SUPPORTING INFORMATION

A Single Point Mutation Converts GH84 *O*-GlcNAc Hydrolases into Phosphorylases. Experimental and Theoretical Evidence.

David Teze[1,2,5*], Joan Coines[3,5], Lluís Raich[3], Valentina Kalichuk[2], Claude Solleux[2], Charles Tellier[2], Corinne Andre-Miral[2], Birte Svensson[1*] and Carme Rovira[3,4*]

[1]Department of Biotechnology and Biomedicine, Technical University of Denmark, Kongens Lyngby, Denmark. [2]UFIP, CNRS, Université de Nantes, Nantes, France. [3]Departament de Química Inorgànica i Orgànica and Institut de Química Teòrica i Computacional (IQTCUB), Universitat de Barcelona, Barcelona, Spain. [4]Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain. [5]These authors contributed equally. Correspondence should be addressed to D.T. (davtez@dtu.dk), B.S. (bis@bio.dtu.dk) or C.R. (c.rovira@ub.edu).

CONTENTS

1. Experimental setup

1.1 Protein production and purification

pET-28a(+) plasmids encoding the genes of choice codon-optimised for *Escherichia coli* production were ordered from Biomatik (Ontario, Canada). Gene sequences can be found in Uniprot,[1] under the codes Q99YP8 (*Sp*OGA) and D1CDN2 (*Tt*OGA). Precultures of *E. coli* BL21(DE3) transformed cells were used to inoculate 1 L of lysogeny broth media containing 50 mg kanamycin. Cultures were incubated at 37ºC with shaking at 180 rpm until OD_{600} reached ~0.6. Gene expression was induced with 200 μM isopropyl β-D-1-thiogalactopyranoside and continued overnight at 20ºC. Cultures were subsequently centrifuged, pellets resuspended in cold 50 mM HEPES buffer pH 8 containing 250 U of benzonase nuclease, lysed (Cell Pressure Homogenizer, Stansted, UK), and centrifuged. GH84 variants were purified from the supernatant by Ni^{2+}-affinity chromatography (HisTrapHP, GE Healthcare, Sweden). Eluates were analysed by SDS-PAGE, and protein concentration was evaluated using UV_{280} absorption measured on a Nanodrop (Thermo Scientific) and molar extinction coefficients computed from the ExPASy server (www.expasy.org).

1.2 Nuclear magnetic resonance (NMR)

NMR spectra were recorded on an 800 MHz Bruker Avance III (799.75 MHz for ^{1}H) equipped with a 5 mm TCI cryoprobe using ^{1}H with presaturation. ^{1}H chemical shifts corresponding to published ^{1}H chemical shifts[2] for GlcNAc-1-P were identified as the only additional peaks occurring upon phosphate addition to the reaction mixture. Reactions were carried out with 7.5 mM *p*NP-GlcNAc, 0 or 50 mM phosphate, and 1 $\mu g \cdot mL^{-1}$ *Sp*OGA-D245N (15 nM) in 600 μL HEPES buffer 50 mM with measured pD 7.2 (corresponding to pH 7.8).[3] Time course experiments were obtained using pseudo-2D kinetics experiments, with spectra recorded every 3 or 5 min. Integration of anomeric protons is inaccurate due to the closeness of the presaturated HOD peak.

1.3 Capillary electrophoresis (CE)

The formation of GlcNAc and GlcNAc-1-P, as well as *p*NP-GlcNAc or Glc-ox disappearance over time, were determined by capillary electrophoresis (Beckman P/ACE System 5000 with an uncoated fused silica capillary, 47 cm).[4,5] Media containing 7.5 mM donor (*p*NP-GlcNAc or Glc-ox) and 1 $\mu g \cdot mL^{-1}$ *Sp*OGA-D245N (15 nM) was incubated at 25°C in 200 μL HEPES buffer 50 mM, pH 7.8, within the CE apparatus and analysed every 27 min for 8 h. If specified, 0.5 M phosphate buffer pH 7.8 was added. Separations were performed at 17 kV with 50 mM borax, pH 9.5 as running buffer.

1.4 Kinetics

Kinetics experiments were performed in a PowerWave XS microtiter plate reader monitored by the Gen5 software (Bio-Tek Instruments, Inc.) by measuring OD_{405} each 40 s for 30 min at 25°C. GH84 variants (100 nM final) were added to 0.0025–10 mM *p*NP-GlcNAc (twelve concentrations) in 200 μL of 50 mM HEPES buffer pH 7.8, at 25°C. Phosphate buffer pH 7.8 (0−500 mM final) was added. Initial rates were calculated from slopes of *p*NP formation *vs* time and curves plotted against substrate concentration were

fitted to the Michaelis-Menten equation $v = \frac{k_{cat}\cdot[E]\cdot[S]}{K_M+[S]}$ using OriginPro 2015 (OriginLab, USA) to obtain k_{cat} and K_M.

1.5 High performance anion exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD)

Samples containing 0.05–5 µM of GH84, 10 mM *p*NP-GlcNAc and 0–0.5 M phosphate in 50 mM HEPES buffer pH 7.8 were incubated (2 h, 25°C), then heat denatured (2 min, 98°C), centrifuged, and the obtained supernatant was diluted 100-fold in milliQ H_2O and filtered. Separation was carried out at 30°C on a CarboPac PA-100 (Dionex) in an ICS-3000 (Dionex) monitored by the software Chromeleon (Dionex). A 0.1–0.4 M acetate gradient in 0.1 M NaOH was used as eluent in 40 min separations.

2. Computational details

2.1 System preparation and setup

The initial structure was taken from the X-ray structure of the *Tt*OGA-D120N mutant in complex with the glycopeptide hTab1-*O*-GlcNAc (PDB entry 5DIY).[6] The D120N mutation was reverted to regenerate the wild-type *Tt*OGA enzyme (*Tt*OGA-WT). The peptide was removed, except the GlcNAc residue, which served as a scaffold to build the oxazoline/oxazolinium moiety. In the simulations with a phosphate ion in the active site, its initial position was estimated by substituting the oxygen and carbon atoms of the bound side chain serine residue of the hTab1-*O*-GlcNAc glycopeptide by the O and P phosphate atoms, respectively. The protonation state of all residues was assigned considering the experimental pH of 7.8. The protonation of His residues was further assessed according their chemical environment. Specifically, His residues 20, 48, 112, 127, 200, 206, 242, 248, 250 were considered neutral with their proton located at N_δ; while His236 was protonated at N_ε. Crystallographic water molecules were retained. To compute the distribution of phosphate ions around the active site, different models were tested in which an appropriate number of Na^+ and phosphate ions were added to the system to achieve a concentration of 20 mM phosphate and reach neutrality (see section 2.2.1). Additional water molecules were added to solvate the protein and generate a simulation box of 86.4 × 97.6 × 132.7 $Å^3$. The composition of each system is specified in Table S1.

2.2. Classical MD simulations. General details

All classical molecular dynamics (MD) were run employing the AMBER14 software.[7] Amber ff14SB[8] and TIP3P[9] force fields were used to describe the protein and water molecules, respectively. Phosphate ions (HPO_4^{2-} and $H_2PO_4^-$) were described with parameters adapted from phosphorylated serine parameters.[10] Glc-oxazoline and Glc-oxazolinium ion molecules (Glc-ox and Glc-ox$^+$, respectively) were parametrized employing the antechamber module,[11] in conjunction with GAFF[12] parameters and the

RESP atomic charges obtained from first principles calculations at the HF/6-31G* level of theory using Gaussian09.[13]

All systems were equilibrated in periodic boundary conditions, using the following protocol. First, the energy was minimized with 20,000 minimization cycles, relaxing Na^+ ions and water molecules. Subsequently, the entire system was relaxed with 20,000 minimization cycles. Afterwards, the system was heated gradually to 100, 200, 250 and 300 K in the NVT ensemble at intervals of 50 ps. Spatial restraints to the protein and ligands during the first heating interval, while all restraints were released after reaching 100K. Subsequently, the density was converged up to water density at 300 K during 100 ps in the NPT ensemble. Equilibration runs of 150 ps were performed to equilibrate the RMSD in the NVT ensemble with a time step of 1 fs. The time step was subsequently increased to 2 fs, employing the SHAKE algorithm[14] and the simulations were continued for 20 ns. Analysis of the trajectories was carried out using standard tools of AMBER and VMD.[15]

2.2.1. Distribution of phosphate ions

Specific MD simulations were performed to evaluate the distribution of phosphate ions around the enzyme, in particular to check whether phosphate ions are able to get close to the active site. At pH 7.8, both HPO_4^{2-} and $H_2PO_4^-$ species of phosphate ions are present in the solution ($pK_2 \approx 7.21$), thus both species were considered. To model a phosphate concentration of 20 mM at pH 7.8 in a system box of ~910 nm^3, 9 HPO_4^{2-} and $H_2PO_4^-$ ions were randomly placed in the simulation cell. This phosphate concentration was the one that produced a maximal phosphorylation enhancement employing *Tt*OGA-D120N. The system was equilibrated following the protocol described in the previous section. Production runs of 50 ns were performed for both *Tt*OGA-WT and *Tt*OGA-D120N mutant enzymes at 20 mM phosphate concentration in complex with a Glc-oxazolinium ion. The relative phosphate density (Figure 3a) was obtained by analysing the radial distribution function of both HPO_4^{2-} and $H_2PO_4^-$ ions with respect to the side chain of D120/N120.

2.2.2. Modeling the reactivity of *Tt*OGA-Glc-oxazoline/oxazolinium ion complexes

Several systems were tested to model the phosphorylation of *Tt*OGA and *Tt*OGA-D120 complexes with the Glc-oxazoline/oxazolinium ion. Eight different systems were set up (Figure S9) depending on the protonation state of the intermediate (Glc-oxazoline or Glc-oxazolinium ion), the phosphate form (HPO_4^{2-} and $H_2PO_4^-$) and the enzyme form (*Tt*OGA-WT or *Tt*OGA-D120N). In the case of Glc-ox, relevant active site interactions were not maintained during the simulation, indicating that this is not the most stable form of the reaction intermediate. Further calculations at QM/MM level (section 2.5) confirmed that the reaction intermediate features a (positively charged) oxazolinium ion, thus only Glc-ox^+ ion models were further considered.

To elucidate which phosphate form will most likely react within the Glc-ox^+ substrate of wild-type and mutant of *Tt*OGA, either HPO_4^{2-} or $H_2PO_4^-$ anions were placed near the anomeric carbon (C1) employing a restraint during the first 10 ns of MD simulation (after

system equilibration). Several values for the constant force were subsequently employed for 5 replicas of 1 ns (Figure S10). In the case of *Tt*OGA-WT, the ions were not stable and rapidly diffused away from the active site (this happened faster in the case of the more charged HPO_4^{2-}) unless heavy restraints were used. The exclusion of phosphate ions from the active site is due to two factors: the electrostatic repulsion with D120 and the well-known tendency of force fields to oversolvate the phosphate ions.[16]. To solve these issues, and be able to obtain a reaction free energy profile, the QM/MM simulations were initiated from a snapshot of a MD simulation with a restrain in the distance between the anomeric carbon (C1) and the closest oxygen atom of the active site phosphate ion (Figure S10). In the case of the wild-type enzyme, the phosphate monoanion, for which electrostatic repulsion with D120 is lower than the dianion, was considered. In the case of the D120N mutant, the electrostatic repulsion with D120 vanishes and the phosphate anion could be accommodated in the active site employing just a weak restraint to counterbalance the limitations of the phosphate force field. All restraints released during the subsequent QM/MM simulations, in which the active site phosphate ion was treated at QM level. The more nucleophilic HPO_4^{2-} ion was considered for the QM/MM simulations of the D120N mutant. This is also consistent with the higher density of this ion close to the active site that was observed in the simulations with 20 mM phosphate concentration (Figure 3a).

2.3. QM/MM MD simulations

Quantum mechanics/molecular mechanics (QM/MM) MD calculations were performed employing the method developed by Laio et al.[17] that combines classical MD with *ab initio* MD (Car-Parrinello method).[18] Three systems were set up: wild-type *Tt*OGA with either $H_2PO_4^-$ or water molecules located above the Glc-ox$^+$ substrate (to study phosphorylation and hydrolysis reactions, respectively); and *Tt*OGA-D120N in complex with the HPO_4^{2-} ion in a reactive configuration. Snapshots from the classical MD simulations were taken as starting configurations. The QM region for all systems included the Glc-ox$^+$ substrate, the side chain of D119 and D120/N120 (saturating the C_α atom of the C_α-C_β bond) and the molecule that performs the nucleophilic attack on the anomeric carbon (either water, HPO_4^{2-} or $H_2PO_4^-$). The MM region included the rest of the protein and the solvent. The parameters for each system are shown in Table S2 and were selected following the protocol described in reference [19]. Density Functional Theory (DFT) was used to describe the electronic structure, employing the Perdew, Burke, and Ernzerhoff generalized gradient-corrected approximation (PBE),[20] which has been proven to give good results in the description of isolated carbohydrates[21] and carbohydrates in complex with glycoside hydrolases.[22] A plane wave basis set was employed to expand the Kohn-Sham orbitals, with a kinetic energy cutoff of 70 Ry. Norm-conserving *ab initio* pseudopotentials were used, generated within the Troullier-Martins Scheme.[23]

QM/MM MD calculations were performed according to the following protocol. The structure was optimized with annealing of the nuclei until reaching a maximal component of the nuclear gradient of $5 \cdot 10^{-4}$ a.u. Afterwards, 3 ps of MD were performed to re-

equilibrate the system at 300 K using the Nosé–Hoover thermostat.[24,25] A snapshot of this simulation was used as a starting point for further metadynamics simulations.

2.4. QM/MM metadynamics simulation of enzymatic reactions by *Tt*OGA and *Tt*OGA-D120N

Metadynamics calculations were performed employing the PLUMED plugin.[26] All reactions were activated using atomic distances as collective variables (CVs), which are schematically represented in Figure S11. In the case of the reactions catalysed by wild-type *Tt*OGA, we used two distances as CVs. The first CV (CV_1) accounts for the nucleophilic attack of the oxygen atom of either water or $H_2PO_4^-$ (hydrolysis and phosphorylation, respectively) on the anomeric carbon of Glc-oxazolinium ion (C1). The second CV (CV_2) accounts for the proton transfer from water or $H_2PO_4^-$ to the carboxylic oxygen of the acid/base residue (D120). The reconstructed free energy landscapes and main states along the minimum energy pathway are shown in Figure S12 (Figure 3b shows a projection into CV_1). In the case of the phosphorylation reaction catalysed by *Tt*OGA-D120N, only CV_1 was used, since no proton transfer event is necessary in this case (moreover, N120 cannot accept a proton from the phosphate).

The parameters used in the metadynamics simulations are collected in Table S3. Relevant distances and charges of the species along the enzymatic reactions are listed in Tables S4-S6. The conformation of the sugar ring at the oxazolinium ion reaction intermediate, transition state and reaction products are shown in the Mercator representation of Figure S13. The resulting conformational itineraries are: $^4C_1/^4E \rightarrow [^4E]^{\ddagger} \rightarrow {}^{1,4}B/^4E$ (phosphorylation by either *Tt*OGA-WT or *Tt*OGA-D120N) and $^4H_5/^4E \rightarrow [^4H_5]^{\ddagger} \rightarrow {}^4H_5$ (hydrolysis by *Tt*OGA).

2.5. Analysis of substrate protonation state

Starting from the equilibrated structure of *Tt*OGA in complex with Glc-ox$^+$, we computed the free energy change with respect to transferring the proton from the nitrogen of the oxazolinium ion to the assisting residue D119. We used metadynamics with one CV (CV_3 in Figure S11) that accounts for the aforementioned proton transfer. The resulting free energy profile shows that the pair D119-COOH/Glc-ox is 4.9 kcal·mol^{-1} higher in energy than D119-COO$^-$/Glc-ox$^+$ (Figure S15). Thus the proton prefers to stay in the substrate rather than being transferred to the D119 assisting residue. In fact, a simple structure optimization from D119-COOH/Glc-ox (Figure S15) shows that the proton jumps from D119 to Glc-ox. The substrate remained protonated >99.7% in all the simulations performed (three enzymatic reactions). All these results disqualify the possibility that a neutral oxazoline as the intermediate of the enzymatic reaction, hence supporting the proposed oxazolinium ion intermediate (Figure 1b).

3. References

(1) Apweiler, R.; Bairoch, A.; Wu, C. H.; Barker, W. C.; Boeckmann, B.; Ferro, S.; Gasteiger, E.; Huang, H.; Lopez, R.; Magrane, M.; Martin, M. J.; Natale, D. A.; O'Donovan, C.; Redaschi, N.; Yeh L.L. UniProt: The Universal Protein Knowledgebase. *Nucleic Acids Res.* **2004**, *32*, D115-9.

(2) Macdonald, S. S.; Blaukopf, M.; Withers, S. G. N-Acetylglucosaminidases from CAZy Family GH3 Are Really Glycoside Phosphorylases, Thereby Explaining Their Use of Histidine as an Acid/Base Catalyst in Place of Glutamic Acid. *J. Biol. Chem.* **2015**, *290* (8), 4887–4895.

(3) Covington, A. K.; Paabo, M.; Robinson, R. A.; Bates, R. G. Use of the Glass Electrode in Deuterium Oxide and the Relation between the Standardized PD (PaD) Scale and the Operational PH in Heavy Water. *Anal. Chem.* **1968**, *40* (4), 700–706.

(4) Teze, D.; Hendrickx, J.; Czjzek, M.; Ropartz, D.; Sanejouand, Y. H.; Tran, V.; Tellier, C.; Dion, M. Semi-Rational Approach for Converting a GH1 β-Glycosidase into a β-Transglycosidase. *Protein Eng. Des. Sel.* **2014**, *27* (1), 13–19.

(5) Teze, D.; Daligault, F.; Ferrières, V.; Sanejouand, Y.-H. H.; Tellier, C.; Ferrieres, V.; Sanejouand, Y.-H. H.; Tellier, C. Semi-Rational Approach for Converting a GH36 α-Glycosidase into an α-Transglycosidase. *Glycobiology* **2015**, *25* (4), 420–427.

(6) Ostrowski, A.; Gundogdu, M.; Ferenbach, A. T.; Lebedev, A. A.; Van Aalten, D. M. F. Evidence for a Functional O-Linked N-Acetylglucosamine (O-GlcNAc) System in the Thermophilic Bacterium Thermobaculum Terrenum. *J. Biol. Chem.* **2015**, *290* (51), 30291–30305.

(7) D.A. Case, V. Babin, J.T. Berryman, R.M. Betz, Q. Cai, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E.Duke, H. Gohlke, A.W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T.S. Lee, S. LeGrand, T. Luchko, R. Luo, B. M, X. W. and P. A. K. AMBER 14. *University of California,*. 2014, p San Francisco.

(8) 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.

(9) 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.

(10) Homeyer, N.; Horn, A. H. C.; Lanig, H.; Sticht, H. AMBER Force-Field Parameters for Phosphorylated Amino Acids in Different Protonation States: Phosphoserine, Phosphothreonine, Phosphotyrosine, and Phosphohistidine. *J. Mol. Model.* **2006**, *12* (3), 281–289.

(11) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, *25* (2), 247–260.

(12) Wang, J. M.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. a; Case, D. a. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.

(13) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg,

M. Had, J. C. and D. J. F. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT. **2013**.

(14) 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.

(15) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38.

(16) Steinbrecher, T.; Latzer, J.; Case, D. A. Revised AMBER Parameters for Bioorganic Phosphates. *J. Chem. Theory Comput.* **2012**, *8* (11), 4405–4412.

(17) Laio, A.; VandeVondele, J.; Rothlisberger, U. A Hamiltonian Electrostatic Coupling Scheme for Hybrid Car-Parrinello Molecular Dynamics Simulations. *J. Chem. Phys.* **2002**, *116* (16), 6941–6947.

(18) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55* (22), 2471–2474.

(19) Raich, L.; Nin-Hill, A.; Ardèvol, A.; Rovira, C. *Enzymatic Cleavage of Glycosidic Bonds: Strategies on How to Set Up and Control a QM/MM Metadynamics Simulation*, 1st ed.; Elsevier Inc., 2016; Vol. 577.

(20) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(21) Marianski, M.; Supady, A.; Ingram, T.; Schneider, M.; Baldauf, C. Assessing the Accuracy of Across-the-Scale Methods for Predicting Carbohydrate Conformational Energies for the Examples of Glucose and α-Maltose. *J. Chem. Theory Comput.* **2016**, *12* (12), 6157–6168.

(22) 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.

(23) Troullier, N.; Martins, J. L. Efficient Pseudopotentials for Plane-Wave Calculations. II. Operators for Fast Iterative Diagonalization. *Phys. Rev. B* **1991**, *43* (11), 8861–8869.

(24) Nosé, S. A Molecular Dynamics Method for Simulations in the Canonical Ensemble. *Mol. Phys.* **1984**, *52* (2), 255–268.

(25) Hoover, W. G. Canonical Dynamics: Equilibrium Phase-Space Distributions. *Phys. Rev. A* **1985**, *31* (3), 1695–1697.

(26) 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.

4. Figures S1-S15

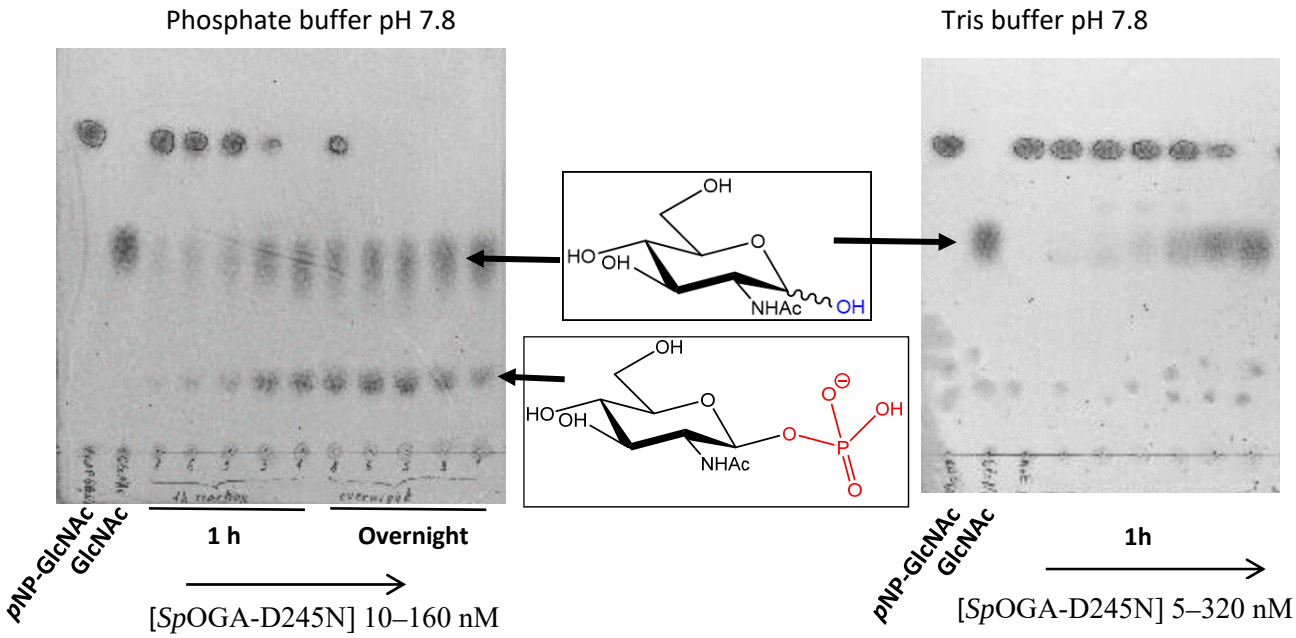

Figure S1. Analysis of reaction mixtures by thin layer chromatography, in presence of 10 mM *p*NP-GlcNAc, 5–320 nM *Sp*OGA-D245N (0.33–22 $\mu g \cdot mL^{-1}$) and 50 mM phosphate buffer (left) or 50 mM Tris buffer (right), pH 7.8, 37°C.

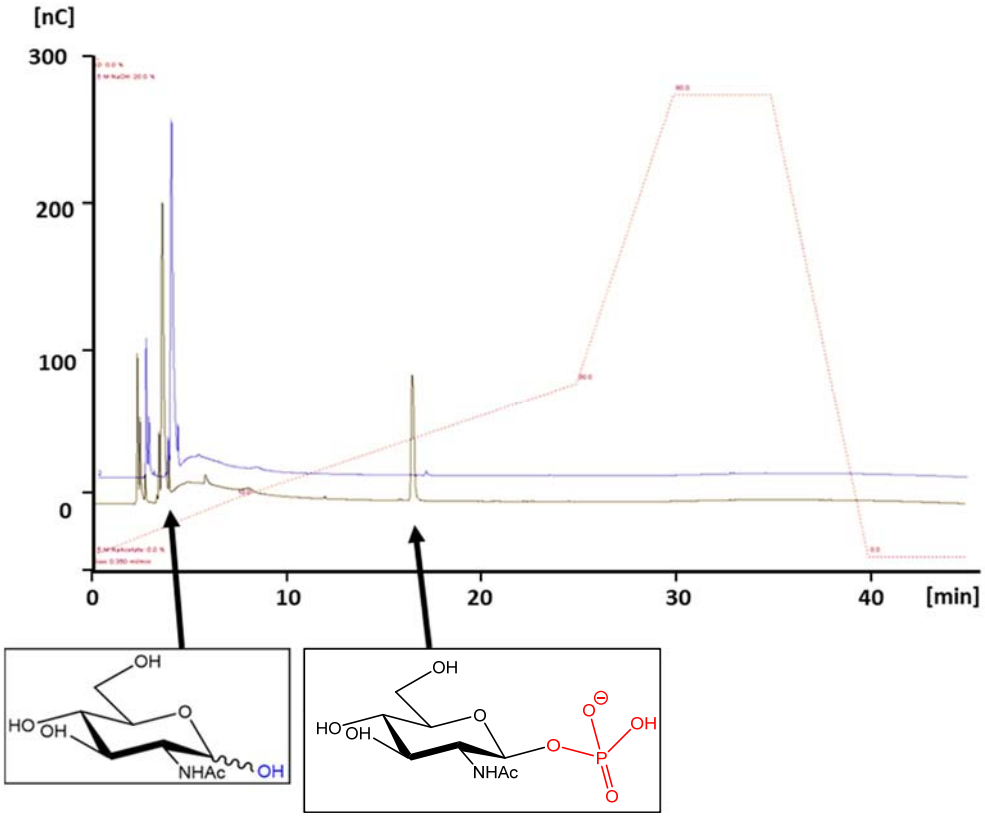

Figure S2. Chromatographic analysis of reaction mixture after 1 h at 37 °C by HPAEC-PAD, in presence of 10 mM *p*NP-GlcNAc, *Tt*OGA-D120N and 50 mM HEPES (blue) or 50 mM phosphate (black) buffer, pH 7.8, 37 °C.

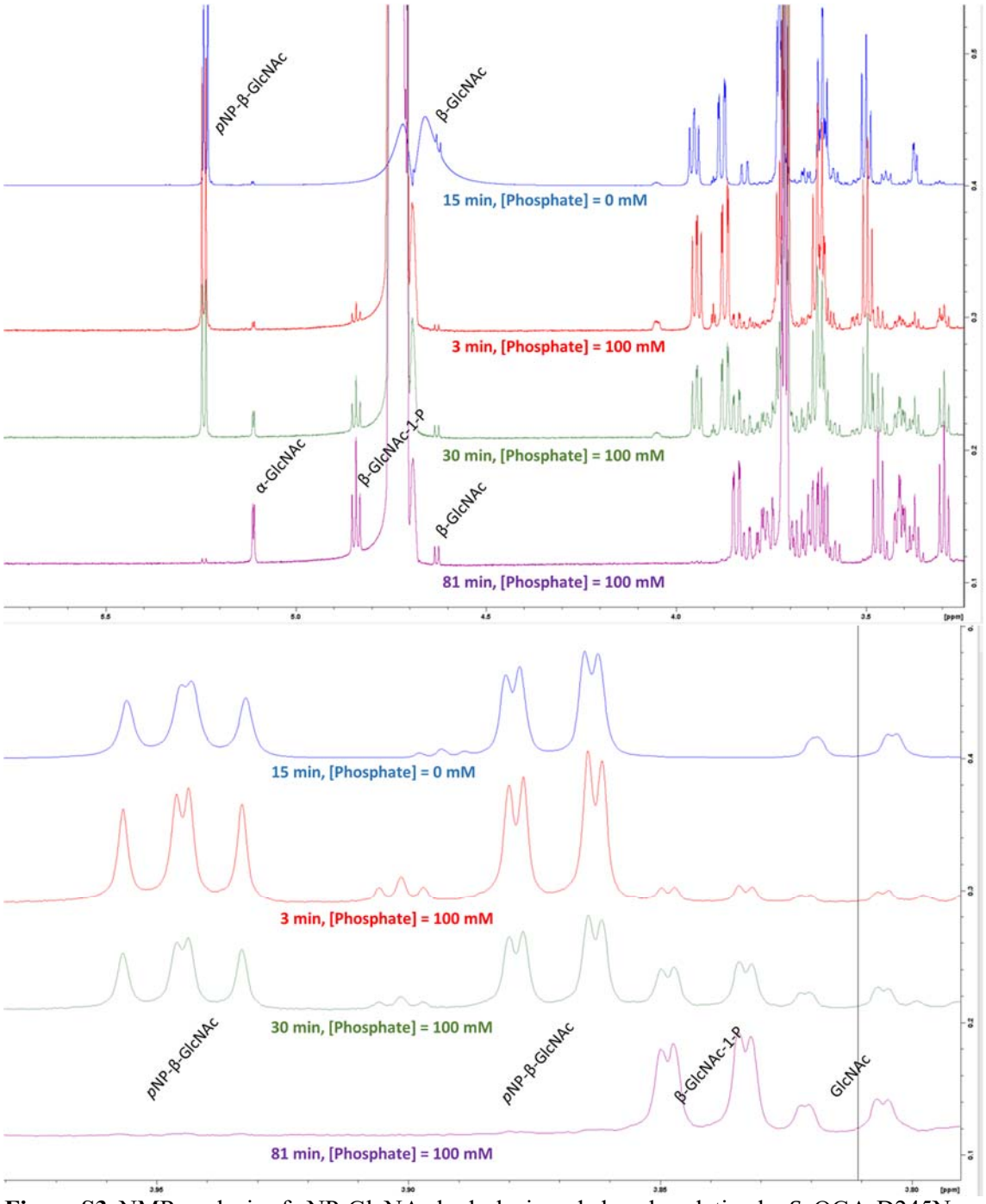

Figure S3. NMR analysis of *p*NP-GlcNAc hydrolysis and phosphorylation by *Sp*OGA-D245N. Spectra recorded in presence of 2 mM *p*NP-GlcNAc, *Sp*OGA-D245N and 50 mM HEPES (blue, top spectra, 500 µg *Sp*OGA-D245N) or 100 mM phosphate (red, green and purple, 40 µg *Sp*OGA-D245N), pD 7.8, 25°C. Top, generic view of the spectra; bottom, zoom on the region used for GlcNAc-1-P monitoring, particularly the peak at δ = 4.84 ppm. The peaks labelled "β-GlcNAc-1-P" do not appear in the reaction without phosphate, and correspond to the chemical shifts observed for the GH3 GlcNAc phosphorylases.[2]

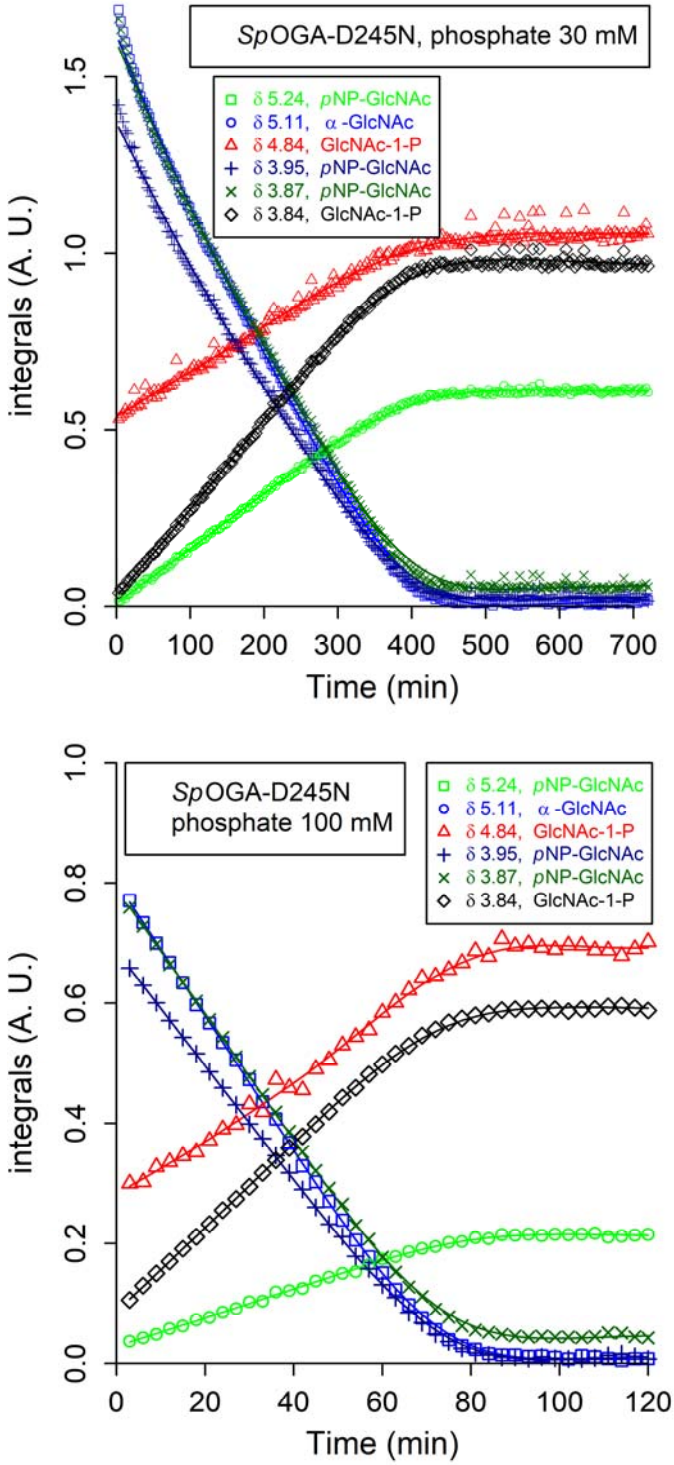

Figure S4. *p*NP-GlcNAc hydrolysis and phosphorylation by *Sp*OGA-D245N according to phosphate concentration. Spectra recorded in presence of 2 mM *p*NP-GlcNAc, *Sp*OGA-D245N and phosphate 30 mM (top, 10 µg *Sp*OGA-D245N) or 100 mM phosphate (bottom, 40 µg *Sp*OGA-D245N), pD 7.8, 25°C. Due to the proximity with the HOD peak and its presaturation, the integrals of the peak at δ = 4.84 ppm (in red) cannot be used to quantify GlcNAc-1-P, and the peak at δ = 3.84 ppm is used instead (see Figure S3, bottom).

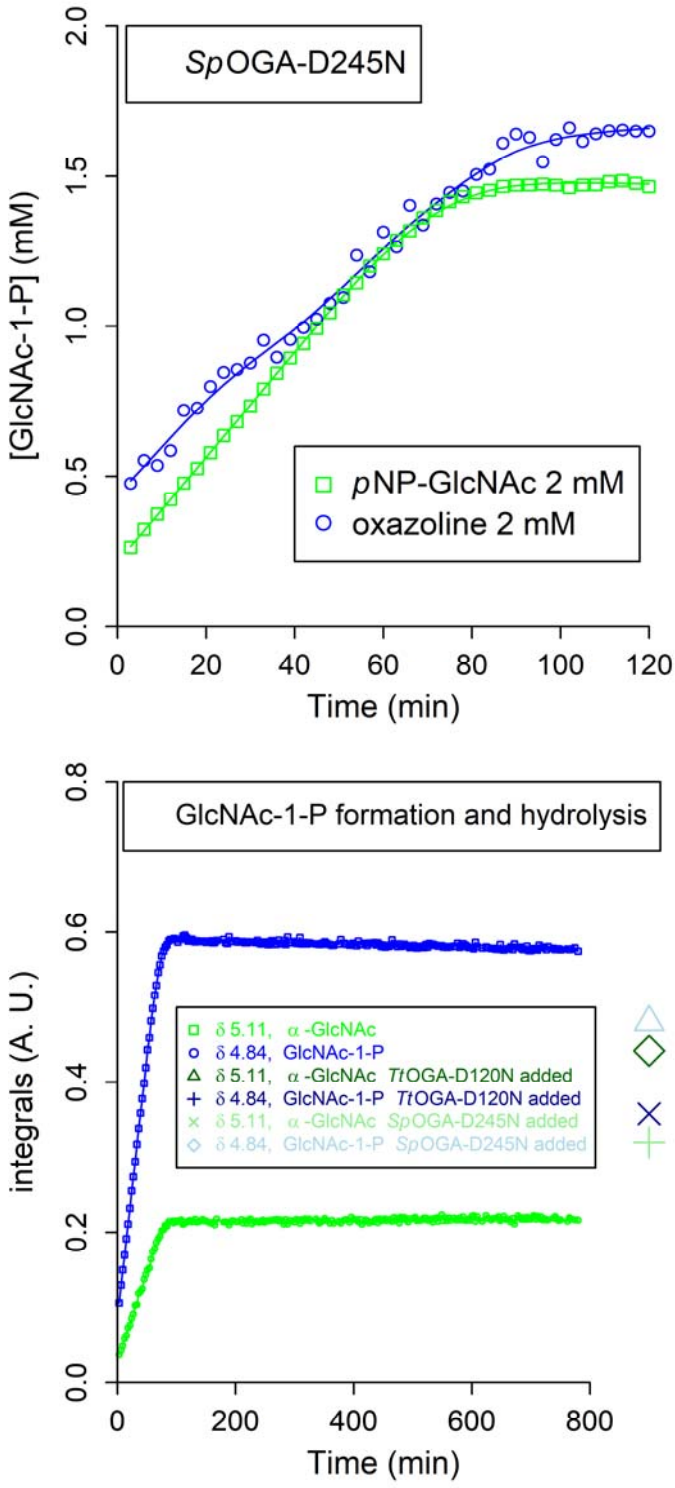

Figure S5. Top, *p*NP-GlcNAc and oxazoline phosphorylation by *Sp*OGA-D245N. Spectra recorded in presence of 2 mM *p*NP-GlcNAc or oxazoline, *Sp*OGA-D245N 40 µg and phosphate 100 mM. Bottom, GlcNAc-1-P phosphorolysis. Data on the right were recorded 4 h after addition of 40 µg *Sp*OGA-D245N or 40 µg *Tt*OGA-D120N), pD 7.8, 25°C. From the slopes, GlcNAc-1-P hydrolysis by *Sp*OGA-D245N is ≈ 2 orders of magnitudes slower than GlcNAc-1-P formation by the same enzyme. *Tt*OGA-D120N is approximatively 3-fold faster than *Sp*OGA-D245N for GlcNAc-1-P hydrolysis.

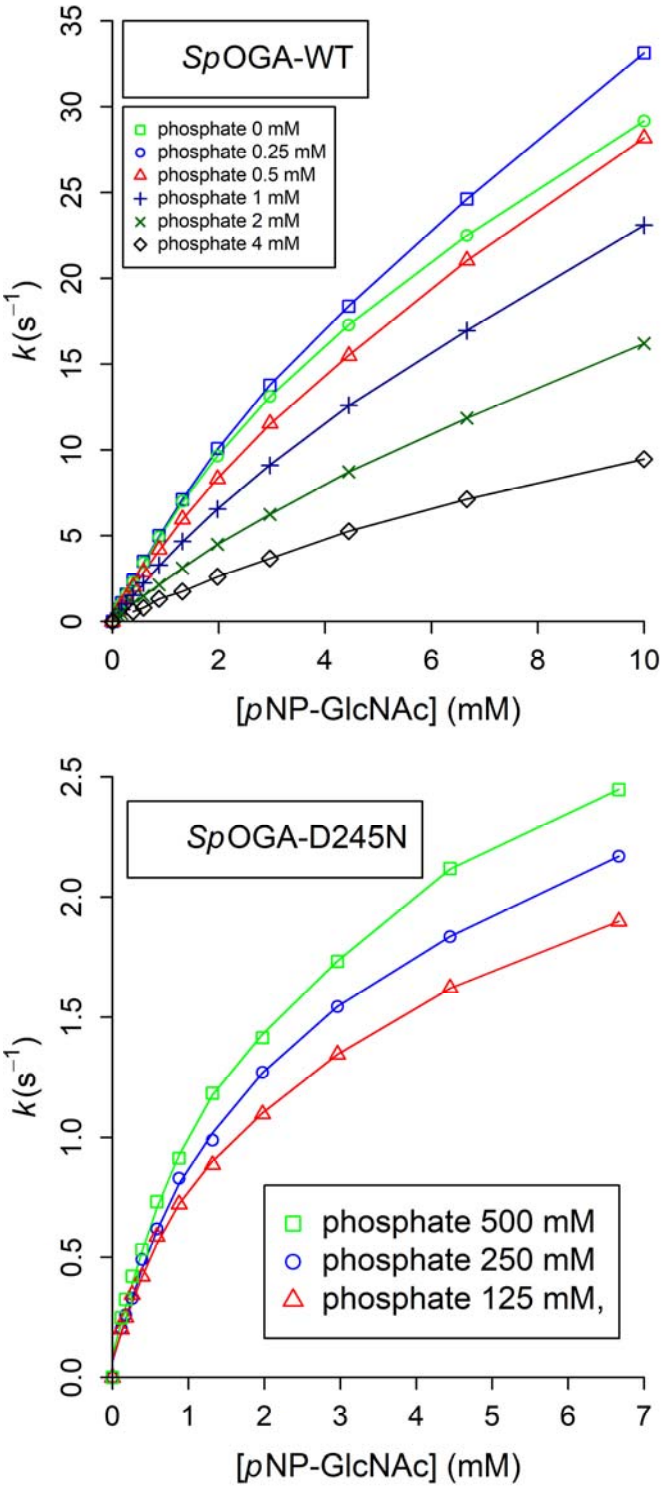

Figure S6. Kinetics of *p*NP release by *Sp*OGA-WT and *Sp*OGA-D245N in function of phosphate concentration, at pH 7.8, 25°C. Top, *Sp*OGA-WT inhibition by phosphate. Bottom, *Sp*OGA-D245N activation by phosphate.

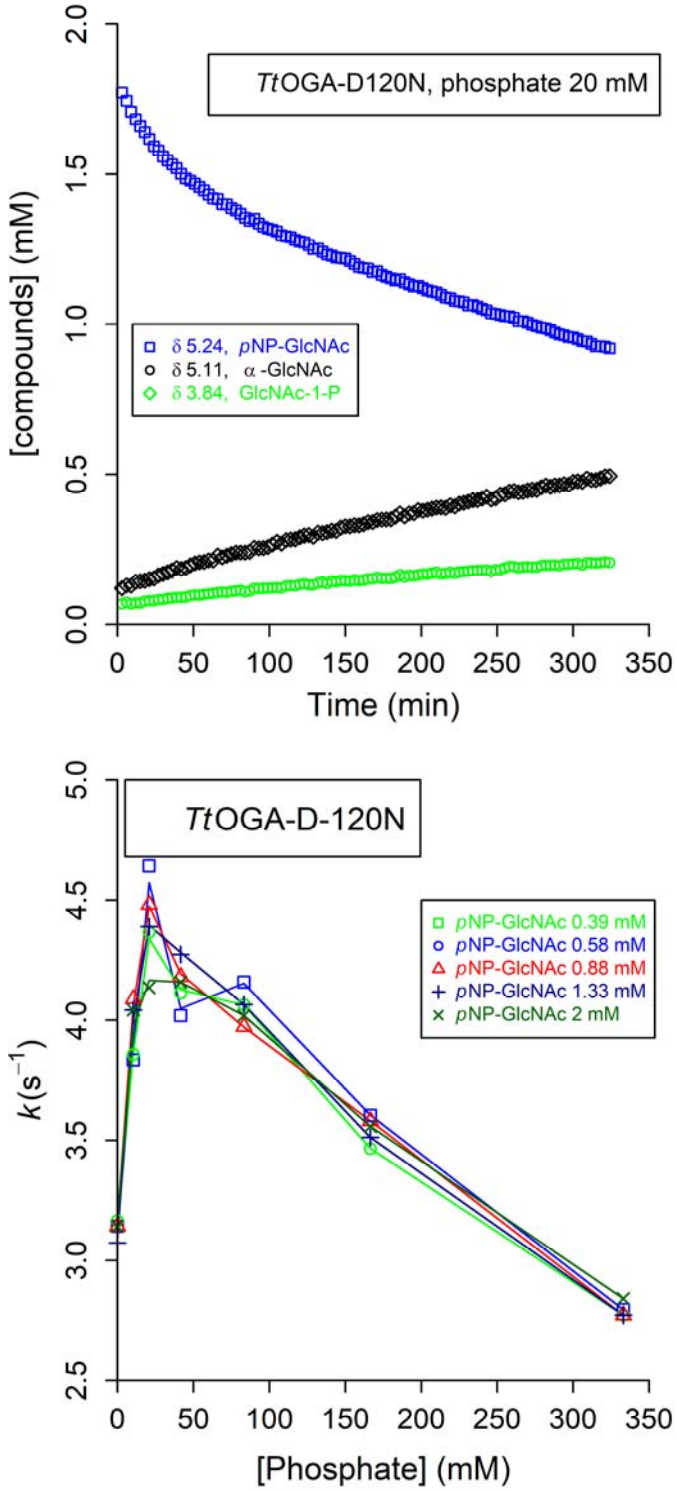

Figure S7. *Tt*OGA-D120N. Top, *p*NP-GlcNAc hydrolysis and phosphorylation by 40 µg *Tt*OGA-D120N. Bottom, kinetics of *p*NP release by *Tt*OGA-D120N as function of phosphate concentration, at pH 7.8, 25°C.

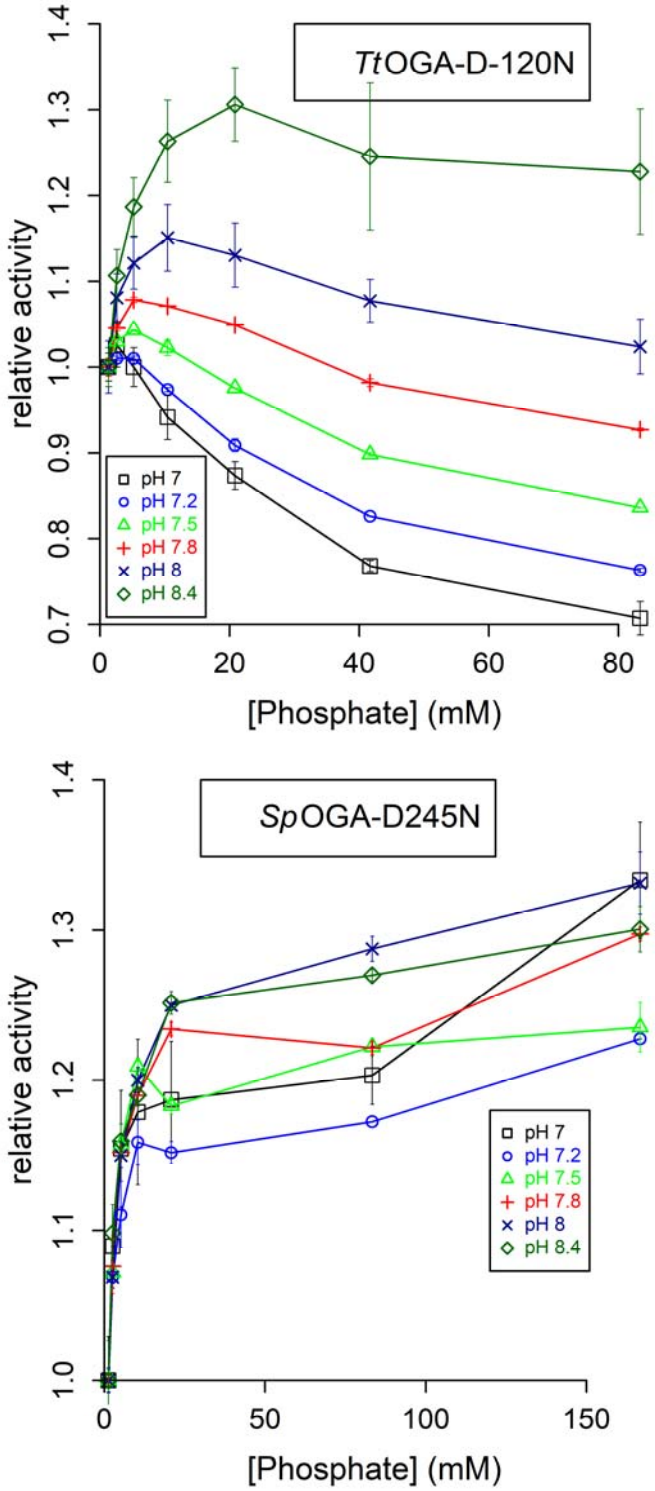

Figure S8. pH influence on phosphate activation of *p*NP-release in presence of 2 mM *p*NP-GlcNAc at 25 °C. Top, *Tt*OGA-D120N; bottom, *Sp*OGA-D245N.

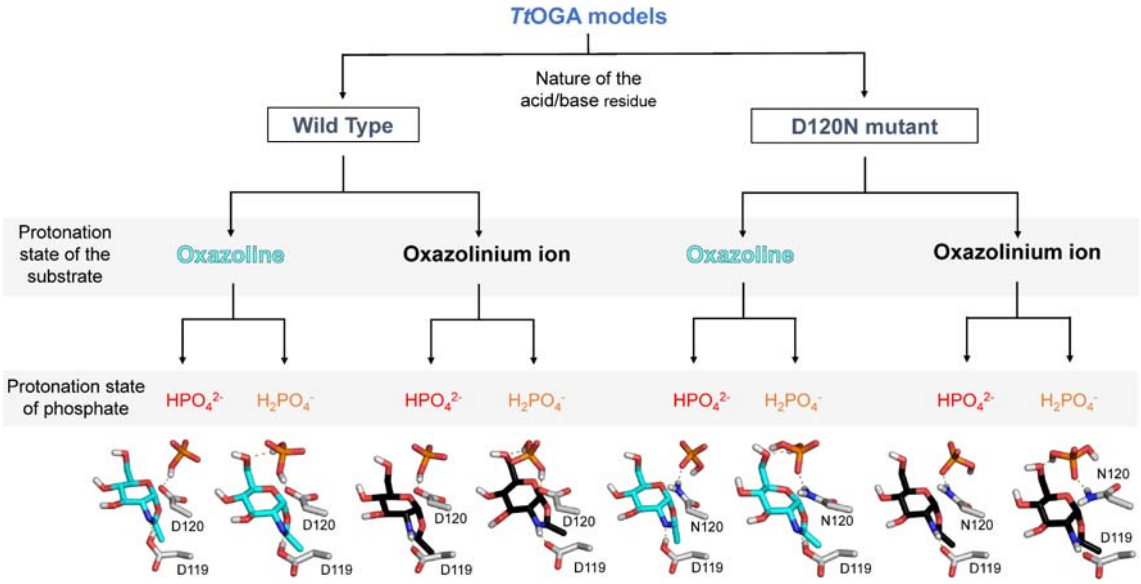

Figure S9. Schematic representation of the models for the *Tt*OGA enzyme considered in this work, depending on the nature of the acid/base residue (wild-type or D120N mutant), and the protonation state of the intermediate/phosphate anion (oxazoline or oxazolinium ion/HPO_4^- and $H_2PO_4^{2-}$).

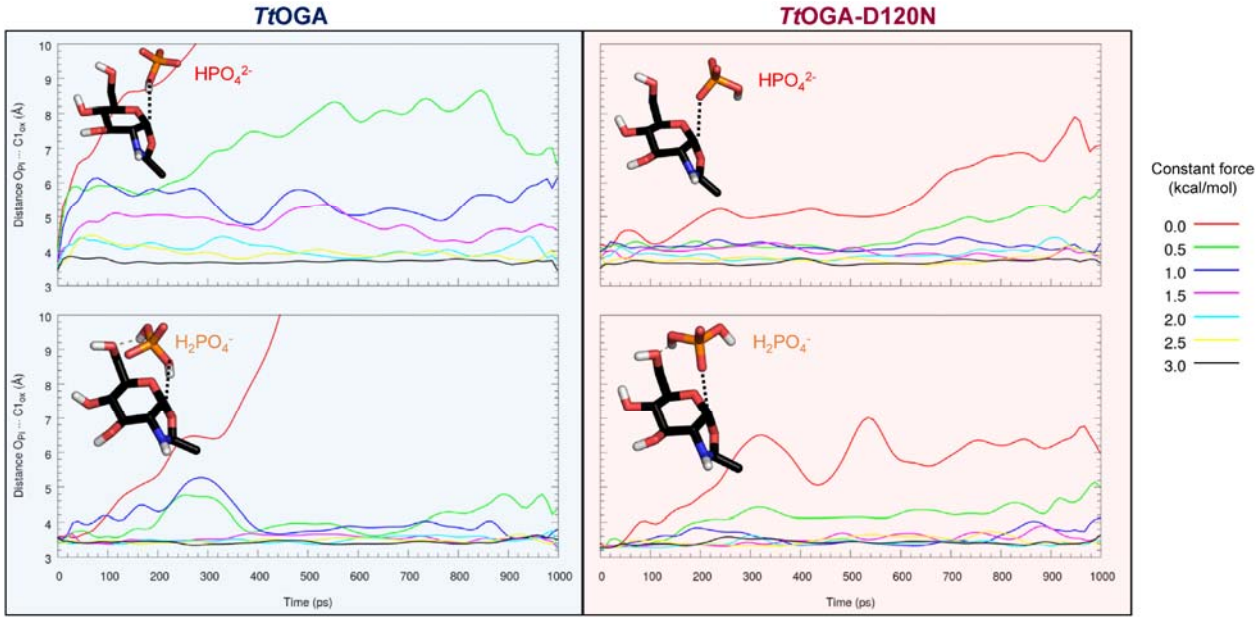

Figure S10. Evolution of the phosphate-anomeric carbon distance during the restrained MD simulations of *Tt*OGA and *Tt*OGA-D120N model complexes. Different values for the constant force of the restraint between anomeric carbon (C1) and the oxygen of the phosphate are used. Results are averaged over five MD replicas that were performed for each applied restrain.

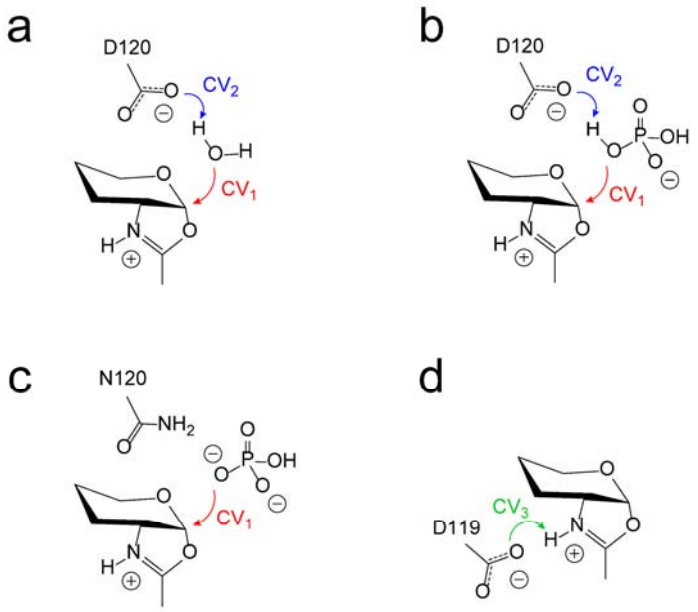

Figure S11. Collective variables (CVs) used in this work to study: **a)** hydrolysis, **b)** phosphorylation by wild-type *Tt*OGA, **c)** phosphorylation by *Tt*OGA-D120N, and **d)** the protonation state of the bicyclic reaction intermediate in wild-type *Tt*OGA.

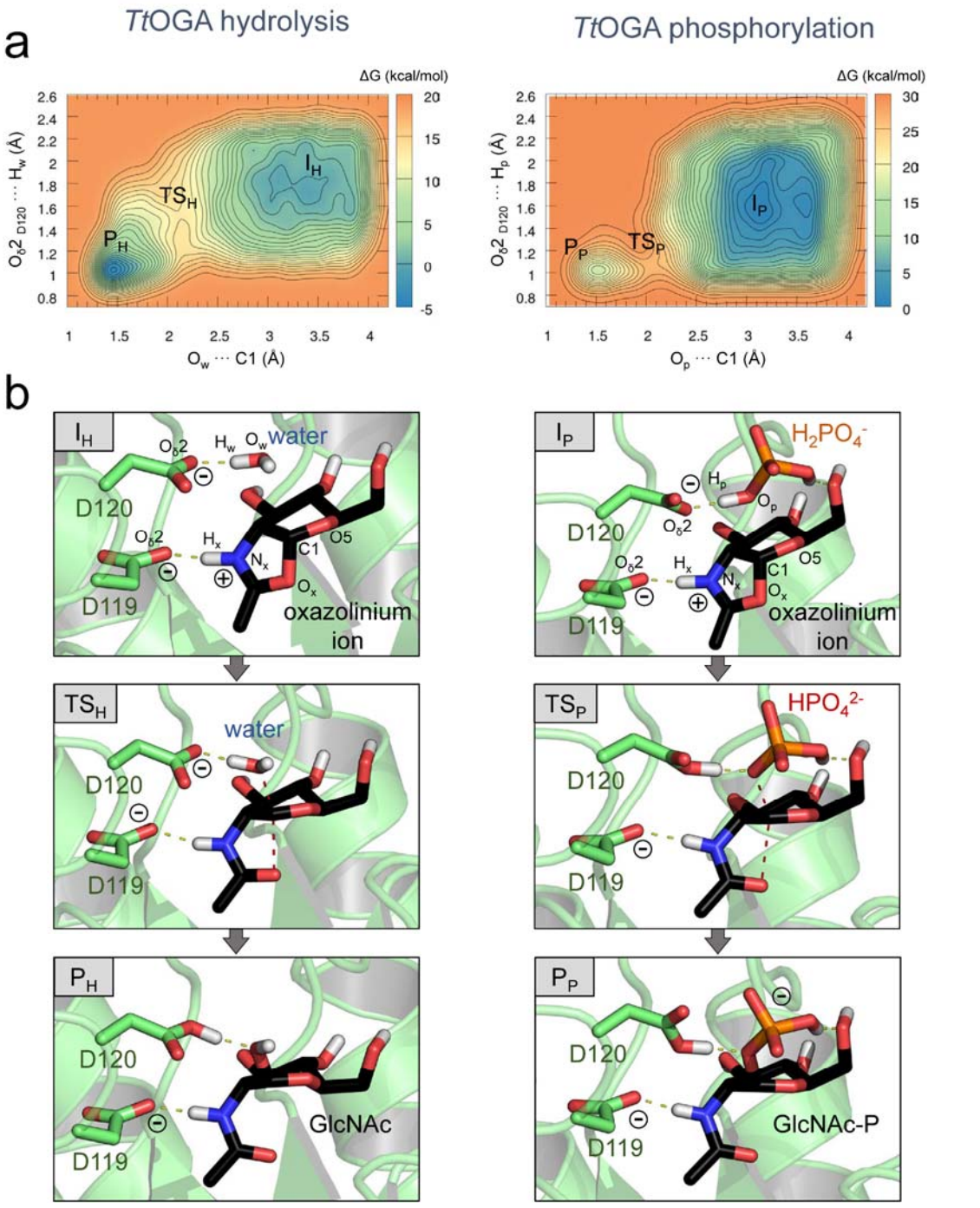

Figure S12. Glc-ox$^+$ hydrolysis (left) and phosphorylation (right) by wild-type *Tt*OGA. **a)** Free energy surfaces reconstructed from the metadynamics simulations of hydrolysis (left) and phosphorylation (right) reactions of Glx-ox$^+$ by *Tt*OGA. Isolines are at 1 kcal·mol^{-1}. The two employed CVs account for the deprotonation of either water or phosphate by D120, and the nucleophilic attack of either the water or phosphate oxygens on C1, respectively. **b)** Average structures of the main states along the reaction coordinate of Glc-ox$^+$ hydrolysis (left) and phosphorylation (right). Hydrogen atoms have been omitted for clarity, except those attached to heteroatoms. Atom labels are represented in the intermediate state. **I** = Glc-ox$^+$ intermediate, **TS** = transition state, **P** = products. **H** and **P** subscripts correspond to hydrolysis and phosphorylation, respectively. Hydrogen bonds and ionic interactions are indicated with yellow dashed lines, whereas bonds being formed/broken are indicated with red dashed lines.

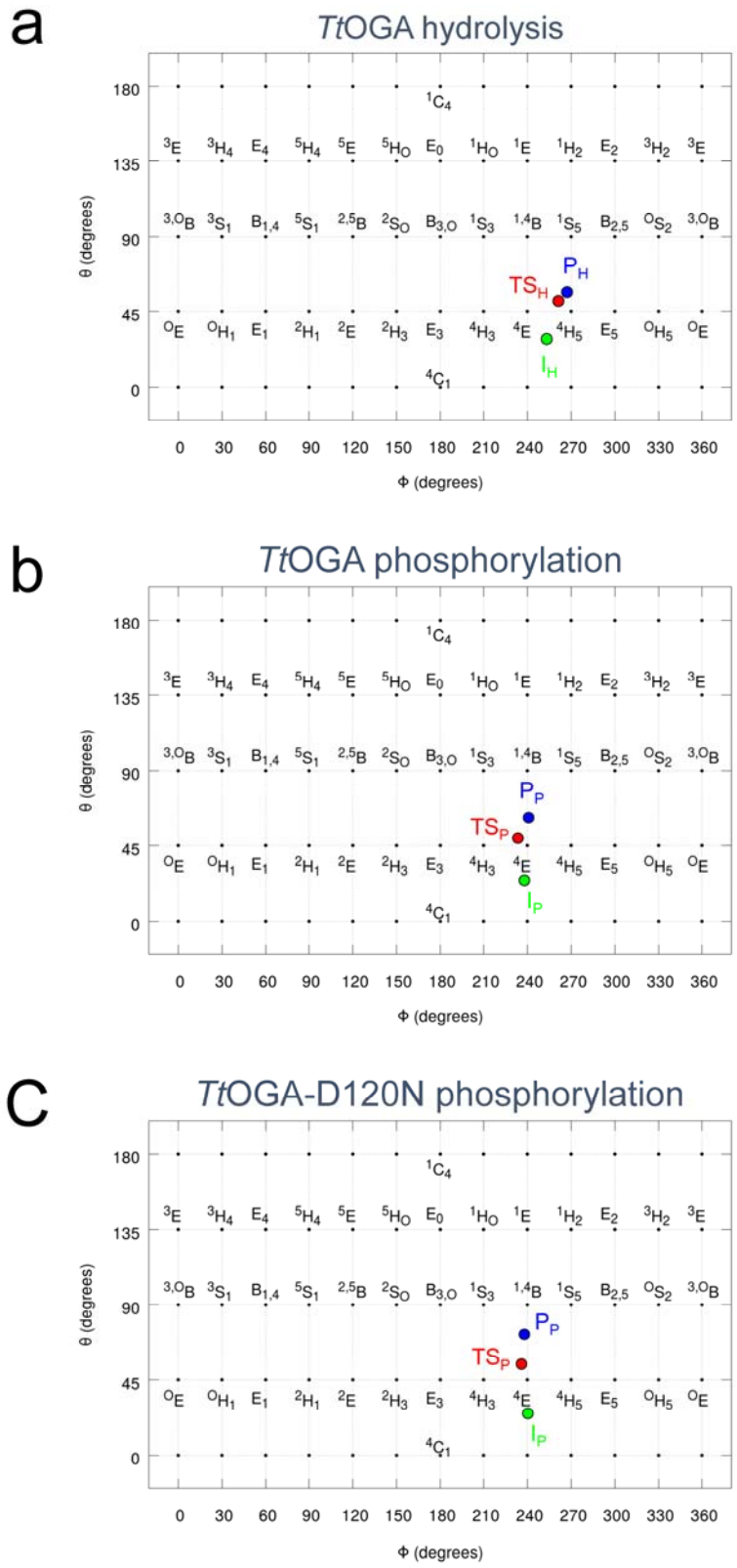

Figure S13. Mercator representation of the puckering coordinates calculated for the averaged structures of the main states along the enzymatic reactions: **a)** hydrolysis, **b)** phosphorylation by wild-type *Tt*OGA, and **c)** phosphorylation by *Tt*OGA-D120N. **I** = Glc-ox^{+} intermediate, **TS** = transition state, **P** = products. **H** and **P** subscripts correspond to hydrolysis and phosphorylation, respectively.

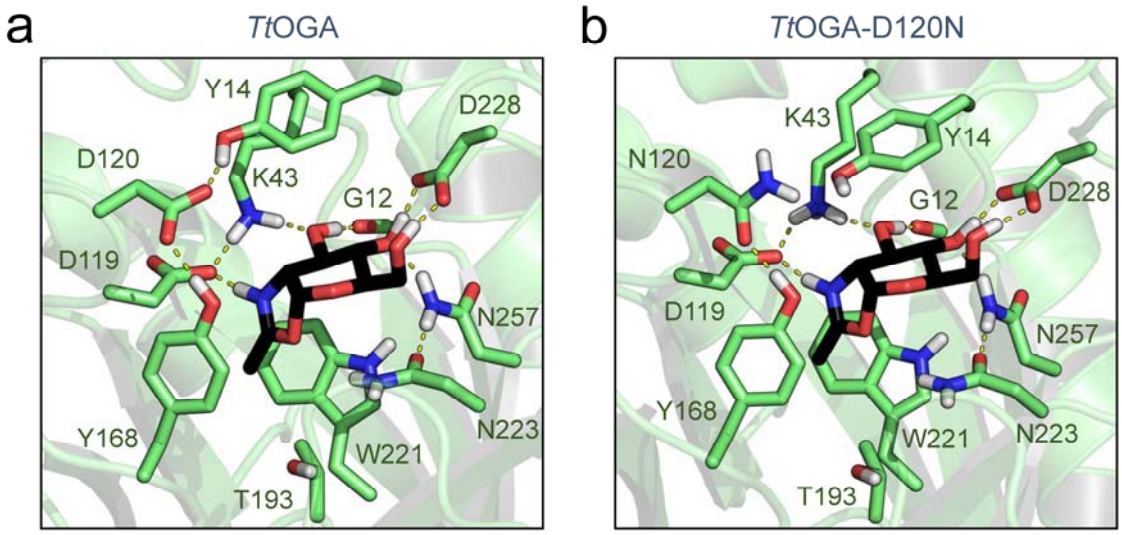

Figure S14. Active site architecture of **a)** *Tt*OGA wild-type enzyme and **b)** *Tt*OGA-D120N mutant. The main interactions between the enzyme and the Glc-ox^{+} substrate are shown. Side chains of Y168, T193 and W221 provide a hydrophobic cavity for the methyl group of the substrate. The OH3 of oxazolinium ion interacts with K43 and the backbone carbonyl of G12, while OH4 and OH6 interacts by hydrogen bonding with both oxygens of the carboxylate group of D228. The assisting residue D119 establishes ionic interactions with the nitrogen of the oxazolinium ion ring and the side chain of K43. The main difference between *Tt*OGA and *Tt*OGA-D120N enzymes lies in the environment of the acid/base residue (D120 and N120, respectively). **a)** In the wild-type *Tt*OGA, the acid/base residue D120 is oriented for catalysis and stabilized by the phenolic hydroxyls of both the Y14 and Y168 residues. **b)** In *Tt*OGA-D120N, N120 interacts only with Y168. The NH group of the amide N120 can form a hydrogen bond bonding with the incoming HPO_4^{2-}.

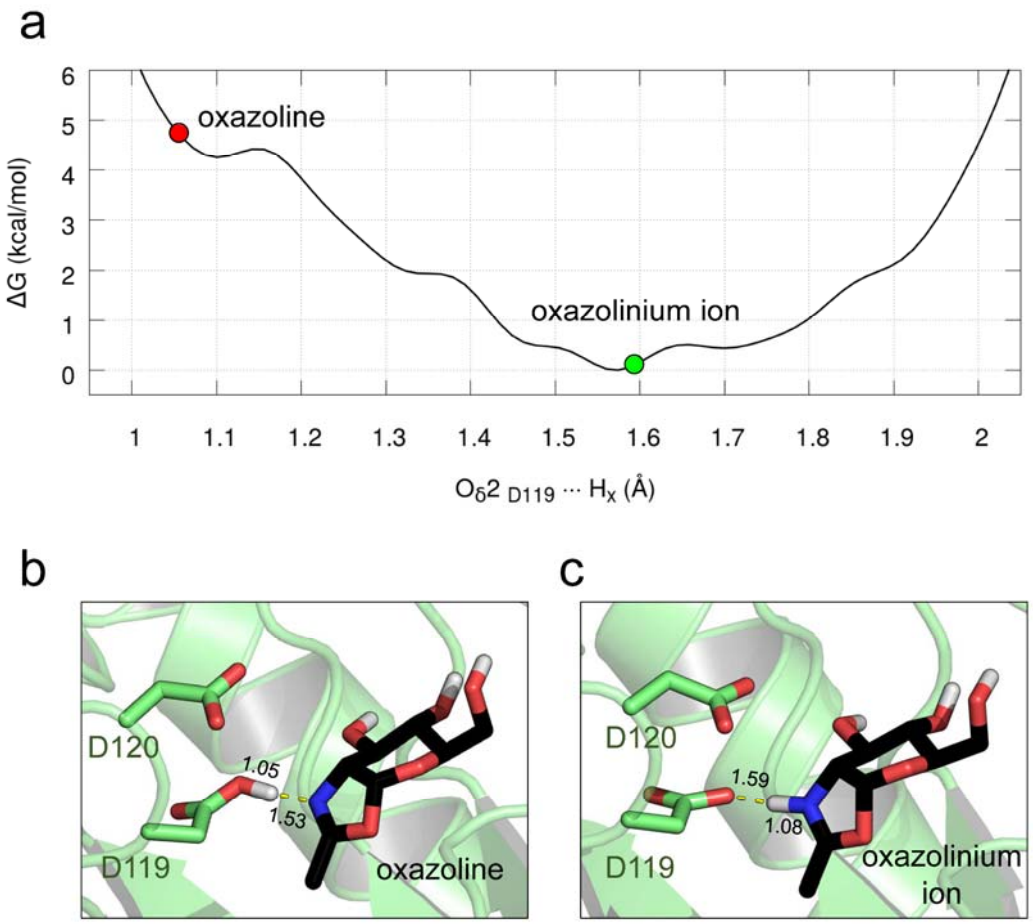

Figure S15. Analysis of the protonation state of the reaction intermediate. **a)** Free energy profile reconstructed from a metadynamics simulation using as a CV the distance between the oxygen of D119 and the hydrogen from the oxazoline ring nitrogen. **b)** Initial structure for geometry optimization: the proton is bound to D119, the oxazoline is neutral. **c)** The optimized structure, in which the proton has transferred to the N atom.

5. Tables S1-S6

Table S1. Computational set up for all systems analyzed in this work.

Simulation type	Enzyme	Reaction intermediate	HPO_4^-/ $H_2PO_4^{2-}$/Na^+	Total nº atoms	Production run (ns)
Analysis of phosphate distribution	WT	Glc-ox	9/2/35	93555	50
	D120N	Glc-ox	9/2/34	93559	50
Reaction coordinate analysis	WT	Glc-ox	1/0/24	93553	20
	WT	Glc-ox	0/1/23	93559	20
	WT	Glc-ox$^+$	1/0/24	93553	20
	WT	Glc-ox$^+$	0/1/23	93559	50
	D120N	Glc-ox	1/0/23	93557	20
	D120N	Glc-ox	0/1/22	93560	20
	D120N	Glc-ox$^+$	1/0/23	93557	50
	D120N	Glc-ox$^+$	0/1/22	93560	20

Table S2. QM/MM MD parameters for the three systems studied in the present work.

System	Total nº atoms	QM nº atoms	Box (Å)	NN/ MIX/ESP radii (Å)	Timestep (fs)	Electronic mass (a.u.)
*Tt*OGA + H_2O	93561	45	14.26×14.17×18.45	9.00/13.23/15.88	0.12	600
*Tt*OGA + $H_2PO_4^{2-}$	93561	49	15.45×15.56×18.33	7.94/10.58/15.88	0.12	600
*Tt*OGA-D120N + HPO_4^-	93559	50	13.26×17.08×18.22	6.35/10.58/15.88	0.12	600

Table S3. Metadynamics parameters for hydrolysis/phosphorylation by *Tt*OGA and phosphorylation by *Tt*OGA-D120N mutant.

System	Collective variable (CVs)	Gaussian height (kcal·mol^{-1})	Gaussian width (Å)	Deposition time (MD steps)
*Tt*OGA + H_2O	$CV_1 = O_w$-C1; $CV_2 = H_w$-O_{D120}	0.6	0.10; 0.10	300
*Tt*OGA + $H_2PO_4^{2-}$	$CV_1 = O_p$-C1; $CV_2 = H_p$-O_{D120}	1.0	0.10; 0.10	400
*Tt*OGA-D120N + HPO_4^-	$CV_1 = O_p$-C1	0.6	0.10	400

Table S4. Values of the most relevant active site distances during the hydrolysis of Glc-ox$^+$ by *Tt*OGA.

Distance	Reaction intermediate	Transition state	Product
C1-O_w	3.39 ± 0.05	2.10 ± 0.05	1.46 ± 0.02
C1-O_x	1.60 ± 0.07	2.40 ± 0.23	3.05 ± 0.26
C1-O5	1.35 ± 0.03	1.32 ± 0.03	1.43 ± 0.05
O_w-H_w	1.01 ± 0.03	1.07 ± 0.02	1.60 ± 0.06
H_w-$O_\delta 2_{D120}$	1.75 ± 0.06	1.53 ± 0.03	1.03 ± 0.02
H_x-$O_\delta 2_{D119}$	1.58 ± 0.10	1.90 ± 0.28	1.80 ± 0.03
H_x-N_x	1.08 ± 0.05	1.03 ± 0.06	1.06 ± 0.03
O_x- $H_\delta 2_{N223}$	2.78 ± 0.28	2.11 ± 0.11	2.30 ± 0.20
$H_{\eta Y14}$- $O_\delta 2_{D120}$	1.84 ± 0.12	1.84 ± 0.08	1.93 ± 0.10
$H_{\eta Y168}$- $O_\delta 1_{D120}$	1.88 ± 0.16	1.84 ± 0.12	1.91 ± 0.12

Table S5. Relevant distances involved in the phosphorylation of Glc-ox$^+$ by *Tt*OGA.

Distance	Reaction intermediate	Transition state	Product
C1-O_p	3.15 ± 0.02	2.10 ± 0.04	1.51 ± 0.02
C1-O_x	1.63 ± 0.07	2.59 ± 0.08	2.85 ± 0.07
C1-O5	1.34 ± 0.03	1.29 ± 0.03	1.40 ± 0.03
O_p-H_p	1.05 ± 0.04	1.53 ± 0.12	1.96 ± 0.09
H_p-$O_\delta 2_{D120}$	1.56 ± 0.04	1.05 ± 0.01	1.03 ± 0.02
H_x-$O_\delta 2_{D119}$	1.60 ± 0.14	1.94 ± 0.13	1.79 ± 0.18
H_x-N_x	1.09 ± 0.03	1.05 ± 0.04	1.03 ± 0.02
O_x- $H_\delta 2_{N223}$	2.34 ± 0.22	1.91 ± 0.07	1.87 ± 0.10
$H_{\eta Y14}$- $O_\delta 2_{D120}$	2.61 ± 0.35	2.64 ± 0.12	2.75 ± 0.13
$H_{\eta Y168}$- $O_\delta 1_{D120}$	2.98 ± 0.69	3.08 ± 0.47	3.22 ± 0.12

Table S6. Relevant distances involved in the phosphorylation of Glc-ox$^+$ by *Tt*OGA-D120N.

Distance	Reaction intermediate	Transition state	Product
C1-O_p	3.40 ± 0.08	2.17 ± 0.01	1.47 ± 0.03
C1-O_x	1.66 ± 0.09	2.55 ± 0.12	3.07 ± 0.13
C1-O5	1.34 ± 0.03	1.29 ± 0.02	1.40 ± 0.04
O_p-$H_{\eta Y14}$	1.94 ± 0.11	2.03 ± 0.21	2.12 ± 0.16
O_p-$H_\delta 2_{N120}$	1.95 ± 0.33	2.95 ± 0.19	3.11 ± 0.23
H_x-$O_\delta 2_{D119}$	1.68 ± 0.15	2.19 ± 0.13	2.38 ± 0.27
H_x-N_x	1.07 ± 0.04	1.04 ± 0.03	1.03 ± 0.03
O_x-$H_\delta 2_{N223}$	2.59 ± 0.38	2.11 ± 0.14	2.01 ± 0.14
$H_{\eta Y168}$- $O_\delta 2_{N120}$	1.92 ± 0.15	1.81 ± 0.09	1.89 ± 0.11
$O_p 2$-$H_\delta 2_{N120}$	3.63 ± 0.27	2.30 ± 0.71	2.37 ± 0.49