# Cell-surface exposure of a hybrid 3-cohesin scaffoldin allowing the functionalization of *Escherichia coli* envelope.

**Nicolas Vita[1*], Romain Borne[1$] and Henri-Pierre Fierobe**

Aix-Marseille université, CNRS, LCB, Marseille, France.

[1]: NV and RB should be considered joint first author.

*: correspondence to this author: LCB-CNRS, UMR7283, 31, chemin Joseph Aiguier, F-13402, Marseille cedex 20, France; Email: nvita@imm.cnrs.fr; Tel: + 33 (0)4 91 16 41 43.

[$]: Present address: Laboratoire Chrono-environnement - UMR 6249 UFC/CNRS

16, route de Gray, F-25030 Besançon cedex, France

This research was supported by a grant from the « Agence Nationale de la Recherche » (Grant number: ANR-14-CE05-0019-01)

**Running Title:** *E. coli* surface exposure of a scaffoldin

**Abstract**

Cellulosomes are large plant cell wall degrading complexes secreted by some anaerobic bacteria. They are typically composed of a major scaffolding protein containing multiple receptors called cohesins which tightly anchor a small complementary module termed dockerin harbored by the cellulosomal enzymes. In the present study we have successfully cell-surface exposed in *Escherichia coli* a hybrid scaffoldin, Scaf6, fused to the curli protein CsgA, the latter is known to polymerize at the surface of *E. coli* to form extracellular fibres under stressful environmental conditions. The C-terminal part of the chimera encompasses the hybrid scaffoldin composed of three cohesins from different bacterial origins and a carbohydrate binding module (CBM) targeting insoluble cellulose. Using three cellulases hosting the complementary dockerin modules and labeled with different fluorophores, we have shown that the hybrid scaffoldin merged to CsgA is massively exposed at the cell surface of *E. coli* and that each cohesin module is fully operational. Altogether these data open a new route for a series of biotechnological applications exploiting the cell surface exposure of CsgA-Scaf6 in various industrial sectors such as vaccines, biocatalysts or bioremediation, simply by grafting the small dockerin module to the desired proteins prior to incubation with the engineered *E. coli*.

Key words: Cellulosome, Scaffoldin, CsgA, cell-surface exposure, *Escherichia coli*.

## INTRODUCTION

The design of microbial cell factories has been raising huge interest for years. It aims to create new and complex microorganisms by gathering of useful and simple phenotypes from diverse origins. This synthetic biology process allows to synthetize products of interest by constructing non-natural cascade reactions catalyzed by various enzymes. It is based on the combination of small components in a unique organism to generate new systems that should work better than their natural counterparts. First, an appropriate host has to be selected as "chassis" able to cope with the new introduced complex pathways or functions while remaining easy to handle.

*Escherichia coli* secretes only few endogenous proteins and is generally not considered as a good secreting microorganism of heterologous proteins as compared to Gram+ bacteria like *Bacillus subtilis* or yeasts. Cell surface display of desired proteins or antigens is also challenging in *E. coli*, though some exposure was obtained for specific heterologous proteins by fusion with the ice nucleation protein InaK from *Pseudomonas syringae* (Gu et al., 2013) the endogenous lipocalin Bcl (Tanaka et al., 2011), or the use of the autotransporter Ag43 (Ramesh et al., 2012). Nevertheless, it was quite recently reported (van Gerven et al., 2014) that its naturally present secretion pathway, the curli system forming the extracellular fibers in response to stressful environmental conditions, could be hijacked to display fusion proteins at the surface of the Gram- bacterium. This pathway is encoded by two adjacent operons encompassing genes coding for a transcriptional activator CsgD (Römling et al., 1998), a curli-specific translocation channel in the outer membrane CsgG (Robinson et al., 2006), a periplasmic chaperon CsgE (Nenninger et al., 2011), a chaperon-like protein CsgF (Nenninger et al., 2009), as well as the major and minor structural subunits, CsgA and CsgB, respectively (van Gerven et al., 2014). CsgA is secreted through the outer membrane translocation channel as a monomer and its polymerization at the extracellular surface is nucleated by CsgB (Hammer et al., 2007). Studies of the characteristics of the passengers that could be fused to C-terminus

CsgA suggest that only small proteins or domains displaying elongated folding (i.e. single antibody domain, RNAse I or fimbrial lectin domain) could be successfully displayed at the surface, a restriction likely be due to the relatively narrow 0.9 nm transverse CsgG pore diameter (Goyal et al., 2014).

Despite this reported limitation, we challenged the curli pathway for cell surface exposure of a rather large protein, the hybrid scaffoldin Scaf6 (Fierobe et al., 2005). Scaffoldins are essential components of extracellular large multienzymatic complexes called cellulosomes produced by most cellulolytic bacteria living in anaerobic biotopes (Artzi et al., 2017). The cellulosomes are nanomachines that efficiently deconstruct the plant cell wall crystalline cellulose fibres (Ravachol et al., 2014; Ravachol et al., 2015), the surrounding heterogeneous polysaccharides forming the hemicellulose (Mroueh et al., 2019; Ravachol et al., 2016) and the pectins (Pagès et al., 2003). "Simple" cellulosomes, such as those produced by *Ruminiclostridium cellulolyticum*, are composed of a single major scaffoldin containing an optional potent carbohydrate binding module (CBM3a) targeting cellulose, and a variable number of receptor modules called cohesins (Pagès et al., 1999). The cellulosomal cellulases, hemicellulases and other plant cell wall degrading enzymes all contain a complementary module named the dockerin which tenaciously binds to the cohesin with $K_D$ values in the $10^{-9}$- $10^{-11}$ M range (Cameron et al., 2015; Fierobe et al., 1999). The cohesin/dockerin interaction is thus one of the strongest bimolecular interactions reported (Stahl et al., 2012). Nevertheless, the cohesin/dockerin interaction tends to be non-specific within a given bacterial species, i.e. an enzyme bearing its dockerin module can bind to any of the cohesins displayed by the scaffoldin with similar affinity (Pagès et al., 1999; Ravachol et al., 2015; Yaron et al., 1995), whereas this strong protein/protein interaction is usually species-specific (Pagès et al., 1997). For instance, the dockerin of a cellulosomal enzyme from *R. cellulolyticum* cannot bind to the cohesins borne by the major scaffoldin of *R. thermocellum* and vice versa. This property was exploited to

design artificial cellulosomes composed of a hybrid scaffoldin displaying cohesins from different bacterial species, and enzymes engineered to exhibit the cognate dockerins (Fierobe et al., 2001; Fierobe et al., 2002; Fierobe et al., 2005). In these designer cellulosomes, the position of each enzyme onto the hybrid scaffoldin is strictly controlled thereby leading to a highly homogenous complex, in contrast to natural cellulosomes which are heterogeneous in terms of enzyme composition, stoichiometry and location. This technology was first used to dissect the molecular mechanisms underlying the strong activity of cellulosomes towards plant cell walls, but proteins or enzymes that do not participate to cellulosomes in Nature can easily be integrated *in vitro* into artificial cellulosomes by grafting an appropriate dockerin module, to serve various biotechnological goals (Arfi et al., 2014; Li et al., 2016; Mingardon et al., 2007). In this respect, we attempted to engineer an *E. coli* strain displaying at the surface the scaffoldin Scaf6, containing three cohesins from three different bacterial origins and a CBM (Fierobe et al., 2005) targeting insoluble cellulose. Although this protein (71.8 kDa) is globally much larger than the proteins successfully displayed using the curli pathway (van Gerven et al., 2014), it is composed of a chain of rather small modules (cohesins and CBM) connected by linkers. These modules exhibit similar β-sandwich or jelly roll structures (Bule et al., 2017; Shimon et al., 1997; Spinelli et al., 2000; Tormo et al., 1996). Thus, the CsgA-Scaf6 fusion, despite its high molecular mass and the transverse diameter (2-3 nm range) of its constituting modules, was challenged as a substrate for the curli machinery, and its delivery at the surface of *E. coli*. If successful, it would then be easy to functionalize the surface of the engineered *E. coli* strain by grafting the small dockerin module to any type of antigens, nanobodies or enzymes that would bind to the various cohesins of the surface-exposed Scaf6, to create for instance living vaccines, biochelators or biocatalysts.

## MATERIALS AND METHODS

**Strains and plasmids**

The pET28* vector was generated by replacement of the ATA**CCATGG** sequence (NcoI in bold) by the **GGATCC**ATGA sequence (BamHI in bold). The pET9dScaf6 without the BamHI restriction site in the Scaf6 insert (scaf6 – BamHI) was generated by site-directed mutagenesis using the primers MutBamHIScaf6up (GTTGGAGATATA**GGATCA**GCCGGTGGTTTAT, mutated BamHI site in bold) and MutBamHIScaf6do (ATAAACCACCGGC**TGA TCC**TATATCTCCAAC, mutated BamHI site in bold) and the vector pET9dScaf6H (Fierobe et al., 2005) as template. The resulting plasmid was then used as template to amplify the "scaf6 – BamHI" amplicon using the primers NheIScaf6up (ggggg**GCTAGC**GGCGATTCTCT TAAAGTTACAG, NheI in bold) and XhoIScaf6do (ggggg**CTCGAG**CTTAACAAT GATAGCGCCATCAG, XhoI in bold). The CsgA amplicon was produced by PCR using the primers BamHICsgAup (ggggg**GGATCC**ATGAAACTTTTAAAAGTAGCAGC, BamHI in bold) and NheICsgAdo (ggggg**GCTAGC**GTACTGATGAGCGGTCGCGT, NheI in bold) and the genomic DNA from *E. coli* MG1655 as template. The CsgA and "scaf6 – BamHI" amplicons and the pET28* were digested with the appropriate restriction enzymes. The fragments of interest were purified and the BamHI-XhoI pET28* was subsequently dephosphorylated. The plasmid pET28CsgAScaf6 coding for the chimeric protein CsgA-Scaf6 was constructed by ligation of the dephosphorylated BamHI-XhoI pET28* with the BamHI-NheI CsgA and the NheI-XhoI "scaf6 – BamHI" fragments. To construct the plasmid pET28CsgA coding for the native CsgA protein, the overlapping oligomers DBstopcarrierNheI-XhoIup (**CTAG**<u>C</u>TGATGACCGCGGC, NheI overhang in bold, overlapping sequence underlined) and DBstopcarrierNheI-XhoIdo (**TCGAG**<u>CCGCGGTCATCAG,</u> XhoI overhang in bold, overlapping sequence underlined) were annealed to generate the DBstopcarrierNheI-XhoI fragment containing 2 stop codons in series (TGATGA) and the overhangs extremities

normally generated by the NheI and XhoI enzyme digestion. pET28CsgA was obtained by ligation of the annealed DBstopcarrierNheI-XhoI fragment with the BamHI-NheI CsgA fragment and the dephosphorylated BamHI-XhoI pET28*. The NEB5α *E. coli* strain (NEB, Ipswich, MA) was used as the host strain for cloning and positive clones were verified by DNA sequencing. The MG1655 (DE3) *E. coli* expression strain was generated obtained using the λDE3 Lysogenization Kit (Novagen, Madison, WI).

**Cellulases and scaffoldin Scaf6 production and purification.**

The chimeric proteins (cellulases coupled with species specific dockerins) were produced and purified as formerly reported: Cel48Ft which is the *R. cellulolyticum* cellulase Cel48F coupled with a *R. thermocellum* dockerin (Fierobe et al., 2001), Cel5Af which is the *R. cellulolyticum* cellulase Cel5A engineered to bear an *Ruminococcus flavefaciens* dockerin (Borne et al., 2013) and Cel9Gc whose native *R. cellulolyticum* dockerin was replaced by the native *R. cellulolyticum* dockerin of Cel48F (Borne et al., 2013). The production and purification of Scaf6 have been described previously (Fierobe et al., 2005).

***In vivo* production and localization of CsgA-Scaf6.**

For MG1655 (DE3) *E. coli* strains harbouring pET28CsgA and pET28CsagAScaf6, the cells were grown in 100 mL-flasks in Lysogeny Broth supplemented with kanamycin (50 µg/mL) at 37°C until $OD_{600}$ reached 0.8-1. The cultures were then induced at 37°C with 500 µM isopropyl thio-β-D-galactoside (IPTG) for 16 h. The equivalents of 500 µL of culture at $OD_{600}$ = 1, were centrifuged at 6,000 *g* for 5 min and the cell-containing pellet was resuspended in SDS-PAGE loading buffer and boiled for 5 min. Five mL of each culture were also centrifuged at 6,000 *g* for 10 min at 4°C. The cell pellets were resuspended in 250 µL of the B-PER reagent (Thermo Scientific, Waltham, MA) supplemented with Lysozyme (0.4 mg/mL) and DNAse I (10 µg/mL,

Roche, Basel, Switzerland), and incubated at room temperature for 1 h. The curli fibrils which are resistant to heating in SDS and require formic acid treatment to depolymerize were extracted essentially as formerly described (Collinson et al., 1991; van Gerven et al., 2014). Briefly, the cell lysates were centrifuged at 20,000 *g* for 10 min. The pellets were resuspended twice in 1 % SDS, boiled for 10 min and centrifuged. The final pellets were resuspended in 70 µL of formic acid, prior removal of the acid using a Speed Vac Plus (Savant instruments, Midland, Mi) concentrator operating at 45°C for 1 h. The samples were then resuspended in SDS-PAGE loading buffer and boiled for 10 min. The samples were loaded onto a miniprotean precast gel 4-15% (Biorad, Hercules, Ca). After transfer onto a nitrocellulose membrane, immunodetection of CsgA-Scaf6 was performed using a mouse monoclonal antibody anti-$His_6$-peroxydase (Roche).

**Protein labeling, determination of the labeling degree and *in vitro* complexation.**

Cel48Ft, Cel5Af and Cel9Gc were concentrated and buffer exchanged by ultrafiltration (3 times) with cold 20 mM potassium phosphate pH 8. The concentration of the proteins was estimated by absorbance at 280 nm in this buffer using the free program ProtParam tool (www.expasy.org/tools/protparam.html) to determine the specific extinction coefficient for each protein. The cellulases were labeled in the dark following suppliers' protocols: Cel48Ft and Cel9Gc with Alexa Fluor 350 protein Labeling kit (Thermo Fischer Scientific, excitation wavelength [Ex] = 346nm, emission wavelength [Em] = 442 nm) and Alexa Fluor 594 protein Labeling kit (Thermo Fischer Scientific, Ex = 590 nm, Em = 617 nm) respectively, and Cel5Af with DyLight 488 NHS Ester (Thermo Fisher Scientific, Ex = 493 nm, Em = 518 nm). Labeled proteins (Cel48Ft*, Cel5Af* and Cel9Gc*) were buffer exchanged by ultrafiltration (5 times) in the dark with cold 10 mM Tris-HCl pH 8, 1 mM $CaCl_2$ in order to remove dyes in excess. The determination of the labeling degree was performed according to suppliers' protocols. *In*

*vitro* complexation between Scaf6 and labeled or unlabeled Cel48Ft, Cel5Af and Cel9Gc were performed as previously described (Borne et al., 2013): samples (10 µM final concentration) were mixed and incubated 5 min at room temperature in 20 mM Tris/maleate (pH 6.0) and 1 mM $CaCl_2$ and 4 µL of each mix was subjected to non-denaturing PAGE (4–15% gradient gel) with a Phastsystem apparatus (GE Healthcare, Pittsburgh, PA). Complexation was checked by Coomassie blue staining or fluorescence scanning using suitable filters (LPG-filter and LPR-filter for Cel5Af* and Cel9Gc* imaging, respectively) with a Typhoon FLA 9500 (GE Healthcare).

***In vivo* complexation, UV-visible microscopy and fluorescence microscopy.**

For *in vivo* complexation, the strains carrying pET28CsgA and pET28CsgAScaf6 were grown as follows: the MG1655 (DE3) *E. coli* strains overproducing CsgA and CsgAScaf6 were grown in 100 mL- flasks at 37°C in Lysogeny Broth supplemented with 50 µg/mL kanamycin at 150 rpm until $OD_{600}$ = 0.7-0.9. The induction of the expression was performed for 3 or 24 h at 37°C with 500 µM isopropyl thio-β-D-galactoside (IPTG). At the end of the induction phase, an equivalent of 100µl of cells at $OD_{600}$ = 4.5 was harvested by centrifugation (8,000 *g*, 5 min at 4°C) and resuspended in 500 µL of cold 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 200 mM NaCl. A control was performed in the same conditions using an MG1655 (DE3) *E. coli* strain untransformed without induction and without antibiotics. The cells were coated into the channels of IBIDI Flow Chambers µ-Slides (1µ-Slide $VI^{0.4}$ Poly-L-Lysine, IBIDI, Gräfelfing, Germany). Fifty µL of resuspension were loaded into the poly-L-lysine treated microscope slides and incubated for 30 min at room temperature. Fixed cells were washed 3 times as described by the supplier with 100 µL of cold 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 200 mM NaCl. After removal of all liquid, 50 to 100 µL of labeled proteins (10 µM final concentration for each) were applied on the fixed cells and incubated for 30 min at room temperature in the

dark. The excess labeling was removed by 3 washes with 100 µL of cold 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 200 mM NaCl and the stained µ-Slides were kept in the dark. For the imaging, slides were examined by microscopy using a Nikon Eclipse TE2000 E PFS inverted epifluorescence microscope (Tokyo, Japan) with a 100 x magnification oil immersion lens (objective NA 1.3 Phase Contrast). *E. coli* cells UV-visible and fluorescence micrographs were acquired in the visible range and using the appropriate filters and settings for the EPI fluorescence microscopy: GFP (for Cel5Af*), DAPI (for Cel48Ft*) and Texas Red (for Cel9Gc*). Image analysis and comparison were processed using the Fiji/ImageJ package. A statistical analysis of the corresponding micrographs was performed to determine the proportions of cells displaying no, single, dual, and triple fluorescence signal in each case (see supplemental Table 2) using a Macro whose script is reported in the supplemental material.

***In vivo* estimation of cell surface exposed CsgA-Scaf6**

Cells overproducing CsgA and CsgA-Scaf6 were grown and induced as described above. The equivalent of 200 µL at $OD_{600}$ = 0.5 were harvested by 5 min of centrifugation (6,000 *g*) at 4°C, and washed twice with 200 µL of cold and sterile buffer 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$, 200 mM NaCl (5 min, 6,000 *g*, 4°C). The pellets were then resuspended in the same buffer containing 0.5, 1 or 4 µM of Cel9Gc labeled with Alexa Fluor 594 protein Labeling kit (see above). After 1 hour of incubation at room temperature under mild stirring, the cells were centrifuged (5 min, 6,000 *g*, 4°C), and the pellets were washed twice with 200 µL of cold PBS buffer. The last cell-containing pellets were resuspended in 200 µL of cold PBS and distributed in a 96-wells microplate together with 12 solutions of 200 µL of labeled Cel9Gc in PBS buffer at concentrations ranging from 2 $10^{-10}$ to 4.1 $10^{-7}$ M. The fluorescence was monitored with a Spark 10M device (Tecan, Männedorf, Switzerland), using 10 flashes of an excitation wavelength at 580 nm (bandwith 20 nm) and an emission wavelength at 610 nm (bandwith 20

nm) (see supplemental Table 1). The experiments were performed in triplicate and dilutions of the cellular suspensions were then plated on LB-Agar to accurately determine the number of Colony Forming Units (CFU). The content of exposed hybrid scaffoldin/cell was calculated as follows:

The standard curve using known concentrations of Cel9Gc* provided the following relationship, where $F$ is the fluorescence at 610 nm of Cel9Gc* corrected with PBS buffer fluorescence.

$$[Cel9Gc*] = \frac{F - 99}{9.02 \times 10^{10}}$$

The calculation of the amount of exposed Scaf6 molecules per cell was determined using the following relationship where $F_{CsgA\text{-}Scaf6}$ and $F_{CsgA}$ are the measured fluorescence at 610 nm for CsgA-Scaf6 and CsgA overproducing strains, respectively. CFU represents the number of colony forming units per L in the sample.

$$Content\ of\ Scaf6\ per\ cell = \frac{(F_{CsgA\text{-}Scaf6} - F_{CsgA})}{9.02 \times 10^{10}} \times \frac{6.023 \times 10^{23}}{CFU}$$

**Cellulolytic activity measurements of the cells.**

One hundred-mL cultures of the strains MG1655(DE3), MG1655(DE3) (pET28CsgA) and MG1655(DE3) (pET28CsgAScaf6) were performed in Lysogeny Broth supplemented with 50 µg/mL of kanamycin if required, at 37°C until $OD_{600}$ = 0.45-0.9, and induced overnight using 500 µM IPTG. At the end of the induction, several samples of each strain corresponding to 10 mL at $OD_{600}$ = 1 were aliquoted and centrifuged at 6000 *g* for 5 min at 4°C. The cell-containing pellets were then washed twice with 1 mL of the buffer 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$ and 100 mM NaCl at 6000 *g* for 5 min at 4°C, prior to a 20-minute incubation at room temperature under mild stirring (40 rpm) with 1 mL of the same buffer without cellulase, or

containing a mix of Cel5Af, Cel48Ft and Cel9Gc at 0.1 or 1 µM. At the end of the incubation, the cells were harvested by centrifugation as described above, and washed three times using 1 mL of 10 mM Tris-HCl pH 8.0, 1 mM $CaCl_2$ and 100 mM NaCl. The last cell-containing pellets were subsequently resuspended in 1 mL of 25 mM of Tris-Maleate pH 6.1, 1 mM $CaCl_2$ and 50 mM NaCl. Ten µL of the suspensions were pipetted to visualize the activity of the cells on soluble cellulose (see below), whereas the rest of the suspension was centrifuged again and resuspended in 1 mL of the same buffer containing 3.5 g/L of crystalline cellulose Avicel (Fluka, Buchs, Switzerland). The suspensions were then incubated at 37°C under mild stirring (60 rpm). Three hundred-µL samples were pipetted after 1, 6 and 24 h of incubation, cooled down on ice and centrifuged at 4°C for 10 min at 20,000 *g*. Two hundreds µL of the supernatants were mixed with 50 µL of 0.5 M NaOH, and the released soluble cellodextrins were identified and quantified by High Pressure Anion Exchange Chromatography coupled with Pulsed Amperometric Detection (HPAEC-PAD), using a Dionex ICS 3000 (Thermofisher, Waltham, MA): 25 µL of each sample were applied to a Dionex CarboPac PA1 column (4 x 250 mm) and the corresponding guard column (4 x 50 mm) at 30°C. Sugars were eluted with the buffers 0.1 M NaOH and 0.5 M sodium acetate + 0.1 M NaOH as the eluents A and B, respectively. For glucose and cellodextrins quantifications, the following multi-step procedure was used: isochratic separation (5 min, 95 % A + 5 % B), separation gradient (8 min, 10 to 37 % B), column wash (2 min, 99 % B) and subsequent column equilibration (2.5 min, 95 % A + 5 % B). The flow rate was kept at 1 mL/min. Injections of samples containing glucose, cellobiose, cellotriose, cellotetraose, and cellopentaose (Sigma, Saint Louis, MO) at known concentrations (ranging from 4 to 100 µM) were used to identify and quantify the released sugars.

The visualization of the activity on soluble cellulose was performed by placing a cell suspension drop (10 µL in 1 mL of 25 mM of Tris-Maleate pH 6.1, 1 mM $CaCl_2$ and 50 mM NaCl, see above) onto an Agar (15 g/L) plate containing 5 g/L of CarboxyMethyl Cellulose medium

viscosity (Sigma), which was subsequently incubated overnight at 37°C. The activity on the soluble cellulose was visualized by incubating the plates with 1 % (w/v) Congo red for 15 minutes, followed by 3 washes with 1 M NaCl.

## RESULTS AND DISCUSSION

### *In vitro* assembly of trivalent heterogeneous minicellulosomes and fluorescence labeling.

The hybrid scaffoldin Scaf6 (Fierobe et al., 2005) contains a family 3a CBM framed by one cohesin module from *R. cellulolyticum* at the N-terminus, and two cohesin modules from *R. thermocellum* and *Ruminococcus flavefaciens*, respectively, at the C-terminus (Fig. 1A). It was already shown that mixing this hybrid scaffoldin with equimolar amounts of Cel48F from *R. cellulolyticum* harboring an *R. thermocellum* dockerin module (Cel48Ft), Cel9G from *R. cellulolyticum* appended with a *R. cellulolyticum* dockerin module (Cel9Gc) and Cel5A from *R. cellulolyticum* bearing a *R. flavefaciens* dockerin module (Cel5Af), generates a single and homogeneous well-ordered complex with defined enzymatic composition (Fig. 1B) (Borne et al., 2013; Fierobe et al., 2005).

To evaluate the effect of the specific grafting of a fluorescent molecular probe to each engineered cellulase on the formation of the complex *in vitro* (Fig. 1C, Table 1), the interaction of Scaf6 with the labeled and unlabeled chimeric cellulases was verified by non-denaturing polyacrylamide gel electrophoresis (PAGE) (Fig. 2). Coomassie blue staining shows that the addition of the labeled proteins (Cel48Ft*, Cel5Af* and Cel9Gc*) to Scaf6 in an equimolar ratio induced a migration shift (Fig. 2A, lanes 1 and 5) slightly different from the one observed when Scaf6 is mixed with the unlabeled partners (Cel48Ft, Cel5Af and Cel9Gc, Fig. 2A, lane 6). The fluorescence imaging (Fig. 2B) demonstrates that the complex formed upon addition of the labeled proteins contains at least Scaf6, Cel5Af* and Cel9Gc* (Fig. 2A, lanes 2, 4 and 5).

Unfortunately, complexation of the Cel48Ft* could not be monitored by fluorescence imaging due to the absence of a suitable filter in the Typhoon FLA 9500 imager. However, we confirmed the formation of the tri-functional complex by analyzing, the sequential and incremental diminution of the complex migration observed in non-denaturing PAGE upon successive addition and complexation of each cellulase partner (Fig. 2C).

These results demonstrate the labeling of the three cellulases has little or no impact on the formation *in vitro* of a single and well-ordered complex with Scaf6. We only observed a slight difference in migration between the Scaf6-based complexes containing unlabeled and labeled cellulases (Figure 2A lanes 5 and 6), which is probably due to the modified mobility of the cellulases upon labeling with the fluorophore. Altogether, these *in vitro* observations confirmed that Cel48Ft*, Cel5Af* and Cel9Gc* can be used as fluorescent probes to test the cell surface exposure of the CsgA-Scaf6 chimera.

### *In vivo* assembly and cell surface exposition of trivalent heterologous minicellulosomes

We first examined the production and localization of CsgA-Scaf6 by western blot analysis using a monoclonal antibody targeting the C-terminal His-tag of the fusion protein. As shown in Figure 3, the full-length fusion protein was detected in total extract of cells producing CsgA-Scaf6 (lane 4), and in the corresponding SDS-insoluble fraction treated with formic acid expected to contain the depolymerized curli fibrils constituents (lane 8). A truncated form of the CsgA-Scaf6 fusion was detected in the supernatant of the cell lysate of the corresponding strain (lane 6) displaying similar mobility as the purified Scaf6 (lane 9) suggesting that some proteolysis occurred around the CsgA and the hybrid scaffoldin fusion site of the chimera. In contrast, nearly no degradation fragments of the CsgA-Scaf6 fusion protein contained in the SDS-insoluble fraction (curli fibrils constituents, lane 8) was detected by the monoclonal

antibody anti His-tag borne by the chimera, thereby indicating that the intact fusion CsgA-Scaf6 was integrated in the curli fibrils.

We then examined whether the CsgA-fused passenger protein Scaf6 can attain the cell surface in its native fold and consequently be exposed and functional (Fig. 4). As described above *in vitro*, the species specific cohesin–dockerin interactions of the 3 cohesin modules from different bacterial origins borne by the CsgA-Scaf6 (Fig. 4C) should allow *in vivo* the recruitment of Cel48Ft*, Cel5Af* and Cel9Gc* (Fig. 4F), thereby leading to fluorescence emission at the surface of the strain producing this fusion. On the contrary, no fluorescent labeling on cell overproducing or not CsgA should be detected (Fig. 4 A, B, D and E). After 24 h-induction, the overlays of visible and fluorescence micrographs of bacteria producing the CsgA-Scaf6 fusion protein without the 3 fluorescents probes revealed no significant autofluorescence (Fig. 4I), as for the bacteria transformed or not with a plasmid encoding CsgA (Fig. 4G and H). Contrary to the MG1655 (DE3) *E. coli* cells harboring or not the plasmid pET28CsgA (Fig. 4J and K respectively), white spots are observed for the strain harboring the plasmid pET28CsgAScaf6 upon incubation with Cel48Ft*, Cel5Af* and Cel9Gc* (Fig. 4L) corresponding to the combination of blue (Cel48Ft*), green (Cel5Af*) and red (Cel9Gc*) artificial colors. As expected, the three fluorescent labeled cellulases probes are likely co-localized on the marked cell surface probably through their recruitment by the efficiently exposed CsgA-Scaf6 chimeras in the corresponding cells (Fig. 4L). However, not all the cells expressing CsgA-Scaf6 are labeled (Fig. 4L), a phenomenon already reported for the heterogeneous expression of genes cloned in pET vectors in a cell population using IPTG as inducer (Choi et al., 2010). A statistical analysis of the micrographs (supplemental Table 2) revealed that around 21 % of the cells carrying the vector pET28CsgAScaf6 were fluorescently labeled, among which, cells exhibiting simultaneously the three fluorescence signals (white spots) reach a proportion of 64 %. Moreover, nearly no fluorescence is detected for cells expressing or not CsgA (Fig. 4 J, 4

K, and supplemental Table 2) suggesting that non-specific interactions between the fluorescent cellulases and the cells with or without CsgA surface exposed were kept to a minimum.

All these observations support a recruitment of the fluorescent cellulases by the CsgA-Scaf6 platform at the *E. coli* cell surface, via specific interactions between the functional cohesins displayed by CsgA-Scaf6 and the cognate dockerins borne by the cellulases. This is also supported by the presence in Fig. 4L of blue, red, green, magenta, yellow and cyan spots corresponding to the individual or dual recruitment of the labeled cellulases and our results obtained by single or double labeling *in vivo* (supplemental Figure 1).

The very high levels of fluorescence detected for the marked cells in Fig. 4L also noteworthy suggests a truly massive exposure of CsgA-Scaf6 at the cell surface level, a fact supported by the analysis of the raw micrographs of the fluorescence intensity of Csga-Scaf6 producing cells (and control strain producing CsgA) incubated with Cel5Af* (supplemental Figure 2).

To confirm this massive exposure, we have investigated the average content of cell-surface exposed Scaf6 by the strain overproducing CsgA-Scaf6 using a Spark 10M microplate reader and the probe Cel9Gc*. This labeled cellulase was selected because its dockerin targets the innermost cohesin (Fig. 4F) which is presumably the least accessible, and this protein was suitably labeled for detection with our in-house Spark device. Compared to the control performed simultaneously on the strain overproducing CsgA, a value of 86,000 ± 17,000 CsgA-Scaf6 per cell interacting with the labeled cellulase was determined at saturating concentrations of Cel9Gc* (1 and 4 µM, supplemental Table 1). The profusion of Scaf6 at the surface of the engineered *E. coli* suspected by fluorescent microscopy experiments (Figure 4 and supplemental Figure 2) is thus completely confirmed: the hybrid scaffoldin is massively exposed at the surface of the strain overproducing CsgA-Scaf6.

Altogether, our data demonstrate that CsgA-Scaf6 protein is secreted and cell-surface exposed. This result is surprising since the Type VIII secretion pathway implicated in the curli formation

is mediated by a channel composed of CsgG subunits and whose diameter appears to control the transport efficiency of the CsgA-fusion (Robinson et al., 2006). It has been proposed that the diameter size of the folded conformation of the introduced polypeptides should not exceed roughly 1 nm to be translocated across the outer membrane (Goyal et al., 2014). Despite a molecular weight of 86.8 kDa for the CsgA-Scaf6 fusion (15 kDa for CsgA + 71.8 kDa for Scaf6), the chaplet structure of the Scaf6 moiety could explain why the CsgA-Scaf6 protein is efficiently secreted as the Scaf6 component should adopt a suitable conformation allowing its secretion in fusion with CsgA via the natural *E. coli* Type VIII secretion pathway, and is integrated into the extracellular fibres generated by CsgA polymerization (Figure 3). Nevertheless, the dimensions of the modules composing the Scaf6 determined from their structures exceed the proposed maximum pore size of 1 nm mentioned above: 45 x 30 x 30 Å for CBM3a (Tormo et al., 1996), 50 x 31 x 26 Å and 46 x 28 x 21 Å for the *R. cellulolyticum* (Spinelli et al., 2000) and *R. thermocellum* (Shimon et al., 1997), respectively. This suggests that the CsgG channel can to some extent accommodate larger proteins/modules during the translocation across the outer membrane, as previously reported (van Gerven et al., 2014). Moreover, each specific cohesin module of the Scaf6 passenger has attained its native fold as each of them was able to specifically interact with the corresponding dockerin module borne by the fluorescent labeled cellulases.

The cellulase activity of the strains MG1655(DE3), MG1655(DE3) (pET28CsgA) and MG1655(DE3) (pET28CsgAScaf6) after incubation with or without an equimolar mixture of Cel5Af, Cel9Gc and Cel48Ft at 0.1 or 1 µM was also monitored on crystalline cellulose Avicel and soluble cellulose CarboxyMethyl Cellulose (CMC). As shown in Figures 5A and B, the cells harboring CsgA-Scaf6 trapped a significant amount of the cellulases during the incubation and therefore displayed a considerably higher activity on crystalline cellulose compared to the two other strains, though some unspecific binding of the cellulases onto the cells overproducing

CsgA was detected at elevated enzyme concentration (1 µM, Figure 5A). Similarly, with respect to the activity on soluble cellulose, a clear CMC degradation halo was only observed for the strain MG1655(DE3) (pET28CsgAScaf6) incubated with the cellulases at 1 and 0.1 µM (Figures 5C and D). Altogether, the activities of the engineered strain MG1655(DE3) (pET28CsgAScaf6) on both types of cellulose after incubation with the cellulases confirm that the cell surface exposed hybrid scaffoldin successfully anchored the cellulases, thereby conferring cellulolytic activity to the strain.

## CONCLUSION

In conclusion we have successfully combined two presumably incompatible technologies (CsgA fusion and artificial hybrid cellulosomes) to develop a novel strain of non-pathogenic *E. coli* MG1655 which massively displays at the cell surface a hybrid scaffoldin whose three cohesins can robustly and specifically anchor three proteins/enzymes simply engineered to bear the cognate dockerin modules. This new highly flexible tool could be exploited in a large variety of biotechnology applications. For instance, the engineered *E. coli* cells could host three enzymes involved in three successive steps in a specific pathway and form a biocatalyst that could be easily removed from the sample when the desired product is obtained, by centrifugation or by using the affinity of the hosted CBM3a for cellulose. Bioremediation using appropriate proteins to capture/oxidize the toxic compound or living vaccines exhibiting three different antigens of the same pathogenic microbe on the cell surface exposed Scaf6 could also be planned. Furthermore, as the repertoire of specific cohesin/dockerin devices continues to expand, it should be possible to cell surface display scaffoldins containing at least up to 7 different cohesins at the same time (Stern et al., 2016).

## ACKNOWLEDGEMENTS

The authors are indebted to Leon Espinoza from the LCB for his help during fluorescence microscopy experiments. The authors thank Stephanie Perret, Chantal Tardif, Pascale de Philip and Sandrine Pagès for helpful discussions, and critically reading the manuscript. This work was supported by a grant from the "Agence Nationale de la Recherche" (Grant number: ANR-14-CE05-0019-01).

## REFERENCES

Arfi Y, Shamshoum M, Rogachev I, Peleg Y, Bayer EA. 2014. Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc. Natl. Acad. Sci. U. S. A.* **111**:9109–9114.

Artzi L, Bayer EA, Moraïs S. 2017. Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* **15**:83–95.

Borne R, Bayer EA, Pagès S, Perret S, Fierobe H-P. 2013. Unraveling enzyme discrimination during cellulosome assembly independent of cohesin-dockerin affinity. *FEBS J.* **280**:5764–5779.

Bule P, Alves VD, Israeli-Ruimy V, Carvalho AL, Ferreira LMA, Smith SP, Gilbert HJ, Najmudin S, Bayer EA, Fontes CMGA. 2017. Assembly of *Ruminococcus flavefaciens* cellulosome revealed by structures of two cohesin-dockerin complexes. *Sci. Rep.* **7**:759.

Cameron K, Weinstein JY, Zhivin O, Bule P, Fleishman SJ, Alves VD, Gilbert HJ, Ferreira LMA, Fontes CMGA, Bayer EA, Najmudin S. 2015. Combined Crystal Structure of a Type I Cohesin: mutation and affinity binding studies reveal structural determinants of cohesin-dockerin specificities. *J. Biol. Chem.* **290**:16215–16225.

Choi YJ, Morel L, Le François T, Bourque D, Bourget L, Groleau D, Massie B, Míguez CB. 2010. Novel, versatile, and tightly regulated expression system for *Escherichia coli* strains. *Appl. Environ. Microbiol.* **76**:5058–5066.

Collinson SK, Emödy L, Müller KH, Trust TJ, Kay WW. 1991. Purification and characterization of thin, aggregative fimbriae from *Salmonella enteritidis*. *J. Bacteriol.* **173**:4773–4781.

Fierobe HP, Mechaly A, Tardif C, Belaich A, Lamed R, Shoham Y, Belaich JP, Bayer EA. 2001. Design and production of active cellulosome chimeras. Selective incorporation of dockerin-containing enzymes into defined functional complexes. *J. Biol. Chem.* **276**:21257–21261.

Fierobe HP, Pagès S, Bélaïch A, Champ S, Lexa D, Bélaïch JP. 1999. Cellulosome from *Clostridium cellulolyticum*: molecular study of the Dockerin/Cohesin interaction. *Biochemistry* **38**:12822–12832.

Fierobe H-P, Bayer EA, Tardif C, Czjzek M, Mechaly A, Bélaïch A, Lamed R, Shoham Y, Bélaïch J-P. 2002. Degradation of cellulose substrates by cellulosome chimeras. Substrate targeting versus proximity of enzyme components. *J. Biol. Chem.* **277**:49621–49630.

Fierobe H-P, Mingardon F, Mechaly A, Bélaïch A, Rincon MT, Pagès S, Lamed R, Tardif C, Bélaïch J-P, Bayer EA. 2005. Action of designer cellulosomes on homogeneous versus complex substrates: controlled incorporation of three distinct enzymes into a defined trifunctional scaffoldin. *J. Biol. Chem.* **280**:16325–16334.

Goyal P, Krasteva PV, Van Gerven N, Gubellini F, Van den Broeck I, Troupiotis-Tsaïlaki A, Jonckheere W, Péhau-Arnaudet G, Pinkner JS, Chapman MR, Hultgren SJ, Howorka S, Fronzes R, Remaut H. 2014. Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG. *Nature* **516**:250–253.

Gu M-Z, Wang J-C, Liu W-B, Zhou Y, Ye B-C. 2013. Expression and displaying of β-glucosidase from *Streptomyces coelicolor* A3 in *Escherichia coli*. *Appl. Biochem. Biotechnol.* **170**:1713–1723.

Hammer ND, Schmidt JC, Chapman MR. 2007. The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization. *Proc. Natl. Acad. Sci. U. S. A.* **104**:12494–12499.

Li M, Yue Y, Zhang Z-J, Wang Z-Y, Tan T-W, Fan L-H. 2016. Site-Specific and High-Loading Immobilization of Proteins by Using Cohesin-Dockerin and CBM-Cellulose Interactions. *Bioconjug. Chem.* **27**:1579–1583.

Mingardon F, Chanal A, López-Contreras AM, Dray C, Bayer EA, Fierobe H-P. 2007. Incorporation of fungal cellulases in bacterial minicellulosomes yields viable, synergistically acting cellulolytic complexes. *Appl. Environ. Microbiol.* **73**:3822–3832.

Mroueh M, Aruanno M, Borne R, de Philip P, Fierobe H-P, Tardif C, Pagès S. 2019. The xyl-doc gene cluster of *Ruminiclostridium cellulolyticum* encodes GH43- and GH62-α-l-arabinofuranosidases with complementary modes of action. *Biotechnol. Biofuels* **12**:144.

Nenninger AA, Robinson LS, Hammer ND, Epstein EA, Badtke MP, Hultgren SJ, Chapman MR. 2011. CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. *Mol. Microbiol.* **81**:486–499.

Nenninger AA, Robinson LS, Hultgren SJ. 2009. Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proc. Natl. Acad. Sci. U. S. A.* **106**:900–905.

Pagès S, Bélaïch A, Bélaïch JP, Morag E, Lamed R, Shoham Y, Bayer EA. 1997. Species-specificity of the cohesin-dockerin interaction between *Clostridium thermocellum* and *Clostridium cellulolyticum*: prediction of specificity determinants of the dockerin domain. *Proteins* **29**:517–527.

Pagès S, Bélaïch A, Fierobe HP, Tardif C, Gaudin C, Bélaïch JP. 1999. Sequence analysis of scaffolding protein CipC and ORFXp, a new cohesin-containing protein in *Clostridium*

*cellulolyticum*: comparison of various cohesin domains and subcellular localization of ORFXp. *J. Bacteriol.* **181**:1801–1810.

Pagès S, Valette O, Abdou L, Bélaïch A, Bélaïch J-P. 2003. A rhamnogalacturonan lyase in the *Clostridium cellulolyticum* cellulosome. *J. Bacteriol.* **185**:4727–4733.

Ramesh B, Sendra VG, Cirino PC, Varadarajan N. 2012. Single-cell characterization of autotransporter-mediated *Escherichia coli* surface display of disulfide bond-containing proteins. *J. Biol. Chem.* **287**:38580–38589.

Ravachol J, Borne R, Meynial-Salles I, Soucaille P, Pagès S, Tardif C, Fierobe H-P. 2015. Combining free and aggregated cellulolytic systems in the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. *Biotechnol. Biofuels* **8**:114.

Ravachol J, Borne R, Tardif C, de Philip P, Fierobe H-P. 2014. Characterization of all family-9 glycoside hydrolases synthesized by the cellulosome-producing bacterium *Clostridium cellulolyticum*. *J. Biol. Chem.* **289**:7335–7348.

Ravachol J, de Philip P, Borne R, Mansuelle P, Maté MJ, Perret S, Fierobe H-P. 2016. Mechanisms involved in xyloglucan catabolism by the cellulosome-producing bacterium *Ruminiclostridium cellulolyticum*. *Sci. Rep.* **6**:22770.

Robinson LS, Ashman EM, Hultgren SJ, Chapman MR. 2006. Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol. Microbiol.* **59**:870–881.

Römling U, Bian Z, Hammar M, Sierralta WD, Normark S. 1998. Curli fibers are highly conserved between *Salmonella typhimurium* and *Escherichia coli* with respect to operon structure and regulation. *J. Bacteriol.* **180**:722–731.

Shimon LJ, Bayer EA, Morag E, Lamed R, Yaron S, Shoham Y, Frolow F. 1997. A cohesin domain from *Clostridium thermocellum*: the crystal structure provides new insights into cellulosome assembly. *Struct. Lond. Engl. 1993* **5**:381–390.

Spinelli S, Fierobe HP, Belaïch A, Belaïch JP, Henrissat B, Cambillau C. 2000. Crystal structure of a cohesin module from *Clostridium cellulolyticum*: implications for dockerin recognition. *J. Mol. Biol.* **304**:189–200.

Stahl SW, Nash MA, Fried DB, Slutzki M, Barak Y, Bayer EA, Gaub HE. 2012. Single-molecule dissection of the high-affinity cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A.* **109**:20431–20436.

Stern J, Moraïs S, Lamed R, Bayer EA. 2016. Adaptor scaffoldins: an original strategy for extended designer cellulosomes, inspired from Nature. *mBio* **7**:e00083.

Tanaka T, Kawabata H, Ogino C, Kondo A. 2011. Creation of a cellooligosaccharide-assimilating *Escherichia coli* strain by displaying active beta-glucosidase on the cell surface via a novel anchor protein. *Appl. Environ. Microbiol.* **77**:6265–6270.

Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y, Steitz TA. 1996. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. *EMBO J.* **15**:5739–5751.

van Gerven N, Goyal P, Vandenbussche G, De Kerpel M, Jonckheere W, De Greve H, Remaut H. 2014. Secretion and functional display of fusion proteins through the curli biogenesis pathway. *Mol. Microbiol.* **91**:1022–1035.

Yaron S, Morag E, Bayer EA, Lamed R, Shoham Y. 1995. Expression, purification and subunit-binding properties of cohesins 2 and 3 of the *Clostridium thermocellum* cellulosome. *FEBS Lett.* **360**:121–124.

**Figure legends:**

**Figure 1: Schematic representation of the proteins used and assembled complexes**. A) the hybrid scaffoldin Sca6, B) the Scaf6-based complex containing unlabeled Cel9Gc, Cel48Ft and Cel5Af, and C) Scaf6-based complex containing Cel48Ft labeled with Alexafluor 350 (Cel48Ft*, blue), Cel5Af labeled with Dylight 488 (Cel5Af*, green) and Cel9Gc labeled with Alexafluor 594 (Cel9Gc*, red). Proteins/modules from *R. cellulolyticum*, *R. thermocellum* and *R. flavefaciens* are shown in grey, black and white, respectively.

**Figure 2: *In vitro* verification of labeled cellulases-Scaf6 complexation by non-denaturing PAGE analysis**. A) Coomassie blue stained gel showing the migration of Scaf6 (lane 1), Cel5Af* (lane 2), Cel48Ft* (lane 3), Cel9Gc* (lane 4), an equimolar mix of Scaf6, Cel5Af*, Cel48Ft* and Cel9Gc* (lane 5) and an equimolar mix of Scaf6, Cel5Af, Cel48Ft and (lane 6). B) Fluorescence imaging of the non-denaturing PAGE gel loaded as in A) using the LPG filter (left) and the LPR filter (middle) from the Typhoon FLA 9500 imager and both pictures superimposed (right). Displayed colors are artificial: green (Cel5Af*) and red (Cel9Gc*) for the LPG and LPR filters, respectively. C) Coomassie blue stained gel showing the migration of Scaf6 (lane 1), an equimolar mix of Scaf6 and Cel48Ft* (lane 2), an equimolar mix of Scaf6, Cel48Ft* and Cel9Gc* (lane 3) and an equimolar mix of Scaf6, Cel48Ft*, Cel9Gc* and Cel5Af* (lane 4). For all panels, 4 µL containing 10 µM of each protein or complex were loaded in each lane.

**Figure 3**: ***In vivo* production and localization of CsgA-Scaf6 by western blot analysis.** The production and cell surface exposition of CsgA-Scaf6 was checked for *E. coli* MG1555(DE3) strain transformed with pET28CsgAScaf6 or pET28CsgA (as control) using an antibody

targeting the C-terminal His-tag of the CsgA-Scaf6 fusion. Lane 1, molecular mass markers (molecular masses indicated on the left); lane 2, not induced MG1555(DE3) (pET28CsgAScaf6) cells; lane 3, induced MG1555(DE3) (pET28CsgA) cells; lane 4, induced MG1555(DE3) (pET28CsgAScaf6) cells; lane 5, SDS-soluble fraction of MG1555(DE3) (pET28CsgA) lysate; lane 6, SDS-soluble fraction of MG1555(DE3) (pET28CsgAScaf6) lysate; lane 7, SDS-insoluble fraction of MG1555(DE3) (pET28CsgA) lysate; lane 8, SDS-insoluble fraction of MG1555(DE3) (pET28CsgAScaf6) lysate; lane 9, purified Scaf6. * indicates the band corresponding to the expected size of the fusion protein CsgA-Scaf6. Ten µL were loaded in each sample lane.

**Figure 4: analysis of the complexation between CsgA-Scaf6 with fluorescent cellulases**. A-C) Schematic representation of the examined *E. coli* strains. D-F) expected labeling after incubation of the examined strains with the three fluorescent cellulases. The autofluorescence of MG1655(DE3), MG1655(DE3) (pET28CsgA) and MG1655(DE3) (pET28CsgAScaf6) without incubation with the labeled cellulases, is shown in Figures G, H and I, respectively (phase contrast and fluorescence superimposed). Fluorescence of MG1655(DE3), MG1655(DE3) (pET28CsgA) and MG1655(DE3) (pET28CsgAScaf6) after incubation with a mix of the three labeled cellulases at 10 µM is shown in Figures J, K and L, respectively (phase contrast and fluorescence superimposed). G-L) Displayed colors are artificial: blue, green and red for Cel48Ft*, Cel5Af* and Cel9Gc* emission, respectively. The yellow rectangle (Figures B, C, E and F) symbolizes a curli fibril. Legend same as Figure 1.

**Figure 5: Activities on crystalline cellulose Avicel and soluble cellulose CMC of the various engineered strains after incubation with a mixture of Cel5Af, Cel9Gc and**

**Cel48Ft.** The Avicelase activity of the cells after incubation with buffer, or an equimolar mix of Cel5Af, Cel9Gc and Cel48Ft at 1 µM (A) or 0.1 µM (B) was performed at 37°C and the cellodextrins released in the supernatants after 1, 6 and 24 h of incubation were monitored using HPAEC-PAD. The curves are labelled as follows: grey solid line, strain MG1655(DE3) (pET28CsgAScaf6) incubated with cellulases; grey dotted line, strain MG1655(DE3) (pET28CsgAScaf6) incubated with buffer; orange solid line, strain MG1655(DE3) (pET28CsgA) incubated with cellulases; orange dotted line, strain MG1655(DE3) (pET28CsgA) incubated with buffer; blue solid line, strain MG1655(DE3) incubated with cellulases; blue dotted line, strain MG1655(DE3) incubated with buffer; The data show the mean and standard deviations of two independent experiments. The activity on CMC (C and D) was visualized by applying 10 µL of cell suspension incubated with buffer or a mix of Cel5Af, Cel9Gc and Cel48Ft at 1 µM (C) or 0.1 µM (D) onto CMC-agar plates. Plates were incubated for 16 hours at 37°C, and the CMCase activity was subsequently detected using Congo red. The names of the strains and presence (+) or absence (-) of cellulases during the incubation are indicated on top of each snapshot.

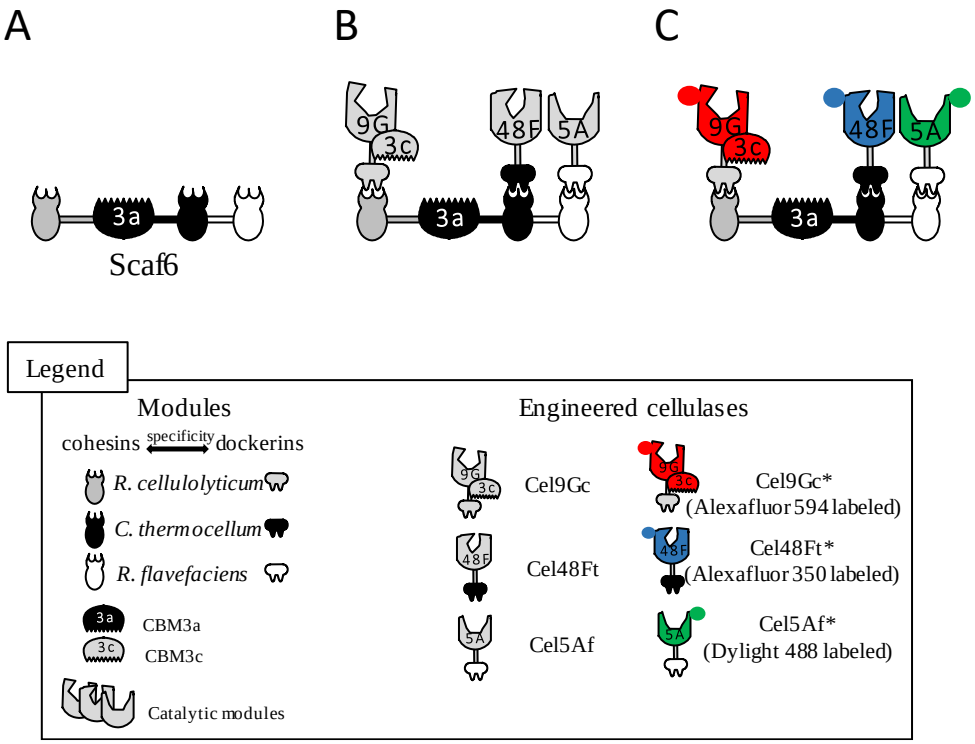

A
B
C
9G
3c
48F
5A
3a
Scaf6
Legend
Modules
cohesins
specificity
dockerins
R. cellulolyticum
C. thermocellum
R. flavefaciens
CBM3a
CBM3c
Catalytic modules
Engineered cellulases
Cel9Gc
Cel48Ft
Cel5Af
Cel9Gc*
(Alexafluor 594 labeled)
Cel48Ft*
(Alexafluor 350 labeled)
Cel5Af*
(Dylight 488 labeled)

**Figure 1**

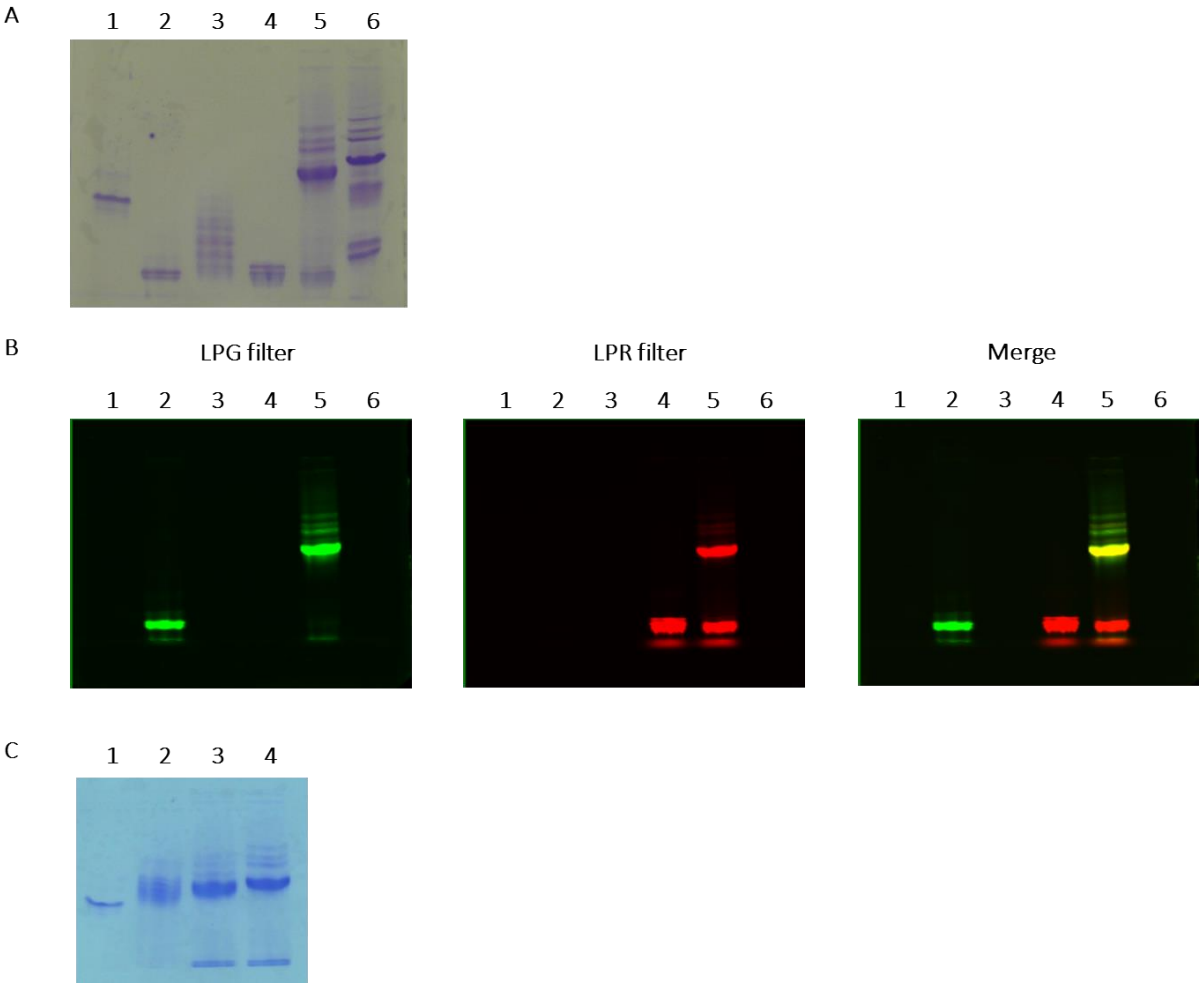

A
1 2 3 4 5 6
B
LPG filter
LPR filter
Merge
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
C
1 2 3 4

**Figure 2**

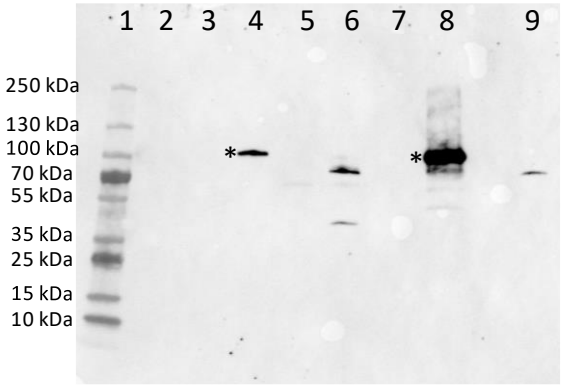

1 2 3 4 5 6 7 8 9
250 kDa
130 kDa
100 kDa
70 kDa
55 kDa
35 kDa
25 kDa
15 kDa
10 kDa
*
*

**Figure 3**

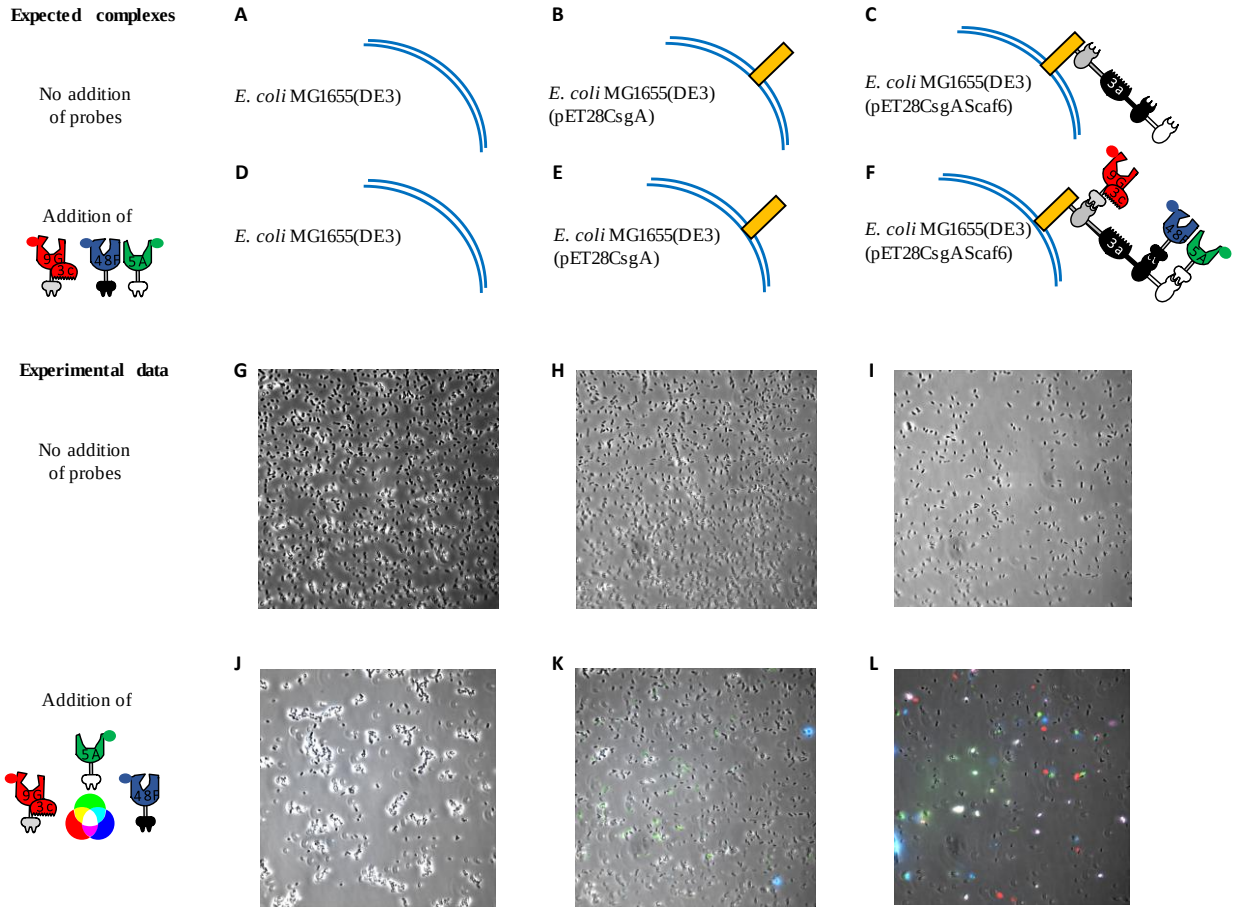

Expected complexes
A
B
C
No addition of probes
E. coli MG1655(DE3)
E. coli MG1655(DE3) (pET28CsgA)
E. coli MG1655(DE3) (pET28CsgAScaf6)
D
E
F
Addition of
E. coli MG1655(DE3)
E. coli MG1655(DE3) (pET28CsgA)
E. coli MG1655(DE3) (pET28CsgAScaf6)
Experimental data
G
H
I
No addition of probes
J
K
L
Addition of

Figure 4

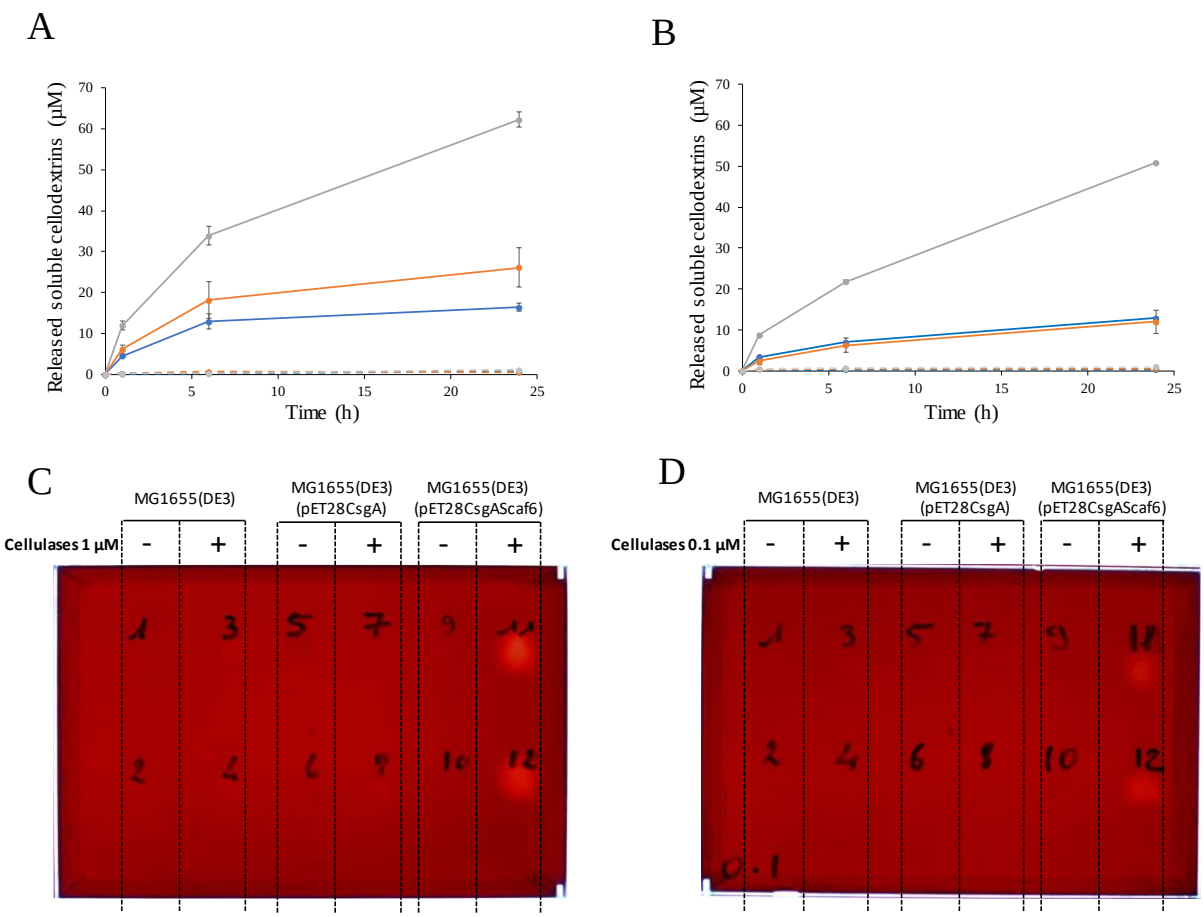

A
Released soluble cellodextrins (µM)
Time (h)
B
Released soluble cellodextrins (µM)
Time (h)
C
MG1655(DE3)
MG1655(DE3) (pET28CsgA)
MG1655(DE3) (pET28CsgAScaf6)
Cellulases 1 µM
- + - + - +
D
MG1655(DE3)
MG1655(DE3) (pET28CsgA)
MG1655(DE3) (pET28CsgAScaf6)
Cellulases 0.1 µM
- + - + - +

Figure 5