Deletion of the Cel48S cellulase from *Clostridium thermocellum*

Daniel G. Olson[a,b,c], Shital A. Tripathi[a,c], Richard J. Giannone[c,d], Jonathan Lo[c,e], Nicky C. Caiazza[a,c], David A. Hogsett[a,c], Robert L. Hettich[c,d], Adam M. Guss[b,c], Genia Dubrovsky[b,c], and Lee R. Lynd[a,b,c,e,1]

[a]Mascoma Corporation, Lebanon, NH 03766; [b]Thayer School of Engineering and [e]Department of Biological Sciences, Dartmouth College, Hanover, NH 03755; [c]BioEnergy Science Center, Oak Ridge, TN 37830; and [d]Oak Ridge National Laboratory, Oak Ridge, TN 37830

Edited* by Lonnie O'Neal Ingram, University of Florida, Gainesville, FL, and approved August 16, 2010 (received for review April 9, 2010)

Clostridium thermocellum **is a thermophilic anaerobic bacterium that rapidly solubilizes cellulose with the aid of a multienzyme cellulosome complex. Creation of knockout mutants for Cel48S (also known as CelS, S_S, and S8), the most abundant cellulosome subunit, was undertaken to gain insight into its role in enzymatic and microbial cellulose solubilization. Cultures of the Cel48S deletion mutant (S mutant) were able to completely solubilize 10 g/L crystalline cellulose. The cellulose hydrolysis rate of the S mutant strain was 60% lower than the parent strain, with the S mutant strain also exhibiting a 40% reduction in cell yield. The cellulosome produced by the S mutant strain was purified by affinity digestion, characterized enzymatically, and found to have a 35% lower specific activity on Avicel. The composition of the purified cellulosome was analyzed by tandem mass spectrometry with APEX quantification and no significant changes in abundance were observed in any of the major (>1% of cellulosomal protein) enzymatic subunits. Although most cellulolytic bacteria have one family 48 cellulase,** ***C. thermocellum*** **has two, Cel48S and Cel48Y. Cellulose solubilization by a Cel48S and Cel48Y double knockout was essentially the same as that of the Cel48S single knockout. Our results indicate that solubilization of crystalline cellulose by** ***C. thermocellum*** **can proceed to completion without expression of a family 48 cellulase.**

cellulosome | CelS | family 48 | exoglucanase

A key obstacle to the cost-effective production of cellulosic biofuels is the recalcitrance of cellulosic biomass (1). Consolidated bioprocessing (CBP), featuring cellulase production, cellulose solubilization, and fermentation in a single, integrated step, is a promising strategy for reducing processing costs (2). Development of microorganisms capable of mediating consolidated bioprocessing can proceed either by conferring cellulolytic capability to microbes that have strong product formation properties or by improving product formation in microbes that have strong cellulolytic capability. *Clostridium thermocellum* has received consideration in the context of the latter strategy. This anaerobic, thermophilic bacterium exhibits one of the highest rates of cellulose solubilization among described microbes (3, 4) and produces a cellulase enzyme complex called a cellulosome that is noted for being highly effective at solubilizing crystalline cellulose (5, 6).

Family 48 glycoside hydrolase (GH48) enzymes are highly expressed in truly cellulolytic bacteria (7) including *Clostridium cellulolyticum*, *Clostridium cellulovorans*, *Clostridium josui*, *Clostridium phytofermentans*, and *C. thermocellum* (8–11). Cellulases from family 48, along with family 9, are up-regulated during growth of C. *thermocellum* on crystalline cellulose as compared with cellobiose (12, 13). In light of these and other considerations, family GH48 enzymes are thought to play an essential role in bacterial cellulolytic systems (5, 7). Whereas most cellulolytic bacteria have one family 48 enzyme, *C. thermocellum* has two, Cel48S and Cel48Y. Cel48Y, which does not have a dockerin domain and therefore is not part of the cellulosome, is believed to form a separate, soluble, cellulolytic system in combination with a handful of other noncellulosomal cellulases (7). Cel48S is the most abundant enzymatic subunit in the cellulosome (12–15); however, the extent to which Cel48S, or for that matter any catalytic subunit, contributes to cellulosome function is not known.

Targeted gene deletion followed by biochemical and microbiological characterization is a well-established method for understanding complex biological systems but has not been reported for *C. thermocellum* until recently due to methodological limitations. Disruption of CipA by random integration of insertion elements has been reported in *C. thermocellum* (16), and targeted gene inactivation mediated by group II introns has been used to evaluate family 9 cellulase function in the mesophile *C. phytofermentans* (11). Targeted gene deletion in *C. thermocellum* has recently been reported using positive and negative selection based on uracil auxotrophy (17) but has not been used previously to investigate the cellulosome system.

Here we report targeted deletion of *cel48S* and *cel48Y* in *C. thermocellum*, characterize the phenotype of the Cel48S mutant from both a microbial and enzymatic perspective, and draw inferences with respect to the role of Cel48S in the cellulosome.

Results

S Mutant Strain Construction. Plasmid pDGO-01 (*SI Appendix*, Fig. S1) was transformed into the parent strain (*SI Appendix*, Table S1) by electroporation using recently described methods (17), and, after overnight recovery, cells harboring the plasmid were selected by the addition of thiamphenicol (Tm) (Fig. 1*A*). Cells were subcultured into media containing both 5-fluoroorotic acid (FOA) and Tm to select for cells where integration of the chloramphenicol acetyl transferase (*cat)* gene had replaced the *cel48S* gene on the chromosome (Fig. 1*B*). This also selected for loss of the pDGO-01 plasmid. To confirm deletion of *cel48S*, clones resistant to Tm and FOA were screened at the *cel48S* locus using diagnostic PCR, showing a 4.6-kb amplicon for the *cel48S* region and a 3.6-kb amplicon for the *Δcel48S::PgapDH-cat* region (Fig. 1*A*). Additional PCR reactions showed the presence of the *cat* gene and the absence of two internal fragments of *cel48S* in the S mutant strain (*SI Appendix*, Figs. S2 and S3). The amplicon in the S mutant strain (*SI Appendix*, Table S1) was sequenced and the *cel48S* gene was found to have been replaced by the *cat* cassette from plasmid pDGO-01 (*SI Appendix*, Dataset S1).

Author contributions: D.G.O., S.A.T., R.J.G., J.L., N.C.C., D.A.H., R.L.H., A.M.G., and L.R.L. designed research; D.G.O., S.A.T., R.J.G., J.L., and G.D. performed research; D.G.O., S.A.T., R.J.G., J.L., N.C.C., D.A.H., R.L.H., and L.R.L. analyzed data; and D.G.O. and L.R.L. wrote the paper.

Conflict of interest statement: Several of the authors are employees or hold a consulting position with the Mascoma Corporation, which has a financial interest in the organism described here.

*This Direct Submission article had a prearranged editor.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. HQ157351–HQ157352).

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed: E-mail: lee.lynd@dartmouth.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1003584107/-/DCSupplemental.

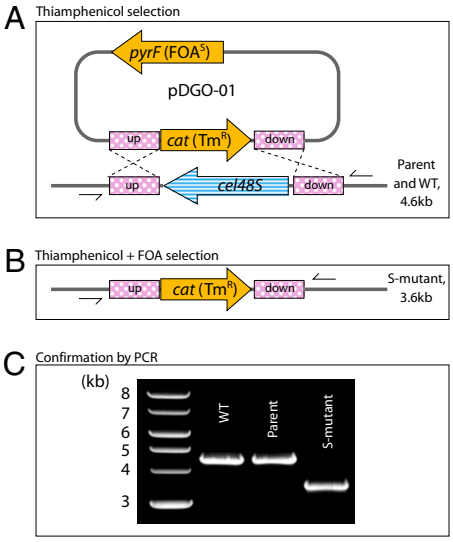

Fig. 1. Steps involved in making Cel48S deletion. (*A*) Plasmid pDGO-01 is transformed into the parent strain. (*B*) Addition of FOA selects for cells where *cat* has replaced *cel48S*, and the plasmid has been lost. (*C*) Diagnostic PCR shows a smaller band at the *cel48S* locus.

SY Mutant Strain Construction. Plasmid pJL2 was transformed into the S mutant strain by electroporation as above, but selection was performed using neomycin (Neo). Selection on media containing Neo and FOA resulted in a strain in which integration of the neomycin resistance (*neo*) gene had replaced the *cel48Y* gene on the chromosome. To confirm deletion of *cel48Y*, clones resistant to Neo and FOA were screened at the *cel48Y* locus using diagnostic PCR, which showed a 5.3-kb amplicon for the *cel48Y* region and a 3.9-kb amplicon for the Δ*cel48Y*::P*gapDH-neo* region (*SI Appendix*, Fig. S6). The amplicon in the SY mutant strain was sequenced and the *cel48Y* gene was found to have been replaced by the *neo* cassette from plasmid pJL2 (*SI Appendix*, Dataset S2).

Microbial Characterization. To investigate the role of Cel48S for growth on cellulose, the WT, parent, and S mutant strains (*SI Appendix*, Table S1) were grown in media with 10 g/L Avicel or cellobiose as growth substrate and the formation of microbial biomass and the consumption of substrate was determined. Cellobiose consumption was similar for all three strains in both rate and extent and was complete in 20 h. The rate of biomass production of cellobiose-grown cells was similar for all three strains as well, but there were slight differences in final biomass titer. The S mutant strain produced 9.9 ± 5.4% less biomass than the parent strain and 15.6 ± 7.1% less than the WT strain (Fig. 2 and Table 1). The similarity of these fermentations was expected because Cel48S plays no known role during growth on cellobiose.

The rate of pellet biomass formation was similar for both the WT and parent strains but 80% lower for the S mutant strain (Table 1). Final pellet biomass was 50% lower in the S mutant strain compared with the WT and parent strains (Fig. 2). The Avicel consumption rate was similar for the WT and parent strains but 60% lower in the S mutant; however, all of the strains eventually solubilized >97% of the Avicel initially present and thus retained true cellulolytic ability (Fig. 2). Growth on Avicel was also measured for the SY mutant strain (*SI Appendix*, Table S1) and was not found to be substantially different from the S mutant (*SI Appendix*, Fig. S7)

Enzymatic Characterization. Initial rates of Avicel solubilization by purified cellulosome preparations were measured over a range of loadings to investigate the implication of the loss of Cel48S on overall enzymatic activity (Fig. 3). Because the *C. thermocellum* cellulosome produces cellodextrin polymers of varying lengths, they were digested to glucose monomers with Novozym 188 beta-glucosidase (Sigma-Aldrich). The resulting monomers were measured and reported on a soluble glucose equivalent (SGE) basis (18).

The slope of the linear part of the graph (0.05–0.2 mg) shows that the cellulosome from the S mutant strain has a substantially lower specific activity (0.60 ± 0.04 U/mg protein^{-1}) than purified cellulosomes from either the WT (1.00 ± 0.07 U/mg protein^{-1}) or parent strain (0.95 ± 0.09 U/mg protein^{-1}). Extrapolating from the saturation region of the graph (0.4–0.8 mg), the saturation rate can be estimated to be 0.4 μmol/min for the parent and WT strains and 0.2 μmol/min for the S mutant strain (Fig. 3).

Proteomic Characterization. Affinity digestion (19) was used to purify cellulosomes collected at the end of Avicel fermentations: total protein concentration was 1.47 ± 0.06 mg/mL for the WT, 1.56 ± 0.09 mg/mL for the parent, and 1.22 ± 0.14 mg/mL for the S mutant. The WT and parent strains exhibited a reproducible banding pattern typical of the *C. thermocellum* cellulosome (Fig. 4) (13), with 14 major bands denoted S1–S14 (6). As expected, the S mutant strain was similar to the WT and parent strains except for the S8 band (~80 kDa), which was much fainter and of a slightly higher molecular weight (Fig. 4). The S8 band has been shown to correspond to the protein Cel48S (20). To determine the identity of the ~80-kDa band, it was excised from a gel and characterized by mass spectrometry (NextGen Sciences). The band from the WT and parent strains was found to be a mixture of Cel48S and Cel9Q at an approximately 2:1 ratio (*SI Appendix*, Table S2) The band from the S mutant strain was exclusively Cel9Q and no Cel48S was detected (*SI Appendix*, Table S2).

The purified cellulosome was further analyzed by tandem mass spectrometry to determine if the composition of the cellulosome had changed in response to the deletion of Cel48S. The most significant change detected was the disappearance of Cel48S in the S mutant strain. In addition to Cel48S, there were four proteins whose abundance had changed significantly ($P < 0.01$) between the parent and S mutant strains (*SI Appendix*, Table S3): Cthe_0452 (OlpC) (20), Cthe_2761 (annotated as GH9-CBM3c-Doc1) (15), Cthe_3079 (Orf2p), and Cthe_3132 (annotated as Unknown-Doc1) (15). All of these showed changes of ±twofold with the notable exception of Cthe_0452, which showed a decrease of approximately eightfold (Fig. 5).

Discussion

Deletion of Cel48S from *C. thermocellum* led to a decrease in the enzymatic hydrolysis rate, a decrease in microbial hydrolysis rate, and a decrease in biomass formation during growth on Avicel.

The similarity of enzyme saturation curves for the WT and parent strains suggests that the Δ*pyrF* mutation in the parent strain has no effect on cellulosome function, as expected. The S mutant strain, however, exhibited a reduction in both specific activity and saturation rate. A reduction in specific activity is indicative of impaired function and consistent with decreased synergy among components of the cellulosome in the absence of Cel48S (3).

The role of GH families in cellulose solubilization is a topic of much debate. Family 48 cellulases are a prominent component of many bacterial cellulase systems and, due to their ubiquity, are thought to play an important role in cellulose solubilization (21). On one hand, disruption of the single family 9 GH in *C. phytofermentans* eliminated its ability to grow on filter paper (11), suggesting that some GH families are essential. On the other hand, down-regulation of *cel48F* in *C. cellulolyticum* gave only a modest effect, with no reduction in growth rate and a 30% reduction in specific activity of the purified cellulosome (22). The deletion of Cel48S from *C. thermocellum* reduced growth rate and specific activity by about half,

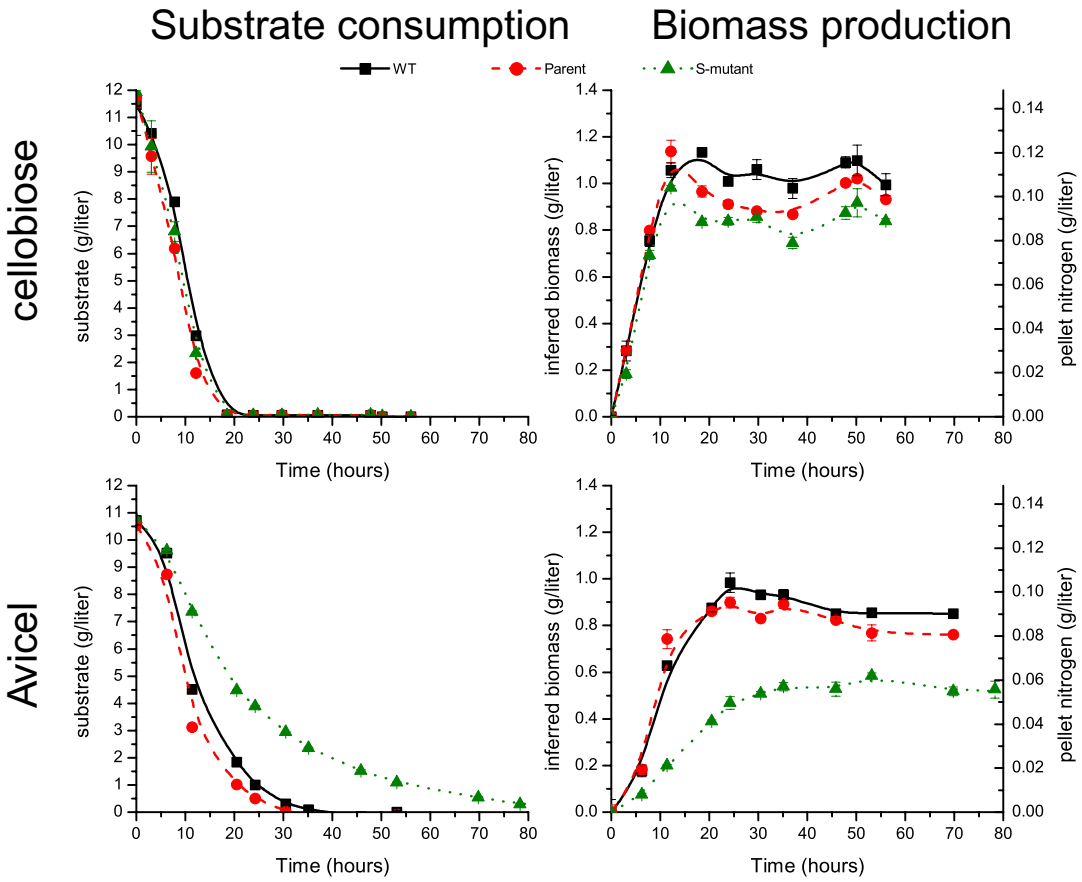

Fig. 2. Batch fermentations of WT, parent, and S mutant strains growing on either Avicel or cellobiose. (*Left*) Substrate consumption; (*Right*) biomass production. Biomass production was inferred based on pellet nitrogen measurements. Growth of all three strains was similar on cellobiose, whereas on Avicel, the S mutant strain consumed the Avicel more slowly and made less biomass. The data at each time point represent the averages of the results from duplicate measurements. Error bars represent SD.

which was more substantial than the *cel48F* phenotype observed in *C. cellulolyticum* but not the complete elimination of cellulolytic ability observed in *C. phytofermentans*. Cel48Y expression was either not detected or detected at extremely low levels in several recent studies of the *C. thermocellum* cellulosome (12, 13, 15) and was not detected in our samples (*SI Appendix*, Table S3). Nevertheless, we deleted it to show that it was not responsible for the residual cellulolytic activity observed in the *S* mutant strain and to further support our claim that family 48 GHs are not necessary for growth of *C. thermocellum* on crystalline cellulose. This raises the question: what is the role of family 48 GH enzymes in bacterial cellulase systems? Based on the enzymatic and microbial data, Cel48S appears to be a rate-limiting enzyme in cellulose solubilization, but, even in the absence of Cel48S, *C. thermocellum* produces a cellulosome with the ability to completely solubilize crystalline cellulose. Understanding the mechanism behind this residual activity is a promising direction for future work.

Mutant characterization was undertaken at a microbial as well as enzymatic level because there is strong technological interest in microbial conversion systems and many additional fundamental phenomena are operative when cellulose solubilization is

Table 1. Maximum rate of substrate consumption and biomass production in the WT, parent, and S mutant strains

	Maximum rate, g/L/h			
	Cellobiose		Avicel	
Strain	Substrate	Biomass	Substrate	Biomass
WT	1.1 ± 0.04	0.06 ± 0.008	0.9 ± 0.01	0.05 ± 0.003
Parent	1.1 ± 0.02	0.06 ± 0.006	1.0 ± 0.01	0.06 ± 0.005
S mutant	1.0 ± 0.09	0.06 ± 0.006	0.4 ± 0.01	0.01 ± 0.001

Biomass production was inferred from pellet nitrogen measurements. The data represent the averages of the results from duplicate measurements. Error represents one SD.

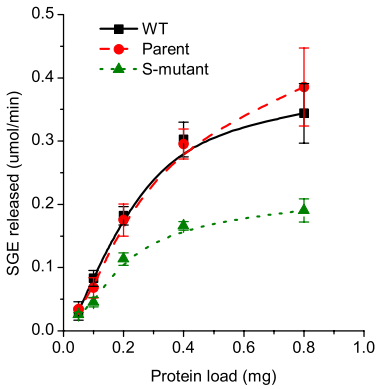

Fig. 3. Enzymatic activity of purified cellulosomes against 0.6 g/L Avicel. Activity is measured in SGE. The data represent the averages of the results from triplicate experiments. Error bars represent SD.

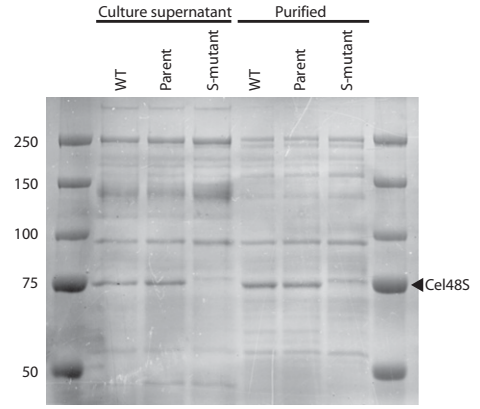

Fig. 4. SDS/PAGE of proteins from the WT, parent, and S mutant strains. The left three lanes are supernatant from 10 g/L Avicel fermentations. The right three lanes are culture supernatant purified by affinity digestion and diluted eightfold. The Cel48S band is indicated by the triangle. The left lane shows molecular mass markers (kDa).

mediated by microbial cultures as compared with enzyme preparations. Neither the *ΔpyrF* modification of the parent strain nor the Δ*cel48S::*P*gapDH-cat* modification in the S mutant strain exhibited a deleterious effect on cellobiose growth (Fig. 2). Avicel consumption, on the other hand, was much slower for the S mutant strain than for either the WT or parent strains with twofold more time required to achieve complete cellulose solubilization under comparable conditions. Slower Avicel consumption by the mutant culture is consistent with the lower activity of the mutant cellulosome observed during in vitro experiments. However, the reduced pellet biomass observed in the S mutant may also be a factor. The relative importance of reduced cellulosome effectiveness and reduced cellulosome production in determining the slower utilization of Avicel by the mutant is unclear at this time. The reduction in cellulosomal protein was partially compensated by an increase in supernatant protein (*SI Appendix*, Figs. S4 and S5). The identity of these proteins is currently unknown but might point to a regulatory effect. The absence of Cel48S did not significantly impact the abundance of any other major component of the cellulosome. Densitometry analysis of the denaturing gel (*SI Appendix*, Fig. S4) shows very little change for bands other than Cel48S. Tandem mass spectrometry revealed changes in four cellulosomal proteins; two of these were noncatalytic (OlpC and Orf2p) and two were minor catalytic components (Cthe_2761 and Cthe_3132, each <1% of cellulosomal protein). Both of the catalytic components that showed significant changes in abundance between the S mutant and parent strains also showed significant changes in abundance between the WT and parent strains, suggesting that these changes do not have a large effect on cellulosome function. The two other proteins that showed significant changes in abundance, OlpC and Orf2p, contain cohesin domains and are therefore thought to be structural components of the cellulosome. The twofold increase in Orf2p abundance and eightfold decrease in OlpC abundance might reflect the action of regulatory mechanisms. *C. thermocellum* has been shown to regulate protein expression by substrate sensing (23) and changes in growth rate (24, 25), both of which may have occurred in the S mutant strain.

The development of gene knockout capability for *C. thermocellum* has provided an opportunity to improve understanding of microbial cellulose utilization and the action of the cellulosome. The creation and characterization of a *cel48S* mutant strain and its cellulosome underscore the value of this approach.

Materials and Methods

Strains and Media. *C. thermocellum* strain DSM 1313 (WT) was grown in modified DSM 122 broth (26) with the addition of 50 mM 3-(N-morpholino) propanesulfonic acid (MOPS) sodium salt and 3 g/L trisodium citrate ($Na_3C_6H_5O_7{\cdot}2\,H_2O$). The parent strain (*SI Appendix*, Table S1) is a deletion of *pyrF*, which exhibits auxotrophy for uracil and needs to be supplemented with uracil at 40 μg/mL (17). The S mutant strain was derived from the parent strain by replacing the *cel48S* gene with a thiamphenicol antibiotic resistance marker. Unless otherwise noted, cells were grown at 55 °C with gentle stirring using 5 g/L cellobiose as the primary carbon source. All manipulations were carried out inside an anaerobic chamber (Coy Laboratory Products Inc.) with an atmosphere of 85% nitrogen, 10% carbon dioxide, 5% hydrogen, and <5 parts per million oxygen.

Molecular Biological Methods. Plasmids were constructed using yeast-mediated ligation (27) or standard cloning techniques (28). Plasmid pMQ87 was a gift from Robert Shanks (University of Pittsburgh Eye Center, Pittsburgh, PA) (27). Plasmids were maintained in *Escherichia coli* TOP10 cells (Invitrogen Corporation) and prepared using QIAGEN Plasmid Mini kit (QIAGEN Inc.). Sequences of chromosomal DNA were obtained by PCR using genomic DNA from *C. thermocellum* strain DSM 1313 as the template and primers designed using the *C. thermocellum* ATCC 27405 genome published by the Joint Genome Institute (http://www.jgi.doe.gov/). Plasmid pDGO-01 (*SI Appendix*, Fig. S1 and Table S1) was based on the *C. thermocellum, E. coli, Saccharomyces cerevisiae* shuttle plasmid pMU749 (17). Approximately1-kb regions of homology flanking the *cel48S* gene on the *C. thermocellum* chromosome were added upstream and downstream of a thiamphenicol resistance cassette. The native *pyrF* gene, under control of the *C. thermocellum* cellobiose phosphorylase (*cbp*) promoter, was cloned outside the homologous flanks (*SI Appendix*, Fig. S1). Plasmid pJL2 was constructed in a similar manner, with ~1-kb regions of homology flanking the *cel48Y* gene on the *C. thermocellum* chromosome and neomycin resistance provided by the *kan* marker from plasmid pIKM1 (29). PCR was performed using either Taq or Phusion DNA polymerase (New England Biolabs Inc.) according to the directions provided by the manufacturer. When using whole cells as the PCR

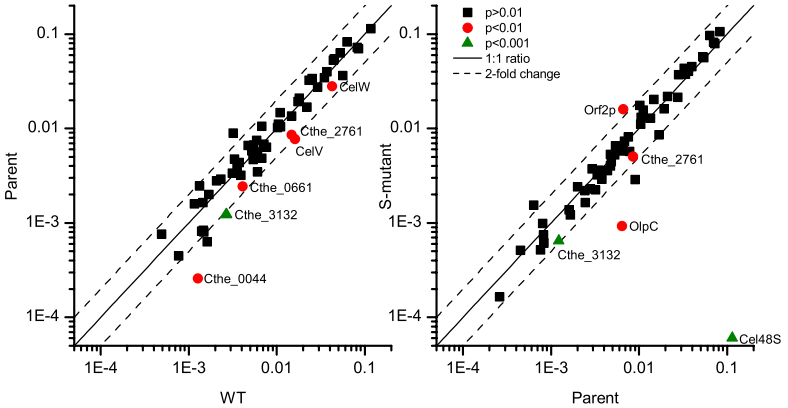

Fig. 5. Protein abundance determined by APEX. Dashed lines represent a twofold change. The limit of detection for Cel48S was 5×10^{-5} unscaled APEX units.

template, a 10-min heating step was included at the beginning of the thermocycling protocol to lyse the cells. When using Taq DNA polymerase, the lysing temperature was 95 °C. When using Phusion DNA polymerase, the lysing temperature was 98 °C. DNA sequencing was performed using standard techniques with an ABI Model 3100 genetic analyzer (Applied Biosystems).

Transformation. Transformation was performed according to protocol. Cells were prepared for transformation by inoculation from a freezer stock into uracil-supplemented media with 5 g/L cellobiose as the primary carbon source (17). The culture was incubated at 55 °C until the optical density at 600 nm (OD600) reached 0.4–0.8 absorbance units. Cells were washed in reverse-osmosis purified (18 MΩ) water that had been autoclaved to remove oxygen. Twenty microliters of cell suspension were added to each cuvettete along with 1–8 μL of DNA (10–2,000 ng) eluted in water. Standard 0.1-cm gap electroporation cuvettetes were used. A series of 60 square pulses, each of 30-μs duration, were applied to the sample. The period of the pulses was 300 μs and the amplitude was 1.9 kV, resulting in an applied field strength of 19 kV/cm. After pulsing, cells were allowed to recover at 51 °C in 3–5 mL uracil-supplemented media overnight (15–18 h) before being subjected to selective pressure. Selection for the *cat* marker was performed by the addition of Tm at a final concentration of 48 μg/mL to the culture medium. Selection for *neo* was performed by the addition of Neo at a final concentration of 250 μg/mL to the culture medium. Selection against the *pyrF* gene was performed by the addition of FOA at a final concentration of 500 μg/mL. When used in conjunction, the FOA concentration was 500 μg/mL and the thiamphenicol concentration was 6 μg/mL.

Microbial Growth and Hydrolysis Analysis. Microbial growth and hydrolysis analysis were determined by batch fermentation in uracil-supplemented media with 10 g/L Avicel PH-105 microcrystalline cellulose (Sigma-Aldrich) as the primary carbon source. Fermentations were performed in a 1-L volume at 55 °C in a Sartorius Q+ fermentation system with pH control provided by the addition of 5N potassium hydroxide. Fermentations were determined to be complete when no further base addition occurred. The culture was stirred at 100 rpm, which was sufficient to keep Avicel particles suspended. Thirty-milliliter aliquots were drawn at intervals throughout the fermentation to determine substrate consumption and biomass formation. Cellobiose consumption was measured by HPLC. Pellet nitrogen was measured with a Shimadzu TOC-V CPH elemental analyzer with TNM-1 and ASI-V modules (Shimadzu Corp.) on 1-mL aliquots that had been washed twice with water and then centrifuged. Dry weight was measured by washing an 8-mL aliquot twice with water, followed by centrifugation. The washed sample was then dried at 60 °C to constant weight. For cellobiose-grown cultures, dry weight was composed exclusively of biomass. For Avicel-grown cultures, the dry weight represents the sum of Avicel and biomass. Biomass was calculated from pellet nitrogen data assuming that nitrogen makes up a constant 10.6% of cell mass (30). Residual Avicel was determined by subtracting biomass from dry weight.

Enzymatic Analysis. After fermentation, cellulosomes were purified from 200 mL of broth, using the affinity digestion protocol from Morgenstern et al. (18) Total protein was measured with Bio-Rad Bradford protein assay with BSA (BSA) as a standard. Initial hydrolysis rate measurements were performed at 55 °C in a 50-mL volume in a 150-mL serum bottle with constant shaking following the protocol of Bernardez et al. (17) The activity buffer contained 50 mM sodium acetate (pH 5.0), 5 mM cysteine-HCl, 12 mM $CaCl_2$, 40 μg/mL tetracycline (to prevent microbial growth), and 0.02% vol/vol Novozym 188 [to convert cellobiose to glucose for later analysis and to prevent product inhibition from affecting initial rate measurements (Sigma-Aldrich)]. This concentration of Novozym 188 was >10-fold in excess of what would have been necessary to convert all of the cellobiose generated by the cellulosome into glucose. One-microliter aliquots were taken hourly during the first 3 h for SGE analysis. Glucose was measured using the Hexokinase Glucose Assay kit (Sigma-Aldrich). The slope of these measurements was used to determine the initial hydrolysis rate (μmol/min) for a range of different cellulosome loadings from 0.05 to 0.8 mg. One unit (U) of enzyme activity releases 1 μmol SGE/min. Specific activity is measured in U/mg protein.

Denaturing Gel Electrophoresis. Ten-microliter aliquots of the fermentation supernatant and 1.25-μL aliquots of purified cellulosome were run on a 7.5% Tris-HCl SDS/PAGE gel (Bio-Rad) to visualize the individual subunits (Fig. 4). Bands from the three purified lanes were excised, and the major components were analyzed by mass spectrometry (NextGen Science) (*SI Appendix*, Table S2).

Protein Measurements. Samples were prepared for protein measurement following the method of Zhang and Lynd (31). Supernatant protein was measured with the Bradford assay (Thermo Scientific), and pellet protein was measured with the BCA assay (Thermo Scientific). BSA was used as the standard.

Protein Data Collection for Tandem Mass Spectrometry. Purified cellulosomes were processed for 2D LC-MS/MS analysis as follows. Proteins were precipitated overnight at −20 °C by adding trichloroacetic acid (TCA) to a final concentration of 20%. The resulting protein pellets were washed with ice-cold acetone, resolublized in denaturing buffer (8 M urea, 100 mM Tris, pH 8.0), and reduced with 20 mM DTT. Samples were diluted to 4 M urea with 100 mM Tris, 10 mM $CaCl_2$, pH 8.0, and digested via two additions of modified trypsin (Promega) at a 1:75 enzyme to protein ratio (wt/wt). Resulting peptides were protonated with 0.1% formic acid, spin filtered (Ultrafree-MC; Millipore), and 25 μg was loaded onto a MudPIT (32, 33) back column packed with strong cation exchange (SCX, Luna; Phenomenex) and C18 reversed phase (RP, Aqua; Phenomenex) resins, as previously described (34), and separated by charge (salt pulses of 0, 10, 25, and 100% of 500 mM ammonium acetate) and hydrophobicity (100-min aqueous to organic gradient) using an HPLC pump (u3000; Dionex) coupled to an LTQ XL mass spectrometer (Thermo Scientific). Eluting peptides were measured, isolated, and fragmented by the LTQ XL operating in data-dependent mode. Each sample was analyzed in triplicate.

The resulting tandem mass spectra were searched with SEQUEST (35) against the *C. thermocellum* ATCC 27405 proteome concatenated with reversed FASTA protein entries to assess false-discovery rates (FDR), common contaminants, and the *cat* gene to assess the fidelity of the Cel48S deletion and determine the composition of the resulting cellulosome. As urea was used as the denaturant, searches were performed with the inclusion of carbamylation (+43 Da) as a dynamic modification potentially occurring on lysines, arginines, and peptide N-termini.

Protein Data Analysis. The DTASelect data were converted into pepXML format with the freely available software program Out2XML (Institute for Systems Biology). Protein abundance was analyzed with the APEX Quantitative Proteomics Tool (36) (J. Craig Venter Institute) using a 1% false-detection threshold. After quantification, proteins that were not identified in all nine samples (three replicates per strain, three strains) were eliminated. Only cellulosomal proteins (12, 15) were included in abundance normalization.

APEX is a technique for quantification of protein data from mass spectrometry experiments that uses machine learning to correct for sequence-specific detection bias (37). The 40 most abundant proteins (*SI Appendix*, Table S3) were selected for training the APEX classifier. Pair-wise comparisons were made between data sets (WT, parent, and S mutant) and proteins whose abundance had changed significantly were identified by *t* test ($P < 0.01$) of log-transformed data. Log-transformed data were found to be normal by the Shapiro–Wilk test at the 0.05 level.

ACKNOWLEDGMENTS. We thank Dr. Erin Wiswall for assistance with protein purification and Meryl B. Olson for her assistance with statistical calculations. This research was supported by Chemical Sciences, Geosciences and Biosciences Division, Office of Basic Energy Sciences, Department of Energy Grant DE-FG02-02ER1535 and grants from the Mascoma Corporation and the BioEnergy Science Center, Oak Ridge National Laboratory, a Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the Department of Energy Office of Science.

1. Lynd LR, et al. (2008) How biotech can transform biofuels. *Nat Biotechnol* 26:169–172.
2. Lynd LR, van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of cellulosic biomass: An update. *Curr Opin Biotechnol* 16:577–583.
3. Lynd LR, Weimer PJ, van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: Fundamentals and biotechnology. *Microbiol Mol Biol Rev* 66:506–577.
4. Johnson EA, Sakajoh M, Halliwell G, Madia A, Demain AL (1982) Saccharification of complex cellulosic substrates by the cellulase system from *Clostridium thermocellum*. *Appl Environ Microbiol* 43:1125–1132.
5. Demain AL, Newcomb M, Wu JHD (2005) Cellulase, clostridia, and ethanol. *Microbiol Mol Biol Rev* 69:124–154.
6. Lamed R, Setter E, Bayer EA (1983) Characterization of a cellulose-binding, cellulase-containing complex in *Clostridium thermocellum*. *J Bacteriol* 156:828–836.
7. 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.

8. Reverbel-Leroy C, Pages S, Belaich A, Belaich JP, 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.
9. Liu CC, Doi RH (1998) Properties of exgS, a gene for a major subunit of the Clostridium cellulovorans cellulosome. *Gene* 211:39–47.
10. Kakiuchi M, et al. (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.
11. Tolonen AC, Chilaka AC, Church GM (2009) Targeted gene inactivation in *Clostridium phytofermentans* shows that cellulose degradation requires the family 9 hydrolase Cphy3367. *Mol Microbiol* 74:1300–1313.
12. Raman B, et al. (2009) Impact of pretreated Switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: A quantitative proteomic analysis. *PLoS ONE* 4:e5271.
13. Gold ND, Martin VJJ (2007) Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J Bacteriol* 189:6787–6795.
14. Wu JHD, Orme-Johnson WH, Demain AL (1988) Two components of an extracellular protein aggregate of *Clostridium thermocellum* together degrade crystalline cellulose. *Biochemistry* 27:1703–1709.
15. Zverlov VV, Kellermann J, Schwarz WH (2005) Functional subgenomics of *Clostridium thermocellum* cellulosomal genes: Identification of the major catalytic components in the extracellular complex and detection of three new enzymes. *Proteomics* 5: 3646–3653.
16. Zverlov VV, Klupp M, Krauss J, Schwarz WH (2008) 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.
17. Tripathi SA, et al. (2010) Development of pyrF-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a pta mutant. *Appl Environ Microbiol*, AEM.01484-01410.
18. Bernardez TD, Lyford KA, Lynd LR (1994) Kinetics of the extracellular cellulases of *Clostridium thermocellum* acting on pretreated mixed hardwood and Avicel. *Appl Microbiol Biotechnol* 41:620–625.
19. Morgenstern EM, Bayer EA, Lamed R (1992) Affinity digestion for the near-total recovery of purified cellulosome from *Clostridium thermocellum*. *Enzyme Microb Technol* 14:289–292.
20. Morag E, Bayer EA, Hazlewood GP, Gilbert HJ, Lamed R (1993) Cellulase S_s (CelS) is synonymous with the major cellobiohydrolase (subunit S8) from the cellulosome of *Clostridium thermocellum*. *Appl Biochem Biotechnol* 43:147–151.
21. Pinheiro BA, et al. (2009) Functional insights into the role of novel type I cohesin and dockerin domains from *Clostridium thermocellum*. *Biochem J* 424:375–384.
22. Izquierdo JA, Sizova MV, Lynd LR (2010) Diversity of bacteria and glycosyl hydrolase family 48 genes in cellulolytic consortia enriched from thermophilic biocompost. *Appl Environ Microbiol* 76:3545–3553.
23. Perret S, Maamar H, Bélaich JP, Tardif C (2004) Use of antisense RNA to modify the composition of cellulosomes produced by Clostridium cellulolyticum. *Mol Microbiol* 51:599–607.
24. Newcomb M, Chen CY, Wu JHD (2007) Induction of the *celC* operon of *Clostridium thermocellum* by laminaribiose. *Proc Natl Acad Sci USA* 104:3747–3752.
25. Dror TW, et al. (2003) Regulation of the cellulosomal *CelS* (*cel48A*) gene of *Clostridium thermocellum* is growth rate dependent. *J Bacteriol* 185:3042–3048.
26. Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2003) Regulation of expression of scaffoldin-related genes in *Clostridium thermocellum*. *J Bacteriol* 185:5109–5116.
27. Tyurin MV, Desai SG, Lynd LR (2004) Electrotransformation of *Clostridium thermocellum*. *Appl Environ Microbiol* 70:883–890.
28. Shanks RMQ, Caiazza NC, Hinsa SM, Toutain CM, O'Toole GA (2006) *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative bacteria. *Appl Environ Microbiol* 72:5027–5036.
29. Ausubel FA, et al. (1990) *Current Protocols in Molecular Biology* (Wiley Interscience, New York).
30. Mai V, Lorenz WW, Wiegel J (1997) Transformation of *Thermoanaerobacterium sp.* strain JW/SL-YS485 with plasmid pIKM1 conferring kanamycin resistance. *FEMS Microbiol Lett* 148:163–167.
31. Lynd LR, Grethlein HE, Wolkin RH (1989) Fermentation of cellulosic substrates in batch and continuous culture by *Clostridium thermocellum*. *Appl Environ Microbiol* 55: 3131–3139.
32. Zhang YH, Lynd LR (2003) Quantification of cell and cellulase mass concentrations during anaerobic cellulose fermentation: Development of an enzyme-linked immunosorbent assay-based method with application to Clostridium thermocellum batch cultures. *Anal Chem* 75:219–227.
33. Washburn MP, Wolters D, Yates JR, 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19: 242–247.
34. Hayes McDonald W, Ryoma O, Miyamoto DT, Mitchison TJ, Yates JR, III (2002) Comparison of three directly coupled HPLC MS/MS strategies for identification of proteins from complex mixtures: Single-dimension LC-MS/MS, 2-phase MudPIT, and 3-phase MudPIT. *Int J Mass Spectrom* 219:245–251.
35. Giannone RJ, et al. (2007) Dual-tagging system for the affinity purification of mammalian protein complexes. *Biotechniques* 43:296–302, 298, 300 passim.
36. Eng JK, McCormack AL, Yates JR (1994) An approach to correlate tandem mass-spectral data of peptides with amino-acid-sequences in a protein database. *J Am Soc Mass Spectrom* 5:976–989.
37. Braisted JC, et al. (2008) The APEX Quantitative Proteomics Tool: Generating protein quantitation estimates from LC-MS/MS proteomics results. *BMC Bioinformatics* 9:529.
38. Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. *Nat Biotechnol* 25:117–124.