



Minimalistic Cellulosome of the Butanogenic Bacterium *Clostridium saccharoperbutylacetonicum*

Bosmat Levi Hevroni,^a Sarah Morais,^{a,b} Yonit Ben-David,^a Ely Morag,^a  Edward A. Bayer^a

^aDepartment of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, Israel

^bDepartment of Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel

ABSTRACT *Clostridium saccharoperbutylacetonicum* is a mesophilic, anaerobic, butanol-producing bacterium, originally isolated from soil. It was recently reported that *C. saccharoperbutylacetonicum* possesses multiple cellulosomal elements and would potentially form the smallest cellulosome known in nature. Its genome contains only eight dockerin-bearing enzymes, and its unique scaffoldin bears two cohesins (Cohs), three X2 modules, and two carbohydrate-binding modules (CBMs). In this study, all of the cellulosome-related modules were cloned, expressed, and purified. The recombinant cohesins, dockerins, and CBMs were tested for binding activity using enzyme-linked immunosorbent assay (ELISA)-based techniques. All the enzymes were tested for their comparative enzymatic activity on seven different cellulosic and hemicellulosic substrates, thus revealing four cellulases, a xylanase, a mannanase, a xyloglucanase, and a lichenase. All dockerin-containing enzymes interacted similarly with the second cohesin (Coh2) module, whereas Coh1 was more restricted in its interaction pattern. In addition, the polysaccharide-binding properties of the CBMs within the scaffoldin were examined by two complementary assays, affinity electrophoresis and affinity pulldown. The scaffoldin of *C. saccharoperbutylacetonicum* exhibited high affinity for cellulosic and hemicellulosic substrates, specifically to microcrystalline cellulose and xyloglucan. Evidence that supports substrate-dependent *in vivo* secretion of cellulosomes is presented. The results of our analyses contribute to a better understanding of simple cellulosome systems by identifying the key players in this minimalistic system and the binding pattern of its cohesin-dockerin interaction. The knowledge gained by our study will assist further exploration of similar minimalistic cellulosomes and will contribute to the significance of specific sets of defined cellulosomal enzymes in the degradation of cellulosic biomass.

IMPORTANCE Cellulosome-producing bacteria are considered among the most important bacteria in both mesophilic and thermophilic environments, owing to their capacity to deconstruct recalcitrant plant-derived polysaccharides (and notably cellulose) into soluble saccharides for subsequent processing. In many ecosystems, the cellulosome-producing bacteria are particularly effective “first responders.” The massive amounts of sugars produced are potentially amenable in industrial settings to further fermentation by appropriate microbes to biofuels, notably ethanol and butanol. Among the solvent-producing bacteria, *Clostridium saccharoperbutylacetonicum* has the smallest cellulosome system known thus far. The importance of investigating the building blocks of such a small, multifunctional nanomachine is crucial to understanding the fundamental activities of this efficient enzymatic complex.

KEYWORDS glycoside hydrolases, enzymatic profiling, scaffoldin, cohesin-dockerin specificity, CBM, cohesin specificity

Citation Levi Hevroni B, Morais S, Ben-David Y, Morag E, Bayer EA. 2020. Minimalistic cellulosome of the butanogenic bacterium *Clostridium saccharoperbutylacetonicum*. mBio 11:e00443-20. <https://doi.org/10.1128/mBio.00443-20>.

Editor María Mercedes Zambrano, Corporación CorpoGen

Copyright © 2020 Levi Hevroni et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Edward A. Bayer, ed.bayer@weizmann.ac.il.

This article is a direct contribution from Edward A. Bayer, a Fellow of the American Academy of Microbiology, who arranged for and secured reviews by Isaac Cann, University of Illinois at Urbana-Champaign, and Brian Davison, Oak Ridge National Laboratory.

Received 25 February 2020

Accepted 5 March 2020

Published 31 March 2020

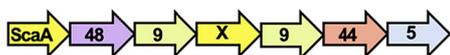
Cellulosomes are high-molecular-weight, multienzyme complexes, which are deemed important for efficient bacterial degradation of recalcitrant polysaccharides, notably cellulose. These highly efficient complexes are produced by specific cellulolytic anaerobic bacteria (1). These specialized bacteria can be found in various environments in nature, such as soil, compost piles, sewage sludge, and the digestive tracts of insects and herbivorous animals (2). The cellulosome system is based on several unique and interacting functional and structural subunits. The functional catalytic components (i.e., the enzymatic system) are attached to noncatalytic, structural scaffoldin proteins. The assembly between the scaffoldin and enzymes occurs via the complementary interaction of noncatalytic modules called cohesins (located on the scaffoldin) and dockerins (modular component of the enzymes). Frequently, scaffoldin subunits can also contain an additional functional component, a carbohydrate-binding module (CBM). The CBM assists the bacterium in attaching to cellulose, so that the cellulosome gains proximity to both its substrate and the bacterial cell wall (1).

Cellulosome architecture can vary from “complex” cellulosomes, composed of multiple interacting scaffoldins which assemble into a much more intricate cellulosome complex (1, 2), to “simple” cellulosomes that contain a single major enzyme-bearing scaffoldin subunit. Complex cellulosome systems can be found both in thermophilic bacteria, e.g., *Clostridium clariflavum*, and in the mesophilic *Bacteroides* (*Pseudobacteroides*) *cellulosolvens*. The *C. clariflavum* cellulosomal system has 13 scaffoldins and 75 dockerin-containing proteins (mainly enzymes) (3), whereas *B. cellulosolvens* has 212 dockerin-bearing proteins and 32 scaffoldins (4). On the other hand, much simpler and smaller cellulosome systems can be found in the mesophilic bacteria *Clostridium bornimense* and *Clostridium acetobutylicum*, which have 5 dockerins and 2 scaffoldins in the former and 10 dockerins and 2 scaffoldins in the latter (5). In addition to the major scaffoldin, which bears multiple cohesins and one or two CBMs, the latter mesophilic cellulosome-producing bacteria each contains a gene termed *orfX* encoding a single cohesin. This type of protein can thus also be considered a scaffoldin, but its precise role in cellulosome assembly is unclear (6).

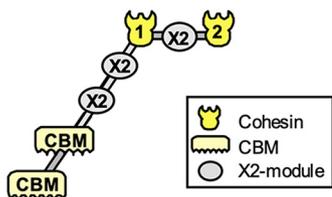
In recent years, genome sequencing of microorganisms has become more accessible and standardized. This progress contributes to the discovery of new cellulosomal species, thus expanding our knowledge of the cellulosomal paradigm in nature. In a recent review, a genome-wide analysis of a dozen mesophilic clostridial species revealed the presence of heretofore undescribed cellulosome systems in three mesophilic clostridia (5). In particular, the cellulosomal system of *Clostridium saccharoperbutylacetonicum* was of special interest to us. *C. saccharoperbutylacetonicum* is a mesophilic, anaerobic, spore-forming, butanol-producing Gram-positive bacterium originally isolated from soil (7), and its potential to use lignocellulosic substrates was never reported. There are several other known solvent-producing clostridia, such as *C. acetobutylicum*, *C. saccharobutylicum*, *C. beijerinckii*, etc. Among these solventogenic clostridia, only *C. acetobutylicum* and *C. saccharoperbutylacetonicum* are known to possess cellulosomal elements. However, despite the presence of major dockerin-bearing cellulases, the cellulosome of *C. acetobutylicum* was reported to be inactive on cellulosic substrates, and the extracellular cellulolytic activity was limited and/or regulated by the available substrate (8–10). Surprisingly, the cellulosome system of *C. saccharoperbutylacetonicum* was never properly examined. Thus, we were intrigued by this bacterium and its cellulosomal system, which is so small and simple and might shed light on the basic elements of a cellulosome.

The complete and fully annotated genome of *C. saccharoperbutylacetonicum* strain N1-4 (HMT) was published in 1999, and it was found to comprise two replicons, a chromosome, and a circular megaplasmid (11). The solventogenic *sol* operon of *C. saccharoperbutylacetonicum* is located on the chromosome, as in the noncellulosome producer, *Clostridium beijerinckii*, but in contrast to that of *Clostridium acetobutylicum*, which is located on the megaplasmid. Like *C. beijerinckii*, *C. saccharoperbutylacetonicum* possesses an aldehyde dehydrogenase (*ald*) gene, whereas *C. acetobutylicum* has an alcohol/aldehyde dehydrogenase-encoding gene (*adhE*) together with acetoacetate

A. Cellulosomal gene cluster



B. Cip scaffoldin



C. Dockerin sequences



FIG 1 The cellulosome system of *C. saccharoperbutylacetonicum* (5). (A) Schematic representation of the cellulosome gene cluster. The gene encoding the major scaffoldin is followed downstream by genes coding for dockerin-containing cellulolytic enzymes. The major scaffoldin gene is designated *scaA*; the numbers denote the GH family; X stands for the *orfX* gene (a similar gene encoding a cohesin-containing protein is present in all cellulosome-producing bacteria). (B) Schematic representation of the modular composition of the primary scaffoldin protein. (C) Conserved sequence features of dockerin modules, visualized by WebLogo. Ca^{2+} -binding residues are shown on cyan background, and putative cohesin recognition positions are shown on yellow background. Cohesin recognition residues (XXQK/TARK) are not repeated in symmetric fashion, thus suggesting a single rather than a dual binding mode of interaction. The recognition residues are unique in the current collection of dockerin sequences.

decarboxylase (*adc*), which are located on separate operons (12). The recently reported cellulosomal elements of *C. saccharoperbutylacetonicum* (5) are located in the cellulosome gene cluster of the chromosome. Theoretically, its cellulosome is among the smallest known in nature, comprising a “minimalistic” cellulosome with only eight dockerin-bearing enzymes (five of which are located in the major cellulosome gene cluster (Fig. 1A), and the remaining three are distributed elsewhere on the chromosome) and only two cohesins in its single major scaffoldin (Fig. 1B). To date, this represents the smallest number of cellulosomal scaffoldin-borne cohesins among the mesophilic clostridia (5). For comparison, *C. acetobutylicum* which is considered to have a small and simple cellulosome, has 10 dockerin-containing enzymes and 5 cohesins on its major scaffoldin. Interestingly, *C. saccharoperbutylacetonicum* major scaffoldin contains two copies of CBM3 at the N terminus—similar to the major scaffoldin in *C. bornimense* but unlike all other known mesophilic cellulosome-producing bacteria which have only a single scaffoldin-borne CBM3. The *C. saccharoperbutylacetonicum* scaffoldin also bears three X2 domains (a module of ca. 100 amino acid residues of unknown structure and function) (13) between its CBM3 and cohesin modules (similar only to the major scaffoldin of *C. acetobutylicum*) (5). In addition, the *C. saccharoperbutylacetonicum* genome encodes 146 free glycoside hydrolase (GH) enzymes.

In this study, we aimed to reveal the full architecture and function of *C. saccharoperbutylacetonicum* cellulosome system. Therefore, each cellulosome-related module (cohesin, dockerin, catalytic, and carbohydrate-binding modules) was cloned, expressed in *Escherichia coli*, and purified, and the recombinant components were tested for enzymatic activity on cellulosic and hemicellulosic substrates. The binding capacity of each cohesin and dockerin module was tested using enzyme-linked immunosorbent

TABLE 1 Putative dockerin-containing glycoside hydrolases of *C. saccharoperbutylacetonicum*^a

GH family	Current name	Modular organization	Mol wt ^b	NCBI:protein accession no.
GH5	GH5A	GH5A-Doc	54,258	WP_015391959
	GH5B	GH5B-Doc	50,917	WP_015391954
GH9	GH9A	CBM4-Ig-GH9A-Doc	93,651	WP_015391950
	GH9B	GH9B-CBM3c-Doc	77,973	WP_015391952
GH26	GH26A	CBM6-GH26A-Doc	40,989	WP_015391960
GH44	GH44A	GH44A-Doc	65,499	WP_015391953
GH48	GH48A	GH48A-Doc	81,337	WP_015391949
GH74	GH74A	GH74A-Doc	90,254	WP_015391947

^aAbbreviations: GH, glycoside hydrolase; Doc, dockerin; CBM, carbohydrate-binding module; Ig, immunoglobulin-like domain of unknown function.

^bThe molecular weight based on the known amino acid composition of the desired protein using the ProtParam tool (15).

assay (ELISA)-based techniques in order to evaluate their pattern of interaction. In addition, the CBMs of the scaffoldin and of the cellulosomal enzymes were explored for their binding affinities.

RESULTS

The cellulosomal system of *C. saccharoperbutylacetonicum*. The bioinformatic analysis of *C. saccharoperbutylacetonicum* revealed a set of only eight putative dockerin-containing glycoside hydrolases (GHs) (Table 1), and one major scaffoldin, ScaA, which possess two family 3 CBMs, three X2 modules, and two cohesins. The eight putative enzymes belong to GH families 5, 9, 26, 44, 48, and 74. All putative dockerin-containing enzymes and the scaffoldin protein have an N-terminal signal peptide sequence, suggesting that these proteins are secreted.

In order to characterize the *C. saccharoperbutylacetonicum* cellulosomal components, the following proteins were cloned (without their signal peptide sequence) and expressed: the eight dockerin-containing enzymes GH5A, GH5B, GH9A, GH9B, GH26A, GH44A, GH48A, and GH74A and the putative scaffoldin (ScaA). Since the X2 modules could have a stabilizing role for the neighboring cohesin (14), the two ScaA cohesins were also cloned and expressed separately with and without their X modules as follows: Coh1, X2-Coh1, Coh2, and X2-Coh2. All proteins were purified to homogeneity with a major band on sodium dodecyl sulfate-polyacrylamide gel (SDS-PAGE) at the expected calculated molecular weights (see Fig. S1 in the supplemental material).

Activity profiling of *C. saccharoperbutylacetonicum* cellulosomal enzymes. In order to identify the preferred substrate(s) of the *C. saccharoperbutylacetonicum* dockerin-containing enzymes, all enzymes were tested for their activity on various cellulosic and hemicellulosic substrates, according to their putative activity (as reported for their respective GH family [15]). This approach was used previously to gain initial assessment of enzyme activity in *Ruminococcus champanellensis* (16). Since the activity of an enzyme is influenced by environmental conditions (e.g., pH, temperature) (17), an optimization process was first performed to determine optimal conditions of activity.

Optimizing the activity assay parameters. For the optimization process of the enzymatic activity assays, we chose two representative enzymes, GH44A (originally considered a putative cellulase) and GH5B (a putative mannanase). Each enzyme was tested with its corresponding substrate: GH44A was tested for its ability to cleave carboxymethyl cellulose (CMC) and GH5B to cleave locust bean gum (mannan). The enzymes were active on each substrate in accordance with the other characterized members of the same GH families (15). In order to estimate the optimal parameters for enzyme activity (e.g., pH, temperature), four temperatures were tested: 30, 40, 50, and 60°C. We also tested three buffers at different pH values (according to the buffer type): acetate buffer pH 4 and 5, citrate buffer pH 5 and 6, and phosphate buffer pH 6 and 7. Our results revealed that the highest activity obtained for both enzymes was with the acetate buffer under conditions of pH 5 at 40°C; therefore, these conditions were

selected for the biochemical characterization of the other enzymes. The results for GH5B and GH44A are presented in Fig. S2A and S2B, respectively.

Enzymatic activity profile. The substrates tested were in accordance with the putative activity of the enzymes according to the CAZy classification (15), and all of the purified recombinant enzymes were tested for their activity(ies) on all of the tested substrates. The enzymes GH5A, GH9A, GH9B, GH44A, and GH48A were all predicted to be cellulases according to CAZy and bioinformatic analysis. Therefore, they were expected to show activity on some or all of the tested cellulosic substrates: Avicel (microcrystalline cellulose), phosphoric acid-swollen cellulose (PASC), and CMC. Indeed, all of the putative cellulases exhibited activity on Avicel and PASC (Fig. 2A and B). GH5A gave the highest activity on both substrates. GH44A also presented high activity, whereas GH9A and GH9B showed moderate activity. GH48A showed very low activity on all cellulosic substrates, as expected for this particular putative exoglucanase, based on previous performance of the GH48s from other bacterial species (18–21). GH9B exhibited higher relative activity on Avicel versus PASC. Among the five predicted cellulases, three were highly active on the soluble CMC substrate: GH5A, GH9B, and GH44A (Fig. 2C). Therefore, these three can be classified as endoglucanases or processive endoglucanases (the GH9B enzyme). In contrast, GH9A and GH48A showed only nominal activity on CMC versus Avicel and PASC and are thus considered exoglucanases.

The two putative xyloglucanases, GH44A and GH74A, were both active on xyloglucan, but GH74A was more active than GH44A (Fig. 2E). GH44A was also very active on all cellulosic substrates tested and exhibited the highest activity on CMC of all of the *C. saccharoperbutylacetonicum* enzymes. Nevertheless, GH44A was the only enzyme that was highly active on beech wood xylan (Fig. 2F) and was therefore classified as a xylanase, while GH74A was classified as a xyloglucanase. Intriguingly, GH44A exhibits a very broad activity pattern with the ability to cleave all of the tested hemicellulosic and cellulosic substrates tested, with the exception of mannan (Fig. 2). The observed xylanase activity in the *C. saccharoperbutylacetonicum* GH44 enzyme reinforces previous reports (22, 23), thus extending the CAZy annotation (15) and previously reported work (24) that indicate endoglucanase and xyloglucanase activities for this family of enzymes.

Surprisingly, of the two predicted mannanases (according to their sequence and GH family affiliation), GH5B and GH26A, only GH5B was able to degrade locust bean gum mannan and showed extremely high activity on this substrate (Fig. 2D). The other putative mannanase, GH26A, was not active on the mannan substrate; however, it exhibited substantial xyloglucanase activity (Fig. 2E). We thus explored its amino acid sequence more thoroughly and constructed a phylogenetic tree (Fig. S3) to examine its genetic relationship to various characterized GH26s. The vast majority of family GH26 enzymes are indeed annotated as mannanases and the rest are annotated as xylanases. Interestingly, we discovered that GH26A has the closest sequence similarity to Lic26A from *C. thermocellum*, the single reported lichenase from family GH26 (25). Therefore, we tested the ability of all our enzymes to cleave lichenan (Fig. 2G), and indeed, GH26A presented the highest activity among all the enzymes on this substrate and could thus be classified as a lichenase but it exhibited high xyloglucanase activity as well. Endoglucanases GH5A, GH9B, and GH44A also presented substantial activity on lichenan.

All of the cellulosomal enzymes from *C. saccharoperbutylacetonicum* were thus characterized for their activity(ies) on specific substrate(s), and four cellulases, a xylanase, a mannanase, a xyloglucanase, and a lichenase, were revealed, with the GH44 xylanase exhibiting broad specificity for several other polysaccharide substrates. In Fig. 3, the schematic modular architecture of these enzymes is shown, along with their enzymatic activities, their dockerin groupings (see paragraph below) and their proposed nomenclature.

Cohesin-dockerin recognition. *C. saccharoperbutylacetonicum* contains two cohesins on its scaffoldin and eight dockerin-containing enzymes that can theoretically

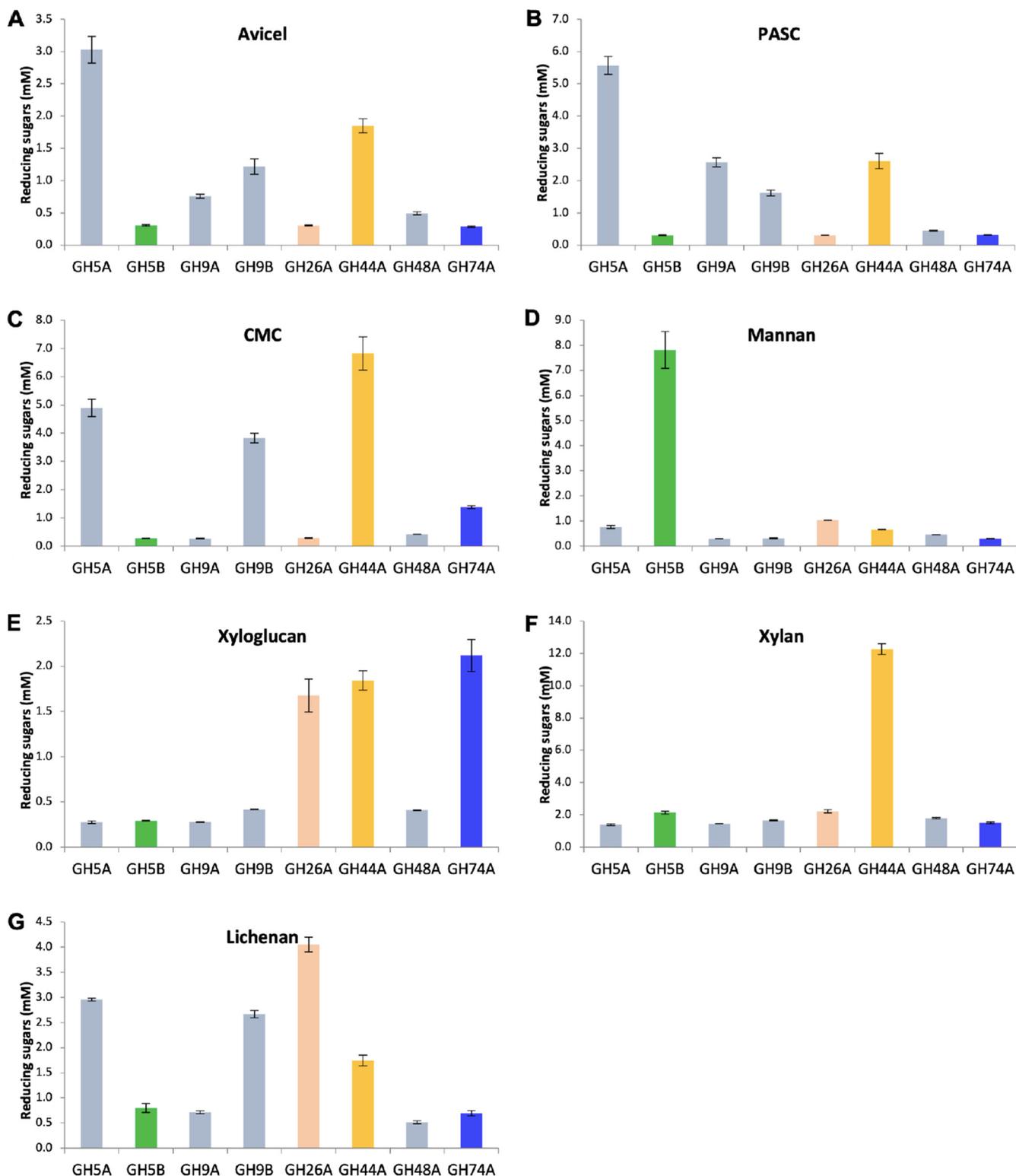


FIG 2 Enzymatic profiling of *C. saccharoperbutylacetonicum* cellulosomal glycoside hydrolases. The comparative enzymatic activities were carried out at a concentration of 0.5 μ M in acetate buffer (pH 5) at 40°C. (A) Cellulase activity on 5% Avicel for 24 h. (B) Cellulase activity on PASC for 24 h. (C) Cellulase activity on 1% CMC for 2 h. (D) Mannanase activity on 1% locust bean gum for 2 h. (E) Xyloglucanase activity on 0.5% xyloglucan for 1 h. (F) Xylanase activity on 1% beech wood xylan for 2 h. (G) Lichenase activity on 0.25% lichenan for 1 h. Reactions were performed at least twice in triplicate; standard deviations are indicated by the error bars.

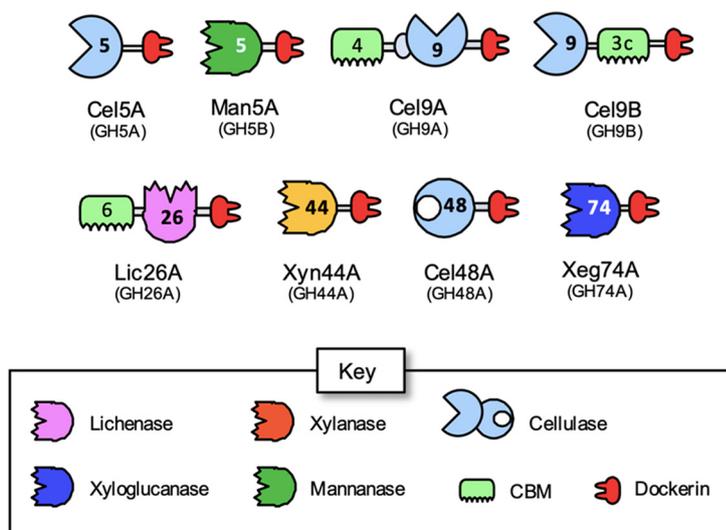


FIG 3 Schematic representation and proposed nomenclature of the experimentally defined dockerin-containing enzymes of *C. saccharoperbutylacetonicum*. Enzymatic activities are color coded according to the key. GH and CBM families are indicated by the numbers.

self-assemble into cellulosomal architectures. In order to draw a complete map of the potential cohesin-dockerin interactions, all of the dockerin-containing enzymes were examined for their interactions with each of the two *C. saccharoperbutylacetonicum* cohesins (with or without their original adjoining X2 module—i.e., X2-Coh1, Coh1, X2-Coh2, and Coh2) using an ELISA-based affinity assay system (26). The two individual cohesin-containing genes were each fused to a CBM cassette, both for antibody recognition and for increased solubility of the cohesin modules when expressed in *E. coli* (27). The presence of the CBM could also be used for an additional purification step.

At first, we examined the effect of the presence of the adjacent X2 module on the binding capacities of the cohesins with their dockerin-containing enzymes. The results showed that the interactions of the dockerins were higher when the X2 module was present adjacent to both cohesins, and the effect was higher for X2-Coh1 (Fig. S4). In view of these findings, in subsequent experiments, we used cohesins fused to X2 modules for analysis. Each immobilized dockerin was tested for its binding capacity with increasing concentrations of either X2-Coh1 or X2-Coh2 (Fig. 4).

The results indicated different patterns of interaction for the two X2-cohesin modular dyads. We thus calculated the logarithmically transformed 50% effective concentration (pEC_{50}) values for each Coh-Doc pair (Fig. 5) from the appropriate concentration-response curve presented in Fig. 4. Our data indicate that all dockerin-containing enzymes interacted strongly with the X2-Coh2 module, with Xyn44A and Xeg74A presenting the highest affinities (pEC_{50} values of 4.0) and Cel9B exhibiting the lowest affinity ($pEC_{50} = 2.4$). Nevertheless, it appeared that X2-Coh1 was more restricted in its interaction pattern. The dockerins from Man5A, Cel48A, and Xyn44A interacted strongly with X2-Coh1 with pEC_{50} values of 2.5, 3.0, and 4.0, respectively. The rest of the tested dockerins interacted weakly with X2-Coh1, where Cel9B exhibited the lowest binding affinity ($pEC_{50} = 0.7$). For a negative control, a dockerin-containing enzyme from *Clostridium thermocellum* (Cel5D-Ct) was used.

CBM affinity. Another interesting feature of the cellulosome is the presence of the two CBM3s on the *C. saccharoperbutylacetonicum* scaffoldin, which mediate the binding to polysaccharide substrate(s). These particular CBMs could shed light on the preferred nutritional substrate of the bacterium. *C. saccharoperbutylacetonicum* scaffoldin has two copies of the CBM3 at the N terminus of its scaffoldin gene (5). In order to experimentally evaluate the binding affinity of *C. saccharoperbutylacetonicum* scaffoldin, we used an affinity electrophoresis assay (28). We used various soluble polysaccharides (PASC,

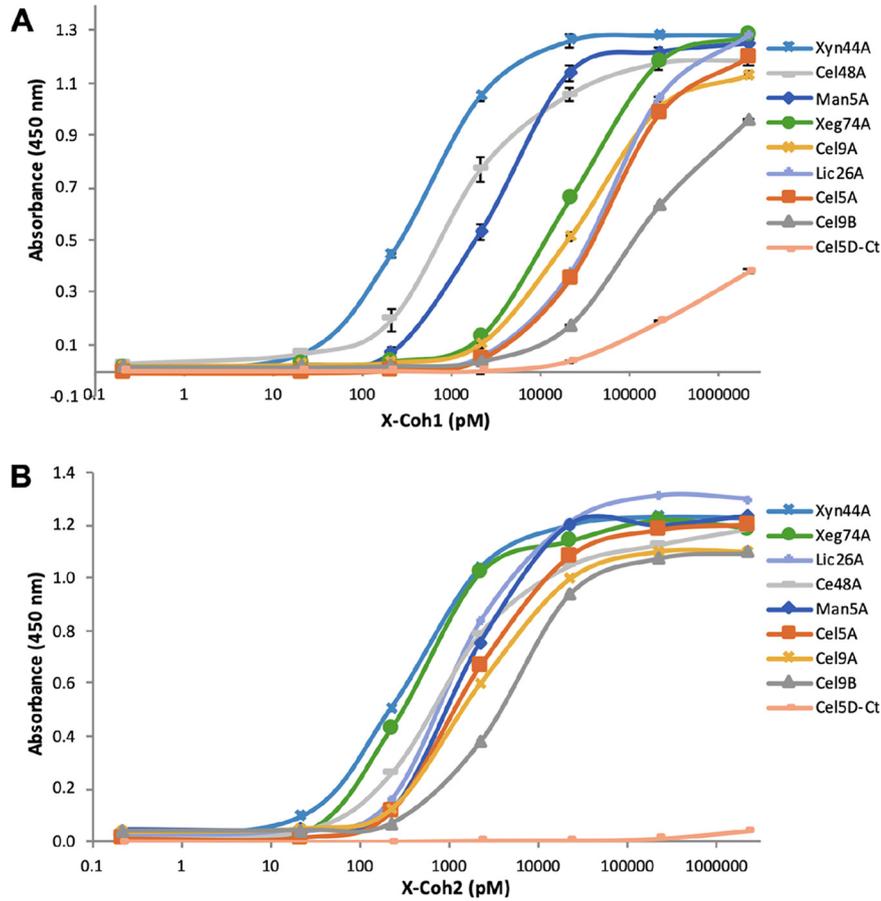


FIG 4 Affinity-based ELISA of cohesins of the *C. saccharoperbutylacetonicum* scaffoldin on immobilized dockerin. (A) X2-Coh1 interaction pattern. (B) X2-Coh2 interaction pattern. Cel5D-Ct (a dockerin-containing enzyme from *C. thermocellum*) was used as a negative control.

xyloglucan, xylan, starch, mannan, and lichenan) in nondenaturing polyacrylamide gels containing one of the latter polysaccharides in soluble form to test the binding abilities of a CBM. In addition to the scaffoldin CBMs, three out of the eight cellulosomal enzymes of *C. saccharoperbutylacetonicum* contain a CBM: Cel9A contains a CBM4 at its

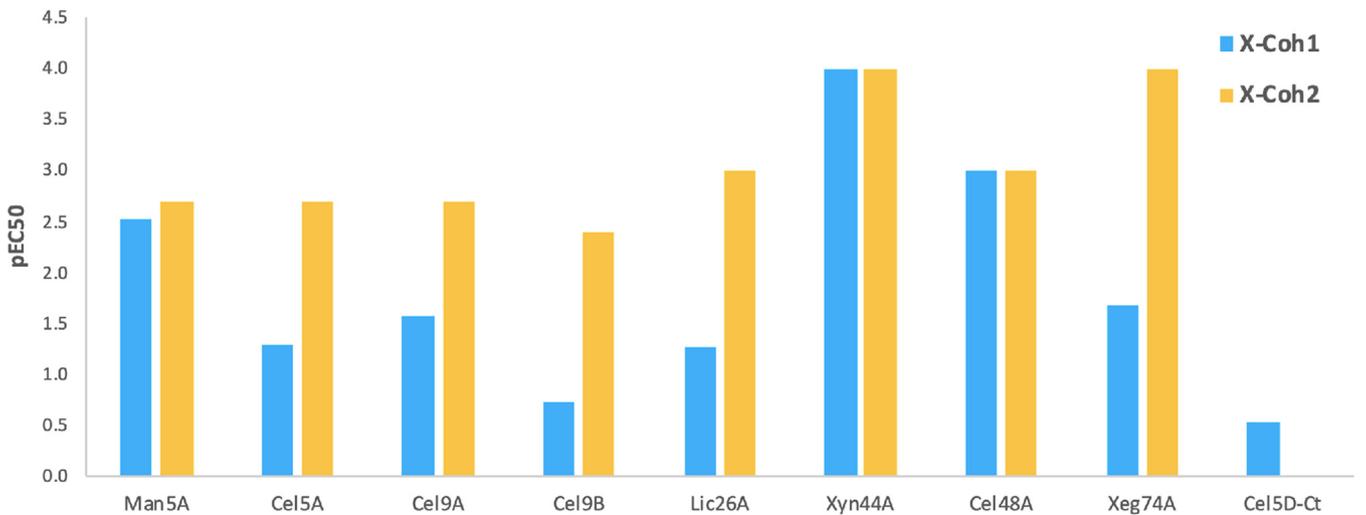


FIG 5 pEC₅₀ values calculated for each Coh-Doc pair from their concentration-response curves presented in Fig. 4.

TABLE 2 Relative mobility values of *C. saccharoperbutylacetonicum* dockerin-containing enzymes and scaffoldin (normalized to BSA) with various soluble polysaccharides

		Xylan 0.2%	Starch 0.2%	PASC 0.1%	Mannan 0.1%	Xyloglucan	
						0.1%	0.2%
BSA (control)		1.00	1.00	1.00	1.00	1.00	1.00
Scaffoldin		1.00	0.88	0.44	1.00	0.56	0.33
Man5A		0.98	0.94	1.00	0.24	0.81	0.55
Lic26A		0.97	0.98	0.99	1.00	0.90	0.81
Xeg74A		0.86	0.88	1.01	0.86	0.86	0.53
Xyn44A		1.00	0.97	1.00	0.78	0.51	0.41
Cel48A		1.00	0.90	0.94	1.00	0.53	0.48
Cel9A		1.00	0.94	0.74	1.00	0.44	0.28
Cel9B		1.00	0.89	1.00	1.00	0.88	0.84
Cel5A		1.01	1.02	1.03	1.00	1.00	0.93

N terminus, Cel9B contains a CBM3c between the catalytic site and the dockerin, and Lic26A contains a CBM6 (Table 1 and Fig. 3).

In this binding assay, the CBM binding ability to the tested polysaccharide is evaluated by the delayed migration of the protein through the gel containing the polysaccharide compared to a comparable gel that lacks the polysaccharide. A protein lacking a CBM could also be delayed due to the inherent affinity of its catalytic site to the polysaccharide (29, 30). Relative mobilities (migration distance of the protein on the gel compared to its migration on a gel with the tested polysaccharide [28]) of each protein was determined with a control gel that does not contain any polysaccharide, and bovine serum albumin (BSA) was used as a negative nonbinding control and served to normalize the relative mobilities. Gel results are presented in Fig. S5, and relative mobility values are presented in Table 2. The results show that the mobility of the *C. saccharoperbutylacetonicum* scaffoldin is hampered when run on gels with PASC and xyloglucan, meaning its CBMs demonstrated high cellulose- and xyloglucan-binding affinity. Very low relative mobility values of the scaffoldin were calculated when run on gels with 0.1% PASC (Fig. S5B) and 0.1% xyloglucan (Fig. S5D) with 0.44 and 0.56 values, respectively. Upon increasing the percentage of xyloglucan in the gel, the relative mobility of the scaffoldin further decreased to 0.33 with 0.2% xyloglucan (Fig. S5E). Although these two CBMs are classified as CBM3s (according to sequence), and would bind primarily to crystalline cellulose according to the literature (15), it seems that one or both of these CBMs are capable of binding cellulosic substrates as well as xyloglucan. Binding of CBM3s to xyloglucan was indeed demonstrated recently (31). The data showed that the CBMs bound to both crystalline cellulose and xyloglucan, thus demonstrating a broader specificity pattern than previously recognized.

Among the enzymes, Cel5A, Cel9B (with CBM3c), and GH26A (which contains an N-terminal CBM6) did not display any significant binding affinity on any of the tested polysaccharides. Cel9A (which contains an N-terminal CBM4) showed low mobility on 0.1% PASC and even lower mobility on xyloglucan, with relative mobility values of 0.74 and 0.28 (on 0.2% xyloglucan), respectively. Some of the enzymes that do not contain a CBM also exhibited some retention on the various gels. For example, Xyn44A and Xeg74A both demonstrated low mobility on xyloglucan, with relative mobility values of 0.41 and 0.53, respectively (on 0.2% xyloglucan), and both enzymes were found to be highly active on this substrate (Fig. 2E). Man5A was also delayed on the xyloglucan gel, with a relative mobility of 0.55 (on 0.2% xyloglucan). The relative mobility of Man5A on its main substrate, mannan, was very low—0.24. Cel48A, which was classified earlier as an exoglucanase, exhibited low mobility on PASC and xyloglucan with relative mobility

values of 0.35, and 0.48 (on 0.2% xyloglucan), respectively. As expected, BSA (the negative control) did not bind to any of the polysaccharides.

Affinity pulldown assay. To complement the affinity electrophoresis assays that shed light on the affinity of the CBMs and related modules to the soluble polysaccharides, we used an affinity pulldown assay to provide insights on binding to insoluble polysaccharides. Using this method, the tested CBM is incubated with an insoluble polysaccharide, and after a centrifugation step, the concentration of bound (pellet) and unbound protein (supernatant fluids) is measured. Here, we tested the ability of the *C. saccharoperbutylacetonicum* scaffoldin to bind the insoluble cellulosic substrate Avicel. The CBM3 of *C. thermocellum* was used as a positive control. Our results clearly demonstrated that the *C. saccharoperbutylacetonicum* scaffoldin has the ability to bind Avicel (Fig. S6), with 46% bound protein versus 50% for CBM3. These results are in agreement with the affinity electrophoresis results (Table 2).

Detection of *C. saccharoperbutylacetonicum* scaffoldin *in vivo*. Next, we wanted to explore the ability of *C. saccharoperbutylacetonicum* to secrete a cellulosome *in vivo*. The results we described demonstrated that the *C. saccharoperbutylacetonicum* cellulosome targets mainly cellulose but also other hemicellulose substrates, specifically mannan, xylan, xyloglucan, and lichenin (Fig. 2 and 3 and Table 2). It has been reported that *C. saccharoperbutylacetonicum* and other butanologenic bacteria do not ferment purified cellulose (32, 33). Therefore, in order to explore *in vivo* cellulosome secretion by *C. saccharoperbutylacetonicum*, we chose to grow the bacterium on two different and contrasting carbohydrate sources. These two sources included a simple soluble saccharide substrate, i.e., the disaccharide cellobiose (CB) which is the major breakdown product of cellulose, and a natural complex insoluble cellulosic substrate, i.e., brewer's spent grain (BSG) (34, 35).

In order to detect the secreted cellulosome, we applied a modification of the affinity-based ELISA approach (26) used above. Here, we used the concentrated spent cell culture media and immobilized their contents to an ELISA plate. In order to confirm the presence of the native scaffoldin in the immobilized material, we interacted the samples with a recombinant dockerin-bearing probe (XynDoc44) that exhibited the highest affinity toward both cohesins in the cohesin-dockerin recognition section (Fig. 4 and 5). The extent of interacting dockerin was determined immunochemically using an antixylanase primary antibody and horseradish peroxidase (HRP)-labeled secondary antibody as reported previously (26). Results presented in Fig. 6 show the relative amount of scaffoldin (cellulosome) in the samples, compared to a positive control (the recombinant scaffoldin ScaA). The results display a larger amount of secreted scaffoldin when the bacterium was grown on the natural cellulosic substrate compared to growth of the bacterium on the simple soluble disaccharide, thus providing experimental evidence that *C. saccharoperbutylacetonicum* actually secretes cellulosomal components *in vivo*.

In order to validate the results obtained from the affinity-based ELISA approach, we also employed an affinity pulldown assay used above, where we exploited the ability of the recombinant CBM-bearing scaffoldin to bind microcrystalline cellulose (Avicel). Using this approach, we employed a modified version of the pulldown assay, where we incubated the concentrated spent cell culture media with Avicel to "fish out" the native scaffoldin from the tested samples. The results (Fig. S7) confirm the presence of a larger amount of scaffoldin secreted when the bacterium was grown on BSG versus CB. These results correlate well with those obtained by the affinity-based ELISA (Fig. 6).

DISCUSSION

Current lignocellulosic biofuel research mainly focuses on bioethanol, aiming at replacing starch- and other simple-sugar-based raw materials, especially from corn, which has caused the "Food versus Fuel" controversy (36). However, ethanol has many disadvantages as a biofuel. It has a low energy density (less than two-thirds that of gasoline) (37), it is highly hygroscopic and corrosive, and thus cannot be transported and stored in existing pipelines and storage facilities. Moreover, vehicles cannot be

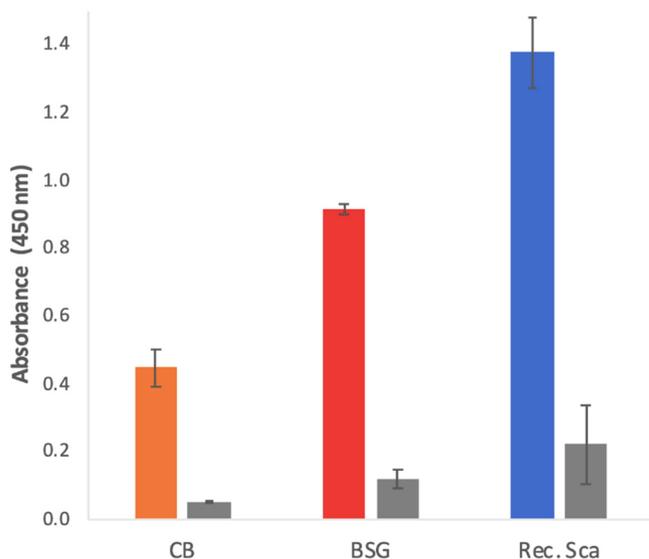


FIG 6 *In vivo* affinity-based ELISA of *C. saccharoperbutylacetonicum* secreted scaffoldin (cellulosome) when grown on media containing cellobiose (CB) or brewer's spent grain (BSG). The recombinant scaffoldin (blue column) was used as a positive control. Absorbance at 450 nm represents the amount of XynDoc44, which corresponds to the relative amount of scaffoldin in the samples. Negative-control experiments (gray columns) were performed without the XynDoc44 interaction step.

powered by pure ethanol without redesigning the engines in order to use the low-energy fuel. Therefore, the prospect for large-scale use of ethanol to replace petroleum fuels is not promising. On the other hand, biobutanol has superior fuel properties (higher energy density, lower volatility, etc.), and can fit better into existing fuel infrastructures (with low water miscibility and corrosiveness). Biobutanol is safer (low flammability) than ethanol and can directly replace gasoline in car engines without modification. It is therefore considered an advanced biofuel, superior to bioethanol (38). In addition, butanol is also a bulk fundamental material with a variety of uses, including as a precursor in organic synthesis for producing dyes, paints, plastics, resins, and rubber, as a solvent for producing antibiotics and synthetic drugs, and as a food-grade extraction agent in the food and fragrance industries. Therefore, there is a need for biobutanol production from renewable biomass.

The solvent-producing clostridia are known for their ability to grow on a variety of polysaccharides. Although *C. saccharoperbutylacetonicum* does not ferment cellulose, it can utilize other simple and more complex sugars (7, 39), including different cellulosic biowastes as carbon sources (40–42), and use them to produce biobutanol. Another unique feature of these strains is their ability to use starch directly without any additional hydrolysis step. Under these conditions, *C. saccharoperbutylacetonicum* produced 16.9 g/liter butanol from cassava starch (43).

Since the discovery of the cellulosome, many cellulosomal systems have been investigated from “complex” systems such as the cellulosome of the extensively researched *C. thermocellum* (44) and the largest and highly complexed cellulosome of *B. cellulosolvans* (4) to much smaller and “simple” systems (5) such as those of *C. josui* (45) and the solvent-producing *C. acetobutylicum* (10). Here, we conducted a broad analysis of the minimal cellulosome of *C. saccharoperbutylacetonicum*, which currently comprises the smallest known cellulosome thus far.

The *C. saccharoperbutylacetonicum* cellulosome contains eight dockerin-containing enzymes that were characterized in the present study. In addition, the putative *C. saccharoperbutylacetonicum* scaffoldin was also characterized as a complete unit as well as its individual components, i.e., the cohesin modules separately (i.e., Coh1 and Coh2 with and without their adjacent X modules). Initially, according to sequence, we inferred that the eight cellulosomal enzymes would exhibit the predicted activities of

two mannanases, four cellulases, and two xyloglucanases. However, after conducting broad activity profiling using model substrates, we discovered that some of the experimentally derived activities deviated from initial predictions, thus revealing five different types of enzyme activities and not three (with the addition of lichenase and xylanase activities).

C. saccharoperbutylacetonicum possesses a set of eight distinct enzymes with predicted cellulase (endoglucanase, processive endoglucanase, and exoglucanase) and hemicellulase activities. The hemicellulases, GH26, GH44, GH74, and the exoglucanase GH48, each appear once in the genome, and they are all cellulosomal (containing a dockerin module). The GH9 and GH5 enzymes have three and four members, respectively, in the *C. saccharoperbutylacetonicum* genome, where two of each family are cellulosome related. Among the four cellulosomal cellulases, two were characterized as a standard endoglucanase and a processive endoglucanase (Cel5A and Cel9B, respectively), whereas the other two were classified as exoglucanases (Cel9A and Cel48A). Intriguingly, Lic26A, Xyn44A, and Xeg74A were all able to digest xyloglucan, where Xeg74A exhibited the highest activity. It was also surprising to discover a lichenase among this small selection of enzymes. *C. saccharoperbutylacetonicum* Lic26A is the second known published lichenase of the GH26 family. Interestingly, Xyn44A might be an interesting candidate for designer cellulosome (46–48), in view of its high activity on xylan while also being active on a diverse set of cellulolytic and hemicellulolytic substrates. It also interacted strongly with both cohesins. Moreover, Xyn44A was active and stable at 60°C (see Fig. S2 in the supplemental material). One of the most encouraging results of this paper is the measurable activity shown by Cel48A. While it is known that Cel48 of *C. acetobutylicum* is inactive or has minimal activity (10), it seems that this key cellulase is indeed active in *C. saccharoperbutylacetonicum*.

Although *C. saccharoperbutylacetonicum* scaffoldin has only two cohesins, they exhibited remarkably different binding patterns. Coh1 presented varied affinities toward the dockerin-containing enzymes, with the highest affinity toward Xyn44A ($pEC_{50} = 4.0$) and the lowest affinity toward Cel9B ($pEC_{50} = 0.7$). Coh2 was much less selective and exhibited higher affinity toward all dockerins (pEC_{50} range of 2.4 to 4.0). In addition, we can conclude that the Man5A, the important broad-activity Xyn44A, and the key endoglucanase Cel48A all bind very similarly to either of the two cohesin modules, whereas the other dockerins bind preferentially to Coh2 compared to Coh1. This may reflect the likelihood that two different enzymes will pair together on the same scaffoldin. Therefore, while Man5A, Xyn44A, and Cel48A will probably bind to Coh1, the other cellulases (Cel5A, Cel9A, and Cel9B), the mannanase (Man5A) and the lichenase (Lic26A) will tend to bind to Coh2.

As mentioned above, the *C. saccharoperbutylacetonicum* dockerin modules exhibit a unique feature. While the majority of the known type I dockerins interact with their corresponding cohesins via a dual binding mode (49), the *C. saccharoperbutylacetonicum* dockerins apparently display a single mode of binding (50), owing to their unique asymmetric sequences (Fig. 1C). The functional significance of this single binding mode to the *C. saccharoperbutylacetonicum* cellulosome is not currently known.

Another unusual feature of *C. saccharoperbutylacetonicum* is the presence of two copies of CBMs on its scaffoldin. Among the recently evaluated mesophilic cellulolytic clostridia (5), *C. saccharoperbutylacetonicum* and *C. bornimense* are the only known bacteria that exhibit two copies (rather than one) of a CBM3 at the N terminus of their scaffoldin gene. Interestingly, the cohesin sequences of the two latter bacterial strains are also very similar (5). Multiple CBMs may reflect a necessity to bind to longer portions of the polysaccharide chain. Affinity electrophoresis revealed that the *C. saccharoperbutylacetonicum* scaffoldin can bind both Avicel (cellulose) and xyloglucan (hemicellulose), which further supports a recently published report (31) that demonstrated that the *C. thermocellum* CBM3a does not exclusively bind to crystalline cellulose as previously thought but that it also binds to xyloglucan. In the latter report, the authors showed that the same aromatic residues of the polysaccharide recognition site are responsible for binding both crystalline cellulose and xyloglucan.

In addition to the 8 cellulosomal dockerin-bearing GHs discussed in this paper, *C. saccharoperbutylacetonicum* encodes an additional 138 free GHs (5). Among them are two conspicuous families, GH1 (putative β -glycosidases) with 20 members and GH13 (putative amylases or pullulanases) with 14 members. Since synergy between free and cellulosomal paradigms within the same microorganism has been demonstrated (51, 52), we assume that the minimalistic *C. saccharoperbutylacetonicum* cellulosome plays a supportive role to the much larger collection of free enzymes in this bacterium in deconstructing fiber biomass.

In this work, we also provide initial evidence that *C. saccharoperbutylacetonicum* secretes a cellulosome *in vivo*. Since the bacterium does not ferment cellulose, it was grown on two different types of substrates in the growth media, which were chosen according to the cellulosome function *in vitro*, as described above, in order to examine its influence on cellulosome secretion. The results suggest that the secretion is substrate dependent, whereby larger amounts of cellulosome were obtained when the bacterium was grown on the natural cellulosic substrate versus cellobiose. This phenomenon correlates with our previously published work, which demonstrated that cellulosome production in other bacterial species was regulated by the type of substrate in the growth medium (3, 44, 53).

Our analyses described in the present communication contribute to a better understanding of simple cellulosome systems with regard to the “key players” in this reduced system and the binding pattern of its cohesin-dockerin interaction. The list of eight cellulosomal cellulase and hemicellulase enzymes in *C. saccharoperbutylacetonicum* provides a basis for the minimalistic cellulosome system that would act on cellulosic biomass. A divalent scaffoldin would by definition provide a minimal requirement for enzyme proximity and targeting effects (54), which can be further examined by assembly of even smaller artificial systems, such as designer cellulosomes (46–48). In future work, it will be important to determine the significance of such a minimalistic cellulosome to this bacterium, to its growth on complex polysaccharide biomass, and to its production of butanol as a biofuel. In other words, why is it there, what is it doing, and why is it so small?

MATERIALS AND METHODS

Cloning of dockerins and CBM-fused cohesins. All related cellulosomal components, dockerin-containing enzymes, cohesins, and scaffoldin genes, were cloned from *C. saccharoperbutylacetonicum* N1-4 (HMT) genomic DNA (DSMZ 14923) using appropriate primers (Sigma-Aldrich, Israel) (see Table S1 in the supplemental material). They were all cloned without their signal peptide, which was predicted by the SignalP 3.0 server (55). GH5A, GH26A, GH44A, GH48A, GH74A, and the putative scaffoldin gene were amplified by PCR using Phusion DNA polymerase (Thermo Scientific, USA). The scaffoldin modules Coh1, X-Coh1, Coh2, X-Coh2, as well as the XynDoc44 were amplified using KAPA HiFi HotStart ReadyMix PCR kit (KAPA Biosystems, USA). PCR products were then cleaved using FastDigest enzymes (Thermo Scientific, USA): dockerin genes were cut with NcoI and XhoI, while cohesin genes were cleaved with BamHI and XhoI restriction enzymes, respectively. Next, these PCR products were purified using the HiYield gel/PCR DNA fragments extraction kit (Real Genomics, RBC Bioscience, Taiwan). The purified restricted PCR products were then ligated to a suitable plasmid, i.e., dockerin-containing genes were ligated into pET28a (Novagen, Madison, WI, USA), and cohesin modules were ligated into CBM-fused pET28a cassette (26), using T4 DNA ligase (New England BioLabs, USA). The XynDoc gene cassette consists of xylanase T6 from *Geobacillus stearothermophilus* with an N-terminal His tag cloned into plasmid pET9d (Novagen Inc., Madison, WI), into which a dockerin-encoding sequence was introduced between the KpnI and BamHI restriction sites of the plasmid (26). GH5B and GH9A had the restriction sites mentioned above in their desired sequence; therefore, they were cloned using a restriction free method (56). For GH5B, first, its gene was amplified using KAPA HiFi kit in order to form a mega primer. Second, it was ligated to a closed pET28a using T4 DNA ligase to create a megaplasmid. The remaining unreacted methylated plasmids were digested by DpnI (New England BioLabs, USA). The GH9A gene was cloned using the NEBuilder HiFi DNA assembly cloning Kit (New England BioLabs, USA). The constructs were designed to contain a His tag for subsequent purification. The sequences of all plasmids were verified by DNA sequencing using the appropriate primers (Table S1) or by using pETrev and T7 primers. For longer genes, internal primers were also designed (Table S1) in order to examine the full desired cloned sequences. All plasmids were transformed to competent *Escherichia coli* DH5 α . Polyclonal antibody against xylanase-T6 was prepared as described earlier (27). Secondary antibody-enzyme conjugate (horseradish peroxidase [HRP]-labeled goat anti-rabbit IgG) was a product of Jackson ImmunoResearch Laboratories Inc. (West Grove, PA).

Protein expression and purification. For protein expression, the plasmids containing GH5B, GH9A, GH9B, GH74A, XynDoc44, the cohesin modules, and the putative scaffoldin, were all transformed into *E. coli* BL21(DE3) cells. The rest of the plasmids were transformed into BL21 Star (DE) cells. All the cells, besides GH48A, were grown in 1 liter of Luria-Bertani (LB) medium (GH48A was grown in 2 liters of LB), at 37°C to an A_{600} of ≈ 0.8 to 1. For dockerin-containing enzymes, 2 mM CaCl_2 was added to ensure proper folding of the dockerins. Subsequently, 0.1 mM (final concentration) isopropyl- β -D-thiogalactopyranoside (IPTG) was added to induce protein expression. After overnight growth at 16°C, the cells were harvested by centrifugation at 6,000 rpm for 15 min at 10°C.

The cells were resuspended in 30 ml Tris-buffered saline (TBS) (137 mM NaCl, 2.7 mM KCl, 25 mM Tris-HCl [pH 7.4]) with 5 mM imidazole and protease inhibitor cocktail and incubated on ice. The cells were then sonicated at 45% amplitude, with 50-s pulses for 3.5 min. The cells were then centrifuged (16,000 rpm, 320 min, 10°C), and the supernatant was further purified. The CBM-cohesin-containing protein was added to 2 g of preswollen macroporous beaded cellulose (Iontosorb, Czech Republic) and incubated while rotating for 1 h at 4°C. The mixture was then loaded onto a gravity column, washed with 100 ml of TBS containing 1 M NaCl, and then washed with 100 ml of TBS. The desired proteins were eluted with 1% triethanolamine, and three fractions of 5 ml each were collected. The His-tagged proteins were loaded onto an immobilized nickel nitrilotriacetic acid (Ni-NTA) column (Qiagen, Netherlands), as reported earlier (57). The purity of the proteins was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and then dialyzed against TBS with 5 mM CaCl_2 overnight at 4°C (purified enzymes are presented in Fig. S1 in the supplemental material). Protein concentration was estimated by absorbance (280 nm) based on the known amino acid composition of the protein using the PROTPARAM tool (58). Proteins were stored in 50% glycerol (vol/vol) at -20°C .

Enzymatic activity assay. (i) Enzyme substrates. Avicel, locust bean gum (for mannanase activity assay), and xylan from beech wood were purchased from Sigma-Aldrich (Darmstadt, Germany). Xyloglucan and lichenan were from Megazyme (Bray, Ireland). CMC was from Fisher Chemical (Thermo Fisher Scientific, Waltham, MA, USA). Phosphoric acid swollen cellulose (PASC) was prepared from Avicel in a manner similar to the method of Wood (59). From these substrates, fresh stock solutions were prepared with the following substrate concentration: 10% Avicel, 2% CMC, 0.6% PASC, 2% mannan, 0.5% lichenan, 2% xylan, and 1% xyloglucan.

(ii) Activity assay. All activity assays were performed under optimized reaction conditions (as described above): 0.5 M acetate buffer (pH 5) at 40°C. The final concentration of enzyme was 0.5 μM per reaction mixture in a total volume of 200 μl , and the final buffer concentration was 50 mM. All examined proteins were tested for their activity on all seven substrates, in different final concentrations and different incubation times: CMC, xylan, and mannan at 1% final concentration for 2 h, Avicel at 5% for 24 h, PASC at 0.3% for 24 h, lichenan at 0.25% for 1 h, and xyloglucan at 0.5% for 1 h. All assays were performed at least twice in triplicate with an agitation rate of 300 rpm. The quantification of the reduced soluble sugars released from the polysaccharide substrates was done by using the 3,5-dinitrosalicylic acid (DNS) method (60, 61). At the end of the assay, the reaction tubes were transferred to an ice-water bath and then centrifuged for 4 min at 13,000 rpm at room temperature (RT). Then, 100 μl of each sample was added to 150 μl of the DNS solution. This mixture was then incubated for 10 min at 100°C, 200 μl of this mixture was transferred to each well on a 96-well plate, and the absorbance at 540 nm was measured. The sugar concentration was determined using a glucose standard curve.

Affinity-based ELISA of recombinant proteins. In order to determine the cohesin-dockerin specificity of interaction, the standard affinity-based ELISA procedure was used as described previously (26). Briefly, 5 $\mu\text{g/ml}$ of each dockerin-containing enzyme was immobilized on a MaxiSorp 96-well ELISA plate (Greiner Bio-One, Belgium). Subsequently, CBM-Coh modules were applied on the plate in a concentration gradient of 0.01 ng/ml to 100 $\mu\text{g/ml}$. Rabbit anti-CBD diluted 1:5,000 was used as the primary antibody preparation. The negative control was Cel5D-Ct (a dockerin-containing enzyme from *C. thermocellum*) (62, 63). The pEC_{50} values were calculated using the “Quest Graph EC50 calculator” (<https://www.aatbio.com/tools/ec50-calculator/>) of AAT Bioquest, Inc. (Sunnyvale, CA, USA).

Affinity electrophoresis. The affinity electrophoresis assay was performed by the method of Morais et al. (28). In this assay, we tested six different gels that contained various polysaccharides in different percentages: (i) 0.1% (wt/vol) PASC, (ii) 0.1% (wt/vol) xyloglucan, (iii) 0.2% (wt/vol) xyloglucan, (iv) 0.2% (wt/vol) xylan, (v) 0.2% (wt/vol) starch, and (vi) 0.1% (wt/vol) mannan. Each gel was run for 2.5 h at 100 V, versus a control gel (with no polysaccharide). Relative mobilities were calculated (and normalized) versus BSA mobility on the control gel.

Affinity pulldown. The affinity pulldown assay is based on measuring the soluble protein concentrations in a sample containing an insoluble polysaccharide versus a reference sample without the polysaccharide. Each sample was prepared in a 1.5-ml tube and contained 25 mg Avicel, a protein with a final concentration of 0.1 mg/ml, 100 μl TBS $\times 10$, and double distilled water up to 1 ml. A reference sample was prepared the same but without Avicel. All samples were shaken on rotary rotator at 4°C for 1 h. The samples were then centrifuged at 13,000 rpm for 5 min, and the supernatant fluids (“unbound” fraction) were separated from the pellet (“bound” fraction). Each pellet was washed three times using TBS $\times 1$ buffer (pH 7.4), and resuspended with 225 μl TBS $\times 1$. The protein concentration was measured from the supernatant, pellet, and reference fractions, using the Pierce BCA protein assay kit (Thermo Scientific, USA). The positive control was CBM3 of *C. thermocellum* (27). All the measurements were performed at least three times. Reduction in protein concentration compared to the reference sample indicates binding to Avicel.

Anaerobic fermentation of *Clostridium saccharoperbutylacetonicum*. *C. saccharoperbutylacetonicum* strain N1-4 (HMT) (DSM 14923) was obtained from the German Collection of Microorganisms (DSMZ,

Leibniz, Germany). The microorganism was grown on a tryptone-yeast extract-acetate (TYA) medium under anaerobic conditions by the method of Al-Shorgani et al. (64) with either 4% (wt/vol) cellobiose (CB) (Sigma Chemical Co., St. Louis, MO) or 5% (wt/vol) brewer's spent grain (BSG), kindly provided by Instituto Tecnológico Agrario de Castilla y León (ITACyL) Valladolid, Spain (34), as the sole carbon and energy source. Growth on each of the two carbon sources was performed in two biological repeats. *C. saccharoperbutylacetonicum* cell culture supernatant fluids were filtered through 0.2- μ m sterile plastic filters (Thermo Fisher Scientific, Waltham, MA, USA) and concentrated 15-fold using a Vivaspin concentrator (30-kDa cutoff; Sartorius Stedim Biotech GmbH, Göttingen, Germany).

Affinity-based ELISA for detection of *in vivo* cellulosome. Affinity-based ELISAs were performed as described above with minor modifications. For affinity-based ELISA of material absorbed to plates, 96-well ELISA plates were coated with concentrated spent cell culture media from CB- or BSG-grown *C. saccharoperbutylacetonicum* cells or with recombinant scaffoldin (used as the positive control). Next, 10 μ g/ml XynDoc44 was used to examine the presence of the scaffoldin (via cohesin-dockerin interaction). Rabbit anti-XynT6 diluted 1:5,000 was then used as the primary antibody, followed by HRP-labeled anti-rabbit diluted 1:10,000 as the secondary antibody (26).

For the affinity pulldown assay, the concentrated spent cell culture media from CB- and BSG-grown *C. saccharoperbutylacetonicum* were incubated with 25 mg of Avicel. Here, the washed incubated "Avicel pellet" was interacted with XynDoc44 followed by the above-described immunochemical procedure. Negative-control experiments were performed without the XynDoc44 interaction step. All experiments were performed at least twice.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

FIG S1, JPG file, 0.2 MB.

FIG S2, JPG file, 0.3 MB.

FIG S3, JPG file, 0.2 MB.

FIG S4, JPG file, 0.3 MB.

FIG S5, JPG file, 0.4 MB.

FIG S6, JPG file, 0.1 MB.

FIG S7, TIF file, 0.4 MB.

TABLE S1, PDF file, 0.1 MB.

ACKNOWLEDGMENTS

This research was supported by the United States–Israel Binational Science Foundation (BSF grant 2013284) (Jerusalem, Israel), the Israel Science Foundation (ISF grant 1349/13), the European Union NMP.2013.1.1-2: CellulosomePlus project 604530, and a European Union Horizon 2020 contract for sustainable production of next-generation biofuels from waste streams: "Waste2Fuels," proposal 654623-2. The research was also supported by a research grant from the Yotam Project via the Sustainability and Energy Research Initiative (SAERI) at the Weizmann Institute of Science.

REFERENCES

- Bayer EA, Belaich J-P, Shoham Y, Lamed R. 2004. The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* 58:521–554. <https://doi.org/10.1146/annurev.micro.57.030502.091022>.
- Artzi L, Bayer EA, Morais S. 2017. Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat Rev Microbiol* 15:83–95. <https://doi.org/10.1038/nrmicro.2016.164>.
- Artzi L, Morag E, Barak Y, Lamed R, Bayer EA. 2015. *Clostridium clariflavum*: key cellulosome players are revealed by proteomic analysis. *mBio* 6:e00411-15. <https://doi.org/10.1128/mBio.00411-15>.
- Zhivin O, Dassa B, Morais S, Utturkar SM, Brown SD, Henrissat B, Lamed R, Bayer EA. 2017. Unique organization and unprecedented diversity of the *Bacteroides* (*Pseudobacteroides*) *cellulosolvans* cellulosome system. *Biotechnol Biofuels* 10:211. <https://doi.org/10.1186/s13068-017-0898-6>.
- Dassa B, Borovok I, Lombard V, Henrissat B, Lamed R, Bayer EA, Morais S. 2017. Pan-cellulosomics of mesophilic clostridia: variations on a theme. *Microorganisms* 5:74. <https://doi.org/10.3390/microorganisms5040074>.
- Pagès S, Bélaïch A, Fierobe H-P, Tardif C, Gaudin C, Bélaïch J-P. 1999. Sequence analysis of scaffolding protein CipC and ORFXp, a new cohesin-containing protein in *Clostridium cellulolyticum*: comparison of various cohesin domains and subcellular localization of ORFXp. *J Bacteriol* 181:1801–1810. <https://doi.org/10.1128/JB.181.6.1801-1810.1999>.
- Motoyoshi H. July 1960. Process for producing butanol by fermentation. US patent US2945786A.
- López-Contreras AM, Gabor K, Martens AA, Renckens BA, Claassen PA, Van Der Oost J, De Vos WM. 2004. Substrate-induced production and secretion of cellulases by *Clostridium acetobutylicum*. *Appl Environ Microbiol* 70:5238–5243. <https://doi.org/10.1128/AEM.70.9.5238-5243.2004>.
- López-Contreras AM, Martens AA, Szijarto N, Mooibroek H, Claassen PA, van der Oost J, de Vos WM. 2003. Production by *Clostridium acetobutylicum* ATCC 824 of CelG, a cellulosomal glycoside hydrolase belonging to family 9. *Appl Environ Microbiol* 69:869–877. <https://doi.org/10.1128/aem.69.2.869-877.2003>.
- Sabathé F, Bélaïch A, Soucaille P. 2002. Characterization of the cellulosytic complex (cellulosome) of *Clostridium acetobutylicum*. *FEMS Microbiol Lett* 217:15–22. <https://doi.org/10.1111/j.1574-6968.2002.tb11450.x>.
- Poehlein A, Krabben P, Durre P, Daniel R. 2014. Complete genome sequence of the solvent producer *Clostridium saccharoperbutylacetonicum* strain DSM 14923. *Genome Announc* 2:e01056-14. <https://doi.org/10.1128/genomeA.01056-14>.
- Toth J, Ismaiel AA, Chen J-S. 1999. The ald gene, encoding a coenzyme A-acylating aldehyde dehydrogenase, distinguishes *Clostridium beijerinckii* and two other solvent-producing clostridia from *Clostridium acetobutylicum*. *Appl Environ Microbiol* 65:4973–4980. <https://doi.org/10.1128/AEM.65.11.4973-4980.1999>.

13. Mosbah A, Belaïch A, Bornet O, Belaïch JP, Henrissat B, 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. <https://doi.org/10.1006/jmbi.2000.4192>.
14. Adams JJ, Pal G, Jia Z, Smith SP. 2006. Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc Natl Acad Sci U S A* 103:305–310. <https://doi.org/10.1073/pnas.0507109103>.
15. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The Carbohydrate-Active Enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42:D490–D495. <https://doi.org/10.1093/nar/gkt1178>.
16. Morais S, David YB, Bensoussan L, Duncan SH, Koropatkin NM, Martens EC, Flint HJ, Bayer EA. 2016. Enzymatic profiling of cellulosomal enzymes from the human gut bacterium, *Ruminococcus champanellensis*, reveals a fine-tuned system for cohesin-dockerin recognition. *Environ Microbiol* 18:542–556. <https://doi.org/10.1111/1462-2920.13047>.
17. Cheng C. 1998. Cellulase activity in different buffering media during waste paper hydrolysis by HPLC. *J Chin Chem Soc* 45:679–688. <https://doi.org/10.1002/jccs.199800103>.
18. Vazana Y, Morais S, Barak Y, Lamed R, Bayer EA. 2010. Interplay between *Clostridium thermocellum* family 48 and family 9 cellulases in cellulosomal versus noncellulosomal states. *Appl Environ Microbiol* 76:3236–3243. <https://doi.org/10.1128/AEM.00009-10>.
19. Berger E, Zhang D, Zverlov VV, Schwarz WH. 2007. Two noncellulosomal cellulases of *Clostridium thermocellum*, Cel9I and Cel48Y, hydrolyse crystalline cellulose synergistically. *FEMS Microbiol Lett* 268:194–201. <https://doi.org/10.1111/j.1574-6968.2006.00583.x>.
20. Reverbel-Leroy C, Pages S, Belaïch A, Belaïch J-P, Tardif C. 1997. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: purification and characterization of the recombinant form. *J Bacteriol* 179:46–52. <https://doi.org/10.1128/jb.179.1.46-52.1997>.
21. Morag E, Halevy I, Bayer E, Lamed R. 1991. Isolation and properties of a major cellobiohydrolase from the cellulosome of *Clostridium thermocellum*. *J Bacteriol* 173:4155–4162. <https://doi.org/10.1128/jb.173.13.4155-4162.1991>.
22. Warner CD, Hoy JA, Shilling TC, Linnen MJ, Ginder ND, Ford CF, Honzatko RB, Reilly PJ. 2010. Tertiary structure and characterization of a glycoside hydrolase family 44 endoglucanase from *Clostridium acetobutylicum*. *Appl Environ Microbiol* 76:338–346. <https://doi.org/10.1128/AEM.02026-09>.
23. Warner CD, Go RM, Garcia-Salinas C, Ford C, Reilly PJ. 2011. Kinetic characterization of a glycoside hydrolase family 44 xyloglucanase/endoglucanase from *Ruminococcus flavefaciens* FD-1. *Enzyme Microb Technol* 48:27–32. <https://doi.org/10.1016/j.enzmictec.2010.08.009>.
24. Ariza A, Eklöf JM, Spadiut O, Offen WA, Roberts SM, Besenmatter W, Friis EP, Skjöt M, Wilson KS, Brumer H, Davies G. 2011. Structure and activity of *Paenibacillus polymyxa* xyloglucanase from glycoside hydrolase family 44. *J Biol Chem* 286:33890–33900. <https://doi.org/10.1074/jbc.M111.262345>.
25. Taylor EJ, Goyal A, Guerreiro CIPD, Prates JAM, Money VA, Ferry N, Morland C, Planas A, Macdonald JA, Stick RV, Gilbert HJ, Fontes CMGA, Davies GJ. 2005. How family 26 glycoside hydrolases orchestrate catalysis on different polysaccharides. Structure and activity of a *Clostridium thermocellum* lichenase, CtLic26A. *J Biol Chem* 280:32761–32767. <https://doi.org/10.1074/jbc.M506580200>.
26. Barak Y, Handelsman T, Nakar D, Mechaly A, Lamed R, Shoham Y, Bayer EA. 2005. Matching fusion protein systems for affinity analysis of two interacting families of proteins: the cohesin-dockerin interaction. *J Mol Recognit* 18:491–501. <https://doi.org/10.1002/jmr.749>.
27. Morag E, Lapidot A, Govorko D, Lamed R, Wilchek M, Bayer EA, Shoham Y. 1995. Expression, purification, and characterization of the cellulose-binding domain of the scaffoldin subunit from the cellulosome of *Clostridium thermocellum*. *Appl Environ Microbiol* 61:1980–1986. <https://doi.org/10.1128/AEM.61.5.1980-1986.1995>.
28. Morais S, Lamed R, Bayer EA. 2012. Affinity electrophoresis as a method for determining substrate-binding specificity of carbohydrate-active enzymes for soluble polysaccharides. *Methods Mol Biol* 908:119–127. https://doi.org/10.1007/978-1-61779-956-3_12.
29. Vazana Y, Morais S, Barak Y, Lamed R, Bayer EA. 2012. Designer cellulosomes for enhanced hydrolysis of cellulosic substrates. *Methods Enzymol* 510:429–452. <https://doi.org/10.1016/B978-0-12-415931-0.00023-9>.
30. Stern J, Artzi L, Morais S, Fontes CMGA, Bayer EA. 2017. Carbohydrate depolymerization by intricate cellulosomal systems. *Methods Mol Biol* 1588:93–116. https://doi.org/10.1007/978-1-4939-6899-2_8.
31. Hernandez-Gomez MC, Rydahl MG, Rogowski A, Morland C, Cartmell A, Crouch L, Labourel A, Fontes CMGA, Willats WGT, Gilbert HJ, Knox JP. 2015. Recognition of xyloglucan by the crystalline cellulose-binding site of a family 3a carbohydrate-binding module. *FEBS Lett* 589:2297–2303. <https://doi.org/10.1016/j.febslet.2015.07.009>.
32. Nakayama S, Kiyoshi K, Kadokura T, Nakazato A. 2011. Butanol production from crystalline cellulose by cocultured *Clostridium thermocellum* and *Clostridium saccharoperbutylacetonicum* N1-4. *Appl Environ Microbiol* 77:6470–6475. <https://doi.org/10.1128/AEM.00706-11>.
33. Xin F, Dong W, Zhang W, Ma J, Jiang M. 2019. Biobutanol production from crystalline cellulose through consolidated bioprocessing. *Trends Biotechnol* 37:167–180. <https://doi.org/10.1016/j.tibtech.2018.08.007>.
34. Plaza PE, Gallego-Morales LJ, Peñuela-Vásquez M, Lucas S, García-Cubero MT, Coca M. 2017. Biobutanol production from brewer's spent grain hydrolysates by *Clostridium beijerinckii*. *Bioresour Technol* 244:166–174. <https://doi.org/10.1016/j.biortech.2017.07.139>.
35. Carvalheiro F, Esteves M, Parajó J, Pereira H, Grijo F. 2004. Production of oligosaccharides by autohydrolysis of brewer's spent grain. *Bioresour Technol* 91:93–100. [https://doi.org/10.1016/S0960-8524\(03\)00148-2](https://doi.org/10.1016/S0960-8524(03)00148-2).
36. Zhao XQ, Zi LH, Bai FW, Lin HL, Hao XM, Yue GJ, Ho NW. 2012. Bioethanol from lignocellulosic biomass. *Adv Biochem Eng Biotechnol* 128:25–51. https://doi.org/10.1007/10_2011_129.
37. US Energy Information Administration. 2014. How much ethanol is in gasoline, and how does it affect fuel economy? US Energy Information Administration, Washington, DC.
38. Xue C, Zhao XQ, Liu CG, Chen LJ, Bai FW. 2013. Prospective and development of butanol as an advanced biofuel. *Biotechnol Adv* 31:1575–1584. <https://doi.org/10.1016/j.biotechadv.2013.08.004>.
39. Tracy BP, Jones SW, Fast AG, Indurthi DC, Papoutsakis ET. 2012. Clostridia: the importance of their exceptional substrate and metabolite diversity for biofuel and biorefinery applications. *Curr Opin Biotechnol* 23:364–381. <https://doi.org/10.1016/j.copbio.2011.10.008>.
40. Ellis JT, Hengge NN, Sims RC, Miller CD. 2012. Acetone, butanol, and ethanol production from wastewater algae. *Bioresour Technol* 111:491–495. <https://doi.org/10.1016/j.biortech.2012.02.002>.
41. Li Y, Tang W, Chen Y, Liu J, Chia-Fon FL. 2019. Potential of acetone-butanol-ethanol (ABE) as a biofuel. *Fuel* 242:673–686. <https://doi.org/10.1016/j.fuel.2019.01.063>.
42. Shukor H, Al-Shorgani NKN, Abdeshahian P, Hamid AA, Anuar N, Rahman NA, Kalil MS. 2014. Production of butanol by *Clostridium saccharoperbutylacetonicum* N1-4 from palm kernel cake in acetone-butanol-ethanol fermentation using an empirical model. *Bioresour Technol* 170:565–573. <https://doi.org/10.1016/j.biortech.2014.07.055>.
43. Thang VH, Kanda K, Kobayashi G. 2010. Production of acetone-butanol-ethanol (ABE) in direct fermentation of cassava by *Clostridium saccharoperbutylacetonicum* N1-4. *Appl Biochem Biotechnol* 161:157–170. <https://doi.org/10.1007/s12010-009-8770-1>.
44. Yoav S, Barak Y, Shamshoum M, Borovok I, Lamed R, Dassa B, Hadar Y, Morag E, Bayer EA. 2017. How does cellulosome composition influence deconstruction of lignocellulosic substrates in *Clostridium (Ruminoclostridium) thermocellum* DSM 1313? *Biotechnol Biofuels* 10:222. <https://doi.org/10.1186/s13068-017-0909-7>.
45. Kakiuchi M, Isui A, Suzuki K, Fujino T, Fujino E, Kimura T, Karita S, Sakka K, Ohmiya K. 1998. Cloning and DNA sequencing of the genes encoding *Clostridium josui* scaffolding protein CipA and cellulase CelD and identification of their gene products as major components of the cellulosome. *J Bacteriol* 180:4303–4308. <https://doi.org/10.1128/JB.180.16.4303-4308.1998>.
46. Vazana Y, Barak Y, Unger T, Peleg Y, Shamshoum M, Ben-Yehzekel T, Mazor Y, Shapiro E, Lamed R, Bayer EA. 2013. A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. *Biotechnol Biofuels* 6:182. <https://doi.org/10.1186/1754-6834-6-182>.
47. Gunnoo M, Cazade P-A, Galera-Prat A, Nash MA, Czjzek M, Cieplak M, Alvarez B, Aguilar M, Karpol A, Gaub H, Carrión-Vázquez M, Bayer EA, Thompson D. 2016. Nanoscale engineering of designer cellulosomes. *Adv Mater* 28:5619–5647. <https://doi.org/10.1002/adma.201503948>.
48. Bayer EA. 2017. Cellulosomes and designer cellulosomes: why toy with Nature? *Environ Microbiol Rep* 9:14–15. <https://doi.org/10.1111/1758-2229.12489>.
49. Nash MA, Smith SP, Fontes CM, Bayer EA. 2016. Single versus dual-binding conformations in cellulosomal cohesin-dockerin complexes. *Curr Opin Struct Biol* 40:89–96. <https://doi.org/10.1016/j.sbi.2016.08.002>.
50. Brás JL, Alves VD, Carvalho AL, Najmudin S, Prates JA, Ferreira LM, Bolam

- DN, Romão MJ, Gilbert HJ, Fontes CM. 2012. Novel *Clostridium thermocellum* type I cohesin-dockerin complexes reveal a single binding mode. *J Biol Chem* 287:44394–44405. <https://doi.org/10.1074/jbc.M112.407700>.
51. Tamaru Y, Miyake H, Kuroda K, Nakanishi A, Matsushima C, Doi RH, Ueda M. 2011. Comparison of the mesophilic cellulosome-producing *Clostridium cellulovorans* genome with other cellulosome-related clostridial genomes. *Microb Biotechnol* 4:64–73. <https://doi.org/10.1111/j.1751-7915.2010.00210.x>.
52. Resch MG, Donohoe BS, Baker JO, Decker SR, Bayer EA, Beckham GT, Himmel ME. 2013. Fungal cellulases and complexed cellulosomal enzymes exhibit synergistic mechanisms in cellulose deconstruction. *Energy Environ Sci* 6:1858–1867. <https://doi.org/10.1039/c3ee00019b>.
53. Zhivin-Nissan O, Dassa B, Morag E, Kupervaser M, Levin Y, Bayer EA. 2019. Unraveling essential cellulosomal components of the (*Pseudo*)*Bacteroides cellulosolvens* reveals an extensive reservoir of novel catalytic enzymes. *Biotechnol Biofuels* 12:115. <https://doi.org/10.1186/s13068-019-1447-2>.
54. Fierobe H-P, Bayer EA, Tardif C, Czjzek M, Mechaly A, Bélaich A, Lamed R, Shoham Y, Bélaich J-P. 2002. Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. *J Biol Chem* 277:49621–49630. <https://doi.org/10.1074/jbc.M207672200>.
55. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786. <https://doi.org/10.1038/nmeth.1701>.
56. Unger T, Jacobovitch Y, Dantes A, Bernheim R, Peleg Y. 2010. Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J Struct Biol* 172:34–44. <https://doi.org/10.1016/j.jsb.2010.06.016>.
57. Caspi J, Irwin D, Lamed R, Shoham Y, Fierobe H-P, Wilson DB, Bayer EA. 2006. *Thermobifida fusca* family-6 cellulases as potential designer cellulosome components. *Biocatal Biotransform* 24:3–12. <https://doi.org/10.1080/10242420600598046>.
58. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A. 2005. Protein identification and analysis tools on the ExPASy server, p 571–607. *In* Walker JM (ed), *The proteomics protocols handbook*. Humana Press Inc, Totawa, NJ.
59. Wood TM. 1988. Preparation of crystalline, amorphous, and dyed cellulase substrates. *Methods Enzymol* 160:19–25. [https://doi.org/10.1016/0076-6879\(88\)60103-0](https://doi.org/10.1016/0076-6879(88)60103-0).
60. Miller GL. 1959. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal Chem* 31:426–428. <https://doi.org/10.1021/ac60147a030>.
61. Ghose T. 1987. Measurement of cellulase activities. *Pure Appl Chem* 59:257–268. <https://doi.org/10.1351/pac198759020257>.
62. Chen C, Cui Z, Xiao Y, Cui Q, Smith SP, Lamed R, Bayer EA, Feng Y. 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. <https://doi.org/10.1016/j.jsb.2014.09.006>.
63. Kahn A, Morais S, Galanopoulou AP, Chung D, Sarai NS, Hengge N, Hatzinikolaou DG, Himmel ME, Bomble YJ, Bayer EA. 2019. Creation of a functional hyperthermostable designer cellulosome. *Biotechnol Biofuels* 12:44. <https://doi.org/10.1186/s13068-019-1386-y>.
64. Al-Shorgani NKN, Ali E, Kalil MS, Yusoff W. 2012. Bioconversion of butyric acid to butanol by *Clostridium saccharoperbutylacetonicum* N1-4 (ATCC 13564) in a limited nutrient medium. *Bioenerg Res* 5:287–293. <https://doi.org/10.1007/s12155-011-9126-6>.