Direct determination of protonation states and visualization of hydrogen bonding in a glycoside hydrolase with neutron crystallography

Classification: BIOLOGICAL SCIENCES, Biophysics and Computational Biology.

Qun Wan,[1] Jerry M. Parks,[2] B. Leif Hanson,[3] Suzanne Zoe Fisher,[4] Andreas Ostermann,[5] Tobias E. Schrader,[6] David E. Graham,[7] Leighton Coates,[8] Paul Langan,[8] and Andrey Kovalevsky[8*]

[1]Department of Physics, College of Science, Nanjing Agricultural University, Nanjing 210095, P. R. China.

[2]UT/ORNL Center for Molecular Biophysics, Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

[3]Chemistry Department, University of Toledo, Toledo, OH 43606, USA.

[4]Scientific Activities Division, European Spallation Source, PO Box 176, Lund 22100, Sweden

[5] Heinz Maier-Leibnitz Zentrum (MLZ), Technische Universität München, 85748 Garching, Germany.

[6]Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, 85747 Garching, Germany.

[7]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

[8]Biology and Soft Matter Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA.

* To whom correspondence should be addressed: kovalevskyay@ornl.gov (AK)

This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

Abstract

Glycoside hydrolase (GH) enzymes apply acid/base chemistry to catalyze the decomposition of complex carbohydrates. These ubiquitous enzymes accept protons from solvent and donate them to substrates at close to neutral pH by modulating the pK_a values of key side chains during catalysis. However, it is not known how the catalytic acid residue acquires a proton and transfers it efficiently to the substrate. To better understand GH chemistry, we used macromolecular neutron crystallography to directly determine protonation and ionization states of the active site residues of a family 11 GH at multiple pD values. The general acid glutamate (Glu) cycles between two conformations, "upward" and "downward", but is protonated only in the "downward" orientation. We performed continuum electrostatics calculations to estimate the pK_a values of the catalytic Glu residues in both the apo- and substrate-bound states of the enzyme. The calculated pK_a of the Glu increases substantially when the side chain moves down. The energy barrier required to rotate the catalytic Glu residue back to the "upward" conformation, where it can protonate the glycosidic oxygen of the substrate, is 4.3 kcal/mol according to free energy simulations. These findings shed light on the initial stage of the glycoside hydrolysis reaction in which molecular motion enables the general acid catalyst to obtain a proton from the bulk solvent and deliver it to the glycosidic oxygen.

Keywords: glycoside hydrolase, protonation state, macromolecular neutron crystallography, xylanase

Significance statement

Most enzymatic reactions involve hydrogen or proton transfer among the enzyme, substrate, and water at physiological pH. Thus, enzyme catalysis cannot be fully understood without accurate mapping of hydrogen atom positions in these macromolecular catalysts. Direct information on the location of hydrogen atoms can be obtained using neutron crystallography. We used neutron crystallography and biomolecular simulation to characterize the initial stage of the glycoside hydrolysis reaction catalyzed by a family 11 glycoside hydrolase. We provide evidence that the catalytic glutamate residue alternates between two conformations bearing different basicities, first to obtain a proton from the bulk solvent, and then to deliver it to the glycosidic oxygen to initiate the hydrolysis reaction.

/body

INTRODUCTION

Glycoside hydrolases (GHs) catalyze the cleavage of the glycosidic bond, one of the most stable linkages in biological macromolecules, with rate enhancements on the order of 10^{17}-10^{18} relative to the uncatalyzed reaction in neutral solution (1), placing them among the most efficient biochemical catalysts known. GHs naturally decompose billions of tons of organic matter worldwide annually, and they produce commercially valuable chemicals from lignocellulosic plant biomass (2,3). They are also important targets for the development of therapeutic agents against bacterial and viral infections, diabetes, and other diseases (4,5). Understanding their catalytic mechanisms and ligand binding characteristics requires knowledge of the protonation states and hydrogen bonding at the atomic level, which is of paramount importance for protein engineering and drug design (6).

Glycosides are acetals that are spontaneously cleaved through a specific acid hydrolysis mechanism involving protonation of the glycosidic oxygen of the scissile C-O bond by H_3O^+ to generate an oxocarbenium cation intermediate, followed by the nucleophilic attack of a water molecule that leads to a racemic hemiacetal product (7). GH enzymes enhance the rate of this slow hydrolysis reaction using a general acid, usually the carboxylic acid of a Glu residue, which donates a proton to the leaving group oxygen in the rate-limiting step (8-10). Two models for the catalytic mechanism of β-glycoside hydrolases have been proposed: a) a double displacement pathway first proposed by Koshland, essentially proceeding as an S_N2 reaction (11,12), and b) an oxocarbenium ion intermediate pathway originally proposed by Phillips (13), occurring as an S_N1 reaction (Figure S1) (14,15). The Koshland mechanism suggests a glycosylation step, in which a covalent glycosyl-enzyme intermediate bearing inverted stereochemistry at the anomeric C1 is formed, followed by the deglycosylation step, in which the latter is hydrolyzed by an incoming H_2O activated by the first Glu residue and switched its role to act as a general base. This canonical double-displacement mechanism has been challenged by proponents of the oxocarbenium ion intermediate mechanism, in which the general acid-assisted departure of an aglycon leaving group leads to a short-lived, planar carbocation on C1 of the glycon, which is then hydrated by an incoming H_2O with an estimated rate constant of 10^{12} s^{-1} (16). Each

mechanism draws support from different types of experimental and computational evidence (17-22), with the double-displacement mechanism being generally accepted.

Two catalytic Glu residues, one serving as an acid/base catalyst and the other as a nucleophile, are nearly ubiquitous in GH enzymes (23). Such a scenario necessitates the former being a weaker acid than the latter, and implies that in the reactant state the general acid must be protonated and neutral (i.e., -COOH), with a pK_a greater than 6, and the nucleophile must be deprotonated and negatively charged (i.e., $-COO^-$), with a pK_a less than 5 (24). However, the protonation states of the key residues in the active site as well as the distribution of hydrogen bonds have not been directly observed in GH enzymes. Thus the changes in protonation predicted by the reaction mechanism have only been inferred from indirect measurements. H atoms are difficult to visualize by X-ray crystallography, even in structures determined at sub-Ångström resolution. Whereas X-rays are scattered from electron density, neutrons are scattered by atomic nuclei through the strong force which is independent of atomic number. The smallest atom, H, and its heavy isotope deuterium (D) scatter neutrons as well as other atoms constituting biological molecules. As a result, locating H and D atoms in macromolecular neutron structures is straightforward even at medium resolutions of 2.0-2.5 Å (25).

GH11 enzymes are small (~20 kDa) proteins called xylanases with a conserved jelly-roll fold that catalyze the hydrolysis of the most abundant component of hemicellulose, xylan, into smaller oligosaccharides (26-29). The active-site cavity of xylanase, XynII, from *Trichoderma reesei* adopts a half-pipe shape and is sculpted by an extensive β-sheet that twists on itself (Figure S2). On the basis of previous X-ray structures (30), the active site cleft has six subsites that bind the xylose units of xylan, spanning positions -3 to +3, with hydrolysis occurring between subsites -1 and +1 (31). Two catalytic Glu residues, Glu177 and Glu86, face each other from opposite sides of the cleft and their side chains are strategically positioned within about 6 Å from each other (32). Glu177 is believed to serve as the acid/base catalyst and must be protonated before the reaction, whereas Glu86 is the nucleophile and must be deprotonated.

To directly determine protonation states in a GH, and specifically to locate the H atom on carboxylic acid of the catalytic Glu, we obtained five neutron structures of wild-type and mutant (N44D) XynII at high resolutions of 1.7-2.0 Å in the apo- and 2-(*N*-morpholino)ethanesulfonic acid (MES)-bound forms. The N44D mutant displays altered pK_a values of key residues in the

active site. These structures were obtained from H/D exchanged crystals formed at pD values from 4.8 to 8.9, spanning the range of xylanase activity and the pK_a values of key residues (pD = pH + 0.4; see Supporting Text). To provide additional insight into active site protonation states, selected structures were further characterized with pK_a calculations and free energy simulations.

RESULTS

Apo-XynII at different pD values

Structures of the native XynII and N44D variant at pD values 6.2 and 6.4, respectively, which are close to the optimum pD for XynII activity, were expected to confirm protonation of Glu177 and the lack of proton on Glu86. Surprisingly, in the neutron structure of native XynII at pD 6.2, named **pD6.2** henceforth, the side chains of both Glu177 and Glu86 are deprotonated and therefore are negatively charged. Glu177 is in the "upward" conformation facing Glu86, with the closest distance between the side chains being 5.9 Å. The carboxylates of Glu177 and Glu86 accept D atoms from the phenolic oxygens of Tyr88 and Tyr77 in two hydrogen bonds with D…O distances of 1.9 Å and 1.7 Å, respectively (Figure 1a). Tyr88 is also hydrogen bonded to a water molecule, whereas Tyr77 makes a weak interaction with Trp79. Glu177 accepts two more hydrogen bonds with nearby water molecules, but is 4.4 Å away from the side-chain amide of Asn44. The Glu86 carboxylate in **pD6.2** forms a 1.9 Å hydrogen bond with the adjacent Gln136 side chain, and is also exposed to the bulk solvent.

Previous ^{13}C-NMR experiments using the homologous *Bacillus circulans* xylanase suggested that replacing an Asn residue with Asp near the acidic Glu residue increases the pK_a of Glu due to the possibility of formation of a short hydrogen bond (33). Therefore we expected that an equivalent N44D variant of the *T. reesei* enzyme would have a Glu177 pK_a above 8. However, we still did not observe a D atom on the carboxylate of Glu177. The active site of the N44D variant at pD 6.4 (**N44D**) is very similar to that in **pD6.2** (Figure 1b). In both structures, Asp44 remains over 4 Å away from the deprotonated Glu177 positioned in the "upward" conformation. The omit F_O-F_C neutron scattering length density maps clearly demonstrate that the D atoms of the Tyr88 and Tyr77 phenolic groups are covalently bound to the oxygen atoms,

indicating normal hydrogen bonds with Glu177 and Glu86 in both **pD6.2** and **N44D** (Figure 1a and 1b).

In the neutron structure at the basic pD of 8.9, **pD8.9**, the omit F_O-F_C nuclear density map showed that the D atom taking part in the hydrogen bond between Glu177 and Tyr88 is located halfway between the two oxygens, with D…O distances of 1.5 and 1.6 Å, respectively (Figure 1c). This observation indicates the possible presence of a low-barrier hydrogen bond (LBHB) induced by the increased pD. These changes in the hydrogen bonding at the active site might explain low catalytic activity of XynII in basic conditions.

Under acidic conditions, a significant change occurs in the XynII active site of the neutron structure at pD 4.8, **pD4.8** (Figure 1d). The Glu177 side chain rotates ~90° about the $\chi 2$ dihedral angle so that the carboxylate moves away from Glu86 to assume a "downward" conformation. The distance between the side chains of Glu177 and Glu86 increases to 9.3 Å. In this conformation the Glu177 side chain is protonated and positioned in a pocket lined by the backbone of Ser72 and Gly178, and by the side chain groups of Trp18, Asn44, Val46, Asn71, Tyr73, Tyr96, Ala175, and Tyr179, but makes no hydrogen bonding contacts with any of them. The COOH is hydrogen bonded at a distance of 1.5 Å to a D_2O molecule, which is quite mobile as indicated by its high B-factor of ~60 $Å^2$. In **pD4.8**, Asn44 is located closer to Glu177 than in **pD6.2**, with the side-chain amide N being 3.7 Å away from the carboxylate O, but this distance is still too long for a hydrogen bond. There is a concerted movement of the aromatic rings of Tyr73, Tyr77 and Tyr88 in the opposite direction of Glu177 rotation upon decreasing the pD (Figure 2). Surprisingly, the hydroxyl group of Tyr88 is oriented similarly in both the **pD4.8** and **pD6.2**, even though its phenol group moves closer to and is positioned within hydrogen bonding distance of the Glu86 carboxylate in **pD4.8** (O…O distance = 3.2 Å). The atomic positions are very similar in the low-temperature X-ray structures at low pH of the recombinant wild-type XynII at pH 4.0 and N44D variant at pH 4.6 (the RMSD of the main-chain atoms for the two structures is 0.1 Å). Interestingly, in the pH 4.0 X-ray structure a Na^+ cation was found instead of a water molecule observed in **pD4.8** between Asn71 and Glu177 in the "downward" conformation. This observation can be explained by the large amount of NaCl (~ 1M) used in the crystallization conditions for the recombinant enzyme. Na^+ is coordinated to the oxygen of the amide side chain of Asn71 and a water molecule and is 3.75 Å away from Glu177.

Complexes with MES

We were not able to grow large crystals of a XynII complex with a bound substrate (30); however, we obtained a neutron structure with MES to mimic proton donation by Glu177 to the glycosidic oxygen of the substrate in the first step of glycoside hydrolysis. A MES molecule was found in the active site of the native XynII neutron structure at pD 6.4, denoted **WT-MES**, and in the X-ray structure of N44E variant at high concentrations of MES buffer (>0.5 M). The active sites of **WT-MES** and N44E superimpose almost perfectly, with the ligand occupying almost identical positions. The only difference is in the orientation of the side chain of Glu44 relative to Asn44 in the two structures. Asn44 is oriented toward Glu177 in **WT-MES**. The longer side chain of Glu44 cannot be accommodated in the same orientation. Thus, it rotates about 90° around the $\chi 2$ dihedral angle to face the bulk solvent, making a short C-H…O contact of 3.1 Å with the morpholinic ring of MES (Figure S3). In **WT-MES**, MES forms three hydrogen bonds with the active-site residues. The morpholinic nitrogen is protonated and donates its D atom in a strong hydrogen bond with Oε2 of Glu177 (D…O distance = 1.7 Å, Figure S3). The sulfonate makes two hydrogen bonds with the side-chain groups of Tyr73 and Gln136 (D…O distances = 1.9 Å). Additionally, the CH_2 group vicinal to the nitrogen of the morpholinic ring makes a C-H…O contact with the Glu86 carboxylate. These interactions are similar to those made by xylohexaose substrate in E177Q-X6 (30). Oε1 of Glu177 is hydrogen bonded to Tyr88, while Glu86 forms two hydrogen bonds with Tyr77 and Gln136 in a similar fashion as in **pD6.2**. These two complexes are good proxies for the substrate-bound XynII protein, illustrating the "upward" conformation of the deprotonated Glu177 residue.

H/D exchange analysis

Figure S4 depicts the hydrogen/deuterium (H/D) exchange pattern for the main-chain amides in **pD6.2** (see Figure S5 for the analysis of the other neutron structures). Overall, xylanase exhibits moderate levels of H/D exchange, with about half of the residues showing full exchange and the rest being fully non-exchanged. Only a few residues are partially exchanged.

Notably, β strands B3-B6 close to the catalytic glutamates are almost fully non-exchanged. β sheets are rigid structures, stabilized by extensive hydrogen bonding between the strands, which typically undergo very little H/D exchange (34). The sole α helix in the XynII structure contains fully exchanged amides on the outer side, whereas those on the inner side facing β strands B7 and B8 are completely non-exchanged. The H/D exchange patterns in all other neutron structures are similar to that in **pD6.2**, except for **pD8.9** with slightly higher exchange in the A2-A5 β strands and the α helix. This observation suggests that **pD8.9** is more dynamic than the other structures, which may be due to the increased pD. The H/D exchange patterns for the studied structures are indicative of an enzyme whose backbone is well ordered (Figures S4 and S5), in agreement with the previous NMR measurements (35).

pK_a calculations and MD simulations

Using Poisson-Boltzmann continuum electrostatics calculations (see Supporting Text for details) we observed a conformational dependence in the pK_a of Glu177, with the "downward" state bearing a higher pK_a than the "upward" state. We used **pD4.8** and **pD6.2** to estimate the pK_a values of selected ionizable residues (Tables S3, S4). To investigate the dependence of the computed pK_a's on the internal dielectric constant (ε_{int}), we computed their values with $\varepsilon_{int} = 10$ and $\varepsilon_{int} = 4$. The pK_a of Glu86 with $\varepsilon_{int} = 10$ is similar for both **pD4.8** and **pD6.2**, with values of 3.5 and 3.9, respectively. Reducing ε_{int} to 4 had a minor effect for both structures, lowering the pK_a by only 0.6 and 0.2 units, respectively, indicating that Glu86 is relatively insensitive to the choice of ε_{int}. The computed pK_a values are consistent with the assumption that Glu86 is deprotonated throughout the reaction.

The pK_a values of Glu177 in **pD4.8** and **pD6.2** were computed to be 6.0 and 5.3, respectively, when ε_{int} was set equal to 10. For **pD4.8**, the Glu177 carboxylate occupies a relatively buried, hydrophobic site when oriented "downward", as evidenced by the computed desolvation penalty of ~2.0 pK units. The desolvation penalty for the same residue in **pD6.2** is 1.35 pK units, indicating that the carboxylate occupies a more hydrophilic site when in the "upward" conformation. To account for the apparent hydrophobicity increase of the Glu177 side chain in the "downward" state, we reduced the internal dielectric constant from 10 to 4 and

recomputed the pK_a's. The result is a relatively large shift of +2.4 pK units, increasing the pK_a of Glu177 in the "downward" conformation to 8.3. However, when the same procedure was carried out for **pD6.2**, the pK_a was perturbed upward by only 0.6 pK units, indicating a noticeable difference in hydrophobicity between the two conformations. Thus, on the basis of the upper and lower limiting values considered here for the internal dielectric constant, we estimate that the pK_a of Glu177 is 0.7–2.4 pK units higher in the "downward" than in the "upward" orientation.

To investigate whether a bound substrate alters the pK_a of Glu177, we computed pK_a's using the 1.15 Å resolution X-ray crystal structure E177Q-X6 (Tables S3, S4). We replaced Gln with Glu and positioned its side chain in an energy-minimized "downward" conformation to mimic the Michaelis complex. The pK_a of Glu177 calculated with $\varepsilon_{int} = 4$ increased by 3.7 pK units relative to the apo-enzyme, whereas the pK_a of Glu86 shifted down by -1.9 pK units. Although we expect errors in the absolute pK_as, the direction of the pK_a shifts should be meaningful. Thus, in the Michaelis complex the substrate appears to perturb the pK_a's of Glu177 up and of Glu86 down relative to the apo-enzyme.

We used the same Michaelis complex model to estimate the free energy profile for converting the side chain of Glu177 from a "downward" to an "upward" conformation by performing classical umbrella sampling MD simulations (see Supporting Text for details). The χ_2 dihedral of the Glu177 side chain was used as the reaction coordinate in the simulations, and Glu177 was modeled in the neutral, protonated state. Consistent with **pD4.8**, these simulations revealed a free energy minimum at $\chi_2 = -78°$ (Figure S6), which corresponds to the "downward" conformation of Glu177. For the rotation of χ_2 toward the "upward" conformation, we obtained a free energy barrier of 4.3 kcal/mol with a maximum at $\chi_2 = -135°$. A second minimum at χ_2 of -177°, corresponding to the "upward" state, was located at 1.4 kcal/mol above the first minimum.

DISCUSSION

Enzymes function much more efficiently than hydronium (H_3O^+) or hydroxide (OH^-) acid/base catalysts (36). GHs achieve their rate enhancements by converting the specific acid-catalyzed bimolecular hydrolysis of glycosides in solution into a unimolecular general acid

catalyst in the active site of the enzyme (10). In GH11 enzymes the general acid catalyst is a glutamic acid (Glu177 in XynII), which is initially protonated; a second Glu residue (Glu86 in XynII) serves the role of the nucleophile and is supposed to be negatively charged. To date, the protonation of the general acid catalyst has not been directly observed in any GH enzyme.

The interplay of the two catalytic Glu residues is crucial for initiating the hydrolysis reaction. As a consequence, the general acid must be a significantly weaker acid than the nucleophile, with a microscopic pK_a above 6. The apparent pK_a values for Glu177 and Glu86 in XynII have been estimated to be around 7 and 4 on the basis of the pH dependence of activity (37). The pK_a values of corresponding Glu residues in *B. circulans* xylanase (BCX) have been determined using infrared spectroscopy and ^{13}C nuclear magnetic resonance (NMR) spectroscopy to be very similar to those in XynII, at 6.7 and 4.6, respectively (33,38,39). The active site residues are conserved in both enzymes; therefore, similar pK_a values are expected. These pK_a values indicate that in slightly acidic conditions ($pH \leq 6$) the protonation of Glu177 can be anticipated. Surprisingly, we did not observe a D atom on Oε2 of Glu177 in **pD6.2** when the side chain adopts the "upward" conformation (Figure 1a). The other oxygen of the carboxylate, Oε1, accepts D atoms in two hydrogen bonds with Tyr88 and a D_2O, and therefore cannot be protonated. We note that Gln177 occupies the same position in the E177Q-X6 Michaelis complex mimic, in which Nε2 of the side-chain amide makes a hydrogen bond with the glycosidic oxygen (30). Hence, Oε2 of Glu177 must be the oxygen atom that accepts a proton from the bulk solvent and delivers it to the glycosidic oxygen. Furthermore, no protonation of Glu177 was found in **N44D** at pD 6.4 (Figure 1b), which is even more surprising because NMR measurements indicated that the pK_a of the general acid increases by about 1.5 pK units in the equivalent N35D variant of BCX relative to the wild-type enzyme (33,40).

To reconcile the apparent inconsistencies between the experimental pK_a values of the catalytic Glu residues requiring Glu177 protonation in XynII at pD ~ 6, and the lack of D atoms on this glutamate in the **pD6.2** and **N44D** neutron structures, we analyzed in more detail how the surroundings of both Glu177 and Glu86 may affect their pK_a values. Both carboxylates are exposed to solvent and are similarly hydrated. Glu86 makes weak interactions with nearby mobile water molecules, and is tightly hydrogen bonded to Tyr77 and Gln136. The hydrogen bonding lowers the pK_a of Glu86, as does the 5 Å proximity to the positively charged Arg122 as

was previously proposed (41). In the "upward" conformation Glu177 is hydrogen bonded to two D_2O molecules, makes weak contacts with several mobile waters, and is hydrogen bonded with Tyr88. The hydrogen bond between Tyr88 and Glu177 is an important feature of the catalytic site. Site-directed mutagenesis experiments demonstrated that substituting the equivalent Tyr80 in BCX with Phe crippled enzyme activity (33). In addition, the strength of the Glu177···Tyr88 hydrogen bond appears to be pH dependent, with Tyr88 possibly having a pK_a lower than the standard value of ~10. When the pD was increased to 8.9 this interaction transformed into a possible low-barrier hydrogen bond as observed in **pD8.9**, with the D atom of the phenolic oxygen situated halfway between the latter and Oε1 of Glu177 (Figure 1c). Glu86 and Glu177 have comparable hydrophobic surroundings. Glu86 has hydrophobic interactions with Trp79 and Phe134, whereas Glu177 contacts Val46 and Tyr73. Consequently, the reason for the very different experimental pK_a values of Glu86 and Glu177 (and the equivalent residues in other GH11 enzymes) appears unclear. The only features that set the carboxylates apart are one versus two hydrogen bonds made by Glu177 and Glu86, respectively, with enzyme residues, and the proximity of Arg122 to Glu86. Our pK_a calculations revealed that when Glu177 is in the "upward" conformation its pK_a is below 6, which would explain why protonation was not observed in our neutron structures, whereas Glu86 is expectedly more acidic than Glu177 (Table 1). The same trend in computed pK_a values was found for the catalytic residues in BCX and *Bacillus agaradhaerens* xylanases (41).

Törrönen and Rouvinen previously proposed that Glu177 changes its conformation "downward" at pH ≤ 6 by rotating about χ2, placing it in a different environment that can alter its pK_a (28,29). We have directly observed that Glu177 is indeed protonated when it occupies the "downward" conformation in the **pD4.8** neutron structure, and we argue that the "downward" conformation is functionally relevant. In this orientation the side-chain carboxylic group is positioned in a hydrophobic pocket, where it donates a D atom in a short hydrogen bond with a mobile D_2O, but makes no other hydrogen bonding interactions (Figures 1d and 2). Its computed pK_a increases by 0.7-2.4 pK units relative to the value in the "upward" conformation, demonstrating that the carboxyl is more basic when positioned "downward". The pK_a of Glu86 is decreased by 0.4-0.8 pK units due to the negative charge on Glu177 moving away from it, which weakens their electrostatic interaction. Thus, the apparent pK_a value of the general acid catalyst measured previously by NMR can be considered as an average of the intrinsic pK_a's of the two

glutamate conformations. Sequence analysis of GH11 enzymes reveals that residues are highly conserved in the pocket surrounding the "downward" conformation of Glu177 (42). In acidophilic xylanases active well below pH 5.0, the Asn71, Gly178, and Tyr179 residues from this pocket are frequently replaced by Ser, Ala, and Trp, respectively, whereas Asn44 is universally substituted with Asp.

Based on the neutron structures and pK_a calculations reported here we propose that Glu177 obtains a proton from the solvent when in the "downward" conformation. In our X-ray structure of the recombinant wild-type XynII obtained using a high concentration of NaCl, a Na^+ cation is located in the site occupied by the D_2O hydrogen bonded to the protonated Glu177 in **pD4.8**. It is possible that this is a specific cation-binding site that could be occupied by an H_3O^+ ion that protonates Glu177 when it rotates downward. The side chain then rotates up to assume the "upward" conformation, in which Glu177 can act as a general acid catalyst to transfer its proton to the glycosidic oxygen of the substrate (Figure 3). This conformational change requires ~4 kcal/mol of free energy according to umbrella sampling simulations of the Michaelis complex, with the protonated Glu177 being slightly more stable in the "downward" than the "upward" conformation (Figure S6). The energy barrier is low enough, however, that the "downward" and "upward" conformations of the side chain of Glu177 are in dynamic equilibrium before the hydrolysis reaction begins. Interestingly, upon downward rotation of Glu177 there is a significant movement of nearby tyrosine residues, Tyr73, Tyr77, and Tyr88, whose side chains shift in the opposite direction to that of the glutamate motion (Figure 2). This movement brings the phenolic oxygen of Tyr88 within hydrogen bonding distance of the carboxylate of Glu86 ($O^{\cdots}O$ distance = 3.0 Å). However, the O-D group of Tyr88 does not rotate toward Glu86, but remains oriented toward Glu177. This behavior of Tyr88 would ensure that when Glu177 returns back to the "upward" conformation there is no extra energy cost associated with breaking a hydrogen bond with Glu86, and that Tyr88 always stays correctly positioned to quickly capture the moving Glu177 in the correct orientation for the reaction to be initiated.

Changes in the computed pK_a values of the catalytic glutamates due to the rotation of Glu177 in the Michaelis complex are similar to those calculated for apo-XynII (Tables S3). Glu177 becomes a stronger acid and Glu86 a weaker acid when the side-chain orientation of Glu177 changes from the "downward" to the "upward" conformation. In the **WT-MES** neutron

structure, Glu177 forms a short hydrogen bond with the protonated amino nitrogen of MES, whereas Glu86 is 3.3 Å from the carbon atom vicinal to the nitrogen (Figure S3). These two contacts represent interactions that the catalytic glutamates would make in the real reactant complex, although the hydrogen would be donated by Glu177 to the glycosidic oxygen. In **WT-MES**, Oε1 of Glu177 makes an identical hydrogen bond with Tyr88 to that in **pD6.2**. Asn44 moves 0.7 Å closer to Oε2 relative to its position in **pD6.2**, but does not form a hydrogen bond with Glu177 as observed in E177Q-X6 (30). Repulsive interactions of Asn44 with the morpholinic ring, whose orientation is different from the xylose unit at position -1, may be responsible for its inability to slide closer to Glu177 in **WT-MES**.

It has been previously demonstrated that a single amino acid substitution of Asn35 to Asp in BCX reduces the pH optimum of activity from 5.7 to 4.6 (40). It was proposed that the N35D variant follows a reverse protonation mechanism (43). In that scenario, the catalytically active state of the enzyme would contain the minor protonated species of Asp35 with a pK_a of 3.7, acting as the general acid instead of Glu172, and the minor deprotonated species of the nucleophilic Glu78 with a pK_a of 5.7. Because the pK_a of the general acid Glu172 was measured to be 8.4 in the N35D variant, both residues, Asp35 and Glu172, were considered to be protonated in the catalytic site, serving the role of the general acid in unison because Glu172 would be too weak an acid to function alone. We did not observe protonation of the corresponding general acid Glu177 in our **N44D** neutron structure, indicating that the "upward" conformation has a pK_a lower than 6.0, as it does in the native XynII. We suggest that in **N44D** (corresponding to N35D of BCX and other similarly substituted xylanases) Glu177 can still act as the general acid catalyst, while maintaining dynamic equilibrium between the "downward" conformation with the higher pK_a and the "upward" conformation with the lower pK_a. The "downward" conformation may have a higher pK_a than that in the native enzyme due to the presence of a nearby negative charge on the Asp, explaining the value of 8.4 for the general acid determined by NMR in the *B. circulans* N35D variant. With the general acid being able to cycle its pK_a by adopting two different conformations, and its intrinsic pK_a in the "upward" conformation being close to that in the native enzyme according to our neutron structure, a reverse protonation mechanism is unnecessary to explain the activity of the Asn-to-Asp variant at pH values above 5. However, protonation of this Asp may become important at pH values below 5 when the population of the neutral species can be sufficient to affect the rate of catalysis.

CONCLUSIONS

To our knowledge, this is the first study to report direct determination of proton positions in a GH enzyme using neutron crystallography. We have obtained five neutron structures of a GH11 xylanase at different pD values in the apo- and ligand-bound forms, and performed pK_a calculations and free energy simulations. Our findings have led us to propose that in GH11 enzymes the Glu general acid side chain dynamically moves between the "downward" conformation, with its high pK_a facilitating its protonation, and the "upward" conformation, with its low pK_a facilitating the delivery of a proton to the glycosidic oxygen (Figure 3). We believe that our findings resolve decades of confusion surrounding the question of why the general acid and nucleophilic glutamates in GH11 enzymes, while having similar hydrogen bonding interactions and solvent accessibilities, exhibit very different pK_a values and assume their corresponding catalytic roles. On the basis of the H/D exchange analysis of XynII and the dynamics of BCX measured with NMR, we hypothesize that GH11 enzymes do not require palm dynamics for function, but the conformational flexibility of the thumb and fingers may be crucial for substrate binding and product release.

METHODS

For neutron diffraction single crystals of XynII and recombinant **N44D** variant of 2-4 mm^3 in volume were obtained using vapor diffusion methods in sitting drops. Neutron crystallographic data were collected at the PCS, LANSCE, LANL, BioDiff (FRMII), and MaNDi (SNS, ORNL). Corresponding room-temperature X-ray diffraction data for joint X-ray/neutron refinement were collected in-house. Low-temperature X-ray data were obtained on beamline ID19 at 100K at the Advanced Photon Source, Argonne, IL, USA. The neutron structures were refined in nCNS and PHENIX using joint X-ray/neutron refinement. The H++ web server version 3.1, which uses a Poisson-Boltzmann continuum electrostatics approach to approximate the electrostatic environment of biological macromolecules, was used to compute the pK_as of ionizable residues in the pD 4.8 and pD 6.2 neutron structures of xylanase, **pD4.8** and **pD6.2**, respectively. Methods are described in detail in *Supporting Text*.

Acknowledgements

The research was partly supported by the LDRD Program at ORNL, which is managed by UT-Battelle LLC for the US Department of Energy (DOE) under DOE Contract No. DE-AC05-00OR22725. We thank the Center for Structural Molecular Biology (CSMB) at ORNL supported by the DOE Office of Biological and Environmental Research (BER) for using their facilities. MaNDi beamline at the SNS was supported by the Scientific User Facilities Division, DOE BES. PCS is funded by the DOE BER. We thank ID19 beamline at the Advanced Photon Source at Argonne National Laboratory for assistance with data collection. Argonne is operated by UChicago Argonne LLC for the DOE Office of Science under contract DE-AC02-06CH11357. We thank Demian M. Riccardi for his contributions to pK_a calculations. LC, PL and AK were partly supported by the DOE BES. QW is partly supported by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry and the Innovation Fund of Yangzhou University (No. 2014CXJ058).

Author contributions

QW, PL, DEG and AK designed research; QW, JMP, and AK performed research; BLH collected X-ray diffraction data; ZF, AO, TES and LC collected neutron diffraction data; QW, JMP, and AK analyzed data; QW, JMP, DEG, PL and AK wrote the paper.

References

1. Wolfenden R, Lu X, Young, G. (1998) Spontaneous hydrolysis of glycosides. *J Am Chem Soc* 120:6814-6815.

2. Wilson DB (2009) Cellulases and biofuels. *Curr Opin Biotechnol* 20:295-299.

3. Dodd D, Cann I.K.O. (2009) Enzymatic deconstruction of xylan for biofuel production. *GCB Bioenergy* 1:2-17.

4. Sharma V, Nagpal M, Jain UK, Mangotia A, Kumar R. (2013) Antidiabetic drug and combination therapy. *Adv. Res. Pharm. Biol.* 3:389-394.

5. Gloster TM, Davies GJ (2010) Glycosidase inhibition: Assessing mimicry of the transition state. *Org Biomol Chem* 8:305-320.

6. Speciale G, Thompson AJ, Davies GJ, Williams JA Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr Opin Struct Biol* 28:1-13.

7. Smith MB, March J. (2007) March's advanced organic chemistry: Reactions, mehcanisms, and structure. Wiley Interscience, John Wiley and Sons, Inc. New Jersey, pp 2357.

8. Selwood T, Sinnott ML (1990) A solvent-isotope-effect study of proton transfer during catalysis of *Escherichia coli* (lacZ) beta-galactosidase. *Biochem J* 268:317-323.

9. Rosenberg S, Kirsch JF (1981) Oxygen-18 leaving group kinetic isotope effects on the hydrolysis of nitrophenyl glycosides. 1. Beta-galactosidase-catalyzed hydrolysis. *Biochemistry* 20:3189-3196.

10. Legler G (1990) Glycoside hydrolases: Mechanistic information from studies with reversible and irreversible inhibitors. *Adv Carbohydr Chem Biochem* 48:319-384.

11. Koshland DE, Jr (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev Camb Soc* 28:416-435.

12. Zechel DL, Withers SG (2000) Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc Chem Res* 33:11-18.

13. Phillips DC (1967) The Henn egg-white lysozyme molecule. *Proc Natl Acad Sci USA* 57:483-495.

14. Rupley JA, Gates V (1967) Studies of the enzymatic activity of lysozyme, II. The hydrolysis and transfer reactions of N-acetylglucosamine oligosaccharides. *Proc Natl Acad Sci USA* 57:496-510.

15. Chiba S (2012) A historical perspective for the catalytic reaction mechanism of glycosidase; So as to bring about breakthrough in confusing situation. *Biosci Biotechnol Biochem* 76:215-231.

16. Banait NS, Jencks WP (1991) Reaction of anionic nucleophiles with α-D-glucopyranosyl fluoride in aqueous solution through a concerted, A_ND_N (S_N2) mechanism. *J Am Chem Soc* 113:7951-7958.

17. Notenboom V, Birsan C, Nitz M, Rose DR, Warren RAJ, Withers SG (1998) Insights into transition state stabilization of the β-1,4-glycosidase Cex by covalent intermediate accumulation in active site mutants. *Nat Struct Mol Biol* 5:812-818.

18. Davies GJ, Mackenzie L, Varrot A, Dauter M, Brzozowski AM, Schülein M, Withers SG (1998) Snapshots along an enzymatic reaction coordinate: Analysis of a retaining β-glycoside hydrolase. *Biochemistry* 37:11707-11713.

19. Suzuki R, Fujimoto Z, Ito S, Kawahara S-I, Kaneko S, Taira K, Hasegawa T, Kuno A (2009) Crystallographic snapshots of an entire reaction cycle for a retaining xylanase from *Streptomyces olivaceoviridis* E-86. *J Biochem* 146:61-70.

20. Tanaka Y, Tao W, Blanchard JS, Hehre EJ (1994) Transition state structures for the hydrolysis of alpha-D-glucopyranosyl fluoride by retaining and inverting reactions of glycosylases. *J Biol Chem* 269:32306-32312.

21. Huang X, Tanaka KSE, Bennet AJ (1997) Glucosidase-catalyzed hydrolysis of α-glucopyranosyl pyridinium salts: Kinetic evidence for nucleophilic involvement at the glucosidation transition state. *J Am Chem Soc* 119:11147-11154.

22. Mori H, Lee J-H, Okuyama M, Nishimoto M, Ohguchi M, Kim D, Kimura A., Chiba S (2009) Catalytic reaction mechanism based on α-secondary deuterium isotope effects in hydrolysis of trehalose by European honeybee trehalase. *Biosci Biotechnol Biochem* 73:2466-2473.

23. Vuong TV, Wilson DB (2010) Glycoside hydrolases: Catalytic base/nucleophile diversity. *Biotechnol Bioeng* 107:195-205.

24. White A, Rose DR (1997) Mechanism of catalysis by retaining β-glycosyl hydrolases. *Curr Opin Struct Biol* 7:645-651.

25. Niimura N, Podjarny A (2011) Neutron protein crystallography. Hydrogen, protons, and hydration in bio-macromolecules. Oxford Press, pp 232.

26. Paës G, Berrin J-G, Beaugrand J (2012) GH11 xylanases: Structure/function/properties relationships and applications. *Biotech Adv* 30:564-592.

27. Törrönen A, Mach RL, Messner R, Gonzalez R, Kalkkinen N, Harkki A, Kubicek CP (1992) The two major xylanases from *Trichoderma reesei*: Characterization of both enzymes and genes. *Nat Biotechnol* 10:1461-1465.

28. Törrönen A, Harkki A, Rouvinen J (1994) Three-dimensional structure of endo-1,4-β-xylanase II from *Trichoderma reesei*: two conformational states in the active site. *EMBO J* 13:2493-2501.

29. Törrönen A, Rouvinen J (1995) Structural comparison of two major endo-1,4-xylanases from Trichoderma reesei. *Biochemistry* 34:847-856.

30. Wan Q, Zhang Q, Hamilton-Brehm S, Weiss K, Mustyakimov M, Coates L, Langan P, Graham D, Kovalevsky A (2014) X-ray crystallographic studies of family 11 xylanase Michaelis and product complexes: Implications for the catalytic mechanism. *Acta Cryst* D70:11-23.

31. Davies GJ, Wilson KS, Henrissat B (1997) Nomenclature for sugar-binding subsites in glycosyl hydrolsases. *Biochem J* 321:557-559.

32. Mhlongo NN, Skelton AA, Kruger G, Soliman MES, Williams IHA (2014) Critical survey of average distances between catalytic carboxylic groups in glycoside hydrolases. *Prot Struct Funct Bioinf* 82:1747-1755.

33. Joshi MD, Sidhu G, Nielsen JE, Brayer GD, Withers SG, McIntosh LP (2001) Dissecting the electrostatic interactions and pH-dependent activity of a family 11 glycosidase. *Biochemistry* 40:10115-10139.

34. Blum M-M, Mustyakimov M, Rüterjans H, Kehe K, Schoenborn BP, Langan P, Chen JC-H (2009) Rapid determination of hydrogen positions and protonation states of diisopropyl fluorophosphatase by joint neutron and X-ray diffraction refinement. *Proc Natl Acad Sci USA* 106:713-718.

35. Connelly GP, Withers SG, McIntosh LP (2000) Analysis of the dynamic properties of *Bacillus circulans* xylanase upon formation of a covalent glycosyl-enzyme intermediate. *Prot Sci* 9:512-524.

36. Cox MM, Phillips GN (2007) Enzymes: The active site. Handbook of Proteins 1, John Wiley and Sons Ltd., Chichester, UK, 396-403.

37. Tenkanen M, Puls J, Poutanen K (1992) Two major xylanases from Trichoderam reesei. *Enz Microb Technol* 14:566-574.

38. Davoodi J, Wakarchuk WW, Campbell RL, Carey PR, Surewicz WK (1995) Abnormally high pKa of an active-site glutamic acid residue in *Bacillus circulans* xylanase. The role of electrostatic interactions. *Eur J Biochem* 232:839-843.

39. McIntosh LP, Hand G, Johnson PE, Joshi MD, Körner M, Plesniak LA, Ziser L, Wakarchuk WW, Withers SG (1996) The pKa of the general acid/base carboxyl group of a glycosidase cycles during catalysis: A ^{13}C-NMR study of *Bacillus circulans* xylanase. *Biochemistry* 35:9958-9966.

40. Joshi MD, Sidhu G, Pot I, Brayer GD, Withers SG, McIntosh LP (2000) Hydrogen bonding and catalysis: A novel explanation for how a single amino acid substitution can change the pH optimum of a glycosidase. *J Mol Biol* 299:255-279.

41. Kongsted J, Ryde U, Wydra J, Jensen JH (2007) Prediction and rationalization of the pH dependence of the activity and stability of family 11 xylanases. *Biochemistry* 46:13581-13592.

42. Sapag A, Wouters J, Lambert C, de Ioannes P, Eyzaguirre J, Depiereux E (2002) The endoxylanases from family 11: Computer analysis of protein sequences reveals important structural and phylogenetic relationships. *J Biotechnol* 95:109-131.

43. Mock WL (1992) Theory of enzymatic reverse protonation catalysis. *Bioorg Chem* 20:377-381.

44. Wan, Q., Kovalevsky, A., Zhang, Q., Hamilton-Brehm, S., Upton, R., Weiss, K.L., Mustyakimov, M., Graham, D., Coates, L. and Langan, P. Heterologous expression, purification, crystallization and preliminary X-ray analysis of *Trichoderma reesei* xylanase II and four variants. *Acta Cryst.* **F69**, 320-323 (2013).

45. Minor, W., Cymborowski, M., Otwinovski, Z. and Chruszcz, M. HKL3000: The integration of data reduction and structure solution – from diffraction images to an initial model in minutes. *Acta Cryst.* **D62**, 859-866 (2006).

46. Sheldrick, G. M. and Schneider, T. R. High-resolution refinement. *Methods Enzymol.* **277**, 319-343 (1997).

47. Adams, P.D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst.* **D66**, 213-221 (2010).

48. Kovalevsky, A.Y., Hanson, B.L., Seaver, S., Fisher, S.Z., Mustyakimov, M. and Langan, P. Preliminary joint X-ray and neutron protein crystallographic studies of endoxylanase II from the fungus *Trichoderma longibrachiatum*. *Acta Cryst.* **F67**, 283-286 (2011).

49. Otwinoski, Z. and Minor, W. *Methods in Enzymology*, vol. 276, *Macromolecular Crystallography*, Part A, C. W. Carter Jr., R. M. Sweet, Eds., Academic Press, New York, , pp. 307–326 (1997).

50. Coates, L., Stoica, A.D, Hoffmann, C., Richards, J. and Cooper, R. The macromolecular neutron diffractometer (MaNDi) at the Spallation Neutron Source, Oak Ridge: enhanced

optics design, high-resolution neutron detectors and simulated diffraction. *J. Appl. Cryst.* **43**, 570-577 (2010).

51. Adams, P.D., Mustyakimov, M., Afonine, P.V. and Langan, P. Generalized X-ray and neutron crystallographic analysis: more accurate and complete structures for biological macromolecules. *Acta Cryst.* **D65**, 567-573 (2009).

52. Anandakrishnan, R., Aguilar, B., and Onufriev, A. V. H++ 3.0: automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* **40**, W537 (2012).

53. Kumar, S., Bouzida, D., Swendsen, R.H., Kollman, P.A. and Rosenberg. J.M. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comp. Chem.* **13**, 1011-1021 (1992).

54. Phillips, J.C. et al. K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781-1802 (2005).

55. Brooks, B.R. et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **30**, 1545-1614 (2009).

56. Guvench, O., Hatcher, E. R., Venable, R. M., Pastor, R. W. and MacKerell, Jr., A. D. Additive empirical CHARMM Force Field for glycosyl-linked hexopyranoses. *J. Chem. Theory Comput.* **5**, 2353-2370 (2009).

Figure legends

Figure 1. Neutron scattering length density maps for the active sites in XynII structures. (a) **pD6.2** , (b) **N44D** , (c) **pD8.9** , and (d) **pD4.8**. $2F_O$-F_C maps are shown as green meshes contoured at 1.2 σ. F_O-F_C omit maps are shown as red meshes contoured at 3.0 σ in (a), (b), and (c), and at 2.5 σ in (d). Hydrogen bonds are shown as blue dotted lines. D atoms are dark cyan, while H atoms are light grey. Distances are in Å.

Figure 2. Superposition of **pD4.8** (atoms colored by atom type, carbon is yellow and cyan) and **pD6.2** (atoms colored by atom type, carbon is green and magenta).

Figure 3. Proposed cycling of the general acid between two conformations controlling pKa values in GH11 enzymes that ensure Glu protonation and delivery of the proton to the substrate's glycosidic oxygen. A Tyr residue that is hydrogen bonded to the general acid moves toward the nucleophile when the general acid changes its conformation downward, but the phenolic group of the Tyr makes no hydrogen bond with the nucleophile.