Received Date: 27-Jul-2012
Revised Date : 09-Oct-2012
Accepted Date: 10-Oct-2012
Article type : Research Letter
Editor : Andre Klier

Exchange of Type II Dockerin-Containing Subunits of the *C. thermocellum* Cellulosome as Revealed by SNAP-tags

Benjamin H. Waller[2,4], Daniel G. Olson[1,4], Devin H. Currie[2,4], Adam M. Guss[3,4], and Lee R. Lynd[1,2,4*]

[1]Thayer School of Engineering, Dartmouth College, Hanover, NH 03755
[2]Department of Biological Sciences, Dartmouth College, Hanover, NH 03755
[3]Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831
[4]BioEnergy Science Center, Oak Ridge, TN 37830

*To whom correspondence should be addressed
Phone: 603-646-2231
Fax: (603) 646-2277
Email: Lee.R.Lynd@Dartmouth.edu

Keywords: Dockerin, Cohesin, Cellulosome, *Clostridium thermocellum*, SNAP-tag
Running Title: Visualization of type II cohesins in *Clostridium thermocellum*

Abstract:

Clostridium thermocellum is a thermophilic anaerobic bacterium which efficiently hydrolyzes and metabolizes cellulose to ethanol through the action of its cellulosome, a multiprotein enzymatic complex. A fluorescent protein probe, consisting of a type II dockerin-module fused to a SNAP-tag, was developed in order to gain insight into the quaternary configuration of the cellulosome and to investigate the effect of deleting *cipA,* the protein scaffold on which the cellulosome is built. Fluorescence microscopy suggested that the probe had localized to polycellulosomal protuberances on the cell surface. Surprisingly, fluorescence intensity did not substantially change in the *cipA* deletion mutants. Sequential labeling experiments suggested that this was a result of bound type II dockerins from CipA being replaced by unbound type II dockerins from the fluorophore-SNAP-XDocII probe. This mechanism of dockerin exchange could represent an efficient means for modifying cellulosome composition.

Introduction:

Clostridium thermocellum is a thermophilic, gram-positive bacterium which is of interest for biofuel production due to its high rate of cellulose utilization (Lynd et al., 2002). This ability is due in part to its cellulosome, a multiprotein enzymatic complex tethered to the cell surface. The cellulosome consists of many repeated enzymatic subunits organized around a noncatalytic polypeptide, the primary scaffoldin, CipA. CipA has nine type I cohesin modules, one type II dockerin module, and a cellulose binding module that mediates attachment of the cellulosome to its substrate. The type I cohesins of CipA bind to type I dockerin modules on enzymatic subunits that possess diverse hydrolytic activities. The type II dockerin of CipA binds to a type II cohesin on secondary anchoring scaffoldins tethered to the cell surface by an S-layer protein which interacts noncovalently with the peptidoglycan layer of the bacterial cell wall. Anchoring scaffoldins SdbA, Orf2p and OlpB bind 1, 2, and 7 CipAs respectively, allowing incorporation of up to 63 enzymatic subunits into a single complex that acts synergistically at the cell surface (Bayer et al., 2008).

The expression of both catalytic and structural components of the cellulosome change during growth on different substrates, indicating that *C. thermocellum* regulates its cellulosome composition in response to the available substrate and that the ability to exchange these subunits is important for efficient metabolism (Gold & Martin, 2007; Raman et al., 2009). A bicistronic system of carbohydrate-sensing anti-sigma and sigma factors has been shown to be able to regulate cellulase gene expression and respond to changes in substrate (Nataf et al., 2010).

Polypeptide sequences of the cellulosome components contain typical surface signal peptides, suggesting that the components are secreted individually and the cellulosome is assembled on the cell surface (Beguin & Aubert, 1994). The cellulosome subunits are invariably found in the complexed form, suggesting a strong interaction between enzymes and scaffoldin proteins(Bayer et al., 1985). The interaction between cohesins and dockerins is one of the strongest reported in nature with disassociation constants < 10^{-9}M (Mechaly et al 2001). During active growth, the cellulosome tightly adheres to the cell surface and also to the solid substrate forming a complex between cells, cellulosome and cellulose. However, *C. thermocellum* is known to release cellulosomes throughout growth and en-masse in late-stationary phase (Demain et al., 2005; Raman et al., 2009). This turnover and release of cellulosomes during fermentation may be necessary to allow for the creation of new cellulosomes with modified composition. It has also been suggested that the controlled release of cellulosomes during growth may function as a mechanism to release *C. thermocellum* from its substrate, leaving deployed cellulosomes to continue hydrolyzing cellulose (Bayer & Lamed, 1986).

Although extensive work has been done analyzing the composition of purified cellulosomes, the composition of the cellulosome in its native microbial context is not well understood. There is increasing interest in building artificial cellulosomes, which is currently limited by a lack of understanding of structural elements in native cellulosomes (Krauss et al., 2012).

In order to increase understanding of the cellulosome in its native microbial context, we undertook work to develop a fluorescent probe for labeling type II cohesins based on the commercially available SNAP-tag labeling system (Keppler et al. 2003). The SNAP-tag system

was developed by Keppler et al. as a method of covalently labeling fusion proteins *in*vivo. SNAP-tag is a mutant of the O^6-alkylguanine-DNA alkyl transferase human DNA repair protein which has increased activity against its substrate O^6-benzylguanine. The mutated protein binds covalently with benzylguanine-derived fluorophores. To create the probe, we fused a type II dockerin with the commercially available SNAP-tag. We then used this probe to visualize localization of type II cohesin modules in the cellulosome for both wild type and mutants of the *cipA* scaffolding protein (Supplemental figure 1).

Methods:

Strains and media

C. thermocellum DSM 1313 (WT) was grown in modified DSM 122 broth (Olson et al., 2010) with the addition of 50 mM 3-(N-morpholino) propanesulfonic acid (MOPS) sodium salt and 3 g/L trisodium citrate (Na3-C6H5O7*2H2O). All manipulations of *C. thermocellum* 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. *C. thermocellum* was grown at 55°C using 5 g/L cellobiose as the primary carbon source. The genotype of strains used in this work are listed in Table 1. Strain construction was performed as described previously (Argyros et al., 2011; Guss et al., 2012; Olson et al., 2012) using plasmids listed in Table 2. Briefly, the regions annotated as "5' flank" and "3' flank" are present on both the plasmid and the chromosome. By a series of recombination events, the region flanked by the "5' flank" and "3' flank" on the chromosome is replaced by the corresponding region from the plasmid. Plasmid sequences are available from Genbank (accession number in Table 2).

Expression and purification of the fusion protein

A SNAP-XDocII fusion protein was created by cloning the XDocII region from the *cipA* gene into the pSNAP-tag® (T7)-2 Vector in-frame with the SNAP-tag coding sequence using the In-Fusion cloning kit (Clontech) to create plasmid pDGO-54. This plasmid was transformed into T7 Express lysY/Iq competent *Escherichia coli* (New England BioLabs). A 0.1% inoculum was used, and cell cultures were incubated aerobically at 37°C with vigorous shaking. When the optical density (600 nm) reached a value of 0.6, the incubation temperature was reduced to 30°C and expression of the fusion protein was induced with 0.1 mM IPTG for 15 minutes. Cells were collected by centrifugation for 30 min at 6000 x g, supernatant was discarded and pellets were frozen overnight. Cell pellets were resuspended 10X in Lysis Buffer containing 25 mM Tris-HCl, pH 7.4, 250 mM NaCl, 8 M Urea. A final volume of 3 ml was sonicated for 5 minutes total process time (30s on, 30s off) using Misonix S-4000 (Misonix Inc.) with the amplitude set to 55%. Cell debris was removed by centrifugation for 30 min at 6000 x g and the supernatant containing the fusion protein was collected for further analysis.

Supernatant was dialyzed to Dockerin Reaction Buffer (25 mM Tris-HCl, pH 7.4, 50 mM NaCl, 1 mM CaCl2, 1 mM DTT, 0.1% Tween 20). The sample was centrifuged for 30 min at 6000g to remove precipitates formed during dialysis and pellet was discarded. Supernatant containing the SNAP-XDocII fusion protein in Dockerin Reaction Buffer was used in all subsequent labeling experiments.

Expression of the SNAP-XDocII fusion protein was optimized to include a short induction period of 15 minutes at 30°C. Protocols for recovery of the SNAP-XDocII fusion protein were

adapted from Adams et al. 2004 (Adams et al., 2004). Under these conditions, the soluble SNAP-XDocII fusion protein was recovered at a final concentration of 2.5 mM. The SNAP-XDocII fusion protein exhibited covalent binding to the SNAP fluorophore, as determined by SDS-PAGE analysis. Optimized parameters for labeling the fusion protein with SNAP fluorophore resulted in complete labeling of the fusion protein, with unbound fluorophore remaining in solution at less than 50% of the concentration of the fusion protein.

Fluorescent labeling of the SNAP-XDocII fusion protein

Fusion proteins for flow cytometry and microscopy were labeled with SNAP-Surface® Alexa Fluor® 647 and SNAP-Cell® 505 fluorescent dyes (New England BioLabs) by incubation of 2.5 mM fluorescent dye with fusion protein at 37°C for 1 hour. The resulting fluorescent proteins are referred to as 505-SNAP-XDocII and 647-SNAP-XDocII (Supplemental figure 1). Before incubation with *C. thermocellum*, the labeling reaction was centrifuged to remove non-fluorescent precipitates that formed during 37°C incubation. For fluorescent SDS-PAGE analysis, fusion protein was labeled with SNAP-Vista® Green according to the manufacturer's instructions (New England BioLabs).

Labeling *C. thermocellum* with fluorescent fusion protein

Volumes of *C. thermocellum* culture, grown to an OD_{600} of 0.5 were harvested by centrifugation for 2 min at 15,000 x g. Cell pellets were resuspended with an equal volume of 0.4 mM fusion protein in Dockerin Reaction Buffer and incubated for < 1 minute at room temperature (25°C) under aerobic conditions. Cells were washed three times with Dockerin Reaction Buffer. In negative control experiments, cells were labeled with mixtures containing purified SNAP-tag® protein (missing the XDocII fusion partner) or fluorescent dye (with no fusion protein) at 0.4 mM concentration under the same conditions. Varying the ratio of *C. thermocellum* cells to fluorescent fusion protein showed complete saturation at 0.83 pmol of fluorescent fusion protein per µl cells at an approximate 600 nm optical density of 0.5.

Microscopy

Microscopy was performed using a Nikon Optiphot-2 microscope. Fluorescent microscopy used a Prior Lumen 2000 for illumination set at 100%. A Nikon G-2A filter (EX 510-560, DM 575, EF 590) was used for visualizing SNAP-Cell® 505 fluorescence. A Chroma 49006 filter (EX 620, DM 660, EF 700) was used for visualizing SNAP-Surface® Alexa Fluor® 647. Images were captured using NIS-Elements Basic Research version 3.07 software Auto-Capture settings. Exposure time was kept constant for all images in a series.

Flow cytometry

Non-sorting flow cytometry experiments were performed using a Becton Dickinson 5-Color FacScan™. Flow cytometry sorting was performed using a Becton Dickinson FacsAria™. Data was collected using Becton Dickinson CellQuest™ software. Flow cytometry data was further analyzed using Flowing Software 2 (www.flowingsoftware.com). Graphs were prepared using Origin Labs Origin Pro 8.6 software.

SDS-PAGE

Samples were mixed with an equal volume of Novex 2X SDS Sample Buffer and incubated at 99°C for 5 minutes. 25 uL of sample was loaded into each well. Gels were 4–20% Mini-PROTEAN® TGX™ precast gels (Bio-Rad). SDS-PAGE gels were stained with SimplyBlue™

SafeStain (Invitrogen) according to the manufacturer's instructions. SDS-PAGE gels with samples labeled with SNAP-Vista® Green were visualized using 302 nm UV transillumination on a Bio-Rad XR+ system. Images were captured and analyzed with Quantity One version 4.6.9 software (Bio-Rad).

Results & Discussion:

In order to test the specificity of labeling type II cohesins with our 505-SNAP-XDocII protein, we attempted to label both *C. thermocellum* and *E. coli* cells. *C. thermocellum* cells were labeled by SNAP-XDocII, but not the *E. coli* cells, indicating that our protein binds specifically to *C. thermocellum* (Fig.1). Although fluorescent signals were observed in the labeling reactions containing *E. coli* cells, they did not correspond with the position of cells, as determined by phase contrast microscopy. Instead, they may represent aggregations of the SNAP-XDocII protein, since the XDocII module is known to form homodimers in solution (Adams et al., 2010).

The ability of SNAP-XDocII to bind to *C. thermocellum* suggests that type II cohesins are available for binding in the wild type strain. However, it was unclear whether this availability was due to a subpopulation of unoccupied anchor proteins or whether CipA was being displaced from occupied anchors. Therefore, we examined whether the SNAP-XDocII probe bound differentially to the *C. thermocellum* wild type strain and *ΔcipA*. Comparison of cells with and without *cipA* did not show any clear differences in fluorescent labeling (Fig. 1). In both cases some cells were labeled quite strongly and some cells were not labeled at all.

To focus on the effects of the removing the XDocII module, instead of the whole CipA protein, we extended our investigation to a strain where just the XDocII module of CipA had been deleted. Unfortunately, *cipA* contains extensive regions of DNA repeats (Gerngross et al., 1993), making genetic manipulation problematic. Therefore, the wild type allele of *cipA* was synthesized with extensive synonymous mutations, such that the regions of DNA identity were removed while maintaining the amino acid sequence. Two forms of this allele were created: *cipA** and *cipA*ΔxdocII* (*cipA** with the DocII module deleted).These alleles were used to replace the wild type *cipA* allele on the chromosome, resulting in *C. thermocellum* strains LL347 (*cipA**) and LL348 (*cipA*ΔXDocII*). These strains provide a more controlled platform for testing the role of the dockerin because they differ only by the presence or absence of the XDocII module. Similar to the comparison between wild type and *ΔcipA*, microscopy of strains *cipA** and *cipA*ΔXDocII* did not reveal any clear differences in fluorescent labeling (Fig. 1).

It is difficult to get quantitative data from microscopy experiments; therefore, the labeling intensity of the wild type and *ΔcipA* strains was measured by flow cytometry. Both strains displayed similarity in distribution of fluorescence intensity. The relative mean fluorescence intensity (RMFI) of wild type cells was 1,014 ± 40 (99% confidence interval) and the RMFI of *ΔcipA* cells was 1,011 ±44 (99% confidence interval).

Interestingly, microscopy revealed that the label was not evenly distributed along the length of the cell, but localized to specific regions including cell extremities and some cell-cell interfaces(Fig. 2). Cellulosome protuberances have been observed to protract and form fibrous corridors between cells and between cell and substrate under certain conditions (Bayer and Lamed, 1986). The size and shape of the labeled regions is similar to that of polycellulosomal

protuberances (Lemaire et al., 1995), although it is notable that most cells contain dozens of polycellulosomes but fewer labeled regions.

Next, the specificity of the labeling was quantified by flow cytometry. We attempted to label *C. thermocellum* cells with SNAP-XDocII protein and SNAP protein missing the XDocII module. Labeling cells with the SNAP protein missing the XDocII module did not result in labeling of *C. thermocellum* cells, indicating that binding was mediated by the XDocII module, as expected (Fig. 3). In the absence of the fluorophore, the SNAP protein or the XDocII module, a mean fluorescence intensity of ~10 was observed. In the presence of all 3 components, a mean fluorescence intensity of ~1000 was observed, further indicating the binding specificity of the SNAP-XDocII protein and ruling out possible background fluorescence from components of the *E. coli* lysate, non-specific binding of the SNAP protein or SNAP-fluorophore. Taken together with the subcellular localization of the labeling, these data indicate that the SNAP-XDocII fusion protein fluorescently labeled *C. thermocellum* via the cohesin-dockerin interaction.

Three mechanisms could explain why the presence of native CipA protein did not affect fluorescent labeling intensity. First, a significant excess of type II cohesins in proportion to CipA could mask the differences in cohesin availability between wild type and *ΔcipA*. Indeed, transcript and proteomic analyses have suggested that *C. thermocellum* has an excess of type II cohesin modules at the cell surface in relation to the number of CipA scaffoldins (Dror et al., 2003; Raman et al., 2009). A second possibility is that levels of cohesin-containing proteins were different in wild type and *ΔcipA*. A third possibility is that SNAP-XDocII fusion proteins could displace native CipA proteins in the wild type by competitive dockerin-replacement, masking the differences in cohesin availability between wild type and *ΔcipA*. We refer to this third possibility as the "dockerin-replacement" hypothesis.

To investigate the possibility of dockerin-replacement, wild-type *C. thermocellum* cells were subjected to sequential incubations in the presence of SNAP-XDocII fusion protein bound to different fluorophores. The fluorescent intensity of the labeled cells was analyzed by flow cytometry. The relative mean fluorescent intensity (RMFI) of the population was normalized to 1.00 based on the single labeling reaction, using either the SNAP-Cell 505 or 674 fluorophores. After labeling the cells with SNAP-Cell 505, a second labeling reaction was performed with the SNAP-Cell 647 fluorophore. The RMFI of the SNAP-Cell 647 label was 1.63 and the RMFI of the SNAP-Cell 505 label had decreased to 0.67. A third labeling reaction (with the same SNAP-Cell 505 label used in the first labeling reaction) resulted in an increase in the RMFI of the SNAP-Cell 505 label to 1.46 and a decrease of the RMFI of the SNAP-Cell 647 label to 0.73. Each additional label substantially decreased the intensity of the previous label (Fig. 4), indicating that the SNAP-XDocII proteins were capable of displacing each other, and supporting a role for the dockerin-replacement hypothesis. It is interesting that subsequent labeling reactions increased the fluorescence intensity of supposedly saturated samples (RMFI values > 1 in Fig. 4). One possible explanation is that cellulosomal protuberances may prolapse during the washing procedure exposing additional unbound cohesins that were not accessible to the SNAP-XDocII probe during the initial reaction.

In all samples, 30-70% of the flow cytometry "events" did not display any fluorescence, which is in agreement with the proportion of cells that did not display fluorescence in the microscopy experiment. In multiple labeling experiments, however, this value changed by less than 2 percentage points between labeling reactions, suggesting that the unlabeled populations

are stable. The results do not rule out the possibility of the other hypotheses. Resolving which mechanism is predominant remains an unresolved question. However, dockerin replacement may explain the surprising result that cells with and without the *cipA* gene showed similar levels of fluorescence after labeling with the SNAP-XDocII fusion protein, since the necessity of displacing CipA protein in the wild type and *cipA** strains did not reduce fluorescence intensity.

We have shown that the SNAP-tag system can be used to fluorescently label *C. thermocellum* via the cohesin-dockerin interaction. Previous studies have visualized cellulosomes by transmission electron microscopy (Bayer et al., 1985), however the ability to specifically label the cellulosome in aqueous solution could lead to the ability to observe cellulosome operation in-vivo.

Although much is known about the interaction between free dockerins and free cohesins, the interaction between free dockerins and bound cohesin-dockerin pairs has been less well-studied. Dockerin exchange suggests a mechanism for compositional change of the cellulosome. *C. thermocellum* is known to release cellulosomes in the late-stationary phase of growth, as well as optimize the composition of cellulosomes attached to its surface in response to substrate changes (Bayer & Lamed, 1986; Raman et al., 2009). It has been suggested that detachment of intact cellulosomes in these processes is achieved by proteolytic cleavage of the cohesin-II containing anchor proteins (Raman et al., 2009). The results of this study suggest an alternate or complementary mechanism, wherein the mere production of CipA molecules can effect turnover by dockerin-exchange. Similar experiments could be used to probe interactions between type I cohesins and dockerins.

In this study we have demonstrated displacement of bound dockerin-containing proteins with free dockerin-containing proteins. This result sheds light on a possible mechanism for the natural turnover and reordering of cellulosome subunits within the polycellulosome. Furthermore, the methods of this paper have established the SNAP-tag system as a valuable tool for labeling components and sub-components of the cellulosome.

Acknowledgments:

The authors would like to thank Gary Ward for assistance with flow cytometry studies and Kenneth Orndorff for microscopy research. This research was supported by 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 and a Dartmouth College Dean of Faculty Undergraduate Research Grant. We would like to declare one competing interest. Lee R. Lynd is a founder of the Mascoma Corporation which has a financial interest in *Clostridium thermocellum*.

References:

Adams JJ, Currie MA, Ali S, Bayer EA, Jia Z, & Smith SP (2010) Insights into higher-order organization of the cellulosome revealed by a dissect-and-build approach: crystal structure of interacting Clostridium thermocellum multimodular components. *J Mol Biol* **396**:833-839.

Adams JJ, Jang CJ, Spencer HL, Elliott M, & Smith SP (2004) Expression, purification and structural characterization of the scaffolding hydrophilic X-module from the cellulosome of Clostridium thermocellum. *Protein Expres Purif* **38:** 258-263.

Argyros DA, Tripathi SA, Barrett TF, et al. (2011) High ethanol titers from cellulose using metabolically engineered thermophilic, anaerobic microbes. *Appl Environ Microbiol* 10.1128/AEM.00646-11.

Bayer EA & Lamed R (1986) Ultrastructure of the cell surface cellulosome of *Clostridium thermocellum* and its interaction with cellulose. *J Bacteriol* **167:** 828–836.

Bayer E.A, Setter E, & Lamed R (1985) Organization and distribution of the cellulosome in *Clostridium thermocellum*. *J Bacteriol* **163:**552–559.

Bayer EA, Lamed R, White BA, & Flint HJ (2008) From cellulosomes to cellulosomics. *Chem Rec* **8:** 364–377.

Beguin P, & Aubert JP (1994) The biological degradation of cellulose. *FEMS Microbiol Rev* **13:** 25–58.

Demain AL, Newcomb M, & Wu JHD (2005) Cellulase, clostridia, and ethanol. *Microbiol Mol Biol R* **69:** 124–154.

Dror T, Rolider A, Bayer EA, Lamed R, & Shoham Y (2003) Regulation of expression of scaffoldin-related genes in *Clostridium thermocelllum*. *J Bacteriol* **185:** 5109-5116.

Gerngross UT, Romaniec MPM, Kobayashi T, Huskisson NS & Demain AL (1993) Sequencing of a *Clostridium thermocellum* Gene (*cipA*) Encoding the Cellulosomal S(L) Protein Reveals an Unusual Degree of Internal Homology. *Mol Microbiol* **8**: 325-334.

Gold ND, & Martin VJJ (2007) Global view of the *Clostridium thermocellum* cellulosome revealed by quantitative proteomic analysis. *J Bacteriol* **189:** 6787–6795.

Guss A, Olson D, Caiazza N & Lynd L (2012) Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. *Biotechnology for Biofuels* **5**: 30.Krauss J, Zverlov VV, & Schwarz WH (2012) *In vitro* reconstitution of the complete *Clostridium thermocellum* cellulosome and synergistic activity on crystalline cellulose. *Appl Environ Microbiol* 10.1128/AEM.07959-11.

Keppler A, Genreeizig S, Gronemeyer T, Pick H, Vogel H & Johnsson K (2003) A general method for the covalent labeling of fusion proteins with small mocules *in vivo*. *Nat Biotechnol* **21**:86-89.

Krauss, J., V. V. Zverlov, et al. (2012). "In Vitro Reconstitution of the Complete Clostridium thermocellum Cellulosome and Synergistic Activity on Crystalline Cellulose." Applied and Environmental Microbiology **78**(12): 4301-4307.

Lemaire M, Ohayon H, Gounon P, Fujino T, & Beguin P (1995) OlpB, a New Outer Layer Protein of *Clostridium thermocellum*, and Binding of Its S-Layer-Like Domains to Components of the Cell-Envelope. *J Bacteriol* **177:**2451-2459.

Lynd LR, Weimer PJ, van Zyl WH, & Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol R* **66:** 506-577.

Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL, & Shoham Y (2010) *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. *PNAS* **107:** 18646–18651.

Olson DG & Lynd LR (2012) Chapter seventeen - transformation of *Clostridium thermocellum* by electroporation. *Methods Enzymol, Vol. 510* (Gilbert HJ, ed.), pp. 317-330. Academic Press, Waltham, MA.

McCarthy AJ (1989) Thermomonospora. *Bergey's Manual of Systematic Bacteriology, Vol. 4*(Williams ST, Sharpe ME & Holt JG, eds), pp. 2552-2572. Williams and Wilkins, Baltimore, MD.

Mechaly, A., H. P. Fierobe, et al. (2001). "Cohesin-dockerin interaction in cellulosome assembly - A single hydroxyl group of a dockerin domain distinguishes between nonrecognition and high affinity recognition." Journal of Biological Chemistry **276**(13): 9883-9888.

Olson DG, Tripathi SA, Giannone RJ, Lo J, Caiazza NC, Hogsett DA, Hettich RL, Guss AM, Dubrovsky G,& Lynd LR (2010) Deletion of the Cel48S cellulase from *Clostridium thermocellum*. *Proc Natl Acad Sci U S A* **107:** 17727-17732.

Raman B, Pan C, Hurst GB, Rodriguez Jr. M, McKeown CK, Lankford PK, Samatova NF, & Mielenz JR (2009) Impact of pretreated switchgrass and biomass carbohydrates on *Clostridium thermocellum* ATCC 27405 cellulosome composition: a quantitative proteomic analysis. *PLoS ONE* **4:** e5271.

Raman B, McKeown CK, Rodriguez Jr. M, Brown SD, & Mielenz JR. (2011) Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. *BMC Microbiol* **11:** 134.

Table 1: Strains

Strain	Genotype	Description/source
DSM1313	Wild type	from DSMZ
M1354	Δ*hpt*	from Argyros et al. 2011
DS16	Δ*hpt* Δ*cipA*	M1354 with deletion of *cipA* ORF and upstream promoter region
LL347	Δ*hpt cipA**::*cipA*	M1354 with allelic replacement of *cipA* with *cipA**
LL348	Δ*hpt cipA**Δ*xdocII*::*cipA*	M1354 with allelic replacement of *cipA* with *cipA** missing the *xdocII* module
Eco_SNAP	T7 Express *lysY*/Iq pDGO-54	T7 Express *lysY*/Iq *E. coli* from NEB with the pDGO-54 plasmid

Table 2: Plasmids

Plasmid name	Genbank accession number	Description
pDGO-03	JX489218	CipA deletion vector used for making strain DS11
pDGO-34	JX489219	CipA deletion vector used for making strain DS16
pAMG270	JX477172	Allelic replacement vector used for making strain LL347
pAMG269	JX477171	Allelic replacement vector used for making strain LL348
pDGO-54	JX500710	*E. coli* SNAP-XDocII expression vector. The type II dockerin from *C. thermocellum* (including the X-module required for solubility) were inserted downstream of and in frame with the SNAP-tag protein from the pSNAP-tag (T7)-2 Vector from New England Biolabs.

Figure 1: Fluorescence microscopy of *E. coli* and *C. thermocellum* cells labeled with 505-SNAP-XDocII fusion protein. Images of cells after incubation with 505-SNAP-XDocII fusion protein. Column 1, phase contrast; column 2, fluorescence; column 3, merged image of the phase contrast and fluorescence with the fluorescence layer false-colored red. Bars indicate 10 μm.

Figure 2: Fluorescence microscopy of *C. thermocellum* cells labeled with 505-SNAP-XDocII fusion protein, adjusted for increased contrast. Fluorescent images of *C. thermocellum cipA*ΔXDocII* cells after incubation with 505-SNAP-XDocII fusion protein. Cells were chosen which displayed little movement and high contrast of fluorescent labeling on their cell surface. **(A)** Phase contrast image, **(B)** Fluorescent image of the same area, **(C)** Digital zoom of fluorescent image in (B). Arrows denote localization of fluorescence. Bars indicate 2 μm.

Figure 3: Relative mean fluorescence intensity of wild type cell populations labeled with 505-SNAP-XDocII protein as determined by flow cytometry. Wild type *C. thermocellum* cells were incubated with various preparations containing different components of the SNAP-XDocII protein labeling scheme. Bars indicate the mean fluorescence, at 505 nm, of wild type cell populations analyzed by flow cytometry, measured in relative units of fluorescence (RFUs). Error bars indicate 99.9% confidence intervals. "*E. coli* lysate" indicates lysate from *E. coli* BL21(DE3) cells and was used to control for non-specific fluorescence from the lysate. "505-SNAP" indicates purified SNAP protein bound to the SNAP-Cell 505 fluorophore, and was used to control for non-specific binding of the SNAP protein to *C. thermocellum*. "E. coli lysate with 505-SNAP-XDocII" indicates a lysate of *E. coli* BL21(DE3) cells where expression of the SNAP-XDocII protein had been induced. Subsequently, the lysate was labeled with SNAP-Cell 505 fluorophore (to generate 505-SNAP-XDocII) before being added to *C. thermocellum* cells.

Figure 4: Relative mean fluorescence intensity of wild type cell populations sequentially labeled with 505-SNAP-XDocII or 647-SNAP-XDocII fusion protein, as determined by flow cytometry. Wild type *C. thermocellum* cells were sequentially labeled two and three times with alternating incubations of either 505 nm fluorescently-labeled or 647 nm fluorescently-labeled fusion proteins. Bars indicate the mean fluorescence of wild type cell populations determined by flow cytometry. RMFIs are scaled units of fluorescence. RMFI of negative controls (unlabeled wild type *C. thermocellum* cells) was scaled to zero. RMFI of singly-labeled positive controls was scaled to 1. Error bars denote 99.9% confidence intervals.

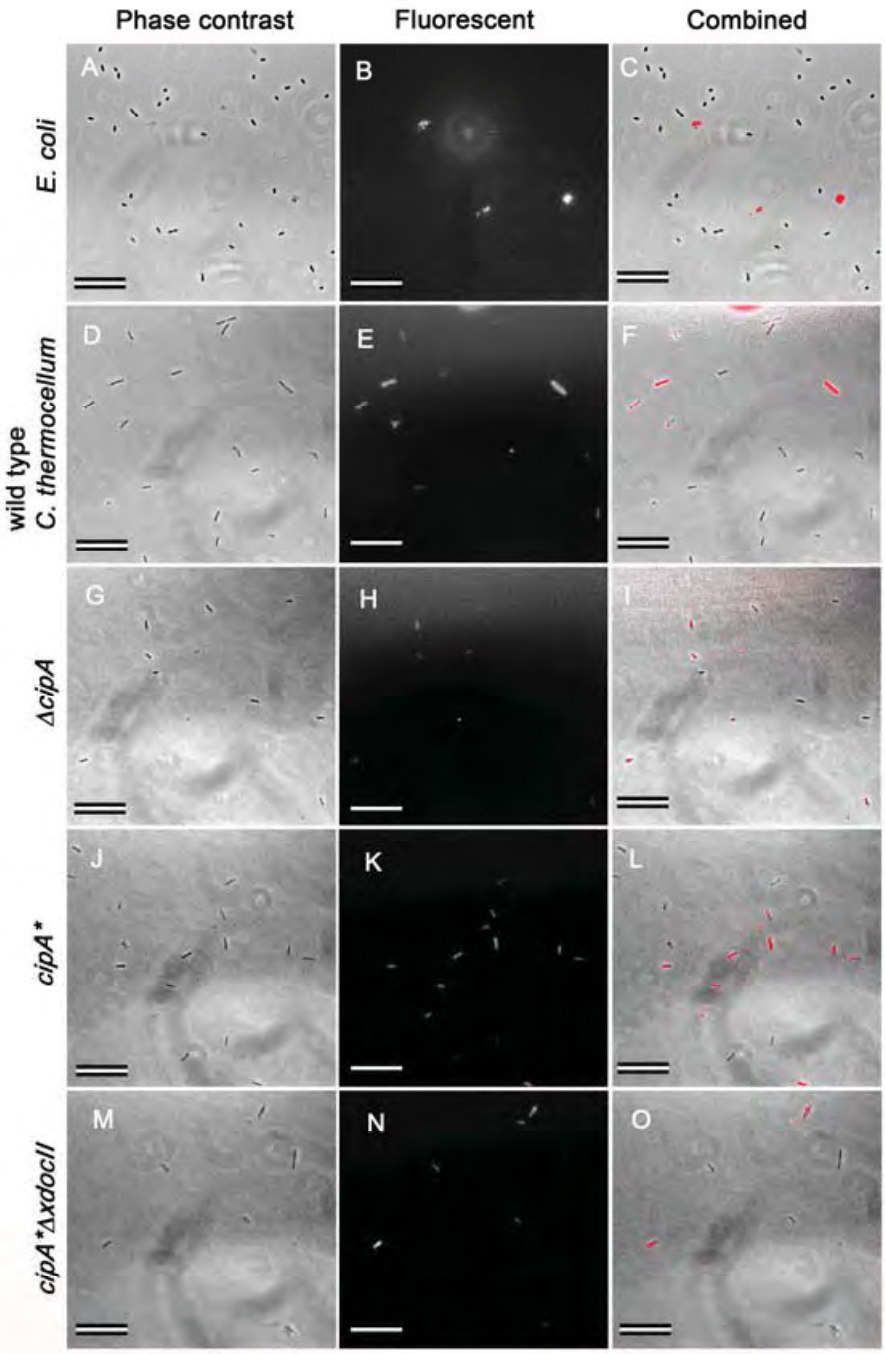

Fig. 1

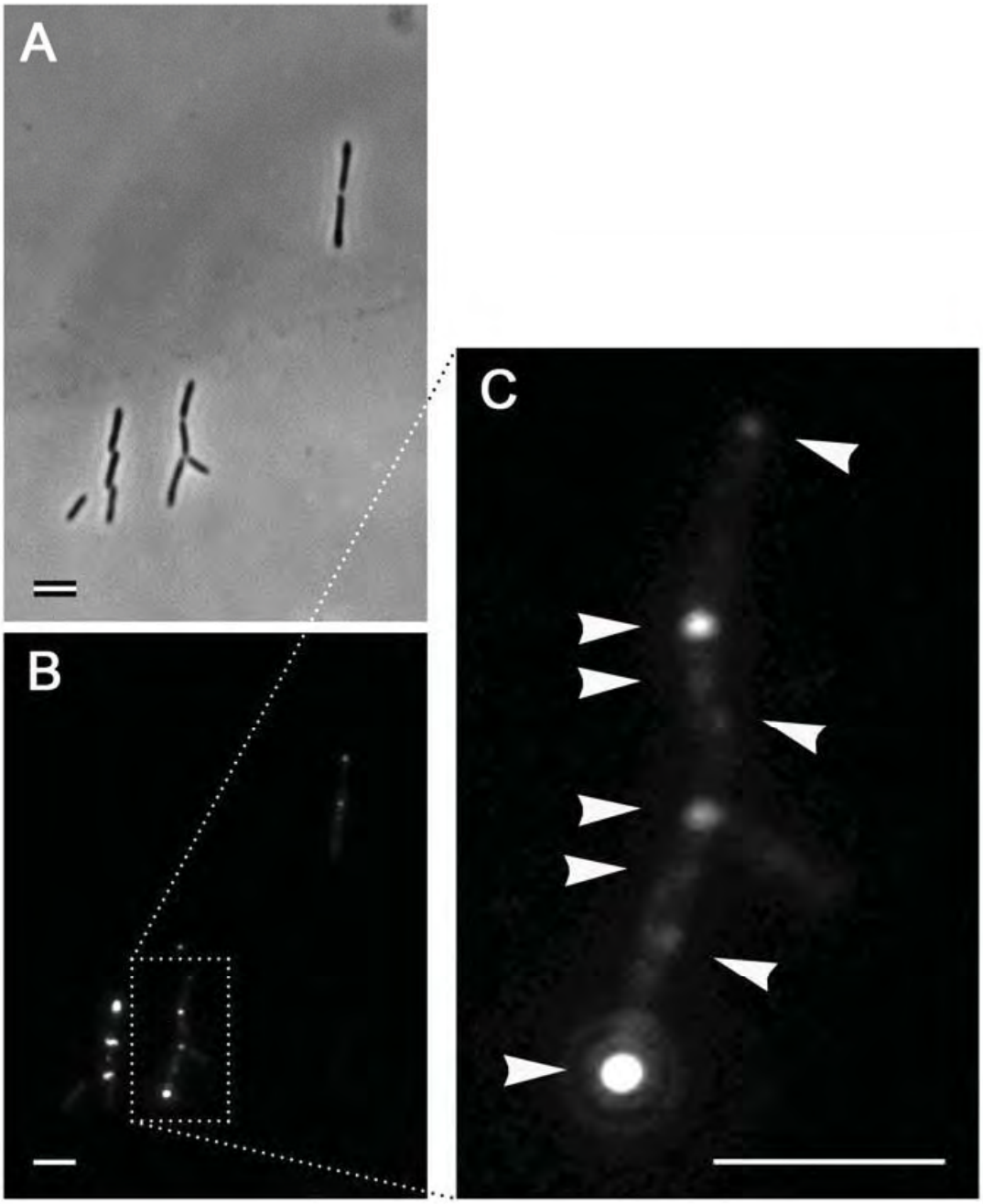

Fig. 2

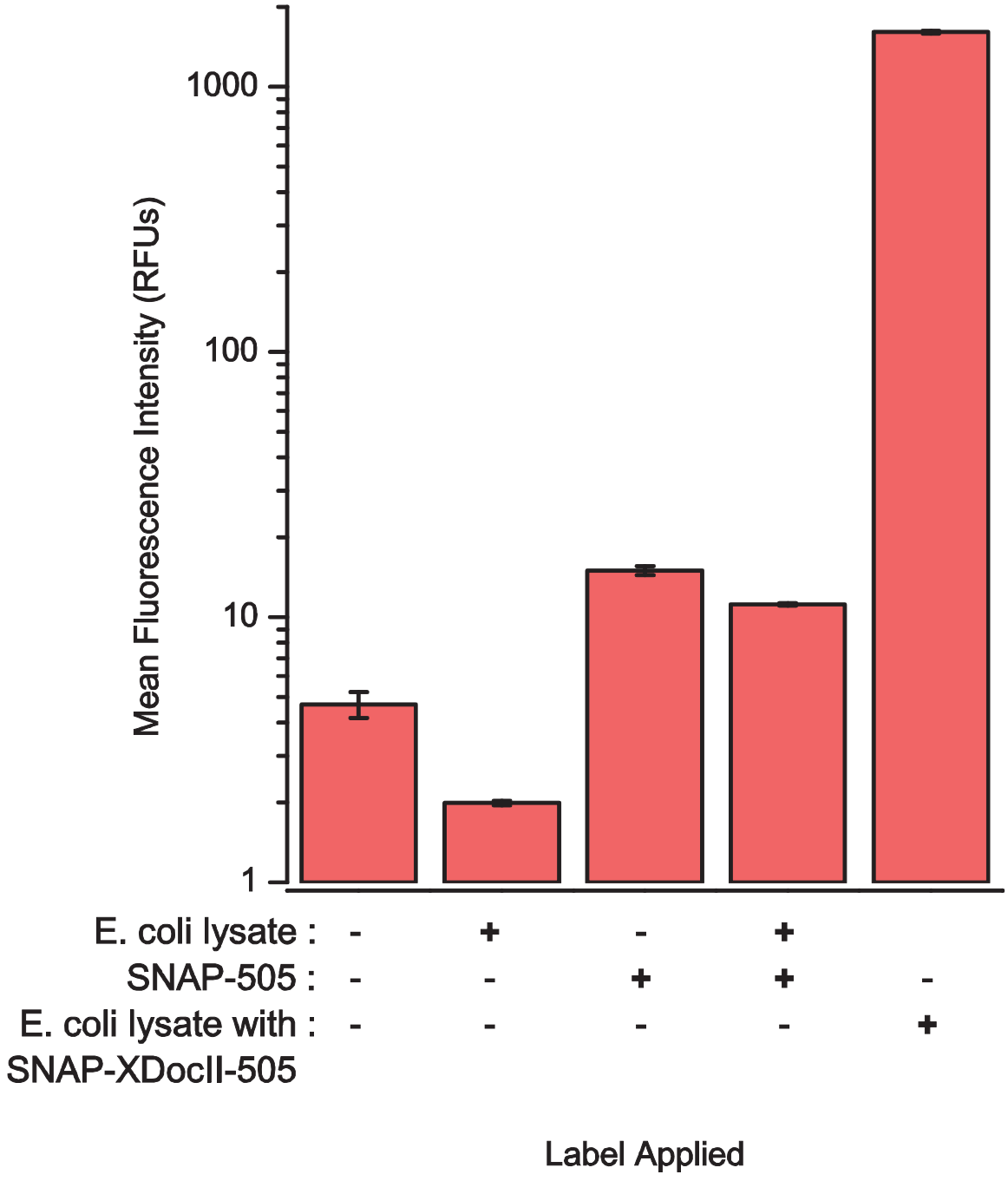

Fig. 3

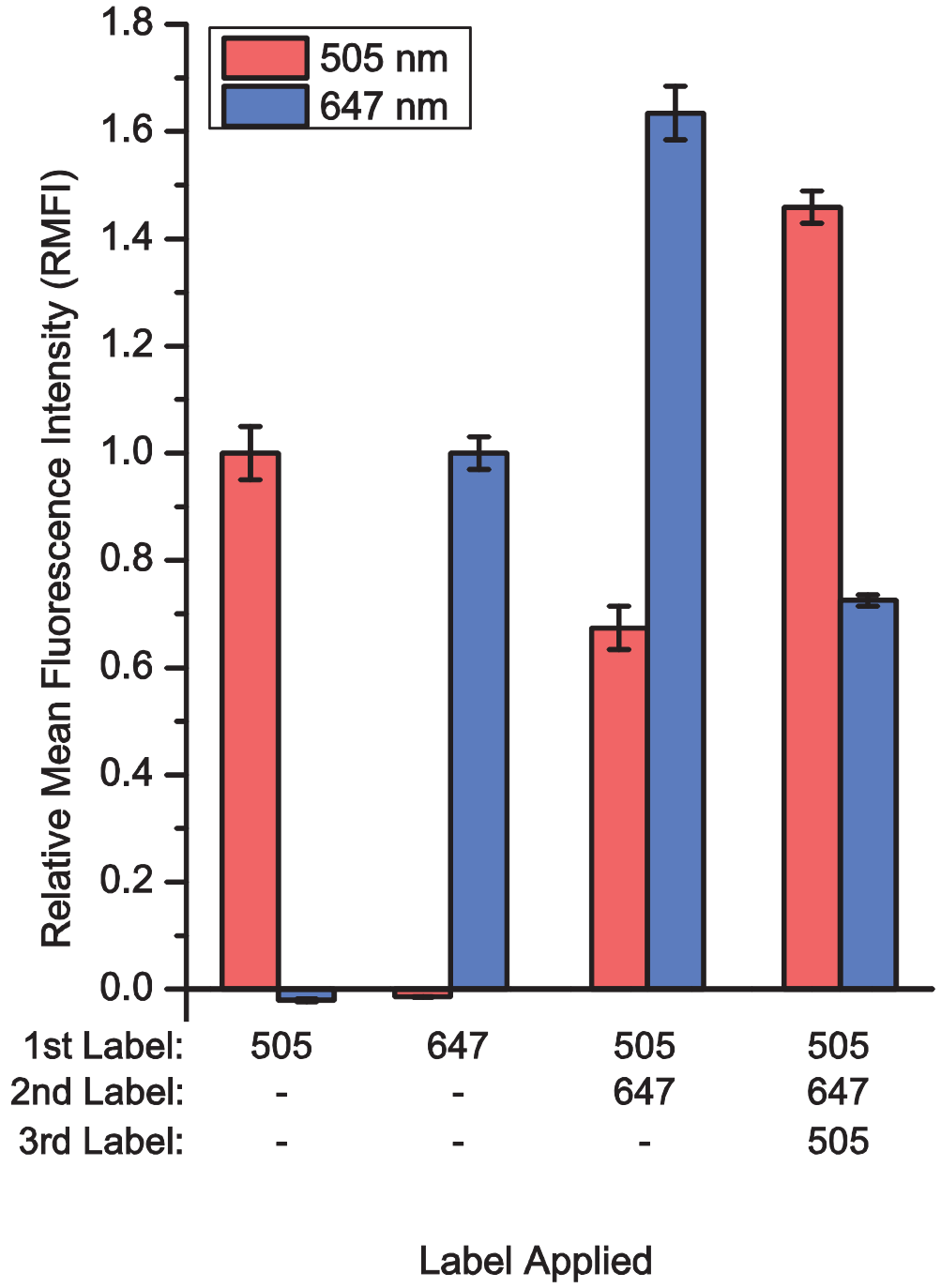

Fig. 4