as well as the NH_2 group of the relevant Asn175 residue. This interaction, not fully characterized in the X-ray structure (likely due to the Glu176Gln mutation), also contributes to TS stabilization (by 6.5 kcal/mol). This suggest that mutations affecting the 2-position are expected to increase energy barriers. Given this scenario, one can envisage that the use of these mutants with activated substrates—aryl- or fluor-substituted donors—to generate the GEI, followed by addition of suitable acceptors to intercept the intermediate, could result in high yields of transglycosylation products, avoiding secondary hydrolysis. This could explain the observed enhancement in the hydrolysis/transglycosylation ratio recently found for a GH1 enzyme:[6] the Asn163Ala variant (Asn163 is equivalent to Asn175 in *Sc*Gas2) leads to higher transglycosylation yields (from 30% in the wild-type to 80% for the variant). Interestingly, this interaction pattern is conserved among several GHs:[56] Asn126 in Cex (PDB 2HIS), His108 in CtLic26A (PDB 2CIP), Asn163 in Ttβ-gly (PDB 1UG6), Asn175 in Cel7A (PDB 4C4C), Asn175 in TxAbf (PDB 2VRQ), or Asn127 in E-82 xylanase (PDB 2D24). Thus, targeting the 2-OH interacting residue may be a promising strategy to rationally convert a GH into a TG. Experiments along this line are currently underway in our laboratory.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b10092.

Additional experimental and computational details, including Figures S1−S9, Schemes S1 and S2, and Tables S1 and S2 (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*d.m.f.vanaalten@dundee.ac.uk
*rhurtado@bifi.es
*c.rovira@ub.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by grants CTQ2013-44367-C2-2-P (to R.H.-G.) and CTQ2014-55174 (to C.R.) from MINECO, as well as grant 2014SGR-987 (to C.R.) from AGAUR. We acknowledge the computer support, technical expertise, and assistance provided by the Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS). L.R. thanks the University of Barcelona for an APIF fellowship.

■ REFERENCES

(1) Seeberger, P. H.; Werz, D. B. *Nature* **2007**, *446*, 1046.
(2) Sears, P.; Wong, C. H. *Science* **2001**, *291*, 2344.
(3) Feng, H. Y.; Drone, J.; Hoffmann, L.; Tran, V.; Tellier, C.; Rabiller, C.; Dion, M. *J. Biol. Chem.* **2005**, *280*, 37088.
(4) Osanjo, G.; Dion, M.; Drone, J.; Solleux, C.; Tran, V.; Rabiller, C.; Tellier, C. *Biochemistry* **2007**, *46*, 1022.
(5) Arnold, F. H. *Nature* **2001**, *409*, 253.
(6) Teze, D.; Hendrickx, J.; Czjzek, M.; Ropartz, D.; Sanejouand, Y.-H.; Tran, V.; Tellier, C.; Dion, M. *Protein Eng., Des. Sel.* **2014**, *27*, 13.
(7) Frutuoso, M. A.; Marana, S. R. *Protein Pept. Lett.* **2013**, *20*, 102.
(8) Mackenzie, L. F.; Wang, Q.; Warren, R. A. J.; Withers, S. G. *J. Am. Chem. Soc.* **1998**, *120*, 5583.
(9) Malet, C.; Planas, A. *FEBS Lett.* **1998**, *440*, 208.
(10) Shaikh, F. A.; Withers, S. G. *Biochem. Cell Biol.* **2008**, *86*, 169.
(11) Vocadlo, D. J.; Davies, G. J.; Laine, R.; Withers, S. G. *Nature* **2001**, *412*, 835.
(12) Bissaro, B.; Monsan, P.; Faure, R.; O'Donohue, M. J. *Biochem. J.* **2015**, *467*, 17.
(13) Tewari, Y. B.; Goldberg, R. N. *J. Biol. Chem.* **1989**, *264*, 3966.
(14) Baumann, M. J.; Eklof, J. M.; Michel, G.; Kallas, A. M.; Teeri, T. T.; Czjzek, M.; Brumer, H., 3rd *Plant Cell* **2007**, *19*, 1947.
(15) Monsan, P.; Remaud-Simeon, M.; Andre, I. *Curr. Opin. Microbiol.* **2010**, *13*, 293.
(16) Montagna, G.; Cremona, M. L.; Paris, G.; Amaya, M. F.; Buschiazzo, A.; Alzari, P. M.; Frasch, A. C. *Eur. J. Biochem.* **2002**, *269*, 2941.
(17) Bissaro, B.; Durand, J.; Biarnés, X.; Planas, A.; Monsan, P.; O'Donohue, M. J.; Fauré, R. *ACS Catal.* **2015**, *5*, 4598.
(18) Piens, K.; Faure, R.; Sundqvist, G.; Baumann, M. J.; Saura-Valls, M.; Teeri, T. T.; Cottaz, S.; Planas, A.; Driguez, H.; Brumer, H. *J. Biol. Chem.* **2008**, *283*, 21864.
(19) Namchuk, M. N.; Withers, S. G. *Biochemistry* **1995**, *34*, 16194.
(20) White, A.; Tull, D.; Johns, K.; Withers, S. G.; Rose, D. R. *Nat. Struct. Biol.* **1996**, *3*, 149.
(21) Mazan, M.; Ragni, E.; Popolo, L.; Farkas, V. *Biochem. J.* **2011**, *438*, 275.
(22) Latge, J. P. *Mol. Microbiol.* **2007**, *66*, 279.
(23) Hartland, R. P.; Fontaine, T.; Debeaupuis, J. P.; Simenel, C.; Delepierre, M.; Latge, J. P. *J. Biol. Chem.* **1996**, *271*, 26843.

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332

(24) Hurtado-Guerrero, R.; Schuttelkopf, A. W.; Mouyna, I.; Ibrahim, A. F.; Shepherd, S.; Fontaine, T.; Latge, J. P.; van Aalten, D. M. *J. Biol. Chem.* **2009**, *284*, 8461.
(25) 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.
(26) Notenboom, V.; Birsan, C.; Nitz, M.; Rose, D. R.; Warren, R. A.; Withers, S. G. *Nat. Struct. Biol.* **1998**, *5*, 812.
(27) Hovel, K.; Shallom, D.; Niefind, K.; Belakhov, V.; Shoham, G.; Baasov, T.; Shoham, Y.; Schomburg, D. *EMBO J.* **2003**, *22*, 4922.
(28) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 2126.
(29) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1997**, *53*, 240.
(30) Schuttelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1355.
(31) Laskowski, R. A.; Moss, D. S.; Thornton, J. M. *J. Mol. Biol.* **1993**, *231*, 1049.
(32) Colonna, B.; Harding, V. D.; Nepogodiev, S. A.; Raymo, F. M.; Spencer, N.; Stoddart, J. F. *Chem. - Eur. J.* **1998**, *4*, 1244.
(33) Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M. F.; Walker, R.; Zhang, W.; Merz, K. M.; Wang, B.; Hayik, S.; Roitberg, A. E.; Seabra, G.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wu, X.; Brozell, S.; Steinbrecher, T.; Gohlke, H.; Yang, L.; Tan, C.; Mongan, J.; Hornak, V.; Cui, G.; Mathews, D. H.; Seetin, M. G.; Sagui, C.; Babin, V.; Kollman, P. *AMBER 11*; University of California, San Francisco, 2010.
(34) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins: Struct., Funct., Genet.* **2006**, *65*, 712.
(35) 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.
(36) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.
(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(38) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.
(39) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941.
(40) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471.
(41) Troullier, N.; Martins, J. L. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1991**, *43*, 1993.
(42) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(43) Ireta, J.; Neugebauer, J.; Scheffler, M. *J. Phys. Chem. A* **2004**, *108*, 5692.
(44) Ardèvol, A.; Rovira, C. *J. Am. Chem. Soc.* **2015**, *137*, 7528.
(45) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302.
(46) Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. *J. Phys. Chem. B* **2005**, *109*, 6676.
(47) Biarnés, X.; Ardèvol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. *J. Am. Chem. Soc.* **2007**, *129*, 10686.
(48) Williams, S. J.; Withers, S. G. *Carbohydr. Res.* **2000**, *327*, 27.
(49) Sinnott, M. L.; Jencks, W. P. *J. Am. Chem. Soc.* **1980**, *102*, 2026.
(50) Ardèvol, A.; Rovira, C. *Angew. Chem., Int. Ed.* **2011**, *50*, 10897.
(51) Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J. *Curr. Opin. Struct. Biol.* **2014**, *28C*, 1.
(52) Biarnés, X.; Ardèvol, A.; Iglesias-Fernández, J.; Planas, A.; Rovira, C. *J. Am. Chem. Soc.* **2011**, *133*, 20301.
(53) Schramm, V. L.; Shi, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 657.
(54) Davies, G. J.; Planas, A.; Rovira, C. *Acc. Chem. Res.* **2012**, *45*, 308.
(55) Bueren-Calabuig, J. A.; Pierdominici-Sottile, G.; Roitberg, A. E. *J. Phys. Chem. B* **2014**, *118*, 5807.
(56) Money, V. A.; Smith, N. L.; Scaffidi, A.; Stick, R. V.; Gilbert, H. J.; Davies, G. J. *Angew. Chem., Int. Ed.* **2006**, *45*, 5136.

DOI: 10.1021/jacs.5b10092
J. Am. Chem. Soc. 2016, 138, 3325−3332