UNIVERSITY *of York*

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Computational design-of-experiment unveils the conformational reaction coordinate of GH125 α-mannosidases

Santiago Alonso-Gil,[1] Alexandra Males,[2] Pearl Z. Fernandes,[3] Spencer J Williams,[3] Gideon J Davies[2,*] & Carme Rovira[1,4,*]

AUTHOR ADDRESS 1. 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. 2. York Structural Biology Laboratory, Department of Chemistry, The University of York, YO10 5DD, United Kingdom. 3. School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Victoria 3010, Australia. 4. Institució Catala de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08020 Barcelona (Spain)

Supporting Information Placeholder

ABSTRACT: The conformational analysis of enzyme-catalysed mannoside hydrolysis has revealed two predominant conformational itineraries through $B_{2,5}$ or $^3H_4$ transition state conformations. A prominent unassigned catalytic itinerary is that of the *exo*-1,6-α-mannosidases belonging to CAZy family 125. A published complex of the *Clostridium perfringens* GH125 enzyme with a non-hydrolysable 1,6-α-thiomannoside substrate mimic bound across the active site revealed an undistorted $^4C_1$ conformation and provided no insight into the catalytic pathway of this enzyme. Here we show, through a purely computational approach (QM/MM metadynamics) that sulfur-for-oxygen substitution in the glycosidic linkage fundamentally alters the energetically accessible conformational space of a thiomannoside when bound within the GH125 active site. Thus, while modelling of the conformational free energy landscape (FEL) of a thioglycoside strongly favors a mechanistically uninformative $^4C_1$ conformation within the GH125 enzyme active site, the FEL of the corresponding *O*-glycoside substrate reveals a preference for a Michaelis complex in an $^OS_2$ conformation (consistent with catalysis through a $B_{2,5}$ transition state). This prediction was tested experimentally by determination of the 3-D X-ray structure of the pseudo-Michaelis complex of an inactive (D220N) variant of the *C. perfringens* GH125 enzyme in complex with 1,6-α-mannobiose. This complex revealed unambiguous distortion of the –1 subsite mannoside to an $^OS_2$ conformation, matching that predicted by theory, and supporting an $^OS_2 \rightarrow B_{2,5} \rightarrow {}^1S_5$ conformational itinerary for GH125 α-mannosidases. This work highlights the power of the QM/MM approach and identified potential shortcomings in the use of non-hydrolysable substrate analogues for conformational analysis of enzyme-bound species.

The conformational itineraries employed by glycoside hydrolases to perform nucleophilic substitution reactions at the anomeric center of glycosides have been the topic of sustained interest since the mid-1990s (reviewed in Refs[1-2]). Physical organic studies have provided compelling evidence that glycosidase-catalyzed glycoside cleavage occurs through oxocarbenium-ion-like transition states with significant partial double-bond character between the anomeric carbon and the ring oxygen.[3-4] Sinnott postulated that glycosidases must therefore react through transition states in

one of 4 major conformations: $^4H_3$ and $^3H_4$ half chairs (or their related envelopes), or $B_{2,5}$ and $^{2,5}B$ boats. The topological relationships of such conformations are conveniently visualized through plotting the conformations as a Mercator projection (Figure 1).
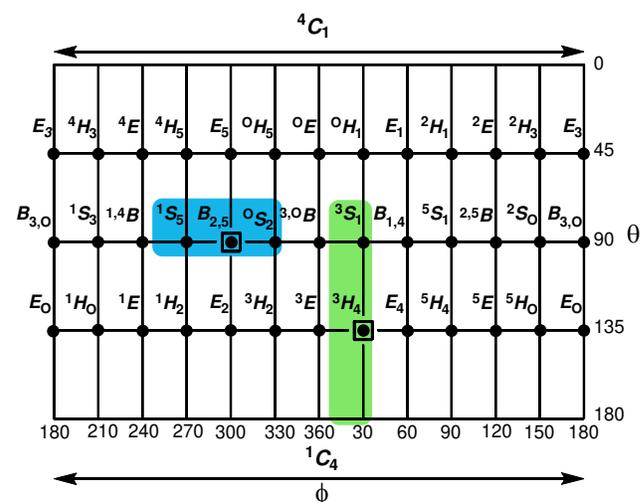
Figure 1. Mercator plot of major canonical conformations of a pyranose ring. The transition state conformations (boxed) and associated ground-state conformations of mannosidase conformational itineraries through transition states with $B_{2,5}$ (blue) and $^3H_4$ (green) conformations.

According to the principle of least nuclear motion,[5] the conformations of the ground states of the enzymatic Michaelis complex, products, and (if relevant) associated intermediates must flank the transition states. While early analyses focused on the $^4H_3$ transition state conformation, studies over the last 20 years have identified that all four major transition state conformations are co-opted by various enzymes working across the breadth of stereochemically-diverse carbohydrate substrates.[1-2] As transition state mimicry provides a practical blueprint for the development of tight binding inhibitors, analysis of these reaction coordinates is proving invaluable in the design and application of transition-state mimicking species as mechanistic probes and therapeutic agents.[6]

The canvas upon which nature's treasure-chest of glycosidases is depicted is the carbohydrate-active enzymes (CAZy) classifica-

tion.[7] Enzymes are classified into families according to amino-acid sequence (and hence 3-D structural) similarity. Of particular interest are the diverse α- and β-mannanases and mannosidases, which catalyse the sterically-challenged reaction at the crowded anomeric carbon of mannose, for which mechanistic insights can inform and enlighten key challenges involved in the chemical synthesis of mannosides.[8] α- and β-mannanases are involved in glycan processing within important industrial and biological processes. In the latter case assorted α-mannosidases are involved in N-glycan maturation and processing,[9-10] fungal cell-wall biosynthesis[11] and catabolism,[12] and other cellular reactions of high interest for therapeutic intervention.

According to the CAZy classification, α- and β-mannosidases (both *exo-* and *endo-*acting) populate a large number of GH families: (α) 38, 47, 76, 92, 99 and 125, and (β) 2, 5, 26, 113, 130 and 134, respectively. Systematic analysis of the conformational itineraries of these enzyme families, primarily through crystallography of stable species flanking or mimicking the reaction transition-state(s), has revealed two predominant strategies employed by these catalysts to overcome the challenges of mannoside chemistry (note: the GH99 α-mannosidases are believed to react through an epoxide intermediate[13] and are not discussed further). One group of α- and β-mannosidases belonging to GH families 2,[14] 5,[15] 26,[16] 38,[17] 76,[18] 92,[19] 113,[20] and 130[21] perform catalysis through a pathway around the $^OS_2$—$B_{2,5}$—$^1S_5$ region of the conformational space (Figure S2). The other group of GHs include the family GH47 α-mannosidases[22-24] and the GH134 β-mannanases,[25] which react in a 'ring-flipped' (southern hemisphere) $^3S_1$—$^3H_4$—$^1C_4$ conformational arena (Figure S2).

In seminal work Gregg and colleagues reported the creation of GH family 125 based on the discovery of 1,6-α-mannosidase activity for enzymes from *Clostridium perfringens* (*Cp*GH125) and *Streptococcus pneumonia*.[26] This family was shown to operate through an inverting mechanism, and insight into the active site residues was provided through X-ray structures of these enzymes in complex with the non-hydrolyzable substrate analogue 1,6-α-thiomannobiose (PDB entry 3QT9), and deoxymannojirimycin (PDB 3QRY). Surprisingly, despite the 1,6-α-thiomannobiose substrate mimic spanning the active site, the mechanistically informative –1 subsite mannose residue was observed in an undistorted, ground-state $^4C_1$ conformation, providing no insight into the conformational itinerary of this family of α-mannosidases. Intrigued by this surprising but uninformative result, we were motivated to investigate further. Although the distortion-free binding of the thiomannoside is surprising, it is not unprecedented – a similar situation was noted in the case of 1,2-α-thiomannobiose bound to a GH92 1,2-α-mannosidase; in that case a complex with the transition state mimic mannoimidazole provided evidence in support of an $^OS_2 \rightarrow B_{2,5} \rightarrow ^1S_5$ conformational itinerary.[19] However, in the GH125 case the same approach cannot be applied as the general acid residue is not appropriately situated to allow lateral protonation of the basic mannoimidazole nitrogen, whereas in family GH92 enzymes the orientation of the general acid residue is '*anti*'[27] to the C1-O5 bond, which enables lateral protonation and binding of this inhibitor. The inability to assign a conformational itinerary to GH family 125 prevents rational application and design of conformationally-locked or biased inhibitors selective for this family of biomedically important enzymes. To understand the conformational preferences of thioglycosides within the active site of *Cp*GH125, we first adopted a computational approach (*ab-initio* QM/MM metadynamics)[28-29] to map the conformational free energy landscape of the –1 mannoside ring as a function of the Cremer-Pople ring puckering coordinates;[30] an approach that has been applied to other GH families.[31] We first calculated the free energy surface for isolated 1-thio-α-mannopyranose (see computational details in the Supporting Information, SI, and Figure S1). As previously found for α-mannopyranose,[22] the sugar has a preference for $^4C_1$ conformation, but with other regions of the conformational energy surface energetically accessible – most notably the region around the $^OS_2$ conformation.
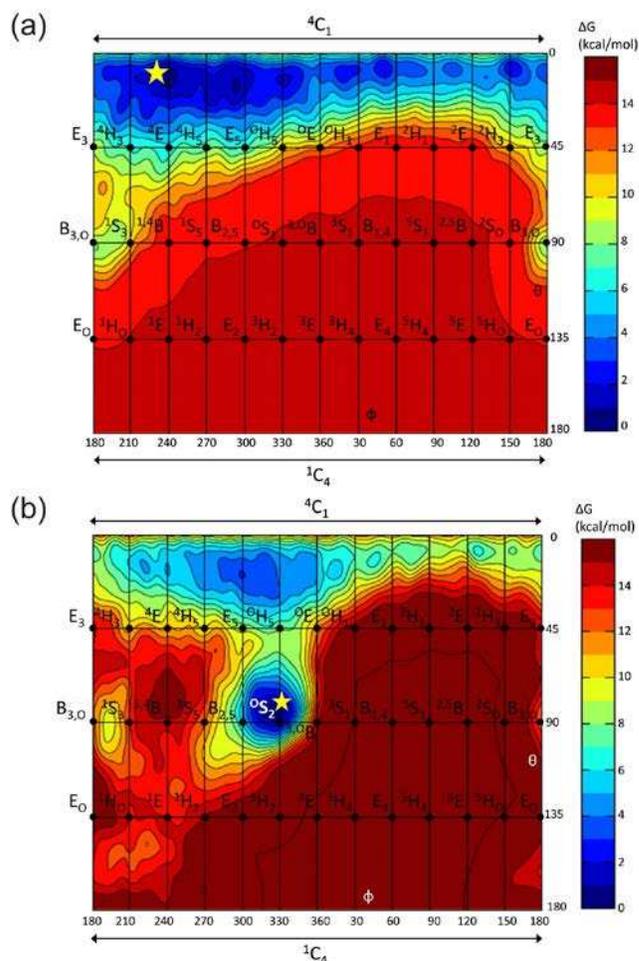
Figure 2. (a) Conformational FEL of the 1-thio-α-mannosyl residue at the –1 subsite of *Cp*GH125 in complex with 1,6-α-thiomannobiose. The star symbol plots the conformation observed experimentally.[26] (b) Conformational FEL of the α-mannosyl residue at the –1 subsite of *Cp*GH125 in complex with 1,6-α-mannobiose. The star symbol plots the new conformation subsequently observed experimentally (see below). Contour lines at 1 kcal mol⁻¹.

When the QM/MM metadynamics approach was applied to the 'on-enzyme' complex of 1-thio-α-mannopyranose and *Cp*GH125, the free energy landscape is transformed such that the accessible conformational surface is dramatically restricted, Figure 2a. Thus, the use of an S-linked substrate analogue results in a strong bias to a $^4C_1$ conformation, matching that observed in the original report of Gregg *et al.*,[26] with other, more mechanistically relevant conformations not energetically accessible. Thus, in this case, while the thiomannoside substrate mimic is informative on the gross details of the catalytic apparatus and the ligand interactions, it is silent in terms of conformational insight.

We next sought to establish whether a solely computational approach could make testable predictions for the catalytic itinerary consistent with that previously observed for α- and β-mannosidases. Starting with the experimentally determined *Cp*GH125 1,6-α-thiomannobiose complex, the glycosidic sulfur was substituted for oxygen *in silico* to generate a catalytically-viable Michaelis complex, which was subjected to minimization to generate a lower energy form, followed by MD equilibration.

The full conformational landscape of the –1 sugar ring of this competent substrate containing an O-glycosidic linkage was then calculated by QM/MM metadynamics, using the same procedure as in the case of the 1,6-α-thiomannobiose complex. Figure 2b shows that within this Michaelis complex (on-enzyme), an O-glycoside strongly favors an $^OS_2$ conformation, consistent with the α-mannosidase performing catalysis through an $^OS_2 \rightarrow B_{2,5} \rightarrow {}^1S_5$ conformational itinerary. QM/MM simulations of the reaction mechanism (Figures 3 and S2) starting from the $^OS_2$ conformation led to a $B_{2,5}$ transition state, in a dissociative reaction pathway generating a β-mannose product bound to *Cp*GH125 with a $^1S_5/B_{2,5}$ conformation. Overall, this computational data, derived from the coordinate of the *Cp*GH125 1,6-α-thiomannobiose complex, matches that proposed for GH families 2, 5, 26, 38, 76, 92 113, and 130 (reviewed in Refs[1-2, 6]).

In order to validate, experimentally, the *in silico* prediction, an inactive variant in which the general acid (D220) of *Cp*GH125 was mutated to a non-acidic asparagine residue was engineered. This catalytically-inactive variant was crystallized and soaked with the native O-glycosides, 1,6-α-mannobiose and -mannotriose to obtain pseudo-Michaelis complexes. Comparison of the structures of the ligand-free *Cp*GH125 wildtype and ligand-bound D220N enzymes revealed no changes in the position of the amino acid side-chain or other residues, providing confidence that the observed ligand conformation was not a result of non-isomorphism. The *Cp*GH125 D220N complexes, solved at resolutions of 2.10 and 1.55 Å (Supplementary Table 1), unambiguously reveal the –1 subsite mannoside distorted to a $^OS_2$ conformation (Figure 4a; for 1,6-α-mannotriose complex see Figure S3), matching that predicted *a priori* by computation.

The *Cp*GH125 complexes highlight the molecular basis for catalysis, with a nucleophilic water poised for in-line nucleophilic attack at the anomeric carbon and with E393 positioned to act as the catalytic Brønsted base in an inverting mechanism, essentially as proposed previously.[26] Interestingly, the nucleophilic water molecule is engaged in a hydrogen-bonding interaction with O3, rather than with O2, as was instead observed in the *Cp*GH125 1,6-α-thiomannobiose complex. This interaction with O3 is reminiscent of that seen for the nucleophilic residue for a GH family 76 retaining 1,6-α-mannanase from *Bacillus circulans*,[18] and is thus a feature of the non-metal dependent, family 76 and 125 α-mannosidases. Overlay of the *Cp*GH125 D220N 1,6-α-mannobiose complex with the previously determined 1,6-α-thiomannobiose complex (Figure 4b) highlights the structural basis for the conformational differences; whilst the +1 (leaving group) subsite mannoses are essentially identical in terms of conformation and interactions, the –1 subsite mannoside moieties adopt different conformations, and match those predicted by computation. One major contributor to these different conformations is the longer C—S bond (1.89 vs. 1.48 Å for C—O); presumably as a result of this key structural difference the –1 thiomannoside in a $^4C_1$ conformation with an axial O2 group makes similar interactions to the pseudo-axial O3 of the mannoside in an $^OS_2$ conformation (Figure S4). Both the theory-based calculations and the subsequent experimental observation support a conformational itinerary for the inverting GH125 α-mannosidases that proceeds through a (near) $B_{2,5}$ transition-state conformation. This transition state is accessed following binding of the substrate in the ES complex in an $^OS_2$ conformation, Figure 3.

GH family 125 joins the growing list of mannose-active enzymes that follows a latitudinal pathway around a $B_{2,5}$ transition state in which a key "feature" is the near-eclipsed 40 degree torsional angle between O3 and O2 that positions a *manno*-configured O2 pseudo-equatorial and stabilized through H-bonding on-enzyme. Strikingly, there is a remarkable connection to the Crich β-

mannosylation methodology wherein judicious choice of a 4,6-*O*-benzylidene protecting group favors a similar pathway.[8]
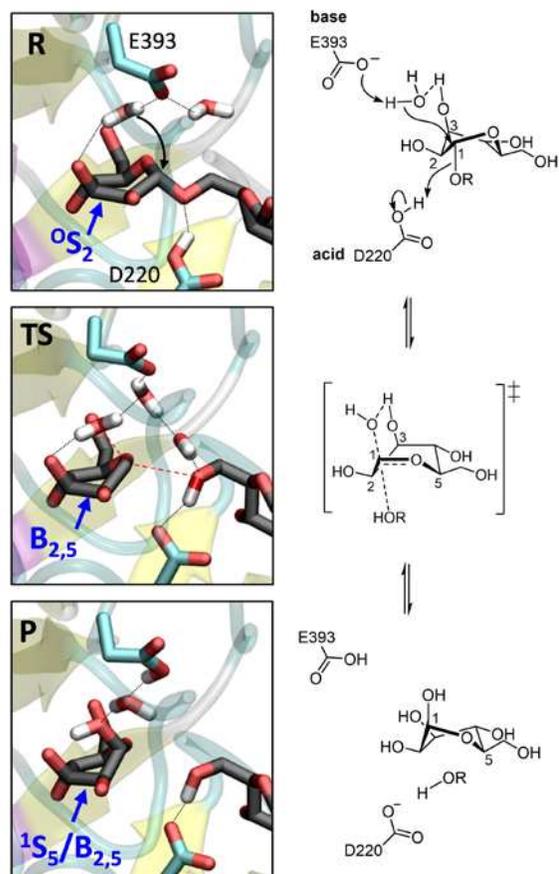
Figure 3. Reaction coordinate for *Cp*GH125 inverting 1,6-α-mannosidase obtained by QM/MM metadynamics with four collective variables. Hydrogen atoms have been omitted for clarity, except those of the carboxylate groups and water molecules.
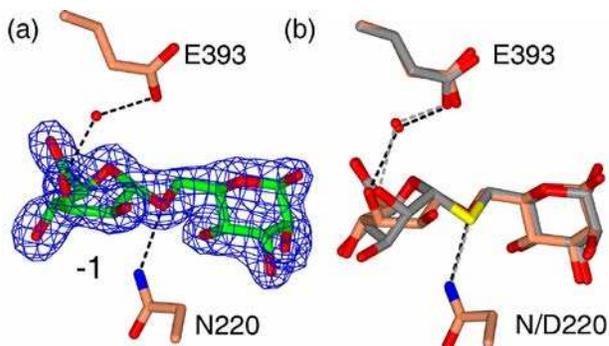
Figure 4. (a) Observed electron density ($2F_{obs}$-F_{calc}, σ_A and maximum likelihood weighted) for the D220N 1,6-α-mannobiose complex of *Cp*GH125, contoured at 0.31 electrons / Å^3. (b) Comparison of the *C. perfringens* GH125 complexes with 1,6-α-mannobiose (this work, brick red) with the 1,6-α-thiomannobiose complex (grey, PDB 3QT9, ref.[26]). D220 is the general acid; E393 is the general base; the proposed nucleophile water is shown.

Thiooligosaccharide substrate mimics have been widely used in X-ray crystallographic studies where they have provided mechanistically relevant insight into the conformations possible on enzyme, most notably in the case of distorted thiocellopentaoside bound to *Fusarium oxysporum* cellulase of GH family 7.[32] Could other thioglycoside complexes be misleading? In the case of an-

3

other inverting α-mannosidase, from GH family 47, a 1,2-α-thiomannobioside was ring flipped and distorted to the $^3S_1$ conformation suggesting that in this case it is mechanistically-relevant.[22-23] Indeed, other complexes, notably with mannoimidazole in a $^3H_4$ conformation, and kifunensine in a $^1C_4$ conformation, as well as subsequent QM/MM analysis of the FEL of α-mannose 'on-enzyme', collectively support the $^3S_1 \rightarrow {}^3H_4 \rightarrow {}^1C_4$ pathway for that enzyme.[22-23] In contrast, as discussed earlier, the 1,2-α-thiomannobioside complex observed on family GH92, as observed here for GH125, was also observed undistorted and again silent to conformational pathways; in that case distortion of enzyme-bound mannoimidazole to a boat conformation allowed assignment of a $^OS_2 \rightarrow B_{2,5} \rightarrow {}^1S_5$ pathway for that enzyme.[19]

This work highlights the predictive power of computational methods to use preliminary enzyme-ligand complexes to explore conformational space and generate testable predictions that can provide mechanistic insight using X-ray structural methods. Here these approaches predicted ES distortion for *Cp*GH125 and informed an experimental approach that enabled direct observation of a distorted pseudo-Michaelis complex. This combined *in silico*-experimental approach could be applied to identify catalytic itineraries for other GH families that are presently unknown or for which unusual conformations have been proposed guiding inhibitor design and leading to the development of mechanistic probes, cellular probes and ultimately therapeutic agents. In this latter context, the ultimate goal is to obtain conformationally selective and thus specific inhibition of just one enzyme family, as has been achieved through the inhibition of ($^3S_1 \rightarrow {}^3H_4 \rightarrow {}^1C_4$ pathway) GH47 α-mannosidases by kifunensine; a $^1C_4$ chair mimic.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website. Coordinates have been deposited with PDB codes 5M7Y and 5M7I.

AUTHOR INFORMATION

Corresponding Authors

* (C.R.) E-mail: c.rovira@ub.edu
* (G.J.D.) E-mail: gideon.davies@york.ac.uk

Author Contributions

CR, SJW and GJD designed experiments. S A-G performed computational work, AM structural work and PF organic synthesis. GJD, SJW and CR wrote the manuscript.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

REFERENCES

1. Davies, G. J.; Planas, A.; Rovira, C., *Acc. Chem. Res.* **2012,** *45*, 308-16.
2. Speciale, G.; Thompson, A. J.; Davies, G. J.; Williams, S. J., *Curr. Opin. Struct. Biol.* **2014,** *28*, 1-13.
3. Koshland, D. E., *Biol. Rev.* **1953,** *28*, 416 - 436.
4. Sinnott, M. L., Glycosyl group transfer. In *Enzyme mechanisms*, Page, M. I.; Williams, A., Eds. Royal Society of Chemistry: London, 1987; pp 259 - 297.
5. Sinnott, M. L., The Principle of Least Nuclear Motion and the Theory of Stereoelectronic Control. In *Advances in Physical Organic Chemistry*, Bethell, D., Ed. Academic Press: 1988; Vol. Volume 24, pp 113-204.
6. Davies, G. J.; Williams, S. J., *Biochem. Soc. Trans.* **2016,** *44*, 79-87.
7. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B., *Nucleic Acids Res.* **2014,** *42*, D490-5.
8. Crich, D., *Acc. Chem. Res.* **2010,** *43*, 1144-1153.
9. Kornfeld, R.; Kornfeld, S., *Annu. Rev. Biochem.* **1985,** *54*, 631-664.
10. Helenius, A.; Aebi, M., *Annu. Rev. Biochem.* **2004,** *73*, 1019-1049.
11. Free, S. J., *Adv. Gen.* **2013,** *81*, 33-82.
12. Cuskin, F.; Lowe, E. C.; Temple, M. J.; Zhu, Y.; Cameron, E. A.; Pudlo, N. A.; Porter, N. T.; Urs, K.; Thompson, A. J.; Cartmell, A.; Rogowski, A.; Hamilton, B. S.; Chen, R.; Tolbert, T. J.; Piens, K.; Bracke, D.; Vervecken, W.; Hakki, Z.; Speciale, G.; Munoz-Munoz, J. L.; Day, A.; Pena, M. J.; McLean, R.; Suits, M. D.; Boraston, A. B.; Atherly, T.; Ziemer, C. J.; Williams, S. J.; Davies, G. J.; Abbott, D. W.; Martens, E. C.; Gilbert, H. J., *Nature* **2015,** *517*, 165-9.
13. Thompson, A. J.; Williams, R. J.; Hakki, Z.; Alonzi, D. S.; Wennekes, T.; Gloster, T. M.; Songsrirote, K.; Thomas-Oates, J. E.; Wrodnigg, T. M.; Spreitz, J.; Stutz, A. E.; Butters, T. D.; Williams, S. J.; Davies, G. J., *Proc. Natl. Acad. Sci. USA* **2012,** *109*, 781-6.
14. 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., *Nat. Chem. Biol.* **2008,** *4*, 306-312.
15. Vincent, F.; Gloster, T. M.; Macdonald, J.; Morland, C.; Stick, R. V.; Dias, F. M. V.; Prates, J. A. M.; Fontes, C. M. G. A.; Gilbert, H. J.; Davies, G. J., *ChemBioChem* **2004,** *5*, 1596-1599.
16. Ducros, V.; Zechel, D. L.; Murshudov, G. N.; Gilbert, H. J.; Szabo, L.; Stoll, D.; Withers, S. G.; Davies, G. J., *Angew. Chem. Int. Ed.* **2002,** *41*, 2824-2827.
17. Numao, S.; Kuntz, D. A.; Withers, S. G.; Rose, D. R., *J. Biol. Chem.* **2003,** *278*, 48074-83.
18. Thompson, A. J.; Speciale, G.; Iglesias-Fernández, J.; Hakki, Z.; Belz, T.; Cartmell, A.; Spears, R. J.; Chandler, E.; Temple, M. J.; Stepper, J.; Gilbert, H. J.; Rovira, C.; Williams, S. J.; Davies, G. J., *Angew. Chem. Int. Ed.* **2015,** *54*, 5378–5382.
19. Zhu, Y.; Suits, M. D.; Thompson, A. J.; Chavan, S.; Dinev, Z.; Dumon, C.; Smith, N.; Moremen, K.; Xiang, Y.; Siriwardena, A.; Williams, S. J.; Gilbert, H. J.; Davies, G. J., *Nat. Chem. Biol.* **2010,** *6*, 125-132.
20. Williams, R. J.; Iglesias-Fernández, J.; Stepper, J.; Jackson, A.; Thompson, A. J.; Lowe, E. C.; White, J. M.; Gilbert, H. J.; Rovira, C.; Davies, G. J.; Williams, S. J., *Angew. Chem. Int. Ed.* **2014,** *53*, 1087-1091.
21. Cuskin, F.; Baslé, A.; Ladevèze, S.; Day, A. M.; Gilbert, H. J.; Davies, G. J.; Potocki-Véronèse, G.; Lowe, E. C., *J. Biol. Chem.* **2015,** *290*, 25023-25033.
22. Thompson, A. J.; Dabin, J.; Iglesias-Fernandez, J.; Ardevol, A.; Dinev, Z.; Williams, S. J.; Bande, O.; Siriwardena, A.; Moreland, C.; Hu, T. C.; Smith, D. K.; Gilbert, H. J.; Rovira, C.; Davies, G. J., *Angew. Chem. Int. Ed.* **2012,** *51*, 10997-11001.
23. Karaveg, K.; Siriwardena, A.; Tempel, W.; Liu, Z. J.; Glushka, J.; Wang, B. C.; Moremen, K. W., *J. Biol. Chem.* **2005,** *280*, 16197-207.
24. Vallée, F.; Karaveg, K.; Herscovics, A.; Moremen, K. W.; Howell, P. L., *J. Biol. Chem.* **2000,** *275*, 41287-41298.
25. Jin, Y.; Petricevic, M.; John, A.; Raich, L.; Jenkins, H.; Souza, L. P. D.; Cuskin, F.; Gilbert, H. J.; Rovira, C.; Goddard-Borger, E. D.; Williams, S. J.; Davies, G. J., *ACS Central Sci* **2016,** *in press*.
26. Gregg, K. J.; Zandberg, W. F.; Hehemann, J. H.; Whitworth, G. E.; Deng, L.; Vocadlo, D. J.; Boraston, A. B., *J. Biol. Chem.* **2011,** *286*, 15586-96.
27. Heightman, T. D.; Vasella, A. T., *Angew. Chem. Int. Ed.* **1999,** *38*, 750-770.
28. Barducci, A.; Bonomi, M.; Parrinello, M., *WIREs Comput. Mol. Sci.* **2011,** *1*, 826-843.
29. Laio, A.; Parrinello, M., *Proc. Natl. Acad. Sci. USA* **2002,** *99*, 12562-12566.
30. Cremer, D.; Pople, J. A., *J. Am. Chem. Soc.* **1975,** *97*, 1354-1358.
31. Ardèvol, A.; Rovira, C., *J. Am. Chem. Soc.* **2015,** *137*, 7528-47.
32. Sulzenbacher, G.; Driguez, H.; Henrissat, B.; Schulein, M.; Davies, G. J., *Biochemistry* **1996,** *35*, 15280-7.

QM/MM conformational
free energy landscape

Experimental observation