# Constrained Catalytic Itinerary of a Retaining 3,6-Anhydro-D-Galactosidase, a Key Enzyme in Red Algal Cell Wall Degradation

Michael Wallace, Irene Cuxart, Thomas Roret, Laura Guée, Aleksandra Debowski, Mirjam Czjzek, Carme Rovira, Keith Stubbs, Elizabeth Ficko-Blean

How to cite: *Angew. Chem. Int. Ed.* **2024**, *63*, e202411171
doi.org/10.1002/anie.202411171

*Glycoside Hydrolases*

# Constrained Catalytic Itinerary of a Retaining 3,6-Anhydro-D-Galactosidase, a Key Enzyme in Red Algal Cell Wall Degradation

*Michael D. Wallace[+], Irene Cuxart[+], Thomas Roret, Laura Guée, Aleksandra W. Debowski, Mirjam Czjzek, Carme Rovira,* Keith A. Stubbs,* and Elizabeth Ficko-Blean**

**Abstract:** The marine Bacteroidota *Zobellia galactanivorans* has a polysaccharide utilization locus dedicated to the catabolism of the red algal cell wall galactan carrageenan and its unique and industrially important α-3,6-anhydro-D-galactose (ADG) monosaccharide. Here we present the first analysis of the specific molecular interactions that the exo-(α-1,3)-3,6-anhydro-D-galactosidase *Zg*GH129 uses to cope with the strict steric restrictions imposed by its bicyclic ADG substrate — which is ring flipped relative to D-galactose. Crystallographic snapshots of key catalytic states obtained with the natural substrate and novel chemical tools designed to mimic species along the reaction coordinate, together with quantum mechanics/molecular mechanics (QM/MM) metadynamics methods and kinetic studies, demonstrate a retaining mechanism where the second step is rate limiting. The conformational landscape of the constrained 3,6-anhydro-D-galactopyranose ring proceeds through enzyme glycosylation $B_{1,4}\rightarrow[E_4]^{\ddagger}\rightarrow E_4/{}^1C_4$ and deglycosylation $E_4/{}^1C_4\rightarrow[E_4]^{\ddagger}\rightarrow B_{1,4}$ itineraries limited to the *Southern* Hemisphere of the Cremer–Pople sphere. These results demonstrate the conformational changes throughout catalysis in a non-standard, sterically restrained, bicyclic monosaccharide, and provide a molecular framework for mechanism-based inhibitor design for anhydro-type carbohydrate-processing enzymes and for future applications involving carrageenan degradation. In addition, our study provides a rare example of distinct niche-based conformational itineraries within the same carbohydrate-active enzyme family.

Red algae are ancient,[1] photosynthetic, mainly marine eukaryotes. Their biomass is composed largely of cell wall components, including the complex carrageenan galactan (Figure 1A). Carrageenans have valuable gelling and stabilizing properties (e.g. the main gelling carrageenans, β-, ι- and κ-carrageenans),[2] though natural carrageenans are heterogeneous structures. Furthermore, carrageenan oligosaccharides are potential high-value products, interacting with human anti-Gal antibodies and galectin-3,[3] showing anti-tumour,[4] antihyperlipidemic[5] and antiviral[6] activities, and acting as agricultural biostimulants.[7] The galactose units in carrageenan are linked by alternating β-1,4 and α-1,3 glycosidic bonds, with further variation introduced through the addition of sulfate and other chemical groups to the free hydroxy groups.[8] The unique bicyclic monosaccharide α-3,6-anhydro-D-galactose (ADG) is formed in the polymer through the action of algal sulfurylases.[9] The 3,6-anhydro-bridge introduces strict structural constraints on the α-galactosyl unit resulting in a ring inversion (to a $^1C_4$

[*] Dr. M. D. Wallace,[+] Dr. A. W. Debowski, Dr. K. A. Stubbs
School of Molecular Sciences
The University of Western Australia
35 Stirling Hwy, Crawley, 6009, Australia
E-mail: keith.stubbs@uwa.edu.au

I. Cuxart,[+] Prof. C. Rovira
Department of Inorganic and Organic Chemistry (Section of Organic Chemistry) Institute of Computational and Theoretical Chemistry (IQTCUB)
Martí i Franquès 1, Barcelona, 08028, Spain
E-mail: c.rovira@ub.edu

T. Roret, Dr. M. Czjzek
CNRS, Sorbonne Université,
FR2424, Station Biologique de Roscoff
Roscoff, 29688, France

L. Guée, Dr. M. Czjzek, Dr. E. Ficko-Blean
CNRS, Sorbonne Université, UMR8227
Laboratory of Integrative Biology of Marine Models, Station Biologique de Roscoff
Roscoff, 29688, France
E-mail: efickoblean@sb-roscoff.fr

Dr. A. W. Debowski
School of Biomedical Sciences
The University of Western Australia
35 Stirling Hwy, Crawley, 6009, Australia

Prof. C. Rovira
Institució Catalana de Recerca i Estudis Avançats (ICREA)
Passeig Lluís Companys, 23, Barcelona, 08010, Spain

Dr. K. A. Stubbs
ARC Training Centre for Next-Gen Technologies in Biomedical Analysis, School of Molecular Sciences
The University of Western Australia
35 Stirling Hwy, Crawley, 6009, Australia

[+] These authors contributed equally to the work.

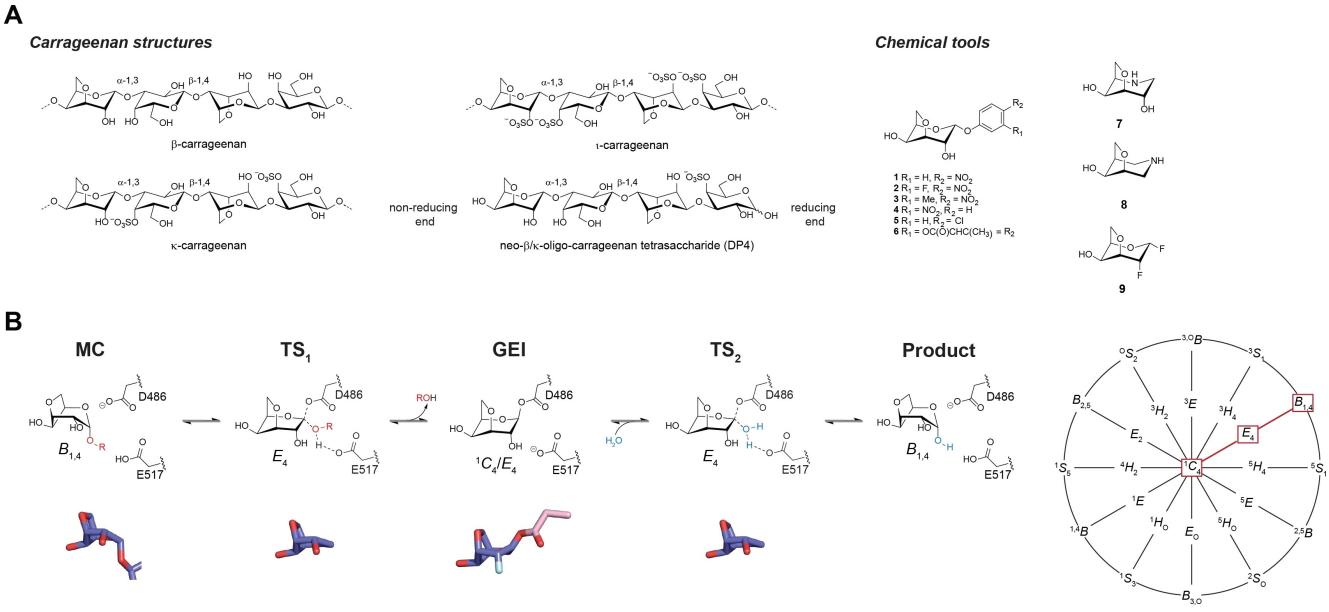


***Figure 1.*** A. Structures of the main gelling carrageenans β-, κ- and ι-carrageenan, neo-β/κ-oligo-carrageenan tetrasaccharide (DP4) and the chemical tools used to probe the catalytic mechanism and reaction geometry of *Zg*GH129: pNP-ADG **1**, 3-F-pNP-ADG **2**, 3-MepNP-ADG **3**, mNP-ADG **4**, 4-ClP-ADG **5**, MeUMB-ADG **6**, AD-DGJ **7**, ADG-IF **8**, 1,2-diF-ADG **9**. B. The catalytic mechanism of *Zg*GH129, where the conformational itinerary of the reaction coordinate passes through the *Southern* Hemisphere of the Cremer–Pople sphere,[27] shown are the Michaelis complex, MC; transition state 1, $TS_1$; glycosyl-enzyme intermediate, GEI; transition state 2, $TS_2$; and product. Snapshots of the −1 subsite trapped crystal structure complexes representing stages along this conformational itinerary are shown below the chemdraw structures of the reaction mechanism (From left to right: neo-β/κ-oligo-carrageenan DP4, AD-DGJ **7**, 1,2-diF-ADG **9**, and AD-DGJ **7**).

conformation), and consequently the α-linkage and 4-OH group are maintained equatorially rather than axially within the galactan.[10] Importantly, ADG drives polymer gel formation,[11] a rheological property capitalized extensively by industry.[12]

Red algal polysaccharides, including carrageenan, are carbon sources used by marine heterotrophic bacteria in the coastal carbon cycle.[13] Marine bacteria have developed highly specialized mechanisms for carrageenan catabolism, such as the organization of carrageenan-utilization genes into carrageenan polysaccharide utilization loci (CarPULs).[13b–e] In *Zobellia galactanivorans*, a marine model Bacteroidota isolated from a red alga,[14] many of the carrageenan-catabolism genes, such as the recently discovered 3,6-anhydro-D-galactosidases, are localized in a CarPUL, while the other necessary genes are coded elsewhere in the genome and are co-regulated with the CarPUL through a carrageenan regulon.[13b] The initial steps of carrageenan depolymerisation are carried out by highly specific endo-carrageenases[3,15] that hydrolyze the β-1,4-linkages to produce the neo-series of oligo-carrageenans, with ADG on the non-reducing terminus. *Z. galactanivorans* produces four exo-(α-1,3)-3,6-anhydro-D-galactosidases, glycoside hydrolases (GHs) in the established GH127[16] and GH129[17] carbohydrate-active enzyme (CAZy) families,[18] that hydrolyse non-reducing end ADG from carrageenan neo-oligosaccharides.[13b] The 3,6-anhydro-D-galactosidases are predicted to be retaining GHs for both families,[17,19] which are exemplified by the formation of a glycosyl-enzyme intermediate (GEI) during catalysis (Figure 1B). As the 3,6-anhydro-D-galactosidases have only been recently discovered,[13b] limited structural and enzymatic data is available, with no literature examples of real enzyme complexes, and both the catalytic residues and the catalytic mechanism have remained experimentally unassigned. To remediate this, we combined chemical, biochemical, structural and computational methods to provide insight into the specificity for *Zg*GH129 (Genbank ID CAZ97291.1) and the conformational changes ADG undergoes during the retaining catalytic mechanism.

To investigate the enzyme-substrate interactions and reactivity of 3,6-anhydro-D-galactosidases, we designed and synthesized compounds that could act as putative substrates and inhibitors, adding to the limited carrageenan-based chemical tools available.[20] To complement the known artificial substrate, 4-nitrophenyl 3,6-anhydro-α-D-galactoside **1**,[20b] suitable putative aryl 3,6-anhydroglycoside-based substrates **2**–**6** were prepared with different leaving groups (LGs) (Figure 1A). Next, two putative small molecule inhibitors were prepared, 3,6-anhydro-D-1-deoxygalactonojirimycin (AD-DGJ) **7**[21] and 3,6-anhydro-D-galacto-isofagomine (ADG-IF) **8**, as they may mimic the build-up of partial positive charge on the O5 or C1 of the putative transition state (TS), respectively. Finally, a proposed retaining 2-deoxy-2-fluoro mechanism-based inhibitor,[22] 1,2-diF-ADG **9**, was prepared (See Supporting Information).

With these compounds in hand, attention was turned to their evaluation against the 3,6-anhydro-D-galactosidase *Zg*GH129 (Figure S1). Michaelis–Menten parameters for

the putative substrates **2**–**6** were determined at the pH optimum from the analysis of reaction rates at steady-state conditions (Table S1). Brønsted plots of $\log(k_{cat})$ and $\log(k_{cat}/K_m)$ versus $pK_a$ of the aryl LG were generated (Figure 2A). The Brønsted coefficient of $\beta_{lg}(k_{cat}) \approx 0$ demonstrates that there is no dependence of rate on LG ability, which suggests that the rate determining step is not glycosylation and is either the deglycosylation step or a non-chemical step after the formation of the GEI.[23] For the plot $\log(k_{cat}/K_m)$, there is a good correlation ($R^2 = 0.88$) however, the value for **6** does not fit the linear regression well, likely due to a poor approximation of the relevant second order rate constant. The steric bulk of the MeUMB group can affect the observed kinetics and linear free-energy relationships should be conducted with substrates in which the only perturbation is electronic, thus this may be confounding the results.[24] If **6** is removed the Brønsted coefficient $\beta_{lg}(k_{cat}/K_m) = -0.16 \pm 0.03$ is obtained and the fit is improved ($R^2 = 0.95$) (Figure S2). The small negative $\beta_{lg}(k_{cat}/K_m)$ indicates that there is a build-up of negative charge on the exocyclic oxygen during the TS, suggestive of a late TS where cleavage of the glycosidic bond and proton donation are advanced[25] which is consistent with the general structures of TSs in the GH superfamily.[26]

The putative inhibitors demonstrated vastly different potencies when evaluated using **1** as a substrate. AD-DGJ **7** demonstrated poor inhibition of *Zg*GH129 even at a [**7**] of 8 mM and the pH optimum of the enzyme (Figure S3). This is in contrast to Sato et al.[28] who found 2-acetamido-1,2-dideoxy-D-*galacto*-nojirimycin to be a potent inhibitor of *Bb*GH129 ($K_i = 51$ nM), a gut bacterial exo-α-*N*-acetyl-galactosaminidase from *Bifidobacterium bifidum* which hydrolyzes non-reducing end $^4C_1$ α-*N*-acetyl-D-galactosamine (D-GalNAc) moieties.[17] These observations are likely related to the substantially different natural substrates which would affect their conformational itineraries and the efficacy of putative TS-based inhibitors.[29] Interestingly, these enzymes fall into two niche-based phylogenetic clades in the GH129 family, human host bacterial enzymes (including *Bb*GH129) and marine bacterial enzymes (including *Zg*GH129).[20b]

Pleasingly, ADG-IF **8** was a potent inhibitor of *Zg*GH129 with a $K_i$ value of $1.50 \pm 0.06$ μM (Figure 2B). This result is consistent with the general observation that azasugars with the nitrogen at the anomeric position are potent inhibitors of glycosidases that catalyse the cleavage of substrates with equatorial aglycones (typically β-glycosides, but here an equatorial α-linkage).[30] Incubation of *Zg*GH129 with 1,2-diF-ADG **9** resulted in time-dependent inactivation of the enzyme. (Figure 2CD). The kinetic parameters governing inactivation ($k_i$ and $K_i$) were calculated by plotting $k_{obs}$ versus [**9**]. A $K_i$ value of $68.0 \pm 5.6$ μM and a $k_i$ value of $0.20 \pm 0.01$ min$^{-1}$ was obtained for **9** which is consistent with inhibition through enzyme inactivation. This supports that *Zg*GH129 and family GH129 enzymes in general indeed proceed through a GEI.

The active site of *Zg*GH129 was first identified from the native structure (PDB ID 5OPQ) based on mutagenesis and its CAZy family relationship to *Bb*GH129.[13b,17] To gain more insight into inhibitor interactions with *Zg*GH129, as well as the mechanism of catalysis, we successfully determined the structure in complex with AD-DGJ **7**, ADG-IF **8**, 1,2-diF-ADG **9**, a natural neo-β/κ-oligo-carrageenan tetrasaccharide (neo-β/κ-DP4) as well as ADG alone (Figure 3A–E, Table S2). *Zg*GH129 crystallizes as a tetramer of two biological homodimers in the asymmetric unit.[13b] Both monomers contribute to two active sites sitting at the bottom of a deep crevasse (Figure 3F). The active site residues include the predicted catalytic nucleophile D486 (D486N in the Michaelis complex, MC), the predicted catalytic acid/base residue E517 and other residues lining the active site: N218, T220, P346, H347,

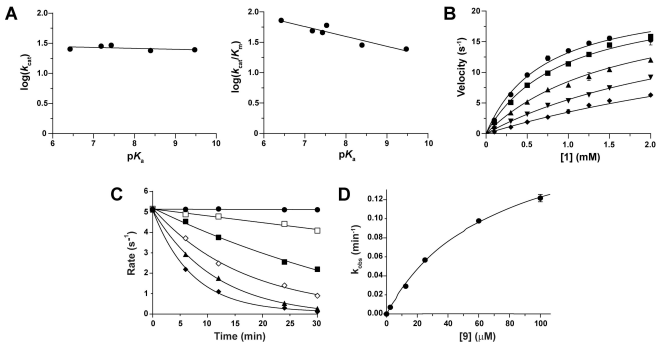


***Figure 2.*** A. Plots of $\log(k_{cat})$ and $\log(k_{cat}/K_m)$ against the p*K*a of the leaving group phenol for substrates **1**–**6**; B. $K_i$ determination for the inhibition of *Zg*GH129 with **8**. [**8**] used in $K_i$: 0 (●), 0.5 (■), 2 (▲), 5 (▼) and 10 (♦) μM in pH 5 buffer; C. Time-dependent inactivation of *Zg*GH129 by **9**. [**9**] used: 0 (●), 2.5 (□), 12.5 (■), 25 (◇), 60 (▲) and 100 (♦) μM. Curves are the nonlinear fits of data to a single exponential decay equation. D. Plot of the inactivation rate constants ($k_{obs}$) as a function of the concentration of **9** for *Zg*GH129. All data points were done in technical triplicates and the error bars represent standard error of the mean.

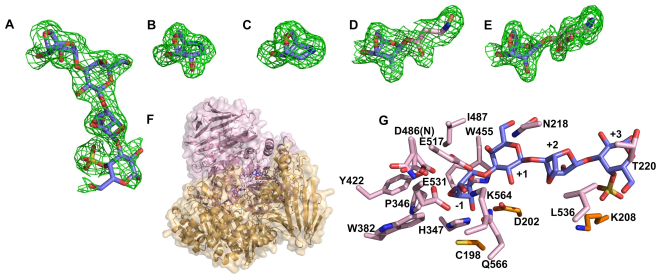


***Figure 3.*** Electron density omit maps (maximum-likelihood/$\sigma_A$-weighted[31] Fobs-Fcalc contoured at 3 σ) were generated from the X-ray crystallography refinements (Table S2) by omitting the following molecules: A. neo-β/κ-oligo-carrageenan tetrasaccharide (neo-β/κ-DP4); B. AD-DGJ **7**; C. ADG-IF **8**; D. 1,2-di-F-ADG **9**; E. ADG; F. *Zg*GH129 surface and secondary structure representation showing the two homodimeric active sites at the bottom of the deep crevasse, pictured with the neo-β/κ-DP4; G. Zoom on the ZgGH129_D486N active site residues that interact with neo-β/κ-DP4. Amino acids are shown in light pink for monomer A and light orange for monomer B. The two additional conformations of D486 are from the superposed ADG complex. The −1, +1, +2, and +3 subsite sugars are labelled.

W382, Y422, W455, I487, E531, L536, K564 and Q566 from the first monomer of the homodimer, and C198, D202 and K208 from the second monomer (Figure 3G).

The MC was obtained with a catalytically inactive nucleophile mutant *Zg*GH129_D486N[13b] and neo-β/κ-DP4 (Figure 3AFG). The terminal non-reducing ADG moiety is bound in the −1 subsite, a D-galactosyl, a second ADG and a 4-sulfo-D-galactosyl moiety are found in the +1, +2 and +3 subsites, respectively. The residue W455 is integral to the −1 subsite, forming a hydrophobic support alongside the cyclic O-C5-C6 edge of the ADG. The −1 subsite ADG is in a $B_{1,4}$ conformation with the C1 and C4 positioned below the pyranose ring plane. This conformation places the glycosidic bond and LG into an axial orientation, which reduces steric hinderance from neighbouring groups, presumably lining up the D486 residue (in a wild-type setting) for nucleophilic attack towards the σ* anti-bonding orbital of C1.[32] The MC D486N residue is shifted relative to the D486 residue in the other complexes; however, even the unmodified D486 is capable of movement, as demonstrated in the ADG complex which has two clear D486 conformations (Figure 3EG). The MC also confirms structurally that *Zg*GH129 prefers neo-β-carrabiose on the non-reducing end in the −1 and +1 subsites.[20b] The +2 subsite, which holds a $^1C_4$ ADG, appears capable of accommodating a 2-sulfo modification on the ADG residue and the +3 subsite Lys208 forms a salt bridge with the 4-sulfo-moiety of the D-galactosyl unit, these results support prior MS/MS analyses.[20b] Intriguingly, two oligosaccharides were captured in the depths of each of the active sites at the bottom of the dimer crevasse which is 40 Å deep. The reducing ends incline upwards and towards each other as they reach for the crevasse opening (Figure 3F). It is tempting to speculate that the deep subsite binding interactions function to separate and stabilize the long helix-forming (and gellifying) carrageenan strands for enzyme accessibility.[33]

The inactivation inhibitor 1,2-DiF-ADG **9** was successfully used to trap the GEI. A covalent bond is observed between the anomeric carbon and the catalytic nucleophile D486 (Figure 3D). The conformation of the pyranose is distorted between an $E_4$ and $^1C_4$ conformation. This distortion might reduce 1,3-diaxial interactions between the C1-D486 axial bond and the 3,6-anhydro-bridge.[32] The glycosyl-enzyme glycosidic linkage is axially oriented to reduce steric hinderance prior to water nucleophilic attack for the deglycosylation step of the enzyme reaction. In this complex, as well as in the competitive inhibitor complexes, there is a conserved water which sits primed for nucleophilic attack directly below C1. It is coordinated by the putative acid/base catalytic residue E517. The AD-DGJ **7**- and ADG-IF **8** complexes were also obtained with both in the −1 subsite, despite **7** being a poor inhibitor of *Zg*GH129 (Figure 3BC). AD-DGJ **7** adopts an $E_4$ conformation, consistent with the putative planar TS. ADG-IF **8** adopts an $E_4/^1C_4$ conformation with the C1 pointing above the ring plane. The observed ADG-IF **8** conformation is similar to the trapped GEI, suggesting it is not a TS mimic.

Surprisingly, soaks with ADG resulted in a partial GEI complex (Figure 3E). The catalytic nucleophile, D486, shows significant movement between the covalent GEI and the absence of a covalent bond with the ADG. The electron density shown is an averaged representation of the equilibrium between the GEI, product and off state form of ADG. Thus, the refinement of the monosaccharide into the electron density represents three states in equilibrium and results in a 'flattened' glycosidic bond appearance modelled for the $^1C_4$ ADG GEI complex. In the carrageenan polymer, due to the steric restraints imposed by the 3,6-anhydro-bridge, ADG motifs are found in the $^1C_4$ conformation.[10] These structural results suggest a D-galactosidase *Southern* Hemisphere conformational catalytic itinerary (Figure 1B), normally associated with some mannosidase and neuraminidase families or with enzymes processing L-sugars like GH29 L-fucosidases.[34]

A structural overlay of *Bb*GH129[28] with *Zg*GH129 (RMSD 2.8 Å) shows an almost perfect overlap of *Bb*D435 (catalytic nucleophile) and *Bb*E478 (catalytic acid/base) with the corresponding residues in *Zg*GH129 (Figure 4ABC, see Supporting Information and Figure S4). The active site of *Bb*GH129 contains a divalent cation coordinating the 2-acetamido group, there are no active site metal ions in *Zg*GH129. *Bb*W398 lies perpendicular in relation to *Zg*W455 and provides a hydrophobic platform maintaining the axial C4-OH of $^4C_1$ D-GalNAc. *Bb*D330 provides an H-bond to the axial C4-OH. *Zg*W455 is replaced by *Bb*D371 which is within H-bonding distance of the D-GalNAc C6-OH. Evidently this interaction is unnecessary in *Zg*GH129 as the C6 forms part of the 3,6-anhydro-bridge. *Zg*D202 from the second *Zg*GH129 monomer is found in a similar position to *Bb*W398. Along with *Zg*H347, *Zg*D202 provides an H-bond with the equatorial C4-OH. This H-bond seems to be essential as a conservative mutant (*Zg*D202N) abolishes enzyme activity.[13b] Notably, the active-site participating helix from the second monomer of *Zg*GH129, which supports the *Zg*D202 residue, is not found in *Bb*GH129 (Figure 4DEF). Relatedly, *Zg*GH129 forms a homodimer in the crystal structure, *Bb*GH129 exists as a monomer. This comparison shows how GH129 uses an enzyme scaffold with small modifications to the active site and larger modifications to secondary structure to alter substrate preference and provide niche-based recognition of different sugar conformations.

We next turned to quantum mechanics/molecular mechanics (QM/MM) metadynamics simulations to further unveil the molecular mechanism of catalysis (See Supporting Information). Using the crystal structure of *Zg*GH129 neo-β/κ-DP4, we reconstructed the MC, reverted the D486N mutation, and equilibrated the structure for 200 ns of MD simulations to show that both subunits accommodate the neo-β/κ-DP4 in four well-defined subsites (Figure S5).

To gain insight into what conformations are accessible to ADG in the MC, we computed the conformational free energy landscape (FEL) of the ADG unit of the native neo-β/κ-DP4 using QM/MM metadynamics simulations, along with Cremer–Pople ring puckering coordinates[27] as collec-

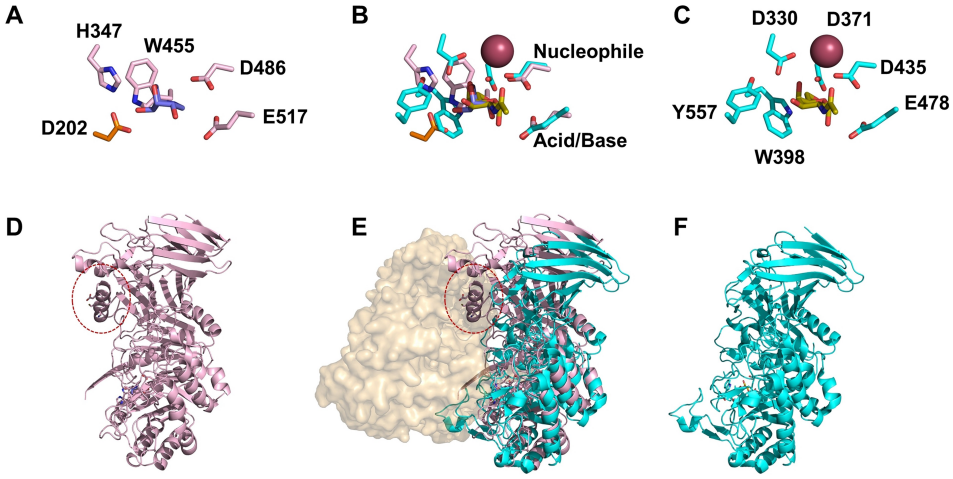


***Figure 4.*** Structural overlay of *Zg*GH129 in complex with AD-DGJ **7** and *Bb*GH129 PDB ID 5WZN in complex with D-GalNAc show that while there are similarities, such as in the positions of the catalytically active residues, there are major structural differences. A. Active site of *Zg*GH129, amino acids are shown in light pink for monomer A and light orange for monomer B. B. Active site overlay of *Zg*GH129 and *Bb*GH129. C. Active site of *Bb*GH129 shown in light blue. D. Secondary structure representation of one *Zg*GH129 monomer with the active site participating helix circled in red. E. Overlay of *Zg*GH129 and *Bb*GH129, the other monomer in the homodimer of *Zg*GH129 is shown in surface representation. F. Secondary structure representation of *Bb*GH129.

tive variables. The results (Figure 5A) show that that the enzyme allows ADG to adopt a $B_{1,4}$ conformation preferentially, consistent with the conformation experimentally observed in the MC (a $^{1}C_{4}$ conformation is also possible, Figure S6).

To model the glycosylation mechanism of *Zg*GH129, we selected a structure corresponding to the most stable $B_{1,4}$ free energy minimum, in which the LG is axial (Figure S7). Here, the D486 nucleophile residue is close to the anomeric carbon of the ADG unit (C1-$O_{D486}$ = 3.20 ± 0.17 Å, Table S3) and E517 forms an H-bond with the glycosidic oxygen of the scissile glycosidic bond (O1-$H_{E517}$ = 1.51 ± 0.09 Å, Table S3), indicating that the substrate is prepared for catalysis. To model the glycosylation step, two collective variables were used to drive the reaction (Figure S8). The first CV (CV1) was defined as the difference of C1-$O_{D486}$ and C1-O1, describing the nucleophilic attack, whereas the second one (CV2) was taken as the $H_{E517}$-$O_1$ distance, to account for the protonation of the LG by the catalytic acid residue E517. The FEL obtained from the simulation shows two main minima and a single TS, indicating a concerted one-step mechanism. The computed free energy barrier (14.2 kcal/mol) agrees with the one that can be estimated from the experimental rate constant (for the substrate **1** it was calculated as 15.6 kcal/mol) (Table S1). There is an additional minimum corresponding to a pre-catalytic complex, MC', separated from the MC by a small energy barrier. It corresponds to $H_{E517}$ switching from forming an H-bond with the 2-OH of ADG to the glycosidic oxygen (Figure S9).

Analysis of the reaction coordinate provides an atomistic view of the glycosylation step (Figure 5B). Once the acid/base residue starts forming an H-bond with the glycosidic oxygen, the glycosidic bond starts to break and the ADG evolves towards an envelope conformation ($E_4$). Afterwards, the D486 nucleophilic residue approaches the C1 atom and a proton transfers from the E517 general acid catalyst to the glycosidic oxygen O1. At the TS ($TS_1$, Figure 5B), the ADG motif is in a planar $E_4$ conformation, consistent with the stereoelectronic properties stabilizing an oxocarbenium ion-like TS. The C1-O5 bond decreases from 1.39 ± 0.03 Å at MC to 1.29 ± 0.01 Å at $TS_1$ (Table S3), reflecting the partial double bond character between the anomeric carbon and the endocyclic oxygen. The C1-O1 is almost broken at $TS_1$ (C1-O1 = 2.19 ± 0.04 Å) but the glycosyl-enzyme bond is not yet formed (C1-$O_{D486}$ = 2.62 ± 0.05 Å), indicative of a dissociative mechanism. H-bond interactions help to maintain a tight active site during the glycosylation step, namely those between D202 and the 4-OH of ADG and the interaction of the Y422 side chain with D486. There is also a water-mediated interaction between the 2-OH of the ADG and E531 (Table S3). The ADG keeps these interactions while evolving towards the GEI, following a $B_{1,4} \rightarrow [E_4]^{\ddagger} \rightarrow E_4/^{1}C_4$ conformational itinerary (Supporting Video File 1 and Figure S8B).

To model the deglycosylation step of catalysis, we reconstructed the natural GEI from the crystal structure of the enzyme covalently bound to 1,2-diF-ADG **9** and replaced the fluorine at C2 of **9** by a hydroxyl group (Figure S10). The corresponding complex (GEI' in Figure 5B), equilibrated by 300 ns of classical MD simulations, agreed with the GEI complex obtained in the QM/MM metadynamics simulations of the glycosylation step. The simulations show that a water molecule enters the active site and is accommodated underneath the anomeric carbon and coordinated by the general base residue E517, which is in its unprotonated state at this stage of the mechanism and in an orientation suitable for hydrolysis. The water-mediated interaction between the 2-OH of ADG and the

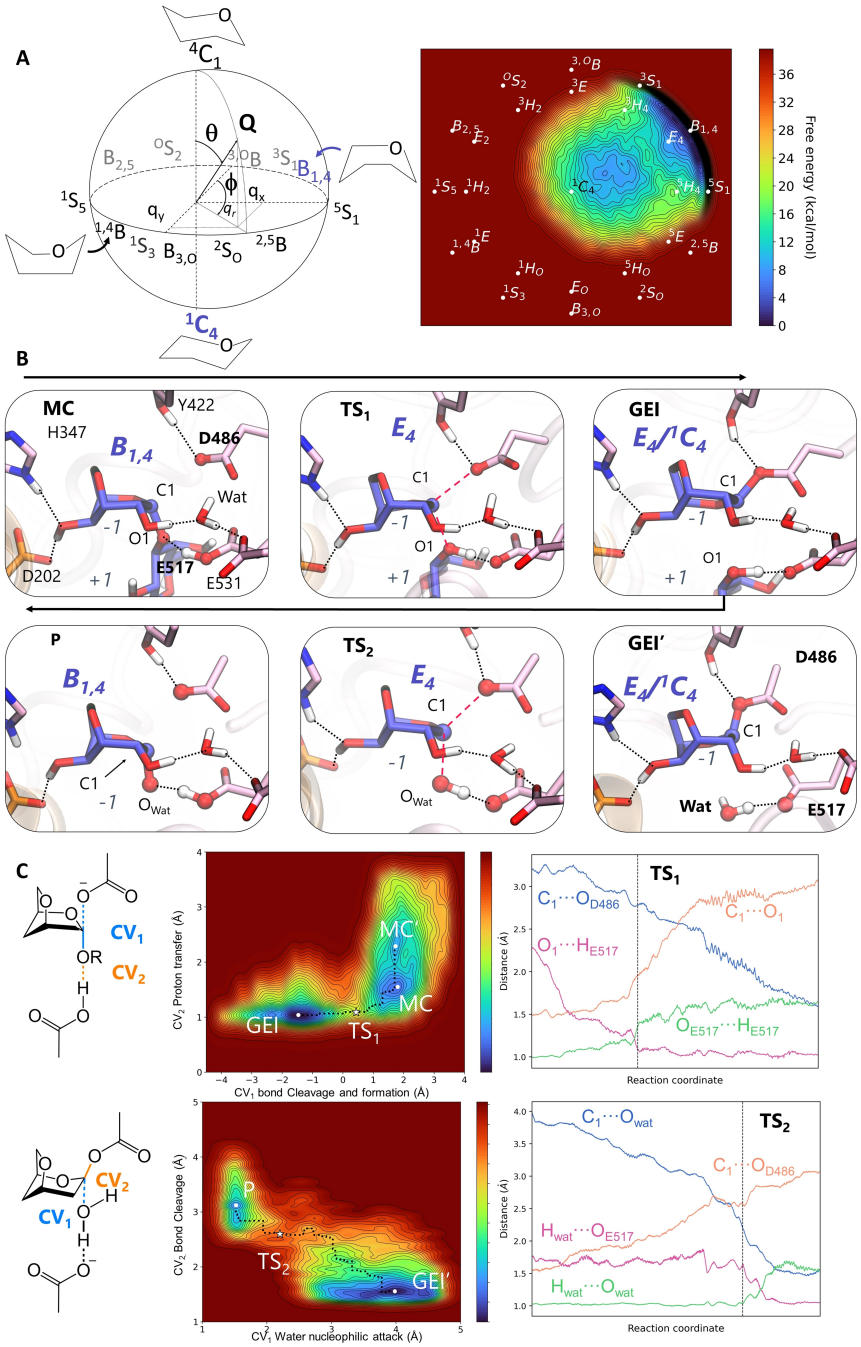


***Figure 5.*** A. Left: Cremer–Pople sphere depicting the possible conformations of a pyranose ring according to polar coordinates (Q, θ and φ) projected in the cartesian $q_x$ and $q_y$. *Zg*GH129 operates in the *Southern* Hemisphere. Right: Conformational free energy landscape of ADG at the −1 subsite of *Zg*GH129 projected into Stoddard's diagram. Contour lines at 1 kcal/mol. B. Representative structures along the minimum free energy pathway (reaction coordinate) of the glycosylation and deglycosylation reactions. C. Reaction free energy landscapes (glycosylation and deglycosylation) and evolution of the main catalytic distances along the reaction coordinate. Contour lines at 1 kcal/mol.

other glutamate of the active site, E531, was also observed during deglycosylation. E531 is thus key for binding the substrate but does not participate in the reaction.

During the deglycosylation reaction the GEI (GEI' in Figure 5B) shows the water molecule, described just above, forming an H-bond with the catalytic base ($H_{wat}$–$O_{E517}$ = 1.71 ± 0.13 Å), close to the reactive anomeric carbon (C1-$O_{wat}$ = 3.86 ± 0.25 Å) and in a favorable configuration for nucleophilic attack. Two CVs were defined to drive the reaction towards product state (Figure S11). Similar to the glycosylation reaction, the deglycosylation reaction exhibits a single TS ($TS_2$, Figure 5B), in which the ADG adopts an $E_4$ conformation. The reaction is dissociative, as the C1-$O_{D486}$ covalent bond is already broken, while the new bond with $O_{wat}$ is not yet formed (C1-$O_{wat}$ = 2.17 ± 0.05 Å, C1-$O_{D486}$ = 2.61 ± 0.06 Å). The C1-O5 bond shrinks from 1.39 ±

0.03 Å at GEI' to 1.31 ± 0.01 Å at $TS_2$, again indicative of the formation of an oxocarbenium ion-like species. The reaction evolves towards the product complex, following a $E_4/^1C_4 \rightarrow [E_4]^{\ddagger} \rightarrow B_{1,4}$, conformational itinerary that mirrors that of the glycosylation step (Supporting Video File 2). The computed reaction free energy barrier (Figure 5C) of the deglycosylation step (17.2 kcal/mol) is higher than that of the glycosylation step (14.2 kcal/mol). Therefore, both the kinetic data and our simulations predict that the glycosylation reaction is not the rate-limiting step during catalysis.

A combination of chemical and structural analyses coupled with QM/MM metadynamics simulations support a two-step, retaining catalytic mechanism for the hydrolysis of (α-1,3)-3,6-anhydro-D-galactose by *Zg*GH129 wherein the residues D486 and E517 have the catalytic roles of nucleophile and general acid/base, respectively (Figure 1B). To the best of our knowledge, all characterized catalytic itineraries of D-galactosidases are located across the *Northern* Hemisphere[35] (see Supporting Information and Table S4), despite the isolated D-galactose free energy landscape showing that the inverted $^1C_4$ chair is one of its accessible conformations.[36] *Zg*GH129 and the rigid ADG pyranose ring restrict the conformational landscape, only the distorted $B_{1,4}$ conformation is used for catalysis which proceeds through $B_{1,4} \rightarrow [E_4]^{\ddagger} \rightarrow E_4/^1C_4$ and $E_4/^1C_4 \rightarrow [E_4]^{\ddagger} \rightarrow B_{1,4}$ itineraries (glycosylation and deglycosylation, respectively) across the *Southern* Hemisphere of the Cremer–Pople sphere, with the deglycosylation step being rate-limiting. Overall, we demonstrate novel insights into glycosidase catalysis by establishing the fine molecular details of glycan conformational changes in a non-standard, sterically restrained, bicyclic monosaccharide. This provides an example of how distinct conformational itineraries are accommodated within the same CAZy family and molecularly explains the strict partitioning of the GH129 family into two niche-based clades.[20b] It also establishes a framework to assist rational engineering for mechanism-based inhibitor design of anhydro-type carbohydrate-processing enzymes and for future applications involving these and other carrageenan-interacting enzymes.

## Supporting Information

The authors have cited additional references within the Supporting Information (Ref. [37–88]).

Supporting Information File 1.

Supporting Video File 1. Simulation trajectories corresponding to the glycosylation step showing the reactive ADG and catalytic residues D486 and E517. The atoms involved in the reactive events are highlighted in ball representation.

Supporting Video File 2. Simulation trajectories corresponding to the deglycosylation step showing the reactive ADG, catalytic residues D486 and E517 and the catalytic water molecule. The atoms involved in the reactive events are highlighted in ball representation.

## Acknowledgements

MDW and KAS acknowledge the facilities, and the scientific and technical assistance of Microscopy Australia at the Centre for Microscopy, Characterisation and Analysis, The University of Western Australia, a facility funded by the University, State and Federal Commonwealth Governments. MDW is supported by a Research Training Program Scholarship provided by the Australian Federal Government and The University of Western Australia. MDW also thanks the School of Molecular Sciences for the Dr Wayne Best Travel Award, and Campus France and the Franco-Australian Hubert Curien Programme (FASIC) for funding. CR acknowledges the Spanish Ministry of Science and Innovation (MICINN/AEI/FEDER, UE, PID2020-118893GB-100), the Spanish Structures of Excellence María de Maeztu (CEX2021-001202-M), the Agency for Management of University and Research Grants of Catalonia (AGAUR, 2021-SGR-00680) and the European Research Council (ERC-2020-SyG-951231 "Carbocentre"). IC and CR would like to thank the technical support provided by the Barcelona Supercomputing Center (BSC) and Red Nacional de Supercomputación (RES) for computer resources at MareNostrum IV and CTE-Power supercomputers. EF-B acknowledges support from the Agence National de la Recherche (ANR) through the ANR Mirror Mirror (ANR-22-CE11-0025-01). EF-B would like to thank Genomer Platform and the Biogenouest Crystallography Platform at the Station Biologique de Roscoff and the SOLEIL synchrotron staff on the Proxima 1 and 2 beamlines for their important technical support.

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

Data is available in Supporting Information File 1, Supporting Video File 1, Supporting Video File 2 and on Zenodo DOI 10.5281/zenodo.10789035.

PDB files can be found on the RCSB Protein Data Bank page https://www.rcsb.org/.

ZgGH129 ADG: pdb id 8RZG
ZgGH129 AD-DGJ: pdb id 8RZH
ZgGH129 1,2-diF-ADG: pdb id 8RZI
ZgGH129 ADG-IF: pdb id 8RZJ
ZgGH129 neo-Î²/Î°-DP4: pdb id 8RZK

**Keywords:** enzyme mechanism · oligosaccharides · structural biology · Cremer–Pople sphere · glycoside hydrolase

[1] S. Bengtson, T. Sallstedt, V. Belivanova, M. Whitehouse, *PLoS Biol.* **2017**, *15*, e2000735.
[2] M. Ciancia, M. C. Matulewicz, R. Tuvikene, *Front. Plant Sci.* **2020**, *11*, 559986.

[3] E. Sokolova, D. Jouanneau, A. Chevenier, M. Jam, N. Desban, P. Colas, E. Ficko-Blean, G. Michel, *Carbohydr. Polym.* **2024**, *324*, 121563.
[4] a) X. K. Hu, X. L. Jiang, E. Aubree, P. Boulenguer, A. T. Critchley, *Pharm. Biol.* **2006**, *44*, 646–650; b) H. J. Mou, X. L. Jiang, H. S. Guan, *J. Appl. Phycol.* **2003**, *15*, 297–303.
[5] E. V. Sokolova, L. N. Bogdanovich, T. B. Ivanova, A. O. Byankina, S. P. Kryzhanovskiy, I. M. Yermak, *PharmaNutrition* **2014**, *2*, 33–37.
[6] M. Ludwig, E. Enzenhofer, S. Schneider, M. Rauch, A. Bodenteich, K. Neumann, E. Prieschl-Grassauer, A. Grassauer, T. Lion, C. A. Mueller, *Respir. Res.* **2013**, *14*, 124.
[7] a) C. Lemonnier-Le Penhuizic, C. Chatelet, B. Kloareg, P. Potin, *Plant Sci.* **2001**, *160*, 1211–1220; b) A. Moenne, A. Gonzalez, *Carbohydr. Res.* **2021**, *503*, 108298.
[8] Z. A. Popper, G. Michel, C. Herve, D. S. Domozych, W. G. T. Willats, M. G. Tuohy, B. Kloareg, D. B. Stengel, *Annu. Rev. Plant Biol.* **2011**, *62*, 567–590.
[9] S. Genicot-Joncour, A. Poinas, O. Richard, P. Potin, B. Rudolph, B. Kloareg, W. Helbert, *Plant Physiol.* **2009**, *151*, 1609–1616.
[10] E. Ficko-Blean, C. Hervé, G. Michel, *PiP* **2015**, *2*, 51–64.
[11] a) G. A. De Ruiter, B. Rudolph, *Trends Food Sci. Technol.* **1997**, *8*, 389–395; b) F. van de Velde, S. H. Knutsen, A. I. Usov, H. S. Rollema, A. S. Cerezo, *Trends Food Sci. Technol.* **2002**, *13*, 73–92.
[12] F. Ferdouse, S. L. Holdt, R. Smith, P. Murua, Z. Yang, *Food and Agriculture Organization of the United Nations* **2018**, *124*.
[13] a) M. Martin, D. Portetelle, G. Michel, M. Vandenbol, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 2917–2935; b) E. Ficko-Blean, A. Prechoux, F. Thomas, T. Rochat, R. Larocque, Y. T. Zhu, M. Stam, S. Genicot, M. Jam, A. Calteau, B. Viart, D. Ropartz, D. Perez-Pascual, G. Correc, M. Matard-Mann, K. A. Stubbs, H. Rogniaux, A. Jeudy, T. Barbeyron, C. Medigue, M. Czjzek, D. Vallenet, M. J. McBride, E. Duchaud, G. Michel, *Nat. Commun.* **2017**, *8*; c) L. Christiansen, D. Pathiraja, P. K. Bech, M. Schultz-Johansen, R. Hennessy, D. Teze, I. G. Choi, P. Stougaard, *mSphere* **2020**, *5*; d) M. Schultz-Johansen, P. K. Bech, R. C. Hennessy, M. A. Glaring, T. Barbeyron, M. Czjzek, P. Stougaard, *Front. Microbiol.* **2018**, *9*, 839; e) A. G. Hettle, J. K. Hobbs, B. Pluvinage, C. Vickers, K. T. Abe, O. Salama-Alber, B. E. McGuire, J. H. Hehemann, J. P. M. Hui, F. Berrue, A. Banskota, J. Zhang, E. M. Bottos, J. Van Hamme, A. B. Boraston, *Commun. Biol.* **2019**, *2*, 474; f) A. Gobet, T. Barbeyron, M. Matard-Mann, G. Magdelenat, D. Vallenet, E. Duchaud, G. Michel, *Front. Microbiol.* **2018**, *9*, 2740.
[14] T. Barbeyron, F. Thomas, V. Barbe, H. Teeling, C. Schenowitz, C. Dossat, A. Goesmann, C. Leblanc, F. Oliver Glockner, M. Czjzek, R. Amann, G. Michel, *Environ. Microbiol.* **2016**, *18*, 4610–4627.
[15] a) E. Rebuffet, T. Barbeyron, A. Jeudy, M. Jam, M. Czjzek, G. Michel, *Biochemistry* **2010**, *49*, 7590–7599; b) H. J. Mou, X. L. Jiang, Z. H. Liu, H. Guan, *J. Food Biochem.* **2004**, *28*, 245–260; c) M. Matard-Mann, T. Bernard, C. Leroux, T. Barbeyron, R. Larocque, A. Prechoux, A. Jeudy, M. Jam, P. N. Collen, G. Michel, M. Czjzek, *J. Biol. Chem.* **2017**, *292*, 19919–19934.
[16] K. Fujita, S. Sakamoto, Y. Ono, M. Wakao, Y. Suda, K. Kitahara, T. Suganuma, *J. Biol. Chem.* **2011**, *286*, 5143–5150.
[17] M. Kiyohara, T. Nakatomi, S. Kurihara, S. Fushinobu, H. Suzuki, T. Tanaka, S. I. Shoda, M. Kitaoka, T. Katayama, K. Yamamoto, H. Ashida, *J. Biol. Chem.* **2012**, *287*, 693–700.
[18] E. Drula, M. L. Garron, S. Dogan, V. Lombard, B. Henrissat, N. Terrapon, *Nucleic Acids Res.* **2022**, *50*, D571–D577.
[19] K. Fujita, Y. Takashi, E. Obuchi, K. Kitahara, T. Suganuma, *J. Biol. Chem.* **2014**, *289*, 5240–5249.
[20] a) M. D. Wallace, E. Ficko-Blean, K. A. Stubbs, *J. Org. Chem.* **2020**, *85*, 16182–16195; b) M. D. Wallace, L. Guee, D. Ropartz, M. Fanuel, G. Lannuzel, G. Correc, K. A. Stubbs, E. Ficko-Blean, *Int. J. Biol. Macromol.* **2020**, *163*, 1471–1479.
[21] P. Szolcsanyi, T. Gracza, *Tetrahedron* **2006**, *62*, 8498–8502.
[22] B. P. Rempel, S. G. Withers, *Glycobiology* **2008**, *18*, 570–586.
[23] D. J. Vocadlo, S. G. Withers, *Biochemistry* **2005**, *44*, 12809–12818.
[24] a) M. M. Mader, P. A. Bartlett, *Chem. Rev.* **1997**, *97*, 1281–1302; b) I. P. Street, J. B. Kempton, S. G. Withers, *Biochemistry* **1992**, *31*, 9970–9978.
[25] N. Cetinbas, M. S. Macauley, K. A. Stubbs, R. Drapala, D. J. Vocadlo, *Biochemistry* **2006**, *45*, 3835–3844.
[26] G. J. Davies, M. L. Sinnott, S. G. Withers, in *Comprehensive Biological Catalysis: A Mechanistic Reference, Vol. 1* (Ed.: M. L. Sinnott), Academic Press, San Diego, **1998**, pp. 119–208.
[27] D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354–1358.
[28] M. Sato, D. Liebschner, Y. Yamada, N. Matsugaki, T. Arakawa, S. S. Wills, M. Hattie, K. A. Stubbs, T. Ito, T. Senda, H. Ashida, S. Fushinobu, *J. Biol. Chem.* **2017**, *292*, 12126–12138.
[29] V. A. Money, N. L. Smith, A. Scaffidi, R. V. Stick, H. J. Gilbert, G. J. Davies, *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 5136–5140.
[30] D. L. Zechel, S. G. Withers, *Acc. Chem. Res.* **2000**, *33*, 11–18.
[31] R. J. Read, *Acta Crystallogr. Sect. A* **1986**, *A42*, 140–149.
[32] D. J. Vocadlo, G. J. Davies, *Curr. Opin. Chem. Biol.* **2008**, *12*, 539–555.
[33] M. Diener, J. Adamcik, A. Sanchez-Ferrer, F. Jaedig, L. Schefer, R. Mezzenga, *Biomacromolecules* **2019**, *20*, 1731–1739.
[34] a) G. J. Davies, A. Planas, C. Rovira, *Accounts Chem Res* **2012**, *45*, 308–316; b) C. Rovira, A. Males, G. J. Davies, S. J. Williams, *Curr. Opin. Struct. Biol.* **2020**, *62*, 79–92; c) Y. Jin, M. Petricevic, A. John, L. Raich, H. Jenkins, L. Portela De Souza, F. Cuskin, H. J. Gilbert, C. Rovira, E. D. Goddard-Borger, S. J. Williams, G. J. Davies, *ACS Cent. Sci.* **2016**, *2*, 896–903; d) A. Lammerts van Bueren, A. Ardevol, J. Fayers-Kerr, B. Luo, Y. Zhang, M. Sollogoub, Y. Bleriot, C. Rovira, G. J. Davies, *J. Am. Chem. Soc.* **2010**, *132*, 1804–1806.
[35] J. L. de Meirelles, F. C. Nepomuceno, J. Pena-Garcia, R. R. Schmidt, H. Perez-Sanchez, H. Verli, *J. Chem. Inf. Model.* **2020**, *60*, 684–699.
[36] A. Nin-Hill, C. Rovira, *ACS Catal.* **2020**, *10*, 12091–12097.

Manuscript received: June 13, 2024
Accepted manuscript online: July 18, 2024
Version of record online: September 17, 2024