CHEMICAL BIOLOGY

Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose

Wiriya Thongsomboon,[1] Diego O. Serra,[2] Alexandra Possling,[2] Chris Hadjineophytou,[2,3]* Regine Hengge,[2]† Lynette Cegelski[1]†

Cellulose is a major contributor to the chemical and mechanical properties of plants and assumes structural roles in bacterial communities termed biofilms. We find that *Escherichia coli* produces chemically modified cellulose that is required for extracellular matrix assembly and biofilm architecture. Solid-state nuclear magnetic resonance spectroscopy of the intact and insoluble material elucidates the zwitterionic phosphoethanolamine modification that had evaded detection by conventional methods. Installation of the phosphoethanolamine group requires BcsG, a proposed phosphoethanolamine transferase, with biofilm-promoting cyclic diguanylate monophosphate input through a BcsE-BcsF-BcsG transmembrane signaling pathway. The *bcsEFG* operon is present in many bacteria, including *Salmonella* species, that also produce the modified cellulose. The discovery of phosphoethanolamine cellulose and the genetic and molecular basis for its production offers opportunities to modulate its production in bacteria and inspires efforts to biosynthetically engineer alternatively modified cellulosic materials.

Cellulose is the most abundant biopolymer on Earth. Plants rely on the tensile strength and mechanical properties of cellulose to stand upright (*1*). Chemically, cellulose is a linear polysaccharide composed of β-1,4–linked glucosyl residues. Individual strands participate in strong hydrogen-bonding networks with neighboring strands and contribute to the physical and chemical integrity of plant cell walls and cellulosic materials (*2*). Microorganisms are also major producers of cellulose (*3*). The essential genetic and protein machinery for cellulose production in bacteria include the cellulose synthase genes, termed *bcsA* and *bcsB*, which encode cellulose synthase subunits BcsA and BcsB (*4*). BcsA is an integral membrane protein containing the catalytic active site. BcsB interacts with BcsA at the periplasmic face of the inner membrane in Gram-negative bacteria, with the two subunits forming a channel for cosynthetic secretion of cellulose. Cellulose biosynthesis requires activation by the ubiquitous bacterial second messenger cyclic diguanylate monophosphate (c-di-GMP) (*5*), which directly binds to BcsA (*6*). Intense curiosity has emerged in understanding the diversity of additional genes in cellulose biosynthesis operons that are present in many microorganisms (*3*). Here we report on the determination of the structure of a modified cellulose, phosphoethanolamine (pEtN) cellulose, produced naturally by *Escherichia coli* and other Gram-negative bacteria. We provide the genetic basis for its production and the functional implications of gene-directed pEtN cellulose synthesis.

E. coli and *Salmonella* are among the best-studied microorganisms reported to produce cellulose. These include human pathogens such as uropathogenic and enterohemorrhagic *E. coli*. Functionally, the exopolysaccharide cellulose is a major component of the self-produced extracellular matrix in biofilms, which represent physiologically heterogeneous and spatially structured bacterial communities (*7*, *8*). Biofilm formation is of high medical relevance, as it confers enhanced resistance to antibiotics and host defenses during infection (*9*). Within the biofilm matrix, cellulose forms a nanocomposite with amyloid curli fibers that encapsulates individual cells in supramolecular basketlike structures, enmeshes the bacterial community, and confers cohesion and elasticity that allow biofilms to fold and buckle up in a tissuelike manner (*10*–*12*). Biochemical and solid-state nuclear magnetic resonance (NMR) measurements with the clinically important uropathogenic *E. coli* strain UTI89 established that the matrix was composed of curli fibers and cellulosic material in a 6:1 ratio by mass. During this bottom-up analysis involving ^{13}C and ^{15}N NMR analysis of the purified components, we also discovered that the cellulose portion appears to be modified in some way with an aminoethyl functionality (*11*).

Solid-state NMR analysis of the intact cellulosic material, complemented by solution-state NMR and mass spectrometry analysis of acid-digested material, has now enabled the determination of the chemical structure of the modified cellulose as a polymer containing glucose and glucose-6-phosphoethanolamine (Fig. 1A). The ^{13}C cross-polarization/magic angle spinning (CPMAS) NMR spectrum of isolated cellulosic material contains carbon contributions from the glucose backbone plus two additional carbons associated with the modification (Fig. 1B). A comparison of this spectrum with that of the cellulosic material isolated with the aid of the dye Congo red (CR) is provided in Fig. 1C, as the extraction and yield of cellulosic material is enhanced in the presence of CR and is sometimes used for the preparation of larger samples in subsequent analysis (*13*). CR is also commonly used as a supplement in nutrient agar plates for evaluation of *E. coli* and *Salmonella* community phenotypes because both curli and cellulosic polymers bind the dye (*14*). The use of CR does not result in any changes to the cellulosic carbon composition (Fig. 1C). C{N} rotational-echo double-resonance (REDOR) NMR was used to experimentally select directly bonded carbon-nitrogen pairs and identified the 41–parts per million (ppm) carbon peak as the C-8 carbon, the only carbon directly bonded to nitrogen (fig. S1). ^{31}P CPMAS revealed the presence of ^{31}P as a phosphate in the polymer (fig. S2), and we hypothesized that the attachment site was at the C-6 position because of the downfield shift of a carbon in the C-6 region (as in glucose-6-phosphate). C{P} REDOR NMR revealed that phosphorous was indeed positioned closest to the C-6′ and C-7 carbons and next closest to the C-5 and C-8 carbons (Fig. 1D), suggesting the full structural assignment as pEtN cellulose. A ^{13}C CP-array NMR experiment enabled the quantitative accounting of carbon contributions to the pEtN ^{13}C CPMAS spectrum and determined that approximately one-half of the cellulose glucose units in the intact polymer are modified (fig. S3).

Additional solution-based analyses were performed to complement the solid-state NMR analysis, although this required acid digestion of the pEtN cellulose to release soluble components into solution. Solution-state NMR analysis of acid-digested pEtN cellulose supported the assignments from solid-state NMR and confirmed that the modification occurs at the C-6 position. However, a standard 48-hour HCl hydrolysis leads to degradation of the modification, and only soluble glucose, glucose-6-phosphate, and ethanolamine were detected (figs. S4 to 7), observations that explain the difficulty of identifying the modified cellulose using digestions and solution-based methods. To attempt to capture an intact modified glucose unit, shorter hydrolysis times were used and liquid chromatography–mass spectrometry was used for analysis. The intact pEtN glucose was detected, as well as released glucose, glucose-6-phosphate, and ethanolamine (fig. S8). The detection of pEtN glucose likely escaped detection in previous research because of the instability of the modification upon acid digestion and isolation protocols designed to specifically detect cellulose. Typical approaches do not use the careful purification protocol we developed and, instead, rely on cell lysis and polysaccharide enrichment followed by harsh hydrolysis and either colorimetric detection upon reaction with a sulfuric acid solution of anthrone (*15*) or chemical derivatization and mass spectrometry detection (*16*) to support the presence

[1]Department of Chemistry, Stanford University, Stanford, CA 94305, USA. [2]Institute of Biology, Microbiology, Humboldt–Universität zu Berlin, 10115 Berlin, Germany. [3]Department of Chemistry and Molecular Biology, University of Gothenburg, 41296 Gothenburg, Sweden.
*Present address: Department of Biosciences, University of Oslo, 0371 Oslo, Norway.
†Corresponding author. Email: cegelski@stanford.edu (L.C.); regine.hengge@hu-berlin.de (R.H.)

of significant glucose content as a reporter for cellulose production. The presence of glucose-6-phosphate, ethanolamine, and other compounds could escape detection or be ascribed to residual cellular components.

A biosynthetically modified cellulose has wide-ranging implications and potential applications. Among these, the specifically modified cellulose could be essential for the formation and function of bacterial biofilms containing the polymer, could exhibit attractive properties for new cellulosic materials, and could potentially be introduced into other organisms, if gene directed. Thus, we sought to identify the genes involved in the installation of the cellulose modification. The *bcsEFG* operon, which is part of the cellulose gene cluster in *E. coli*, had not been ascribed a definitive role in cellulose synthesis. The ^{13}C CPMAS NMR comparison of the isolated cellulose from a *bcsG* deletion mutant (Δ*bcsG*) revealed that the *bcsG* gene was indispensable for the cellulose modification (Fig. 1E). The spectrum lacks the contributions from the 41-ppm C-8 carbon and the 63-ppm C-7 carbon. As expected, the sugar C-6 carbon contribution appears only at the upfield ^{13}C position of 63 ppm, corresponding to unmodified glucose. Complementation of the Δ*bcsG* mutant with *bcsG* on a plasmid restored production of the modified cellulose (fig. S9). The prevalence of the modification was reduced in the in-frame nonpolar Δ*bcsF* mutant and more strongly reduced or abolished in the Δ*bcsE* mutant, indicating that BcsE and BcsF may play accessory, and possibly regulatory, roles in the installation of pEtN by BcsG (fig. S10).

Similar macrocolony phenotypes of nonpolar Δ*bcsE*, Δ*bcsF*, and Δ*bcsG* mutants (fig. S11), as well as their coexpression from a single operon in the *bcs* gene cluster, also suggested functional cooperation of these three proteins. Notably, BcsG and BcsF are membrane inserted, whereas BcsE is a soluble c-di-GMP–binding protein (*17*). To further elucidate the molecular basis of BcsG function and the roles of BcsE and BcsF, we tested for potential direct interactions between these proteins as well as with the cellulose synthase subunits BcsA and BcsB. Using a bacterial two-hybrid assay that is based on the reconstitution of adenylate cyclase from two separate domains fused to proteins that potentially interact (*18*), we observed strong interactions in vivo of BcsG with BcsF as well as with BcsA (Fig. 2A). Thus, BcsG operates in close proximity to the cellulose synthesizing BcsA-BcsB complex. Moreover, BcsF also showed strong interaction with BcsE (Fig. 2A), suggesting a BcsE-BcsF-BcsG pathway that controls cellulose modification through direct protein-protein interactions. Qualitative evaluation of cellulose production by CR binding in *E. coli* strain AR3110Δ*csgBA*, which produces no curli fibers but cellulose only, and its Δ*bcsE*, Δ*bcsF*, and Δ*bcsG* mutant derivatives additionally revealed comparable CR binding for AR3110Δ*csgBA* and AR3110Δ*csgBA*Δ*bcsG*, but reduced CR binding for AR3110Δ*csgBA*Δ*bcsE* and AR3110Δ*csgBA*Δ*bcsF* (fig. S12). The decrease in cellulose production by the Δ*bcsE* and Δ*bcsF* mutants was corroborated by a low yield in isolatable material for NMR analysis and could indicate that BcsE and BcsF contribute to stability of the BcsA-BcsB machinery in AR3110, influencing efficiency of cellulose synthesis.

BcsG is composed of 559 amino acids and has been predicted to be an integral membrane protein. A hydropathy plot analysis of BcsG supported the presence of several putative transmembrane-spanning regions in the N-terminal 160 amino acids followed by a large hydrophilic C-terminal domain. However, charge distribution flanking the hydrophobic amino acid stretches did not allow us to unequivocally predict the number and orientation of the transmembrane regions (*19*) and thus the localization of the C-terminal domain. Therefore, we generated a set of translational reporter fusions to β-galactosidase (LacZ) and alkaline phosphatase (PhoA) to identify cytoplasmic and periplasmic regions of BcsG. This approach is based on the observation that the signal sequence–driven attempt to export the normally cytoplasmic LacZ results in the jamming of secretion machinery and interferes with LacZ activity, whereas PhoA becomes active only upon its secretion to the periplasm, where DsbAB-mediated disulfide bond formation occurs (*20*). Accordingly, fusions inserted after codon 1 of *bcsG* showed high LacZ activity, but low PhoA activity (Fig. 2B). By contrast, when the entire N-terminal BcsG domain of approximately 160 amino acids—including all hydrophobic stretches—was present in the hybrid proteins, low LacZ and high PhoA activities were observed, consistent with the large C-terminal domain of BcsG residing on the periplasmic side of the inner membrane (Fig. 2, B and C). A similar analysis with BcsF, a 63–amino acid peptide, showed that its N-terminal single-transmembrane region, which is preceded by negative charges and followed by positive charges, inserts into the membrane such that the small hydrophilic C-terminal domain remains on the cytoplasmic side of the membrane (Fig. 2, B and C).

We addressed the question of the substrate for BcsG-mediated PEtN modification of cellulose emerging from the BcsA-BcsB complex. We noticed that, with respect to overall size (559 and 563

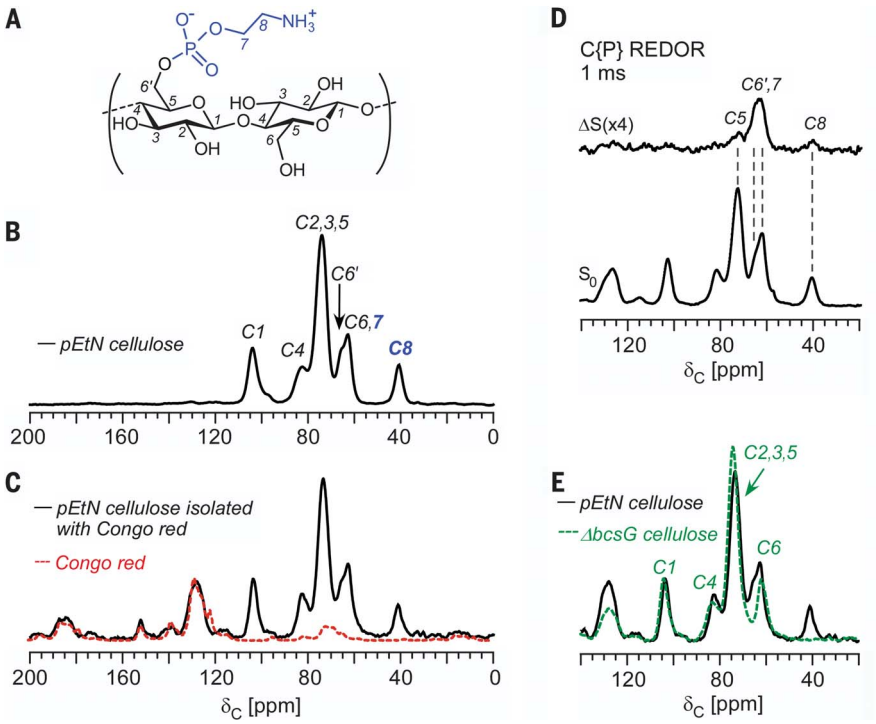

Fig. 1. *E. coli* produces phosphoethanolamine cellulose. (**A**) Representation of the chemical structure of glucose and pEtN glucose units in pEtN cellulose. (**B**) ^{13}C CPMAS solid-state NMR spectra of the pure modified cellulose with two additional carbon contributions, C-7 (63 ppm) and C-8 (41 ppm). The C-6 carbon appears at 62 ppm for the unmodified glucose units and at 66 ppm for the modified glucose units (figs. S1 and S4). (**C**) ^{13}C CPMAS spectra of the modified cellulose compared with that of the modified cellulose isolated from cells grown in the presence of CR. The pure CR spectrum is provided as an overlay (dashed red line). The comparison demonstrates that purification with CR does not influence the polysaccharide composition. δ_C, carbon chemical shift. (**D**) The C-6′ and C-7 carbon chemical-shift region exhibited the strongest dephasing in the 1-ms C{P} REDOR NMR measurement, followed by that of the C-5′ and C-8 carbons, suggesting the full structural assignment as pEtN cellulose, further confirmed by solution-state NMR and mass spectrometry (figs. S4 to S8). S_0, REDOR full-echo spectrum; Δ*S*, REDOR difference spectrum. (**E**) The ^{13}C CPMAS spectrum of the cellulosic material isolated from the *bcsG* derivative lacked modification carbons and contained only the ^{13}C chemical shifts expected for standard amorphous cellulose.

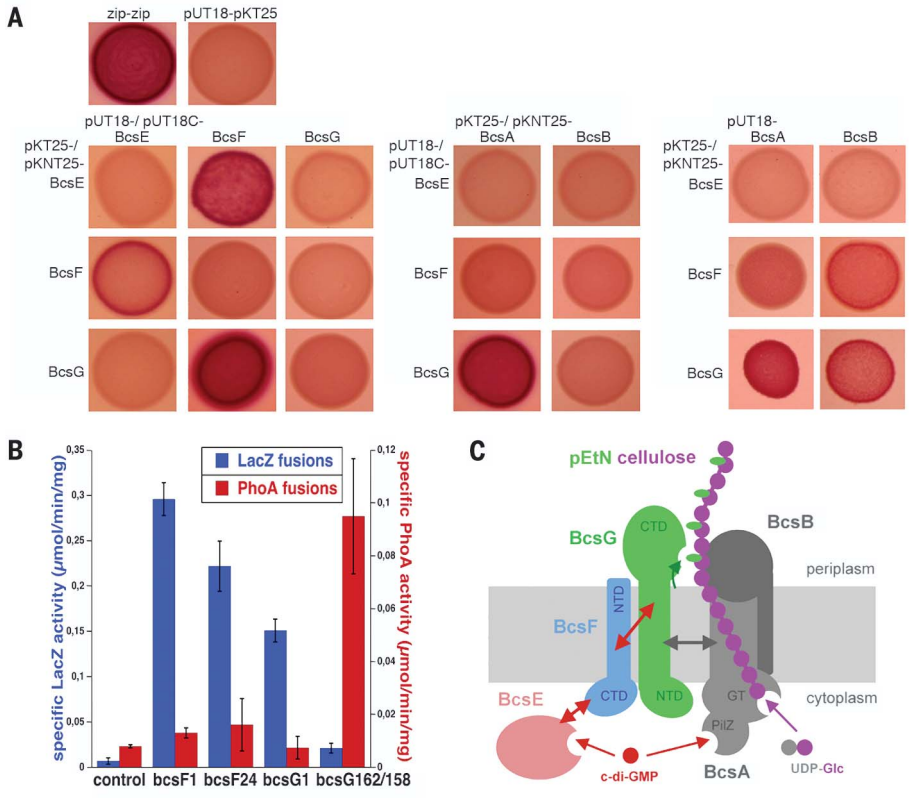

Fig. 2. BcsG directly interacts with cellulose synthase and communicates with the c-di-GMP–binding BcsE via the transmembrane peptide BcsF. (**A**) Interactions of the indicated proteins were tested using a bacterial two-hybrid (2H) system based on the reconstitution of adenylate cyclase (AC) (*18*), which allows the utilization of maltose by W3110Δ*cya*, resulting in red color on MacConkey agar plate. The 2H vector plasmids allow the attachment of the respective AC domain tags (*18*, *25*), either at the N terminus (pKT25, pUT18c) or the C terminus (pKNT25, pUT18) of a protein. For BcsA, BcsB, BcsE, and BcsF, the tags were located at the C terminus; for BcsG, the tags were located at the N terminus. Zip-zip, leucine zipper domain of the yeast GCN4 protein, used as a positive control. (**B**) Transmembrane orientation of BcsF and BcsG was determined by assaying enzymatic activities of hybrid proteins between N-terminal parts from BcsF and BcsG fused to LacZ and PhoA expressed from low–copy number plasmids in strains W3110Δ*lac (I-A)* and W3110Δ*phoA*. Fusion joints were after codon 1 (all combinations), codon 24 (of *bcsF* fused to both reporter genes), codon 162 (*bcsG::lacZ*), or codon 158 (*bcsG::phoA*). (**C**) A schematic model of the directly interacting modules for cellulose synthesis (BcsAB) and modification (BcsEFG) summarizes the protein-protein interactions (double-headed arrows) detected in (A), the transmembrane orientation of BcsG and BcsF as tested in (B), and dual control by the second messenger c-di-GMP, which binds to both BcsA (*6*) and BcsE (*17*). NTD, N-terminal domain; CTD, C-terminal domain; UDP-Glc, uridine diphosphate–glucose.

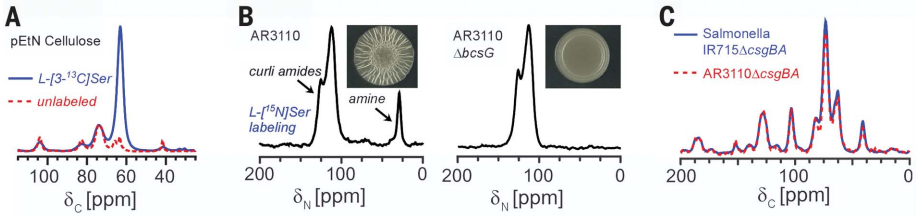

Fig. 3. Phosphoethanolamine cellulose production is detected in curli-integrated *E. coli* biofilm matrices with isotopic serine labeling and is also produced by *Salmonella enterica*. (**A**) Isotopic labeling with L-[3-^{13}C]Ser–supplemented YESCA nutrient medium resulted in enrichment of the pEtN cellulose C-7 carbon in an isolated pEtN sample, consistent with routing through a possible substrate such as phosphatidylethanolamine. (**B**) Isotopic labeling with L-[^{15}N]Ser was evaluated by ^{15}N CPMAS NMR on extracellular matrix samples containing both curli and cellulosic material. The ^{15}N-amide signals correspond to curli amides. The loss of the ^{15}N-amine signal in the *bcsG* derivative (right) confirmed the amine nitrogen assignment as that from pEtN cellulose. Loss of the modification was accompanied by loss of the wrinkled macrocolony morphology (inset photographs). (**C**) The ^{13}C CPMAS spectrum of the cellulosic material isolated from *Salmonella enterica* serovar Typhimurium strain IR715Δ*csgBA* matched that of pEtN cellulose from AR3110Δ*csgBA*.

residues, respectively) and length and transmembrane orientation of domains, BcsG resembles EptB, a PEtN transferase that uses the phospholipid phosphatidylethanolamine (PE) to modify bacterial lipopolysaccharide (*21*). Two similar PEtN transferases, EptC from *Campylobacter jejuni* and LptA from *Neisseria meningitis*, also have their active site domains on the periplasmic side of the cytoplasmic membrane; they equally use PE as a substrate, and their crystal structures show two or three histidine residues coordinating zinc, which is essential for activity (*22*). We therefore hypothesized that histidine residues in BcsG may play a similar role and that pEtN modification by BcsG may also originate from PE. We isolated a mutant version of BcsG (BcsG^{mut}) in which His^{396}, His^{400}, and His^{443} in the periplasmic domain were replaced by alanine residues. BcsG^{mut} did not complement the Δ*bcsG* mutation with respect to macrocolony formation (fig. S13) and was inactive with respect to cellulose modification (fig. S14), suggesting that BcsG enzymatic activity resides in the periplasm and is related to these known PEtN transferases, despite the absence of clear primary-sequence similarity. If BcsG also uses PE as a substrate, the modified cellulose should have atoms derived from serine, which serves as a direct precursor for the ethanolamine moiety of PE. Thus, pEtN cellulose was prepared from cells grown on agar medium supplemented with 25 mg/liter of L-[3-^{13}C]Ser to detect whether pEtN cellulose would be enriched through incorporation of the serine label. The expected C-7 carbon in the pEtN cellulose spectrum was indeed enhanced as a result of label incorporation from serine (Fig. 3A). The routing of serine into the modification is consistent with PE serving as a substrate for BcsG. Labeling with L-[^{15}N]Ser was also successful and is additionally valuable in identifying the pEtN cellulose modification in the context of complex extracellular matrix samples that contain both pEtN cellulose and curli. Labeling with L-[^{15}N]Ser yielded the anticipated ^{15}N-amine contribution from pEtN cellulose and the shift-resolved curli ^{15}N-amide contributions from serine as well as glycine residues that result from serine conversion in glycine biosynthesis (Fig. 3B). In this way, the potential presence of pEtN cellulose can be determined in intact extracellular matrix preparations from different *E.coli* strains and different organisms.

Overall, we propose a model in which BcsG acts as a pEtN transferase with its catalytic domain in the periplasm, modifying cellulose after its emergence from the BcsA-BcsB machinery (Fig. 2C). Given the association of BcsG with the c-di-GMP–binding cytoplasmic protein BcsE and the transmembrane peptide BcsF, we also propose that cellulose modification is controlled by c-di-GMP in a transmembrane signaling pathway that involves BcsE and BcsF, with BcsF serving as a direct link between BcsE and BcsG (Fig. 2C). Thus, the biofilm-promoting second messenger c-di-GMP plays a dual role by activating both the synthesis and the modification of cellulose via the PilZ domain of BcsA and the BcsE-BcsF-BcsG transmembrane signaling pathway, respectively.

Notably, the K_d values for c-di-GMP binding to BcsA and BcsE are 8 (*23*) and 2.4 μM (*17*), respectively, which ensures efficient modification whenever c-di-GMP levels are high enough to support cellulose synthesis.

Production of the biofilm matrix by wild-type *E. coli* involves the coproduction and tight association of amyloid curli fibers and what has been considered to be cellulose (*8*, *12*). Yet, we have now determined that *E. coli* strains such as strain AR3110, which is a direct derivative of the widely studied K-12 laboratory strain W3110 (*12*), as well as the classical uropathogenic UTI89 (*24*) produce pEtN cellulose. Thus, we sought to test whether this cellulose modification is functionally important for biofilm-matrix architecture and function. Macrocolony morphotypes, the matrix fine structure as analyzed by in situ fluorescence and electron microscopy, and the multicellular cohesion of macrocolonies when challenged by shear stress were evaluated. BcsG was required for the buckling into radial ridges and wrinkles typically observed for the otherwise very flat AR3110 macrocolonies (Fig. 4A, compare panels i and iii). The ring-shaped curli-only–driven macrocolony architecture of AR3110Δ*bcsG* (Fig. 4A, panel iii) resembles that of cellulose-free strains such as AR3110Δ*bcsA* (Fig. 4A, panel ii) or W3110 (*12*). Furthermore, fluorescence microscopy of vertical cryosections through macrocolonies grown in the presence of thioflavin S, acting as a matrix dye that binds to cellulose and curli, revealed that BcsG was required for cellulose to assemble into long, thick, and straight filaments. These were most clearly observed in the absence of curli, i.e., in AR3110Δ*csgB* (Fig. 4A, panel iv). Compared to these long filaments, only short, thin, and curled filaments were detected in AR3110Δ*csgB*Δ*bcsG* macrocolonies (Fig. 4A, panel v, and fig. S15). Scanning electron microscopy further supported a role of modified cellulose in matrix architecture. The matrix at the surface of a AR3110Δ*bcsG* macrocolony resembles that of the curli-only strain W3110 (Fig. 4B). Thus, cellulose modification is required to form the extended composite structure with curli fibers that nearly fully covers the surface of a macrocolony biofilm (as visible with strain AR3110 in Fig. 4B). Finally, an extracellular matrix consisting of either the cellulose-curli nanocomposite or of cellulose alone is known to provide cohesion to the cellular community. The resulting tissuelike behavior includes the ability not only to fold and buckle up but also to resist shear stress as a cohesive community (*12*). The latter can be tested by submerging a macrocolony in liquid and exposing it to gentle shaking. In this assay, the curli-free AR3110Δ*csgBA* strain detached from the agar phase as an entire macrocolony, whereas colonies of the corresponding cellulose modification–deficient strain just dissolved into flares of loose cells (fig. S16). Thus, pEtN modification of cellulose is required for community behavior based on the formation of the connective matrix consisting of either long-range cellulose fibers or the curli-cellulose nanocomposite network that envelops and connects bacterial cells during biofilm formation.

The production of pEtN cellulose may also have consequences for virulence of pathogenic strains. Amyloid curli fibers are proinflammatory (*25*, *26*) and a virulence factor for various types of pathogenic *E. coli* (*27*, *28*), yet the tight association of curli with "cellulose" was reported to counteract these properties (*27*, *29*). Therefore, mutations in the *bcsEFG* operon resulting in nonmodified cellulose that cannot form the nanocomposite with curli fibers may enhance the contribution of curli to virulence of pathogenic *E. coli*. Consistent with this notion, the 2011 European outbreak O104:H4 strain, an enteroaggregative and enterohemorrhagic *E. coli* which not only produced Shiga toxin but also high amounts of curli at 37°C while being a *bcsE* mutant (*30*), was of unprecedented virulence for this pathotype of *E. coli* (*31*).

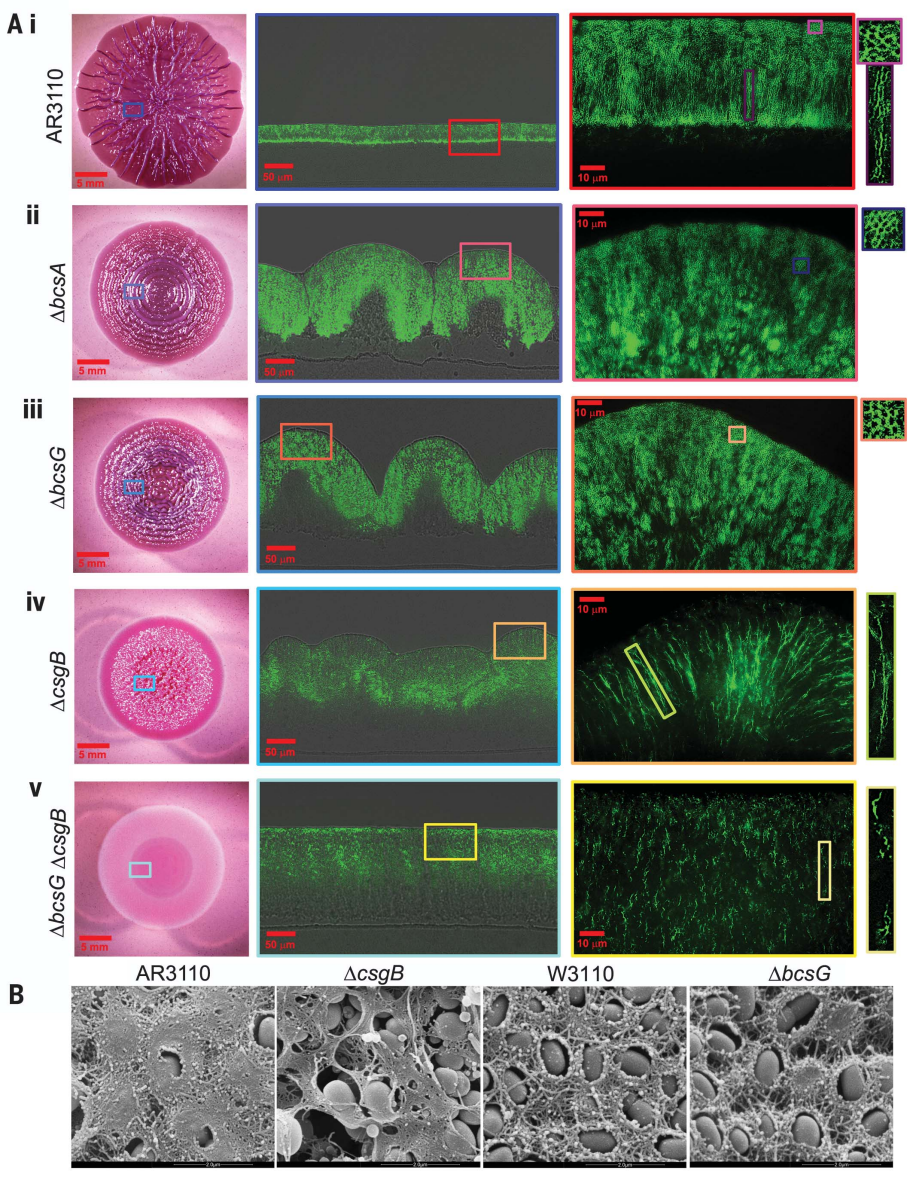

Fig. 4. Eliminating BcsG changes macroscopic morphology and microscopic matrix architecture of *E. coli* macrocolony biofilms. (**A**) Macrocolonies of strain AR3110, which produces both cellulose and curli fibers, and the indicated mutant derivatives (i to v) were grown for three days on salt-free LB agar plates containing either Congo red or the green-fluorescent thioflavin S, which stain cellulose and curli without affecting the overall matrix architecture and colony morphotype. The microscopic architecture of thioflavin S–stained matrix was visualized in thin cross sections of macrocolonies, with color-coded boxed areas being further enlarged adjacently. (**B**) The surface of macrocolonies was visualized at high resolution by scanning electron microscopy. The classical *E. coli* K-12 lab strain W3110 is isogenic to AR3110, except for a $bcsQ^{stop}$ mutation, which eliminates the ability to produce cellulose (*12*). Because of polarity, the Δ*csgB* mutation also eliminates the expression of CsgA (from the *csgBA* operon), i.e., both curli subunits encoded by *csgBA* are not produced.

Finally, together with core cellulose genes, *bcsEFG* genes have been shown to occur in many γ- and β-proteobacteria (*3*). We isolated the cellulose material from *Salmonella enterica* serovar Typhimurium strain IR715Δ*csgBA* (*32*), a curli mutant, and discovered that it also produces pEtN cellulose. The ^{13}C CPMAS NMR spectrum of isolated cellulose from *Salmonella* matches that of pEtN cellulose from *E. coli* (Fig. 3C). Thus, the pEtN modification of cellulose is likely to be common in the γ and β branches of proteobacteria. Notably, however, *Komagataeibacter xylinus* (formerly known as *Acetobacter xylinum*), which produces excessive amounts of cellulose such that PEtN modification would predictably lead to a depletion of the headgroups of the phospholipid membrane, does not possess the *bcsEFG* genes (*3*). Other bacterial species that do not possess *bcsEFG* genes, but feature accessory *bcs* genes of unknown function, could possibly use alternative modes of cellulose modification.

Modified cellulose is produced by strains that have been assumed in the literature to be producing standard amorphous cellulose on the basis of simple calcofluor-white staining procedures and conventional isolation methods designed for the detection of glucose from hydrolyzed cellulose. However, these methods involve harsh hydrolysis protocols and crude purification or enrichment methods, followed by chromatography and mass spectrometry, and, thus, a complete accounting of the intact material has not been attempted. Solid-state NMR analysis of the relevant intact polysaccharide was able to identify this biologically important pEtN modification that evaded detection by conventional approaches. PEtN cellulose is a newly identified zwitterionic polymer, and, to our knowledge, our study provides the first definitive evidence so far of a naturally postsynthetically modified cellulose. In the extracellular matrix of bacterial biofilms, pEtN modification of cellulose seems to be multifunctional: It is required for the formation of long cellulose fibrils and a tight nanocomposite with amyloid curli fibers that generate tissuelike cohesive and elastic behavior of biofilms, it may confer resistance against attacks by cellulase-producing microorganisms (e.g., fungi) in the environment, and it can prevent amyloid curli fibers from hyperstimulating immune responses, which, in the long run, may contribute to pathogen fitness in the host. Moreover, cellulose biosynthesis and postsynthetic modification are coregulated by the ubiquitous biofilm-promoting second messenger c-di-GMP via a transmembrane c-di-GMP signaling pathway. Inhibition of BcsG could offer new opportunities to control biofilm formation, in particular by Gram-negative pathogens associated with chronic infections. Furthermore, the identification of the gene-directed biosynthetic machinery also inspires the generation of engineered systems to produce alternately modified cellulosic materials.

REFERENCES AND NOTES

1. D. Klemm, B. Heublein, H. P. Fink, A. Bohn, *Angew. Chem. Int. Ed.* **44**, 3358–3393 (2005).
2. Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, *J. Am. Chem. Soc.* **125**, 14300–14306 (2003).
3. U. Römling, M. Y. Galperin, *Trends Microbiol.* **23**, 545–557 (2015).
4. O. Omadjela *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 17856–17861 (2013).
5. U. Jenal, A. Reinders, C. Lori, *Nat. Rev. Microbiol.* **15**, 271–284 (2017).
6. J. L. Morgan, J. T. McNamara, J. Zimmer, *Nat. Struct. Mol. Biol.* **21**, 489–496 (2014).
7. P. S. Stewart, M. J. Franklin, *Nat. Rev. Microbiol.* **6**, 199–210 (2008).
8. D. O. Serra, A. M. Richter, G. Klauck, F. Mika, R. Hengge, *mBio* **4**, e00103-13 (2013).
9. G. G. Anderson, G. A. O'Toole, *Curr. Top. Microbiol. Immunol.* **322**, 85–105 (2008).
10. C. Hung *et al.*, *mBio* **4**, e00645-13 (2013).
11. O. A. McCrate, X. Zhou, C. Reichhardt, L. Cegelski, *J. Mol. Biol.* **425**, 4286–4294 (2013).
12. D. O. Serra, A. M. Richter, R. Hengge, *J. Bacteriol.* **195**, 5540–5554 (2013).
13. C. Reichhardt *et al.*, *Anal. Bioanal. Chem.* **408**, 7709–7717 (2016).
14. U. Römling, *Cell. Mol. Life Sci.* **62**, 1234–1246 (2005).
15. D. M. Updegraff, *Anal. Biochem.* **32**, 420–424 (1969).
16. X. Zogaj, M. Nimtz, M. Rohde, W. Bokranz, U. Römling, *Mol. Microbiol.* **39**, 1452–1463 (2001).
17. X. Fang *et al.*, *Mol. Microbiol.* **93**, 439–452 (2014).
18. G. Karimova, J. Pidoux, A. Ullmann, D. Ladant, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 5752–5756 (1998).
19. J. Nilsson, B. Persson, G. von Heijne, *Proteins* **60**, 606–616 (2005).
20. C. Manoil, J. Beckwith, *Science* **233**, 1403–1408 (1986).
21. C. M. Reynolds, S. R. Kalb, R. J. Cotter, C. R. H. Raetz, *J. Biol. Chem.* **280**, 21202–21211 (2005).
22. C. D. Fage, D. B. Brown, J. M. Boll, A. T. Keatinge-Clay, M. S. Trent, *Acta Crystallogr. D Biol. Crystallogr.* **70**, 2730–2739 (2014).
23. I. S. Pultz *et al.*, *Mol. Microbiol.* **86**, 1424–1440 (2012).
24. S. L. Chen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **103**, 5977–5982 (2006).
25. Z. Bian, A. Brauner, Y. Li, S. Normark, *J. Infect. Dis.* **181**, 602–612 (2000).
26. C. Tükel *et al.*, *Cell. Microbiol.* **12**, 1495–1505 (2010).
27. Y. Kai-Larsen *et al.*, *PLOS Pathog.* **6**, e1001010 (2010).
28. L. Cegelski *et al.*, *Nat. Chem. Biol.* **5**, 913–919 (2009).
29. X. Wang *et al.*, *Cell. Mol. Life Sci.* **63**, 2352–2363 (2006).
30. A. M. Richter, T. L. Povolotsky, L. H. Wieler, R. Hengge, *EMBO Mol. Med.* **6**, 1622–1637 (2014).
31. H. Karch *et al.*, *EMBO Mol. Med.* **4**, 841–848 (2012).
32. C. Tükel *et al.*, *Mol. Microbiol.* **58**, 289–304 (2005).

ACKNOWLEDGMENTS

We thank J. Schaefer for assistance with ^{31}P NMR measurements; S. Lynch and A. Banerjee for assistance with solution-state NMR; and the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University Mass Spectrometry (T. McLaughlin) for mass spectrometry assistance. We thank F. Yildiz and D. Zamorano-Sanchez for providing the pMMB956 plasmid; J. Rolff (Freie Universität Berlin) for the ability to use the scanning electron microscope in his laboratories; and C. Tükel for the *Salmonella enterica* serovar Typhimurium strain IR715Δ*csgBA*. Financial support was provided by the NSF CAREER award 1453247 (to L.C.), the Precourt Institute for Energy Seed Grant (to L.C.), and the Deutsche Forschungsgemeinschaft grants He1556/17-1 and He1556/21-1 (to R.H.). C.H. was supported by an ERASMUS fellowship from the European Union. W.T. and L.C. have filed a U.S. patent application relating to the work in this manuscript, no. PCT/US2017/047511 filed 18 August 2017 (production and applications of pEtN cellulose).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/359/6373/334/suppl/DC1
Materials and Methods
Figs. S1 to S16
Table S1
References (*33*–*38*)

17 July 2017; accepted 13 November 2017
10.1126/science.aao4096

Phosphoethanolamine cellulose: A naturally produced chemically modified cellulose

Wiriya Thongsomboon, Diego O. Serra, Alexandra Possling, Chris Hadjineophytou, Regine Hengge and Lynette Cegelski

Science **359** (6373), 334-338.
DOI: 10.1126/science.aao4096

A naturally modified cellulose

Cellulose is the most abundant biopolymer on Earth and an important component of bacterial biofilms. Thongsomboon *et al.* used solid-state nuclear magnetic resonance spectroscopy to identify a naturally derived, chemically modified cellulose, phosphoethanolamine cellulose (see the Perspective by Galperin and Shalaeva). They went on to identify the genetic basis and molecular signaling involved in introducing this modification in bacteria, which regulates biofilm matrix architecture and function. This discovery has implications for understanding bacterial biofilms and for the generation of new cellulosic materials.

Science, this issue p. 334; see also p. 276