A special issue of Essays in Biochemistry on current advances about CAZymes and their impact and key role in human health and environment

Mirjam Czjzek, Elizabeth Ficko-Blean, Jean-Guy Berrin

Editorial for Essays in Biochemistry

A special issue of *Essays in Biochemistry* on current advances about CAZymes and their impact and key role in human health and environment

Mirjam Czjzek[1‡], Elizabeth Ficko-Blean[1] and Jean-Guy Berrin[2]

[1]Sorbonne Université, CNRS, Integrative Biology of Marine Models (LBI2M), Station Biologique de Roscoff (SBR), 29680 Roscoff, Bretagne, France

[2]INRAE, Aix Marseille Université, UMR1163 Biodiversité et Biotechnologie Fongiques, 13009 Marseille, France.

‡Correspondence: Mirjam Czjzek (czjzek@sb-roscoff.fr)

Abstract

Carbohydrate active enzymes (CAZymes) and their biochemical characterization have been the subject of extensive research over the past ten years due to their importance to carbohydrate metabolism in different biological contexts. For instance, the understanding that "polysaccharide utilizing loci" (PUL) systems hosted by specific "carbohydrate degraders" in the intestinal microbiota play key roles in health and disease, such as Crohn's disease, ulcerative colitis or colorectal cancer to name the most well-characterized [1], has led to an outstanding effort in trying to decipher the molecular mechanisms by which these processes are organized and regulated.
The past ten years has also seen the expansion of CAZymes with auxiliary activities, such as lytic polysaccharide monooxygenases (LPMOs) or even sulfatases, and interest has grown in general about the enzymes needed to remove the numerous decorations and modifications of complex biomass, such as carbohydrate esterases (CE). Today, the characterization of these "modifying" enzymes allows us to tackle a much more complex biomass, which presents sulfations, methylations, acetylations or interconnections with lignin.
This special issue about CAZyme biochemistry covers all these aspects, ranging from implications in disease to environmental and biotechnological impact, with a varied collection of twenty-four review articles providing current biochemical, structural and mechanistic insights into their respective topics.

Article

With the discovery of "polysaccharide utilizing loci" (PULs) in Bacteroidota [2] and their importance for carbohydrate degradation in gut microbiota [3], the field of CAZymes (www.cazy.org) [4] and their biochemistry has received a significant/important boost. Indeed, these operon-like gene clusters, in general contiguous and co-regulated, code for entire and complex carbohydrate degrading systems that can contain up to roughly 30 genes, among them the hydrolytic CAZymes, that are needed to depolymerize, uptake and catabolize a given carbohydrate substrate. Some Bacteroidota have almost 20% of their genome content dedicated to carbohydrate catabolism and contain up to 100 PULs. Astonishingly, these PULs contained a significant number of genes annotated as 'hypothetical proteins' or 'proteins of unknown function,' the characterization of which has led to extensive discovery and expansion of CAZyme families with sometimes novel activities.
At the same time, it has become increasingly evident that, via their dietary fiber catabolizing activity, Bacteroidota play a key role in influencing the composition of the human gut microbiome (HGM). Variations in diet and metabolic status can contribute to impacting changes in the gut microbiota,

deregulating the orchestrated interactions of the intestinal microbiota with host immunity [5], and as a consequence leading to inflammatory diseases, ranging from metabolic, to neurodegenerative diseases up to cancer [1, 6]. Understanding these connections at the level of the "holobiont" thus also passes by understanding the key molecular actors, namely CAZymes.
One recent aspect of these discoveries concerns the link between mucin degradation by bacteria present in the intestinal microbiome and human health and disease. In this context, the review article by Labourel et al. [7] discusses current technologies that aim at boosting the discovery of mucin-degrading CAZymes using microfluidics combined with metagenomics and culturomics, while the review presented by Raba & Luis [8] summarizes the actual knowledge about the biochemistry of CAZymes active on mucin capping and on mucin O-glycan backbones. While this knowledge starts to put together the picture of the enzymes needed for the breakdown of highly O-glycosylated mucins, questions remain about the interplay and regulation of the different bacteria acting together on this complex substrate, and how and when these processes give rise to disease. To tackle diseases that are related to CAZymes it is necessary to develop potential inhibitors, implying understanding the fine details of enzymatic mechanisms, as well as intermediates and transition states of substrate molecules. Molecular modeling, classical methods or enhanced sampling techniques, attempts to catch these complex pictures of the reaction pathway, and the review by Nin-Hill et al. [9] exemplifies modelling strategies to explore disease related enzymes and their reaction pathways.

Another class of components that can represent an important substrate to members of the human gut microbiome (HGM) are N-glycans. Besides being relevant in humans, N-glycans from dietary sources such as those derived from other animals, plants, insects, and fungi can be relevant resources for our gut microbiota. The review by Crouch [10] explores the activities and specificities of N-glycan active CAZymes, many of which originate from bacterial members of the HGM, but also including some eukaryotic systems, highlighting their importance within a healthy equilibrium of the human gut microbiota. Alginate is frequently used as food stabilizer and is equally a relevant compound for which some organisms of the HGM have acquired enzymatic tools to utilize this marine environmental substrate; the review by Rønne et al. [11] provides insight into structure and function of alginate degrading and metabolizing enzymes from the HGM. Fucose is an important component of mucin but also other animal, microbial and plant glycans. The review by Wu & Juge [12] discusses the specific case of exo-fucosidases, giving an overview of the enzymatic and structural properties of α-L-fucosidases and insights into their biological function and biotechnological applications.

But Bacteroidota do not act alone! Another important bacterial phylum present in the HGM is the one of Firmicutes, which are also recognized for being strongly associated with a healthy immune-homeostasis and protection from inflammatory disorders, and in particular those producing butyrate. This molecule, which is considered to be the reason of the benefic effect, is the terminal electron sink of glycan fermentation by prevalent and abundant colonic Firmicutes from the *Lachnospiraceae* and *Oscillospiraceae* families. The review by Leth et al. [13] discusses recent findings on the strategies that colonic butyrate producers have evolved to harvest energy from major dietary fibers but also from human milk oligosaccharides (HMOs). These findings rely on an unexpected discovery of a conserved protein apparatus that confers the growth of butyrate producers on HMOs, which are unique to mother's milk, attesting to the adaptation of this group to both infant and adult guts.

The PUL systems have also led to the characterization of enzymes that have not classically been considered CAZymes as they have enzyme activities that reflect non-carbohydrate components on

the sugars, such as sulfate groups (sulfatlas.sb-roscoff.fr/sulfatlas) [14]. For some bacterial members of the HGM, degradation of mucin requires the action of appropriate carbohydrate sulfatases to access and use O-glycan sugars. In line with this, the review article by Luis et al. [15] treats the special case of mucin-active sulfatases, describing the molecular features that govern substrate recognition and specificity. La Rosa et al. [16] focuses on carbohydrate esterases, from Firmicutes, Bacteroidoita and other microbes found in the HGM. The carbohydrate esterases specifically target O-acetylations on food ingredients, from plant-polysaccharides but also from bacterial exopolysaccharides used as stabilizers or from microbes directly used in food and beverage production.

Plant pathogens also rely on sophisticated molecular strategies to overcome the physical barrier through the secretion of CAZymes targeting the plant cell wall. In this context, the current knowledge and specific role of CAZymes in *Xanthomonas* that connect their function with pathogenicity and lifestyle is reviewed by Guiseppe et al [17]. In mycorrhizal fungi, the modification of plant cell wall polysaccharides is much more subtle to establish mutually beneficial relationships with terrestrial plants. The review by Gong et al [18] describes the key role played by specific plant cell wall active CAZymes to assure symbiotic and healthy interaction with plants. Puchart & Biely [19] give an overview of microbial esterases from saprotrophic microorganisms acting in concert with glycanases for the degradation of the major plant hemicellulose, xylan. A complimentary article by Larsbrink & Lo Leggio [20] focuses on the structure-function relationships of bacterial and fungal CE15 glucuronoyl esterases that specifically target bonds between lignin and carbohydrates found in the plant cell wall. Finally, microbial extracellular glycans also represent an important source of carbon, within a biofilm for example, and some microorganisms, including bacteria (e.g., *Flavobacterium johnsoniae, Mucilaginibacter mallensis, Microbacterium dextranolyticum,* etc.) and fungi (e.g., *Streptomyces* sp. and *Trichoderma viride)*, have evolved dedicated degradation systems specific to these different glycan structures that are not encountered in plants or animals. The review by Miyasaki [21] depicts the recent progress in microbial α-glucan degrading enzymes, among which many have recently been discovered.

The importance of CAZymes to the circular bioeconomy is emphasized in the reviews by Khamassi & Dumon [22] and Ostby & Varnai [23]. While Khamassi & Dumon [22] depict the importance of mastering enzyme synergies for cellulose and hemicelluloses deconstruction, Ostby & Varnai [23] provide a general overview of CAZymes required to upgrade hemicellulose processing to a broad range of value-added products. In this biorefinery context, the discovery of lytic polysaccharide monooxygenases (LPMOs) has revolutionized the understanding of recalcitrant substrate degradation and several reviews of the collection treat various aspects of their mechanisms and interplay with GHs. Sorlie et al [24] review the literature on the interplay between LPMOs and GHs toward improvement of chitin or cellulose degradation. The presence of carbohydrate-binding modules (CBMs), other domains and linker regions is known to have a great impact on the action of LPMOs. Forsberg & Curtade [25] summarize recent literature with a particular focus on comparative LPMO truncation studies. The oxidative mechanism of these copper enzymes and the importance and nature of electron donors have been the subject of a controversial debate. In particular, the question whether LPMOs act as peroxidases rather than monooxygenases, involving the production of highly reactive hydroxyl-radical species is presented in the review by Bissaro & Eijsink [26]. In light of this new paradigm, Hemsworth [27] revisits the importance and role of electron donors in the LPMO reaction. The essay by Fong & Brumer [28] enlarges the discussion about other redox enzymes, specifically fungal AA5 copper radical oxidases (CROs). The recent and significant increase in

biochemically and structurally characterized CROs has expanded the diversity of activities in the AA5 family widening their potential as “green” oxidative biocatalysts in biotechnological applications.

Since the very early descriptions of cellulases [30], the different types of multimodularity of CAZymes has intrigued. This feature, used by a large variety of enzymes, has extensively been explored but still has not revealed all its secrets of why and how multimodularity often leads to enhanced enzyme efficiency. In their contribution, Cedillo & Montanier [31] depict the state-of-the-art of knowledge and discuss the different hypotheses attempting to explain the synergistic effect of multimodularity. Mannuronan-C5-epimerases are another class of often multimodular enzymes, active on the marine brown algal polysaccharide alginate, but yet not officially classified as CAZymes. The structural resemblance of their core catalytic domain with some the fold of some polysaccharide lyases supposes an evolutionary link. Additional domains, called R-domains in the bacterial mannuronan-C5-epimerases, still hamper the understanding of their mechanism at the molecular level. Petersen et al. [29] review the current state of knowledge about the mechanism and potential use of these complex multimodular enzymes.

The final review among this collection, presented by Zhang et al [32], focuses on a peculiar family of glycosyl-transferases, namely the GT47s, important to heparan sulfate biosynthesis in animals and mannan, pectin, xylan and xyloglucan biosynthesis in plants. Though the GT47s share significant sequence similarity, they have evolved members with a great diversity in both donor and acceptor substrate specificity, even for those enzymes that are classified in the same phylogenetic clade.

Taken together, the various topics addressed in this special issue on the state-of-the-art knowledge of CAZymes show that the importance of these classes of enzymes is still growing. Main open questions in the field that remain concern the detailed interplay of these molecular (CAZyme) players, for example, to understand the link between inflammation, nutrition, CAZymes and human disease. Other challenges concern the efficient transformation of complex biomass, either for understanding the environmental impact on the carbon-cycle or for the purpose of biotechnological applications; these will require **integrative studies** to tackle the question simultaneously at different scales (from the ecological to cellular, tissue and to the molecular) and levels of biological processes. We believe that the concerted effort of the strongly interconnected “CAZy community” will be a major asset in the future, and we hope that you enjoy this special issue of *Essays in Biochemistry*!

1. Juge N. Relationship between mucosa-associated gut microbiota and human diseases. *Biochemical Society Transactions* **2022** 50 1225–1236. doi.org/10.1042/BST20201201

2. Anderson K L, Salyers A A. Biochemical evidence that starch breakdown by Bacteroides thetaiotaomicron involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J Bacteriol*. **1989** Jun;171(6):3192-8. doi: 10.1128/jb.171.6.3192-3198.1989.

3. Xu J, Bjursell MK, Himrod J, Deng S, Carmichael LK, Chiang HC, Hooper LV, Gordon JI. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science*. **2003** Mar 28;299(5615):2074-6. doi: 10.1126/science.1080029.PMID: 12663928.

4. Drula E, Garron M L, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* **2022** 50: D571–D577.

5. Zheng D, Liwinski T, Elinav E. Interaction between microbiota and immunity in health and disease. *Cell Res.* **2020**, 30(6):492–506. doi: 10.1038/s41422-020-0332-7.

6. Ziegler S, Bereswill S, Heimesaat MM. Modulation of the intestinal microbiota impacts the efficacy of immunotherapy in cancer patients - A recent literature survey. *Eur J Microbiol Immunol* (Bp). **2022** 12(3):63–72. doi: 10.1556/1886.2022.00017.

7. Labourel A et al. O-mucins degrading CAZymes and their possible implication in inflammatory bowel diseases

8. Raba G & Luis A.S. Mucin utilization by gut microbiota - recent advances on characterization of key enzymes

9. Nin-Hill A et al. *In-silico* modeling of the function of disease-related CAZymes

10. Crouch L I. N-glycan degradation by bacterial CAZymes

11. Rønne M E et al. Gut bacterial alginate degrading enzymes

12. Wu H & Juge N. Microbial fucosidases: a mini review

13. Leth M L et al. Butyrate producing colonic clostridia: picky glycan utilization specialists

14. Stam M, Lelièvre P, Hoebeke M, Corre E, Barbeyron T, Michel G. SulfAtlas, the sulfatase database: state of the art and new developments. *Nucleic Acids Res.* **2023** Jan 6;51(D1):D647-D653. doi: 10.1093/nar/gkac977.

15. Luis A S et al. Functions and specificity of bacterial carbohydrate sulfatases targeting host glycans

16. La Rosa S L et al. Carbohydrate esterases involved in deacetylation of food components by the human gut microbiota

17. Giuseppe P O et al. Enzymatic systems for carbohydrate utilization 1 and biosynthesis in Xanthomonas and their role in pathogenesis and lifestyle

18. Gong Y et al. Role of carbohydrate-active enzymes in mycorrhizal symbioses

19. Puchart V & Biely P. Microbial xylanolytic carbohydrate esterases

20. Larsbrink J & Lo Leggio L. Glucuronoyl esterases – enzymes to decouple lignin and carbohydrates and enable better utilization of renewable plant biomass

21. Miyasaki T. Glycoside 1 hydrolases active on microbial exopolysaccharide α-glucans: structures and function

22. Khamassi A & Dumon C. Enzyme synergy for polysaccharide degradation

23. Østby H & Várnai A. Hemicellulolytic enzymes in lignocellulose processing

24. Sorlie M et al. The interplay between lytic polysaccharide monooxygenases and glycoside hydrolases

25. Forsberg Z & Courtade G. On the impact of CBMs in lytic polysaccharide monooxygenases (LPMOs)

26. Bissaro B. & Eijsink V G H. Lytic polysaccharide monooxygenases (LPMOs) – Enzymes for controlled and site-specific Fenton-like chemistry

27. Hemsworth G R. Revisiting the Role of Electron Donors in Lytic Polysaccharide Monooxygenase Biochemistry

28. Fong J K & Brumer H. Copper Radical Oxidases: galactose oxidase, glyoxal oxidase, and beyond!

29. Petersen A B et al. Alginate-active epimerases: tailoring alginate properties

30. Tomme P, Van Tilbeurgh H, Pettersson G, Van Damme J, Vandekerckhove J, Knowles J, Teeri T, Claeyssens M. Studies of the cellulolytic system of Trichoderma reesei QM 9414. Analysis of domain function in two cellobiohydrolases by limited proteolysis. *Eur J Biochem*. **1988** Jan 4;170(3):575-81. doi: 10.1111/j.1432-1033.1988.tb13736.x.

31. Barba-Cedillo V & Montanier C Y. Spatial organization within multi-modular CAZymes

32. Zhang L et al. Glycosyltransferase family 47 (GT47) proteins in plants and animals