Cell-Surface-Anchoring Role of N-Terminal Surface Layer Homology Domains of Clostridium cellulovorans EngE

Akihiko Kosugi, Koichiro Murashima, Yutaka Tamaru, and Roy H. Doi

Section of Molecular and Cellular Biology, University of California Davis, Davis, California 95616, and Department of Life Science, Mie University, Tsu 514-8507, Japan

Received 27 August 2001/Accepted 13 November 2001

engE, coding for endoglucanase E, one of the three major subunits of the Clostridium cellulovorans cellulosome, has been cloned and sequenced (Y. Tamaru and R. H. Doi, J. Bacteriol. 181:3270–3276, 1999). The N-terminal-half region of EngE possesses three repeated surface layer homology (SLH) domains, which are homologous to those of some bacterial S-layer proteins. Also, the C-terminal-half region consists of a catalytic domain of glycosyl hydrolase family 5 and a duplicated sequence (dockerin) for binding EngE to scaffolding protein CbpA. Our hypothesis is that the SLH domains serve in the role of anchoring to the cell surface. This model was investigated by using recombinant EngEs (rEngE) with and without SLH domains that were synthesized in Escherichia coli and cell wall preparations from C. cellulovorans. When rEngE and SLH polypeptides of EngE were incubated with cell wall fragments prepared by sodium dodecyl sulfate treatment, these proteins bound strongly to the cell wall. However, rEngEs without SLH domains lost their ability to bind to cell walls. When rEngE was incubated with mini-CbpA, consisting of two cohesin domains, and cell wall fragments, the mini-CbpA was able to bind to the cell wall with rEngE. However, the binding of mini-CbpA was dramatically inhibited by addition of a chelating reagent, such as EDTA, which prevents cohesin-dockerin interactions. These results suggest not only that the SLH domains of EngE can bind to the cell surface but also that EngE plays an anchoring role for cellulosomes through the interaction of its dockerin domain with a CbpA cohesin.

Clostridium cellulovorans is a mesophilic, anaerobic, spore-forming bacterium which stains as gram negative and which utilizes not only cellulose but also xylan, pectin, and several other carbon sources (5, 27). C. cellulovorans produces an extracellular cellulolytic multi-enzyme complex, which has been called the cellulosome (1, 5), that has a total molecular weight of about 1,000,000 and that is capable of hydrolyzing crystalline cellulose (5). The cellulosome of C. cellulovorans comprises three major subunits, scaffolding protein CbpA (26), exoglucanase ExgG (14), and endoglucanase EngE (29). In addition, the cellulosome contains several other enzyme subunits. We recently cloned and sequenced several cellulosomal subunits (5). One of these subunits, EngE, had a unique structure on the basis of the derived amino acid sequence (29), i.e., the N-terminal-half region of EngE possessed three repeated-sequence domains which were partially homologous to those of bacterial S-layer proteins. Also, the C-terminal-half region consisted of a catalytic domain of glycosyl hydrolase family 5 and a duplicated sequence (dockerin) for binding to scaffolding protein CbpA.

Specifically, the bacterial surface layer homology (SLH) domains appeared to have the important role of mediating binding between extracellular proteins and the cell surface. Several results showing cell wall association of SLH domains in vivo and in vitro were reported recently; e.g., a mutant having an SLH-defective protein of the S-layer (SlpA) in Thermus thermophilus lost its ability to bind to cell wall components such as peptidoglycan (18). Lemaire et al. reported that recombinant proteins containing SLH domains of C. thermocellum outer layer protein OlpB and S-layer protein SlpA could be bound to cell wall preparations (12, 13). Therefore, the interesting SLH domain structure of EngE suggests not only that EngE binds to scaffolding protein CbpA through its dockerin but also that it may bind to the cell surface through its SLH domains. To investigate this hypothesis, we carried out binding assays using recombinant EngEs (rEngEs) and mini-CbpA, consisting of a cellulose binding domain, one hydrophilic domain, and two cohesin domains, and two cohesin domains of CbpA, synthesized in Escherichia coli, and cell wall preparations from C. cellulovorans.

In this paper, we provide evidence that the SLH domains of EngE are necessary to anchor EngE to the C. cellulovorans cell surface and are able to anchor mini-CbpA to the cell surface through its cohesin-dockerin interaction. This is the first report that indicates that a mesophilic cellulosomal enzyme may be able to anchor the cellulosome to the cell surface.

MATERIALS AND METHODS

Bacterial strains and media. C. cellulovorans ATCC 35296 (27) was used as the source of chromosomal DNA and cell wall preparations. C. cellulovorans was grown anaerobically at 37°C in serum bottles containing a previously described medium (27) which included 5 g of cellobiose (Sigma) per liter. E. coli Novablue and BL21(DE3) (Novagen) were used as cloning hosts for production of recombinant proteins and were grown at 30 or 37°C in Luria-Bertani medium containing ampicillin (100 μg/ml, Sigma).

Construction of pENG, pSLH, pDSLH, and pEMCBP. Four fragments containing 3,014, 1,470, 1,629, and 1,665 bp were amplified by PCR to create expression plasmids pENG, pSLH, pDSLH, and pEMCBP, respectively. The forward and reverse primers for pENG, pSLH, pDSLH, and pEMCBP were designed to carry artificial restriction enzyme sites of NcoI and XhoI, respectively.

Copyright © 2002, American Society for Microbiology. All Rights Reserved.
Expression and purification of the recombinant proteins. Four recombinant proteins were purified from each of the *E. coli* BL21 strains harboring expression plasmids. When the cultures reached an optical density at 600 nm of 0.5, isolated from each of the C-terminal end with the six-histidine tag from the vector. This expression plasmids were designed to allow fusion proteins at the N-terminal end with the PelA signal peptide and at the C-terminal end with the six-histidine tag from the vector.

Preparation of cell wall fragments of *C. cellulovorans*. *C. cellulovorans* cells from 500 ml of a mid-exponential-phase culture were harvested by centrifugation at 12,100 × g for 10 min and were washed twice with 100 ml of 50 mM phosphate buffer (pH 7.0) and resuspended with 10 ml of the same buffer. An aliquot of the cell suspension was labeled as the whole-cell fraction (fraction F1). The cell suspension was disrupted by sonication, intact cells were removed by twice centrifugation at 1,940 × g for 5 min, and the suspensions were recentlyrifuged at 39,200 × g for 20 min. The resulting supernatant was labeled as the cell extract fraction (fraction F2), and the pellet, which was suspended in 5 ml of 50 mM phosphate buffer (pH 7.0), contained the crude cell wall fraction (fraction F3). Fraction F3 was treated with 1% sodium dodecyl sulfate (SDS) by heating at 100°C in a water bath for 20 min and was centrifuged at 16,000 × g for 20 min at room temperature. The supernatant consisted of the cell wall-associated proteins (fraction F4). The pellet was resuspended in 5 ml of 50 mM phosphate buffer (pH 7.0) after being washed three times with the phosphate buffer. This suspension consisted of cell wall fragments containing peptidoglycan and polysaccharide (fraction F5). To remove covalently bound cell wall polymers from the peptidoglycan layer, the method of Ries et al. (21) was used. Fraction F5 was treated with 48% hydrofluoric acid (HF) for 24 h at 4°C. After centrifugation at 16,000 × g for 20 min at 4°C, the pellet was washed three times with 50 mM phosphate buffer (pH 7.0).

Interaction of each recombinant protein with cell wall fragments of *C. cellulovorans*. Binding experiments were carried out by incubating and coincubating each of the recombinant proteins with cell wall fragments (12). Each polypeptide (40 to 60 μg) was mixed with 100 μl of fraction F5 (75.1 μM diaminopimelate [DAPA]), and the reaction volume was brought to 300 μl with 50 mM phosphate buffer (pH 7.0). The reaction mixtures were incubated for 4 h at 30°C with shaking. For the binding test for mini-CbpA, after mini-CbpA and rEngE were incubated for 1 h at 30°C, the cell wall fragments were added to the reaction mixture. The bound and free polypeptides were separated by centrifugation at 16,000 × g for 20 min at room temperature. The supernatant consisted of the soluble fraction. The bound and free polypeptides were separated by centrifugation at 16,000 × g for 20 min at room temperature. The supernatant consisted of the soluble fraction. A wash fraction was obtained after washing the cell wall with 50 mM phosphate buffer (pH 7.0). The pellets, consisting of the insoluble cell wall fragments and attached proteins, were washed with the same buffer and then resuspended with 300 μl of the phosphate buffer. Each fraction was analyzed by SDS-polyacrylamide gel electrophoresis (PAGE).

Analysis of the peptidoglycan of *C. cellulovorans*. To determine the amino acids and the DAPA content in peptidoglycan, fraction F5 was analyzed. Fraction F5 was vacuum dried and hydrolyzed in 6 N HCl for 24 h at 110°C. Aliquots of the hydrolysate were analyzed by a Beckman 6300 amino acid analyzer.

**RESULTS AND DISCUSSION**

Domain structure of EngE. Analysis of the deduced amino acid sequence of EngE revealed a multidomain structure with a unique feature, i.e., a triplicated SLH domain at the N-terminal followed, by a family 5 catalytic domain, a domain of unknown function, and a duplicated sequence (donor). The N-terminal half of EngE (residues 32 to 491) contained three highly conserved repeats of 163, 126, and 163 amino acids (29). Moreover, the latter half of the repeated region (residues 304 to 426) of EngE had high similarity to the N-terminal domain (residues 1 to 154) of S-layer protein RsaA from *Caulobacter crescentus* (21.1% identity and 69.5% similarity to the N-terminal region of RsaA). The N-terminal domain of RsaA was reported to be involved in cell surface anchoring (2, 29). Also, the N-terminal region of EngE (residues 63 to 272) possessed high similarity to the C-terminal region of RsaA (residues 784 to 907), which is involved in a secretion signal (2, 29). On the other hand, a comparison of the SLH domain of EngE with the SLH domains of several structural proteins from *Clostridium thermocellum* outer layer proteins OlpB, SlpA, and SdbA (7, 11, 13) and *Bacillus anthracis* outer layer proteins (21) showed no significant homology. It was reported that SLH binding sites in different bacterial species are different and that these sites are not unique within the same species (4). We think that the homology between SLH domains may depend on the bacterial species and surface layer structures. Therefore, to determine whether or not the SLH domains of EngE have functions similar to those of RsaA, i.e., the functions of anchoring and targeting to the cell surface, we decided to test the binding ability of the SLH domains of EngE for *C. cellulovorans* cell wall preparation.

Localization of EngE in *C. cellulovorans*. To identify the distribution of EngE in *C. cellulovorans*, we carried out immunoblot analyses using antiserum against EngE on prepared fractions. When fraction F3 was treated with SDS, many proteins associated with the cell wall were solubilized by the SDS (Fig. 1A). Such a treatment is able to remove the proteins that are noncovalently associated with the cell surface (16, 24). Thus, several proteins belonging to the *C. cellulovorans* cell envelope or present on the cell surface could be identified in fractions F3 and F4. On the other hand, no proteins were detected after Coomassie blue staining of the material ex-
tration by treatment with SDS (Fig. 1A). Therefore, fraction F5 appears to consist of protein-free cell wall polymers containing peptidoglycan and polysaccharides. Furthermore, to know the proportion of components in the peptidoglycan, we analyzed the quantities of DAPA and amino acids in fraction F5. Alanine, glutamate, and DAPA were detected in hydrolysis products. The composition coincided with the peptidoglycan composition of clostridia (25). In immunoblot analysis using an EngE antibody, fractions F2 and F5 did not contain EngE; major bands, corresponding to 110 kDa, could be found in fractions F1, F3, and F4 (Fig. 1B). The signals were in good agreement with the molecular mass of EngE (29). These results indicated that EngE was associated with the C. cellulovorans cell envelope.

**Binding of the SLH domains of EngE to the C. cellulovorans cell wall.** It was known that the SLH domains of several bacterial surface layer proteins had the ability to mediate attachment of proteins to cell walls (24). In addition, it was proposed that the attachment of SLH domains to cell walls in *Bacillus stearothermophilus* PV72/p2 and *Thermoanaerobacterium thermosulfurigenes* EM1 was mediated by secondary cell wall polysaccharides (3, 21). To appraise the interaction of the SLH domain of EngE with *C. cellulovorans* cell walls, we constructed rEngE with or without the SLH domains in *E. coli*. rEngF32-1030 and rSLH32-515, containing EngE SLH domains, and rdSLH494-1030- lacking EngE SLH domains, were constructed as fusion proteins with a six-histidine tag at the C terminus (Fig. 2). After purification by Ni²⁺ affinity chromatography, these recombinant proteins were tested for their ability to bind to the cell wall fraction, F5. As shown in Fig. 3, although rSLH32-515 and rEngE32-1030 with SLH domains were able to bind to the cell wall fragments, rdSLH494-1030 without its SLH domains lost the ability to bind to cell walls. These results strongly indicated that the SLH domains of EngE performed the role of anchoring to the *C. cellulovorans* cell surface. In addition, when the cell wall fragments were treated with HF, the binding of rSLH32-515 and rEngE32-1030 was abolished (data not shown). Similar effects of HF have been reported (3, 4, 12). HF is able to release secondary polymers, such as material containing GlcNAc, ManNAc (N-acetylmannosamine), MurNAc (N-acetylmuramic acid), glucose, galactose, and DAPA, by cleavage of phosphodiester linkages from peptidoglycan layers (3, 21). Thus, the results suggest that accessory polymers, acting as the adhesion component in fraction F5, are needed for attachment of SLH domains of EngE to the *C. cellulovorans* cell surface.

Large glycohydrolase CAC3469, which has high homology to EngE (17), has been identified recently from the sequence analysis of the *Clostridium acetobutylicum* genome. Unlike EngE, CAC3469 has an additional cell adhesion domain, which corresponds to the unknown domain (residues 840 to 970) of EngE (17, 29). Moreover, rdSLH494-1030, included in the unknown domain, was able to bind slightly to fraction 5 (Fig. 3). Thus, it appears that the unknown domain of EngE could be contributing a little to binding to the cell envelope.

The SLH domains of EngE were also similar to four hydrophilic domains (HLDs) of scaffolding protein CbpA of *C. cellulosolvens* (5, 29). Scaffolding proteins in several cellulolytic clostridia, such as CipA of *C. thermocellum* (8), CipP of *Clostridium cellulolyticum* (20), and CipA of *Clostridium josui* (9), also possess unknown domains called X (1, 20), which are very similar to the HLDs (SHOs) of *C. cellulosolvens* CbpA (5). So far, the function of these HLD domains remains obscure. We have proposed that the four HLDs in scaffolding protein CbpA may also have a role in anchoring to the cell surface, especially since there is similarity between the HLDs of CbpA and the SLH domains of EngE (5, 29).

**Anchoring of mini-CbpA to the cell surface by SLH and...**
anchorin and dockerin domains of EngE. A duplicated sequence (dockerin; residues 978 to 1030) resides in the C-terminal region of EngE (29). The dockerins, consisting of 22-amino-acid repeats, are well conserved in cellulosomal subunits from C. cellulovorans and other Clostridium species (1, 5). It is also known that the interaction between cohesins of the scaffolding protein CbpA and dockerin domains of enzymatic subunits is necessary to form cellulosomes in C. cellulovorans (5, 28, 30). Thus, we proposed that, through an interaction between the dockerin of EngE and the cohesin of CbpA, the SLH domains of EngE could help CbpA attach to the cell surface (29) and in this process could bind the whole cellulosome to the cell surface.

To test this hypothesis, we constructed a mini-CbpA that consisted of a cellulose binding domain, one HLD, and two cohesin domains (Fig. 4A). When rEngE32-1030 was incubated with mini-CbpA and fraction F5, the mini-CbpA was able to bind to the cell wall with rEngE32-1030 (Fig. 4B, a and b). The weak binding of mini-CbpA by itself to the cell wall (Fig. 4B, b and c) may be due to the single HLD present in mini-CbpA. However, the binding ability was dramatically inhibited by the addition of chelating reagent EDTA, which prevents the dockerin-cohesin interaction (Fig. 4B, c). It has been known that the dockerin domain of cellulosomal subunits has a putative calcium binding motif and that calcium ions are necessary for the binding between dockerin and cohesin domains of C. cellulovorans, C. cellulolyticum, and C. thermocellum (5, 19, 30).

Therefore, these results indicated not only that the SLH domains of EngE can bind to the cell surface but also that EngE plays an anchoring role for cellulosomes through its interaction with the CbpA cohesins. The presence of EDTA influenced not only the cohesin-dockerin interaction but also indirectly the ability of rEngE32-1030 to bind to cell wall fragments (Fig. 4B, a and c).

Surface-associated proteins can be solubilized with agents that break hydrogen bonds (e.g., guanidine hydrochloride and urea) and with metal-chelating agents (e.g., EDTA and EGTA) (16, 24). When fraction F5 was treated with EDTA and subjected to the binding test, the cell wall binding ability of rEngE32-1030 was drastically decreased (data not shown). In Caulobacter crescentus, calcium ions were involved with the S-layer assembly process, which was disrupted by metal ion chelators, such as EGTA (31). Thus, the interaction between SLH domains of EngE and cell wall components may also require metal ions that are attached to the cell surface layer.

In C. thermocellum, several proteins mediating attachment to the cell surface were found from sequence analyses (7). The manner of anchoring cellulosomes to the cell surface has been known to involve several proteins, such as OlpB, ORF2p, and SdbA, which also possess repeated SLH domains and type II cohesin domains (7). Although the type I cohesin domain in OlpA was unable to bind the C-terminal type II dockerin domain of CipA, the type II cohesin domains of OlpB, ORF2p, and SdbA were necessary to bind to the type II dockerin of CipA (11, 22, 23). The mechanism for anchoring mini-CbpA to the cell surface by interactions between cohesin and dockerin domains and the SLH domains of EngE is similar to that for CipA and OlpB in C. thermocellum. On the other hand, so far, the mechanism for anchoring the cellulosome to the cell surface in mesophilic clostridia such as C. cellulovorans, C. cellulyticum, and C. josui has been unclear. Interestingly, mini-CbpAs were able to bind to cell wall fragments not only through the cohesin-dockerin interaction but also slightly by themselves (Fig. 4B, b). Therefore, C. cellulovorans may possess several systems for anchoring the cellulosomes to the cell surface: (i) through the SLHs of EngE; (ii) through the interaction by the HLDs (SLHs) in CbpA, and (iii) through anchoring proteins such as C. thermocellum OlpB and SdbA. Further studies will be required for us to clearly understand the mechanisms for anchoring cellulosomes to the C. cellulovorans cell surface.

In Fig. 5, we present an updated anchoring model for cell surface attachment of scaffolding protein CbpA through the EngE molecule. Note that EngE binds to CbpA through the cohesin-dockerin interaction and to the cell surface via its SLHs.
ACKNOWLEDGMENT
This research was supported in part by grant DE-DDF03-92ER20069 from the U.S. Department of Energy.

REFERENCES