Glycoside Hydrolase Catalysis: Do Substrates and Mechanism-Based Covalent Inhibitors React via Matching Transition States?

Oluwafemi Akintola,[1] Marco Farren-Dai,[1] Weiwu Ren,[1] Sandeep Bhosale,[1] Robert Britton,[1] Katarzyna Świderek,[2] Vicent Moliner,[2,*] and Andrew J. Bennet[1,*]

[1] Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, Canada, V5A 1S6

[2] BioComp Group, Institute of Advanced Materials (INAM), Universitat Jaume I, 12071 Castellón, Spain.

Corresponding authors Andrew J. Bennett, e-mail bennet@sfu.ca; Vicent Moliner, e-mail moliner@uji.es.

Keywords: Transition State Analogy, Mechanism-Based, Glycosides, Glycoside Hydrolase, Linear Free Energy Relationships, Carbasugar Inhibitors, QM/MM MD simulations.

Abstract

In this study we look at how a catalytically efficient α-galactosidase stabilizes transition state charge delocalization for substrate hydrolysis. We then assess whether covalent inhibition of the enzyme by three types of mechanism-based covalent inhibitors occurs via similar modes of transition state stabilization. We show, using Bartlett-type linear free energy relationships, that good correlations are obtained between the catalytic efficiencies (k_{cat}/K_m and/or k_{inact}/K_i) for enzyme catalyzed reactions of natural and activated galactoside substrates, and of representatives of three families of classical mechanism-based inhibitors: a 2-deoxy-2-fluoroglycoside, allylic carbasugars, and an epoxy carbasugar. Of note, we show that glycoside natural substrates and allylic carbasugars display log(rate)-log(rate) correlations that are unity (slope ≈ 1), an observation consistent with them having identical positive charge stabilization at the S_N1-like glycosylation and pseudo-glycosylation transition states, respectively. In contrast, 2-deoxy-2-fluoroglycoside mechanism-based inhibitors react via a different enzyme-catalyzed mechanism (S_N2), while the strained epoxy carbasugar inactivates the α-galactosidase by traversing a transition state in which the glycoside hydrolase stabilizes the inactivation transition state that has a significantly lower degree of charge stabilization to those for the natural glycoside substrates. To add weight to these conclusions, we computed free energy landscapes and their associated galactosylation and pseudo-galactosylation transition states using QM/MM molecular dynamics methods with the whole solvated enzyme.

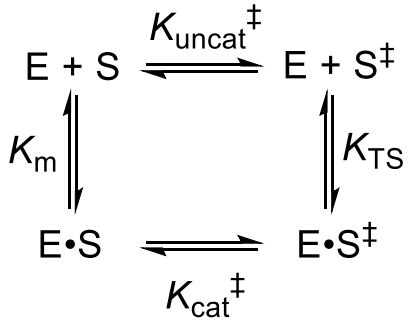
Introduction

For well over a century and starting with the "lock and key" model proposed by Emil Fischer,[1] researchers have advanced a variety of theories on how enzymes catalyze chemical reactions. The first model envisioned a perfectly designed fit between the three dimensional shape of a substrate and a corresponding hollow in the enzyme.[1] Later, Haldane proposed that enzymes and substrates form a precise complementary alignment at the transition state for the reaction rather than at the initially formed Michaelis complex of enzyme and substrate.[2-3] Koshland's induced-fit model provided further insight as to how enzymes act to maximize catalytic efficiency.[4] In the 1970s, Albery and Knowles proposed that, through the process of evolution,[5] the catalytic efficiency (k_{cat}/K_m) of an enzyme can be maximized by biomolecular alterations that affect the free energies of the bound enzyme-substrate and transition states.[6] That is, the enzyme preferentially stabilizes the transition state(s) of kinetically significant step(s) relative to all of the other bound states.

Enzymes are flexible entities and, as such, protein vibrational modes may contribute to the catalytic process.[7-8] In a more general sense, proteins in their native state may exist as a conformational ensemble that undergoes vibrational fluctuations to give a thermodynamic distribution of energy states, with ground states that are optimized for biological activity.[9] In this regard, as a guide to redesign enzymes, some researchers have proposed that allosteric regulation and modulation by dynamic conformational ensembles function to enhance the populations of those reactive conformational states.[10-11] Despite the unquestionable property of flexibility and its role in several of the steps of the overall catalytic process of enzyme, the impact of these effects on reducing the activation energy barrier are still under debate.[12-13]

Focusing on the transition state (TS),Wolfenden in 1969, laid the foundation for the theory of TS analogy as it pertains to tight binding competitive inhibitors that mimic the structure of the

enzyme•substrate complex at the point where it traverses the reaction TS.[14] In particular, an enzyme's catalytic proficiency (CP = K_{TS}) can be expressed as its pseudo-equilibrium constant for binding the reaction TS (Scheme 1),[15-16] which is associated with the equilibrium constants for substrate binding (K_m) and the two pseudo-equilibrium constants between ground state and the TSs for enzyme-catalyzed and uncatalyzed reactions ($K_{cat}^{\ddagger}$ and $K_{uncat}^{\ddagger}$, equation 1).[16] This reduction in reaction free energy is given by the second-order rate constant (k_{cat}/K_m) divided by the spontaneous rate constant for the water-promoted reaction (k_{uncat}).[15-18] Bartlett expanded on these ideas and showed that a linear free energy relationship (LFER) should exist between the rate constants for enzymatic proficiency [$k_{cat}/(K_m \times k_{uncat})$] and the dissociation constant of a competitive transition state analog (TSA) inhibitor (K_i, equation 2).[16, 19-20]



Scheme 1. Thermodynamic cycle for the relationships between K_{TS} and K_m, $K_{cat}^{\ddagger}$ and $K_{uncat}^{\ddagger}$.

$$K_{TS} = \frac{K_m K_{uncat}^{\ddagger}}{K_{cat}^{\ddagger}} \quad \text{equation 1}$$

$$K_i = \delta K_{TS} = log\left(\frac{K_m}{k_{cat}}\right) + log(\delta k_{uncat}) \qquad \text{equation 2}$$

In addition to producing catalytically proficient enzymes, natural selection has generated natural products that increase the likelihood of an organism surviving.[21] Some of these secondary metabolites operate by competitively inhibiting critical enzymes present in rival organisms. Such compounds are often referred to as TSA inhibitors based on the assumption that their structures

resemble the TS for substrate turnover at the key chemical step. Yet for most of these compounds experimental evidence to support these assumptions is missing.

Several LFER studies have been performed on glycoside hydrolases to determine if competitive inhibitors are TSAs.[20] However, this approach has not been applied to mechanism-based glycoside hydrolase (GH) inhibitors. Such inhibitors covalently label an enzymatic active site residue and are increasingly used as activity-based probes in chemical biology studies.[22-24] These mechanism-based inhibitors are assumed to react via TSs that are analogous to those located along the reaction coordinate for enzyme-catalyzed glycosidic C–O bond cleavage, but this has not been verified experimentally. Indeed, the design of more proficient and selective mechanism-based probes hinges on improving our understanding of the structures, electron density and, in particular, how positive charge is delocalized at the TS structures for GH-catalyzed covalent bond formation with the inhibitor.

The study herein focuses on three classes of mechanism-based covalent inhibitors that require TS charge stabilization, with one of them additionally benefiting from the release of epoxide ring strain. We present a robust LFER approach to correlate, for the first time, the catalytic efficiencies (k_{cat}/K_m) for hydrolysis of glycoside substrates with the second-order rate constants for either i) covalent inhibitors where the intermediate hydrolyses (k_{cat}/K_m), or ii) covalent inactivators, where it does not (k_{inact}/K_i). We used a natural unactivated pyranoside substrate (**1**), a chromogenic substrate (**2**), and five mechanism-based covalent inhibitors (**3**–**7**, Figure 1). The structural motif in inhibitor **3**, a 2-deoxy-2-fluoroglycoside, was designed based on the notion of TS destabilization by incorporation of a fluorine,[25] inhibitors **4**–**6** are structural analogs of the natural product valienamine,[26-27] which is a sub-structure of acarbose, an antidiabetic drug that may[28] or may not be[29] a TSA competitive inhibitor; and **7** is the α-*galacto*-stereoisomer of cyclophellitol, a natural

product inactivator of glycoside hydrolases.[30] We show that the retaining α-galactosidase from *Thermotoga maritima* (*Tm*GalA) stabilizes the positive charge development at the hydrolytic and covalent inhibition TSs equally for pyranoside and allylic carbasugars that have identical leaving groups (both S_N1-like mechanisms). However, in contrast to glycoside substrates the fluorine containing mechanism-based inhibitors **3** and **6** react via concerted mechanisms (S_N2), while the enzyme inactivation by the structurally-strained *galacto*-cyclophellitol involves TS traversal with a significantly lower degree of charge stabilization. To strengthen these conclusions, we computed free energy landscapes and their associated galactosylation and pseudo-galactosylation transition states using molecular dynamics (MD) simulations with Quantum Mechanics / Molecular Mechanics (QM/MM) multiscale potentials with the whole solvated enzyme.

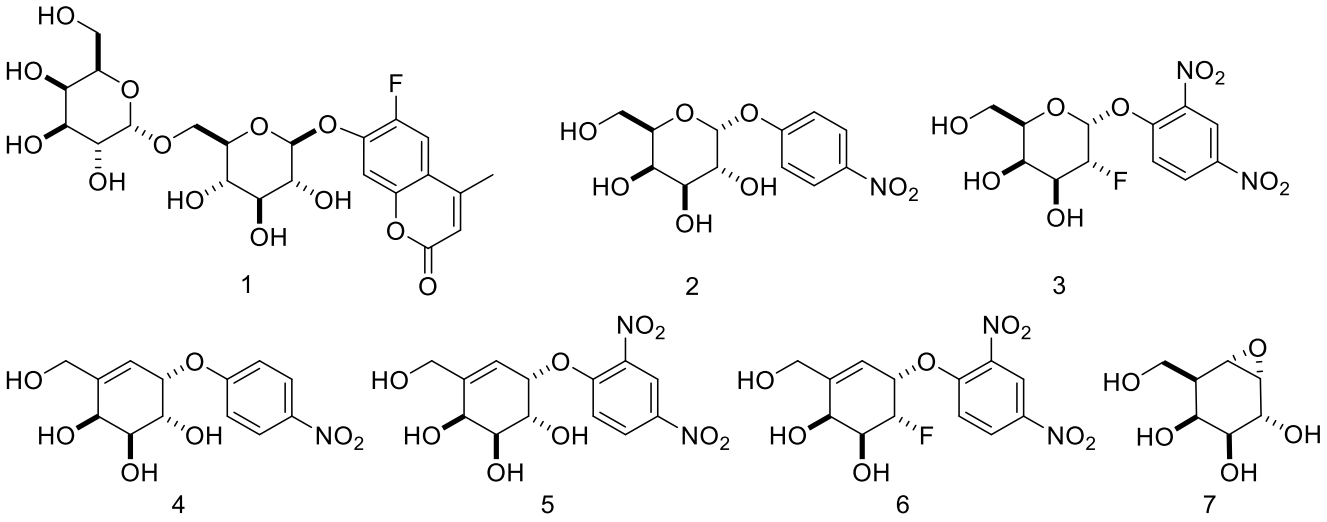


Figure 1. Substrates and inhibitors used in this study. 6-fluoro-4-methylumbelliferyl β-D-melibioside **1**, 4-nitrophenyl α-D-galactopyranoside **2**, 2,4-dinitrophenyl 2-deoxy-2-fluoro-α-D-galactopyranoside **3**, aryl 5,5a-didehydro-5a-carba-β-L-*arabino*-hexopyranosides **4** and **5**, 2,4-dinitrophenyl 2-deoxy-5,5a-didehydro-2-fluoro-5a-carba-β-L-*arabino*-hexopyranoside **6**, and α-*galacto*-cyclophellitol **7**.

Experimental Details

Mutagenesis, Gene Expression, and Protein Purification

A series of 14 *Tm*GalA variant enzymes were generated, each variant possesses one non-catalytic amino acid residue change in the substrate binding pocket. Mutant plasmids required to produce variant *Tm*GalA were generated using the QuikChange site-directed mutagenesis method using the forward primers listed in Table S7 (Supporting Information) and their complimentary reverse primers.[31] Plasmid clones with WT and mutant *gal*A genes were transformed into *E. coli*. Plasmid DNA was isolated from single bacterial colonies and the expected mutation was verified by DNA sequencing. Recombinant *Tm*GalA and its derived mutants were expressed and purified as previously described.[32] All derived variants and WT *Tm*GalA were subjected to thermal denaturation by differential scanning fluorimetry (DSF) to determine their apparent melting temperatures (T_m).

Differential scanning fluorimetry

Differential Scanning Fluorimetry was performed using an Applied Biosystems StepOnePlus Real-Time PCR machine equipped with a 96-well heating block. Reactions were performed in MicroAmp 96-well optical PCR plates (Applied Biosystems) at a final volume of 20 µL per well. Each reaction contained between 0.3–0.5 mg/mL protein, 1· Protein Thermal Shift Dye (Thermo Fisher Scientific) in 1× Protein thermal shift buffer (Thermo Fisher Scientific) and sealed with MicroAmp Optical Adhesive Film (Thermo Fisher Scientific). The reactions were kept on ice prior to the assay. The temperature was ramped in two steps. In step 1, the temperature was ramped at 100% ramp rate (equivalent to 1.6 °C/s) from room temperature to 25 °C, and in step 2 the temperature was ramped from 25 to 99 °C at 1% ramp rate (equivalent to 0.05 °C/s). Changes in

fluorescence intensity were followed with the 580 nm excitation/ 623 nm emission filter set. Each reaction was performed in duplicate, and data were analyzed using QuantStudio Design and Analysis software. All obtained melt profiles were fit to a Boltzmann sigmoidal curve using Graphpad Prism 7 and apparent T_m values were obtained from the inflection point of the curve.

Measurement of Michaelis-Menten kinetic parameters

The k_{cat}/K_m values for the pH-rate profiles of *Tm*GalA with 4-nitrophenyl α-D-galactopyranoside and 6-fluoro-4-methylumbelliferyl β-D-melibioside, were determined by using a continuous assay. Michaelis–Menten kinetic parameters for the hydrolysis of the substrates were determined from a minimum of eight initial rate measurements using a concentration range of at least K_m /3 to 3× K_m. The progress of each reaction was monitored continuously at 400 nm for 4-nitrophenol and 365 nm excitation and 445 nm emission wavelengths for 6-fluoro-4-methylumbelliferone using a Biotek Neo2 multimode plate reader. Each reaction mixture was prepared by addition of the appropriate buffer (50mM MES buffer for pH 6.5, 50 mM HEPES buffer for pH 7.0–7.5, 50 mM glycylglycine buffer for pH 8.0–8.5, 50 mM glycine-NaOH buffer for pH 9.0–10.0), 0.1% BSA, substrate, and enzyme. All reactions were performed at 37 °C.

All other kinetic measurements (used to construct the LFER plots) were performed at pH 8.0 in 50 mM glycylglycine, 0.1% BSA, and 37 °C and monitored as noted above unless specified otherwise. The second-order rate constants (k_{cat}/K_m) for the hydrolysis of the substrates 6-fluoro-4-methylumbelliferyl β-D-melibioside **1**, 4-nitrophenyl α-D-galactopyranoside **2**, 2,4-dinitrophenyl 2-deoxy-2-fluoro-α-D-galactopyranoside **3**, and carbasugars **4**, **5**, and **6** by WT and the nine variants of *Tm*GalA were determined from a minimum of six initial rate measurements. The release of 4-nitrophenol, and 2,4-dinitrophenol were followed at 400 nm using a Cary 300 UV-vis spectrometer equipped with a Peltier device temperature controller. For the 6-fluoro-4-

methylumbelliferyl β-D-melibioside substrate (**1**), a coupled assay was used to follow the progress of each reaction by using an optimized excess amount (from 20 μg/mL at a pH of 6.5 to 400 μg/mL at a pH of 10.0) of β-glucosidase from almonds (Sigma). The release of 6-fluoro-4-methylumbelliferyl β-D-glucoside from the melibioside substrate was continuously monitoring by following the β-glucosidase-catalyzed release of 6-fluoro-4-methylumbelliferone at an excitation wavelength of 365 nm with a 445 nm emission using a Biotek Neo2 multimode plate reader.

The calculated rate constants (k_{obs}) for each reaction as a function of substrate concentration were fit to a standard Michaelis–Menten equation using a standard nonlinear least-squares program (Prism 8.0).

Measurement of covalent inactivation rate constants for α-*galacto*-cyclophellitol

The kinetic parameters for the pH-rate profile for the inactivation of WT *Tm*GalA by α-*galacto*-cyclophellitol **7** were determined using a classical dilution assay that involved preincubation of the enzyme with varying concentrations of inhibitor at 37 °C in appropriate buffer (MES buffer for pH 6.5, HEPES buffer for pH 7.0–7.5, glycylglycine buffer for pH 8.0–8.5, glycine-NaOH buffer for pH 9.0–10.0) with 0.1% BSA. The remaining enzyme activity was measured periodically by removing an aliquot (10 μL) and adding it to a 490 μL pre-equilibrated solution (37 °C) containing 4-nitrophenyl α-D-galactopyranoside (250 μM) in the same buffer conditions used for preincubation. The kinetic parameters for inactivation of WT and nine variants of *Tm*GalA by α-*galacto*-cyclophellitol **7** used to construct the LFER plots were performed similarly at pH 8.0 in glycylglycine buffer. The pseudo first-order rate constants for inactivation (k_{obs}) were determined by fitting the absorbance versus time data to a standard first-order rate equation. The first- and second-order rate constants (k_{inact} and k_{inact}/K_i) for inactivation were calculated by fitting the rate

constant data versus the inactivator concentration to a standard Michaelis–Menten equation using a standard nonlinear least-squares program (Prism 7.0).

Computational details

a) Molecular model set up: The starting structure for the computer simulations of *Thermotoga maritima* α-galactosidase (*Tm*GalA) was adapted from the X-Ray structure in Protein Data Bank under code 5M12[33] for substrate **3** and 6GTA[34] for inhibitor **4**. The structure of *Tm*GalA in complex with an inhibitor was modified to correspond to substrate **3** and inhibitor **4** at the active site. The missing atoms of Lys77 and Glu80 residues in the X-ray structure were incorporated with Accelrys Discovery Studio Visualizer ver. 4.5.[35] Charges and parameters for **3** and **4** were calculated using Antechamber software package[36] with a general AMBER force field (GAFF),[37] listed in Tables S8 and S9 (Supporting Information). Hydrogen atoms were added to the protein structure using the tLEAP[37] module of Amber Tools program. The protonation state of titratable amino acids at pH 7.4 was previously determined using pK_a values calculated with PROPKA ver. 3.1,[38] which is available on the PDB2PQR server.[39] The computed results indicate that residues Asp387 and Glu459 are present in their protonated form. Furthermore, residue Glu224 was protonated to allow more favorable hydrogen bonding between it and neighboring Asp221. Additionally, after structural inspection two histidines, His30 and His273 were protonated at the δ-position, while all others were protonated on the ε-nitrogen. The total charge of the system was neutralized by incorporation of 17 sodium cations (Na^+) in the most electrostatically favorable positions. Subsequently, the system was placed in orthorhombic box of TIP3P[40] water molecules with size of 90 × 101 × 80 $Å^3$ for substrate **3** and 90 × 97 × 72 $Å^3$ for inhibitor **4**; geometries of the remaining water molecules were then optimized. The full system consists of the protein (8453 atoms), the substrate/inhibitor (37 atoms in **E•3** and 36 atoms in **E•4**), 17 sodium ions, and solvation water

molecules (19467 residues/58401 atoms for glycoside **3**, 15967 residues/47901 atoms for carbasugar **4**).

b) Molecular dynamics (MD): Using the NAMD molecular dynamics program,[41] the prepared computational model was heated from 0 to 310 K with 0.001 K temperature increments, the systems was then equilibrated for the **E•S** or the **E•I** state using the Langevin-Verlet algorithm,[42] and finally 10 ns of classical MD simulation (at temperature 310 K) was carried out in the NVT ensemble. Periodic boundary conditions (PBC) using the particle mesh Ewald method were applied. To improve calculation time, a nonbonding interaction cut-off was applied using a smooth switching function between 14.5 to 16.0 Å. The time dependence of RMSD and total energy confirms that the systems are equilibrated after 10 ns of the MD simulation (Figures S17 and S18, Supporting Information).

c) QM/MM simulations: In this work, an additive hybrid QM/MM scheme was employed for the construction of the total Hamiltonian where the total energy is obtained from the sum of each contribution to the energy (Equation 3). Here, E_{QM} describes the atoms in the QM part, $E_{QM/MM}$ defines the interaction between the QM and MM regions and E_{MM} describes the rest of the MM part. As shown in Figure S19 (Supporting Information), the two active site aspartate residues, Asp327 and Asp387, together with the full substrate **3** or inhibitor **4** were described at the QM level in the QM/MM simulations. To saturate the valence of the QM/MM frontier atoms, two link atoms[43] were inserted where the QM/MM boundary intersected covalent bonds in the positions indicated on Figure S18 (Supporting Information). The Austin Model 1 (AM1)[44] semiempirical method and the Minnesota Functional M06-2X[45] with the standard 6-31+G(d,p) basis set were used to treat the QM sub-set of atoms, as implemented in Gaussian09 program.[46] The protein and solvent water molecules were described by AMBER and TIP3P force fields, as implemented in the

fDynamo library.[47-48] For glycoside **3**, the structure after 9.4 ns of MM MD simulation was used to run the QM/MM calculations while for inhibitor **4** the last structure after 10 ns MM MD simulation was used in the QM/MM calculations. To reduce time of calculations, positions of atoms presented beyond 20 Å from substrate **3** or inhibitor **4** were fixed.

d) Free Energy Surfaces: To generate the free energy surfaces (FESs), in terms of 2D-potential of mean forces (PMFs), potential energy surfaces (PESs) were computed first to generate the required grid of structures. The two distinguished reaction coordinates correspond to the leaving group oxygen-anomeric carbon distance [d(C1–O^{LG})] and to the nucleophile aspartate oxygen-anomeric carbon distance [d(C1–O^{Asp327})] (Scheme S3, Supporting Information). Then, FESs were generated using the Umbrella Sampling approach[49-50] combined with the Weighted Histogram Analysis Method (WHAM).[51] For each structure of the grid generated to construct the PES, MD simulations were performed with a total of 5 ps of equilibration (with 1 fs time step) and 20 ps of production at 310 K using the Langevin-Verlet algorithm[42] with a time step of 0.5 fs and an umbrella force constant of 5000 $kJ \cdot mol^{-1} \cdot Å^{-2}$ to constrain the key interatomic distances defining the reaction coordinates. Because many structures must be sampled during the QM/MM MD simulations, the AM1 Hamiltonian was selected to describe the QM sub-set of atoms.

e) Spline corrections: To improve the quality of the FESs due to possible limitations associated with the semiempirical method, the FESs were corrected at DFT/MM level. Thus, based on the work of Truhlar and co-workers for reactions in solution[52-54] a spline under tension[55-56] was used to interpolate this correction term at any value of the distinguished reaction coordinate, ξ_1 and ξ_2 selected to generate the 2D FES.[57-58] A continuous energy function is used to obtain the corrected PMFs (Equation 4) where S is the two-dimensional spline function and ΔE_{LL}^{HL} is the difference between the energies obtained at low-level (LL) and high-level (HL) of theory of the QM part. The

AM1 semiempirical Hamiltonian was used as LL method, while the hybrid M06-2X[45] functional with the standard 6-31+G(d,p) basis set was selected for the HL energy calculation. These calculations were carried out using the Gaussian09 program.[46] From the DFT corrected FESs, structures were selected for GS/TS localization with Baker's algorithm[59] at M06-2X/OPLS-AA/TIP3P level, using the 6-31+G(d,p) basis set for the treatment of the QM subset of atoms. The fDynamo library, in combination with Gaussian 09 were used for these calculations. Localized TS structures were further characterized confirming existence of one imaginary frequency and the minimum energy path that was traced to reactants and products using the Intrinsic Reaction Coordinate (IRC) method.[60]

$$E_{QM/MM} = E_{QM} + E_{QM/MM}^{elect} + E_{QM/MM}^{vdW} + E_{MM} \qquad \text{Equation 3}$$

$$E = E_{LL/MM} + S[\Delta E_{LL}^{HL}(\xi_1, \xi_2)] \qquad \text{Equation 4}$$

Results and Discussion

Although enzymes are flexible entities and protein vibrational modes are proposed to contribute to catalytic function [7-8] the impact of such effects in reducing reaction coordinate free energy barriers is under debate.[12-13] Glycoside hydrolase catalysis involves positively charged pyranosylium ion-like transition states (TSs); accordingly, we based the analysis of our LFERs on the computational insights that enzymatic TSs[61-62] – relative to the reaction TS in solution[61-64] – are best stabilized by electrostatic and hydrogen bonding interactions.

With regard to the catalytic mechanism of the retaining α-galactosidase *Tm*GalA (the catalyst studied herein), Comfort et al. showed that the second-order rate constants for the enzyme-catalyzed hydrolysis of aryl α-D-galactopyranosides (k_{cat}/K_m) are independent of the leaving group ability of the substituted phenol.[31] Such observations are consistent with a non-chemical step as

the limiting step for catalysis; this step occurs prior to cleavage of the glycosidic bond.[65] To compare the TSs for the chemical bond cleavage step that gives the glycosylated enzyme intermediate from a Michaelis complex, we (i) modified the reaction conditions for *Tm*GalA-catalyzed hydrolysis of substrates and then (ii) identified conditions for which glycosidic C–O bond cleavage is kinetically significant. Specifically, we synthesized an unactivated natural disaccharide substrate (**1**, Scheme S1, Supporting Information) in which cleavage of the unactivated glycosidic bond can be monitored using a coupled enzyme assay (Scheme S2, Supporting Information), and we measured the second-order rate constants (k_{cat}/K_m) for the cleavage of **1** and 4-nitrophenyl α-D-galactopyranoside **2** as a function of pH. Figure 2 shows the logarithms of the measured rate constants (k_{cat}/K_m) for these two *Tm*GalA-catalyzed reactions with values determined for pH conditions between 6.5 and 10.0 (rate constant values are given in Table S1, Supporting Information).

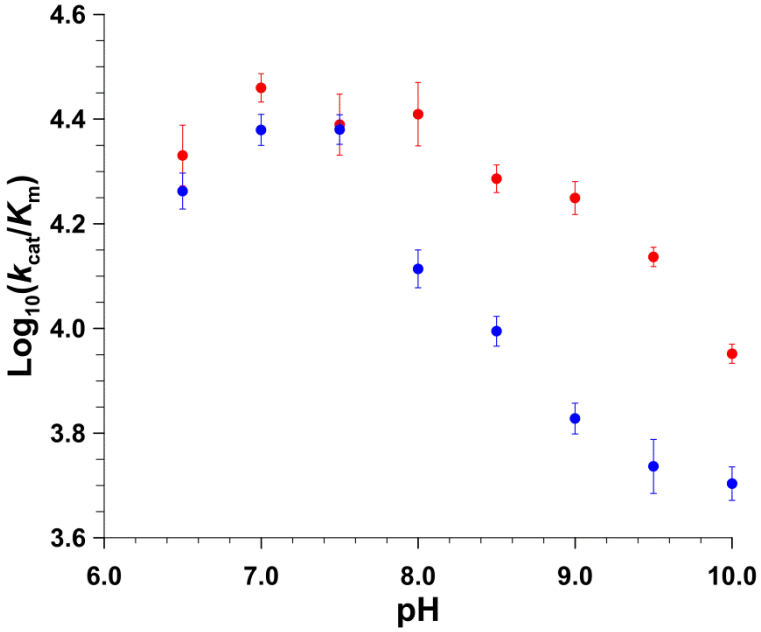


Figure 2. $Log_{10}(k_{cat}/K_m)$ versus pH plots for the wild-type *Tm*GalA-catalyzed hydrolyses of substrates **1** (blue circles) and **2** (red circles).

We found a similar lack of dependence on leaving group ability at pH values between 6.5 to 7.5, however, as the pH is raised above 7.5, there is a clear difference in the k_{cat}/K_m values for the *Tm*GalA-catalyzed hydrolyses of **1** and **2** (Figure 2). We therefore followed the reactions of *Tm*GalA at pH 8.0, a value where chemistry is at least partially rate limiting and enzyme activity with our natural substrate requires a protonated enzymatic residue (likely the general-acid catalytic residue D387) to traverse the TS for glycosidic C–O bond cleavage. A limitation of this experiment is that it provides no information on whether transfer of this proton to the aglycone initiates before or after the system crosses the glycosylation TS.

We considered two major experimental designs for a protocol to perform Bartlett-type LFER correlations between substrates and mechanism-based inhibitors, those that measure: i) wild-type enzyme-catalyzed reactions using a series of substrate and inhibitor dyads, or ii) several variant enzyme-catalyzed reactions on a single substrate and inhibitor pair.[20] We opted not to probe LFERs with a panel of substrate and inhibitor pairs, because structural changes are generally made to a single functionality within the panel of substrate/inhibitor dyads and this normally gives derived LFER slopes of close to one for TS mimics, with k_{uncat} either assumed to be invariant[66-67] or explicitly included in the fit.[68] Such experiments generally modulate a limited subset of substrate/inhibitor interactions that occur within the enzymatic active site and report on whether these particular interactions are used by the catalyst to stabilize the enzymatic TS. However, no information is provided on whether other structural elements of the inhibitor mimic TS stabilization from the corresponding active site regions in the enzyme-catalyzed reaction of substrates.

We therefore performed Bartlett LFER-type studies using multiple variant *Tm*GalA enzymes in which amino acid changes are commonly made to residues that surround and interact with

substrates bound in the active site, and thus the catalytic parameters recorded from a single substrate/inhibitor pair are the result of altering a range of heterogeneous TS stabilizing interactions. Such experiments often yield good LFER correlations with the data points displaying intrinsic scatter around the best linear fit.[28, 69-71]

For these experiments we prepared a series of variant *Tm*GalA enzymes, each of which contained a single amino acid residue change. Making one amino acid change per enzyme variant, we made changes to a total of nine active site residues located in the enzyme region that surrounds bound substrate; the two key catalytic aspartic acids (D327 and D387) were not modified. We made two variants for some residues and in all, generated fourteen variant *Tm*GalA enzymes. We purified the fourteen enzyme variants and then tested each one for temperature stability and catalytic activity (Table S2, Supporting Information). The active site locations of the amino acid residues that were modified by site directed mutagenesis are shown in Figure 3 (drawn using the coordinates from PDB file 6GTA);[72] the changed residues that gave active *Tm*GalA enzyme are shown in panel A, while the three mutagenized amino acid residues that gave inactive enzyme are shown in panel B (D220, K325, and D427). Notably, the three residues that yielded inactive *Tm*GalA are involved in H-bonding to the substrate via hydroxyl groups on carbon-4 (D220 and K325) and carbon-3 (D427 via a H-bond chain through Y191).

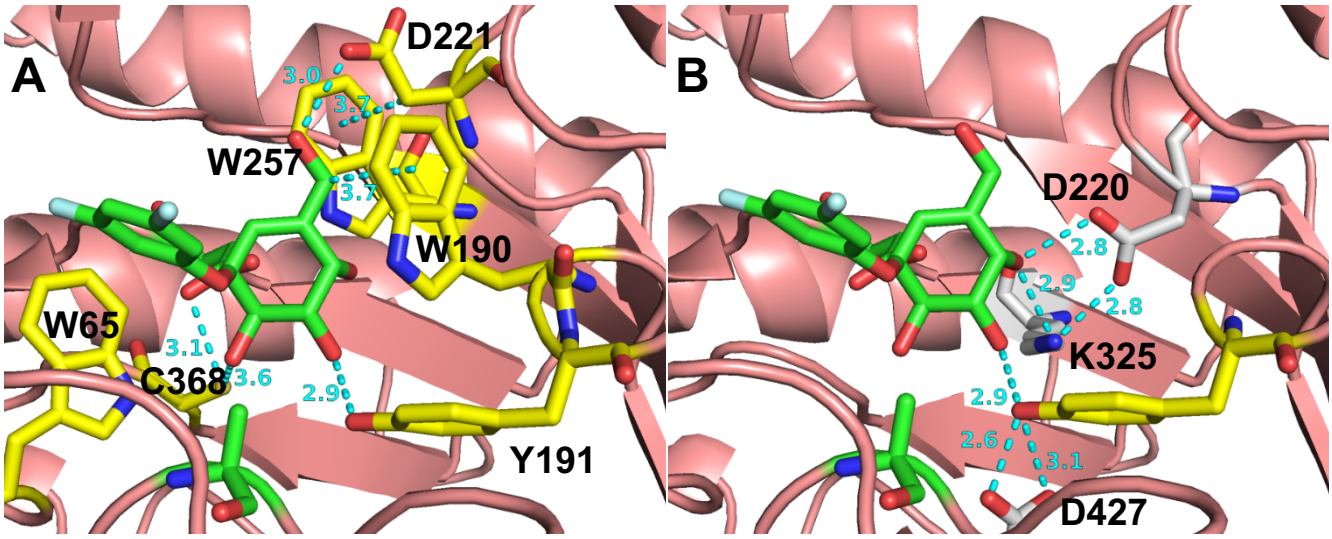


Figure 3. Positions of amino acid substitutions in the variant *Tm*GalA enzymes, structure is of the general-acid catalyst variant (D387A) of *Tm*GalA that has a bound intact carbasugar containing a 3,5-difluorophenyl aglycone. Panel A) shows the positions of substitution that resulted in partially active enzymes (yellow carbon atoms); and panel B) shows the three mutation sites that gave severely compromised catalysts (gray carbon atoms). Cyan dashed lines and distances (Å) show key interactions, H-bonding, and C–H to π.

LFERs for melibioside natural substrate analogue and aryl galactosides. Our initial Bartlett-type LFER involved examining whether a correlation exists between the rate constants for the catalyzed hydrolyses of a natural substrate analogue (**1**) and a galactoside with an activated aryloxy leaving group (**2**) using the nine active variant *Tm*GalA enzymes (Figure 4, Table S3 Supporting Information). For the LFER correlation between the log(K_m/k_{cat}) values for *Tm*GalA-catalyzed hydrolyses of **1** and **2**, we excluded the wild-type enzyme data point in this and all subsequent plots involving the catalyzed reactions of **2**, because cleavage of the C–O glycosidic bond in 4-nitrophenyl α-D-galactopyranoside is not kinetically significant under our assay conditions (Figure

2). Interestingly, the two variants of tryptophan-65 (W65F and W65Y)—a residue that interacts with the aglycone—show larger rate reductions for the β-D-glucopyranoside leaving group in **1** than for the 4-nitrophenoxide leaving group produced in the reactions of **2**. We note that this active site tryptophan is located in the aglycone binding site (Figure 3A), which suggests to us that W65 assists with the departure of carbohydrate aglycones found in natural substrate(s); indeed, such tryptophan–carbohydrate interactions are common binding motifs found in lectins and glycoside hydrolases.[73-74] We analyzed the LFER correlation between the kinetic parameters for the *Tm*GalA-catalyzed hydrolyses of **1** and **2** (Figure 4A) by fitting the data points for the seven other variant enzymes to probe for TS similarities and we obtained a slope of 0.88 ± 0.14 (r^2 = 0.8898). Also, as amino acid residue W257 interacts with one of the two oxygen atoms from the active site nucleophile D327 (Figure 3), we expect a greater rate reduction in the rate of reaction for cleavage of an anomeric C–O bond with an unactivated aglycone if the H-bonding interaction between these two residues (D327–W257) is important in positioning the either nucleophilic attack or electrostatic TS stabilization. This is indeed the case for the catalyzed hydrolysis of the unactivated substrate (**1**), as the linear correlation (six data points) has a slope of 0.85 ± 0.08 (r^2 = 0.9652). Regardless, our derived slope (< 0.9) for the LFER correlation of *Tm*GalA-catalyzed hydrolysis of **1** and **2** is consistent with catalyzed hydrolysis of **2** having an earlier TS with less positive charge

development, and thus a lower degree of electrostatic TS stabilization, than the corresponding TS for **1**.

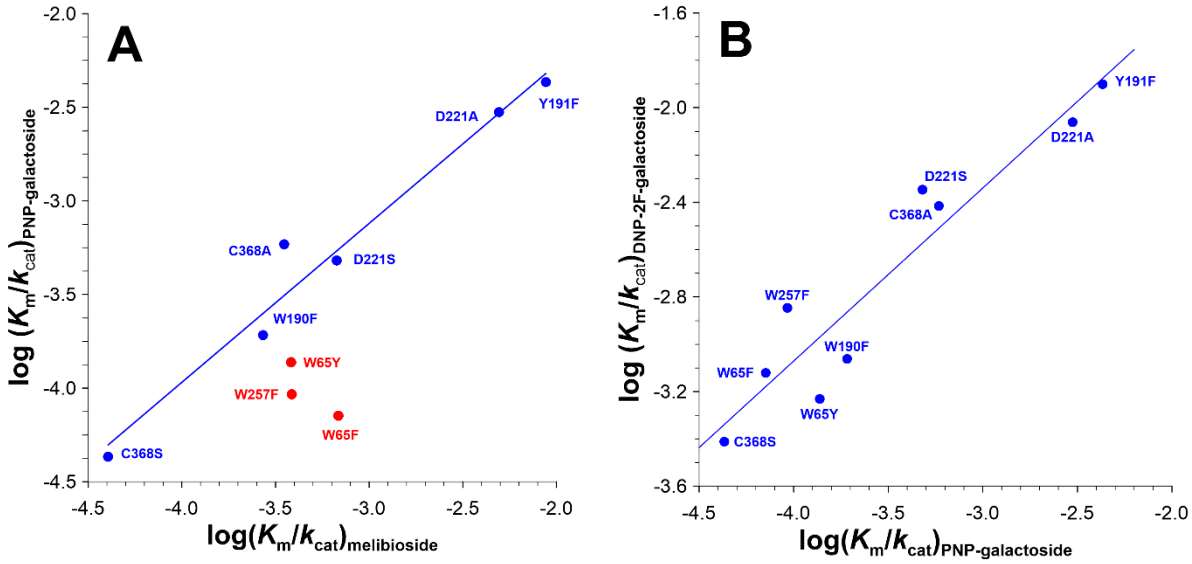


Figure 4. Bartlett-type LFER plots showing correlations between the second-order rate constants for *Tm*GalA-catalyzed hydrolyses. Panel A) substrates **1** and **2**; and panel B) glycosides **2** and **3**.

Bartlett LFERs for a 2-deoxy-2-fluoroglycoside mechanism-based inhibitor. Given the widespread use of C2-fluorine substituted glycosides as mechanism-based GH inactivators,[25, 75-76] we applied the LFER approach to examine how the substitution of the C2-hydroxyl group by a fluorine atom perturbs the enzymatic glycosylation TS. For our analysis, we plotted the measured values for log(K_m/k_{cat}) for the *Tm*GalA-catalyzed hydrolysis of 4-nitrophenyl α-D-galactopyranoside (**2**) against those for 2,4-dinitrophenyl 2-deoxy-2-fluoro-α-D-galactopyranoside (**3**). The resultant Bartlett LFER plot has a slope of 0.79 ± 0.09 ($r^2 = 0.8972$, Figure 4B, Table S3 Supporting Information), indicating a good mechanistic correlation for the *Tm*GalA-catalyzed hydrolysis of these two substrates. Interestingly, the differences in TS stabilization for the W65 variants that we noted above for the melibioside substrate **1** are not apparent in this LFER; the

similarity in TS stabilization between the W65 variants likely results from both pyranoside substrates having activated electron deficient aromatic leaving groups, 4-nitrophenoxide for **2** and 2,4-dinitrophenoxide for **3**, that interact with tryptophan-65 in the aglycone binding pocket during enzyme catalysis. We expected that the introduction of a C2-fluorine-substitution (a change that raises the free energy of the pyranosylium ion) would make the TS later (Hammond effect so that more positive charge is developed at the TS).[77] However, the less than unit LFER correlation (slope = 0.79 ± 0.09) indicates a change in mechanism, from S_N1-like to S_N2, has occurred and now enzyme-catalyzed C–O bond cleavage occurs with assistance by the active site nucleophile (Asp327) with a concomitant lower degree of TS charge development in the reactions of **3**.

To add weight to this conclusion, we computed a viable TS for the *Tm*GalA-catalyzed glycosylation of **3** using QM/MM MD simulations with the whole solvated enzyme system (See Methods in Supporting Information). Our computed free energy surface (FES) for the *Tm*GalA-catalyzed glycosylation reaction of **3** (Figure 5A) shows that covalent bond formation to the active site nucleophile (Asp327) occurs via an exploded S_N2 TS (Figure 5B), and that no discrete 2-deoxy-2-fluoro-D-galactopyranosylium ion intermediate is evident along the reaction coordinate. In other words, the *Tm*GalA-catalyzed hydrolysis of **3** involves a tighter TS with more S_N2 character and less positive charge development than for the corresponding reaction of **2**, a conclusion that is consistent with the less than unit LFER slope of 0.79 ± 0.09.

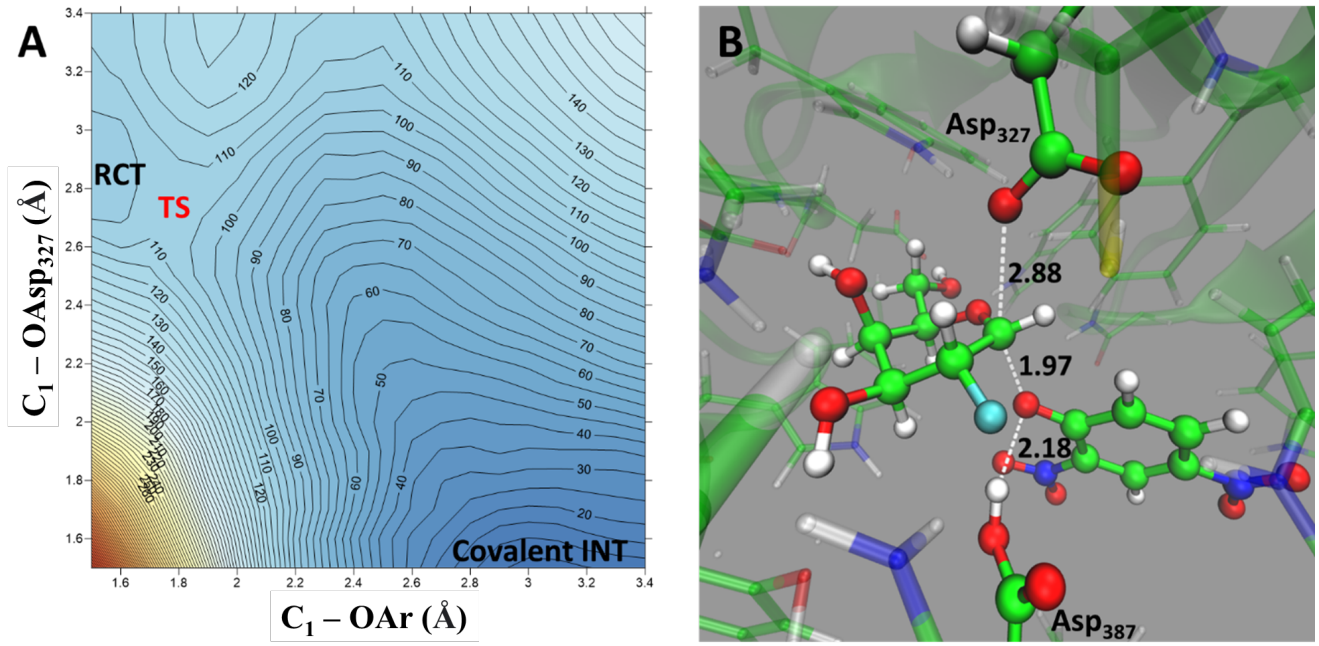
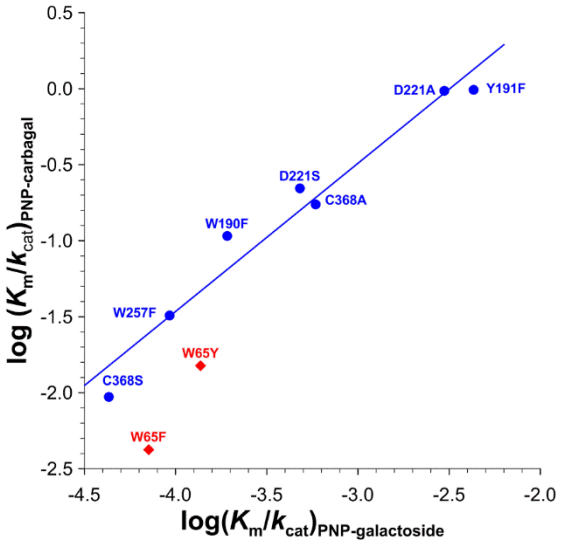


Figure 5 Computational results of the catalytic itinerary for the hydrolysis of **3** catalyzed by *Tm*GalA. Panel A) M06-2X:AM1/MM FES corresponding to the α-galactosidase-catalyzed cleavage of the C−OAr bond in **3** from the Michaelis complex (RCT), concomitant with the nucleophilic attack by Asp327, to give covalently bound glycosylated enzyme (Covalent INT). Energies of iso-energetic lines in kJ·mol^{-1}. Activation and reaction free energies for the reaction of the Michaelis complex (RCT) to the covalent intermediate are 6.0 and –99.4 kJ·mol^{-1}, respectively; and panel B) Detail of the active site in the transition state optimized at M06-2X/MM level. Key distances in the located transition state are indicated as dashed white lines and reported in Å.

Bartlett LFERs for allylic carbasugar mechanism-based covalent inhibitors. Next, we measured the second-order rate constants for the panel of variant *Tm*GalA-catalyzed reactions with allylic carbasugar **4**, **5**, and **6** (Table S4, Supporting Information). Figure 6 shows the resulting



Bartlett-type LFER correlation for the enzyme-catalyzed log(K_m/k_{cat}) values for reactions of galactoside **2** and carbasugar **4**, both of which possess identical 4-nitrophenoxide leaving groups.

Figure 6. Bartlett-type LFER plot between the second-order rate constants for the catalyzed hydrolyses of substrate **2** and covalent inhibitor **4**. The blue line shows the best linear fit to the blue circles, while the W65 variants (red diamonds) are omitted from the fit, see text for full details.

Based on literature precedence for several GH families (which cover enzyme groups that include β-glucosidases, β-mannosidases, β-galactosidases, α-glucosidases, and α-galactosidases), a consensus was reached that glycosidic bond cleavage occurs via a $^4H_3$ pyran ring TS conformation;[78-79] we suggest that *Tm*GalA catalysis involves substrates traversing an equivalent $^4H_3$ TS. Intriguingly, the W65 variants again show deviations from the best linear fit. We propose that the interactions between the aglycone and tryptophan-65, which facilitate 4-nitrophenoxide departure, result in a greater TS stabilization during the requisite distortion of galactoside **2** (solution $^4C_1$ ground state to $^4H_3$ TS)[78-80] than in the allylic carbasugar **4** (solution $^2H_3$ ground state to E_3 TS)[32, 72] As a result, the W65 variants are markedly less active with pyranoside substrate **2**,

than with the corresponding covalent inhibitor **4** (Figure 6), with the least electron-rich phenylalanine variant exhibiting the largest difference in activity between cleavage of pyranoside **2** and carbasugar **4**.[81] We note that the excellent correlation for the seven other variant enzymes displays a slope indistinguishable from unity (0.97 ± 0.09, $r^2 = 0.9591$), an observation that is consistent with *Tm*GalA being able to stabilize a similar degree of positive charge development at pyranosylium ($^4H_3$) and allylic (E_3) cation-like TSs that both lead to the formation of covalent intermediates.

To provide further insight into the reaction coordinate for *Tm*GalA-catalyzed covalent inhibition by carbasugar **4**, we computed the FES for the pseudo-glycosylation step. Figure 7A shows the computed FES, while the TS for cleavage of the pseudo-glycosidic bond in **4**, which occurs via a S_N1-type mechanism in which the allylic cation is in a shallow energy well, is displayed in Figure 7B. The enzymatic reaction path passes over the TS and through this shallow energy well to give the covalent adduct directly. For this reaction, the TS – which has an almost fully developed positive charge – is S_N1-like, with nucleophilic attack occurring as the reaction coordinate exits the shallow energy well of the cation. This S_N1-type reaction path is similar to that for the inverting GH15 amylase, where the catalyzed hydrolysis of the natural substrate analogue α-D-glucopyranosyl fluoride proceeds through a pyranosylium ion intermediate.[82] That is, both the pyranoside and carbasugar reactions proceed via S_N1-like mechanisms with a similar degree of positive charge stabilization at the glycosylation and pseudo-glycosylation TSs.

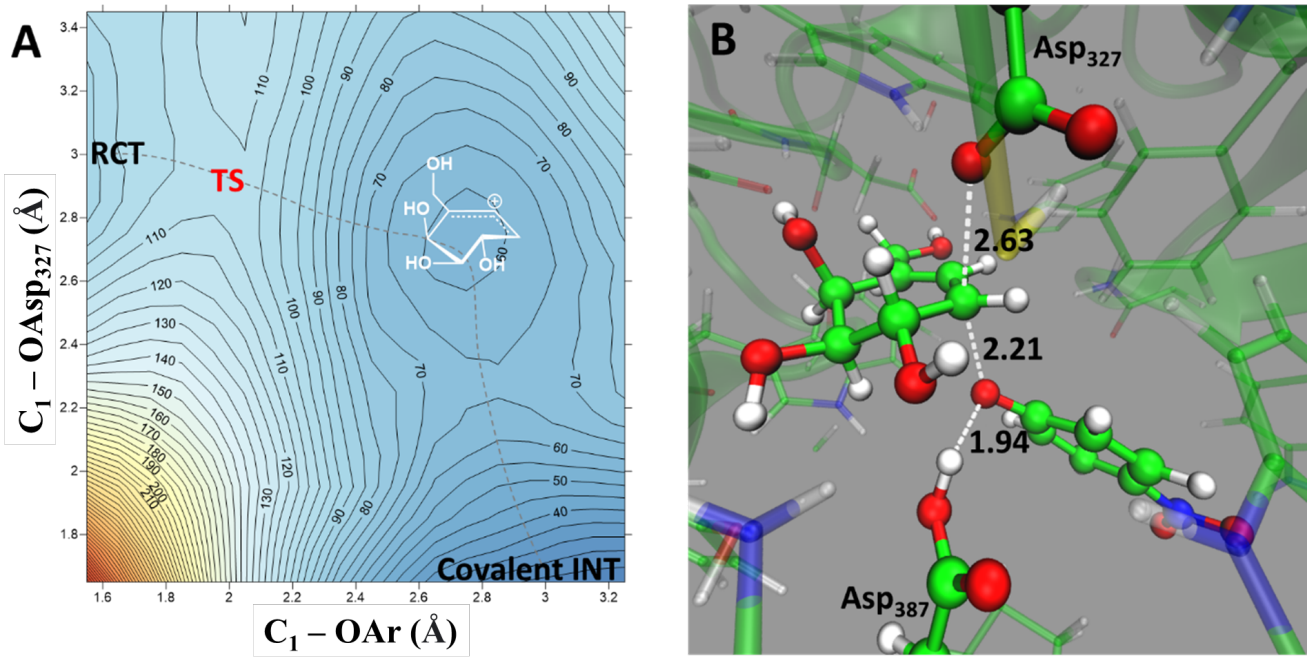


Figure 7. Computational results of the catalytic itinerary for the hydrolysis of **4** catalyzed by *Tm*GalA. Panel A) M06-2X:AM1/MM FES corresponding to the *Tm*GalA-catalyzed cleavage of the C−OAr bond in **4** from the Michaelis complex (RCT), followed by the nucleophilic attack by Asp327, to give covalently bound glycosylated enzyme (Covalent INT). Energies of iso-energetic lines in kJ·mol^{-1}. Activation and reaction free energies for the reaction of the Michaelis complex (RCT) to the covalent intermediate are16.7 and –90.5 kJ·mol^{-1}, respectively; and panel B) Detail of the active site in the transition state optimized at M06-2X/MM level. Key distances in the located transition state are indicated as dashed white lines and reported in Å.

Transition state similarity between the two mechanism-based covalent inhibitors 3 and 6. Greater scatter is apparent in LFER plots that correlate the enzymatic parameters for reaction of the two C2-fluoro containing mechanism-based inhibitors **3** and **6** (Figures 4B and 8). We hypothesize that incorporation of a C2-fluorine substituent lowers the active site H-bonding network cooperativity, a requisite for proficient catalysis of unactivated substrates. During the

stabilization of positive charge at the TSs for glycosidic or allylic C–O bond cleavage, each amino acid residue is more independent, rather than participating as a unified ensemble, which involves the extensive H-bonding shown in Figure S14 (Supporting Information).[32]

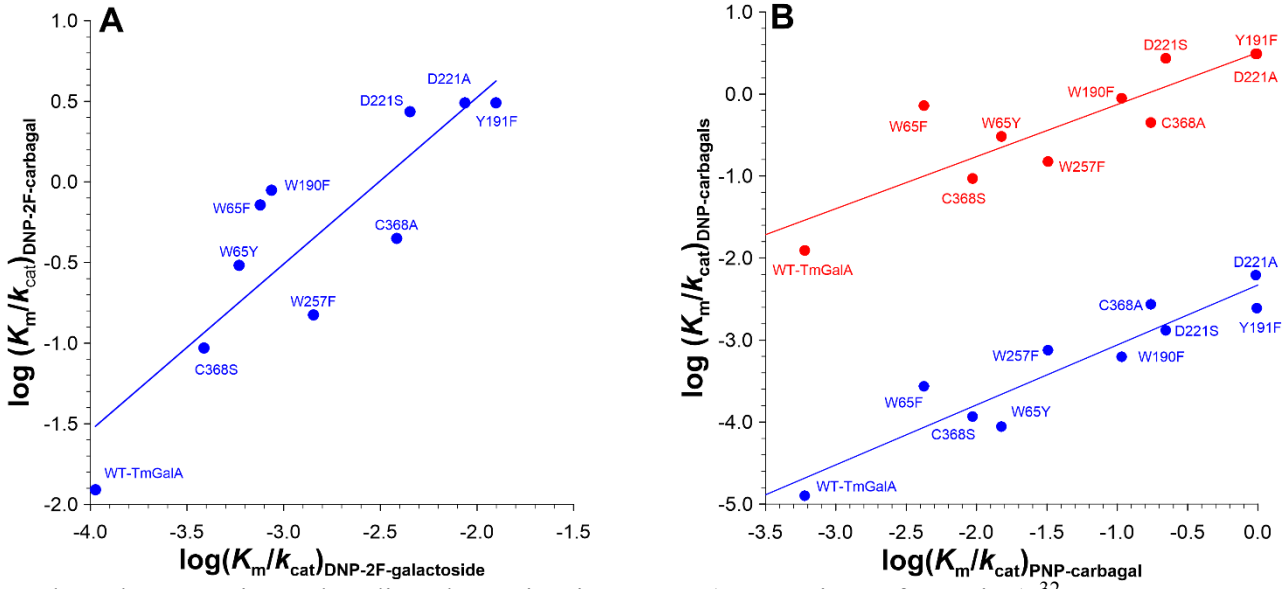


Figure 8. Selected Bartlett-type LFER correlations between the second-order rate constants for the catalyzed hydrolyses of galactoside **3** and covalent inhibitors **4**–**6**. Panel A) LFER for the *Tm*GalA-catalyzed reactions of the fluorinated galactoside **3** versus the fluorinated covalent inhibitor **6**; and panel B) LFER plots for the *Tm*GalA-catalyzed hydrolyses of covalent inhibitors, correlation between the negative logarithms of the second-order rate constants [log(K_m/k_{cat})] for **4** versus **5** (blue circles) and **4** versus **6** (red circles).

Despite the scatter, the catalyzed reactions of 2-deoxy-2-fluorogalactoside **3** and allylic carbasugar **6**, both of which possess 2,4-dinitrophenyl aglycones, display a unit correlation (1.03 ± 0.20; r^2 = 0.7648). Specifically, formation of covalent intermediates during turnover of 2-fluoro galactoside **3** and covalent inhibitor **6** proceeds via analogous "exploded" S_N2 TSs (Figure 5B and

reported for **6**)[32] where the corresponding enzymatic TSs have the same degree of electrostatic TS stabilization for these fluorine-containing mechanism-based inhibitors.

Next, we analyzed two Bartlett LFER correlations, one that compares the 4-nitrophenyl containing reference carbasugar **4** with carbasugar **5** (which possesses a better leaving group) and the other between reference compound **4** and carbasugar **6** (which is deactivated by the C2 fluoro-substituent). Figure 8B shows the correlation data between the log(K_m/k_{cat}) values for carbasugars **4** and **5** (blue), and for **4** and **6** (red). The derived LFER slopes are 0.73 ± 0.10 (r^2 = 0.8695) for **4** versus **5**, and 0.63 ± 0.13 (r^2 = 0.7430) for **4** versus **6**. We conclude that covalent inhibition of *Tm*GalA by **5** has less allylic positive charge development at an earlier TS because **5** has a better leaving group (2,4-dinitrophenoxide) than **4** (4-nitrophenoxide). For the variant *Tm*GalA-catalyzed C–O bond cleavage reactions, the LFER correlation between carbasugars **4** and **6** is comparable to that for galactosides **2** and **3** (see Figure 4B). That is, the less than unit LFER correlation (slope = 0.63 ± 0.13) for the introduction of a positive charge destabilizing C2-fluorine atom results in the assistance of C–O bond cleavage by the active site nucleophile (S_N2) while the 2-hydroxy covalent inhibitor reacts via an S_N1-like mechanism (Figure 7A).

LFERs for melibioside substrate and mechanism-based inactivator *galacto*-cyclophellitol. We turned our attention to a natural product, the well-known oxirane containing inactivator cyclophellitol.[83] We first studied the effect of pH on the rate of *Tm*GalA inactivation by α-D-*galacto*-cyclophellitol (**7**) and found that the rate constant for covalent labeling (k_{inact}/K_i) decreases as the pH increases (Figure 9A). This observation is consistent with the mechanism of *Tm*GalA-catalyzed inactivation to require a protonated form of the enzyme. We therefore measured the rate constants for inactivation of our panel of *Tm*GalA variants at the same pH used for hydrolysis of the melibioside natural substrate analogue **1**. Surprisingly, one variant (W65F) inactivated nearly

two orders of magnitude faster than any of the other enzymes used in this study. The LFER correlation between melibioside **1** and cyclophellitol **7** has a slope = 0.59 ± 0.06 (R^2 = 0.9406; Figure 9B) when the outlier W65F is excluded. This result is consistent with an early TS for nucleophilic ring opening of the oxirane, due to the release of substantial angle strain. Such an early TS would have little charge development on either the carbasugar ring or the oxygen of the oxirane, which is consistent with the shallow slope (= 0.59) of the LFER plot (Figure 9B).

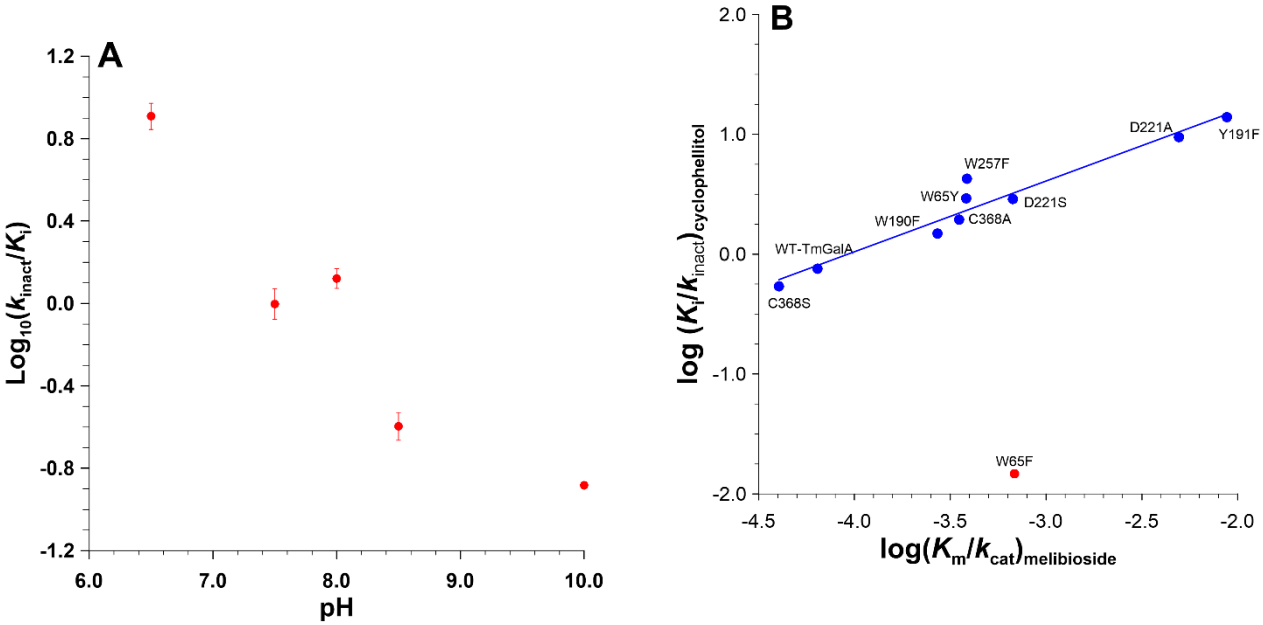


Figure 9. Inactivation of *Tm*GalA and its variant enzymes by **7**. A) $Log_{10}(k_{inact})$ versus pH profile for the inactivation of wild-type *Tm*GalA by α-*galacto*-cyclophellitol (**7**); B) Bartlett-type LFER plot of the logarithms of the second-order rate constants for the *Tm*GalA variant-catalyzed inactivation by inhibitor **7** (K_i/k_{inact}) and hydrolysis of the melibiose substrate **1** (K_m/k_{cat}), blue circular points included in the linear fit (blue line), while the red circular point (W65F) was excluded from the fit.

Conclusions

We have shown that the three classes of mechanism-based inhibitors studied herein react via different degrees of TS charge stabilization. Specifically, i) 2-deoxy-2-fluoroglycosides react via S_N2-like TSs while the glycoside natural substrates react via a S_N1 TS, ii) mechanism-based allylic carbasugars react via analogous TSs to those of the glycosides, and iii) the epoxy carbasugar inactivator **7** reacts via a ring-opening TS that has a reduced requirement for charge stabilization than the TS for reaction of a natural substrate disaccharide. We propose that the design of efficient mechanism-based covalent inhibitors ideally should mimic the flexibility of GHs with respect to stabilizing a variety of charge distributions at the pseudo-glycosylation transition state.

AUTHOR INFORMATION

Corresponding Author

* E-mail: bennet@sfu.ca, Tel.: (778)782–8814, Fax: (778)782–3765.

Present Addresses

† Key Laboratory of Marine Drugs, Chinese Ministry of Education, School of Medicine and Pharmacy, Ocean University of China, 5 Yushan Road, Qingdao 266003, China, and Laboratory for Marine Drugs and Bioproducts, Qingdao National Laboratory for Marine Science and Technology (QNLM), Qingdao 266237, China.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Funding Sources

This work was supported by the Natural Sciences and Engineering Research Council of Canada (AJB Discovery Grant: 2017–04910) and by the Spanish Ministerio de Ciencia, Innovación y Universidades (PGC2021-23332OB-C21 and PID2019-107098RJ-I00), the Generalitat Valenciana (PROMETEO with ref. CIPROM/2021/079 and SEJI/2020/007), Universitat Jaume I (UJI-A2019-04 and UJI-B2020-03). K.Ś. thanks Ministerio de Ciencia, Innovación y Universidades for a Ramón y Cajal contract (REF: RYC2020-030596-I). The simulations were enabled by computational resources provided by Compute Canada (www.computecanada.ca).

Notes

The authors declare no competing financial interest.

Supporting Information

Acknowledgements

The authors thank Dr. T. Kitos for helpful editorial suggestions.

References.

1. Fischer, E., Influence of configuration on the action of enzymes. *Ber. Dt. Chern. Ges.* **1894**, *27*, 2985−2993.
2. Haldane, J. B. S., *Enzymes*. Longmans, Green: London, New York, 1930; p 235.
3. Pauling, L., Molecular structure and biological specificity. *Chem. Eng. News* **1948**, *24*, 1375−1377.
4. Koshland Jr, D. E., Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1958**, *44*, 98–104.
5. Benner, S. A., Enzyme kinetics and molecular evolution. *Chem. Rev.* **1989**, *89*, 789−806.
6. Albery, W. J.; Knowles, J. R., Evolution of enzyme function and development of catalytic efficiency. *Biochemistry* **1976**, *15*, 5631−5640.

7. Silva, R. G.; Murkin, A. S.; Schramm, V. L., Femtosecond dynamics coupled to chemical barrier crossing in a Born-Oppenheimer enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 18661−18665.
8. Toney, M. D.; Castro, J. N.; Addington, T. A., Heavy-enzyme kinetic isotope effects on proton transfer in alanine racemase. *J. Am. Chem. Soc.* **2013**, *135*, 2509−2511.
9. Wei, G. H.; Xi, W. H.; Nussinov, R.; Ma, B. Y., Protein ensembles: how does Nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell. *Chem. Rev.* **2016**, *116*, 6516-6551.
10. Calvo-Tusell, C.; Maria-Solano, M. A.; Osuna, S.; Feixas, F., Time Evolution of the Millisecond Allosteric Activation of Imidazole Glycerol Phosphate Synthase. *J. Am. Chem. Soc.* **2022**, *144*, 7146-7159.
11. Guo, J. J.; Zhou, H. X., Protein Allostery and Conformational Dynamics. *Chemical Reviews* **2016**, *116*, 6503-6515.
12. Krzemińska, A.; Moliner, V.; Świderek, K., Dynamic and Electrostatic Effects on the Reaction Catalyzed by HIV-1 Protease. *J. Am. Chem. Soc.* **2016**, *138*, 16283–16298.
13. Luk, L. Y. P.; Ruiz-Pernia, J. J.; Dawson, W. M.; Roca, M.; Loveridge, E. J.; Glowacki, D. R.; Harvey, J. N.; Mulholland, A. J.; Tuñón, I.; Moliner, V.; Allemann, R. K., Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 16344–16349.
14. Wolfenden, R., Transition state analogues for enzyme catalysis. *Nature* **1969**, *223*, 704−705.
15. Wolfenden, R., Analog approaches to structure of transition state in enzyme reactions. *Acc. Chem. Res.* **1972**, *5*, 10−18.
16. Mader, M. M.; Bartlett, P. A., Binding energy and catalysis: The implications for transition-state analogs and catalytic antibodies. *Chem. Rev.* **1997**, *97*, 1281−1301.
17. Lienhard, G. E., Enzymatic catalysis and transition state theory. *Science* **1973**, *180*, 149−154.
18. Wolfenden, R., Thermodynamic and extrathermodynamic requirements of enzyme catalysis. *Biophys. Chem.* **2003**, *105*, 559−572.
19. Schramm, V. L., Enzymatic transition states and transition state analogues. *Curr. Opin. Struct. Biol.* **2005**, *15*, 604−613.
20. Colombo, C.; Bennet, A. J., Probing transition state analogy in glycoside hydrolase catalysis. *Adv. Phys. Org. Chem.* **2017**, *51*, 99−127.
21. Firn, R. D.; Jones, C. G., Natural products - a simple model to explain chemical diversity. *Natl. Prod. Rep.* **2003**, *20*, 382−391.
22. Lahav, D.; Liu, B.; van den Berg, R.; van den Nieuwendijk, A.; Wennekes, T.; Ghisaidoobe, A. T.; Breen, I.; Ferraz, M. J.; Kuo, C. L.; Wu, L.; Geurink, P. P.; Ovaa, H.; van der Marel, G. A.; van der Stelt, M.; Boot, R. G.; Davies, G. J.; Aerts, J.; Overkleeft, H. S., A fluorescence polarization activity-based protein profiling assay in the discovery of potent, selective inhibitors for Human nonlysosomal glucosylceramidase. *J. Am. Chem. Soc.* **2017**, *139*, 14192−14197.
23. Wu, L.; Armstrong, Z.; Schroder, S. P.; de Boer, C.; Artola, M.; Aerts, J.; Overkleeft, H. S.; Davies, G. J., An overview of activity-based probes for glycosidases. *Curr. Opin. Chem. Biol.* **2019**, *53*, 25-36.
24. 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.; Codee, 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*, 4648-4662.

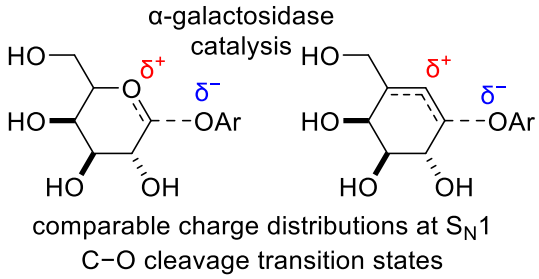
25. Withers, S. G.; Street, I. P.; Bird, P.; Dolphin, D. H., 2-Deoxy-2-fluoroglucosides: A novel class of mechanism-based glucosidase inhibitors. *J. Am. Chem. Soc.* **1987**, *109*, 7530−7531.
26. Kameda, Y.; Horii, S., The unsaturated cyclitol part of the new antibiotics, the validamycins. *J. Chem. Soc., Chem. Commun.* **1972**, 746−747.
27. Kameda, Y.; Asano, N.; Yoshikawa, M.; Matsui, K., Valienamine as an α-glucosidase inhibitor. *J. Antibiot.* **1980**, *33*, 1575−1576.
28. Mosi, R.; Sham, H.; Uitdehaag, J. C. M.; Ruiterkamp, R.; Dijkstra, B. W.; Withers, S. G., Reassessment of acarbose as a transition state analogue inhibitor of cyclodextrin glycosyltransferase. *Biochemistry* **1998**, *37*, 17192−17198.
29. Berland, C. R.; Sigurskjold, B. W.; Stoffer, B.; Frandsen, T. P.; Svensson, B., Thermodynamics of inhibitor binding to mutant forms of glucoamylase from *Aspergillus niger* determined by isothermal titration calorimetry. *Biochemistry* **1995**, *34*, 10153−10161.
30. Atsumi, S.; Iinuma, H.; Nosaka, C.; Umezawa, K., Biological activities of cyclophellitol. *J. Antibiot.* **1990**, *43*, 1579−1585.
31. Comfort, D. A.; Bobrov, K. S.; Ivanen, D. R.; Shabalin, K. A.; Harris, J. M.; Kulminskaya, A. A.; Brumer, H.; Kelly, R. M., Biochemical analysis of *Thermotoga maritima* GH36 α-galactosidase (*Tm*GalA) confirms the mechanistic commonality of Clan GH-D glycoside hydrolases. *Biochemistry* **2007**, *46*, 3319−3330.
32. Ren, W.; Farren-Dai, M.; Sannikova, N.; Świderek, K.; Wang, Y.; Akintola, O.; Britton, R.; Moliner, V.; Bennet, A. J., Glycoside hydrolase stabilization of transition state charge: New directions for inhibitor design. *Chem. Sci.* **2020**, *11*, 10488-10495.
33. Adamson, C.; Pengelly, R. J.; Shamsi Kazem Abadi, S.; Chakladar, S.; Draper, J.; Britton, R.; Gloster, T. M.; Bennet, A. J., Structural Snapshots for Mechanism-Based Inactivation of a Glycoside Hydrolase by Cyclopropyl Carbasugars. *Angewandte Chemie* **2016**, *128*, 15202-15206.
34. Ren, W.; Pengelly, R.; Farren-Dai, M.; Shamsi Kazem Abadi, S.; Oehler, V.; Akintola, O.; Draper, J.; Meanwell, M.; Chakladar, S.; Świderek, K.; Moliner, V.; Britton, R.; Gloster, T. M.; Bennet, A. J., Revealing the mechanism for covalent inhibition of glycoside hydrolases by carbasugars at an atomic level. *Nature Communications* **2018**, *9*, 3243-3243.
35. Accelrys Software, I. *Discovery Studio Modeling Environment, Release 4.5*, San Diego, 2015.
36. Wang, J. M.; Wang, W.; Kollman, P. A.; Case, D. A., Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph.* **2006**, *25*, 247-260.
37. Case, D. A.; Darden, T. A.; Cheatham, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Götz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Hsieh, M. J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A. *AMBER 12*, San Francisco, 2012.
38. Olsson, M. H. M.; Sondergaard, C. R.; Rostkowski, M.; Jensen, J. H., PROPKA3: Consistent treatment of internal and surface residues in empirical pK_a predictions. *J. Chem. Theory. Comput.* **2011**, *7*, 525-537.
39. Dolinsky, T. J.; Nielsen, J. E.; McCammon, J. A.; Baker, N. A., PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **2004**, *32*, W665-W667.
40. 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*, 926-935.

41. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802.
42. Verlet, L., Computer experiments on classical fluids. 1. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **1967**, *159*, 98-103.
43. Field, M. J.; Bash, P. A.; Karplus, M., A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *J. Comput. Chem.* **1990**, *11*, 700-733.
44. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., The development and use of quantum-mechanical molecular-models. 76. AM1 - a new general-purpose quantum-mechanical molecular-model. *J. Am. Chem. Soc.* **1985**, *107*, 3902-3909.
45. Zhao, Y.; Truhlar, D. G., The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215-241.
46. Frisch, M. J.; Trucks, G. W.; Schlegel, G. E.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.; P; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Lyengar, S. S.; Tomasi, J.; Cossi, M.; N, R.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.02*, Wallingford, CT, 2009.
47. Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A., The Dynamo library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. *J. Comput. Chem.* **2000**, *21*, 1088-1100.
48. Krzemińska, A.; Paneth, P.; Moliner, V.; Świderek, K., Binding isotope effects as a tool for distinguishing hydrophobic and hydrophilic binding sites of HIV-1 RT. *J. Phys. Chem. B* **2015**, *119*, 917-927.
49. Roux, B., The calculation of the potential of mean force using computer-simulations. *Comput. Phys. Commun.* **1995**, *91*, 275-282.
50. Torrie, G. M.; Valleau, J. P., Non-physical sampling distributions in Monte-Carlo free-energy estimation - umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187-199.
51. Kumar, S.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; Rosenberg, J. M., The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method. *J. Comput. Chem.* **1992**, *13*, 1011-1021.
52. Chuang, Y. Y.; Corchado, J. C.; Truhlar, D. G., Mapped interpolation scheme for single-point energy corrections in reaction rate calculations and a critical evaluation of dual-level reaction path dynamics methods. *J. Chem. Phys. A* **1999**, *103*, 1140-1149.
53. Corchado, J. C.; Coitino, E. L.; Chuang, Y. Y.; Fast, P. L.; Truhlar, D. G., Interpolated variational transition-state theory by mapping. *J. Chem. Phys. A* **1998**, *102*, 2424-2438.

54. Nguyen, K. A.; Rossi, I.; Truhlar, D. G., A dual-level shepard interpolation method for generating potential-energy surfaces for dynamics calculations. *J. Chem. Phys.* **1995**, *103*, 5522-5530.
55. Renka, R. J., Interpolatory tension splines with automatic selection of tension factors. *SIAM J. Stat. Comput.* **1987**, *8*, 393-415.
56. Renka, R. J., Algorithm-716 tspack - tension spline curve-fitting package. *ACM Trans. Math. Software* **1993**, *19*, 81-94.
57. Ruiz-Pernia, J. J.; Silla, E.; Tunon, I.; Marti, S., Hybrid quantum mechanics/molecular mechanics simulations with two-dimensional interpolated corrections: Application to enzymatic processes. *J. Phys. Chem. B* **2006**, *110*, 17663-17670.
58. Ruiz-Pernia, J. J.; Silla, E.; Tuñón, I.; Marti, S.; Moliner, V., Hybrid QM/MM potentials of mean force with interpolated corrections. *J. Phys. Chem. B* **2004**, *108*, 8427-8433.
59. Baker, J., An algorithm for the location of transition states. *J. Comput. Chem.* **1986**, *7*, 385-395.
60. Fukui, K., The path of chemical-reactions – The IRC approach. *Acc. Chem. Res.* **1981**, *14*, 363-368.
61. Gao, J. L.; Ma, S. H.; Major, D. T.; Nam, K.; Pu, J. Z.; Truhlar, D. G., Mechanisms and free energies of enzymatic reactions. *Chem. Rev.* **2006**, *106*, 3188−3209.
62. Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H. B.; Olsson, M. H. M., Electrostatic basis for enzyme catalysis. *Chem. Rev.* **2006**, *106*, 3210−3235.
63. Garcia-Viloca, M.; Gao, J.; Karplus, M.; Truhlar, D. G., How enzymes work: Analysis by modern rate theory and computer simulations. *Science* **2004**, *303*, 186−195.
64. Marti, S.; Roca, M.; Andres, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertran, J., Theoretical insights in enzyme catalysis. *Chem. Soc. Rev.* **2004**, *33*, 98−107.
65. Hosie, L.; Sinnott, M. L., Effects of deuterium substitution α- and β- to the reaction centre, ^{18}O substitution in the leaving group, and aglycone acidity on hydrolyses of aryl glucosides and glucosyl pyridinium ions by yeast α-glucosidase. *Biochem. J.* **1985**, *226*, 437−446.
66. Cekic, N.; Heinonen, J. E.; Stubbs, K. A.; Roth, C.; He, Y.; Bennet, A. J.; McEachern, E. J.; Davies, G. J.; Vocadlo, D. J., Analysis of transition state mimicry by tight binding aminothiazoline inhibitors provides insight into catalysis by human O-GlcNAcase. *Chem. Sci.* **2016**, *7*, 3742−3750.
67. Whitworth, G. E.; Macauley, M. S.; Stubbs, K. A.; Dennis, R. J.; Taylor, E. J.; Davies, G. J.; Greig, I. R.; Vocadlo, D. J., Analysis of PUGNAc and NAG-thiazoline as transition state analogues for human O-GlcNAcase: Mechanistic and structural insights into inhibitor selectivity and transition state poise. *J. Am. Chem. Soc.* **2007**, *129*, 635−644.
68. Shidmoossavee, F. S.; Watson, J. N.; Bennet, A. J., Chemical insight into the emergence of influenza virus strains that are resistant to Relenza. *J. Am. Chem. Soc.* **2013**, *135*, 13254−13257.
69. Bartlett, P. A.; Marlowe, C. K., Phosphonamidates as transition-state analog inhibitors of thermolysin. *Biochemistry* **1983**, *22*, 4618−4624.
70. Tailford, L. E.; Offen, W. A.; Smith, N. L.; Dumon, C.; Morland, C.; Gratien, J.; Heck, M. P.; Stick, R. V.; Bleriot, Y.; Vasella, A.; Gilbert, H. J.; Davies, G. J., Structural and biochemical evidence for a boat-like transition state in β-mannosidases. *Nat. Chem. Biol.* **2008**, *4*, 306−312.
71. Balachandran, N.; To, F.; Berti, P. J., Linear Free Energy Relationship analysis of transition state mimicry by 3-deoxy-D-*arabino*-heptulosonate-7-phosphate (DAHP) oxime, a DAHP synthase inhibitor and phosphate mimic. *Biochemistry* **2017**, *56*, 592−601.
72. Ren, W.; Pengelly, R.; Farren-Dai, M.; Shamsi Kazem Abadi, S.; Oehler, V.; Akintola, O.; Draper, J.; Meanwell, M.; Chakladar, S.; Świderek, K.; Moliner, V.; Britton, R.; Gloster, T. M.;

Bennet, A. J., Revealing the mechanism for covalent inhibition of glycoside hydrolases by carbasugars at an atomic level. *Nat. Commun.* **2018**, *9*, 3243.
73. Hudson, K. L.; Bartlett, G. J.; Diehl, R. C.; Agirre, J.; Gallagher, T.; Kiessling, L. L.; Woolfson, D. N., Carbohydrate-aromatic interactions in proteins. *J. Am. Chem. Soc.* **2015**, *137*, 15152−15160.
74. Spiwok, V., CH/π interactions in carbohydrate recognition. *Molecules* **2017**, *22*, 1038.
75. McCarter, J. D.; Adam, M. J.; Braun, C.; Namchuk, M.; Tull, D.; Withers, S. G., Syntheses of 2-deoxy-2-fluoro mono- and oligo-saccharide glycosides from glycals and evaluation as glycosidase inhibitors. *Carbohydr. Res.* **1993**, *249*, 77−90.
76. Shaikh, F. A.; van Bueren, A. L.; Davies, G. J.; Withers, S. G., Identifying the catalytic acid/base in GH29 α-L-fucosidase subfamilies. *Biochemistry* **2013**, *52*, 5857−5864.
77. Hammond, G. S., A correlation of reaction rates. *J. Am. Chem. Soc.* **1955**, *77*, 334−338.
78. Davies, G. J.; Planas, A.; Rovira, C., Conformational analyses of the reaction coordinate of glycosidases. *Acc. Chem. Res.* **2012**, *45*, 308−316.
79. Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J., Dissecting conformational contributions to glycosidase catalysis and inhibition. *Curr. Opin. Struct. Biol.* **2014**, *28*, 1−13.
80. Vocadlo, D. J.; Davies, G. J., Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539-555.
81. Adabala, P. J. P.; Shamsi Kazem Abadi, S.; Akintola, O.; Bhosale, S.; Bennet, A. J., Conformationally controlled reactivity of carbasugars uncovers the choreography of glycoside hydrolase catalysis. *J. Org. Chem.* **2020**, *85*, 3336-3348.
82. Farren-Dai, M.; Sannikova, N.; Świderek, K.; Moliner, V.; Bennet, A. J., Fundamental insight into glycoside hydrolase-catalyzed hydrolysis of the universal Koshland substrates–glycopyranosyl fluorides. *ACS Catal.* **2021**, *11*, 10383–10393.
83. Atsumi, S.; Umezawa, K.; Iinuma, H.; Naganawa, H.; Nakamura, H.; Iitaka, Y.; Takeuchi, T., Production, isolation and structure determination of a novel β-glucosidase inhibitor, cyclophellitol, from *Phellinus sp. J. Antibiot.* **1990**, *43*, 49−53.

TOC Graphic



comparable charge distributions at S_N1
C−O cleavage transition states