COMMUNICATION

Cysteine Nucleophiles in Glycosidase Catalysis: Application of a Covalent β-L-Arabinofuranosidase Inhibitor

Nicholas G.S. McGregor‡,[a] Joan Coines‡,[b] Valentina Borlandelli,[c] Satoko Amaki,[d] Marta Artola,[c] Alba Nin-Hill,[b] Daniël Linzel,[c] Chihaya Yamada,[d] Takatoshi Arakawa,[d] Akihiro Ishiwata,[e] Yukishige Ito,[e][f] Gijsbert A. van der Marel,[c] Jeroen D.C. Codée,[c] Shinya Fushinobu,[d] Herman S. Overkleeft,[c] Carme Rovira,*[b][g] Gideon J. Davies*[a]

[a] Dr. Nicholas G.S. McGregor, Prof. Gideon J. Davies, FMedSci, FRS
York Structural Biology Laboratory, Department of Chemistry
The University of York
Heslington, York, YO10 5DD, United Kingdom
E-mail: gideon.davies@york.ac.uk

[b] Dr Joan Coines, Alba Nin-Hill, Prof. Carme Rovira
Departament de Química Inorgànica i Orgànica (Secció de Química Orgànica) & Institut de Química Teòrica i Computacional (IQTCUB)
Universitat de Barcelona
Martí i Franquès 1, 08028 Barcelona, Spain
E-mail: c.rovira@ub.edu

[c] Valentina Borlandelli, Dr. Marta Arola, Daniël Linzel, Prof. Gijsbert A. van der Marel, Dr. Jeroen D.C. Codée, Prof. Herman S. Overkleeft
Leiden Institute of Chemistry
Leiden University
Einsteinweg 55, 2300 RA Leiden, The Netherlands

[d] Satoka Amaki, Dr. Chihaya Yamada, Dr. Takatoshi Arakawa, Prof. Shinya Fushinobu
Department of Biotechnology
The University of Tokyo
Bunkyo-ku, Tokyo 113-8657, Japan

[e] Akihiro Ishiwata, Yukishige Ito
Synthetic Cellular Chemistry Laboratory, RIKEN, 2-1 Hirosawa
Wako, Saitama 351-0198, Japan

[f] Yukishige Ito
Graduate School of Science, Osaka University
Toyonaka, Osaka 560-0043, Japan

[g] Prof. Carme Rovira
Institució Catalana de Recerca i Estudis Avançats (ICREA)
Passeig Lluís Companys 23, 08020 Barcelona, Spain

‡ Equal contribution. Supporting information for this article is given via a link at the end of the document.

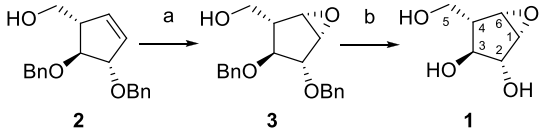
Abstract: The recent discovery of zinc-dependent retaining glycoside hydrolases (GHs), with active sites built around a $Zn(Cys)_3(Glu)$ coordination complex, has presented unresolved mechanistic questions. In particular, the proposed mechanism, depending on a Zn-coordinated cysteine nucleophile and passing through a thioglycosyl enzyme intermediate, remains controversial primarily due to the expected stability of the intermediate C-S bond. To facilitate the study of this atypical mechanism, we report the synthesis of a cyclophellitol-derived β-L-arabinofuranosidase inhibitor, hypothesised to react with the catalytic nucleophile to form non-hydrolysable adduct analogous to the mechanistic covalent intermediate. This β-L-arabinofuranosidase inhibitor reacts exclusively with the proposed cysteine thiol catalytic nucleophiles of representatives of GH families 127 and 146. X-ray crystal structures determined for the resulting adducts enable MD and QM/MM simulations, which provide insight into the mechanism of thioglycosyl enzyme intermediate breakdown. Leveraging the unique chemistry of cyclophellitol derivatives, the structures and simulations presented here support the assignment of a zinc-coordinated cysteine as the catalytic nucleophile and illuminate the finely tuned energetics of this remarkable metalloenzyme clan.

First identified in 2014 and 2018 respectively, glycoside hydrolase families[1] 127 and 146 share a common active site structure centred around an unusual $Zn(Cys)_3(Glu)$ coordination complex.[2,3] Though hydrolysis of β-L-arabinofuranosides was shown to be stereochemically retaining[4], the structures of the active site and substrate preclude all known retaining mechanisms, including the classical Koshland double displacement mechanism[5,6], the NAD^+-dependent mechanism[7], the neighbouring group participation mechanism[8,9], and the recently described 1,2-epoxide intermediate mechanism[10].

The first crystallographic complex of a GH127 enzyme with L-arabinofuranose was solved in 2014.[3] This structure revealed that *Bifidobacterium longum* GH127 (HypBA1) recognizes L-arabinofuranose with the anomeric carbon positioned adjacent to C417, as if hydrolysed from a thioglycosidic linkage. A structure solved for *Bacteroides thetaiotaomicron* GH146 (*Bt*GH146) found similar positioning of L-arabinofuranose adjacent to C416, homologous to C417 in HypBA1.[2]

Site-directed mutagenesis of HypBA1 supported the status of C417 as the catalytic nucleophile, but was confounded by the profound loss of activity resulting from mutagenesis of any residue forming part of the $Zn(Cys)_3(Glu)$ complex.[3] Theoretical analysis suggested that a previously proposed hydrolytic mechanism passing through a thioglycosyl enzyme intermediate (tGEI) is plausible. However, the tGEI was found to be more stable than the product complex, and experimental evidence for the formation

of this intermediate has remained lacking. To determine experimentally if the enzyme possesses a viable catalytic nucleophile, and to enable better computational study, we prepared a cyclophellitol-derived inhibitor.[11,12] Cyclophellitol and its derivatives are glycoside mimics which react with the catalytic nucleophiles of retaining GHs to form non-hydrolysable adducts analogous to the mechanistic covalent intermediate.



Scheme 1. Synthesis of β-L-arabinofuranosyl-configured epoxide **1**. Reagents and conditions: a) *m*-CPBA, H_2NaPO_4/HNa_2PO_4, CH_2Cl_2, 4 °C, 4 days, 74%; b) H_2, $Pd(OH)_2$, MeOH, r.t., 19 h, 18%.

The synthesis of β-L-arabinofuranose-configured epoxide **1** relied on the stereoselective epoxidation of key cyclopentene **2**.[13] This was achieved by exploiting the directing effect of the neighbouring primary alcohol through hydrogen bonding with *m*-CPBA at low temperature (Scheme 1).[13] Final β-L-arabinofuranosyl epoxide **1** was obtained by hydrogenation of partially benzylated epoxide **3** with Pearson's catalyst.

To determine whether **1** had the capacity to label the $Zn(Cys)_3(Glu)$ active site, the covalent inactivation of HypBA1 (GH127) by **1** was followed by intact mass spectrometry and residual activity measurement. Samples taken following 0, 10, 60, and 1300 minutes of incubation with 0.1 mM epoxide **1** at 37°C in 50 mM pH 4.5 NaOAc buffer confirmed time-dependent single labelling, giving a species with a mass difference of 146 Da (Figure 1A). The ~55% labelling measured following 10 minutes of incubation allows estimation of a performance constant (k_i/K_I) of ~20 $M^{-1}s^{-1}$ (assuming K_I>>0.1 mM). This was in good agreement with the inhibition performance constant determined via residual activity measurements (k_i/K_I = 14.5±0.7 $M^{-1}s^{-1}$, figure S1), confirming that labelling was concordant with enzyme inhibition. The measured performance constant for the interaction between HypBA1 and compound **1** is comparable to the inhibition by β-D-xylo-configured cyclophellitol epoxide of *Anid*XlnD (GH3) (170 $M^{-1}s^{-1}$) or the inhibition by α-L-arabinofurano-configured cyclophellitol aziridine of *An*AbfA (GH51, 39 $M^{-1}s^{-1}$) or *Ak*AbfB (GH54, 28 $M^{-1}s^{-1}$).[13,14] An activity assay using the synthetic 4-nitrophenyl-β-L-arabinofuranoside substrate also confirmed a loss of activity over time in the crystallised enzyme sample (Figure S2). This result demonstrates that the electrophilic addition typical of appropriately configured cyclophellitol derivatives can occur efficiently within the $Zn(Cys)_3(Glu)$ active site, inhibiting enzymatic activity.

To determine whether labelling occurs exclusively at the putative catalytic nucleophile (C417), A search of MS/MS spectra collected from the tryptic digest of HypBA1 labelled with **1** for 1 hour, identified 458 MS/MS spectra which corresponded to peptides derived from HypBA1 (78% coverage, FDR 0.9%). Labelling with **1** was detected exclusively at C417 (Figure 1B). Furthermore, intact MS was run under identical conditions with HypBA1 C417S, giving no detectable labelling after 1300 minutes (Figure 1A).

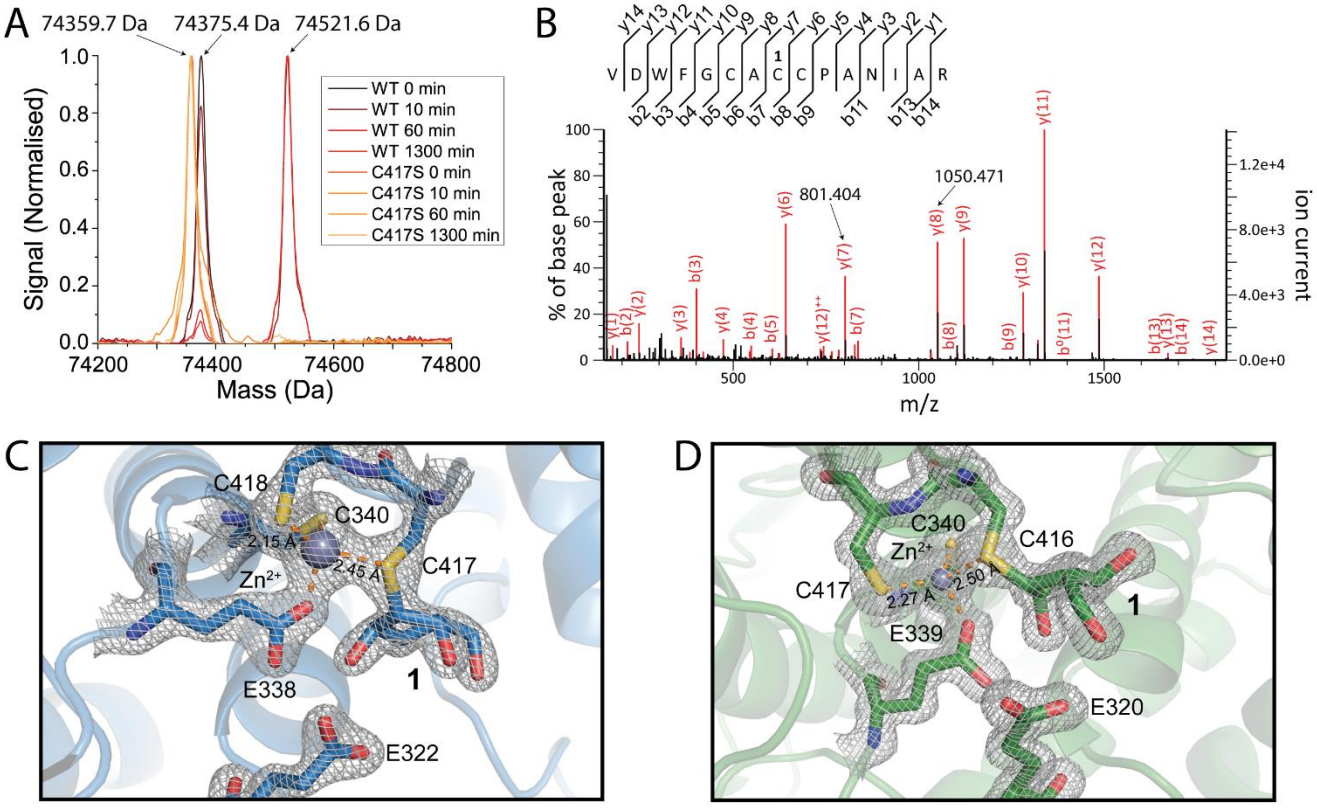


Figure 1. Labelling of β-L-arabinofuranosidases by **1**. A) Overlay of intact MS spectra collected for HypBA1 and HypBA1 C417S following different incubation periods with **1**. Expected mass (native) = 74376.5 Da, with **1** = 74522.6 Da, C417S = 74360.5 Da. B) Active site peptide MS/MS spectra from **1**-labelled HypBA1. m/z values for the major diagnostic y_7 and y_8 fragment signals are given (see table S2 for all y and b fragment m/z values). C) Active site structure of HypBA1 bound to **1**. The $2F_o$-F_c map is shown contoured to 2σ for key residues. The bonds coordinating the zinc ion are shown as orange dashed lines with key bond lengths shown. D) Active site structure of *Bt*GH146 following reaction with **1** with density and coordination shown as in panel C.

Similar experiments were run with BtGH146, a recently characterized GH146 β-L-arabinofuranosidase from a distinct phylogenetic clade of retaining glycoside hydrolases, which also possess a $Zn(Cys)_3(Glu)$ active site.[2] Samples taken for intact MS following 0, 60, 240, and 1200 minutes of incubation with 0.1 mM epoxide **1** at 37°C in 20 mM HEPES pH 7.5 confirmed similar, though slower, labelling (Figure S3A). Digestion of the labelled enzyme followed by LC-MS/MS of the resulting peptides confirmed exclusive labelling of the putative active site peptide at C416 (homologous to C417 in HypBA1, Figure S3B). Thus, the reactivity of **1** exclusively with the putative catalytic cysteine nucleophile appears to be a general phenomenon.

Crystallization of the complex between HypBA1 and **1** (crystallographic data and refinement statistics given in Table S1) yielded a structure which contains clear ligand density bound to the $Zn(Cys)_3(Glu)$ cluster through C1 (Figure 1C). Key hydrogen bonding interactions between O2 and both E338 and H270, between O3 and H194, and between O5 and H142 are preserved between the covalent complex with **1** and the previously determined complex with L-arabinose (PDB ID: 3WKX, Figure S4A). Relative to the L-arabinose complex, C1 has undergone electrophilic migration towards C417, forming a covalent bond with this putative nucleophile and shifting the ring conformation to the somewhat higher energy ^4E (Figure S5).

Crystallization of *Bt*GH146 labelled with compound **1** yielded an unexpected structure, in which the C416 sulphur had attacked C6 of the inhibitor (Figure 1D). Comparison to the L-arabinofuranose complex (PDB ID: 5OPJ) shows that binding occurred with a net rotation of the ring by ~75° such that O5 formed a hydrogen bond with E320 instead of E217 (Figure S4B), preserving other key interactions. The epoxide oxygen formed an apparent hydrogen bond with E339, which retained its dative bond to zinc.

To understand the origin and mechanistic relevance of our observed complexes, molecular dynamics (MD) simulations were performed to reconstruct the binding of **1** in unreacted Michaelis complexes. The crystallised complexes between each enzyme and **1** were taken as initial structures for MD simulations (Figures S6-S9),[15] after manually breaking the bond between C417/C416 and the anomeric carbon, and re-building the epoxide group (see methods). The unreacted epoxide was found to be stable in the active site of HypBA1 with E322 protonated and E338 deprotonated (Figure S6A, S10). The sulphur atom of C417 is closest to C1, being at proper distance and orientation to perform a nucleophilic attack (Figure S6C). This explains the formation of the C1-S bond in the covalent complex. The simulations also show that the protonated acid/base side chain is very mobile (Figure S7). Due to the positioning of the epoxide oxygen away from the normal glycosidic oxygen position, the acid/base does not interact with it directly. Instead, it is likely that epoxide protonation by E322 is mediated by a water molecule.

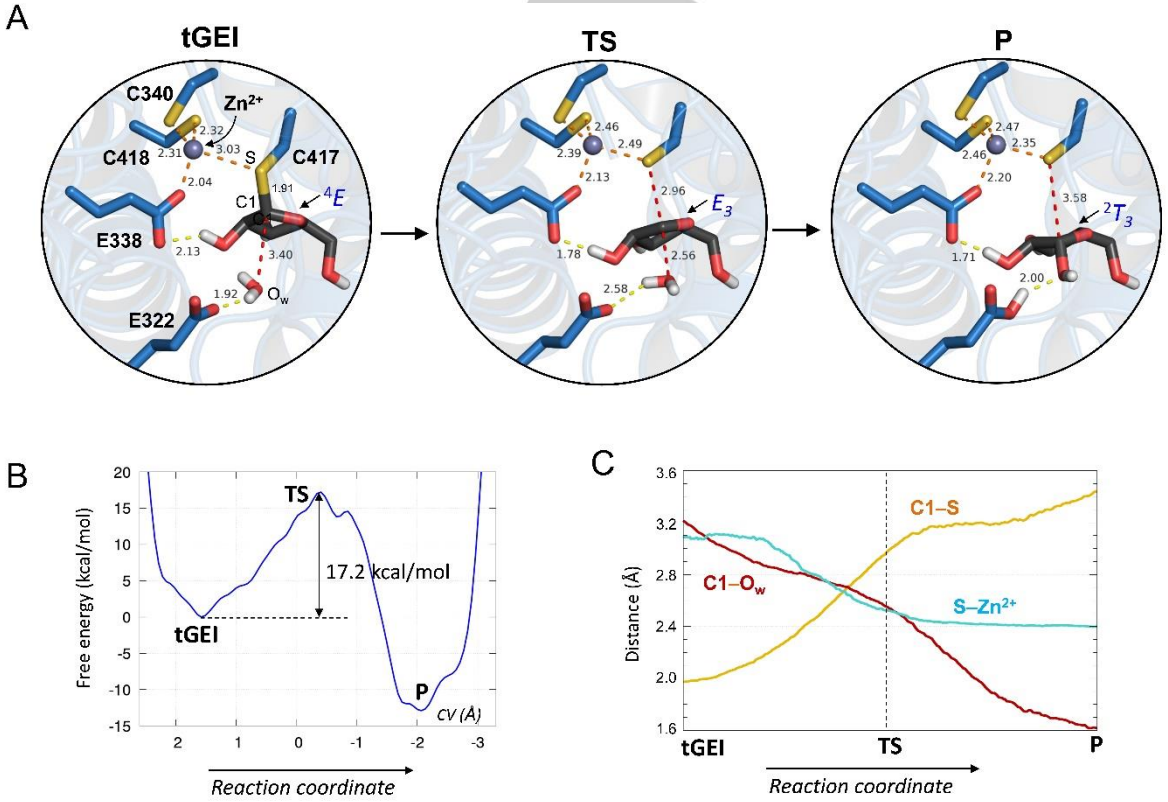


Figure 2. Deglycosylation reaction in HypBA1. A) Active site configuration of the thioglycosyl enzyme intermediate (tGEI), the reaction transition state (TS) and the reaction products (P) obtained from QM/MM metadynamics simulations informed by the X-ray structures obtained in this work. (B) Reaction free energy profile. C) Evolution of relevant distances along the reaction coordinate.

Analysis of the Michaelis complex for *Bt*GH146 reveals a different scenario. The reconstructed Michaelis complex for *Bt*GH146 interacting with **1** showed that this complex is generally unstable unless both E320 and E339 were protonated (Figure S11). While

it is unlikely that protonation of both of these groups would occur at pH 7, this type of binding, in an unstrained E_3 conformation[16,17], shows how this unexpected binding mode favours reactivity with C6 over C1; the sulphur atom of C416 is closer to C6 than to C1 throughout the simulation (Figure S6B,C). Thus, we believe that the observed covalent *Bt*GH146:**1** complex is derived from an atypical binding mode, in which the epoxide binds across the C416-Zn-E339 axis priming C6 for attack by the cysteine nucleophile. We speculate that this is made possible by the lack of steric bulk in the inhibitor structure since any extension from the glycosidic or epoxide oxygen would clash with H264 and E320 in the observed binding position, and thus the epoxide reacts "erroneously"

In the study of the mechanism of $Zn(Cys)_3(Glu)$-dependent glycoside hydrolases, the absence of a covalent intermediate mimic had previously precluded informed computational studies on the controversial deglycosylation step of the reaction. Enabled by the observed complex between **1** and HypBA1, we simulated the deglycosylation step of the enzyme reaction mechanism. We reconstructed the natural tGEI by replacing the CHOH group at C6 of the covalently bound inhibitor (Figure 1C) with an oxygen atom and then equilibrated the complex with MD and QM/MM MD simulations. A large QM region, including the Zn^{2+} cation and its coordination shell was described with density functional theory (DFT), using the Becke–Lee–Yang–Parr (BLYP) functional[18,19], previously used in the study of Zn^{2+}-containing GHs[20], whereas the MM region was described with the Amber force field[21] (further details in Supporting Information). The simulations show that the complex is stable and, most interestingly, the active site accommodates a water molecule bridging the anomeric carbon and the side chain of the (deprotonated) acid/base residue, in a suitable orientation for deglycosylating nucleophilic attack (tGEI in Figure 2A). This is an indication that the cyclophellitol-labelled complex is a *bona fide* mimic of the covalent intermediate of the chemical reaction.

The deglycosylation reaction was modelled using the metadynamics approach, starting from a snap-shot of the equilibration QM/MM MD simulation. The reaction was driven from the tGEI to the product of the reaction (β-L-arabinofuranose) using one collective variable (CV), which combines the main bonds formed and cleaved during the reaction. The CV was defined as the difference of C1-S_{Cys} and C1-O_w distances (O_w is the oxygen atom of the nucleophilic water). Analysis of the free energy profile of the simulated reaction provides an atomistic picture of the HypBA1 catalytic mechanism. The nucleophilic attack of the water molecule takes place via a dissociative transition state, in which the C1-S bond is significantly elongated (changing from 1.91 Å at the tGEI to 2.96 Å at the TS; Figure 2A and 2C) relative to the nucleophilic attack distance (the C1-O_w distance shortening from 3.40 Å at the tGEI to 2.56 Å at the TS). The TS shows a clear oxocarbenium ion-like character: the C1-O5 distance of the β-L-arabinofuranose shrinks from 1.44 Å at the tGEI to 1.31 Å at the TS, and the anomeric charge increases by 0.78 e^-. The β-L-arabinofuranose ring changes conformation during the reaction, following a $^4E \rightarrow [E_3]^{\ddagger} \rightarrow {}^2T_3$ itinerary, in excellent agreement with the conformations observed in the product complex of HypBA1 (PDB 3WKX, see also Figure S5). The computed reaction free-energy profile (Figure 2B) exhibits a single transition state, indicative of a concerted mechanism, with a free energy barrier (17.2 kcal/mol) consistent with that estimated from the experimental rate constant (~16.4 kcal/mol, estimated using the Eyring-Polanyi equation[22] computed with the k_{cat} measured for the reaction with *p*-nitrophenyl-β-Araf,[23] under the assumption that deglycosylation is the rate-limiting step in this reaction). This barrier is also comparable to the 15.5 kcal/mol barrier calculated for the hydrolysis of the canonical glycosyl-enzyme intermediate of *E. coli* LacZ β-galactosidase using similar methods.[24] Sensibly, the computed reaction is exergonic, i.e. the products of the reaction are more stable than the tGEI.

Analysis of the evolution of the C1-S and Zn^{2+}-S distances along the reaction coordinate (Figure 2C) reveals an inverse relationship, reflecting the assistance of the metal cation during the cleavage of the C1-S bond. At the tGEI, C417 is weakly coordinated to the metal cation (Zn^{2+}-S = 3.03 Å). However, the Zn^{2+}-S distance shortens as soon as the C1-S distance starts to elongate until C417 recovers its lowest energy coordination distance (Figure S12). The Zn^{2+}-S distance in the simulated product complex (2.4 Å) is in excellent agreement with the one observed in the X-ray structure of the product complex (2.4 Å, PDB 3WKX). Our results suggest that the presence of the metal ion significantly decreases the nucleophilicity of C417, simultaneously ensuring that the C1-S bond can be efficiently hydrolysed. Altogether, the simulations support the role of C417 as the nucleophile of the reaction catalyzed by HypBA1 and demonstrate that participation of the sulphur nucleophile in the $Zn(Cys)_3(Glu)$ complex is essential for glycoside hydrolysis. They furthermore demonstrate the value of covalent species obtained from cyclophellitol-labeling experiments as starting points to model the reactivity of glycosyl enzyme intermediates.

In summary, here we have presented the first known inhibitor of a β-L-arabinofuranosidase and shown that it covalently modifies GH127 and GH146 active sites to form a tGEI mimic. Structural analysis of prepared complexes revealed an unexpected flexibility in binding mode giving rise to two different reacted structures, which both support the proposed identity of C417 (HypBA1) or C416 (*Bt*GH146), part of the $Zn(Cys)_3(Glu)$ complex, as the catalytic nucleophile. Reconstruction of the tGEI from the coordinates of the complex between HypBA1 and its cyclophellitol-derived inhibitor facilitated analysis of the cleavage mechanism of the C-S bond within the tGEI. QM/MM metadynamics simulations illuminate the finely tuned energetics of this active site, revealing that the Zn^{2+}-S interaction significantly destabilises the C-S bond, facilitating the attack of water onto C1 and driving the reaction towards a lower energy product state.

Acknowledgements

We thank the Natural Sciences and Engineering Research Council of Canada (Post-Doctoral Fellowship to N.G.S.M.), the Spanish Ministry of Science, Innovation and Universities (MICINN/AEI/FEDER, grant CTQ2017-85496-P to C.R. and predoctoral fellowship FPI-BES-2015- 072055 to J.C.), the Spanish Structures of Excellence María de Maeztu (grant MDM-2017-0767 to C.R.), the Agency for Management of University and Research Grants of Generalitat de Catalunya (AGAUR, grant 2017SGR-1189 to C.R. and FI-AGAUR PhD scholarship to A.N.-H), the Royal Society (Ken Murray Research Professorship to

G.J.D.), the Biotechnology and Biological Sciences Research Council (BBSRC) (grant BB/R001162/1 to G.J.D.), the Netherlands Organization for Scientific Research (NWO TOP grant 2018-714.018.002 to H.S.O.), the European Research Council (ERC-2011-AdG-290836 “Chembiosphing” to H.S.O.), JSPS KAKENHI (15H02443, 19H00929 and 26660083 to S.F.), and the Marie Skłodowska-Curie Innovative Training Networks (H2020-MSCA-ITN-2018-814102 “Sweet crosstalk” to H.S.O and C.R.). The authors would like to thank the technical support provided by the Barcelona Supercomputing Center (BCV-2018-3-0014) for computer resources MareNostrum IV and MinoTauro, as well as the computational support from the University of York High Performance Computing service, Viking and the Research Computing team (chem-menz-2019). We thank the Diamond Light Source and KEK-PF for beamtime (proposals 18598 and 2019G017, respectively), and the staff of beamlines I03 and NW12A for assistance with crystal testing and data collection.

Keywords: cyclophellitol, arabinofuranosidase, zinc, glycoside hydrolase, enzyme mechanism, metalloenzyme

[1] V. Lombard, H. Golaconda Ramulu, E. Drula, P. M. Coutinho, B. Henrissat, *Nucleic Acids Res.* **2014**, *42*, D490–D495.

[2] A. S. Luis, J. Briggs, X. Zhang, B. Farnell, D. Ndeh, A. Labourel, A. Baslé, A. Cartmell, N. Terrapon, K. Stott, et al., *Nat. Microbiol.* **2018**, *3*, 210–219.

[3] T. Ito, K. Saikawa, S. Kim, K. Fujita, A. Ishiwata, S. Kaeothip, T. Arakawa, T. Wakagi, G. T. Beckham, Y. Ito, et al., *Biochem. Biophys. Res. Commun.* **2014**, *447*, 32–37.

[4] K. Fujita, Y. Takashi, E. Obuchi, K. Kitahara, T. Suganuma, *J. Biol. Chem.* **2014**, *289*, 5240–5249.

[5] G. Davies, B. Henrissat, *Structure* **1995**, *3*, 853–859.

[6] D. E. Koshland, *Biol. Rev.* **1953**, *28*, 416–436.

[7] V. L. Y. Yip, A. Varrot, G. J. Davies, S. S. Rajan, X. Yang, J. Thompson, W. F. Anderson, S. G. Withers, *J. Am. Chem. Soc.* **2004**, *126*, 8354–8355.

[8] A. C. Terwisscha van Scheltinga, S. Armand, K. H. Kalk, A. Isogai, B. Henrissat, B. W. Dijkstra, *Biochemistry* **1995**, *34*, 15619–15623.

[9] B. L. Mark, D. J. Vocadlo, S. Knapp, B. L. Triggs-Raine, S. G. Withers, M. N. G. James, *J. Biol. Chem.* **2001**, *276*, 10330–10337.

[10] L. F. Sobala, G. Speciale, S. Zhu, L. Raich, N. Sannikova, A. J. Thompson, Z. Hakki, D. Lu, S. Shamsi Kazem Abadi, A. R. Lewis, et al., *ACS Cent. Sci.* **2020**, *6*, 760–770.

[11] L. I. Willems, H. S. Overkleeft, S. I. Van Kasteren, *Bioconjug. Chem.* **2014**, *25*, 1181–1191.

[12] L. Wu, Z. Armstrong, S. P. Schröder, C. de Boer, M. Artola, J. M. Aerts, H. S. Overkleeft, G. J. Davies, *Curr. Opin. Chem. Biol.* **2019**, *53*, 25–36.

[13] N. G. S. McGregor, M. Artola, A. Nin-Hill, D. Linzel, M. Haon, J. Reijngoud, A. Ram, M. N. Rosso, G. A. Van Der Marel, J. D. C. Codeé, et al., *J. Am. Chem. Soc.* **2020**, *142*, 4648–4662.

[14] S. P. Schröder, C. De Boer, N. G. S. McGregor, R. J. Rowland, O. Moroz, E. Blagova, J. Reijngoud, M. Arentshorst, D. Osborn, M. D. Morant, et al., *ACS Cent. Sci.* **2019**, *5*, 1067–1078.

[15] D. A. Case, R. C. Walker, T. E. Cheatham, C. Simmerling, A. Roitberg, K. M. Merz, R. Luo, T. Darden, *Univ. California, San Fr.* **2018**, Univ. California, San Fr.

[16] X. Biarnés, A. Ardèvol, A. Planas, C. Rovira, A. Laio, M. Parrinello, *J. Am. Chem. Soc.* **2007**, *129*, 10686–10693.

[17] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 12562–12566.

[18] A. D. Becke, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[19] C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785–789.

[20] L. Petersen, A. Ardèvol, C. Rovira, P. J. Reilly, *J. Am. Chem. Soc.* **2010**, *132*, 8291–8300.

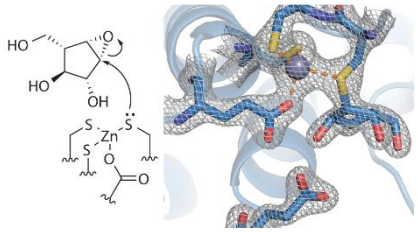
[21] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, C. Simmerling, *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713.

[22] H. Eyring, *J. Chem. Phys.* **1935**, *3*, 63–71.

[23] S. Kaeothip, A. Ishiwata, T. Ito, S. Fushinobu, K. Fujita, Y. Ito, *Carbohydr. Res.* **2013**, *382*, 95–100.

[24] N. F. Bráa, P. A. Fernandes, M. J. Ramos, *J. Chem. Theory Comput.* **2010**, *6*, 421–433.

Entry for the Table of Contents



β-L-arabinofuranosidases are thought to make use of a unique zinc-associated cysteine nucleophile to catalyse glycoside hydrolysis. Using a bespoke synthetic covalent inhibitor, this proposed mechanism is put to the test. Mass spectrometry, X-ray crystallography, and mechanistic simulations support the formation of a thioglycosyl enzyme intermediate and reveal how the enzyme facilitates C-S bond cleavage.

Supplemental Materials
for
Cysteine Nucleophiles in Glycosidase Catalysis: Application of a Covalent β-L-Arabinofuranosidase Inhibitor

Nicholas G.S. McGregor‡,[a] Joan Coines‡,[b] Valentina Borlandelli,[c] Satoko Amaki,[d] Marta Artola,[c] Alba Nin-Hill,[b] Daniël Linzel,[c] Chihaya Yamada,[d] Takatoshi Arakawa,[d] Akihiro Ishiwata,[e] Yukishige Ito,[e][f] Gijsbert A. van der Marel,[c] Jeroen D.C. Codée,[c] Shinya Fushinobu,[d] Herman S. Overkleeft,[c] Carme Rovira,*[b][g] Gideon J. Davies*[a]

[a] Dr. Nicholas G.S. McGregor, Prof. Gideon J. Davies, FMedSci, FRS
York Structural Biology Laboratory, Department of Chemistry
The University of York
Heslington, York, YO10 5DD, United Kingdom
E-mail: gideon.davies@york.ac.uk

[b] Dr Joan Coines, Alba Nin-Hill, Prof. Carme Rovira
Departament de Química Inorgànica i Orgànica (Secció de Química Orgànica) & Institut de Química Teòrica i Computacional (IQTCUB)
Universitat de Barcelona
Martí i Franquès 1, 08028 Barcelona, Spain
E-mail: c.rovira@ub.edu

[c] Valentina Borlandelli, Dr. Marta Arola, Daniël Linzel, Prof. Gijsbert A. van der Marel, Dr. Jeroen D.C. Codée, Prof. Herman S. Overkleeft
Leiden Institute of Chemistry
Leiden University
Einsteinweg 55, 2300 RA Leiden, The Netherlands

[d] Satoka Amaki, Dr. Chihaya Yamada, Dr. Takatoshi Arakawa, Prof. Shinya Fushinobu
Department of Biotechnology
The University of Tokyo
Bunkyo-ku, Tokyo 113-8657, Japan

[e] Akihiro Ishiwata, Yukishige Ito
Synthetic Cellular Chemistry Laboratory, RIKEN, 2-1 Hirosawa
Wako, Saitama 351-0198, Japan

[f] Yukishige Ito
Graduate School of Science, Osaka University
Toyonaka, Osaka 560-0043, Japan

[g] Prof. Carme Rovira
Institució Catalana de Recerca i Estudis Avançats (ICREA)
Passeig Lluís Companys 23, 08020 Barcelona, Spain

‡ Equal contribution.

Materials and Methods

All chemicals were purchased from Sigma Aldrich unless otherwise specified. β-L-arabinofuranose-configured cyclophellitol was synthesized as described in the supplemental synthetic methods and compound characterization section.

Recombinant enzyme production

*Bt*GH146 was synthesized and cloned into pET28a(+) with an N-terminal TEV protease-cleavable His_6 tag by Genscript (Hong Kong). The enzyme was produced in transformed BL21(DE3) gold *E. coli* (Agilent) grown at 37°C in autoinduction medium (1% tryptone, 0.5% yeast extract, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.05% glucose, 0.5% glycerol, 0.2% lactose) overnight. Cells were collected by centrifugation at 4000xg for 15 minutes, resuspended in buffer A (50 mM NaP_i, 300 mM NaCl, 20 mM imidazole, pH 7.5) and lysed by a single pass through a French press. Lysate was clarified by centrifugation at 40000 x *g* for 15 minutes then passed through a 5 mL histrap FF crude column (GE Healthcare). The major UV-active fractions eluted during a 10 CV gradient from 20-500 mM imidazole were pooled, concentrated to ~10 mg/mL using a 30 kDa MWCO centrifugal concentrator and purified using a Superdex 200 column (GE Healthcare) equilibrated with 20 mM HEPES pH 7.5 (SEC buffer). Eluted *Bt*GH146 was pooled, concentrated to ~10 mg/mL and frozen at -80°C.

HypBA1 and HypBA1 C417S were produced and purified as previously described.[1]

Enzyme labelling detected by intact mass spectrometry

*Bt*GH146 was diluted to 1 mg/mL in SEC buffer. Inhibitor **1** was added to a final concentration of 0.1 mM and the reaction was incubated at 37°C. 2 µL samples taken following 60, 240, and 1440 minutes of incubation were diluted with 48 µL of 1% formic acid, 10% acetonitrile and analysed as described previously.[2] HypBA1 and HypBA C417S were diluted to 0.1 mg/mL in 50 mM NaPi pH 4.5 with 0.1 mM **1**. 10 uL samples taken following 10, 60, and 1300 minutes of incubation were mixed with 10 uL of 1% formic acid, 10% acetonitrile and stored at -20°C until intact MS could be analysed in the same manner as the *Bt*GH146 samples.

Enzyme labelling detected by activity measurement

HypBA1 (20 mg/mL stored at -80°C) was thawed and diluted to 1 mg/mL in 50 mM Na-acetate (pH 4.5) supplemented with 1 mM DTT (assay buffer). Inhibitor **1** was dissolved in water at 5 mM and used to prepare a dilution series from 1 mM to 16 µM in assay buffer alongside a no inhibitor buffer control. A working enzyme solution was prepared at 20 µg/mL in assay buffer. To 35 µL of pre-warmed inhibitor solution was added 35 µL of enzyme solution and the inhibition reactions were incubated at 37°C. Following 5, 10, 20, 30, 40, and 60 minutes of incubation, 7.5 µL of the enzyme+inhibitor solution was diluted into 142.5 µL of pre-warmed 0.25 mM 4-nitrophenyl-β-L-arabinofuranoside in assay buffer and incubated at 37°C. 40 µL samples were taken from this substrate hydrolysis reaction following 1, 3, and 8 minutes of incubation and mixed with 40 µL of stop solution (200 µM Na_2CO_3) in a 384-well plate. A_{405} was read using an Epoch plate reader (Biotek). Hydrolysis rates were determined as the slope of a linear fit of A405 vs. time. Slope values were converted into rates using a 40 µL calibration series of 4-nitrophenol in assay buffer mixed with 40 µL of stop solution. To account for slow enzyme activity loss in the no inhibitor control, rates were converted into residual activity by dividing each measured rate by the inhibitor-free hydrolysis rate measured at that incubation point. Residual activities were then plotted against incubation time using OriginPro graphing software (OriginLab) for each inhibitor concentration and fitted with exponential decay curves having y offset (y_0) values fixed at 0 (with the exception of the uninhibited rates where y_0=1). Extracted apparent decay constant (k_{app}) values were then plotted again the concentration of compound **1** with error estimated from the standard fitting error. Since no inflection was observed in the k_{app} vs [I] curve, an error-weighted linear fit was performed to determine k_i/K_I.

HypBA1 Enzyme-Inhibition Kinetics

HypBA1 was diluted to 3.6 mg/mL in 50 mM Na-acetate (pH 4.5) and 150 mM NaCl. Inhibitor **1** was added to a final concentration of 0.1 mM and incubated at 37°C. Aliquots of the labelled protein solution was used for activity assay as described previously.[1,3] The assay mixture consisted of 50 mM Na-acetate (pH 4.5), 10 mM dithiothreitol, 0.2 mM 4-nitrophenyl-β-L-arabinofuranoside, and the enzyme. The hydrolytic reaction was carried out at 37°C for about 8 min. At every 1 min interval, a portion was taken from the assay mixture and mixed with equal volume of stop solution (0.5 M NaOH) in a 384 well plate with flat bottom. Initial velocities were calculated from amounts of liberated 4-nitrophenol, which was measured by the absorbance at 400 nm using a Synergy H1 Hybrid Multi-Mode Microplate Reader (BioTek).

Enzyme trypsinization and LC-MS/MS for modified peptide identification

Two 10 µL samples of 1 mg/mL *Bt*GH146 in SEC buffer were prepared. To one sample was added 1 µL of 1 mM inhibitor **1** in water and to the other was added 1 µL of water. Following 24 hours at 37°C, both samples were diluted with 12.5 µL of water, buffered with 1 µL of 1 M TEAB (pH 8.5, Pierce), supplemented with 1.5 µL of 0.1 M DTT, and denatured by heating to 95°C for 5 minutes. 3 µL of 0.1 M IAA was then added and the samples were incubated at RT in the dark for 20 minutes to alkylate free cysteine residues. 1 µL of trypsin (1 mg/mL, Sigma) in 50 mM acetic acid was then added and the sample was further incubated at 37°C for 2 hours to generate peptides.

Samples were then acidified by mixing 1:1 with 1% formic acid in 10% acetonitrile. Peptides from 2 µL of samples were then collected on a 180 µm x 20 mm 5 µm Symmetry C18 trap column (Waters) flowing at 2500 nL/min and subsequently separated over a 75 µm x 250 mm 1.7 µm Peptide CSH C18 (C18) flowing at 300 nL/min using a nanoAcquity M-Class LC system (Waters). The column was maintained at 60°C. Solution A was 0.1% formic acid in LC-MS grade water and solution B was 0.1% formic acid in LC-MS grade acetonitrile. The separation gradient was 3 minutes of isocratic 2.5% B followed by a 7 minute gradient to 8% B, then a 30 minute gradient to 30% B, a 5 minute gradient to 80% B, a 4 minute gradient to 95% B, a 1 minute gradient to 2.5% B, and 15 minutes of isocratic 2.5% B. MS/MS peaks were picked using Compass (Bruker). Searches were performed against the predicted proteome of *B. thetaiotaomicron* supplemented with common contaminants using Mascot (Matrix science) with a mass tolerance of 5 ppm and a false discovery rate of 1%. Variable modifications including cysteine carbamidomethylation, methionine oxidation, and cysteine or glutamate modification with **1** were included in the search.

Enzyme crystallization and diffraction

Crystals of *Bt*GH146 were grown from a 1.2 µL sitting drop containing a 1:1 mixture of 10 mg/mL *Bt*GH146 in SEC buffer (pre-incubated at 37°C for 16 hours following supplementation with 200 µM **1**) with 20% PEG3350, 0.2 M ammonium formate, 2.5% (170 mM) arabinose, 0.1 M pH 6.0 MES (Figure S13). Crystals were transferred into mother liquor supplemented with 20% PEG400 prior to cryocooling in LN_2. Diffraction data were collected out to 1.41 Å on beamline I03 at Diamond Light Source (Harwell) and processed using the Xia2[4] pipeline with Dials[5].

HypBA1 protein labelled with **1** was used for crystallization. After purification by Ni affinity chromatography, the protein sample was incubated with 0.1 mM inhibitor **1** at 37°C for 1 hour in 50 mM Na-acetate (pH 4.5) buffer. The labelled sample was further purified by size exclusion chromatography using HiLoad 16/60 Superdex 200 pg column in 10 mM HEPES-NaOH (pH 7.5) and 150 mM NaCl. The purified sample was concentrated and buffer-exchanged to 10 mM HEPES-NaOH (pH 7.5) with an ultrafiltration centrifugal membrane unit (Vivaspin Turbo 15, 10 kDa molecular weight cutoff; Sartorius Stedim Biotech, Göttingen, Germany). Crystals were grown at 20°C for 2 days using the sitting drop vapor diffusion method by mixing 0.5 µL of protein solution containing 40 mg/mL HypBA1:**1** with equal volume of a reservoir solution, which contained 0.7 M Na-citrate and 0.1 M MES-NaOH, pH 6.5 (Figure S14). For cryoprotectant, 15% (w/v) trehalose was used. Diffraction data were collected at the Photon Factory of the High Energy Accelerator Research Organization (KEK-PF, Tsukuba, Japan) and processed using XDS[6].

Structure solution and refinement

The structure of the *Bt*GH146:**1** complex was solved by molecular replacement using Phaser[7] with the known complex with L-arabinofuranose (PDBID: 5OPJ) as the search model. The resulting solution showed clear density for the bound ligand within the enzyme active site. Ligand coordinates and dictionaries were generated using jLigand[8] and built into the model using Coot[9], followed by alternating rounds of manual model building and density refinement using Coot and REFMAC[10] within the CCP4 suite[11].

The structure of the HypBA1:**1** complex was determined by molecular replacement using Molrep[12] with the HypBA1 structure (PDBID: 3WKX) as the search model. The resulting solution showed clear density for the ligand that is covalently linked to the side chain of C417. Ligand restraint generation and crystallographic refinement were carried out as above.

Computational methods

Conformational free energy landscapes of β-L-arabinofuranose and compound 1

Conformational free energy landscapes (FELs) were computed for β-L-arabinofuranose and compound **1** using Density Functional Theory-based molecular dynamics (MD), according to the Car−Parrinello (CP) method.[13] Each molecule was enclosed in an isolated cubic box of 12.5 $Å^3$. A fictitious electron mass of 500 atomic units (a.u.) was used for the CP Lagrangian and a time step of 0.12 fs was used in all CPMD simulations. This is the same setup used in previous work on α-L-arabinofuranose inhibitors.[2] The Kohn−Sham orbitals were expanded in a plane wave basis set with a kinetic energy cutoff of 70 Ry. *Ab initio* pseudopotentials, generated within the Troullier-Martins scheme[14], were employed. The Perdew, Burke, and Ernzerhoff (PBE) generalized gradient-corrected approximation[15] was selected in view of its good performance in previous work on isolated sugars, glycosidases, and glycosyltransferases[16–18]. The metadynamics algorithm[19], provided by the Plumed 2 plugin[20], was used to explore the conformational FEL of the systems, taking as collective variables (CVs) the pseudorotational phase (φ) puckering coordinate[21], as well as a dihedral angle accounting for the rotation of the sugar hydroxymethyl group. The energy was projected into the φ coordinate for representation purposes (Figure S5). The height of these Gaussian terms was set at 0.6 kcal/mol and a new Gaussian-like potential was added every 500 MD steps. The width of the CVs was set to 0.035 and 0.1 rad for φ and the hydroxymethyl dihedral angle, respectively, according to the oscillations of the CVs in the free dynamics. The simulations were stopped when energy differences among wells remain constant and a diffusive behaviour was observed in both CVs, which was further confirmed by a time-independent free energy estimator.[22] The energy error, taken from the standard deviation within the last 30 ps, is below 0.6 kcal/mol. The exploration of the phase space was extended up to 240 ps for both β-L-arabinofuranose and compound **1**, which corresponds to 4000 added Gaussian functions.

Quantum chemical description of the $Zn(Cys)_3(Glu)$ coordination complex

For analysis, the initial structure of the $Zn(Cys)_3(Glu)$ coordination complex was taken from the X-ray structure of *Bt*GH146 obtained in this work (PDB 6YQH). The $Zn(Cys)_3(Glu)$ cluster included the Zn^{2+} atom and the surrounding C341, C416, C417 and E339 residues. C_α carbon atoms from residues E339, C341 and the backbone carbonyl group of W415 were saturated with additional hydrogen atoms, the peptide bond between C416 and C417 was retained and the nitrogen atom from V418 backbone was substituted by a methyl group (see Figure S12). The electronic structure was calculated within density functional theory (DFT), using the Becke-3-Lee-Yang-Parr (B3LYP)[23] exchange correlation functional and a 6-311G** basis set to expand Kohn−Sham orbitals. This approach was used in previous Zn^{2+}-$(Cys)_x$ cluster studies,[24–26] providing reliable results. Geometry optimizations were performed by varying the $Zn\cdots S_{Cys416}$ distance from the starting configuration (distance of 2.51 Å) in intervals of 0.01-0.03 Å. The corresponding potential energy profile is shown in Figure S12. All calculations were performed with Gaussian09.[27]

Classical molecular dynamics simulations

The initial structures for the simulations of unreacted HypBA1 and *Bt*GH146 enzyme complexes were taken from their respective X-ray structures after labelling with **1**, reported in this work (PDBIDs 7BZL and 6YQH, respectively). Molecular dynamics (MD) simulations were performed using Amber 18.[28] The Amber ff14SB[29] and TIP3P[30] force fields were used to describe proteins and water molecules, respectively. The inhibitor **1** was parametrised employing the antechamber module,[31] in conjunction with GAFF[32] parameters. Atomic charges were obtained from *ab initio* calculations at the HF/6-31G* level of theory using Gaussian09.[27] The coordinates of residues 246 and 247, which are missing in the structure of the HypBA1:**1** complex, were taken from a previous higher resolution X-ray structure (PDB 3WKW[1]). Two small loops comprising residues 495-498 and 629-631, missing in the *Bt*GH146:**1** structure, were modelled using MODELLER.[33] Crystallographic water molecules were retained in all complexes. A water box within a radius of 15 Å from the enzyme surface was used to further solvate each system. Sodium ions (25/26 Na^+, depending on whether two or one active site Glu residues were considered as protonated, respectively) were added to neutralize the total charge of the HypBA1:**1** complex. Similarly, chloride anions were used to neutralize the *Bt*GH146:**1** complex (2/3 Cl^- ions were required). Upon solvation, the total number of atoms of the system is 100659 /119818 (HypBA1:**1**/*Bt*GH146:**1**). All simulations were performed in periodic boundary conditions, with cell dimensions of 108.2 × 91.9 × 119.8 $Å^3$/107.7 × 114.3 × 113.9 $Å^3$ (HypBA1:**1**/*Bt*GH146:**1**).

The protonation states of HypBA1 and *Bt*GH146 residues were assigned considering neutral pH, using the PROPKA 3.0 software,[34] as well as careful analysis of the chemical environment of each residue. All Glu and Asp residues were taken as deprotonated (i.e. negatively charged), unless otherwise specified. Residues E196 and E368 in HypBA1 were considered as protonated. Special attention was taken with two active site Glu residues that are in proximity. These are the acid/base residue (E322 and E320, in HypBA1 and *Bt*GH146, respectively) and the Glu residue being closest to the inhibitor epoxide oxygen (E338 and E339, in HypBA1 and *Bt*GH146, respectively). Several protonation states were tested for these residues (Figures S9-S10), with the objective of finding the most likely protonation state that is compatible with the formation of a stable complex and with the position of the inhibitor in the X-ray structure. In the case of the HypBA1:**1** complex, the most favourable configuration resulted when E322 is in its protonated form, whereas E338 is deprotonated (Figure S10). In the case of the *Bt*GH146:**1** complex, the most favourable configuration was obtained when both E320 and E339 are protonated (Figure S11). Histidine residues 149, 270, 404, 408 and 446 (HypBA1) were considered neutral with one proton located at N_ε, whereas residues 142 and 194 were protonated at N_δ. Histidine residues 52, 69, 92, 158, 189, 245, 283, 319, 403, 472 and 613 were considered positively charged. In the case of *Bt*GH146, residues 25, 51, 89, 122, 169, 183, 246, 264, 296, 308, 381, 403, 425, 435, 508, 535, 582, 606, 642, 710 and 755 were considered as N_ε-protonated, while residues 93, 112, 218, 321, 323, 585, 609 and 706 were taken as N_δ-protonated. Histidine residues 290 and 694 were taken as positively charged.

MD simulations were performed using the following protocol. First, ions and water molecules were relaxed with 20,000 minimization cycles, followed by relaxation of the whole system with additional 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 were applied to the protein and ligands during the first heating step, thus only water molecules and ions were able to diffuse. All restraints were released after reaching 100K. Subsequently, the density was converged up to water density at 300 K by running 250 ps in the NPT ensemble. An equilibration run of 500 ps was performed to equilibrate the RMSD in the NVT ensemble with a time step of 1 fs. At this point, the time step was increased to 2 fs, employing the SHAKE algorithm[35] and the simulations were extended up to 100 ns. Two MD replicas were performed for each system. Analysis of the trajectories was carried out using standard tools of AMBER and VMD.[36] The time evolution of the RMSD of the protein is provided in Figures S8 and S9.

QM/MM metadynamics simulation of the deglycosylation reaction

The deglycosylation step of the double-displacement reaction of HyBA1 was modelled from the structure obtained in this work for HyBA1 labelled with **1** (PDB entry 7BZL). The natural thioglycosyl enzyme intermediate (tGEI) was built by replacing the CH-OH at carbon 6 with an oxygen atom. The system was equilibrated following the MD protocol described in the previous section. The cluster composed by the Zn^{2+}, the tGEI and the residues around the cation were parametrized using MCPB.py[37] from AmberTools[28] with default parameters. This structure was kept

fixed during the minimization step. After 100 ns of production MD, a snapshot of the trajectory was taken as initial structure for quantum mechanics/molecular mechanics (QM/MM) calculations.

QM/MM MD simulations were performed using the method developed by Laio et al.,[38] which combines *ab initio* MD (Car–Parrinello MD)[13], with classical MD. The QM region included the sugar and the Zn^{2+} ion, together with all the coordinating residue neighbours (51 atoms). That is, the side chains of E322, E338, C340, C417 and C418, which were further saturated at each C_α atom of the C_α–C_β bonds with capping hydrogens to set the QM-MM border. The QM region was enclosed in an isolated supercell of size 16.7 × 18.5 × 16.0 $Å^3$. The NN, MIX and ESP radii were set at 7.45 Å, 9.57 Å and 13.30 Å in order to efficiently describe the electrostatic interactions of the system, following the protocol established in previous studies.[39] The Becke–Lee–Yang–Parr (BLYP) gradient-corrected exchange-correlation functional was used to treat the quantum atoms at DFT level of theory, which has been shown to provide a reliable description of glycosidases that contain a Zn^{2+} cluster in the active site.[40] Kohn–Sham orbitals were expanded in a plane wave basis set with a kinetic energy cut-off of 80 Ry. Norm–conserving Troullier–Martins *ab initio* pseudopotentials have been used for all elements.[14]

The system was equilibrated according the following protocol. First, the geometry was optimized with annealing of the nuclei until a maximal component of the nuclear gradient reached a value of $5 \cdot 10^{-4}$ a.u. Afterwards, 5 ps of QM/MM MD were performed at 300 K employing the Nosé-Hoover thermostat.[41,42] A time step of 0.12 fs and a fictitious electronic 800 mass were used to ensure adiabaticity. A snapshot of the resulting equilibrated structure was chosen as starting point for the QM/MM metadynamics simulations.

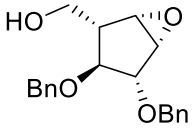
Hydrolysis of tGEI by HyBA1 was investigated using QM/MM metadynamics[19] with the PLUMED 2 plugin.[20] The collective variable (CV) employed to drive the reaction, from the tGEI to the reaction products, was taken as the difference between two distances: the distance from the oxygen from the nucleophilic water (O_w) and the anomeric carbon (C1) of the substrate; and the distance from the C1 atom to the S atom of C417 (S_{cys}). The width and height of the biasing Gaussian functions were taken as 0.12 Å and 1.0 kcal/mol, respectively, according to the oscillations obtained in free dynamics. For better convergence of the free energy profile, the Gaussian height was decreased to 0.5 kcal/mol when the system was close to the TS. The deposition time was set to 350 MD steps. The simulation was stopped when a re-crossing over the TS occurred, as recommended for chemical reactions.[43] At this point, 630 Gaussian terms were already added. The corresponding free energy profile is shown in Figure 2 of the manuscript.

Supplemental Synthetic Methods and Compound Characterization

All synthetic reagents were of a commercial grade and were used as received unless stated otherwise. Reactions were monitored by analytical thin-layer chromatography (TLC) using Merck aluminum sheets pre-coated with silica gel 60 with detection by UV absorption (254 nm) and by spraying with a solution of $(NH_4)_6Mo_7O_{24}\cdot H_2O$ (25 g/L) and $(NH_4)_4Ce(SO_4)_4\cdot H_2O$ (10 g/L) in 10% sulfuric acid followed by charring at ~150 ˚C. Column chromatography was performed manually using either Baker or Screening Device silica gel 60 (0.04 - 0.063 mm) or a Biotage Isolera™ flash purification system using silica gel cartridges (Screening devices SiliaSep HP, particle size 15-40 µm, 60Å) in the indicated solvents. ^{1}H NMR and ^{13}C NMR spectra were recorded on Bruker AV-400 (400/101 MHz) and Bruker AV-500 (500/126 MHz) spectrometer in the given solvent. Chemical shifts are given in ppm relative to the chloroform residual solvent peak or tetramethylsilane (TMS, used in MeOD solvent) as internal standard. The following abbreviations are used to describe peak patterns when appropriate: s (singlet), d (doublet), dd (doublet of doublet), dddd (doublet of doublet of doublet of doublet), tdd (triplet of doublet of doublet), m (multiplet), br (broad). 2D NMR experiments (HSQC, COSY and NOESY) were carried out to assign protons and carbons of the new structures and assignation follows the general numbering shown in final epoxide **1**. High-resolution mass spectra (HRMS) of compounds were recorded with an LTQ Orbitrap (Thermo Finnigan). Optical rotations were measured on an Anton Paar MCP automatic polarimeter (Sodium D-line, λ = 589 nm).

Synthesis of β-L-*arabinofuranosyl* epoxide 1

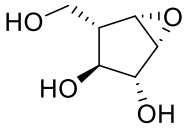
((1*S*,2*S*,3*S*,4*R*,5*S*)-3,4-bis(benzyloxy)-6-oxabicyclo[3.1.0]hexan-2-yl)methanol 3



To a buffered solution of NaH_2PO_4/ Na_2HPO_4 (1:1 ratio, 13 mL) under vigorous stirring at 0 °C, cyclohexene 2[2] (200 mg, 0.64 mmol) dissolved in CH_2Cl_2 (6 mL) was added. Afterwards, a solution of *m*-CPBA (200 mg, 1.16 mmol) in CH_2Cl_2 (6 mL) was added and the reaction was stirred at 4 °C overnight. As TLC indicated the presence of starting material after 19 h, more *m*-CPBA (105 mg, 0.61 mmol) was added. After stirring the reaction mixture for 3 additional days, the layers were separated, and the aqueous phase was washed with EtOAc (2 x 100 mL). The combined organic layers were dried over Mg_2SO_4, filtered and concentrated under reduced pressure. Purification of the crude by automatic silica gel column chromatography (25 g silica, 0→70% EtOAc in pentane) afforded the title compound 3 as major product (155 mg, 74%). $[\alpha]^D_{20}$ = –12.6 (*c* = 1, $CHCl_3$). ^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.45 – 7.27 (m, 10H, CH Ar), 4.79 (d, J = 11.9 Hz, 1H, C*H*HPh), 4.69 (d, *J* = 11.3 Hz, 2H, CH_2Ph), 4.55 (d, *J* = 11.6 Hz, 1H, C*H*HPh), 4.11 (dd, *J* = 5.3, 1.3 Hz, 1H, CH-3), 3.91 (dd, *J* = 10.6, 5.4 Hz, 1H, C*H*HOH), 3.84 (dd, *J* = 10.6, 7.5

Hz, 1H, C*H*HOH), 3.56 (m, 3H, CH-1, CH-2, CH-4), 2.23 (tdd, *J* = 7.2, 5.4, 1.4 Hz, 1H, CH-5). ^{13}C NMR (101 MHz, $CDCl_3$) δ = 138.2, 136.9 ($2C_q$ Ar), 128.7, 128.6, 128.1, 128.0, 127.9 (10CH Ar), 85.7 (CH-3), 81.0 (CH-4), 72.7, 71.9 ($2CH_2Ph$), 62.2 (CH_2OH), 55.1, 54.0 (CH-1, CH-2), 46.8 (CH-5). HRMS: calcd. for $[C_{20}H_{22}NaO_4]^+$ 349.3812; found 349.1420.

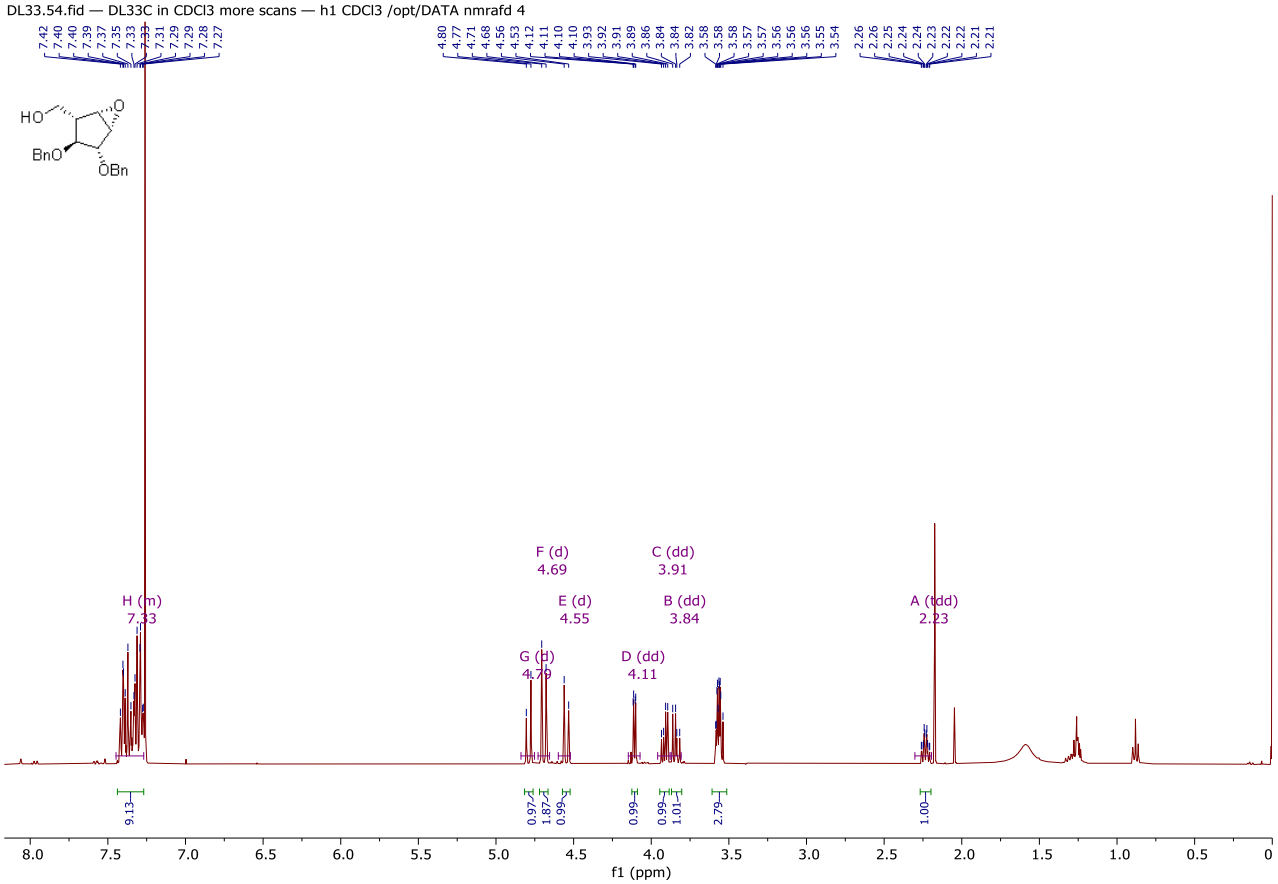
(1*R*,2*R*,3*S*,4*S*,5*S*)-4-(hydroxymethyl)-6-oxabicyclo[3.1.0]hexane-2,3-diol 1

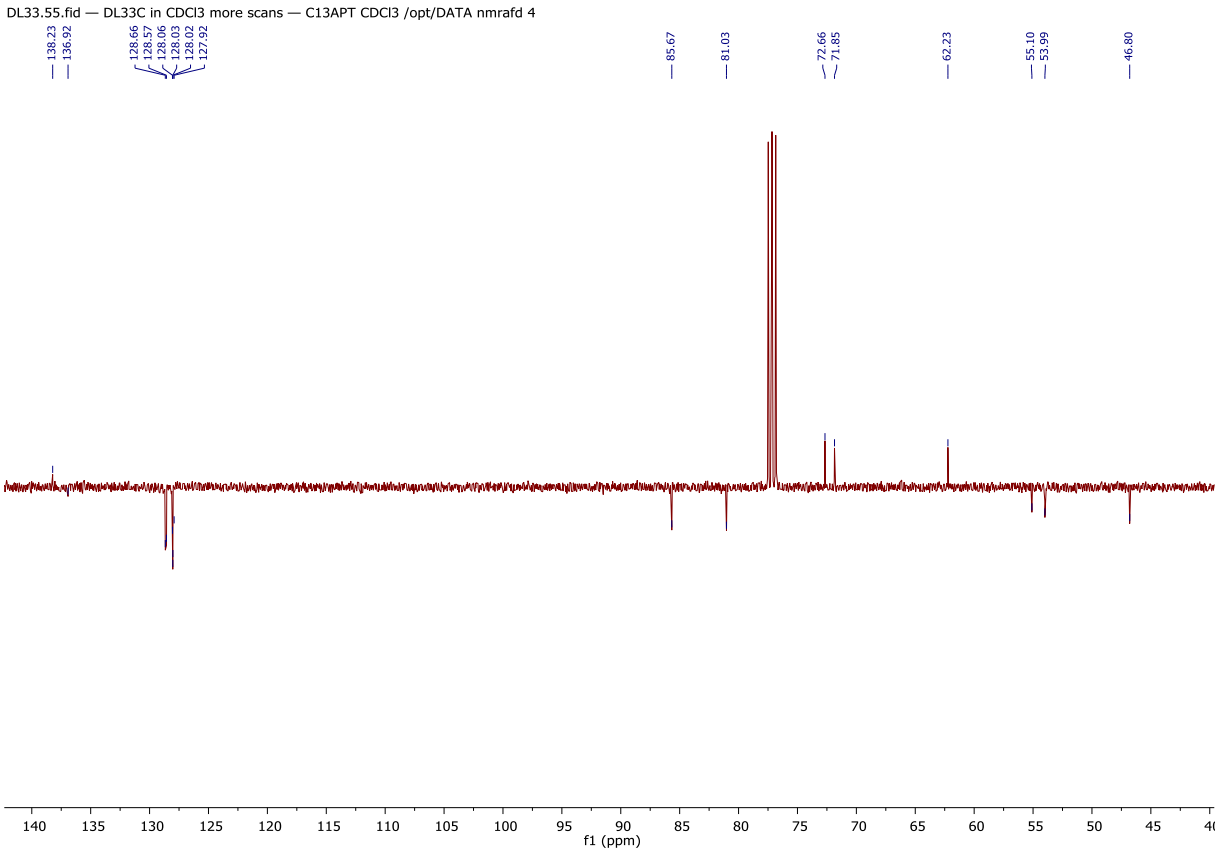


Compound **3** (200 mg, 0. 64 mmol) was dissolved in MeOH (6 mL) and thoroughly degassed under argon atmosphere. $Pd(OH)_2$ (20% on carbon, 2.6 mg, 0.018 mmol) was added to the solution and the mixture was further degassed with argon. H_2 gas was bubbled through the reaction mixture with a balloon at rt overnight while stirring. The reaction mixture was then bubbled with argon and filtered over Celite®, and the filter was washed several times with MeOH. The filtrate was concentrated under reduced pressure and purified by silica gel column chromatography (0%→20% MeOH in DCM), and the purified product was freeze-dried to afford final β-L-*arabinofuranose*-configured epoxide **1** as white solid (2.54 mg, 18%). ^{1}H NMR (500 MHz, MeOD) δ = 3.94 (dd, *J* = 6.0, 1.4 Hz, 1H, CH-3), 3.81 (dd, *J* = 10.6, 4.7 Hz, 1H, C*H*HOH), 3.71 – 3.63 (m, 1H, C*H*HOH), 3.55 (dd, *J* = 3.2, 1.2 Hz, 1H, CH), 3.46 (dd, *J* = 3.0, 1.6 Hz, 1H, CH), 3.18 (dd, *J* = 7.5, 6.1 Hz, 1H, CH-4), 1.97 (dddd, *J* = 9.3, 7.6, 4.7, 1.3 Hz, 1H, CH-5). ^{13}C NMR (126 MHz, MeOD) δ = 80.4 (CH-3), 76.0 (CH-4), 61.7 (CH_2), 57.6, 54.9 (CH-1, CH-2), 49.6 (CH-5).

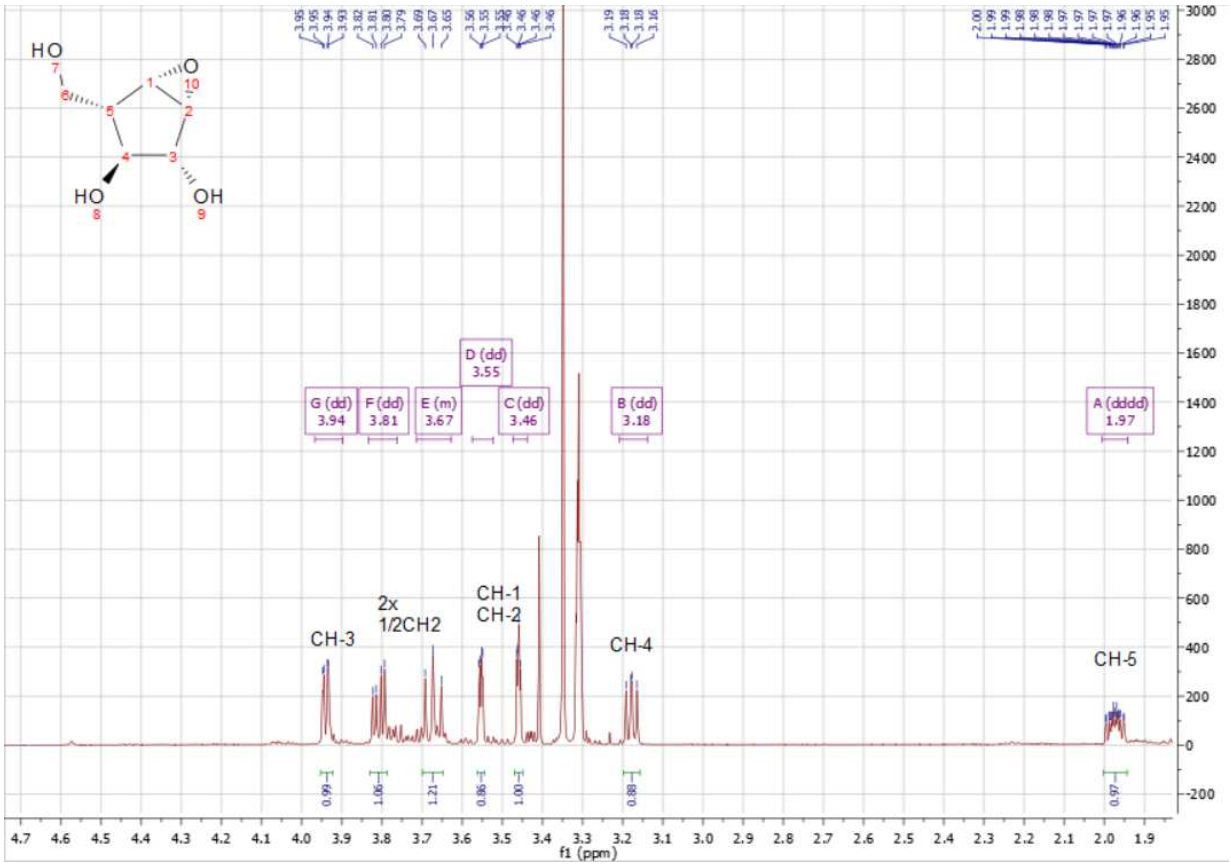
NMR Spectra

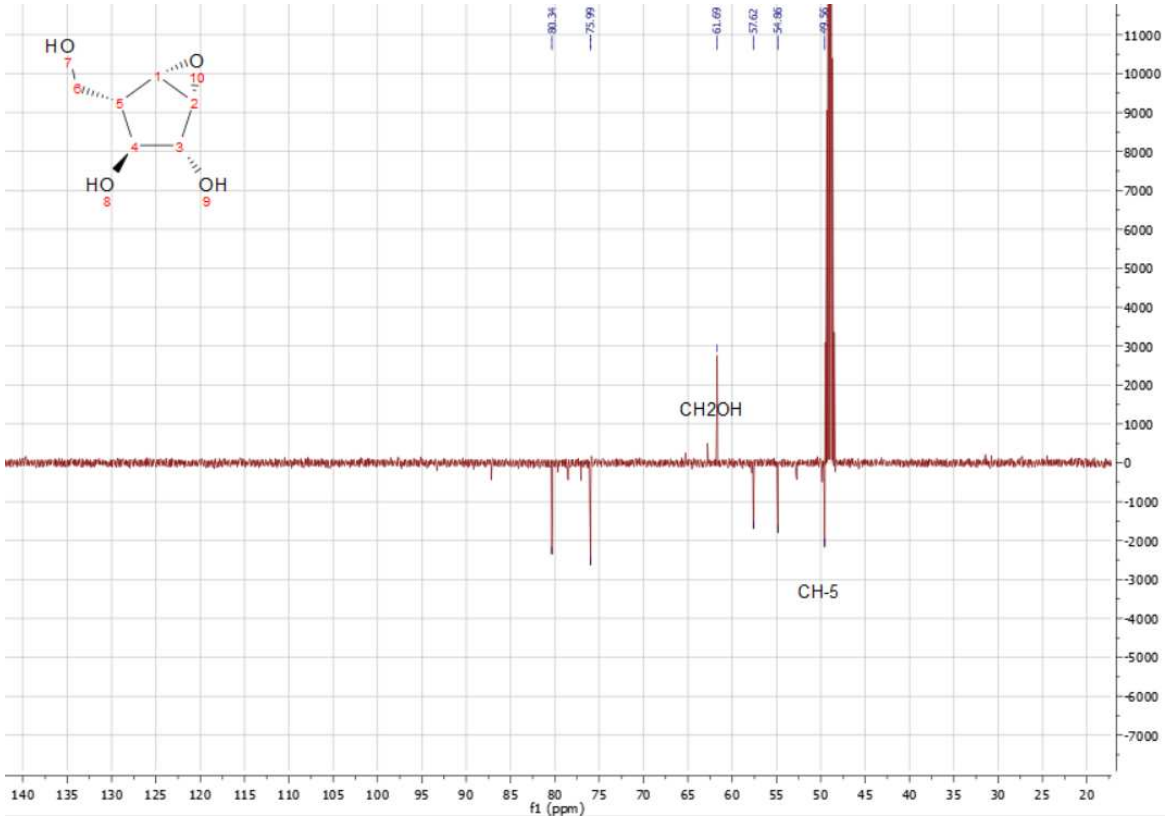
^{1}H-NMR and ^{13}C-NMR spectra of **3** in $CDCl_3$





^{1}H-NMR and ^{13}C-NMR spectra of **1** in MeOD





Supplemental Figures and Tables

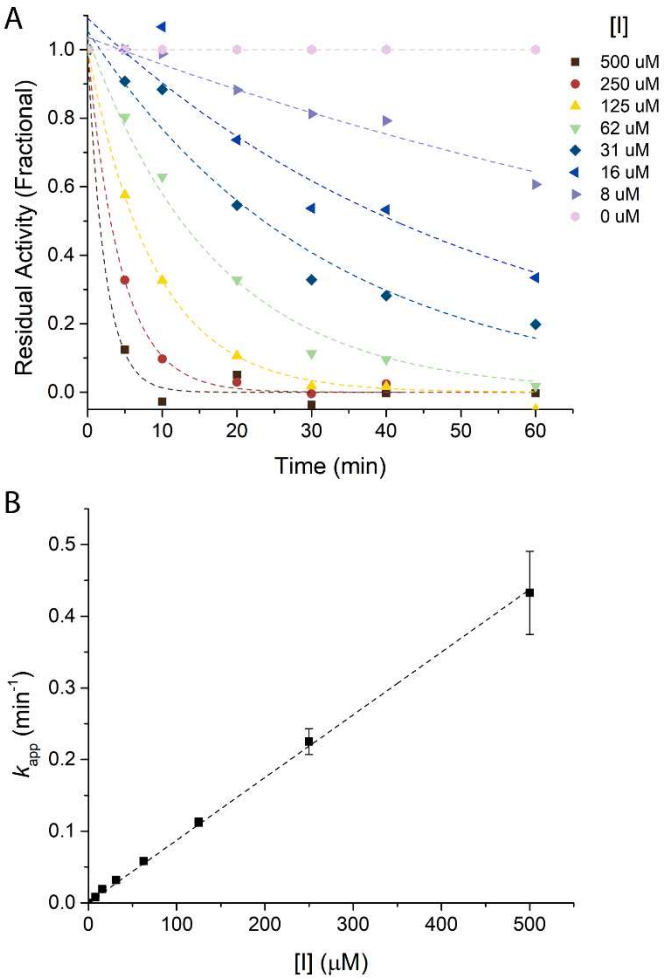


Figure S1. Irreversible inhibition kinetics for the inhibition of HypBA1 with compound **1**. A) Residual activity values are plotted against incubation time for each concentration of inhibitor. Fitted exponential decay curves are plotted as dashed lines. B) Apparent decay constant (k_{app}) values are plotted against inhibitor concentration. Error bars represent the standard error of the k_{app} value from the fit in panel A. The linear fit used to estimate k_i/K_I is displayed as a dashed line.

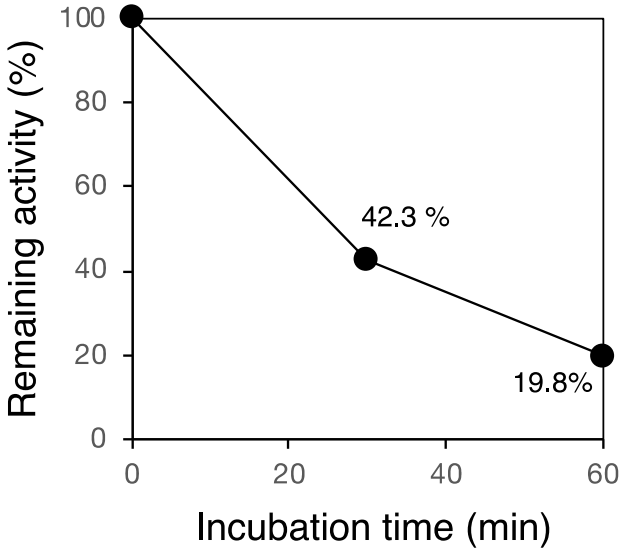


Figure S2. Activity of HypBA1 following incubation at 37°C, pH 4.5 in the presence of 0.1 mM **1** prior to crystallisation.

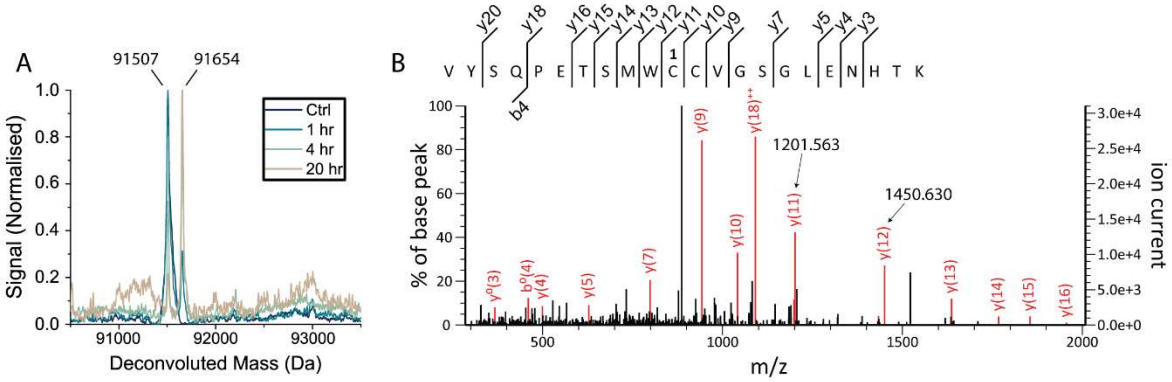


Figure S3. Mass spectrometry of *Bt*GH146 labelled with **1**. A) Intact MS of *Bt*GH146 following incubation with 0.1 mM **1** in pH 7.5 HEPES buffer at 37°C for different lengths of time. Expected mass (native) = 91512 Da, (with **1**) = 91658 Da. B) MS/MS spectrum collected for the labelled active site peptide of *Bt*GH146. The m/z values for the diagnostic y_{11} and y_{12} fragments are given. See table S3 for all observed y and b fragment m/z values.

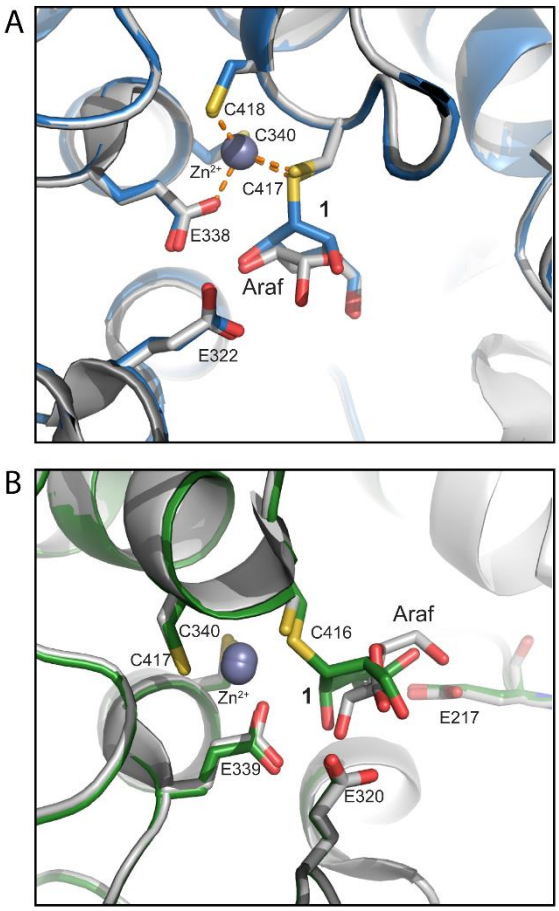


Figure S4. Overlay of the structures of HypBA1 and BtGH146 in complex with either L-arabinofuranose (Araf) or compound **1**. A) Superimposition of HypBA1 following labelling with **1** (blue) and HypBA1 bound to Araf (white). B) Superimposition of the *Bt*GH146 following labelling with **1** (green) and *Bt*GH146 bound to Araf (white).

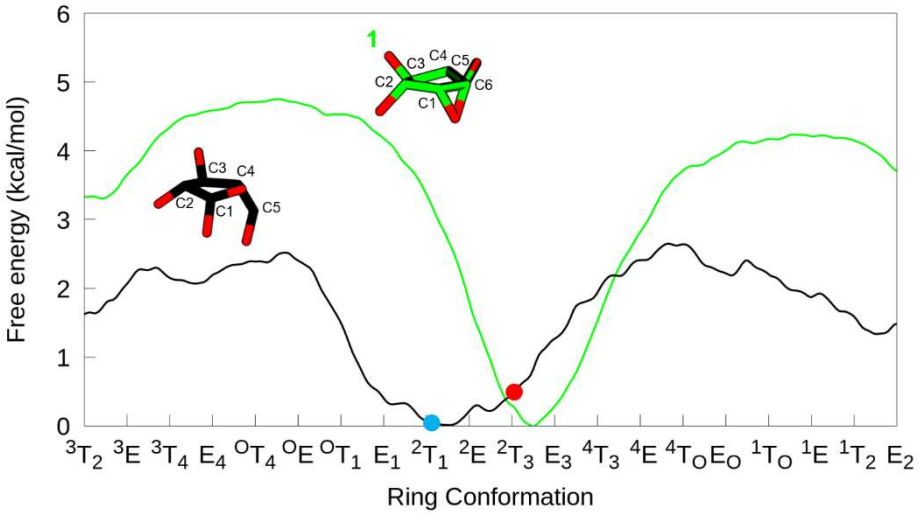


Figure S5. Conformational free energy profile of isolated β-L-arabinofuranose (black line) and compound **1** (green line). Conformations observed in product complexes of β-L-arabinofuranosidases are represented with either a red circle (PDB 3WKX for HypBA1) or a blue circle (PDB 5OPJ for *Bt*GH146).

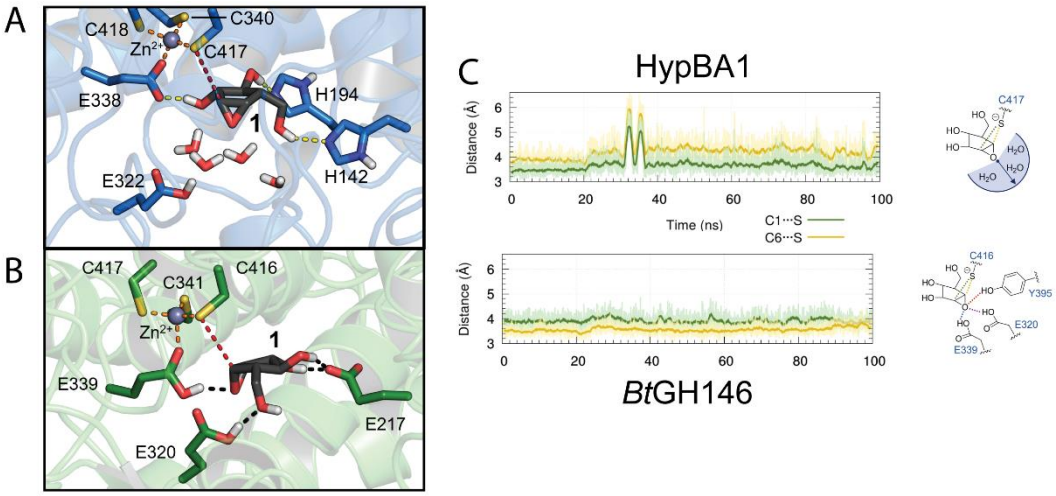


Figure S6. Unreacted Michaelis complexes obtained from MD simulation. A) Complex of HypBA1 by **1**. B) Complex of *Bt*GH146 with **1**. Interactions are shown with dashed lines. C) Distances between the putative catalytic nucleophile sulphur atom and C1 or C6 over the course of the simulation.

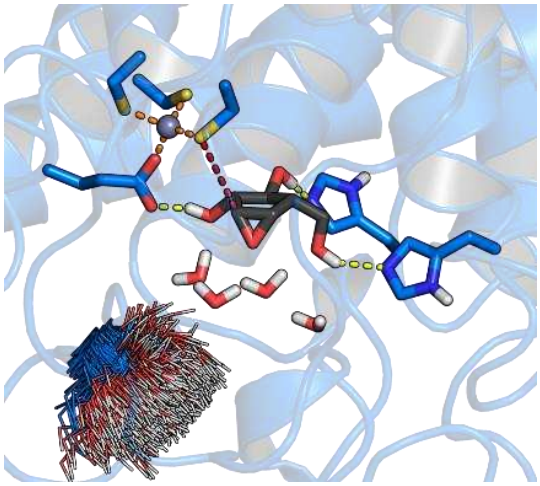


Figure S7. Conformations of the acid/base residue (E322) sampled during the MD simulation of the unreacted complex of HypBA1 with **1**.

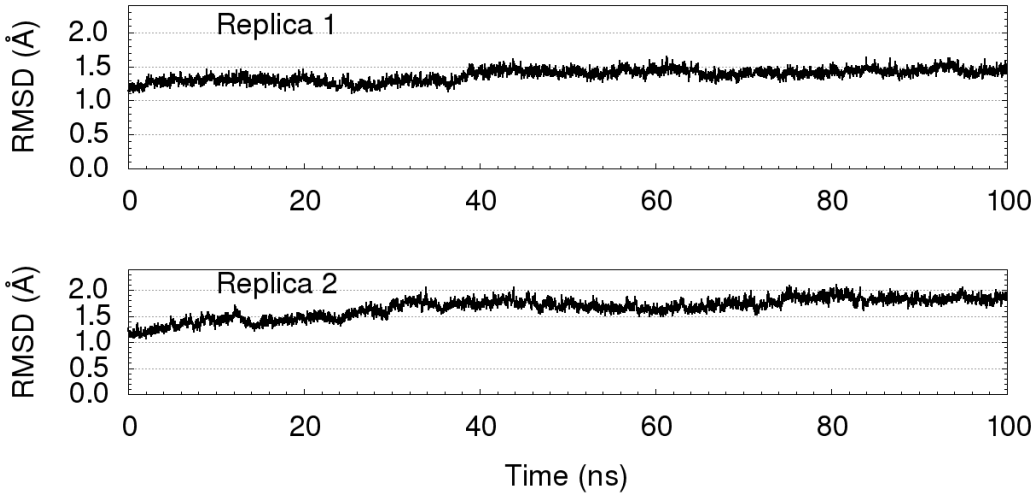


Figure S8. RMSD (in Å) of the protein backbone during the production run of the unreacted complex of HypBA1 with **1**.

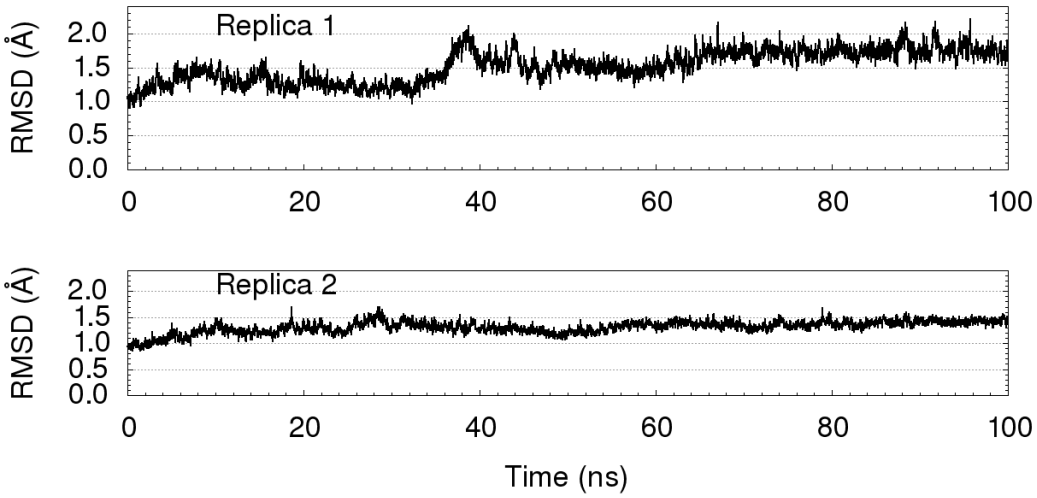


Figure S9. RMSD (in A) of the protein backbone during the production run of the unreacted complex of *Bt*GH146 with **1**.

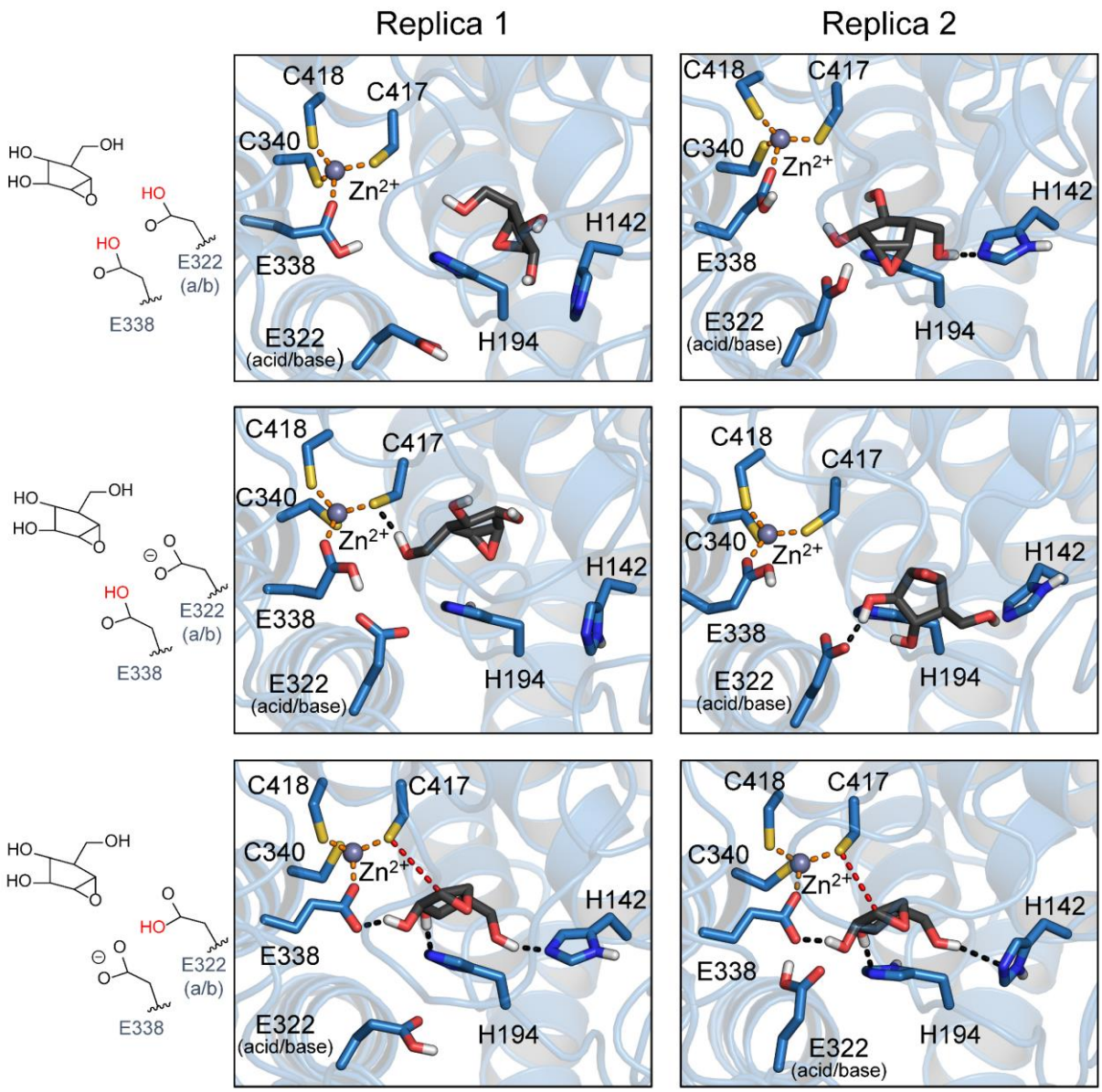


Figure S10. Distinct protonation states of the active site glutamate residues E320 (acid/base) and E339 (glutamate coordinating Zn^{2+}) that were tested for the binding mode of compound **1** in the active site of the HypBA1. Unless otherwise indicated, the pictures show structures after 100 ns of MD simulation. The coordination distances of the Zn^{2+} cluster are shown in dashed orange lines, while hydrogen bonds between the inhibitor and the enzyme are shown in black dashed lines. A red dashed line indicates the nucleophilic attack from the nucleophilic cysteine to the carbon of the inhibitor **1**. The structure shown for models with both glutamate residues protonated (top panels) is an early one (taken before 25 and 65 ns for replica 1 and 2, respectively), since complete unbinding was observed later in the MD simulation.

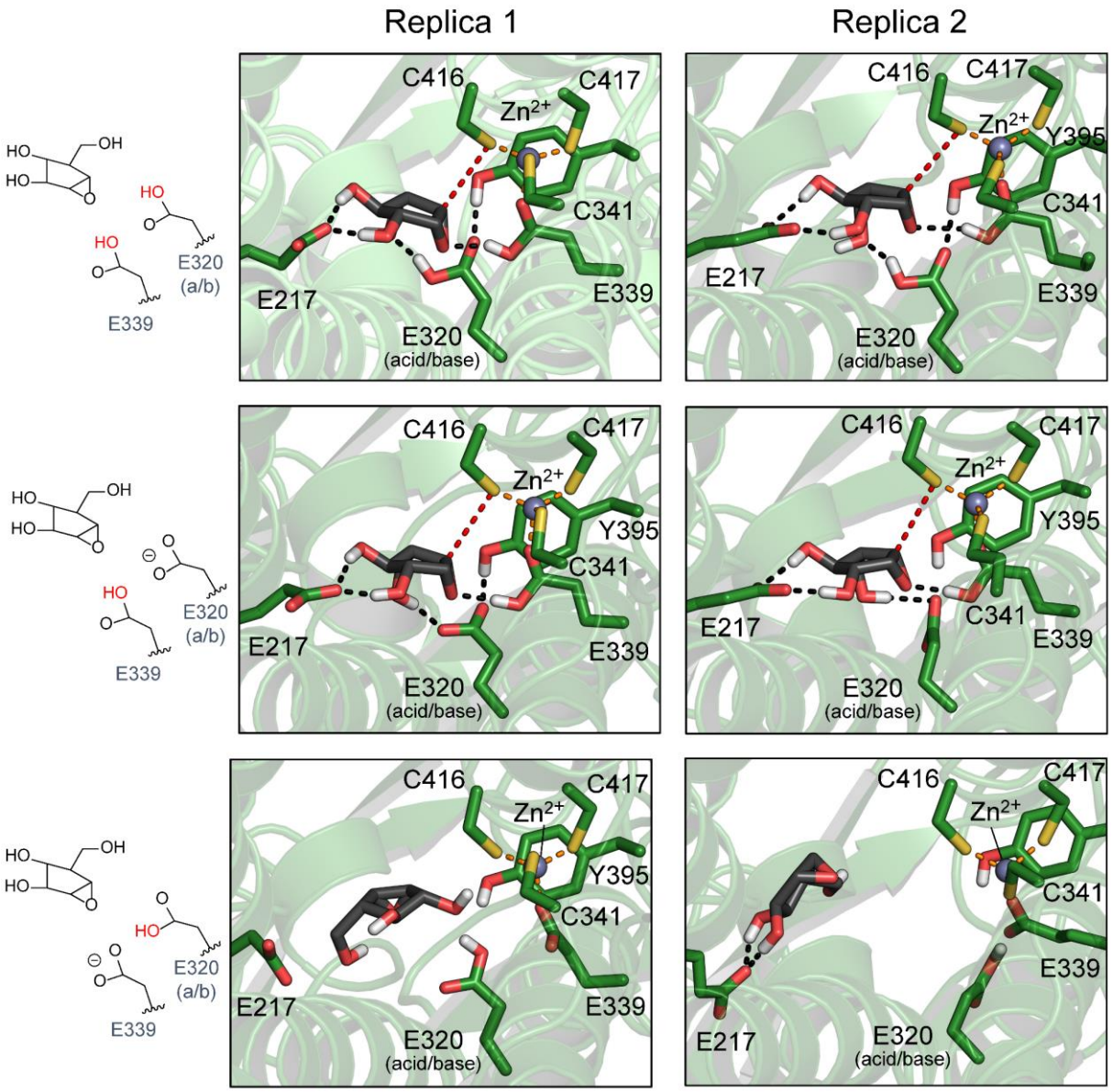


Figure S11. Distinct protonation states of the active site glutamate residues E320 (acid/base) and E339 (glutamate coordinating Zn^{2+}) that were tested for the binding mode of compound **1** in the active site of the *Bt*GH146. The pictures show structures after 100 ns of MD simulation. The coordination distances of the Zn^{2+} cluster are shown in dashed orange lines, while hydrogen bonds between the inhibitor and the enzyme are shown in black dashed lines. A red dashed line indicates the nucleophilic attack from the nucleophilic cysteine to the carbon of the inhibitor **1**.

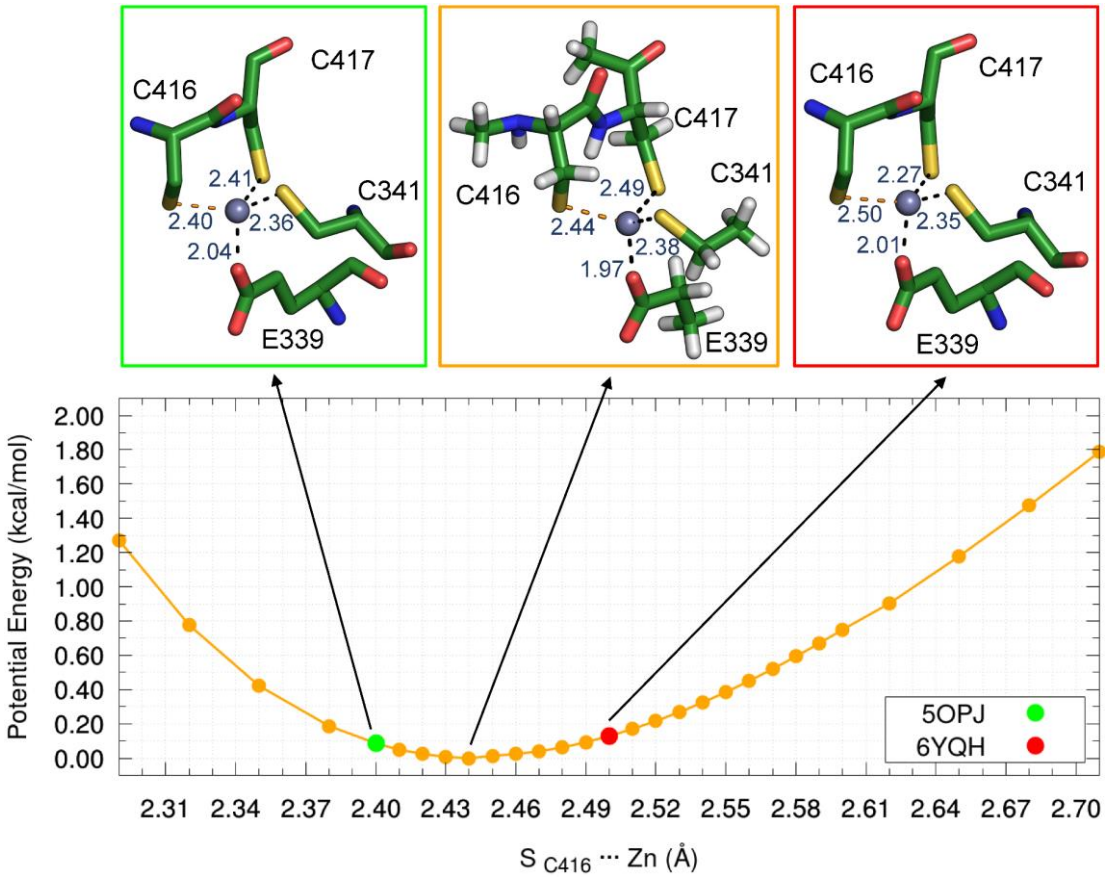


Figure S12. Potential energy profile of the Zn···S_{C416} distance (shown in orange dashed lines in the molecular models) within the $Zn(Cys)_3(Glu)$ coordination complex in *Bt*GH146, computed at the B3LYP/6-311G** level. Energy values corresponding to the distance in the X-ray structure of the covalent intermediate mimic complex and the reaction products are highlighted in green and red, respectively. The structure corresponding to the minimum of the in the potential energy profile is shown in the middle (orange) panel.

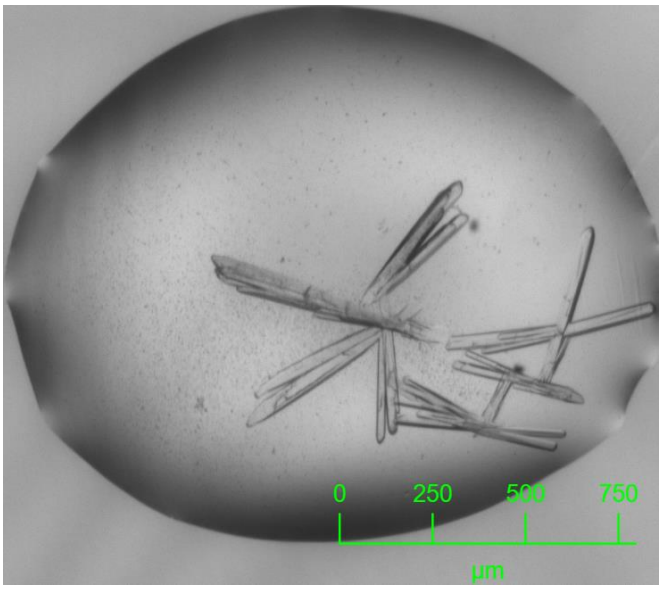


Figure S13. Appearance of the BtGH146:**1** crystals used in this study.

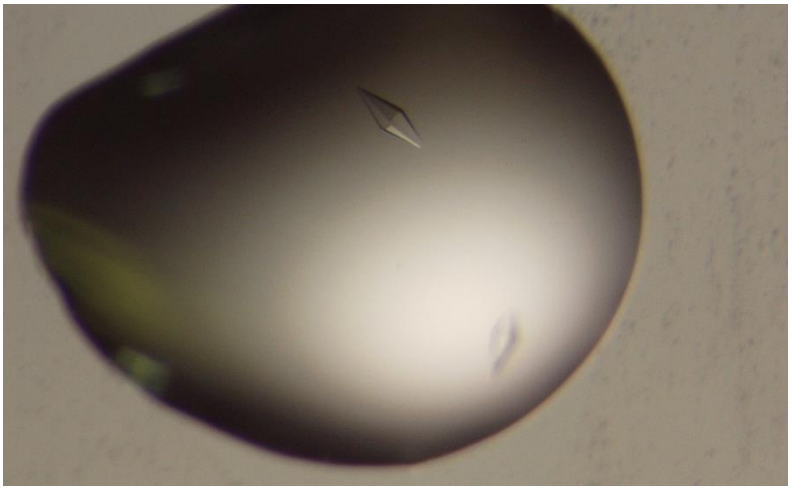


Figure S14. Appearance of the HypBA1:**1** crystals used in this study.

Table S1. Data collection and refinement statistics (molecular replacement)

	*Bt*GH146	HypBA1
	Complex with **1** (PDB 6YQH)	Complex with **1** (PDB 7DIF)
Data collection		
Space group	$P4_32_12$	$P3_221$
a, *b*, *c* (Å)	95.42, 95.42, 189.07	77.71, 77.71, 252.09
α, β, γ (°)	90, 90, 90	90, 90, 120
Resolution (Å)	95.42-1.41 (1.45-1.41)	46.00-1.75 (1.78-1.75)
R_{meas}	0.120 (2.957)	0.083 (3.325)
R_{pim}	0.030 (0.719)	0.028 (1.089)
$I / \sigma I$	11.7 (1.1)	10.8 (0.8)
Completeness (%)	100.0 (99.9)	100.0 (100.0)
Redundancy	16.2 (16.4)	9.9 (10.1)
$CC_{1/2}$	0.999 (0.484)	0.999 (0.670)
Refinement		
No. reflections	167656	85691
R_{work} / R_{free}	0.13/0.17	0.22/0.24
No. atoms		
Protein	6211	5136
Ligand/ion	87	12
Water	575	233
B-factors (Å^2)		
Protein	21	46
Ligand/ion	40	40
Water	31	43
R.m.s. deviations		
Bond lengths (Å)	0.015	0.0141
Bond angles (°)	1.88	1.86

*Values in parentheses are for highest-resolution shell.

Table S2. m/z values for the MS/MS spectrum of the labelled active site peptide of HypBA1. All expected b and y fragment m/z values are given for the sequence of the expected tryptic active site peptide modified with compound 1 on the putative catalytic cysteine. Values coloured black were observed in the measured MS/MS spectrum and values coloured grey were not.

Residue	AA Seq	b (m/z)	y (m/z)
1	V	100.0757	
2	D	215.1026	1786.7346
3	W	401.1819	1671.7076
4	F	548.2504	1485.6283
5	G	605.2718	1338.5599
6	C	765.3025	1281.5384
7	A	836.3396	1121.5078
8	C-1	1085.407	1050.4707
9	C	1245.437	801.4036
10	P	1342.49	641.3729
11	A	1413.527	544.3202
12	N	1527.57	473.2831
13	I	1640.654	359.2401
14	A	1711.691	246.1561
15	R		175.119

Table S3. m/z values for the MS/MS spectrum of the labelled active site peptide of *Bt*GH146. All expected b and y fragment m/z values are given for the sequence of the expected tryptic active site peptide modified with compound 1 on the putative catalytic cysteine. Values coloured black were observed in the measured MS/MS spectrum and values coloured grey were not.

#	Seq.	b	y
1	V	100.0757	
2	Y	263.139	2560.079
3	S	350.171	2397.016
4	Q	478.2296	2309.984
5	P	575.2824	2181.925
6	E	704.325	2084.872
7	T	805.3727	1955.83
8	S	892.4047	1854.782
9	M	1023.445	1767.75
10	W	1209.525	1636.709
11	C-1	1458.592	1450.63
12	C	1618.622	1201.563
13	V	1717.691	1041.532
14	G	1774.712	942.4639
15	S	1861.744	885.4425
16	G	1918.766	798.4104
17	L	2031.85	741.389
18	E	2160.892	628.3049
19	N	2274.935	499.2623
20	H	2411.994	385.2194
21	T	2513.042	248.1605
22	K		147.1128

Supplemental References

(1) Ito, T.; Saikawa, K.; Kim, S.; Fujita, K.; Ishiwata, A.; Kaeothip, S.; Arakawa, T.; Wakagi, T.; Beckham, G. T.; Ito, Y.; Fushinobu, S. Crystal Structure of Glycoside Hydrolase Family 127 β-l- Arabinofuranosidase from Bifidobacterium Longum. *Biochem. Biophys. Res. Commun.* **2014**, *447* (1), 32–37.

(2) McGregor, N. G. S.; Artola, M.; Nin-Hill, A.; Linzel, D.; Haon, M.; Reijngoud, J.; Ram, A.; Rosso, M. N.; Van Der Marel, G. A.; Codeé, J. D. C.; Van Wezel, G. P.; Berrin, J. G.; Rovira, C.; Overkleeft, H. S.; Davies, G. J. Rational Design of Mechanism-Based Inhibitors and Activity-Based Probes for the Identification of Retaining α-l-Arabinofuranosidases. *J. Am. Chem. Soc.* **2020**, *142* (10), 4648–4662.

(3) Kaeothip, S.; Ishiwata, A.; Ito, T.; Fushinobu, S.; Fujita, K.; Ito, Y. Preparation of P-Nitrophenyl β-l-Arabinofuranoside as a Substrate of β-l-Arabinofuranosidase. *Carbohydr. Res.* **2013**, *382*, 95–100.

(4) Winter, G.; Lobley, C. M. C.; Prince, S. M. Decision Making in Xia2. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69* (7), 1260–1273.

(5) Winter, G.; Waterman, D. G.; Parkhurst, J. M.; Brewster, A. S.; Gildea, R. J.; Gerstel, M.; Fuentes-Montero, L.; Vollmar, M.; Michels-Clark, T.; Young, I. D.; Sauter, N. K.; Evans, G. DIALS: Implementation and Evaluation of a New Integration Package. *Acta Crystallogr. Sect. D Struct. Biol.* **2018**, *74*, 85–97.

(6) Kabsch, W. XDS. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 125–132.

(7) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. Phaser Crystallographic Software. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(8) Lebedev, A. A.; Young, P.; Isupov, M. N.; Moroz, O. V.; Vagin, A. A.; Murshudov, G. N. JLigand: A Graphical Tool for the CCP4 Template-Restraint Library. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68* (4), 431–440.

(9) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and Development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (4), 486–501.

(10) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 355–367.

(11) Ballard, C.; Keegan, R.; Krissinel, E.; Lebedev, A.; Uski, V.; Waterman, D.; Wojdyr, M. CCP4: A Resource for Macromolecular Crystallography. *Acta Crystallogr. Sect. A Found. Adv.* **2014**, *70* (a1), C1723–C1723.

(12) Vagin, A.; Teplyakov, A. Molecular Replacement with MOLREP. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 22–25.

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

(14) 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.

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

(16) 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.

(17) Biarnés, X.; Ardèvol, 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.

(18) 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.

(19) Laio, A.; Parrinello, M. Escaping Free-Energy Minima. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (20), 12562–12566.

(20) 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.

(21) Huang, M.; Giese, T. J.; Lee, T. S.; York, D. M. Improvement of DNA and RNA Sugar Pucker Profiles from Semiempirical Quantum Methods. *J. Chem. Theory Comput.* **2014**, *10* (4), 1538–1545.

(22) Tiwary, P.; Parrinello, M. A Time-Independent Free Energy Estimator for Metadynamics. *J. Phys. Chem. B* **2015**, *119* (3), 736–742.

(23) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.

(24) Malik, A.; Lin, Y. F.; Pratihar, S.; Angel, L. A.; Hase, W. L. Direct Dynamics Simulations of Fragmentation of a Zn(Ii)-2cys-2his Oligopeptide. Comparison with Mass Spectrometry Collision-Induced Dissociation. *J. Phys. Chem. A* **2019**, *123* (32), 6868–6885.

(25) Tjörnhammar, R.; Edholm, O. Molecular Dynamics Simulations of Zn2+ Coordination in Protein Binding Sites. *J. Chem. Phys.* **2010**, *132* (20).

(26) Brandt, E. G.; Hellgren, M.; Brinck, T.; Bergman, T.; Edholm, O. Molecular Dynamics Study of Zinc Binding to Cysteines in a Peptide Mimic of the Alcohol Dehydrogenase Structural Zinc Site. *Phys. Chem. Chem. Phys.* **2009**, *11* (6), 975–983.

(27) 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**.

(28) Case, D. A.; Walker, R. C.; Cheatham, T. E.; Simmerling, C.; Roitberg, A.; Merz, K. M.; Luo, R.; Darden, T. Amber 18. *Univ. California, San Fr.* **2018**, Univ. California, San Fr.

(29) 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.

(30) 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.

(31) 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.

(32) 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.

(33) Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, *2016*, 5.6.1-5.6.37.

(34) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical PKa Predictions. *J. Chem. Theory Comput.* **2011**, *7* (2), 525–537.

(35) 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.

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

(37) Li, P.; Merz, K. M. MCPB.Py: A Python Based Metal Center Parameter Builder. *J. Chem. Inf. Model.* **2016**, *56* (4), 599–604.

(38) 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.

(39) 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.

(40) 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 (8), 8291–8300.

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

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

(43) Ensing, B.; Laio, A.; Parrinello, M.; Klein, M. L. A Recipe for the Computation of the Free Energy Barrier and the Lowest Free Energy Path of Concerted Reactions. *J. Phys. Chem. B* **2005**, *109* (14), 6676–6687.