Check for updates

Structural and functional insights into cellulosomes: masters of plant cell wall degradation

Nataša Lindič[1]* and Maša Vodovnik[2]

[1]Department of Biochemistry, Molecular and Structural Biology, Jožef Stefan Institute, Ljubljana, Slovenia, [2]Department of Microbiology, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

Cellulosomes are complex multi-enzyme systems that enable efficient cellulose breakdown in some anaerobic bacteria and fungi. Understanding cellulosome functionality plays a crucial role in expanding their potential for industrial plant biomass degradation and valorization. While knowledge on these intricate structures has been accumulating for several decades, recent insights into their modular architecture, dynamic assembly mechanisms, and potential for synthetic biology-driven redesign for biotechnological applications call for a comprehensive re-evaluation of their structural and functional complexity. This review explores recent advances in understanding these cellulolytic nanomachines, focusing on substrate recognition and binding mechanisms, including the roles of carbohydrate-binding modules and cohesin-dockerin interactions. Cell-surface mechanisms that allow these complexes to attach to and effectively degrade plant biomass are also reviewed. Furthermore, structural adaptations to diverse substrates and environmental conditions are discussed, highlighting the flexibility and the interplay between the cellulosomal components, both catalytic and non-catalytic, and their impact on optimizing cellulose degradation, including carbon source sensing, and its role in modulating cellulosome architecture and activity.

KEYWORDS

cellulosome, cellulose, cellulase, CBM, cohesin, dockerin, scaffoldin, SLH

1 Introduction

Degradation of cellulose, the most abundant organic polymer on Earth, is instrumental in the global carbon cycle. Cellulosomes are large multienzyme complexes that break down cellulose-rich plant material. They are found in natural environments like soil, compost, and the rumen of herbivores, as well as in artificial systems such as anaerobic digesters. They are primarily produced by anaerobic bacteria, although similar structures have also been detected in some anaerobic fungi (Artzi et al., 2017; Lillington et al., 2021; Minor et al., 2024; Vodovnik and Lindič, 2025). In a recent genomic study involving the analysis of 305,693 bacterial genomes, 33 bacterial species with the genomic capacity to produce cellulosomes (including 10 previously unreported) were identified. This analysis revealed that cellulosome-producing capacity originates mostly from four bacterial genera: *Acetivibrio*, *Ruminococcus*, *Ruminiclostridium* and *Clostridium* (Minor et al., 2024). Cellulosomal complexes have been shown to enhance biomass degradation by optimizing enzyme synergy and substrate targeting. This offers an interesting option for different industrial applications, particularly second-generation biofuels. Furthermore, the modular architecture of cellulosomes has inspired the design of artificial complexes [designer cellulosomes (DCs)], tailored for even more efficient and targeted degradation of plant biomass (Vodovnik and Lindič, 2025).
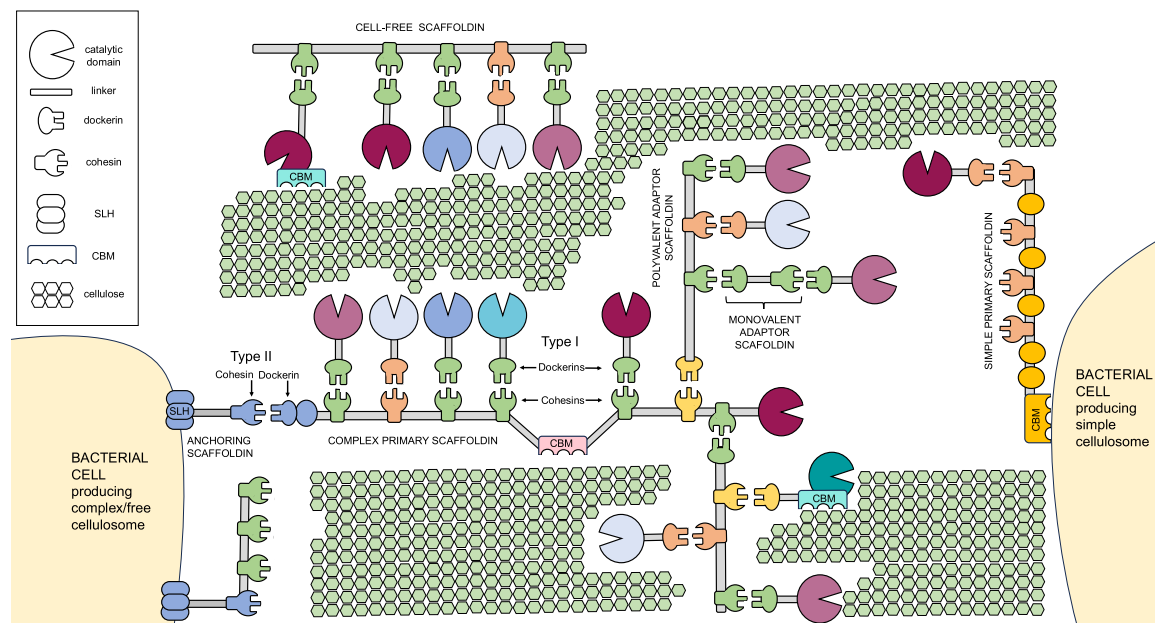
FIGURE 1
The architectural diversity of the cellulosomes. Major structural components of different types of cellulosomes and their attachment to the host cells are shown. CBM, carbohydrate binding module; SLH, S-layer homology domain.
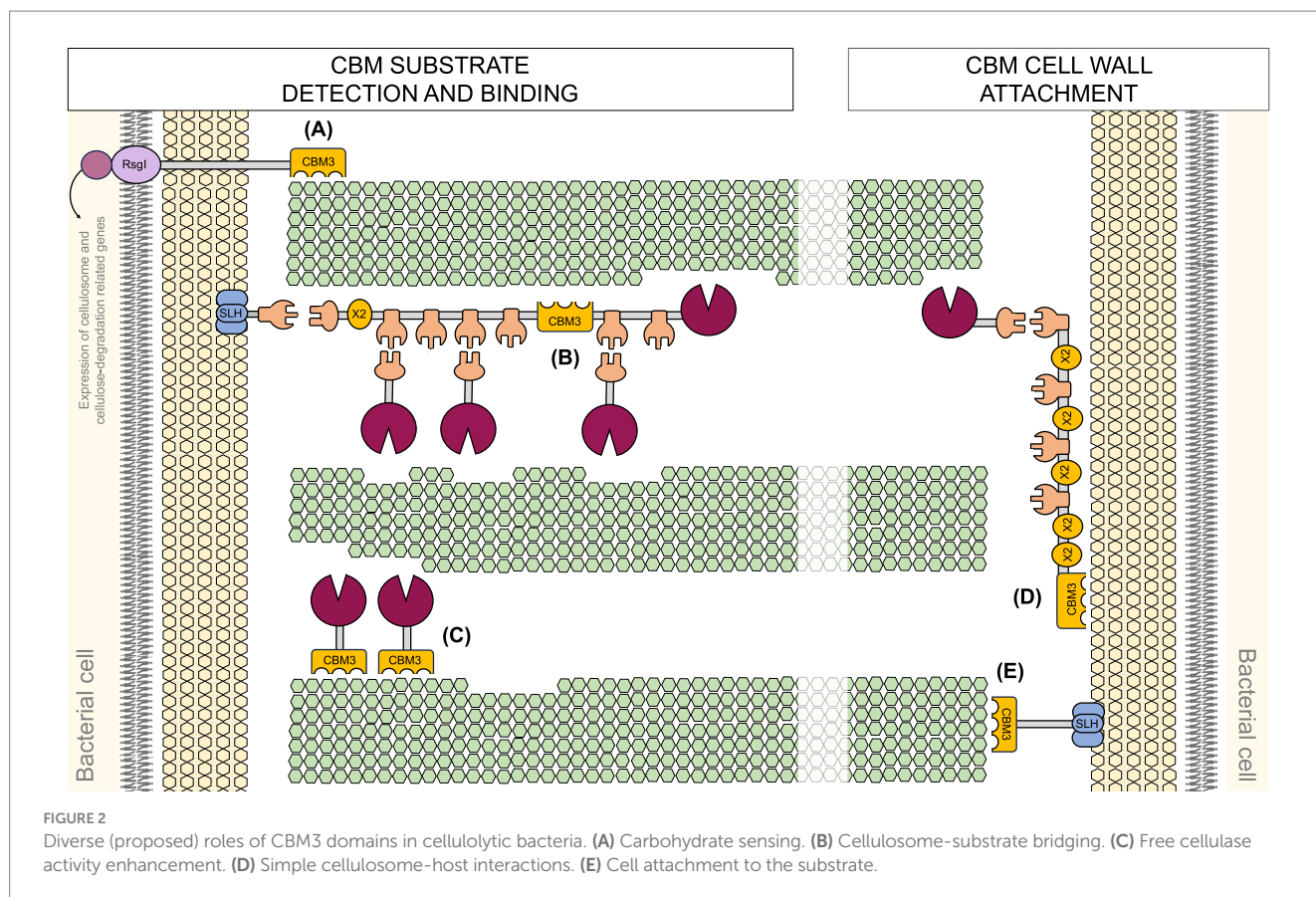
Cellulosomes are architecturally versatile complexes (Figure 1) that play a pivotal role in the process of cellulose degradation by integrating carbohydrate-degrading enzymes onto a flexible structural backbone - the scaffoldin. Scaffoldins are large non-catalytic proteins that organize and anchor catalytic components via multiple cohesin domains. Cohesin domains interact specifically and strongly with the dockerin domains of catalytic subunits. Catalytic domains typically include cellulases, hemicellulases, pectinases, and other carbohydrate-active enzymes (reviewed in Artzi et al., 2017). Additionally, accessory proteins such as expansin-type proteins (Chen C, et al., 2016; Cosgrove, 2017), proteases (Levy-Assaraf et al., 2013) and protease inhibitors (Ó Cuív et al., 2013) are also present. Interestingly, proteins with CotH-like domains have also been identified within some cellulosomes (Vodovnik et al., 2013). Although their role is not clear yet, findings suggest that CotH proteins in cellulosomes may serve in two ways: as structural components facilitating the assembly of the complex as well as regulatory elements via their kinase activity (Ben David et al., 2015; Blouzard et al., 2010; Lillington et al., 2023; Zverlov et al., 2003). Scaffoldins form the structural backbone of the cellulosomes. The most common scaffoldins, called primary scaffoldins, usually have multiple cohesin domains that bind dockerin-tagged enzymes. Most scaffoldins also have a special dockerin that helps attach them to the cell surface by binding to cohesins on anchoring scaffoldins. These protein connect the complex to the cell surface either non-covalently via S-layer homology (SLH) domains or covalently through sortase-mediated attachment. In more complex

cellulosomes, adaptor scaffoldins have also been found, linking two scaffoldins or a scaffoldin and an enzyme (reviewed in Artzi et al., 2017, 2014). Cellulosomes also work as a link between the bacterial cell and the substrate. The attachment is mediated via scaffoldin's carbohydrate-binding modules (CBMs) and results in concentrating enzymes directly on the substrate, additionally enhancing its degradation efficiency (Galera-Prat et al., 2018; Valbuena et al., 2009; Verdorfer et al., 2017). The cellulosome-mediated sensing and degradation of cellulose biomass is a result of several complex mechanisms, including substrate recognition and binding, structural adaptation, cell-surface anchoring, enzyme synergy and proximity effects. Recent advances in studying these mechanisms with emphasis on functional and structural aspects of cellulosomes are reviewed below.

2 Carbohydrate-binding modules and their interplay with other cellulosomal domains

Carbohydrate binding modules (CBMs) are non-catalytic domains found on cellulosomal scaffoldins as well as various carbohydrate-active enzymes (CAZymes). CBMs play diverse roles in cellulolytic bacteria (Figure 2). One of the crucial ones is substrate recognition, i.e., binding of the cellulosomal complex to specific carbohydrate structures within the plant biomass. CBMs are an integral part of the enzymes in structure and function. As such, they have also been shown to improve the thermostability of the enzymes (De Chellis et al., 2024b; Li et al., 2023), and disrupt the substrate to facilitate attack by the catalytic domain or other enzymes (Arantes and Saddler, 2010; Li and Liu, 2024; Reinikainen et al., 1992; You et al.,

---

FIGURE 2
Diverse (proposed) roles of CBM3 domains in cellulolytic bacteria. **(A)** Carbohydrate sensing. **(B)** Cellulosome–substrate bridging. **(C)** Free cellulase activity enhancement. **(D)** Simple cellulosome–host interactions. **(E)** Cell attachment to the substrate.

2024). In addition, CBMs in simple cellulosomes were proposed to be involved in cell-surface anchorage, described in section 3 in more detail. Furthermore, non-cellulosomal CBMs in cellulolytic bacteria have been shown to be a part of the unique carbohydrate-biosensing system. Herein, an extracellular CBM as a part of an anti-sigma factor RsgI was shown to detect polysaccharides. This leads to signal transduction and subsequently activation of gene transcription of cellulosomal systems and cellulase-related genes, potentially enabling a dynamic cellulosome assembly for efficient polysaccharide hydrolysis (Chen C, et al., 2023; Dong et al., 2024; Hershko Rimon et al., 2021; Kahel-Raifer et al., 2010; Sand et al., 2015; Yaniv et al., 2014).

2.1 CBM types

The diversity of CBM structures allows cellulosomes and CAZymes for a more specific substate recognition within heterogeneous plant biomass. Based on their tertiary structure/function similarities, particularly their ligand-binding sites, CBMs are classified into three types, type A, B and C. Type-A CBMs have a flat or platform-like binding site composed of aromatic residues that recognize the surface of crystalline polysaccharides (surface-binding CBMs). Type-B CBMs have a groove or cleft-like binding that can bind internally to single polysaccharide chains (endo-type CBMs). Type-C CBMs lack the extended binding-site grooves of type B CBMs and bind the shorter substrates at polysaccharide termini (exo-type; Boraston et al., 2004; Creagh et al., 1996; Gilbert et al., 2013; Liu et al., 2022). Recently, a novel ligand preference-based classification has

been proposed, describing four types of CBMs. Herein, type I and II include CBMs specific for non-branched carbohydrates, while type III and IV CBMs recognize and attach to branched polysaccharides and enhance the hydrolysis of substrates containing side chains (Liu et al., 2021).

2.2 CBM3 structure

According to their primary structure, CBMs are currently classified into 106 families (Carbohydrate Active Enzymes Database 2025;[1] Boraston et al., 2004; Creagh et al., 1996; Drula et al., 2022; Gilbert et al., 2013; Liu et al., 2022; You et al., 2024). Among these, CBM3 modules (Figure 2) are particularly frequently found in cellulosomes. CBM3a are known for strong binding to crystalline cellulose, which proved essential for its degradation (Yaniv et al., 2012; Yaniv et al., 2011). CBM3s are approximately 150 amino acids long. They typically adopt a β-jelly roll fold with nine β-strands forming one flat and one curved β-sheet. While the role of the concave surface remains unclear (Hershko Rimon et al., 2021; Yaniv et al., 2012), the flat surface composed of aromatic residues (His, Trp, Tyr and an Arg-Asp ion pair) in a linear strip aligns with the glucose rings of the cellulose chain (Ding et al., 2006; Tormo et al., 1996). Mutations herein have been shown to affect substrate specificity. Single mutations can result in the

---

1   http://www.cazy.org/

loss of binding to cellulose (Cai et al., 2011), restoration of binding (Yaniv et al., 2012) or altered adhesion properties (von Schantz et al., 2012). In a non-cellulosomal CBM3 from *Clostridium thermocellum (C. thermocellum),* proposed to serve as a surface-anchored polysaccharide biosensor, the lack of the required number of key aromatic residues might explain the preference for xylan binding (Hershko Rimon et al., 2021). Furthermore, multimodal binding of the CBM to the substrate was revealed using recently developed acoustic force spectroscopy-based methods (Chundawat et al., 2024; De Chellis et al., 2024a; Hackl et al., 2022). Specifically, the pulling on the wild-type CBM3a and consequently its dissociation from the fiber revealed that the planar aromatic binding motif contains two carbohydrate binding regions (Hackl et al., 2022).

In addition to multimodal binding CBM characteristics, multivalent CBMs have been discovered. These exhibit dual-specificity binding of both, cellulose and xyloglucan, which is influenced by substrate cleft topology (Hernandez-Gomez et al., 2015; Levi Hevroni et al., 2020; Venditto et al., 2016). Notably, certain multivalent CBMs, such as recently discovered family 92, feature unique three-site binding modes that were proposed to facilitate polysaccharide cross-linking in nature (Hao et al., 2024). In contrast to CBMs discussed so far that are functionally and structurally autonomous from their associated catalytic domains, there are some exceptions like group c CBM3s (CBM3cs). CBM3cs are integral to the substrate-binding clefts of glycoside hydrolases 9 (GH9) and lack conserved ligand-binding residues. Thus, they do not bind cellulose directly. Instead, CBM3cs play a structural or catalytic-assisting role. Here, aromatic residues are replaced with polar amino acids, that disrupt cellulose inter-chain hydrogen bonds, leading the cellulose chains into the catalytic domain's active site for hydrolysis (Gilad et al., 2003; Jindou et al., 2006; Kuch et al., 2023; Li et al., 2010; Sakon et al., 1997). All in all, since the substrate specificity of CBMs is intricately tied to critical residues and their spatial structure, targeted modification of these residues is a promising strategy for directed evolution and diversification of these modules (Zhou et al., 2024).

2.3 X2 modules

All simple primary scaffoldins found in *C. cellulovorans, Ruminiclostridium cellulolyticum (R. cellulolyticum), Clostridium josui* and *Clostridium saccharoperbutylacetonicum* (Dassa et al., 2017; Levi Hevroni et al., 2020) contain X2 modules. X2 modules are hydrophilic domains of approximately 100 amino acids (Mosbah et al., 2000) widely distributed among cellulosome-producing bacteria (Dassa et al., 2017; Doi et al., 1994; Lamed and Bayer, 1988) and free cellulases (Pasari et al., 2017; Ravachol et al., 2015; Vita et al., 2019). The function of X2 modules remains unclear, though they have been proposed to contribute to cellulose binding and hydrolysis (Aburaya et al., 2015; Pasari et al., 2017; Tamaru et al., 2011), to stabilize adjacent cohesin domains (Levi Hevroni et al., 2020) or interact with bacterial cell walls (Kosugi et al., 2004; Tarraran et al., 2021). Nevertheless, direct binding of a cell wall component has never been experimentally demonstrated. Moreover, recent studies demonstrate that the X2 modules within the CipC scaffoldin lack direct binding affinity to both, the bacterial cell surface and cellulose (Tao et al., 2022). The structural analysis of the X2 module shows a predominantly hydrophilic surface with a shallow hydrophobic groove (Mosbah et al.,

2000) which might explain the lack of direct cell wall binding. When the loop-located motif (NGNT), conserved in all X2 modules, is removed, this groove widens (Tao et al., 2022) into conformation that resembles CBM structures (Luís et al., 2013; Pasari et al., 2017) and presumably allows weak cell wall interactions (Tao et al., 2022). This motif was found to be crucial for cellulose hydrolysis in *Clostridium cellulolyticum (C. cellulolyticum)* (Tao et al., 2022). All in all, the X2 modules are located next to the CBM3a modules in free cellulases and the CBM3-X2 module has shown a better binding affinity to crystalline cellulose compared to CBM3 alone (Pasari et al., 2017; Zhang et al., 2018). Thus, the mechanism was proposed where X2 module interacts with CBM3a to promote its binding function (Tao et al., 2022). These findings offer a compelling framework for further investigations.

2.4 CBM-based engineering applications

CBMs have long been known to enhance enzymatic activity of their partner catalytic domains by ensuring their proximity to the substrate. In DCs, CBMs with different specificities have been used to direct complexes to different substrates (Lamote et al., 2023; Vanderstraeten et al., 2022a). Fusing various *Ruminoclostridium thermocellum* CBMs with specific binding activities to multifunctional cellulosomal enzyme CelE enabled its activity on diverse substrates like lichenan, xylan, and mannan (Walker et al., 2015). Similarly, CBM binding specificity affected the potency of chimeric endoglucanases to degrade milled lignocellulosic materials (Ichikawa et al., 2016). In *Paenibacillus polymyxa* A18, the *X2-CBM3 module* in association with xyloglucanase and endoglucanase enhanced enzyme activity up to 4.6-fold (Pasari et al., 2017). Similar increase in enzymatic activity was observed for *Acetivibrio thermocellus* CBM3 *fused to GH5* endoglucanase from *Trichoderma viride* (Ran et al., 2024). Furthermore, CBM1 fused with lytic polysaccharide monooxygenases from *Neurospora crassa* enhanced fiber oxidation (Støpamo et al., 2024). Using CBMs with cutinase even increased PET plastic degradation, showcasing the potential of CBMs in developing environmentally friendly recycling strategies (Chen Y, et al., 2023). Surprisingly, reducing CBM-binding affinity by mutagenesis of aromatic residues in the planar binding motif while maintaining binding sites conformation can enhance the catalytic activity of CelE–CBM mutants on some cellulose substrates. This supports the hypothesis that non-productive binding hinders efficiency (Nemmaru et al., 2021) and provides additional confirmation that the optimal binding strength is critical for catalysis (Kari et al., 2018), as observed in studies on cellulases and expansins (Hepler and Cosgrove, 2019). The surface-charged interactions between slightly negative lignin and CAZymes were recently explored. Introducing and fine-tunning mutations to supercharge the surface of CBM and/or its endocellulase partners has been shown to enhance endocellulase thermostability, binding, and activity on cellulosic biomass (De Chellis et al., 2024b). This is in line with the observation that electrostatic interactions are involved in the binding of CBM3s to hemicellulose and that the xylan-binding ability of the CBM3s is affected by the ionic strength of the environment (Hershko Rimon et al., 2021). Altogether, these findings highlight the critical need to further investigate the complex interplay between the CBMs and their catalytic domain partners.

Incorporation or removal of non-catalytic domains such as CBMs from the polypeptide can modulate thermostability, binding specificity,

and catalytic efficiency of the enzymatic module (Hettle et al., 2017; Kuch et al., 2023; Nakamura et al., 2016; Sammond et al., 2012; Tang et al., 2013; Walker et al., 2015). While most cellulase-associated CBMs are located N- or C-terminally, insertion of the CBM sequence within the sequence of the catalytic domain is less common. So far, it was identified in only a few xylanases (Flint et al., 1997; Wu et al., 2011). Additional example identified recently is the endoglucanase from thermophilic bacterium *Meiothermus taiwanensis* WR-220 (Ye et al., 2022). Structural analysis showed that this GH5-family endo-*β*-1,4-glucanase has a bipartite architecture featuring a Cel5A-like domain with an inserted CBM29-like domain. Deletion of this CBM domain significantly reduced the activity of the enzyme, while its insertion into the *Thermotoga maritima* Cel5A significantly enhanced affinity of the chimera to longer polysaccharides (Ye et al., 2022). Furthermore, superimposition of MtGlu5 with homologs and substrate-binding modeling suggest that the polysaccharide chain interacts with a continuous groove from the catalytic domain to the CBM. Mutagenesis experiments showed six tryptophans (three in the CBM and three in the catalytic domain) that are crucial for sugar binding. These findings highlight the synergistic effect of a CBM insertion into a GH domain of the enzyme and its potential for extending the substrate binding groove to increase affinity to longer substrates (Ye et al., 2022).

In a recent study by Vita et al. (2019), the specific features required to generate an efficient free or cellulosomal family-9 cellulase were identified. Specifically, removing the two C-terminal X2 modules and the CBM3b moiety from free-state Cel9A reduced its activity on crystalline cellulose. In contrast, adding these elements to cellulosomal Cel9G resulted in increased activity on crystalline cellulose (Vita et al., 2019). X2 and CBM3 together were demonstrated to be highly efficient in enhancing the activity of GHs toward the insoluble substrate (Pasari et al., 2017). Furthermore, in a DC system, the influence of CBM copy number within the scaffoldin (constructs containing two, one, or no CBMs) was evaluated. The results demonstrated that increasing the number of CBMs significantly enhances cellulosome efficiency by improving substrate binding. The construct with two CBMs achieved the highest sugar release from Avicel. A less pronounced effect was observed during the degradation of PASC (Anandharaj et al., 2020). All in all, tandem CBMs can coordinate catalytic domains by leveraging differences in binding affinity, ligand preference, and spatial arrangement of multiple enzymes into a single multimodular unit for complex substrate degradation (Kuch et al., 2023; Liu et al., 2024).

Variability in glycan recognition was also suggested to be an important aspect of efficient polysaccharide degradation and regulation of the transcription of cellulosomal proteins in response to different polysaccharides. For example, in the CBM-ome of *Ruminococcus flavefaciens* (*R. flavefaciens*), six novel CBM families binding β-glucans, β-mannans, and the pectic polysaccharide homogalacturonan were identified (Venditto et al., 2016). Similarly, in (*Pseudo*)*bacteroides cellulosolvens*, which was reported to contain the most complex known cellulosome system to date, a novel type B CBM has been shown to displays similarities to fibronectin type III (Fn3) domain and acts as an extracellular substrate biosensor with broad polysaccharide binding profile. This module exhibited extensive binding to chitosan and arabinoxylan, medium binding to amorphous cellulose and xylan, and weak binding to Avicel (Dong et al., 2024). Recent genomic and proteomic studies further emphasizes the variability of CBMs across different species (Dassa et al., 2017; Minor et al., 2024; Zhivin-Nissan et al., 2019).

Altogether, CBMs within cellulosomes and those in free CAZymes play essential roles in substrate targeting, enzymatic efficiency, and stability. Among them, cellulosome-prevalent CBM3a is well known for its strong affinity to crystalline cellulose, while X2 modules are thought to support CBM function and/or scaffoldin stability. Recent advances reveal additional layers of CBM functionality, including multimodal binding, dual substrate specificity, and roles in carbohydrate sensing and transcriptional regulation. From an engineering standpoint, CBMs serve as modifiable elements for tailoring enzymes and DCs with improved binding, activity, and specificity.

3 Cohesin-dockerin interactions

The primary interaction within the cellulosomal subunits is the interaction between dockerin and cohesin domains (Artzi et al., 2017; Bule et al., 2018a; Bule et al., 2017). It resembles a plug-and-socket mechanism, and can be classified based on their origin, specificity and structural features. According to their primary sequence, three types of cohesin-dockerin pairs have so far been identified: type I, type II, and type III. Type I pairs are commonly involved in enzyme organization within the cellulosome, while type II pairs facilitate cellulosome attachment to the cell surface (Artzi et al., 2017; Pinheiro et al., 2009). Exceptions where the situation is reversed have also been described. Namely, in *Bacteroides cellulosolvens* (*B. cellulosolvens*), enzymes are associated with type II dockerins, while scaffoldins are associated with type I dockerins (Bayer et al., 2008; Cameron et al., 2015; Duarte et al., 2021; Noach et al., 2005). Type III pairs are specific to ruminococcal cellulosomes and differ from the type I and II found in *Clostridium* species (Duarte et al., 2023; Voronov-Goldman et al., 2015). Understanding these modular and specific cohesin-dockerin interactions lays the foundation for harnessing and redesigning cellulosomes as modular synthetic scaffolds, opening avenues for customizable enzyme complexes in synthetic biology.

3.1 Cohesin-dockerin binding modes and their structural determinants

The binding mode of a given dockerin is dictated by its structural features and the conservation or divergence of key interacting residues. Type I dockerins are well-documented for their symmetrical binding surface, which supports dual binding modes. They contain two EF-hand-like Ca^{2+}-binding motifs, each with symmetrical antiparallel *α*-helices flanking the conserved Ca^{2+}-binding residues (aspartate and asparagine) and coordination patterns (Adams et al., 2005; Pagès et al., 1997; Pinheiro et al., 2009). This duplication creates a symmetrical cohesin-binding surface with residues critical for species-specific recognition, enabling dual binding in two orientations (180° apart). Dual binding mode increases flexibility and efficiency in enzyme assembly within the cellulosome and thus provides enhanced function, avoidance of steric clashes, and adaptation to substrates (Carvalho et al., 2007; Carvalho et al., 2003; Chen et al., 2014; Pagès et al., 1997; Pinheiro et al., 2008). In this mode, the cohesin structure remains mostly unchanged, while the dockerin's loop–helix–helix–loop–helix motif undergoes conformational changes, yet retaining its EF-hand coordination with Ca^{2+} ions. Dual-binding mode was recently proposed to be allosterically regulated. Two alternative binding conformations in

type-I cohesin-dockerin pairs were identified, and their binding mode mediated by the isomerization state of a single proline residue. While the exact determinants of binding mode equilibrium remain unclear, they likely depend on the specific cohesin-dockerin pair (Vera et al., 2021).

Until recently, type II cohesin-dockerin pairs (typically involved in cell wall attachment) were considered to adopt a single-binding mode. In single binding mode, only one of the dockerin interfaces supports the formation of the cohesin-dockerin complex, as seen for type II *C. thermocellum* dockerin (Adams et al., 2006) or adaptor scaffoldins in *R. flavefaciens* (Bule et al., 2016). However, recent work revealed a dual-binding mode of type II dockerin from *B. cellulosolvens* (Duarte et al., 2021). This finding challenges the previous belief that a dockerin's binding mode depends on its type or function. In fact, both type I and type II complexes, whether involved in enzyme recruitment or cell-surface anchoring and scaffoldin assembly, exhibit dual-binding modes. This was further confirmed in *Ruminococcus*-specific type III complexes (Duarte et al., 2023). The structure of type III dockerins exhibits less symmetry compared to type I and II dockerins found in *Clostridium* species and lacks the second canonical Ca^{2+}-binding loop (Bule et al., 2018b; Duarte et al., 2023; Rincon et al., 2007; Rincon et al., 2005; Rincon et al., 2003; Salama-Alber et al., 2013; Voronov-Goldman et al., 2015). Nevertheless, it demonstrates similar behavior, maintaining responsiveness to Ca^{2+} and a strong affinity for the cohesin on the scaffoldin (Karpol et al., 2013). Type III-specific recognition "hot spots" were identified by switching the specificity of ScaA cohesin to mimic type I cohesin interaction (Weinstein et al., 2015). Notably, Duarte et al. (2023) described the crystal structure of the adaptor scaffoldin ScaH-borne dockerin bound to the cohesin of the anchoring scaffoldin ScaE, demonstrating an unexpected dual-binding mode of a type III complex (Duarte et al., 2023).

All in all, the dual-binding mode is much more widely distributed than initially proposed (Bule et al., 2018a). There is no definitive evidence for selective pressure favoring the dual-binding mode. Species like *R. flavefaciens* can assemble functional cellulosomes using only single-binding modes, suggesting the dual-binding mechanism may lack universal evolutionary relevance (Brás et al., 2012; Vera et al., 2021; Yao et al., 2020). Overall, the active regulation of dockerin-binding orientation and the complexity of systems like *Bacteroides cellulosolvens* suggest that single vs. dual-binding modes in cohesin-dockerin complexes are independent of type or function (enzyme recruitment vs. cell attachment/scaffoldin assembly; Duarte et al., 2021). Instead, these modes may be linked to cellulosome size and complexity, with larger systems requiring greater flexibility for assembly and substrate access. The effect of cellulosome size on catalytic efficiency was investigated recently using recombinantly assembled cellulosomes. Primary and secondary scaffoldins containing 1, 3, or 5 type I/II cohesin domains were synthesized and assembled with 1, 3, 5, 9, 15, or 25 cellulase molecules from GH families 5, 9, and 48. These complexes were attached to the cellulose by CBM3a. It was found that increasing cellulosome complexity enhances hydrolysis efficiency. Cellulosomes with 9, 15, or 25 cellulases showed similar activity, suggesting synergy saturated at 9 subunits in this system (Chen and Ge, 2018). However, in cellulolytic bacteria like *C. thermocellum*, which produce dozens of diverse cellulases, synergy continues to increase with higher enzyme diversity—up to 40 enzymes have shown enhanced effects (Hirano et al., 2016), highlighting the need for further investigation.

3.2 Multi-dockerin modules and cellulosome complexity

Generally, cellulosomal scaffoldins are composed of multiple tandem cohesin modules, each capable of binding different enzymes, most of which have only a single dockerin (Carvalho et al., 2007). Nevertheless, rare double- and multiple-dockerin modules have been identified in some cellulosome-producing bacteria. For example, in *C. thermocellum*, double-dockerins were identified, either with no partner or bearing a protease domain (Chen et al., 2024; Chen et al., 2019). Structural analysis of the S8 protease-associated double-dockerin module dDoc_0689 from *C. thermocellum* revealed that its two dockerins form a stable interface through hydrophobic interactions and hydrogen bonds with limited mobility. The first dockerin module displays a unique intramolecular clasp and exhibits binding preference to the cohesin of the cell-bound scaffoldin ScaD, while its second dockerin module remains functionally ambiguous. These findings suggest that dDoc_0689 may facilitate protease anchoring to the bacterial cell surface rather than integrating into primary enzyme-scaffoldin complexes (Chen et al., 2024). These double-dockerin proteins are conserved across different *C. thermocellum* strains and their expression has been confirmed by transcriptomic and proteomic studies (Raman et al., 2011; Wilson et al., 2013). Furthermore, genes encoding triple, sextuple and septuple dockerins have been identified in the genomes of some other bacterial species, for example *R. flavefaciens* and *Oscillospiraceae bacterium* J5864_00540 (Chen et al., 2024; Dassa et al., 2014) and also fungi (Gilmore et al., 2020). These multi-dockerin proteins might provide even more complexity for the assembly of cellulosomes, yet their structure and function are still unknown. Unraveling the conformation of these regions will not only deepen our understanding of cellulosomes but also contribute novel elements to synthetic biology and biotechnology (Chen et al., 2024).

3.3 Specificity and cross-species interactions of cohesin-dockerin complexes

The cohesin-dockerin interaction has generally been considered species-specific. The specificity of type I and type II cohesin-dockerin interactions is believed to prevent cross-reactivity, ensuring precise cellulosome assembly and cell surface attachment. This selectivity is governed by subtle differences in binding site topology and the lack of sequence identity in β-strand sequences that comprise the core of the dockerin binding site (Alber et al., 2009; Cameron et al., 2015; Carvalho et al., 2005; Carvalho et al., 2003; Noach et al., 2005), with a few key residues governing cohesin-dockerin specificity (Nakar et al., 2004). In type I cohesin-dockerins from *C. cellulolyticum* and *C. thermocellum*, species-specific recognition has been attributed to apolar interactions between the key residues centered around a hydrophobic pocket on the surface of the cohesin. In *C. cellulolyticum*, this pocket is formed by Leu87 and Leu89, which is occupied, in the two binding modes, by the dockerin residues Phe19 and Leu50, respectively (Pinheiro et al., 2008). Using a combination of computational and experimental methods, two types of hot spot residues were identified in *C. thermocellum* cohesin: affinity hot spots

like Leu83, which enhance binding strength, and specificity hot spots like Asn37, which determine partner selectivity (Slutzki et al., 2015). When Asn37 mutants were investigated, it was found that N37A binds both *C. thermocellum* and *C. cellulolyticum* dockerins promiscuously, while N37L shifts specificity toward *C. cellulolyticum*, reducing binding to the native partner (Slutzki et al., 2015). Nevertheless, in contrast to the type I, type II cohesin-dockerin pairs have a relatively extensive cross-species plasticity, which is common among the simple cellulosome systems of mesophilic clostridia (Haimovitz et al., 2008). Recently, a higher affinity for a cohesin from *C. cellulolyticum* than for those in *C. thermocellum* was detected in a study where a rare double-dockerin module from *C. thermocellum* was tested (Chen et al., 2024). An exception in species specificity was also observed in *C. thermocellum* dockerin interaction with *C. josui* cohesin (Jindou et al., 2004). However, biological relevance of this lack of species specificity is still unclear. Inter-species cohesin-dockerin interactions support the pan-cellulosome concept (Dassa et al., 2017), allowing microbial communities in shared ecosystems to assemble "symbiotic cellulosomes." This proposed cooperative strategy enhances enzymatic diversity, particularly under extreme conditions, without additional resource investment (Chen et al., 2024).

3.4 Biotechnological applications and engineering potential of cohesin-dockerin systems

As already detailed in the previous sections, the structural flexibility (3.1), multi-dockerin complexity (3.2), and specificity patterns (3.3) of cohesin-dockerin interactions offer powerful tools for synthetic biology. These features enable precise design of modular, multi-enzyme assemblies with controllable binding and functionality, paving the way for advanced biotechnological applications and engineered biomaterials.

Recent genomic analysis of 305,693 sequenced bacterial genomes identified 33 bacterial species with the genomic potential to produce cellulosomes (Minor et al., 2024). A high number of cohesins and dockerins identified are not only valuable as building blocks for DCs but also hold promise for the broader design of protein scaffolds enabling multi-enzyme assembly (reviewed in Gad and Ayakar, 2021). By combining cohesin-dockerin pairs from different species, it is possible to precisely control the enzyme composition, spatial arrangement, and stoichiometry within the DC (Vanderstraeten et al., 2022a). Advancements in protein engineering supported by computational modeling have led to the development of cohesin-dockerin pairs with tailored specificity, improved binding affinities, and increased stability under environmental conditions (Anandharaj et al., 2020; Bomble et al., 2011; Borne et al., 2020; Dorival et al., 2022; Duarte et al., 2023; Gilmore et al., 2020; Vazana et al., 2013; Vera et al., 2021; Wojciechowski et al., 2018). One notable achievement in this area is the construction of the largest known cellulosome complex, displaying 63 enzymes on the surface of *Kluyveromyces marxianus*. This system supports the modular addition of enzymes, creating a cellulolytic host for ethanol production. In this study, scaffoldins containing three, six, or nine cohesin modules were tested. Experiments revealed that enzyme synergism—and thus sugar release from highly crystalline substrates such as Avicel—was significantly influenced by cohesin number. In contrast, the impact was less pronounced when phosphoric acid-swollen cellulose (PASC) was used (Anandharaj et al., 2020).

Further insights into cellulosome dynamics were provided by Borne et al. (2020), who demonstrated that a fully assembled cellulosome can be functionally reprogrammed. They showed that the stability of preassembled enzyme–scaffoldin complex is governed by a key dockerin aminoacid residue at position 22. This position determines whether the interaction with the cohesin is reversible or nearly irreversible. Site-directed mutagenesis of this residue enables the conversion between these two binding states (Borne et al., 2020). Leucine residues at positions 22 and 58 are highly conserved among dockerin-bearing proteins across multiple bacterial species (Dassa et al., 2017). They are located at the termini of the α-helices in both dockerin segments (Pinheiro et al., 2008). Due to the dual-binding mode of cohesin–dockerin interactions, one of these residues is consistently positioned at the cohesin interface (Borne et al., 2020). Their conservation suggests a general mechanism underlying cellulosome dynamics across bacterial species and highlights the biotechnological potential of replacing specific enzymatic subunits through competition between dockerin variants (Borne et al., 2020).

Further, pH-dependent dual-binding-site switch in *C. acetobutylicum* cohesin-dockerin pair was recently discovered. This marks the first report of a pH-regulated protein–protein interaction switch and represents a model of biological regulation and a novel strategy for designing pH-responsive protein devices and biomaterials for biotechnology (Yao et al., 2020). In addition, dockerins's electrostatic profile alteration can enhance module affinity, providing a versatile platform for designing high-affinity technologies in industrial and research applications (Duarte et al., 2023). The cohesin-dockerin interaction is among the strongest non-covalent interactions in nature (Duarte et al., 2021). Its mechanical stability also influences cellulosomes enzymatic activity (Stahl et al., 2012). Mechanical forces have been shown to influence cohesins located on the scaffoldin (in the connecting region between the cell and cellulose anchoring points). This might lead to forced unfolding of cohesins and therefore to dockerin (and enzyme) release, resulting in a negative effect on cellulosome activity (Galera-Prat et al., 2020; Galera-Prat et al., 2018). In a study by Galera-Prat et al. (2020), the cellulolytic activity of bound cellulosomes was compared to that of a free complex after the exposure to mechanical stress using magnetic stirring under high agitation conditions. Specifically, monovalent cellulosomes were designed bearing four different single cohesins with known mechanical stability. They were attached to polystyrene microparticles to mimic cell surface attachment. It was found that a cohesin with a low mechanical stability in a scaffoldin's connecting region decreases the cellulosome's activity compared to one with the higher stability. This highlights mechanical stability as an emerging industrial parameter in biotechnology. Specifically, in two-point attachment systems, cohesins between anchoring points should have high mechanical stability to preserve activity of the cellulosome. Notably, these principles may also explain the evolution of natural cellulosome architectures (Galera-Prat et al., 2020).

4 Cell-surface anchorage

For the majority of cellulosomes described so far, attachment to the bacterial cell surface relies on calcium-mediated interactions

between type II dockerin modules of the cellulosomal scaffoldin and cohesin modules of cell-surface proteins anchored in the peptidoglycan layer. Structural and biochemical studies of clostridial cell surface dockerin-cohesin complex revealed two cohesin-binding interfaces with differential specificities (Adams et al., 2006). Type II dockerins of primary scaffoldins from *C. thermocellum* and *C. cellulolyticum* ScaA and CipB feature two distinct cohesin-binding surfaces that recognize different cohesin partners, while in *A. cellulolyticus*, the dual-binding mode of the type II dockerin module provides enhanced flexibility in cohesin recognition and allows greater versatility for efficient cellulosome assembly (Brás et al., 2016).

4.1 Anchoring scaffoldins: sortase- and SLH module-based mechanisms and beyond

Cell surface-anchoring scaffoldins are typically embedded in the peptidoglycan either through sortase-mediated covalent attachment or by noncovalent binding of scaffoldins via S-layer homology (SLH) modules. In the sortase-mediated mechanism, as described for the scaffoldin ScaE from *R. flavefaciens*, a sortase recognition motif (LPXTG) is cleaved by the sortase at the cell surface and the protein is covalently linked to the cell wall (Rincon et al., 2005). This well-characterized process is common in gram-positive bacteria (reviewed in Bhat et al., 2021). The second mechanism involves non-covalent interactions between polysaccharides protruding from the peptidoglycan and SLH modules on the scaffoldins. Herein, cohesin domains are connected to SLH modules through flexible linkers (Zhao et al., 2006). In *C. thermocellum,* four novel anchoring scaffoldins ScaB, ScaC, ScaD and ScaF were identified (Brás et al., 2016). Interestingly, ScaF is the only one that contains an additional module of unknown function (ScaF-X) between the cohesin and SLH modules, with no homologs found in the Protein Data Bank (PDB) database. NMR-based studies suggest that ScaF-X is a well-folded and potentially functional protein module, holding promise as a new element in synthetic biology and biotechnology application, although further research is needed to elucidate its role (Li et al., 2021).

Complex cellulosome producers like *Acetivibrio cellulolyticus* (*A. cellulolyticus*) harbor numerous genes encoding CBM3-SLH fusion proteins (Figure 2), which were proposed to facilitate microbial attachment to the cellulose (Minor et al., 2024; Wang et al., 2022). In addition to sortase recognition and SLH motifs, other domains may play a role in cellulosome attachment to the cell wall. In *B. cellulosolvens* genome, several novel domains within its scaffoldins like PA14 domain in ScaU, cadherin in ScaD, VCBS in ScaH3, and PPC domains in ScaO and ScaP, were recently identified (Zhivin et al., 2017). These domains are typically associated with functions like carbohydrate binding, cell adhesion, and protein–protein interactions (Finn et al., 2017). Their presence suggests potential mechanisms for cell wall anchoring and substrate recognition beyond the classical SLH-based system.

Unlike the more complex cellulosomes, simple scaffoldins such as those found in *R. cellulolyticum* typically lack classical cell wall-binding modules like SLH domains. While they do contain cohesin and dockerin domains necessary for enzyme assembly, their mode of anchoring to the cell surface has remained unclear. Recent findings suggest that the N-terminal CBM3a domain of the primary scaffoldin

CipC can bind to the bacterial surface and may contribute to cellulosome localization (Tao et al., 2022). Additionally, while X2 domains do not directly bind to the cell wall or cellulose in *R. cellulolyticum*, they are thought to influence the spatial organization of the cellulosome. As such, they may modulate interactions with cellulosic substrates or other components of the cell surface (Kosugi et al., 2004; Tao et al., 2022). In other species such as *C. cellulovorans*, X2 domains have been implicated in scaffoldin surface display, especially when present alongside SLH modules (Tarraran et al., 2021). These findings indicate that CBM3a and X2 domains may contribute to alternative modes of surface localization in simple cellulosomes, though further research is needed to clarify their precise roles and potential cooperative functions.

4.2 Cell-free scaffoldins

Most cellulosomes were found to be attached to cell surfaces, but cell-free scaffoldins have also been observed in species like *C. thermocellum*, *Clostridium clariflavum*, and *A. cellulolyticus* (Artzi et al., 2015; Raman et al., 2009; Xu et al., 2016; Zhivin-Nissan et al., 2019). Some secrete cell-free cellulosomes with multi-cohesin scaffoldins and dockerin-tagged GH enzymes to degrade lignocellulosic biomass. Certain species, including *R. sufflavum*, *Clostridium* sp. HBUAS56017 and *Clostridium bornimense*, have been reported to produce scaffoldins without dockerin or cell wall-binding modules, suggesting secretion into the environment (Minor et al., 2024). Cell-free scaffoldins were suggested to play an important role in degradation of remote cellulosic substrates (Zhivin-Nissan et al., 2019). While the potential for free cellulosomes to facilitate cooperative lignocellulose breakdown is intriguing, the regulatory mechanisms governing their release remain unclear. Further research is needed to determine the mechanism(s) that control the secretion of free cellulosomes, and how this regulation might optimize their deployment for synergistic degradation.

4.3 Biotechnological applications of cell surface display systems

Surface-display technologies enable the functional assembly of DCs on non-cellulolytic microbial hosts, generating whole-cell biocatalysts capable of lignocellulosic biomass degradation and consolidated bioprocessing. This strategy integrates enzymatic hydrolysis with downstream metabolic conversion in a single microbial chassis. In food and health-related applications, lactic acid bacteria such as *Lactococcus lactis* and *Lactobacillus plantarum* offer the advantages of GRAS (Generally Recognized As Safe) status and probiotic potential. For instance, scaffoldins from *C. cellulovorans* containing different combinations of X2 and SLH domains were successfully displayed on the surface of *Lactococcus lactis* (Tarraran et al., 2021). This suggests a shared binding mechanism mediated by components of the thick peptidoglycan layer (Tarraran et al., 2021), which is known to be interwoven with proteins and glycopolymers such as teichoic and lipoteichoic acids (Desvaux et al., 2006). Similarly, DCs have been assembled on the surface of *Lactobacillus plantarum* (Stern et al., 2018). *Bacillus subtilis*, another GRAS-designated and industrially robust Gram-positive host, was used to

anchor a trifunctional minicellulosome via a cell wall-binding domain derived from its endogenous hydrolase LytE (You et al., 2012). Yeasts, particularly suited for biofuel production, have also been extensively explored (reviewed in Lamote et al., 2023; Li et al., 2024). For example, in *Saccharomyces cerevisiae*, a tetravalent DC displayed via adaptor scaffoldins led to a several-fold enhancement in cellulose hydrolysis and ethanol yield relative to free enzymes (Tsai et al., 2013). In a study by Tang et al. (2018), a yeast surface display system was developed in which cellulosomal components were anchored via disulfide bonds to a yeast cell wall protein, significantly improving display efficiency in *S. cerevisiae* (Tang et al., 2018). Recently, *Pichia pastoris* has been engineered as a whole-cell biocatalyst for carboxymethyl cellulose to ethanol conversion. Herein, an ultra-high-affinity IM7/CL7 system from *E. coli* was used to display minicellulosomes assembled *in vitro* from endoglucanase, exoglucanase, β-glucosidase and CBM. To date, this is the first example of engineered yeast enabling efficient and direct conversion of carboxymethyl cellulose to ethanol with titers reaching 5.1 g/L (Dong et al., 2020). Further, probiotic yeast *Kluyveromyces marxianus* was engineered to produce and display the largest cellulosome complex that can accommodate up to 63 enzymes. Herein, the surface attachment of the designed scaffoldin was achieved by replacing its SLH domain with the glycosylphosphatidylinositol (*Sc*GPI), a cell-surface protein from *S. cerevisiae* (Anandharaj et al., 2020). Finally, Gram-negative bacteria have also been adapted for surface display, despite structural limitations posed by their outer membrane. For example, a cohesin-dockerin pair from *Archaeoglobus fulgidus* was employed to attach a dimeric α-neoagarobiose hydrolase to the surface of *E. coli* cells (Ko et al., 2021). Altogether, these platforms extend cellulosome display across diverse microbial hosts and hold strong promise for industrial biotechnology, sustainable bioenergy production, and the development of next-generation probiotics.

5 Catalytic subunits and the functional interplay of cellulosomal components

One of the key advantages of cellulosomes over free enzymes lies in the close spatial proximity of their enzymatic components. This proximity enables coordinated and efficient degradation of cellulose and other plant cell wall polysaccharides. It also facilitates substrate channeling, minimizes product inhibition and enhances overall efficiency. Compared to free enzymes, cellulosomes demonstrate superior substrate conversion due to the synergistic action of their enzyme assemblies (Artzi et al., 2017; Vazana et al., 2010).

5.1 Catalytic repertoire of cellulosomes

Cellulosomes typically incorporate a wide array of carbohydrate-active enzymes (CAZymes) that contribute to the degradation of plant cell walls (Table 1). These include: (1) GHs—notably from families such as GH5, GH9, GH10, GH11, GH26, and GH48—which hydrolyze glycosidic bonds in cellulose, hemicellulose, and other glycans (e.g., xylanases, mannanases, xyloglucanases, glucanases), (2) carbohydrate esterases (CEs)—which remove ester-linked substituents (e.g., acetyl or feruloyl groups) that can hinder GH

access to their substrates, (3) polysaccharide lyases, that cleave polysaccharides through non-hydrolytic mechanisms, such as β-elimination, facilitating the degradation of pectins and other acidic polysaccharides, and in some cases, (4) non-CAZy accessory proteins, such as proteases, protease inhibitors, and expansin-like proteins, which may modulate enzyme accessibility or cell wall structure (Artzi et al., 2017; Gharechahi et al., 2023; Lombard et al., 2025). Altogether, this complex and modular enzymatic architecture underlies the remarkable versatility and efficiency of the cellulosomal system.

Efficient hydrolysis of β-glycosidic bonds in cellulose relies on the coordinated, synergistic action of multimodular enzymes with two types of catalytic domains: endoglucanases and exoglucanases. Endoglucanases cleave internal β-1,4-bonds within the cellulose polymer generating shorter oligosaccharides. Exoglucanases cleave the chains progressively from either the reducing or non-reducing end, releasing cellobiose or cellotetraose units. β-glucosidases then hydrolyze cellobiose into glucose, alleviating product inhibition and making glucose available for microbial metabolism (reviewed in Datta, 2024; Ranjan et al., 2023; Sidar et al., 2020). This synergy is quantitatively reflected in a degree of synergistic effect often exceeding the sum of individual enzyme activities, with optimal enzyme ratios, substrate properties, and intermolecular proximity further enhancing hydrolysis efficiency.

The complexity of hemicellulose requires a diverse set of hemicellulases. Endoxylanase and xylosidases degrade the xylan backbone, while accessory enzymes such as α-L-arabinofuranosidases remove arabinose side chains from arabinoxylans. Acetyl xylan esterases remove acetyl groups from xylan chains, and ferulic acid esterases cleave ester linkages between hemicellulose and lignin. Additionally, α-glucuronidases remove glucuronic acid side chains from glucuronoxylans, and mannanases and β-mannosidases facilitate the degradation of mannan rich hemicelluloses.

Alongside (hemi)cellulose and pectins, lignin represents a key structural component of plant cell walls. Unlike cellulolytic enzymes, lignin-degrading oxidative enzymes, such as laccases, lignin peroxidases and manganese peroxidases are exclusively produced by aerobic organisms like white rot fungi and certain bacteria and have therefore not been identified within native bacterial cellulosomes (reviewed in Artzi et al., 2017; Gharechahi et al., 2023; Wang et al., 2019). Nevertheless, advances in synthetic biology have enabled the engineering of "designer cellulosomes" that incorporate lignin modifying enzymes (Davidi et al., 2016).

A diverse array of proteins beyond traditional carbohydrate-active enzymes has been identified within cellulosomes, contributing to enzymatic synergy and functional regulation. Notably, expansin-like proteins have been discovered in the cellulosome system of *Clostridium clariflavum*. Two such proteins, CclEXL1 (Clocl_1862) and CclEXL2 (Clocl_1298), contain type I dockerin modules, suggesting their incorporation into the cellulosome complex. Proteomic analyses have confirmed the expression of CclEXL1 under specific growth conditions, with its dockerin module exhibiting selective binding to type I cohesins, particularly the cohesin of scaffoldin ScaG (Artzi et al., 2016). Functionally, CclEXL1 binds preferentially to microcrystalline cellulose and demonstrates a pronounced loosening effect on cellulose fibers. This enhances the enzymatic hydrolysis of cellulose by both native cellulosomes and individual cellulases such as GH48 and GH9 (Artzi et al., 2016). The

TABLE 1 Catalytic subunits identified as parts of native bacterial cellulosomes.

Substrate	Enzyme function	EC number	Family examples	References
Cellulose	Endoglucanase (endo-1,4-β-glucanase)	EC 3.2.1.4	GH5, GH8, GH9, GH24, GH44	Aburaya et al. (2019); Blouzard et al. (2010); Esaka et al. (2015); Gold and Martin (2007); Hirano et al. (2016); Matsui et al. (2013); Morisaka et al. (2012); Raman et al. (2009); Vodovnik et al. (2013); Zhivin-Nissan et al. (2019)
	Exoglucanase (cellobiohydrolase)	EC 3.2.1.91	GH9, GH48	
Hemicellulose	Endo-β-1,4-xylanase	EC 3.2.1.8	GH5, GH10, GH11, GH30	Aburaya et al. (2019); Blouzard et al. (2010); Esaka et al. (2015); Gold and Martin (2007); Han et al. (2004); Hirano et al. (2016); Matsui et al. (2013); Morisaka et al. (2012); Murashima et al. (2002); Raman et al. (2009); Vodovnik et al. (2013); Zhivin-Nissan et al. (2019)
	Mannanase	EC 3.2.1.78	GH5, GH26	
	β-xylosidase	EC 3.2.1.37	GH3, GH43	
	α-arabinofuranosidase	EC 3.2.1.55	GH43, GH51, GH62	
	α-glucuronidase	EC 3.2.1.139	GH67, GH115	
	Acetyl xylan esterase	EC 3.1.1.72	CE1, CE3, CE4, CE6, CE12	
	Feruloyl esterase	EC 3.1.1.73	CE1, CE6, CE7	
	Xyloglucanase	EC 3.2.1.151	GH74	
Pectin	Rhamnogalacturonan hydrolase	EC 3.2.1.171	GH28	Aburaya et al. (2019); Blouzard et al. (2010); Gold and Martin (2007); Hirano et al. (2016); Morisaka et al. (2012); Pagès et al. (2003); Raman et al. (2009)
	Polygalacturonase	EC 3.2.1.15	GH28	
	Endo-1,4-β-galactanase	EC 3.2.1.89	GH53	
	Rhamnogalacturonan lyase	EC 4.2.2.23	PL4, PL11	
	Pectate lyase	EC 4.2.2.2	PL1, PL9	
	Rhamnogalacturonan acetyl esterase	EC 3.1.1.86	CE12	
β-Glucans	Lichenase (β-1,3-1,4-glucanase)	EC 3.2.1.73	GH16	Gold and Martin (2007); Hirano et al. (2016)
Proteins / peptides	Protease/peptidase	EC 3.4.21.-EC 3.4.22.-	Peptidase S8 (subtilisin-like) Cysteine peptidase	Levy-Assaraf et al. (2013); Morisaka et al. (2012)
Chitin	Chitinase	EC 3.2.1.14	GH18	Hirano et al. (2016); Zverlov et al. (2002)

biological role of bacterial expansins remains under investigation (de Sandozequi et al., 2022), but their incorporation into cellulosomes suggests they play a supportive role in optimizing substrate accessibility (Cosgrove, 2017). Furthermore, both free and dockerin-containing expansin-like proteins from *Bacillus subtilis* and *C. clariflavum* have been shown to promote cellulose degradation by native cellulosomes. This synergistic effect is amplified when these proteins are integrated into trivalent DCs, further enhancing their cellulolytic performance (Artzi et al., 2016; Chen C, et al., 2016).

Further on, a protease domain in double dockerin-containing proteins was identified in various cellulosome-producing bacterial species (Chen et al., 2024). It was proposed that it may facilitate the maturation, activation, and turnover of cellulosomal components, ensuring optimal enzyme function. In *R. flavefaciens*, putative cysteine peptidase that binds to the surface-anchoring ScaE cohesin via its X-dockerin modular dyad was identified (Levy-Assaraf et al., 2013). In addition, the structure of putative S8 protease with a tandem bimodular double-dockerin from *C. thermocellum* was determined. These dockerins were able to bind to various scaffoldins from different species, implying that rather than being directly involved in biomass degradation, cellulosomal peptidases may play a role in cellulosomal posttranslational processing or cellular nitrogen recycling from secreted proteins (Chen et al., 2024).

Interestingly, cellulosome-localized protease inhibitors were also identified in *Clostridia*. In *C. thermocellum*, serpins were suggested to play a protective role in cellulosomes by inhibiting subtilisin-like proteases, safeguarding the integrity of the complex (Kang et al., 2006;

Ó Cuív et al., 2013). Similarly, in cellulosomes of *C. cellulolyticum*, a dockerin-containing protease inhibitor Dpi was discovered that protects key cellulosomal cellulases from proteolysis (Xu et al., 2014). On the other hand, proteolytic mechanisms may allow cellulosome-producing bacteria to recruit specialized enzymes from other organisms in the same ecosystem. This saves energy and resources and could also regulate cellulosome assembly in different species (Chen et al., 2024; Haimovitz et al., 2008). However, the exact roles of these proteases and their specific locations within the cellulosome need further experimental validation.

5.2 Spatial and structural principles guiding designer cellulosome construction

A deep understanding of cellulosomal architecture is essential for designing efficient, tailored DCs. The diverse enzyme composition in cellulosomes enables effective plant polysaccharide degradation, with the cellulosomal architecture minimizing the diffusion and enhancing the uptake of the degradation products (Arora et al., 2015; Vodovnik and Marinšek-Logar, 2010). For enhanced synergistic activity, the position of the enzymes in the scaffoldin, enzymes functionality and flexibility of scaffoldin and enzyme-dockerin linkers are of significant importance. Optimal distances between the active sites of different enzymes on the protein scaffold enable effective substrate channeling, promoting sequential reactions over diffusion (Deng et al., 2020). In a study by Stern et al. (2015), the optimal position of recombinant

processive endoglucanase Cel9A from *T. fusca* was determined relative to two additional enzymes: the *C. thermocellum* exoglucanase Cel48A and the *C. cellulolyticum* endoglucanase Cel5A. The optimal order was determined as follows: processive endoglucanase Cel9A (GH9), exoglucanase Cel48A (GH48) and endoglucanase Cel5A (GH5). A GH9 endoglucanase proximity to the GH5 endoglucanase negatively affected overall degradation efficiency. This is in accordance with the fact that these two enzymes possess similar functionalities. Thus, their immediate proximity may cause unproductive competition between the two catalytic sites of the two functionally similar enzymes (Stern et al., 2015). This is in line with the results of a similar study on *C. cellulolyticum*, where a successful enzyme combination was obtained using a chimeric scaffoldin in which the GH9 processive endoglucanase was placed on the scaffoldin together with the GH48 exoglucanase and away from the GH5 endoglucanase (i.e., GH5-GH48-GH9; Fierobe et al., 2005). To determine if docking enzyme size affects DC activity, Vanderstraeten et al. (2022a) used a designed cellulosome complex with three enzymes: a smaller mannanase (47 kDa), a galactosidase (54 kDa), and a larger mannosidase (104 kDa). Placing the larger mannosidase between the smaller enzymes did not reduce mannose monomer production (Vanderstraeten et al., 2022a). Similarly, Vazana et al. (2013) found no preferred modular arrangement when incorporating three *C. thermocellum* enzymes of varying sizes—an exoglucanase Cel48S (82 kDa), an endoglucanase Cel8A (52 kDa), and an endoglucanase Cel9K (99 kDa)—into a DC (Vazana et al., 2013). These results suggest that docking enzyme size is not a primary factor influencing substrate degradation efficiency of the complex. The position of the CBM of the scaffoldin also proved to be an important factor in cellulosome functionality. In case of the cellulosomes incorporating recombinant enzymes originating from *T. fusca*, it was shown to be optimal when placed close to the processive endoglucanase and at the extremity of the scaffoldin (either at the N- or at the C- terminus), allowing more freedom in movement (Hammel et al., 2004; Stern et al., 2015).

In order to adapt to the complexity of the substrate, cellulosome requires maximal flexibility, which is mainly provided by the intermodular linkers of scaffoldin subunits. Structural analyses, including crystal structures and SAXS studies (Currie et al., 2012; Hammel et al., 2004; Molinier et al., 2011; Noach et al., 2009), have revealed that these linkers confer the necessary conformational adaptability to scaffoldins, facilitating efficient substrate degradation. Vazana et al. (2013) investigated the impact of intermodular linker length on DC activity. Scaffoldins with no linker, short linkers (5 aminoacids) and long native linkers (27–35 amino acids) were compared, and DCs with the long linker were found to achieve higher levels of activity on both pure and complex substrates. This enhancement was attributed to increased flexibility and optimal spatial arrangement of catalytic modules (Vazana et al., 2013). In addition, ScaH, a *R. flavefaciens* adaptor scaffoldin, can work as a variable length spacer to avoid clashes between cellulosomal units on the bacterial surface, thereby promoting efficient assembly of complex cellulosomal structures (Duarte et al., 2023). In modular GHs, the length and flexibility of the linker between the CBM and the catalytic modules were found to be critical for the catalysis. Studies on GH5 family cellulase 5A from *B. subtilis* (Ruiz et al., 2016) and GH45 family endoglucanase from *Rhizopus stolonifera* (Tang et al., 2015) have demonstrated that deviations from optimal linker properties can impair enzyme function, underscoring the importance of precise

inter-domain spacing and flexibility. However, the influence of linker properties can be enzyme-specific. Caspi et al. (2009) converted the modular *T. fusca* cellulase Cel5A to a cellulosomal configuration by replacing its CBM with a dockerin module from either *C. thermocellum* or *R. flavefaciens*. They found that varying the linker length between the catalytic domain and the dockerin had minimal impact on enzymatic activity (Caspi et al., 2009). This suggests that some enzymes are less sensitive to linker variations. Kahn et al. (2019) explored the effects of introducing long linkers between catalytic domains and dockerin modules in *Caldicellulosiruptor bescii* enzymes. They found that a longer linker enhanced GH9 activity but had no effect on GH48-containing complexes. The differential impact was attributed to the absence of glycosylation in recombinant linkers, which, unlike in native enzymes (Chung et al., 2015), may affect flexibility and protection from proteolysis, influencing enzyme stability and function (Kahn et al., 2019).

Recently, a trivalent DCs carrying mannanase, mannosidase and galactosidase was constructed and optimized for galactomannan degradation. By testing linkers that differ in flexibility and length on the three enzymes, it was found that direct enzyme-dockerin fusion with no linker significantly decreased the enzyme activity. In addition, the optimal linker type and length was found to be enzyme-dependent (Vanderstraeten et al., 2022b). Furthermore, a lytic polysaccharide monooxygenase (LPMO) catalytic domain combined with a linker region and CBM domain was used as a model to investigate the effect of the linker on enzyme activity. Herein, the linker region was found to be important for protein thermostability, binding of and activity toward recalcitrant polysaccharide substrate (Srivastava et al., 2022).

Altogether, these results highlight the importance of architectural parameters like linker flexibility and length, positioning of the catalytic units within the scaffoldin, and enzymes functionality in enhancing the performance of cellulases and/or cellulosomes. Understanding how these factors influence cellulosomal function provides a crucial foundation for the rational design of DCs. By strategically engineering these features, it is possible to optimize enzyme synergy, improve substrate accessibility, and increase catalytic efficiency. Insights gained from natural systems and synthetic assemblies enable the construction of tailored DCs suited for specific substrates or industrial applications.

6 Advances in designer cellulosome construction

Enzymatic efficiency in the degradation of complex biomass can be significantly enhanced by incorporating heterologous enzymes or cellulosomal components into DC assemblies. These engineered complexes enable precise control over enzyme composition, spatial organization, and overall stability, thereby achieving catalytic performances that surpass those of native systems. In addition to the strategies discussed above, and summarized in Figure 3, several notable approaches have been developed to further enhance the enzymatic efficiency of DCs. One effective strategy involves increasing the enzyme load within a single cellulosomal complex. For instance, in *Acetivibrio cellulolyticus*, the adaptor scaffoldin ScaB mediates the incorporation of up to four ScaA scaffoldins into a single cellulosome, substantially expanding the enzymatic repertoire contained within one assembly (Brás et al., 2016). Another innovative approach was demonstrated by
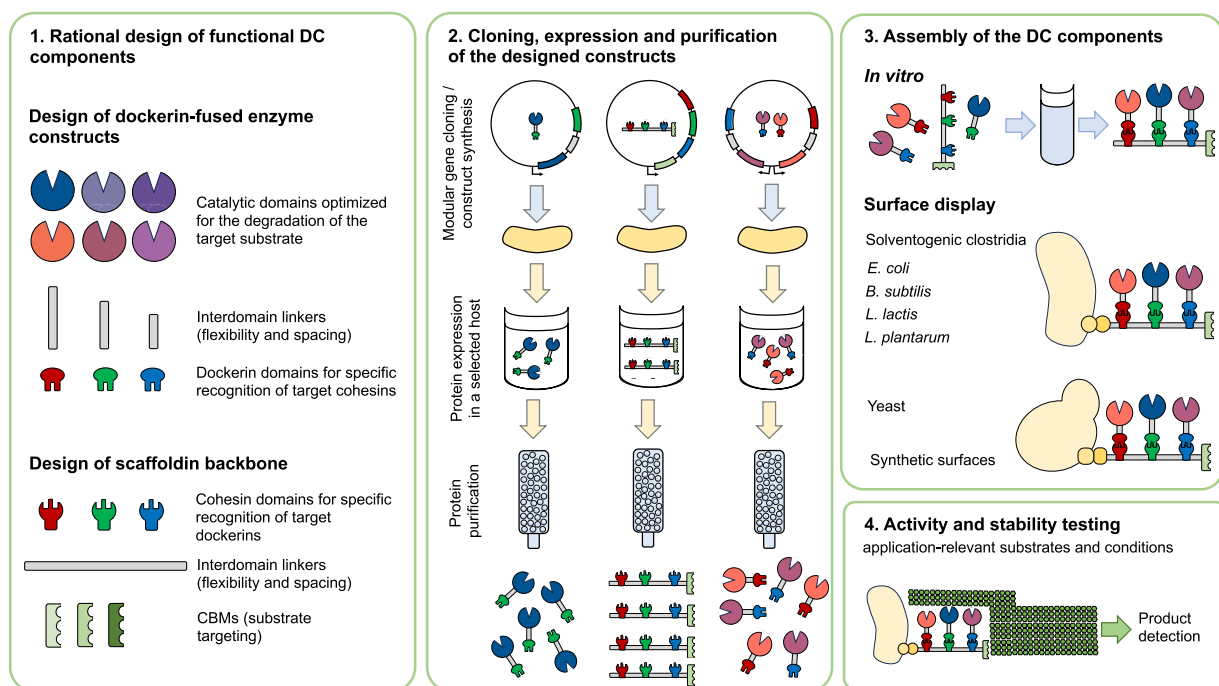
FIGURE 3
Overview of designer cellulosome construction and assembly. A modular workflow for the design, production, and assembly of a (synthetic) designer cellulosome. Scaffoldins with species-specific cohesins and CBMs are combined with dockerin-tagged catalytic domains to form multifunctional complexes. Following expression and purification, components are assembled either *in vitro* or via surface display in industrially-relevant strains. The performance of the resulting complexes is than assessed under application-relevant conditions. CBM: Carbohydrate-binding module, DC- designer cellulosome.

Gilmore et al. (2020), who engineered chimeric enzymes by fusing fungal dockerin domains with GH catalytic modules from the hyperthermophilic bacterium *Thermotoga maritima*. These chimeric enzymes retained activity at elevated temperatures (70–90°C) and displayed stability within the 50–70°C range, conditions relevant to industrial biomass conversion. This thermostability reduces enzyme dosage requirements and enhances process efficiency (Gilmore et al., 2020). Furthermore, Kahn et al. (2019) recently developed a hyperthermostable DC system, active at 75°C. Enzymes from *Caldicellulosiruptor bescii*, a cellulolytic hyperthermophilic bacterium, were adapted to the cellulosomal mode by attaching dockerins that matched the thermostable cohesins within a chimeric scaffoldin. Three cohesin-dockerin pairs from three different thermophiles (*C. thermocellum, C. clariflavum* and hyperthermophilic archaeon *Archaeoglobus fulgidus*) were tested and remained stable at 75°C for 72 h. The resulting cellulosome complex showed superior enzymatic activity on microcrystalline cellulose at 75°C, outperforming *C. thermocellum*-based DC and the native *C. thermocellum* cellulosome (Kahn et al., 2019; Stern et al., 2015). Glycosylation of cellulosomal components has also been shown to improve stability and efficiency in cellulose hydrolysis. Inclusion of the glycosylated components yielded an active cellulosome system that exhibited long-term stability at higher temperatures and significantly improved activity compared to the enzymatic components alone (Kahn et al., 2020). These properties are particularly desirable for industrial biomass degradation, where stability and efficiency under harsh process conditions are essential.

Further taking advantage of the modularity of DCs, Vanderstraeten et al. (2022b) successfully reconstituted the free xyloglucan degradation system from *Cellvibrio japonicus* into a cellulosomal format. They engineered and purified constructs combining four catalytic domains appended with dockerin modules from diverse origins and assembled these onto a tetravalent scaffoldin. The resulting "xyloglucanosome" efficiently hydrolyzed xyloglucan into oligosaccharides, galactose, xylose, and glucose, demonstrating the functional integration of complex enzyme systems into DCs (Vanderstraeten et al., 2022b).

When designing novel docking enzymes, the CBM domain of the selected multimodular enzyme is often replaced with a dockerin. This strategy preserves the construct's natural order of the modules, and, consequently, the optimal catalytic performance of the enzymatic module (Vanderstraeten et al., 2022a). For single-module enzymes, the optimal dockerin position must be determined empirically, as it is difficult to predict. In the native proteins, dockerins are usually positioned C-terminally from enzymatic domain. Nevertheless, they can also be located N-terminally (Mondal et al., 2021; Raghothama et al., 2001) or even internally, between the two enzymatic modules (Grépinet et al., 1988). In DCs, dockerins were found to be functional when placed either N- or C- terminally (Caspi et al., 2009; Caspi et al., 2008; Vanderstraeten et al., 2022a).

The discovery of novel catalytic modules continues to expand the cellulosomal toolbox. For instance, the endoglucanase Rf GH16_21 from *R. flavefaciens* was found to possess a N-terminal dockerin module and three tandem GH16 domains. The enzyme was shown to exhibit notable thermostability and exclusively *β*-1,3–1,4-endoglucanase activity. This activity is likely essential for coordinating the *R. flavefaciens* cellulosome's degradation of β-1,3–1,4-glucans (Ahmed et al., 2024; Mondal et al., 2021). Similarly, a recent proteomic analysis of the (Pseudo)*Bacteroides cellulosolvens* secretome revealed

expression of 24 scaffoldins and 166 dockerin-bearing components (predominantly enzymes), as well as free enzymatic subunits (Zhivin-Nissan et al., 2019), highlighting a rich resource of modular parts for biotechnological applications.

Taken together, these advances show that DCs provide a modular and scalable platform for industrial biomass conversion. Their adaptability—through thermostable, glycosylated, and chimeric components—supports efficient performance under demanding conditions. Combined with improved recombinant expression and assembly methods, these features enable their practical application in large-scale bioprocessing.

7 Regulation and adaptive plasticity of cellulosomal activity in response to substrate availability and environmental conditions

Cellulosomes exhibit a remarkable ability to adapt their composition in response to environmental conditions. Carbon sources, pH, temperature, ionic strength, chemical compounds (like inhibitors or enhancers) and chemical or biological pretreatment of the substrate can all affect their activity (reviewed in Wang et al., 2019 and Datta, 2024).

The composition and activity of cellulosomes can vary, allowing microorganisms to assemble different enzyme complexes depending on the available carbon source (Figure 4). Upregulation of genes encoding CAZymes, cell motility, chemotaxis, quorum sensing, and some GH regulation was recently detected in a study on

C. thermocellum B8 during growth on sugarcane bagasse and straw versus purified microcrystalline cellulose (de Camargo et al., 2023). Differential expression of gene clusters targeting different carbohydrate sources, i.e., cellulose or hemicellulose, has also been observed in other bacteria, for example *C. cellulolyticum* (Blouzard et al., 2010; Raman et al., 2009).

Differential composition of catalytic and structural subunits was also detected within cellulosome complexes isolated from *C. thermocellum* grown on different carbon sources. Microcrystalline cellulose- and glucose-derived cellulosome samples exhibited higher endoglucanase-to-exoglucanase ratios and greater catalytic subunit-per-scaffoldin ratios than those derived from lignocellulose (Yoav et al., 2017). A proteome-wide study investigating *B. cellulosolvens* cellulosomes revealed distinct cellulolytic enzyme profiles depending on the carbon source. When grown on microcrystalline cellulose, the bacteria produced cellulosomes displaying the highest expression of the structural and enzymatic subunits. These cellulosomes also displayed the highest degradation activity compared to the ones derived from the bacteria grown on four other cellulosic and hemicellulosic substrates (Zhivin-Nissan et al., 2019). Furthermore, research on *R. cellulolyticum* has revealed that the enzymatic subunits of cellulosomes can readily be exchanged even after assembly, indicating dynamics and flexibility. The strength of dockerin-cohesin interaction is based on specific residues in the dockerin sequence (described in more detail in chapter 2), which can either lock the enzymes in place or allow for reversible binding. This adaptability allows cellulosomes to be refunctionalized over time, enabling them to adjust their catalytic capabilities as needed, a feature that highlights their potential for dynamic environmental response (Borne et al., 2020).
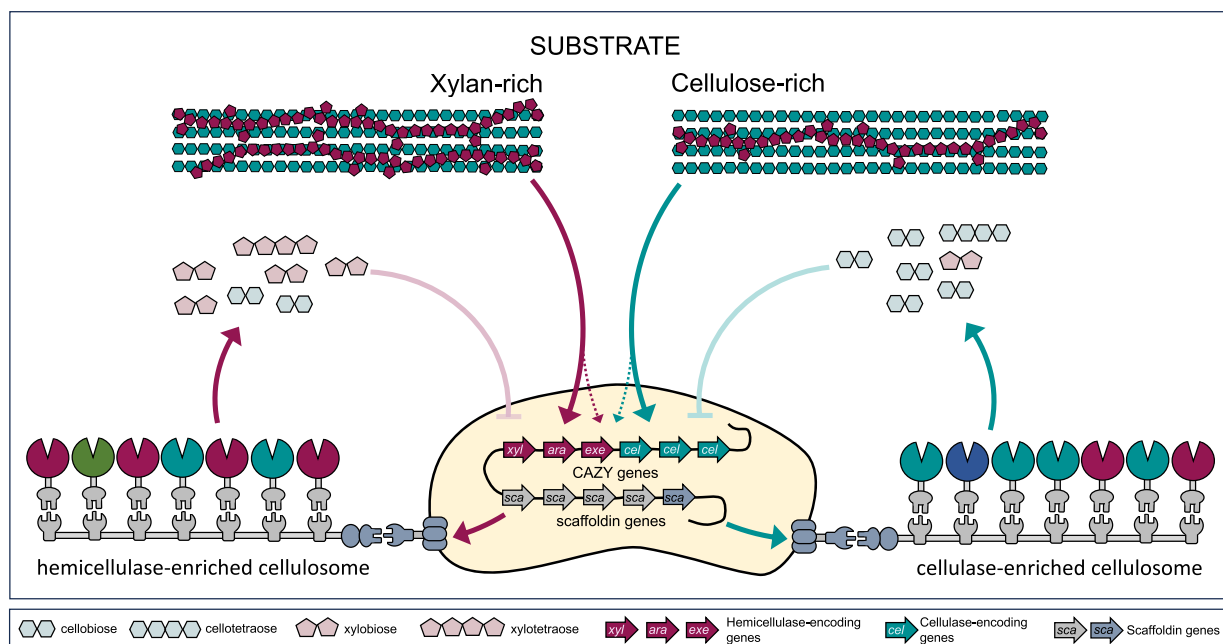

FIGURE 4
Substrate-mediated regulation of cellulosomal gene expression and complex composition. Adaptive remodeling of cellulosome composition and activity is driven by key regulatory mechanisms, including differential gene expression, enzyme exchangeability, and feedback inhibition. Substrate sensing leads to differential expression of hemicellulase (*xyl, ara, exe*) genes (e.g., endo-β-1,4-xylanase, α-L-arabinofuranosidase, β-xylosidase, acetyl xylan esterase, feruloyl esterase), cellulase (*cel*) genes (endo- and exoglucanases), and scaffoldin (*sca*) genes. On the other hand, accumulation of higher concentration of the degradation products can lead to feedback inhibition and decreased complex activity.

Feedback inhibition in cellulosomal activity occurs when the accumulated end products, such as cellobiose or glucose. These products inhibit the activity of cellulases and other enzymes within the cellulosome, thereby reducing biomass degradation efficiency. Wild-type cellulosomes and their producing strains often fall short of industrial requirements due to such inhibition (Lamote et al., 2023). To address this challenge, strains like *C. thermocellum* DSM1313 have been genetically modified to express extracellular β-glucosidase, which helps to hydrolyze cellobiose and alleviate feedback inhibition (Qi et al., 2021; Zhang et al., 2017). Despite these advancements, challenges remain. For example, the key exoglucanase Cel48S is known to be inhibited by soluble xylan present in lignocellulosic hydrolysates, indicating the importance of timely xylan removal during saccharification (Chen et al., 2022). Strategies to overcome feedback inhibition include engineering enzymes with reduced sensitivity to end products (Atreya et al., 2016), introducing alternative pathways for product removal, or optimizing reaction conditions to minimize product accumulation (Qi et al., 2021).

The modular architecture of cellulosomes, the functional exchangeability of their enzymatic components, and the carbon source–responsive regulation of cellulosomal genes together provide a strong basis for engineering dynamic and environmentally adaptive systems. Using an integrative strategy that leverages structural biology, protein engineering, computational and synthetic biology, cellulosomes can be engineered to dynamically adapt their composition in response to environmental cues such as pH, temperature, substrate type, mechanical stress or inhibitors. A key strategy involves the engineering of cohesin–dockerin pairs with enhanced environmental tolerance (Bugada et al., 2025; Duarte et al., 2023; Galera-Prat et al., 2020; Galera-Prat et al., 2018; Kahn et al., 2019; Moraïs et al., 2016; Vera et al., 2021) including dockerins with dual-binding modes that enable reversible enzyme exchange (Duarte et al., 2021; Kahn et al., 2019). Combining cohesin–dockerin pairs from different species allows precise control of enzyme composition and spatial organization (Borne et al., 2013; Vanderstraeten et al., 2022a). Environmental responsiveness can be built in through pH-dependent binding switches, as demonstrated in *C. acetobutylicum* (Yao et al., 2020), while electrostatic tuning offers additional control over dockerin affinity (Chen et al., 2019; Duarte et al., 2023). Scaffoldins can be modularly designed with tunable linker lengths (Bomble et al., 2011; Dorival et al., 2022) and multidockerin domains for conditional and/or sequential enzyme recruitment (Chen et al., 2024; Chen et al., 2019). Although not yet widely realized, design concepts envision cellulosomes with inducible or environmental-sensing modules that could dynamically control assembly (e.g., via riboswitch-linked dockerins or stimulus-responsive cohesins). Adaptation to inhibitors can be engineered by including enzymes that remove inhibitors (Cao et al., 2018; Qi et al., 2021; Zhang et al., 2017) or enzymes with improved tolerance to inhibitors like furfural and 5-HMF (de Andrades et al., 2024). Additionally, CBMs with dual substrate specificity and putative regulatory functions can act as substrate sensors, potentially influencing enzyme targeting and recruitment based on substrate composition (Boraston et al., 2004). At the genetic level, synthetic regulatory elements such as modular promoters, inducible operons, and riboswitches offer strategies for conditional expression of scaffoldins and enzyme components, enabling dynamic adjustment of catalytic activity in response to

environmental signals (e.g., nutrient status, pH, inhibitors; see Müller et al., 2025; Topp et al., 2010). Furthermore, computational modeling and high-throughput screening can facilitate the rational design, prediction, and optimization of these complex multienzyme assemblies (Bomble et al., 2011; Bozkurt et al., 2025; Chen M, et al., 2016; Dorival et al., 2022; Slutzki et al., 2015; Vanderstraeten et al., 2022a). Together, these strategies lay the foundation for construction of environmentally responsive, high-performance cellulosomes tailored for sustainable and flexible biomass processing.

8 Ultrastructural studies of cellulosome complexes

Cellulosome is heterogenous and dynamic in nature, which is reflected in its overall structure. The complex can assume different forms. Transmission electron microscopy studies revealed that *C. thermocellum* cellulosomes at early stages of growth appear compact, but take on a more relaxed conformation during the later stages of cultivation (Dorival et al., 2022; Mayer et al., 1987; Tatli et al., 2022). Recent research has shown that upon binding to cellulose, cellulosome structure changes to an elongated, even filamentous shape and morphs dynamically at below 1 min time scale according to requirements of the substrate surface (Eibinger et al., 2020). Similarly, multiple conformations were observed in solution as revealed by small angle X-ray scattering (SAXS) and molecular modeling analyses. It was observed that compact forms maintain structural integrity, while extended forms stabilize interactions with solid substrates (Dorival et al., 2022). To date, no crystallographic structure of an entire cellulosome has been successfully solved; cellulosomal unstructured linkers, scaffoldin flexibility, enzyme heterogeneity, and their glycosylation in most species and the dual mode of binding are probably the reasons why. Thus, researchers must resort to alternative methods and complementary approaches to understand the structural basis for their high efficiency.

Early microscopic studies, SAXS and cryo-electron microscopy (cryo-EM) have been used to study cellulosomal components under near-*in vivo* conditions. These studies revealed the flexibility of cellulosomes to adopt tight or loose conformations (Mayer et al., 1987; Smith and Bayer, 2013). Using a "dissect and build" approach, researchers reconstituted 75% of the scaffoldin CipA from *C. thermocellum*. They found that it adopts compact yet flexible conformations with catalytic domains projecting in alternating directions (Bule et al., 2018a; Currie et al., 2012; García-Alvarez et al., 2011; Smith and Bayer, 2013). These findings highlight an efficient spatial organization within the cellulosome, balancing compactness and flexibility. Interestingly, SAXS and biochemical studies indicated that length and composition of inter-cohesins linkers have limited or no impact on the synergy and activity of cellulosomal cellulases (Molinier et al., 2011). It was also observed that loading scaffoldins with enzymes influences the flexibility of linker regions. Specifically, the more enzymes added, the more compact the structure becomes (Dorival et al., 2022).

Real-time movement of the *C. thermocellum* cellulosome complex on a cellulosome fiber (Figure 5) was investigated using atomic force microscopy (AFM) in two recent studies. In the first one, imaging of dynamic structural changes of the cellulosomes interacting with its

crystalline cellulose substrate revealed that cellulosomes use a distinct, microfibril cutting and shortening mechanism, a "sit-and-dig" mode, to degrade crystalline cellulose (Eibinger et al., 2020). This is in contrast to a "slide-and-peel" mode of free cellulases (surface ablation; Eibinger et al., 2020; Eibinger et al., 2017; Igarashi et al., 2011; Payne et al., 2015; Resch et al., 2013). These studies also provide a mechanistic explanation for cooperativity observed between cellulosomes and cellulases in cellulose degradation (Eibinger et al., 2020; Resch et al., 2014; Resch et al., 2013). They also explain why in a complex biomass like the plant cell wall, the relatively small cellulases can penetrate into the fibrillar matrix of cell walls, whereas cellulosomes remain primarily on the surface due to their large size (Ding et al., 2012). In the second AFM study, it was discovered that on rough cellulose surfaces, *C. thermocellum* cellulosomes exhibit random conformational changes, accommodating the local nanomorphologies during vertical substrate degradation (D-type). On smooth surfaces, directional movement (W- and S-type moving) depended on the mechanochemical coupling of catalysis and motion via processive exocellulase Cel48S. Both W- and S-type moving cellulosomes functioned as molecular motors using a Brownian ratchet mechanism, in which random motions were directed by Cel48S-mediated cellulose hydrolysis. The "burnt bridge" model explains this movement: as the cellulosome advances, it degrades its track, preventing backward diffusion. S-type cellulosomes achieved multilayer degradation through synergy between Cel48S exocellulases and endocellulases in their aligned-elongated conformation. In contrast, W-type cellulosomes degraded single layers due to inefficient endo-exo synergy, likely caused by the inaccessible local nanostructure of crystalline cellulose (Zajki-Zechmeister et al., 2024).

Recent nanoscale structural investigations of intact *C. thermocellum* cells provided critical insight into the organization and regulation of cellulosomes at near-physiological conditions (Tatli et al., 2022). Using cryo-EM, cryo-ET, and confocal immunofluorescence, the study found that cellulosomes form a distinct, uniform layer surrounding the bacterial cell at constant distance (~64 nm) from the cell wall (Figure 6). This "cellulosome capsule" physically encases and engages cellulose-degradation intermediates and likely provides proximity between the cellulosomal enzymatic machinery and the substrate. Importantly, the study also uncovered significant phenotypic heterogeneity within the *C. thermocellum* cell population in terms of cellulosome density. Two distinct phenotypes were identified within the same environment and under the same conditions: cells with high-density and cells with low-density cellulosome coverage (Figure 6). Under the stationary phase conditions, the majority of the cells exhibited low-density phenotype, while upon transition into a cellulose-only environment, the proportion of high-density cells markedly increased. These findings reflect the dynamic response linked to the soluble sugar availability. Furthermore, in the presence of soluble sugars (5 mM glucose or cellobiose), the expression of key cellulosomal components, such as the exocellulase Cel48S, was downregulated, suggesting a feedback inhibition of cellulosome production by products from the cellulose-degradation process. Even further, the presence of both phenotypes in the population and changes in their ratio also suggest a division of labor within the bacterial population. Herein, a bet-hedging strategy of costly cellulosome expression even when soluble sugars are available keeps the population prepared for future cellulose-rich environments (Tatli et al., 2022).

Altogether, these findings highlight how the cellulosomal structural flexibility enables functional adaptability. Compact conformations support enzyme clustering and structural stability, while extended or motile states optimize substrate engagement, spatial exploration, and degradation efficiency. This adaptability is essential for navigating heterogeneous substrates and fluctuating environmental conditions.
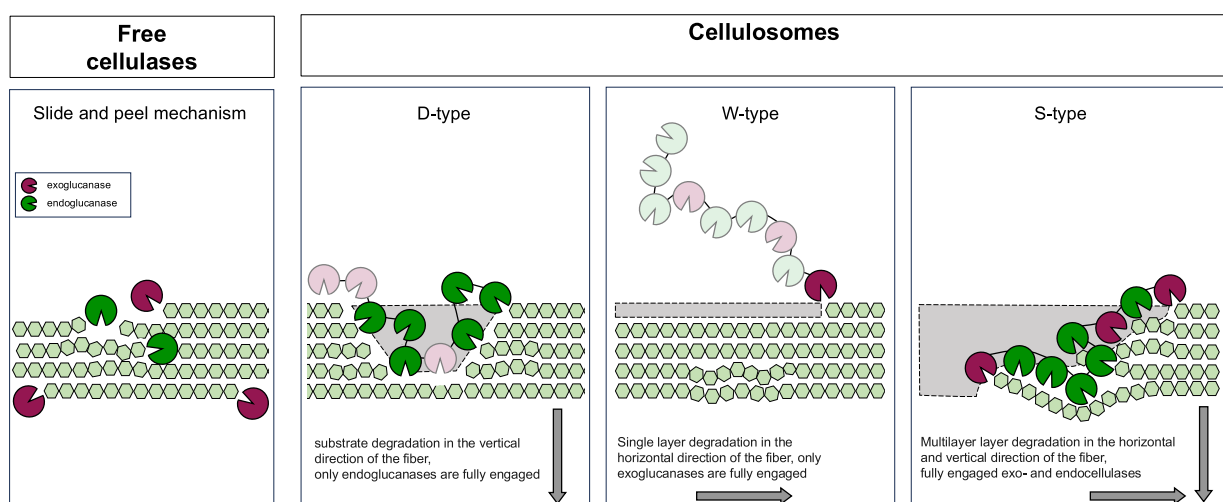

FIGURE 5
Movement of non-complexed cellulases versus cellulosomes on cellulose microfibrils. Free non-complexed cellulases penetrate the cellulose substrate and move via "slide and peel" mechanism. In contrast, cellulosomes operate in a distinct "sit-and-dig" manner. Functioning as molecular machines, cellulosomes degrade cellulose fibrils through three modes: D-type, W-type, and S-type. The engagement of exo- (magenta) and endoglucanases (green) is indicated by filled circles, and the extent of cellulose removal (gray) reflects the contribution of each mode. The S-type shows full subunit engagement and bidirectional fibril degradation (gray arrows). Adapted from Zajki-Zechmeister et al. (2024; CC BY 4.0).
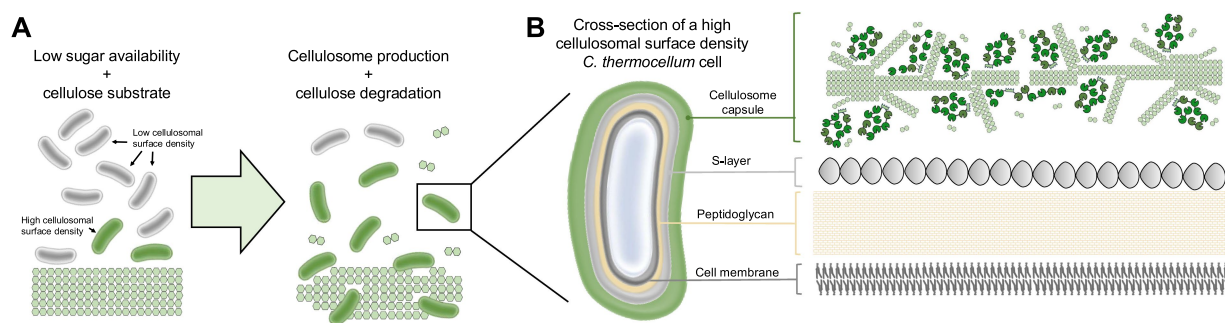
FIGURE 6
Phenotypic and ultrastructural adaptations of *C. thermocellum* during cellulose degradation. **(A)** *C. thermocellum* cell population displays dynamic phenotypic heterogeneity regulated by sugar availability. Gray and green rods represent cells with low and high density of surface associated cellulosomes, respectively. Under low-sugar conditions and upon cellulose exposure, cellulosome expression increases, expanding the high-density population to enhance cellulose degradation. As sugar accumulates, expression is downregulated, restoring the low-density majority. **(B)** Cross-section of an high-cellulosome density phenotype *C. thermocellum* cell showing the cell membrane, peptidoglycan layer, S-layer and surrounding cellulosome capsule. Cellulosomes are engaged with cellulose fibers, and when in interaction, the compact fibrillar structure is locally disrupted, and the fiber degraded. Adapted from Tatli et al. (2022; CC0 1.0).

9 Limitations and challenges of cellulosome-based strategies in biorefinery applications

The integration of natural and designer cellulosomes into biorefinery platforms holds significant promise for efficient lignocellulosic biomass conversion, yet their industrial deployment faces several critical bottlenecks. Firstly, despite the cellulosome's inherently synergistic architecture, effective hydrolysis still relies on costly pretreatment methods to enhance substrate accessibility, often negating the economic advantage of consolidated bioprocessing strategies (Liu et al., 2019; Wang et al., 2019; Yu et al., 2020).

Natural cellulosomes, such as those produced by *C. thermocellum*, also suffer from compositional limitations. For example, the absence of *β*-glucosidase leads to cellobiose accumulation and feedback inhibition (Hirano et al., 2019; Kruus et al., 1995). DCs attempt to overcome such constraints through synthetic scaffolding of multiple enzymes, but their performance on pretreated biomass frequently remains inferior to optimized free-enzyme cocktails, achieving only 33–42% of native cellulosome efficiency (Chundawat et al., 2016; Gefen et al., 2012; Hirano et al., 2019; Moraïs et al., 2012). Despite proven efficacy of proximity-driven synergy between cellulases and accessory enzymes such as LPMOs, xylanases, laccases, or expansin-like proteins, defining optimal enzyme combinations within the expansive design space remains challenging. Furthermore, the production and assembly of complex multienzyme designer systems are technically demanding, costly, and not yet scalable. The immense combinatorial possibilities of scaffoldin architectures and enzyme modules make high-throughput screening and rational design exceptionally difficult to implement (Lamote et al., 2023; Tsai et al., 2022).

In consolidated bioprocessing strategies contexts, metabolic burden imposed by heterologous enzyme expression in engineered hosts such as *Saccharomyces cerevisiae* often compromises fermentation efficiency (Dong et al., 2020b; Sharma et al., 2022; Song et al., 2025; Tsai et al., 2022). Additionally, host organisms that naturally produce cellulosomes are difficult to genetically manipulate (Olson et al., 2023; Riley et al., 2019) and may lack robust industrial

traits (Mazzoli and Olson, 2020). On the biochemical level, trade-offs between (thermo)stability and enzymatic activity (Kahn et al., 2019) as well as steric hindrance during enzyme immobilization (Vazana et al., 2013) limit the practical stability and reusability of cellulosomal systems. Altogether, these barriers, ranging from substrate challenges and enzyme composition to host engineering and process economics, highlight the need for continued advances in synthetic biology, pretreatment innovation, and techno-economic optimization to enable the viable industrial application of cellulosome-based technologies in biorefineries.

10 Conclusion and outlook

Despite recent advances highlighting the vast potential of optimized cellulosome design for improved polysaccharide processing, several challenges remain. The dynamic modularity and component heterogeneity of cellulosomes continue to hinder high-resolution structural characterization and mechanistic understanding of their assembly and function in native environments.

One major area of interest is the optimization of CBMs, which are pivotal for substrate targeting and enzyme–substrate proximity. However, challenges persist in improving thermostability and substrate specificity, as well as optimizing the spatial arrangement and functional integration of CBMs with catalytic domains. Furthermore, the multivalency and cooperativity of CBMs in the context of native or engineered cellulosomes remain incompletely understood, limiting the rational design of more efficient cellulolytic systems.

Equally important is the need to better understand cohesin–dockerin interactions, particularly dual-binding modes that underlie the dynamic adaptability of cellulosomes. These interactions remain poorly understood, especially in wild-type systems. In addition, less-studied type III cohesin-dockerin interactions, together with advances in genomics, structural biology, and bioinformatics, offer opportunities to uncover new functions and engineer next-generation cellulosomal assemblies. Notably, some cohesin-dockerin systems have even been proposed to function beyond classical polysaccharide degradation,

potentially contributing to extracellular protein or lipid breakdown, microbial community interactions, and environmental sensing.

Another key challenge lies in understanding and optimizing cell surface attachment mechanisms. This includes identification of novel anchoring mechanisms beyond classical SLH and dockerin-cohesin mediated interactions, and optimization of anchorage efficiency, robustness, and flexibility, particularly for compatibility with the non-native hosts. Moreover, the dynamic regulation of attachment systems to adapt to changes in the environmental conditions warrants further investigation.

Realizing the full potential of DCs will also require a deeper understanding of enzymatic synergy between cellulases and hemicellulases, the mechanisms of product and intermediate inhibition, and the functional roles of associated enzymes with atypical catalytic activities, such as proteases or redox enzymes, that may contribute to biomass deconstruction.

DCs offer a promising strategy to reduce the environmental footprint and improve the economic viability of lignocellulosic biorefineries. By enhancing catalytic efficiency, DCs lower enzyme usage, reduce energy demand and transport emissions and support local bioeconomies. Realizing their full potential will require a coordinated research roadmap that merges mechanistic insight with rational engineering: (1) applying high-resolution cryo-EM, NMR, and single-molecule techniques (e.g., FRET, optical tweezers) to capture the transient cohesin–dockerin, CBM, and catalytic-domain interactions; (2) mining genomes and metagenomes to uncover novel DC-building blocks; (3) using molecular-dynamics simulations to refine domain architecture, dockerin-cohesin interactions, linker flexibility, substrate binding and surface-display; (4) integrating transcriptomic, proteomic, and metabolomic data in systems-biology models to understand and engineer regulatory mechanisms controlling cellulosome expression and anchoring; and (5) iterating computational design, modular assembly, and high-throughput screening to generate DCs tailored for specific microbial chassis and lignocellulosic substrates. By uniting these experimental and computational tools within a sustainability-driven framework, the field is poised to deliver high-performance, application-specific cellulosomal platforms that can power the next generation of biomass-valorization technologies.

Author contributions

NL: Writing – review & editing, Writing – original draft, Visualization, Conceptualization. MV: Writing – review & editing, Visualization, Writing – original draft, Conceptualization.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript. During the preparation of this work, the generative AI language model ChatGPT, version GPT-4o-mini, developed by OpenAI (https://openai.com) was used in order to improve the text (language). After using this tool/service, we reviewed and edited the content and take full responsibility for the content of the publication.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aburaya, S., Aoki, W., Kuroda, K., Minakuchi, H., and Ueda, M. (2019). Temporal proteome dynamics of *Clostridium cellulovorans* cultured with major plant cell wall polysaccharides. *BMC Microbiol.* 19:118. doi: 10.1186/s12866-019-1480-0

Aburaya, S., Esaka, K., Morisaka, H., Kuroda, K., and Ueda, M. (2015). Elucidation of the recognition mechanisms for hemicellulose and pectin in *Clostridium cellulovorans* using intracellular quantitative proteome analysis. *AMB Express* 5:29. doi: 10.1186/s13568-015-0115-6

Adams, J. J., Pal, G., Jia, Z., and Smith, S. P. (2006). Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl. Acad. Sci.* 103, 305–310. doi: 10.1073/pnas.0507109103

Adams, J. J., Webb, B. A., Spencer, H. L., and Smith, S. P. (2005). Structural characterization of type II Dockerin module from the Cellulosome of *Clostridium thermocellum*: calcium-induced effects on conformation and target recognition. *Biochemistry* 44, 2173–2182. doi: 10.1021/bi048039u

Ahmed, J., Poonia, A., and Goyal, A. (2024). Deciphering the structure of a distinctive trimodular cellulosomal licheninase (*Rf* GH16_21), a family 16 glycoside hydrolase from *Ruminococcus flavefaciens* by computational and experimental methods. *J. Biomol. Struct. Dyn.* 42, 3094–3107. doi: 10.1080/07391102.2023.2212076

Alber, O., Noach, I., Rincon, M. T., Flint, H. J., Shimon, L. J. W., Lamed, R., et al. (2009). Cohesin diversity revealed by the crystal structure of the anchoring cohesin from *Ruminococcus flavefaciens*. *Proteins* 77, 699–709. doi: 10.1002/prot.22483

Anandharaj, M., Lin, Y.-J., Rani, R. P., Nadendla, E. K., Ho, M.-C., Huang, C.-C., et al. (2020). Constructing a yeast to express the largest cellulosome complex on the cell surface. *Proc. Natl. Acad. Sci.* 117, 2385–2394. doi: 10.1073/pnas.1916529117

Arantes, V., and Saddler, J. N. (2010). Access to cellulose limits the efficiency of enzymatic hydrolysis: the role of amorphogenesis. *Biotechnol. Biofuels* 3:4. doi: 10.1186/1754-6834-3-4

Arora, R., Behera, S., Sharma, N. K., and Kumar, S. (2015). Bioprospecting thermostable cellulosomes for efficient biofuel production from lignocellulosic biomass. *Bioresour. Bioprocess.* 2:38. doi: 10.1186/s40643-015-0066-4

Artzi, L., Bayer, E. A., and Moraïs, S. (2017). Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* 15, 83–95. doi: 10.1038/nrmicro.2016.164

Artzi, L., Dassa, B., Borovok, I., Shamshoum, M., Lamed, R., and Bayer, E. A. (2014). Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum. Biotechnol. Biofuels* 7:100. doi: 10.1186/1754-6834-7-100

Artzi, L., Morag, E., Barak, Y., Lamed, R., and Bayer, E. A. (2015). *Clostridium clariflavum*: key cellulosome players are revealed by proteomic analysis. *MBio* 6:e00411-15. doi: 10.1128/mBio.00411-15

Artzi, L., Morag, E., Shamshoum, M., and Bayer, E. A. (2016). Cellulosomal expansin: functionality and incorporation into the complex. *Biotechnol. Biofuels* 9:61. doi: 10.1186/s13068-016-0474-5

Atreya, M. E., Strobel, K. L., and Clark, D. S. (2016). Alleviating product inhibition in cellulase enzyme Cel7A. *Biotechnol. Bioeng.* 113, 330–338. doi: 10.1002/bit.25809

Bayer, E. A., Lamed, R., White, B. A., and Flint, H. J. (2008). From cellulosomes to cellulosomics. *Chem. Rec.* 8, 364–377. doi: 10.1002/tcr.20160

Ben David, Y., Dassa, B., Borovok, I., Lamed, R., Koropatkin, N. M., Martens, E. C., et al. (2015). Ruminococcal cellulose systems from rumen to human. *Environ. Microbiol.* 17, 3407–3426. doi: 10.1111/1462-2920.12868

Bhat, A. H., Nguyen, M. T., Das, A., and Ton-That, H. (2021). Anchoring surface proteins to the bacterial cell wall by sortase enzymes: how it started and what we know now. *Curr. Opin. Microbiol.* 60, 73–79. doi: 10.1016/j.mib.2021.01.013

Blouzard, J., Coutinho, P. M., Fierobe, H., Henrissat, B., Lignon, S., Tardif, C., et al. (2010). Modulation of cellulosome composition in *Clostridium cellulolyticum*: adaptation to the polysaccharide environment revealed by proteomic and carbohydrate-active enzyme analyses. *Proteomics* 10, 541–554. doi: 10.1002/pmic.200900311

Bomble, Y. J., Beckham, G. T., Matthews, J. F., Nimlos, M. R., Himmel, M. E., and Crowley, M. F. (2011). Modeling the self-assembly of the cellulosome enzyme complex. *J. Biol. Chem.* 286, 5614–5623. doi: 10.1074/jbc.M110.186031

Boraston, A. B., Bolam, D. N., Gilbert, H. J., and Davies, G. J. (2004). Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochem. J.* 382, 769–781. doi: 10.1042/BJ20040892

Borne, R., Bayer, E. A., Pagès, S., Perret, S., and Fierobe, H. (2013). Unraveling enzyme discrimination during cellulosome assembly independent of cohesin – dockerin affinity. *FEBS J.* 280, 5764–5779. doi: 10.1111/febs.12497

Borne, R., Dao Ti, M., Fierobe, H., Vita, N., Tardif, C., and Pagès, S. (2020). Catalytic subunit exchanges in the cellulosomes produced by *Ruminiclostridium cellulolyticum* suggest unexpected dynamics and adaptability of their enzymatic composition. *FEBS J.* 287, 2544–2559. doi: 10.1111/febs.15155

Bozkurt, E. U., Ørsted, E. C., Volke, D. C., and Nikel, P. I. (2025). Accelerating enzyme discovery and engineering with high-throughput screening. *Nat. Prod. Rep.* doi: 10.1039/D4NP00031E

Brás, J. L. A., Alves, V. D., Carvalho, A. L., Najmudin, S., Prates, J. A. M., Ferreira, L. M. A., et al. (2012). Novel *Clostridium thermocellum* type I Cohesin-Dockerin complexes reveal a single binding mode. *J. Biol. Chem.* 287, 44394–44405. doi: 10.1074/jbc.M112.407700

Brás, J. L. A., Pinheiro, B. A., Cameron, K., Cuskin, F., Viegas, A., Najmudin, S., et al. (2016). Diverse specificity of cellulosome attachment to the bacterial cell surface. *Sci. Rep.* 6:38292. doi: 10.1038/srep38292

Bugada, L. F., Prabhu, P., Bailey, R., and Wen, F. (2025). Covalent and orthogonal cohesin–dockerin interactions enabled by intermolecular disulfide bonds for hyperthermostable cellulase assembly. *Ind. Eng. Chem. Res.* 64, 100–112. doi: 10.1021/acs.iecr.4c03484

Bule, P., Alves, V. D., Israeli-Ruimy, V., Carvalho, A. L., Ferreira, L. M. A., Smith, S. P., et al. (2017). Assembly of *Ruminococcus flavefaciens* cellulosome revealed by structures of two cohesin-dockerin complexes. *Sci. Rep.* 7:759. doi: 10.1038/s41598-017-00919-w

Bule, P., Alves, V. D., Leitão, A., Ferreira, L. M. A., Bayer, E. A., Smith, S. P., et al. (2016). Single binding mode integration of hemicellulose-degrading enzymes via adaptor Scaffoldins in *Ruminococcus flavefaciens* Cellulosome. *J. Biol. Chem.* 291, 26658–26669. doi: 10.1074/jbc.M116.761643

Bule, P., Pires, V. M. R., Alves, V. D., Carvalho, A. L., Prates, J. A. M., Ferreira, L. M. A., et al. (2018b). Higher order scaffoldin assembly in *Ruminococcus flavefaciens* cellulosome is coordinated by a discrete cohesin-dockerin interaction. *Sci. Rep.* 8:6987. doi: 10.1038/s41598-018-25171-8

Bule, P., Pires, V. M., Fontes, C. M., and Alves, V. D. (2018a). Cellulosome assembly: paradigms are meant to be broken! *Curr. Opin. Struct. Biol.* 49, 154–161. doi: 10.1016/j.sbi.2018.03.012

Cai, S., Zheng, X., and Dong, X. (2011). CBM3d, a novel subfamily of family 3 carbohydrate-binding modules identified in Cel48A Exoglucanase of *Cellulosilyticum ruminicola. J. Bacteriol.* 193, 5199–5206. doi: 10.1128/JB.05227-11

Cameron, K., Weinstein, J. Y., Zhivin, O., Bule, P., Fleishman, S. J., Alves, V. D., et al. (2015). Combined crystal structure of a type I Cohesin: mutation and affinity binding studies reveal structural determinants of cohesin-dockerin specificities. *J. Biol. Chem.* 290, 16215–16225. doi: 10.1074/jbc.M115.653303

Cao, L., Li, S., Huang, X., Qin, Z., Kong, W., Xie, W., et al. (2018). Enhancing the Thermostability of highly active and glucose-tolerant β-glucosidase Ks5A7 by directed evolution for good performance of three properties. *J. Agric. Food Chem.* 66, 13228–13235. doi: 10.1021/acs.jafc.8b05662

Carvalho, A. L., Dias, F. M. V., Nagy, T., Prates, J. A. M., Proctor, M. R., Smith, N., et al. (2007). Evidence for a dual binding mode of dockerin modules to cohesins. *Proc. Natl. Acad. Sci.* 104, 3089–3094. doi: 10.1073/pnas.0611173104

Carvalho, A. L., Dias, F. M. V., Prates, J. A. M., Nagy, T., Gilbert, H. J., Davies, G. J., et al. (2003). Cellulosome assembly revealed by the crystal structure of the cohesin–dockerin complex. *Proc. Natl. Acad. Sci.* 100, 13809–13814. doi: 10.1073/pnas.1936124100

Carvalho, A. L., Pires, V. M. R., Gloster, T. M., Turkenburg, J. P., Prates, J. A. M., Ferreira, L. M. A., et al. (2005). Insights into the structural determinants of Cohesin—Dockerin specificity revealed by the crystal structure of the type II Cohesin from *Clostridium thermocellum* SdbA. *J. Mol. Biol.* 349, 909–915. doi: 10.1016/j.jmb.2005.04.037

Caspi, J., Barak, Y., Haimovitz, R., Irwin, D., Lamed, R., Wilson, D. B., et al. (2009). Effect of linker length and dockerin position on conversion of a *Thermobifida fusca* endoglucanase to the cellulosomal mode. *Appl. Environ. Microbiol.* 75, 7335–7342. doi: 10.1128/AEM.01241-09

Caspi, J., Irwin, D., Lamed, R., Li, Y., Fierobe, H.-P., Wilson, D. B., et al. (2008). Conversion of *Thermobifida fusca* free exoglucanases into cellulosomal components: comparative impact on cellulose-degrading activity. *J. Biotechnol.* 135, 351–357. doi: 10.1016/j.jbiotec.2008.05.003

Chen, M., Bu, L., Alahuhta, M., Brunecky, R., Xu, Q., Lunin, V. V., et al. (2016). Strategies to reduce end-product inhibition in family 48 glycoside hydrolases. *Proteins* 84, 295–304. doi: 10.1002/prot.24965

Chen, C., Cui, Z., Song, X., Liu, Y.-J., Cui, Q., and Feng, Y. (2016). Integration of bacterial expansin-like proteins into cellulosome promotes the cellulose degradation. *Appl. Microbiol. Biotechnol.* 100, 2203–2212. doi: 10.1007/s00253-015-7071-6

Chen, C., Cui, Z., Xiao, Y., Cui, Q., Smith, S. P., Lamed, R., et al. (2014). Revisiting the NMR solution structure of the Cel48S type-I dockerin module from *Clostridium thermocellum* reveals a cohesin-primed conformation. *J. Struct. Biol.* 188, 188–193. doi: 10.1016/j.jsb.2014.09.006

Chen, C., Dong, S., Yu, Z., Qiao, Y., Li, J., Ding, X., et al. (2023). Essential autoproteolysis of bacterial anti-σ factor RsgI for transmembrane signal transduction. *Sci. Adv.* 9:eadg4846. doi: 10.1126/sciadv.adg4846

Chen, L., and Ge, X. (2018). Correlation between size and activity enhancement of Recombinantly assembled Cellulosomes. *Appl. Biochem. Biotechnol.* 186, 937–948. doi: 10.1007/s12010-018-2786-3

Chen, C., Qi, K., Chi, F., Song, X., Feng, Y., Cui, Q., et al. (2022). Dissolved xylan inhibits cellulosome-based saccharification by binding to the key cellulosomal component of *Clostridium thermocellum. Int. J. Biol. Macromol.* 207, 784–790. doi: 10.1016/j.ijbiomac.2022.03.158

Chen, C., Yang, H., Dong, S., You, C., Moraïs, S., Bayer, E. A., et al. (2024). A cellulosomal double-dockerin module from *Clostridium thermocellum* shows distinct structural and cohesin-binding features. *Protein Sci.* 33:e4937. doi: 10.1002/pro.4937

Chen, C., Yang, H., Xuan, J., Cui, Q., and Feng, Y. (2019). Resonance assignments of a cellulosomal double-dockerin from *Clostridium thermocellum. Biomol. NMR Assign.* 13, 97–101. doi: 10.1007/s12104-018-9859-7

Chen, Y., Zhang, S., Zhai, Z., Zhang, S., Ma, J., Liang, X., et al. (2023). Construction of fusion protein with carbohydrate-binding module and leaf-branch compost Cutinase to enhance the degradation efficiency of polyethylene terephthalate. *Int. J. Mol. Sci.* 24:2780. doi: 10.3390/ijms24032780

Chundawat, S. P. S., Hackl, M., Lopez, C. A., Santo, K. P., Gnanakaran, S., and Neimark, A. (2024). Acoustic force spectroscopy reveals subtle differences in interfacial protein-polysaccharide binding interactions. *Biophys. J.* 123:143a. doi: 10.1016/j.bpj.2023.11.992

Chundawat, S. P. S., Paavola, C. D., Raman, B., Nouailler, M., Chan, S. L., Mielenz, J. R., et al. (2016). Saccharification of thermochemically pretreated cellulosic biomass using native and engineered cellulosomal enzyme systems. *React. Chem. Eng.* 1, 616–628. doi: 10.1039/C6RE00172F

Chung, D., Young, J., Bomble, Y. J., Vander Wall, T. A., Groom, J., Himmel, M. E., et al. (2015). Homologous expression of the *Caldicellulosiruptor bescii* CelA reveals that the extracellular protein is glycosylated. *PLoS One* 10:e0119508. doi: 10.1371/journal.pone.0119508

Cosgrove, D. J. (2017). Microbial Expansins. *Ann. Rev. Microbiol.* 71, 479–497. doi: 10.1146/annurev-micro-090816-093315

Creagh, A. L., Ong, E., Jervis, E., Kilburn, D. G., and Haynes, C. A. (1996). Binding of the cellulose-binding domain of exoglucanase Cex from *Cellulomonas fimi* to insoluble microcrystalline cellulose is entropically driven. *Proc. Natl. Acad. Sci.* 93, 12229–12234. doi: 10.1073/pnas.93.22.12229

Currie, M. A., Adams, J. J., Faucher, F., Bayer, E. A., Jia, Z., and Smith, S. P. (2012). Scaffoldin conformation and dynamics revealed by a ternary complex from the *Clostridium thermocellum* cellulosome. *J. Biol. Chem.* 287, 26953–26961. doi: 10.1074/jbc.M112.343897

Dassa, B., Borovok, I., Lombard, V., Henrissat, B., Lamed, R., Bayer, E. A., et al. (2017). Pan-Cellulosomics of mesophilic Clostridia: variations on a theme. *Microorganisms* 5:74. doi: 10.3390/microorganisms5040074

Dassa, B., Borovok, I., Ruimy-Israeli, V., Lamed, R., Flint, H. J., Duncan, S. H., et al. (2014). Rumen Cellulosomics: divergent Fiber-degrading strategies revealed by comparative genome-wide analysis of six Ruminococcal strains. *PLoS One* 9:e99221. doi: 10.1371/journal.pone.0099221

Datta, R. (2024). Enzymatic degradation of cellulose in soil: a review. *Heliyon* 10:e24022. doi: 10.1016/j.heliyon.2024.e24022

Davidi, L., Moraïs, S., Artzi, L., Knop, D., Hadar, Y., Arfi, Y., et al. (2016). Toward combined delignification and saccharification of wheat straw by a laccase-containing designer cellulosome. *Proc. Natl. Acad. Sci.* 113, 10854–10859. doi: 10.1073/pnas.1608021113

de Andrades, D., Alnoch, R. C., Alves, G. S., Salgado, J. C. S., Almeida, P. Z., Berto, G. L., et al. (2024). Recombinant GH3 β-glucosidase stimulated by xylose and tolerant to furfural and 5-hydroxymethylfurfural obtained from aspergillus nidulans. *Bioresour. Bioprocess.* 11:77. doi: 10.1186/s40643-024-00784-2

de Camargo, B. R., Steindorff, A. S., da Silva, L. A., de Oliveira, A. S., Hamann, P. R. V., and Noronha, E. F. (2023). Expression profiling of *Clostridium thermocellum* B8 during the deconstruction of sugarcane bagasse and straw. *World J. Microbiol. Biotechnol.* 39:105. doi: 10.1007/s11274-023-03546-y

De Chellis, A., Murphy, R., Hackl, M., Ganesan, G., and Chundawat, S. P. S. (2024a). Characterizing multivalent carbohydrate binding modules interactions at polysaccharide interfaces with acoustic force spectroscopy. *Biophys. J.* 123:144a. doi: 10.1016/j.bpj.2023.11.996

De Chellis, A., Nemmaru, B., Sammond, D., Douglass, J., Patil, N., Reste, O., et al. (2024b). Supercharging carbohydrate binding module alone enhances endocellulase thermostability, binding, and activity on cellulosic biomass. *ACS Sustain. Chem. Eng.* 12, 3500–3516. doi: 10.1021/acssuschemeng.3c06266

de Sandozequi, A., Salazar-Cortés, J. J., Tapia-Vázquez, I., and Martínez-Anaya, C. (2022). Prevalent association with the bacterial cell envelope of prokaryotic expansins revealed by bioinformatics analysis. *Protein Sci.* 31:e4315. doi: 10.1002/pro.4315

Deng, D., Meng, Q., Li, Z., Ma, R., Yang, Y., Wang, Z., et al. (2020). Enzyme-inspired assembly: incorporating multivariate interactions to optimize the host–guest configuration for high-speed enantioselective catalysis. *ACS Appl. Mater. Interfaces* 12, 47966–47974. doi: 10.1021/acsami.0c13802

Desvaux, M., Dumas, E., and Chafsey, I. (2006). Protein cell surface display in gram-positive bacteria: from single protein to macromolecular protein structure. *FEMS Microbiol. Lett.* 256, 1–15. doi: 10.1111/j.1574-6968.2006.00122.x

Ding, S.-Y., Liu, Y.-S., Zeng, Y., Himmel, M. E., Baker, J. O., and Bayer, E. A. (2012). How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? *Science* 338, 1055–1060. doi: 10.1126/science.1227491

Ding, S.-Y., Xu, Q., Ali, M. K., Baker, J. O., Bayer, E. A., Barak, Y., et al. (2006). Versatile derivatives of carbohydrate-binding modules for imaging of complex carbohydrates approaching the molecular level of resolution. *BioTechniques* 41, 435–443. doi: 10.2144/000112244

Doi, R. H., Goldstein, M., Hashida, S., Park, J.-S., and Takagi, M. (1994). The *Clostridium cellulovorans* Cellulosome. *Crit. Rev. Microbiol.* 20, 87–93.

Dong, S., Chen, C., Li, J., Liu, Y.-J., Bayer, E. A., Lamed, R., et al. (2024). Unique Fn3-like biosensor in σI/anti-σI factors for regulatory expression of major cellulosomal scaffoldins in *pseudobacteroides cellulosolvens*. *Protein Sci.* 33:e5193. doi: 10.1002/pro.5193

Dong, C., Qiao, J., Wang, X., Sun, W., Chen, L., Li, S., et al. (2020). Engineering Pichia pastoris with surface-display minicellulosomes for carboxymethyl cellulose hydrolysis and ethanol production. *Biotechnol. Biofuels* 13:108. doi: 10.1186/s13068-020-01749-1

Dorival, J., Moraïs, S., Labourel, A., Rozycki, B., Cazade, P.-A., Dabin, J., et al. (2022). Mapping the deformability of natural and designed cellulosomes in solution. *Biotechnol. Biofuels Bioprod.* 15:68. doi: 10.1186/s13068-022-02165-3

Drula, E., Garron, M.-L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045

Duarte, M., Alves, V. D., Correia, M., Caseiro, C., Ferreira, L. M. A., Romão, M. J., et al. (2023). Structure-function studies can improve binding affinity of cohesin-dockerin interactions for multi-protein assemblies. *Int. J. Biol. Macromol.* 224, 55–67. doi: 10.1016/j.ijbiomac.2022.10.102

Duarte, M., Viegas, A., Alves, V. D., Prates, J. A. M., Ferreira, L. M. A., Najmudin, S., et al. (2021). A dual cohesin–dockerin complex binding mode in *Bacteroides*

cellulosolvens contributes to the size and complexity of its cellulosome. *J. Biol. Chem.* 296:100552. doi: 10.1016/j.jbc.2021.100552

Eibinger, M., Ganner, T., Plank, H., and Nidetzky, B. (2020). A biological Nanomachine at work: watching the Cellulosome degrade crystalline cellulose. *ACS Cent Sci* 6, 739–746. doi: 10.1021/acscentsci.0c00050

Eibinger, M., Sattelkow, J., Ganner, T., Plank, H., and Nidetzky, B. (2017). Single-molecule study of oxidative enzymatic deconstruction of cellulose. *Nat. Commun.* 8:894. doi: 10.1038/s41467-017-01028-y

Esaka, K., Aburaya, S., Morisaka, H., Kuroda, K., and Ueda, M. (2015). Exoproteome analysis of *Clostridium cellulovorans* in natural soft-biomass degradation. *AMB Express* 5:2. doi: 10.1186/s13568-014-0089-9

Fierobe, H.-P., Mingardon, F., Mechaly, A., Bélaïch, A., Rincon, M. T., Pagès, S., et al. (2005). Action of designer Cellulosomes on homogeneous versus complex substrates. *J. Biol. Chem.* 280, 16325–16334. doi: 10.1074/jbc.M414449200

Finn, R. D., Attwood, T. K., Babbitt, P. C., Bateman, A., Bork, P., Bridge, A. J., et al. (2017). InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res.* 45, D190–D199. doi: 10.1093/nar/gkw1107

Flint, H. J., Whitehead, T. R., Martin, J. C., and Gasparic, A. (1997). Interrupted catalytic domain structures in xylanases from two distantly related strains of *Prevotella ruminicola*. *Biochim. Biophys. Acta Protein Struct. Mol. Enzymol.* 1337, 161–165. doi: 10.1016/S0167-4838(96)00213-0

Gad, S., and Ayakar, S. (2021). Protein scaffolds: a tool for multi-enzyme assembly. *Biotechnol. Reports* 32:e00670. doi: 10.1016/j.btre.2021.e00670

Galera-Prat, A., Moraïs, S., Vazana, Y., Bayer, E. A., and Carrión-Vázquez, M. (2018). The cohesin module is a major determinant of cellulosome mechanical stability. *J. Biol. Chem.* 293, 7139–7147. doi: 10.1074/jbc.RA117.000644

Galera-Prat, A., Vera, A. M., Moraïs, S., Vazana, Y., Bayer, E. A., and Carrión-Vázquez, M. (2020). Impact of scaffoldin mechanostability on cellulosomal activity. *Biomater. Sci.* 8, 3601–3610. doi: 10.1039/C9BM02052G

García-Alvarez, B., Melero, R., Dias, F. M. V., Prates, J. A. M., Fontes, C. M. G. A., Smith, S. P., et al. (2011). Molecular architecture and structural transitions of a *Clostridium thermocellum* Mini-Cellulosome. *J. Mol. Biol.* 407, 571–580. doi: 10.1016/j.jmb.2011.01.060

Gefen, G., Anbar, M., Morag, E., Lamed, R., and Bayer, E. A. (2012). Enhanced cellulose degradation by targeted integration of a cohesin-fused β-glucosidase into the *Clostridium thermocellum* cellulosome. *Proc. Natl. Acad. Sci. USA* 109, 10298–10303. doi: 10.1073/pnas.1202747109

Gharechahi, J., Vahidi, M. F., Sharifi, G., Ariaeenejad, S., Ding, X.-Z., Han, J.-L., et al. (2023). Lignocellulose degradation by rumen bacterial communities: new insights from metagenome analyses. *Environ. Res.* 229:115925. doi: 10.1016/j.envres.2023.115925

Gilad, R., Rabinovich, L., Yaron, S., Bayer, E. A., Lamed, R., Gilbert, H. J., et al. (2003). CelI, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *J. Bacteriol.* 185, 391–398. doi: 10.1128/JB.185.2.391-398.2003

Gilbert, H. J., Knox, J. P., and Boraston, A. B. (2013). Advances in understanding the molecular basis of plant cell wall polysaccharide recognition by carbohydrate-binding modules. *Curr. Opin. Struct. Biol.* 23, 669–677. doi: 10.1016/j.sbi.2013.05.005

Gilmore, S. P., Lillington, S. P., Haitjema, C. H., de Groot, R., and O'Malley, M. A. (2020). Designing chimeric enzymes inspired by fungal cellulosomes. *Synth Syst Biotechnol* 5, 23–32. doi: 10.1016/j.synbio.2020.01.003

Gold, N. D., and Martin, V. J. J. (2007). Global view of the *Clostridium thermocellum* Cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.* 189, 6787–6795. doi: 10.1128/JB.00882-07

Grépinet, O., Chebrou, M. C., and Béguin, P. (1988). Nucleotide sequence and deletion analysis of the xylanase gene (xynZ) of *Clostridium thermocellum*. *J. Bacteriol.* 170, 4582–4588. doi: 10.1128/jb.170.10.4582-4588.1988

Hackl, M., Contrada, E. V., Ash, J. E., Kulkarni, A., Yoon, J., Cho, H.-Y., et al. (2022). Acoustic force spectroscopy reveals subtle differences in cellulose unbinding behavior of carbohydrate-binding modules. *Proc. Natl. Acad. Sci.* 119:e2117467119. doi: 10.1073/pnas.2117467119

Haimovitz, R., Barak, Y., Morag, E., Voronov-Goldman, M., Shoham, Y., Lamed, R., et al. (2008). Cohesin-dockerin microarray: diverse specificities between two complementary families of interacting protein modules. *Proteomics* 8, 968–979. doi: 10.1002/pmic.200700486

Hammel, M., Fierobe, H.-P., Czjzek, M., Finet, S., and Receveur-Bréchot, V. (2004). Structural insights into the mechanism of formation of Cellulosomes probed by small angle X-ray scattering. *J. Biol. Chem.* 279, 55985–55994. doi: 10.1074/jbc.M408979200

Han, S. O., Cho, H.-Y., Yukawa, H., Inui, M., and Doi, R. H. (2004). Regulation of expression of Cellulosomes and Noncellulosomal (hemi)cellulolytic enzymes in *Clostridium cellulovorans* during growth on different carbon sources. *J. Bacteriol.* 186, 4218–4227. doi: 10.1128/JB.186.13.4218-4227.2004

Hao, M.-S., Mazurkewich, S., Li, H., Kvammen, A., Saha, S., Koskela, S., et al. (2024). Structural and biochemical analysis of family 92 carbohydrate-binding modules uncovers multivalent binding to β-glucans. *Nat. Commun.* 15:3429. doi: 10.1038/s41467-024-47584-y

Hepler, N. K., and Cosgrove, D. J. (2019). Directed *in vitro* evolution of bacterial expansin BsEXLX1 for higher cellulose binding and its consequences for plant cell wall-loosening activities. *FEBS Lett.* 593, 2545–2555. doi: 10.1002/1873-3468.13528

Hernandez-Gomez, M. C., Rydahl, M. G., Rogowski, A., Morland, C., Cartmell, A., Crouch, L., et al. (2015). Recognition of xyloglucan by the crystalline cellulose-binding site of a family 3a carbohydrate-binding module. *FEBS Lett.* 589, 2297–2303. doi: 10.1016/j.febslet.2015.07.009

Hershko Rimon, A., Livnah, O., Rozman Grinberg, I., Ortiz de Ora, L., Yaniv, O., Lamed, R., et al. (2021). Novel clostridial cell-surface hemicellulose-binding CBM3 proteins. *Acta Crystallogr F Struct Biol Commun* 77, 95–104. doi: 10.1107/S2053230X21002764

Hettle, A., Fillo, A., Abe, K., Massel, P., Pluvinage, B., Langelaan, D. N., et al. (2017). Properties of a family 56 carbohydrate-binding module and its role in the recognition and hydrolysis of β-1,3-glucan. *J. Biol. Chem.* 292, 16955–16968. doi: 10.1074/jbc.M117.806711

Hirano, K., Kurosaki, M., Nihei, S., Hasegawa, H., Shinoda, S., Haruki, M., et al. (2016). Enzymatic diversity of the *Clostridium thermocellum* cellulosome is crucial for the degradation of crystalline cellulose and plant biomass. *Sci. Rep.* 6:35709. doi: 10.1038/srep35709

Hirano, K., Saito, T., Shinoda, S., Haruki, M., and Hirano, N. (2019). *In vitro* assembly and cellulolytic activity of a β -glucosidase-integrated cellulosome complex. *FEMS Microbiol. Lett.* 366:fnz209. doi: 10.1093/femsle/fnz209

Ichikawa, S., Yoshida, M., Karita, S., Kondo, M., and Goto, M. (2016). Carbohydrate-binding modules influence substrate specificity of an endoglucanase from *Clostridium thermocellum*. *Biosci. Biotechnol. Biochem.* 80, 188–192. doi: 10.1080/09168451.2015.1069696

Igarashi, K., Uchihashi, T., Koivula, A., Wada, M., Kimura, S., Okamoto, T., et al. (2011). Traffic jams reduce hydrolytic efficiency of cellulase on cellulose surface. *Science* 333, 1279–1282. doi: 10.1126/science.1208386

Jindou, S., Soda, A., Karita, S., Kajino, T., Béguin, P., Wu, J. H. D., et al. (2004). Cohesin-Dockerin interactions within and between Clostridium josui and *Clostridium thermocellum*. *J. Biol. Chem.* 279, 9867–9874. doi: 10.1074/jbc.M308673200

Jindou, S., Xu, Q., Kenig, R., Shulman, M., Shoham, Y., Bayer, E. A., et al. (2006). Novel architecture of family-9 glycoside hydrolases identified in cellulosomal enzymes of *Acetivibrio cellulolyticus* and *Clostridium thermocellum*. *FEMS Microbiol. Lett.* 254, 308–316. doi: 10.1111/j.1574-6968.2005.00040.x

Kahel-Raifer, H., Jindou, S., Bahari, L., Nataf, Y., Shoham, Y., Bayer, E. A., et al. (2010). The unique set of putative membrane-associated anti-σ factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. *FEMS Microbiol. Lett.* 308, 84–93. doi: 10.1111/j.1574-6968.2010.01997.x

Kahn, A., Moraïs, S., Chung, D., Sarai, N. S., Hengge, N. N., Kahn, A., et al. (2020). Glycosylation of hyperthermostable designer cellulosome components yields enhanced stability and cellulose hydrolysis. *FEBS J.* 287, 4370–4388. doi: 10.1111/febs.15251

Kahn, A., Moraïs, S., Galanopoulou, A. P., Chung, D., Sarai, N. S., Hengge, N., et al. (2019). Creation of a functional hyperthermostable designer cellulosome. *Biotechnol. Biofuels* 12:44. doi: 10.1186/s13068-019-1386-y

Kang, S., Barak, Y., Lamed, R., Bayer, E. A., and Morrison, M. (2006). The functional repertoire of prokaryote cellulosomes includes the serpin superfamily of serine proteinase inhibitors. *Mol. Microbiol.* 60, 1344–1354. doi: 10.1111/j.1365-2958.2006.05182.x

Kari, J., Olsen, J. P., Jensen, K., Badino, S. F., Krogh, K. B. R. M., Borch, K., et al. (2018). Sabatier principle for interfacial (heterogeneous) enzyme catalysis. *ACS Catal.* 8, 11966–11972. doi: 10.1021/acscatal.8b03547

Karpol, A., Jobby, M. K., Slutzki, M., Noach, I., Chitayat, S., Smith, S. P., et al. (2013). Structural and functional characterization of a novel type-III dockerin from *Ruminococcus flavefaciens*. *FEBS Lett.* 587, 30–36. doi: 10.1016/j.febslet.2012.11.012

Ko, H.-J., Song, H., and Choi, I.-G. (2021). Development of a novel cell surface attachment system to display multi-protein complex using the Cohesin-Dockerin binding pair. *J. Microbiol. Biotechnol.* 31, 1183–1189. doi: 10.4014/jmb.2105.05022

Kosugi, A., Amano, Y., Murashima, K., and Doi, R. H. (2004). Hydrophilic domains of scaffolding protein CbpA promote Glycosyl hydrolase activity and localization of Cellulosomes to the cell surface of *Clostridium cellulovorans*. *J. Bacteriol.* 186, 6351–6359. doi: 10.1128/JB.186.19.6351-6359.2004

Kruus, K., Andreacchi, A., Wang, W. K., and Wu, J. H. D. (1995). Product inhibition of the recombinant CelS, an exoglucanase component of the *Clostridium thermocellum* cellulosome. *Appl. Microbiol. Biotechnol.* 44, 399–404.

Kuch, N. J., Kutschke, M. E., Parker, R., Bingman, C. A., and Fox, B. G. (2023). Contribution of calcium ligands in substrate binding and product release in the Acetovibrio thermocellus glycoside hydrolase family 9 cellulase CelR. *J. Biol. Chem.* 299:104655. doi: 10.1016/j.jbc.2023.104655

Lamed, R., and Bayer, E. A. (1988). The Cellulosome of *Clostridium thermocellum*, in *Advances in Applied Microbiology*. (eds.) A. I. Laskin (San Diego, CA, USA: Academic Press), 1–46. doi: 10.1016/S0065-2164(08)70203-X

Lamote, B., da Fonseca, M. J. M., Vanderstraeten, J., Meert, K., Elias, M., and Briers, Y. (2023). Current challenges in designer cellulosome engineering. *Appl. Microbiol. Biotechnol.* 107, 2755–2770. doi: 10.1007/s00253-023-12474-8

Levi Hevroni, B., Moraïs, S., Ben-David, Y., Morag, E., and Bayer, E. A. (2020). Minimalistic cellulosome of the butanologenic bacterium *Clostridium saccharoperbutylacetonicum*. *MBio* 11:e00443-20. doi: 10.1128/mBio.00443-20

Levy-Assaraf, M., Voronov-Goldman, M., Rozman Grinberg, I., Weiserman, G., Shimon, L. J. W., Jindou, S., et al. (2013). Crystal structure of an uncommon Cellulosome-related protein module from *Ruminococcus flavefaciens* That resembles papain-like cysteine peptidases. *PLoS One* 8:e56138. doi: 10.1371/journal.pone.0056138

Li, J., Chen, C., Liu, Y.-J., Cui, Q., Bayer, E. A., and Feng, Y. (2021). NMR chemical shift assignments of a module of unknown function in the cellulosomal secondary scaffoldin ScaF from *Clostridium thermocellum*. *Biomol. NMR Assign.* 15, 329–334. doi: 10.1007/s12104-021-10025-8

Li, Y., Irwin, D. C., and Wilson, D. B. (2010). Increased crystalline cellulose activity via combinations of amino acid changes in the family 9 catalytic domain and family 3c cellulose binding module of *Thermobifida fusca* Cel9A. *Appl. Environ. Microbiol.* 76, 2582–2588. doi: 10.1128/AEM.02735-09

Li, S., and Liu, G. (2024). Harnessing cellulose-binding protein domains for the development of functionalized cellulose materials. *Bioresour. Bioprocess.* 11:74. doi: 10.1186/s40643-024-00790-4

Li, H., Lu, Z., Hao, M.-S., Kvammen, A., Inman, A. R., Srivastava, V., et al. (2023). Family 92 carbohydrate-binding modules specific for β-1,6-glucans increase the thermostability of a bacterial chitinase. *Biochimie* 212, 153–160. doi: 10.1016/j.biochi.2023.04.019

Li, Y., Wang, X., Zhou, N.-Y., and Ding, J. (2024). Yeast surface display technology: mechanisms, applications, and perspectives. *Biotechnol. Adv.* 76:108422. doi: 10.1016/j.biotechadv.2024.108422

Lillington, S. P., Chrisler, W., Haitjema, C. H., Gilmore, S. P., Smallwood, C. R., Shutthanandan, V., et al. (2021). Cellulosome localization patterns vary across life stages of anaerobic fungi. *MBio* 12:e00832-21. doi: 10.1128/mBio.00832-21

Lillington, S. P., Hamilton, M., Cheng, J.-F., Yoshikuni, Y., and O'Malley, M. A. (2023). Expression and characterization of spore coat CotH kinases from the cellulosomes of anaerobic fungi (Neocallimastigomycetes). *Protein Expr. Purif.* 210:106323. doi: 10.1016/j.pep.2023.106323

Liu, S., Liu, Y.-J., Feng, Y., Li, B., and Cui, Q. (2019). Construction of consolidated bio-saccharification biocatalyst and process optimization for highly efficient lignocellulose solubilization. *Biotechnol. Biofuels* 12:35. doi: 10.1186/s13068-019-1374-2

Liu, J., Shi, J., Gao, J., Shi, R., Zhu, J., Jensen, M. S., et al. (2024). Functional studies on tandem carbohydrate-binding modules of a multimodular enzyme possessing two catalytic domains. *Appl. Environ. Microbiol.* 90:e0088824. doi: 10.1128/aem.00888-24

Liu, J., Sun, D., Zhu, J., Liu, C., and Liu, W. (2021). Carbohydrate-binding modules targeting branched polysaccharides: overcoming side-chain recalcitrance in a non-catalytic approach. *Bioresour. Bioprocess.* 8:28. doi: 10.1186/s40643-021-00381-7

Liu, Y., Wang, P., Tian, J., Seidi, F., Guo, J., Zhu, W., et al. (2022). Carbohydrate-binding modules of potential resources: occurrence in nature, function, and application in Fiber recognition and treatment. *Polymers (Basel)* 14:1806. doi: 10.3390/polym14091806

Lombard, V., Henrissat, B., and Garron, M.-L. (2025). CAZac: an activity descriptor for carbohydrate-active enzymes. *Nucleic Acids Res.* 53, D625–D633. doi: 10.1093/nar/gkae1045

Luís, A. S., Venditto, I., Temple, M. J., Rogowski, A., Baslé, A., Xue, J., et al. (2013). Understanding how noncatalytic carbohydrate binding modules can display specificity for xyloglucan. *J. Biol. Chem.* 288, 4799–4809. doi: 10.1074/jbc.M112.432781

Matsui, K., Bae, J., Esaka, K., Morisaka, H., Kuroda, K., and Ueda, M. (2013). Exoproteome profiles of *Clostridium cellulovorans* grown on various carbon sources. *Appl. Environ. Microbiol.* 79, 6576–6584. doi: 10.1128/AEM.02137-13

Mayer, F., Coughlan, M. P., Mori, Y., and Ljungdahl, L. G. (1987). Macromolecular Organization of the Cellulolytic Enzyme Complex of *Clostridium thermocellum* as revealed by Electron microscopy. *Appl. Environ. Microbiol.* 53, 2785–2792.

Mazzoli, R., and Olson, D. G. (2020). *Clostridium thermocellum*: A microbial platform for high-value chemical production from lignocellulose, *Advances in Applied Microbiology*, (eds.) G. M. Gadd and S. Sariaslani, (San Diego, CA, USA: Academic Press) 113, 111–161. doi: 10.1016/bs.aambs.2020.07.004

Minor, C. M., Takayesu, A., Ha, S. M., Salwinski, L., Sawaya, M. R., Pellegrini, M., et al. (2024). A genomic analysis reveals the diversity of cellulosome displaying bacteria. *Front. Microbiol.* 15:1473396. doi: 10.3389/fmicb.2024.1473396

Molinier, A.-L., Nouailler, M., Valette, O., Tardif, C., Receveur-Bréchot, V., and Fierobe, H.-P. (2011). Synergy, structure and conformational flexibility of hybrid Cellulosomes displaying various inter-cohesins linkers. *J. Mol. Biol.* 405, 143–157. doi: 10.1016/j.jmb.2010.10.013

Mondal, S., Thakur, A., Fontes, C. M. G. A., and Goyal, A. (2021). A trimodular family 16 glycoside hydrolase from the cellulosome of *Ruminococcus flavefaciens* displays highly specific licheninase (EC 3.2.1.73) activity. *Microbiology* 167:167. doi: 10.1099/mic.0.001055

Moraïs, S., Morag, E., Barak, Y., Goldman, D., Hadar, Y., Lamed, R., et al. (2012). Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes. *MBio* 3:e00508-12. doi: 10.1128/mBio.00508-12

Moraïs, S., Stern, J., Kahn, A., Galanopoulou, A. P., Yoav, S., Shamshoum, M., et al. (2016). Enhancement of cellulosome-mediated deconstruction of cellulose by improving enzyme thermostability. *Biotechnol. Biofuels* 9:164. doi: 10.1186/s13068-016-0577-z

Morisaka, H., Matsui, K., Tatsukami, Y., Kuroda, K., Miyake, H., Tamaru, Y., et al. (2012). Profile of native cellulosomal proteins of *Clostridium cellulovorans* adapted to various carbon sources. *AMB Express* 2:37. doi: 10.1186/2191-0855-2-37

Mosbah, A., Belaïch, A., Bornet, O., Belaïch, J.-P., Henrissat, B., and Darbon, H. (2000). Solution structure of the module X2_1 of unknown function of the cellulosomal scaffolding protein CipC of *Clostridium cellulolyticum*. *J. Mol. Biol.* 304, 201–217. doi: 10.1006/jmbi.2000.4192

Müller, M. M., Arndt, K. M., and Hoffmann, S. A. (2025). Genetic circuits in synthetic biology: broadening the toolbox of regulatory devices. *Front. Synth. Biol.* 3:1548572. doi: 10.3389/fsybi.2025.1548572

Murashima, K., Kosugi, A., and Doi, R. H. (2002). Determination of subunit composition of *Clostridium cellulovorans* Cellulosomes That degrade plant cell walls. *Appl. Environ. Microbiol.* 68, 1610–1615. doi: 10.1128/AEM.68.4.1610-1615.2002

Nakamura, A., Tasaki, T., Ishiwata, D., Yamamoto, M., Okuni, Y., Visootsat, A., et al. (2016). Single-molecule imaging analysis of binding, Processive movement, and dissociation of Cellobiohydrolase Trichoderma reesei Cel6A and its domains on crystalline cellulose. *J. Biol. Chem.* 291, 22404–22413. doi: 10.1074/jbc.M116.752048

Nakar, D., Handelsman, T., Shoham, Y., Fierobe, H.-P., Belaich, J.-P., Morag, E., et al. (2004). Pinpoint mapping of recognition residues on the Cohesin surface by progressive homologue swapping. *J. Biol. Chem.* 279, 42881–42888. doi: 10.1074/jbc.M407350200

Nemmaru, B., Ramirez, N., Farino, C. J., Yarbrough, J. M., Kravchenko, N., and Chundawat, S. P. S. (2021). Reduced type-a carbohydrate-binding module interactions to cellulose I leads to improved endocellulase activity. *Biotechnol. Bioeng.* 118, 1141–1151. doi: 10.1002/bit.27637

Noach, I., Frolow, F., Alber, O., Lamed, R., Shimon, L. J. W., and Bayer, E. A. (2009). Intermodular linker flexibility revealed from crystal structures of adjacent cellulosomal cohesins of acetivibrio cellulolyticus. *J. Mol. Biol.* 391, 86–97. doi: 10.1016/j.jmb.2009.06.006

Noach, I., Frolow, F., Jakoby, H., Rosenheck, S., Shimon, L. J. W., Lamed, R., et al. (2005). Crystal structure of a type-II cohesin module from the *Bacteroides cellulosolvens* cellulosome reveals novel and distinctive secondary structural elements. *J. Mol. Biol.* 348, 1–12. doi: 10.1016/j.jmb.2005.02.024

Ó Cuív, P., Gupta, R., Goswami, H. P., and Morrison, M. (2013). Extending the cellulosome paradigm: the modular *Clostridium thermocellum* cellulosomal serpin PinA is a broad-spectrum inhibitor of subtilisin-like proteases. *Appl. Environ. Microbiol.* 79, 6173–6175. doi: 10.1128/AEM.01912-13

Olson, D. G., Maloney, M. I., Lanahan, A. A., Cervenka, N. D., Xia, Y., Pech-Canul, A., et al. (2023). Ethanol tolerance in engineered strains of *Clostridium thermocellum*. *Biotechnol. Biofuels Bioprod.* 16:137. doi: 10.1186/s13068-023-02379-z

Pagès, S., Bélaïch, A., Bélaïch, J. P., Morag, E., Lamed, R., Shoham, Y., et al. (1997). Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. *Proteins* 29, 517–527.

Pagès, S., Valette, O., Abdou, L., Bélaïch, A., and Bélaïch, J.-P. (2003). A Rhamnogalacturonan Lyase in the *Clostridium cellulolyticum* Cellulosome. *J. Bacteriol.* 185, 4727–4733. doi: 10.1128/JB.185.16.4727-4733.2003

Pasari, N., Adlakha, N., Gupta, M., Bashir, Z., Rajacharya, G. H., Verma, G., et al. (2017). Impact of module-X2 and carbohydrate binding Module-3 on the catalytic activity of associated glycoside hydrolases towards plant biomass. *Sci. Rep.* 7:3700. doi: 10.1038/s41598-017-03927-y

Payne, C. M., Knott, B. C., Mayes, H. B., Hansson, H., Himmel, M. E., Sandgren, M., et al. (2015). Fungal Cellulases. *Chem. Rev.* 115, 1308–1448. doi: 10.1021/cr500351c

Pinheiro, B. A., Gilbert, H. J., Sakka, K., Sakka, K., Fernandes, V. O., Prates, J. A. M., et al. (2009). Functional insights into the role of novel type I cohesin and dockerin domains from *Clostridium thermocellum*. *Biochem. J.* 424, 375–384. doi: 10.1042/BJ20091152

Pinheiro, B. A., Proctor, M. R., Martinez-Fleites, C., Prates, J. A. M., Money, V. A., Davies, G. J., et al. (2008). The *Clostridium cellulolyticum* Dockerin displays a dual binding mode for its Cohesin partner. *J. Biol. Chem.* 283, 18422–18430. doi: 10.1074/jbc.M801533200

Qi, K., Chen, C., Yan, F., Feng, Y., Bayer, E. A., Kosugi, A., et al. (2021). Coordinated β-glucosidase activity with the cellulosome is effective for enhanced lignocellulose saccharification. *Bioresour. Technol.* 337:125441. doi: 10.1016/j.biortech.2021.125441

Raghothama, S., Eberhardt, R. Y., Simpson, P., Wigelsworth, D., White, P., Hazlewood, G. P., et al. (2001). Characterization of a cellulosome dockerin domain from the anaerobic fungus Piromyces equi. *Nat. Struct. Biol.* 8, 775–778. doi: 10.1038/nsb0901-775

Raman, B., McKeown, C. K., Rodriguez, M., Brown, S. D., and Mielenz, J. R. (2011). Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. *BMC Microbiol.* 11:134. doi: 10.1186/1471-2180-11-134

Raman, B., Pan, C., Hurst, G. B., Rodriguez, M., McKeown, C. K., Lankford, P. K., et al. (2009). Impact of pretreated switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 Cellulosome composition: a quantitative proteomic analysis. *PLoS One* 4:e5271. doi: 10.1371/journal.pone.0005271

Ran, Q., Zhang, D., Jiang, W., Zhang, H., Cheng, W., Li, H., et al. (2024). How carbohydrate-binding module affects the catalytic properties of endoglucanase. *Int. J. Biol. Macromol.* 278:134653. doi: 10.1016/j.ijbiomac.2024.134653

Ranjan, R., Rai, R., Bhatt, S. B., and Dhar, P. (2023). Technological road map of cellulase: a comprehensive outlook to structural, computational, and industrial applications. *Biochem. Eng. J.* 198:109020. doi: 10.1016/j.bej.2023.109020

Ravachol, J., Borne, R., Meynial-Salles, I., Soucaille, P., Pagès, S., Tardif, C., et al. (2015). Combining free and aggregated cellulolytic systems in the cellulosome-producing bacterium Ruminiclostridium cellulolyticum. *Biotechnol. Biofuels* 8:114. doi: 10.1186/s13068-015-0301-4

Reinikainen, T., Ruohonen, L., Nevanen, T., Laaksonen, L., Kraulis, P., Jones, T. A., et al. (1992). Investigation of the function of mutated cellulose-binding domains of *Trichoderma reesei* cellobiohydrolase I. *Proteins* 14, 475–482. doi: 10.1002/prot.340140408

Resch, M. G., Donohoe, B. S., Baker, J. O., Decker, S. R., Bayer, E. A., Beckham, G. T., et al. (2013). Fungal cellulases and complexed cellulosomal enzymes exhibit synergistic mechanisms in cellulose deconstruction. *Energy Environ. Sci.* 6:1858. doi: 10.1039/c3ee00019b

Resch, M. G., Donohoe, B. S., Ciesielski, P. N., Nill, J. E., Magnusson, L., Himmel, M. E., et al. (2014). Clean fractionation pretreatment reduces enzyme loadings for biomass saccharification and reveals the mechanism of free and cellulosomal enzyme synergy. *ACS Sustain. Chem. Eng.* 2, 1377–1387. doi: 10.1021/sc500210w

Riley, L. A., Ji, L., Schmitz, R. J., Westpheling, J., and Guss, A. M. (2019). Rational development of transformation in *Clostridium thermocellum* ATCC 27405 via complete methylome analysis and evasion of native restriction–modification systems. *J. Ind. Microbiol. Biotechnol.* 46, 1435–1443. doi: 10.1007/s10295-019-02218-x

Rincon, M. T., Cepeljnik, T., Martin, J. C., Barak, Y., Lamed, R., Bayer, E. A., et al. (2007). A novel cell surface-anchored cellulose-binding protein encoded by the sca gene cluster of *Ruminococcus flavefaciens*. *J. Bacteriol.* 189, 4774–4783. doi: 10.1128/JB.00143-07

Rincon, M. T., Čepeljnik, T., Martin, J. C., Lamed, R., Barak, Y., Bayer, E. A., et al. (2005). Unconventional mode of attachment of the *Ruminococcus flavefaciens* Cellulosome to the cell surface. *J. Bacteriol.* 187, 7569–7578. doi: 10.1128/JB.187.22.7569-7578.2005

Rincon, M. T., Ding, S.-Y., McCrae, S. I., Martin, J. C., Aurilia, V., Lamed, R., et al. (2003). Novel organization and divergent dockerin specificities in the cellulosome system of *Ruminococcus flavefaciens*. *J. Bacteriol.* 185, 703–713. doi: 10.1128/JB.185.3.703-713.2003

Ruiz, D. M., Turowski, V. R., and Murakami, M. T. (2016). Effects of the linker region on the structure and function of modular GH5 cellulases. *Sci. Rep.* 6:28504. doi: 10.1038/srep28504

Sakon, J., Irwin, D., Wilson, D. B., and Karplus, P. A. (1997). Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. *Nat. Struct. Biol.* 4, 810–818.

Salama-Alber, O., Jobby, M. K., Chitayat, S., Smith, S. P., White, B. A., Shimon, L. J. W., et al. (2013). Atypical cohesin-dockerin complex responsible for cell surface attachment of cellulosomal components: binding fidelity, promiscuity, and structural buttresses. *J. Biol. Chem.* 288, 16827–16838. doi: 10.1074/jbc.M113.466672

Sammond, D. W., Payne, C. M., Brunecky, R., Himmel, M. E., Crowley, M. F., and Beckham, G. T. (2012). Cellulase linkers are optimized based on domain type and function: insights from sequence analysis, biophysical measurements, and molecular simulation. *PLoS One* 7:e48615. doi: 10.1371/journal.pone.0048615

Sand, A., Holwerda, E. K., Ruppertsberger, N. M., Maloney, M., Olson, D. G., Nataf, Y., et al. (2015). Three cellulosomal xylanase genes in *Clostridium thermocellum* are regulated by both vegetative SigA (σ A) and alternative SigI6 (σ I6) factors. *FEBS Lett.* 589, 3133–3140. doi: 10.1016/j.febslet.2015.08.026

Sharma, J., Kumar, V., Prasad, R., and Gaur, N. A. (2022). Engineering of *Saccharomyces cerevisiae* as a consolidated bioprocessing host to produce cellulosic ethanol: recent advancements and current challenges. *Biotechnol. Adv.* 56:107925. doi: 10.1016/j.biotechadv.2022.107925

Sidar, A., Albuquerque, E. D., Voshol, G. P., Ram, A. F. J., Vijgenboom, E., and Punt, P. J. (2020). Carbohydrate binding modules: diversity of domain architecture in amylases and Cellulases from filamentous microorganisms. *Front. Bioeng. Biotechnol.* 8:871. doi: 10.3389/fbioe.2020.00871

Slutzki, M., Reshef, D., Barak, Y., Haimovitz, R., Rotem-Bamberger, S., Lamed, R., et al. (2015). Crucial roles of single residues in binding affinity, specificity, and promiscuity in the Cellulosomal Cohesin-Dockerin Interface. *J. Biol. Chem.* 290, 13654–13666. doi: 10.1074/jbc.M115.651208

Smith, S. P., and Bayer, E. A. (2013). Insights into cellulosome assembly and dynamics: from dissection to reconstruction of the supramolecular enzyme complex. *Curr. Opin. Struct. Biol.* 23, 686–694. doi: 10.1016/j.sbi.2013.09.002

Song, X., Zhang, J., Fu, S., Liu, Z., Chen, Y., and Zhu, T. (2025). Improving cellulosic ethanol production by an engineered yeast consortium displaying a pentafunctional mini-cellulosome. *FEMS Yeast Res.* 25:foaf022. doi: 10.1093/femsyr/foaf022

Srivastava, A., Nagar, P., Rathore, S., and Adlakha, N. (2022). The linker region promotes activity and binding efficiency of modular LPMO towards polymeric substrate. *Microbiol Spectr* 10:e0269721. doi: 10.1128/spectrum.02697-21

Stahl, S. W., Nash, M. A., Fried, D. B., Slutzki, M., Barak, Y., Bayer, E. A., et al. (2012). Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. USA* 109, 20431–20436. doi: 10.1073/pnas.1211929109

Stern, J., Kahn, A., Vazana, Y., Shamshoum, M., Moraïs, S., Lamed, R., et al. (2015). Significance of relative position of cellulases in designer cellulosomes for optimized cellulolysis. *PLoS One* 10:e0127326. doi: 10.1371/journal.pone.0127326

Stern, J., Moraïs, S., Ben-David, Y., Salama, R., Shamshoum, M., Lamed, R., et al. (2018). Assembly of synthetic functional Cellulosomal structures onto the cell surface of *Lactobacillus plantarum*, a potent member of the gut microbiome. *Appl. Environ. Microbiol.* 84:e00282-18. doi: 10.1128/AEM.00282-18

Støpamo, F. G., Sulaeva, I., Budischowsky, D., Rahikainen, J., Marjamaa, K., Kruus, K., et al. (2024). The impact of the carbohydrate-binding module on how a lytic polysaccharide monooxygenase modifies cellulose fibers. *Biotechnol. Biofuels Bioprod.* 17:118. doi: 10.1186/s13068-024-02564-8

Tamaru, Y., Miyake, H., Kuroda, K., Nakanishi, A., Matsushima, C., Doi, R. H., et al. (2011). Comparison of the mesophilic cellulosome-producing *Clostridium cellulovorans* genome with other cellulosome-related clostridial genomes. *Microb. Biotechnol.* 4, 64–73. doi: 10.1111/j.1751-7915.2010.00210.x

Tang, C.-D., Li, J.-F., Wei, X.-H., Min, R., Gao, S.-J., Wang, J.-Q., et al. (2013). Fusing a carbohydrate-binding module into the aspergillus usamii β-Mannanase to improve its Thermostability and cellulose-binding capacity by in silico design. *PLoS One* 8:e64766. doi: 10.1371/journal.pone.0064766

Tang, H., Wang, J., Wang, S., Shen, Y., Petranovic, D., Hou, J., et al. (2018). Efficient yeast surface-display of novel complex synthetic cellulosomes. *Microb. Cell Factories* 17:122. doi: 10.1186/s12934-018-0971-2

Tang, B., Wei, L., Tang, W., Li, S., and Zhou, R. (2015). Effect of linker flexibility on the catalytic features of a glycoside hydrolase family 45 endoglucanase from Rhizopus stolonifer. *Appl. Biochem. Biotechnol.* 176, 2242–2252. doi: 10.1007/s12010-015-1711-2

Tao, X., Liu, J., Kempher, M. L., Xu, T., and Zhou, J. (2022). In vivo functional characterization of hydrophilic X2 modules in the Cellulosomal scaffolding protein. *Front. Microbiol.* 13:861549. doi: 10.3389/fmicb.2022.861549

Tarraran, L., Gandini, C., Luganini, A., and Mazzoli, R. (2021). Cell-surface binding domains from *Clostridium cellulovorans* can be used for surface display of cellulosomal scaffoldins in *Lactococcus lactis*. *Biotechnol. J.* 16:e2100064. doi: 10.1002/biot.202100064

Tatli, M., Moraïs, S., Tovar-Herrera, O. E., Bomble, Y. J., Bayer, E. A., Medalia, O., et al. (2022). Nanoscale resolution of microbial fiber degradation in action. *eLife* 11:e76523. doi: 10.7554/eLife.76523

Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., et al. (2010). Synthetic riboswitches That induce gene expression in diverse bacterial species. *Appl. Environ. Microbiol.* 76, 7881–7884. doi: 10.1128/AEM.01537-10

Tormo, J., Lamed, R., Chirino, A. J., Morag, E., Bayer, E. A., Shoham, Y., et al. (1996). Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.* 15, 5739–5751.

Tsai, S.-L., DaSilva, N. A., and Chen, W. (2013). Functional display of complex Cellulosomes on the yeast surface via adaptive assembly. *ACS Synth. Biol.* 2, 14–21. doi: 10.1021/sb300047u

Tsai, S.-L., Sun, Q., and Chen, W. (2022). Advances in consolidated bioprocessing using synthetic cellulosomes. *Curr. Opin. Biotechnol.* 78:102840. doi: 10.1016/j.copbio.2022.102840

Valbuena, A., Oroz, J., Hervás, R., Vera, A. M., Rodríguez, D., Menéndez, M., et al. (2009). On the remarkable mechanostability of scaffoldins and the mechanical clamp motif. *Proc. Natl. Acad. Sci.* 106, 13791–13796. doi: 10.1073/pnas.0813093106

Vanderstraeten, J., da Fonseca, M. J. M., De Groote, P., Grimon, D., Gerstmans, H., Kahn, A., et al. (2022a). Combinatorial assembly and optimisation of designer cellulosomes: a galactomannan case study. *Biotechnol. Biofuels Bioprod.* 15:60. doi: 10.1186/s13068-022-02158-2

Vanderstraeten, J., Lamote, B., da Fonseca, M. J. M., De Groote, P., and Briers, Y. (2022b). Conversion of the free *Cellvibrio japonicus* xyloglucan degradation system to the cellulosomal mode. *Appl. Microbiol. Biotechnol.* 106, 5495–5509. doi: 10.1007/s00253-022-12072-0

Vazana, Y., Barak, Y., Unger, T., Peleg, Y., Shamshoum, M., Ben-Yehezkel, T., et al. (2013). A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. *Biotechnol. Biofuels* 6:182. doi: 10.1186/1754-6834-6-182

Vazana, Y., Moraïs, S., Barak, Y., Lamed, R., and Bayer, E. A. (2010). Interplay between *Clostridium thermocellum* family 48 and family 9 Cellulases in Cellulosomal versus Noncellulosomal states. *Appl. Environ. Microbiol.* 76, 3236–3243. doi: 10.1128/AEM.00009-10

Venditto, I., Luis, A. S., Rydahl, M., Schückel, J., Fernandes, V. O., Vidal-Melgosa, S., et al. (2016). Complexity of the *Ruminococcus flavefaciens* cellulosome reflects an expansion in glycan recognition. *Proc. Natl. Acad. Sci.* 113, 7136–7141. doi: 10.1073/pnas.1601558113

Vera, A. M., Galera-Prat, A., Wojciechowski, M., Różycki, B., Laurents, D. V., Carrión-Vázquez, M., et al. (2021). Cohesin-dockerin code in cellulosomal dual binding modes and its allosteric regulation by proline isomerization. *Structure* 29, 587–597.e8. doi: 10.1016/j.str.2021.01.006

Verdorfer, T., Bernardi, R. C., Meinhold, A., Ott, W., Luthey-Schulten, Z., Nash, M. A., et al. (2017). Combining *in vitro* and *in silico* single-molecule force spectroscopy to characterize and tune Cellulosomal Scaffoldin mechanics. *J. Am. Chem. Soc.* 139, 17841–17852. doi: 10.1021/jacs.7b07574

Vita, N., Borne, R., Perret, S., de Philip, P., and Fierobe, H. (2019). Turning a potent family-9 free cellulase into an operational cellulosomal component and vice versa. *FEBS J.* 286, 3359–3373. doi: 10.1111/febs.14858

Vodovnik, M., Duncan, S. H., Reid, M. D., Cantlay, L., Turner, K., Parkhill, J., et al. (2013). Expression of cellulosome components and type IV pili within the extracellular proteome of *Ruminococcus flavefaciens* 007. *PLoS One* 8:e65333. doi: 10.1371/journal.pone.0065333

Vodovnik, M., and Lindič, N. (2025). Towards the application of nature's catalytic nanomachines: Cellulosomes in 2nd generation biofuel production. *Biotechnol. Adv.* 79:108523. doi: 10.1016/j.biotechadv.2025.108523

Vodovnik, M., and Marinšek-Logar, R. (2010). Cellulosomes - promising supramolecular machines of anaerobic cellulolytic microorganisms. *Acta Chim. Slov.* 57, 767–774

von Schantz, L., Håkansson, M., Logan, D. T., Walse, B., Österlin, J., Nordberg-Karlsson, E., et al. (2012). Structural basis for carbohydrate-binding specificity—a comparative assessment of two engineered carbohydrate-binding modules. *Glycobiology* 22, 948–961. doi: 10.1093/glycob/cws063

Voronov-Goldman, M., Yaniv, O., Gul, O., Yoffe, H., Salama-Alber, O., Slutzki, M., et al. (2015). Standalone cohesin as a molecular shuttle in cellulosome assembly. *FEBS Lett.* 589, 1569–1576. doi: 10.1016/j.febslet.2015.04.013

Walker, J. A., Takasuka, T. E., Deng, K., Bianchetti, C. M., Udell, H. S., Prom, B. M., et al. (2015). Multifunctional cellulase catalysis targeted by fusion to different carbohydrate-binding modules. *Biotechnol. Biofuels* 8:220. doi: 10.1186/s13068-015-0402-0

Wang, Y., Leng, L., Islam, M. K., Liu, F., Lin, C. S. K., and Leu, S.-Y. (2019). Substrate-related factors affecting Cellulosome-induced hydrolysis for lignocellulose valorization. *Int. J. Mol. Sci.* 20:3354. doi: 10.3390/ijms20133354

Wang, Y., Li, L., Xia, Y., and Zhang, T. (2022). Reliable and scalable identification and prioritization of putative cellulolytic anaerobes with large genome data. *Front. Bioinfo.* 2:813771. doi: 10.3389/fbinf.2022.813771

Weinstein, J. Y., Slutzki, M., Karpol, A., Barak, Y., Gul, O., Lamed, R., et al. (2015). Insights into a type III cohesin–dockerin recognition interface from the cellulose-degrading bacterium *Ruminococcus flavefaciens*. *J. Mol. Recognit.* 28, 148–154. doi: 10.1002/jmr.2380

Wilson, C. M., Rodriguez, M., Johnson, C. M., Martin, S. L., Chu, T. M., Wolfinger, R. D., et al. (2013). Global transcriptome analysis of *Clostridium thermocellum* ATCC 27405 during growth on dilute acid pretreated Populus and switchgrass. *Biotechnol. Biofuels* 6:179. doi: 10.1186/1754-6834-6-179

Wojciechowski, M., Różycki, B., Huy, P. D. Q., Li, M. S., Bayer, E. A., and Cieplak, M. (2018). Dual binding in cohesin-dockerin complexes: the energy landscape and the role of short, terminal segments of the dockerin module. *Sci. Rep.* 8:5051. doi: 10.1038/s41598-018-23380-9

Wu, T.-H., Huang, C.-H., Ko, T.-P., Lai, H.-L., Ma, Y., Chen, C.-C., et al. (2011). Diverse substrate recognition mechanism revealed by *Thermotoga maritima* Cel5A structures in complex with cellotetraose, cellobiose and mannotriose. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1814, 1832–1840. doi: 10.1016/j.bbapap.2011.07.020

Xu, T., Li, Y., He, Z., and Zhou, J. (2014). Dockerin-containing protease inhibitor protects key cellulosomal cellulases from proteolysis in *Clostridium cellulolyticum*. *Mol. Microbiol.* 91, 694–705. doi: 10.1111/mmi.12488

Xu, Q., Resch, M. G., Podkaminer, K., Yang, S., Baker, J. O., Donohoe, B. S., et al. (2016). Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv.* 2:e1501254. doi: 10.1126/sciadv.1501254

Yaniv, O., Fichman, G., Borovok, I., Shoham, Y., Bayer, E. A., Lamed, R., et al. (2014). Fine-structural variance of family 3 carbohydrate-binding modules as extracellular biomass-sensing components of *Clostridium thermocellum* anti-σ [1] factors. *Acta Crystallogr. D Biol. Crystallogr.* 70, 522–534. doi: 10.1107/S139900471302926X

Yaniv, O., Petkun, S., Shimon, L. J. W., Bayer, E. A., Lamed, R., and Frolow, F. (2012). A single mutation reforms the binding activity of an adhesion-deficient family 3 carbohydrate-binding module. *Acta Crystallogr. D Biol. Crystallogr.* 68, 819–828. doi: 10.1107/S0907444912013133

Yaniv, O., Shimon, L. J. W., Bayer, E. A., Lamed, R., and Frolow, F. (2011). Scaffoldin-borne family 3b carbohydrate-binding module from the cellulosome of *Bacteroides cellulosolvens*: structural diversity and significance of calcium for carbohydrate binding. *Acta Crystallogr. D Biol. Crystallogr.* 67, 506–515. doi: 10.1107/S0907444911011322

Yao, X., Chen, C., Wang, Y., Dong, S., Liu, Y.-J., Li, Y., et al. (2020). Discovery and mechanism of a pH-dependent dual-binding-site switch in the interaction of a pair of protein modules. *Sci. Adv.* 6:eabd7182. doi: 10.1126/sciadv.abd7182

Ye, T. J., Huang, K. F., Ko, T. P., and Wu, S. H. (2022). Synergic action of an inserted carbohydrate-binding module in a glycoside hydrolase family 5 endoglucanase. *Acta Crystallogr D Struct Biol* 78, 633–646. doi: 10.1107/S2059798322002601

Yoav, S., Barak, Y., Shamshoum, M., Borovok, I., Lamed, R., Dassa, B., et al. (2017). How does cellulosome composition influence deconstruction of lignocellulosic substrates in Clostridium (Ruminiclostridium) thermocellum DSM 1313? *Biotechnol. Biofuels* 10:222. doi: 10.1186/s13068-017-0909-7

You, Y., Kong, H., Li, C., Gu, Z., Ban, X., and Li, Z. (2024). Carbohydrate binding modules: compact yet potent accessories in the specific substrate binding and performance evolution of carbohydrate-active enzymes. *Biotechnol. Adv.* 73:108365. doi: 10.1016/j.biotechadv.2024.108365

You, C., Zhang, X.-Z., Sathitsuksanoh, N., Lynd, L. R., and Zhang, Y.-H. P. (2012). Enhanced microbial utilization of recalcitrant cellulose by an *ex vivo* Cellulosome-microbe complex. *Appl. Environ. Microbiol.* 78, 1437–1444. doi: 10.1128/AEM.07138-11

Yu, G., Liu, S., Feng, X., Zhang, Y., Liu, C., Liu, Y.-J., et al. (2020). Impact of ammonium sulfite-based sequential pretreatment combinations on two distinct saccharifications of wheat straw. *RSC Adv.* 10, 17129–17142. doi: 10.1039/D0RA01759K

Zajki-Zechmeister, K., Eibinger, M., Kaira, G. S., and Nidetzky, B. (2024). Mechanochemical coupling of catalysis and motion in a cellulose-degrading multienzyme Nanomachine. *ACS Catal.* 14, 2656–2663. doi: 10.1021/acscatal.3c05653

Zhang, K.-D., Li, W., Wang, Y.-F., Zheng, Y.-L., Tan, F.-C., Ma, X.-Q., et al. (2018). Processive degradation of crystalline cellulose by a multimodular endoglucanase via a Wirewalking mode. *Biomacromolecules* 19, 1686–1696. doi: 10.1021/acs.biomac.8b00340

Zhang, J., Liu, S., Li, R., Hong, W., Xiao, Y., Feng, Y., et al. (2017). Efficient whole-cell-catalyzing cellulose saccharification using engineered *Clostridium thermocellum*. *Biotechnol. Biofuels* 10:124. doi: 10.1186/s13068-017-0796-y

Zhao, G., Ali, E., Sakka, M., Kimura, T., and Sakka, K. (2006). Binding of S-layer homology modules from *Clostridium thermocellum* SdbA to peptidoglycans. *Appl. Microbiol. Biotechnol.* 70, 464–469. doi: 10.1007/s00253-005-0079-6

Zhivin, O., Dassa, B., Moraïs, S., Utturkar, S. M., Brown, S. D., Henrissat, B., et al. (2017). Unique organization and unprecedented diversity of the Bacteroides (Pseudobacteroides) cellulosolvens cellulosome system. *Biotechnol. Biofuels* 10:211. doi: 10.1186/s13068-017-0898-6

Zhivin-Nissan, O., Dassa, B., Morag, E., Kupervaser, M., Levin, Y., and Bayer, E. A. (2019). Unraveling essential cellulosomal components of the (Pseudo)*Bacteroides cellulosolvens* reveals an extensive reservoir of novel catalytic enzymes. *Biotechnol. Biofuels* 12:115. doi: 10.1186/s13068-019-1447-2

Zhou, P., Gao, C., Song, W., Wei, W., Wu, J., Liu, L., et al. (2024). Engineering status of protein for improving microbial cell factories. *Biotechnol. Adv.* 70:108282. doi: 10.1016/j.biotechadv.2023.108282

Zverlov, V. V., Fuchs, K.-P., and Schwarz, W. H. (2002). Chi18A, the Endochitinase in the Cellulosome of the thermophilic, cellulolytic bacterium *Clostridium thermocellum*. *Appl. Environ. Microbiol.* 68, 3176–3179. doi: 10.1128/AEM.68.6.3176-3179.2002

Zverlov, V. V., Velikodvorskaya, G. A., and Schwarz, W. H. (2003). Two new cellulosome components encoded downstream of celI in the genome of *Clostridium thermocellum*: the non-processive endoglucanase CelN and the possibly structural protein CseP. *Microbiology* 149, 515–524. doi: 10.1099/mic.0.25959-0