



HAL
open science

Innovating glycoside hydrolase activity on a same structural scaffold

Mirjam Czjzek, Gurvan Michel

► **To cite this version:**

Mirjam Czjzek, Gurvan Michel. Innovating glycoside hydrolase activity on a same structural scaffold. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114 (19), pp.4857-4859. 10.1073/pnas.1704802114 . hal-03758586

HAL Id: hal-03758586

<https://hal.science/hal-03758586>

Submitted on 6 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Innovating glycoside hydrolase activity on a same structural scaffold

^{a,b}Mirjam Czjzek¹ and ^{a,b}Gurvan Michel

Carbohydrates play many fundamental roles in the cell physiology and development of plants, animals and microbes. They can take the form of glycoproteins, glycolipids and polysaccharides and represent the largest reservoir of carbon resources that are fuelling microbial communities as well as free-living microorganisms. The structural diversity of naturally occurring carbohydrate compounds is matched by an equally diverse class of enzymes, tailored to breakdown specifically each and every glycosidic bond. The enzymes responsible for the hydrolysis of these glycosides are termed glycoside hydrolases, or glycosidases, and are grouped and classified in sequence-based families within the 'carbohydrate active enzymes' (CAZy) database (1). In general they are very efficient catalysts that can enhance the reaction rate by the order of 10^{17} -fold over the non catalytic reaction rate (2). The mechanism by which they achieve this feat has thus been subject of extensive work (reviewed in ref 3 & 4). The outcome of the hydrolytic reaction, in most cases involving a couple of carboxylic acid residues positioned to activate a water molecule that is added to the glycosidic bond, occurs either with net retention or inversion of the stereochemistry at the anomeric center, and the canonical mechanisms explaining these were described as early as 1953, in a seminal paper by Koshland (5). However, in recent years the flow of new (bacterial) sequences and families has provided several new and unusual mechanisms (6), highlighting the variability of these catabolic processes but also the knowledge gap that still exists in the field of glyco-biochemistry. In particular, the recent advent of systems encoded by Polysaccharide Utilization Loci (PULs) in *Bacteroidetes* has revealed an elegant way of discovering new families and functions (7). It appears that variations of the classical reaction mechanisms depend on the nature of the substrate and are in general observed for particularly recalcitrant substrates, substrate assisted reactions or stereochemically challenging reactions, such as those involving mannose or rhamnose (since the position of the hydroxyl group next to the anomeric carbon is axial). In PNAS, Munoz-Munoz and coworkers add an astonishing exception to the list of non-canonical glycoside hydrolases (8). They describe the discovery of a new glycoside hydrolase (GH) family that, surprisingly, in addition possesses a very uncommon catalytic apparatus. Even more intriguingly, the catalytic active site is not, as usual, located in the center of the deep

pocket formed on the anterior side of the beta-propeller fold, but instead is found on the 'back-side' of the enzyme.

Seeking for enzymes in *Bacteroides thetaiotaomicron* that are able to degrade rhamnose-containing complex carbohydrates, such as arabino-galactan proteins (AGPs) or rhamnogalacturonans I and II (RGI and RGII), the authors inspected the up-regulation of specific PULs in the presence of these substrates, and screened those proteins for activities that had no annotated functions. Doing this, they discovered that BT3686 released α -L-rhamnose units from the AGP substrate named Gum Arabic (GA). As part of the thorough biochemical characterization, Munoz-Munoz et al. determined the 3D crystal structure of the newly discovered enzyme followed by site directed mutagenesis, with the aim to elucidate the structural determinants responsible of the observed glycoside hydrolase activity and α -L-rhamnose specificity. Unexpectedly, activity was not abolished when deleting the side-chains of highly conserved residues within the anterior pocket, formed by the seven-bladed beta-propeller fold, suggesting that this was not the location of the active site. The authors thus adopted an alternative strategy, taking advantage of the crystal structure of the enzyme, by co-crystallizing the enzyme with substrate and product molecules. Fortunately this approach was successful with D-glucuronic acid, one of the monosaccharide products released by the hydrolytic reaction. But instead of being located in the usually conserved active site pocket on the anterior side of the beta-propeller, the monosaccharide was bound in a shallow cleft on the posterior side of the enzyme. By site-directed mutagenesis, Munoz-Munoz and coworkers confirmed that this posterior cleft was indeed the active site of this new family of glycoside hydrolases. This result is outstanding, since in the majority of cases, the positions of active sites are conserved within a given fold (9,10). Some folds, such as the TIM barrel or the Rossmann fold, are more susceptible to functional innovations because their catalytic machinery is usually located on loops loosely connected to the more structured protein core (11,12). But here the posterior active site is located in a well-structured region.

The second major surprise is that the catalytic machinery is not constituted by the canonical pair of carboxylic acid residues, but is composed of a single histidine residue (His48). A catalytic histidine remains a very uncommon catalytic residue for glycoside hydrolases and has previously been suggested to play a role in members of family GH117 α -1,3-(3,6-anhydro)-L-galactosidases (13,14), where it replaces the catalytic acid-base that protonates

the oxygen of the glycosidic linkage. In the mechanism proposed by Munoz-Munoz and coworkers (8), however, the histidine residue is the only catalytic residue and abstracts the proton at the O2 position, implicating that an epoxy-intermediate is formed. This reaction intermediate, to date, is rare in carbohydrates and has been proposed to occur also in family GH99 α -mannosidases (15). The mechanism of these enzymes is dependent of the stereochemical configuration of the substrate molecules that have in common the axial position of the hydroxyl-group next to the anomeric carbon, opposite to the glycosidic bond that is cleaved in the catalytic reaction (15,8). Other unique and rare catalytic reaction mechanisms of glycoside hydrolases that depend on the nature of the substrate have also been observed in sialidases, where a conserved tyrosine plays a crucial role in the nucleophilic attack of the anomeric carbon (16).

But we are not at an end with surprises provided by this work. Whereas the catalytic residues are normally conserved within GH families, the catalytic histidine is absent in ~30% of BT3686 homologues! Three of such homologues were shown to be indeed proteins without obvious enzymatic activities. Spectacularly, the introduction of a histidine at the equivalent position in these inactive homologues restored an α -L-rhamnosidase activity, demonstrating that the function of substrate recognition is intact in these proteins. Finally, phylogenetic analyses support that the α -L-rhamnosidase activity was unlikely the ancestral function of this new GH family and rather resulted from a secondary, recent innovation that specifically occurred in the *Bacteroidetes* phylum. A seducing scenario is that the common ancestor of this protein family had an active site located in the anterior side as observed in the vast majority of the β -propeller fold proteins and that this classical catalytic machinery was lost early in the evolution of this family. The relaxed constraints of starting from a non-enzymatic structural scaffold has probably facilitated the emergence of a new active site location in β -propeller.

The work by Munoz-Munoz et al. thus represents an interesting example illustrating how recent evolutionary events may tailor new activities onto a given fold; the work also provides a lesson about questioning dogmas and rules, since discovery of exceptions are potentially missed when strictly following the ways paved by these rules.

^a CNRS, France; and ^b Station Biologique, France. ¹To whom correspondence should be addressed. Email: mirjam.czjzek@sb-roscoff.fr. Author contributions: M.C. and G.M. wrote the paper.

1. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42, D490-495.
2. Wolfenden R. (2011) Benchmark reaction rates, the stability of biological molecules in water, and the evolution of catalytic power in enzymes. *Annu. Rev. Biochem.*, 80, 645–667.
3. Zechel D, Withers S. (2000) Glycosidase mechanisms: Anatomy of a finely tuned catalyst. *Acc. Chem. Res.*, 33, 11–18.
4. Vocadlo DJ, Davies GJ. (2008) Mechanistic insights into glycosidase chemistry. *Curr. Opin. Chem. Biol.*, 12, 539–555.
5. Koshland DE. (1953) Stereochemistry and the mechanism of enzymatic reactions. *Biol. Rev.*, 28, 416–436.
6. Jongkees SAK, Withers SG. (2013) Unusual Enzymatic Glycoside Cleavage Mechanisms. *Acc. Chem. Res.*, 47, 226–235.
7. Ndeh D, Rogowski A, Cartmell A, Luis AS, Baslé A, Gray J, Venditto I, Briggs J, Zhang X, Labourel A, Terrapon N, Buffetto F, Nepogodiev S, Xiao Y, Field RA, Zhu Y, O'Neill MA, Urbanowicz BR, York WS, Davies GJ, Abbott DW, Ralet MC, Martens EC, Henrissat B, Gilbert HJ. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature*. 2017 Mar 22. doi: 10.1038/nature21725
8. Munoz-Munoz J, Cartmell A, Terrapon N, Henrissat B, Gilbert HJ, PNAS (2017) this issue
9. Arcus V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol.*, 12, 794-801.
10. Theobald DL, Wuttke DS. (2005) Divergent Evolution Within Protein Superfolds Inferred from Profile-based Phylogenetics. *J. Mol. Biol.* 345, 722-737.
11. Dellus-Gur E, Toth-Petroczy A, Elias M, Tawfik DS (2013) What makes a protein fold amenable to functional innovation? Fold polarity and stability trade-offs. *J. Mol. Biol.* 425, 2609-2621.
12. Furnham N, Dawson NL, Rahman SA, Thornton JM, Orengo CA (2016) Large-Scale analysis exploring evolution of catalytic machineries and mechanisms in enzyme superfamilies. *J. Mol. Biol.* 428, 253-267.
13. Rebuffet E, Groisillier A, Thompson A, Jeudy A, Barbeyron T, Czjzek M, Michel G. (2011) Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ Microbiol.* 13, 1253-70.

14. Hehemann JH, Smyth L, Yadav A, Vocadlo DJ, Boraston AB (2012) Analysis of keystone enzyme in Agar hydrolysis provides insight into the degradation (of a polysaccharide from) red seaweeds. *J Biol Chem.* 287(17), 13985-13995.
15. Thompson AJ, Williams RJ, Hakki Z, Alonzi DS, Wennekes T, Gloster TM, Songsrirote K, Thomas-Oates JE, Wrodnigg TM, Spreitz J, Stütz AE, Butters TD, Williams SJ, Davies GJ. (2011) Structural and mechanistic insight into N-glycan processing by endo- α -mannosidase. *Proc Natl Acad Sci U S A.* 109(3), 781-6.
16. Xu, G.; Kiefel, M. J.; Wilson, J. C.; Andrew, P. W.; Oggioni, M. R.; Taylor, G. L. Three *Streptococcus pneumoniae* sialidases: Three different products. *J. Am. Chem. Soc.* 2011, 133, 1718–1721.

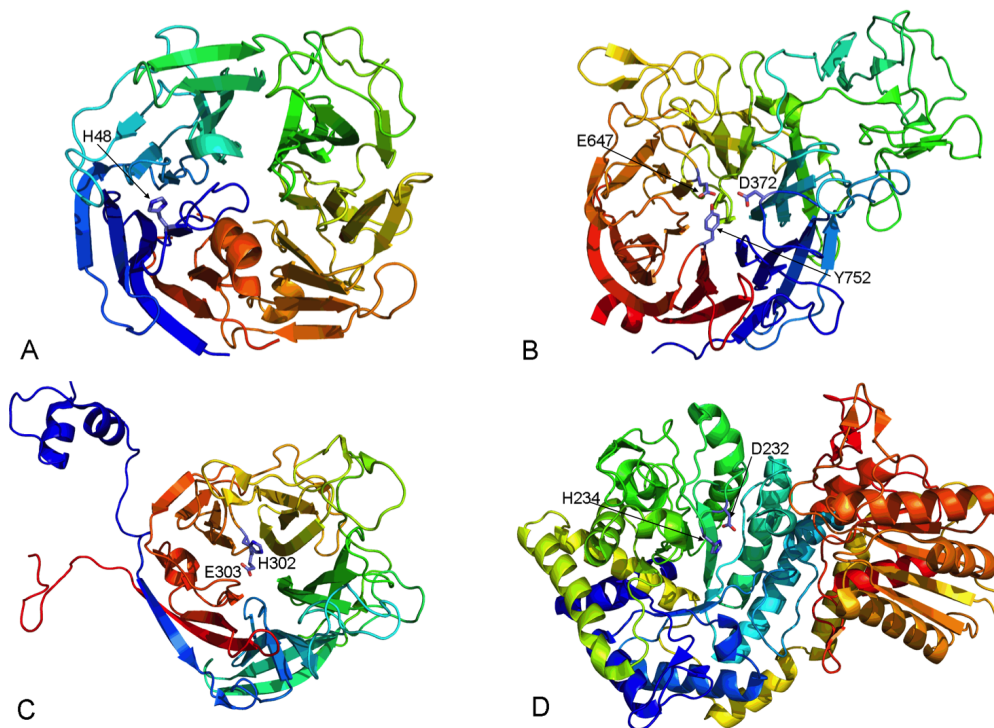


Figure legend

Ribbon representations of selected glycoside hydrolase families with non-canonical catalytic residues. In each panel, the catalytic residues are labeled. A. The α -L-rhamnosidase from the new glycoside hydrolase family described in this issue by Munoz-Munoz et al. (8 ; PDB code 4IRT). B. The sialidase NanA from family GH33 (PDB code 2VVZ). C. The α -1,3-(3,6-anhydro)-L-galactosidase from family GH117 (PDB code 4AK5). D. The β -N-acetylglucosaminidase from family GH3 (PDB code 3BMX).