

Crystallographic snapshot of cellulose synthesis and membrane translocation

Jacob L.W. Morgan, Joanna Strumillo^a and Jochen Zimmer*

*Center for Membrane Biology
Department of Molecular Physiology and Biological Physics
University of Virginia*

480 Ray C. Hunt Dr.
Snyder Bldg. #362
Charlottesville, VA 22908, USA
Phone: 434 243 6506, Fax: 434 982 1616
Email: jochen_zimmer@virginia.edu

* To whom correspondence should be addressed.

^a Faculty of Biology and Environmental Protection, University of Lodz, Pilarskiego Street 14/16, 90-231 Lodz, Poland

Supplementary Information:

Supplementary Table

Supplementary Figures 1 - 10

Supplementary Table. Crystallographic data collection and refinement statistics.

	Wild-type	Se-Met-BcsA-B	SmCl₃-soaked	EMTS-soaked
Data collection				
Space group	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2	P4 ₃ 2 ₁ 2
Cell dimensions <i>a, b, c</i> (Å)	103.1, 103.1, 468.3	104.3, 104.3, 470.5	103.6, 103.6, 470.3	103.0, 103.0, 469.5
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	35-3.25 (3.42- 3.25) *	50-4.64 (4.89- 4.64)	50-5.0 (5.27- 5.0)	50-3.97 (4.19- 3.97)
Wavelength (Å)	0.97949	0.97949	1.84527	1.00595
<i>R</i> _{meas}	0.1 (0.54)	0.157 (>1.0)	0.18 (0.96)	0.089 (0.28)
Mean <i>I</i> / σI	17 (4.8)	15.1 (3.1)	15.9 (4.2)	29.6 (13.4)
Completeness (%)	99.9 (100.0)	97.9 (100.0)	100.0 (100.0)	99.9 (100.0)
Redundancy	15.2 (15.1)	22.4 (21.5)	26.9 (27.8)	26.8 (27.6)
Refinement				
Resolution (Å)	35-3.25			
No. reflections				
Total	39,037			
<i>R</i> _{free}	1,951			
<i>R</i> _{work} / <i>R</i> _{free}	21.28 / 28.17			
No. atoms				
Protein	10,788			
β -1,4 glucan	199			
UDP	25			
LDAO	25			
<i>B</i> -factors				
Chain A	135.8			
Chain B	148.0			
β -1,4 glucan	119.0			
UDP	89.0			
R.m.s deviations				
Bond lengths (Å)	0.007			
Bond angles (°)	1.37			

*Values in parentheses refer to the highest-resolution shell.

Supplementary Figures

```

At_CESA1 1 MEASAGLVAGSYRRNELVRI RHESDGGTKPLKNMNGQICQICGDDVGLAETGDVVFVACNE
At_CESA3 1 MESEG-----ETAG--KPMKNIVPQTCQICSDNVGKTVGDGRFVACDI
At_CESA6 1 MNTGGRLIAGSHRNEFVLINADENARIRSVQELSGQTCQICRDEIELTVDGEPFVACNE
Rs_BcsA 1 -----

At_CESA1 61 CAFPVCRPCYEYERKDGTOCCPQCKTRFRHRHRSRVEGDEDEDDVDDIEN-EFNYAOGA
At_CESA3 42 CSFPVCRPCYEYERKDGNSQCPQCKTRYKRLKGS PAIPGDKDEGLADEGTVEFNYPQKE
At_CESA6 61 CAFPVCRPCYEYERREGNQACPQCKTRFKRLKGS RVEGDEEEDDIDLDN-EFEYGNNG
Rs_BcsA 1 -----

At_CESA1 120 NKAR-----HQRHGEFSSS--SRHESQ-PIPLLTHGHTVS GEIRTPDTSVTRTSGP
At_CESA3 102 KISERMLGWLTRGKGEEEMGEPQYDKEVSHNHL PRLTSRQDTSGEFSAASPERLSVSS-T
At_CESA6 120 IGFDQVSEGMSISRNSGFPQSDLDSAPP GSGQIPLLTYGD---EDVEISSDRHALIVPPS
Rs_BcsA 1 -----

At_CESA1 170 LGPSDRNAISSPYIDPRQVPVRI VDP SKDLNSYGLGNVDWKERVEGWK LKQEK NMLQMT
At_CESA3 161 IAGGKRLPYSS---DVNQSPNRRIVDP-----VGLGNVAWKERVDGWKMKQEKNTGPVS
At_CESA6 177 LGGHGNRVHPVLSLSDPTVA AHPRMPVQKDLAVYGYGSVAWKDRMEEWKRKQNEKLQVVR
Rs_BcsA 1 -----MTVRAKARSPLRVVPVLLFLLWVALLVPPFGLLAA

At_CESA1 230 GKYHEGKGG-EIEG-TGSNGEELQ MADDTRL PMSRVVPI PSSRLTPYRVV IILRLIILCF
At_CESA3 212 TQAASERGGVDIDASTDILADEALLNDEARQ PLSRKVSI PSSRINPYRMVIMLRLVILCL
At_CESA6 237 -----HEGDDPFEDGDD---ADFPMMDEGRQ PLSRKIPIKSSKINPYRMLIVLRLVILGL
Rs_BcsA 35 AP-----VAPSAQGLIALSAVVLVALLKPFADKMVPRFLLLSAASMLVMR--YFWWR

At_CESA1 288 FLQYRTHPVKNAYPLWLT SVICEIWF AFSWLLDQFPKWYPI NRETYLDR LALRYDRDGE
At_CESA3 272 FLHYRITNPVPNAFALWLV SVICEIWF ALSWILDQFPKWF PVNRETYLDR LALRYDREGE
At_CESA6 289 FFHYRILHPVKDAYALWLV SVICEIWF AVSWVLDQFPKWYPI ERETYLDR LSLRYEKEGK
Rs_BcsA 85 LFETLPPPALDASFLFALLLFAVETFSISIFFLNGFLSADPTDRP-----FPRPLQ

At_CESA1 348 PSQLVPVDV FVSTVDPLKEPPLVTANTVLSILSVDY PVDKVACIYVSDDG SAMLTFESLSE
At_CESA3 332 PSQLAAVDIFVSTVDPLKEPPLVTANTVLSILAVD YPVDKVACIYVSDDGAAMLSFESLAE
At_CESA6 349 PSGLSPVDV FVSTVDPLKEPPLITANTVLSILAVD YPVDKVACIYVSDDGAAMLTFEALSE
Rs_BcsA 136 PEELPTVDILVPSYN--EPADMLSVTLAAAKNMIY PARLRTVVLCDG-----

At_CESA1 408 TAEFAKKWVPFCKKFNIEPRAPEFYFAQKIDY LKDKIQPSFVKERRAMKREYEEFKVRIN
At_CESA3 392 TSEFARKWVPFCKKYSIEPRAPEWYFAAKIDY LKDKVQTSFVKDRRAMKREYEEFKIRIN
At_CESA6 409 TAEFARKWVPFCKKYCIEPRAPEWYFCHKMDY LKNKVHPAFVRRRAMKR DYEEFKVKIN
Rs_BcsA 182 -----GTDQRCMS P-----DPELAQKAQERRRELQQLCRELG

At_CESA1 468 ALVAKAQKIPEEGWTMQDGPWP GNNTRDHPGMIQVFLGHSGGLD TDGNELPRLIYVVSRE
At_CESA3 452 ALVSKALCPEEGWVMQDGPWP GNNTRDHPGMIQVFLGQNGGLDAEGNELPRLVYVVSRE
At_CESA6 469 ALVATAQKVPEDGWTMQDGPWP GNSVRDHPGMIQVFLGSDGVRDVENNELPRLVYVVSRE
Rs_BcsA 214 -----VVYSTRE

At_CESA1 528 KRPGFQH H K K A G A M N A L I R V S A V L T N G A Y L L N V D C D H Y F N N S K A I K E A M C F M M D P A I G K K
At_CESA3 512 KRPGFQH H K K A G A M N A L V R V S A V L T N G P F I L N L D C D H Y I N N S K A L R E A M C F L M D P N L G K Q
At_CESA6 529 KRPGFDH H K K A G A M N S L I R V S G V L S N A P Y L L N V D C D H Y I N N S K A L R E A M C F M M D P Q S G K K
Rs_BcsA 221 R---NEHAKAGNMSAALER---LKGELVVVF DADHVPSRDFLARTVGYFVEDPDLFLV
&
At_CESA1 588 CCYVQFPQRF D G I D L-----H D R Y A N R N I V F F D I N M K G L D G I Q G P V Y V G T G C C F N R Q A
At_CESA3 572 CCYVQFPQRF D G I D K-----N D R Y A N R N T V F F D I N L R G L D G I Q G P V Y V G T G C V F N R T A
At_CESA6 589 ICYVQFPQRF D G I D R-----H D R Y S N R N V V F F D I N M K G L D G I Q G P I Y V G T G C V F R R Q A
Rs_BcsA 273 ---QTPHFIFINDP IQRNLALGDRCPPE NEMFYGKIHRGLDRWGGAFFCGSAAVLRRRA
IF1

At_CESA1 641 LYGYDPVLTEEDLEP-----N I I V K S C C G S R K K G K S S K K Y N Y E K R R G I N R S D S N A P L F N M
At_CESA3 625 LYGYEPPIKVHKHKP-----S L L S K L C G G S R K K N S K A K K E S D K K K S G R - H T D S T V P V F N L
At_CESA6 642 LYGFDA P K K K K G P R K T C N C W P K W C L L C F G S R K N R K A K T V A A D K K K K N R ---E A S K Q I H A L
Rs_BcsA 329 LD-----

At_CESA1 696 EDIDEGFEG--YDDERSILMSORSVEKRFGQSPVFIAATFMEQGGIPPTTNPATLLKEAI
At_CESA3 679 DDIEEGVEGAGFDDEKALLMSQMSLEKRFGQSAV FVASTLMENGGVPPSATPENLLKEAI
At_CESA6 699 ENIEEGRVTKGSNVEQSTEAMQMKLEKKFGQSPV FVASARMENGGMARNASPACLLKEAI
Rs_BcsA 331 -----EAGGFAG-----

At_CESA1 754 HVISCGYEDKTEWGKEIGWIYGSVTE D I L T G F K M H A R G W I S I Y C N P P R P A F K G S A P I N L S
At_CESA3 739 HVISCGYEDKSDWGMEIGWIYGSVTE D I L T G F K M H A R G W R S I Y C M P K L P A F K G S A P I N L S
At_CESA6 759 QVISCGYEDKTEWGKEIGWIYGSVTE D I L T G F K M H S H G W R S V Y C T P K L A A F K G S A P I N L S
Rs_BcsA 338 -----ETI TEDAETALEI HSRGWKSLYIDR--AMIAGLQPETFA

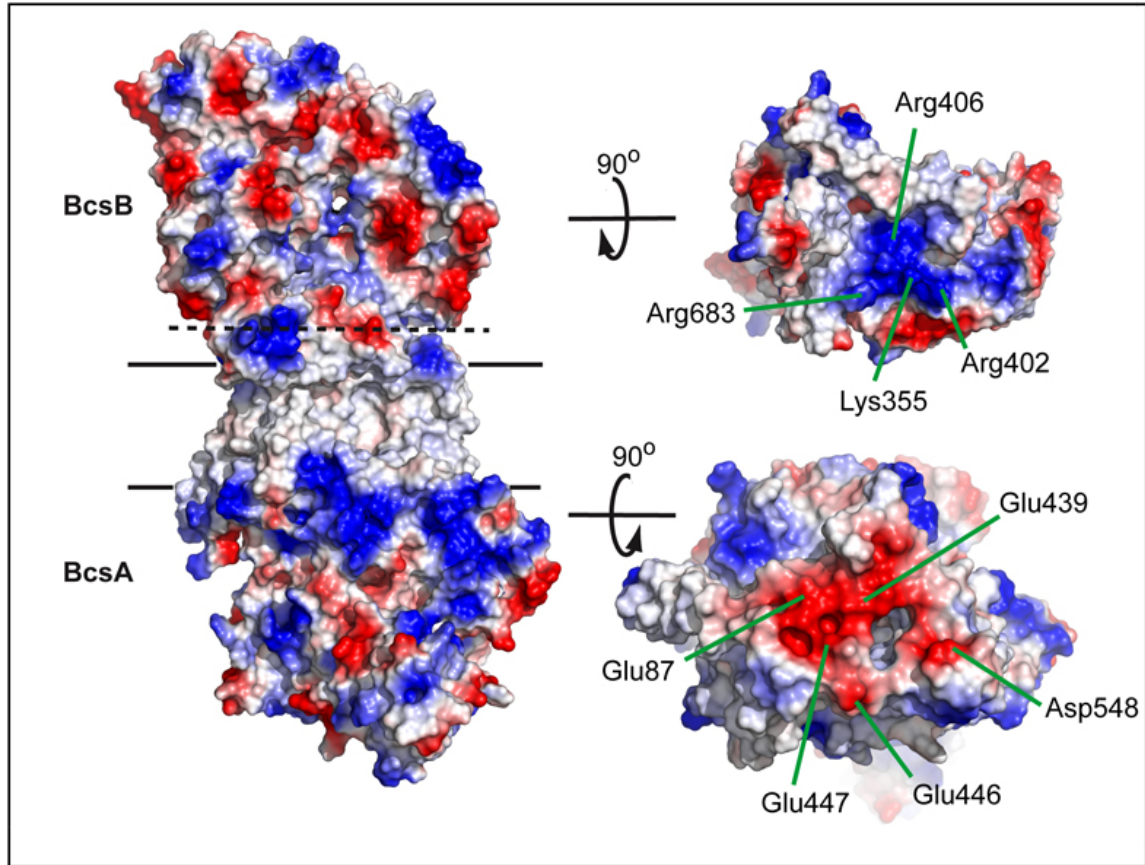
```

```

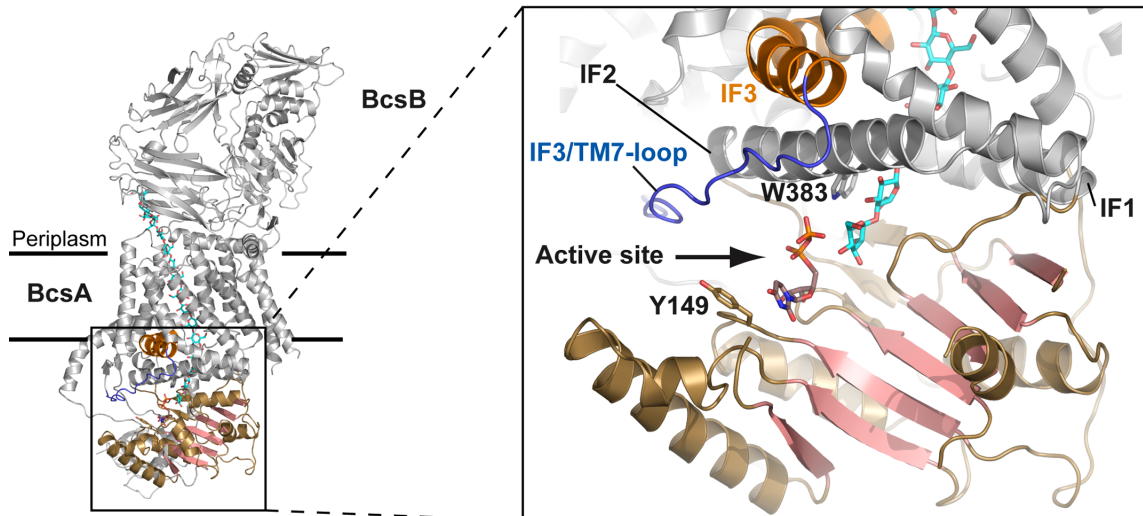
                                     IF2                                     TM5
At_CESA1  814  DRLNQVLRWALGSIEILLRHCPIWYGYHGRLRLLERIA&YINTIVYPITSIPLIAYCILP
At_CESA3  799  DRLNQVLRWALGSVEILFSRHCPIWYGYNGRLKFLERFAYVNTTIYPITSIPLLMYCTLP
At_CESA6  819  DRLHQVLRWALGSVEIFLSRHCPIWYGYGGGLKWLERLSY*INSVVYPWTSPLIVYCSLP
Rs_BcsA    375  SFIQQRGRWATGMMQMLLLKNPLFRRGLG----IAQRLCYLNSMSFWFFPLVRMMFLVAP
                                     TM6                                     IF3
At_CESA1  874  AFCLITDRFI#IPEISNYASIW&FILLFISIAVTGILELRWSGVSIEDWWRNEQFWVIGGTS
At_CESA3  859  AVCLFTNQFI#IPQISNIASIWFLSLFLSIFATGILEMRWSGVGIDEDWWRNEQFWVIGGVS
At_CESA6  879  AICLLTGKFI#IVPEISNYASILFMALFSSIAITGILEMQWGKVGIDEDWWRNEQFWVIGGVS
Rs_BcsA    431  LIY#LVFFGIEIF-----VATFEEVLAYMPGYLAVSFLVQNALFARQRWPLVSEVYEVAQAP
                                     TM7
At_CESA1  934  AHLFAVFQGLLKVL#LAGIDTNFTVTSKATDEDGDFAE#LYIFKWTALLIPPTTVLLVNLIGI
At_CESA3  919  AHLFAVFQGLILKVL#LAGIDTNFTVTSKASDEDGDFAE#LYLFKWTLLIPPTLLIVNLVGV
At_CESA6  939  AHLFALFQGLLKVL#AGVDTNFTVTSKAAD-DGEFSDLYLFKWTSL#LIPMTLLIINVIGV
Rs_BcsA    486  YLARAI#VTTLLRPRS---ARFAVTA#KDETLS-----ENYISPIYRPLLETFLLC
                                     TM8
At_CESA1  994  VAGVSYAVNSGYQSWG#PFLGK#LFALWVIAHLYPFLKGLLGRQNRTP#TIVIVWSVLLASI
At_CESA3  979  VAGVSYA#INSGYQSWG#PFLGK#LF#AFWVIVHLYPFLKGLMGRQNRTP#TIVVWSVLLASI
At_CESA6  998  IVGVSDAISNGYDSWG#PFLGRLFFALWV#IHLYPFLKGLLGGQDRMPT#IIVVWSVLLASI
Rs_BcsA    532  LSGV#LATLVR---WVAFPG-----DRSVLLLVGGWAVLNVLL
                                     TM1
At_CESA1  1054  FSL#LWVRINPFVDANPNANNFNGKGGVF
At_CESA3  1039  FSL#LWVRIDPFTSRVTGPDILECGINC-
At_CESA6  1058  LTL#LWVRVNPFFVAKG-GPILEICGLDCL
Rs_BcsA    566  VGFALRAVAEKQRR-----

```

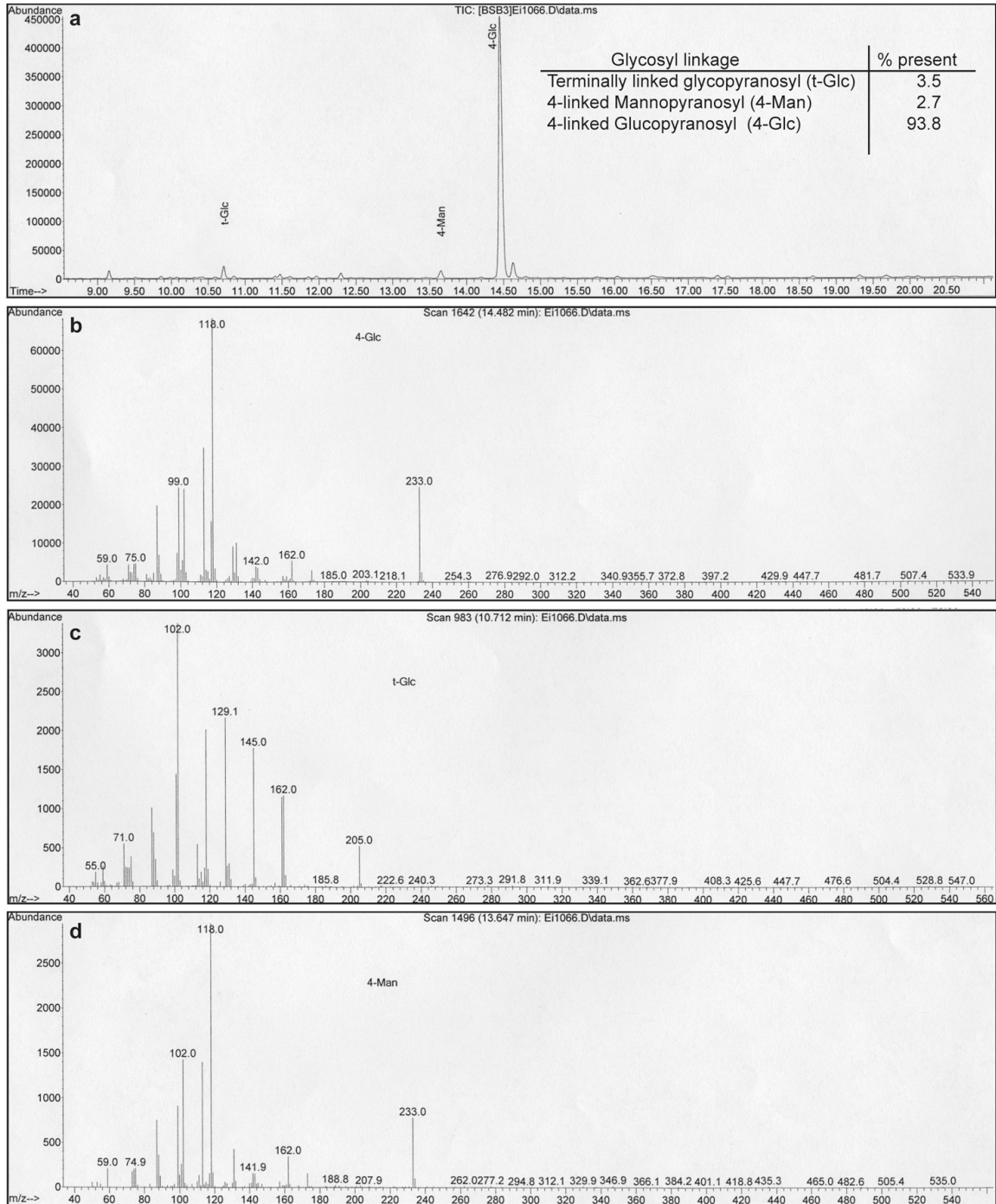
Supplementary Figure 1 | Sequence alignment of *Rhodobacter sphaeroides* BcsA and *Arabidopsis thaliana* CESAs. Residues 1 to 580 of *R. sphaeroides* (Rs) BcsA are aligned with the transmembrane (TM)- and glycosyltransferase domains of *A. thaliana* (At) CESA1, 3 and 6. BcsA's C-terminal PilZ-domain was omitted from the alignment. The TM- and cytoplasmic interface helices of BcsA are shaded green and blue, respectively. The predicted CESA TM-helices are indicated with a black dashed line above the alignment. The locations of the dominant negative *thanatos* mutation (Pro578Ser) and the quinoxypen and isoxaben resistance conferring mutations (Ala903Val and Thr942Ile) in *At* are framed blue and are indicated with the symbols “&”, “*” and “#” above the alignment, respectively. TM1 and 2 of BcsA are not present in eukaryotic CESAs and are indicated with a green box. We note that TM5 of CESA aligns with IF3 of BcsA and CESA TM7 partially aligns with the periplasmic loop connecting BcsA's TM7 and 8. Additional experiments are required to determine whether these predicted TM-helices indeed span the membrane or form interface helices on the intra- and extracellular sides of the membrane.



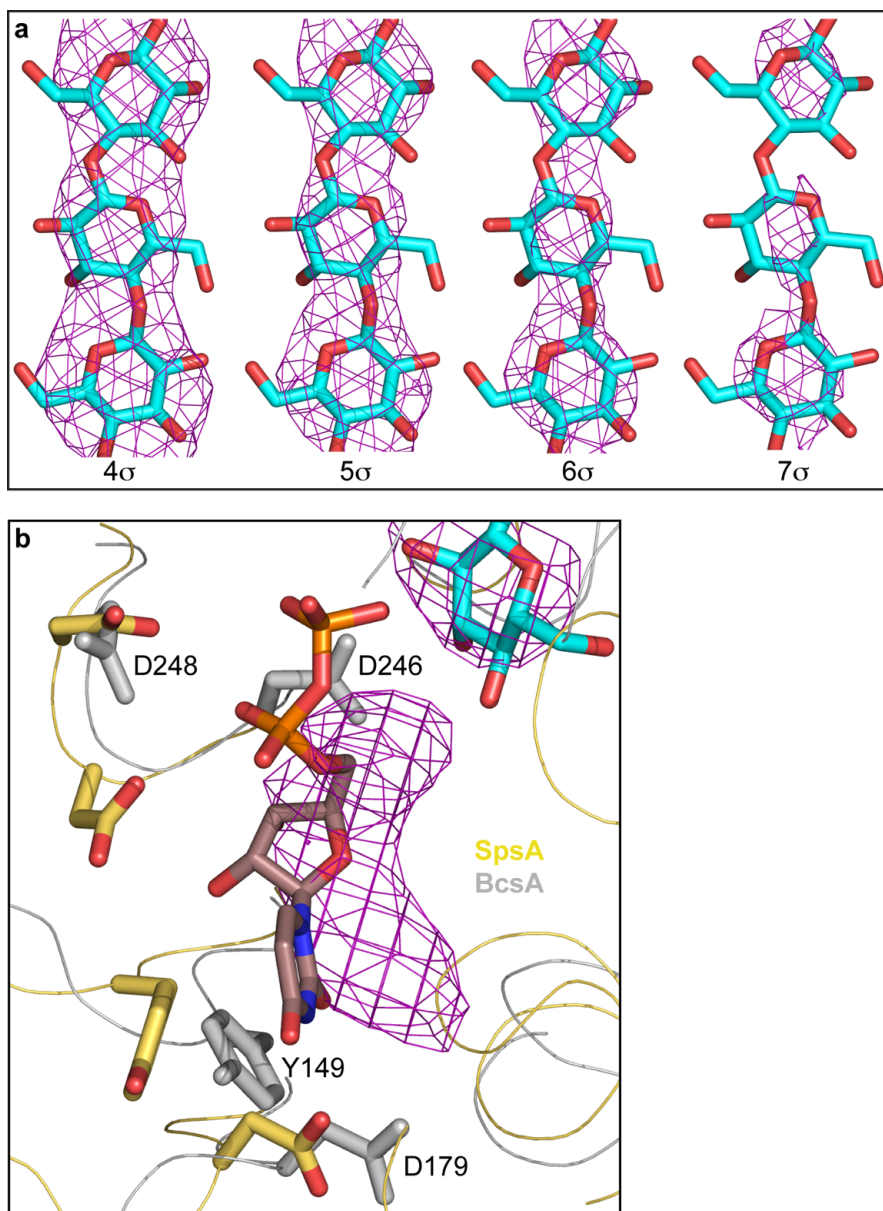
Supplementary Figure 2| Vacuum electrostatics of BcsA-B. The interaction of BcsA and BcsB is stabilized by a 4500\AA^2 large interface that includes a cluster of negatively and positively charged residues on BcsA and BcsB, respectively. In addition, the TM-helix of BcsB interacts with TM-helices 1 to 3 of BcsA. The dashed line indicates the periplasmic BcsA-B interface shown on the right for the individual subunits. Horizontal bars indicate the membrane boundaries.



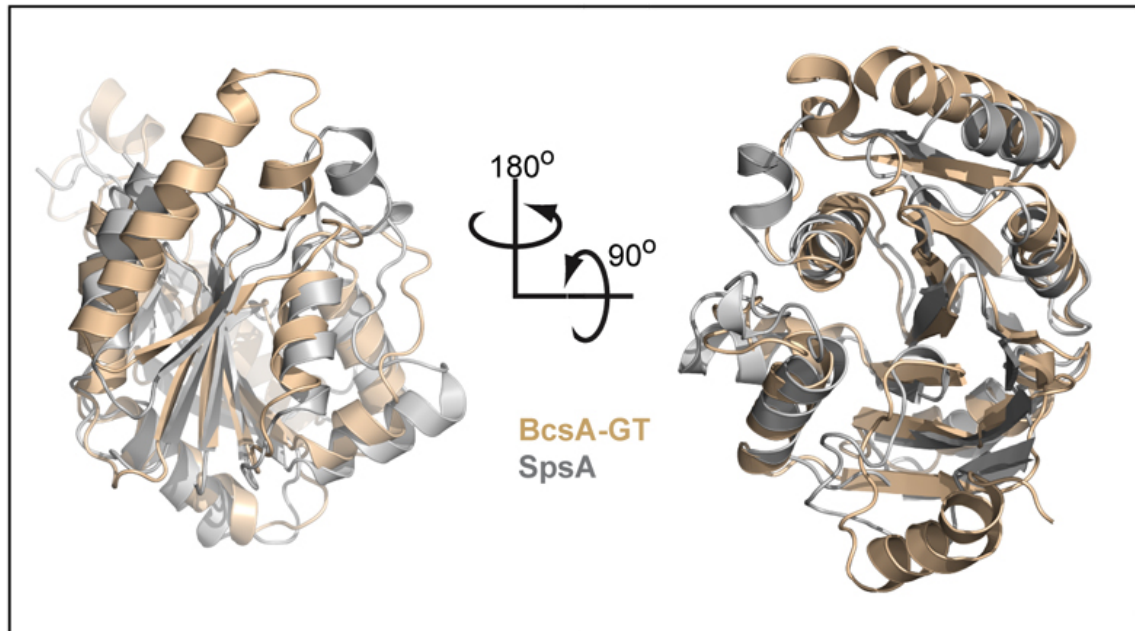
Supplementary Figure 3| The glycosyltransferase domain of BcsA. The loop connecting BcsA's TM4 and -5 (colored sand and light red) adopts a GT-A fold and forms the catalytic domain of the synthase. Access to the active site is likely controlled by the position of the conserved IF3/TM7-loop (shown in blue). Residues 223-238 and 334-354 of the GT-domain are omitted for clarity, UDP and the translocating glucan are shown as violet and cyan sticks, respectively. Trp383 of the "Q(Q/R)xRW" motif and Tyr149 are shown as sticks.



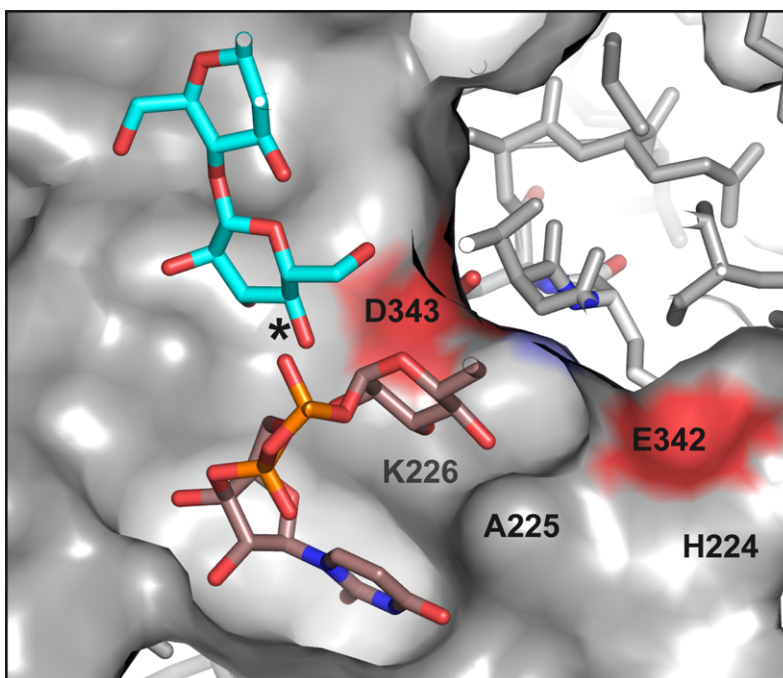
Supplementary Figure 4| Glycosyl linkage analysis of *in vitro* synthesized cellulose. The detergent solubilized *Rhodobacter* BcsA-B complex was incubated with UDP-Glc and cd-GMP at 37°C. The synthesized polymer was sedimented, washed and subjected to linkage analysis by the Complex Carbohydrate Research Center, University of Georgia. The analysis involved permethylation, depolymerization, reduction and acetylation of the polymer followed by coupled gas chromatography (a) and mass spectrometry (b-d) to identify the partially methylated alditol acetates.¹



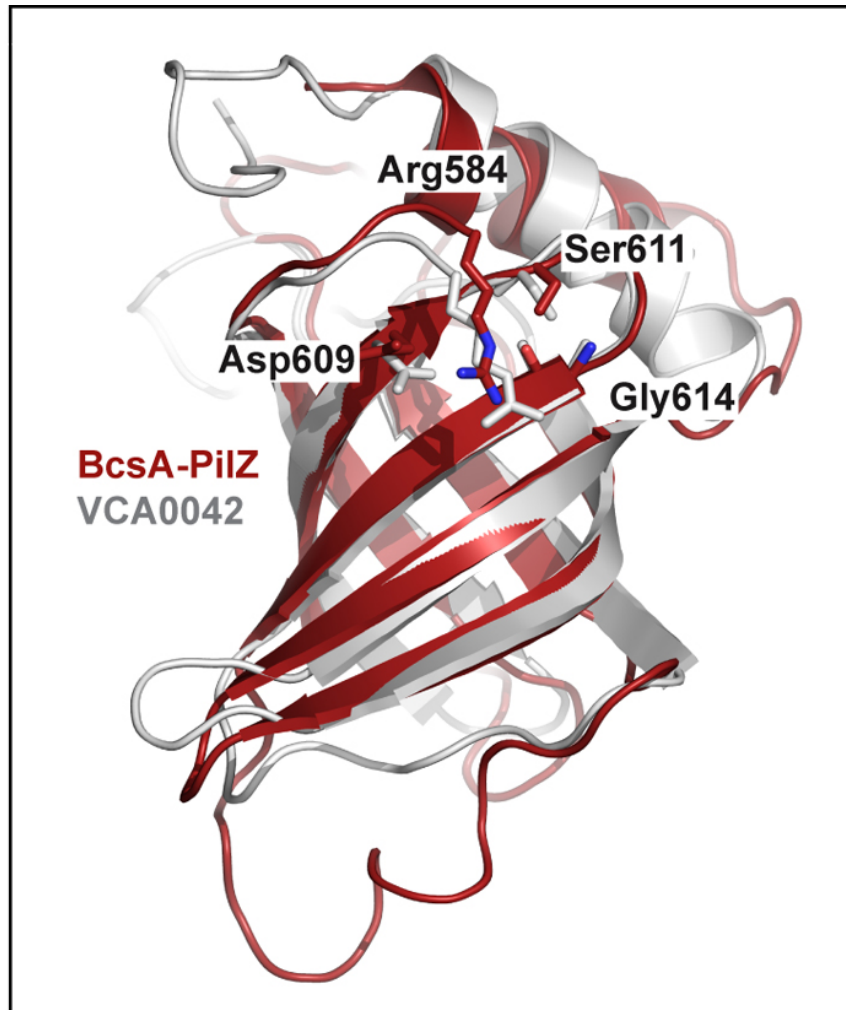
Supplementary Figure 5| Difference Fourier electron density of the translocating glucan and UDP. a, The glucose units of the translocating glucan occupy distinct positions along the translocation path. The unbiased, positive FoFc-difference Fourier electron density was contoured at the indicated levels and reveals the position of the individual glucose units. The density was calculated with phases obtained from a BcsA-B model prior to placing and refining the glucan. **b,** BcsA was superimposed with the UDP-bound structure of SpsA (pdb entry 1QGS). BcsA is shown in grey, SpsA is colored yellow and the SpsA-bound UDP molecule is shown as violet sticks. Conserved residues coordinating UDP in SpsA and BcsA are shown as sticks and are labeled for BcsA. The coordination of the hook-shaped density and the good agreement with the SpsA-UDP complex support its interpretation as a weakly bound UDP molecule.



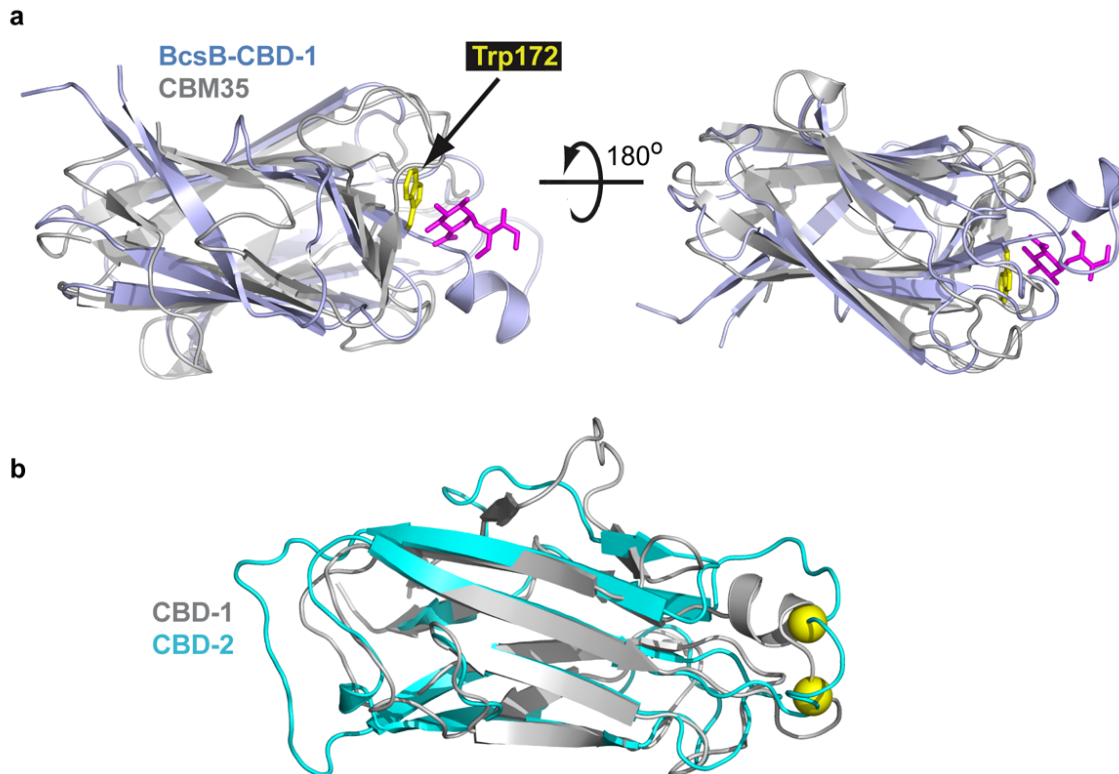
Supplementary Figure 6 | Superimposition of BcsA's glycosyltransferase domain with SpsA. Residues 142 to 367 of BcsA were superimposed with residues 2 to 217 of SpsA (pdb entry 1QGS) in Coot by secondary structure matching, SSM. Both proteins adopt a GT-A fold and align with an r.m.s.d. of 2.15Å between C α atoms.



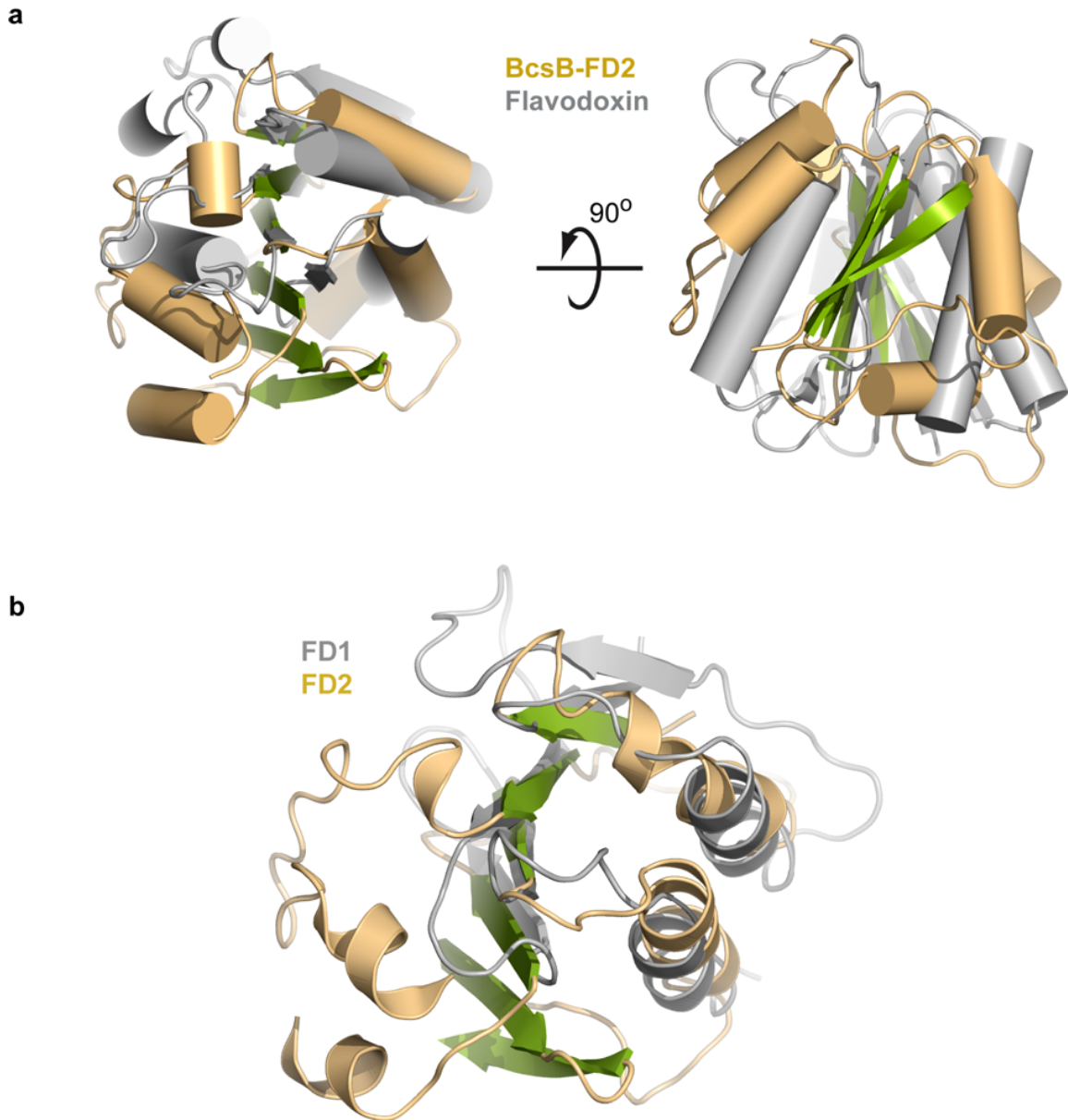
Supplementary Figure 7 | Proposed donor glucose binding site. Surface representation of BcsA's GT-domain. Residues 224-226 of the "HAKAGN" and 341-343 of the "TED" motifs might form a binding site for the donor Glc next to the UDP binding pocket. This would position the β -face of the anomeric carbon towards the 4' hydroxyl of the acceptor (*). The acceptor glucan and UDP-Glc are shown as cyan- and violet sticks. UDP-Glc was manually docked into the binding pocket.



Supplementary Figure 8 | Superimposition of BcsA's PilZ-domain with the VCA0042 protein. Residues 584 to 693 of BcsA were superimposed with residues 139 to 238 of the cyclic-di-GMP binding protein VCA0042 (pdb entry 3KYG) in Coot by secondary structure matching, SSM. The C α atoms of both β -barrels align with an r.m.s.d. of 2.2Å. Conserved residues of BcsA likely implicated in cyclic-di-GMP binding are labeled and shown as sticks.



Supplementary Figure 9 | Superimposition of BcsB's CBD-1 with the carbohydrate binding module family 35. **a**, Residues 54 to 187 of BcsA were superimposed with the carbohydrate binding module 35 from *Cellvibrio japonicus* (pdb entry 2W87) in Coot by secondary structure matching, SSM. The glucuronic acid disaccharide observed in pdb 2W87 is shown as pink- and the conserved Trp172 of BcsB as yellow sticks. The disaccharide in 2W87 makes CH- π interactions with a Trp residue. **b**, Superimposition of BcsB's CBD-1 and -2. Conserved cysteines forming a disulfide bond between CBD-1 and -2 are indicated with a yellow sphere for their C α atoms.



Supplementary Figure 10| BcsB's FD-domains adopt a flavodoxin fold. a, Flavodoxin from *Desulfovibrio desulfuricans* (pdb entry 3KAP) was aligned with FD2 of BcsB. The structures share a 4-stranded β -sheet that is framed by two α -helices on either side. **b,** Superimposition by secondary structure matching of BcsB's FD1 and -2.

Supplementary Reference

1. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. T. & Albersheim, P. Isolation and Characterization of Plant Cell Walls and Cell Wall Components. *Methods Enzymol* **118**, 1-38 (1986).