Published in final edited form as:

Nature. 2014 December 11; 516(7530): 250–253. doi:10.1038/nature13768.

Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG

Parveen Goyal^{1,2}, Petya V. Krasteva^{3,4}, Nani Van Gerven^{1,2}, Francesca Gubellini^{3,4}, Imke Van den Broeck^{1,2}, Anastassia Troupiotis-Tsaïlaki⁵, Wim Jonckheere^{1,2}, Gérard Péhau-Arnaudet⁴, Jerome S. Pinkner⁶, Matthew R. Chapman⁷, Scott J. Hultgren⁶, Stefan Howorka⁸, Rémi Fronzes^{3,4}, and Han Remaut^{1,2}

¹Structural and Molecular Microbiology, Structural Biology Research Center, VIB, Pleinlaan 2, 1050 Brussels, Belgium

²Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

³Unité G5 Biologie structurale de la sécrétion bactérienne, Institut Pasteur, 25–28 rue du Docteur Roux, 75015 Paris, France

⁴UMR 3528, CNRS, Institut Pasteur, 25–28 rue du Docteur Roux, 75015 Paris, France

⁵Structure et Fonction des Membranes Biologiques (SFMB), Université Libre de Bruxelles, 1050 Brussels, Belgium

⁶Department of Molecular Microbiology and Microbial Pathogenesis, Washington University in Saint Louis School of Medicine, St Louis, Missouri 63110-1010, USA

⁷Department of Molecular, Cellular and Developmental Biology, University of Michigan, Ann Arbor, Michigan 48109-1048, USA

⁸Department of Chemistry, Institute for Structural and Molecular Biology, University College London, London WC1H 0AJ, UK

Abstract

Curli are functional amyloid fibres that constitute the major protein component of the extracellular matrix in pellicle biofilms formed by Bacteroidetes and Proteobacteria (predominantly of the α

©2014 Macmillan Publishers Limited. All rights reserved

Correspondence and requests for materials should be addressed to H.R. (han.remaut@vib-vub.be) or R.F. (remi.fronzes@pasteur.fr).

Author Contributions P.G. produced, purified and crystallized CsgG and CsgG_{C1S}, and determined their X-ray structures. Single-particle EM was performed by P.V.K., F.G. and G.P.H. and supervised by R.F. P.V.K., F.G. and W.J. performed the *in vitro* characterization of the CsgG–CsgE complex. N.V.G. and W.J. performed mutagenesis and phenotyping experiments. I.V.d.B. conducted the single-channel recordings, and S.H. supervised the acquisition and analysis of the recordings. A.T.T. and P.G. recorded and analysed FTIR spectra. J.S.P., M.R.C. and S.J.H. conceived the study and contributed expression constructs and protein. H.R. conceived and supervised the study, analysed data and wrote the paper with contributions from all authors.

Coordinates and structure factors for CsgG_{C1S} and CsgG have been deposited in the Protein Data Bank under accession codes 4uv2 and 4uv3, respectively. The cryo-EM map for CsgG–CsgE has been deposited in the EMDataBank under accession code EMD-2750.

Reprints and permissions information is available at www.nature.com/reprints.

The authors declare no competing financial interests.

Readers are welcome to comment on the online version of the paper.

and γ classes)^{1–3}. They provide a fitness advantage in pathogenic strains and induce a strong pro-inflammatory response during bacteraemia^{1,4,5}. Curli formation requires a dedicated protein secretion machinery comprising the outer membrane lipoprotein CsgG and two soluble accessory proteins, CsgE and CsgF^{6,7}. Here we report the X-ray structure of *Escherichia coli* CsgG in a non-lipidated, soluble form as well as in its native membrane-extracted conformation. CsgG forms an oligomeric transport complex composed of nine anticodon-binding-domain-like units that give rise to a 36-stranded β -barrel that traverses the bilayer and is connected to a cage-like vestibule in the periplasm. The trans-membrane and periplasmic domains are separated by a 0.9-nm channel constriction composed of three stacked concentric phenylalanine, asparagine and tyrosine rings that may guide the extended polypeptide substrate through the secretion pore. The specificity factor CsgE forms a nonameric adaptor that binds and closes off the periplasmic face of the secretion channel, creating a 24,000 Å³ pre-constriction chamber. Our structural, functional and electrophysiological analyses imply that CsgG is an ungated, non-selective protein secretion channel that is expected to employ a diffusion-based, entropy-driven transport mechanism.

Curli are bacterial surface appendages that have structural and physical characteristics of amyloid fibrils, best known from human degenerative diseases^{7–9}. However, the role of bacterial amyloids such as curli are to facilitate biofilm formation^{4,10}. Unlike pathogenic amyloids, which are the product of protein misfolding, curli formation is coordinated by proteins encoded in two dedicated operons, *csgBAC* (*curli specific genes BAC*) and *csgDEFG* in *Escherichia coli* (Extended Data Fig. 1)^{6,7}. After secretion, CsgB nucleates CsgA subunits into curli fibres^{7,11,12}. Secretion and extracellular deposition of CsgA and CsgB are dependent on two soluble accessory factors, respectively CsgE and CsgF, as well as CsgG, a 262-residue lipoprotein located in the outer membrane^{13–16}. Because of the lack of hydrolysable energy sources or ion gradients at the outer membrane, CsgG falls into a specialized class of protein translocators that must operate through an alternatively energized transport mechanism. In the absence of a structural model, the dynamic workings of how CsgG promotes the secretion and assembly of highly stable amyloid-like fibres in a regulated fashion across a biological membrane has so far remained enigmatic.

Before insertion in to the outer membrane, lipoproteins are piloted across the periplasm by means of the lipoprotein localization (Lol) pathway¹⁷. We observed that non-lipidated CsgG (CsgG_{C1S}) could be isolated as a soluble periplasmic intermediate, analogous to the pre-pore forms observed in pore-forming proteins and toxins¹⁸. CsgG_{C1S} was found predominantly as monomers, in addition to a minor fraction of discrete oligomeric complexes (Extended Data Fig. 2)¹⁹. The soluble CsgG_{C1S} oligomers were crystallized and their structure was determined to 2.8 Å, revealing a hexadecameric particle with eight-fold dihedral symmetry (D_8), consisting of two ring-shaped octameric complexes (C_8) that are joined in a tail-to-tail interaction (Extended Data Fig. 2 and Fig. 1). The CsgG_{C1S} protomer shows an anticodon-binding domain (ABD)-like fold that is extended with two α -helices at the amino and carboxy termini (α N and α C, respectively; Fig. 1 and Extended Data Fig. 3a–c). Additional CsgG-specific elements are an extended loop linking β 1 and α 1, two insertions in the loops connecting β 3– β 4 and β 5– α 3, and an extended α 2 helix that is implicated in CsgG oligomerization by packing between adjacent monomers (Fig. 1b). Further inter-protomer

contacts are formed between the back of the $\beta 3$ – $\beta 5$ sheet and the extended $\beta 1$ – $\alpha 1$ loop (Extended Data Fig. 3d, e).

In the CsgG_{C15} structure, residues 1–17, which would link $\alpha 1$ to the N-terminal lipid anchor, are disordered and no obvious transmembrane (TM) domain can be discerned (Fig. 1). Attenuated total reflection Fourier transform infrared spectroscopy (ATR–FTIR) of CsgG_{C15} and native, membrane-extracted CsgG revealed that the latter has a higher absorption in the β -sheet region (1,625–1,630 cm^{-1}) and a concomitant reduction in the random coil and α -helical regions (1,645–1,650 cm^{-1} and 1,656 cm^{-1} , respectively; Fig. 2a), suggesting that membrane-associated CsgG contains a β -barrel domain. Candidate sequence stretches for β -strand formation are found in the poorly ordered, extended loops connecting $\beta 3$ – $\beta 4$ (residues 134–154) and $\beta 5$ – $\alpha 3$ (residues 184–204); deletion of these resulted in the loss of curli formation (Fig. 2b). The crystal structure of detergent-extracted CsgG confirmed a conformational rearrangement of both regions into two adjacent β -hairpins, extending the β -sheet formed by $\beta 3$ – $\beta 4$ (TM1) and $\beta 5$ – $\alpha 3$ (TM2) (Fig. 2c). Their juxtaposition in the CsgG oligomer gave rise to a composite 36-stranded β -barrel (Fig. 2d). Whereas the crystallized CsgG_{C15} oligomers showed a D_8 symmetry, the CsgG structure showed D_9 symmetry, with CsgG protomers retaining equivalent interprotomer contacts, except for a 5° rotation relative to the central axis and a 4 Å translation along the radial axes (Extended Data Fig. 2). This observation is reconciled in the in-solution oligomeric states revealed by single-particle electron microscopy, which exclusively found C_9 and D_9 symmetries for membrane-extracted CsgG (Extended Data Fig. 2). The predominant presence of monomers in the non-lipidated sample and the symmetry mismatch with the membrane-bound protein argue that before membrane insertion, CsgG is targeted to the outer membrane in a monomeric, LolA-bound form and that the C_8 and D_8 particles are an artefact of highly concentrated solutions of CsgG_{C15}. Furthermore, we show that the C_9 nonamer rather than the D_9 complex forms the physiologically relevant particle, because in isolated *E. coli* outer membranes, cysteine substitutions in residues enclosed by the observed tail-to-tail dimerization are accessible to labelling with maleimide-polyethylene glycol (PEG, 5 kDa; Extended Data Fig. 4).

Thus, CsgG forms a nonameric transport complex 120 Å in width and 85 Å in height. The complex traverses the outer membrane through a 36-stranded β -barrel with an inner diameter of 40 Å (Fig. 2e). The N-terminal lipid anchor is separated from the core domain by an 18-residue linker that wraps over the adjacent protomer (Extended Data Fig. 3d). The diacylglycerol- and amide-linked acyl chain on the N-terminal Cys are not resolved in the electron density maps, but on the basis of the location of Leu 2 the lipid anchor is expected to flank the outer wall of the β -barrel. On the periplasmic side, the transporter forms a large solvent-accessible cavity with an inner diameter of 35 Å and a height of 40 Å that opens to the periplasm in a 50 Å mouth formed by helix 2 (Fig. 2e). At its apex, this periplasmic vestibule is separated from the TM channel by a conserved 12-residue loop connecting $\beta 1$ to $\alpha 1$ (C-loop; Figs 2e and 3a, b), which constricts the secretion conduit to a solvent-excluded diameter of 9.0 Å (Fig. 3a, c). These pore dimensions would be compatible with the residence of one or two (for example a looped structure) extended polypeptide segments, with five residues spanning the height of the constriction (Extended Data Fig. 5). The

luminal lining of the constriction is composed of three stacked concentric rings formed by the side chains of residues Tyr 51, Asn 55 and Phe 56 (Fig. 3a, b). In the anthraxPA₆₃ toxin, a topologically equivalent concentric Phe ring (referred to as a ϕ -clamp) lines the entry of the translocation channel and catalyses polypeptide capture and passage^{20–22}. Multiple sequence alignment of CsgG-like translocators shows the absolute conservation of Phe 56 and the conservative variation of Asn 55 to Ser or Thr (Extended Data Fig. 6). Mutation of Phe 56 or Asn 55 to Ala leads to a near loss of curli production (Fig. 3d), whereas a Asn 55→Ser substitution retains wild-type secretion levels, together alluding to the requirement of the stacked configuration of a ϕ -clamp followed by a hydrogen-bond donor/acceptor in the CsgG constriction (Fig. 3b and Extended Data Fig. 6).

Single-channel current recordings of CsgG reconstituted in planar phospholipid bilayers led to a steady current of 43.1 ± 4.5 pA ($n = 33$) or -45.1 ± 4.0 pA ($n = 13$) using standard electrolyte conditions and a potential of +50 mV or –50 mV, respectively (Fig. 3e, f and Extended Data Fig. 7). The observed current was in good agreement with the predicted value of 47 pA calculated on the basis of a simple three-segment pore model and the dimensions of the channel lumen seen in the X-ray structure (Fig. 3c). A second, low-conductance conformation can also be observed under negative electrical field potential (-26.2 ± 3.6 pA ($n = 13$); Extended Data Fig. 7). It is unclear, however, whether this species is present under physiological conditions.

Our structural data and single-channel recordings imply that CsgG forms an ungated peptide diffusion channel. In PA₆₃, a model peptide diffusion channel, polypeptide passage depends on a pH-driven Brownian ratchet that rectifies the diffusive steps in the translocation channel^{20–22}. However, such proton gradients are not present at the outer membrane, requiring an alternative driving force. Whereas at elevated concentrations CsgG facilitates a non-selective diffusive leakage of periplasmic polypeptides, secretion is specific for CsgA under native conditions and requires the periplasmic factor CsgE^{16,23}. In the presence of excess CsgE, purified CsgG forms a more slowly migrating species on native PAGE (Fig. 4a). SDS–PAGE analysis shows this new species consists of a CsgG–CsgE complex that is present in an equimolar stoichiometry (Fig. 4b). Cryo-electron microscopy (cryo-EM) visualization of CsgG–CsgE isolated by pull-down affinity purification revealed a nine-fold symmetrical particle corresponding to the CsgG nonamer and an additional capping density at the entrance to the periplasmic vestibule, similar in size and shape to a C₉ CsgE oligomer also observed by single-particle EM and size-exclusion chromatography (Fig. 4c–e and Extended Data Fig. 8). The location of the observed CsgG–CsgE contact interface was corroborated by blocking point mutations in CsgG helix 2 (Extended Data Fig. 8). In agreement with a capping function, single-channel recordings showed that CsgE binding to the translocator led to the specific silencing of its ion conductance (Fig. 4f and Extended Data Fig. 7). This CsgE capping of the channel seemed to be an all-or-none response in function of CsgE nonamer binding. At saturation, CsgE binding induced full blockage of the channel, independent of voltage sign, ruling out the possibility that purely electrophoretically or electroosmotically driven CsgE proteins clog the pore. At about 10 nM, an equilibrium between CsgE binding and dissociation events resulted in an

intermittently blocked or fully open translocator. At 1 nM or below, transient (<1 ms) partial blockage events may have stemmed from short-lived encounters with monomeric CsgE.

Thus, CsgG and CsgE seem to form an encaging complex enclosing a central cavity of $\sim 24,000 \text{ \AA}^3$, reminiscent in appearance to the substrate-binding cavity and encapsulating lid structure seen in the GroEL chaperonin and GroES co-chaperonin²⁴. The CsgG–CsgE enclosure would be compatible with the full or partial entrapment of the 129-residue CsgA. The caging of a translocation substrate has recently been observed in ABC toxins²⁵. Spatial confinement of an unfolded polypeptide leads to a decrease in its conformational space, creating an entropic potential that has been shown to favour polypeptide folding in the case of chaperonins^{24,26}. Similarly, we speculate that in curli transport the local high concentration and conformational confinement of curli subunits in the CsgG vestibule would generate an entropic free-energy gradient over the translocation channel (Fig. 4g). On capture into the constriction, the polypeptide chain is then expected to move progressively outwards by Brownian diffusion, rectified by the entropic potential generated from the CsgE-mediated confinement and/or substrate trapping near the secretion channel. For full confinement in the pre-constriction cavity, the escape of an unfolded 129-residue polypeptide to the bulk solvent would correspond to an entropic free-energy release of up to $\sim 80 \text{ kcal mol}^{-1}$ (about 340 kJ mol^{-1} ; ref. 27). The initial entropic cost of substrate docking and confinement are likely to be at least partly compensated for by binding energy released during assembly of the CsgG–CsgE–CsgA complex and an already lowered CsgA entropy in the periplasm. On theoretical grounds, three potential routes of CsgA recruitment to the secretion complex can be envisaged (Extended Data Fig. 9).

Curli-induced biofilms form a fitness and virulence factor in pathogenic Enterobacteriaceae^{4,5}. Their unique secretion and assembly properties are also rapidly gaining interest for (bio)technological application^{23,28,29}. Our structural characterization and biochemical study of two key secretion components provide a tentative model of an iterative mechanism for the membrane translocation of unfolded protein substrates in the absence of a hydrolysable energy source, a membrane potential or an ion gradient (Fig. 4e and Extended Data Fig. 9). The full validation and deconstruction of the contributing factors in the proposed secretion model will require the *in vitro* reconstitution of the translocator to allow transport kinetics to be followed accurately at the single-molecule level.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

METHODS

Cloning and strains

Expression constructs for the production of outer membrane localized C-terminally StrepII-tagged CsgG (pPG1) and periplasmic C-terminally StrepII-tagged CsgG_{C1S} (pPG2) have been described in ref. 19. For selenomethionine labelling, StrepII-tagged CsgG_{C1S} was expressed in the cytoplasm because of increased yields. Therefore, pPG2 was altered to remove the N-terminal signal peptide using inverse PCR with primers 5'-TCT TTA AC

CGC CCC GCC TAA AG-3' (forward) and 5'-CAT TTT TTG CCC TCG TTA TC-3' (reverse) (pPG3). For phenotypic assays, a *csgG* deletion mutant of *E. coli* BW25141 (*E. coli* NVG2) was constructed by the method described in ref. 30 (with primers 5'-AAT AAC TCA ACC GAT TTT TAA GCC CCA GCT TCA TAA GGA AAA TAA TCG TGT AGG CTG GAG CTG CTT C-3' and 5'-CGC TTA AAC AGT AAA ATG CCG GAT GAT AAT TCC GGC TTT TTT ATC TGC ATA TGA ATA TCC TCC TTA G-3'). The various CsgG substitution mutants used for Cys accessibility assays and for phenotypic probing of the channel constriction were constructed by site-directed mutagenesis (QuikChange protocol; Stratagene) starting from pMC2, a pTRC99a vector containing *csgG* under control of the *trc* promoter¹⁴.

Protein expression and purification

CsgG and CsgG_{C1S} were expressed and purified as described¹⁹. In brief, CsgG was recombinantly produced in *E. coli* BL 21 (DE3) transformed with pPG1 and extracted from isolated outer membranes with the use of 1% n-dodecyl- β -D-maltoside (DDM) in buffer A (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM dithiothreitol (DTT)). StrepII-tagged CsgG was loaded onto a 5 ml Strep-Tactin Sepharose column (Iba GmbH) and detergent-exchanged by washing with 20 column volumes of buffer A supplemented with 0.5% tetraethylene glycol mono-octyl ether (C8E4; Affymetrix) and 4 mM lauryldimethylamine-*N*-oxide (LDAO; Affymetrix). The protein was eluted by the addition of 2.5 mM D-desthiobiotin and concentrated to 5 mg ml⁻¹ for crystallization experiments. For selenomethionine labelling, CsgG_{C1S} was produced in the Met auxotrophic strain B834 (DE3) transformed with pPG3 and grown on M9 minimal medium supplemented with 40 mg l⁻¹ L-selenomethionine. Cell pellets were resuspended in 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, supplemented with cOmplete Protease Inhibitor Cocktail (Roche) and disrupted by passage through a TS series cell disruptor (Constant Systems Ltd) operated at 20×10^3 lb in⁻². Labelled CsgG_{C1S} was purified as described¹⁹. DTT (5 mM) was added throughout the purification procedure to avoid oxidation of selenomethionine.

CsgE was produced in *E. coli* NEBC2566 cells harbouring pNH27 (ref. 16). Cell lysates in 25 mM Tris-HCl pH 8.0, 150 mM NaCl, 25 mM imidazole, 5% (v/v) glycerol were loaded on a HisTrap FF (GE Healthcare). CsgE-his was eluted with a linear gradient to 500 mM imidazole in 20 mM Tris-HCl pH 8.0, 150 mM NaCl, 5% (v/v) glycerol buffer. Fractions containing CsgE were supplemented with 250 mM (NH₄)₂SO₄ and applied to a 5 ml HiTrap Phenyl HP column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 250 mM (NH₄)₂SO₄, 5% (v/v) glycerol. A linear gradient to 20 mM Tris-HCl pH 8.0, 10 mM NaCl, 5% (v/v) glycerol was applied for elution. CsgE containing fractions were loaded onto a Superose 6 Prep Grade 10/600 (GE Healthcare) column equilibrated in 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 5% (v/v) glycerol.

In-solution oligomeric state assessment

About 0.5 mg each of detergent-solubilized CsgG (0.5% C8E4, 4 mM LDAO) and CsgG_{C1S} were applied to a Superdex 200 10/300 GL analytical gel filtration column (GE Healthcare) equilibrated with 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT, 4 mM LDAO and

0.5% C8E4 (CsgG) or with 25 mM Tris-HCl pH 8.0, 200 mM NaCl (CsgG_{C1S}), and run at 0.7 ml min⁻¹. The column elution volumes were calibrated with bovine thyroglobulin, bovine γ -globulin, chicken ovalbumin, horse myoglobin and vitamin B₁₂ (Bio-Rad) (Extended Data Fig. 2). Membrane-extracted CsgG, 20 μ g of the detergent-solubilized protein was also run on 3–10% blue native PAGE using the procedure described in ref. 31 (Extended Data Fig. 2). NativeMark (Life Technologies) unstained protein standard (7 μ l) was used for molecular mass estimation.

Crystallization, data collection and structure determination

Selenomethionine-labelled CsgG_{C1S} was concentrated to 3.8 mg ml⁻¹ and crystallized by sitting-drop vapour diffusion against a solution containing 100 mM sodium acetate pH 4.2, 8% PEG 4000 and 100 mM sodium malonate pH 7.0. Crystals were incubated in crystallization buffer supplemented with 15% glycerol and flash-frozen in liquid nitrogen. Detergent-solubilized CsgG was concentrated to 5 mg ml⁻¹ and crystallized by hanging-drop vapour diffusion against a solution containing 100 mM Tris-HCl pH 8.0, 8% PEG 4000, 100 mM NaCl and 500 mM MgCl₂. Crystals were flash-frozen in liquid nitrogen and cryoprotected by the detergent present in the crystallization solution. For optimization of crystal conditions and screening for crystals with good diffraction quality, crystals were analysed on beamlines Proxima-1 and Proxima-2a (Soleil, France), PX-I (Swiss Light Source, Switzerland), I02, I03, I04 and I24 (Diamond Light Source, UK) and ID14eh2, ID23eh1 and ID23eh2 (ESRF, France). Final diffraction data used for structure determination of CsgG_{C1S} and CsgG were collected at beamlines I04 and I03, respectively (see Extended Data Fig. 10a for data collection and refinement statistics). Diffraction data for CsgG_{C1S} were processed using Xia2 and the XDS package^{32,33}. Crystals of CsgG_{C1S} belonged to space group *P1* with unit cell dimensions of $a = 101.3$ Å, $b = 103.6$ Å, $c = 141.7$ Å, $\alpha = 111.3^\circ$, $\beta = 90.5^\circ$, $\gamma = 118.2^\circ$, containing 16 protein copies in the asymmetric unit. For structure determination and refinement, data collected at 0.9795 Å wavelength were truncated at 2.8 Å on the basis of an $I/\sigma I$ cutoff of 2 in the highest-resolution shell. The structure was solved using experimental phases calculated from a single anomalous dispersion (SAD) experiment. A total of 92 selenium sites were located in the asymmetric unit by using ShelxC and ShelxD³⁴, and were refined and used for phase calculation with Sharp³⁵ (phasing power 0.79, figure of merit (FOM) 0.25). Experimental phases were density modified and averaged by non-crystallographic symmetry (NCS) using Parrot³⁶ (Extended Data Fig. 10; FOM 0.85). An initial model was built with Buccaneer³⁷ and refined by iterative rounds of maximum-likelihood refinement with Phenix refine³⁸ and manual inspection and model (re)building in Coot³⁹. The final structure contained 28,853 atoms in 3,700 residues belonging to 16 CsgG_{C1S} chains (Extended Data Fig. 2), with a molprobity40 score of 1.34; 98% of the residues lay in favoured regions of the Ramachandran plot (99.7% in allowed regions). Electron density maps showed no unambiguous density corresponding to possible solvent molecules, and no water molecules or ions were therefore built in. Sixteen fold NCS averaging was maintained throughout refinement, using strict and local NCS restraints in early and late stages of refinement, respectively.

Diffraction data for CsgG were collected from a single crystal at 0.9763 Å wavelength and were indexed and scaled, using the XDS package^{32,33}, in space group *C2* with unit-cell

dimensions $a = 161.7 \text{ \AA}$, $b = 372.3 \text{ \AA}$, $c = 161.8 \text{ \AA}$ and $\beta = 92.9^\circ$, encompassing 18 CsgG copies in the asymmetric unit and a 72% solvent content. Diffraction data for structure determination and refinement were elliptically truncated to resolution limits of $3.6 \text{ \AA}$, $3.7 \text{ \AA}$ and $3.8 \text{ \AA}$ along reciprocal cell directions a^* , b^* and c^* and scaled anisotropically with the Diffraction Anisotropy Server⁴¹. Molecular replacement using the CsgG_{C1S} monomer proved unsuccessful. Analysis of the self rotation function revealed D_9 symmetry in the asymmetric unit (not shown). On the basis of on the CsgG_{C1S} structure, a nonameric search model was generated in the assumption that after going from a C_8 to C_9 oligomer, the interprotomer arc at the particle circumference would stay approximately the same as the interprotomer angle changed from 45° to 40° , giving a calculated increase in radius of about $4 \text{ \AA}$. Using the calculated nonamer as search model, a molecular replacement solution containing two copies was found with Phaser⁴². Inspection of density-modified and NCS-averaged electron density maps (Parrot³⁶; Extended Data Fig. 10) allowed manual building of the TM1 and TM2 and remodelling of adjacent residues in the protein core, as well as the building of residues 2–18, which were missing from the CsgG_{C1S} model and linked the $\alpha 1$ helix to the N-terminal lipid anchor. Refinement of the CsgG model was performed with Buster-TNT⁴³ and Refmac5 (ref. 44) for initial and final refinement rounds, respectively. Eighteen fold local NCS restraints were applied throughout refinement, and Refmac5 was run employing a jelly-body refinement with sigma 0.01 and hydrogen-bond restraints generated by ProSMART⁴⁵. The final structure contained 34,165 atoms in 4,451 residues belonging to 18 CsgG chains (Extended Data Fig. 2), with a molprobity score of 2.79; 93.0% of the residues lay in favoured regions of the Ramachandran plot (99.3% in allowed regions). No unambiguous electron density corresponding the N-terminal lipid anchor could be discerned.

Congo red assay

For analysis of Congo red binding, a bacterial overnight culture grown at 37°C in Lysogeny Broth (LB) was diluted in LB medium until a D_{600} of 0.5 was reached. A $5 \mu\text{l}$ sample was spotted on LB agar plates supplemented with ampicillin (100 mg l^{-1}), Congo red (100 mg l^{-1}) and 0.1% (w/v) isopropyl β -D-thiogalactoside (IPTG). Plates were incubated at room temperature ($20\text{--}22^\circ \text{C}$) for 48 h to induce curli expression. The development of the colony morphology and dye binding were observed at 48 h.

Cysteine accessibility assays

Cysteine mutants were generated in pMC2 using site-directed mutagenesis and expressed in *E. coli* LSR12 (ref. 7). Bacterial cultures grown overnight were spotted onto LB agar plates containing 1 mM IPTG and 100 mg l^{-1} ampicillin. Plates were incubated at room temperature and cells were scraped after 48 h, resuspended in 1 ml of PBS and normalized using D_{600} . The cells were lysed by sonication and centrifuged for 20 s at $3,000g$ at 4°C to remove unbroken cells from cell lysate and suspended membranes. Proteins in the supernatant were labelled with 15 mM methoxypolyethylene glycolmaleimide (MAL-PEG 5 kDa) for 1 h at room temperature. The reaction was stopped with 100 mM DTT and centrifuged at $40,000 \text{ r.p.m.}$ ($\sim 100,000g$) in a 50.4 Ti rotor for 20 min at 4°C to pellet total membranes. The pellet was washed with 1% sodium lauroyl sarcosinate to solubilize cytoplasmic membranes and centrifuged again. The resulting outer membranes were

resuspended and solubilized using PBS containing 1% DDM. Metal-affinity pull downs with nickel beads were used for SDS-PAGE and anti-His western blots. *E. coli* LSR12 cells with empty pMC2 vector were used as negative control.

ATR-FTIR spectroscopy

ATR-FTIR measurements were performed on an Equinox 55 infrared spectrophotometer (Bruker), continuously purged with dried air, equipped with a liquid-nitrogen-refrigerated mercury cadmium telluride detector and a Golden Gate reflectance accessory (Specac). The internal reflection element was a diamond crystal (2 mm × 2 mm) and the beam incidence angle was 45°. Each purified protein sample (1 µl) was spread at the surface of the crystal and dried under a gaseous nitrogen flow to form a film. Each spectrum, recorded at 2 cm⁻¹ resolution, was an average of 128 accumulations for improved signal-to-noise ratio. All the spectra were treated with water vapour contribution subtraction, smoothed at a final resolution of 4 cm⁻¹ by apodization and normalized on the area of the Amide I band (1,700–1,600 cm⁻¹) to allow their comparison⁴⁶.

Negative stain EM and symmetry determination

Negative stain EM was used to monitor in-solution oligomerization states of CsgG, CsgG_{C15} and CsgE. CsgE, CsgG_{C15} and amphipol-bound CsgG were adjusted to a concentration of 0.05 mg ml⁻¹ and applied to glow-discharged carbon-coated copper grids (CF-400; Electron Microscopy Sciences). After 1 min incubation, samples were blotted, then washed and stained in 2% uranyl acetate. Images were collected on a Tecnai T12 BioTWIN LaB6 microscope operating at a voltage of 120 kV, at a nominal magnification of ×49,000 and defocus between 800 and 2,000 nm. Contrast transfer function (CTF), phase flipping and particle selection were performed as described for cryo-EM. For membrane-extracted CsgG, octadecameric particles (1,780 in all) were analysed separately from nonamers and top views. For purified CsgE a total of 2,452 particles were analysed. In all cases, after normalization and centring, images were classified using IMAGIC-4D as described in the cryo-EM section. The best classes corresponding to characteristic views were selected for each set of particles. Symmetry determination of CsgG top views was performed using the best class averages with roughly 20 images per class. The rotational autocorrelation function was calculated using IMAGIC and plotted.

CsgG-CsgE complex formation

For CsgG-CsgE complex formation, the solubilizing detergents in purified CsgG were exchanged for Amphipols A8-35 (Anatrace) by adding 120 µl of CsgG (24 mg ml⁻¹ protein in 0.5% C8E4, 4 mM LDAO, 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT) to 300 µl of detergent-destabilized liposomes (1 mg ml⁻¹ 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 0.4% LDAO) and incubating for 5 min on ice before the addition of 90 µl of A8-35 amphipols at 100 mg ml⁻¹ stock. After an additional 15 min incubation on ice, the sample was loaded on a Superose 6 10/300 GL (GE Healthcare) column and gel filtration was performed in 200 mM NaCl, 2.5% xylitol, 25 mM Tris-HCl pH 8, 0.2 mM DTT. An equal volume of purified monomeric CsgE in 200 mM NaCl, 2.5% xylitol, 25 mM Tris-HCl pH 8, 0.2 mM DTT was added to the amphipol-solubilized CsgG at

final protein concentrations of 15 and 5 μ M for CsgE and CsgG, respectively, and the sample was run at 125 V at 18 °C on a 4.5% native PAGE in 0.5 $\times$ TBE buffer. For the second, denaturing dimension, the band corresponding to the CsgG–CsgE complex was cut out of unstained lanes run in parallel on the same gel, boiled for 5 min in Laemmli buffer (60 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 5% 2-mercaptoethanol, 0.01% bromophenol blue) and run on 4–20% SDS–PAGE. Purified CsgE and CsgG were run alongside the complex as control samples. Gels were stained with InstantBlue Coomassie for visual inspection or SYPRO orange for stoichiometry assessment of the CsgG–CsgE complex by fluorescence detection (Typhoon FLA 9000) of the CsgE and CsgG bands on SDS–PAGE, yielding a CsgG/CsgE ratio of 0.97.

CsgG–CsgE Cryo-EM

Cryo-electron microscopy was used to determine the in-solution structure of the C_9 CsgG–CsgE complex. CsgG–CsgE complex prepared as described above was bound and eluted with buffer supplemented with 100 mM imidazole from a TALON cobalt metal affinity resin to remove unbound CsgG, and on elution it was immediately applied to Quantifoil R2/2 carbon coated grids (Quantifoil Micro Tools GmbH) that had been glow-discharged at 20 mA for 30 s. Samples were plunge-frozen in liquid nitrogen using an automated system (Leica) and observed under a FEI F20 microscope operating at a voltage of 200 kV, a nominal magnification of $\times 50,000$ under low-dose conditions and a defocus range of 1.4–3 μ m. Image frames were recorded on a Falcon II detector. The pixel size at the specimen level was 1.9 Å per pixel. The CTF parameters were assessed using CTFFIND3 (ref. 47), and the phase flipping was done in SPIDER⁴⁸. Particles were automatically selected from CTF-corrected micrographs using BOXER (EMAN2; ref. 49). Images with an astigmatism of more than 10% were discarded. A total of 4,881 particles were selected from 75 micrographs and windowed into 128-pixel $\times$ 128-pixel boxes. Images were normalized to the same mean and standard deviation and high-pass filtered at a low-resolution cut-off of ~ 200 Å. They were centred and then subjected to a first round of MSA. An initial reference set was obtained using reference free classification in IMAGIC-4D (Image Science Software). The best classes corresponding to characteristic side views of the C_9 cylindrical particles were used as references for the MRA. For CsgG–CsgE complex, the first three-dimensional model was calculated from the best 125 characteristic views (with good contrast and well-defined features) encompassing 1,221 particles of the complex with orientations determined by angular reconstitution (Image Science Software). The three-dimensional map was refined by iterative rounds of MRA, MSA and anchor set refinement. The resolution was estimated to be 24 Å by Fourier shell correlation (FSC) according to the 0.5 criteria level (Extended Data Fig. 7). Visualization of the map and figures was performed in UCSF Chimera⁵⁰.

Bile salt toxicity assay

Outer-membrane permeability was investigated by decreased growth on agar plates containing bile salts. Tenfold serial dilutions of *E. coli* LSR12 (ref. 7) cells (5 μ l) harbouring both pLR42 (ref. 16) and pMC2 (ref. 14) (or derived helix 2 mutants) were spotted on McConkey agar plates containing 100 μ g l⁻¹ ampicillin, 25 μ g l⁻¹ chloramphenicol, 1 mM IPTG with or without 0.2% (w/v) L-arabinose. After incubation overnight at 37 °C, colony growth was examined.

Single-channel current recordings

Single-channel current recordings were performed using parallel high-resolution electrical recording with the Orbit 16 device from Nanion. In brief, horizontal bilayers of 1,2-diphytanoyl-*sn*-glycero-3-phosphocholine (Avanti Polar Lipids) were formed over microcavities (of subpicolitre volume) in a 16-channel multielectrode cavity array (MECA) chip (Ionera)⁵¹. Both the *cis* and *trans* cavities above and below the bilayer contained 1.0 M KCl, 25 mM Tris-HCl pH 8.0. To insert channels into the membrane, CsgG dissolved in 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 1 mM DTT, 0.5% C8E4, 5 mM LDAO was added to the *cis* compartment to a final concentration of 90–300 nM. To test the interaction of the CsgG channel with CsgE, a solution of the latter protein dissolved in 25 mM Tris-HCl pH 8.0, 150 mM NaCl was added to the *cis* compartment to final concentrations of 0.1, 1, 10 and 100 nM. Transmembrane currents were recorded at a holding potential of +50 mV and –50 mV (with the *cis* side grounded) using a Tecella Triton 16-channel amplifier at a low-pass filtering frequency of 3 kHz and a sampling frequency of 10 kHz. Current traces were analysed using the Clampfit of the pClamp suite (Molecular Devices). Plots were generated using Origin 8.6 (Microcal)⁵².

Measured currents were compared with those calculated based on the pore dimensions of the CsgG X-ray structure, modelled to be composed of three segments: the transmembrane section, the periplasmic vestibule, and the inner channel constriction connecting the two. The first two segments were modelled to be of conical shape; the constriction was represented as a cylinder. The corresponding resistances R_1 , R_2 and R_3 , respectively, were calculated as

$$\begin{aligned} R_1 &= L_1 / (\pi D_1 d_1 \kappa) \\ R_2 &= L_2 / (\pi D_2 d_2 \kappa) \\ R_3 &= L_3 / (\pi d_1 d_2 \kappa) \end{aligned}$$

where L_1 , L_2 and L_3 are the axial lengths of the segments, measuring 3.5, 4.0 and 2.0 nm, respectively, and D_1 , d_1 , D_2 and d_2 are the maximum and minimum diameters of segments 1 and 2, measuring 4.0, 0.8, 3.5 and 0.8 nm, respectively. The conductivity κ has the macroscopic bulk value of 10.6 S m^{-1} for the wider conical segments. The conductivity was half this value for the narrow central constriction, owing to the reduced ion mobility, in line with findings for the OmpF pore of similar dimensions⁵³. The current was calculated by inserting R_1 , R_2 and R_3 and voltage $V = 50 \text{ mV}$ into

$$I = V / (R_1 + R_2 + R_3)$$

Access resistance also included in the calculations.

Extended Data

Extended Data Figure 1. Curli biosynthesis pathway in *E. coli*

The major curli subunit CsgA (light green) is secreted from the cell as a soluble monomeric protein. The minor curli subunit CsgB (dark green) is associated with the outer membrane (OM) and acts as a nucleator for the conversion of CsgA from a soluble protein to amyloid deposit. CsgG (orange) assembles into an oligomeric curli-specific translocation channel in the outer membrane. CsgE (purple) and CsgF (light blue) form soluble accessory proteins required for productive CsgA and CsgB transport and deposition. CsgC forms a putative oxidoreductase of unknown function. All curli proteins have putative Sec signal sequences for transport across the cytoplasmic (inner) membrane (IM).

Extended Data Figure 2. In-solution oligomerization states of CsgG and CsgG_{C1S} analysed by size-exclusion chromatography and negative-stain electron microscopy

a, Raw negative-stain EM image of C8E4/LDAO-solubilized CsgG. Arrows indicate the different particle populations as labelled in the size exclusion profile shown in **g**, being (I) aggregates of CsgG nonamers, (II) CsgG octadecamers and (III) CsgG nonamers. Scale bar, 20 nm. **b**, Representative class average for top and side views of the indicated oligomeric states. **c**, Rotational autocorrelation function graph of LDAO-solubilized CsgG in top view, showing nine-fold symmetry. **d**, Raw negative-stain EM image of CsgG_{C1S}. Arrows indicate the hexadecameric (IV) and octameric (V) particles observed by size-exclusion chromatography in **g**. **e**, Representative class average for side views of CsgG_{C1S} oligomers. No top views were observed for this construct. **f**, Table of elution volumes (EV) of CsgG_{C1S} and CsgG particles observed by size-exclusion chromatography shown in **g**, calculated

molecular mass (MW_{calc}), expected molecular mass (MW_{CsgG}) corresponding CsgG oligomerization state ($CsgG_n$) and the particles' symmetry as observed by negative-stain EM and X-ray crystallography. **g**, Size-exclusion chromatogram of CsgG_{C1S} (black) and C8E4/LDAO-solubilized CsgG (grey) run on Superdex 200 10/300 GL (GE Healthcare). **h, i**, Ribbon representation of crystallized oligomers in top and side view, showing the D_8 hexadecamers for CsgG_{C1S} (**h**) and D_9 octadecamers for membrane-extracted CsgG (**i**). One protomer is coloured in rainbow from N terminus (blue) to C terminus (red). The two C_8 octamers (CsgG_{C1S}) or C_9 nonamers (CsgG) that form the tail-to-tail dimers captured in the crystals are coloured blue and tan. r and θ give radius and interprotomer rotation, respectively.

Extended Data Figure 3. Comparison of CsgG with structural homologues and interprotomer contacts in CsgG

a, b, Ribbon diagram for the CsgG_{C1S} monomer (for example CsgG in pre-pore conformation) (**a**) and the nucleotide-binding-domain-like domain of TolB (**b**) (PDB 2hqs), both coloured in rainbow from N terminus (blue) to C terminus (red). Common secondary structure elements are labelled equivalently. **c**, CsgG_{C1S} (grey) in superimposition with, from left to right, *Xanthomonas campestris* rare lipoprotein B (PDB 2r76, coloured pink), *Shewanella oneidensis* hypothetical lipoprotein DUF330 (PDB 2iqi, coloured pink) and *Escherichia coli* TolB (PDB 2hqs, coloured pink and yellow for the N-terminal and β -propeller domains, respectively). CsgG-specific structural elements are labelled and coloured as in the upper left panel. **d, e**, Ribbon diagram of two adjacent protomers as found

in the CsgG structure, viewed along the plane of the bilayer, either from outside (**c**) or inside (**d**) the oligomer. One protomer is shown in rainbow (dark blue to red) from N terminus to C terminus; a second protomer is shown in light blue (core domain), blue (helix 2) and tan (TM domain). Four main oligomerization interfaces are apparent: $\beta 6$ – $\beta 3'$ main-chain interactions inside the β -barrel, the constriction loop (CL), side-chain packing of helix 1 ($\alpha 1$) against $\beta 1$ – $\beta 3$ – $\beta 4$ – $\beta 5$, and helix–helix packing of helix 2 ($\alpha 2$). The 18-residue N-terminal loop connecting the lipid anchor (a magenta sphere shows the Ca position of Leu 2) and N-terminal helix (αN) is also seen to wrap over the adjacent two protomers. The projected position of the lipid anchor is expected to lie against the TM1 and TM2 hairpins of the +2 protomer (not shown for clarity).

Extended Data Figure 4. Cys accessibility assays for selected surface residues in the CsgG oligomers

a–c, Ribbon representation of CsgG nonamers shown in periplasmic (**a**), side (**b**) and extracellular (**c**) views. One protomer is coloured in rainbow from N terminus (blue) to C terminus (red). Cysteine substitutions are labelled and the equivalent locations of the S atoms are shown as spheres, coloured according to accessibility to MAL-PEG (5,000 Da) labelling in *E. coli* outer membranes. **d**, Western blot of MAL-PEG reacted samples analysed on SDS–PAGE, showing 5 kDa increase on MAL-PEG binding of the introduced cysteine. Accessible (++ and +++), moderately accessible (+) and inaccessible (–) sites are coloured green, orange and red, respectively, in **a–c**. For Arg 97 and Arg 110 a second species at 44 kDa is present, corresponding to a fraction of protein in which both the introduced and native cysteine became labelled. Data are representative of four independent

experiments from biological replicates. **e**, Side view of the dimerization interface in the D_9 octadecamer as present in the X-ray structure. Introduced cysteines in the dimerization interface or inside the lumen of the D_9 particle are labelled. In membrane-bound CsgG, these residues are accessible to MAL-PEG, demonstrating that the D_9 particles are an artefact of concentrated solutions of membrane-extracted CsgG and that the C_9 complex forms the physiologically relevant species. Residues in the C-terminal helix (α C; Lys 242, Asp 248 and His 255) are found to be inaccessible to poorly accessible, indicating that α C may form additional contacts with the *E. coli* cell envelope, possibly the peptidoglycan layer.

Extended Data Figure 5. Molecular dynamics simulation of CsgG constriction with model polyaniline chain

a, b, Top (**a**) and side (**b**) views of the CsgG constriction modelled with a polyaniline chain threaded through the channel in an extended conformation, here shown in a C-terminal to N-terminal direction. Substrate passage through the CsgG transporter is itself not sequence specific^{16,23}. For clarity, a polyaniline chain was used for modelling the putative interactions of a passing polypeptide chain. The modelled area is composed of nine concentric CsgG C-loops, each comprising residues 47–58. Side chains lining the constriction are shown in stick representation, with Phe 51 coloured slate blue, Asn 55 (amide-clamp) cyan, and Phe 48 and Phe 56 (ϕ -clamp) in light and dark orange, respectively. N, O and H atoms (only hydroxyl or side-chain amide H atoms are shown) are coloured blue, red and white, respectively. The polyaniline chain is coloured green, blue, red and white for C, N, O and H atoms, respectively. Solvent molecules (water) within 10 Å of the polyaniline residues inside the constriction (residues labelled +1 to +5) are shown as red dots. **c**, Modelled solvation of the polyaniline chain, position as in **b** and with C-loops removed for clarity (shown solvent molecules are those within 10 Å of the full polyaniline chain). At the height of the amide-clamp and ϕ -clamp, the solvation of the polyaniline chain is reduced to a single water shell that bridges the peptide backbone and amide-clamp side chains. Most side chains in the Tyr 51 ring have rotated towards the solvent in comparison with their inward, centre-pointing position observed in the CsgG (and the CsgG_{C1S}) X-ray structure. The model is the result of a 40 ns all-atom explicit solvent molecular dynamics simulation with GROMACS⁵⁴ using the AMBER99SB-ILDN⁵⁵ force field and with the Ca

atoms of the residues at the extremity of the C-loop (Gln 47 and Thr 58) positionally restricted.

Extended Data Figure 6. Sequence conservation in CsgG homologues

a, Surface representation of the CsgG nonamer coloured according to sequence similarity (coloured yellow to blue from low to high conservation score)⁵⁶ and viewed from the periplasm (far left), the side (middle left), the extracellular milieu (middle right) or as a cross-sectional side view (far right). The figures show that the regions of highest sequence conservation map to the entry of the periplasmic vestibule, the vestibular side of the constriction loop and the luminal surface of the TM domain. **b**, Multiple sequence alignment of CsgG-like lipoproteins. The selected sequences were chosen from monophyletic clades across the phylogenetic tree of CsgG-like sequences (not shown), to give a representative view of sequence diversity. Secondary structure elements are shown as arrows or bars for β -strands and α -helices, respectively, and are based on the *E. coli* CsgG crystal structure. **c**, **d**, CsgG protomer in secondary structure representation (**c**) and a cross-sectional side view (**d**) of the CsgG nonamer in surface representation, both coloured grey and with three continuous blocks of high sequence conservation coloured red (HCR1), blue (HCR2) and yellow (HCR3). HCR1 and HCR2 shape the vestibular side of the constriction loop; HCR3 corresponds to helix 2, lying at the entry of the periplasmic vestibule. Inside the constriction, Phe 56 is 100% conserved, whereas Asn 55 can be conservatively replaced by Ser or Thr, for example by a small polar side chain that can act as hydrogen-bond donor/acceptor. The concentric side-chain ring at the exit of the constriction (Tyr 51) is not conserved. The

presence of the Phe-ring at the entrance of the constriction is topologically similar to the Phe 427-ring (referred to as the φ -clamp) in the anthrax protective antigen PA₆₃, in which it was shown to catalyse polypeptide capture and passage²⁰. MST of *tox*B superfamily proteins reveals a conserved motif D(D/Q)(F)(S/N)S at the height of the Phe-ring. This is similar to the S(Q/N/T)(F)ST motif seen in curli-like transporters. Although an atomic resolution structure of PA₆₃ in pore conformation is not yet available, available structures suggest the Phe-ring may similarly be followed by a conserved hydrogen-bond donor/acceptor (Ser/Asn 428) as a subsequent concentric ring in the translocation channel (note that the orientation of the element is inverted in both transporters).

Extended Data Figure 7. Single-channel current analysis of CsgG and CsgG + CsgE pores

a, Under negative field potential, CsgG pores show two conductance states. The upper left and right panels show a representative single-channel current trace of, respectively, the normal (measured at +50, 0 and -50 mV) and the low-conductance forms (measured at 0, +50 and -50 mV). No conversions between both states were observed during the total observation time ($n = 22$), indicating that the conductance states have long lifetimes (second to minute timescale). The lower left panel shows a current histogram for the normal and

low-conductance forms of CsgG pores acquired at +50 and –50 mV ($n = 33$). *I–V* curves for CsgG pores with regular and low conductance are shown in the lower right panel. These data represent averages and standard deviations from at least four independent recordings. The nature or physiological existence of the low-conductance form is unknown. **b**, Electrophysiology of CsgG channels titrated with the periplasmic factor CsgE. The plots display the normalized occurrence, that is, the fractions of open, closed and intermediate-state channels, as a function of CsgE concentration. Open and closed states of CsgG are illustrated in Fig. 4f. Increasing the concentration of CsgE to more than 10 nM leads to the closure of CsgG pores. The effect occurs at +50 mV (left) and –50 mV (right), ruling out the possibility that the pore blockade is caused by electroosmosis or electrophoresis of CsgE (calculated pI 4.7) into the CsgG pore. An infrequent (<5%) intermediate state has roughly half the conductance of the open channel. It may represent CsgE-induced incomplete closures of the CsgG channel. Alternatively, it could represent the temporary formation of a CsgG dimer caused by the binding of residual CsgG monomer from the electrolyte solution to the membrane-embedded pore. The fractions for the three states were obtained from all-point histogram analysis of single-channel current traces. The histograms yielded peak areas for up to three states, and the fraction for a given state was obtained by dividing the corresponding peak area by the sum of all other states in the recording. Under negative field potential, two open conductance states are discerned, similar to the observations for CsgG (see **a**). Because both open channel variations were blocked by higher CsgE concentrations, the ‘open’ traces in **b** combine both conductance forms. The data in the plot represent averages and standard deviations from three independent recordings. **c**, The crystal structure, size-exclusion chromatography and EM show that detergent extracted CsgG pores form non-native tail-to-tail stacked dimers (for example, two nonamers as D_9 particle; Extended Data Fig. 2) at higher protein concentration. These dimers can also be observed in single-channel recordings. The upper panel shows the single-channel current trace of stacked CsgG pores at +50, 0 and –50 mV (left to right). The lower left panel shows a current histogram of dimeric CsgG pores recorded at +50 and –50 mV. The experimental conductances of $+16.2 \pm 1.8$ and -16.0 ± 3.0 pA ($n = 15$) at +50 and –50 mV, respectively, are near the theoretically calculated value of 23 pA. The lower right panel shows an *I–V* curve for the stacked CsgG pores. The data represent averages and standard deviations from six independent recordings. **d**, The ability of CsgE to bind and block stacked CsgG pores was tested by electrophysiology. Shown are single-channel current traces of stacked CsgG pore in the presence of 10 or 100 nM CsgE at +50 mV (upper) and –50 mV (lower). The current traces indicate that otherwise saturating concentrations of CsgE do not lead to pore closure for stacked CsgG dimers. These observations are in good agreement with the mapping of the CsgG–CsgE contact zone to helix 2 and the mouth of the CsgG periplasmic cavity as discerned by EM and site-directed mutagenesis (Fig. 4 and Extended Data Fig. 7).

Extended Data Figure 8. CsgE oligomer and CsgG-CsgE complex

a, Size-exclusion chromatography of CsgE (Superose 6, 16/600; running buffer 20 mM Tris-HCl pH 8, 100 mM NaCl, 2.5% glycerol) shows an equilibrium of two oligomeric states, 1 and 2, with an apparent molecular mass ratio of 9.16:1. Negative-stain EM inspection of peak 1 shows discrete CsgE particles (five representative class averages are shown in the inset, ordered by increasing tilt angles) compatible in size with nine CsgE copies. **b**, Selected class average of CsgE oligomer observed in top view by cryo-EM and its rotational autocorrelation show the presence of C_9 symmetry. **c**, FSC analysis of CsgG-CsgE cryo-EM model. Three-dimensional reconstruction achieved a resolution of 24 Å as determined by FSC at a threshold of 0.5 correlation using 125 classes corresponding to 1,221 particles. **d**, Overlay of CsgG-CsgE cryo-EM density and the CsgG nonamer observed in the X-ray structure. The overlays are shown viewed from the side as semi-transparent density (left) or as a cross-sectional view. **e**, Congo red binding of *E. coli* BW25141Δ*csgG* complemented with wild-type *csgG* (WT), empty vector (*ΔcsgG*) or *csgG* helix 2 mutants (single amino acid replacements labelled in single-letter code). Data are representative of four biological replicates. **f**, Effect of bile salt toxicity on *E. coli* LSR12 complemented with *csgG* (WT) or on *csgG* carrying different helix 2 mutations, complemented with (+) or without (-) *csgE*. Tenfold serial dilution starting from 10⁷ bacteria were spotted on McConkey agar plates. Expression of the CsgG pore in the outer membrane leads to an increased bile salt sensitivity that can be blocked by co-expression of CsgE ($n = 6$, three biological replicates, with two repetitions each). **g**, Cross-sectional view of CsgG X-ray structure in molecular surface representation. CsgG mutants without an effect on Congo red binding or toxicity are shown in blue; mutants that interfere with CsgE-mediated rescue of bile salt sensitivity are indicated in red.

Extended Data Figure 9. Assembly and substrate recruitment of the CsgG secretion complex

The curli transporter CsgG and the soluble secretion cofactor CsgE form a secretion complex with 9:9 stoichiometry that encloses a $\sim 24,000 \text{ \AA}^3$ chamber that is proposed to entrap the CsgA substrate and facilitate its entropy-driven diffusion across the outer membrane (OM; see the text and Fig. 4). On theoretical grounds, three putative pathways (a–c) for substrate recruitment and assembly of the secretion complex can be envisaged. **a**, A ‘catch-and-cap’ mechanism entails the binding of CsgA to the apo CsgG translocation channel (1), leading to a conformational change in the latter that exposes a high-affinity binding platform for CsgE binding (2). CsgE binding leads to capping of the substrate cage. On secretion of CsgA, CsgG would fall back into its low-affinity conformation, leading to CsgE dissociation and liberation of the secretion channel for a new secretion cycle. **b**, In a ‘dock-and-trap’ mechanism, periplasmic CsgA is first captured by CsgE (1), causing the latter to adopt a high-affinity complex that docks onto the CsgG translocation pore (2), enclosing CsgA in the secretion complex. CsgA binding could be directly to CsgE oligomers or to CsgE monomers, the latter leading to subsequent oligomerization and CsgG binding. Secretion of CsgA leads CsgE to fall back into its low-affinity conformation and to dissociate from the secretion channel. **c**, CsgG and CsgE form a constitutive complex, in which CsgE conformational dynamics cycle between open and closed forms in the course of CsgA recruitment and secretion. Currently published or available data do not allow us to discriminate between these the putative recruitment modes or derivatives thereof, or to put forward one of them.

Extended Data Figure 10. Data collection statistics and electron density maps of CsgG_{C1S} and CsgG

a, Data collection statistics for CsgG_{C1S} and CsgG X-ray structures. **b**, Electron density map at 2.8 Å for CsgG_{C1S} calculated using NCS-averaged and density-modified experimental SAD phases, and contoured at 1.5 σ . The map shows the region of the channel construction (CL; a single protomer is labelled) and is overlaid on the final refined model. **c**, Electron density map (resolutions 3.6, 3.7 and 3.8 Å along reciprocal vectors *a**, *b** and *c**, respectively) in the CsgG TM domain region, calculated from NCS-averaged and density-modified molecular replacement phases (TM loops were absent from the input model); *B*-factor sharpened by -20 Å² and contoured at 1.0 σ . The figure shows the TM1 (Lys 135–Leu 154) and TM2 (Leu 182–Asn 209) region of a single CsgG protomer, overlaid on the final refined model.

Acknowledgments

This research was supported by VIB through project grant PRJ9 (P.G., N.V.G. and H.R.), by Hercules Foundation through equipment grant UABR/09/005, by National Institutes of Health RO1 grants AI099099 and AI048689 (J.P. and S.J.H.) and AI073847 (M.R.C.), and by Institut Pasteur and Centre national de la recherche scientifique (F.G., G.P.A. and R.F.). S.H. is funded by the Engineering and Physical Sciences Research Council (Institutional Sponsorship Award), the National Physical Laboratory and University College London Chemistry. F.G. is the recipient of a 'Bourse Roux' from Institut Pasteur. P.V.K. was supported by the European Research Council (ERC). We acknowledge Diamond Light Source for time on beamlines I02, I03, I04 and I24 under proposal mx7351, and the Soleil synchrotron for access to Proxima-1 and Proxima-2a under proposals 20100734, 20110924 and 20121253.

References

1. Olsen A, Jonsson A, Normark S. Fibronectin binding mediated by a novel class of surface organelles on *Escherichia coli*. *Nature*. 1989; 338:652–655. [PubMed: 2649795]
2. Collinson SK, et al. Thin, aggregative fimbriae mediate binding of *Salmonella enteritidis* to fibronectin. *J Bacteriol*. 1993; 175:12–18. [PubMed: 8093237]
3. Dueholm MS, Albertsen M, Otzen D, Nielsen PH. Curli functional amyloid systems are phylogenetically widespread and display large diversity in operon and protein structure. *PLoS ONE*. 2012; 7:e51274. [PubMed: 23251478]
4. Cegelski L, et al. Small-molecule inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nature Chem Biol*. 2009; 5:913–919. [PubMed: 19915538]
5. Herwald H, et al. Activation of the contact-phase system on bacterial surfaces—a clue to serious complications in infectious diseases. *Nature Med*. 1998; 4:298–302. [PubMed: 9500602]
6. Hammar M, Arnqvist A, Bian Z, Olsen A, Normark S. Expression of two *csg* operons is required for production of fibronectin- and Congo red-binding curli polymers in *Escherichia coli* K-12. *Mol Microbiol*. 1995; 18:661–670. [PubMed: 8817489]
7. Chapman MR, et al. Role of *Escherichia coli* curli operons in directing amyloid fiber formation. *Science*. 2002; 295:851–855. [PubMed: 11823641]
8. Wang X, Smith DR, Jones JW, Chapman MR. *In vitro* polymerization of a functional *Escherichia coli* amyloid protein. *J Biol Chem*. 2007; 282:3713–3719. [PubMed: 17164238]
9. Dueholm MS, et al. Fibrillation of the major curli subunit CsgA under a wide range of conditions implies a robust design of aggregation. *Biochemistry*. 2011; 50:8281–8290. [PubMed: 21877724]
10. Hung C, et al. *Escherichia coli* biofilms have an organized and complex extracellular matrix structure. *MBio*. 2013; 4:e00645–13. [PubMed: 24023384]
11. Hammar M, Bian Z, Normark S. Nucleator-dependent intercellular assembly of adhesive curli organelles in *Escherichia coli*. *Proc Natl Acad Sci USA*. 1996; 93:6562–6566. [PubMed: 8692856]
12. Bian Z, Normark S. Nucleator function of CsgB for the assembly of adhesive surface organelles in *Escherichia coli*. *EMBO J*. 1997; 16:5827–5836. [PubMed: 9312041]
13. Loferer H, Hammar M, Normark S. Availability of the fibre subunit CsgA and the nucleator protein CsgB during assembly of fibronectin-binding curli is limited by the intracellular concentration of the novel lipoprotein CsgG. *Mol Microbiol*. 1997; 26:11–23. [PubMed: 9383186]
14. Robinson LS, Ashman EM, Hultgren SJ, Chapman MR. Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein. *Mol Microbiol*. 2006; 59:870–881. [PubMed: 16420357]
15. Nenninger AA, Robinson LS, Hultgren SJ. Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF. *Proc Natl Acad Sci USA*. 2009; 106:900–905. [PubMed: 19131513]
16. Nenninger AA, et al. CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation. *Mol Microbiol*. 2011; 81:486–499. [PubMed: 21645131]
17. Okuda S, Tokuda H. Lipoprotein sorting in bacteria. *Annu Rev Microbiol*. 2011; 65:239–259. [PubMed: 21663440]
18. Iacovache I, Bischofberger M, van der Goot FG. Structure and assembly of pore-forming proteins. *Curr Opin Struct Biol*. 2010; 20:241–246. [PubMed: 20172710]
19. Goyal P, Van Gerven N, Jonckheere W, Remaut H. Crystallization and preliminary X-ray crystallographic analysis of the curli transporter CsgG. *Acta Crystallogr F*. 2013; 69:1349–1353.
20. Krantz BA, et al. A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science*. 2005; 309:777–781. [PubMed: 16051798]
21. Janowiak BE, Fischer A, Collier RJ. Effects of introducing a single charged residue into the phenylalanine clamp of multimeric anthrax protective antigen. *J Biol Chem*. 2010; 285:8130–8137. [PubMed: 20061382]
22. Feld GK, Brown MJ, Krantz BA. Ratcheting up protein translocation with anthrax toxin. *Protein Sci*. 2012; 21:606–624. [PubMed: 22374876]

23. Van Gerven N, et al. Secretion and functional display of fusion proteins through the curli biogenesis pathway. *Mol Microbiol.* 2014; 91:1022–1035. [PubMed: 24417346]
24. Brinker A, et al. Dual function of protein confinement in chaperonin-assisted protein folding. *Cell.* 2001; 107:223–233. [PubMed: 11672529]
25. Busby JN, Panjikar S, Landsberg MJ, Hurst MR, Lott JS. The BC component of ABC toxins is an RHS-repeat-containing protein encapsulation device. *Nature.* 2013; 501:547–550. [PubMed: 23913273]
26. Takagi F, Koga N, Takada S. How protein thermodynamics and folding mechanisms are altered by the chaperonin cage: molecular simulations. *Proc Natl Acad Sci USA.* 2003; 100:11367–11372. [PubMed: 12947041]
27. Zhou HX. Protein folding in confined and crowded environments. *Arch Biochem Biophys.* 2008; 469:76–82. [PubMed: 17719556]
28. Chen AY, et al. Synthesis and patterning of tunable multiscale materials with engineered cells. *Nature Mater.* 2014; 13:515–523. [PubMed: 24658114]
29. Sivanathan V, Hochschild A. A bacterial export system for generating extracellular amyloid aggregates. *Nature Protocols.* 2013; 8:1381–1390.
30. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci USA.* 2000; 97:6640–6645. [PubMed: 10829079]
31. Swamy, M.; Siegers, GM.; Minguet, S.; Wollscheid, B.; Schamel, WWA. Blue native polyacrylamide gel electrophoresis (BN-PAGE) for the identification and analysis of multiprotein complexes; *Sci STKE.* 2006. p. pl4<http://dx.doi.org/10.1126/stke.3452006pl4>
32. Winter G. xia2: an expert system for macromolecular crystallography data reduction. *J Appl Crystallogr.* 2010; 43:186–190.
33. Kabsch W. Xds. *Acta Crystallogr D.* 2010; 66:125–132. [PubMed: 20124692]
34. Sheldrick GM. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D.* 2010; 66:479–485. [PubMed: 20383001]
35. Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D.* 2003; 59:2023–2030. [PubMed: 14573958]
36. Cowtan K. Recent developments in classical density modification. *Acta Crystallogr D.* 2010; 66:470–478. [PubMed: 20383000]
37. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D.* 2006; 62:1002–1011. [PubMed: 16929101]
38. Adams PD, et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D.* 2010; 66:213–221. [PubMed: 20124702]
39. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallogr D.* 2010; 66:486–501. [PubMed: 20383002]
40. Davis IW, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res.* 2007; 35(Suppl 2):W375–W383. [PubMed: 17452350]
41. Strong M, et al. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci USA.* 2006; 103:8060–8065. [PubMed: 16690741]
42. McCoy AJ, et al. Phaser crystallographic software. *J Appl Crystallogr.* 2007; 40:658–674. [PubMed: 19461840]
43. Smart OS, et al. Exploiting structure similarity in refinement: automated NCS and target-structure restraints in BUSTER. *Acta Crystallogr D.* 2012; 68:368–380. [PubMed: 22505257]
44. Murshudov GN, et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr D.* 2011; 67:355–367. [PubMed: 21460454]
45. Nicholls RA, Long F, Murshudov GN. Low-resolution refinement tools in REFMAC5. *Acta Crystallogr D.* 2012; 68:404–417. [PubMed: 22505260]
46. Goormaghtigh E, Ruyschaert JM. Subtraction of atmospheric water contribution in Fourier transform infrared spectroscopy of biological membranes and proteins. *Spectrochim Acta.* 1994; 50A:2137–2144.

47. Mindell JA, Grigorieff N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol.* 2003; 142:334–347. [PubMed: 12781660]
48. Shaikh TR, et al. SPIDER image processing for single-particle reconstruction of biological macromolecules from electron micrographs. *Nature Protocols.* 2008; 3:1941–1974.
49. Tang G, et al. EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol.* 2007; 157:38–46. [PubMed: 16859925]
50. Pettersen EF, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; 25:1605–1612. [PubMed: 15264254]
51. Del Rio Martinez, JM.; Zaitseva, E.; Petersen, S.; Baaken, G.; Behrends, JC. [13 August 2014] Automated formation of lipid membrane microarrays for ionic single molecule sensing with protein nanopores. *Small.* <http://dx.doi.org/10.1002/sml.201402016>
52. Movileanu L, Howorka S, Braha O, Bayley H. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nature Biotechnol.* 2000; 18:1091–1095. [PubMed: 11017049]
53. Im W, Roux B. Ions and counterions in a biological channel: a molecular dynamics simulation of OmpF porin from *Escherichia coli* in an explicit membrane with 1 M KCl aqueous salt solution. *J Mol Biol.* 2002; 319:1177–1179. [PubMed: 12079356]
54. Pronk S, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics.* 2013; 29:845–854. [PubMed: 23407358]
55. Lindorff-Larsen K, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins.* 2010; 78:1950–1958. [PubMed: 20408171]
56. Capra JA, Singh M. Predicting functionally important residues from sequence conservation. *Bioinformatics.* 2007; 23:1875–1882. [PubMed: 17519246]

Figure 1. X-ray structure of CsgG_{C15} in pre-pore conformation

a, Ribbon diagram of the CsgG_{C15} monomer coloured as a blue-to-red rainbow from N terminus to C terminus. Secondary structure elements are labelled according to the ABD-like fold, with the additional N-terminal and C-terminal α -helices and the extended loop connecting β 1 and α 1 labelled α N, α C and C-loop (CL), respectively. **b**, Side view of the CsgG_{C15} C₈ octamer with subunits differentiated by colour and one subunit oriented and coloured as in **a**.

Figure 2. Structure of CsgG in its channel conformation

a, Amide I region (1,700–1,600 cm⁻¹) of ATR-FTIR spectra of CsgG_{C1S} (blue) and membrane-extracted CsgG (red). **b**, TM1 and TM2 sequence (bilayer-facing residues in blue) and Congo red binding of *E. coli* BW25141 *csgG* complemented with wild-type *csgG* (WT), empty vector or *csgG* lacking the underlined fragments of TM1 or TM2. Data are representative of three biological replicates. **c**, Overlay of CsgG monomer in pre-pore (light blue; TM1 pink, TM2 purple) and channel conformation (tan; TM1 green, TM2 orange). CL, C-loop. **d**, **e**, Side view (**d**) and cross-sectional view (**e**) of CsgG nonamers in ribbon and surface representation; helix 2, the core domain and TM hairpins are shown in blue, light blue and tan, respectively. A single protomer is coloured as in Fig. 1a. Magenta spheres show the position of Leu 2. OM, outer membrane.

Figure 3. CsgG channel constriction

a, Cross-section of CsgG channel constriction and its solvent-excluded diameters. **b**, The constriction is composed of three stacked concentric side-chain layers: Tyr 51, Asn 55 and Phe 56, preceded by Phe 48 from the periplasmic side. **c**, CsgG channel topology. **d**, Congo red binding of *E. coli* BW25141 *csgG* complemented with *csgG* (WT), empty vector or *csgG* carrying indicated constrictions mutants. Data are representative of six biological replicates. **e**, **f**, Representative single-channel current recordings (**e**) and conductance histogram (**f**) of CsgG reconstituted in planar phospholipid bilayers and measured under an electrical field of +50 mV ($n = 33$) or -50 mV ($n = 13$).

Figure 4. Model of CsgG transport mechanism

a, Native PAGE of CsgE (E), CsgG (G) and CsgG supplemented with excess CsgE (E + G), showing the formation of a CsgG–CsgE complex (E–G*). Data are representative of seven experiments, encompassing four protein batches. **b**, SDS–PAGE of CsgE (E), CsgG (G) and the E–G* complex recovered from native PAGE. Data are representative of two repetitions. M, molecular mass markers. **c**, Selected class averages of CsgG–CsgE particles. From left to right: averaged top and side views visualized by cryo-EM, and comparison of negatively stained side views of CsgG–CsgE and CsgG nonamers. **d**, Cryo-EM averages of top and tilted side-viewed CsgE particles. Rotational autocorrelation shows nine-fold symmetry. **e**, Three-dimensional reconstruction of CsgG–CsgE (24 Å resolution, 1,221 single particles) shows a nonameric particle comprising CsgG (blue) and an additional density assigned as a CsgE nonamer (orange). **f**, Single-channel current recordings of PPB-reconstituted CsgG at +50 mV or –50 mV and supplemented with incremental concentrations of CsgE. Horizontal scale bars lie at 0 pA. **g**, Tentative model for CsgG-mediated protein secretion. CsgG and CsgE are proposed to form a secretion complex that entraps CsgA (discussed in Extended Data Fig. 9), generating an entropic potential over the channel. After capture of CsgA in the channel constriction, a ΔS -rectified Brownian diffusion facilitates the progressive translocation of the polypeptide across the outer membrane.