



**HAL**  
open science

## Effect of multi-modularity and spatial organization of glycoside hydrolases on catalysis

Víctor Barba-Cedillo, Cédric Y Montanier

► **To cite this version:**

Víctor Barba-Cedillo, Cédric Y Montanier. Effect of multi-modularity and spatial organization of glycoside hydrolases on catalysis. *Essays in Biochemistry*, 2023, 67 (3), pp.629-638. 10.1042/EBC20220167 . hal-04038640

**HAL Id: hal-04038640**

**<https://hal.science/hal-04038640>**

Submitted on 21 Mar 2023

**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.

# Effect of multi-modularity and spatial organization of glycoside hydrolases on catalysis

Víctor Barba-Cedillo and Cédric Y. Montanier

TBI, Université de Toulouse, CNRS, INRAE, INSA, Toulouse, France

Correspondence: Cédric Y. Montanier ([cedric.montanier@insa-toulouse.fr](mailto:cedric.montanier@insa-toulouse.fr))

## Abstract

The wide diversity among the carbohydrate active enzymes (CAZymes) reflects the equally broad versatility in terms of composition and chemical bonds found in the plant cell wall polymers on which they are active. This diversity, is also expressed through the various strategies developed to circumvent the recalcitrance of these substrates to biological degradation. Glycoside hydrolases (GHs) are the most abundant of the CAZymes and are expressed as isolated catalytic modules or in association with carbohydrate binding module (CBM), acting in synergism within complex arrays of enzymes. This multi-modularity can be even more complex. The cellulosome presents a scaffold protein immobilized to the outer membrane of some microorganisms on which enzymes are grafted to prevent their dispersion and increase catalytic synergism. In Polysaccharide-Utilization Loci (PUL), GHs are also distributed across the membranes of some bacteria to coordinate the deconstruction of polysaccharides and the internalization of metabolizable carbohydrates. Although the study and characterization of these enzymatic activities need to take into account the entirety of this complex organization - in particular because of the dynamics involved in it - technical problems limit this study to isolated enzymes. However, these enzymatic complexes also have a spatio-temporal organization, whose still neglected aspect must be considered. In this review, the different levels of multi-modularity that can occur in GHs will be reviewed, from its simplest forms to the most complex. In addition, attempts to characterize or study the effect on catalytic activity of the spatial organization within GHs will be addressed.

## 28 **Introduction**

29           Spatial arrangement of multi-enzyme cascades is very common in cell, especially regarding the metabolic  
30 pathways [1]. In such a case, enzymes are close to each other and facilitate substrate channelling, consisting of  
31 transferring the metabolite product from one enzyme to the next enzyme in order to reduce diffusion in the bulk  
32 aqueous solvent. For this reason, enzymes must be close enough and correctly oriented [2]. In a way, carbohydrate-  
33 active enzymes or CAZymes, because of the diversity of their mode of action and their synergy, must be considered  
34 as multi-enzyme cascades. Carbon is central to the metabolism of microorganisms and some of them, the  
35 lignocellulolytic organisms, have developed a wide range of enzymes to deconstruct the complex plant cell walls  
36 and release carbon molecules contained in carbohydrate-based polymers. Among these enzymes, glycoside  
37 hydrolases (GHs), which are responsible for the hydrolysis of the glycosidic bond between adjacent carbohydrates  
38 or between a carbohydrate and a non-carbohydrate moiety, have been widely studied and numerous examples are  
39 found in the bibliography. In the course of evolution and to counteract the increasing recalcitrance of plant biomass  
40 towards enzymatic deconstruction, microorganisms have developed several strategies. In this review, we will  
41 consider a transient association as an event concerning the spatial proximity and/or the spatial topology between  
42 domains connected to each other through flexible linkers or distinct GHs interacting in a multi-enzyme complex. It  
43 will also concern the transient association of enzyme clusters, in a cascade of events leading to the  
44 depolymerization of plant polysaccharides. All these aspects should be challenged in terms of catalysis efficiency.  
45 The first part of this review will focus on the different aspects of the modularity of GHs, and how this modularity  
46 may impact the organization of these enzymes in relation to each other during polysaccharide degradation. GHs are  
47 often found in multi-modular enzymes in order to bring catalytic and non-catalytic domains close to each other [3].  
48 These enzymes could also be grafted to a megadalton-size complex known as cellulosome, discovered in the 80's  
49 [4] or in the case of what has been named Polysaccharide Utilization Loci (PUL) in the early 2000s [5], the GHs  
50 are localized on both sides of the bacterial periplasm or the cytoplasm, in close vicinity to membrane bound  
51 proteins. And even more recently, clustering of enzymes on complex substrate can also be seen as transient  
52 physical associations [6]. However, instead of acting on small molecules as in the metabolic pathways, GHs have to  
53 deal with oligosaccharides, soluble polymers and even insoluble substrates. During catalysis, products are released  
54 and constitute new substrates with new properties. As a consequence, the enzymes would have to adapt constantly  
55 their spatial distribution as the plant cell complex matrix is degraded. It is therefore not surprising that GHs require

56 plasticity, either because of the presence of linkers or because of the intrinsic flexibility of the cellulosome scaffold,  
57 one of the main reasons of its efficiency [7]. As dynamic events, it is not possible or very difficult, to explore the  
58 spatial geometry and organization of such enzymes at different moments of the catalytic process. Not considering  
59 these dynamic aspects and the multiple interactions between enzymes and a raw natural substrate presents several  
60 drawbacks. One first disadvantage of such an approach has to do with the specific recognition interactions that exist  
61 between the enzymes and their real substrates and that triggers the catalysis. When using artificial and/or isolated  
62 substrates, the conditions for this recognition are not satisfied which can lead to a wrong estimation/understanding  
63 of the enzymatic activity towards a real substrate [8]. The second downside has to do with the accessibility of the  
64 enzymes to their substrates, which is dramatically facilitated using purified polysaccharides as opposed to  
65 polysaccharides that are embedded in the plant cell wall. This is the reason why it seems to be important to  
66 consider also spatial topology and distances between the different domains of individual enzymes to characterize  
67 such enzymatic complexes. Unfortunately, very few works question these different aspects. Thus, the second part  
68 of this review will propose to summarize information available for the different type of GH organizations and  
69 attempts for their characterization.

## 70 **A large diversity of modularity and spatial organization**

71 In a first level of spatial organization, clusters of GHs have been recently described [6,9]. The cellulases  
72 Cel7A and Cel7B from *Trichoderma reesei* can generate molecular proximity working as a collective of individual  
73 dispersed hydrolases. Thus, the enzymes create transient spatio-temporally coordinated and mechanistically  
74 concerted exo- and endo-cellulase activity clusters for efficient interfacial catalysis for the degradation of cellulose  
75 in a multilayer processive way (Fig 1.A). The authors revealed that besides desorption and re-adsorption, enzyme  
76 dispersion involves molecular diffusions as well as direct movement of the cellulases on the cellulose allowing its  
77 deconstruction by “surface ablation” as opposed to a “fibril cutting” tactic develop by the cellulosome (see below),  
78 which maximizes the effect a of local fragmentation [6]. This strategy results in a synergy that makes the exo-  
79 cellulase to act 100 times faster on an organized polymer material than when it acts alone. Most of the GHs present  
80 several modules (Fig 1.B), with one or more catalytic modules [10]. And in order to promote the association of the  
81 catalytic domain and its prolonged contact with the substrate, carbohydrate-binding modules (CBMs) provide a  
82 proximity effect and a targeting function [11]. The later role directs the GHs to accessible regions of the plant that  
83 are either undergoing biological degradation or cell wall rearrangement [8]. However, the presence of a CBM is  
84 not always required since a so-called substrate binding sites (SBSs) located on the catalytic domain of some GHs

85 may play a similar role [12]. Sometimes, the presence of a CBM can differentiate the activity of the catalytic  
86 domain towards a substrate, as exemplified in the glycoside hydrolase family 5 subfamily 8 (GH5\_8) mannanases  
87 [13]. The authors postulated that GH5\_8 mannanases with CBMs target insoluble mannose-based polymers while  
88 monomodular members prefer soluble oligosaccharides. In other cases, the presence of a CBM is essential for  
89 enzyme activity, as recently described for two modular proteins wtsFae1A and wtsFae1B [14]. These proteins are  
90 feruloyl esterases with a CBM family 48 (CBM48) tethered. This CBM48 interacts with the catalytic domain in  
91 such a way as to extend the cleft where the arabinoxylan backbone is housed. In the absence of the CBM, both  
92 enzymes fail to bind arabinoxylan and are also unable to release ferulic acid from it. Multimodularity could also  
93 create dedicated shape, as reported for AbnA, an extracellular GH43  $\alpha$ -L-arabinanase from *Geobacillus*  
94 *stearothermophilus* [15]. The architecture of this enzyme is composed of four domains; the catalytic domain and  
95 two uncharacterized CBMs spaced from each other by a fourth domain seeming to function as a hinge around  
96 which the other domains revolve. Ultimately, the multimodular enzyme is capable of large conformational changes,  
97 where the catalytic domain seems to interact with the CBM, hydrolysing its substrate through a “harpoon”  
98 mechanism. Nonetheless, most of the time, catalytic domain and CBM are linked to each other by a flexible linker  
99 [16]. Its role is not fixed, but some studies investigated the composition, length and position of the linker as  
100 exemplified by in the exoglucanase Cel6B from *Thermobifida fusca*, which has a CBM2 at its *N*-terminus [17].  
101 While the length of the linker connecting the two domains did not impact the catalytic efficiency, the position of  
102 the CBM or even a replacement by another produces a different inferred plasticity in the interaction between the  
103 CBM and the catalytic module.

104 It seems obvious that nature has to make a lot of efforts to optimize the activity of GHs, with or without  
105 CBMs, to make sure that these activities are consistent with the complexity of the targeted substrate. And well  
106 beyond the simple question of the architecture of enzymes into distinct domains, there remains the question of the  
107 more general organization of all enzymes in relation to each other during the hydrolysis of a complex substrate, and  
108 the dynamics of this organization during catalysis.

109 In a higher level of organization, combining several GHs, cellulosomes are dynamic elaborate complexes  
110 produced and generally bounded and exposed on the cell surface by anaerobic Gram-positive bacteria for the  
111 efficient deconstruction of lignocellulosic biomass. Considered as true nanomachines, these cellulosomes are  
112 organised around a multi-modular architecture based on a non-catalytic subunit called scaffoldin, containing a  
113 variable number of cohesin modules separated by intermodular linker regions. Dockering modules, in fusion with a

114 large variety of catalytic domains, are strongly associated to the cohesins (Fig 1.C). The spatial proximity between  
115 the numerous catalytic domains of the cellulosome allows a synergistic effect and reduce the dispersion of soluble  
116 sugars [18]. The way in which the different docking bearing enzymes are assembled with their cohesin  
117 counterparts are of great importance. Far from being purely the result of chance, studies developing cellulosome  
118 chimeras have demonstrated that the distribution of the different catalytic domains is in fact driven by the length of  
119 the inter-cohesin linker segments of the scaffoldin rather than by the cohesin-dockerin interactions [19]. This  
120 observation suggests that the spatial proximity between enzymes is relevant to increase synergy and how it affects  
121 also the overall conformational flexibility. Furthermore, the enzymatic composition and abundance of cellulosomes  
122 depend on the carbon source and its evolution during the deconstruction of plant cell wall polymers, as revealed by  
123 the studies on *Clostridium clariflavum* cellulosomes [20]. The bacterium is able to express 13 different scaffoldins,  
124 resulting in cell-free and cell-bound cellulosomes. Their expression levels are higher when the bacterium is  
125 cultivated in presence of complex substrates. As a consequence, the cellulosomes produced display a higher  
126 amount of scaffoldins compared with those expressed in presence of oligosaccharides, exposing up to 56 different  
127 active GHs on complex substrate. In this context, research on cellulosomes from *Ruminiclostridium cellulolyticum*  
128 has revealed that the conformation of the two domains in dockerins involved in the interaction with cohesins allows  
129 the re-functionalization of the cellulosome, and so the exchange between catalytic domains as an adaptation to  
130 heterogeneous plant cell wall substrates in active degradation [21]. Furthermore, spatial expression and exposition  
131 of cellulosomes on the cell surface varies on substrate used for growth, as exemplified by the cellulosomes  
132 produced by *Clostridium clariflavum* [22]. The results highlight how cellulosome levels displayed on the cell  
133 surface of bacteria cultivated on a soluble substrate are significantly lower than those grown in recalcitrant,  
134 insoluble complex substrates. In the first situation, cellulosomes are organized as restricted “islands” since most of  
135 the produced cellulosome are released into the growth medium. In the second situation, the cellulosome clusters are  
136 more widely distributed on the surface of the bacterium, allowing the intimate hydrolysis of the recalcitrant  
137 substrates and assimilation of the soluble sugar break down products immediately.

138         Again, the internal organization of cellulosomes and their clustering as protuberances on the surface of the  
139 bacterial cell may raise questions about how the spatial organization of the hundreds of enzymes that act together  
140 and/or sequentially when working on an insoluble substrate conditions the end-product.

141         In recent years, another multi-enzymatic organization called PUL (Fig 1.D) have attracted the attention of a  
142 whole part of the glycobiology community. According to PUL DataBase (PULDB, [23]), a PUL is a set of

143 physically-linked genes coding for proteins related to the starch utilization system (Sus), which includes the sugar  
144 transport and capture proteins SusC and SusD. PULs are widespread in the *Bacteroidota* phylum and confer an  
145 evolutionary advantage to these species by orchestrating the degradation of complex glycans through encoded GHs.

146

147 These enzymes are dedicated to a particular glycan [24]. Because of the presence of SusC-like and SusD-like, all  
148 PULs are studied by analogy with the archetypical Sus locus from *Bacteroides thetaiotaomicron* [25], initially  
149 described by Anderson and Salyers in 1989 [26]. The number of PULs within a genome is strongly correlated with  
150 the number of polysaccharides metabolized by a particular specie [27], so some *Bacteroidota* possess over 100  
151 different PULs [28]. PUL encode also for specific sensor/regulator and surface glycan-binding proteins (SGBPs).  
152 Glycan import into the periplasm is mediated by a SusC-like protein (a TonB dependent transporter) that works in  
153 concert with a SusD-like binding lipoprotein [29], creating a protein complex with a “pedal bin”-like mechanism  
154 [30,31]. Genetic studies on deletion of these spatially related components eliminate or reduce PUL function,  
155 supporting its importance for polysaccharide degradation [29]. Currently, new systems similar to PULs are being  
156 discovered in bacteria from different *phylum* including gram positive bacteria [24].

157 The complex arrangement of PUL-encoded proteins is not yet fully elucidated, and although this system is  
158 very different from the cellulosome, we can hypothesize that the set of enzymatic processes that guide the  
159 deconstruction of polysaccharides into metabolized oligosaccharides involves transient associations and spatial  
160 organization between proteins active on sugars.

### 161 **Attempts to characterize or study the effect of spatial organization between individual GHs**

162 As described above, GHs exhibit a large variety of architectures, which has been shown to be important for  
163 catalytic function. The catalytic activity of these GHs on plant cell wall polysaccharides implies a cooperative  
164 mode of action meaning a spatio-temporal organization of a large amount of individual GHs in close vicinity to the  
165 substrate. However, the relationship between the catalytic activity and the spatial organization of the different  
166 modules of GHs is, to our knowledge, rarely the focus of research and very partially addressed. This is also the  
167 reason why we have extended our examples to another type of carbohydrate-acting enzyme whose activity is  
168 complementary to that of GHs.

169 Whether encoded by PULs or not, GHs produced by microorganisms degrading plant biomass are often  
170 composed of several catalytic domains and/or CBMs, connected to each other *via* a linker more or less lengthy. If  
171 the role of the linker plays a fundamental function in the activity of processive cellulases [32], thermal adaptation  
172 of some GHs [33] or promote activity and binding efficiency [34], it could also be of importance in spatial module  
173 organization. The effects of length or composition of the linker on the activity of a multi-modular GH5 cellulase  
174 was explored [35]. The authors designed several linkers connecting GH5 to its native CBM3 targeting cellulose and  
175 assessed the enzymatic activity. In a more original way, they studied the effects of linkers on the molecular  
176 architecture of the GH5 in solution, using small angle X-ray scattering (SAXS), enabling to consider the distance  
177 between the two domains and the total length of the enzyme. The linker length induces molecular rearrangements,  
178 supporting the notion that the inter-modular distance and catalytic efficiency are correlated in non-processive GH5.  
179 For example, a linker twice the length of the wild-type enzyme resulted in the highest catalytic efficiency. In that  
180 case, SAXS analysis revealed an optimal distance of 60 Å between the two domains. The relative position of both  
181 domains is also important as inter-domain distance is critical for the diffusion of substrates. More recently, nuclear  
182 magnetic resonance (NMR) spectroscopy was used to provide a mechanistic basis to investigate the role of the  
183 dynamic generated by a linker connecting two adjacent domains [36]. The subject of the study is a Lytic  
184 Polysaccharide MonoOxygenase (LPMO) from *Streptomyces coelicolor* acting on cellulose. It is composed of the  
185 catalytic domain connected to a CBM2 targeting cellulose by a 30-residue linker region. The authors solved the  
186 structure of the LPMO in solution, providing some evidences on how the linker is disordered and extended,  
187 maintaining the distance between the two domains. As a consequence, the CBM helps the catalytic domain to bind  
188 in the proximity of a previous cut for a prolonged time (Fig 2.A). The flexibility of the linker allows the CBM and  
189 the catalytic domain to move independently of each other. This flexibility promotes also the catalytic domain to  
190 repeatedly oxidize the substrate in a larger surface area. Another study used multiple computational methods to  
191 investigate the properties of disordered linkers in cellulosomal systems and compared with synthetic linkers  
192 displaying differences in length and composition [37]. The authors demonstrated that the length is the main factor,  
193 conditioning the proximity between adjacent domains. Furthermore, short linkers affect the structure of the  
194 domains they are linked to, potentially affecting the efficiency of the catalytic site.

195 Several types of cellulosomal systems are present in nature, directly bound to the cell surface, even with  
196 intermediate adapter scaffolds, or existing in an intrinsically free scaffold. In *Clostridium clariflavum*, the  
197 successive assembly of different protein structures amplifies the number of enzymatic subunits, resulting in a



198 surface-bound cellulosomal complex of 160 enzymes [38]. In such protein complex, the question of the spatial  
199 organization is essential, knowing that this system is very flexible [39], and that this flexibility is essential for the  
200 efficiency of this complex [7]. Works have addressed this issue, mainly using reduced models of cellulosome or  
201 designer cellulosomes [40]. The question whether the position of defined cellulolytic enzymes is important for the  
202 optimized degradation of a cellulosic substrate was assessed by developing 10 chimaeric scaffoldins incorporating  
203 three cellulases [41]. The authors evidenced that the relative position of each enzyme impacts the catalytic  
204 efficiency of the chimaeric scaffoldins, highlighting undoubtedly an optimal spatial organization (Fig 2.B). A more  
205 recent work questioned the spatial proximity effect of a  $\beta$ -glucosidase mixed with cellulosomes [42]. In order to  
206 reduce the inhibition effect of hydrolysate cellobiose restricting the hydrolytic ability of cellulosomes, the authors  
207 designed a chimaeric  $\beta$ -glucosidase fused to a CBM targeting cellulose or to a dockerin type II in order to bind the  
208 enzyme close to the cellulosome. Hence, spatial distance between the cellulosome and the  $\beta$ -glucosidase seems to  
209 be an important factor to increase degradation of the cellulosic substrate, even if the raw and complex substrate's  
210 topology might be also involved in the kinetic. Beyond the synergistic effect of the cellulosomal enzymes [43], the  
211 size and specific activity of cellulosome was also investigated [44]. The authors constructed several recombinant  
212 cellulosomes containing various mixtures of three cellulases from family GH5, 9 and 48. The activity increases  
213 with cellulosome complexity and reaches its maximum when the cellulosome is composed of 9 GHs, far from the  
214 maximum of 25 enzymes carried. The authors postulated that the diversity of enzymatic activities is the bottleneck  
215 on the overall cellulosome activity, instead of the protein proximity.

216 Since it is very challenging to work with a designer cellulosome, even more with a native cellulosome,  
217 several strategies have been used to develop artificial cellulosomes based on the immobilization of enzymes using  
218 various carriers [45–50]. Despite some uncertainty as to the correct orientation or not of the enzymes, the findings  
219 reveal that generally, proximity between enzymes increases the catalytic efficiency in comparison to the free  
220 enzymes, likely reflecting the primary role of the cellulosome in the deconstruction of complex lignocellulosic  
221 substrates. However, some works are more informative about the spatial organization of catalytic domains. A  
222 polymeric micelle grafted with nitrilotriacetic acid moieties on the corona layer was developed to immobilize  
223 cellulases expressed with a His<sub>6</sub>-tag opposite to their catalytic pocket [51]. In this way, the authors controlled the  
224 spatial orientation of immobilized enzymes, the distribution, and synergistic interactions of different types of  
225 cellulases. Using fluorescent proteins as probe and a Fluorescence Resonance Energy Transfer (FRET) experiment  
226 to assess the distance between enzymes, the authors set different ratios of cellulases to be immobilized. The

227 resulting synergistic catalysis was enhanced over two-fold compare to the cellulases free in solution. In the same  
228 vein, a xylanase from family GH11 was immobilized at precisely controllable surface densities on magnetic beads  
229 [52]. The authors were able to orientate the enzyme optimally with an average centre-to-centre distance between  
230 adjacent enzymes from 4.4 nm to 30.0 nm. They demonstrated that grafting density was the only parameter  
231 affecting not only the specific activity but also the product profile released from a polymeric substrate. Thus, high  
232 density grafting preferentially releases shorter oligosaccharides compared to the same amount of enzymes in  
233 solution (Fig 2.C).

234 In the era of exponentially growing numbers of sequenced genomes, PULs and their clusters of physically  
235 linked genes targeting carbohydrate degradation are considered an ideal source of new enzyme discovery [28].  
236 However, beyond the different functions of these proteins, one could postulate that the relatively large amount of  
237 enzymes and transporters encoded by PULs should be spatially organized in a way to efficiently act together.  
238 Studies of the archetypical Sus from *Bacteroides thetaiotaomicron* indicate transient, substrate-induced complex  
239 formation involving SusC/D, two SGBPs and a surface amylase GH13 [53]. In addition, the xyloglucan utilization  
240 locus from *Bacteroides ovatus* contains 12 coding protein genes, including 8 for different GHs. Most of them are  
241 structurally and functionally characterized, making this PUL a suitable model [29]. Among these GHs a cell surface  
242 endo-(xylo)glucanase GH9 may have evolved toward a non-catalytic, substrate-binding function and clearly acts in  
243 synergy with the SusD-like and SGBPs components, facilitating substrate capture and import through the  
244 transporter [54]. The authors suggest an assembly of a discrete multiprotein complex possibly facilitating the  
245 orientation of the xyloglucan to be transferred across the outer membrane. Recently, single-particle cryogenic  
246 electron microscopy (cryo-EM) revealed the organization of the SusCD-like proteins from a levan utilization loci  
247 found in *Bacteroides thetaiotaomicron* [55]. The co-purified complex exhibited the presence of two lipoproteins,  
248 including an endo-levanase GH32, and revealed large and concerted conformational changes during nutrient import  
249 in the cell. It is important to underscore that , in general, PULs encode also proteins of unknown function or  
250 hypothetical proteins (see Polysaccharide-Utilization Loci DataBase, [23]), and not all PUL proteins are involved  
251 in the complexes described above.

252 In any case, there is no doubt that PULs offer the possibility to discover new GHs with an interesting  
253 organization and spatial regulation, strongly favoured by the distribution of these enzymes on both sides of the cell  
254 membrane, interacting with more or less complex substrates.

## 255 **Future perspectives**

256 Although the molecular basis by which the spatial organization of GHs influences enzymatic catalysis is  
257 not fully understood, research in this area may open up new opportunities to improve and develop new  
258 synergistically acting enzymatic systems capable of degrading complex carbohydrates. So far, most of the studies  
259 developed have focused on biochemical aspects that could explain the effect of spatial organization, but few have  
260 really addressed this topic in depth with orthogonal, mainly structural, approaches. There is therefore a need to  
261 advance the field using an orthogonal "integrative structure determination" approach, integrating structural,  
262 biochemical and bioinformatics studies as proposed by Lansky *et al.* [15]. In addition, only a few studies aimed at  
263 inducing spatial proximity using protein scaffolds have been developed [56,57]. In these studies, the authors used a  
264 covalent association process involving two protein fragments to systematically study spatial organizations within  
265 multi-modular GHs, correlating biochemical and structural studies. It is clear that biological systems are complex,  
266 at several scales of observation, and it is important to take into account this complexity and in particular the  
267 transient interactions between GHs involved in the hydrolysis of raw substrates. Thus, single-molecule atomic force  
268 microscopy (AFM) is a very promising way to understand the dynamics of GHs during polymer deconstruction  
269 [58], while *in situ* cryo-electron tomography (Cryo-ET) proved to be a relevant tool to investigate the structure of  
270 cellulosomes in a near-physiological state, and to probe the interactions of the bacterium with cellulose fibres [59].  
271 In addition, the use of nanometre-scale superresolution imaging in real time in live cells can reveal dynamic  
272 interactions between the proteins encoded by PULs [60].

## 273 **Summary**

- 274 • The fundamental unit of GHs is the catalytic domain. But depending on the complexity of the substrate,  
275 protein architecture, in terms of catalytic and non-catalytic domains, can diverge dramatically, from the  
276 simplest to the most sophisticated spatial organization.
- 277 • Due to technical limitations, most of the enzymatic activities of GHs are characterised using non-natural  
278 and/or chromogenic substrates, far away from the intrinsic complexity of plant cell wall raw polymers.
- 279 • Emerging works combine structural studies and enzymatic activities to investigate the relationship between  
280 the spatial organization of GHs and the specificity of their mode of action.

## 281 **Competing Interests**

282 The authors declare that there are no competing interests associated with this manuscript.

### 283 **Funding**

284 This work was supported by a grant to VBC from the European Commission Horizon 2020 Marie Skłodowska-  
285 Curie Action Individual Fellowship Program to develop the project EvoXUL (grant n° 897655).

### 286 **Open Access**

287

### 288 **Acknowledgements**

289

### 290 **Abbreviations**

291 CAZymes, carbohydrate-active enzymes; GHs, glycoside hydrolases; GH5, glycoside hydrolase family 5; LPMO,  
292 lytic polysaccharide monooxygenase; CBM, carbohydrate binding module; CBM2, carbohydrate binding module  
293 family 2; PUL, polysaccharide utilization loci; TBDT, TonB dependent transporter; SGBPs, surface glycan-binding  
294 proteins; SBS, substrate binding site; SAXS, small angle X-ray scattering; NMR, nuclear magnetic resonance;  
295 FRET, fluorescence resonance energy transfer; CryoET, Cryo-electron tomography; AFM, atomic force  
296 microscopy.

### 297 **References**

- 298 1 Quin, M. B., Wallin, K. K., Zhang, G. and Schmidt-Dannert, C. (2017) Spatial organization of multi-enzyme  
299 biocatalytic cascades. *Org. Biomol. Chem.* **15**, 4260–4271, <https://doi.org/10.1039/C7OB00391A>
- 300 2 Lim, S. I., Yang, B., Jung, Y., Cha, J., Cho, J., Choi, E.-S., et al. (2016) Controlled orientation of active sites in  
301 a nanostructured multienzyme complex. *Sci. Rep.* **6**, 39587, <https://doi.org/10.1038/srep39587>
- 302 3 Conway, J. M., Crosby, J. R., Hren, A. P., Southerland, R. T., Lee, L. L., Lunin, V. V., et al. (2018) Novel  
303 multidomain, multifunctional glycoside hydrolases from highly lignocellulolytic *Caldicellulosiruptor* species.  
304 *AIChE J.* **64**, 4218–4228, <https://doi.org/10.1002/aic.16354>
- 305 4 Lamed R, Setter E, and Bayer E A. (1983) Characterization of a cellulose-binding, cellulase-containing  
306 complex in *Clostridium thermocellum*. *J. Bacteriol.* **156**, 828–836, [https://doi.org/10.1128%2Fjb.156.2.828-](https://doi.org/10.1128%2Fjb.156.2.828-836.1983)  
307 [836.1983](https://doi.org/10.1128%2Fjb.156.2.828-836.1983)

- 308 5 Bjursell, M. K., Martens, E. C. and Gordon, J. I. (2006) Functional genomic and metabolic studies of the  
309 adaptations of a prominent adult human gut symbiont, *Bacteroides thetaiotaomicron*, to the suckling period. *J.*  
310 *Biol. Chem.* **281**, 36269–36279, <https://doi.org/10.1074/jbc.m606509200>
- 311 6 Zajki-Zechmeister, K., Eibinger, M. and Nidetzky, B. (2022) Enzyme synergy in transient clusters of endo-  
312 and exocellulase enables a multilayer mode of processive depolymerization of cellulose. *ACS Catal.* **12**,  
313 10984–10994, <https://doi.org/10.1021/acscatal.2c02377>
- 314 7 Himmel, M. E., Ding, S.-Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., et al. (2007) Biomass  
315 recalcitrance: engineering plants and enzymes for biofuels production. *Science* **315**, 804–807,  
316 <https://doi.org/10.1126/science.1137016>
- 317 8 Montanier, C., van Bueren, A. L., Dumon, C., Flint, J. E., Correia, M. A., Prates, J. A., et al. (2009) Evidence  
318 that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proc. Natl.*  
319 *Acad. Sci.* **106**, 3065–3070, <https://doi.org/10.1073/pnas.0808972106>
- 320 9 Zajki-Zechmeister, K., Kaira, G. S., Eibinger, M., Seelich, K. and Nidetzky, B. (2021) Processive enzymes  
321 kept on a leash: how cellulase activity in multienzyme complexes directs nanoscale deconstruction of cellulose.  
322 *ACS Catal.* **11**, 13530–13542, <https://doi.org/10.1021/acscatal.1c03465>
- 323 10 Janeček, Š. and Svensson, B. (2022) How many  $\alpha$ -amylase GH families are there in the CAZy database?  
324 *Amylase* **6**, 1-10. <https://doi.org/10.1515/amylase-2022-0001>
- 325 11 Boraston, A. B., Bolam, D. N., Gilbert, H. J. and Davies, G. J. (2004) Carbohydrate-binding modules: fine-  
326 tuning polysaccharide recognition. *Biochem. J.* **382**, 769–781. <https://doi.org/10.1042%2FBJ20040892>
- 327 12 Cuyvers, S., Dornez, E., Delcour, J. A. and Courtin, C. M. (2012) Occurrence and functional significance of  
328 secondary carbohydrate binding sites in glycoside hydrolases. *Crit. Rev. Biotechnol.*, **32**, 93–107.  
329 <https://doi.org/10.3109/07388551.2011.561537>
- 330 13 Møller, M. S. (2022) Impact of modular architecture on activity of glycoside hydrolase family 5 subfamily 8  
331 mannanases. *Molecules* **27**, 1915, <https://doi.org/10.3390/molecules27061915>
- 332 14 Holck, J., Fredslund, F., Møller, M. S., Brask, J., Krogh, K. B. R. M., Lange, L., et al. (2019) A carbohydrate-  
333 binding family 48 module enables feruloyl esterase action on polymeric arabinoxylan. *J. Biol. Chem.* **294**,  
334 17339–17353, <https://doi.org/10.1074/jbc.ra119.009523>

- 335 15 Lansky, S., Salama, R., Biarnés, X., Shwartstein, O., Schneidman-Duhovny, D., Planas, A., et al. (2022)  
336 Integrative structure determination reveals functional global flexibility for an ultra-multimodular arabinanase.  
337 *Commun. Biol.* **5**, 465, <https://doi.org/10.1038/s42003-022-03054-z>
- 338 16 Czjzek, M. and Ficko-Blean, E. (2017) Probing the complex architecture of multimodular carbohydrate-active  
339 enzymes using a combination of small angle X-Ray scattering and X-Ray crystallography. In *Protein-  
340 Carbohydrate Interactions: Methods and Protocols* (Abbott, D. W., and Lammerts van Bueren, A., eds.), pp  
341 239–253, Springer New York, NY, [https://doi.org/10.1007/978-1-4939-6899-2\\_19](https://doi.org/10.1007/978-1-4939-6899-2_19)
- 342 17 Setter-Lamed, E., Morais, S., Stern, J., Lamed, R. and Bayer, E. A. (2017) Modular organization of the  
343 *Thermobifida fusca* exoglucanase cel6B impacts cellulose hydrolysis and designer cellulosome efficiency.  
344 *Biotechnol. J.* **12**, 1700205, <https://doi.org/10.1002/biot.201700205>
- 345 18 Fontes, C. M. G. A. and Gilbert, H. J. (2010) Cellulosomes: highly efficient nanomachines designed to  
346 deconstruct plant cell wall complex carbohydrates. *Annu. Rev. Biochem.* **79**, 655–681,  
347 <https://doi.org/10.1146/annurev-biochem-091208-085603>
- 348 19 Borne, R., Bayer, E. A., Pagès, S., Perret, S. and Fierobe, H.-P. (2013) Unraveling enzyme discrimination  
349 during cellulosome assembly independent of cohesin–dockerin affinity. *FEBS J.* **280**, 5764–5779,  
350 <https://doi.org/10.1111/febs.12497>
- 351 20 Artzi Lior, Morag Ely, Barak Yoav, Lamed Raphael, and Bayer Edward A. (2015) *Clostridium clariflavum*:  
352 key cellulosome players are revealed by proteomic analysis. *mBio* **6**, e00411-15,  
353 <https://doi.org/10.1128/mBio.00411-15>
- 354 21 Borne, R., Dao Ti, M.-U., Fierobe, H.-P., Vita, N., Tardif, C. and Pagès, S. (2020) Catalytic subunit exchanges  
355 in the cellulosomes produced by *Ruminiclostridium cellulolyticum* suggest unexpected dynamics and  
356 adaptability of their enzymatic composition. *FEBS J.* **287**, 2544–2559, <https://doi.org/10.1111/febs.15155>
- 357 22 Artzi Lior, Dadosh Tali, Milrot Elad, Morais Sarah, Levin-Zaidman Smadar, Morag Ely, et al. (2018)  
358 Colocalization and disposition of cellulosomes in *Clostridium clariflavum* as revealed by correlative  
359 superresolution imaging. *mBio* **9**, e00012-18, <https://doi.org/10.1128/mBio.00012-18>
- 360 23 Terrapon, N., Lombard, V., Drula, É., Lapébie, P., Al-Masaudi, S., Gilbert, H. J. et al. (2018) PULDB: the  
361 expanded database of Polysaccharide Utilization Loci. *Nucleic Acids Res.* **46**, D677–D683,  
362 <https://doi.org/10.1093/nar/gkx1022>

- 363 24 Grondin Julie M., Tamura Kazune, Déjean Guillaume, Abbott D. Wade, and Brumer Harry. (2017)  
364 Polysaccharide utilization loci: fueling microbial communities. *J. Bacteriol.* **199**, e00860-16,  
365 <https://doi.org/10.1128/jb.00860-16>
- 366 25 Foley, M. H., Cockburn, D. W. and Koropatkin, N. M. (2016) The Sus operon: a model system for starch  
367 uptake by the human gut *Bacteroidetes*. *Cell. Mol. Life Sci.* **73**, 2603–2617,  
368 <https://doi.org/10.1007%2Fs00018-016-2242-x>
- 369 26 Anderson K L and Salyers A A. (1989) Biochemical evidence that starch breakdown by *Bacteroides*  
370 *thetaiotaomicron* involves outer membrane starch-binding sites and periplasmic starch-degrading enzymes. *J.*  
371 *Bacteriol.*, ASM **171**, 3192–3198. <https://doi.org/10.1128/jb.171.6.3192-3198.1989>
- 372 27 Martens, E. C., Lowe, E. C., Chiang, H., Pudlo, N. A., Wu, M., McNulty, N. P., et al. (2011) Recognition and  
373 degradation of plant cell wall polysaccharides by two human gut symbionts. *PLOS Biol.* **9**, e1001221,  
374 <https://doi.org/10.1371/journal.pbio.1001221>
- 375 28 Lapébie, P., Lombard, V., Drula, E., Terrapon, N. and Henrissat, B. (2019) *Bacteroidetes* use thousands of  
376 enzyme combinations to break down glycans. *Nat. Commun.* **10**, 2043, [https://doi.org/10.1038/s41467-019-](https://doi.org/10.1038/s41467-019-10068-5)  
377 [10068-5](https://doi.org/10.1038/s41467-019-10068-5)
- 378 29 Larsbrink, J., Rogers, T. E., Hemsworth, G. R., McKee, L. S., Tauzin, A. S., Spadiut, O., et al. (2014) A  
379 discrete genetic locus confers xyloglucan metabolism in select human gut *Bacteroidetes*. *Nature* **506**, 498–502,  
380 <https://doi.org/10.1038/nature12907>
- 381 30 Glenwright, A. J., Pothula, K. R., Bhamidimarri, S. P., Chorev, D. S., Baslé, A., Firbank, S. J., et al. (2017)  
382 Structural basis for nutrient acquisition by dominant members of the human gut microbiota. *Nature* **541**, 407–  
383 411, <https://doi.org/10.1038/nature20828>
- 384 31 Bolam, D. N. and van den Berg, B. (2018) TonB-dependent transport by the gut microbiota: novel aspects of  
385 an old problem. *Eng. Des. New Appl. Membr.* **51**, 35–43, <https://doi.org/10.1016/j.sbi.2018.03.001>
- 386 32 Nakamura, A., Tasaki, T., Ishiwata, D., Yamamoto, M., Okuni, Y., Visootsat, A., et al. (2016) Single-molecule  
387 imaging analysis of binding, processive movement, and dissociation of cellobiohydrolase *Trichoderma reesei*  
388 cel6A and its domains on crystalline cellulose. *J. Biol. Chem.* **291**, 22404–22413,  
389 <https://doi.org/10.1074/jbc.m116.752048>

- 390 33 Wu, X., Shi, Z., Tian, W., Liu, M., Huang, S., Liu, X., et al. (2022) A thermostable and CBM2-linked GH10  
391 xylanase from *Thermobifida fusca* for paper bleaching. *Front. Bioeng. Biotechnol.* **10**,  
392 <https://doi.org/10.3389/fbioe.2022.939550>
- 393 34 Srivastava Aishwarya, Nagar Pragya, Rathore Sumit, and Adlakha Nidhi. (2022) The linker region promotes  
394 activity and binding efficiency of modular LPMO towards polymeric substrate. *Microbiol. Spectr.* **10**, e02697-  
395 21, <https://doi.org/10.1128/spectrum.02697-21>
- 396 35 Ruiz, D. M., Turowski, V. R. and Murakami, M. T. (2016) Effects of the linker region on the structure and  
397 function of modular GH5 cellulases. *Sci. Rep.* **6**, 28504, <https://doi.org/10.1038/srep28504>
- 398 36 Courtade, G., Forsberg, Z., Heggset, E. B., Eijsink, V. G. H. and Aachmann, F. L. (2018) The carbohydrate-  
399 binding module and linker of a modular lytic polysaccharide monooxygenase promote localized cellulose  
400 oxidation. *J. Biol. Chem.* **293**, 13006–13015, <https://doi.org/10.1074/jbc.ra118.004269>
- 401 37 Różycki, B., Cazade, P.-A., O'Mahony, S., Thompson, D. and Cieplak, M. (2017) The length but not the  
402 sequence of peptide linker modules exerts the primary influence on the conformations of protein domains in  
403 cellulosome multi-enzyme complexes. *Phys. Chem. Chem. Phys.* **19**, 21414–21425,  
404 <https://doi.org/10.1039/C7CP04114D>
- 405 38 Artzi, L., Bayer, E. A. and Moraïs, S. (2017) Cellulosomes: bacterial nanomachines for dismantling plant  
406 polysaccharides. *Nat. Rev. Microbiol.* **15**, 83–95, <https://doi.org/10.1038/nrmicro.2016.164>
- 407 39 Dorival, J., Moraïs, S., Labourel, A., Rozycki, B., Cazade, P.-A., Dabin, J., et al. (2022) Mapping the  
408 deformability of natural and designed cellulosomes in solution. *Biotechnol. Biofuels Bioprod.* **15**, 68,  
409 <https://doi.org/10.1186/s13068-022-02165-3>
- 410 40 Vazana, Y., Moraïs, S., Barak, Y., Lamed, R. and Bayer, E. A. (2012) Designer cellulosomes for enhanced  
411 hydrolysis of cellulosic substrates. Chapter twenty-three, In *Methods in Enzymology* (Gilbert, H. J., ed.), pp  
412 429–452, Academic Press., <https://doi.org/10.1016/b978-0-12-415931-0.00023-9>
- 413 41 Stern, J., Kahn, A., Vazana, Y., Shamshoum, M., Moraïs, S., Lamed, R., et al. (2015) Significance of relative  
414 position of cellulases in designer cellulosomes for optimized cellulolysis. *PLOS ONE*, **10**, e0127326,  
415 <https://doi.org/10.1371/journal.pone.0127326>
- 416 42 Li, X., Xiao, Y., Feng, Y., Li, B., Li, W. and Cui, Q. (2018) The spatial proximity effect of beta-glucosidase  
417 and cellulosomes on cellulose degradation. *Enzyme Microb. Technol.* **115**, 52–61,  
418 <https://doi.org/10.1016/j.enzmictec.2018.04.009>



- 419 43 Xu, Q., Ding, S.-Y., Brunecky, R., Bomble, Y. J., Himmel, M. E. and Baker, J. O. (2013) Improving activity of  
420 minicellulosomes by integration of intra- and intermolecular synergies. *Biotechnol. Biofuels* **6**, 126,  
421 <https://doi.org/10.1186/1754-6834-6-126>
- 422 44 Chen, L. and Ge, X. (2018) Correlation between size and activity enhancement of recombinantly assembled  
423 cellulosomes. *Appl. Biochem. Biotechnol.* **186**, 937–948, <https://doi.org/10.1007/s12010-018-2786-3>
- 424 45 Bae Jungu, Kuroda Kouichi, and Ueda Mitsuyoshi. (2015) Proximity effect among cellulose-degrading  
425 enzymes displayed on the *Saccharomyces cerevisiae* cell surface. *Appl. Environ. Microbiol.* **81**, 59–66,  
426 <https://doi.org/10.1128/aem.02864-14>
- 427 46 Chundawat, S. P. S., Paavola, C. D., Raman, B., Nouailler, M., Chan, S. L., Mielenz, J. R., et al. (2016)  
428 Saccharification of thermochemically pretreated cellulosic biomass using native and engineered cellulosomal  
429 enzyme systems. *React. Chem. Eng.* **1**, 616–628, <https://doi.org/10.1039/C6RE00172F>
- 430 47 Gilbert, C., Howarth, M., Harwood, C. R. and Ellis, T. (2017) Extracellular self-assembly of functional and  
431 tunable protein conjugates from *Bacillus subtilis*. *ACS Synth. Biol.* **6**, 957–967,  
432 <https://doi.org/10.1021/acssynbio.6b00292>
- 433 48 Leis, B., Held, C., Andreeßen, B., Liebl, W., Graubner, S., Schulte, L.-P., et al. (2018) Optimizing the  
434 composition of a synthetic cellulosome complex for the hydrolysis of softwood pulp: identification of the  
435 enzymatic core functions and biochemical complex characterization. *Biotechnol. Biofuels* **11**, 220,  
436 <https://doi.org/10.1186/s13068-018-1220-y>
- 437 49 Zholobko, O., Hamed, A., Zakharchenko, A., Borodinov, N., Luzinov, I., Urbanowicz, B., et al. (2021)  
438 Biomimetic cellulosomes assembled on molecular brush scaffolds: random complexes vs enzyme mixtures.  
439 *ACS Appl. Polym. Mater.* **3**, 1840–1853, <https://doi.org/10.1021/acsapm.0c01407>
- 440 50 Dong, H., Zhang, W., Zhou, S., Huang, J. and Wang, P. (2021) Engineering bioscaffolds for enzyme assembly.  
441 *Biotechnol. Adv.* **53**, 107721, <https://doi.org/10.1016/j.biotechadv.2021.107721>
- 442 51 Lu, L., Zhang, L., Yuan, L., Zhu, T., Chen, W., Wang, G., et al. (2019) Artificial cellulosome complex from  
443 the self-assembly of Ni-NTA-functionalized polymeric micelles and cellulases. *ChemBioChem* **20**, 1394–1399,  
444 <https://doi.org/10.1002/cbic.201900061>
- 445 52 Montanier, C. Y., Fanuel, M., Rogniaux, H., Ropartz, D., Di Guilmi, A.-M. and Bouchoux, A. (2019)  
446 Changing surface grafting density has an effect on the activity of immobilized xylanase towards natural  
447 polysaccharides. *Sci. Rep.* **9**, 5763, <https://doi.org/10.1038/s41598-019-42206-w>

- 448 53 Tuson, H. H., Foley, M. H., Koropatkin, N. M. and Biteen, J. S. (2018) The Starch Utilization System  
449 Assembles around Stationary Starch-Binding Proteins. *Biophys. J.* **115**, 242–250,  
450 <https://doi.org/10.1016/j.bpj.2017.12.015>
- 451 54 Foley, M. H., Déjean, G., Hemsworth, G. R., Davies, G. J., Brumer, H. and Koropatkin, N. M. (2019) A Cell-  
452 Surface GH9 Endo-Glucanase Coordinates with Surface Glycan-Binding Proteins to Mediate Xyloglucan  
453 Uptake in the Gut Symbiont *Bacteroides ovatus*. *J. Mol. Biol.* **431**, 981–995,  
454 <https://doi.org/10.1016/j.jmb.2019.01.008>
- 455 55 Gray, D. A., White, J. B. R., Oluwole, A. O., Rath, P., Glenwright, A. J., Mazur, A., et al. (2021) Insights into  
456 SusCD-mediated glycan import by a prominent gut symbiont. *Nat. Commun.* **12**, 44,  
457 <https://doi.org/10.1038/s41467-020-20285-y>
- 458 56 Enjalbert, T., De La Mare, M., Roblin, P., Badruna, L., Vernet, T., Dumon, C. et al. (2020) Characterisation of  
459 the Effect of the Spatial Organisation of Hemicellulases on the Hydrolysis of Plant Biomass Polymer. *Int. J.*  
460 *Mol. Sci.* **21**, <https://doi.org/10.3390/ijms21124360>
- 461 57 Badruna, L., Burlat, V., Roblin, P., Enjalbert, T., Lippens, G., Venditto, I., et al. (2021) The Jo-In protein  
462 welding system is a relevant tool to create CBM-containing plant cell wall degrading enzymes. *New*  
463 *Biotechnol.* **65**, 31–41, <https://doi.org/10.1016/j.nbt.2021.07.004>
- 464 58 Eibinger, M., Ganner, T., Plank, H. and Nidetzky, B. (2020) A Biological nanomachine at work: watching the  
465 cellulosome degrade crystalline cellulose. *ACS Cent. Sci.* **6**, 739–746,  
466 <https://doi.org/10.1021/acscentsci.0c00050>
- 467 59 Tatli, M., Moraïs, S., Tovar-Herrera, O. E., Bomble, Y. J., Bayer, E. A., Medalia, O., et al. (2022) Nanoscale  
468 resolution of microbial fiber degradation in action. *eLife* **11**, e76523, <https://doi.org/10.7554/eLife.76523>
- 469 60 Karunatilaka K. S., Cameron E. A., Martens E. C., Koropatkin N. M., and Biteen J. S. (2014) Superresolution  
470 Imaging Captures Carbohydrate Utilization Dynamics in Human Gut Symbionts. *mBio*, ASM **5**, e02172-14,  
471 <https://doi.org/10.1128/mBio.02172-14>

472  
473

474 **Figure 1: Levels of spatial organization in GHs and other CAZymes.** Different degrees of structural and spatial  
475 organization are shown, from simplest to more complex ones. (A) Clusters of GHs. Recently some cellulases from  
476 *T. reesei* have been described forming these transient associations during cellulose degradation. (B) Multi-

477 modular and multifunctional enzymes. In nature microorganisms produce a large variety of these kind of enzymes  
478 exhibiting in many cases complex architectures, the presence of CBMs is common in a great part of them. These  
479 CBMs affect enzymes' activities both targeting enzymes to their complex substrates and interacting spatially with  
480 the catalytic modules. (C) Cellulosome complexes. Depending on the bacteria specie and conditions, cellulosomes  
481 can be released to the environment or be attached to the cell surface by special anchoring systems. (D) PULs.  
482 While a great variety of these gene clusters have been described by metagenomics approaches, all share in their  
483 organization the presence of the SusC/D-like proteins transporter complexes; SGBPs and some GHs expressed as  
484 lipoproteins can interact between them and with SusC/D-like proteins facilitating the sequestering and hydrolysis of  
485 complex substrates. The image shows the system used by *Bacteroides ovatus* for xylan utilization (adapted from  
486 [24]). Notice that any of the existing catalytic modules from the different CAZymes hydrolytic families (glycoside  
487 hydrolases (GHs) and carbohydrate esterases (CEs)) can appear in the different levels and are represented by the  
488 different colours and shapes proposed. Different families of CBMs are also showed with different colours. BCW,  
489 bacterial cell wall; AP, anchoring protein; OM, outer membrane; P, periplasm; IM, inner membrane; C, cytoplasm;  
490 TBDT, TonB dependent transporter; HTCS, hybrid two-component systems. Monosaccharides are represented by  
491 Consortium for Functional Glycomics symbols.

492

493 **Figure 2: Spatial organization and enzymatic activity of GHs.** (A) Full-length model of *ScLPMO10C* composed  
494 of the LPMO module (*ScAA10*, in green; PDB code 4OY7), the linker (in blue; Biomagnetic Resonance Data Bank  
495 accession number 27078) and the CBM (*ScCBM2*, in red; PDB code 6F7E). The corresponding amino acid  
496 sequences of the C-ter of *ScAA10*, the linker and the N-ter of *ScCBM2* is illustrated above using the corresponding  
497 labeling colors. At low substrate concentration (from 2 to 10 g/L of avicel), the full-length *ScLPMO10C* (upper  
498 panel) is able to release more solubilized oxidized products than *ScAA10* (lower panel) (adapted from [36]). (B)  
499 Comparative enzymatic degradation of cellulosic substrate by recombinant enzymes either in the free state (Free,  
500 light blue), bound to respective monovalent scaffoldins (CBM-restored, orange) or bound to scaffoldins at different  
501 positions (Scaf, light purple and dark purple for the most efficient GHs combinations). Details of the constructions  
502 are provided on the right, with number referring to GH families (adapted from [41]). The position of enzymes in a  
503 scaffold can be critical for optimal substrate degradation. (C) A GH11 xylanase *NpXyn11A* was immobilized on a  
504 solid surface at a controlled protein surface density to regulate the distance between enzymes. In this study, the  
505 proteins Jo and In were used as self-grafting inducers. The left panel illustrates the increasing protein surface

506 density of grafting from the higher density (square 1:bead1) to the lower (square 5: bead 5). The hypothetical  
507 arrangement of immobilized enzymes at maximum occupancy is schematically shown. The right panel illustrates  
508 the amount of oligosaccharides (XOS) released after 5 h of incubation: the closer the enzymes are to each other, the  
509 more XOS are released relative to the equivalent amount of free enzyme (in orange) (adapted from [52]).