CELL BIOLOGY

10.1126/sciadv.1501254

Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities

Qi Xu,[1,2] Michael G. Resch,[2,3] Kara Podkaminer,[1,2] Shihui Yang,[3] John O. Baker,[1,2] Bryon S. Donohoe,[1,2] Charlotte Wilson,[2,4] Dawn M. Klingeman,[2,4] Daniel G. Olson,[2,5] Stephen R. Decker,[1,2] Richard J. Giannone,[2,4] Robert L. Hettich,[2,4] Steven D. Brown,[2,4] Lee R. Lynd,[2,5] Edward A. Bayer,[6] Michael E. Himmel,[1,2] Yannick J. Bomble[1,2]*

Clostridium thermocellum is the most efficient microorganism for solubilizing lignocellulosic biomass known to date. Its high cellulose digestion capability is attributed to efficient cellulases consisting of both a free-enzyme system and a tethered cellulosomal system wherein carbohydrate active enzymes (CAZymes) are organized by primary and secondary scaffoldin proteins to generate large protein complexes attached to the bacterial cell wall. This study demonstrates that *C. thermocellum* also uses a type of cellulosomal system not bound to the bacterial cell wall, called the "cell-free" cellulosomal system. The cell-free cellulosome complex can be seen as a "long range cellulosome" because it can diffuse away from the cell and degrade polysaccharide substrates remotely from the bacterial cell. The contribution of these two types of cellulosomal systems in *C. thermocellum* was elucidated by characterization of mutants with different combinations of scaffoldin gene deletions. The primary scaffoldin, CipA, was found to play the most important role in cellulose degradation by *C. thermocellum*, whereas the secondary scaffoldins have less important roles. Additionally, the distinct and efficient mode of action of the *C. thermocellum* exoproteome, wherein the cellulosomes splay or divide biomass particles, changes when either the primary or secondary scaffolds are removed, showing that the intact wild-type cellulosomal system is necessary for this essential mode of action. This new transcriptional and proteomic evidence shows that a functional primary scaffoldin plays a more important role compared to secondary scaffoldins in the proper regulation of CAZyme genes, cellodextrin transport, and other cellular functions.

INTRODUCTION

Cellulosic biomass is the largest source of organic matter on earth, providing a promising renewable feedstock for production of biofuels and chemicals. The main bottleneck in biofuel production is the low efficiency of cellulose conversion, leading to high production costs (*1*, *2*). Understanding and improving cellulose degradation by microorganisms is essential to developing cost-competitive biofuels and chemicals. In nature, some microorganisms are able to degrade cellulose very efficiently using varying degrees of complexity in their cellulase systems. There are two major types of known cellulase systems: the cellulosome, which consists of large multienzyme complexes produced by some anaerobic bacteria and fungi, and a system of free, noncomplexed individual enzymes, typically produced by fungi and aerobic bacteria (*3*, *4*). Some microorganisms have both free enzymes and cellulosomes. For example, *C. thermocellum* has 72 proteins in their cellulosomes as well as 25 free enzymes.

C. thermocellum is the most efficient single biomass degrader characterized to date (*5*). The *C. thermocellum* cellulosome was discovered in the early 1980s, and its proposed structural model consisting of multienzyme complexes has been supported by biochemical studies (*6*, *7*) and reviewed recently (*8*). The main structural component of the *C. thermocellum* cellulosomes is CipA, a protein scaffold comprising of nine type I cohesin modules, a type II dockerin module, and a family III carbohydrate binding module known for strong affinity for cellulose. The CipA type I cohesins can bind through strong noncovalent interactions to type I dockerins on enzymatic subunits, thereby integrating various catalytic activities into the complex. In turn, the type II dockerin on CipA can bind to type II cohesins of the so-called secondary scaffoldins that adhere to the microbial cell surface through a surface layer homology (SLH) module that interacts with peptidoglycans (*9*–*11*). Several of these secondary scaffoldins have been identified, such as OlpB, which is the most abundant secondary scaffoldin in *C. thermocellum* and can bind up to seven CipAs. The smaller secondary scaffoldins, SdbA and Orf2p, bind one and two CipA molecules, respectively (*12*–*14*) (Fig. 1). However, new potential secondary scaffoldins have recently been discovered by detailed proteomic analysis and include a putative "free" secondary scaffold, Cthe_0736, that lacks SLH binding modules (*15*).

Thus far, investigations have revealed two main types of cellulosomal systems in terms of organization on scaffoldins. One is the simple cellulosomal system, which uses a single primary scaffoldin, and has been reported in *Clostridium cellulolyticum, Clostridium cellulovorans*, and *Clostridium papyrosolvens* (*13*, *16*, *17*); the other is the multiplex cellulosome, composed of primary scaffoldins (and their bound enzymes), further complexed with secondary scaffoldins, as described earlier for *C. thermocellum* and also found in *Acetivibrio cellulolyticus*, *Clostridium clariflavum*, and other species (*13*, *14*, *18*–*20*).

A genetic approach has been reported for the study of the *C. thermocellum* cellulosomal system. For example, by characterizing mutants in which CipA was either disrupted or deleted, it has been demonstrated that CipA molecules play a critical role in cellulose degradation (*21*–*23*). Some individual secondary scaffoldins also appear to

[1]Biosciences Center, National Renewable Energy Laboratory, Golden, CO 80401, USA. [2]BioEnergy Science Center, Oak Ridge, TN 37831, USA. [3]National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO 80401, USA. [4]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA. [5]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755, USA. [6]The Weizmann Institute of Science, Rehovot 7610001, Israel.
*Corresponding author. E-mail: yannick.bomble@nrel.gov

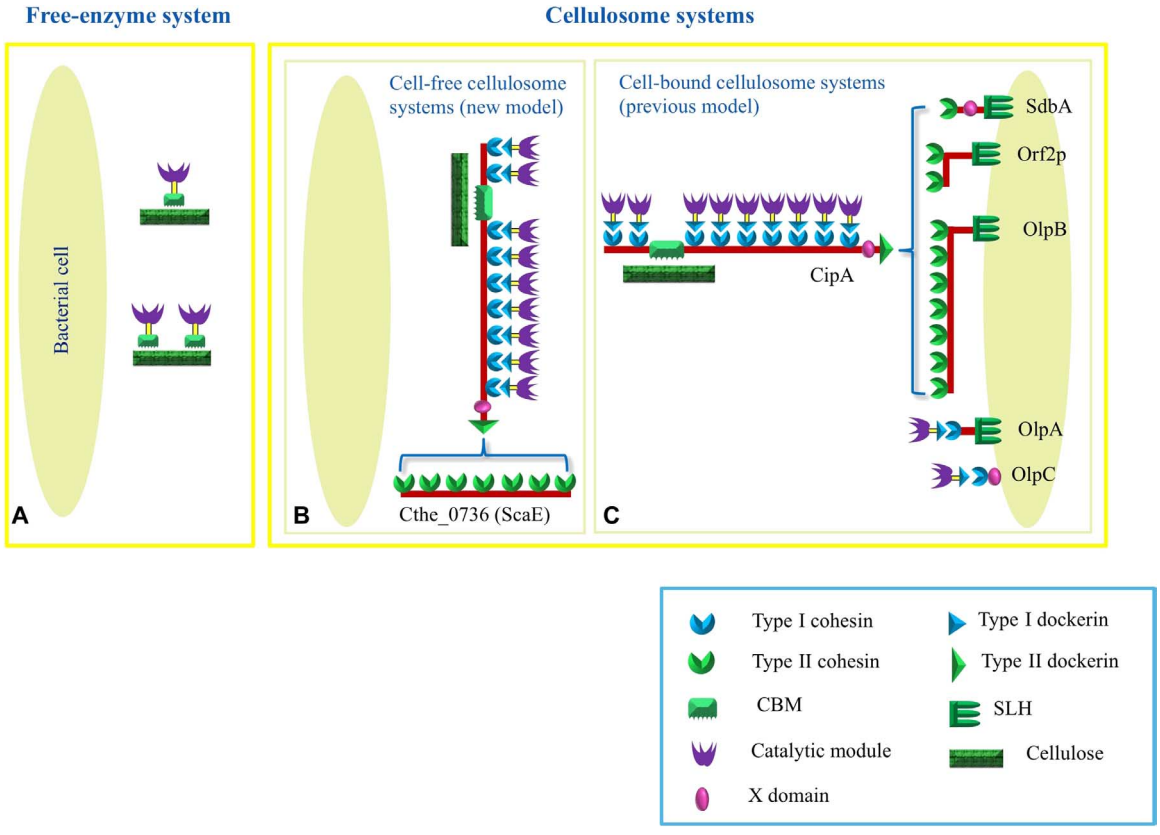


Fig. 1. Model of *C. thermocellum* cellulase systems. (**A** to **C**) *C. thermocellum* cellulase systems consist of a free-enzyme system (A) and cellulosomal systems including both a cell-free cellulosomal system (B), which was discovered in this study, and cell-bound cellulosome systems (C), which were reported previously.

contribute to cellulose digestion by *C. thermocellum*. However, all reported work has pertained to the deletion or disruption of single genes. Simultaneous silencing of scaffolding genes can be expected to provide insights not revealed by single gene silencing, but no such work has been reported before this study.

We report the discovery of a new cell-free cellulosomal system in *C. thermocellum* that is not tethered to the bacterial cell wall and is independent of the primary (tethered) cellulosomes. To elucidate the contribution of these two cellulosomal systems (cell-free and cell-bound) during cellulose degradation, as well as the function of the multiple scaffoldins and their interactions, we systematically engineered mutants by deleting individual scaffoldins or combinations of scaffoldins. These mutants were evaluated for their ability to degrade crystalline cellulose and biomass through activity assays and imaging and were further characterized by transcriptomic and proteomic analysis.

RESULTS

A new model of *C. thermocellum* cellulosomal systems

A putative secondary scaffoldin gene, Clo1313_1487 (Cthe_0736), encoding a polypeptide chain consisting of seven type II cohesin modules, was identified earlier in the *C. thermocellum* genome (*15*). Orthologous proteins bearing seven type II cohesins have also been identified in two other multiplex cellulosome–producing species, *A. cellulolyticus* and *C. clariflavum*, and were termed ScaE (*24*). We adopted the hypothesis that each of the cohesin modules of this gene product, which we shall refer to as ScaE in this study, can specifically bind to a type II dockerin of a CipA protein to form a large cellulosomal complex. Because ScaE lacks SLH modules and other modules that facilitate binding to the *C. thermocellum* cell wall, the cellulosome complex integrated by this scaffoldin is presumed to be unable to specifically bind to the bacterial cell wall. We thus termed this cellulosome as a "cell-free cellulosomal system" to distinguish it from the previously proposed cell-bound cellulosomal system.

The hypothesis that cohesin modules of ScaE bind type II dockerin molecules, as in *C. clariflavum* (*20*) and *A. cellulolyticus* (*25*), was confirmed by in vitro protein binding assays. We expressed and purified two different cohesin modules (cohesin 1 and cohesin 7) of the putative secondary scaffoldin Cthe_0736 (ScaE) (Fig. 2). Using native gel electrophoresis, we have observed that these proteins can bind to a modular dyad (XDoc) composed of the unknown functional domain (X domain) and an adjacent type II dockerin of the primary scaffoldin CipA when mixed at equimolar ratios (figs. S1 and S2). The strong binding exhibited between the cohesin and XDoc modules implies that the putative scaffoldin Cthe_0736 is capable of assembling arrays of CipA molecules and is therefore a scaffoldin, termed ScaE.

The existence of a cell-free cellulosomal system was demonstrated in vivo by proteomic analysis of the scaffoldin mutant CTN5 (see Scheme 1 and Table 1). For this mutant, all secondary scaffoldin genes

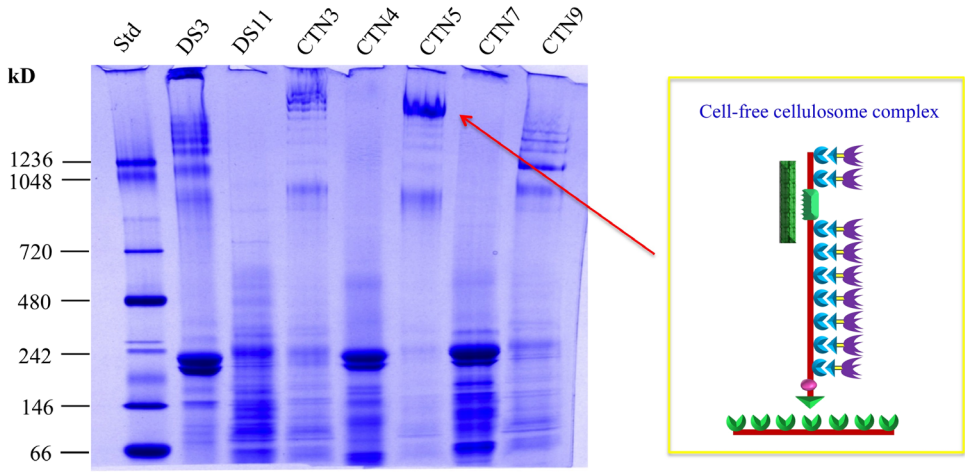
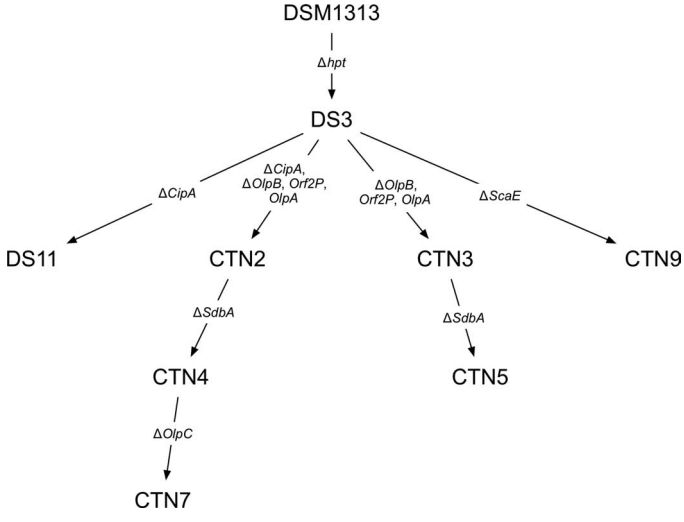


Fig. 2. Cellulosome complexes revealed by native PAGE. Parent and mutants were grown on cellobiose as a carbon source, and their culture supernatants were concentrated and analyzed by native PAGE. The band indicated by the arrow was excised and used for proteomic analysis. This band has been identified as the cell-free cellulosome on the basis of proteomics data. CipA and ScaE as well as 50 other proteins were identified in the complex, with 48 belonging to cellulosomes. The most abundant proteins (for example, CipA, CelS) in the complex are consistent with previous reports, and the scaffoldin deletions were also confirmed.

with SLH modules (SdbA, Orf2P, OlpA, and OlpB) were deleted, whereas CipA and ScaE remained in the genome, which should result in the expression of cell-free cellulosomes. From the CTN5 exoproteome, we identified one large protein complex with a molecular weight (MW) near 2000 kD (Fig. 2). This large protein complex was analyzed by liquid chromatography–tandem mass spectrometry (LC-MS/MS). CipA and ScaE were identified within this complex along with 50 other proteins, 48 of which were dockerin-bearing cellulosomal proteins. Most of these 48 cellulosomal proteins were related to lignocellulose degradation. These results confirm that the primary scaffoldin CipA populated with cellulosomal proteins can bind to the secondary scaffoldin ScaE to generate a large protein complex. The proteins in the large complex may be involved in the degradation of crystalline cellulose, given that cellulases (including Cel48S, Cbh9A, Cel8A, Cel9/44J, and Cel9R) were detected in this protein complex. The other secondary scaffoldins (SdbA, Olp2p, OlpB, and OlpA) were not detected, thus demonstrating correct scaffoldin gene deletion in CTN5. On the basis of these results, we have proposed a new *C. thermocellum* cellulosome model consisting of both cell-free and cell-bound cellulosomal systems (Fig. 1).



Scheme 1. The genetic relationship between the *C. thermocellum* parent strain and mutants.

Contribution of the two types of cellulosomal systems to *C. thermocellum* exoproteome activity

To probe the contribution of scaffoldins and the two types of cellulosomal systems to cellulose degradation by *C. thermocellum*, we created and characterized mutants with deletions of various scaffoldins and their combinations.

Scaffoldin gene deletions and their verification. The DS3 strain (M1354), a mutant produced from wild-type *C. thermocellum* DSM1313 by deletion of the hypoxanthine phosphoribosyltransferase (hpt) gene, was created earlier and is used in the present study as the parent strain in a *C. thermocellum* genetic manipulation system wherein the selection markers are recyclable (*26*). With this genetic system, deletion mutants of the primary and the secondary scaffoldin genes have been created. Deleted genes and their genetic relationships are listed in Table 1 and Scheme 1. Deletion of targeted genes was confirmed by polymerase chain reaction (PCR) (table S1 and figs. S3 and S4) and by DNA sequencing with the exception of the *olpC* deletion. Several mutants were also characterized by transcription profiles (tables S2 and S3 and fig. S5).

Cellulosomal complex formation revealed by native gel. Mutants were investigated for cellulosomal formation by native polyacrylamide gel electrophoresis (PAGE) (Fig. 2), and the results revealed that, (i) after deletion of the individual primary scaffoldin CipA (DS11) or deletion of the gene encoding CipA in combination with those coding for SLH bearing secondary scaffoldins (CTN4 and CTN7), no cellulosome complexes were detected, indicating that the cellulosomal system was not assembled upon deletion of *cipA*. Upon deletion of all SLH-borne secondary scaffoldins and retention of *cipA* and *scaE* in the genome (CTN5), large cellulosome complexes were observed as described above. (ii) The sole deletion of the secondary scaffoldin, ScaE (CTN9), resulted in gel patterns resembling those of parent

Table 1. *C. thermocellum* strains used in this study. See Scheme 1 for relationships between parent and mutants. Genes and their accession numbers: *CipA*, Cthe_3077, Clo1313_0627; *OlpB*, Cthe_3078, Clo1313_0628; *Orf2*, Cthe_3079, Clo1313_0629; *OlpA*, Cthe_3080, Clo1313_0630; *SdbA*, Cthe_1307, Clo1313_0950; *OlpC*, Cthe_0452, Clo1313_1768; *ScaE*, Clo1313_1487, Cthe_0736; *hpt*, hypoxanthine phosphoribosyltransferase, Cthe_2254, Clo1313_2927. Δ indicates that the gene was deleted. ND, not determined.

Strain	Cellulosome system present	Short name used in characterization description	Gene								Verification of gene deletion		
			CipA	*OlpB*	*Orf2*	*OlpA*	*SdbA*	*OlpC*	Cthe_0736 (*ScaE*)	*hpt*	PCR	DNA sequencing	Transcription
DS11	No cellulosome formation	Δ1°	Δ							Δ	Verified	Verified	Verified
CTN3				Δ	Δ	Δ				Δ	Verified	Verified	
CTN5	Cell-free cellulosome system	Δ2°		Δ	Δ	Δ	Δ			Δ	Verified	Verified	Verified
CTN2			Δ	Δ	Δ	Δ				Δ	Verified	Verified	
CTN4	No cellulosome formation		Δ	Δ	Δ	Δ	Δ			Δ	Verified	Verified	Verified
CTN7	No cellulosome formation	Δ1° and Δ2°	Δ	Δ	Δ	Δ	Δ	Δ		Δ	Verified	Not Determined	Verified
CTN9	Cell-bound cellulosome systems	ΔScaE							Δ	Δ	Verified	Verified	
DS3	Cell-free and cell-bound cellulosome systems	Parent								Δ			

(DS3) where cellulosome complexes were generated. Additionally, protein compositions in exoproteomes of the various mutants were analyzed by SDS-PAGE (fig. S6). One band on SDS-PAGE of about 190 kD (the calculated MW of CipA without its signal peptide was 193.8 kD) was not found in the mutants with deletion of the primary scaffoldin (DS11), as well as in the cases of deletions of the combinations of the primary and second scaffoldins (CTN4 and CTN7), whereas this band was present in the parent strain and in mutants (CTN3, CTN5, and CTN9) with deletions only of the secondary scaffoldins (fig. S6).

Contribution of cell-free and cell-bound cellulosomal systems to cellulose degradation

Microbial degradation and utilization of cellulose. To determine the contribution of scaffoldins and the two cellulosomal systems to microbial degradation of cellulose, the knockout strains were grown in the presence of Avicel or deacetylated, dilute acid–pretreated corn stover, and the residual (cellulose or biomass) was measured during the fermentations. The results shown in Fig. 3 demonstrate the overwhelming importance of CipA to both cellulosomal systems. Knocking out CipA disrupts both the cell-bound and free cellulosomal systems, resulting in cells that require nearly 500 hours to achieve the same extent of conversion of Avicel that the parent strain achieves in 45 hours. As indicated by the rightmost two curves in Fig. 3A, it does not matter whether CipA alone is deleted, as in DS11, or whether CipA is deleted along with all of the SLH region–containing secondary scaffoldins (CTN7) because either case results in a drastic and identical decrease in cellulose-degrading capability. The leftmost three curves track the cellulose-degrading capabilities of the parent strain and two deletion-mutants (CTN5 and CTN9) in which the primary scaffoldin CipA is retained and different classes of secondary scaffoldins have been deleted.

In mutant CTN5, all of the SLH-borne secondary scaffoldins have been deleted, resulting in the elimination of the cell-bound cellulosomal system. Mutant CTN9, on the other hand, was created from the parent by the single gene deletion of ScaE, thereby preventing the formation of the cell-free cellulosomal system. The disruption of either of these systems alone is insufficient to significantly reduce the cellulose-degrading capability of the microorganism; both the cell-bound system deletion and the cell-free system deletion strains show cellulose-degrading capabilities essentially equal to that of the parent strain (Fig. 3B). As long as the critical carbohydrate active enzymes can be organized on CipA, further larger-scale organization by secondary scaffoldins does not appear to be crucial.

Exoproteome activity. *C. thermocellum* strains were grown on cellobiose, and their exoproteomes were isolated to determine the contribution of the scaffoldins to the degradation of cellulose and pretreated biomass. The exoproteomes were incubated with Avicel and with deacetylated and subsequently dilute acid–pretreated corn stover (DACS) to measure the enzymatic conversion of glucan. We observed similar patterns of biomass deconstruction independent of the substrates used (Fig. 4, A and B). The exoproteome from the parent strain hydrolyzed Avicel and DACS the fastest. The most detrimental knockouts to enzymatic hydrolysis were observed with mutants in which the primary scaffoldin CipA was deleted and all cellulosomal systems were disrupted. The contribution of multiple secondary scaffoldins to glucan deconstruction is clearly shown by the intermediate hydrolysis rate shown in Fig. 4A, a trend that was also apparent using more complex substrates, such as DACS (Fig. 4B). Deleting the secondary scaffoldins resulted in a reduction in activity of about 40% on Avicel and about

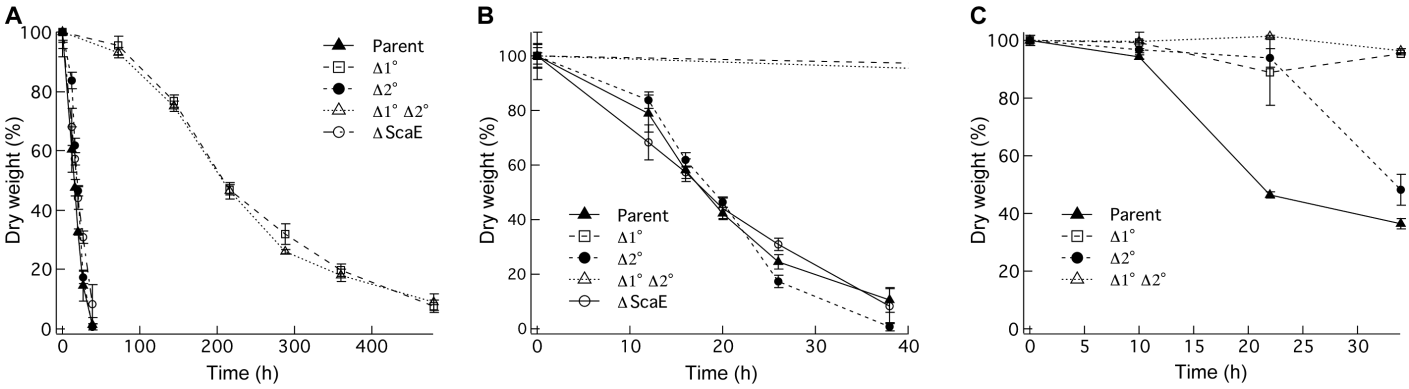


Fig. 3. (**A** to **C**) Comparison of *C. thermocellum* parent and deletion strains grown on Avicel (5 g/liter) (A and B) and deacetylated, dilute acid–pretreated corn stover (C) showing the amount of substrate utilization during fermentation of the parent strain (closed triangle), Δ1° (open square), Δ2° (closed circle), and Δ1° and Δ2° (open triangle), and ΔScaE (open circle; not determined for DACS). The residual biomass was weighed to determine biomass utilization during fermentation.

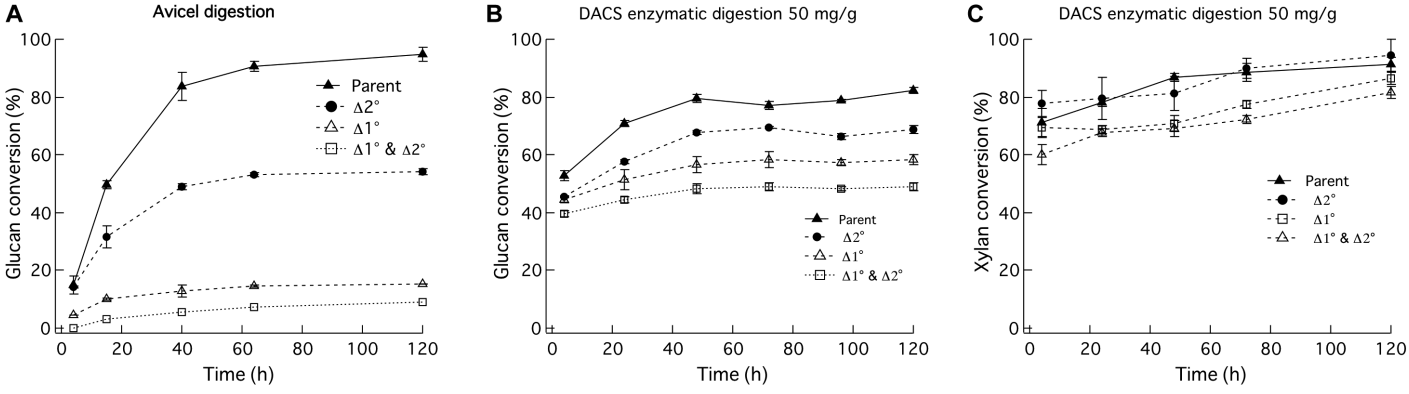


Fig. 4. Enzymatic digestion of cellulose by *C. thermocellum* exoproteomes. (**A** to **C**) *C. thermocellum* exoproteomes were combined with Avicel PH101 (A) and DACS (B and C) in a 1.4-ml slurry of 1% cellulose content. Samples were taken at 4, 12, 48, 96 and 120 hours. Enzymes were loaded equally based on milligram of protein per gram of glucan. Percent glucan (A and B) and xylan (C) conversion was calculated by quantifying the amount of cellobiose plus glucose, or xylose by HPLC, converting to polysaccharide equivalents, and dividing the amount by the initial glucan or xylan contents, respectively.

20% on DACS. After disruption of both the primary and secondary cellulosomal systems, the exoproteome activity was reduced by about 90% on Avicel and about 40% on DACS.

It is interesting that the degradation of xylan in DACS was much less affected by the deletion of scaffoldins. Xylan degradation by mutant exoproteomes with both the primary and secondary scaffoldins deleted is not greatly different from degradation demonstrated by the parent exoproteome, indicating that xylan degradation was not significantly affected by disruption of cellulosome organization (see Fig. 4C). This could be explained by the noncrystalline nature of xylan, which may not be especially susceptible to enzymes arranged in large assemblies, as is the case for the degradation of crystalline cellulose.

Morphology of residual substrates. Isolated cellulosomes were previously observed to separate and splay out the ends of Avicel cellulose particles (*27*). Evidence for the same deconstruction mechanism was observed using the parent strains incubated with Avicel particles in this study (Fig. 5, A and B). However, this deconstruction pattern was not observed in strains without CipA (Fig. 5, C and D). Strains containing cell-free cellulosomes in solution (that is, in which OlpB, Orf2p, and SdbA have been deleted but ScaE and CipA are still expressed) did appear to separate the ends of some Avicel particles but not nearly to the extent of the parent strain (Fig. 5, E and F). Deletion of primary and secondary scaffoldins resulted in low cellulolytic activity, which shows in the transmission electron micrograph where the Avicel particles lack the surface roughness or partitioning of particle segments seen in some of the other samples. These morphological differences suggest that the cellulosomes bound to secondary scaffoldins contribute to the ability of the cellulosome to physically divide or maintain separation within cellulose particles as observed using the parent strain.

Impact of scaffoldin *cipA* deletion on gene expression and translation

To gain further insights into the contributions of the cell-free and cell-bound cellulosomal systems, we conducted a whole genome microarray and extracellular proteomics study for deletion strains DS3, DS11, CTN4, CTN5, and CTN7 grown on Avicel.

In each case, only background levels of DNA microarray signal were detected for probes that spanned deleted gene regions for respective deletion strains, confirming PCR and Sanger sequencing results (table S1). Proteomic results further confirmed these deletions (Fig. 6). Principal

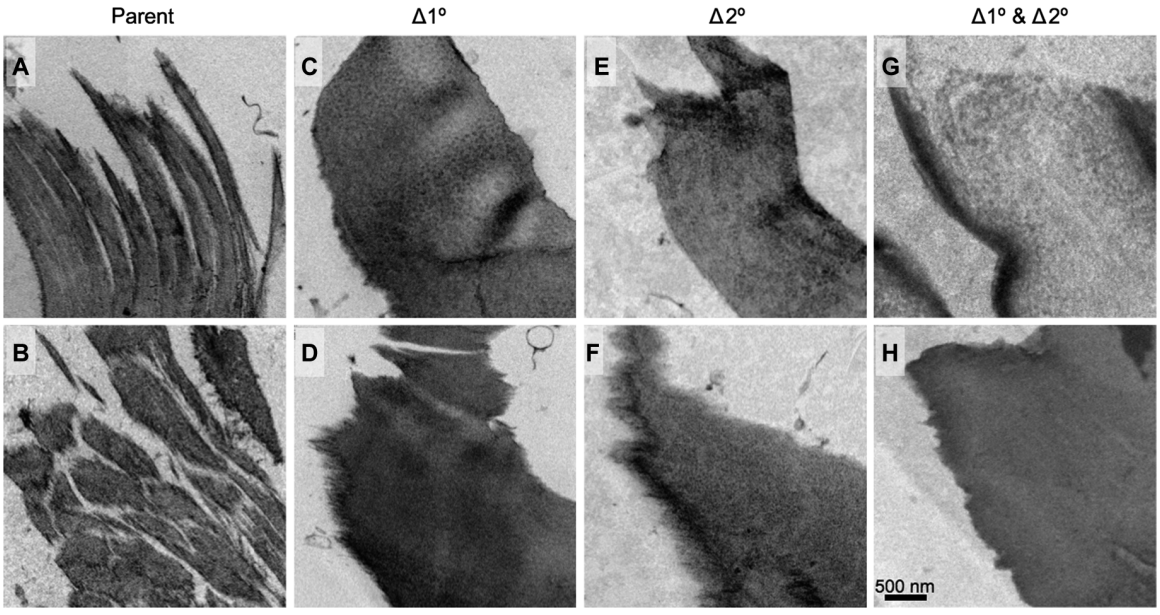


Fig. 5. Transmission electron micrographs of Avicel PH101 particles after incubation with *C. thermocellum* strains. (**A** and **B**) Sectioned Avicel particles digested by the parent bacterial strain displayed a divided or splayed morphology. This pattern suggests that the parent strain cellulosomes with intact scaffoldins are capable of creating or at least maintaining increased substrate surface area during deconstruction. (**C** to **H**) The Avicel particles that had been digested by the 1° scaffoldin deletion strain (C and D), the 2° scaffoldin deletion strain (E and F), or the combined 1° and 2° scaffoldin deletion strains (G and H) appeared very similar, with all three displaying evidence of surface erosion but lacking the internal division. Scale bar, 500 nm.

components analysis (PCA) of transcript profiles showed that biological and technical replicates clustered together for the respective strains (fig. S5 and tables S2 and S3). For like samples (that is, replicates), Pearson *R* correlations for gene-wise comparisons of transcriptome data ranged from 0.89 to 99. Faster-growing strains (that is, the Δ*hpt* parent DS3 and the secondary scaffoldin deletion strain CTN5) clustered with each other, as did *cipA* and derivative deletions strains. Similar results were observed for the extracellular proteomics.

Few altered gene expression patterns were observed in the scaffoldin genes when each strain was compared to the parental DS3 strain (figs. S7 to S9 and table S4). Deletion of *cipA* in strain DS11 yielded an increase in *olpB* gene expression, either as compensation for *cipA* loss or as a polar effect. However, this compensatory effect did not translate to increased protein production because OlpB levels measured in the supernatant were 10- to 40-fold decreased relative to DS3 (based on protein matched-ion intensities) in 100% versus 50% substrate utilization during microbial growth, respectively). The *olpC* scaffoldin had twofold lower gene expression in the *cipA* deletion strain (DS11) compared to the parental strain. When only the 2° scaffoldins were deleted (*olpB*, *orf2*, *olpA*, and *sdbA*) in strain CTN5, an increase in the expression of *scaE* occurred perhaps contributing to the formation of cell-free cellulosomes. These results are corroborated at the protein level where OlpC was not detected in any strain where *cipA* was deleted [CTN4, CTN7 (minimal signal/carry-over), and DS11] but was increased in abundance (more than threefold) in CTN5 but only at 100% substrate utilization. At 50% substrate utilization, OlpC was found to be roughly equivalent to the levels measured in DS3.

The cellulosomal genes Clo1313_0177 (*xynD*), Clo1313_0350 (*lecA*), Clo1313_0349 (*celV*), Clo1313_0500, Clo1313_0693, Clo1313_1424 (*celX*), Clo1313_1425 (*celE*), Clo1313_1494, Clo1313_2188 (*cseP*), Clo1313_2216 (*ctGH43*), Clo1313_2794, and Clo1313_2795 were up-regulated by more than twofold in all three strains (DS11, CTN4, and CTN7) that had *cipA* deleted relative to the parental strain. XynD, CelV, CelE, and Clo1313_2795 were also found to be consistently up-regulated (more than twofold) at the protein level across both substrate utilization levels. Exoproteome proteomics identified seven additional proteins with increased abundance relative to DS3 when *cipA* was knocked out, including Clo1313_0420, Clo1313_1477 (CelW), Clo1313_1808 (CbhA), Clo1313_2043, Clo1313_2635 (XynZ), Clo1313_2747 (CelS), and Clo1313_2861. Of the catalytic components of the cellulosome, four exhibited a more than fivefold increase in abundance compared to DS3, with two containing ample peptide information (>20) to stabilize the abundance measurement: XynD (58-fold) and CelW (sixfold). Conversely, proteins consistently down-regulated (more than twofold) in these same strains (Δ*cipA*) included Clo1313_0689 (CprA), Clo1313_1396 (CelD), Clo1313_2530 (XynC), Clo1313_2793 (carbohydrate-binding family 6), and Clo1313_2843 (pectate lyase). Considering only those enzymes involved in the free enzyme system, Clo1313_0333 was down-regulated in the same three strains (Δ*cipA*), an observation corroborated in both transcriptomic and proteomic measurements.

Distinguishing between a loss-of-function effect due to the absence of *cipA* and altered transcription due to changes in growth rate is difficult because the latter has been shown to affect a number of genes encoding cellulolytic enzymes (*28*). However, only one gene in the current study, *celX* (Clo1313_1424), exhibited similar responses with the previous report, suggesting that the remaining genes could have *cipA*-dependent patterns of gene expression (*28*). There were few genes with consistent changes in the expression of cellulosome-related genes when 2° scaffoldins were deleted irrespective of whether *cipA* was present or absent. Clo1313_0563 (*ctxynGH30*) and Clo1313_2202 (*ctmanF*) were the only two genes differentially expressed (more than

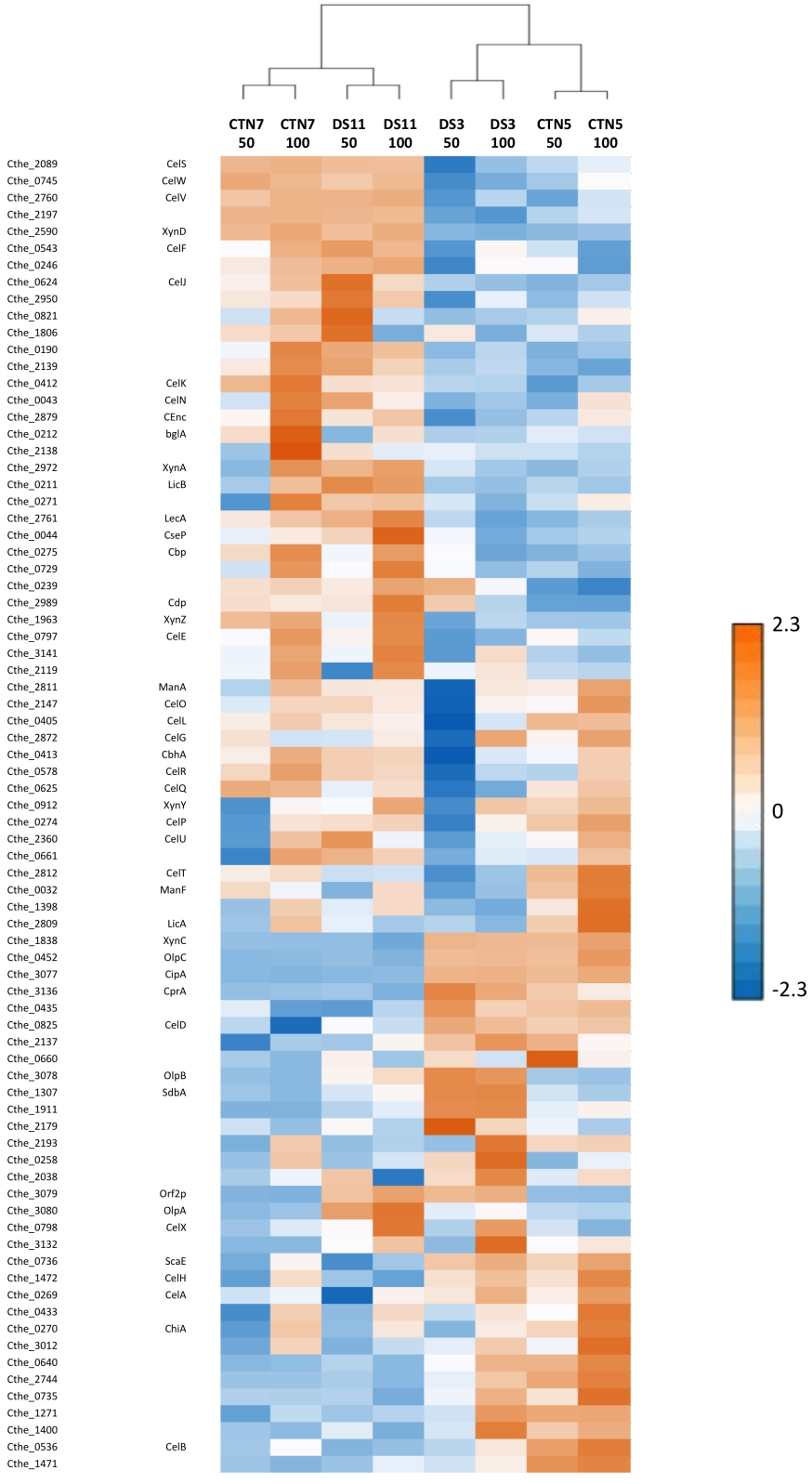


Fig. 6. Heat map depicting cellulosome-associated and nonassociated proteins identified in the supernatant by LC-MS/MS. Protein abundances measured across strains (CTN7, DS11, DS3, and CTN5) and substrate utilization levels (50 and 100%) were standardized by *z*-score (−2.3 blue → 0.00 white → +2.3 orange) and hierarchically clustered (method = Ward; distance = correlation) to group together proteins (rows) and samples (columns) with similar abundance profiles. Sample/strain relationships and their distances are depicted above the heat map.

twofold higher) in CTN4, CTN5, and CTN7 compared to the parental DS3 strain. No cellulosomal genes were differentially expressed when CTN4 and CTN7 strains were directly compared, indicating that the loss of *olpC* in CTN7, (the only differences between the two strains) had little influence on the cellulosome and cellulase gene expression once the 1° and other 2° scaffoldins were deleted.

The deletion of *cipA* affected the expression of genes related to cellodextrin transport and phosphorylation. ABC (adenosine triphosphate–binding cassette) transporters for cellodextrins of G2 to G5 in length (Clo1313_0077-9) and G3-specific transport (Clo1313_1828-30) (*29*) were markedly down-regulated in both Δ1° and Δ1°Δ2° mutants, but not the Δ2° mutant CTN5, compared to wild-type DS3. The transporter for G1 to G5 cellodextrins, Clo1313_2783-86, was up-regulated in all mutant strains relative to the parental DS3, except for Clo1313_2786, the solute binding protein, which was only up-regulated by more than twofold in strain CTN5 (Δ2° mutant). Eleven proteins related to sigma factors, RsgI, occur in *C. thermocellum* and have been associated with cellulosomal gene expression (*30*). The deletion of scaffoldins altered the expression of two pairs of sigma factor–related genes. Clo1313_1961/2 and Clo1313_0524/5 were up-regulated in the Δ1°, Δ1°Δ2°, and Δ2° mutants relative to the parental strain. Clo1313_1961/2 (σ^{I2}-RsgI2) was shown to respond to cellulose in *C. thermocellum* ATCC 27405 (Cthe_0267/8). The other putative sigma factor Clo1313_0524/5 was not as well characterized.

DISCUSSION

Most cellulosome models have not yet been genetically analyzed, specifically by in vivo genetic engineering (*19*, *20*, *31*, *32*). Here, *C. thermocellum* scaffoldin genes have been systematically deleted both individually and in strategic combinations. Large cellulosome complexes have been identified in the exoproteomes of various mutant and parental strains by nondenaturing gel electrophoresis, which not only provides direct evidence of the previously proposed cell-bound cellulosomal systems but also reveals the existence of the newly proposed cell-free cellulosomal system. Here, the unbound *C. thermocellum* scaffoldin, ScaE, was characterized both in vitro and in vivo, where we found that the cellulosomes formed on ScaE contrasted with previously proposed models in which all of the cellulosomes were cell-bound. The cell-free cellulosome complex can thus be regarded as a longer-range cellulosome, in that it can diffuse away from and degrade polysaccharide substrates relatively distant from the bacterial cell. A new model of *C. thermocellum* cellulosomal systems, including both the cell-free and cell-bound cellulosomal systems, was therefore proposed. The study of this cell-free cellulosomal system will provide a more complete understanding of the functioning mechanism of cellulose degradation by *C. thermocellum*.

A cell-free cellulosomal system has been proposed in the anaerobic bacteria *C. clariflavum* and *A. cellulolyticus* but has not been characterized or studied in vivo (*20*). The cell-free cellulosome complex from the cell-free exoproteome in our study was purified by deleting all secondary scaffoldins having SLH (SdbA, Orf2P, and OlpB) and then visualized by nondenaturing gel electrophoresis. The results of these deletions were characterized by transcriptomic and proteomic analysis. This is the first time a study of the primary and secondary scaffoldins as discrete enzyme categories has been conducted systematically.

In terms of the different cellulase systems in *C. thermocellum*—the cellulosomal systems (including both the cell-free and cell-bound systems) and the free, individual enzyme system—it would appear that the cellulosomal systems are the major contributors to cellulose degradation and that the free enzymes are much less important. This difference in importance may be even greater than indicated by the data shown in Figs. 3 and 4, in that the contribution of the free-enzyme system (CTN7, Δ1°Δ2°) may be exaggerated by the presence not only of naturally free (non–dockerin-containing) enzymes but also of non-complexed dockerin-bearing cellulosomal enzymes. The wild-type system is undoubtedly optimized for the substrates normally encountered by the growing cell, which is not the same substrate as the cellulose-enriched biomass used for modern biofuel production.

Previously proposed cellulosomal systems assumed that a fundamental property of cellulosome complexes is that they are anchored to bacterial cell walls. Cell-free exoproteomes, however, always contain large cellulosome complexes, and indeed they were observed in this study as well (Fig. 2). It was also earlier assumed that the cellulosome displacement from the cell surface was probably caused by proteolytic degradation of surface-attached proteins. Serpins (a type of protease inhibitor carrying a type I dockerin) were seen as being included in the complex to inhibit this proteolysis and keep the cellulosome intact. The balance of serpins/proteases could probably determine the proportions of cellulosome being anchored to the cell wall to that released to the medium (*33*, *34*). However, a natural cell-free cellulosomal system was found in this study and was revealed to be fundamentally independent of the cell wall (not being equipped with an SLH domain on the secondary scaffoldin). This result provides another explanation for some of the large cellulosome complexes found in the exoproteome; that is, these species are cell-free cellulosomes and not inherently designed to bind to the microorganism.

In an earlier study, individual scaffoldin genes were disrupted, and the contributions of individual secondary scaffoldins to cellulose degradation were estimated (*22*). Owing to the multiplicity of secondary scaffoldin genes in *C. thermocellum* cellulosomal systems, the conclusions from this previous study were limited because only individual scaffoldin genes were deleted. The lack of a single secondary scaffoldin could then be compensated by the presence of the remaining ones. In the present study, we deleted multiple secondary scaffoldins to focus on the contributions of types and classes of scaffoldins, rather than of individual scaffoldins. We demonstrated here that the secondary scaffoldins play a lesser role than the primary scaffoldin in cellulose degradation. However, transmission electron micrographs of Avicel particles after incubation with *C. thermocellum* showed divided or splayed morphology. This pattern suggests that the parental strain cellulosomes, which have intact scaffoldins, are capable of creating or at least maintaining increased substrate surface area during deconstruction. The Avicel particles that had been digested by the 1°, 2°, and combined 1° and 2° scaffoldin deletion strains appeared very similar in surface morphology—with all three types of particles showing evidence of surface erosion but lacking the divided or spayed morphology associated with wild-type cellulosome action.

Although the effects of scaffoldin gene deletion or disruption on the cellulolytic activity of the *C. thermocellum* exoproteome and cell performance have been investigated (*21*–*23*), no studies have been done on the impact of scaffoldin deletion on the expression and translation of other genes related to cellulose degradation and utilization. Here, transcriptomic and proteomic data from mutants formed by deletion of CipA and other scaffoldins were studied. *cipA* was found to be related to the transcription regulation of numerous global genes,

especially those involved in cellulose degradation, sugar transportation, glycolysis, and energy metabolism, providing new information to further understand the essential role of CipA in cellulose degradation by *C. thermocellum*.

MATERIALS AND METHODS

Gene cloning, gene overexpression, and recombinant protein purification

All primers used for this study are listed in table S5. The target gene (along with coding for a C-terminal His tag) was inserted into the pET22b plasmid (Novagen). The plasmid with gene insertion was then transformed into *Escherichia coli* BL21(DE3) (Stratagene). The gene was overexpressed and induced by 0.3 mM isopropyl-β-D-thiogalactopyranoside at 37°C in LB medium. Cells were broken by sonication, and the recombinant protein was purified using a nickel–nitrilotriacetic acid preparatory column (Qiagene) following the manufacturer's protocol (*31*).

Genetic manipulation system used for *C. thermocellum* gene knockout

The strategy and procedure of the genetic manipulation system for *C. thermocellum* gene knockout followed that of Olson and colleagues (*21*, *35*). The genetic relationship among mutants used in this study is presented in Scheme 1. The correct gene deletion in the mutants was confirmed by a variety of methods including PCR, DNA sequencing, transcription profiling, and proteomics.

Exoproteome protein identification for protein analysis by SDS-PAGE and nondenaturing gel

Cells were grown overnight in CTFUD (*21*, *35*) or MTC (*36*) medium containing 0.5% cellobiose at 55°C. Cell-free supernatants were collected by centrifugation at 10,000*g* and stored in a buffer consisting of 50 mM tris, 150 mM NaCl, 5 mM $CaCl_2$ and 0.02% NaN_3. Supernatant was concentrated by filtration with a membrane of 10 kD molecular weight cutoff (MWCO) (Amicon, EMD Millipore). Protein composition was analyzed by SDS-PAGE (4 to 12%) or native gel (3 to 12%) (Invitrogen).

Microbial cellulose degradation and utilization

C. thermocellum strains were grown in 50-ml anaerobic serum bottles on Avicel PH-101 (5 g/liter) (Sigma-Aldrich) in MTC medium (*36*), with the pH maintained at 6.5 with Mops (5 g/liter). Cultures were grown at 60°C and shaken at 180 rpm. Bottles were inoculated with 2% volume of seed cultures taken during logarithmic growth, stored at −80°C, and thawed on ice. Residual dry weight was assessed by taking 1-ml samples over the course of the fermentation and immediately spinning down the pellet at 20,000*g*. The pellet was washed three times with water without disturbing the pellet. The samples were then dried for 1 hour at 105°C, and the weight of the pellets was measured.

Cell performance test in vivo

To test the performance of *C. thermocellum* strains on cellulose, strains were grown in anaerobic serum bottles with a working volume of 25 ml. Avicel PH-101 (Sigma-Aldrich) was added to each serum bottle to provide a final loading of 5 g glucan per liter plus water. The bottles were then purged with nitrogen and sterilized by autoclaving at 121°C. The remaining MTC medium components were then added anaerobically, and then Mops was added. Cultures were grown as described in the preceding section. For each strain, two bottles were sacrificed at each time point to determine the residual dry weight and glucan composition of the pellet. Each pellet was washed three times with water, dried for 1 hour at 105°C, and then weighed. The ≥400-hour fermentations were grown in a volume of 100 ml in 250-ml anaerobic serum bottles, with all the other conditions the same as described earlier. Ten-milliliter samples were drawn at each time point, from which three 1.0-ml samples were taken. The 1.0-ml samples were pelleted and washed three times with water and then dried at 100°C to determine pellet dry weight.

Preparation of the cellulose substrates

Avicel PH-101 was suspended in deionized water under vacuum overnight at 4°C and washed three times with deionized water by centrifugation at 500*g*. Pellets were resuspended at a concentration of 20 mg/ml (w/w) in 30 mM sodium acetate buffer (pH 5.0) containing 0.001% (w/v) sodium NaN_3. The dried corn stover was milled and deacetylated and pretreated according to Chen *et al.* (*37*).

Exoproteome preparation for cellulase activity assay

C. thermocellum cells were grown in 2.5-liter fermentation volumes, with cellobiose (10 g/liter) in MTC medium with 100 μl of Antifoam 3721. Water and resazurin were sterilized by autoclaving and purged with nitrogen and CO_2 before adding the remaining filter-sterilized medium components. Cultures were inoculated with 50 ml of exponentially growing cultures. The pH was maintained at 7.0 with 2 M KOH. The cells were harvested at the stationary phase. Supernatant and pellets were separated by centrifugation at 13,000*g* for 10 min. The supernatant was then applied to a hollow fiber concentrator with a MWCO of 10 kD (EMD Millipore). The resulting concentrate was filtered using a 0.2-μm vacuum filter. The exoproteomes were analyzed by both denaturing and native PAGE, and protein concentration was determined by the Pierce BCA assay (Thermo Fisher Scientific).

Exoproteome activity assay

C. thermocellum exoproteome activity was determined at 60°C at pH 5.5 in 20 mM sodium acetate buffer containing 5 mM $CaCl_2$, 100 mM NaCl, 2 mM EDTA, and 10 mM cysteine along with *Aspergillus niger* β-glucosidase (5 mg/g). Digestions were conducted in sealed 2.0-ml high-performance liquid chromatography (HPLC) vials with continuous mixing by inversion at 10 rpm. Unless otherwise noted, substrates were loaded at 10 mg of cellulose per milliliter of digestion mixture, with enzymes loaded based on milligram of protein per gram of glucan in 1.4-ml reaction volumes. Representative (with respect to both solid and liquid phases of the digestion slurry) 0.1-ml samples were withdrawn from well-mixed digestion mixtures at selected time points during the digestions. Released cellobiose, glucose, and xylose were determined by HPLC analysis.

Transmission electron microscopy

Samples were centrifuged for 1 min in a desktop microcentrifuge. A slurry (1.5 μl) was placed into the No. 707899 type Leica planchets and cryopreserved in a Leica EMPact2. Cryopreserved samples were then placed in cryotube vials and transferred into vials containing 1-ml aliquots of freeze substitution cocktails for freeze substitution processing. Freeze substitution was carried out in a Leica AFS2 (Leica) unit in

1% OsO_4 in dry acetone with the following temperature regime: −90°C for 72 hours, ramp to −30°C over 3 hours, hold at −30°C for 21 hours, ramp to 3°C over 3 hours, hold at 3°C for 21 hours, ramp to 24°C over 3 hours, hold at 24°C for ~1 hour, rinse three times in dry acetone at room temperature. Samples were removed from the planchets using fine-tipped forceps and minimal agitation before proceeding with infiltration. Samples were infiltrated in a graded series (7.5, 15, 30, 60, 90%, 3 × 100%) of Eponate 812 (EMS) resin over 3 days. Samples were transferred and oriented into Easy Molds (EMS) for polymerization. Resin was polymerized for 48 hours in a vacuum oven at 60°C. Resin-embedded samples were sectioned to ~75-nm thickness with a Diatome diamond knife on a Leica EM UTC ultramicrotome (Leica). Sections were collected on 0.5% formvar–coated slot grids (SPI Supplies). Grids were poststained for 6 min with 2% aqueous uranyl acetate and 3 min with Reynolds lead citrate. Images were taken with a 4-megapixel Gatan UltraScan 1000 camera (Gatan) on a FEI Tecnai G2 20 Twin 200 kV LaB6 transmission electron microscope (FEI). The transmission electron microscopy analysis and representative images were based on the analysis of 166 individual micrographs.

LC-MS/MS analysis

Exoproteome proteins were analyzed by 3 to 12% native gel electrophoresis, and single bands were excised and submitted to the Proteomics and Metabolomics Facility, Colorado State University, Fort Collins, CO, for LC-MS/MS analysis.

MS analysis. Peptides were purified and concentrated using an on-line enrichment column [Thermo Scientific, 5 µm, 100 µm inside diameter (ID) x 2 cm C18 column]. Subsequent chromatographic separation was performed on a reversed-phase nanospray column (Thermo Scientific EASYnano-LC, 3 µm, 75 µm ID × 100 mm C18 column) using a 25-min linear gradient from 5 to 30% buffer B (100% acetonitrile, 0.1% formic acid) at a flow rate of 400 nl/min. Peptides were directly eluted into the mass spectrometer (Thermo Scientific Orbitrap Velos Pro), and spectra were collected over an *m*/*z* (mass/charge ratio) range of 400 to 2000 daltons using a dynamic exclusion limit of two MS/MS spectra of a given peptide mass for 30 s (exclusion duration of 90 s). Compound lists of the resulting spectra were generated using Xcalibur 2.2 software (Thermo Scientific) with a signal-to-noise ratio (S/N) threshold of 1.5 and 1 scan/group.

Data analysis. MS/MS spectra were searched against the National Center for Biotechnology Information nonredundant (NCBI nr) database with a fungus taxonomy filter (version 10/16/2013) using the Mascot database search engine (version 2.3). Search parameters were as follows: monoisotopic mass, parent ion mass tolerance of 20 ppm, fragment ion mass tolerance of 0.8 dalton, fully tryptic peptides with one missed cleavage, variable modification of oxidation of M, and fixed modification of carbamidomethylation of C.

Search results for each independently analyzed sample were imported into the Scaffold software (version 4, Proteome Software). Custom peptide thresholds were based on search-specific Mascot ion scores and Scaffold probabilities; protein probability thresholds of 99% and a minimum of two unique peptides were required. Manual validation of MS/MS spectra was performed for all protein identifications above the probability thresholds that were based on only two unique peptides. The criteria for manual validation were (i) a minimum 80% coverage of theoretical y or b ions (at least five in consecutive order); (ii) absence of prominent unassigned peaks greater than 5% of the maximum intensity; and (iii) indicative residue-specific fragmentation, such as intense ions N-terminal to proline and immediately C-terminal to aspartate and glutamate (used as additional parameters of confirmation.)

Transcription microarray

Cells were grown in anaerobic serum bottles at 60°C with shaking at 150 rpm. Cells were harvested at the time points required to achieve about 50% conversion. To determine the appropriate time points, samples were taken from "extra" bottles, and the dry weight was immediately measured by heat drying at 105°C (the assay of dry weight requiring about 30 min). The cells were harvested from the parallel actual "experimental" bottles at the designated time points by centrifugation at 10,000*g* for 2 min. The supernatant and pellet were immediately frozen separately by liquid nitrogen and then stored at −80°C. The pellet was used for transcription studies, and the supernatant was used for the HPLC (table S6) and proteomics analyses.

RNA was isolated from 25 ml of culture using the TRIzol reagent and a bead beating method, as previously described by Yang *et al.* (*38*). Labeled complementary DNA was prepared and hybridized to a NimbleGen 12-plex array (Roche NimbleGen Inc.) following the manufacturer's protocols for gene expression. Probe intensity values were $\log_2$-transformed on importation into JMP Genomics version 6 (SAS Institute). Data were normalized using the standard LOESS normalization algorithm in JMP Genomics version 6, and significant differential expression was determined by the analysis of variance method with a 5% false discovery rate.

Array data have been deposited at the NCBI Gene Expression Omnibus (GEO) database under accession number GSE63883. *C. thermocellum* transcript data were generated using an established *C. thermocellum* strain 27405 DNA microarray platform that contains five to seven unique probes per gene and with three technical replicates for each unique probe, as described previously by Yang *et al.* (*38*) and Wilson *et al.* (*39*, *40*). *C. thermocellum* strains DSM1313 and ATCC 27405 are closely related, with average nucleotide identities (ANIs) of 99.6 and 99.3% in reciprocal genome comparisons, indicating that the strains are very closely related. Strain ATCC 27405 has a high affinity phosphate transport system that DSM1313 lacks (*39*), and ATCC 27405 contains additional prophage and restriction-modification sequences (*41*); hence, the ATCC 27405 DNA microarray was suitable for the assessment of *C. thermocellum* strain DSM1313.

Fermentation supernatant analysis by HPLC

Cell-free fermentation supernatants collected at the different time points as described above were analyzed by HPLC to identify and quantify soluble fermentation products (acetic acid, lactic acid, and ethanol) and residual carbohydrates (glucose, cellobiose, and xylose). HPLC analysis results are shown in table S6. Quantification was performed by separation on an Aminex HPX-87H 300 × 7.8 mm column (Bio-Rad) using the LaChrom Elite HPLC system (Hitachi High Technologies America) equipped with a refractive index detector (model L-2490) with a flow rate of 0.5 ml/min in 5 mM H_2SO_4 and a run time set at 35 min, as previously described by Wilson *et al.* (*39*). The metabolites and residual carbohydrates were identified by comparison with retention times and peak areas of corresponding standards.

Supernatant proteome analysis

Exoproteome analysis by LC-MS/MS was performed on precleared supernatants obtained from cultures of the various strains harvested at both 50 and 100% substrate (Avicel) utilization. Fifty milliliters

of culture supernatant was concentrated to 1.0-ml protein with trichloroacetic acid, as previously described by Blumer-Schuette *et al.* (*42*). The acetone-washed protein pellet was resolubilized in urea, and proteins were processed with dithiothreitol and iodoacetamide and digested with trypsin. Tryptic peptides were quantified by BCA assay (Pierce), and 10 μg was loaded through pressure cell onto a biphasic MudPIT column for online two-dimensional HPLC separation (strong-cation exchange and reversed-phase) and concurrent analysis by nanospray MS/MS using a hybrid LTQ-Orbitrap XL mass spectrometer (Thermo Scientific) operating in data-dependent acquisition (one full scan at 15k resolution followed by 10 MS/MS scans in the LTQ, all one μscan). Three salt cuts of 50, 100, and 500 mM ammonium acetate were performed per sample run, with each followed by a 120-min organic gradient to separate the peptides.

Resultant peptide fragmentation spectra (MS/MS) were searched against the *C. thermocellum* 1313 proteome database concatenated with common contaminants and reversed sequences to control false discovery rates using Myrimatch v.2.1 (*43*). Peptide spectrum matches (PSMs) were filtered by IDPicker v.3 (*44*) and assigned matched-ion intensities (MITs) based on observed peptide fragment peaks. PSM MITs were summed on a per-peptide basis, and only those uniquely and specifically matching a particular protein were moved onto subsequent analysis with InfernoRDN (*45*, *46*). Briefly, peptide intensity distributions were log-transformed, normalized across biological replicates by LOESS, and standardized by median absolute deviation and mean centering across samples as suggested. Peptides were then filtered to maintain at least two hits in one replicate set, and missing values were imputed using a random distribution of low-level values. Protein abundances were determined using the RRollup method whereby peptide abundance trends for each protein were scaled to a specific, well-sampled reference peptide. Sample-to-sample variation was visualized by PCA, Pearson's correlation and hierarchically clustered using the Ward agglomeration method to generate a heat map of protein abundance trends normalized by *z*-score. Protein abundances were then compared across strains and culture harvest points to identify proteins with differential expression.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/2/2/e1501254/DC1

Table S1. Expected and actual PCR sizes (kb) of various strains.

Table S2. Data set of transcriptomic profiling results for *C. thermocellum* DSM1313 scaffoldin deletion strains detected by microarray.

Table S3. Table of the ranked abundance for selected cellulosome and cellulase genes.

Table S4. Supernatant proteomic profiling results for *C. thermocellum* DSM1313 scaffoldin deletion strains measured by LC-MS/MS.

Table S5. Primers used in this study.

Table S6. HPLC analysis of fermentation products.

Fig. S1. Modular architectures of the putative secondary scaffoldin Cthe_0736 (ScaE) and the primary scaffoldin CipA, as well as their three purified recombinant proteins expressed in *E. coli*.

Fig. S2. Cohesins of the putative secondary scaffoldin ScaE binding to type II dockerin of CipA revealed by native PAGE.

Fig. S3. Identification of correct gene deletion in various mutants by PCR.

Fig. S4. Identification of *scaE* deletion by PCR.

Fig. S5. Correlations between biological and technical replicates analyzed by microarray.

Fig. S6. Identification of CipA protein in the exoproteomes of mutants by SDS-PAGE.

Fig. S7. Exoproteome protein abundances cluster into two distinct groups: samples with an intact CipA (DS3 and CTN5) and those without (DS11, CTN4, and CTN7).

Fig. S8. Sample-to-sample correlation at the protein level (actual values).

Fig. S9. PCA depicts two major sample groups: those with an intact CipA (DS3 and CTN5) and those with CipA deleted (DS11, CTN4, and CTN7).

REFERENCES AND NOTES

1. M. E. Himmel, S.-Y. Ding, D. K. Johnson, W. S. Adney, M. R. Nimlos, J. W. Brady, T. D. Foust, Biomass recalcitrance: Engineering plants and enzymes for biofuels production. *Science* **315**, 804–807 (2007).
2. L. R. Lynd, P. J. Weimer, W. H. van Zyl, I. S. Pretorius, Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol. Mol. Biol. Rev.* **66**, 506–577 (2002).
3. M. E. Himmel, Q. Xu, Y. Luo, S.-Y. Ding, R. Lamed, E. A. Bayer., Microbial enzyme systems for biomass conversion: Emerging paradigms. *Biofuels* **1**, 323–341 (2010).
4. H. Wei, Q. Xu, L. E. Taylor II, J. O. Baker, M. P. Tucker, S.-Y. Ding, Natural paradigms of plant cell wall degradation. *Curr. Opin. Biotechnol.* **20**, 330–338 (2009).
5. J. Paye, A. Guseva, S. Hammer, M. Balch, M. Davis, E. Gjersing, B. Donohoe, B. H. Davison, M. G. Hahn, S. Pattathil, L. Lynd, Biological lignocellulose solubilization: Comparative evaluation of biocatalysts and enhancement via cotreatment. *Biotechnol. Biofuels* **9**, 8 (2016).
6. R. Lamed, R. Kenig, E. Setter, E. A. Bayer, Major characteristics of the cellulolytic system of *Clostridium thermocellum* coincide with those of the purified cellulosome. *Enzyme Microb. Tech.* **7**, 37–41 (1985).
7. R. Lamed, E. Setter, R. Kenig, E. A. Bayer, The cellulosome–a discrete cell-surface organelle of *Clostridium thermocellum* which exhibits separate antigenic, cellulose-binding and various cellulolytic activities. *Biotechnol. Bioeng.* **13**, 163–181 (1983).
8. S. E. Blumer-Schuette, S. D. Brown, K. B. Sander, E. A. Bayer, I. Kataeva, J. V. Zurawski, J. M. Conway, M. W. W. Adams, R. M. Kelly, Thermophilic lignocellulose deconstruction. *FEMS Microbiol. Rev.* **38**, 393–448 (2014).
9. S. Chauvaux, M. Matuschek, P. Beguin, Distinct affinity of binding sites for S-layer homologous domains in *Clostridium thermocellum* and *Bacillus anthracis* cell envelopes. *J. Bacteriol.* **181**, 2455–2458 (1999).
10. M. Lemaire, I. Miras, P. Gounon, P. Béguin, Identification of a region responsible for binding to the cell wall within the S-layer protein of *Clostridium thermocellum*. *Microbiology* **144**, 211–217 (1998).
11. G. Olabarría, J. L. Carrascosa, M. A. dePedro, J. Berenguer, A conserved motif in S-layer proteins is involved in peptidoglycan binding in *Thermus thermophilus*. *J. Bacteriol.* **178**, 4765–4772 (1996).
12. E. A. Bayer, L. J. W. Shimon, Y. Shoham, R. Lamed, Cellulosomes—structure and ultrastructure. *J. Struct. Biol.* **124**, 221–234 (1998).
13. E. A. Bayer, J.-P. Belaich, Y. Shoham, R. Lamed, The cellulosomes: Multienzyme machines for degradation of plant cell wall polysaccharides. *Annu. Rev. Microbiol.* **58**, 521–554 (2004).
14. E. A. Bayer, H. Chanzy, R. Lamed, Y. Shoham, Cellulose, cellulases and cellulosomes. *Curr. Opin. Struct. Biol.* **8**, 548–557 (1998).
15. B. Raman, C. Pan, G. B. Hurst, M. Rodriguez Jr., C. K. McKeown, P. K. Lankford, N. F. Samatova, J. R. Mielenz, Impact of pretreated switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: A quantitative proteomic analysis. *PLOS One* **4**, e5271 (2009).
16. R. H. Doi, A. Kosugi, Cellulosomes: Plant-cell-wall-degrading enzyme complexes. *Nat. Rev. Microbiol.* **2**, 541–551 (2004).
17. M. Desvaux, The cellulosome of *Clostridium cellulolyticum*. *Enzyme Microb. Tech.* **37**, 373–385 (2005).
18. Q. Xu, W. Gao, S.-Y. Ding, R. Kenig, Y. Shoham, E. A. Bayer, R. Lamed, The cellulosome system of *Acetivibrio cellulolyticus* includes a novel type of adaptor protein and a cell surface anchoring protein. *J. Bacteriol.* **185**, 4548–4557 (2003).
19. Q. Xu, Y. Barak, R. Kenig, Y. Shoham, E. A. Bayer, R. Lamed, A novel *Acetivibrio cellulolyticus* anchoring scaffoldin that bears divergent cohesins. *J. Bacteriol.* **186**, 5782–5789 (2004).
20. L. Artzi, B. Dassa, I. Borovok, M. Shamshoum, R. Lamed, E. A. Bayer, Cellulosomics of the cellulolytic thermophile *Clostridium clariflavum*. *Biotechnol. Biofuels* **7**, 100 (2014).
21. D. G. Olson, R. J. Giannone, R. L. Hettich, L. R. Lynd, Role of the CipA scaffoldin protein in cellulose solubilization, as determined by targeted gene deletion and complementation in *Clostridium thermocellum*. *J. Bacteriol.* **195**, 733–739 (2013).
22. W. Hong, J. Zhang, Y. Feng, G. Mohr, A. M. Lambowitz, G.-Z. Cui, Y.-J. Liu, Q. Cui, The contribution of cellulosomal scaffoldins to cellulose hydrolysis by *Clostridium thermocellum* analyzed by using thermotargetrons. *Biotechnol. Biofuels* **7**, 80 (2014).
23. V. V. Zverlov, M. Klupp, J. Krauss, W. H. Schwarz, Mutations in the scaffoldin gene, *cipA*, of *Clostridium thermocellum* with impaired cellulosome formation and cellulose hydrolysis: Insertions of a new transposable element, IS*1447*, and implications for cellulase synergism on crystalline cellulose. *J. Bacteriol.* **190**, 4321–4327 (2008).
24. B. Dassa, I. Borovok, R. Lamed, B. Henrissat, P. Coutinho, C. L. Hemme, Y. Huang, J. Zhou, E. A. Bayer, Genome-wide analysis of *acetivibrio cellulolyticus* provides a blueprint of an elaborate cellulosome system. *BMC Genomics* **13**, 210 (2012).
25. Y. Hamberg, V. Ruimy-Israeli, B. Dassa, Y. Barak, R. Lamed, K. Cameron, C. M. G. A. Fontes, E. A. Bayer, D. B. Fried, Elaborate cellulosome architecture of *Acetivibrio cellulolyticus* revealed by selective screening of cohesin–dockerin interactions. *PeerJ* **2**, e636 (2014).
26. D. A. Argyros, S. A. Tripathi, T. F. Barrett, S. R. Rogers, L. F. Feinberg, D. G. Olson, J. M. Foden, B. B. Miller, L. R. Lynd, D. A. Hogsett, N. C. Caiazza, High ethanol titers from cellulose by

using metabolically engineered thermophilic, anaerobic microbes. *Appl. Environ. Microbiol.* **77**, 8288–8294 (2011).
27. M. G. Resch, B. S. Donohoe, J. O. Baker, S. R. Decker, E. A. Bayer, G. T. Beckham, M. E. Himmel, Fungal cellulases and complexed cellulosomal enzymes exhibit synergistic mechanisms in cellulose deconstruction. *Energy Environ. Sci.* **6**, 1858–1867 (2013).
28. A. Riederer, T. E. Takasuka, S.-i. Makino, D. M. Stevenson, Y. V. Bukhman, N. L. Elsen, B. G. Fox, Global gene expression patterns in *Clostridium thermocellum* as determined by microarray analysis of chemostat cultures on cellulose or cellobiose. *Appl. Environ. Microbiol.* **77**, 1243–1253 (2011).
29. Y. Nataf, S. Yaron, F. Stahl, R. Lamed, E. A. Bayer, T.-H. Scheper, A. L. Sonenshein, Y. Shoham, Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. *J. Bacteriol.* **191**, 203–209 (2009).
30. Y. Nataf, L. Bahari, H. Kahel-Raifer, I. Borovok, R. Lamed, E. A. Bayer, A. L. Sonenshein, Y. Shoham, *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 18646–18651 (2010).
31. Q. Xu, S.-Y. Ding, R. Brunecky, Y. J. Bomble, M. E. Himmel, J. O. Baker, Improving activity of minicellulosomes by integration of intra- and intermolecular synergies. *Biotechnol. Biofuels* **6**, 126 (2013).
32. Q. Xu, E. A. Bayer, M. Goldman, R. Kenig, Y. Shoham, R. Lamed, Architecture of the *Bacteroides cellulosolvens* cellulosome: Description of a cell surface-anchoring scaffoldin and a family 48 cellulase. *J. Bacteriol.* **186**, 968–977 (2004).
33. W. H. Schwarz, V. V. Zverlov, Protease inhibitors in bacteria: An emerging concept for the regulation of bacterial protein complexes? *Mol. Microbiol.* **60**, 1323–1326 (2006).
34. S. Kang, Y. Barak, R. Lamed, E. A. Bayer, M. Morrison, The functional repertoire of prokaryote cellulosomes includes the serpin superfamily of serine proteinase inhibitors. *Mol. Microbiol.* **60**, 1344–1354 (2006).
35. D. G. Olson, L. R. Lynd, Chapter seventeen – transformation of *Clostridium Thermocellum* by electroporation. *Methods Enzymol.* **510**, 317–330 (2012).
36. E. K. Holwerda, K. D. Hirst, L. R. Lynd, A defined growth medium with very low background carbon for culturing *Clostridium thermocellum*. *J. Ind. Microbiol. Biotechnol.* **39**, 943–947 (2012).
37. X. Chen, E. Kuhn, W. Wang, S. Park, K. Flanegan, O. Trass, L. Tenlep, L. Tao, M. Tucker, Comparison of different mechanical refining technologies on the enzymatic digestibility of low severity acid pretreated corn stover. *Bioresour. Technol.* **147**, 401–408 (2013).
38. S. Yang, R. J. Giannone, L. Dice, Z. K. Yang, N. L. Engle, T. J. Tschaplinski, R. L. Hettich, S. D. Brown, *Clostridium thermocellum* ATCC27405 transcriptomic, metabolomic and proteomic profiles after ethanol stress. *BMC Genomics* **13**, 336 (2012).
39. C. M. Wilson, M. Rodriguez Jr., C. M. Johnson, S. L. Martin, T. M. Chu, R. D. Wolfinger, L. J. Hauser, M. L. Land, D. M. Klingeman, M. H. Syed, A. J. Ragauskas, T. J. Tschaplinski, J. R. Mielenz, S. D. Brown, Global transcriptome analysis of *Clostridium thermocellum* ATCC 27405 during growth on dilute acid pretreated *Populus* and switchgrass. *Biotechnol. Biofuels* **6**, 179 (2013).
40. C. M. Wilson, S. Yang, M. Rodriguez Jr., Q. Ma, C. M. Johnson, L. Dice, Y. Xu, S. D. Brown, *Clostridium thermocellum* transcriptomic profiles after exposure to furfural or heat stress. *Biotechnol. Biofuels* **6**, 131 (2013).
41. S. D. Brown, S. Nagaraju, S. Utturkar, S. De Tissera, S. Segovia, W. Mitchell, M. L. Land, A. Dassanayake, M. Köpke, Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of *Clostridium autoethanogenum* and analysis of CRISPR systems in industrial relevant Clostridia. *Biotechnol. Biofuels* **7**, 40 (2014).
42. S. E. Blumer-Schuette, R. J. Giannone, J. V. Zurawski, I. Ozdemir, Q. Ma, Y. Yin, Y. Xu, I. Kataeva, F. L. Poole II, M. W. W. Adams, S. D. Hamilton-Brehm, J. G. Elkins, F. W. Larimer, M. L. Land, L. J. Hauser, R. W. Cottingham, R. L. Hettich, R. M. Kelly, *Caldicellulosiruptor* core and pangenomes reveal determinants for noncellulosomal thermophilic deconstruction of plant biomass. *J. Bacteriol.* **194**, 4015–4028 (2012).
43. D. L. Tabb, C. G. Fernando, M. C. Chambers, MyriMatch: Highly accurate tandem mass spectral peptide identification by multivariate hypergeometric analysis. *J. Proteome Res.* **6**, 654–661 (2007).
44. Z.-Q. Ma, S. Dasari, M. C. Chambers, M. D. Litton, S. M. Sobecki, L. J. Zimmerman, P. J. Halvey, B. Schilling, P. M. Drake, B. W. Gibson, D. L. Tabb, IDPicker 2.0: Improved protein assembly with high discrimination peptide identification filtering. *J. Proteome Res.* **8**, 3872–3881 (2009).
45. A. D. Polpitiya, W.-J. Qian, N. Jaitly, V. A. Petyuk, J. N. Adkins, D. G. Camp II, G. A. Anderson, R. D. Smith, DAnTE: A statistical tool for quantitative analysis of -omics data. *Bioinformatics* **24**, 1556–1558 (2008).
46. T. Taverner, Y. V. Karpievitch, A. D. Polpitiya, J. N. Brown, A. R. Dabney, G. A. Anderson, R. D. Smith, DanteR: An extensible R-based tool for quantitative analysis of -omics data. *Bioinformatics* **28**, 2404–2406 (2012).

Acknowledgments

Funding: This work was supported by the BioEnergy Science Center (BESC). BESC is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the U.S. DOE Office of Science. E.A.B.'s research is supported by the United States–Israel Binational Science Foundation (BSF), Jerusalem, Israel; the Israel Science Foundation (ISF grant no. 1349); the Israeli Center of Research Excellence (I-CORE Center No. 152/11); the European Union NMP.2013.1.1-2: CellulosomePlus Project number 604530; and the ERA-IB Consortium (EIB.12.022) FiberFuel. **Author contributions:** Q.X. designed the study, created the mutant strains, provided help with the microbial digestions, and wrote the manuscript. M.G.R. conducted the biomass digestion assays and wrote the manuscript. K.P. conducted the microbial digestions and edited the manuscript. S.Y. analyzed the transcriptomics data and edited the manuscript. J.O.B. designed the study, provided help with the biomass digestion assays, and wrote the manuscript. B.S.D. imaged the Avicel samples and edited the manuscript. C.W. conducted the fermentations for transcriptomics analysis, analyzed the data, and edited the manuscript. D.M.K. conducted the fermentations for transcriptomics analysis and edited the manuscript. D.G.O. created the first DS3 mutant strain, provided expertise on the *C. thermocellum* transformation system, and edited the manuscript. S.R.D. provided expertise on fermentations and edited the manuscript. R.J.G. conducted the proteomics analysis, analyzed the data, and edited the manuscript. R.L.H. analyzed the proteomics data and edited the manuscript. S.D.B. analyzed the transcriptomics data and edited the manuscript. L.R.L. provided expertise on the *C. thermocellum* transformation system and edited the manuscript. E.A.B. provided expertise on the cellulosome and *C. thermocellum* and edited the manuscript. M.E.H. designed the study and wrote the manuscript. Y.J.B. designed the study, supervised the research, and wrote the manuscript. All authors discussed the results and implications and commented on the manuscript at all stages. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and the Supplementary Materials or available upon request from the authors.

Submitted 10 September 2015
Accepted 30 November 2015
Published 5 February 2016
10.1126/sciadv.1501254

Citation: Q. Xu, M. G. Resch, K. Podkaminer, S. Yang, J. O. Baker, B. S. Donohoe, C. Wilson, D. M. Klingeman, D. G. Olson, S. R. Decker, R. J. Giannone, R. L. Hettich, S. D. Brown, L. R. Lynd, E. A. Bayer, M. E. Himmel, Y. J. Bomble, Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv.* **2**, e1501254 (2016).

ScienceAdvances
AAAS

Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities

Qi Xu, Michael G. Resch, Kara Podkaminer, Shihui Yang, John O. Baker, Bryon S. Donohoe, Charlotte Wilson, Dawn M. Klingeman, Daniel G. Olson, Stephen R. Decker, Richard J. Giannone, Robert L. Hettich, Steven D. Brown, Lee R. Lynd, Edward A. Bayer, Michael E. Himmel and Yannick J. Bomble (February 5, 2016)

Sci Adv 2016, 2:.
doi: 10.1126/sciadv.1501254