Molecular Microbiology (2014) 91(4), 694–705 ■
doi:10.1111/mmi.12488
First published online 7 January 2014

Dockerin-containing protease inhibitor protects key cellulosomal cellulases from proteolysis in *Clostridium cellulolyticum*

Tao Xu,[1,2] Yongchao Li,[1,2] Zhili He[1,3] and Jizhong Zhou[1,3,4,5]*

[1]*Institute for Environmental Genomics,* [2]*Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73071, USA.*
[3]*Virtual Institute for Microbial Stress and Survival, Berkeley, USA.*
[4]*Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.*
[5]*Department of Environmental Science and Engineering, Tsinghua University, Beijing 100084, China.*

Summary

Cellulosomes are key for lignocellulosic biomass degradation in cellulolytic *Clostridia*. Better understanding of the mechanism of cellulosome regulation would allow us to improve lignocellulose hydrolysis. It is hypothesized that cellulosomal protease inhibitors would regulate cellulosome architecture and then lignocellulose hydrolysis. Here, a dockerin-containing protease inhibitor gene (*dpi*) in *Clostridium cellulolyticum* H10 was characterized by mutagenesis and physiological analyses. The *dpi* mutant had a decreased cell yield on glucose, cellulose and xylan, lower cellulose utilization efficiency, and a 70% and 52% decrease of the key cellulosomal components, Cel48F and Cel9E respectively. The decreased cellulolysis is caused by the proteolysis of major cellulosomal components, such as Cel48F and Cel9E. Disruption of *cel9E* severely impaired cell growth on cellulose while loss of *cel48F* completely abolished cellulolytic activity. These observations are due to the combinational results of gene inactivation and polar effects caused by intron insertion. Purified recombinant Dpi showed inhibitory activity against cysteine protease. Taken together, Dpi protects key cellulosomal cellulases from proteolysis in H10. This study identified the physiological importance of cellulosome-localized protease inhibitors in *Clostridia*.

Accepted 7 December, 2013. *For correspondence. E-mail jzhou@ou.edu; Tel. (+1) 405 325 6073; Fax (+1) 405 325 7552.

Introduction

Consolidated bioprocessing (CBP) of lignocellulosic biomass integrates the microbial activities of hydrolase production, saccharification and fermentation into a single step, and is regarded as a promising approach for production of low-cost biofuels (Lynd *et al.*, 2002). Cellulolytic *Clostridia* (e.g. *Clostridium thermocellum* and *C. cellulolyticum*) as CBP-enabling candidates have been sequenced (Hemme *et al.*, 2010) and are being engineered with higher efficiency in cellulose hydrolysis and biofuel synthesis (Guedon *et al.*, 2002; Argyros *et al.*, 2011; Brown *et al.*, 2011). Like some cellulose-degrading fungi (Dashtban *et al.*, 2009), these bacteria secrete diverse lignocellulose-degrading enzymes to synergistically decompose lignocellulosic biomass. Some of these enzymes are assembled onto cell surface-attached scaffoldin proteins by dockerin-cohesin interactions, generating multi-enzyme complexes called cellulosomes (Bayer *et al.*, 2004; Doi and Kosugi, 2004; Fontes and Gilbert, 2010). Biochemical studies on individual glycoside hydrolases have been widely conducted (Cantarel *et al.*, 2009) with the goal of boosting their industrial applications (Kuhad *et al.*, 2011). To date, however, only limited reports on the *in vivo* roles of glycoside hydrolases in cellulolysis are available (Perret *et al.*, 2004; Tolonen *et al.*, 2009; Olson *et al.*, 2010). Proper cellulosome assembly is critical to efficiently degrade cellulose when compared with free hydrolases (Schwarz, 2001; Maamar *et al.*, 2004). To accomplish CBP in *Clostridia*, bridging several knowledge gaps (e.g. physiological functions of cellulosomal components, regulatory mechanisms of cellulosome maintenance and gene expression) is necessary.

Cellulosomal heterogeneity is reflected in the varied abundance of components in mature cellulosomes when grown on different carbon sources (Han *et al.*, 2005; Gold and Martin, 2007; Blouzard *et al.*, 2010). The synergistic catalysis of glycoside hydrolases is important for efficient cellulolysis (Lynd *et al.*, 2002). For example, of 62 predicted dockerin-containing proteins in the genome of *C. cellulolyticum*, 50 were identified in isolated cellulosomes, 36 when grown on cellulose, 30 on xylan and 48 on hatched wheat straw (Blouzard *et al.*, 2010). The 26 kb *cip-cel* gene cluster in *C. cellulolyticum* containing 12

genes (*cipC, cel48F*, *cel8C*, *cel9G*, *cel9E*, *orfX*, *cel9H*, *cel9J*, *man5K*, *cel9M*, *rgl11Y* and *cel5N*) produces two large transcripts, a 14 kb mRNA carrying the first five coding sequences and a less abundant 12 kb mRNA with the coding sequences of the genes located in the 3′ part of the cluster (Maamar *et al*., 2006). CipC, Cel48F and Cel9E are three major cellulosomal components in *C. cellulolyticum* (Maamar *et al*., 2004; Perret *et al*., 2004). Previous studies showed *cipC* disruption and *cel48F* repression severely impaired cellulolysis (Maamar *et al*., 2004; Perret *et al*., 2004). However, how microorganisms adapt and maintain their cellulosomes under different environmental conditions remains a mystery.

In addition to dockerin-containing glycoside hydrolases, other enzymes (e.g. esterases, polysaccharide lyases, chitinase and peptidases) have been predicted and/or found to be on cellulosomes (Gold and Martin, 2007; Blouzard *et al*., 2010). Kang *et al*. cloned and studied three serine protease inhibitors, Serpin1–3, from *C. thermocellum* (Kang *et al*., 2006). Serpin1 was able to interact with *cipA* cohesion and inhibit subtilisin activity. Several cysteine peptidase inhibitors that are likely cellulosomal components in *Clostridium cellulovorans* also exhibited inhibitory activities against representative plant proteases, papain and ficin (Meguro *et al*., 2011). Proteomics studies on isolated *C. cellulolyticum* cellulosomes also identified a Chagasin_I42 component that might be a cysteine protease inhibitor (Blouzard *et al*., 2010). Bacterial proteases are involved in several biological processes including protein turnover, sporulation and conidial discharge, germination, enzyme modification, nutrition and regulation of gene expression (Rao *et al*., 1998). Considering the localization of cellulosomal protease inhibitors, it was speculated that they might be responsible for self-protection to avoid proteolysis of exogenous proteases (Meguro *et al*., 2011), or for cellulosome remodelling (Schwarz and Zverlov, 2006). So far, the physiological importance of these inhibitors has not been investigated.

Clostridium cellulolyticum as a non-ruminal mesophilic cellulolytic model is relatively susceptible to genetic manipulation (Petitdemange *et al*., 1984; Desvaux, 2005). A dockerin-containing protease inhibitor gene (*dpi*) (*Ccel_1809*) from *C. cellulolyticum* H10 was chosen to determine the *in vivo* functions of this kind of cellulosome-localized protease inhibitor. The protein encoded by the *dpi* gene has been identified in active cellulosomes (Blouzard *et al*., 2010). In this study, we hypothesized that the cellulosomal protease inhibitor Dpi would be enzymatically functional and affect insoluble carbon utilization by regulating cellulosomal components. To test these hypotheses, a *dpi* mutant was characterized at the phenotypic, physiological and protein levels. We discovered that Dpi was able to effectively block cysteine protease inhibitor activity, protect key *C. cellulolyticum* cellulosomal cellulases, and allow cells to maintain high-efficiency cellulolysis. Three additional mutants, a *trans*-complementation strain (*dpi/over*) and *cel48F* and *cel9E* mutants were constructed to further identify the physiological importance of *dpi, cel48F* and *cel9E* genes in degrading cellulose. This study provides new insights into our understanding of cellulosomal protease inhibitor-mediated protection of cellulosomal components from proteolysis in *C. cellulolyticum*.

Results

dpi mutant construction and phenotypic analyses

To examine whether the cellulosome-localized protease inhibitor Dpi plays an important role in carbon utilization, a *dpi* mutant was constructed using a mobile group II intron-based gene inactivation system (Supporting information Results; Fig. S1) (Heap *et al*., 2010; Li *et al*., 2012). Growth of the mutant was examined on both soluble (glucose and cellobiose) and insoluble substrates (cellulose and xylan). With 10 g l^{-1} cellobiose, there was no significant difference observed between the *dpi* mutant and WT in terms of growth rate and maximal biomass (Fig. 1A). However, with 10 g l^{-1} glucose, the mutant showed a 22% decrease in maximal cell density compared to WT although no difference was observed in the growth rate during exponential phase (Fig. 1B). With 10 g l^{-1} cellulose, the mutant presented a slower growth rate and its maximal biomass was 52% of WT (Fig. 1C). Similar results were observed on xylan, which showed a 23.4% decrease in maximal *dpi* mutant biomass (Fig. 1D). Therefore, the inactivation of the *dpi* gene affected cell growth on glucose, cellulose and xylan but not cellobiose.

The cellulose degradation efficiency of the *dpi* mutant and WT was also examined. For all time points tested, the mutant left higher amounts of cellulose residue in the fermentation broth (Fig. 1E). After entering into stationary phase, the residual cellulose in the mutant culture was 46.7 ± 10.2%, which was obviously higher than that in WT. To visualize differences in cellulose consumption between the WT and mutants, a hydrolysis test on cellulose-containing top-agar plates was performed (Maamar *et al*., 2004). In this test, cellulose degradation results in a transparent halo surrounding colonies, with a larger halo indicating higher amounts of cellulose degradation. The WT developed a large halo while the mutant developed a smaller halo (Fig. 1F). *Trans*-complementation of the *dpi* mutant (*dpi/over* strain) restored the cellulolytic phenotype to produce a halo similar to WT (Fig. 1F). These results confirmed that inactivation of *dpi* negatively affected cellulose utilization in *C. cellulolyticum*.

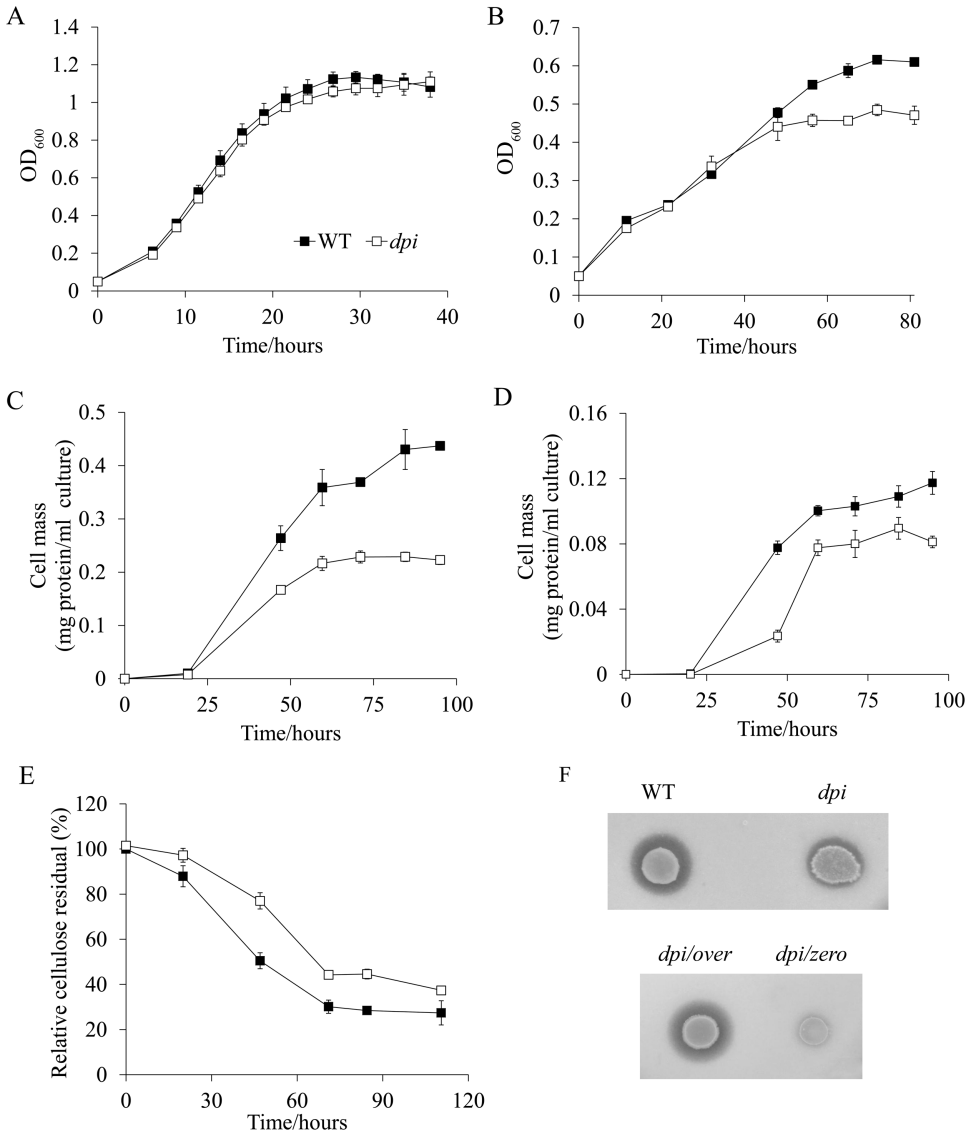

Fig. 1. Growth profiling of WT, *dpi* mutant, *dpi/over* and *dpi/zero* strains. Cell densities of WT and *dpi* mutant on 10 g l^{-1} cellobiose (A) and 10 g l^{-1} glucose (B) were estimated by monitoring OD_{600}. Cell mass obtained on 10 g l^{-1} Avicel cellulose (C) and 10 g l^{-1} xylan (D) were determined by total protein quantification. Cellulose residual percentage (E) was calculated by dividing the cellulose residual amount by the initial cellulose input. The means and standard deviations were calculated from three independent measurements. Avicel degradation tests (F) of WT, *dpi* mutant, *dpi/over* and *dpi/zero* strains were performed on cellulose-containing top-agar plates.

Molecular analysis of the mutant cellulolytic system

To investigate how the mutant reduced cellulose utilization, we isolated cellulosome fractions (Fc) and free extracellular fractions (Ff) from cellulose-grown cultures in the mid-log phase and then compared protein component patterns by SDS-PAGE. There were three prominent components in the WT Fc (Fig. 2A), which was consistent with previous reports that CipC, Cel48F and Cel9E were the three most abundant cellulosomal components (Maamar *et al.*, 2004; Perret *et al.*, 2004). The most significant difference observed in the mutant was the reduced abundance of two major bands in Fc (labelled B1 and B2). There were two additional minor bands also showing decreased density (labelled B3 and B4) in the mutant. B1–B4 bands were verified to be Cel48F, Cel9E, Cel9J and Cel9M, respectively, by mass spectrometry analysis (Table S1). Interestingly, with equal Fc loading, the density of CipC was not altered due to *dpi* disruption. Analysis of Ff also showed that some bands, but not all, were obviously altered in abundance between the WT and mutant. Thus, disruption of *dpi* gene significantly altered key cellulosomal components on the CipC scaffoldin.

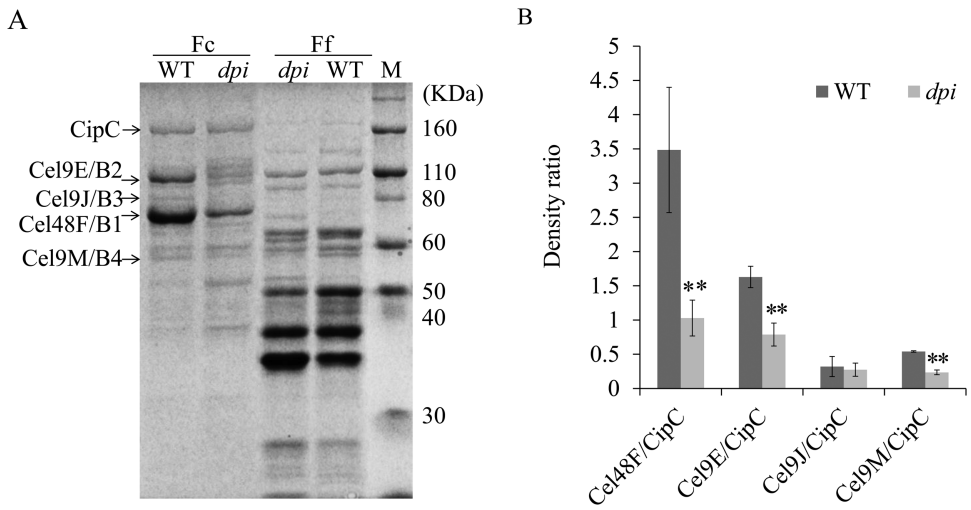

Fig. 2. Composition of the cellulolytic system of the *dpi* mutant and WT.
A. SDS-PAGE analysis of cellulosome fraction (Fc) and free extracellular protein fraction (Ff) isolated from 10 g l^{-1} Avicel cellulose-grown WT and *dpi* mutant at the mid-logarithmic phase. Bands labelled with enzyme names on the left were identified by mass spectrometry.
B. Densitometry analysis of several enzymes in Fc fractions. Ratios of Cel48F/CipC, Cel9E/CipC, Cel9J/CipC and Cel9M/CipC were calculated based on staining intensity for each protein. The means and standard deviations were calculated from gels of three biological replicates. The asterisks denote significant difference between WT and *dpi* mutant (**$P < 0.01$, by Student's *t* test).

CipC as a structural protein has eight cohesion domains for assembly of dockerin-containing enzymes (Desvaux, 2005). The relative abundances of the above four enzymes on CipC were quantified by densitometry analysis (Fig. 2B). The ratio of Cel48F/CipC was 3.48 in WT but dropped to 1.03 in the mutant. Similarly, the *dpi* mutation caused the Cel9E/CipC ratio to drop from 1.63 to 0.79, and the Cel9M/CipC ratio from 0.54 to 0.24. Correspondingly, the ratios of Cel48F/CipC, Cel9E/CipC and Cel9M/CipC decreased by 70%, 52% and 56.5% in the mutant ($P < 0.01$) respectively. However, the Cel9J/CipC ratio did not show a statistically significant change between WT and the mutant ($P > 0.05$). Therefore, the reduced abundance of several cellulosomal components on scaffoldin may further explain the lowered cellulolysis.

Quantification of cipC, cel48F and cel9E transcripts, and cellulosome productivity

The altered abundances of two prominent cellulosomal components (Cel48F and Cel9E) between the WT and mutant could be caused by two possibilities: selective proteolysis and differential transcript levels. *cipC*, *cel48F* and *cel9E* are all located in the *cip-cel* gene cluster and co-transcribed (Maamar *et al.*, 2006), so changes in the level of transcription of these genes would occur simultaneously. qPCR analysis revealed that the mutant had similar amounts of *cipC*, *cel48F* and *cel9E* transcripts during exponential growth, indicating that the difference in component abundance (Fig. 2A) is independent of transcription level ($P < 0.05$) (Fig. 3A). However, the mutant transcription level was around 60% of that observed in the WT. A lower expression level of the *cip-cel* gene cluster would reduce CipC availability and lead to less cellulosomal assembly on the cell surface. To test this, cellulosome productivity, which equals the ratio of isolated cellulosome to total cellular biomass, was determined in cellulose-grown cells at the mid-log phase. The WT and *dpi* mutant both presented similar productivity (Fig. 3B), 0.14 ± 0.01 mg cellulosome complex per mg cellular biomass. Thus, *dpi* inactivation did not affect cellulosome productivity, which further supports the hypothesis that the reduced cellulolysis observed in the *dpi* mutant is caused by the decreased abundance of several major cellulosomal components.

Characterization of cel48F and cel9E mutants

To verify the importance of Cel48F and Cel9E in cellulolysis, *cel48F* and *cel9E* mutants were created (Fig. S2). There was no obvious growth detected with the *cel48F* mutant on 10 g l^{-1} Avicel cellulose (Fig. 4A), indicating that Cel48F might be a pivotal cellulase for cellulolytic activity in this bacterium. This result was consistent with previously reported results using antisense RNA to knock down *cel48F* expression (Perret *et al.*, 2004). Also, direct *cel48F* inactivation had a more obvious effect on cellulolysis than antisense RNA. The *cel9E* mutant showed very weak growth on Avicel cellulose with a 64.5 ± 0.9% decrease in cell mass compared with WT (Fig. 4A). Additionally, although both Cel48F and Cel9E are cellobiohydrolases responsible for degrading cellulose to soluble sugars, dis-

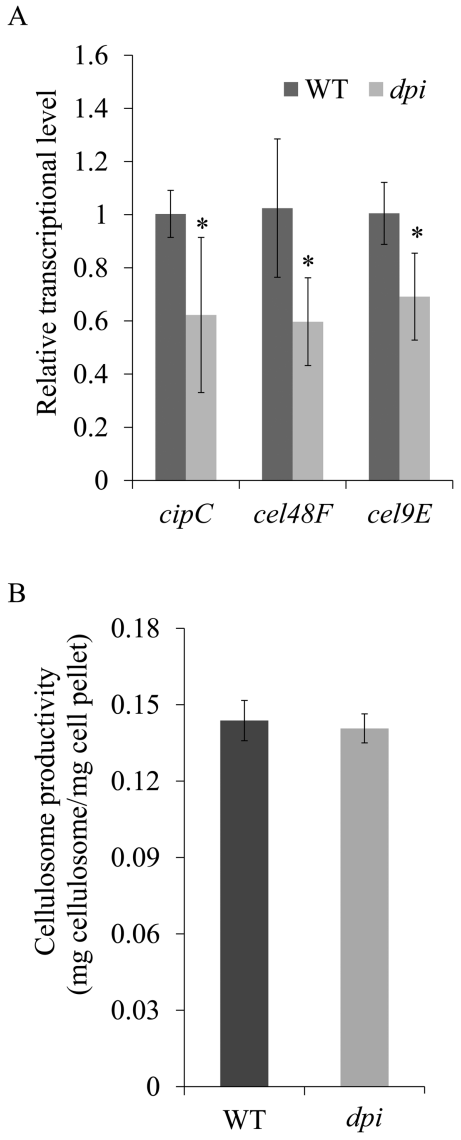

Fig. 3. Comparison of transcript levels and cellulosome productivity. Relative transcript levels (A) of the genes *cipC*, *cel48F* and *cel9E* in WT and *dpi* mutant grown on 10 g l^{-1} Avicel cellulose at the mid-logarithmic phase were compared by normalizing with the expression of the calibrator gene *recA*. Cellulosome productivity (B) of cellulose-grown WT and *dpi* mutant was calculated by dividing the isolated cellulosome amount by the total protein amount in the cell pellet. The means and standard deviations were calculated from the values of three biological replicates. The asterisk means significant difference between WT and *dpi* mutant (*$P < 0.05$, by Student's *t* test).

ruption of *cel48F* and *cel9E* unexpectedly affected cell growth on 5 g l^{-1} cellobiose, producing 29.4 ± 0.7% and 24.9 ± 1.2% less biomass in the stationary phase compared with WT, respectively (Fig. 4B), indicating that Cel48F and Cel9E might also play a role in cell growth on cellobiose.

To evaluate the potential polar effect caused by the insertion of a mobile group II intron, the transcript amounts of *cel8C* and *orfX*, the first downstream genes of *cel48F* and *cel9E*, respectively, were compared between WT and mutants. Since the *cel48F* mutant cannot grow on cellulose and the *cip-cel* gene cluster is expressed on cellobiose (Abdou *et al.*, 2008), the polar effect was evaluated in cellobiose-grown cells. qPCR analyses revealed that transcripts of *cel8C* in the *cel48F* mutant and *orfX* in the *cel9E* mutant were reduced to 42% and 39% respectively (Fig. 5). This means that the polar effect occurred in both mutants, which might also partially contribute to the observed defect in cellulolysis. However, the polar effect was insufficient to cause the total abolishment of *cel48F* mutant growth on cellulose. Therefore, these defects observed in the *cel48F* and *cel9E* mutants are a combinational effect of gene inactivation and polar effect on downstream genes.

Expression and enzymatic activity assay of recombinant Dpi protein

The *dpi* gene putatively encodes a dockerin-containing protease inhibitor (Dpi). Motif scanning of its putative peptide sequence (316 aa) predicted a signal peptide at

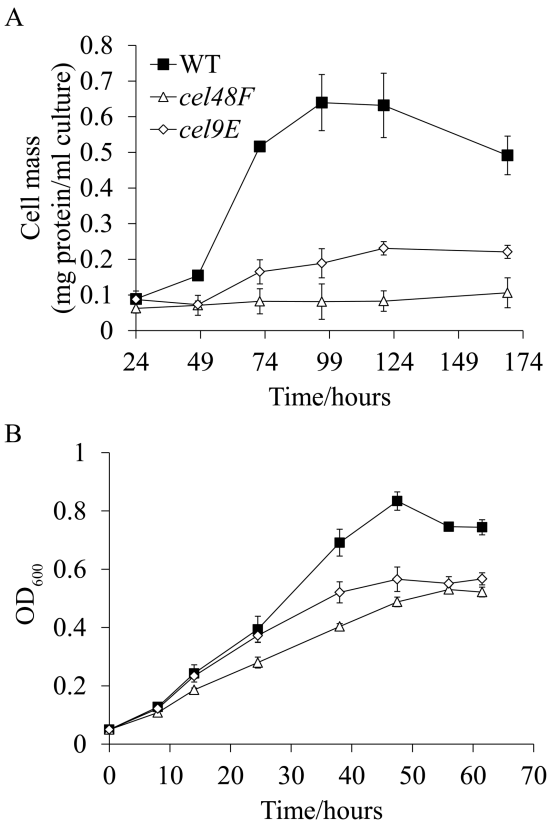

Fig. 4. Growth curves for WT, *cel48F* and *cel9E* mutants. All strains were grown on 10 g l^{-1} Avicel cellulose (A) and 5 g l^{-1} cellobiose (B). The means and standard deviations were calculated from three independent measurements at each time point.

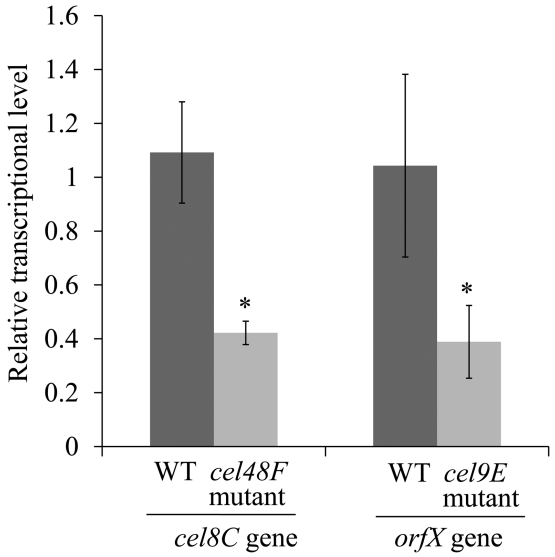

Fig. 5. Detection of the polar effect by quantitative real-time PCR. Relative transcript levels of the first downstream genes, *cel8C* in *cel48F* mutant and *orfX* in *cel9E* mutant were compared with these in WT by normalizing with *recA*. All strains were grown on 5 g l^{-1} cellobiose at the mid-logarithmic phase. The means and standard deviations were calculated from three biological replicates. The asterisk means significant difference between WT and mutants (*$P < 0.05$, by Student's *t* test).

the N-terminal (1–36 aa), a dockerin_1 domain in the middle (88–108 aa), and two Chagasin_I42 domains (135–223 aa and 226–315 aa) (Letunic *et al*., 2012) (http://smart.embl.de/). Using the structure prediction tool Phyre (Kelley and Sternberg, 2009), we constructed a visible model of the Dpi protein (44–316 aa) consisting of a type I dockerin domain and two Chagasin_I42 domains (Fig. 6A). The predicted type I dockerin domain is very similar to reported crystalline structures (Lytle *et al*., 2001; Pinheiro *et al*., 2008). The middle Chagasin_I42 domain has three conserved loops (DE, BC and FG) able to form a flexible wedge that may block the active site of cysteine protease according to previous studies (Figueiredo da Silva *et al*., 2007; Casados-Vazquez *et al*., 2011). However, the C-terminal Chagasin_I42 domain presents a much less compact structure even though conserved amino acids building up these three key loops exist, suggesting that these two Chagasin_I42 domains might have different enzymatic features.

The encoding sequence without the N-terminal signal peptide (44–316 aa) was cloned into pET28a(+) and then expressed in *Escherichia coli*. The recombinant Dpi harbouring a His tag at the N-terminal was produced with high yield (in lane 2, Fig. 6B) and purified with Ni(+) affinity chromatography under native conditions (in lane 4, Fig. 6B). The inhibitory activity of the recombinant Dpi was examined against commercial trypsin, chymotrypsin, papain and pepsin. In this test, only papain was effectively inhibited by Dpi (Table 1). This result is in accordance with functional prediction since papain belongs to the cysteine protease family (Rawlings *et al*., 2012). Moreover, in the reaction with 0.85 nM papain, the residual papain activity was gradually decreased by increasing Dpi dose (Fig. 6C). At a concentration of 1.8 nM of Dpi, half of the maximal papain activity was repressed ($IC_{50} = 1.8$ nM). When the Dpi concentration reached 3.31 nM (the molecular ratio of Dpi to papain is 3.89), less than 2% protease activity remained. Therefore, Dpi is an effective inhibitor of cysteine protease.

Discussion

Blouzard *et al*. identified Dpi as a cellulosomal component upon cell growth on different substrates (cellulose, xylan and wheat straw) of *C. cellulolyticum* (Blouzard *et al*., 2010). The present study reports the physiological functions of this cellulosome-localized protease inhibitor. Disruption of *dpi* affected cell growth on glucose, cellulose and xylan, but not on cellobiose. The *dpi* mutant grown with glucose entered into the stationary phase slightly earlier than WT, which could be caused by a temporal expression of *dpi* or its protease targets. The growth phase-dependent expression of a cell wall-associated cysteine protease has been found in *Staphylococcus epidermidis* (Oleksy *et al*., 2004). If the antagonistic activity of Dpi against protease targets was disrupted or abolished, the resulting hyperactive proteolysis would do damage to functional proteins/enzymes essential for cell growth. Even though glucose and cellobiose are both soluble carbon substrates, the cellobiose-grown mutant did not exhibit obvious differences from WT. It seems like cellobiose catabolism is not associated with cellulosomal Dpi functionalization. In addition to cell growth changes on cellulose and xylan, the mutant also presented lower efficiency in cellulolysis due to impairment of key cellulosomal components. Disturbance of cellulosomal composition negatively affects enzymatic activities to hydrolyse insoluble carbons (Maamar *et al*., 2004; Perret *et al*., 2004), thus reducing the amount of usable sugar available to support cell growth. Dpi is thus

Table 1. Inhibitory activity of the recombinant Dpi against commercial proteases.

Peptidase	Property	% Inhibitory activity
Papain	A cysteine endopeptidase	52.81 ± 1.05
Trypsin	A pancreatic serine protease	2.58 ± 2.15
Chymotrypsin	A serine endopeptidase	ND
Pepsin	A aspartate protease	ND

ND, no inhibition was detected.

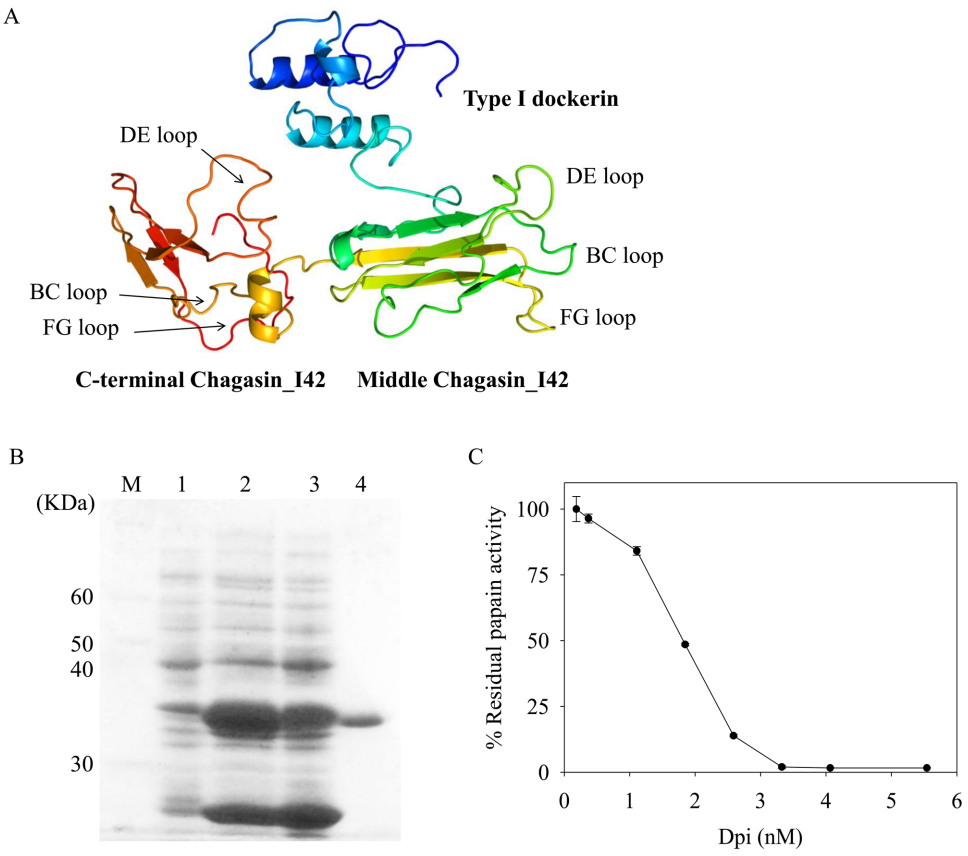

Fig. 6. Characterization of Dpi protein.
A. Modelling structure of Dpi protein (44–316 aa) established by Phyre based on the conserved type I dockerin and chagasin_I42 domains.
B. SDS-PAGE analysis of protein samples from non-induced (lane 1) and induced (lane 2) *E. coli* harbouring pET28a(+)-dpi vector, supernatant of induced cell lysate (lane 3) and purified Dpi (lane 4) after affinity chromatography.
C. Inhibitory efficiency of the purified Dpi against papain. The residual proteolytic activities of 0.85 nM papain were determined with the increase of Dpi dose from 0.18 nM to 5.54 nM. The half maximal inhibitory concentration (IC_{50}) was 1.8 nM at this given condition. The means and standard deviations were calculated from three independent measurements.

physiologically associated with cell growth and biomass utilization in a substrate-dependent manner.

The existence of cellulosomal protease inhibitors raises questions as to how the cellulosome is maintained or modified under diverse environmental conditions (Schwarz and Zverlov, 2006). Similar inhibitors from *C. thermocellum* and *C. cellulovorans* exhibited inhibitory activities against bacterial and plant proteases (Kang *et al*., 2006; Meguro *et al*., 2011). In the present study, *C. cellulolyticum* Dpi was shown to be an effective inhibitor of cysteine protease. It is the sole protease inhibitor encoded by *C. cellulolyticum* according to the MEROPS database (Rawlings *et al*., 2012). Functionalization of cellulosomal Dpi depends on cysteine proteases that can be secreted out of the cell. Whole-genome mining uncovered seven of 25 potential secretory peptidases belonging to the cysteine protease family and can be further divided into three subfamilies, C1A (Ccel_2442), C82 (Ccel_2590) and C40 (Ccel_0747, Ccel_1652, Ccel_1956, Ccel_2128 and Ccel_2940) (Rawlings *et al*., 2012). Interestingly, Ccel_2442 from C1A, carrying a dockerin domain and two Chagasin domains, is a papain-like cysteine protease, which is considered as the most probable target of Dpi but has never been identified in active cellulosomes. Both C82 and C40 are involved in bacterial cell-wall modification. Further efforts will focus on identifying the *in vivo* inhibitory targets of Dpi.

CipC scaffoldin has been used as an internal calibrator to quantify the relative abundance of cellulosomal components (Perret *et al*., 2004). Our analysis determined that the abundances of Cel48F, Cel9E and Cel9M relative to CipC were decreased significantly in the mutant. These changes were largely caused by higher proteolysis induced by Dpi loss, not by different transcript amounts because all of the encoded genes were co-transcribed in the *cip-cel* gene cluster (Maamar *et al*., 2006) and there was no statistically significant difference observed in the transcript amounts of *cipC*, *cel48F* and *cel9E* in the mutant. Additionally, the lack of the *dpi* gene also reduced the amounts of transcripts from the *cip-cel* gene cluster in an

unknown manner. This could be caused by reduced transcriptional activity and/or differential RNA stability. The lowered transcript levels did not significantly reduce cellulosome productivity. The poor correlation between RNA transcript and protein abundance has been reported in both prokaryotic and eukaryotic cells because of various biological factors (e.g. RNA abundance, RNA secondary structure, ribosome occupancy, codon bias, amino acid usage and protein half-lives) and methodological constraints (e.g. detection sensitivity and experimental error and noise) (Maier *et al.*, 2009). Interestingly, CipC abundance was similar in both the mutant and WT and several bands in the Ff fraction also showed similar abundance in both the mutant and WT. These results suggest that the proteolysis may be non-random and target specific proteins.

Even though glycoside hydrolases are important to cellulose saccharification, the contribution of each family to cellulolysis is still under active investigation. Families 48 and 9 are major cellulosomal components (Maamar *et al.*, 2004; Perret *et al.*, 2004; Blouzard *et al.*, 2010; Olson *et al.*, 2010). The disruption of the *cel48F* gene in *C. cellulolyticum* completely eliminated cell growth on cellulose. This defect is more severe than the report based on RNAi-mediated knock-down of *cel48F* expression (Perret *et al.*, 2004). The deletion of Cel48S from *C. thermocellum* resulted in a 40% decrease in cellular yield and 35% lower activity on Avicel cellulose (Olson *et al.*, 2010). However, the essential role of Cel48 in cellulolytic processes needs further evaluation because of the polar effect caused by intron insertion. A more severe polar effect was observed in a *cipC*Mut1 mutant that was created by IS insertion into the *cipC* gene in *C. cellulolyticum* (Maamar *et al.*, 2004). The polar effect in the *cipC*Mut1 mutant blocked the generation of 7.5 kb-long transcripts that were long enough to carry both *cipC* and the downstream *cel48F*. However, the transcript level of *cel8C*, which is immediately downstream in the *cel48F* mutant, remained at 42%. The disruption of the sole family 9 glycoside hydrolase in *Clostridium phytofermentans* abolished cellulose degradation activity (Tolonen *et al.*, 2009). However, the *C. cellulolyticum cel9E* mutant only showed a 64.5% decrease in cellular biomass on cellulose, which was also a combinational effect of gene inactivation and polar effect. Taken together, these studies showed that the importance of Cel48 and Cel9 varied in cellulose-degrading *Clostridium* species. Interestingly, the loss of Cel48F and Cel9E reduced cell yield on cellobiose. A similar result was also observed due to Cel48S deletion in *C. thermocellum* (Olson *et al.*, 2010). Thus, Cel48F and Cel9E as key cellulosomal cellobiohydrolases could exert broader influences on cellular metabolism.

Combined with previous studies (Kang *et al.*, 2006; Schwarz and Zverlov, 2006; Meguro *et al.*, 2011) and our findings, a conceptual model for Dpi-mediated regulation of cellulosome activity is proposed (Fig. 7). In *C. cellulolyticum,* the *cip-cel* gene cluster and many other genes encode and secret diverse structural proteins and hydrolases. Cellulosomes are assembled on the cell surface with a diversity of components (e.g. glycoside hydrolases 48 family and 9 family, protease inhibitors) by dockerin-cohesin interaction (Carvalho *et al.*, 2003; Bayer *et al.*, 2004; Doi and Kosugi, 2004; Blouzard *et al.*, 2010; Fontes and Gilbert, 2010). Glycoside hydrolases on cellulosomes synergistically degrade diverse insoluble carbon substrates into soluble sugars to support cell growth. Under normal conditions, hydrolase activity is stabilized by protease inhibitors such as Dpi in *C. cellulolyticum*, serpins in *C. thermocellum* (Kang *et al.*, 2006) and cyspins in *C. cellulovorans* (Meguro *et al.*, 2011), which inhibit proteolysis, allowing the cells to continue to degrade available substrates at high efficiency. However, if the protease-inhibitor balance is disrupted by the introduction of exogenous proteases from plant biomass, some glycoside hydrolases (e.g. Cel48F and Cel9E in *C. cellulolyticum*) will be preferentially destroyed by the hyperactive proteolytic activity. Then the evacuated cohesin domains will be occupied by

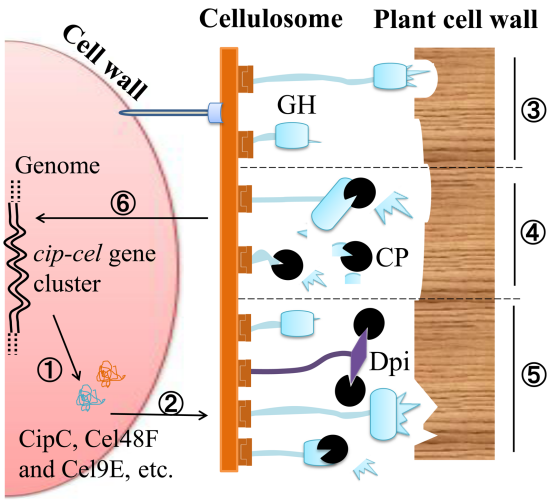

Fig. 7. A conceptual model of Dpi-mediated regulation of cellulosomal activity in *C. cellulolyticum*. The *cip-cel* gene cluster expresses major cellulosomal components (e.g. CipC, Cel48F and Cel9E) (①) which assemble to form cell surface-bound cellulosomes (②) responsible for lignocellulose degradation (③). Bacteria-/plant biomass-derived cysteine proteases attack cellulosomal glycoside hydrolases (e.g. Cel48F and Cel9E) (④), thus reducing cellulolytic activity. Cellulosome-localized Dpi is able to block these proteases and protect cellulosomal components from proteolytic damage (⑤). Conversely, loss of the antagonistic balance due to differential expression or external protease attack will cause proteolysis of key cellulosomal components and simultaneously allow other dockerin-containing components to be incorporated, and also downregulate the expression of the *cip-cel* gene cluster (⑥) with unknown mechanism. CP, cysteine proteases from plant biomass or bacteria; CipC, CipC scaffoldin; Dpi, dockerin-containing protease inhibitor of cysteine proteases; GH, glycoside hydrolases.

Table 2. Bacterial strains and plasmids.

Strain or plasmid	Relevant characteristics	Reference
Strains		
Escherichia coli Top10	F⁻ *mcr*A Δ(*mrr-hsd*RMS- *mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *nup*G *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K *rps*L(Str^R) *end*A1	Invitrogen
E. coli Rosetta™ 2(DE3)pLysS	F⁻ *omp*T *hsd*S_B(r_B⁻ m_B⁻) *gal dcm* (DE3) pLysSRARE2 (Cam^R)	Novagen
Clostridium cellulolyticum H10	Wild type	Petitdemange *et al.* (1984)
dpi	A group II intron was inserted into *dpi* ORF at 171 nt	This study
dpi/zero	*dpi* mutant background with pClostron3-*dpi/zero* plasmid	This study
dpi/over	*dpi* mutant background with pClostron3-*dpi/over* complementary plasmid	This study
cel48F	A group II intron was inserted into *cel48F* ORF at 764 nt	This study
cel9E	A group II intron was inserted into *cel9E* ORF at 653 nt	This study
Plasmids		
pLyc1217Er	Kan^r in *E. coli*, Em^r in *Clostridium*, Fd promoter, pWH199 derivative	Li *et al.* (2012)
pClostron3	CMP^r in *E. coli*, TMP^r in *Clostridium*, Fd promoter, pJIR750a derivative	This study
pClostron3-*dpi/zero*	pClostron3 derivative with deletion of group II intron and LtrA	This study
pClostron3-*dpi/over*	pClostron3 derivative with *dpi* ORF driven by Fd promoter	This study
pET28a(+)-*dpi*	*dpi* coding sequence was ligated into NdeI–NotI-linearized pET28a(+)	This study

Abbreviations: *dpi*, *docterin-containing protease inhibitor* gene; Em^r, erythromycin resistant; Kan^r, kanamycin resistant; Fd, ferredoxin; LtrA, intron-encoded protein.

other available dockerin-containing components, leading to a dynamic change of cellulosomal composition. The disturbance of the protease-inhibitor balance would down-regulate major transcripts of the *cip-cel* gene cluster, but not cellulosome productivity. The hyperactive proteolytic activity would also appear when protease and inhibitor genes are differentially expressed. It is possible that facing new carbon sources, cells may adjust protease-inhibitor expression leading to a proteolysis-dependent removal of the initially incorporated major cellulosomal components, thus allowing new substrate-induced enzymes to be assembled into the cellulosome. It should be noted that although supported by experimental data, further investigation is needed to substantiate this model.

In conclusion, this study uncovered the physiological role of a dockerin-containing protease inhibitor in protecting key cellulosomal cellulases from proteolysis, and identified the *in vivo* importance of two major cellulosomal components, Cel48F and Cel9E in crystalline cellulose degradation. This study suggests a mechanism by which cellulase stability may be enhanced via controlling protease/inhibitor activity or cellulase protein engineering to improve lignocellulose hydrolysis efficiency.

Experimental procedures

Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 2. *E. coli* Top10 (Invitrogen) and Rosetta™ 2(DE3) pLysS strain (Novagen) were used for cloning and protein expression respectively. *E. coli* transformants were grown at 37°C in Luria–Bertani medium supplemented with kanamycin (50 μg ml^{-1}) and/or chloramphenicol (15 μg ml^{-1}) when required.

Clostridium cellulolyticum H10 and the developed strains (including *dpi*, *cel48F, cel9E*, *dpi/over* and *dpi/zero*) were cultured anaerobically at 34°C in modified VM medium supplemented with yeast extract (2.0 g l^{-1}) and various carbon sources (Higashide *et al.*, 2011). The medium was supplemented with erythromycin (15 μg ml^{-1}) or thiamphenicol (15 μg ml^{-1}) as appropriate. Colonies of each strain were isolated on solid VM medium containing 1% (weight/volume) agar and amended with cellobiose (5 g l^{-1}) and erythromycin (15 μg ml^{-1}) or thiamphenicol (15 μg ml^{-1}) as appropriate. For making cellulose-containing top-agar plates, sterile Avicel cellulose mixed with un-solidified VM agar was overlaid on the solidified VM agar. When required, thiamphenicol (15 μg ml^{-1}) was added. Cells (10 μl at OD_{600} = 0.4) were dropped on the plates and then incubated at 34°C.

Plasmid construction and transformation of C. cellulolyticum *H10*

Clostridium cellulolyticum mutants were generated by group II intron insertion. Before transformation, the intron region of *E. coli–C. cellulolyticum* shuttle vector pLyc1217Er was modified (Higashide *et al.*, 2011; Li *et al.*, 2012). Based on the online intron design tool (http://clostron.com/), we chose antisense integration sites at 171 bp, 764 bp and 653 bp downstream of the start codons of *dpi*, *cel48F* and *cel9E* genes, respectively, and then synthesized four PCR primers (Table S2) for each intron modification, including IBS, EBS1d, EBS2 and EBSu. One-step cross-over PCR using these four primers and pLyc1217Er as the template gave intron amplicon which was used to replace the original intron region after digestion with XmaI and BsrGI. The vectors were confirmed by sequencing and then were used for *C. cellulolyticum* transformation (Li *et al.*, 2012), generating *dpi*, *cel48F* and *cel9E* mutants.

For the *dpi* mutant complementation, the plasmid pClostron3-*dpi/over* harbouring the intact *dpi* ORF driven by a *Clostridium pasteurianum* ferredoxin (Fd) promoter was con-

structed (Graves and Rabinowitz, 1986). Using *C. cellulolyticum* genomic DNA as template, primers Dpi-overF and Dpi-overR were used to amplify the ORF (Table S2). PCR product was ligated into pClonstron3. The resulting plasmid confirmed by sequencing was then named pClostron3-*dpi/over*. The empty plasmid without any ORF downstream of the Fd promoter was named pClostron3-*dpi/zero* and was used as a negative control. The *dpi* mutant transformed with pClostron3-*dpi/over* and pClostron3-*dpi/zero* generated *dpi/over* and *dpi/zero* strains respectively. The plasmids were transferred to *C. cellulolyticum* by electroporation as previously described (Li *et al*., 2012).

For expressing the recombinant Dpi in *E. coli*, the coding sequence of the *dpi* gene was cloned into the pET28a(+) vector (Novagen). The coding region was amplified by PCR using primers NtDpiF and NtDpiR (Table S2). An 879 bp amplicon digested with NdeI and NotI were cloned into NdeI–NotI-linearized pET28a(+), resulting in pET28a(+)-*dpi*. The final plasmid carries the *dpi* coding sequence fused in frame at its N-terminus with a sequence encoding hexahistidine residues (His tag). The plasmid was transformed into Rosetta™ 2(DE3)pLysS competent cells to produce recombinant proteins according to the manufacturer's instruction.

Growth and cellulose degradation measurement

Clostridium cellulolyticum growth on glucose (10 g l^{-1}) or cellobiose (10 g l^{-1}) was measured by monitoring OD_{600}. But on Avicel cellulose (10 g l^{-1}) or xylan (10 g l^{-1}), cell growth was determined by measuring bacterial protein content using the Pierce®BCA Protein Assay Kit (Thermo Scientific). Residual cellulose in cultures was estimated by using the phenol-sulphuric acid method, with glucose as the standard (Dubois *et al*., 1956).

RNA isolation and quantitative real-time PCR

Total RNA was extracted from cellulose (10 g l^{-1}) -grown *C. cellulolyticum* cells at the mid-logarithmic phase by TRIzol® Reagent (Invitrogen). The RNA yield and integrity was determined with spectrophotometry and gel electrophoresis respectively. And then reverse transcription was conducted by using SuperScript® III Reverse Transcriptase (Invitrogen). cDNA products were diluted as appropriate and used as the templates. Quantitative real-time PCR was performed using iTaq SYBR Green Supermix with ROX (Bio-Rad) on a Bio-Rad iQ5 thermal cycler. Gene-specific primers used for transcript quantification are listed in Table S2. The thermal cycling conditions were as follows: 95°C for 3 min, 40 cycles of 95°C for 15 s, 55°C for 15 s and 72°C for 45 s. The *recA* gene was used as an internal calibrator (Stevenson and Weimer, 2005). Relative expression level was calculated with the PfaffI Method (Pfaffl, 2001).

Fractionation of extracellular proteins

The *C. cellulolyticum* strains were grown on VM medium with cellulose (10 g l^{-1}). During mid-logarithmic phase the culture was filtered through a 3-μm-pore-size GF/D glass fibre (Whatman). The penetration fluid was centrifuged to collect the supernatant containing the free extracellular protein fraction (Ff) and then concentrated with acetone precipitation. The cellulose retained on the filter was used to isolate bound proteins which mainly contain the cellulosome fraction (Fc) as previously described (Maamar *et al*., 2004). Protein concentration was determined by the Pierce®BCA Protein Assay Kit (Thermo Scientific) according to the manufacturer's instruction.

Expression and purification of recombinant Dpi protein

To express recombinant Dpi protein, Rosetta™ 2(DE3) pLysS strain carrying pET28a(+)-*dpi* vector with an OD_{600} of 0.7 was induced with 0.2 mM IPTG at 25°C for 15 h. The induced cells were harvested by centrifugation and then lysed by using CelLytic™ B 2× (Sigma) according to the manufacturer's protocol. The lysates were centrifuged and filtered with 0.2 μm filters (Sigma). The supernatant lysate was purified using a HisTrap HP 1 ml column (GE Healthcare) according to the manufacturer's instruction. The eluate was fractionized during the washing step and the purity of each fraction was evaluated by SDS-PAGE. The fractions with pure recombinant protein were pooled, dialysed and concentrated with an Amicon concentrator in 50 mM phosphate buffer (pH 7.2). Protein concentration was quantified using the Pierce®BCA Protein Assay Kit (Thermo Scientific). The recombinant protein was supplemented with 50% glycerol and then stored at –20°C for further analysis.

SDS-PAGE analysis and MS identification

Protein samples from *E. coli* and *C. cellulolyticum* were subjected to SDS-PAGE using 10% resolving gels and mini electrophoresis units (Bio-Rad). Gels were stained with Coomassie blue. For densitometry analysis, decoloured gels were scanned and analysed with MYImage (Thermo Scientific).

To identify proteins in the gel, mass spectrometry was performed as follows. Protein bands excised from gel were subjected to in-gel trypsin digestion with reduction and alkylation as previously described (Wilm *et al*., 1996). Then tryptic peptides were applied to HPLC and MS/MS analysis with the DionexUltiMate 3000 and ABI MDS SciexQstar Elite respectively. MS/MS data collected were submitted to in-house MASCOT (Matrix Science) server for protein identification against the NCBInr (02-2012) protein database.

Inhibitory activity test of the recombinant Dpi

The inhibitory activity of the recombinant Dpi was tested on commercial proteases including trypsin, chymotrypsin, papain and pepsin (All from Sigma) by using EnzChek® Protease Assay Kits (Invitrogen). Trypsin, chymotrypsin and pepsin were dissolved in 0.001 N HCl, making stock solutions (0.1 mg ml^{-1}). Papain was dissolved in 50 mM sodium acetate (pH 4.5) with the concentration of 0.5 mg ml^{-1}. Following the manufacturer's instructions, the proteolytic reactions were performed in various working buffers, trypsin (20 μg ml^{-1}) and chymotrypsin (3.75 μg ml^{-1}) in 10 mM Tris-HCl (pH 7.8), pepsin (25 μg ml^{-1}) in 20 mM sodium acetate

(pH 4) and papain (10 μg ml^{-1}) in 10 mM MES (pH 6.2), all of which were supplemented with the native or boiled recombinant Dpi (25 μg ml^{-1}) and substrate casein (5 μg ml^{-1}). All reactions were incubated at 37°C for 1 h before detecting the fluorescence using excitation and emission filters of 595 nm and 630 nm respectively. Percentage of inhibition was calculated by dividing the difference in fluorescence values of reactions with the boiled Dpi from those with intact Dpi by the corresponding control reactions with the boiled Dpi, and then multiplying by 100.

Determination of IC_{50} value of Dpi against papain

The inhibitory capacity of Dpi towards papain activity was determined. Each reaction consists of casein substrate (5 μg ml^{-1}), papain (0.85 nM) and various concentrations of the native or boiled Dpi (0–5.54 nM) in 10 mM MES (pH 6.2). The boiled Dpi was used in control groups at each concentration. Before adding casein substrate, other components were mixed in advance and incubated at 37°C for 15 min. Reactions were incubated at 37°C for 1 h and then the fluorescence was measured using excitation and emission filters of 595 nm and 630 nm respectively. All assays were made in triplicate. IC_{50} was defined as the concentration of Dpi required for achieving 50% inhibition of papain.

Acknowledgements

We thank Professor Laura E. Bartley for suggestions on proteolytic activity test, Dr Chris Hemme and Dr Joy D. Van Nostrand for revising the English in the manuscript and Yue Huang for helpful comments. This work was supported by the NSF EPSCoR award EPS 0814361.

References

Abdou, L., Boileau, C., de Philip, P., Pages, S., Fierobe, H.P., and Tardif, C. (2008) Transcriptional regulation of the *Clostridium cellulolyticum cip-cel* operon: a complex mechanism involving a catabolite-responsive element. *J Bacteriol* **190:** 1499–1506.

Argyros, D.A., Tripathi, S.A., Barrett, T.F., Rogers, S.R., Feinberg, L.F., Olson, D.G., *et al.* (2011) High ethanol titers from cellulose by using metabolically engineered thermophilic, anaerobic microbes. *Appl Environ Microbiol* **77:** 8288–8294.

Bayer, E.A., Belaich, J.P., Shoham, Y., and Lamed, R. (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annu Rev Microbiol* **58:** 521–554.

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

Brown, S.D., Guss, A.M., Karpinets, T.V., Parks, J.M., Smolin, N., Yang, S., *et al.* (2011) Mutant alcohol dehydrogenase leads to improved ethanol tolerance in *Clostridium thermocellum*. *Proc Natl Acad Sci U S A* **108:** 13752–13757.

Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37:** D233–D238.

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

Casados-Vazquez, L.E., Lara-Gonzalez, S., and Brieba, L.G. (2011) Crystal structure of the cysteine protease inhibitor 2 from *Entamoeba histolytica*: functional convergence of a common protein fold. *Gene* **471:** 45–52.

Dashtban, M., Schraft, H., and Qin, W. (2009) Fungal bioconversion of lignocellulosic residues; opportunities & perspectives. *Int J Biol Sci* **5:** 578–595.

Desvaux, M. (2005) *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. *FEMS Microbiol Rev* **29:** 741–764.

Doi, R.H., and Kosugi, A. (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. *Nat Rev Microbiol* **2:** 541–551.

Dubois, M., Gilles, K.A., Hamilton, J.K., Rebers, P.A., and Smith, F. (1956) Colorimetric Method for determination of sugars and related substances. *Anal Chem* **28:** 350–356.

Figueiredo da Silva, A.A., de Carvalho Vieira, L., Krieger, M.A., Goldenberg, S., Zanchin, N.I., and Guimaraes, B.G. (2007) Crystal structure of chagasin, the endogenous cysteine-protease inhibitor from *Trypanosoma cruzi*. *J Struct Biol* **157:** 416–423.

Fontes, C.M., and Gilbert, H.J. (2010) Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annu Rev Biochem* **79:** 655–681.

Gold, N.D., and Martin, V.J. (2007) Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J Bacteriol* **189:** 6787–6795.

Graves, M.C., and Rabinowitz, J.C. (1986) *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferredoxin gene. Evidence for 'extended' promoter elements in gram-positive organisms. *J Biol Chem* **261:** 11409–11415.

Guedon, E., Desvaux, M., and Petitdemange, H. (2002) Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. *Appl Environ Microbiol* **68:** 53–58.

Han, S.O., Yukawa, H., Inui, M., and Doi, R.H. (2005) Effect of carbon source on the cellulosomal subpopulations of *Clostridium cellulovorans*. *Microbiology* **151:** 1491–1497.

Heap, J.T., Kuehne, S.A., Ehsaan, M., Cartman, S.T., Cooksley, C.M., Scott, J.C., and Minton, N.P. (2010) The ClosTron: mutagenesis in *Clostridium* refined and streamlined. *J Microbiol Methods* **80:** 49–55.

Hemme, C.L., Mouttaki, H., Lee, Y.J., Zhang, G., Goodwin, L., Lucas, S., *et al.* (2010) Sequencing of multiple clostridial genomes related to biomass conversion and biofuel production. *J Bacteriol* **192:** 6494–6496.

Higashide, W., Li, Y., Yang, Y., and Liao, J.C. (2011) Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. *Appl Environ Microbiol* **77:** 2727–2733.

Kang, S., Barak, Y., Lamed, R., Bayer, E.A., and Morrison, M.

(2006) The functional repertoire of prokaryote cellulosomes includes the serpin superfamily of serine proteinase inhibitors. *Mol Microbiol* **60:** 1344–1354.
Kelley, L.A., and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4:** 363–371.
Kuhad, R.C., Gupta, R., and Singh, A. (2011) Microbial cellulases and their industrial applications. *Enzyme Res* **2011:** 280696.
Letunic, I., Doerks, T., and Bork, P. (2012) SMART 7: recent updates to the protein domain annotation resource. *Nucleic Acids Res* **40:** D302–D305.
Li, Y., Tschaplinski, T.J., Engle, N.L., Hamilton, C.Y., Rodriguez, M., Jr, Liao, J.C., *et al*. (2012) Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol Biofuels* **5:** 2.
Lynd, L.R., Weimer, P.J., van Zyl, W.H., and Pretorius, I.S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* **66:** 506–577.
Lytle, B.L., Volkman, B.F., Westler, W.M., Heckman, M.P., and Wu, J.H. (2001) Solution structure of a type I dockerin domain, a novel prokaryotic, extracellular calcium-binding domain. *J Mol Biol* **307:** 745–753.
Maamar, H., Valette, O., Fierobe, H.P., Belaich, A., Belaich, J.P., and Tardif, C. (2004) Cellulolysis is severely affected in *Clostridium cellulolyticum* strain *cipC*Mut1. *Mol Microbiol* **51:** 589–598.
Maamar, H., Abdou, L., Boileau, C., Valette, O., and Tardif, C. (2006) Transcriptional analysis of the *cip-cel* gene cluster from *Clostridium cellulolyticum*. *J Bacteriol* **188:** 2614–2624.
Maier, T., Guell, M., and Serrano, L. (2009) Correlation of mRNA and protein in complex biological samples. *FEBS Lett* **583:** 3966–3973.
Meguro, H., Morisaka, H., Kuroda, K., Miyake, H., Tamaru, Y., and Ueda, M. (2011) Putative role of cellulosomal protease inhibitors in *Clostridium cellulovorans* based on gene expression and measurement of activities. *J Bacteriol* **193:** 5527–5530.
Oleksy, A., Golonka, E., Banbula, A., Szmyd, G., Moon, J., Kubica, M., *et al*. (2004) Growth phase-dependent production of a cell wall-associated elastinolytic cysteine proteinase by *Staphylococcus epidermidis*. *Biol Chem* **385:** 525–535.
Olson, D.G., Tripathi, S.A., Giannone, R.J., Lo, J., Caiazza, N.C., Hogsett, D.A., *et al*. (2010) Deletion of the Cel48S cellulase from *Clostridium thermocellum*. *Proc Natl Acad Sci U S A* **107:** 17727–17732.
Perret, S., Maamar, H., Belaich, J.P., and Tardif, C. (2004) Use of antisense RNA to modify the composition of cellulosomes produced by *Clostridium cellulolyticum*. *Mol Microbiol* **51:** 599–607.
Petitdemange, E., Caillet, F., Giallo, J., and Gaudin, C. (1984) *Clostridium cellulolyticum* sp. nov., a cellulolytic, mesophilic species from decayed grass. *Int J Syst Bacteriol* **34:** 155–159.
Pfaffl, M.W. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29:** e45.
Pinheiro, B.A., Proctor, M.R., Martinez-Fleites, C., Prates, J.A., Money, V.A., Davies, G.J., *et al*. (2008) The *Clostridium cellulolyticum* dockerin displays a dual binding mode for its cohesin partner. *J Biol Chem* **283:** 18422–18430.
Rao, M.B., Tanksale, A.M., Ghatge, M.S., and Deshpande, V.V. (1998) Molecular and biotechnological aspects of microbial proteases. *Microbiol Mol Biol Rev* **62:** 597–635.
Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012) MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res* **40:** D343–D350.
Schwarz, W.H. (2001) The cellulosome and cellulose degradation by anaerobic bacteria. *Appl Microbiol Biotechnol* **56:** 634–649.
Schwarz, W.H., and Zverlov, V.V. (2006) Protease inhibitors in bacteria: an emerging concept for the regulation of bacterial protein complexes? *Mol Microbiol* **60:** 1323–1326.
Stevenson, D.M., and Weimer, P.J. (2005) Expression of 17 genes in *Clostridium thermocellum* ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. *Appl Environ Microbiol* **71:** 4672–4678.
Tolonen, A.C., Chilaka, A.C., and Church, G.M. (2009) Targeted gene inactivation in *Clostridium phytofermentans* shows that cellulose degradation requires the family 9 hydrolase Cphy3367. *Mol Microbiol* **74:** 1300–1313.
Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., and Mann, M. (1996) Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. *Nature* **379:** 466–469.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web-site.