Article

QM/MM Simulations of Enzymatic Hydrolysis of Cellulose: Probing the Viability of an Endocyclic Mechanism for an Inverting Cellulase

Caroline S. Pereira, Rodrigo L. Silveira, and Munir S. Skaf*

Cite This: *J. Chem. Inf. Model.* 2021, 61, 1902−1912
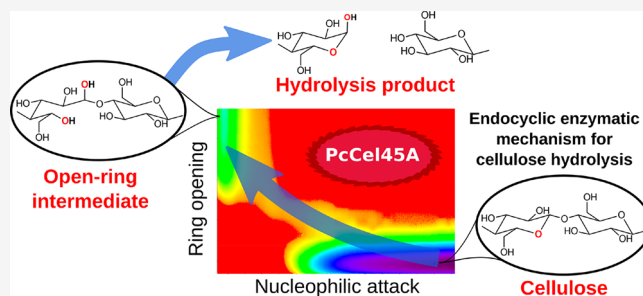
Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Glycoside hydrolases (GH) cleave carbohydrate glycosidic bonds and play pivotal roles in living organisms and in many industrial processes. Unlike acid-catalyzed hydrolysis of carbohydrates in solution, which can occur either via cyclic or acyclic oxocarbenium-like transition states, it is widely accepted that GH-catalyzed hydrolysis proceeds via a general acid mechanism involving a cyclic oxocarbenium-like transition state with protonation of the glycosidic oxygen. The GH45 subfamily C inverting endoglucanase from *Phanerochaete chrysosporium* (PcCel45A) defies the classical inverting mechanism as its crystal structure conspicuously lacks a general Asp or Glu base residue.

Instead, PcCel45A has an Asn residue, a notoriously weak base in solution, as one of its catalytic residues at position 92. Moreover, unlike other inverting GHs, the relative position of the catalytic residues in PcCel45A impairs the proton abstraction from the nucleophilic water that attacks the anomeric carbon, a key step in the classical mechanism. Here, we investigate the viability of an endocyclic mechanism for PcCel45A using hybrid quantum mechanics/molecular mechanics (QM/MM) simulations, with the QM region treated with the self-consistent-charge density-functional tight-binding level of theory. In this mechanism, an acyclic oxocarbenium-like transition state is stabilized leading to the opening of the glucopyranose ring and formation of an unstable acyclic hemiacetal that can be readily decomposed into hydrolysis product. *In silico* characterization of the Michaelis complex shows that PcCel45A significantly restrains the sugar ring to the $^4C_1$ chair conformation at the −1 subsite of the substrate binding cleft, in contrast to the classical exocyclic mechanism in which ring puckering is critical. We also show that PcCel45A provides an environment where the catalytic Asn92 residue in its standard amide form participates in a cooperative hydrogen bond network resulting in its increased nucleophilicity due to an increased negative charge on the oxygen atom. Our results for PcCel45A suggest that carbohydrate hydrolysis catalyzed by GHs may take an alternative route from the classical mechanism.

1. INTRODUCTION

Glycoside hydrolases (GH) are enzymes that hydrolyze glycosidic bonds of carbohydrates. They play major biological roles in all life kingdoms and are applied in a variety of industrial processes, including depolymerization of lignocellulosic biomass for sustainable production of biofuels, chemicals, and valuable materials.[1] The large number of chemical combinations and stereochemical variations of sugar moieties in carbohydrates is accompanied by a large number of GHs in nature, with different chemical and structural specificities.[2,3] Hydrolysis catalyzed by GHs follows a general acid mechanism, with only a few exceptions,[4] involving two key residues: a proton donor (usually Asp or Glu) and a nucleophile/base (usually deprotonated Asp or Glu).[5]

β-Glycoside acid-catalyzed hydrolysis in solution can follow one of two competing pathways: (i) an entropy-favored exocyclic route, which involves protonation of the (exocyclic) glycosidic oxygen and a cyclic oxocarbenium-like transition state (TS), or (ii) an enthalpy-driven endocyclic route, with protonation of an endocyclic oxygen followed by opening of the pyranose ring and an acyclic oxocarbenium-like TS.[6,7]

Previous GH mechanistic studies have considered only GHs that follow the exocyclic route,[3] which, due to stereoelectronic effects,[8] requires puckering of the glycosidic ring from the most stable $^4C_1$ chair conformation to $^1S_3$ or $^2S_0$ skew-boat conformation[9] and the protonation of the glycosidic oxygen atom, which bears a small proton affinity.[10]

In 1986, Post and Karplus proposed an endocyclic route for the lysozyme-catalyzed hydrolysis of carbohydrate.[11] The proposed mechanism was later dismissed after quantum mechanics/molecular mechanics (QM/MM) simulations revealed a prohibitively large energy barrier for this mechanism in lysozyme.[12] Nevertheless, several studies have indicated an endocyclic cleavage of carbohydrates in solution[6,7,13−19] and
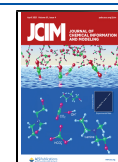
exhibited evidence that endocyclic cleavage is enhanced in β-glycosides with protecting groups that conformationally restrict the ring in $^4C_1$ conformation and increase inner strain.[14,15] Here, QM/MM molecular dynamics (MD) simulations with enhanced sampling were performed on an endoglucanase belonging to family GH45 (subfamily C) from *Phanerochaete chrysosporium* (PcCel45A) and revealed that an endocyclic route for cellulose hydrolysis is a physically feasible mechanism, with an activation free energy comparable to those of other GHs known to proceed via an exocyclic pathway.

GH45 enzymes are widely applied in the textile industry as detergents and in the polishing of denim fabrics.[20,21] There is a keen interest in PcCel45A, particularly, because of its close evolutionary relationship with β-expansins and lytic trans-glycosylase MltA.[3,22] Whereas PcCel45A hydrolytically cleaves β-glucans,[23,24] β-expansins seem to lack any lytic activity[25,26] but are capable of loosening the plant cell wall architecture by a still unknown mechanism.[22,27] Lytic transglycosylase MltA, in turn, cleaves bacterial cell walls nonhydrolytically, and the mechanism is also not established.[28,29] Although they exhibit different functions, PcCel45A, β-expansins, and lytic trans-glycosylase MltA share similar catalytic machinery.[23,25,29,30] Elucidation of the PcCel45A mechanism, therefore, can provide valuable insights toward understanding the long-sought mechanisms of β-expansins and lytic transglycosylase MltA.

PcCel45A is an inverting GH,[30] but its reported crystal structure[23,30] defies the classical mechanism for GH-catalyzed carbohydrate hydrolysis. The canonical inverting mechanism proceeds via a single-displacement reaction, in which an Asp or Glu residue donates a proton to the exocyclic glycosidic oxygen while another Glu or Asp acts as general base to abstract a proton from a nucleophilic water to attack the anomeric carbon (Scheme 1).[3,5] PcCel45A lacks the

Scheme 1. Classical Inverting Mechanism[a]
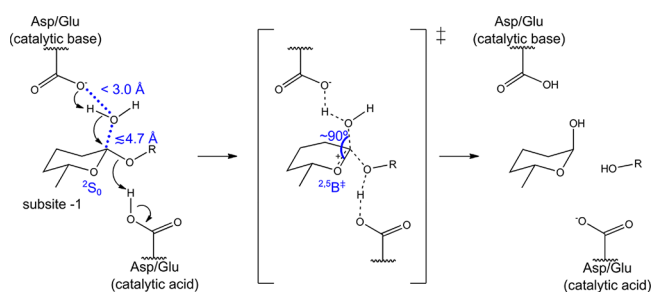
[a]Geometric parameters (distances and angles) and substrate conformation ($^2S_0$ and $^{2,5}B$) required for the single-displacement exocyclic mechanism are shown in blue.

conspicuous Asp or Glu residue at the general base position and has instead an Asn residue at position 92, similar to the assisting Asp residue position observed in GH45 subfamily A from *Humicola insolens* (HiCel45A).[31] Mutation of Asn92 to Asp leads to a significant reduction of activity and an increase of the optimum catalytic pH from 3 (wild type) to 5.5 (Asn92Asp mutant).[30] This indicates that Asn92 is a catalytic residue, unlike all previously characterized GHs which have Asp or Glu as catalytic base.[30] Our previous simulation results indicate that the decreased activity of the Asn92Asp mutant may be attributed to conformational changes of the Asp side

chain relative to Asn at position 92.[23] The reported crystal structures show that Asn92 and the catalytic acid residue (Asp114) are only 5 Å apart in PcCel45A,[23,30] whereas the distance between the catalytic base and acid residues is at least twice as large (∼10 Å) in other inverting GHs, which provides enough room for accommodating a water molecule between the general base and the anomeric carbon required for the nucleophilic substitution.[3,5,32,33] The relative positions of the catalytic residues in PcCel45A (Asn92 and Asp114) do not seem adequate to fit the classical inverting mechanism.

On the basis of high-resolution crystal structures, Nakamura et al.[30] proposed an alternative hydrolysis mechanism for PcCel45A in which Asn92, along with the backbone of several residues, assumes the imidic acid form (amide tautomer) to create a proton transfer relay chain from Asn92 to Asp114 (catalytic acid). However, water-assisted tautomerization of formamide from amide to imidic acid exhibits an activation free energy of ∼20 kcal/mol and a reaction free energy of ∼11 kcal/mol.[34] As PcCel45A lacks any obvious machinery to catalyze this tautomerization, the enzyme would have to undergo multiple amide−imidic acid tautomerization processes involving high barriers and leading to high-energy states in addition to barriers associated with the hydrolysis reaction itself. Here, we present evidence from QM/MM free energy calculations that PcCel45A adopts a much simpler reaction mechanism involving protonation of the endocyclic oxygen.

2. COMPUTATIONAL DETAILS

2.1. Construction of the Initial Structure and Classical MD Simulations. Our departing structure was the high-resolution crystal structure of PcCel45A obtained by Nakamura et al. (PDB ID 3X2M), which contains two cellopentaose molecules bound to the positive and negative subsites of the enzyme.[30] Protonation states of titratable residues were estimated using the H++ server[35,36] considering pH 3, which is the estimated optimal pH of the enzyme.[30] Missing hydrogen atoms were then added to the structure, and all crystallographic waters were maintained. The whole system was immersed in a rectangular water box at least 15 Å thick from the protein, and 23 Na$^+$ and 25 Cl$^-$ ions were added to render the system electrically neutral.

Classical MD simulations were performed using periodic boundary conditions, with long-range interactions treated with particle mesh Ewald (PME)[37] and short-range interactions truncated at a cutoff radius of 8 Å. Chemical bonds involving hydrogen atoms were kept rigid with SHAKE,[38] and a time step of 2 fs was employed to solve the equations of motion. The temperature was kept constant at 310 K with the Langevin thermostat using a collision frequency of 1.0 ps^{-1}. Pressure was kept constant only during the 200 ps equilibration step via the Berendsen barostat with a relaxation time of 1.0 ps. Production simulations were performed at constant volume. The simulations were carried out with Amber16[39] using the Amber ff14SB[40] and Glycam06[41] force fields. Water was modeled with the TIP3P model.[42]

A PcCel45A−cellodecaose model was built for simulations by first equilibrating the 3X2M system comprised of PcCel45A bound to two cellopentaose molecules according to the following steps: (1) 1000 steps of energy minimization (500 steps using the steepest descent algorithm followed by 500 steps of conjugate gradient) followed by 200 ps of MD with the protein and the cellopentaose chains restrained; (2) 2500 steps of energy minimization (1000 steps of steepest descent +

1500 steps of conjugate gradients) followed by 200 ps of MD with only the α-carbons of the protein restrained. Subsequently, one hydroxyl group and one hydrogen atom from the cellopentaose molecules were removed and a glycosidic bond was created to generate a single 10-glucose glucan chain (cellodecaose). The system was then equilibrated according to steps 1 and 2, after which an additional 200 ns of MD was performed with no restraints. The coordinates obtained at the end of this simulation were used as the starting point for the QM/MM simulations, described below.

The rationale behind our choice of structure and protonation states of titratable residues is as follows. In addition to our previous structures with PDB IDs 5KJQ and 5KJO,[23] which correspond to a cellobiose-bound structure and an unliganded structure (thus, inappropriate for our purposes here), there are four other high-resolution wild type PcCel45A structures obtained from neutron diffraction and X-ray crystallography available to date, with PDB IDs[30] 3X2M, 3X2L, 3X2O, and 3X2P. The former two were obtained at cryogenic conditions, whereas the latter two are room temperature structures. The 3X2O structure is unliganded and 3X2P contains a cellopentaose bound to the positive subsites only; both would require docking of a longer substrate chain into the catalytic cleft for reaction mechanism studies. 3X2L, in turn, is bound to noncarbohydrate ligands, whereas 3X2M contains no other ligands but two cellopentaose molecules bound to positive and negative subsites, which renders it the best structure to our study. The catalytic residues assume the same position in both the 3X2P (room temperature) and 3X2M (cryo temperature) structures. The structures reported by Nakamura et al.[30] were obtained at pH 8.0, and the enzyme is inactive at this pH, as the authors pointed out. Thus, the protonation state of titratable residues of the crystal structures should be set to the corresponding optimal pH for atomistic modeling of the reaction mechanism. This is particularly evident for residue Asp114 (general acid), which should be protonated for both the exocyclic and endocyclic mechanisms but appears deprotonated in the room temperature (3X2O and 3X2P) and cryo temperature (3X2M) structures. Furthermore, in our simulations, residue Asn92 (general base) is considered in its typical amide form, since the N atom that would act as a Brønsted base in the imidic acid form is not in the proper position to abstract a proton from the nucleophilic water (see Results and Discussion). In contrast, the O atom of Asn92 in the amide form is, in fact, properly positioned to abstract the proton from the water molecule (see Results and Discussion).

2.2. Hybrid QM/MM Umbrella Sampling Simulations. The equilibrated system was then partitioned into quantum and classical regions for QM/MM simulations. The quantum region comprised the active site (side chain of residues Thr16, Asn92, His112, and Asp114) and the two glucopyranosyl residues linked by the scissile glycosidic bond (Figure S1). Residue Asn92 was considered in the amide form, as already mentioned. The QM region was treated with the self-consistent-charge density-functional tight-binding (SCC-DFTB) method,[43] and the rest of the system was treated with the Amber ff14SB25[40] and Glycam06[41] force fields, along with the TIP3P[42] water model. The computations were conducted using Amber16.[39,44−46] A QM/MM simulation lasting 2.5 ns without any restraint was performed starting from the coordinates of the classical MD simulation. Our choice of the SCC-DFTB method was based on previous

successful studies on cellulose hydrolysis by other GHs (GH6,[47] GH7,[48] and GH45 subfamily A[49]).

Umbrella sampling (US) simulations were first used to generate the distortion free energy surface of the glycosidic ring at the −1 subsite of PcCel45A and the free energy surface of a single cellobiose in aqueous solution for comparison. The simulation of a cellobiose in water was performed at the same conditions as the simulation of the enzyme, except that the edge of the simulation box was 15 Å from the cellobiose and the quantum region comprised just the cellobiose molecule. Two-dimensional free energy profiles were computed along Cremer and Pople puckering coordinates[50,51] q_x and q_y for a six-atom (C1, C2, C3, C4, C5, and O5) ring, available in the Plumed plugin.[52,53] Harmonic potentials of the form $1/2k(r − r_0)^2$ were used with $k = 400$ kcal mol^{-1} Å^{-2}. Different windows were employed to sample q_x and q_y from −1 to +1 Å at intervals of 0.1 Å. The following steps were performed for each umbrella sampling window: (1) 1000 steps of energy minimization with the steepest descent algorithm; (2) 20 ps of equilibration; (3) 60 ps of productive simulation, which was used for obtaining the free energy.

Sampling in the first US window started from an equilibrated structure taken from the QM/MM unrestrained simulation at $^4C_1$ conformation. Initial structures for the other US windows were sequentially taken from previous adjacent and overlapping windows, after the equilibration step. The free energy surfaces were obtained using the weighted histogram analysis method (WHAM)[54] with 210 bins for each reaction coordinate and a convergence tolerance of 10^{-7}.

Restrained simulations were also employed to generate molecular configurations depicting the imminence of the classical inverted mechanism attack with Asn92 in the imidic acid form. A random frame from the previous QM/MM MD without any restraint with Asn92 in amide form was chosen, and then Asn92 was converted into the imidic acid form. An initial QM/MM MD simulation of 50 ps was performed for equilibration without any restraint. Then, a water molecule was added to the QM region. Simulations with the following two-dimensional restraints of interatomic distances (d) were used to carry the system to the configuration close to the nucleophilic attack: $d(C1_{glc}−O_w)$ was sampled from 6 to 1.7 Å with intervals of 0.06 Å and $d(H_{Asp114}−O4_{glc})$ was restrained around 1.0 Å, where $C1_{glc}$ is the anomeric carbon, O_w is the water oxygen, H_{Asp114} is the Asp114 (catalytic acid) acidic proton, and $O4_{glc}$ is the (exocyclic) glycosidic oxygen. Harmonic potentials of the form $k(r − r_0)^2$ were used with $k = 450$ kcal mol^{-1} Å^{-2}, and simulations lasting 5 ps were performed for each window.

Umbrella sampling simulations were also applied to generate the reaction free energy surface along a two-dimensional reaction coordinate (RC1, RC2) in the QM/MM simulations. RC1 = $d(C1_{glc}−O_w)$ was sampled from 1.3 to 4.8 Å, and RC2 = $d(C1_{glc}−O5_{glc})$ was sampled from 1.3 to 4.8 Å, where $O5_{glc}$ is the endocyclic oxygen. Harmonic potentials of the form $k(r − r_0)^2$, with $k = 200$ kcal mol^{-1} Å^{-2}, were employed to restrain RC1 and RC2 at intervals of 0.1 Å for the whole reaction space. Additional potentials of $k = 500$ kcal mol^{-1} Å^{-2} were used to sample RC1 and RC2 around the transition state in intervals of 0.1 Å, with the potentials centered between the windows of $k = 200$ kcal mol^{-1} Å^{-2}. For each umbrella sampling window, the following steps were performed: (1) 1000 steps of energy minimization with the steepest descent algorithm; (2) 20 ps of equilibration; (3) 60 ps of productive
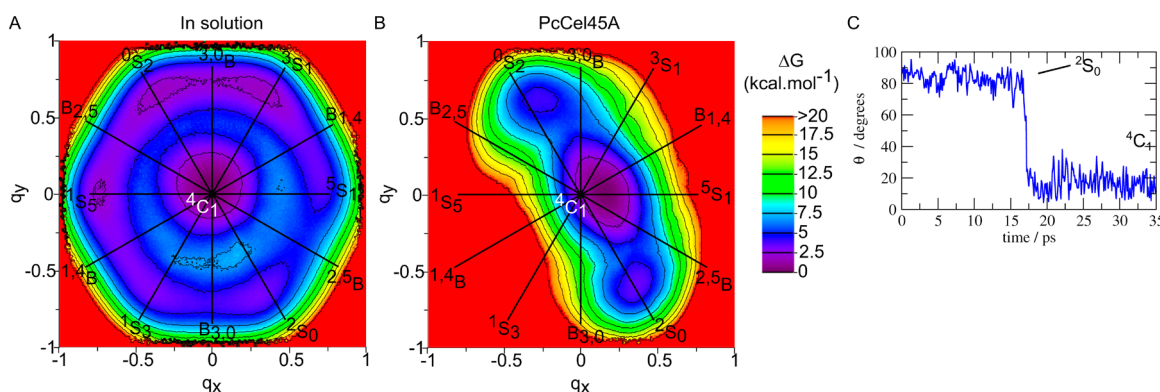
Figure 1. Free energy surface along puckering coordinates of a glycosidic ring of (A) cellobiose in solution and (B) PcCel45A substrate at −1 subsite. (C) Cremer−Pople puckering coordinate (θ) of PcCel45A substrate at −1 subsite along an unrestrained QM/MM simulation.

simulation, which was used for analyses. Similarly, sampling in the first US window started from an equilibrated structure taken from the QM/MM unrestrained simulation of the reactant and the initial structures for the other windows were sequentially taken from previous adjacent and overlapping windows, after equilibration. The free energy surfaces were obtained with WHAM[54] using 400 bins for the reaction coordinate and a convergence tolerance of 10^{-7}.

Two additional unrestrained QM/MM simulations were performed to analyze the stabilization of the positive charge of the protonated Asn92 by nearby residues. These simulations were performed under the conditions described above, except that the quantum region comprised the side chain of residue Gln39 and the backbone of residue Val89, in addition to the residues mentioned above. In one of the simulations, Asn92 was considered neutral, and in the other simulation Asn92 was considered protonated (with positive charge). Each simulation lasted 1.2 ns, and Mulliken charges and hydrogen bond distances were obtained.

2.3. QM/MM Transition Path Sampling Simulations. In order to further corroborate the results obtained from umbrella sampling simulations, fully unbiased reactive trajectories were generated with the aimless shooting (AS)[55,56] version of transition path sampling (TPS).[57] Transition state guesses from restrained QM/MM simulations were used as starting shooting points for AS runs, which were performed with the same conditions and QM region as described in section 2.2. Each AS trajectory crossing the transition state region lasted 1600 fs and Δt was defined as 16 fs, similar to previous studies.[47−49] Three independent AS runs of 100 trajectories, starting from different shooting points, were carried out for each the following reactive processes: (i) exocyclic mechanism with Asn92 in amide form (49 accepted, 201 inconclusive, and 50 rejected); (ii) exocyclic mechanism with Asn92 in imidic acid form (87 accepted, 141 inconclusive, and 72 rejected); and (iii) endocyclic mechanism (136 accepted, 43 inconclusive, and 121 rejected). Each trajectory is accepted when its end points connect reactant and product, rejected when the end points are both reactant or both product, and inconclusive when the trajectory does not evolve to these basin states. For exocyclic mechanisms, a configuration is considered as reactant if $d(C1_{glc}-O_w) \geq 3.0$ Å and $d(C1_{glc}-O4_{glc}) \leq 1.6$ Å and as product if $d(C1_{glc}-O_w) \leq 1.6$ Å and $d(C1_{glc}-O4_{glc}) \geq 2.4$ Å. For the endocyclic mechanism, a configuration is considered as reactant if $d(C1_{glc}-O_w) \geq 3.0$ Å and $d(C1_{glc}-O5_{glc}) \leq 1.6$ Å and as product if $d(C1_{glc}-O_w) \leq$

1.6 Å and $d(C1_{glc}-O5_{glc}) \geq 2.4$ Å. A representative reactive trajectory of each mechanism was then used to generate illustrative movies (Movies S1−S3).

2.4. DFT Calculations. Density functional theory (DFT) calculations were performed to obtain the potential energy difference of Asn92 in amide and imidic acid forms in a reduced model of PcCel45A. The reduced model comprised the side chain of residues Asn92, Gln39, and Pro88, the backbone of residues Pro88, Val89, and Gln90, two glycosyl residues from −1 and +1 subsites, and two water molecules hydrogen bonded the Asn92. The valences of dangling bonds were completed with hydrogen atoms. Each system comprised 103 atoms. Initially, the geometry of the two structures were optimized at the M062X/6-31+G(d,p) level[58] of theory without any restraint, with frequencies obtained at the same level of theory. Then, single point calculations were carried out at the M062X/cc-pVDZ level,[58] from which the ground state energy for each system was obtained. Water solvation effects were introduced in all calculations using the solvation model based on density (SMD).[59] Calculations were carried out using the Gaussian 09 quantum chemistry package.[60]

3. RESULTS AND DISCUSSION

3.1. Ring Conformation at the −1 Subsite. The free energy surface along the puckering coordinates shows that, in solution, the glucopyranose ring exhibits many thermally accessible conformations (Figure 1A; coordinates of the main minima are available in the Supporting Information), with the $^4C_1$ chair conformation being the most stable. In PcCel45A, the ring conformation at the −1 subsite is substantially restrained (Figure 1B; coordinates of the main minima are available in the Supporting Information). Yet, the $^4C_1$ conformation remains the most stable ring conformer. The $^2S_0$ and $^0S_2$ conformations—distorted conformations that can facilitate the classical inverting mechanism—lie approximately 3.4 and 4.0 kcal/mol above $^4C_1$, respectively. Unbiased QM/MM simulation starting with the glycosidic ring in the $^2S_0$ conformation showed that it readily evolves to the $^4C_1$ chair conformation within 17 ps (Figure 1C).

Remarkably, a crystal structure of lytic transglycosylase MltA in the presence of substrate,[29] which is closely related to PcCel45A evolutionarily and shares with it many similarities in the catalytic site, also exhibits the glucopyranose ring at the −1 subsite in the $^4C_1$ conformation. Together, these results suggest that PcCel45A might not require substrate distortion prior to catalysis, in contrast to other inverting GHs, which
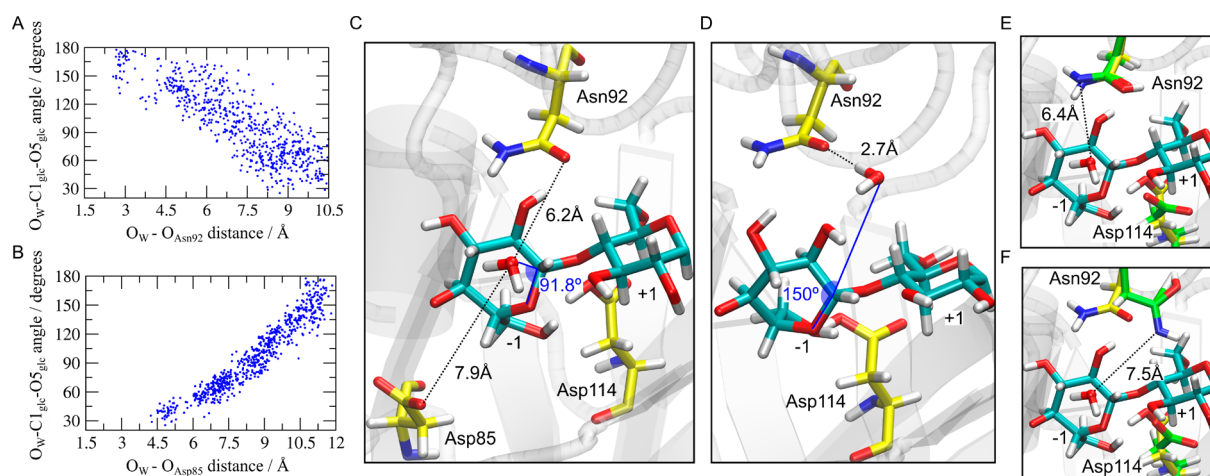
Figure 2. O_w–$C1_{glc}$–$O5_{glc}$ angle versus (A) O_{Asn92} and (B) O_{Asp85} distances for structures having a potential nucleophilic water and with the substrate at −1 subsite in the $^2S_0$ metastable conformation. Representative snapshots where (C) the O_w–$C1_{glc}$–$O5_{glc}$ angle is close to 90° (condition 1), but there is no hydrogen bond between the base and the water molecule (condition 2), and where (D) there is a hydrogen bond between the water molecule and the Asn92 base (condition 2), but the O_w–$C1_{glc}$–$O5_{glc}$ angle of 150° is much greater than that required for the single-displacement nucleophilic attack (condition 1). (E and F) Superposition of simulated PcCel45A−substrate complex (enzyme in yellow and substrate in cyan) with crystal structures (in green) with the Asn92 in the tautomer imidic acid in two conformations. The dotted lines indicate the distances between the Asn92 N atom and the potential nucleophilic water that assumes the ideal the O_w–$C1_{glc}$–$O5_{glc}$ angle of 90°.

induce distortion of the glycosidic ring to $^2S_0$ conformation to reduce the activation barrier in the inverting exocyclic mechanisms.[9]

3.2. Exploring the Possibility of the Exocyclic Mechanism.

Ignoring for a moment our finding that PcCel45A stabilizes the chair $^4C_1$ conformation instead of puckered ones, we assessed the possibility of the classical exocyclic mechanism starting from the metastable $^2S_0$ conformation − which is the distorted conformation closest to the transition state $^{2,5}B^{\ddagger}$ conformation.[9] For that, we selected all structures from the umbrella sampling QM/MM simulation with the glycosidic ring of the −1 subsite in the $^2S_0$ metastable conformation.[9] From this set, we considered all potential nucleophilic water molecules, defined as those having their oxygen atoms within 4.7 Å from the glycosidic ring $C1_{glc}$, which is approximately the distance corresponding to the minimum of free energy (see Figure 5B). In the GH8 inverting endoglucanase crystal structure this distance is 3.8 Å.[61] As shown in Scheme 1, the single-displacement exocyclic mechanism depends on two conditions: (1) the nucleophilic water oxygen (O_w) must form an angle ∼90° with the $C1_{glc}$ and $O5_{glc}$ of the glycosidic ring (O_w–$C1_{glc}$–$O5_{glc}$ angle), and (2) the catalytic base must be well positioned to abstract a proton from the nucleophilic water; i.e., the base must establish a hydrogen bond with the water molecule.[3,9] We then analyzed these two conditions to see if the exocyclic mechanism can occur.

We measured the O_w–$C1_{glc}$–$O5_{glc}$ angle and the distance between O_w and the oxygen atom of the catalytic base in the ensemble of structures of potential nucleophilic water molecules with the sugar ring in the metastable $^2S_0$ conformation at the −1 subsite. We also investigated the possibility of Asp85 to act as the catalytic base, as it is the closest Asp/Glu residue to the catalytic site. As pointed out, the O_w–$C1_{glc}$–$O5_{glc}$ angle must be around 90° (condition 1) and the catalytic base oxygen atom must be within the hydrogen bonding distance of ∼3 Å of O_w (condition 2) for the single-displacement nucleophilic attack to occur.

Parts A and B of Figure 2 show scatter plots of the O_w–$C1_{glc}$–$O5_{glc}$ angle versus the O_w–O_{Asn92} and O_w–O_{Asp85} distances, respectively. We observe that, for either Asn92 or Asp85, conditions 1 and 2 (angle of about 90° and distance of <3 Å) are never simultaneously satisfied along the simulation. If the O_w–$C1_{glc}$–$O5_{glc}$ angle is close to 90°, then O_w does not hydrogen bond the base (Figure 2C). On the other hand, if O_w hydrogen bonds the base, then the O_w–$C1_{glc}$–$O5_{glc}$ angle deviates from 90° (Figure 2D). Thus, any tentative obtaining of the free energy profile of the exocyclic mechanism would result in a nonphysical, prohibitively high barrier, because of these natural geometric constraints the enzyme poses. In fact, all calculations we tried resulted in erratic free energy profiles with no physical sense. Our classical MD simulations show that Asn92 and Asp85 exhibit very low mobility along 200 ns and remain in nearly the same configuration of crystal structure (Figure S2). Therefore, Asp85 and Asn92 do not seem to access conformations that would enable them to act as a general base in the classical inverting mechanism during the simulations.

To further support these results, we generated fully unbiased reactive trajectories of the classic exocyclic mechanism using the aimless shooting version[55,56] of transition path sampling (TPS)[57] (Movie S1). TPS is a powerful technique to investigate rare events, including reaction mechanisms, because it does not require an *a priori* definition of reaction coordinates to drive the system across the barrier and samples the trajectory space in a fully unbiased way.[57] Analysis of the ensemble of all the unbiased reactive trajectories reveals that the nucleophilic water does not get deprotonated, even when the system evolves to the product state with the water oxygen covalently bound to the anomeric carbon (Movie S1 and Figure S3A). These results reinforce then that Asn92 cannot abstract the proton from the nucleophilic water in the classic exocyclic mechanism.

We additionally investigated whether Asn92 in the tautomeric imidic acid form could act as the catalytic base under the above-mentioned conditions. Crystal structures obtained by Nakamura et al. with Asn92 assuming the imidic

acid form show that this residue can assume two different conformations,[30] as depicted in Figure 2E,F. Each of these crystal structures was then aligned to a simulated PcCel45A structure with the substrate in $^2S_0$ conformation and a putative nucleophilic water exhibiting the $O_w-C1_{glc}-O5_{glc}$ angle close to 90° (obtained from the umbrella sampling simulations). We observed that the distance between the Asn92 carbonyl nitrogen and the nucleophilic water is 6 or 7 Å depending on the conformation of Asn92, therefore indicating unfavorable geometric parameters for reaction.

To further explore the possibility of Asn92 in the imidic acid form to act as general base, we performed restrained QM/MM MD simulations with Asn92 in imidic form, applying a moving restraint to bring a water molecule from the bulk to the optimal position for a nucleophilic attack on the anomeric carbon and another fixed restraint to keep the glycosidic oxygen protonated by Asp114 (catalytic acid) (Scheme 1). Such a two-dimensional restraint comprised the interatomic distances (d) $d(C1_{glc}-O_w)$ and $d(H_{Asp114}-O4_{glc})$. Analyses of the trajectories show that the nucleophilic water establishes hydrogen bonds with (transient) water molecules, instead of interacting directly with Asn92, at the imminence of the nucleophilic attack (Figure 3A). As the Asn92 nitrogen atom
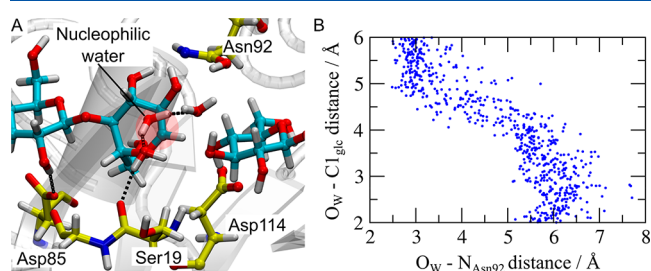
Figure 3. (A) Representative snapshot of restrained QM/MM MD simulations of PcCel45A with Asn92 in imidic acid form, showing the restrained water molecule (highlighted in red shadow) at the optimal position to perform a nucleophilic attack on the anomeric carbon. (B) Scatter plot of the O_w-C1_{glc} distance versus the O_W-O_{Asn92} distance during the restrained QM/MM MD simulations carrying a water molecule to the imminence of nucleophilic attack.

must be within the hydrogen bonding distance (~3 Å) of O_w (condition 2) to act as a base in the exocyclic mechanism, such configurations would not allow the reaction to evolve. Figure 3B shows a scatter plot of $d(O_w-N_{Asn92})$ versus $d(O_w-C1_{glc})$, where N_{Asn92} is the Asn92 nitrogen atom, showing that when the nucleophilic water approaches $C1_{glc}$, it moves away from

Asn92, and when the nucleophilic water is close to Asn92, it is still far from $C1_{glc}$, therefore never reaching the geometric conditions for the reaction. TPS-generated unbiased reactive trajectories of the classic mechanism with the Asn92 in imidic acid tautomer (Movie S2) also show that the nucleophilic water remains intact and does not get deprotonated, indicating that deprotonation by Asn92 does not occur (Movie S2 and Figure S3B). Hence, the Asn92 imidic acid tautomer seems to be in an unfavored position to act as base, just like Asn92 in the amide form.

In addition, DFT calculations performed with a reduced PcCel45A model yield a reaction energy of ~10.4 kcal/mol for the amide to imidic acid tautomerization of Asn92 in PcCel45A, indicating an enthalpic penalty that would contribute significantly to the barrier of the catalyzed hydrolysis reaction if Asn92 operates in its imidic acid form. Together, these results suggest that PcCel45A lacks an adequate residue to act directly as catalytic base within the viewpoint of the classical exocyclic inverting mechanism.

An interesting alternative to this apparent conundrum is to posit that one of the residues would act as base via a Grotthuss-like mechanism,[62] where a proton transfer between the general base residue and the nucleophilic water would take place through additional active site water molecules. An efficient Grotthuss proton-transfer pathway[62] requires the formation of a linear chain of water molecules or a hydrophobic cavity with transient water molecules.[63,64] This mechanism was proposed for an inverting cellulase from family GH6,[47] whose catalytic binding site is tunnel-shaped[65] (Figure 4A). However, unlike GH6 enzymes, PcCel45A possesses a wide and open substrate-binding groove (Figure 4B) with the catalytic site exposed to solvent. Therefore, the potential nucleophilic water molecule is part of a hydrogen bond network connected to the bulk, which would render the proton transfer to the protein unlikely via the Grotthuss mechanism.

3.3. Endocyclic Hydrolysis Mechanism. These results and considerations prompted us to seek alternative mechanistic possibilities for PcCel45A. Our free energy surface along the puckering coordinates (Figure 1B) indicates that the enzyme restrains the glycosidic ring around the $^4C_1$ chair conformation and there is compelling evidence of endocyclic cleavage in solution of β-glycosides conformationally restrained in $^4C_1$.[14,15] We then propose an endocyclic mechanism for hydrolysis catalyzed by PcCel45A, as shown in Scheme 2. According to this proposed mechanism, the reaction starts with the glucopyranose ring at the −1 subsite in the most stable $^4C_1$ chair conformation while a nucleophilic water attacks the
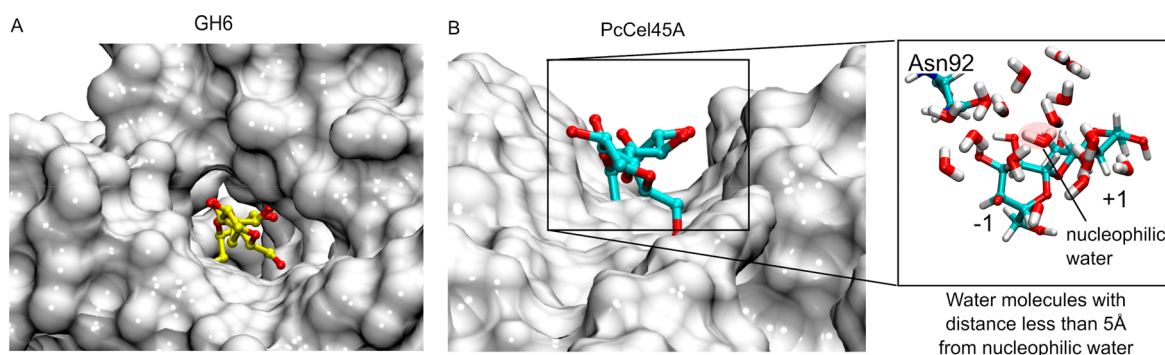
Figure 4. Molecular surface showing the active site architecture of (A) GH6 (a tunnel) and (B) PcCel45A (an open groove) and a representative snapshot showing the potential nucleophilic water surrounded by many water molecules connected to the bulk.

Scheme 2. Proposed Endocyclic Mechanism[a]
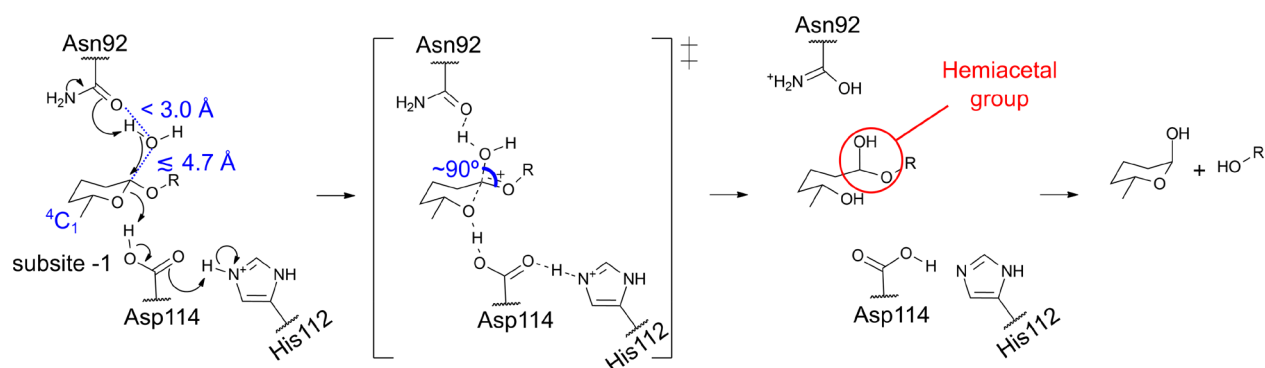
[a]Geometric parameters (distances and angle) and substrate conformation ($^4C_1$) required for the endocyclic mechanism are shown in blue.
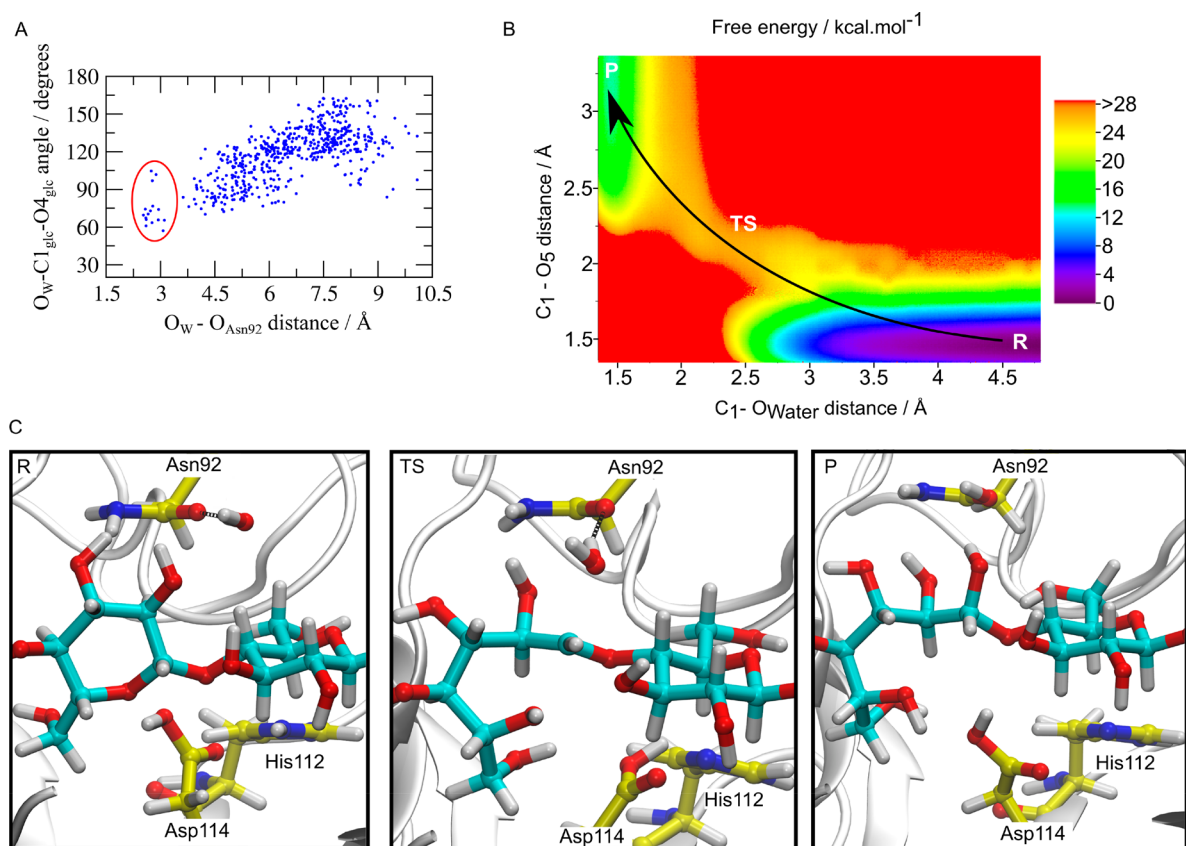
Figure 5. (A) $O_w-C1_{glc}-O4_{glc}$ angle versus O_w-O_{Asn92} distance. The red oval indicates structures visited in the simulation which satisfy the conditions for a single-displacement nucleophilic reaction. (B) Free energy surface of the reaction. Reactants (R) and products (P) are connected by a curved arrow passing through the transition state (TS). (C) Details of the ring-opening step showing the reactants (R), transition state (TS), and products (P).

anomeric carbon to break its bond with the endocyclic oxygen, instead of the exocyclic oxygen, and the nucleophilic water oxygen forms an angle of 90° with the $C1_{glc}$ and $O4_{glc}$ in the transition state.

To see if the system visits conformations that allow the endocyclic mechanism, we analyzed an ensemble of structures taken from the umbrella sampling QM/MM simulations with the glycosidic ring in the stable $^4C_1$ conformation and with a potential nucleophilic water molecule (O_w within 4.7 Å of the anomeric $C1_{glc}$). We then measured the $O_w-C1_{glc}-O4_{glc}$ angle and the distance between the O_w and O_{Asn92} atoms (Figure 5A). The results show that, in fact, there is a set of

configurations in which water molecules simultaneously satisfy conditions 1 and 2 for the single-displacement nucleophilic attack, and therefore, we posit that residue Asn92 is in position to act as a base in the proposed endocyclic cleavage. This is further corroborated by the unbiased TPS simulations of the endocyclic mechanism (Movie S3), where we observe that the nucleophilic water readily gets deprotonated by Asn92 as the system evolves to the product state (Movie S3 and Figure S3C), in contrast to the reactive trajectories of the exocyclic mechanism (Movies S1 and S2 and Figure S2A,B).

The reaction proceeds with (i) a proton transfer from the catalytic acid (Asp114) to the endocyclic oxygen, followed by

the glucopyranose ring opening, (ii) proton abstraction from a water molecule by the catalytic base (Asn92), and (iii) a nucleophilic attack on the anomeric carbon by the activated water (Scheme 2). In order to obtain the two-dimensional free energy landscape for this reaction (Figure 5B), we considered a two-dimensional reaction coordinate, defined as RC1 = $d(C1_{glc}-O_w)$ and RC2 = $d(C1_{glc}-O5_{glc})$. RC1 represents the nucleophilic attack by water, and RC2 represents the glucopyranose ring opening.

As the reaction evolves by sampling RC1 and RC2 in the simulations (Figure 5C; coordinates of the stationary points are available in the Supporting Information), spontaneous proton transfers occur from Asp114 to the endocyclic $O5_{glc}$ atom and from the nucleophilic water to the Asn92 carbonyl oxygen. While residue Asp114 transfers its proton to $O5_{glc}$, His112 (initially protonated) concertedly transfers its proton to Asp114, resulting in a protonated Asp114 and a deprotonated His112 (Scheme 2 and Figure 5C). Thus, His112 plays an assisting role during the protonation of the endocyclic $O5_{glc}$ oxygen. We obtained an activation free energy barrier of $\Delta G^{\ddagger} \sim 25$ kcal/mol and a reaction free energy of $\Delta G \sim 13$ kcal/mol (Figure 5B). This step exhibits a reaction barrier of magnitude comparable to (and even lower than) the steps of exocyclic pathways of GH12 (22 kcal/mol),[66] GH38 (23 kcal/mol),[67] GH43 (24 kcal/mol),[68] GH8 (36 kcal/mol),[69] and GH16 (32 kcal/mol).[70] Recently, Bharadwaj et al.[49] reported a similar activation barrier of 23.6 kcal/mol and reaction free energy of approximately 4 kcal/mol for a GH45 subfamily A inverting endoglucanase from *Humicola insolens* (HiCel45A) using unbiased QM/MM simulations based on the transition path sampling techniques.

The reaction product is a glucan chain, noncovalently bound to the enzyme, with one of its rings open and comprising a hemiacetal group. Because acyclic hemiacetals are unstable and its hydrolysis mechanism is well-established, we computed the free energy surface only for the key step of ring opening and hemiacetal formation.[71,72] The hemiacetal should readily decompose into the main product through an acid-catalyzed mechanism (Figure 5C, Figure S4), because the PcCel45A optimum pH is 3.[72,73] This unstable hemiacetal is probably the main contributor for the positive reaction free energy calculated. The species at completion of this reaction step also consist of PcCel45A with Asn92 protonated at its carbonyl oxygen and a neutral His112.

These results are highly suggestive that PcCel45A catalyzes cellulose hydrolysis via the endocyclic route and differs from the other GHs whose reaction mechanisms were previously studied. However, we point out that our results were obtained from a modeled Michaelis complex, as the glycosidic bond between the residues at −1 and +1 subsites is lacking in the PcCel45A crystal structures currently available.[23,74] Even though we have extensively sampled the substrate conformations at the −1 subsite to obtain this model, its reliability depends on assuming that the starting point of our simulation (the available crystal structure) and the true Michaelis complex (unknown so far) are not significantly different. Otherwise, much longer simulation times would be required to sample large protein motions that could impact the Michaelis complex conformation. Despite our assumption being plausible—since our previous classical MD simulations show that the substrate-bound PcCel45A is very stable during 150 ns runs[23]—we cannot fully rule out the possibility that the true PcCel45A Michaelis complex differs significantly from the available

crystal structures and, therefore, from our model, which could impact the free energy surfaces along the glycosidic ring conformations and along the reaction.

Hence, we do not discard the possibility that, within time scales much longer than our simulations', PcCel45A would visit conformations that stabilize the glycosidic ring at the $^2S_0$ conformation and enable a residue (e.g., Asn92 or Asp85) to act as a base in the classical inverting mechanism (Scheme 1) as in other inverting GHs.[3,9] Another possibility is that, upon substrate binding, PcCel45A undergoes a change in its binding site architecture that buries its catalytic site, thus creating a suitable environment for a Grotthuss-like mechanism, as previously observed for a GH6 cellulase.[47,64] Together, these considerations point to further experimental investigations aimed at corroborating our proposed mechanism.

3.4. Increased Nucleophilicity of Asn92. In the proposed endocyclic mechanism, Asn92 acts as a base in the amide form, not in imidic acid form. Although amides are notoriously weak bases in solution, this may not necessarily be the case in the enzyme environment. We observe that Asn92 is part of a hydrogen bond network involving two other amides: the Gln39 side chain and the Val89 backbone. Amide hydrogen bonding networks are known to resonate to the enol-like structure and redistribute charges (Figure 6A). This results in an increased nucleophilicity due to increased negative charge on the oxygen atom.[75]
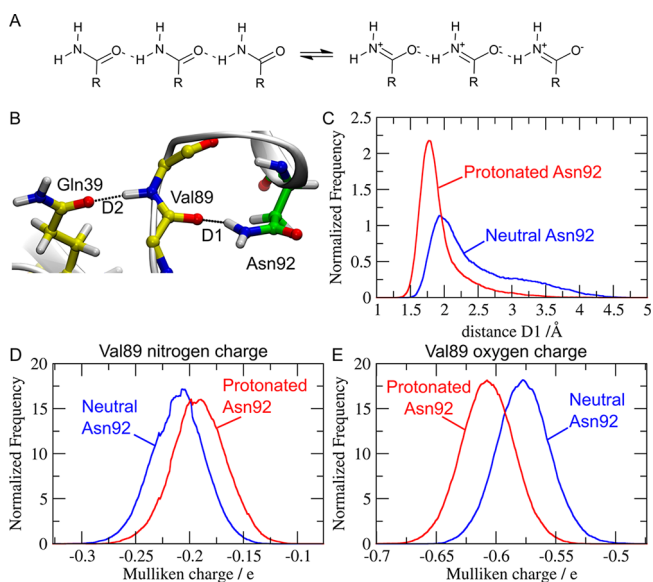
Figure 6. (A) Resonance forms of amides in hydrogen bonding networks. (B) Gln39−Val89−Asn92 hydrogen bonding network. (C) Distribution of the distance (D1) between Asn92 hydrogen and Val89 oxygen. Distribution of Val89 (D) nitrogen and (E) oxygen Mulliken charges, considering both reactant (neutral Asn92) and product (protonated Asn92) states.

We performed conventional unrestrained QM/MM simulations considering the Gln39 side chain and the Val89 backbone in the QM region with (1) Asn92 neutral and (2) Asn92 protonated. Figure 6B illustrates the Asn92−Val89−Gln39 hydrogen bond network in PcCel45A. Figure 6C shows that the Val89−Asn92 hydrogen bond is tighter when Asn92 is protonated than when Asn92 is neutral, suggesting that this hydrogen bond network in PcCel45A contributes to stabilizing the protonated state of the general base. This is related to

changes in the Val89 charges as the reaction proceeds: the Val89 nitrogen atom becomes less negative (Figure 6D) while the Val89 oxygen atom becomes more negative (Figure 6E). We observed only minor changes in the distance between Gln39 oxygen and Val89 backbone hydrogen (Figure S5A). Similarly, Gln39 nitrogen and oxygen atomic charges remained unaltered for protonated and neutral Asn92 (Figure S5B,C). These results show that the structural arrangement of nearby amides in PcCel45A contributes to the Asn92 role as a general base in its standard amide form. It seems, though, that the presence of the weak base Asn92 may be an evolutionary adaptation to the unusually low optimum pH of PcCel45A (around 3).[30] In such a low pH, the Asp or Glu residues would be protonated and unable to act as a general base, which requires unprotonated carboxyl groups.

4. CONCLUDING REMARKS

In this work, we present results of extensive umbrella sampling and transition path sampling QM/MM simulations of the cellulose hydrolysis mechanism catalyzed by PcCel45A, an inverting endoglucanase belonging to the GH45 subfamily C. The catalytic reaction mechanism adopted by this enzyme remains elusive because its catalytic machinery differs from that exhibited by other inverting GHs in important ways. Here, we propose an alternative pathway for the PcCel45A-catalyzed hydrolysis reaction, which involves the protonation of the endocyclic oxygen followed by the opening of the glucopyranose ring that leads to an unstable acyclic hemiacetal, known to readily decompose into hydrolysis products. We provide *in silico* evidence that Asn92 acts as a general base in its amide form, and we estimate an activation free energy barrier of ~25 kcal/mol along the proposed reaction pathway. The estimated barrier lies close to other computational estimates for various GHs from different groups using similar and also different methods. In particular, a very recent investigation on the hydrolysis mechanism of a GH45 subfamily A cellulase reported a similar barrier (23.55 ± 0.3 kcal/mol) using extensive unbiased transition path sampling QM/MM simulations.[49]

Evidently, an experimental investigation of the proposed mechanism would be desirable. One promising technique relies on devising and synthesizing surrogate ligand mimetics that act as mechanism-based inhibitors.[76,77] We conjecture that this mechanism may also play a role in other carbohydrate-active enzymes, such as lytic transglycosylases, expansins, loosenins, and swollenins, which are evolutionarily close to PcCel45A.[22,23,49] We expect our findings motivate further experimental and theoretical studies aimed at probing the inverting endocyclic mechanism in PcCel45A and correlated enzymes.

ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jcim.0c01380.

> Atoms included in QM region for QM/MM calculations and residues around the QM region; superposition and RMSD of residues Asn92 and Asp85 along 200 ns of MD simulation; distances between nucleophilic water oxygen and its bound hydrogen atoms vs distance between nucleophilic water oxygen and anomeric carbon in TPS simulations; mechanism of acid-catalyzed

> hemiacetal hydrolysis; distances between Gln39 oxygen and Val89 backbone hydrogen, and Mulliken charges of Gln39 nitrogen and oxygen (PDF)

> Fully unbiased reactive trajectory of the classic exocyclic mechanism (MP4)

> Fully unbiased reactive trajectory of the classic mechanism with Asn92 in imidic acid tautomer (MP4)

> Fully unbiased reactive trajectory of the endocyclic mechanism (MP4)

> Spatial coordinates of representative configurations of the main minima of the conformational maps and reactant, transition state and product of ring-opening step (ZIP)

AUTHOR INFORMATION

Corresponding Author

Munir S. Skaf − *Institute of Chemistry and Center for Computing in Engineering and Sciences, University of Campinas—Unicamp, Campinas 13084-862, Sao Paulo, Brazil;* ⓞ orcid.org/0000-0001-7485-1228; Email: skaf@unicamp.br

Authors

Caroline S. Pereira − *Institute of Chemistry and Center for Computing in Engineering and Sciences, University of Campinas—Unicamp, Campinas 13084-862, Sao Paulo, Brazil*

Rodrigo L. Silveira − *Institute of Chemistry and Center for Computing in Engineering and Sciences, University of Campinas—Unicamp, Campinas 13084-862, Sao Paulo, Brazil; Institute of Chemistry, Federal University of Rio de Janeiro, Rio de Janeiro 21941-909, Rio de Janeiro, Brazil*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jcim.0c01380

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Chundawat, S. P. S.; Beckham, G. T.; Himmel, M. E.; Dale, B. E. Deconstruction of Lignocellulosic Biomass to Fuels and Chemicals. *Annu. Rev. Chem. Biomol. Eng.* **2011**, *2*, 121−145.

(2) Henrissat, B. A Classification of Glycosyl Hydrolases Based on Amino Acid Sequence Similarities. *Biochem. J.* **1991**, *280* (2), 309−316.

(3) Payne, C. M.; Knott, B. C.; Mayes, H. B.; Hansson, H.; Himmel, M. E.; Sandgren, M.; Stahlberg, J.; Beckham, G. T. Fungal Cellulases. *Chem. Rev.* **2015**, *115* (3), 1308−1448.

(4) Jongkees, S. A. K.; Withers, S. G. Unusual Enzymatic Glycoside Cleavage Mechanisms. *Acc. Chem. Res.* **2014**, *47* (1), 226−235.

(5) Davies, G.; Henrissat, B. Structures and Mechanisms of Glycosyl Hydrolases. *Structure* **1995**, *3* (9), 853−859.

(6) Deslongchamps, P.; Li, S.; Dory, Y. L. Hydrolysis of α-and β-Glycosides. New Experimental Data and Modeling of Reaction Pathways. *Org. Lett.* **2004**, *6* (4), 505−508.

(7) Liras, J. L.; Anslyn, E. V. Exocyclic and Endocyclic Cleavage of Pyranosides in Both Methanol and Water Detected by a Novel Probe. *J. Am. Chem. Soc.* **1994**, *116* (6), 2645−2646.

(8) Deslongchamps, P. *Stereoelectronic Effects in Organic Chemistry*; Pergamon: 1983; Vol. 1.

(9) Ardèvol, A.; Rovira, C. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from Ab Initio Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* **2015**, *137* (24), 7528−7547.

(10) Loerbroks, C.; Rinaldi, R.; Thiel, W. The Electronic Nature of the 1, 4-β-Glycosidic Bond and Its Chemical Environment: DFT Insights into Cellulose Chemistry. *Chem. - Eur. J.* **2013**, *19* (48), 16282−16294.

(11) Post, C. B.; Karplus, M. Does Lysozyme Follow the Lysozyme Pathway? An Alternative Based on Dynamic, Structural, and Stereoelectronic Considerations. *J. Am. Chem. Soc.* **1986**, *108* (6), 1317−1319.

(12) Bowman, A. L.; Grant, I. M.; Mulholland, A. J. QM/MM Simulations Predict a Covalent Intermediate in the Hen Egg White Lysozyme Reaction with Its Natural Substrate. *Chem. Commun.* **2008**, No. 37, 4425−4427.

(13) Gupta, R. B.; Franck, R. W. Direct Experimental Evidence for Cleavage of Both Exo-and Endo-Cyclic Carbon-Oxygen Bonds in the Acid-Catalyzed Reaction of Alkyl. Beta.-Tetrahydropyranyl Acetals. *J. Am. Chem. Soc.* **1987**, *109* (21), 6554−6556.

(14) Satoh, H.; Manabe, S.; Ito, Y.; Lüthi, H. P.; Laino, T.; Hutter, J. Endocyclic Cleavage in Glycosides with 2, 3-Trans Cyclic Protecting Groups. *J. Am. Chem. Soc.* **2011**, *133* (14), 5610−5619.

(15) Manabe, S.; Ishii, K.; Hashizume, D.; Koshino, H.; Ito, Y. Evidence for Endocyclic Cleavage of Conformationally Restricted Glycopyranosides. *Chem. - Eur. J.* **2009**, *15* (28), 6894−6901.

(16) Manabe, S.; Satoh, H.; Hutter, J.; Lüthi, H. P.; Laino, T.; Ito, Y. Significant Substituent Effect on the Anomerization of Pyranosides: Mechanism of Anomerization and Synthesis of a 1, 2-cis Glucosamine Oligomer from the 1, 2-trans Anomer. *Chem. - Eur. J.* **2014**, *20* (1), 124−132.

(17) Crich, D.; Vinod, A. U. 6-O-Benzyl-and 6-O-Silyl-N-Acetyl-2-Amino-2-N, 3-O-Carbonyl-2-Deoxyglucosides: Effective Glycosyl Acceptors in the Glucosamine 4-OH Series. Effect of Anomeric Stereochemistry on the Removal of the Oxazolidinone Group. *J. Org. Chem.* **2005**, *70* (4), 1291−1296.

(18) Geng, Y.; Qin, Q.; Ye, X.-S. Lewis Acids as α-Directing Additives in Glycosylations by Using 2, 3-O-Carbonate-Protected Glucose and Galactose Thioglycoside Donors Based on Preactivation Protocol. *J. Org. Chem.* **2012**, *77* (12), 5255−5270.

(19) Pilgrim, W.; Murphy, P. V. SnCl4-and TiCl4-Catalyzed Anomerization of Acylated O-and S-Glycosides: Analysis of Factors That Lead to Higher α: β Anomer Ratios and Reaction Rates. *J. Org. Chem.* **2010**, *75* (20), 6747−6755.

(20) Nakamura, Y.; Moriya, T.; Baba, Y.; Yanai, K.; Sumida, N.; Nishimura, T.; Murashima, K.; Nakane, A.; Yaguchi, T.; Koga, J. Endoglucanases and Cellulase Preparations Containing the Same. US6921655B1, July 26, 2005.

(21) Schülein, M.; Kauppinen, M. S.; Lange, L.; Lassen, S. F.; Andersen, L. N.; Klysner, S.; Nielsen, J. B. Characterization of Fungal Cellulases for Fiber Modification. *Enzyme Applications in Fiber Processing*; Eriksson, K.-E. L., Cavaco-Paulo, A., Eds.; ACS Symposium Series 687; American Chemical Society: 1998; pp 66−74..

(22) Cosgrove, D. J. Plant Expansins: Diversity and Interactions with Plant Cell Walls. *Curr. Opin. Plant Biol.* **2015**, *25*, 162−172.

(23) Godoy, A. S.; Pereira, C. S.; Ramia, M. P.; Silveira, R. L.; Camilo, C. M.; Kadowaki, M. A.; Lange, L.; Busk, P. K.; Nascimento, A. S.; Skaf, M. S.; Polikarpov, I. Structure, Computational and Biochemical Analysis of PcCel45A Endoglucanase from Phanerochaete Chrysosporium and Catalytic Mechanisms of GH45 Subfamily C Members. *Sci. Rep.* **2018**, *8*, 3678.

(24) Igarashi, K.; Ishida, T.; Hori, C.; Samejima, M. Characterization of an Endoglucanase Belonging to a New Subfamily of Glycoside Hydrolase Family 45 of the Basidiomycete Phanerochaete Chrysosporium. *Appl. Environ. Microbiol.* **2008**, *74* (18), 5628−5634.

(25) Yennawar, N. H.; Li, L.-C.; Dudzinski, D. M.; Tabuchi, A.; Cosgrove, D. J. Crystal Structure and Activities of EXPB1 (Zea m 1), a β-Expansin and Group-1 Pollen Allergen from Maize. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (40), 14664−14671.

(26) Tabuchi, A.; Li, L.; Cosgrove, D. J. Matrix Solubilization and Cell Wall Weakening by B-expansin (Group-1 Allergen) from Maize Pollen. *Plant J.* **2011**, *68* (3), 546−559.

(27) Cosgrove, D. J. Loosening of Plant Cell Walls by Expansins. *Nature* **2000**, *407* (6802), 321−326.

(28) van Straaten, K. E.; Dijkstra, B. W.; Vollmer, W.; Thunnissen, A.-M. W. H. Crystal Structure of MltA from Escherichia Coli Reveals a Unique Lytic Transglycosylase Fold. *J. Mol. Biol.* **2005**, *352* (5), 1068−1080.

(29) van Straaten, K. E.; Barends, T. R. M.; Dijkstra, B. W.; Thunnissen, A.-M. W. H. Structure of Escherichia Coli Lytic Transglycosylase MltA with Bound Chitohexaose IMPLICATIONS FOR PEPTIDOGLYCAN BINDING AND CLEAVAGE. *J. Biol. Chem.* **2007**, *282* (29), 21197−21205.

(30) Nakamura, A.; Ishida, T.; Kusaka, K.; Yamada, T.; Fushinobu, S.; Tanaka, I.; Kaneko, S.; Ohta, K.; Tanaka, H.; Inaka, K.; et al. "Newton's Cradle" Proton Relay with Amide−Imidic Acid Tautomerization in Inverting Cellulase Visualized by Neutron Crystallography. *Sci. Adv.* **2015**, *1* (7), e1500263.

(31) Davies, G. J.; Tolley, S. P.; Henrissat, B.; Hjort, C.; Schulein, M. Structures of Oligosaccharide-Bound Forms of the Endoglucanase V from Humicola Insolens at 1.9. ANG. Resolution. *Biochemistry* **1995**, *34* (49), 16210−16220.

(32) McCarter, J. D.; Stephen Withers, G. Mechanisms of Enzymatic Glycoside Hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4* (6), 885−892.

(33) Zechel, D. L.; Withers, S. G. Glycosidase Mechanisms: Anatomy of a Finely Tuned Catalyst. *Acc. Chem. Res.* **2000**, *33* (1), 11−18.

(34) Constantino, E.; Solans-Monfort, X.; Sodupe, M.; Bertran, J. Basic and Acidic Bifunctional Catalysis: Application to the Tautomeric Equilibrium of Formamide. *Chem. Phys.* **2003**, *295* (2), 151−158.

(35) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. H++: A Server for Estimating pKas and Adding Missing Hydrogens to Macromolecules. *Nucleic Acids Res.* **2005**, *33*, W368−W371.

(36) Myers, J.; Grothaus, G.; Narayanan, S.; Onufriev, A. A Simple Clustering Algorithm Can Be Accurate Enough for Use in Calculations of PKs in Macromolecules. *Proteins: Struct., Funct., Genet.* **2006**, *63* (4), 928−938.

(37) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N·Log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98* (12), 10089−10092.

(38) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J. Comput. Phys.* **1977**, *23* (3), 327−341.

(39) Case, D. A.; Betz, R.; Botello-Smith, W.; Cerutti, D. S.; Cheatham, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S. *Amber 2016*; University of California San Francisco: 2016.

(40) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99SB. *J. Chem. Theory Comput.* **2015**, *11* (8), 3696−3713.

(41) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates. *J. Comput. Chem.* **2008**, *29* (4), 622−655.

(42) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926−935.

(43) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-

Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1998**, *58* (11), 7260−7268.

(44) Seabra, G.; Walker, R. C.; Elstner, M.; Case, D. A.; Roitberg, A. E. Implementation of the SCC-DFTB Method for Hybrid QM/MM Simulations within the Amber Molecular Dynamics Package. *J. Phys. Chem. A* **2007**, *111* (26), 5655−5664.

(45) Walker, R. C.; Crowley, M. F.; Case, D. A. The Implementation of a Fast and Accurate QM/MM Potential Method in Amber. *J. Comput. Chem.* **2008**, *29* (7), 1019−1031.

(46) Case, D. A.; Betz, R. M.; Cerutti, D. S.; Cheatham, T. E., III; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N. *AMBER 2016 Reference Manual*; University of California San Francisco: 2016; pp 1−923.

(47) Mayes, H. B.; Knott, B. C.; Crowley, M. F.; Broadbelt, L. J.; Ståhlberg, J.; Beckham, G. T. Who's on Base? Revealing the Catalytic Mechanism of Inverting Family 6 Glycoside Hydrolases. *Chem. Sci.* **2016**, *7* (9), 5955−5968.

(48) Knott, B. C.; Haddad Momeni, M.; Crowley, M. F.; Mackenzie, L. F.; Götz, A. W.; Sandgren, M.; Withers, S. G.; Ståhlberg, J.; Beckham, G. T. The Mechanism of Cellulose Hydrolysis by a Two-Step, Retaining Cellobiohydrolase Elucidated by Structural and Transition Path Sampling Studies. *J. Am. Chem. Soc.* **2014**, *136* (1), 321−329.

(49) Bharadwaj, V. S.; Knott, B. C.; Ståhlberg, J.; Beckham, G. T.; Crowley, M. F. The Hydrolysis Mechanism of a GH45 Cellulase and Its Potential Relation to Lytic Transglycosylase and Expansin Function. *J. Biol. Chem.* **2020**, *295* (14), 4477−4487.

(50) Cremer, D. t; Pople, J. A. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97* (6), 1354−1358.

(51) Biarnés, X.; Ardevol, A.; Planas, A.; Rovira, C.; Laio, A.; Parrinello, M. The Conformational Free Energy Landscape of β-D-Glucopyranose. Implications for Substrate Preactivation in β-Glucoside Hydrolases. *J. Am. Chem. Soc.* **2007**, *129* (35), 10686−10693.

(52) Bonomi, M.; Branduardi, D.; Bussi, G.; Camilloni, C.; Provasi, D.; Raiteri, P.; Donadio, D.; Marinelli, F.; Pietrucci, F.; Broglia, R. A.; et al. PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput. Phys. Commun.* **2009**, *180* (10), 1961−1972.

(53) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. PLUMED 2: New Feathers for an Old Bird. *Comput. Phys. Commun.* **2014**, *185* (2), 604−613.

(54) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13* (8), 1011−1021.

(55) Peters, B.; Trout, B. L. Obtaining Reaction Coordinates by Likelihood Maximization. *J. Chem. Phys.* **2006**, *125* (5), 054108.

(56) Peters, B.; Beckham, G. T.; Trout, B. L. Extensions to the Likelihood Maximization Approach for Finding Reaction Coordinates. *J. Chem. Phys.* **2007**, *127* (3), 034109.

(57) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53* (1), 291−318.

(58) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1−3), 215−241.

(59) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113* (18), 6378−6396.

(60) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone,

V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Pople, J. A. *Gaussian 09*, rev. C.01; Gaussian Inc.: 1995.

(61) Guérin, D. M. A.; Lascombe, M.-B.; Costabel, M.; Souchon, H.; Lamzin, V.; Béguin, P.; Alzari, P. M. Atomic (0.94 Å) Resolution Structure of an Inverting Glycosidase in Complex with Substrate. *J. Mol. Biol.* **2002**, *316* (5), 1061−1069.

(62) von Grotthuß, T. *Mémoire Sur La Décomposition de l'eau et Des Corps Qu'elle Tient En Dissolution à l'aide de l'électricité Galvanique*; 1805.

(63) Taraphder, S.; Hummer, G. Protein Side-Chain Motion and Hydration in Proton-Transfer Pathways. Results for Cytochrome P450cam. *J. Am. Chem. Soc.* **2003**, *125* (13), 3931−3940.

(64) Cukierman, S. Et Tu, Grotthuss! And Other Unfinished Stories. *Biochim. Biophys. Acta, Bioenerg.* **2006**, *1757* (8), 876−885.

(65) Zou, J.; Kleywegt, G. J.; Ståhlberg, J.; Driguez, H.; Nerinckx, W.; Claeyssens, M.; Koivula, A.; Teeri, T. T.; Jones, T. A. Crystallographic Evidence for Substrate Ring Distortion and Protein Conformational Changes during Catalysis in Cellobiohydrolase Ce16A from Trichoderma Reesei. *Structure* **1999**, *7* (9), 1035−1045.

(66) Lian, P.; Guo, H.-B.; Smith, J. C.; Wei, D.-Q.; Guo, H. Catalytic Mechanism and Origin of High Activity of Cellulase TmCel12A at High Temperature: A Quantum Mechanical/Molecular Mechanical Study. *Cellulose* **2014**, *21* (2), 937−949.

(67) Petersen, L.; Ardèvol, A.; Rovira, C.; Reilly, P. J. Molecular Mechanism of the Glycosylation Step Catalyzed by Golgi α-Mannosidase II: A QM/MM Metadynamics Investigation. *J. Am. Chem. Soc.* **2010**, *132* (24), 8291−8300.

(68) Barker, I. J.; Petersen, L.; Reilly, P. J. Mechanism of Xylobiose Hydrolysis by GH43 β-Xylosidase. *J. Phys. Chem. B* **2010**, *114* (46), 15389−15393.

(69) Petersen, L.; Ardèvol, A.; Rovira, C.; Reilly, P. J. Mechanism of Cellulose Hydrolysis by Inverting GH8 Endoglucanases: A QM/MM Metadynamics Study. *J. Phys. Chem. B* **2009**, *113* (20), 7331−7339.

(70) Biarnés, X.; Ardevol, A.; Iglesias-Fernández, J.; Planas, A.; Rovira, C. Catalytic Itinerary in 1, 3−1, 4-β-Glucanase Unraveled by QM/MM Metadynamics. Charge Is Not yet Fully Developed at the Oxocarbenium Ion-like Transition State. *J. Am. Chem. Soc.* **2011**, *133* (50), 20301−20309.

(71) Cordes, E. H.; Bull, H. G. Mechanism and Catalysis for Hydrolysis of Acetals, Ketals, and Ortho Esters. *Chem. Rev.* **1974**, *74* (5), 581−603.

(72) Clayden, J.; Greeves, N.; Warren, S. *Organic Chemistry*; OUP: Oxford, U.K.: 2012.

(73) Sorensen, P. E.; Jencks, W. P. Acid-and Base-Catalyzed Decomposition of Acetaldehyde Hydrate and Hemiacetals in Aqueous Solution. *J. Am. Chem. Soc.* **1987**, *109* (15), 4675−4690.

(74) Nakamura, A.; Ishida, T.; Kusaka, K.; Yamada, T.; Fushinobu, S.; Tanaka, I.; Kaneko, S.; Ohta, K.; Tanaka, H.; Inaka, K.; Higuchi, Y.; Niimura, N.; Samejima, M.; Igarashi, K. Newton's Cradle" Proton Relay with Amide−Imidic Acid Tautomerization in Inverting Cellulase Visualized by Neutron Crystallography. *Sci. Adv.* **2015**, *1* (7), e1500263.

(75) Zhou, Y.; Deng, G.; Zheng, Y.-Z.; Xu, J.; Ashraf, H.; Yu, Z.-W. Evidences for Cooperative Resonance-Assisted Hydrogen Bonds in Protein Secondary Structure Analogs. *Sci. Rep.* **2016**, *6* (1), 36932.

(76) Fenger, T. H.; Brumer, H. Synthesis and Analysis of Specific Covalent Inhibitors of Endo-Xyloglucanases. *ChemBioChem* **2015**, *16* (4), 575−583.

(77) Jain, N.; Attia, M. A.; Offen, W. A.; Davies, G. J.; Brumer, H. Synthesis and Application of a Highly Branched, Mechanism-Based 2-Deoxy-2-Fluoro-Oligosaccharide Inhibitor of Endo-Xyloglucanases. *Org. Biomol. Chem.* **2018**, *16* (45), 8732−8741.