COMMENTARY

Innovating glycoside hydrolase activity on a same structural scaffold

Mirjam Czjzek[a,b,1] and Gurvan Michel[a,b]

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 fueling 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 specifically to break down each and every glycosidic bond. The enzymes responsible for the hydrolysis of these glycosides are termed glycoside hydrolases (GHs), or glycosidases, and are grouped and classified in sequence-based families within the "Carbohydrate-Active enZYmes" Database (1). In general, they are very efficient catalysts that can enhance the reaction rate by an order of 10^{17}-fold over the noncatalytic reaction rate (2). The mechanism by which they achieve this feat has thus been the subject of extensive work (reviewed in refs. 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 actions 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 glycobiochemistry. 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 generally observed for particularly recalcitrant substrates, substrate-assisted reactions, or stereochemically challenging reactions, such as those reactions involving mannose or rhamnose (because the position of the hydroxyl group next to the anomeric carbon is axial). In PNAS, Munoz-Munoz et al. (8) add an astonishing exception to the list of noncanonical GHs. They describe the discovery of a new 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 is found instead on the "backside" 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. In so doing, they discovered that BT3686 released α-L-rhamnose units from the AGP substrate called gum arabic. As part of the thorough biochemical characterization, Munoz-Munoz et al. (8) determined the 3D crystal structure of the newly discovered enzyme followed by site-directed mutagenesis, with the aim of elucidating the structural determinants responsible for the observed GH 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 cleft 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 cocrystallizing 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. Instead of being located in the usually conserved active site pocket on the anterior side of the beta-propeller, however, the monosaccharide was bound in a shallow cleft on the posterior side of the enzyme. By site-directed mutagenesis, Munoz-Munoz et al. (8) confirmed that this posterior cleft was indeed the active site of this new family of GHs. This result is outstanding, because in the majority of cases, the positions of active sites are conserved within a given fold (9, 10). Some folds, such as the triosephosphateisomerase (TIM) barrel or the

[a]Laboratory for Integrative Biology of Marine Models, Station Biologique, Sorbonne University, Université Pierre et Marie Curie, 29688 Roscoff, France; and [b]Laboratory for Integrative Biology of Marine Models, UMR8227, CNRS, 29688 Roscoff, France

Author contributions: M.C. and G.M. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 4936.

[1]To whom correspondence should be addressed. Email: mirjam.czjzek@sb-roscoff.fr.

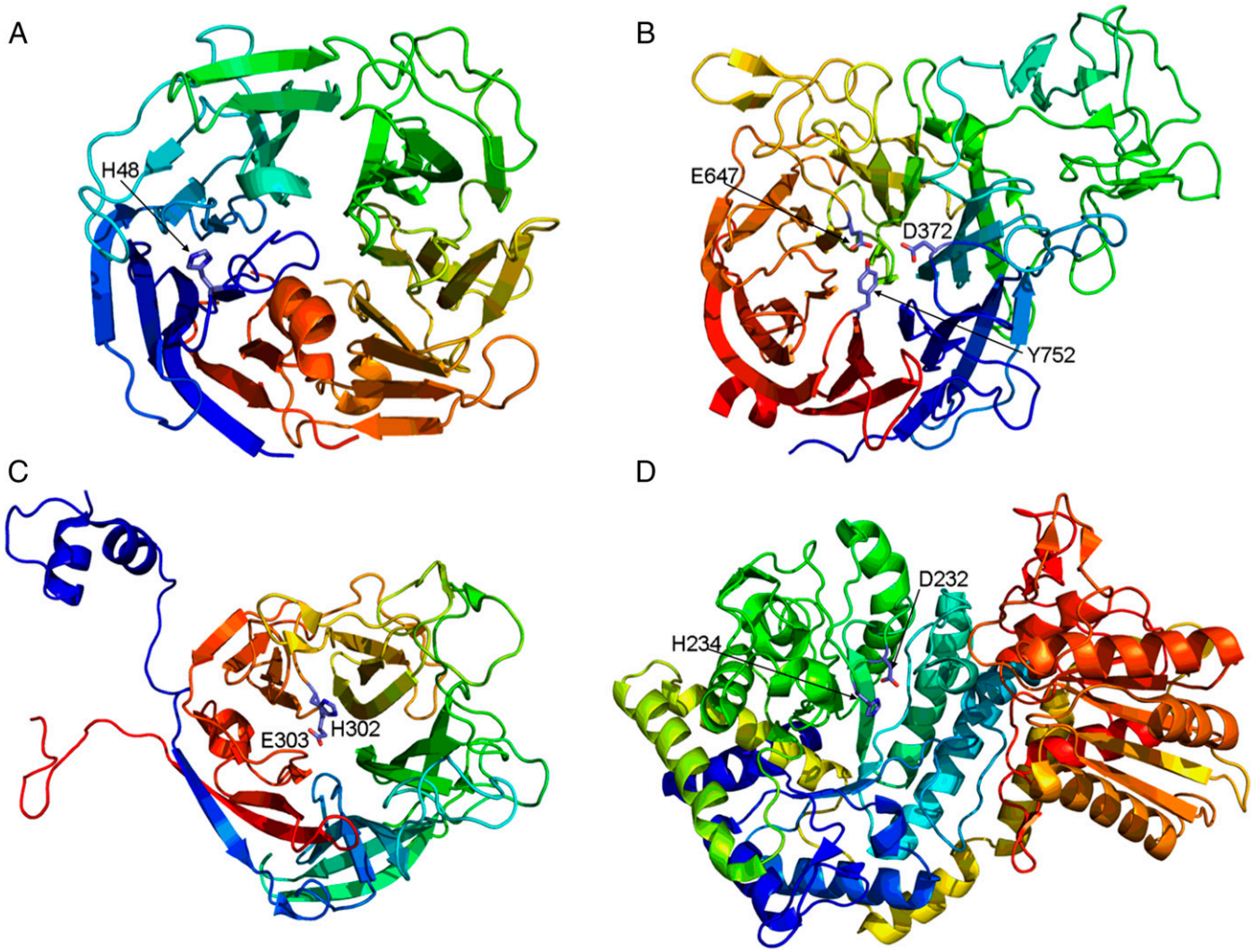

Fig. 1. Ribbon representations of selected GH families with noncanonical catalytic residues. In each panel, the catalytic residues are labeled. (*A*) The α-L-rhamnosidase from the new GH family described by Munoz-Munoz et al. (8) [Protein Data Bank (PDB) ID code 4IRT]. (*B*) Sialidase NanA from family GH33 (PDB ID code 2VVZ). (*C*) The α-1,3-(3,6-anhydro)-L-galactosidase from family GH117 (PDB ID code 4AK5). (*D*) The β-*N*-acetylglucosaminidase from family GH3 (PDB ID code 3BMX).

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). However, 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; Fig. 1). A catalytic histidine remains a very uncommon catalytic residue for GHs 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 et al. (8), however, the histidine residue is the only catalytic residue and abstracts the proton at the O2 position, implying 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 on 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 (8, 15). Other unique and rare catalytic reaction mechanisms of GHs 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).

However, there are yet more surprises provided by this work. Whereas the catalytic residues are normally conserved within GH families, the catalytic histidine is absent in ∼30% of BT3686 homologs. Three such homologs were indeed shown to be proteins without obvious enzymatic activities. Spectacularly, the introduction of a histidine at the equivalent position in these inactive homologs restored 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, instead, resulted from a secondary, recent innovation that specifically occurred in the *Bacteroidetes* phylum. A plausible scenario is that the common ancestor of this protein family had an active site located on 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 nonenzymatic structural

scaffold have probably facilitated the emergence of a new active site location in the β-propeller.

The work by Munoz-Munoz et al. (8) 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, because discovery of exceptions is potentially missed when strictly following the ways paved by these rules.

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–D495.
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 DL, Withers SG (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 Camb Philos Soc* 28:416–436.
6 Jongkees SAK, Withers SG (2014) Unusual enzymatic glycoside cleavage mechanisms. *Acc Chem Res* 47:226–235.
7 Ndeh D, et al. (2017) Complex pectin metabolism by gut bacteria reveals novel catalytic functions. *Nature* 2017:22.
8 Munoz-Munoz J, Cartmell A, Terrapon N, Henrissat B, Gilbert HJ (2017) Unusual active site location and catalytic apparatus in a glycoside hydrolase family. *Proc Natl Acad Sci USA* 114:4936–4941.
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* 354:722–737.
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, et al. (2011) Discovery and structural characterization of a novel glycosidase family of marine origin. *Environ Microbiol* 13:1253–1270.
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* 287:13985–13995.
15 Thompson AJ, et al. (2012) Structural and mechanistic insight into N-glycan processing by endo-α-mannosidase. *Proc Natl Acad Sci USA* 109:781–786.
16 Xu G, et al. (2011) Three Streptococcus pneumoniae sialidases: Three different products. *J Am Chem Soc* 133:1718–1721.