ARTICLE

doi:10.1038/nature11744

Crystallographic snapshot of cellulose synthesis and membrane translocation

Jacob L. W. Morgan[1], Joanna Strumillo[1,2] & Jochen Zimmer[1]

Cellulose, the most abundant biological macromolecule, is an extracellular, linear polymer of glucose molecules. It represents an essential component of plant cell walls but is also found in algae and bacteria. In bacteria, cellulose production frequently correlates with the formation of biofilms, a sessile, multicellular growth form. Cellulose synthesis and transport across the inner bacterial membrane is mediated by a complex of the membrane-integrated catalytic BcsA subunit and the membrane-anchored, periplasmic BcsB protein. Here we present the crystal structure of a complex of BcsA and BcsB from *Rhodobacter sphaeroides* containing a translocating polysaccharide. The structure of the BcsA–BcsB translocation intermediate reveals the architecture of the cellulose synthase, demonstrates how BcsA forms a cellulose-conducting channel, and suggests a model for the coupling of cellulose synthesis and translocation in which the nascent polysaccharide is extended by one glucose molecule at a time.

Polysaccharides perform multiple vital functions in life, as reflected in their great variety in chemical composition and physicochemical properties. Extracellular polysaccharides, such as cellulose, chitin and hyaluronan, primarily perform structural functions and are synthesized inside the cell from nucleotide-activated donor sugars[1–3].

Cellulose is the most abundant biological polymer and consists of glucose molecules that are connected between their C1 and C4 carbons via acetal linkages[4]. It is predominantly generated by vascular plants and a large number of algae, but also by some bacteria[5,6], protists[7] and tunicates[8]. Cellulose synthases (CESAs) are membrane-embedded glycosyltransferases, which use UDP-activated glucose (UDP-Glc) to elongate the nascent polysaccharide processively in a reaction that inverts the configuration at the anomeric carbon of the newly added sugar from α to β[6,9,10]. Prokaryotic and eukaryotic CESAs share a similar predicted topology including eight transmembrane helices and at least one extended intracellular glycosyltransferase (GT) loop (Supplementary Fig. 1).

Bacteria, primarily Gram-negatives, produce and secrete cellulose via a protein complex consisting of at least three subunits (BcsA, BcsB and BcsC)[11]. The inner membrane protein BcsA is the catalytically active subunit and contains a conserved family-2 GT domain between transmembrane helices 4 and 5 (TM4 and TM5)[12]. BcsB is a periplasmic protein that is anchored to the inner membrane via a single, carboxy-terminal transmembrane helix. BcsA and BcsB are fused as a single polypeptide in some species, supporting the genetic observation that BcsB is essential for cellulose synthesis[13,14]. BcsC, required for cellulose synthesis *in vivo* but not *in vitro*[14], is predicted to form an 18-stranded β-barrel in the outer membrane. Some bacteria contain additional, non-essential, periplasmic subunits of unknown function[11,15].

Although some bacteria constitutively produce cellulose[6], bacterial cellulose production is often concomitant with the formation of sessile, multicellular aggregates, termed biofilms[11,16]. Biofilm bacteria are encased in a network of polysaccharides and proteinaceous fibres and are of particular concern to human health due to an increased tolerance to antimicrobial treatments[17]. Biofilm formation is stimulated by the bacterial secondary messenger cyclic-di-GMP, which bacteria sense via PilZ domains[18]. Accordingly, BcsA contains a PilZ domain within its C-terminal, intracellular domain and its activity is strongly stimulated by cyclic-di-GMP[19,20].

In all organisms, the cellulose polymer has to be exported across the plasma membrane. However, the mechanism by which a polymer consisting of thousands of glucose units can be translocated has been unclear. To understand this process, we determined the crystal structure of the catalytically active BcsA–BcsB complex in an intermediate state during cellulose synthesis and translocation at 3.25 Å resolution (Supplementary Table 1).

Architecture of the BcsA–BcsB complex

BcsA and BcsB form a 1:1 stoichiometric complex spanning approximately 150 Å perpendicular and 55 Å parallel to the membrane. The complex is divided into a cuboid-shaped membrane-spanning region sandwiched between large cytoplasmic and periplasmic domains (Fig. 1a). The BcsA–BcsB interface is approximately 4,500 Å^2 including a cluster of negative and positive charges on BcsA and BcsB, respectively (Supplementary Fig. 2).

BcsA contains four amino-terminal and four carboxy-terminal transmembrane helices separated by a large intracellular loop (4/5 loop) that forms a GT domain (amino acids 128–368) (Fig. 1a and Supplementary Fig. 3). TM3–8 form a narrow channel for the translocating polysaccharide, and BcsA's intracellular C terminus (amino acids 575–759) contains a six-stranded β-barrel and a highly curved α-helical region that attaches the β-barrel to the GT domain (Fig. 1a).

BcsB is a dome-shaped, β-strand-rich, periplasmic protein. Its N-terminal region forms the tip of the dome, whereas the C-terminal transmembrane anchor interacts with BcsA (Fig. 1a). Two amphipathic helices further stabilize its interaction with BcsA and the periplasmic water–membrane interface (Fig. 1a).

Throughout model building, we noticed a continuous, strong positive difference Fourier electron density that could not be accounted for by BcsA or BcsB (Fig. 1b). The ribbon-shaped electron density extended from BcsA's GT domain in a 45° angle towards the membrane, kinked near the putative water–lipid interface to run through the centre of the transmembrane region, and bent sharply at the BcsA–BcsB interface to emerge from the complex at approximately

[1]Center for Membrane Biology, Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908, USA. [2]Faculty of Biology and Environmental Protection, University of Lodz, Pilarskiego Street 14/16, 90-231 Lodz, Poland.

©2013 Macmillan Publishers Limited. All rights reserved

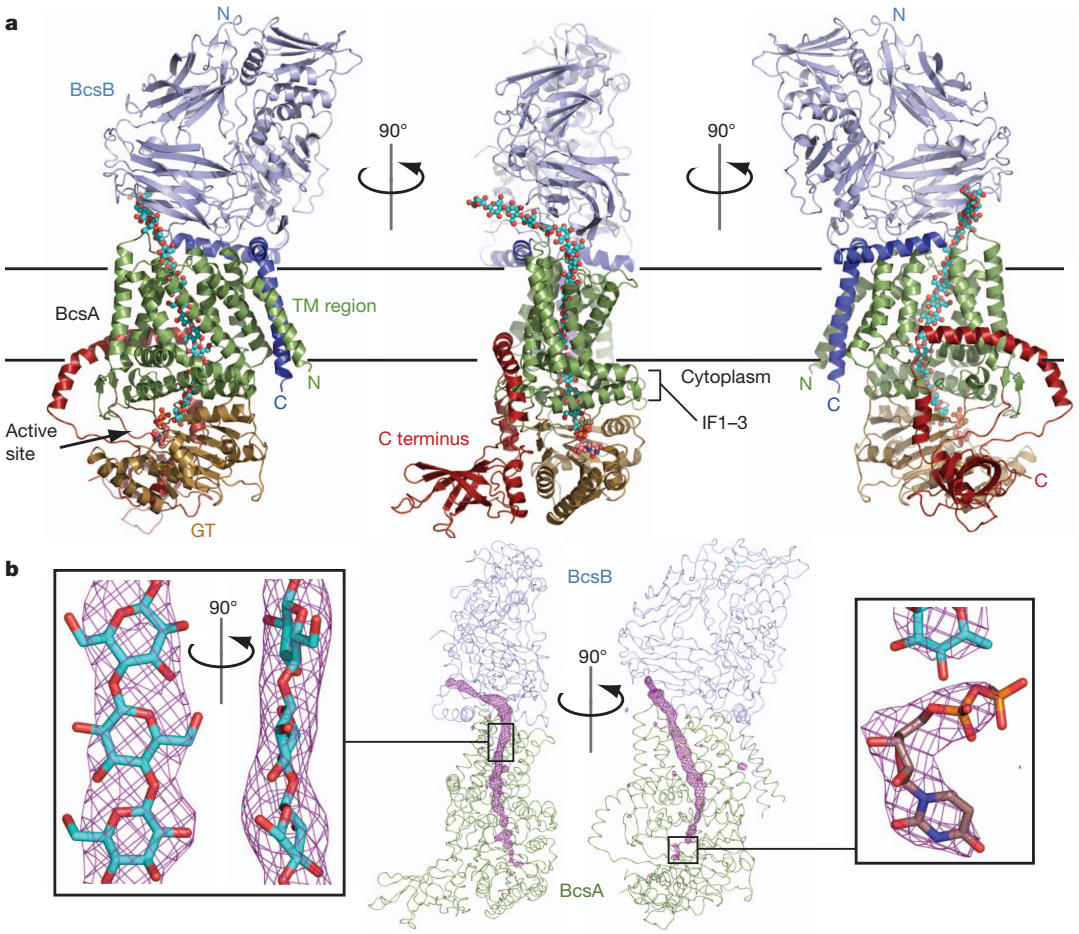

Figure 1 | Architecture of the BcsA–BcsB complex. **a**, BcsA and BcsB form an elongated complex with large cytosolic and periplasmic domains. The transmembrane helices of BcsA are coloured green, the glycosyltransferase (GT) domain beige, and the C-terminal domain red. BcsB is shown in light and dark blue for its periplasmic and membrane-associated regions, respectively. The N and C termini of both subunits are indicated and the translocating glucan and UDP are shown as cyan and violet spheres. Horizontal bars indicate the membrane boundaries. IF, amphipathic interface helices of BcsA. **b**, Unbiased $F_o - F_c$ difference Fourier electron density (pink mesh, contoured at 4.5σ) calculated before modelling the glucan and UDP molecules. The continuous density runs from the intracellular catalytic site to the periplasmic BcsA–BcsB interface and accommodates 18 glucose molecules (cyan sticks). The UDP density was contoured at 3σ (right panel).

145° (Fig. 1b). Past the BcsA–BcsB interface, the electron density extended into the bulk solvent. The characteristic shape of the electron density, its proximity to conserved residues throughout its path (see below), its origin at the glycosyltransferase catalytic site, and extension past the BcsA–BcsB complex strongly suggest that it represents a translocating glucan probably synthesized during or before crystallization.

The catalytic activity of the purified, detergent-solubilized BcsA–BcsB complex was confirmed by incubating it with UDP-Glc and cyclic-di-GMP. The synthesized polymer, that could be completely degraded by β-1,4-glucanase (data not shown), was sedimented, purified and analysed by glycosyl linkage analysis (performed by the Complex Carbohydrate Research Center, University of Georgia), confirming that 94% of it consists of 4-linked glucose (Supplementary Fig. 4).

Although probably reflecting an average of different registers of the translocating glucan, the density discerns the overall position of 18 glucose molecules (Fig. 1b and Supplementary Fig. 5). According to the prevailing model that cellulose synthase elongates the non-reducing end of the growing polymer[21,22], the glucan was modelled with its non-reducing end (glucose-18) at the glycosyltransferase catalytic site. At the active site, we observed additional, weak electron density detached from the glucan. On the basis of its shape and coordination by residues that are predicted to be implicated in nucleotide binding and glycosyl transfer[23], it seemingly represents a weakly bound UDP molecule (Fig. 1b and Supplementary Fig. 5). Therefore, our structure probably represents a cellulose synthase–cellulose translocation intermediate captured after glycosyl transfer and before replacing UDP with UDP-Glc.

The glycosyltransferase domain

The GT domain of BcsA contains the conserved signature D,D,D,Q (Q/R)XRW of three variably spaced aspartic acids and a pentapeptide consisting of a Gln that is often followed by a Gln or Arg, a variable residue, and an Arg and Trp. This signature is shared with other processive β-glycosyltransferases, such as chitin and hyaluronan synthases[2,24].

The GT domain adopts a glycosyltransferase A fold (GT-A)[10] consisting of a mixed, seven-stranded β-sheet that is surrounded by seven α-helices and attaches to the transmembrane region via three amphipathic interface (IF) helices (IF1–3) (Figs 1–3 and Supplementary Fig. 3). IF1 is formed by residues belonging to the 4/5 loop (amino acids 298–312) and packs against the C termini of BcsA's TM8 and TM4 as well as the N terminus of TM5. The second helix, IF2 (amino acids 373–395), directly precedes TM5 and includes the invariant Q(Q/R)XRW sequence of the signature. IF3 (amino acids 470–498) follows TM6 and forms a crescent-shaped helix around the cytosolic entrance to the transmembrane channel (discussed below). It is connected to TM7 via a conserved loop (IF3–TM7 loop) that runs across the entrance to the GT substrate-binding site (Fig. 2b and Supplementary Fig. 3). The GT domain protrudes from these IF helices

©2013 Macmillan Publishers Limited. All rights reserved

at approximately 45° (Fig. 1a). The inverting glycosyltransferase SpsA from *Bacillus subtilis*[23] superimposes with BcsA's GT domain with 2.15 Å root mean square deviation (r.m.s.d.) between Cα atoms (Supplementary Fig. 6).

The structure of the cellulose synthase reveals the function of the signature D,D,D,Q(Q/R)XRW as well as other highly conserved sequence motifs. Consistent with non-processive glycosyltransferases[23,25,26], the first two aspartic acids of the signature, Asp 179 and Asp 246, coordinate UDP (Fig. 2a). The third 'D' (Asp 343), which is part of the invariant 'TED' motif, probably represents the catalytic base owing to its proximity to the non-reducing end of the glucan (Fig. 2a).

The Q(Q/R)XRW part of the signature, together with an equally conserved FFCGS sequence, forms a binding site for the terminal disaccharide of the glucan, the acceptor (Fig. 2a). The Q(Q/R)XRW sequence belonging to IF2 is a part of the cytoplasmic entry to the glucan channel (Fig. 3). The first Q, Gln 379, forms a hydrogen bond to the guanidino group of the neighbouring Arg 380, which in turn forms a salt bridge to the invariant Glu 297. Arg 382, preceding the conserved Trp residue, most likely coordinates the diphosphate of UDP; however, the electron density for its side chain is weak, consistent with the apparent low occupancy of UDP. Trp 383 forms van der Waals interactions with the penultimate glucose molecule at the acceptor site (glucose-17), and packs against the Gln 379-Arg 380 pair towards the distal site of the catalytic pocket (Fig. 2a). The FFCGS pentapeptide (amino acids 316–320) interacts with the opposite side of the terminal disaccharide, primarily via the backbone carbonyls of Cys 318 and Gly 319 with glucose-18.

The dominant-negative *thanatos* mutation (Pro578Ser) in *Arabidopsis thaliana* CESA3[27] maps to the invariant QTPH sequence in bacterial cellulose synthases (Supplementary Fig. 1). His 276 of this motif is within hydrogen-bonding distance to glucose-17 (Fig. 2a) and is thus probably also involved in positioning the non-reducing end with respect to the catalytic base. The motif is part of an extended β-strand and Pro 275 (equivalent to Pro 578 in *A. thaliana* CESA3) induces a 'kink' in the β-strand that positions the side chain of His 276 towards the glucan, which might be disrupted when Pro is replaced by another amino acid.

A cavity next to the UDP-binding pocket might accommodate the glucose donor in the UDP-Glc-bound state. This pocket is flanked on one side by Ala 225 and Lys 226 of the invariant AKAGN motif and, on the other side, by Glu 342 (of the conserved TED motif) and His 224 (Fig. 2a and Supplementary Fig. 7).

Activation by cyclic-di-GMP

The activity of BcsA is stimulated by cyclic-di-GMP[20]. The cyclic-di-GMP-responsive PilZ domain is localized within BcsA's C terminus[18], right next to the GT domain (Fig. 2b). The C terminus of BcsA extends from the last transmembrane helix (TM8) via a short β-strand, which forms a two-stranded β-sheet with the IF3–TM7 loop, before it folds into a six-stranded β-barrel (amino acids 585–675). The β-barrel points away from the GT domain at approximately 90°. Past the β-barrel, the polypeptide chain continues in an α-helical conformation on the surface of the GT domain towards the water–lipid interface. At the interface, it forms an amphipathic helix that interacts with TM6, IF3 and TM8 (Fig. 1a) before it curves away from the membrane towards the GT domain and breaks at Glu 743. After Glu 743, the polypeptide continues in an extended conformation towards the β-barrel, past the substrate-binding pocket.

The β-barrel of BcsA aligns with an r.m.s.d. of 2.2 Å between Cα atoms with the cyclic-di-GMP binding protein VCA0042[28] (Supplementary Fig. 8). The conserved residues Arg 584, Asp 609, Ser 611 and Gly 614 on the β-barrel surface as well as Gln 578 and Arg 579 of the TM8–β-barrel linker are probably implicated in cyclic-di-GMP binding[28] (Fig. 2b and Supplementary Fig. 8). We hypothesize that cyclic-di-GMP binding to BcsA induces conformational changes that allow UDP-Glc to access the catalytic site. A likely candidate for displacement is the IF3–TM7 loop that runs across the entrance to the catalytic pocket and interacts with the TM8–β-barrel linker (Fig. 2b and Supplementary Fig. 3). Notably, replacing Thr

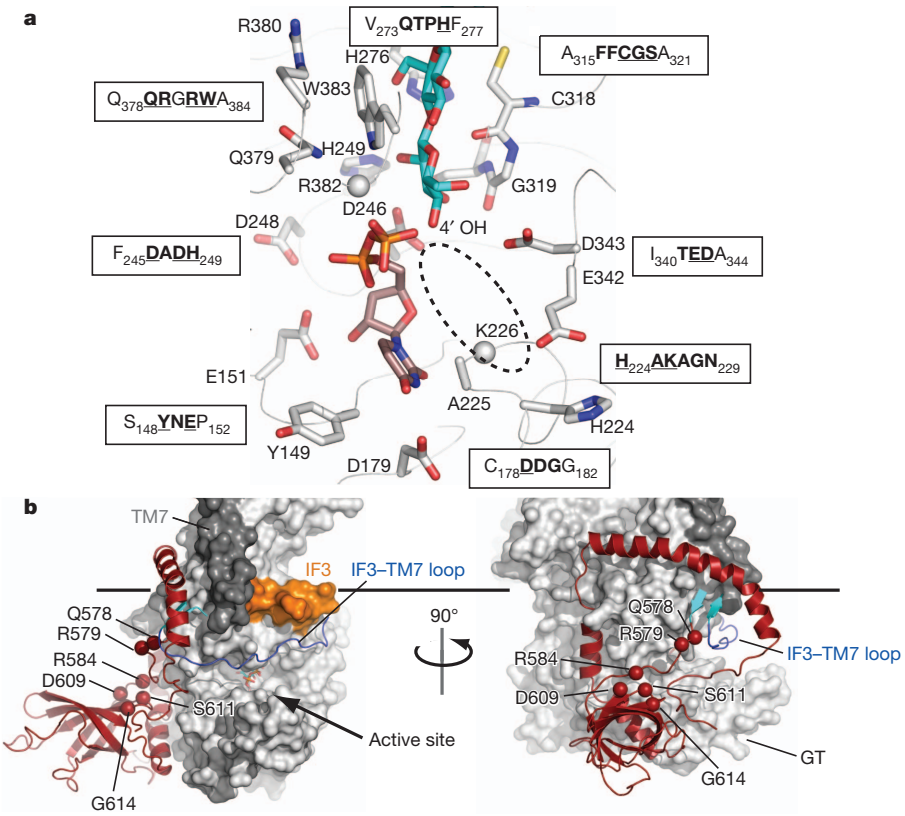

Figure 2 | Organization of BcsA's catalytic site and PilZ domain. a, Conserved residues of BcsA coordinate UDP and the terminal disaccharide of the glucan. Side chains represented in sticks belong to the sequence motifs shown in single letter code. Conserved residues are highlighted in bold and the depicted residues are underlined. All side chains shown could be positioned unambiguously in the electron density map. No density was observed for the side chains of Lys 226 and Arg 382 (shown as spheres for their Cα atoms). The probable position of the donor Glc is indicated by a dashed ellipsoid. **b**, Surface representation of BcsA with its C-terminal domain shown as cartoon in red. The linker connecting TM8 with the β-barrel forms a two-stranded β-sheet (shown in cyan) with the IF3–TM7 loop. TM7 is coloured dark grey and IF3 is orange. The Cα atoms of residues likely involved in cyclic-di-GMP binding are shown as spheres and are labelled. The horizontal bar indicates the cytoplasmic membrane boundary.

©2013 Macmillan Publishers Limited. All rights reserved

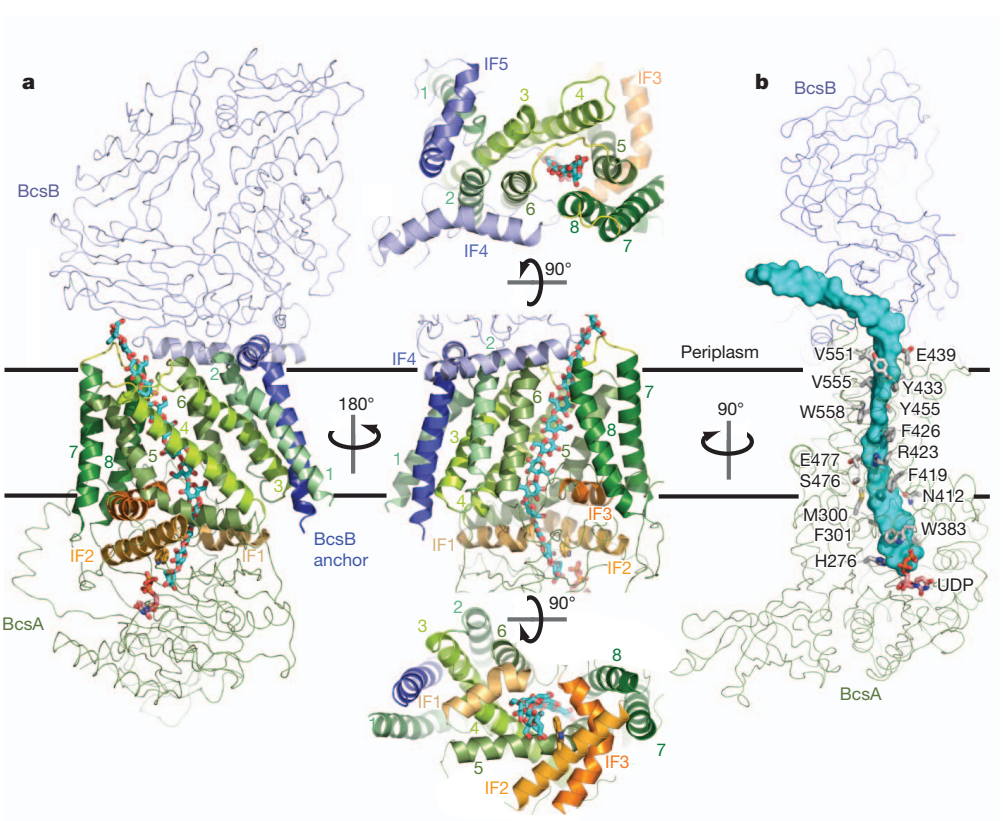

Figure 3 | The membrane-integrated domain of BcsA–BcsB. **a**, The transmembrane region includes nine transmembrane helices and three cytoplasmic and two periplasmic interface helices (IF1–5). BcsA and BcsB are shown as green and blue ribbons, respectively. Transmembrane and interface helices are show in cartoon representation and coloured in shades of green and orange for BcsA's TM1–8 and IF1–3, respectively, and shades of blue for BcsB's IF4 and IF5 and its C-terminal transmembrane anchor. TM3–8 of BcsA form a narrow channel. Its cytoplasmic entrance is formed by IF1–3 as well as the N-terminal half of TM5 and its periplasmic exit is between the periplasmic 5/6 and 7/8 loops of BcsA. Trp 383 of the Q(Q/R)XRW motif is shown in stick representation. **b**, Conserved residues of BcsA interact with the translocating glucan. BcsA and BcsB are shown as ribbons, and side chains of BcsA contacting the glucan (cyan surface) are shown as sticks. Residues of BcsB contacting the polymer are not conserved and are not shown.

942 with Ile in *A. thaliana* CESA3 confers resistance to the herbicide isoxaben[29]. The corresponding, equally conserved residue in our structure is Thr 506 in the IF3–TM7 loop (Supplementary Fig. 1), supporting the hypothesis that this loop is involved in regulating the activity of BcsA.

The transmembrane region and polysaccharide channel

The transmembrane region of the cellulose synthase includes eight transmembrane helices from BcsA (TM1–8) and one from BcsB (Fig. 3a). The helices are framed by three cytoplasmic and two periplasmic IF-helices formed by BcsA and BcsB, respectively. In addition, the N-terminal four turns of TM5 also run parallel to the membrane before the helix bends towards the hydrocarbon core at Pro 420 (Fig. 3a). All transmembrane helices, except TM6 and TM7 and BcsB's transmembrane helix, tilt with respect to the membrane normal and surround a narrow channel approximately 8 Å wide and 33 Å long. Because the transmembrane helices are organized in pairs in one direction (helices 1/2, 3/6, 4/8 and 5/7 of BcsA), their overall arrangement resembles an elongated cuboid, and BcsB's transmembrane anchor packs against one side of the cuboid (Fig. 3a). On the membrane periplasmic side, BcsB forms two amphipathic helices (IF4 and IF5) at the water–lipid interface. IF4 (amino acids 569 to 592) interacts with the periplasmic termini of BcsA's TM2, TM6 and TM8, and IF5 (amino acids 679–693) directly precedes BcsB's membrane anchor (amino acids 695–720) and contacts BcsA's TM1 and the 1/2 loop (Fig. 3a).

The channel formed by BcsA's TM3–8 accommodates ten glucose units of the translocating glucan (Fig. 3). The glucan enters the channel through the cytoplasmic opening formed by IF1–3 and the N-terminal half of TM5, crosses the membrane parallel to TM5 and TM6, and exits on the periplasmic side between BcsA's 5/6 and 7/8 loops at the BcsA–BcsB interface. Here, the glucan kinks to protrude from the complex sideways (Fig. 3b). TM6 and TM8 are separated by about 9 Å and 15 Å (between Cα atoms) at the periplasmic and cytoplasmic sides of the membrane, respectively, and only the tips of their side chains maintain the barrier towards the lipids (Fig. 3a). In the absence of a translocating glucan, transmembrane helices 7 and 8 might move towards the N-terminal transmembrane helices to close the channel.

All transmembrane helices of BcsA, with the exception of TM1 and TM2, directly contact the glucan. The faces of the glucopyranose rings form van der Waals interactions with residues Met 300 and Phe 301 of IF1, Phe 316 directly proceeding IF1, Trp 383 of IF2, Phe 419, Phe 426 and Tyr 433 of TM5, Phe 441 in the 5/6 loop, and Val 551, Val 555 and Trp 558 of TM8 (Fig. 3b). The equatorial hydroxyl groups of the glucose units form hydrophilic contacts with Tyr 80 of TM3, Asn 118 of TM4, His 276 of the GT domain, Asn 412 and Arg 423 of TM5, Glu 439 in the TM5/6 loop, Tyr 455 of TM6, and Ser 476 and Glu 477 of IF3. We note that the quinoxyphen-resistance mutation Ala903Val in *A. thaliana* CESA1[30] aligns with Tyr 455 in TM6, which forms a hydrogen bond to the translocating glucan (Fig. 3b and Supplementary Fig. 1). Therefore, although the sequence alignment in this region is weak, it is likely that quinoxyphen affects the translocation rather than the synthesis of the glucan.

The periplasmic domain

BcsB protrudes approximately 60 Å into the periplasm with its N-terminal half (amino acids 54–307) forming the membrane distal part. BcsB consists primarily of β-strands that are organized into distinct domains, including two jellyroll and two flavodoxin-like folds (Fig. 4). The jellyrolls (amino acids 54–185 and 316–452) show a striking structural similarity with carbohydrate binding modules (Supplementary Fig. 9)[31] and are termed carbohydrate-binding domains (CBD) 1 and 2. Both CBDs contain eight antiparallel β-strands separated by loops 2–16 residues in length and adopt a similar fold, such that their Cα atoms superimpose with an r.m.s.d. of 2.5 Å (Supplementary Fig. 9). The CBDs are positioned at a 45° angle relative to one another and interact via two extended loops, including a disulphide bond between the conserved Cys 163 and Cys 430 (Fig. 4a).

CBD2 runs parallel to the plane of the membrane above the periplasmic exit of the transmembrane channel. It interacts with BcsA's

©2013 Macmillan Publishers Limited. All rights reserved

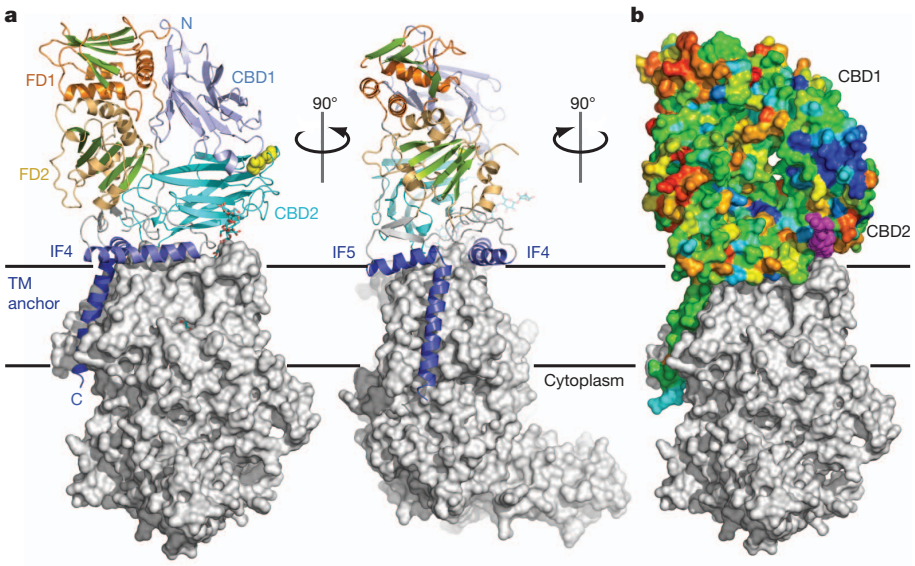

Figure 4 | Organization of the periplasmic domain. **a**, The periplasmic region of the cellulose synthase is primarily formed by BcsB. It consists of two periplasmic carbohydrate binding domains (CBD1 and CBD2) that are connected to two α/β domains (FD1 and FD2). The CBDs are covalently attached via a disulphide bond between the conserved Cys 163 and Cys 430 (yellow spheres). BcsB's transmembrane anchor and interface helices 4 and 5 are coloured in dark blue. The N and C termini of BcsB are indicated. **b**, Surface representation of BcsB coloured according to sequence conservation from red (variable) to deep blue (invariant). The tips of CBD1 and CBD2 form a patch of conserved, primarily hydrophobic residues above the periplasmic exit of the transmembrane channel. BcsA is shown as a grey surface and the glucan as purple spheres.

TM2 and TM3, as well as with its 5/6 loop, thereby accounting for almost 30% of the complex interface. Although BcsB exhibits significant sequence variability, we noticed a cluster of conserved residues including His 159, Arg 160, Ile 161, Leu 171 and Trp 172 at the tip of CBD1, close to its disulphide bond with CBD2 and above the periplasmic exit of the glucan channel (Fig. 4b). A comparison of CBD1 with the bacterial carbohydrate binding module family 35 (CBM35) in complex with a glucuronic acid disaccharide[31] localizes disaccharide binding to the cluster of conserved residues in CBD1 (Supplementary Fig. 9), indicating that this region interacts with the translocating glucan.

Both CBDs are linked to compact α/β domains containing a central β-sheet five or six strands wide that is framed on either side by two α-helices (Fig. 4a). Although probably lacking functional similarity, both domains resemble a flavodoxin fold[32] (Supplementary Fig. 10) and are thus referred to as FD1 (amino acids 192–303) and FD2 (amino acids 457–529 together with 598–666). The repeating structural motif of BcsB therefore contains a CBD linked to an FD domain.

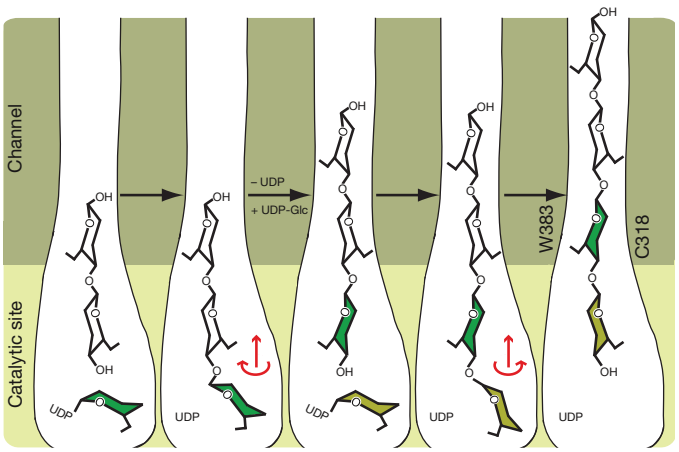

Figure 5 | Proposed model for cellulose synthesis and translocation. After glycosyl transfer, the newly added Glc could rotate around the acetal linkage into the plane of the polymer. The rotation direction would be determined by steric interactions and formation of the β-1,4 glucan characteristic intramolecular O3-H••O5 hydrogen bond. The glucan might translocate into the channel during this relaxation. This process would be repeated with a second UDP-Glc but the rotation direction after glycosyl transfer would be in the opposite direction owing to steric constraints. Alternatively, the glucan might not translocate into the channel until UDP is replaced by UDP-Glc. Trp 383 and Cys 318 mark the entrance to the transmembrane channel (only shown in the right panel).

Implications for cellulose synthesis and translocation

The structure of the cellulose synthase–cellulose translocation intermediate provides unique insights into the mechanism of cellulose synthesis and translocation. On the basis of the large distance between the glycosyltransferase site and the putative water–membrane interface (approximately 25 Å; Fig. 1) as well as BcsA's activity in detergent, a reaction mechanism involving lipid-linked intermediates seems unlikely[5]. The spatial restrictions in the substrate-binding pocket further suggest that the nascent glucan is extended by one[33] and not by two[34] glucose molecules at a time.

The distinct features of the glucan electron density (Fig. 1b) further indicate that the polymer translocates one glucose molecule at a time. The dimensions of the transmembrane channel require that the glucopyranose rings have to be within the same plane to slide into the pore (Fig. 3b). Thus, we speculate that after glycosyl transfer the newly attached terminal glucose molecule will rotate around the acetal linkage[35,36] to align with the glucan in the channel (Fig. 5). Steric interactions might dictate the rotation direction, leading to the β-1,4 glucan characteristic 180° rotation between individual glucose units and intramolecular hydrogen bonding between ring oxygens and 3′ hydroxyl groups of neighbouring units[4,37]. This relaxation might be sufficient to allow the polymer to slide into the channel (Fig. 5). Alternatively, the elongated glucan might not translocate until UDP is replaced with UDP-Glc, necessitating translocation owing to spatial restrictions at the active site. Past the channel, the induced kink of the glucan chain at the BcsA–BcsB interface, the interaction with BcsB's CBDs, or the aggregation with other glucans might further contribute to a unidirectional movement of the polymer.

Cellulose, chitin and hyaluronan are probably synthesized by a conserved mechanism involving a membrane-embedded glycosyltransferase that couples polymer synthesis with its translocation across the cell membrane[3]. This unique mechanism stands in contrast to the translocation of most other biological polymers where 'polymer synthases' function independently of or alongside dedicated translocation machineries[38–41]. Our structure now provides a basis for unravelling the details of this process.

METHODS SUMMARY

The BcsA and BcsB subunits of the *Rhodobacter sphaeroides* cellulose synthase complex were co-expressed and purified by metal affinity and gel filtration chromatography as described in detail in the Methods. The BcsA–BcsB complex was crystallized and the structure was determined after single anomalous dispersion phasing of a data set obtained from samarium III-chloride-soaked crystal. Additional experimental

©2013 Macmillan Publishers Limited. All rights reserved

phases were obtained from selenomethionine-derivatized BcsA–BcsB crystals. The quality of the electron density differed between peripheral and transmembrane regions of the complex. Therefore, modelling of the N-terminal region of BcsB (amino acids 54–177) and amino acids 586–675 of BcsA relied on the position of bulky side chains, Se-Met positions and backbone electron density. The model was built in Coot[42] and includes all residues with the exception of the terminal residues 1–12 and 760–788 of BcsA and 19–53 and 721–725 of BcsB. In addition, a short loop of BcsB consisting of residues 532–543 was not visible in the electron density map.

Full Methods and any associated references are available in the online version of the paper.

Received 12 July; accepted 31 October 2012.
Published online 9 December 2012.

1. Somerville, C. Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* **22,** 53–78 (2006).
2. Merzendorfer, H. Insect chitin synthases: a review. *J. Comp. Physiol. B* **176,** 1–15 (2006).
3. Hubbard, C., McNamara, J. T., Azumaya, C., Patel, M. S. & Zimmer, J. The hyaluronan synthase catalyzes the synthesis and membrane translocation of hyaluronan. *J. Mol. Biol.* **418,** 21–31 (2012).
4. Nishiyama, Y., Langan, P. & Chanzy, H. Crystal structure and hydrogen-bonding system in cellulose Iβ from synchrotron X-ray and neutron fiber diffraction. *J. Am. Chem. Soc.* **124,** 9074–9082 (2002).
5. Matthysse, A. G., Thomas, D. L. & White, A. R. Mechanism of cellulose synthesis in Agrobacterium tumefaciens. *J. Bacteriol.* **177,** 1076–1081 (1995).
6. Bureau, T. E. & Brown, R. M. *In vitro* synthesis of cellulose II from a cytoplasmic membrane fraction of Acetobacter xylinum. *Proc. Natl Acad. Sci. USA* **84,** 6985–6989 (1987).
7. Grimson, M. J., Haigler, C. H. & Blanton, R. L. Cellulose microfibrils, cell motility, and plasma membrane protein organization change in parallel during culmination in Dictyostelium discoideum. *J. Cell Sci.* **109,** 3079–3087 (1996).
8. Kimura, S. & Itoh, T. New cellulose synthesizing complexes (terminal complexes) involved in animal cellulose biosynthesis in the tunicate *Metandrocarpa uedai*. *Protoplasma* **194,** 151–163 (1996).
9. Aloni, Y., Delmer, D. P. & Benziman, M. Achievement of high rates of *in vitro* synthesis of 1,4-beta-D-glucan: activation by cooperative interaction of the *Acetobacter xylinum* enzyme system with GTP, polyethylene glycol, and a protein factor. *Proc. Natl Acad. Sci. USA* **79,** 6448–6452 (1982).
10. Lairson, L. L., Henrissat, B., Davies, G. J. & Withers, S. G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77,** 521–555 (2008).
11. Römling, U. Molecular biology of cellulose production in bacteria. *Res. Microbiol.* **153,** 205–212 (2002).
12. Cantarel, B. L. *et al.* The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* **37,** D233–D238 (2009).
13. Standal, R. *et al.* A new gene required for cellulose production and a gene encoding cellulolytic activity in *Acetobacter xylinum* are colocalized with the bcs operon. *J. Bacteriol.* **176,** 665–672 (1994).
14. Saxena, I. M., Kudlicka, K., Okuda, K. & Brown, R. M. Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. *J. Bacteriol.* **176,** 5735–5752 (1994).
15. Hu, S. Q. *et al.* Structure of bacterial cellulose synthase subunit D octamer with four inner passageways. *Proc. Natl Acad. Sci. USA* **107,** 17957–17961 (2010).
16. Jahn, C. E., Selimi, D. A., Barak, J. D. & Charkowski, A. O. The *Dickeya dadantii* biofilm matrix consists of cellulose nanofibres, and is an emergent property dependent upon the type III secretion system and the cellulose synthesis operon. *Microbiology* **157,** 2733–2744 (2011).
17. Stewart, P. S. & Costerton, J. W. Antibiotic resistance of bacteria in biofilms. *Lancet* **358,** 135–138 (2001).
18. Amikam, D. & Galperin, M. Y. PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* **22,** 3–6 (2006).
19. Römling, U., Gomelsky, M. & Galperin, M. Y. C-di-GMP: the dawning of a novel bacterial signalling system. *Mol. Microbiol.* **57,** 629–639 (2005).
20. Ross, P. *et al.* Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* **325,** 279–281 (1987).
21. Koyama, M., Helbert, W., Imai, T., Sugiyama, J. & Henrissat, B. Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. *Proc. Natl Acad. Sci. USA* **94,** 9091–9095 (1997).
22. Lai-Kee-Him, J. *et al. In vitro* versus *in vivo* cellulose microfibrils from plant primary wall synthases: structural differences. *J. Biol. Chem.* **277,** 36931–36939 (2002).
23. Charnock, S. J. & Davies, G. J. Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms. *Biochemistry* **38,** 6380–6385 (1999).
24. Weigel, P. H. & Deangelis, P. L. Hyaluronan synthases: a decade-plus of novel glycosyltransferases. *J. Biol. Chem.* **282,** 36777–36781 (2007).
25. Kozmon, S. & Tvaroska, I. Catalytic mechanism of glycosyltransferases: hybrid quantum mechanical/molecular mechanical study of the inverting N-acetylglucosaminyltransferase I. *J. Am. Chem. Soc.* **128,** 16921–16927 (2006).
26. Kubota, T. *et al.* Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide α-*N*-acetylgalactosaminyltransferase (pp-GalNAc-T10). *J. Mol. Biol.* **359,** 708–727 (2006).
27. Daras, G. *et al.* The thanatos mutation in *Arabidopsis thaliana* cellulose synthase 3 (AtCesA3) has a dominant-negative effect on cellulose synthesis and plant growth. *New Phytol.* **184,** 114–126 (2009).
28. Benach, J. *et al.* The structural basis of cyclic diguanylate signal transduction by PilZ domains. *EMBO J.* **26,** 5153–5166 (2007).
29. Scheible, W. R., Eshed, R., Richmond, T., Delmer, D. & Somerville, C. Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis* Ixr1 mutants. *Proc. Natl Acad. Sci. USA* **98,** 10079–10084 (2001).
30. Harris, D. M. *et al.* Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A903V and CESA3T942I of cellulose synthase. *Proc. Natl Acad. Sci. USA* **109,** 4098–4103 (2012).
31. Montanier, C. *et al.* Evidence that family 35 carbohydrate binding modules display conserved specificity but divergent function. *Proc. Natl Acad. Sci. USA* **106,** 3065–3070 (2009).
32. Sancho, J. Flavodoxins: sequence, folding, binding, function and beyond. *Cell. Mol. Life Sci.* **63,** 855–864 (2006).
33. Delmer, D. P. Cellulose biosynthesis: Exciting times for a difficult field of study. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50,** 245–276 (1999).
34. Carpita, N. C. Update on mechanisms of plant cell wall biosynthesis: How plants make cellulose and other (1→4)-β-D-glycans. *Plant Physiol.* **155,** 171–184 (2011).
35. Dowd, M. K., French, A. D. & Reilly, P. J. Conformational analysis of the anomeric forms of sophorose, laminarabiose, and cellobiose using MM3. *Carbohydr. Res.* **233,** 15–34 (1992).
36. Momany, F. A. & Schnupf, U. DFTMD studies of β-cellobiose: conformational preference using implicit solvent. *Carbohydr. Res.* **346,** 619–630 (2011).
37. Brown, M. Cellulose structure and biosynthesis: What is in store for the 21st century? *J. Pol. Sci.* **42,** 487–495 (2004).
38. Rapoport, T. A. Protein transport across the endoplasmic reticulum membrane. *FEBS J.* **275,** 4471–4478 (2008).
39. Burton, B. & Dubnau, D. Membrane-associated DNA transport machines. *Cold Spring Harb. Perspect. Biol.* **2,** 1–20 (2010).
40. Whitfield, C. Biosynthesis and assembly of capsular polysaccharides in *Escherichia coli*. *Annu. Rev. Biochem.* **75,** 39–68 (2006).
41. Raetz, C. R. H. & Whitfield, C. Lipopolysaccharide endotoxins. *Annu. Rev. Biochem.* **71,** 635–700 (2002).
42. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60,** 2126–2132 (2004).

Supplementary Information is available in the online version of the paper.

Acknowledgements We are grateful to G. Murshudov for advice on Refmac refinement and U. and Z. Derewenda for discussions. We thank L. Tamm, M. Wiener, A. Walling and T. Rapoport for critical comments on the manuscript. X-ray diffraction data were collected at GM/CA- and Southeast Regional-Collaborative Access Team beamlines at the Advanced Photon Source (APS), Argonne National Laboratory. Use of the APS was supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, contract no. DE-AC02-06CH11357 and W-31-109-Eng-38. GM/CA at APS has been funded in whole or in part with funds from the National Cancer Institute (Y1-CO-1020) and the National Institute of General Medical Sciences (Y1-GM-1104). The University of Georgia CCRC is supported by the Department of Energy funded Center for Plant and Microbial Complex Carbohydrates (DE-FG02-09ER-20097). J.L.W.M. is partially supported by a Peach Fellowship, University of Virginia. J.Z. is support by NIH grant 1R01GM101001 and start-up funds from the University of Virginia School of Medicine.

Author Contributions J.Z. designed the experiments. J.L.W.M. and J.Z. expressed, purified and crystallized the BcsA–BcsB complex. J.L.W.M. and J.Z. analysed the crystallographic data and built the model. J.S. performed *in vitro* cellulose synthesis assays. J.L.W.M. and J.Z. wrote the manuscript.

Author Information The atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession number 4HG6. 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. Correspondence and requests for materials should be addressed to J.Z. (jochen_zimmer@virginia.edu).

©2013 Macmillan Publishers Limited. All rights reserved

METHODS

Expression and purification of BcsA–BcsB. The *bcsA* gene was PCR amplified from *Rhodobacter sphaeroides* 2.4.1. genomic DNA (ATCC) and cloned using NcoI and HindIII restriction sites into an engineered pETDuet expression vector (Novagen) conferring a C-terminal dodeca-histidine tag. The *bcsB* gene was PCR amplified without its native, N-terminal signal sequence. The gene coding for amino acids 21–725 of BcsB was cloned into the pETDuet vector containing the *bcsA* gene, together with an N-terminal PelB signal sequence using NdeI and KpnI restriction sites. After signal sequence cleavage, this generates an additional Met-Gly dipeptide at the N terminus of BcsB. Electro-competent Rosetta 2 cells were transformed with the pETDuetBcsA-12His-BcsB construct and plated on agar plates supplemented with 100 μg ml^{-1} ampicillin and 25 μg ml^{-1} chloramphenicol and grown overnight at 37 °C. The transformed cells were re-suspended and used to inoculate 36 l of ZYP-5052 auto-induction medium[43]. The cells were grown shaking at 37 °C for at least 4 h after the optical density at 600 nm reached 0.8. Subsequently, the cells were collected by centrifugation, re-suspended in RB1-buffer containing 20 mM sodium phosphate buffer pH 7.2, 0.1 M NaCl and 5 mM cellobiose and then lysed in a microfluidizer. The crude membranes were collected by centrifugation for 60 min at 120,000*g* in a Beckman Ti45 rotor and solubilized for 60 min at 4 °C in RB2 buffer containing 20 mM sodium phosphate buffer pH 7.2, 0.3 M NaCl, 5 mM cellobiose, 5 mM $MgCl_2$, 40 mM imidazole, 10% glycerol and 2% Triton X-100. The insoluble material was cleared by centrifugation for 30 min at 120,000*g* in a Beckman Ti45 rotor and the membrane extract was batch incubated with 10 ml Ni-NTA agarose (Qiagen) for 45 min at 4 °C. The resin was packed in a gravity flow chromatography column, washed with 75 ml WB1-buffer (RB2-buffer containing a total of 60 mM imidazole and 5 mM lauryldimethylamine-*N*-oxide, LDAO), 25 ml WB2-buffer (RB2 buffer containing a total of 1 M NaCl and 5 mM LDAO) and the BcsA–BcsB complex was eluted in 50 ml EB-buffer containing 20 mM Tris HCl pH 7.5, 0.1 M NaCl, 5 mM cellobiose, 5 mM $MgCl_2$, 250 mM imidazole, 10% glycerol and 5 mM LDAO. The eluted protein was concentrated and purified over an analytical S200 gel filtration column (GE Healthcare) equilibrated in EB buffer containing 20 mM MES pH 6.5 instead of Tris HCl and lacking $MgCl_2$ and imidazole (GF-buffer). The eluted BcsA–BcsB was concentrated to 50 μM final concentration (using an extinction coefficient at 280 nm of 161,925 M^{-1} cm^{-1}) in a 100-kDa cut-off centrifugal filter (Millipore) and supplemented with 10 mM EDTA (EDTA) and 1 mM uridine diphosphate-glucose (UDP-Glc) before crystallization.

Selenomethionine-derivatized BcsA–BcsB was prepared as described above with the exception that the cells were grown in the auto-inducing PASM-5052 minimal medium supplemented with 125 μg ml^{-1} selenomethionine (Se-Met)[43].

The BcsA–BcsB complex was crystallized in 30% PEG200, 0.1 M MES pH 6.5 and 50 mM NaCl at 4 °C. Initial crystals were observed after approximately 7 days and reached their final size of about 50 × 50 × 100 μm within 3 weeks. The crystals were collected and directly cryo-cooled in liquid nitrogen. Wild-type crystals were soaked with 20 mM samarium III-chloride (Sm) or 1 mM sodium ethylmercuric-thiosalicylate (EMTS) for 2–24 h before collection.

Structure determination. X-ray diffraction data were collected at the Argonne National Laboratory beamlines SER- as well as GM/CA-CAT. The data were indexed and integrated in Mosflm[44], analysed in Pointless and scaled and reduced in space group $P4_32_12$ in Scala from the CCP4 suite of programs[45]. The Sm-derivatized crystals diffracted to 5 Å resolution and were used for single-wavelength anomalous dispersion (SAD) phasing in ShelX[46]. The initial phases allowed localization of all transmembrane helices of BcsA and were further used to calculate an anomalous difference Fourier electron density map for the data set obtained from Se-Met-derivatized BcsA–BcsB. Initially, ten Se-Met positions were identified and refined in Mlphare[45]. The phases were combined with the native data set followed by density modification with phase extension to 3.25 Å in DM[47] using a user-generated solvent mask. Throughout the model building process, Se-Met anomalous difference Fourier electron density maps were calculated, additional Se-Met positions were refined as described above and new phases were calculated and improved by density modification in Parrot[47]. The final experimental phases were calculated from 36 Se-Met-, 4 EMTS- and 8 Sm positions and allowed tracing of most of the backbone of BcsA and the membrane proximal regions of BcsB. The model was built in Coot[42] via an iterative process involving careful model building, combination of experimental and model phases, density modification and restraint refinement in Refmac-5[45] (including TLS parameters[48] for the final model). The model includes a β-1,4-linked glucan containing 18 glucose molecules and a UDP molecule, as well as 2 partially ordered LDAO molecules. The UDP molecule was refined with 50% occupancy owing to weak electron density. The refined model has 91.8% of its residues in the favoured region of the Ramachandran diagram and includes 1.1% outliers, as determined by MolProbity. All figures were prepared in Pymol[49] and the sequence conservation score for BcsB was calculated using CONSURF[50].

***In vitro* cellulose synthesis.** 100 μl of BcsA–BcsB in GF buffer was incubated with 5 mM UDP-Glc, 20 mM $MgCl_2$ and 30 μM cyclic-di-GMP at 37 °C for 90 min. The polymerization reaction was terminated with 40 mM EDTA and the sample was centrifuged at 14,000*g* at room temperature for 20 min. The supernatant was discarded and the pellet was washed twice in 500 μl deionized water, dried and subjected to linkage analysis by the Complex Carbohydrate Research Center, University of Georgia.

43. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41,** 207–234 (2005).
44. Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **62,** 48–57 (2006).
45. CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–763 (1994).
46. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64,** 112–122 (2008).
47. Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. D* **66,** 470–478 (2010).
48. Painter, J. & Merritt, E. A. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D* **62,** 439–450 (2006).
49. PyMol. The PYMOL Molecular Graphics System (DeLano Scientific) (2000).
50. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res.* **38,** W529–W533 (2010).

©2013 Macmillan Publishers Limited. All rights reserved