

Structural Studies of Biopolymer Membrane Transport

Jacob Morgan
Monument, Colorado

B.A. Colorado College, 2008

A Dissertation presented to the Graduate Faculty
of the University of Virginia in Candidacy
for the Degree of Doctor of Philosophy

Interdisciplinary Program in Biophysics

University of Virginia

September 2016

Abstract

The selective transport of specific substrates across the membrane is an essential part of biology. Organisms have evolved a basic mechanism for the transport of ions and small molecule such as sugars across the membrane. This mechanism is termed ‘alternating access’ and involves the alternating exposure of a substrate-binding site to either side of the membrane. While the molecular details of ‘alternating access’ have been revealed for a number of different transporters, this scheme seems infeasible for the transport of long biopolymers such as nucleic acids, proteins, and polysaccharides because ‘alternating access’ requires that the transporter forms a binding site that is large enough to bind the entire substrate. Here, I present novel data addressing the transport mechanism for two bacterial biopolymer transporter proteins: Bacterial Cellulose Synthase which synthesizes and secretes the polysaccharide, cellulose; and PrtD, an ABC transporter which is involved in the secretion of an extracellular protease.

Cellulose is a linear polymer of glucose units that forms a major component of plant cell walls as well as many bacterial biofilms. In bacteria, the components required for cellulose biosynthesis are encoded in a single operon minimally made up of the genes *bcsA*, *bcsB*, *bcsC*, and *bcsZ*. BcsA is an integral inner-membrane (IM) glycosyltransferase enzyme that is responsible for coupling the synthesis of cellulose with its transport across the IM. BcsB is a periplasmic protein with C-terminal IM anchor, and BcsB forms a complex with BcsA (BcsA–B) that is sufficient for *in vitro* cellulose synthesis. BcsZ is a periplasmic cellulase enzyme, and BcsC is an outer-membrane β -barrel that presumably forms a pore for the cellulose polymer to cross the outer membrane.

BcsA–B activity is stimulated by the bacterial signaling molecule, cyclic-di-GMP (c-di-GMP), which is a key regulator of biofilm formation. I present crystal structures of the c-di-GMP-bound BcsA–B complex in the presence and absence of UDP, a competitive inhibitor and substrate mimic. The structures reveal that c-di-GMP releases an auto-inhibited state of the enzyme. A salt bridge stabilizes one of the signature c-di-GMP-binding Arg residues in a position to tether a conserved ‘gating loop’ in front of the active site. The binding of c-di-GMP releases the tether and allows substrate to access the active site. Additionally, the UDP-bound structure reveals an additional role for the ‘gating loop’ in coordinating substrate at the active site. Functional experiments confirm the structural interpretation by revealing that disruption of the auto-inhibitory salt bridge by mutagenesis generates a constitutively-active cellulose synthase. The mechanistic insights presented here represent the first examples of how c-di-GMP allosterically modulates enzymatic functions.

To address the mechanism by which BcsA–B transports cellulose across the IM, I use *in crystallo* enzymology with the c-di-GMP-bound BcsA–B crystals. Because crystallized BcsA–B is catalytically active, these experiments provide a detailed molecular movie of the complete cellulose biosynthesis cycle, from substrate binding to polymer translocation. A substrate-bound structure of BcsA reveals the basis for substrate recognition, and structural snapshots show that BcsA translocates cellulose via a ratcheting mechanism, which involves the upward and downward movement of a ‘finger helix’ that contacts the cellulose polymer’s terminal glucose. The movement of this finger helix is coupled to the insertion and retraction of the ‘gating loop’ in response to substrate binding and polymer extension, respectively. Thus, insertion of the gating loop pushes the ‘finger helix’ upwards, which then pushes the elongated polymer into BcsA’s TM channel.

This mechanism is validated experimentally by tethering BcsA's finger helix, which inhibits polymer translocation but not elongation.

While the BcsA–B enzyme couples the synthesis of the biopolymer with its transport across the IM, polypeptide secretion requires that these two processes are carried out by separate machinery. Type 1 secretion systems (T1SSs) represent a widespread mode of protein secretion across the cell envelope in Gram-negatives. The T1SS is composed of an inner-membrane ABC transporter, a periplasmic membrane fusion protein (MFP), and an outer-membrane TolC-like barrel. These three components assemble into a complex spanning both membranes and providing a conduit for the translocation of unfolded polypeptides. Utilizing the *Dickeya dadantii* PrtDEF (DdPrtDEF) system, I show that ATP hydrolysis and assembly of a complete PrtDEF T1SS complex is necessary for protein translocation. Further, I present a 3.15Å crystal structure of *Aquifex aeolicus* PrtD (AaPrtD), a homologue of DdPrtD. The structure suggests a substrate entry window just above the transporter's nucleotide binding domains (NBDs). In addition, highly kinked transmembrane helices frame a narrow channel not observed in canonical peptide transporters and are likely implicated in substrate translocation. Combined, the AaPrtD structure suggests a polypeptide transport mechanism distinct from alternating access.

List of Abbreviations:

ABC transporter, ATP-binding cassette transporter

Abs₃₄₀, absorbance at 340 nm wavelength

ADP, Adenosine diphosphate

ADP-glucose, ADP-derivatized glucose

Alg44, regulatory subunit of alginate synthase

Alg8, synthetic subunit of alginate synthase

AlgE, the β -barrel subunit of alginate synthase

AlgK, the TPR-containing periplasmic subunit of alginate synthase

ATP, adenosine triphosphate

ADP, adenosine diphosphate

BcsA, Bacterial cellulose synthase subunit A (catalytic)

BcsA–B, complex of Bacterial cellulose synthase subunits A and B

BcsB, Bacterial cellulose synthase subunit B

BcsC, Bacterial cellulose synthase subunit C (outer membrane barrel)

BcsZ, Bacterial cellulose synthase subunit Z (cellulase)

c-di-GMP, cyclic diguanylate (cyclic-di-GMP)

C55-PP, undecaprenyl diphosphate

C_{att}, carbon receiving nucleophilic attack

CBD, carbohydrate-binding domain

CeS, cellulose synthase (can be bacterial or eukaryotic)

CesA, plant cellulose synthase

CFTR, cystic fibrosis transmembrane conductance regulator

CH1, coupling helix 1

CHAPS, (3-((3-cholamidopropyl) dimethylammonio)-1-propanesulfonate)

CLD, C39 peptidase-like domain

CSC, cellulose synthase complex

DAG, diacylglycerol

DANO, anomalous difference

DDM, dodecyl maltoside

DGC, diguanylate cyclase
DMPC, di-myristoyl-phosphatidylcholine
DPM, disintegrations per minute
DTBA, dithiobutylamine
DTT, dithiothreitol
EDTA, ethylenediaminetetraacetic acid
EM, electron microscopy
EMTS, ethylmercuric-thiosalicylate
ER, endoplasmic reticulum
FD, flavodoxin-like domain
GDP, guanosine diphosphate
GPCR, G-protein coupled receptor
GT, glycosyltransferase
GTP, guanosine triphosphate
IF, interface helix
IM, inner membrane
IMV, inverted membrane vesicle
IPTG, isopropyl β -D-1-thiogalactopyranoside
ITC, isothermal titration calorimetry
 K_{act} , activation constant
 k_{cat} , turnover number
 K_{D} , dissociation constant
 K_{M} , Michaelis constant
LDAO, lauryldimethylamine oxide
LDH, lactate dehydrogenase
LFC-E-14, lysofoscholine ether 14
LPS, lipopolysaccharide
MATE, multidrug and toxic compound extrusion
MBP, maltose-binding protein
MFP, membrane fusion protein
MFS, major facilitator superfamily

MR, molecular replacement
MSP, membrane scaffold protein
NAD⁺, oxidized nicotinamide adenine dinucleotide
NADH, reduced nicotinamide adenine dinucleotide
NBD, nucleotide-binding domain
NIH, National Institutes of Health
NTA, nitrilotriacetic acid
OG, octyl glucopyranoside
OM, outer membrane
OPX, outer-membrane polysaccharide export
OST, oligosaccharyltransferase
PCP, polysaccharide co-polymerase
PDE, phosphodiesterase
PEG, polyethylene glycol
PEP, phosphoenolpyruvate
PIP, phosphatidylinositol
PL, proteoliposomes
PLC, protein lipase c
PMF, proton motive force
PMSF, phenylmethylsulfonyl fluoride
PNAG, poly N-acetylglucosamine
POPE, palmitoyl, oleoyl phosphatidylethanolamine
RMSD, root-mean squared deviation
RTX, repeats in toxin
SAD, single-wavelength anomalous diffraction
S_N1, unimolecular nucleophilic substitution
S_N2, bimolecular nucleophilic substitution
S_Ni, unimolecular nucleophilic substitution with internal return
SSS, sodium-solute symporter
T1SS, type-1 secretion system
TCEP, Tris-(2-carboxyethyl)phosphine

TLS, translation, libration, screw

TM, transmembrane

TMD, transmembrane domains

TMH, transmembrane helix

TPR, tetratricopeptide repeat

UDP, uridine diphosphate

UDP-2F-Glc, uridine diphosphate-2-fluoro glucose

UDP-6-thio-Gal, uridine diphosphate-6-thio galactose

UDP-CH₂-Glc, uridine diphosphonate glucose

UDP-Gal, uridine diphosphate galactose

UDP-Glc, uridine diphosphate glucose

UDP-NAG, uridine diphosphate N-acetylglucosamine

UDP-xyl, uridine diphosphate xylose

V_{max}, maximum velocity

WT, wild type

List of Figures:

Figure 1 Cyclization of glucose.....	3
Figure 2 An overview of glycosyl transfer reactions	7
Figure 3 GT folds	9
Figure 4 Chemical structures of prominent polysaccharides made by processive GTs	12
Figure 5 Cellulose	13
Figure 6 Cellulose microfibrils.....	16
Figure 7 Biofilm life cycle	18
Figure 8 C-di-GMP physiology.....	21
Figure 9 Symmetry in transporters	25
Figure 10 Polysaccharide secretion.....	30
Figure 11 BcsA–B activity	33
Figure 12 Activity requires both BcsA and BcsB	35
Figure 13 BcsA–B kinetics.....	37
Figure 14 Product inhibition.....	39
Figure 15 BcsA Substrate specificity	40
Figure 16 Architecture of BcsA–B complex	42
Figure 17 BcsA exhibits a GT–A fold.....	43
Figure 18 BcsA’s GT domain	44
Figure 19 BcsA active site and PilZ domain.....	46
Figure 20 BcsA TM domain.....	48
Figure 21 BcsB structural features	49
Figure 22 BcsB necessary features.....	50

Figure 23 Initial model of cellulose biosynthesis.....	51
Figure 24 Structure of the c-di-GMP-activated BcsA–B complex	58
Figure 25 Correction of a register shift in BcsA	60
Figure 26 Cyclic-di-GMP binding to BcsA	61
Figure 27 Conformational changes of BcsA's gating loop.....	63
Figure 28 Features of BcsA's gating loop.....	65
Figure 29 Stabilization of the gating loop by Arg580.....	66
Figure 30 Insertion of the gating loop into the catalytic pocket.....	68
Figure 31 A conserved salt bridge stabilizes the resting position of BcsA's gating loop.....	69
Figure 32 Comparison of BcsA-catalyzed <i>in vitro</i> cellulose synthesis in the absence and the presence of c-di-GMP	70
Figure 33 Cellulose synthesis activity of BcsA mutants and c-di-GMP binding.....	71
Figure 34 Movement of the finger helix, cellulose translocation, and the acceptor position.....	73
Figure 35 The movement of BcsA's finger helix is supported by a small loop.....	75
Figure 36 Active site signature motifs involved in donor and acceptor coordination	77
Figure 37 <i>In crystallo</i> cellulose biosynthesis.....	94
Figure 38 Conformational flexibility of the gating loop after cellulose extension.....	95
Figure 39 <i>In crystallo</i> translocation of a 6-thio-galactose-containing cellulose polymer	97
Figure 40 Position of the disulfide-tethered finger helix	99
Figure 41 Movement of BcsA's finger helix is essential for cellulose translocation.....	100
Figure 42 The product-bound state.....	102
Figure 43 The substrate-bound state.....	104
Figure 44 Comparison of the UDP conformation in the substrate and UDP-bound states of BcsA	105

Figure 45 Model of cellulose biosynthesis	106
Figure 46 UDP-Glc induced polymer translocation	107
Figure 47 Stabilization of BcsA's finger helix by conserved residues.....	109
Figure 48 <i>In vivo</i> secretion of DdPrtG	132
Figure 49 AaPrtD Properties.....	134
Figure 50 Phasing of AaPrtD	135
Figure 51 Structure of AaPrtD	137
Figure 52 Comparison to peptide transporters	139
Figure 53 AaPrtD conformation.....	142
Figure 54 Cartoon of <i>Aquifex</i> T1SS substrate constructs	164
Figure 55 Cartoon of <i>Pf</i> and <i>Dd</i> T1SS substrate constructs	165

List of Tables:

Table 1 Chapter 2 crystallographic data collection and refinement statistics.....	83
Table 2 Chapter 3 crystallographic data collection and refinement statistics.....	112
Table 3 Chapter 4 crystallographic data collection and refinement statistics.....	148

Table of contents:

Abstract	ii
List of Abbreviations	v
List of Figures	ix
List of Tables	xii
Chapter 1: Introduction	1
1.1 Sugars	1
1.1.1 Biological roles of sugars	1
1.1.2 Chemical properties of sugars	3
1.1.3 Glycosyl transfer reaction	5
1.2 Glycosyltransferase enzymes	6
1.2.1 GT folds	8
1.2.2 Processive GTs	11
1.3 Cellulose structure and properties	13
1.3.1 Cellulose synthase	15
1.4 Bacterial biofilms	17
1.4.1 Composition of biofilm matrix	19
1.4.2 Cyclic-di-GMP signaling	21
1.5 Sugar transport across the membrane	23
1.5.1 Alternating access	23
1.5.2 Peculiarities of biopolymer transport	27
1.5.3 Polysaccharide transport	28
1.5.3.1 Synthase-dependent	28
1.5.3.2 ABC-transporter mediated	29

1.5.3.3 Wzx/Wzy mechanism.....	30
1.6 Bacterial cellulose synthase.....	31
1.6.1 A historical perspective	31
1.6.2 Functional characterization of bacterial cellulose synthase	33
1.6.3 Structure of BcsA–B complex.....	41
1.7 Open questions	52
Chapter 2: Mechanism of activation of bacterial cellulose synthase by cyclic-di-GMP	54
2.1 Introduction	55
2.2 Results	59
2.2.1 Architecture of BcsA–B in complex with c-di-GMP	59
2.2.2 BcsA binds a c-di-GMP dimer on the β -barrel surface	62
2.2.3 Conformational changes of the gating loop	64
2.2.4 C-di-GMP releases an auto-inhibited state of BcsA	70
2.2.5 The TM channel entrance forms the acceptor-binding site	73
2.3 Discussion	78
2.4 Methods	84
2.4.1 Protein purification.....	84
2.4.2 Crystallization	84
2.4.3 Data collection and processing.....	85
2.4.4 Structure determination	85
2.4.5 Inverted membrane vesicle preparation	87
2.4.6 Proteoliposome preparation.....	87
2.4.7 <i>In vitro</i> cellulose synthesis assay.....	88
2.4.8 Western analysis.....	88
2.4.9 Enzyme-coupled activity assay	89
2.4.10 Isothermal titration calorimetry.....	89
Chapter 3: Observing cellulose biosynthesis and membrane translocation <i>in crystallo</i>	90
3.1 Introduction	92
3.2 Results	94

3.2.1 Cellulose synthase elongates the cellulose chain one glucose unit at a time	95
3.2.2 BcsA's finger helix resets in response to polymer extension.....	96
3.2.3 Cellulose translocation by a coordinated movement of BcsA's gating loop and finger helix.....	97
3.2.4 The movement of the finger helix is required for cellulose translocation	100
3.2.5 The product-bound state	102
3.2.6 The donor glucose binds to a hydrophilic pocket underneath the acceptor	104
3.3 Discussion	107
3.3.1 Implications for cellulose biosynthesis and membrane translocation	107
3.4 Methods	113
3.4.1 <i>In crystallo</i> cellulose synthesis.....	113
3.4.2 <i>In crystallo</i> cellulose translocation.....	113
3.4.3 Data collection.....	114
3.4.4 UDP-CH ₂ -Glc soak to generate the donor bound state	115
3.4.5 Finger helix cross-linking and activity assays.....	115
3.4.6 UDP-sugar analogue synthesis	116
Chapter 4: Structure of a type-1 secretion system ABC transporter	125
4.1 Introduction	126
4.2 Results	130
4.2.1 Substrate secretion is dependent on ATP hydrolysis	130
4.2.2 Transport across the inner membrane requires the entire T1SS complex	131
4.2.3 <i>Aquifex aeolicus</i> PrtD exhibits a typical ABC exporter fold	134
4.2.4 AaPrtD contains a narrow, occluded TM channel	138
4.2.5 Comparison to peptide transporters.....	140
4.3 Discussion	143

4.4 Methods	150
4.4.1 Plasmid construction	150
4.4.2 Whole cell secretion assays	150
4.4.3 Spheroplast preparation and secretion assays	151
4.4.4 AaPrtD purification	153
4.4.5 AaPrtD structure determination.....	154
Chapter 5: Perspectives and conclusions.....	156
Appendix 1: T1SS <i>in vitro</i> functional experiments.....	163
App1.1 Expression, purification and reconstitution of AaPrtD	166
App1.2 <i>In vitro</i> transcription and translation of ³⁵ S-Met labeled T1SS substrates	169
App1.3 AaPrtD proteoliposome (PL) protease protection translocation experiments	172
App1.4 DdPrtD (IMV) protease protection translocation experiments	186
Appendix 2: <i>Thermotoga maritima</i> growth protocol	194
References	196

Chapter 1: Introduction

1.1 Sugars

Sugars are important molecules for life. They provide the primary energy source for the vast majority of life on earth. Furthermore, sugars also act in signal transduction and regulation with phosphatidylinositol modifications on lipids and glycosylation of proteins playing important regulatory roles in eukaryotes. Finally, they play important structural roles throughout biology by acting as protective barriers in bacteria and plants cell walls as well as by forming the exoskeleton of insects.

1.1.1 Biological roles of sugars

Sugars are probably best known for their role as energy-storage molecules. Glucose, which is a product of photosynthesis in plants, serves as the initial substrate for glycolysis, the series of reactions that produces pyruvate as a final product and generates 2 ATP molecules and 2 NADH molecules (reducing equivalents) in the process. Pyruvate can then feed into the citric acid cycle, which then generates further reducing equivalents. These can then feed into the electron-transport chain and generate more ATP by oxidative phosphorylation. In the end, a single molecule of aerobically-metabolized glucose produces 30 (eukaryotes)– 32 (prokaryotes) molecules of ATP, illustrating the utility of glucose as an energy carrier (Berg et. al., 2002).

1.1.1 Biological roles of sugars

Sugars are probably best known for their role as energy-storage molecules. Glucose, which is a product of photosynthesis in plants, serves as the initial substrate for

glycolysis, the series of reactions that produces pyruvate as a final product and generates 2 ATP molecules and 2 NADH molecules (reducing equivalents) in the process. Pyruvate can then feed into the citric acid cycle, which then generates further reducing equivalents. These can then feed into the electron-transport chain and generate more ATP by oxidative phosphorylation. In the end, a single molecule of aerobically-metabolized glucose produces 30 (eukaryotes)– 32 (prokaryotes) molecules of ATP, illustrating the utility of glucose as an energy carrier (Berg et. al., 2002).

In eukaryotes, sugars also play key roles in signal transduction and quality control. Phosphoinositides are lipids that are composed of an inositol sugar headgroup attached to a diacylglycerol (DAG) backbone. The inositol headgroup has 3 free hydroxyls that can each be phosphorylated, to give 8 possible different combinations of phosphorylation patterns, each of which can serve separate regulatory functions. In particular, phosphoinositides play key roles in vesicular transport as the lipid anchor allows confinement of a particular species to specific organelles and thus provides a chemical target for specific vesicles to recognize the appropriate target (Balla, 2013). Additionally, activation of phospholipase C (PLC) by G-protein coupled receptor (GPCR) activation is a major signal-transduction pathway in eukaryotic cells. PLC cleaves PIP_2 into DAG and PIP_3 , each of which goes on to signal to specific channels and enzymes, most notably, the inositol triphosphate receptor (InsP_3R), an endoplasmic-reticulum (ER) calcium channel that opens in response to PIP_3 binding and releases ER calcium stores (Putney and Tomita, 2012). Finally, the attachment of sugars to proteins in the ER and the subsequent modification of these sugars is an important component that helps to ensure proper

folding of secreted and plasma-membrane proteins. As a nascent polypeptide emerges in the lumen of the ER, specific Asn residues in the motif N-X-S/T (X represents any amino acid except proline) are recognized by an ER-localized enzyme that transfers a large block of 14 sugar molecules (core) onto the Asn. An additional series of ER-localized enzymes then trims specific sugars from this core while other enzymes add sugars. Because the proper trimming is essential for successful transport through the Golgi, the relative rates of sugar trimming/addition serves to ensure that improperly-folded proteins don't leave the ER (Xu and Ng, 2015).

1.1.2 Chemical properties of sugars

Sugars, also called saccharides or carbohydrates (literally, hydrated carbons), consist of an organic backbone decorated with hydroxyls (Soderberg, 2016). The carbon backbone

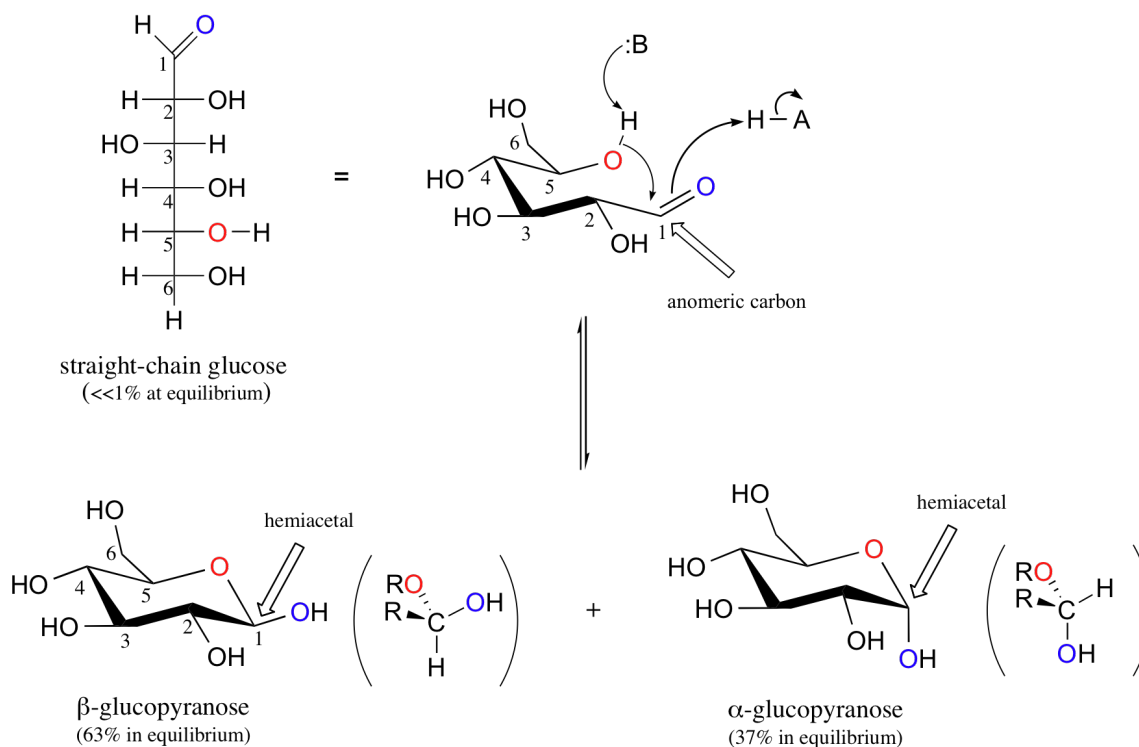


Figure 1. Cyclization of glucose. The carbonyl is shown in blue, the oxygen of the attacking hydroxyl is shown in red. The carbons are numbered, and anomeric carbon is labeled. From (Soderberg, 2016)

can vary in length from 4 to 6 carbons, and one of the carbons in the chain is a carbonyl (anomeric carbon), which can be either an aldehyde (aldose) or a ketone (ketose). Each chiral carbon atom can have either R- or S- stereochemistry, and the sequence of relative stereochemistry along the carbon chain is one of the determinants of the sugar's identity. This stereochemistry can be important for how the sugar behaves; the glucose transporters Glut 3 & 4 transport D-glucose but not L-Glucose, its enantiomer (mirror image isomer) (Hresko et al., 2016). Most aldoses and ketoses exist in an equilibrium between an open-chain form and a cyclic form (**Fig. 1**). The cyclization reaction is due to the presence of the electrophilic carbonyl group and consists of an intramolecular nucleophilic attack on the anomeric carbon by one of the hydroxyls in the sugar chain (**Fig. 1**). Due to the energetic consequences of ring strain, only two cyclization reactions are favorable, one in which the nucleophilic attack comes from a hydroxyl that is 5 carbon units away from the anomeric carbon and one in which the nucleophilic attack comes from a hydroxyl that is 4 carbon units away, giving a 6-membered ring (pyranose) and a 5-membered ring (furanose), respectively. In this reaction, the oxygen of the attacking hydroxyl is incorporated into the ring to give a hemiacetal or hemiketal depending on whether the sugar is an aldose or ketose, respectively (**Fig. 1**). Furthermore, because the nucleophilic attack on the anomeric carbon can come from two possible directions, the resulting hydroxyl after hemi-acetal formation can have two possible conformations, α or β , reflecting the stereochemistry of the anomeric carbon. While the cyclization reaction results in a flat ring for furanose, the ring for pyranose assumes a chair conformation with the majority of hydroxyls in equatorial positions, which face outward as opposed to the downward facing axial positions.

Glucose is a 6-carbon aldose and a hugely important sugar for biology due to its use as the starting substrate in glycolysis. With glucose, the α - and β - conformations of glucopyranos exist in solution at a $\sim 2:1$ equilibrium ratio of $\beta:\alpha$ because the α - conformation places the C1 hydroxyl in an unfavorable axial position while the β - conformation places it in an equatorial position (**Fig. 1**). Galactose is the C4 epimer of glucose, meaning that it has S-stereochemistry at the C4 carbon instead of R- as in glucose. Galactose cannot directly serve as a substrate for glycolysis, illustrating the importance of stereochemistry at the carbon positions in biological function. Interestingly, fructose, a ketose with the anomeric carbon at the C2 position but otherwise similar to glucose, can exist as either fructopyranose or fructofuranose depending on whether the nucleophilic attack comes from the C6 hydroxyl or C5 hydroxyl, respectively. At equilibrium, it exists as 70% β -fructopyranose, 1% α -fructopyranose, 23% β -fructofuranose, and 5% α -fructofuranose, numbers which reflect the effects of ring strain and equatorial vs. axial positions of hydroxyls (Soderberg, 2016).

1.1.3 Glycosyl transfer reaction

Monosaccharides can be linked to proteins or lipids using different substituents of the sugar. Furthermore, monosaccharides can be linked together to generate even further complexity in disaccharides, trisaccharides, up to an estimated 15,000 glucose units in cellulose (McNamara et al., 2015). Sugars are linked to other biological molecules by the glycosyl transfer reaction. Focusing here on the linking of two sugars, the reaction involves a lone pair belonging to the hydroxyl of one sugar (acceptor) performing a

nucleophilic attack on the chemically-modified carbon atom (C_{att}) of another sugar (donor). The C_{att} requires an electronegative substituent in order to exhibit the electrophilicity required for the attack and in order to provide a favorable leaving group. In biology, C_{att} of the donor sugar is substituted with a large electronegative group. For glycosyl transfer reactions occurring in the cytosol, this substituent is often a nucleotide such as uridine diphosphate (UDP) or guanosine diphosphate (GDP). In compartments such as the periplasm in bacteria, where nucleotide sugars are not available, this substituent is often a polyprenyl-phosphate lipid, which has been termed the ‘universal glycan lipid carrier’ due its use across all kingdoms of life (Manat et al., 2014). In bacteria, undecaprenyl-diphosphate (C55-PP) is used ubiquitously as a leaving group for glycosyl transfer reactions occurring in the periplasm (Manat et al., 2014). In addition, because hydroxyls are poor nucleophiles the glycosyl transfer reaction requires a catalytic base in order to increase the nucleophilicity of the acceptor sugar, thereby promoting nucleophilic attack (Lairson et al., 2008).

1.2 Glycosyltransferase enzymes

Following nucleophilic attack, the molecule assumes an ‘oxycarbenium-like intermediate’ (intermediate) with partial bonds from C_{att} to the nucleophile as well as to the leaving group (Lairson et al., 2008). This transition state, which requires both puckering of the sugar away from its favorable ‘chair conformation’ into a more strained conformation as well as assumption of a cation-like character of C_{att} (Schuman et al., 2013; Ardèvol and Rovira, 2015), is quite unfavorable and does not occur in the absence

of a catalyst. However, a broad array of glycosyltransferase (GT) enzymes have evolved to catalyze the specific linkage of a wide variety of sugars.

GTs can be divided into two groups, inverting GTs and retaining GTs, depending on what happens after the intermediate. Inverting GTs are thought to utilize a ‘direct-displacement S_N2 -like reaction’ whereby the bond between the leaving group and C_{att} of the donor sugar deteriorates and allows the formation of a glycosidic bond between the acceptor and donor sugars (**Fig. 2**). Because the nucleophile must attack C_{att} on the side opposite of the leaving group, this reaction mechanism results in the inversion of the stoichiometry at C_{att} (Lairson et al., 2008).

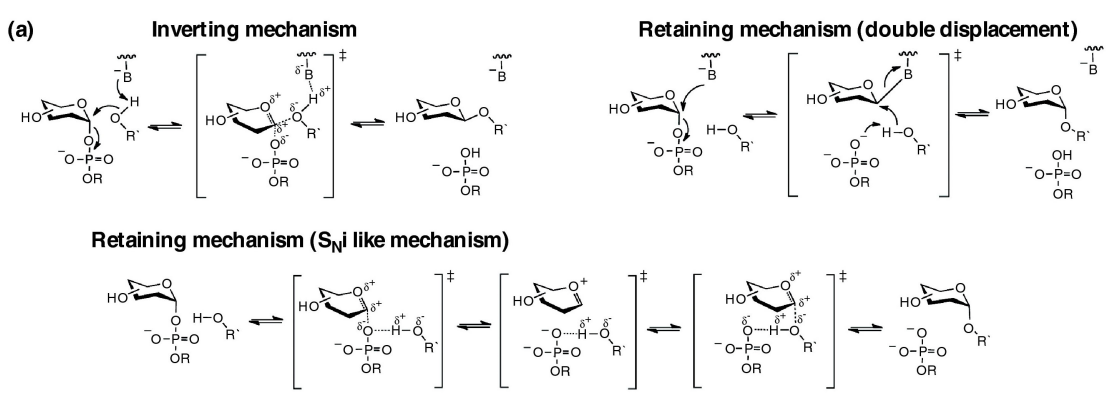


Figure 2 An overview of glycosyl transfer reactions a) Organic representations of inverting and two retaining mechanisms. ($-B$ represents catalytic base). (Chang et. al., 2011)

The mechanism of retaining GTs is more complex. Some think that the initial attack of C_{att} is performed by a side chain on the GT (Lairson et al., 2008). In a similar S_N2 -like reaction, the intermediate deteriorates and leaves a ‘glycosyl–enzyme intermediate’ in which the donor sugar forms a covalent bond to a side chain of the enzyme (**Fig. 2**). As with inverting GTs, this reaction inverts the stoichiometry of C_{att} . A hydroxyl of the

acceptor sugar then attacks C_{att} to generate another intermediate with C_{att} sharing bonds with the acceptor sugar and the enzyme but with the enzyme as the leaving group. The deterioration of this intermediate breaks the bond between the enzyme and the donor sugar which again inverts the stoichiometry at C_{att} , thereby restoring the original configuration (**Fig. 2**). The inverting mechanism is thus termed ‘double-displacement S_N2 -like reaction’. However, other models of the retaining mechanism have also been discussed (Schuman et al., 2013). In particular, the S_Ni mechanism is a variation of the S_N1 reaction with an internal return. In this mechanism, the intermediate decomposes into a pair of ions. Then, a different chemical substituent of the leaving group performs a nucleophilic attack on the same face from which the leaving group was previously attached (Lairson et al., 2008). In the case of GTs, the phosphate of the donor interacts with the proton of the acceptor hydroxyl, resulting in the decomposition into a carbocation at C_{att} and a complex between the acceptor hydroxyl and phosphate leaving group. The acceptor then attacks the carbocation donor at the same face from which the phosphate left, giving the original stereochemistry at C_{att} (Chang et al., 2011; Schuman et al., 2013).

1.2.1 GT folds

For the most part, GT enzymes appear to have three general folds, GT-A, GT-B, and GT-C. Both retaining and inverting GTs contain members with GT-A and GT-B folds. In each fold, the structure is made up of 2 adjacent Rossmann domains with each consisting of a blade of up to 7 beta strands that is surrounded by alpha helices. A GT classification

system has emerged comprising 99 families grouped by sequence similarity (Lombard et al., 2013).

In the GT-A fold, the Rossmann domains are aligned one on top of the other such that there appears to be almost a continuous blade of beta strands down the middle of the protein (**Fig. 3a**). The GT-A folds require a divalent cation for activity (Gloster, 2014). The N-terminal domain

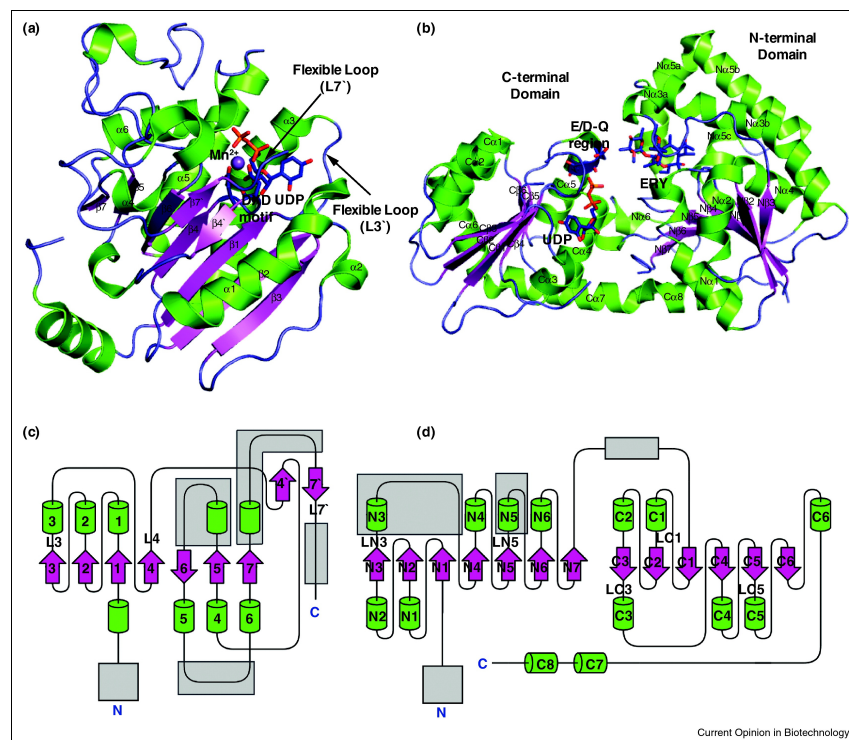


Figure 3. GT folds a) representative tertiary GT-A fold b) representative tertiary GT-B fold c) secondary structure representation of GT-A fold d) secondary structure of GT-B fold. Helices are colored green and β -strands purple (Chang et. al., 2011)

contains a motif known as DxD, which is involved in donor binding. The spacing of these Asp residues, which are either part of or nearby a beta sheet, results in their side chains facing the same direction. This orientation generates a divalent-cation binding site between the two Asp residues. In turn, the cation coordinates the alpha- and beta-phosphates of the nucleotide sugar, thereby stabilizing the donor at the acidic active site (Lairson et al., 2008). The C-terminal domain comprises the acceptor-binding site. While no sequence motifs are clear for this domain, there tends to be an alpha helix near the acceptor binding site ($\alpha 6$ in **Fig. 3a and c**) that contains an N-terminally capping Asp or

Glu residue which forms the catalytic base for the glycosyl transfer reaction (Chang et al., 2011). A prominent example of an enzyme with the GT-A fold is the mammalian xylotransferase, XXYL1, which negatively regulates activation of the Notch receptor (Yu et al., 2015).

In the GT-B fold, the Rossmann domains are aligned side by side with the beta sheets facing each other (**Fig. 3b**). This alignment creates a cleft between the two domains, and donor and acceptor bind in this cleft. GT-B folds do not contain a DxD motif at the donor-binding site, which is instead enriched in basic residues (Lairson et al., 2008). These basic residues act to neutralize the negative charge at the active site and, therefore, to bind substrate and stabilize the intermediate in a metal-independent manner (Lairson et al., 2008; Gloster, 2014). As in GT-A folds, the acceptor-binding site often contains an acidic residue or a histidine at the tip of an alpha helix (**N α 1 in Fig. 3b and d**) to act as the catalytic base; however, the acceptor-binding site in GT-B folds is located on the N-terminal instead of the C-terminal domain (Chang et al., 2011). In GTs that use sugars as the acceptor, an aromatic residue is often found at the acceptor binding site where it stacks on top of the ring and provides stability through CH- π stacking interactions (Chang et al., 2011). Glycogen synthase is an enzyme with a GT-B fold that is involved in the conversion of glucose into glycogen, a branched polymer of glucose molecules that is used for energy storage in almost every living organism except for parasites (Henrissat et al., 2002). Glycogen synthase uses ADP-glucose as a donor and catalyzes the formation of α -1,4-bonds between glucose molecules (Buschiazzo et al., 2004).

The GT-C fold has been predicted for a number of years but has only recently been confirmed structurally with the PglB oligosaccharyltransferase (OST) (Lairson et al., 2008; Lizak et al., 2011). GT-C folds are found in integral membrane proteins involved in protein glycosylation. Two structures are now known, and each contains an N-terminal TM region and a C-terminal GT domain, which contains a DxD motif (Gloster, 2014). Utilizing a lipid–phosphate–linked sugar donor, the GT–C folds appear to require divalent cations for activity (Matsumoto et al., 2013; Lizak et al., 2014).

GTs play a number of important biological roles. The canonical mechanism involves connecting the hydroxyls of two sugars through a glycosidic bond; however, the acceptor can actually comprise a large number of functional groups including nitrogen, sulfur, and carbon (Lairson et al., 2008). Furthermore, the sugar can be transferred onto everything from proteins and nucleic acids to lipids and small molecules.

1.2.2 Processive GTs

While many GTs are soluble non-processive enzymes that catalyze a single round before releasing product, there are also a number of processive GTs that undergo many rounds of substrate binding and glycosyl transfer without releasing the product, thereby producing a long polymer of sugar molecules. Some soluble processive GTs have been described such as GlfT2, an enzyme that synthesizes a galactan polymer composed of alternating β -1,5– and β -1,6–linked galactofuranose units, and which forms an important component of mycobacterial cell walls (Levengood et al., 2011). However, many processive GTs are embedded in the membrane where they couple synthesis with

secretion of the resulting polymer across the membrane. Organisms across all kingdoms of life including some viruses utilize this class of enzymes to secrete some of the most abundant organic molecules on earth (Bi et al., 2015). Some prominent examples include cellulose, a linear polymer of glucose molecules (**Fig. 4**) that forms the primary component of plant cell walls as well as an integral component of many bacterial biofilms (McNamara et al., 2015); chitin, a homopolymer of N-acetylglucosamine units (**Fig. 4**) that forms the exoskeleton of insects as well as a cell wall component in fungi (Merzendorfer and Zimoch, 2003; Lenardon et al., 2010); and hyaluronan, a linear heteropolymer made up of alternating N-acetylglucosamine and glucuronic acid (**Fig. 4**)

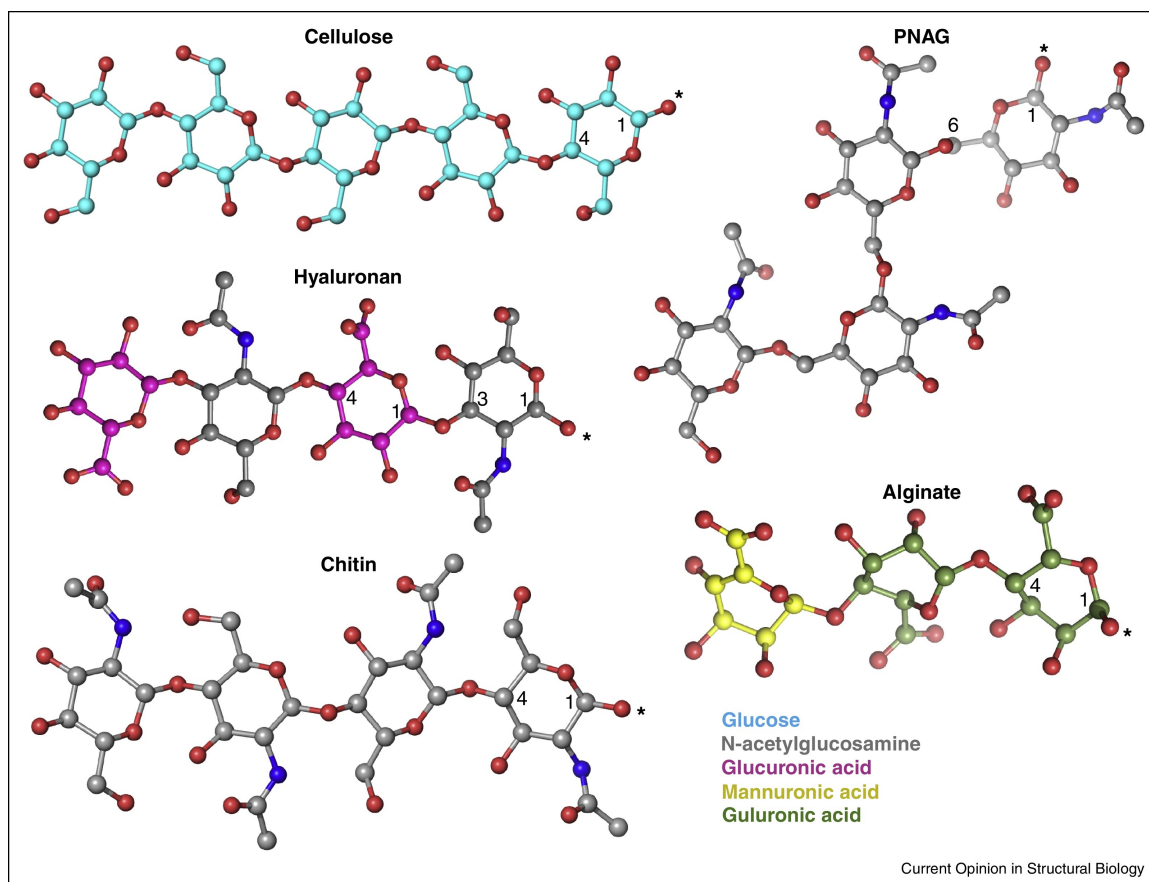
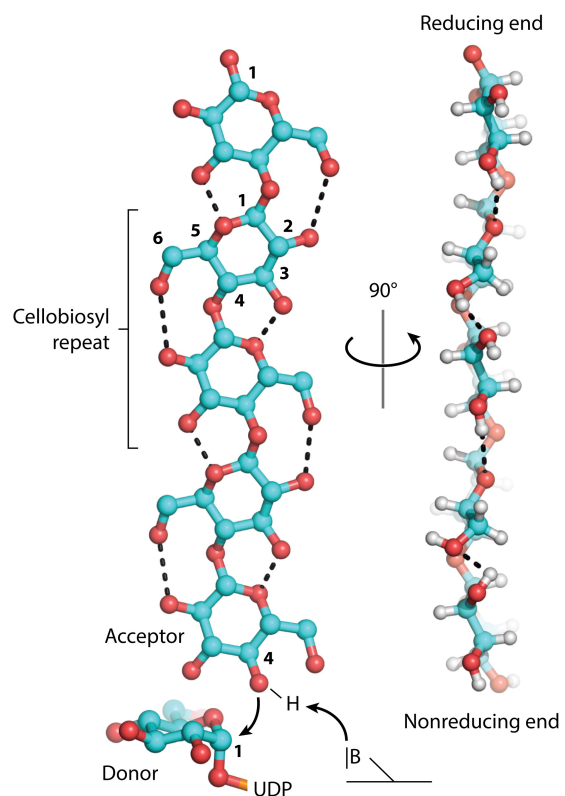


Figure 4 Chemical structures of prominent polysaccharides made by processive GTs. Sugar identities are color coded 1 (Bi et. al., 2015)

which comprises a large volume of the extracellular matrix in vertebrates while also being utilized by a number of bacteria and viruses (Hubbard et al., 2012; Bi et al., 2015). Additionally, Poly-N-acetylglucosamine (PNAG), a homopolymer of β -1,6-linked N-acetylglucosamine sugars, and alginate, a heteropolymer of mannuronic acid and glucuronic acid sugars connected by β -1,4-linkages, are both components of bacterial biofilms, which will be discussed later (Bi et al., 2015).

1.3 Cellulose Structure and Properties

Cellulose is a homopolymer of glucose molecules connected by glycosidic bonds between the C1 carbon of each glucose and the C4 carbon of its neighbor (**Fig. 4 and 5**) (Klemm et al., 2005), thereby resulting in polymers that can reach lengths of 15,000 glucose units (Somerville, 2006). In cellulose, the C1 anomeric carbon is in the beta configuration, which means that the glycosidic bond is equatorial (β -1,4 glycosidic bond) (Klemm et al., 2005). This beta C1 configuration gives a linear conformation to the cellulose polymer (**Fig. 4 and 5**) (Nishiyama et al., 2003). This conformation contrasts sharply with that of amylose (starch), an α -1,4-linked polymer of glucose, where C1 has an axial glycosidic bond, giving each glucose



McNamara JT, et al. 2015.
Annu. Rev. Biochem. 84:895–921

Figure 5 Cellulose. Stick representation of cellulose polymer and GT reaction. B represents catalytic base. (McNamara et. al., 2015)

unit an obtuse angle relative to its neighbors and resulting in an overall helical structure of the polymer (Tester et al., 2004).

The linear conformation of cellulose results in each glucose unit being rotated 180° relative to each of its neighbors. As a result, each glucose unit shares two hydrogen bonds with each of its neighbors via the C2 hydroxyl and ring oxygen forming a hydrogen bond with the C6 and C3 hydroxyls, respectively, of the glucose unit connected to its C1 carbon. In turn, the C6 hydroxyl and C3 hydroxyls form hydrogen bonds with the C3 hydroxyl and ring oxygen of the glucose unit attached at the C4 hydroxyl (**Fig. 5**) (Nishiyama et al., 2003). The 4 hydrogen bonds formed at each glucose unit make the polymer stiff and flat. Furthermore, the flat, linear conformation gives the polymer an amphipathic character with a hydrophobic surface on the face of the glucopyranose rings where the axial hydrogens point upward as well as a hydrophilic surface at the edge of the polymer where the equatorial hydroxyls point outward (**Fig. 5**). The polymer has two distinct free ends, one with an unmodified C1 hydroxyl and anomeric carbon (reducing end) and one with an unmodified C4 hydroxyl (non-reducing end) (**Fig. 5**) (McNamara et al., 2015).

The amphipathic character of cellulose is important because it causes cellulose polymers to aggregate in order to bury the hydrophobic face of the polymer and results in cellulose polymers being insoluble beyond a length of 6-8 glucose units (Gray et al., 2003; Notley et al., 2004). Although plant cell walls also contain pectins, a family of polysaccharides rich in galacturonic acid, as well as hemicelluloses, a family of heterologous and

branched polysaccharide with a backbone of β -1,4-bonded sugars, these insoluble cellulose aggregates, micro- and macrofibrils, are the fundamental load bearing components of the cell wall (Somerville, 2006; Mohnen, 2008; Scheller and Ulvskov, 2010). Some fibrils are further wrapped with lignin, an organic polymer composed of cross-linked phenylpropanoid units (Christopher et al., 2014; McFarlane et al., 2014). Cellulose fibrils and especially lignocellulose are very difficult to break down due to limited accessibility of the glycosidic bonds to enzymes or chemicals (Carroll and Somerville, 2009). Therefore, the breakdown of lignocellulose is a major barrier to the generation of biofuels and has thus become a major focus in biotechnology (Balan, 2014).

1.3.1 Cellulose Synthase

Cellulose is both synthesized and secreted by a membrane-embedded enzyme called cellulose synthase (CeS) (McNamara et al., 2015). CeSs generate cellulose from UDP-activated α -glucose, thus CeS is an inverting, processive family 2 GT that requires a divalent cation (Brown et al., 2012; Lombard et al., 2013; Omadjela et al., 2013). Plants typically encode a large number of CeSs (CesA), which are differentially expressed during the formation of primary and secondary cell walls. *Arabidopsis thaliana*, one of the main model systems for studying cell-wall biosynthesis, encodes 10 *cesA* genes with 1,2,3,5,6, and 9 being expressed during primary cell wall synthesis and *CesA*4, 7, and 8 being expressed during secondary cell wall synthesis (McFarlane et al., 2014). Freeze fracture images of CeS-containing membranes have revealed images of microfibrils emerging from large complexes (**Fig. 6c**). These cellulose synthase complexes (CSCs)

appear to have six-fold symmetry (**Fig. 6c**) (Kimura et al., 1999). Because plant cells generate microfibrils, a hypothesis emerged that these supramolecular ‘rosette’ assemblies lead to a spatial distribution of CesAs that promotes cellulose aggregation into microfibrils upon emerging from the CSC (Somerville, 2006). Much controversy has surrounded the number of CesAs in a rosette with numbers ranging from 12-36 CesAs producing microfibrils with 12-36 cellulose polymers, respectively (McFarlane et al., 2014). However, a recent paper comparing simulated with experimental wide-angle x-ray scattering and NMR data of cellulose microfibrils and another paper comparing simulated with experimental freeze-fracture EM images of rosettes have both supported a model of 6 complexes of CesA trimers in each rosette, thereby giving a total of 18 CesAs per rosette and 18 cellulose chains per microfibril (Newman et al., 2013; Nixon et al., 2016). This model is especially appealing because it suggests that CesA4, 7, and 8 form a trimer that assembles with 6 other trimers to form a rosette, thus it provides a reason as to why all three homologues are necessary for secondary cell wall synthesis.

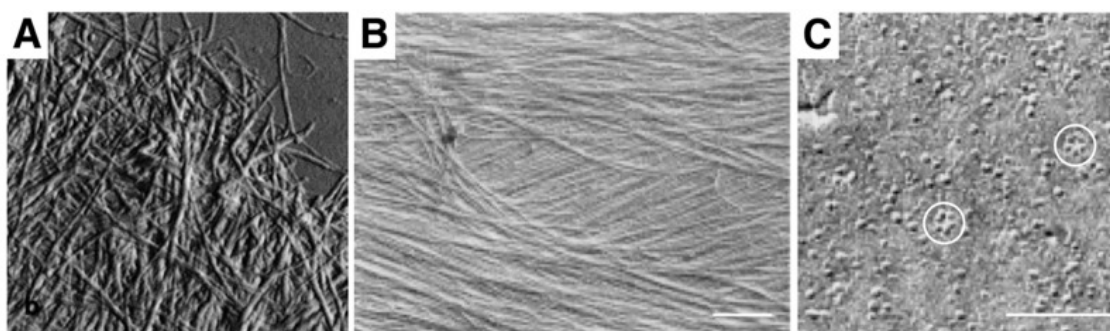


Figure 6 Cellulose microfibrils A) Atomic force microscopy image of cellulose microfibrils from *Arabidopsis*. B) Electron micrograph of cellulose microfibrils from *Arabidopsis*. C) Freeze-fracture image of CSCs in *Arabidopsis* membranes. Note the six-fold symmetry. (Li et. a., 2014)

In contrast to the cellulose microfibrils produced by plants, cellulose produced by bacteria tends to have a different structure. Aside from *Gluconacetobacter xylinus* and

Agrobacterium tumefaciens, all known bacterial cellulose is amorphous, meaning that it is not assembled into microfibrils (McNamara et al., 2015). Interestingly, CeSs from *G. xylinus* appear in freeze–fracture images to assemble into supramolecular complexes (Matthysse et al., 1981; Kimura et al., 2001). While these complexes do not have the six–fold symmetry of plant CSC rosettes and instead appear as linear complexes (linear terminal complexes), their presence nevertheless argues that microfibril formation is a result of spatial distribution of CeSs (McNamara et al., 2015). Bacterial cellulose is a key component of many bacterial biofilms, which will be discussed next.

1.4 Bacterial biofilms

Bacteria can be found in two main states in nature. One is planktonic, which is how laboratory strains of *E. coli* grow in liquid culture (Parsek and Fuqua, 2004). The other is a biofilm, which consists of a culture of bacteria that is not rapidly dividing and is embedded in a dense matrix (Römling and Balsalobre, 2012). Resembling a primitive tissue rich in intercellular signaling as well as horizontal gene transfer (Parsek and Fuqua, 2004; Flemming et al., 2007; Karatan and Watnick, 2009), biofilms are actually much more common of a state in nature than planktonic bacteria (Vu et al., 2009). Biofilm bacteria have received attention as a major public health risk as the NIH has released estimates that up to 80% of all infections are due to biofilm bacteria (Römling and Balsalobre, 2012). The role of the biofilm matrix in infectivity is due to a number of factors. First, the matrix provides resistance to both innate and adaptive immune responses of the infected (del Pozo and Patel, 2007; Römling and Balsalobre, 2012). Additionally, biofilms allow bacteria to attach to surfaces, both biotic and abiotic

(Karatan and

Watnick, 2009).

This feature is a

key contributor to

hospital-acquired

infections as it

allows

pathogenic

bacteria to stick

to medical

devices such as

catheters and stents (Römling and Balsalobre, 2012). Finally, biofilm bacteria exhibit a

greater resistance to antibiotic treatment.

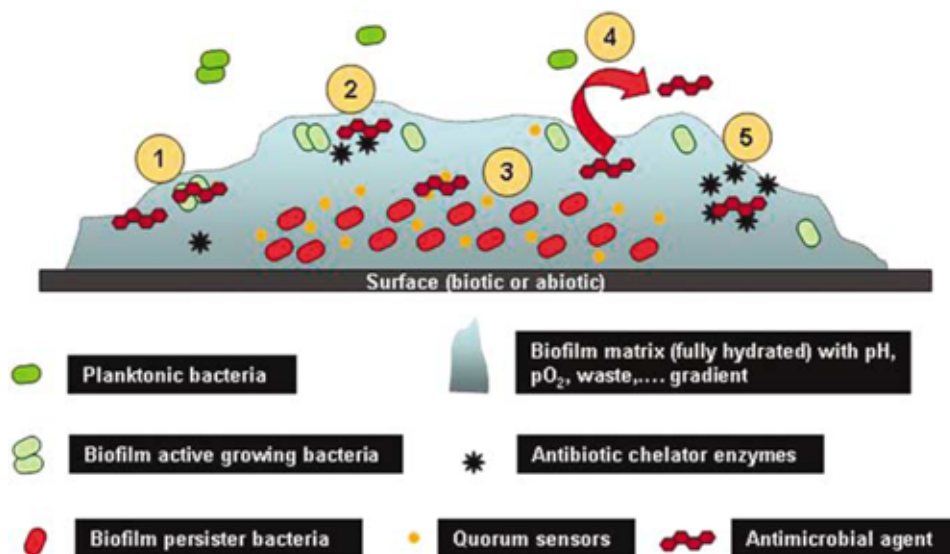


Figure 7 Biofilm life cycle. 1) Bacteria attaches to surface 2) Bacteria multiply to form microcolony 3) gradients of nutrients generate 'persister' populations buried deep in biofilm 4) some bacteria leave biofilm to become planktonic 5) biofilm promotes antibiotic resistance. (del Pozo and Patel, 2007)

The contribution of biofilms to antibiotic resistance has been a major area of research.

The early hypothesis that the biofilm protects bacteria by preventing diffusion of antibiotics into the colony has found mixed results with different antibiotics appearing to have different abilities to diffuse within the matrix (Costerton et al., 1999). So, a simple model of the biofilm matrix providing a physical barrier to antibiotic penetrance doesn't explain everything. Biofilms themselves contain gradients of nutrients including sugars and, especially, oxygen. Related to nutrient levels, released molecules, which signal to promote or prevent growth, by a bacterium differs based on its location within the biofilm (Xu et al., 1998). As a result of these two factors, the periphery of the biofilm, which has

higher nutrient levels, contains more metabolically-active and dividing cells, whereas the interior portions of the biofilm have lower levels of nutrients and, therefore, contain fewer metabolically-active cells (quiescence) (**Fig. 7**) (Parsek and Fuqua, 2004; del Pozo and Patel, 2007). Because antibiotics mainly affect metabolically-active cells, the dormant interior cells ('persisters') are largely unaffected and are able to persist in the presence of antibiotics (**Fig. 7**) (Costerton et al., 1999).

1.4.1 Composition of biofilm matrix

The biofilm matrix is composed of secreted polysaccharides (exopolysaccharides), proteins, nucleic acids, lipids, and even host cells (Sutherland, 2001; Flemming et al., 2007). Though, the exact composition is difficult to determine because of the dynamic nature of the matrix (Parsek and Fuqua, 2004). Even for a specific strain of bacteria, biofilm composition may vary based on levels of oxygen and nitrogen, pH, level of desiccation, temperature, and nutrient level (Vu et al., 2009). Additionally, properties of the surface to which the biofilm is attached, such as hydrophobicity or hydrophilicity, also affect the composition (Ahimou et al., 2007; Vu et al., 2009). Nevertheless, exopolysaccharides are key to the structural integrity of the biofilm, and bacteria that lack polysaccharide secretion machinery are unable to form biofilms (Karatan and Watnick, 2009). Furthermore, treatment of bacterial cultures with purified glycoside hydrolases, which specifically degrade exopolysaccharides, also prevents biofilm formation and can disrupt preexisting biofilms (Itoh et al., 2005; Baker et al., 2016). Alginate is the first polysaccharide that was discovered to be important for biofilms of *Pseudomonas aeruginosa*, which is a model organism for biofilm formation and an opportunistic

pathogen that infects cystic fibrosis patients (Davies, 2002; Parsek and Fuqua, 2004; Franklin et al., 2011; Schurr, 2013). Alginate has received much attention for its role in biofilm formation. However, various studies have challenged alginate's importance citing low levels of alginate in biofilms as well as reports indicating that abolishment of alginate production does not prevent biofilm formation, whereas abolishment of Psl and Pel polysaccharide production does (Friedman and Kolter, 2004; Jackson et al., 2004; Stapper et al., 2004; Schurr, 2013). A consensus now seems to have been reached that different polysaccharides are required at different stages of biofilm formation with Psl and Pel required at early stages and alginate required at later 'chronic infection' stages (Schurr, 2013). While Psl, Pel, and alginate play important roles in *P. aeruginosa* biofilms, cellulose is the primary polysaccharide in many others including the pathogens, *Salmonella typhimurium*, and *E. coli* (Flemming et al., 2007; Karatan and Watnick, 2009).

The lifecycle of a biofilm begins with a free-swimming bacterium approaching a surface that contains a 'conditioning film' of protein which likely comes from the media and is absorbed into the material (Kanematsu and Barry, 2015). The bacterium then attaches to the surface and begins to divide (Römling and Balsalobre, 2012). Through cycles of division, the bacteria forms a 'microcolony' that begins secreting polysaccharides, adhesive pilli, and other components to form the biofilm matrix (**Fig. 7**). The formation of the matrix along with further cell division results in a mature 3-dimensional or 'multilayer' biofilm (Karatan and Watnick, 2009). Finally, the biofilm can disperse and return to a planktonic state upon mechanical disruption of the biofilm or by secretion of

polysaccharide lyases and other enzymes that degrade the matrix (**Fig. 7**). The life cycle of a biofilm is controlled by the coordination of a number of signaling events (Karatan and Watnick, 2009; Römling and Balsalobre, 2012). The main regulatory switch between biofilm formation and dispersion is controlled by cyclic-di-GMP (c-di-GMP), a secondary messenger that is used almost exclusively by bacteria (Simm et al., 2004).

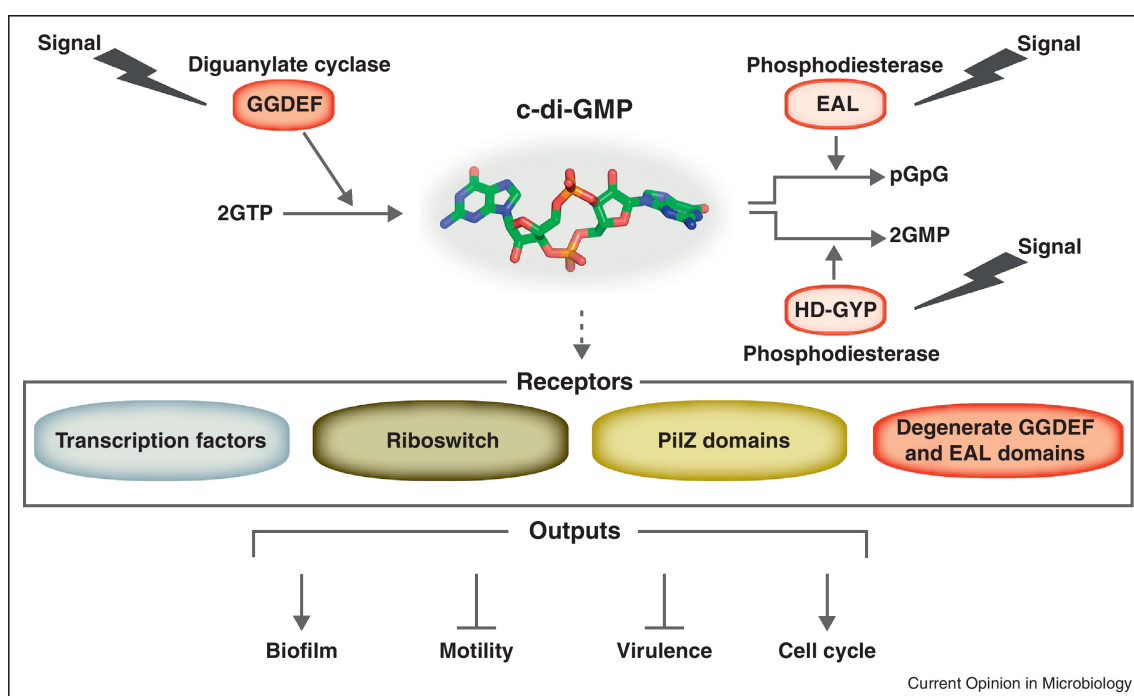


Figure 8 C-di-GMP physiology. Diguanilate cyclases and phosphodiesterases control cytoplasmic [c-di-GMP], which signals to a variety of receptors and results in a number of physiological changes. (Sonderman et. al., 2012)

1.4.2 Cyclic-di-GMP signaling

Interestingly, c-di-GMP was first discovered based on its ability to activate cellulose biosynthesis in membranes of *G. xylinus* (Ross et al., 1987). Since its discovery, c-di-GMP has been found to play a major role in bacterial physiology (thoroughly reviewed in (Römling et al., 2013)). The steady-state cytoplasmic concentration of c-di-GMP is controlled by the opposing actions of diguanilate cyclases (DGCs), which synthesize it

from GTP, and phosphodiesterases (PDEs), which break it down (**Fig. 8**) (Sondermann et al., 2012). DGCs can be identified using bioinformatics by the presence of the motif ‘GGDEF’, and PDEs can be identified by the presence of either ‘EAL’ or ‘HD-GYP’ (Römling et al., 2013). Some bacteria are predicted to encode up to 100 of these c-di-GMP–metabolizing enzymes, illustrating the complexity of c-di-GMP signaling (Römling and Balsalobre, 2012). Many of these c-di-GMP-metabolizing enzymes contain N-terminal regulatory domains in addition to the C-terminal DGC or PDE (Römling et al., 2005). These regulatory domains may respond to a variety of intracellular or environmental signals to activate or inhibit the activity of the DGC or PDE (**Fig. 8**). These signals allow c-di-GMP signaling to be regulated in a ‘chemotaxis-like’ mechanism by such environmental cues as oxygen, nitric oxide, or light (Karatan and Watnick, 2009). The resulting steady-state c-di-GMP concentration signals to a variety of downstream receptors including riboswitches, transcription factors, degenerate DGCs and PDEs, and PilZ domains (**Fig. 8**) (Sondermann et al., 2012). PilZ domains are typically small beta barrels that can be identified using bioinformatics by two motifs that are involved in binding c-di-GMP, the ‘RxxxR’ and ‘DxSxxG’ motifs (Amikam and Galperin, 2006; Benach et al., 2007; Ko et al., 2010). While named for the PilZ protein of *P. aeruginosa*, which is comprised of only the beta-barrel fold, PilZ domains are frequently modular and can be found on a number of proteins, where they presumably regulate the function in a c-di-GMP–dependent manner (Sondermann et al., 2012; Römling et al., 2013).

A key target of c-di-GMP in the formation of biofilms is processive polysaccharide synthases, some of which contain PilZ domains (Whitney and Howell, 2013). The allosteric regulation of bacterial CeS by c-di-GMP is mediated by a C-terminal PilZ domain and will be discussed extensively in Chapter 2. Additionally, Alg8, the synthetic subunit of alginate synthase, forms a complex with Alg44, which contains a PilZ domain at its C-terminus that is responsible for the c-di-GMP-dependent regulation of alginate production in *P. aeruginosa* (Whitney and Howell, 2013). However, not all of c-di-GMP's actions on polysaccharide synthases are mediated by a modular PilZ domain. In the Pel system, the regulatory subunit, PelD, contains a degenerate GGDEF domain (Franklin et al., 2011). Whereas, in the *E. coli* Pga system, which is responsible for the synthesis and secretion of PNAG, the regulatory subunit, PgaD, contains no discernable c-di-GMP binding domain (Whitney and Howell, 2013). Instead, c-di-GMP appears to bind simultaneously to PgaD and PgaC, thereby promoting their interaction and stimulating PgaC's synthetic activity (Steiner et al., 2013).

1.5 Sugar transport across the membrane

1.5.1 Alternating access

The phospholipid membrane is of fundamental importance in biology because it separates cells from the outside world by forming a barrier to ions and large hydrophilic molecules (Alberts et. al., 2008). Therefore, the ability to selectively transport specific molecules across the membrane is what allows a cell to control its internal environment in order to promote certain chemical reactions while preventing others. The mechanism by which specific molecules are selectively transported across the membrane has been a long-

standing research topic. In 1966, Jardetzky proposed a theoretical model whereby a 'polymer molecule' with a large cavity in the interior could have two conformations: one that is open to one side of the membrane and one that is open to the other. Additionally, the polymer would have a binding site for the specific transported substrate in its cavity (Jardetzky, 1966). This model became known as the 'alternating access' model of membrane transport, and the basic principle of a single binding site being alternatively exposed to either side of the membrane has been confirmed in many transporters that have since been characterized. Jardetzky went on to suggest that the transporter could concentrate substrates in the interior based on a difference in affinity of substrate for the binding site in the inward and outward concentrations so that theoretical equilibrium would not be reached when $[\text{substrate}]_{\text{in}} = [\text{substrate}]_{\text{out}}$ but when $[\text{substrate}]_{\text{in}} = K_d$ of inward conformation of transporter and $[\text{substrate}]_{\text{out}} = K_d$ of outward conformation. Charles Tanford (1982) expressed this concept more formally in terms of chemical potentials and argued that $\Delta\mu_b^0 \approx \mu_1 - \mu_2$, where $\Delta\mu_b^0$ represents a difference in binding affinity on each side of the membrane ($\mu_b^0 - \mu_b^0$) and μ_1 and μ_2 represent the chemical potential of substrate on each side of the membrane at equilibrium (Tanford, 1982). Many transporters overcome the difficulties of driving transport uphill by coupling transport of a specific substrate with another energy source. The energy for this process, termed 'active transport', mainly comes from one of two sources, the hydrolysis of ATP or the movement of ions down an electrochemical gradient. Transporters that utilize the energy

of ATP hydrolysis are called ‘primary transporters’ whereas transporters that utilize electrochemical gradients (which are often established by primary transporters) are called ‘secondary transporters’ (Alberts et. al., 2008). Primary transporters fall primarily into three classes: the P-type ATPases, which includes the sodium-potassium pump (Palmgren and Nissen, 2011); rotary ATPases such as the H^+ –pumping vacuolar ATPase (Stewart et al., 2014); and ABC transporters, which includes multidrug transporters such as P-glycoprotein (Rees et al., 2009). Secondary transporters often involve the coupling of substrate transport to sodium gradients in eukaryotes as and the coupling to proton gradients in prokaryotes and mitochondria; however, they can also involve exchange of two

small molecules such as with the ATP/ADP exchanger (Forrest et al., 2011). The coupled ion can either be transported in the same direction across the membrane (symport) or in the opposite direction (antiport) (Forrest et al., 2011). Secondary transporters often exhibit an ‘inverted topology repeat’ fold where a likely gene duplication event creates a two-fold pseudo-symmetry axis across the membrane that relates one half of the transporter to the other half (**Fig. 9**) (Forrest, 2013). The implication of this ‘inverted

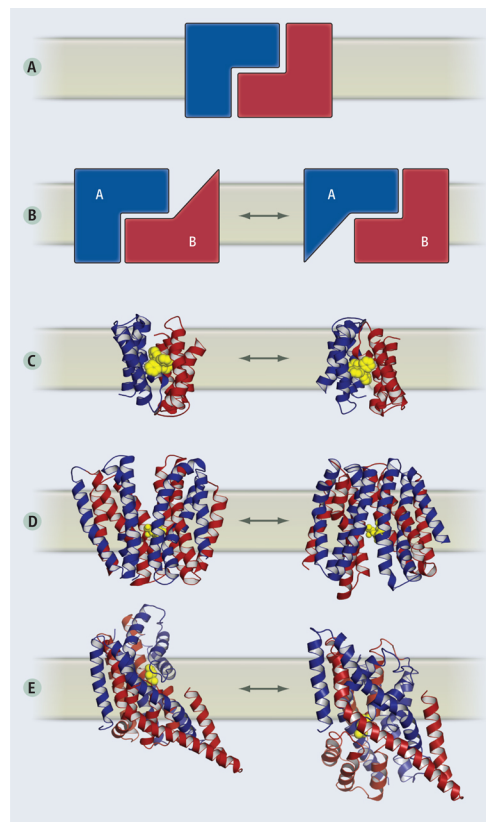


Figure 9 Symmetry in transporters. A) representation of symmetric species. B) representation of pseudosymmetric species. Exchange of subunit conformation exposes cleft to alternating sides of membrane. C), D), E) Pseudo-symmetric examples of EmrE, LacY, and Glt_{ph}, respectively. (Forrest, 2013)

topology repeat' is that the alternating-access transport cycle involves the exchange between two pseudo-symmetrically- (and likely energetically-) equivalent states (**Fig. 9**). Interestingly, this two-fold pseudo-symmetry is also seen in ABC transporters (Dawson and Locher, 2006), and some ABC transporters can actually function as proton-coupled importers if the ATP-binding domains are truncated, suggesting an important role for this 'inverted topology repeat' feature (Venter et al., 2003).

The lipid bilayer is largely impermeable to sugars. Sugars, however, are essential to cellular life, so numerous transporters belonging to many different families have evolved to transport sugars across the membrane. Ronald Kaback extensively studied the *E. coli* lactose permease (LacY). LacY is a proton-coupled importer and the prototype of the major facilitator superfamily (MFS) of transporters. It is made up of 12 TM helices in an inverted topology repeat between TM1-6 and TM7-12 (**Fig. 9d**) (Abramson et al., 2003). Another example of a sugar secondary transporter, the sodium/galactose symporter (SGLT), couples the import of sodium down its concentration gradient with the import of sugars (Wright et al., 1996). It is a paradigm of the solute sodium symporter (SSS) family of transporters and contains 14 TM helices in an inverted repeat topology with TM2-6 related to TM7-11 (Faham et al., 2008). Much like Kaback's work with LacY, Ernest Wright has been studying this family of transporters and has developed a detailed understanding of the mechanism (Wright et al., 2011). As a final example of alternating access in sugar transporters, the *E. coli* maltose transporter, (MalEFGK₂) has been studied extensively and structures are available for the complete transport cycle (Beek et al., 2014). The maltose transporter is the prototype of type-1 ABC importers with a dimer

of MalF and MalG forming the trans-membrane domain (TMD), MalK forms a homodimer that functions as the nucleotide-binding domain (NBD), and MalE is localized on the periplasmic side where it functions as the maltose-binding protein (MBP) (Chen, 2013). In the basic scheme, MalE binds maltose in the periplasm, the maltose-bound conformation can then interact with MalFGK₂. Binding of maltose-loaded MBP at MalFG followed by binding of ATP at MalK₂ opens the binding site at the interface between MalF and G. ATP hydrolysis then promotes the inward-facing conformation, where maltose can diffuse into the cytoplasm (Chen, 2013; Beek et al., 2014; Locher, 2016).

In humans, the major mediators of glucose uptake are the Glut1-4 transporters. These proteins are uniporters that belong to the sugar porter branch of the MFS and facilitate the passive transport of glucose down its concentration gradient (Thorens and Mueckler, 2010). The different isoforms exhibit different affinities for glucose and are, thus, expressed at different times, in different tissues, and in response to different blood glucose levels (Zhao and Keating, 2007). Amazingly, when the first structure of human Glut1 came out, it was surprising to see that it only differed from the *E. coli* xylose transporter (XylE) by 1.2Å RMSD for Cα atoms (Deng et al., 2014).

1.5.2 Peculiarities of biopolymer transport

While the mechanism for small sugar transport is fairly well established for the examples discussed above, alternating access is infeasible for the transport of large polysaccharides such as cellulose and alginate. Alternating access requires that transporter contain a

binding pocket that can accommodate the entire substrate, and it is not possible for a 15,000–unit cellulose polymer to bind in a discrete site. Therefore, an alternative mechanism must exist for the transport of long polysaccharides. Some insight into the general features of a biopolymer transporter can be gained from looking at polypeptide translocation systems, which must overcome the same problem and some of which are well characterized. In Gram-negative bacteria, polypeptides are transported from the cytoplasm to the periplasm either co-translationally or post-translationally (Park and Rapoport, 2012). In both cases, the polypeptide crosses the membrane through a passive channel that is made up of the SecY, –E, and –G subunits with SecY forming the primary channel (Van den Berg et al., 2004). In co-translational translocation, the ribosome docks onto the SecYEG channel during the translation process. Ratcheting of the ribosome during translation then feeds the nascent polypeptide chain through the channel and into the periplasm (Park and Rapoport, 2012; Saraogi and Shan, 2014). In post-translational translocation, the entire nascent polypeptide is synthesized in the cytosol. SecA, a motor ATPase, then binds the nascent polypeptide and docks onto the SecYEG channel (Zimmer et al., 2008; Li et al., 2016). Then, SecA pushes the nascent chain through SecYEG and into the periplasm through cycles of ATP binding and hydrolysis (Bauer et al., 2014).

1.5.3 Polysaccharide transport

1.5.3.1 Synthase-dependent

Somewhat analogous mechanisms exist for the transport of polysaccharides as for the transport of polypeptides. Much like in co-translational protein translocation, synthase-

dependent polysaccharide translocation involves the coupled synthesis and transport of a polysaccharide (Schmid et al., 2015). This system typically consists of a membrane-embedded, processive GT; an outer-membrane beta barrel; and a tetratricopeptide (TPR) repeat-containing periplasmic component, which can sometimes be found on the same polypeptide chain as the beta barrel (**Fig. 10**) (Whitney and Howell, 2013). The GT domain of the synthase is often formed from a large loop between helices of the TM domain; however, they appear to be encoded on separate polypeptide chains in the Pel system (Franklin et al., 2011). The mechanism by which synthesis from nucleotide-activated sugars is coupled to translocation has been a major question and will be addressed in detail in chapter 3. Biofilm exopolysaccharides including cellulose, alginate, and PNAG are the main products of this system (Whitney and Howell, 2013).

1.5.3.2 ABC transporter mediated

Similar to post-translational protein translocation, ABC-transporter-dependent polysaccharide translocation involves polysaccharide synthesis and transport that are carried out in different steps and by different proteins (**Fig. 10**) (Whitney and Howell, 2013; Schmid et al., 2015). Here, the entire polysaccharide is synthesized on a lipid-linked precursor by membrane associated GTs. The polysaccharide is then capped with a chemical group, which is often a methyl or a methylphosphate moiety (Cuthbertson et al., 2010). This cap targets the polysaccharide to the ABC transporter which transports it across the membrane, presumably in a processive manner by cycles of ATP binding and hydrolysis (Whitney and Howell, 2013). As with the synthase-dependent system, the ABC-transporter system encodes a periplasmic polysaccharide co-polymerase (PCP)-like

subunit and an outer-membrane polysaccharide export (OPX) subunit (**Fig. 10**) (Schmid et al., 2015). The primary polysaccharides that are transported by this mechanism are group 2 and 3 capsular polysaccharides, which are secreted by some bacteria in the formation of protective capsules, and O-antigens, which form part of the LPS of gram-negative bacteria (Cuthbertson et al., 2010).

1.5.3.3 Wzx/Wzy mechanism

While the above two systems are responsible for secreting linear polysaccharides across the inner membrane, the Wzx/Wzy-dependent mechanism is often used to secrete branched polysaccharides (Schmid et al., 2015). In this system, the repeating unit of the heterologous polysaccharide (i.e. the building block) is synthesized on a lipid-linked precursor on the cytosolic face of the inner membrane by several GTs. This building

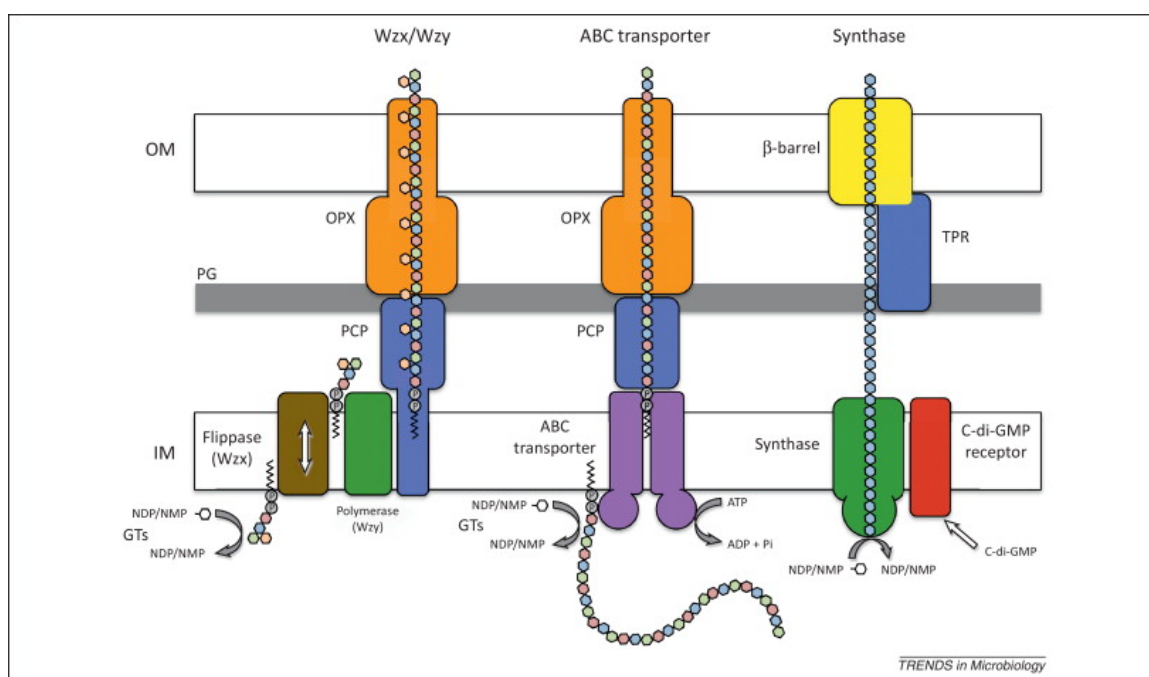


Figure 10 Polysaccharide secretion. Components of Wzx/Wzy, ABC transporter, and Synthase polysaccharide secretion systems are represented. (Whitney and Howell, 2013)

block is then flipped across the inner membrane by Wzx, a member of the multidrug and toxic compound extrusion (MATE) family of transporters (**Fig. 10**) (Islam and Lam, 2014). On the periplasmic side, the polymerase, Wzy, links these building blocks together to generate the full polysaccharide (Whitney and Howell, 2013). The Wzx/Wzy system also contains a PCP-like subunit, which is involved in regulating the length of the polymer, as well as an OPX-like subunit (**Fig. 10**) (Schmid et al., 2015). This mechanism appears to be used to secrete the majority of polysaccharides found on Gram negative bacteria with group 1 capsular polysaccharides being the best characterized examples (Islam and Lam, 2014). Additionally, *P. aeruginosa* Psl appears to be secreted by this mechanism based on homology to the Wzx/Wzy system, although the Psl secretion is largely uncharacterized (Franklin et al., 2011).

1.6 Bacterial Cellulose Synthase

1.6.1 A historical perspective

Bacterial cellulose synthase catalytic subunits (BcsA) synthesize cellulose from UDP-activated glucose (UDP-Glc) and can be identified using bioinformatics based on a motif consisting of three Asp residues variably spaced as well as a pentapeptide containing a conserved Q, R and Trp (D,D,D; QxxRW, respectively) (Saxena and Brown, 1997). These motifs are found on a large cytosolic loop which is predicted to form a GT domain between TM helices 4 and 5 of the 8 TM protein. Using both hydrophobic cluster analysis to predict secondary structure as well as sequence alignments between known processive and non-processive GTs, these motifs were predicted to form two separate domains. The first domain, which consists of the first two Asp residues of the motif was

predicted to be the substrate-binding site, while the second domain, which consists of the final Asp as well as QxxRW, was predicted to form the acceptor-binding site as well as being responsible for imparting processivity (Saxena et al., 1995; Saxena and Brown, 1997). This prediction combined with the 180° rotation between glucose units in the cellulose chain led to an alternative hypothesis that cellulose synthase actually contains two binding sites such that BcsA would bind two UDP-Glc in a 2-fold symmetric orientation while also binding the acceptor. Glycosyl transfer at both sites would then extend the polymer by a cellobiose unit with the glucose molecules aligned with a 180° rotation (Carpita, 2011).

As mentioned previously, cellulose contains a reducing end and a non-reducing end. In principle, either end could feasibly serve as the acceptor during cellulose elongation and, thus, extension could occur at either end. In order to test this possibility, Koyoma (1997) obtained cellulose from the cell wall of *Cladophora* and used silver to specifically label the reactive reducing end of the cellulose polymer. Then, they performed electron micro-diffraction on the cellulose microfibrils in order to correlate the diffraction pattern with the direction of the cellulose chain. With this information, they used a strain of *Acetobacter aceti* to produce cellulose on an electron microscopy grid. Using the position of the bacterium and the orientation of the cellulose microfibril, they determined by micro-diffraction that the non-reducing end is attached to the bacterium and BcsA must therefore extend the polymer from the non-reducing end (Koyama et al., 1997). Similar experiments using cultured cells from blackberries suggest that plant cellulose synthase also extends the polymer from the non-reducing end (Lai-Kee-Him et al., 2002).

1.6.2 Functional Characterization of Bacterial cellulose synthase (summary of

(Omadjela et al., 2013))

The purification of catalytically-active bacterial cellulose synthase subunits represented a major breakthrough in the understanding of how this enzyme functions. If the BcsA gene from *Rhodobacter sphaeroides*, a purple photosynthetic bacterium, containing a 12-histidine tag at its C-terminus is co-expressed in *E. coli* with BcsB, a large periplasmic protein with an inner-membrane anchor, the proteins form a complex that can be purified in a number of detergents. This purified complex is mixed with c-di-GMP, UDP-Glc, MgCl₂, and a small amount of UDP-³H-Glc (reaction mixture) either in Lauryl-dimethylamine oxide (LDAO), a harsh detergent that is useful for structural biology; LysoFos CholineEther 14 (LFC-E-14), a mild phospholipid derivative; or after

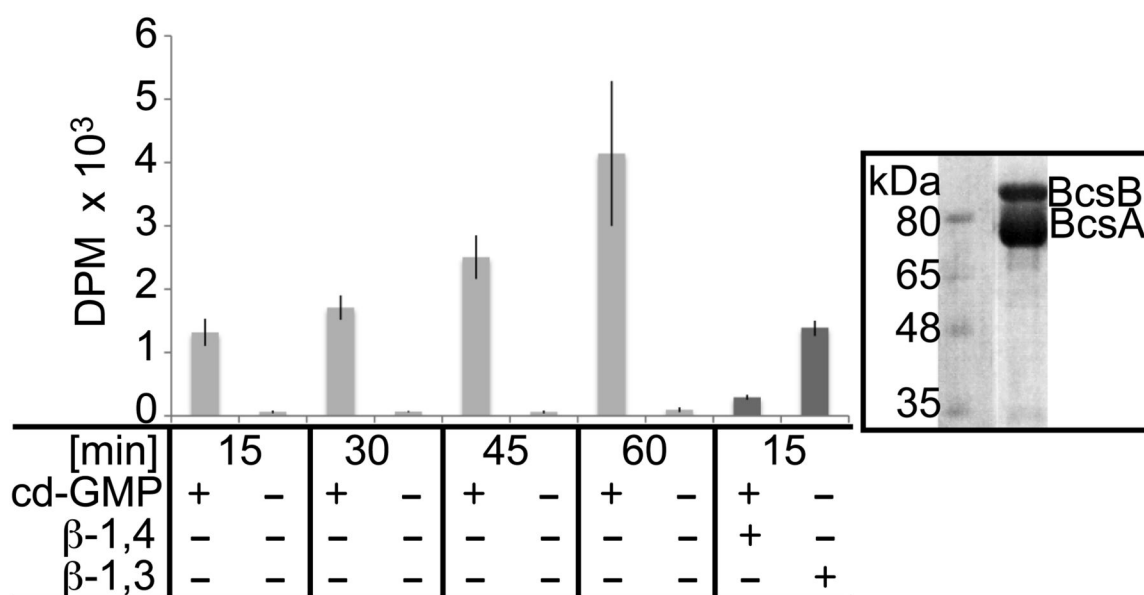


Figure 11 BcsA–B activity. DPM readings (activity) for the paper chromatography origin in the indicated conditions. Inset shows Coomassie-stained gel of purified BcsA–B. (Omadjela et. al., 2013)

reconstitution into proteoliposomes (PLs) made from *E. coli* total lipids (**Fig. 11 and 15**).

The mixture can be incubated at 37°C for a set time and then spotted onto paper and developed by paper chromatography in a buffer containing 60% ethanol. Because the cellulose product is insoluble, it should remain at the origin while the unreacted UDP-³H-Glc should move with the buffer front. Indeed, if the paper is cut into sections following chromatography and each section is quantified by scintillation counting, the buffer front gives a very high signal, the origin gives a moderate signal, and the sections in between give a very low signal. In order to confirm that the signal at the origin is indeed due to cellulose that has incorporated ³H-Glc, the sample is digested with a β-1,4 glucanase, which should only cleave β-1,4 glycosidic bonds, before spotting onto paper. Following treatment, the signal at the origin disappears while treatment with a β-1,3 glucanase does not alter the signal (**Fig. 11**).

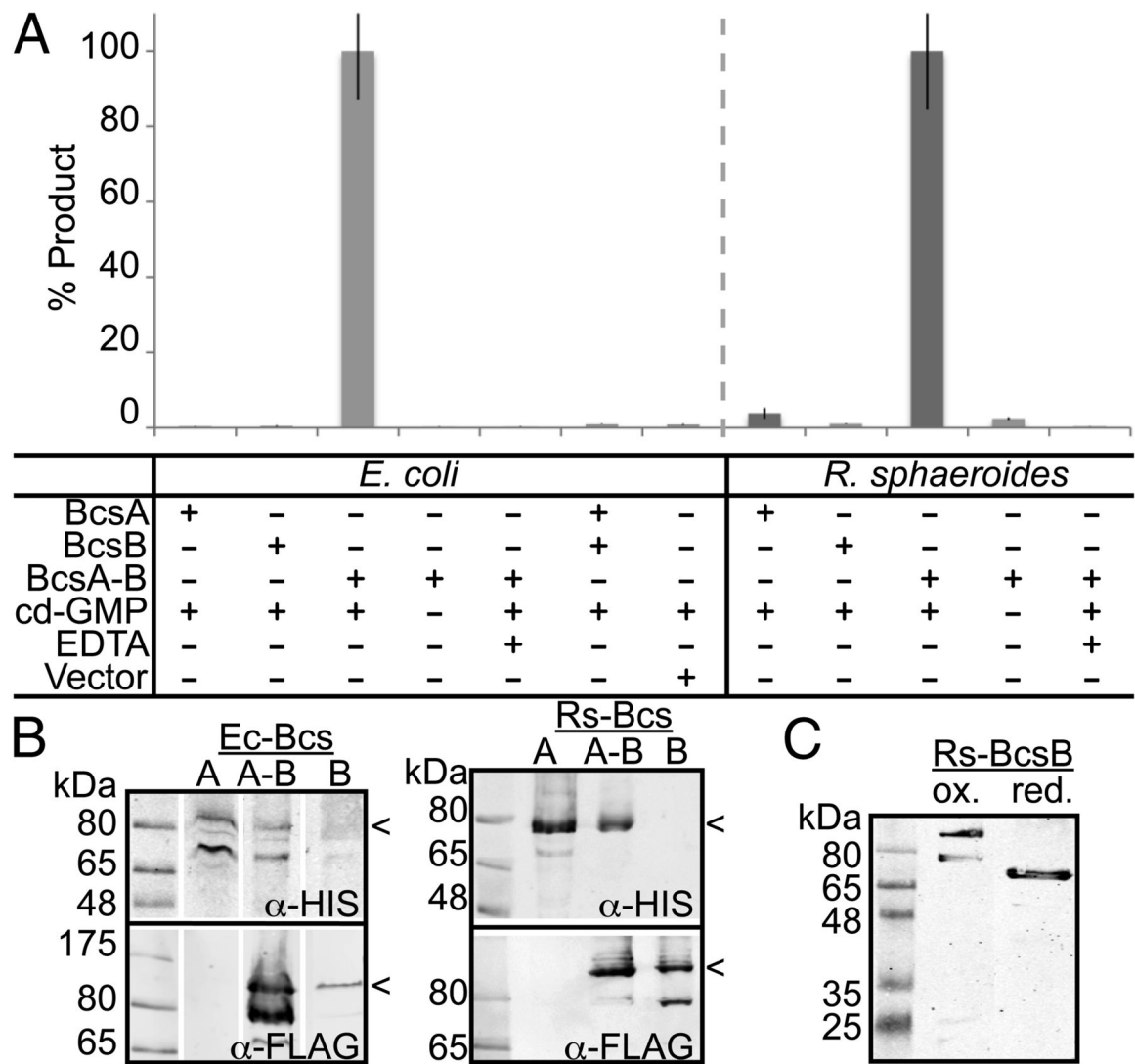


Figure 12 Activity requires both BcsA and BcsB A) % of DPM reading (relative to reading with BcsA–B and c-di-GMP) for the indicated conditions. B) Western blots against His tag on BcsA and FLAG tag on BcsB confirm expression. C) comparison of electrophoretic mobility of BcsB under oxidized and reduced conditions indicates that disulfide bond is formed. (Omadjela et. al., 2013)

Activity strictly depends on the presence of BcsB and c-di-GMP (**Fig. 12**). If membrane vesicles (IMVs) from *E. coli* expressing BcsA and BcsB are incubated with reaction mixture, robust production of cellulose is observed (**Fig. 12A**). However, if IMVs from cells expressing only BcsA or only BcsB are used, no cellulose is produced. This BcsB-dependent activity is observed for both *E. coli* BcsA–B as well as *R. sphaeroides* BcsA–

B, suggesting that it is a general phenomenon (**Fig. 12A**). Furthermore, if c-di-GMP is left out of the reaction mixture, no activity is observed, which reproduces what was seen in *G. xylinus* membranes (**Fig. 12A**) (Ross et al., 1987).

Purified BcsA–B reconstituted into proteoliposomes also displays activity that is strictly dependent on c-di-GMP (**Fig. 11**). This result is important because it demonstrates that c-di-GMP is directly activating BcsA–B as opposed to acting through some endogenous protein, a possibility that could not have been ruled out without purified protein. In order to further characterize the effect of c-di-GMP on BcsA–B, Omadjela utilized another assay to measure BcsA–B activity that is more amenable to kinetic studies. Glycosyl transfer consists of the transfer of a glucose molecule from UDP-Glc to the growing end of the cellulose polymer. This reaction produces a cellulose polymer that is elongated by a single glucose unit as well as UDP as products. Incidentally, UDP is recognized by the glycolysis enzyme pyruvate kinase, which generates ATP from ADP and phosphoenolpyruvate (Berg et. al., 2002). In this reaction scheme, pyruvate kinase transfers a phosphate from phosphoenol pyruvate (PEP) to UDP, thereby generating UTP and pyruvate. The resulting pyruvate molecule is then reduced to lactic acid by lactate dehydrogenase (LDH) using NADH as an electron donor, which results in the formation of lactate and NAD^+ as products. Thus, a single UDP molecule produced by BcsA–B results in a single molecule of NADH oxidized. Because NADH absorbs light at 340 nm, and NAD^+ does not, the formation of UDP can be monitored based on the consumption of NADH. In order to allow substrate to access the entire BcsA–B population, these kinetic experiments are performed on BcsA–B reconstituted into *E. coli* total lipid

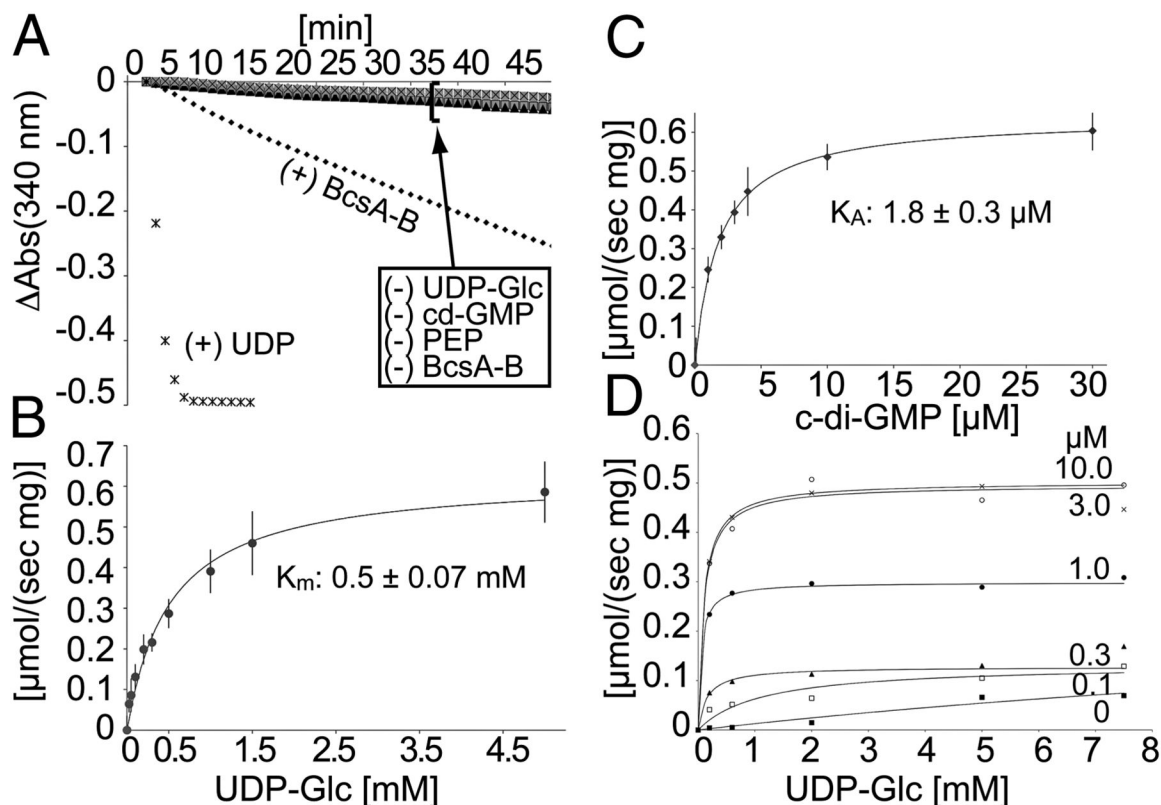


Figure 13 BcsA–B kinetics. A) Kinetic assay monitoring UDP production under indicated conditions B) Plot of activity at different UDP-Glc concentrations C) Plot of activity at different c-di-GMP concentrations D) Plot of UDP-Glc titrations at different c-di-GMP concentrations (Omadjela et. al., 2013)

nanodiscs, which are small patches of lipid bilayer wrapped with a scaffold protein (MSP) to protect the otherwise exposed aliphatic edge. Simple controls confirm that the decrease in Abs_{340} is due to glycosyl transfer by showing that no change in Abs_{340} occurs when BcsA–B, UDP-Glc, PEP, or c-di-GMP are omitted from the reaction mixture, yet Abs_{340} drops dramatically when UDP is added directly to the mixture (**Fig. 13A**). With addition of all the components, Abs_{340} decreases linearly out to 45 minutes (**Fig. 13A**). If the experiment is performed with different c-di-GMP concentrations, the slope of ΔAbs_{340} over time can be plotted against c-di-GMP concentration (**Fig. 13C**). The data can be fit to a saturating, monophasic curve that gives an activation constant (K_{act}) of $\sim 1.8 \mu\text{M}$ (**Fig. 13C**). Holding the c-di-GMP concentration at 30 μM (saturating) and

titrating UDP-Glc gives a curve that can be fit using the Michaelis-Menten equation to give a K_M of ~ 0.5 mM and a maximum turnover rate of ~ 90 molecules of UDP per BcsA–B complex per second (assuming all of the protein is active), which is ~ 10 -fold higher than has been measured in other systems (**Fig. 13B**) (Cifuentes et al., 2010). However, these numbers are difficult to compare because the portion of active enzyme is not known in either system. In order to gain insight into the mechanism by which c-di-GMP activates BcsA, Omadjela titrated UDP-Glc from 0-7.5 mM at concentrations of c-di-GMP from 0-10 μ M (**Fig. 13D**). By using c-di-GMP concentrations around the K_{act} of c-di-GMP, she was able to see the effects of c-di-GMP limitation on the steady state kinetics of BcsA–B. Each UDP-Glc titration curve was fit using the Michaelis-Menten equation revealing that increasing c-di-GMP concentrations increase the V_{max} (**Fig. 13D**). Interestingly, increasing c-di-GMP concentrations had no effect on the K_M for UDP-Glc. Because $K_M = (k_{rev} + k_{cat})/k_{for}$ (where k_{for} and k_{rev} are the on and off rates for substrate in the absence of catalysis, respectively), it seems unlikely that c-di-GMP is altering k_{cat} because either k_{rev} would need to change by equal magnitude in the opposite direction or k_{for} would need to change by equal magnitude in the same direction for K_M to remain constant. Instead, a more plausible model would be that the active site is maintained in a catalytically competent configuration in the absence of c-di-GMP, but UDP-Glc is unable to access it due to occlusion of some sort. C-di-GMP binding would remove this occlusion and allow UDP-Glc to access the active site. This ‘gating-like’ mechanism would then suggest that c-di-GMP titration is essentially increasing the concentration of active enzyme, which fits well with the increased V_{max} and unchanged K_M for UDP-Glc at increasing c-di-GMP concentrations.

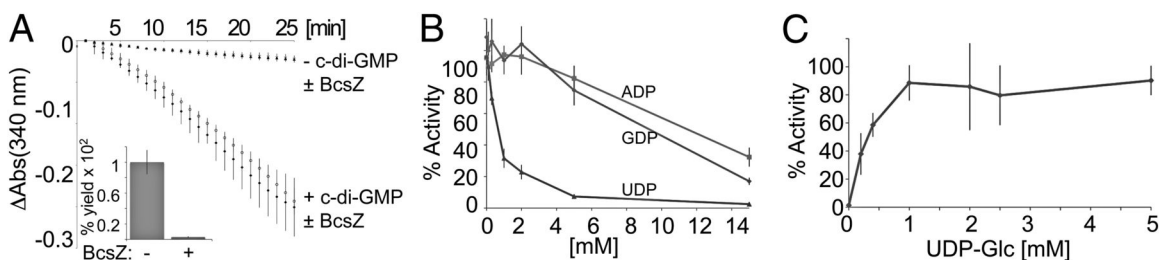


Figure 14 Product inhibition. A) Kinetic assay in presence and absence of BcsZ B) Activity in presence of increasing amounts of nucleotide diphosphates C) rescue of UDP inhibition by adding additional UDP-Glc. (Omadjela et. al., 2013)

As mentioned above, a single round of glycosyl transfer by BcsA results in two products, a cellulose polymer extended by one glucose unit and a molecule of UDP. The next experiments address the question of whether BcsA–B is subject to feedback inhibition by either of the products. To test the effect of increasing polymer length on cellulose biosynthesis, kinetic assays were performed in the presence and absence of BcsZ, a cellulase enzyme. BcsZ will degrade cellulose as it emerges from the BcsA–B complex and thus prevent the accumulation of long cellulose polymers (**Fig. 14A, inset**). Addition of BcsZ to the reaction mixture did not change the kinetics (**Fig. 14A**). However, it is possible that this result is due to the BcsA–B complexes being in solution and not anchored, there may be a greater effect from cellulose aggregation on the activity of membrane-embedded BcsA–B complexes. To test the effect of UDP on BcsA–B activity, it is necessary to use the radioactive assay because UDP is the molecule that is being quantified in the kinetic assay. Because the cellulose biosynthesis is still increasing linearly up to 60 min (**Fig. 11**), the reading at this time point can serve as a kinetic measurement. If increasing concentrations of UDP are added to the reaction mixture, the amount of cellulose produced decreases in a concentration-dependent manner with an IC_{50} of ~0.7 mM UDP (**Fig. 14B**). To test whether this feedback inhibition is competitive in nature, the UDP concentration was held at 0.7 mM, and UDP-Glc was titrated up to 5

mM. Upon increasing UDP-Glc concentration, the activity increased until full activity was reached at about 1 mM UDP-Glc (**Fig. 14C**). This demonstrates that increasing UDP-Glc can outcompete UDP and that UDP inhibits BcsA–B by a competitive mechanism. The purines, GDP and ADP, also showed weak inhibition at higher concentrations.

Next, Omadjela tested the specificity of BcsA–B for UDP-Glc. In the kinetic assay, UDP-glucuronic acid (-GA), -N-acetyl glucosamine (-NAG), -arabinose (-ara), -galactose (-Gal), and -xylose (-Xyl) (see **Fig. 4 for structures**) showed minimal activity (**Fig. 15A**).

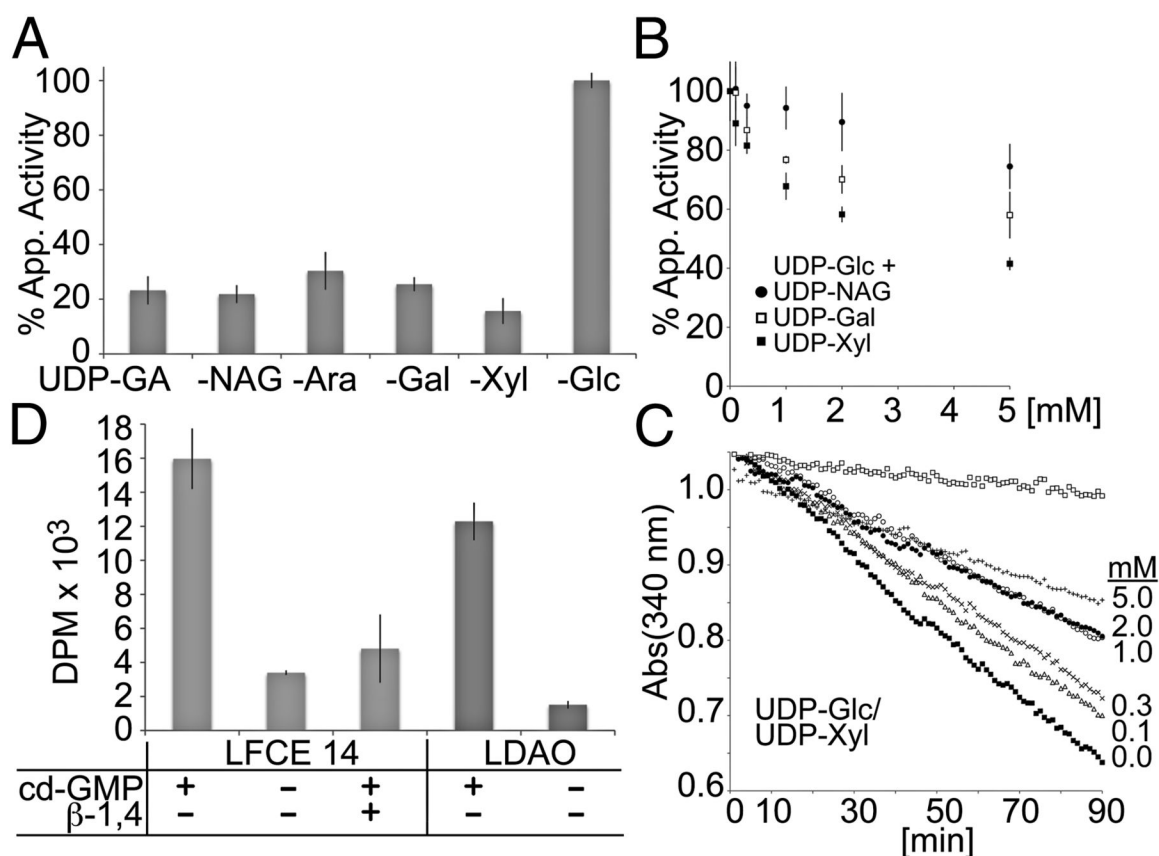


Figure 15 BcsA Substrate specificity. A) % activity using different UDP-sugars as substrate B) % activity using UDP-Glc as substrate in presence of other UDP sugars C) Activity upon increasing concentrations of UDP-xylose D) activity in detergents (Omadjela et. al., 2013)

In a more sensitive approach, these compounds were tested for an ability to inhibit BcsA–B activity in the kinetic assay (**Fig. 15B**). While each of these compounds was much less potent than UDP alone, UDP-NAG, -Gal, and -Xyl all inhibited activity at high concentrations with UDP-Xyl being the best inhibitor (**Fig. 15B and C**).

1.6.3 Structure of BcsA–B complex (a summary of (Morgan et al., 2013))

In order to gain structural insights into the mechanism of bacterial cellulose biosynthesis, we purified BcsA–B in LDAO, a detergent in which the protein is active (**Fig. 15D**). Initial crystallization trials in the presence of EDTA and UDP-Glc produced crystals at 17°C in conditions where low molecular weight polyethylene glycol (PEG) served as the precipitant. The crystals were optimized slightly by altering the buffer and pH to give large rectangular crystals in 0.1M MES pH 6.5, 50 mM NaCl, and 30% PEG200 at 4°C. However, the crystals were not very reproducible, so we relied heavily upon seeding, which involves crushing some formed crystals to generate microcrystals and then adding a small amount of microcrystals to fresh crystal trays in order to promote nucleation of crystals. The best crystals diffracted to 3.25 Å. For phasing, crystals were soaked with a wide variety of heavy-atom compounds. The strongest anomalous signal for phasing came from crystals that were soaked with 20 mM SmCl₃ for 2 hrs. Initial phases were obtained using Single-wavelength anomalous diffraction (SAD), which revealed all of the TM helices of BcsA–B. Following density modification using a custom solvent mask, anomalous difference maps were calculated for datasets from crystals soaked with 1 mM sodium ethylmercuricthiosalicylate (EMTS) for 24 hr as well as for crystals of

selenomethionine-derivatized protein. Experimental phases calculated using the Sm, EMTS, and Se positions were then used for model building.

BcsA and BcsB form a 1:1 complex that buries 4,500 Å² of surface at the protomer

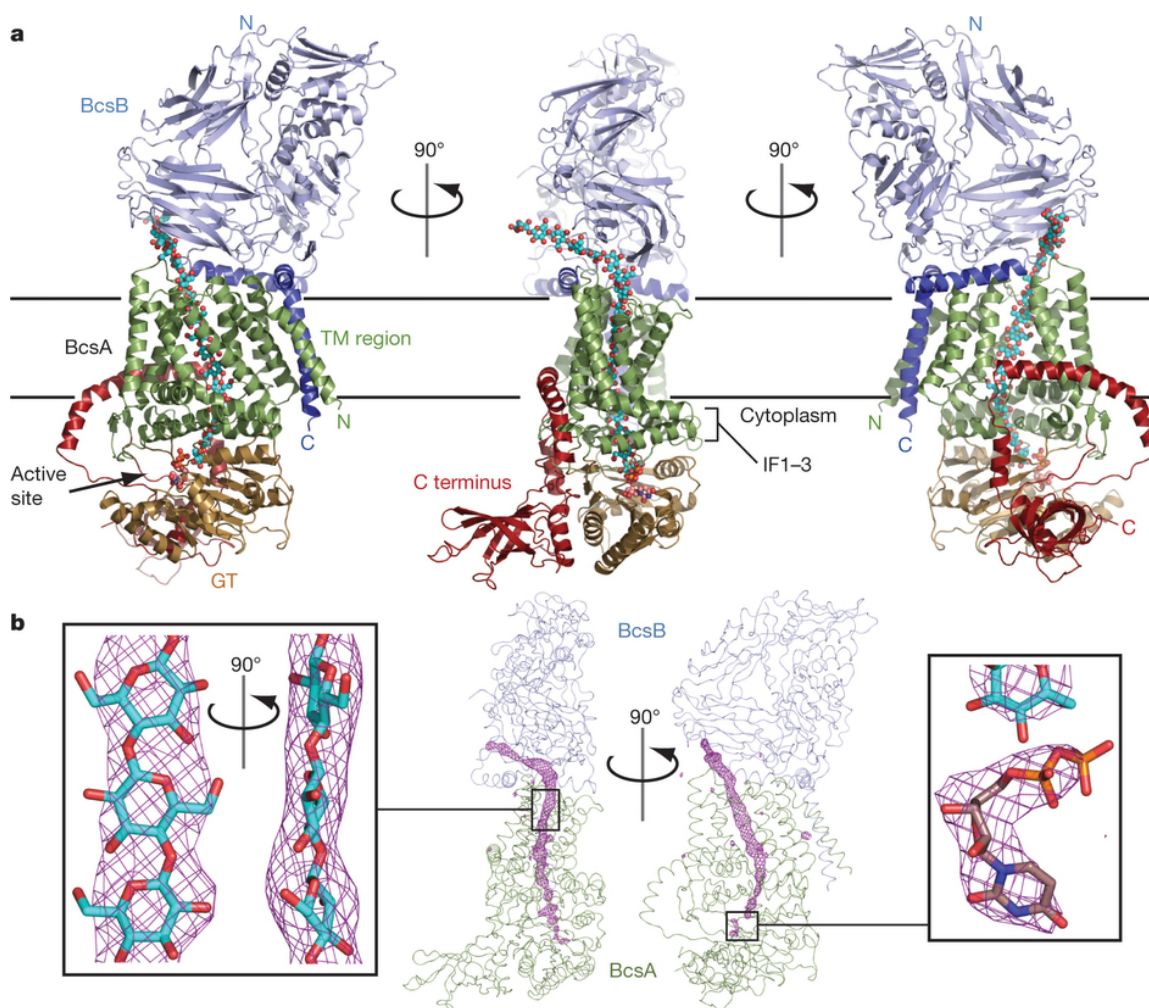


Figure 16 Architecture of BcsA–B complex. A) BcsB is blue. BcsA’s GT, PilZ, TM are shown in gold, green, red, respectively. Plane of membrane and IF helices are indicated B) unbiased difference density shown before modeling cellulose polymer. (Morgan et. al., 2013)

interface. BcsA contains 8 TM helices (**Fig. 16a**). The GT domain is formed by the loop between TM 4 and 5 (**Fig. 16A**) where it forms a tight interaction with the TM domain that is mediated by 3 amphipathic ‘interface helices’ (**Fig. 16A**). The PilZ domain at the C-terminus of BcsA sits to the side of the GT domain such that the beta barrel lies

parallel to the membrane and perpendicular to the opening of the GT domain. A long helical region following the PilZ domain extends upwards towards the membrane interface where it forms an amphipathic helix before looping back down and contacting the PilZ again (**Fig. 16A**). BcsB forms a large, beta-rich periplasmic domain with a C-terminal TM anchor that packs tightly into BcsA's TM domain. BcsB contains two amphipathic helices, both of which pack against BcsA and one of which leads directly into the TM anchor (**Fig. 16A**).

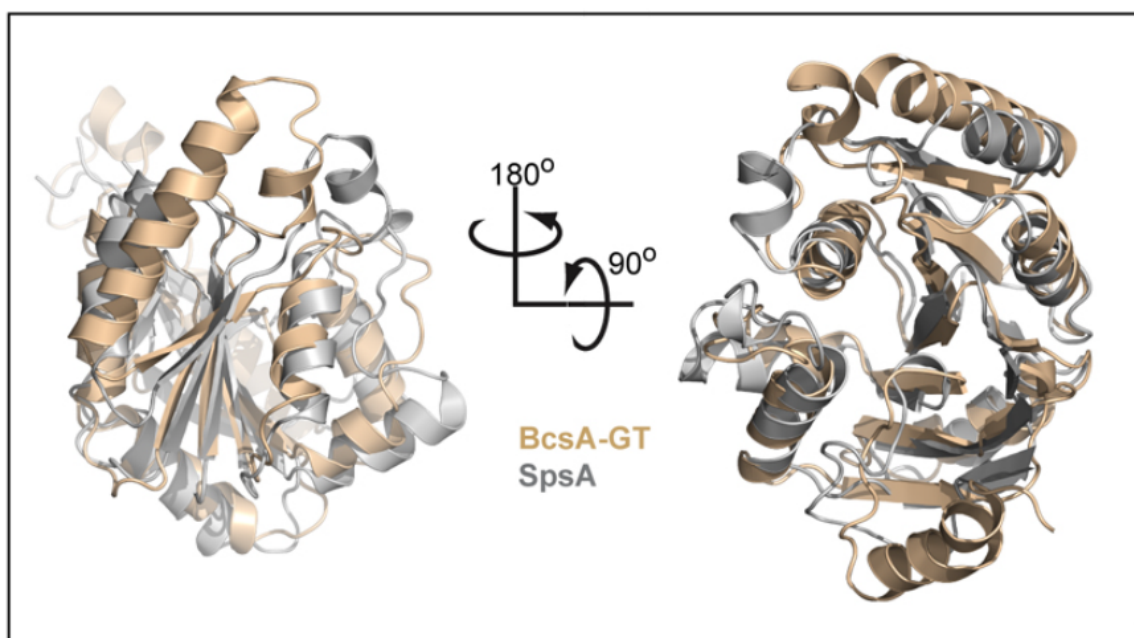


Figure 17 BcsA exhibits a GT-A fold. Superposition of the GT domain of BcsA (gold) with SpsA (gray), a canonical GT-A fold. (Morgan et. al., 2013)

Alarming, the BcsA–B complex contains a long string of density that was apparent throughout model building and which does not belong to either BcsA or BcsB (**Fig. 16B**). As the phases improved, it became clear that this ribbon of additional density is quite flat. Additionally, the density appeared to originate at the active site of BcsA, extending through a pore formed in the TM domain, and exiting on the periplasmic side of the

protein. After deliberation, we concluded that the density must be cellulose. We were able to model an 18-glucose unit cellulose polymer into the density as well as a UDP

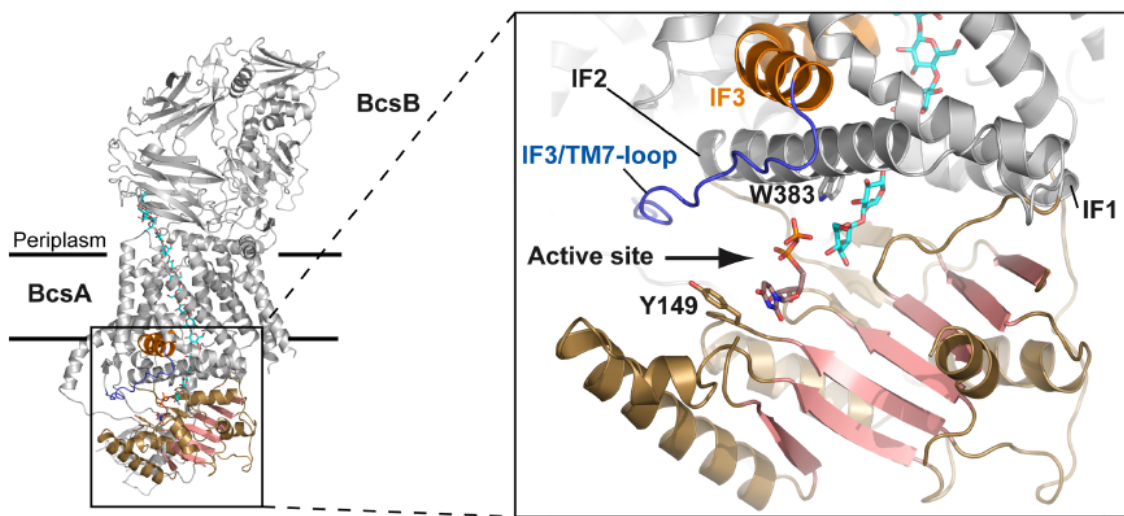


Figure 18 BcsA's GT domain. Beta sheets are colored salmon, alpha helices of GT are colored gold, and IF helices are indicated. UDP and cellulose polymer are shown as sticks. IF3-TM7 loop is blue (Morgan et. al., 2013)

molecule at the active site (**Fig. 16B**). Thus, the structure of the BcsA–B complex is captured in an intermediate state of translocating cellulose across the membrane. The active site of BcsA is located at the GT domain, which forms a GT-A fold that can be aligned with SpsA, a canonical GT-A, with an RMSD of ~ 2.2 Å between C α atoms (**Fig. 17**). The GT domain consists of a 7 beta-sheet blade that is wrapped in alpha helices and extends from the TM domain at a $\sim 45^\circ$ angle, forming a large opening where UDP binds (**Fig. 18**). Because UDP is a competitive inhibitor, it is likely that UDP-Glc binds in a similar way as UDP. Indeed, the UDP binding site is formed from highly conserved motifs (**Fig. 19**). Most notably, D180 and D246 form the first two Asp residues of the D, D, D, QxxRW signature, and these two residues are indeed involved in donor binding as had been predicted (**Fig. 19A, 25, and 36**) (Saxena and Brown, 1997). Additionally,

Y149 and E151 contribute to substrate binding by stacking on the uracil ring and by accepting a hydrogen bond from ribose, respectively. The cellulose polymer's non-reducing end (C4 hydroxyl), which acts as the acceptor in glycosyl transfer, sits in close proximity to UDP in the active site (**Fig. 19A**). Here, W383 stacks on the face of the glucopyranose ring of the penultimate glucose unit. Additionally, D343 lies near hydrogen-bonding distance from the C4 hydroxyl of the terminal glucose unit. D343 and W383 are the third 'D' and the 'W', respectively, of the D,D,D, QxxRW signature, and they are involved in acceptor binding as predicted (**Fig. 19A**) (Saxena and Brown, 1997). Additionally, the structure of the active site strongly argues against the idea that there would be two active sites as there is insufficient space for another donor molecule to bind, yet there is sufficient space for a newly-added sugar to rotate around the glycosidic bond in order to assume the alternating orientation of glucose molecules.

As mentioned above, the orientation of the GT domain relative to the TM domain generates a large opening to the active site. In the BcsA–B structure, this opening is blocked by a loop that connects the third interface helix (IF3) to the 7th TM helix (TM7) (Fig. 18 and 19B). This IF3-TM7 loop contains a number of conserved residues, which suggests that it plays an important role in cellulose biosynthesis; however, the electron density for the loop is weak and shows only the backbone, which indicates that it is not rigid in our structure. Nevertheless, the loop appears to occlude the active site in this

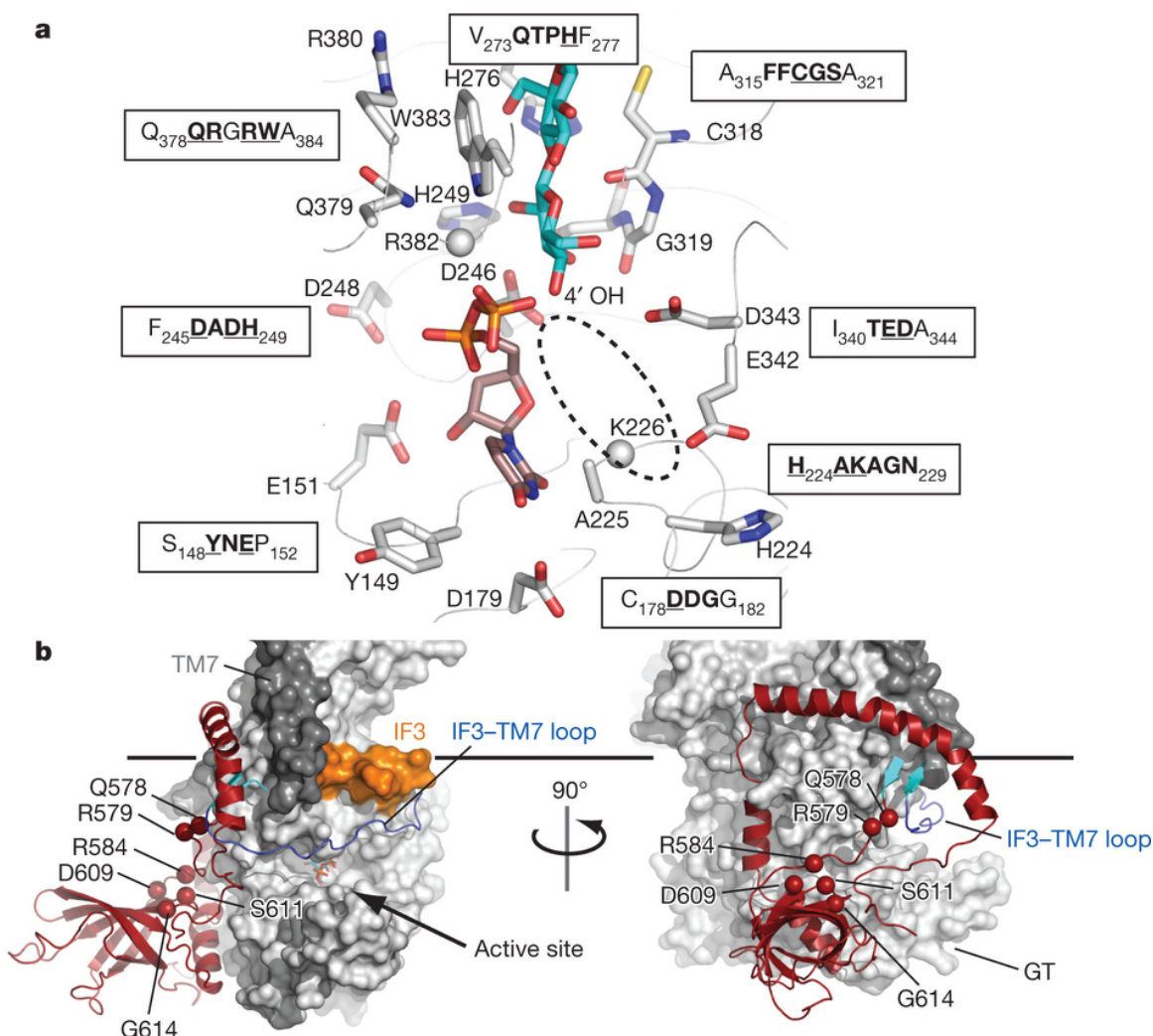


Figure 19. BcsA active site and PilZ domain. A) Indicated cellulose synthase motifs at active site. B) PilZ domain is shown in dark red. Ca atoms of residues implicated in c-di-GMP binding are shown as spheres. IF3-TM7 loop is blue. (Morgan et. al., 2013)

state (**Fig. 19B**). Following TM8 is the PilZ domain which contains the R580 and R584 of the RxxxR c-di-GMP-binding motif on the loop leading from TM8 to the beta barrel. On the surface of the beta barrel is the D609, S611, and G614 of the DxSxxG motif, which is also implicated in c-di-GMP binding (**Fig. 19B**). This cluster of c-di-GMP-binding residues lies in close proximity to the C-terminal end of the IF3–TM7 loop, which actually forms a short beta sheet with the loop directly following TM8 (**Fig. 19B**). We speculated that perhaps c-di-GMP binding could activate cellulose biosynthesis by coupling conformational changes in the RxxxR region of the PilZ domain to conformational changes in the IF3-TM7 loop through the short beta sheet formed between the two. In this way, the binding of c-di-GMP could promote the opening of the active site, which would fit with the functional data suggesting that c-di-GMP gates the active site of BcsA. The mechanism by which c-di-GMP activates BcsA will be addressed in detail in chapter 2.

The TM region of BcsA–B is composed of 8 TM helices of BcsA, 1 TM helix from BcsB, 3 cytoplasmic amphipathic interface helices (IF1-3) from BcsA, and 2 periplasmic amphipathic interface helices (IF4 and 5) from BcsB (**Fig. 20A**). Overall, the TM region forms a rectangular shape that is 2 helices wide and 4 helices long with TM1, 3, 4, and 5 aligned pairwise with TM2, 6, 8, and 7, respectively (**Fig. 20A**). The TM anchor of BcsB sits tightly in a cleft formed by TM1-3 of BcsA, while the periplasmic amphipathic helices form a $\sim 90^\circ$ angle like a bridle joint with IF4 contacting the periplasmic ends of TM 2 and 6 and IF5 contacting the periplasmic ends of TM1-3 (**Fig. 20A**). The pore for the cellulose polymer is framed on the cytosolic side of the membrane by IF1-3 and is

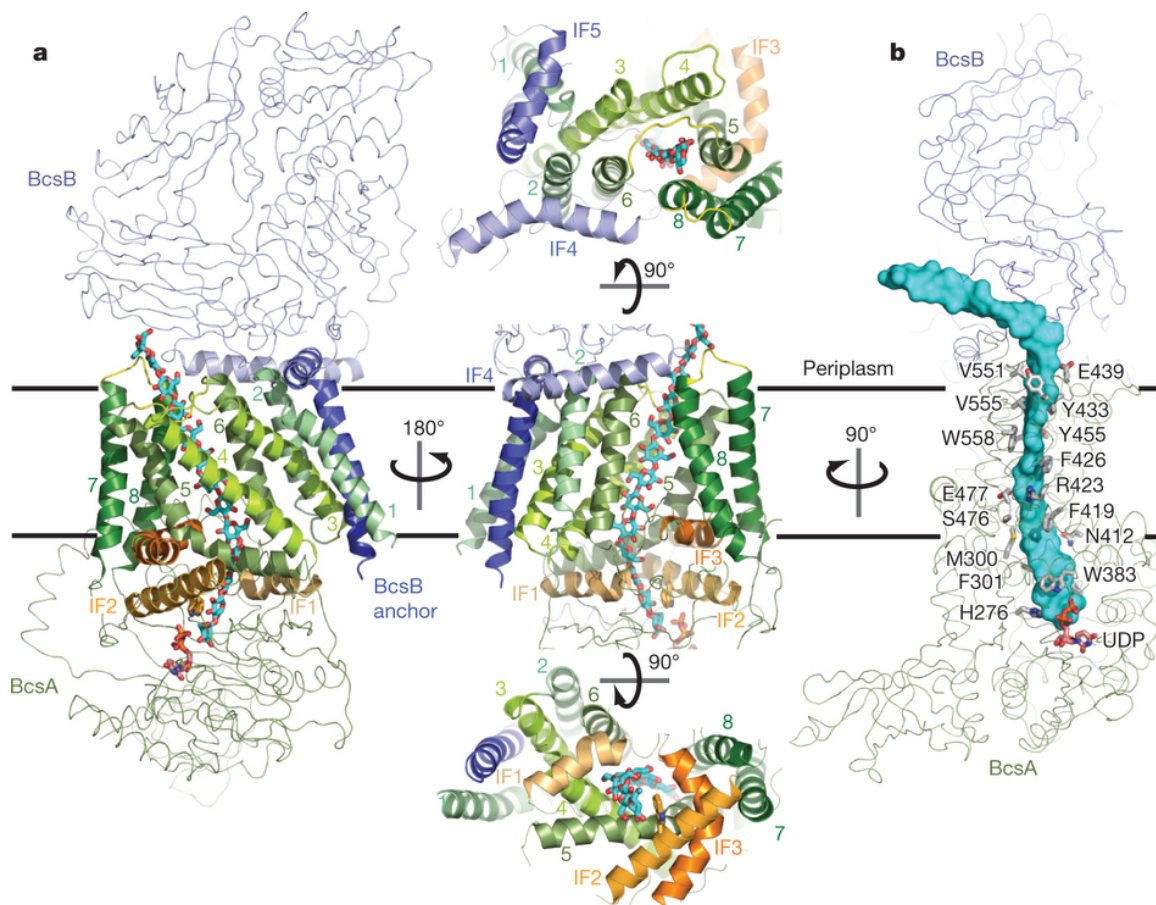


Figure 20 BcsA TM domain. A) TM region is shown with different colors for each helix. B) Residues contacting cellulose polymer (cyan surface) are shown as sticks and labeled (Morgan et. al., 2013)

formed mainly by TM 3, 4, 7, 8, and 6 in a clockwise order as viewed from the periplasm. The periplasmic exit of the pore is located on the side opposite to the BcsB TM anchor but near the N-terminal end of BcsB's IF4 (**Fig. 20A**). The polymer takes a $\sim 100^\circ$ turn upon exiting the pore and emerges from the interface between BcsA and BcsB at a $\sim 20^\circ$ angle to the plane of the membrane surface (**Fig. 20B**). The pore itself has a ribbon-like shape that nicely accommodates the flattened cellulose polymer. Furthermore, it has an amphipathic character that mirrors that of the polymer. The edges of the ribbon, where the equatorial hydroxyls sit, are enriched in hydrogen-bond donors and acceptors while the surface of the ribbon, where the faces of the glucopyranose rings will sit,

contain mainly hydrophobic residues (**Fig. 20B**). The prevalence of aromatic residues that stack on top of the glucopyranose rings of the cellulose polymer throughout the pore is particularly noteworthy because this interaction between the pi electrons of the aromatic residue and the C-H bonds on the face of a sugar (CH-pi stacking) is a common theme in protein-sugar interactions (Asensio et al., 2013). Overall, the limited hydrogen bonding with the polymer suggests that movement of the cellulose polymer through the pore does not appear to present a large energy barrier to translocation.

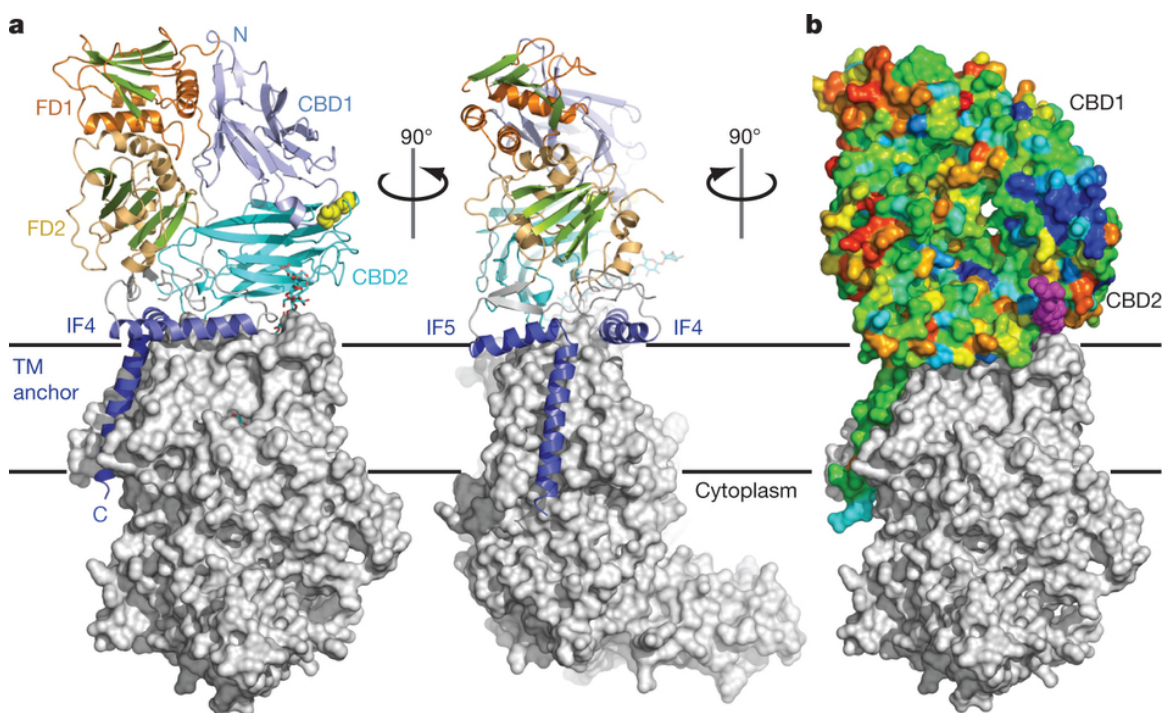


Figure 21 BcsB structural features. A) BcsB is shown as a cartoon with CBDs in light blue and FDs in green/gold, conserved Cys residues are shown as yellow spheres B) BcsB colored according to conservation with blue being most conserved and red being least conserved. (Morgan et. al., 2013)

The BcsB subunit forms a large periplasmic structure that is made up of two repeats of a carbohydrate-binding domain (CBD) followed by a flavodoxin-like domain (FD) with a C-terminal TM anchor. The overall topology of the BcsB subunit from N- to C-terminus

goes CBD1, FD1, CBD2, FD2 (IF4 is an insertion in FD2), IF5, TM anchor (**Fig. 21A**).

A disulfide bond formed between conserved cysteine residues C116 and C430 from CBD1 and CBD2, respectively, connects these two domains (**Fig. 9C and 21A**). Overall, the BcsB subunit has a low level of conservation. However, there is a patch of conserved residues where CBD1 and CBD2 come together near the disulfide bond (**Fig. 21B**).

Interestingly, this region contains the conserved W172, which sits right above where the polymer exits on the periplasmic side (**Fig. 21B**). As discussed previously, aromatic residues often mediate interactions with sugar molecules, suggesting that this conserved region could bind the polysaccharide on the periplasmic side to guide it towards the outer membrane (**Fig. 21B**).

A former undergraduate in the lab tested the role of BcsB on cellulose biosynthesis. Adi Narahari systematically truncated the N-terminus of BcsB by one domain at a time and measured activity for each truncation mutant (**Fig. 22 A and B**). He found that none of

BcsB's domains are

essential and that BcsA

could retain activity

with BcsB truncated all

the way down to the TM

anchor (**Fig. 22B**).

However, lone BcsA

had no activity. The

tight packing of the TM

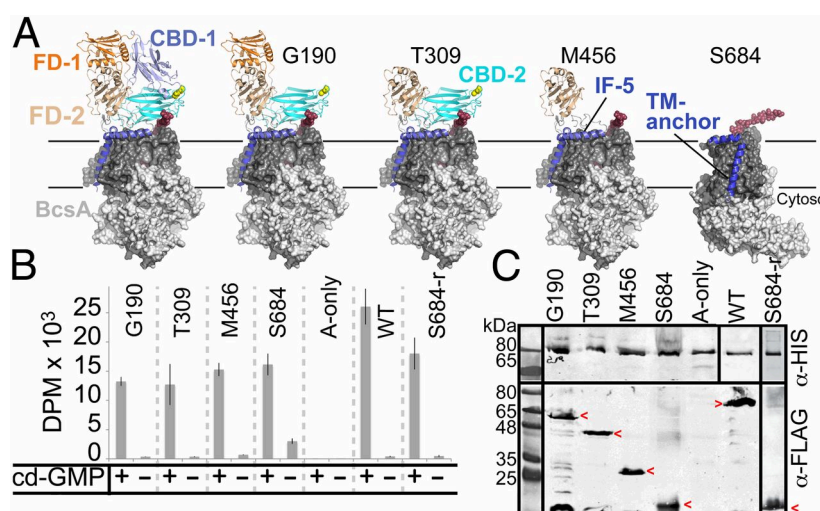


Figure 22 BcsB necessary features. A) cartoon representation of BcsA–BcsB showing BcsB truncations B) Activity of indicated truncations C) Western blot of BcsB truncated mutants (Omadjela et. al., 2013)

anchor of BcsB into the TM region of BcsA may play a role in proper folding of BcsA.

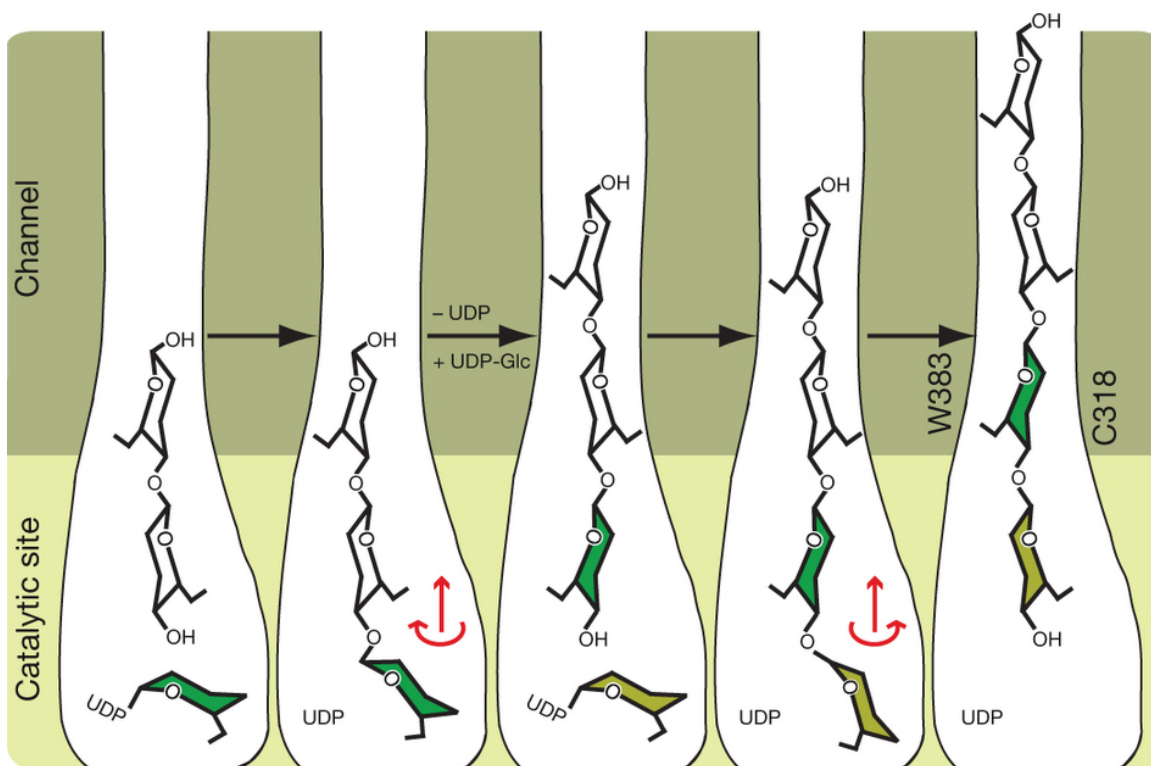


Figure 23 Initial model of cellulose biosynthesis. Channel and active site are indicated. W383 represents channel entrance. Newly-added sugar must rotate into plane of polymer to enter channel. (Morgan et. al., 2013)

At the time of this publication, we suggested the following model of cellulose synthesis and membrane translocation. The acceptor glucose would be positioned 1 unit beyond W383, while D343 would act as the catalytic base (**Fig. 23**). Donor sugar would bind in the active site, and glycosyl transfer would occur. Following glycosyl transfer, the newly-added glucose unit would need to rotate to assume the alternating conformation in line with the polymer (**Fig. 23**). This rotation would be accompanied by the formation of two new hydrogen bonds, and the formation of these hydrogen bonds could provide enough energy to allow the polymer to slip into the channel (**Fig. 23**). With the newly-added

sugar now sitting in the acceptor site, the enzyme is poised for another round of catalysis (Fig. 23). Data showing a complete picture of the cellulose synthesis and translocation cycle will be presented in chapter 3.

1.7 Open questions

The mechanism of cyclic-di-GMP activation of BcsA is a question that has stood since c-di-GMP was discovered in 1987. The structural and functional data presented above gives some insight into how binding at the PilZ domain may be coupled to catalytic activity at the active site. However, the mechanism we present is highly speculative. The most thorough way to study this mechanism would be to determine a structure of BcsA–B bound to c-di-GMP. Chapter two addresses this topic.

Additionally, the mechanism by which the synthesis of cellulose is coupled to its transport across the inner membrane represents a major remaining question for synthase-dependent protein secretion. As mentioned previously, all other biopolymers secretion mechanisms have synthesis and secretion separated and with each task carried out by specialized machineries. The biochemical data and initial structure of BcsA–B give little insight into how this works. Chapter three describes a unique and serendipitous approach that we used to address this problem.

In vivo the cellulose polymer must transverse the periplasm, which is loaded with peptidoglycan, and across the outer membrane through the BcsC subunit. A remaining question involves how cellulose crosses the periplasm. BcsC contains a large N-terminal

soluble domain and a C-terminal domain that is predicted to form an 18-stranded β -barrel (McNamara et al., 2015). Structures of the homologous components in the alginate system, AlgK, the soluble domain, and AlgE, the β -barrel (the domains are encoded on separate polypeptides in this system), are available and show that AlgK forms a superhelical clamshell-like structure of tetratricopeptide repeats (TPR) (Keiski et al., 2010). TPR-containing proteins are frequently involved in mediating protein-protein interactions, suggesting that an interaction between BcsC and BcsB could form a trans-periplasm conduit for the cellulose polymer. A postdoc in the lab is currently testing this hypothesis using experiments with the soluble domains of BcsB and BcsC to test whether they interact.

Finally, all known cellulose synthesis operons in bacteria include a BcsZ, a periplasmic enzyme that hydrolyzes cellulose (Römling, 2002). Because BcsZ is in the same operon, it is likely expressed at the same time as the other components while experiments in *G. xylinus* indicate that it is necessary for cellulose production, yet its role in cellulose production is unclear (Koo et al., 1998). It is possible that either BcsZ acts a quality-control protein to ensure that the periplasm doesn't fill with cellulose if the trans-periplasm conduit breaks down or that it plays a role in forming the conduit. Additional experiments testing for specific interactions between BcsZ and BcsC and/or BcsZ and BcsB will give some insight into whether BcsZ forms a structural role in the complex.

Chapter 2: Mechanism of activation of bacterial cellulose synthase by cyclic-di-GMP

Jacob L. W. Morgan¹, Joshua T. McNamara¹ and Jochen Zimmer¹

¹*Center for Membrane Biology*

Department of Molecular Physiology and Biological Physics

University of Virginia

Charlottesville, VA, USA

In this chapter I will present a paper that we published in *Nature Structural & Molecular Biology* (Morgan et al., 2014) addressing the mechanism of c-di-GMP activation of BcsA–B. I began working on this aspect of the project in early 2013, and we published the paper in early 2014. I was responsible for crystallizing BcsA–B under novel bicelle-mediated conditions, determining the structure, conceiving and testing the mutations, generating all figures (except alignment), and editing/writing of the manuscript. Joshua McNamara, a former technician in the lab and co-author on the paper, made Figure 35 showing the alignment.

2.1 Introduction

Biofilms are sessile multi-cellular bacterial communities that are encased in a 3-dimensional meshwork of biopolymers, such as polysaccharides, proteinaceous filaments and nucleic acids (Gloag et al., 2013; McCrate et al., 2013; Römling et al., 2013). The biofilm matrix provides protection against mechanical stress (Stewart and Costerton, 2001; Wilking et al., 2013) and controls the diffusion of signaling molecules, nutrients and toxic compounds. In fact, biofilm communities exhibit increased tolerance towards conventional anti-microbial treatments and sterilization techniques and are responsible for many chronic infections associated with cystic fibrosis and endocarditis (Pritt et al., 2007; Römling and Balsalobre, 2012) as well as nosocomial infections (Gomes et. al., 2014). In many cases, biofilm formation occurs in response to an elevated cytosolic concentration of cyclic-di-GMP (c-di-GMP) (Cotter and Stibitz, 2007), a bacterial signaling molecule recognized by a wide range of effector proteins, including transcription factors, flagellar components, riboswitches and exopolysaccharide synthases (Römling et al., 2013). Therefore, targeting c-di-GMP-binding effectors has emerged as an attractive new route for the development of urgently needed novel anti-microbial therapeutics.

C-di-GMP activates the synthesis of bacterial cellulose (Ross et al., 1987; Römling et al., 2013), an extracellular polysaccharide often found in biofilms (Zogaj et al., 2003). C-di-GMP monomers and dimers (Zhang et al., 2006; Gentner et al., 2012) are both recognized by effector proteins via PilZ domains, first identified as regulatory components of cell motility (Christen et al., 2007), which comprise an "RxxxR" motif in a flexible linker region followed by a β -sheet or β -barrel that contains a "DxSxxG" motif

(Amikam and Galperin, 2006). Both sequence motifs have been shown to interact with c-di-GMP in structures of isolated PilZ domains (Benach et al., 2007; Ko et al., 2010).

However, the mechanism by which c-di-GMP binding at PilZ domains modulates enzymatic functions is completely unknown to date.

Extracellular polysaccharides of the biofilm matrix, such as cellulose, alginate and poly-N-acetylglucosamine (PNAG), are likely synthesized and secreted by a conserved mechanism (Wang et al., 2004; Merzendorfer, 2006; Hay et al., 2013; Whitney and Howell, 2013; Hay et al., 2014). Bacterial cellulose synthase polymerizes glucose molecules via β -1,4 glycosidic linkages in a multi-step process which requires the presence of a divalent cation, mostly magnesium (Omadjela et al., 2013). First, upon stimulation by c-di-GMP, the enzyme binds its substrate UDP-Glc (donor) at an intracellular glycosyltransferase (GT) domain. Second, the donor glucose is transferred to the 4' hydroxyl group at the non-reducing end of the growing polysaccharide chain (acceptor), thereby extending the polymer and forming UDP as a second reaction product (Omadjela et al., 2013) (Brown et al., 2012). Third, following glycosyl transfer, the elongated polymer has to be translocated by one glucose unit into a transmembrane (TM) channel so that the newly added glucose unit occupies the acceptor site and UDP must be replaced with UDP-Glc for another round of catalysis.

The membrane-integrated bacterial cellulose synthase contains the inner membrane components BcsA and BcsB as well as the outer membrane protein BcsC (Mayer et al., 1991) (Saxena et al., 1994). BcsA, together with the periplasmic membrane-anchored

BcsB subunit, forms a complex that is sufficient for cellulose synthesis and translocation (Morgan et al., 2013; Omadjela et al., 2013). BcsA is homologous to eukaryotic cellulose synthases (Slabaugh et al., 2014) and contains eight TM helices and a cytosolic GT domain between TM helices four and five (Morgan et al., 2013). The enzyme is a processive family-2 GT (Cantarel et al., 2009) that elongates the non-reducing end of the growing polysaccharide chain. This reaction requires a general base, which is likely provided by the Asp residue of a “TED” motif found at the beginning of a short helix within the GT domain and in close proximity to the acceptor’s 4’ hydroxyl (Morgan et al., 2013). BcsA also forms a polysaccharide channel across the membrane, directly above the active site, thereby allowing the coupling of cellulose synthesis and translocation (Hubbard et al., 2012; Morgan et al., 2013).

Bacterial cellulose and alginate synthases are activated by c-di-GMP via PilZ domains (Amikam and Galperin, 2006; Merighi et al., 2007). BcsA forms a PilZ domain within its C-terminal intracellular extension, which consists of a six-stranded β -barrel and a preceding linker region (Amikam and Galperin, 2006; Morgan et al., 2013). The β -barrel rests against the intracellular GT domain and is connected to BcsA's C-terminal TM helix (TM8) via a linker (TM8- β -barrel linker) harboring the “RxxxR” motif involved in c-di-GMP binding (Amikam and Galperin, 2006).

The TM8- β -barrel linker also interacts with BcsA's "gating loop", which runs across the opening of the GT domain towards the cytosol, thereby blocking access to the catalytic pocket in the non-stimulated or “resting” state of the enzyme (Morgan et al., 2013). It

was speculated that substrate binding to the active site requires the repositioning of the gating loop, perhaps induced by c-di-GMP (Morgan et al., 2013). This model is supported by biochemical studies indicating that increasing c-di-GMP concentrations do not alter K_M , but instead increase the fraction of catalytically active enzymes (Omadjela et al., 2013).

In order to unravel the mechanism by which c-di-GMP activates bacterial cellulose synthase, we determined c-di-GMP-bound structures of the *Rhodobacter sphaeroides*

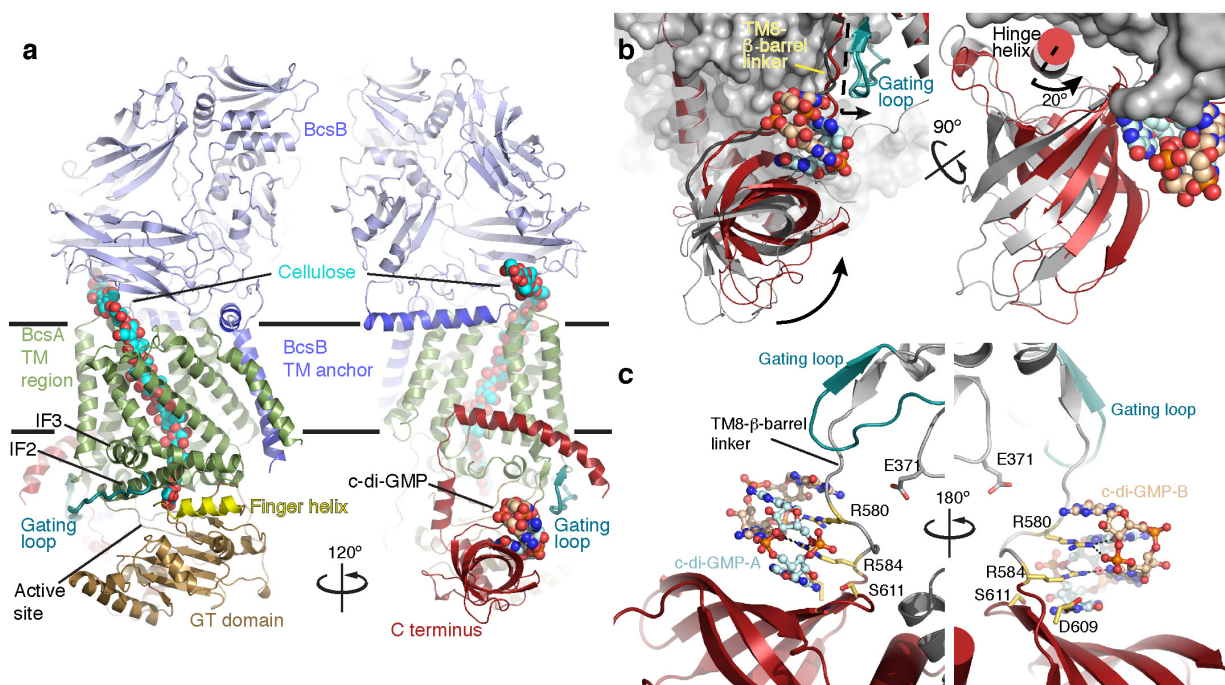


Figure 24 Structure of the c-di-GMP-activated BcsA–B complex. **(a)** Cartoon representation of the BcsA–B structure in the presence of a c-di-GMP homo-dimer. BcsA is colored brown, green and red for its GT domain, TM region, and C terminus, respectively. The 6-stranded b-barrel within BcsA's C terminus forms a c-di-GMP binding PilZ domain. BcsB is shown in blue. The c-di-GMP dimer and translocating cellulose polymer are shown in spheres. BcsA's finger helix and gating loop are colored yellow and steel blue, respectively. IF: Amphipathic interface helices that surround the cytosolic entrance to BcsA's TM channel. Horizontal bars indicate the putative membrane boundaries. **(b)** Comparison of BcsA's PilZ positions in the presence and absence of c-di-GMP. BcsA is shown as a pale gray surface, and BcsA's C terminus is shown as a red cartoon. The position of the b-barrel in the c-di-GMP-free state (pdb 4HG6) is shown as a gray cartoon and c-di-GMP is shown as spheres. **(c)** Interactions of the "RxxxR" and "DxSxxG" motifs with the c-di-GMP dimer. Residues of each motif are shown as yellow sticks and c-di-GMP is shown in sticks and spheres.

BcsA–B complex at intermediate states during cellulose synthesis and translocation. The c-di-GMP-bound structures reveal the architecture of the activated BcsA–B complex and provide unique insights into the mechanism of c-di-GMP signaling. These include the identification of a conserved regulatory salt bridge that auto-inhibits BcsA in the absence of c-di-GMP and the UDP-dependent repositioning of a gating loop to either open the catalytic pocket or to coordinate the nucleotide at the active site. Furthermore, the structures reveal the movement of a “finger helix” of BcsA, which interacts with the acceptor end of the translocating cellulose polymer, towards the TM channel entrance, correlating with the translocation of the cellulose polymer into the channel by one glucose unit. Thus, our data provide the first insights into the mechanism by which c-di-GMP modulates enzymatic functions and represent novel snapshots of cellulose synthesis and membrane translocation.

2.2 Results

2.2.1 Architecture of BcsA–B in complex with c-di-GMP

We purified *Rhodobacter sphaeroides* BcsA–B from *E. coli*, crystallized it in complex with c-di-GMP by the bicelle crystallization method (Faham and Bowie, 2002; Faham et al., 2005) and solved the structure by molecular replacement at a resolution of 2.65 Å. Additionally, we obtained a c-di-GMP- and UDP-bound structure of BcsA–B by soaking crystals with UDP and refining at 3.2 Å resolution (**Table 1**). Both structures contain a translocating cellulose polymer 17 glucose units in length that co-purifies with the BcsA–B complex.

Overall, the c-di-GMP-bound BcsA–B structure is consistent with the previously reported structure obtained from detergent-solubilized complexes (RMSD \approx 1Å for all atoms) (**Fig. 24**) (Morgan et al., 2013). Two register shifts by one residue were identified in regions that were poorly ordered in the previously reported structure of BcsA (residues 171–190) and BcsB (residues 268–280). The corrected register in BcsA positions Asp179 of the "DDG" motif in hydrogen bonding distance with the conserved Tyr216 and Asp180 in hydrogen bonding distance with the uracil moiety of UDP and Arg219 (Fig. 25).

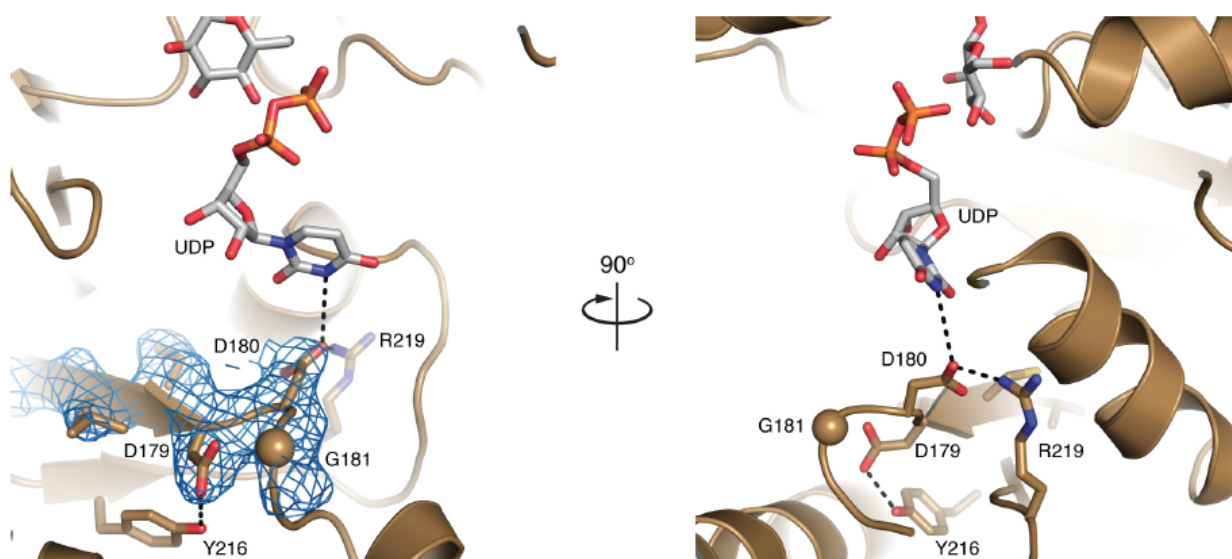


Figure 25 Correction of a register shift in BcsA. A register shift in BcsA (residues 171 to 190) has been corrected in the new structure. The new 2.65 Å electron density map allows the unambiguous assignment of the register in this β -strand. The corrected register positions Asp179 of the conserved "DDG" motif in hydrogen bonding distance to the conserved Tyr216 and Asp180 in hydrogen bonding distance to the UDP uracil moiety and the conserved Arg219. A SigmaA-weighted 2mFo-DFc electron density contoured at 1 σ is shown as a blue mesh. UDP and the translocating glucan as observed in pdb 4HG6 are shown as gray sticks.

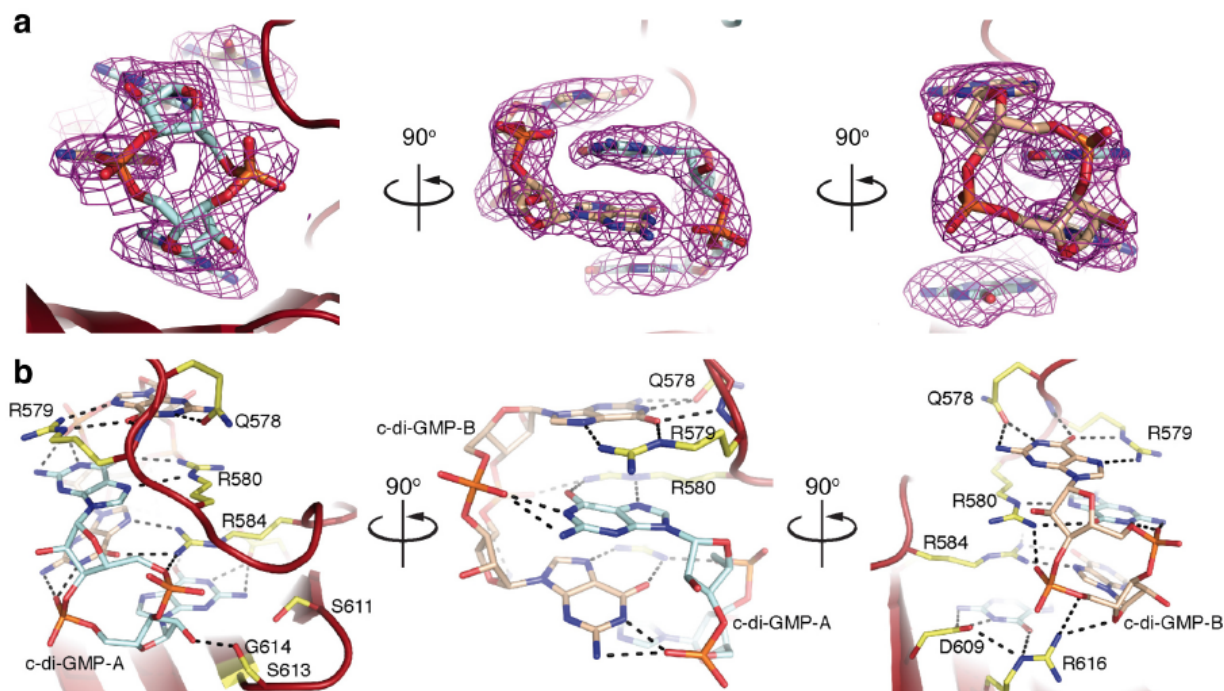


Figure 26 Cyclic-di-GMP binding to BcsA. (a) An intercalated homodimer of c-di-GMP binds to BcsA's PilZ domain. An unbiased SigmaA-weighted mFo-DFc difference electron density for c-di-GMP contoured at 4σ is shown as a magenta mesh. The density was calculated after refining the protein structure and before placing any ligands. The c-di-GMP dimer is shown in sticks colored light blue or pale brown for the carbon atoms, respectively. (b) BcsA's PilZ domain tightly coordinates a c-di-GMP dimer, c-di-GMP-A and -B. One guanylate group of c-di-GMP-A (GA1) packs its guanine group into a pocket on the β -barrel surface formed by the conserved Gly614 and Gly670 where it is further stabilized by side chain interactions with Asp609 of the "DxSxxG" motif as well as Ser613 and Arg616. The guanine interacts with Asp609 via its cyclic N1 atom and exocyclic amine group and its carbonyl oxygen contacts the guanidinium group of Arg616. Ser611 of the "DxSxxG" motif does not directly contact GA1, however, it is likely that its interaction is mediated by an unresolved water molecule. The 2' hydroxyl of the GA1 ribose interacts with the hydroxyl group of Ser613. Arg584 of the TM8- β -barrel linker stacks on top of the GA1 guanine and forms a salt bridge with the phosphate group belonging to the second guanylate of c-di-GMP-A (GA2). The side chain of the invariant Arg580 of the TM8- β -barrel linker is co-planar with the guanine of GA2 and forms hydrogen bonds via its guanidinium group with the GA2's guanine N7 and carbonyl oxygen. Similar to the stacking interactions observed for GA1, the preceding Arg579 stacks on top of the GA2 guanine group. The ring N1 and exocyclic amine group of GA2 interact with the phosphate moiety of the second c-di-GMP molecule (c-di-GMP-B). C-di-GMP-B makes fewer interactions with BcsA and is primarily stabilized by c-di-GMP-A and residues belonging to the TM8- β -barrel linker. Its first guanylate closest to the β -barrel surface (GB1) forms π - π stacking interactions with the guanine of GA2 and hydrogen bonds via its ring N7 and carbonyl oxygen with Arg584, the same residue that stacks on top of the GA1 guanine. As observed for the guanine group of GA2, its ring N1 and exocyclic amine contact the phosphate group of the other c-di-GMP molecule, thereby stabilizing the intercalated c-di-GMP dimer. The second guanylate of c-di-GMP-B (GB2) interacts via its ring carbonyl oxygen with the backbone nitrogen as well as the N ϵ of the co-planar Arg579 and via the guanine's N1 and exocyclic amine with the invariant Gln578 of the TM8- β -barrel linker. In addition, its phosphate group forms a salt bridge with Arg580 that is co-planar with the guanine of GA2.

2.2.2 BcsA binds a c-di-GMP dimer on the β -barrel surface

BcsA's C-terminal PilZ domain binds an intercalated c-di-GMP dimer (Zhang et al., 2006) (**Fig. 24** and **Fig. 26**). The guanine groups of the c-di-GMP dimer stack parallel to the β -barrel surface and perpendicular to the TM8- β -barrel linker. One c-di-GMP molecule (c-di-GMP-A) interacts with the "DxSxxG" motif on the β -barrel surface, while the second (c-di-GMP-B) is stabilized by π - π stacking interactions with c-di-GMP-A as well as by residues within the TM8- β -barrel linker (**Fig. 24c**).

All of the conserved PilZ domain residues mediate interactions with the c-di-GMP dimer (**Fig. 26b**) as also observed with isolated PilZ domains (Benach et al., 2007; Ko et al., 2010). Of note is the interaction of the "RxxxR" motif within the TM8- β -barrel linker with c-di-GMP (**Fig. 24c**). The N-terminal Arg of this motif (Arg580) runs co-planar to the second guanylate of c-di-GMP-A and forms hydrogen bonds via its guanidinium group with the guanine's N7 and carbonyl oxygen. The C-terminal Arg (Arg584) of the "RxxxR" motif also interacts with c-di-GMP-A by stacking on top of the first and forming a salt bridge with the phosphate group of the second guanylate moiety, (**Fig. 24c** and **Fig. 26b**). In the absence of c-di-GMP, Arg580 is rotated by almost 180° towards BcsA's GT domain and forms a salt bridge with Glu371 (Morgan et al., 2013). This interaction is broken upon c-di-GMP binding, leading to increased flexibility of BcsA's gating loop as described below.

The non-conserved Arg579, directly preceding the "RxxxR" motif, runs co-planar to the guanine group of c-di-GMP-B and stacks on top of c-di-GMP-A (**Fig. 26b**). A basic

residue in this position is likely necessary to stabilize the interaction with a c-di-GMP dimer, as demonstrated by mutagenesis studies on isolated PilZ domains (Ko et al., 2010; Fujiwara et al., 2013).

Most structures of β -barrel-containing PilZ domains contain a short α -helix that follows the last strand of the β -barrel and lays flat across its opening (Benach et al., 2007; Ko et al., 2010). In BcsA, this helix (termed hinge helix) is sandwiched at the interface between

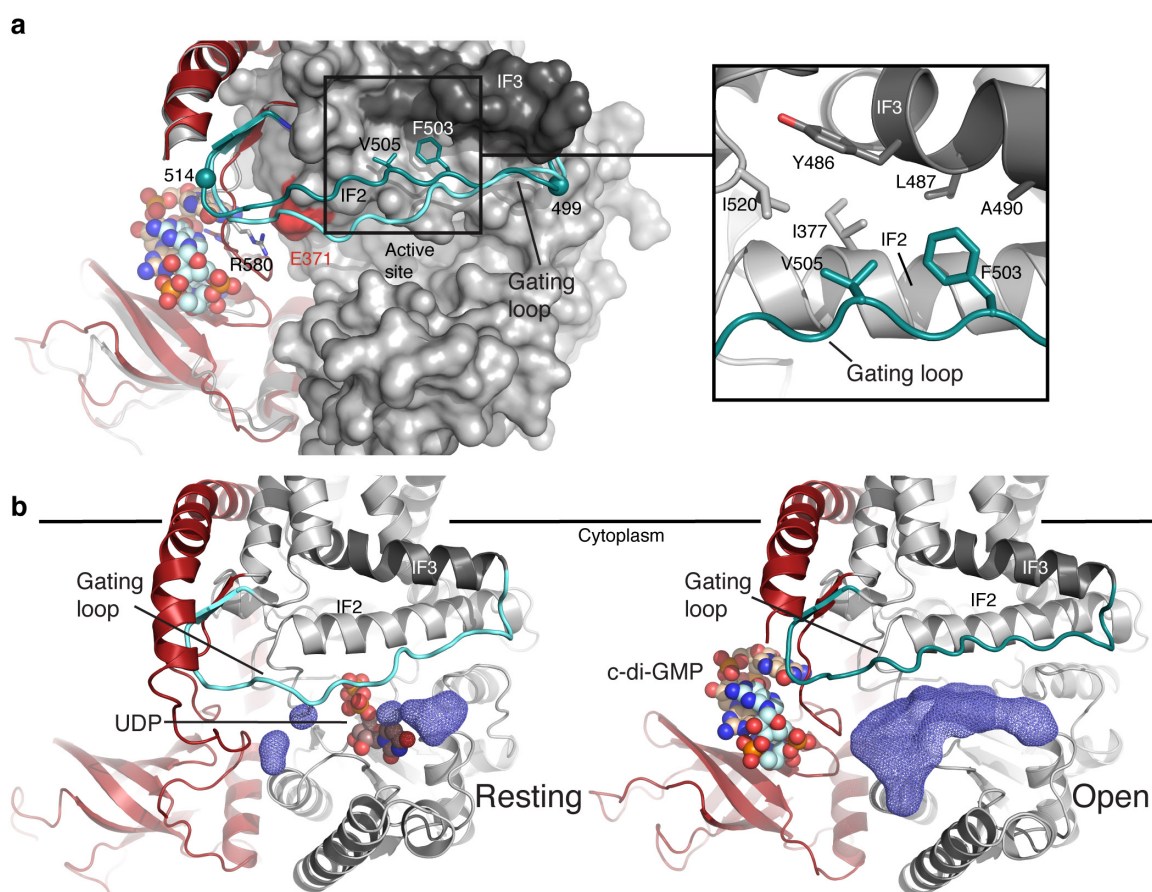


Figure 27 Conformational changes of BcsA's gating loop. **(a)** Gating loop positions in the absence and the presence of c-di-GMP (shown in cyan and steel blue, respectively). Phe503 and Val505 of the "FxVTxK" motif are shown as sticks and the gating loop's pivots, Arg499 and Glu514, are shown as spheres. The C terminus is colored as in Fig. 24. Inset: residues involved in stabilizing the gating loop in the "open" position are shown as sticks. **(b)** Accessible volume at the active site entrance (dark blue mesh) in the absence (left) and the presence (right) of c-di-GMP, calculated with a 3.5 Å probe sphere. UDP in the resting (pdb [4HG6](#)) and c-di-GMP in the open BcsA–B structure are shown as spheres.

the β -barrel and the GT domain (**Fig. 24b**). When the β -barrel interacts with c-di-GMP, it rotates by approximately 20° around the hinge helix towards the GT domain (**Fig. 24b**). This rotation closes a groove between the β -barrel and the GT domain that accommodates a short stretch of BcsA's non-conserved extreme C terminus in the c-di-GMP-free state (Morgan et al., 2013), leading to the disorder of BcsA's C-terminal residues past Arg740.

2.2.3 Conformational changes of the gating loop

C-di-GMP-binding allows BcsA's conserved gating loop (residues 499 to 517) to adopt a new conformation, away from the active site cleft and near the water-lipid interface, (**Fig. 27a** and **Fig. 28**). In this “open” state, the gating loop is stabilized by hydrophobic interactions with BcsA's amphipathic interface helices (IF), which run parallel to the plane of the membrane at the cytosolic water-lipid boundary (**Fig. 24 and 27a**) and form the entrance to BcsA's TM channel (Morgan et al., 2013). Phe503 and Val505 of the gating loop's "FxVTxK" motif (**Fig. 28a**) pack into a conserved hydrophobic pocket formed by Ile377 from IF2, Tyr486, Leu487 and Ala490 of IF3 and Ile520 at the beginning of TM helix 7 (**Fig. 27a**).

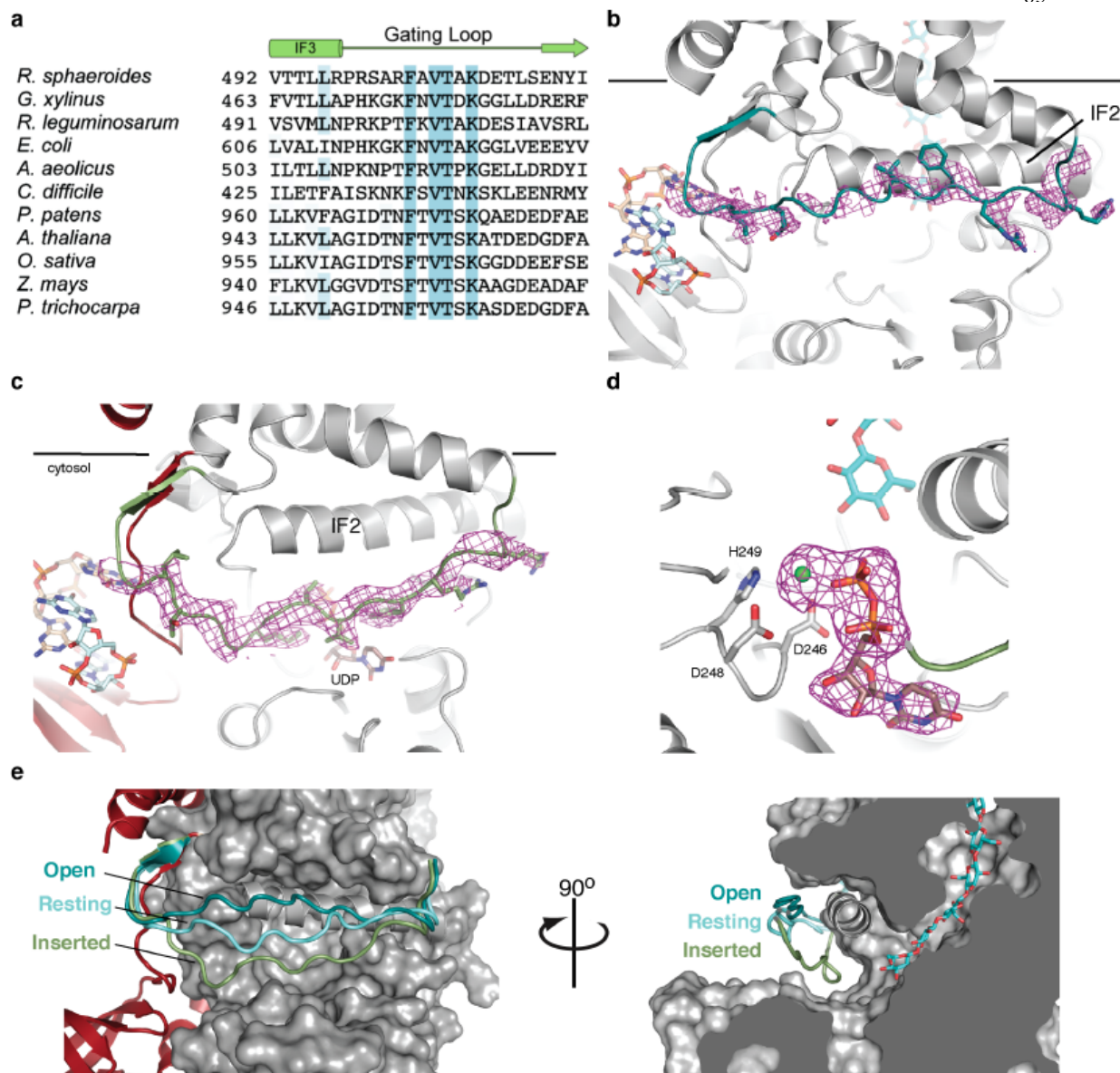


Figure 28 Features of BcsA's gating loop. (a) Sequence conservation of the gating loop. The "FxVTxK" motif is conserved in pro- and eukaryotic cellulose synthases. (b) Unbiased SigmaA-weighted 2mFo-DFc electron density for the gating loop in the "up" position shown as a magenta mesh and contoured at 1σ . The density was calculated before modeling the gating loop. The position of the loop's backbone is well resolved (colored cyan). The side chains of Phe503 and Val505 pack into a hydrophobic pocket on IF2 and are well resolved in the original density map. (c and d) Unbiased SigmaA-weighted mFo-DFc difference electron density for the gating loop in the "down" position and UDP, shown as a magenta mesh. The density was calculated before modeling the gating loop and placing UDP/Mg and is contoured at 2.5σ and 3σ for the gating loop and UDP, respectively. The position of the entire gating loop backbone is well resolved and so are the side chains of the conserved Phe503, Val505, Thr506 and Lys508. Additional electron density between the UDP β -phosphate and BcsA's "DxD" motif is consistent with a bound magnesium ion (shown as a green sphere). UDP is shown in sticks colored violet for the carbon atoms and the gating loop is colored green. (e) Front and side view comparing the three gating loop positions observed in the resting and c-di-GMP bound states. BcsA is shown as a gray surface with the PilZ domain shown as a red cartoon. The three gating loop positions are shown as cartoons and indicated with their respective colors.

The transition of the gating loop from the previously observed resting to the open state is

supported by c-di-GMP-induced conformational changes of the PilZ domain. In the absence of c-di-GMP, the gating loop rests in front of the GT domain entrance, thereby blocking the active site (**Fig. 27** and **Fig. 28**). This “resting” state is stabilized by Arg580

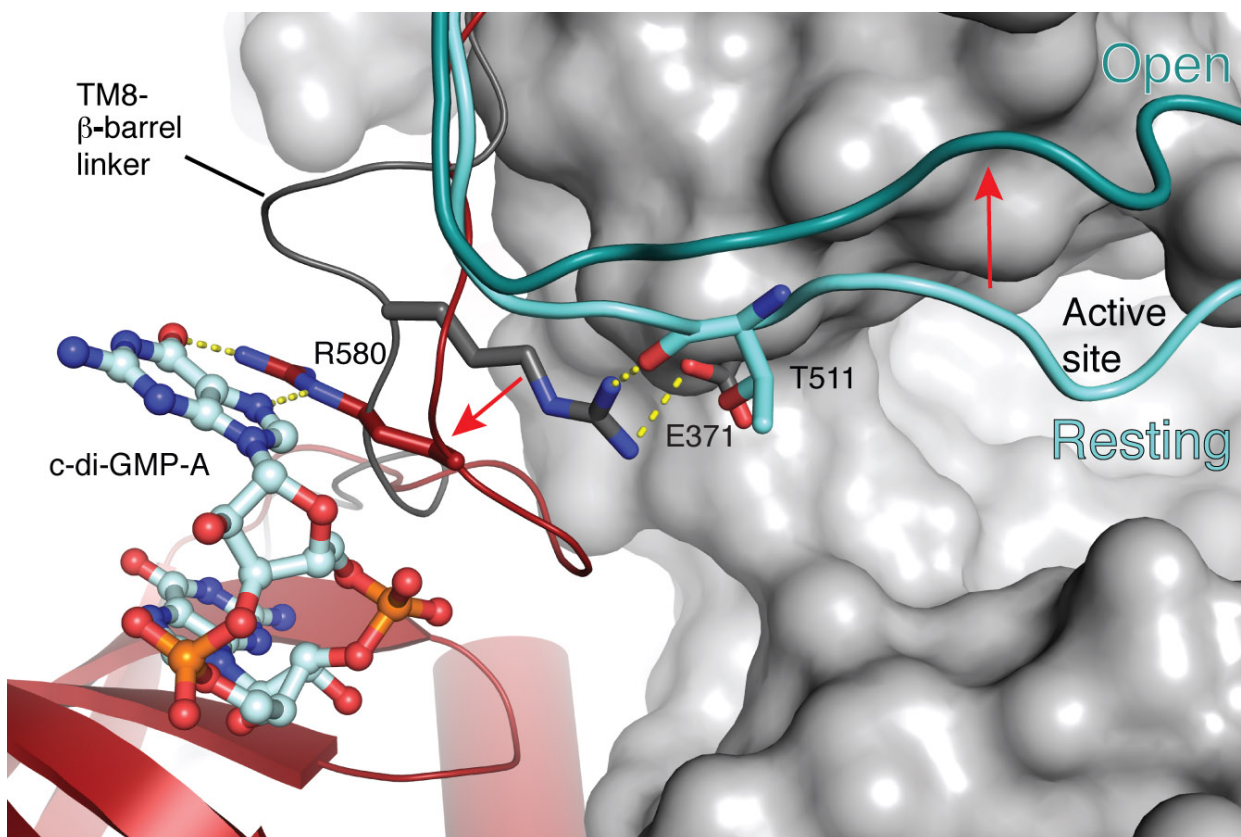


Figure 29 Stabilization of the gating loop by Arg580. A comparison of the Arg580 position in the absence and presence of c-di-GMP. Arg580 is shown as dark gray and red in the absence and presence of c-di-GMP, respectively. The gating loop is shown in cyan and steel blue representing the “resting” and “open” states, respectively. Glu371 is shown in sticks, and putative interactions are indicated. BcsA’s PilZ domain is colored red and the TM8-β-barrel linker in the resting state is shown as a dark gray cartoon.

of the “RxxxR” motif, which contacts the backbone carbonyl of Thr511 near the C-terminal end of the gating loop (**Fig. 29**). Arg580 is positioned in close proximity to Thr511 because it also forms a salt bridge with Glu371 of the GT domain, right next to the gating loop's Thr511 (**Fig. 29**). The gating loop further interacts with the PilZ domain via its C-terminal end, which forms a two-stranded β-sheet with the c-di-GMP-binding

TM8- β -barrel linker (**Fig. 24** and **27**). Accordingly, in the presence of c-di-GMP, the TM8- β -barrel linker together with the interacting gating loop rotates by about 2 Å towards the water-lipid interface (**Fig. 24b**) and, importantly, Arg580 rotates by 180° away from the GT domain to coordinate c-di-GMP, thereby breaking its interaction with the gating loop and Glu371 (**Fig. 29**). This transition releases the gating loop, allowing it to pivot around Arg499 and Glu514 and to swing from its resting state towards the membrane interface. The movement of the gating loop, particularly of residues 504–510, creates a large window at the GT domain entrance approximately 22.5 by 12.5 Å wide, which is sufficient to allow UDP-Glc to enter and UDP to exit the active site (**Fig. 27b**).

In order to mimic a substrate-bound state of BcsA, we soaked crystals with UDP, a product and competitive inhibitor of BcsA (Omadjela et al., 2013). In the presence of UDP the gating loop is found in another conformation, inserted deep into the substrate-binding pocket (**Fig 30a** and **Fig. 28c** and **d**). The loop swings by approximately 15 Å towards the active site, thereby closing the large window formed in its open conformation (**Fig. 30a** and **Fig. 28e**). In this inserted state, each residue of the loop's "FxVTxK" motif is involved in coordinating UDP at the active site. Phe503 and Val505 rest on opposing sides above the uracil moiety while Thr506 and Lys508 coordinate its pyrophosphate (**Fig. 30b**). The pyrophosphate is further stabilized by Gln379 and Arg382 of the "QxxRW" motif, as well as a Mg²⁺ ion coordinated in turn by Asp246 and Asp248 of the "DxD" motif and His249. Thus, the insertion of the gating loop is likely important for positioning the donor for catalysis. Indeed, in the absence of c-di-GMP, its insertion into the active site is prevented due to steric clashes of its backbone with the Arg580-Glu371

salt bridge (**Fig. 31b**), further ensuring a catalytically inactive state.

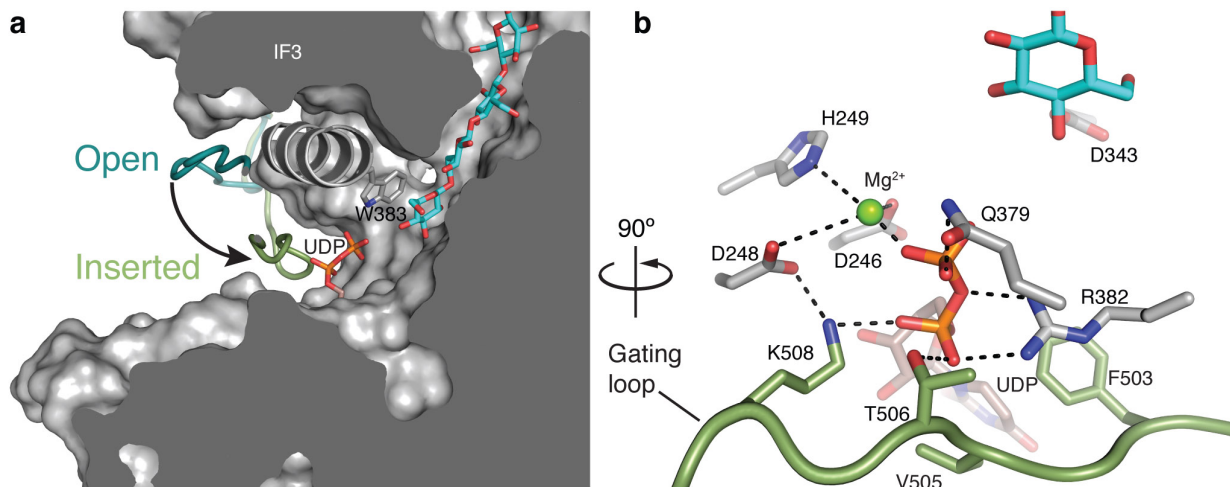


Figure 30 Insertion of the gating loop into the catalytic pocket. **(a)** A comparison of the gating loop positions of c-di-GMP-bound BcsA–B in the absence and presence of UDP (shown in steel blue and green, respectively). The inserted gating loop is colored green, IF2 is shown as a gray cartoon helix, and UDP as well as the translocating cellulose polymer are shown as sticks. Trp383 of the “QxxRW” motif at the entrance to the TM channel is shown as gray sticks. **(b)** Coordination of UDP at the active site by the gating loop’s “FxVTxK” motif. The gating loop is colored green, representing the “inserted” state. UDP, the conserved residues of the gating loop as well as Gln379 and Arg382 of the “QxxRW” motif are shown as sticks. The terminal glucose of the cellulose polymer and the putative catalytic base (Asp343) are shown as cyan and gray sticks, respectively.

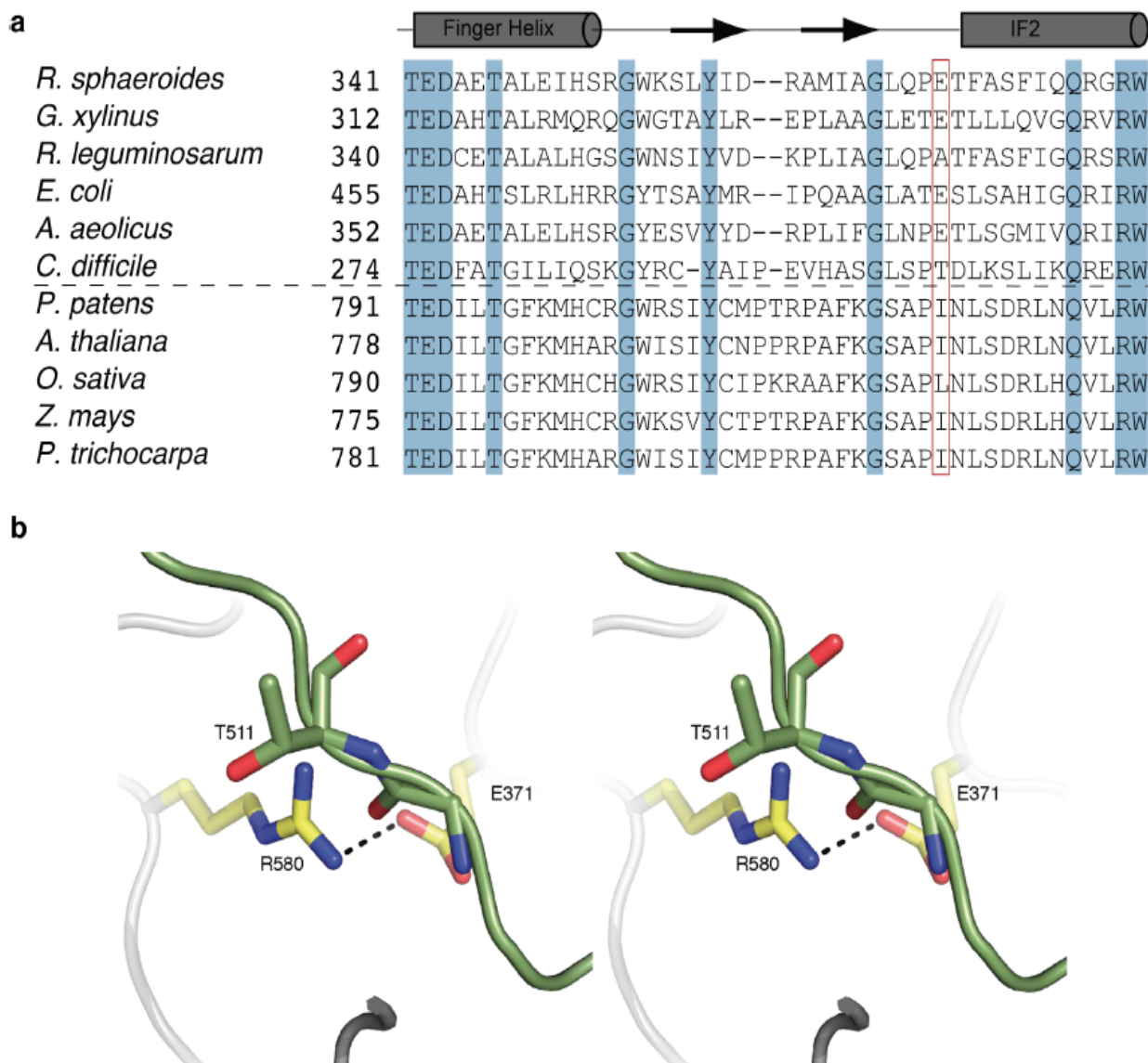


Figure 31 A conserved salt bridge stabilizes the resting position of BcsA's gating loop. (a) Sequence alignment of pro- and eukaryotic cellulose synthases. In the absence of c-di-GMP, the N-terminal Arg of the PilZ domain's "RxxxR" motif forms a salt bridge with a conserved Glu within the GT domain (framed red). Some outliers, such as BcsA from *R. leguminosarum*, contain an Ala at this position, expected to decrease the dependence on c-di-GMP for cellulose biosynthesis. The *C. difficile* sequence might be shifted in this region and the Asp residue next to the aligned Thr might confer a similar functionality. For eukaryotic cellulose synthases, Ile is the most prevalent residue at the corresponding position. The secondary structure of the aligned sequences is shown as a cartoon based on the *R. sphaeroides* BcsA structure. Pro- and eukaryotic sequences are separated by a dashed line. (b) The Glu371-Arg580 salt bridge blocks gating loop insertion in the absence of c-di-GMP. Stereoview of a superposition of pdb 4HG6 and the c-di-GMP/UDP bound structure. Arg580 and Glu371 from 4HG6 are shown as yellow sticks. The inserted state of the gating loop from the c-di-GMP/UDP bound structure is shown in green. A clash between Glu371-Arg580 and the C-terminal end of the gating loop would prevent loop insertion in the absence of c-di-GMP.

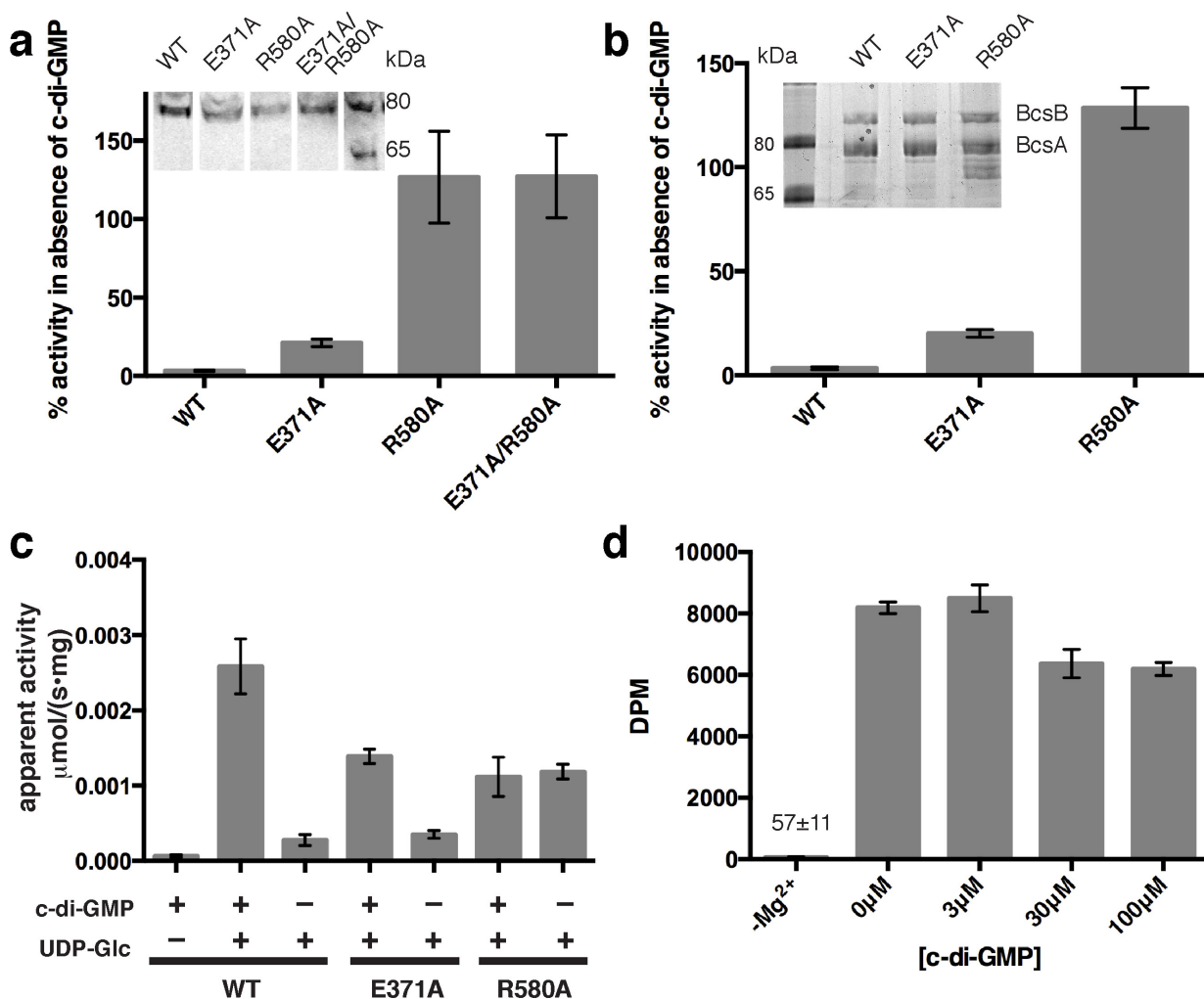


Figure 32 Comparison of BcsA-catalyzed *in vitro* cellulose synthesis in the absence and the presence of c-di-GMP. **(a and b)** Inverted membrane vesicles and proteoliposomes (PL), respectively, containing BcsA–B with the indicated mutations in BcsA were used for cellulose synthesis assays. (WT: wild type). The activity in the absence of c-di-GMP is quantified relative to the activity in the presence of 30 μM c-di-GMP. Insets: Western analysis of IMVs against the C-terminal poly-histidine tag on BcsA (a) and Coomassie-stained SDS-PAGE of the purified BcsA–B complexes (b). **(c)** Catalytic rates of the indicated PL-reconstituted BcsA–B mutants in the presence and absence of 30 μM c-di-GMP as measured by quantifying the formation of UDP. **(d)** Activity of the PL-reconstituted BcsA R580A mutant at increasing c-di-GMP concentrations. No activity is observed in the presence of 30 μM c-di-GMP when magnesium is depleted with 25 mM EDTA (–Mg²⁺). DPM: Disintegrations per minute. (All data represent the means ± SD for 3 technical replicates).

2.2.4 C-di-GMP releases an auto-inhibited state of BcsA

Activation by c-di-GMP is a characteristic of prokaryotic cellulose synthases (Aloni et al., 1982; Ross et al., 1987). Arg580 within the TM8-β-barrel linker either interacts with c-di-GMP or, in the absence of the allosteric activator, is stabilized towards the GT

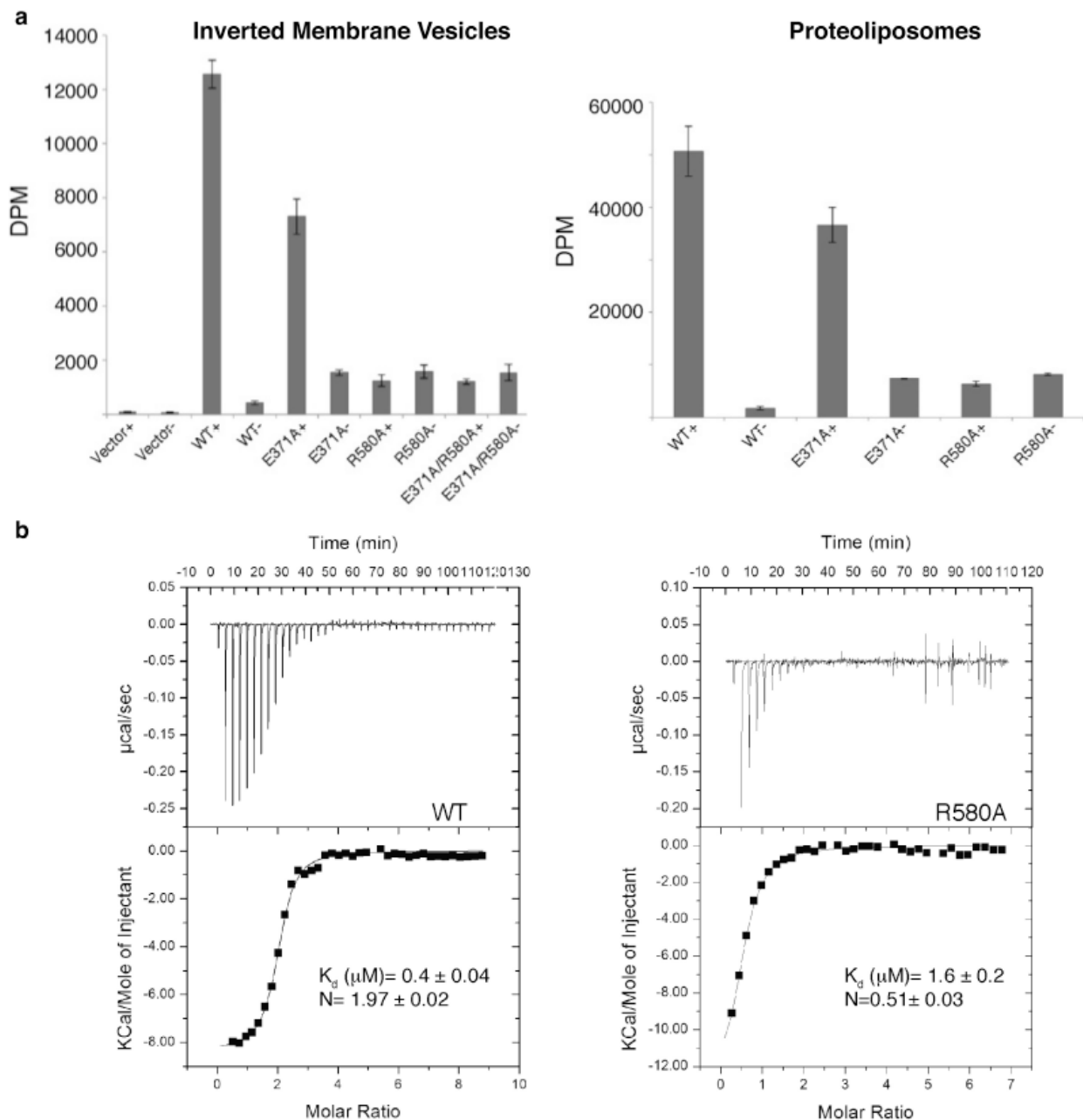


Figure 33 Caption on following page.

domain by forming a salt bridge with Glu371, thereby tethering the gating loop in the resting position (**Fig. 29**). Although belonging to the evolutionarily conserved GT domain, Glu371 is only conserved among prokaryotic, c-di-GMP-responsive cellulose synthases, suggesting a regulatory function for the Arg580-Glu371 interaction (**Fig. 31a**).

Indeed, disrupting this salt bridge by replacing Glu371 with Ala increases the enzyme's catalytic activity in the absence of c-di-GMP approximately 6-fold compared to the wild

type enzyme, (**Fig. 32a and b**). Under these conditions, Arg580 may still be able to interact with the gating loop's backbone. However, replacing Arg580 with Ala, either in the wild type or E371A background, renders BcsA constitutively active as observed by quantifying the formation of each reaction product, cellulose (**Fig. 32a and b and Fig. 33a**) or UDP (**Fig. 32c**).

Importantly, the R580A mutant still binds c-di-GMP, although with slightly reduced affinity (**Fig. 33b**); yet, even at a c-di-GMP concentration more than 50-fold above its dissociation constant, no further stimulation of cellulose biosynthesis is observed (**Fig. 32d**). These observations suggest that the Arg580-Glu371 salt bridge and the subsequent interaction of Arg580 with the gating loop are responsible for auto-inhibiting the synthase. This inhibition is then released when Arg580 rotates away from the GT domain to interact with c-di-GMP.

Figure 33 (cont.) Cellulose synthesis activity of BcsA mutants and c-di-GMP binding. (a) Cellulose synthesis assays were performed in inverted membrane vesicles and proteoliposomes as described in the Experimental Procedures. The amount of ³H-glucose-labeled cellulose produced by each mutant is quantified and graphed. 1 μ l of IMV's were used for each mutant. For PL assays, the protein concentrations were matched based on UV absorbance and SDS-PAGE followed by Coomassie staining. The apparent lower activity of the R580A mutant may be due to differences in the relative orientation of the enzyme in the PLs or its overall stability. +/-: Experiments performed in the presence and absence of 30 μ M c-di-GMP. All data represent the means \pm SD for 3 technical replicates. (b) Binding of c-di-GMP to the BcsA-R580A mutant. The ability of the R580A mutant to bind c-di-GMP was assessed using ITC. Left panel, titrating c-di-GMP into wild type (WT) BcsA-B in 1 mM LFCE14 results in an exothermic curve with a K_d of 0.4 μ M and a c-di-GMP to BcsA-B stoichiometry of \sim 2:1 as expected based on the crystal structure. Right panel, BcsA-R580A under the same conditions exhibits a K_d of 1.6 μ M and a stoichiometry of \sim 0.5. The observation that only 1/2-1/4 of the BcsA-R580A population (depending on whether the mutant binds a c-di-GMP monomer or dimer) interacts with c-di-GMP suggests that a fraction of it is structurally compromised, consistent with the results obtained from functional assays (a).

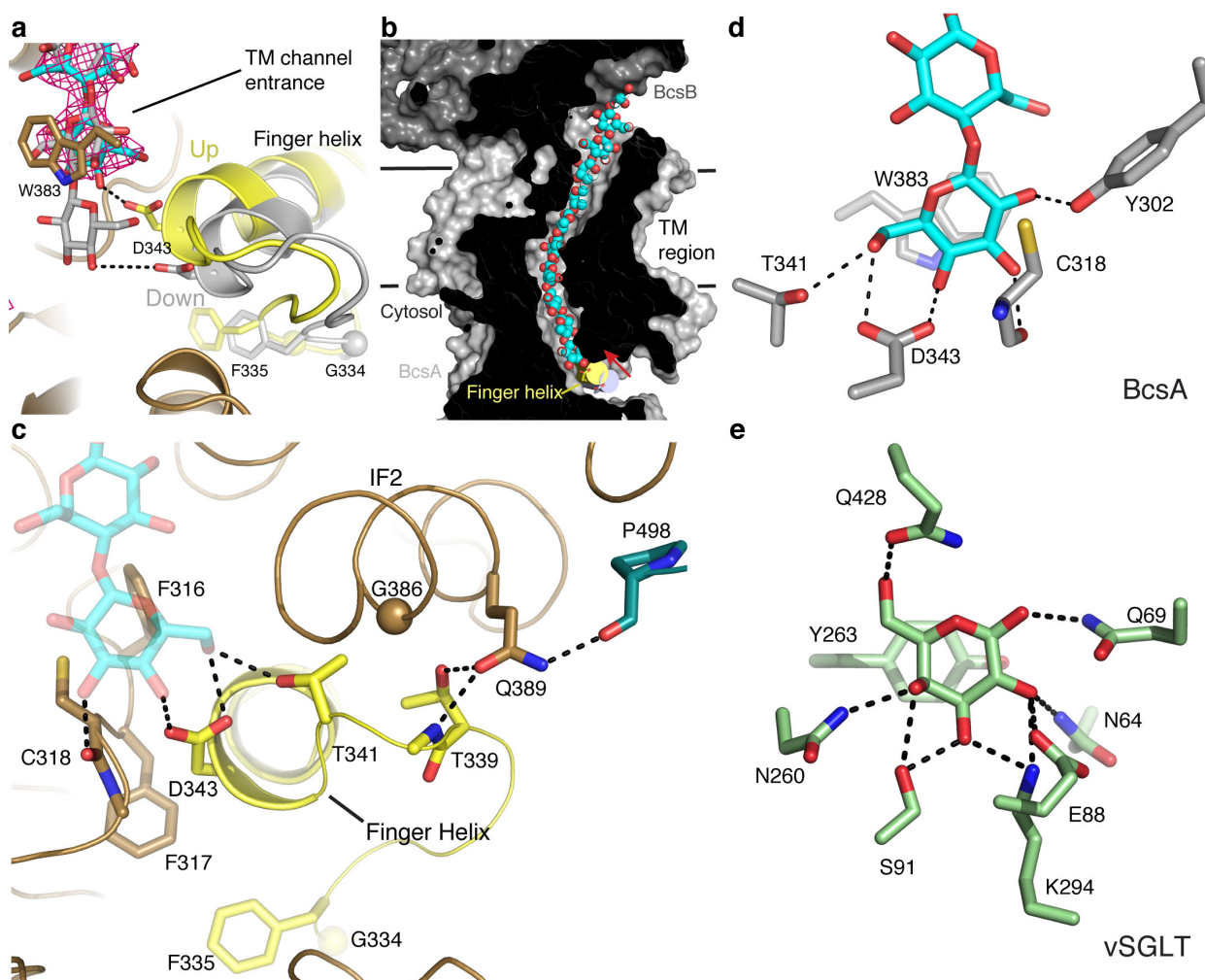


Figure 34 Movement of the finger helix, cellulose translocation, and the acceptor position. **(a)** Comparison of the positions of BcsA's finger helix and translocating glucan in the resting (colored gray, pdb [4HG6](#)) and c-di-GMP-bound states. An unbiased SigmaA-weighted mFo–DFc difference electron density of the translocating cellulose polymer in the UDP-free, c-di-GMP-bound state is contoured at 3.5s and shown as a magenta mesh. The positions of the glucan as observed in the resting state and the c-di-GMP-bound structure are shown as gray and cyan sticks, respectively. BcsA's finger helix and the preceding small loop are colored yellow. **(b)** Cut-away view of BcsA's TM channel with the position of the finger helix in the c-di-GMP-bound and resting states shown as yellow and gray solid cylinders, respectively. The translocating glucan is shown as cyan and red spheres and Asp343 is shown as sticks. Membrane boundaries are indicated by horizontal black lines. **(c)** Conserved residues involved in stabilizing the finger helix in the "up" position are shown as sticks. **(d)** Residues coordinating the polymer's terminal glucose are shown as sticks. **(e)** Stabilization of a single galactose molecule by the sodium-dependent sugar transporter vSGLT. Residues coordinating galactose are shown as sticks (pdb [3DH4](#)).

2.2.5 The TM channel entrance forms the acceptor-binding site

Following sugar transfer, processive GTs, including cellulose, chitin, alginate and hyaluronan synthases, must translocate the elongated polysaccharide, such that the newly formed product sits in a position where it can serve as the acceptor in a subsequent glycosyl transfer reaction. BcsA contacts the acceptor end of the translocating cellulose polymer via a “finger helix” that belongs to the conserved GT domain (**Fig. 24**). The finger helix contains the “TED” motif at its N terminus, of which Asp343 most likely forms the general base for catalysis (Morgan et al., 2013). In contrast to our previously reported structure in which the finger helix points away from the TM channel entrance (“down” state) (Morgan et al., 2013), this helix is bent towards the entrance to the channel in both of our new structures and the cellulose polymer is moved into the channel by one glucose unit (**Fig. 34a**). The finger helix bends near its last C-terminal helical turn and around the conserved His351 (**Fig. 35a**), which is stabilized via side chain interactions with the conserved Ser357 and Tyr410, thereby forming a pivot (**Fig. 35b**). Asp343 at the tip of the finger helix moves by approximately 5 Å towards the TM channel entrance, which is in agreement with the length of a glucopyranose unit (**Fig. 34a and b**).

A network of conserved hydrophilic and hydrophobic interactions stabilizes the “up” position of the finger helix near the TM channel entrance, including residues from the gating loop, IF2 and the TM channel. Phe316 and Phe317 of the “FFCGS” motif (**Fig. 36**) at the TM channel entrance straddle the helix and additional van-der-Waals interactions are mediated via Gly386 and Met390 of IF2 and Tyr410 within the N-terminal amphipathic section of TM helix 5 (**Fig. 34c and Fig. 35b**). In addition, Thr339,

preceding the “TED” motif of the finger helix, hydrogen bonds with the conserved

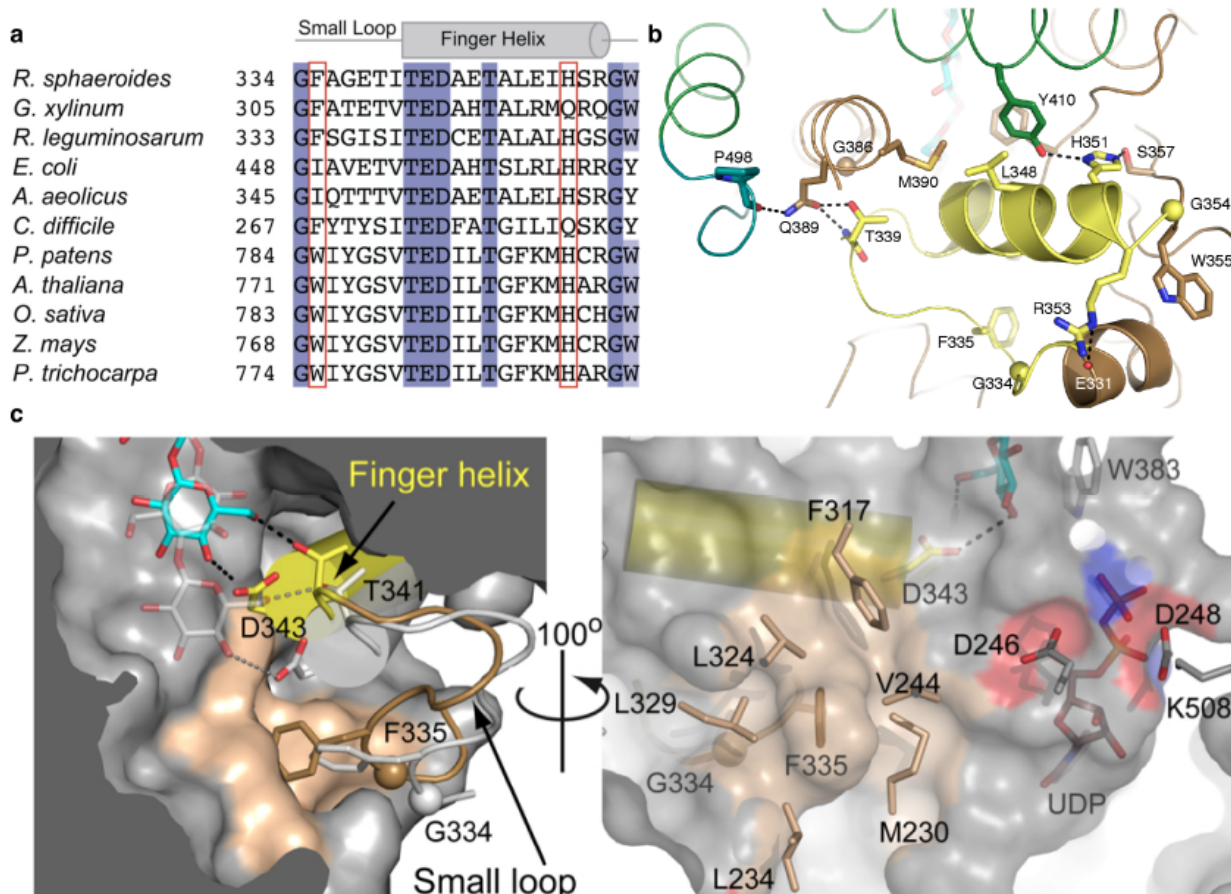


Figure 35 The movement of BcsA's finger helix is supported by a small loop. (a) Sequence alignment of pro- and eukaryotic cellulose synthases. The small loop contains a conserved Gly residue that is followed by a hydrophobic residue, usually Phe or Trp (framed red). The finger helix carries the invariant "TED" motif at its N terminus and contains a conserved His residue (framed red) followed by a Gly-Trp/Tyr motif that stabilizes the helix at its C terminus. (b) The C terminus of BcsA's finger helix (shown as a yellow cartoon) is stabilized by His351, which interacts with the invariant Ser357 and Tyr410. Leu348 packs against Met390 of IF2 and Tyr410 of IF3. Arg353 and Trp355 cap the helix at its C-terminal end. (c) Comparison of BcsA's finger helix and small loop in the presence and absence of c-di-GMP. The translocating glucan is shown before and after finger helix movement as gray and cyan sticks, respectively. In the presence of c-di-GMP, Phe335 of the small loop packs into a hydrophobic pocket (shown as a pale orange surface) underneath the finger helix. The preceding conserved Gly334 is shown as a sphere. The position of the finger helix and small loop in the resting state (pdb 4HG6) is shown as a gray cartoon. The right panel shows the same surface boundary as on the left but viewed from the opposite side.

Gln389 of IF2, which in turn interacts with Pro498 at the N-terminal end of the gating loop (Fig. 34c).

The transition of the finger helix towards the channel entrance is supported by a small peripheral loop (residues 333 to 338) that precedes the finger helix. The loop carries a conserved Gly residue (Gly334) at its midpoint, which is followed by a bulky hydrophobic residue, mostly Phe or Ile (**Fig. 34a** and **c** and **Fig. 35a**). In the c-di-GMP-bound state, the loop moves towards the GT domain and Phe335 plugs into a conserved hydrophobic pocket beneath the active site, where it is surrounded by the side chains of Met230, Leu234, Val244, Phe317, Leu324 and Leu329 (**Fig. 34c** and **Fig. 35c**).

The position of the translocating glucan's terminal glucose unit in the c-di-GMP-bound BcsA–B complex suggests that the acceptor coordination site is located just inside the entrance to the TM channel (**Fig. 34a** and **d**). One face of the acceptors glucopyranose ring forms CH- π stacking interactions (Kumari et al., 2012) with Trp383 of the “QxxRW” motif, which is characteristic of processive GTs, while its opposite side contacts the carbonyl oxygen of Cys318 of the “FFCGS” motif (**Fig. 34c** and **Fig. 36**). The N-terminal part of the finger helix contacts the acceptor via the “TED” motif, of which Thr341 and Asp343 form hydrogen bonds with the acceptor's 2' or 6' (depending on its orientation) and 4' hydroxyl groups (**Fig. 34c** and **d**). In particular, the side chain carboxylate of Asp343 is within 2.6 Å of the acceptor's 4' hydroxyl group, consistent with its putative role as the catalytic base during glycosyl transfer (Morgan et al., 2013). The acceptor further interacts with Tyr302, also located at the entrance to the TM channel (**Fig. 34d**). Accordingly, this implies that cellulose synthase forms the acceptor coordination site with the invariant “QxxRW” and “FFCGS” motifs at the entrance to the TM channel (**Fig. 36**). The stabilization of the terminal glucose unit at this position

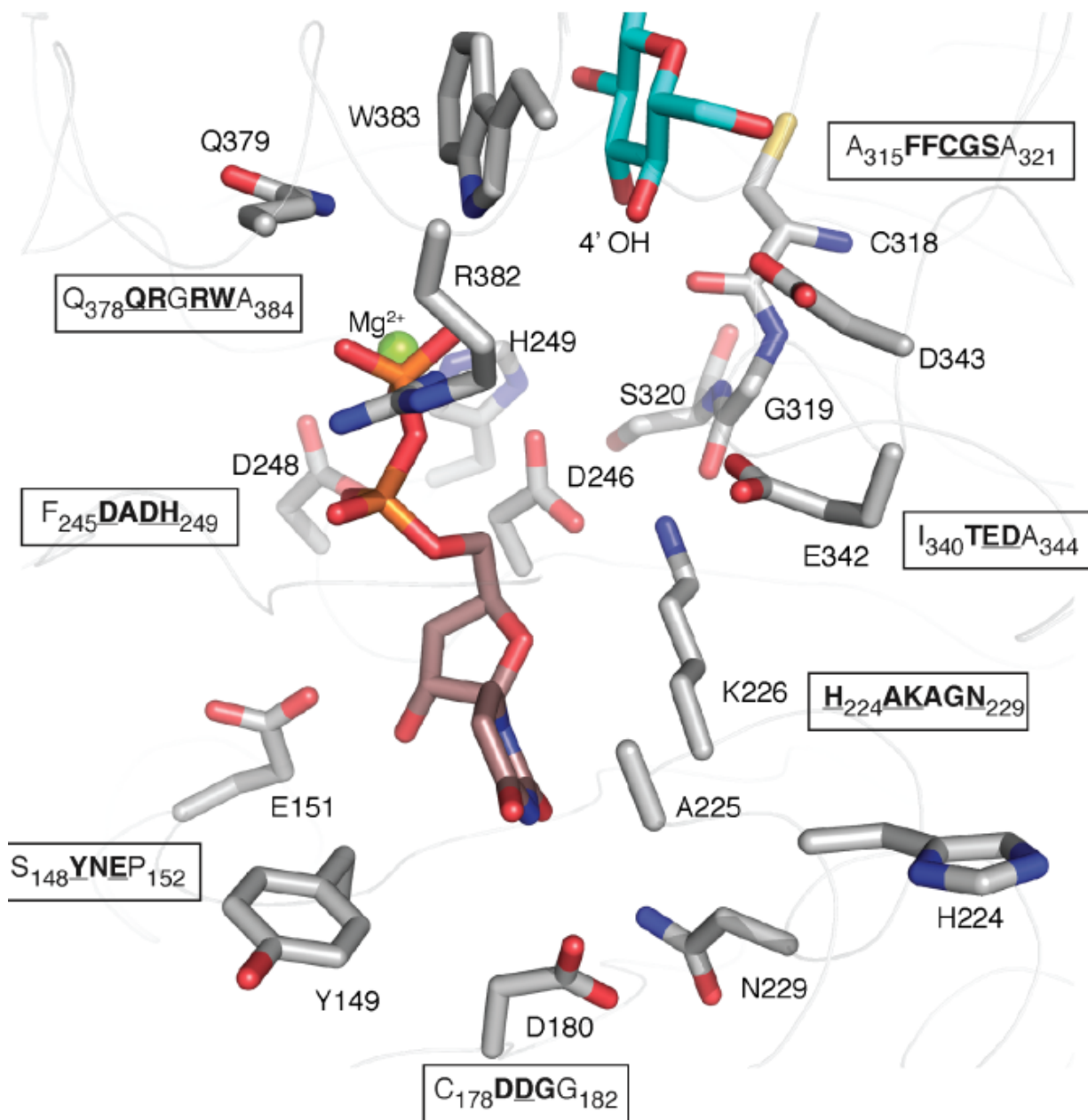


Figure 36 Active site signature motifs involved in donor and acceptor coordination. Conserved residues are shown in sticks. The acceptor glucose is stabilized by interactions with Trp383 belonging to the “QxxRW” motif as well as the backbone carbonyl of Cys318 of the “FFCGS” motif. Residues in BcsA’s gating loop that also contact UDP (Fig. 30b) have been omitted for clarity.

resembles the coordination of a single galactose molecule in the sodium-coupled sugar transporter vSGLT (Faham et al., 2008), (**Fig. 34e**), suggesting that BcsA’s acceptor binding site might also suffice to coordinate a single glucose molecule to initiate cellulose synthesis.

2.3 Discussion

Comparing the structures of the c-di-GMP-activated and resting states of the BcsA–B complex, at intermediate states during cellulose translocation provides unique insights into the mechanism of cellulose biosynthesis. In the absence of c-di-GMP, BcsA is catalytically inactive and its gating loop blocks the entrance to the active site (Morgan et al., 2013; Omadjela et al., 2013). Allosteric activation by c-di-GMP displaces the gating loop from the active site, thereby forming a large opening towards the substrate-binding pocket, wide enough for substrate diffusion. However, opening and closing the active site is unlikely to be the only function of BcsA's gating loop. When UDP binds to the active site, the gating loop inserts deeply into the catalytic pocket and coordinates the nucleotide via conserved residues. Most likely, this also reflects how BcsA interacts with its substrate UDP-Glc, positioning it for catalysis, excluding water from the active site and perhaps also stabilizing the UDP leaving group during glycosyl transfer. A similar mechanism of substrate-dependent loop insertion and de-insertion has been described for non-processive galactosyltransferases (Ramakrishnan et al., 2006; Qasba et al., 2008).

The functional importance of the gating loop is further underlined by its sequence homology with the location of the isoxaben resistance mutation in *Arabidopsis thaliana* cellulose synthase 3 (**Fig. 28a**). Here, Thr942 of the “FxVTxK” motif is mutated to Ile, thereby allowing growth in the presence of the herbicide isoxaben (Scheible et al., 2001). However, because pro- and eukaryotic cellulose synthases differ in their predicted TM

topologies (Slabaugh et al., 2014), further experimental analyses are required to confirm a similar eukaryotic gating loop function.

UDP, the second reaction product of many GTs (Lairson et al., 2008; Brown et al., 2012), competitively inhibits BcsA, which has also been observed for hyaluronan synthases (Tlapak-Simmons et al., 2004; Omadjela et al., 2013). BcsA binds UDP and UDP-Glc with similar affinities (Omadjela et al., 2013), however, the large excess of UDP-Glc over UDP under physiological conditions would favor substrate binding upon gating loop opening (Buckstein et al., 2008). Presumably during or after UDP-Glc binding, the gating loop inserts into the active site to initiate catalysis. Following glycosyl transfer and with the newly extended glucan at the active site, the gating loop may retract from the GT domain, thereby allowing UDP to UDP-Glc exchange. Because the gating loop undergoes its full range of motion in the presence of c-di-GMP, it is likely that the allosteric activator remains bound during catalysis. *In vivo*, c-di-GMP-stimulated cellulose biosynthesis may terminate upon depletion of the activator, whose cytosolic concentration is in turn controlled by the synergy of diguanylate cyclases and diesterases (Römling et al., 2013).

The BcsA–B complex contains a translocating cellulose polymer that spans the distance from the GT domain to the periplasmic BcsA–B interface. In the c-di-GMP activated structure, the polymer's acceptor terminus rests at the entrance to the TM channel, one glucose unit further into the pore compared to its position in the absence of c-di-GMP (Morgan et al., 2013). Thus, while our previously reported structure likely represents a

state post glycosyl transfer but prior to translocation, the c-di-GMP-activated BcsA–B structure is consistent with a state after polymer translocation. Cellulose translocation may be accomplished by BcsA's finger helix, which hydrogen bonds with the acceptor glucose and pivots towards the TM channel entrance in the c-di-GMP-activated complex. In this position, Asp343 of the finger helix is at an ideal distance to facilitate catalysis. Perhaps the finger helix returns to the “down” position after glycosyl transfer to interact with the new polymer terminus. A similar mechanism involving a flexible loop or helical domain has been postulated for the processive translocation of unfolded polypeptide chains (Erlandson et al., 2008; Martin et al., 2008).

C-di-GMP stimulates the biosynthesis of several extracellular polysaccharides important for biofilm formation, including alginate and PNAG (Itoh et al., 2008; Rehm, 2009; Steiner et al., 2013). While the mechanism for activating PNAG biosynthesis most likely differs from BcsA (Steiner et al., 2013), alginate and cellulose synthases share a strikingly similar organization (Keiski et al., 2010). Alginate is a major component of *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients (Keiski et al., 2010; Hay et al., 2013). In contrast to BcsA–B, the alginate synthase's c-di-GMP-binding PilZ domain is located at the intracellular N terminus of Alg44, the non-catalytic subunit that resembles BcsB and likely interacts with the catalytic Alg8 subunit. Thus, c-di-GMP could exert control by a similar mechanism in alginate synthase as revealed for bacterial cellulose synthase.

Our analyses provide the first insights into how enzymatic functions can be modulated by c-di-GMP. A detailed mechanistic characterization of this bacterial signaling system is required for the development of novel anti-microbial therapeutics.

Accession codes. Atomic coordinates and structure factors for the UDP-free and UDP-bound complexes have been deposited at the Protein Data Bank under accession number 4P02 and 4P00, respectively.

Acknowledgments

We are grateful to O. Pornillos, D. Cosgrove, J. Casanova, 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 Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science, under 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). J.L.W.M. is supported by a National Science Foundation Graduate Research Fellowship, Grant No. DGE-1315231. The research was supported by NIH grant R01GM101001 (J.Z.).

Author contributions

J.Z. designed the experiments. J.L.W.M. crystallized the BcsA–B complex and solved the structures and performed biochemical experiments. J.L.W.M., J.T.M. and J.Z. analyzed the data and wrote the manuscript.

Table 1

Chapter 2 crystallographic data collection and refinement statistics

	UDP-free	UDP-bound
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	67.6, 214.7, 220.4	67.5, 216.8, 219.6
α , β , γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	34.75–2.65 (2.70–2.65) *	49.62–3.20 (3.30–3.20)
<i>R</i> _{pim}	0.056 (0.577)	0.079 (0.499)
CC _{1/2} [^]	0.995 (0.553)	0.992 (0.566)
Mean <i>I</i> / σI	7.8 (1.3)	8.7 (2.1)
Completeness (%)	92.3 (94.3)	98.7 (99.5)
Redundancy	4.5 (4.1)	4.7 (4.7)
Refinement		
Resolution (Å)	20–2.65	49–3.2
No. reflections		
Total	85,948	53,241
<i>R</i> _{free}	4,307	2,714
<i>R</i> _{work} / <i>R</i> _{free}	19.9/23.0	20.6/23.8
No. atoms		
Protein	10,673	10,709
β -1,4 glucan	187	187
c-di-GMP	92	92
UDP	–	25
Mg ²⁺	1	1
Lipids	89	58
<i>B</i> -factors		
Chain A	74.5	74.6
Chain B	66.3	68.3
Chain D	77.2	91.3
β -1,4 glucan	78.1	81.2
c-di-GMP-A	59.2	61.5
c-di-GMP-B	64.1	58.8
UDP	–	80.8
Lipids	85.6	113.6
R.m.s deviations		
Bond lengths (Å)	0.003	0.002
Bond angles (°)	0.754	0.760

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

[^] Correlation between intensities from random half-data sets (Karplus and Diederichs, 2012).

2.4 Methods

2.4.1 Protein purification

BcsA–B was purified as previously described (Morgan et al., 2013) with the exception that gel filtration was carried out in 20 mM Tris pH 7.5, 100 mM NaCl, 5 mM MgCl₂, 5 mM cellobiose, 10% glycerol, 5 mM N,N-Dimethyl-N-dodecylamine N-oxide (LDAO), and 0.3 mM LysoFosCholine Ether 12 (LFCE12), (GF buffer). Peak fractions containing BcsA and BcsB were collected and concentrated to ~10 mg/ml and spun at 180,000 *g* for 15 min at 4°C. Bicelles were prepared by mixing 250 µl water with 100 mg of 1,2-dimyristoyl-sn-glycero-3-phosphocholine:1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine:3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate, (DMPC:POPE:CHAPS) at a molar ratio of ~2.34:0.05:1. The concentrated protein was mixed with the bicelles at a 4:1 (v:v) ratio and allowed to equilibrate on ice for at least 1 hour. 2 mM UDP and 1 mM c-di-GMP were added before incubating on ice overnight.

2.4.2 Crystallization

Crystals belonging to space group P2₁2₁2₁ were grown by sitting-drop vapor diffusion at 30°C in 1.65-1.9 M sodium acetate and 100 mM sodium citrate, pH 3–3.5, (final pH≈5.6) at a 1:1 (v:v) ratio of protein/bicelle and well solution. Crystals appeared within 3 days and reached their final size within 14 days. For the UDP-free structure, cryo-protection and dilution of UDP were achieved by gradual addition of a solution containing 20 mM sodium citrate pH 3, 100 mM NaCl, 5 mM MgCl₂, 1.95 M sodium acetate, 20% glycerol, and 20 % bicelles to the drop. For the UDP-bound structure, cryo-protection and soaking with UDP were achieved by gradual addition of a solution containing 20 mM sodium

citrate pH 3.5, 50 mM MgCl₂, 1.95 M sodium acetate, 20% glycerol, 12% bicelles, and 10 mM UDP to the drop. Crystals were flash-cooled in liquid nitrogen for data collection.

2.4.3 Data Collection and Processing

Diffraction data were collected at 100K at a wavelength of 1.0Å. Data in the absence of UDP were collected at the Advanced Photon Source SER-CAT beamline 22-ID and data in the presence of UDP were collected at GM/CA-CAT beamline 23-ID. The data were integrated using Mosflm (Leslie, 2006) and scaled in Aimless as part of the CCP4 program suite (Collaborative Computational Project, 1994).

2.4.4 Structure Determination

Initial phases for the UDP-free structure were determined by molecular replacement (MR) in Phaser (McCoy et al., 2007) using ligand-free pdb 4HG6 with BcsA residues 499–512 (gating loop) and 574–758 (C terminus) truncated as a search model. MR-phases for the UDP-bound structure were determined in MOLREP (Vagin and Teplyakov, 2010) using the UDP-free structure without the gating loop as search model.

The models were refined by rigid body and restrained refinement in Refmac5 (Collaborative Computational Project, 1994) as well as simulated annealing in Phenix (Adams et al., 2010). Phases were improved using density modification in Parrot (Cowtan and Zhang, 1999; Cowtan, 2010), and model building was performed in Coot (Emsley and Cowtan, 2004). Iterative rounds of model building, refinement, and density

modification resulted in a map of sufficient quality to place the missing domains and ligands. In order to minimize model bias, simulated annealing composite omit maps, prime and switch maps, and kicked maps (Terwilliger, 2004; Adams et al., 2010) were calculated and evaluated throughout the model building process. Additionally, TLS parameters determined from the TLSMD server (Painter and Merritt, 2006b) were utilized in later rounds of refinement. The model contains residues 13–740 of BcsA and residue 54–720 of BcsB. Residue 532–543 of BcsB are disordered as previously observed (Morgan et al., 2013) and were omitted from the model. A 10 residue long unidentified peptide likely belonging to either the extended N terminus of BcsB or the C terminus of BcsA is sandwiched by BcsB's flavodoxin-like domain 2 and carbohydrate-binding domain 2 (Morgan et al., 2013) between crystallographic symmetry mates. This peptide was modeled as a poly-alanine with chain identifier "D". The UDP-free model contains 5 partially ordered lipids, one modeled as 1,2-diacyl-sn-glycero-3-phosphoethanolamine and 4 modeled as 1,2-diacyl-sn-glycero-3-phosphocholine. Two of these lipids were also observed in the UDP-bound structure.

The UDP-free BcsA–B structure was refined to an R/R_{free} of 19.9/23.0. 95.8% of residues lie in the favored regions of the Ramachandran plot with no outliers. The UDP-bound complex structure was refined to an R/R_{free} of 20.6/23.8. 95.1% of residues lie in the favored region on a Ramachandran plot with 0.7% outliers. Figures were prepared in Pymol (The PYMOL Molecular Graphics System. *DeLano Scientific*, 2000) and the solvent accessible surface analysis was performed in HOLLOW (Ho and Gruswitz, 2008). Crystallographic software support is provided by SBGrid (Morin et al., 2013).

2.4.5 Inverted Membrane Vesicle Preparation

Inverted membrane vesicles (IMVs) containing wild type BcsB and the indicated BcsA mutants were prepared as previously described (Hubbard et al., 2012). Control IMVs were prepared from *E. coli* transformed with an empty pETDuet vector. In brief, the constructs were expressed as described (Morgan et al., 2013) and the cells were resuspended in RB buffer containing 20 mM sodium phosphate pH 7.3, 100 mM NaCl, and 10% glycerol using 20 ml RB per cell pellet from a 1 L culture. The cells were lysed in a microfluidizer and spun at 12,000 g for 20 min to clear the cell debris. The supernatant was applied on the surface of a 1.8 M sucrose cushion and spun at 150,000 g for 2 hr at 4°C. IMVs were harvested, diluted 3-fold in RB, and spun at 150,000 g overnight. The pellet from a 3 L culture was resuspended in 1 ml RB, homogenized in a dounce and stored in aliquots at –80°C.

2.4.6 Proteoliposome Preparation

Purified BcsA–B complex containing the indicated mutations were reconstituted into proteoliposomes (PLs) as previously described (Omadjela et al., 2013). Briefly, BcsA–B was purified as described above with the exception that 1 mM LysoFosCholine Ether 14 (LFCE14) was used instead of LDAO. The protein was concentrated to 5 µM, incubated with 4 mg/ml *E. coli* total lipid extract (diluted from a 20 mg/ml stock solution in 40 mM LDAO) and allowed to equilibrate on ice for at least 20 min. Bio-Beads (BioRad) were added, and the solution was rotated until it became turbid, indicating the formation of PLs. The samples were then aliquoted, snap-frozen in liquid nitrogen, and stored at –

80°C. The final protein concentration for all mutants was determined by UV absorbance and SDS-PAGE followed by Coomassie staining.

2.4.7 *In vitro* Cellulose Synthesis Assay

IMVs or PLs were added to a solution containing 20 mM sodium phosphate, 100 mM NaCl, 20 mM MgCl₂, 5 mM UDP-glucose, 12.5 µCi/ml UDP-[³H]-glucose as well as 30µM c-di-GMP unless indicated otherwise. The reaction was incubated at 37°C for 45 min with shaking at 350 rpm. 2% SDS was added to terminate the reaction and dissolve the vesicles. The mixture was then spun at 21,000 g for at least 20 min to pellet the insoluble cellulose. The supernatant was carefully removed, and the pellet was resuspended in 50 mM Tris pH 7.5 and 100 mM NaCl and spotted on Whatman 3 mm grid paper. The product was purified in 60% ethanol by descending paper chromatography, with the insoluble cellulose remaining at the origin, and quantified by scintillation counting (Hubbard et al., 2012). All measurements were performed at least in triplicate and error bars represent standard deviations.

2.4.8 Western Analysis

10 µL IMVs were analyzed by SDS-PAGE and transferred to a nitrocellulose membrane using a BioRad Mini-Transfer Cell according to the manufacturer's specifications. The nitrocellulose membrane was blocked in 5% milk/TBS-Tween solution for 30 min and incubated overnight with a mouse anti-penta-His (Qiagen) antibody. The membranes were washed three times in 5% milk/TBS-Tween before incubating with an IRDye800-

conjugated anti-mouse secondary antibody (Rockland) for 45 min at RT. After washing, the membranes were scanned on an Odyssey Infrared Imager (Licor).

2.4.9 Enzyme-coupled activity assay

Enzyme-coupled kinetic assays were carried out as previously described (Omadjela et al., 2013) with the exception that the protein was reconstituted into PLs instead of nanodiscs, the experiments were performed in 150 μ l reaction volume in 96-well flat bottom Microplates (Greiner), and 3 mM UDP-Glc was used.

2.4.10 Isothermal Titration Calorimetry

The protein was purified in 1 mM LFCE14 as described above. Measurements were carried out at 25° C in a MicroCal iTC200 system (GE Healthcare) with 250 μ l of BcsA–B in the cell at 9.3 μ M for WT and 11 μ M for the BcsA–Arg580Ala complex and 400 μ M c-di-GMP in the syringe. An initial 0.5 μ l injection was followed by 39 1 μ l injections spaced 180s apart with stirring at 700 rpm. The data were fit using Origin 7.0 as provided by the manufacturer.

Chapter 3: Observing cellulose biosynthesis and membrane translocation *in crystallo*

Jacob L.W. Morgan^{a*}, Joshua T. McNamara^{a*}, Michael Fischer^{b,c}, Jamie Rich^b, Hong-Ming Chen^b, Stephen G. Withers^b and Jochen Zimmer^{a#}

* These authors contributed equally to this work.

^a University of Virginia School of Medicine, Center for Membrane Biology, Molecular Physiology and Biological Physics, 480 Ray C. Hunt Dr., Charlottesville, VA 22908, USA

^b Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, B.C., Canada V6T 1Z1

^c Present Address: Sandoz GmbH, Biochemiestrasse 10, A-6250 Kundl, Tyrol Austria

In this chapter, I will present a paper that we published in *Nature* (Morgan et al., 2016) addressing the mechanism of cellulose translocation by BcsA–B. I began working on this aspect of the project by accident in mid 2014 when I soaked some c-di-GMP-bound BcsA–B crystals with UDP-Glc and noticed that the electron density corresponding to the cellulose polymer was extended by what looked like a single glucose unit. I then tried to see if I could visualize translocation by getting the electron density for the newly-added sugar to disappear. I was responsible for conceiving and performing all *in crystallo* assays, determining the structures of the ‘product-bound’ state and the ‘pre-translocation’ state, performing all functional assays, and editing/writing the manuscript. Joshua McNamara, a co-first-author, determined the structure of the ‘substrate-bound’ state, generated the cysteine mutants and screened them for changes in activity under redox conditions, identified and crystallized a disulfide tethered BcsA-B complex, and

performed functional assays. Michael Fisher, Jamie Rich, and Hong-Ming Chen were members of Stephen Withers' lab and were responsible for synthesizing the UDP-Glc derivatives that we used in the paper. The paper was published in 2016.

3.1 Introduction

Cellulose is an abundant structural cell component produced by many organisms, including bacteria, vascular plants and animals (Kimura et al., 2001; Römmling, 2002; Keegstra, 2010; Serra et al., 2013). It is a linear polymer of glucose molecules joined between their C1 and C4 carbons (Nishiyama et al., 2003). Cellulose is synthesized by membrane-integrated glycosyltransferases (GTs) that contain 6 to 8 transmembrane helices (TMHs) as well as an intracellular catalytic GT domain (McNamara et al., 2015). These enzymes polymerize UDP-activated glucose (UDP-Glc) (Brown et al., 2012; Omadjela et al., 2013) into chains thousands of glucose units long (Somerville, 2006) and translocate the polymer across the plasma membrane, through a pore formed by their own TM region (Morgan et al., 2013).

Cellulose is also a common biofilm component (Bokranz et al., 2005; Serra et al., 2013) where it is synthesized and secreted via an inner and, in gram-negative bacteria, outer membrane-spanning cellulose synthase complex (Römmling, 2002). At the inner membrane, the catalytic BcsA and membrane-anchored, periplasmic BcsB subunits form a complex sufficient to synthesize and translocate cellulose (Omadjela et al., 2013), while transport across the outer membrane likely occurs through the BcsC subunit (Keiski et al., 2010; Whitney et al., 2011).

Processive GTs, including chitin, alginate and cellulose synthases, transfer the glycosyl moiety from a nucleotide-activated sugar (donor) to a specific hydroxyl group of the growing polysaccharide chain (acceptor) by a nucleophilic S_N2-like substitution reaction

(Lairson et al., 2008). Thereby forming an elongated polymer and nucleoside diphosphate as reaction products. A processive mechanism requires that the elongated polymer is translocated after each glycosyl transfer, such that the polymer's newly added sugar unit becomes the acceptor in a subsequent reaction. Because all known processive GTs are TM channel-forming enzymes (Merzendorfer, 2006; Rehm, 2009; Hubbard et al., 2012), the translocation of the polymer into the TM channel between catalytic steps also gives rise to secretion.

Previous structural and functional analyses of the *Rhodobacter sphaeroides* BcsA-B complex containing a nascent cellulose polymer revealed the architecture of the active site, its close association with the TM channel, as well as the coordination of cellulose within the channel (Morgan et al., 2013). In bacteria, cellulose biosynthesis is activated by the signaling molecule cyclic-di-GMP (c-di-GMP) (Ross et al., 1987), a potent biofilm inducer and allosteric activator of BcsA (Römling et al., 2013). Binding of the activator to BcsA's C-terminal PilZ domain allows a "gating loop" to either insert into the catalytic pocket during substrate binding or to retract from it to release the UDP product (Morgan et al., 2014).

Cellulose synthases contain a short helix within the GT domain, termed "finger helix" (Morgan et al., 2014). The N terminus of the finger helix contacts the polymer's acceptor glucose via an invariant "TED" motif, of which the Asp likely facilitates the deprotonation of the acceptor C4 hydroxyl during catalysis (Morgan et al., 2013; 2014).

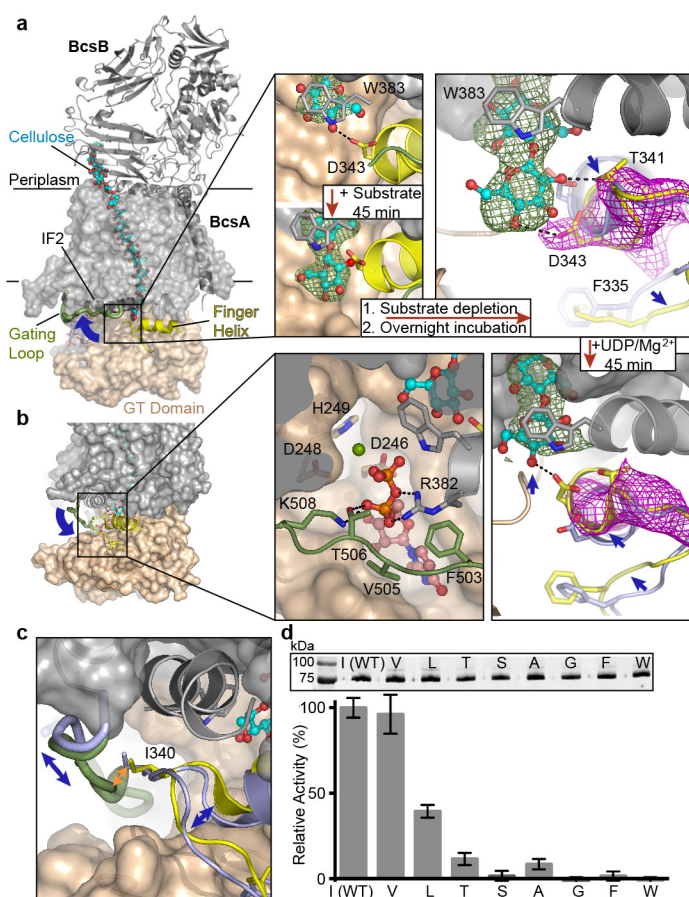


Figure 37 *In crystallo* cellulose biosynthesis. BcsA is shown as a surface with the TM and GT regions colored gray and beige. BcsB is shown as a gray cartoon. The cellulose polymer is shown as cyan sticks, BcsA's finger helix and gating loop are shown as a cartoon colored yellow and green, respectively. Trp383 is represented in gray sticks as a marker for the TM channel entrance. **(a)**, Left Panel: Organization of the BcsA-B complex and formation of a channel for the translocating polymer. Right Panel: The pre-translocation state of BcsA. Unbiased Sigma-A weighted Fo-Fc difference electron densities contoured at 4 and 3σ are shown as green and magenta meshes for the nascent cellulose chain and finger helix, respectively. **(b)** Translocation of cellulose. Crystals described in (a) were subsequently soaked with UDP/Mg²⁺. UDP is shown in sticks and colored violet for the carbon atoms and Mg²⁺ is shown as a green sphere. **(c)** Insertion of the gating loop into the active site is incompatible with the 'down' position of the finger helix, likely due to a clash between Ile340 and the gating loop's backbone. The retracted gating loop and finger helix in the 'down' position are both shown as blue cartoons. **(d)** Cellulose biosynthesis by BcsA I340 mutants. Ile340 was replaced with the indicated residues and *in vitro* cellulose biosynthesis was performed in inverted membrane vesicles (IMVs) as described⁷. All activities are represented relative to the WT activity. All experiments were performed at least in triplicate and error bars represent standard deviations from the means. Inset: Western analysis of the IMVs used, showing equal expression levels of all BcsA mutants.

Crystal structures of the catalytically inactive “resting” state of BcsA-B (in the absence of c-di-GMP) (Morgan et al., 2013) and a c-di-GMP-activated structure (Morgan et al., 2014) provided important insights into the architecture and function of processive GTs. Here, we used *in crystallo* enzymology to obtain structural snapshots of a complete cellulose biosynthesis reaction cycle, providing structures of substrate- and product-bound states, and delineating the mechanism by which the elongated glucan is translocated into BcsA's TM channel.

3.2 Results

3.2.1 Cellulose synthase elongates the cellulose chain one glucose unit at a time

The previously determined c-di-GMP-activated BcsA-B structure (Morgan et al., 2014) contains a nascent cellulose polymer 18 glucose molecules long whose non-reducing terminal glucose unit rests at the entrance to BcsA's TM channel, which is marked by the invariant Trp383 of the "QxxRW" motif (Saxena et al., 2001; Morgan et al., 2014). In

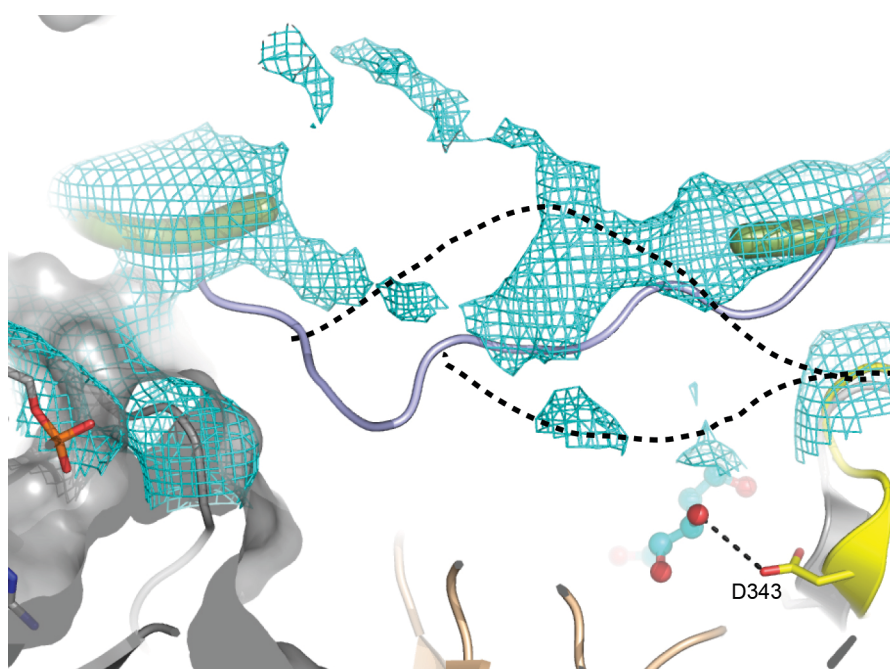


Figure 38 Conformational flexibility of the gating loop after cellulose extension. Unbiased Sigma-A weighted Fo-Fc difference electron density of the gating loop in the pre-translocation state contoured at 2σ . The ordered part of the gating loop is shown as a thick green ribbon and two alternative backbone positions are indicated by a black dashed line. The position of the gating loop in the inserted state in the presence of a UDP molecule as observed in pdb entry 4P00 is shown as a cartoon representation colored blue.

this state, BcsA's finger helix is in an 'up' position where it points toward the entrance of the TM channel, thereby positioning

Asp343, the putative catalytic base, near the C4 hydroxyl of the

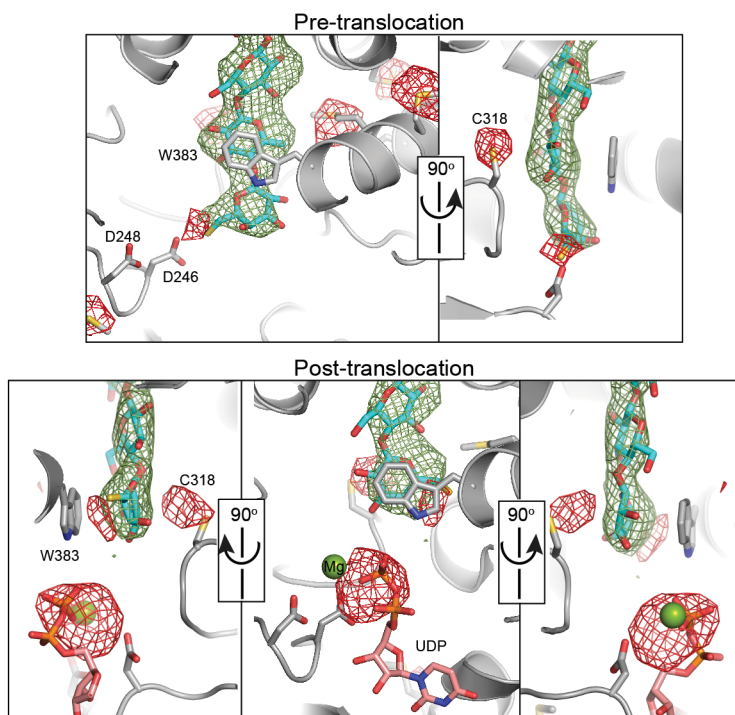
polymer's terminal sugar (**Fig. 37a**). Because BcsA's active site is empty and the gating loop is retracted from it, this structure represents a state in which the enzyme is poised to initiate a new cycle of chain elongation, hereafter referred to as the "post-translocation state".

Strikingly BcsA-B is catalytically active *in crystallo*. Incubating BcsA-B crystals with UDP-glucose in the absence of Mg^{2+} (to slow down the reaction) results in the extension of the polymer's electron density by one glucose unit (**Fig. 37a**). This demonstrates that cellulose elongation occurs via a stepwise addition of glucose units and that Trp383 of the QxxRW motif indeed forms the acceptor-binding site. The elongated polymer points straight into the catalytic pocket, similar to its position in the recently determined resting state of BcsA (Morgan et al., 2013).

3.2.2 BcsA's finger helix resets in response to polymer extension

Processive cellulose biosynthesis requires that the elongated polymer is translocated after each elongation cycle and the above described *in crystallo* cellulose extension demonstrates that glycosyl transfer and polymer translocation are separate steps.

To identify whether the extended cellulose translocates spontaneously over time, we extended the polymer *in crystallo* as described above, then diluted the substrate 65-fold, and incubated the crystals overnight before harvesting. Under these conditions, the density for the extended polymer continues to protrude into the catalytic pocket, suggesting that this state is stable in the absence of substrate (**Fig. 37a**). Strikingly, after extending the cellulose polymer, BcsA's finger helix shifts to a 'down' position, such that Thr341 and Asp343 of its TED motif again form hydrogen bonds with the polymer's terminal glucose unit (**Fig. 37a**).



The finger helix movement is accompanied by the retraction of a small preceding loop with Phe335 at its tip from a hydrophobic pocket underneath the active site (**Fig. 37a**). The resulting "pre-translocation" state of BcsA-B contains an empty catalytic pocket, a gating loop retracted from the active site (**Fig. 38**), an extended polymer, and a downward pointing finger helix (**Fig. 37a**).

3.2.3 Cellulose translocation by a coordinated movement of BcsA's gating loop and finger helix

The observed downward movement of BcsA's finger helix in response to cellulose elongation suggests its role in translocation. If crystals containing the above described pre-translocation step are then soaked with UDP/Mg²⁺, UDP binds and the gating loop inserts into the catalytic pocket (**Fig. 37b**). Additionally, the finger helix returns to the

Figure 39 *In crystallo* translocation of a 6-thio-galactose-containing cellulose polymer. The position of the 6-thio-galactose group at the polymer's non-reducing end was determined after polymer extension (upper panel) and upon subsequent incubation with UDP/Mg²⁺ (lower panel) in an anomalous difference Fourier electron density (DANO) map. DANO peaks detected at a wavelength of 1.74 Å are shown as a red mesh contoured at 3.5σ. Unbiased Sigma-A weighted Fo-Fc difference electron density for the cellulose polymer is shown as a green mesh contoured at 4σ. The cellulose polymer was extended and translocated as described in Fig. 34 with the exception that UDP-6-thio-galactose was used as substrate and Mg²⁺ was included during the initial soaking step. The extended DANO peak around Cys318 in the post-translocation state might arise from overlapping peaks originating from Cys318 and the thio-Gal unit in an opposite orientation. All Cys and Met residues close to BcsA's active site are shown as sticks. UDP is shown as sticks in violet for its carbon atoms

'up' position and the density for the polymer's newly-added glucose unit disappears (**Fig. 37b**) suggesting its translocation into the channel.

To confirm that the elongated polymer is indeed translocated *in crystallo*, we extended the polymer by one unit with the chain-terminating analogue 6-thio-galactose, whose location can be unambiguously determined based on anomalous X-ray scattering of its sulfur atom. After polymer extension, the 6-thio-galactosyl moiety sits inside BcsA's catalytic pocket; then upon UDP/Mg²⁺ binding and gating loop insertion into the active site, the density of the newly-added sugar disappears, and the thio-galactosyl unit moves into the TM channel next to Trp383 (**Fig. 39**), thereby confirming the genuine translocation of cellulose *in crystallo*.

The ability of UDP/Mg²⁺ to induce translocation brings about the question of why translocation doesn't occur immediately after glycosyl transfer when the UDP/Mg²⁺ product is bound at the active site and the gating loop is inserted. All states of BcsA observed thus far either show its gating loop inserted into the active site and the finger helix in the 'up' position (**Fig. 37b**), or the gating loop retracted from the active site and the finger helix in the 'down' position, if the cellulose polymer is extended (**Fig. 37a**) (Morgan et al., 2013; 2014). This suggests that the finger helix cannot move to the 'down' position unless the gating loop is retracted from the active site and that in turn gating loop insertion could induce the upward movement of the finger helix. This coupled movement is likely due to steric clashes between the side chain of Ile340, preceding the TED motif of the finger helix, and the gating loop's backbone (**Fig 37c**). Ile340 is

primarily conserved among bacterial cellulose synthases, eukaryotic enzymes usually contain a valine or occasionally a leucine residue at this position, which could perform a similar function. Indeed, BcsA carrying a Val at position 340 shows indistinguishable catalytic activity compared to the wild type enzyme, while a Ile to Leu substitution

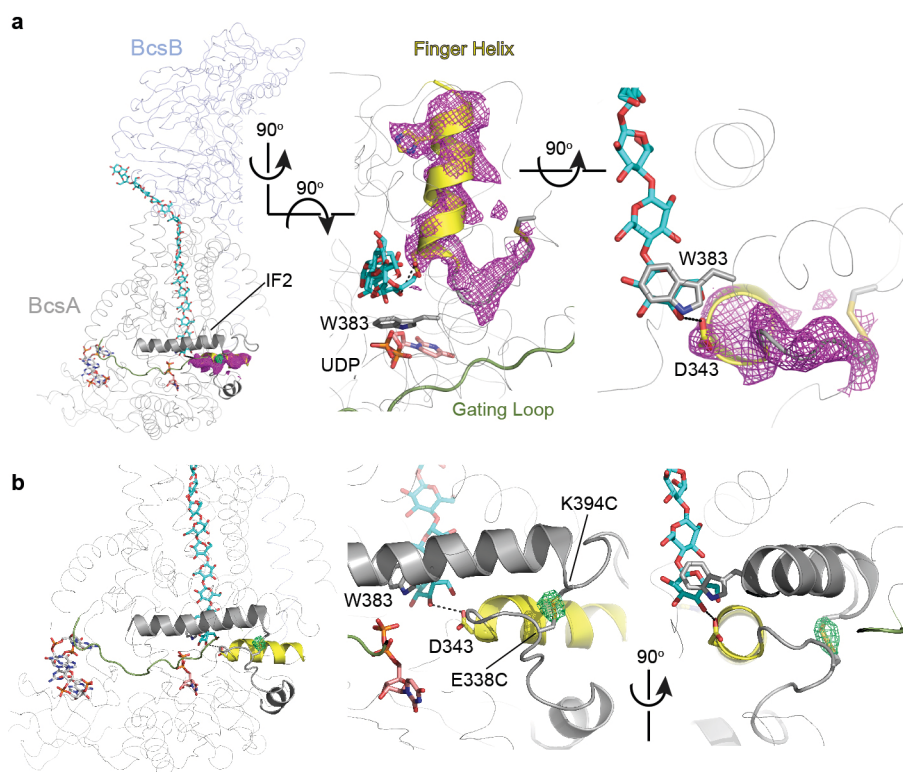


Figure 40 Position of the disulfide-tethered finger helix. The BcsA-2C-B complex was crystallized as described for wild type BcsA-B. **(a)** Unbiased Sigma-A weighted Fo-Fc difference electron density of BcsA's finger helix contoured at 4σ (magenta mesh). Cellulose and BcsA's Trp383 at the entrance to the TM channel are shown as sticks in cyan and gray for their carbon atoms, respectively. The finger helix and IF2 are shown as cartoon helices colored yellow and gray, respectively. **(b)** The finger helix-tethered BcsA-B complex was refined in a resolution range from 34 to 3.2Å to a final R/R_{free} of 19.9/23.9% in Phenix_refine with Ala residues at positions 338 and 394 of BcsA. A strong difference electron density peak indicates the position of the omitted disulfide bond in a Sigma-A weighted Fo-Fc difference electron density map, (green mesh, contoured at 2σ).

reduces the apparent activity by about 50% (**Fig. 37d**). Thr, Ala, Ser, Phe or Trp, however, support only low or background activities, suggesting that gating loop to finger

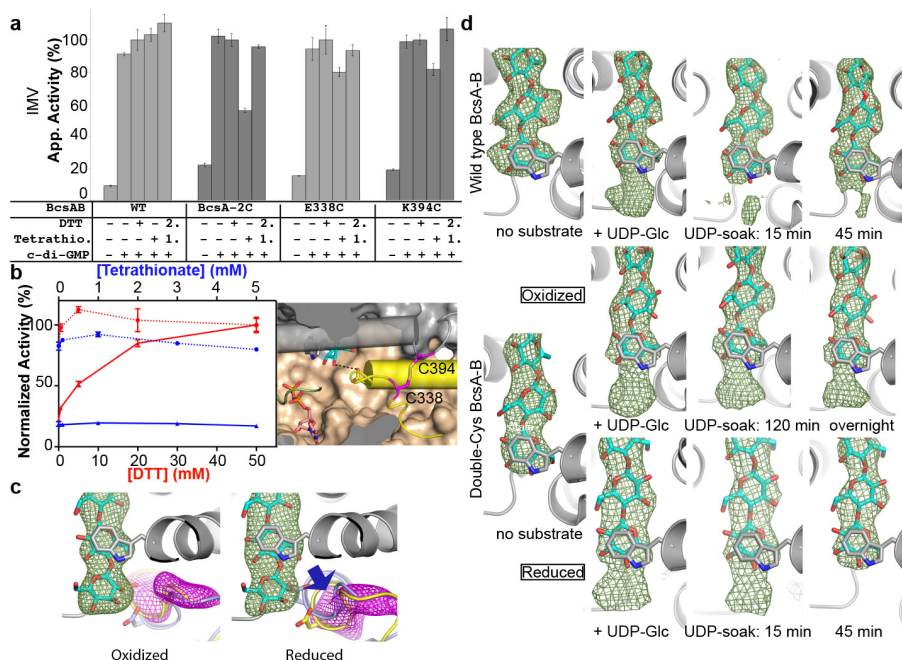


Figure 41 Movement of BcsA's finger helix is essential for cellulose translocation. *In vitro* cellulose formation by wild type and mutant BcsA-B complexes. **(a)** Activity in IMVs under reducing (DTT) and oxidizing (sodium tetrathionate) conditions. For each mutant tested, the apparent activity was normalized to its activity in the presence of 50 mM DTT and (1.) and (2.) indicate the order of DTT and tetrathionate addition. **(b)** DTT and tetrathionate titrations are shown in red and blue, respectively, for wild type BcsA (dashed lines) and BcsA-2C (solid lines). All experiments were performed at least in triplicate and error bars represent the deviations from the means. Right panel: Location of the engineered disulfide bond in BcsA-2C, colored as in Fig. 34. **(c)** Comparison of BcsA-2C finger helix positions following polymer extension under oxidizing and reducing conditions. Upon cellulose elongation, crystals were oxidized or reduced and incubated without substrate for 16 hrs or 6 hrs, respectively. **(d)** Comparison of *in crystallo* cellulose translocation in WT and BcsA-2C. The cellulose polymer was extended, then translocation was initiated as described in Fig. 34, and crystals were harvested after the indicated incubation periods. In all panels, the unbiased electron densities for the glucan and finger helix are contoured and colored as described in Fig. 34. Trp383 of the acceptor-binding site is shown in gray.

helix
coupling
requires a
fairly rigid,
hydrophobic
residue at
position 340
(**Fig. 37d**).

**3.2.4 The
movement
of the finger
helix is
required for
cellulose
translocation**

The
conformational changes

of BcsA described above suggest that its finger helix moves up and down during cellulose translocation. To test this hypothesis, we engineered double Cys BcsA mutants expected to crosslink the finger helix to an amphipathic helix (IF2) above the GT domain (Morgan et al., 2013) (**Fig. 37a and 40**), and screened those mutants for changes in catalytic activity upon oxidation. Introducing Cys residues at positions 338 near the N terminus of the finger helix and 394 at the C-terminal end of IF2 (hereafter referred to as BcsA-2C) results in attenuated catalytic activity compared to the wild type enzyme or single-Cys mutants in inverted membrane vesicles (IMVs). Full activity, however, can be restored upon addition of excess DTT (**Fig 41a**). Upon purification of BcsA-2C and reconstitution into proteoliposomes, its catalytic activity further decreases to ~20% under non-reducing conditions (**Fig. 41b**), likely due to complete disulfide bond formation during purification. Addition of the oxidizing reagent tetrathionate does not further decrease the enzyme's apparent activity, yet the catalytic activity robustly recovers with increasing DTT concentrations. The crystal structure of this BcsA-2C mutant reveals that the disulfide bond forms when the finger helix is in the 'up' position (**Fig. 41b and Fig. 40a and b**), similar to its position in the post-translocation state (Morgan et al., 2014).

Additionally, the ability to observe cellulose elongation and translocation *in crystallo* allows us to further delineate how the engineered disulfide bond affects BcsA's activity. Soaking BcsA-2C crystals with substrate leads to polymer extension as observed for wild type BcsA, demonstrating that glycosyl transfer is not abolished by the mutations (**Fig. 41c and d**). Subsequently, if those crystals are then incubated overnight under oxidizing conditions, the finger helix remains in the 'up' position, revealing that the cross-link

prevents the finger helix from resetting to the ‘down’ position (**Fig. 41c**). Binding of UDP/Mg²⁺ in these crystals, which induces polymer translocation in wild type BcsA within 15 to 45 min, fails to initiate translocation, even after an overnight incubation (**Fig. 41d**). However, reducing the engineered disulfide bond in the BcsA-2C complex restores the capability of the finger helix to move downwards following polymer extension and, most notably, restores polymer translocation (**Fig. 41c and d**), thereby directly correlating the movement of the finger helix with BcsA's ability to translocate the polymer.

3.2.5 The product-bound state

Directly after glycosyl transfer, BcsA contains an elongated glucan plus UDP/Mg²⁺ and an inserted gating loop at the active site, as well as the finger helix in the ‘up’ position. This “product-bound” state is accessible through the BcsA-2C mutant described above.

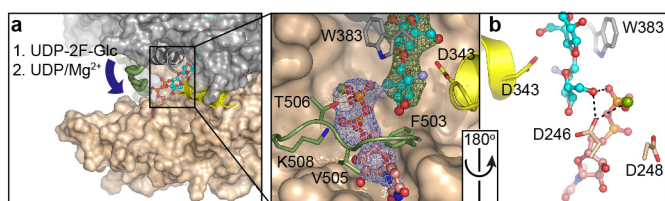


Figure 42 The product-bound state. The product-bound state of BcsA contains an elongated cellulose polymer and an inserted gating loop coordinating UDP/Mg²⁺ at the active site. **(a)** Cellulose was elongated *in crystallo* with a 2-deoxy, 2-fluoro-glucose moiety and UDP/Mg²⁺ was rebound to the active site as described in Fig. 1. Unbiased Sigma-A weighted Fo-Fc difference electron densities contoured at 4 and 4.5 σ are shown for the nascent cellulose polymer and UDP/Mg²⁺ in green and blue, respectively. **(b)** The terminal glucose unit of the extended cellulose polymer forms interactions with the β -phosphate of UDP and Asp246 of the DxD motif. Colors are as in Fig. 37 and fluorine is shown in light blue.

Because the engineered disulfide bond tethers the finger helix, UDP/Mg²⁺ can be

bound to the active site after polymer extension without inducing translocation.

Alternatively, we observed that incorporating 2-fluoro-

substituted glucose into the polymer stabilizes a similar product-bound state. Attempting to trap a “donor-bound” state by using the usually (but not always) unreactive UDP-2-fluoro-glucose as substrate

(Persson et al., 2001; Chan et al., 2013), we observed that the nascent glucan is elongated. However, the subsequent translocation of this polymer is significantly impeded (but not abolished), which may be due to the loss of a hydrogen bond between Thr341 of the finger helix and the polymer's terminal C2 substituent (**Fig. 37a**). Thus, elongating the cellulose polymer with 2-fluoro-glucose in wild type BcsA, then diluting the substrate and soaking in UDP/Mg²⁺ reproduces a similar product-bound state as obtained for the BcsA-2C mutant. Due to higher quality diffraction data (**Table 2**), we discuss the structure obtained with a fluorinated polymer.

The product-bound BcsA structure shows that the catalytic pocket can accommodate both an extended cellulose polymer and UDP/Mg²⁺, suggesting that the newly added glucose unit can align with the polymer before the gating loop retracts from the active site and UDP is released (**Fig. 42**). In this position, the only major interactions of the terminal glucose unit are with UDP's β -phosphate as well as Asp246 of the "DxD" motif (Lairson et al., 2008; McNamara et al., 2015) via its C2 or C6 hydroxyl group (depending on its orientation) (**Fig. 42b**).

Of note, the individual glucose units in cellulose are rotated by approximately 180° relative to their neighbors (Gardner and Blackwell, 1974; Nishiyama et al., 2003). Therefore, during relaxation into the polymer's plane, the newly added glucose moiety

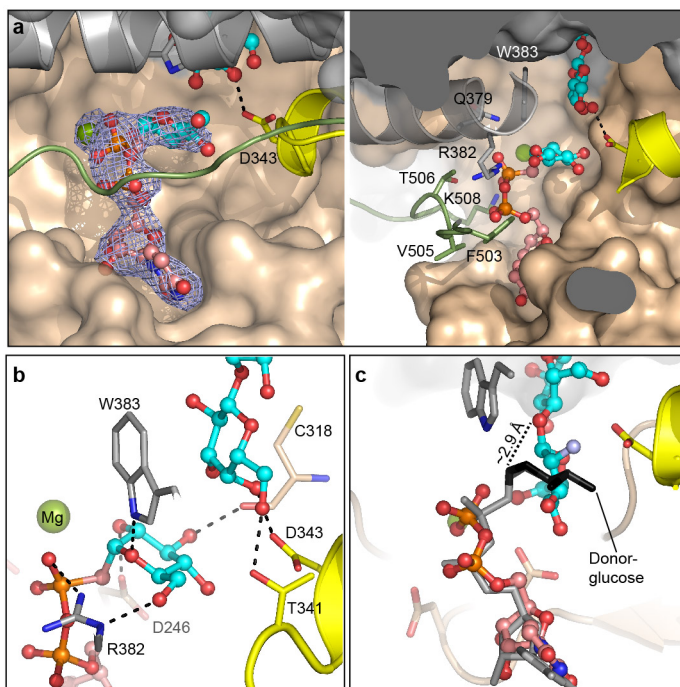


Figure 43 The substrate-bound state. The donor glucose binds in a conserved pocket beneath the acceptor. **(a)** Unbiased Sigma-A weighted Fo-Fc difference electron density of UDP-CH₂-glucose at BcsA's active site, contoured at 4σ. Right panel: Conserved residues of BcsA involved in coordinating the donor glucose are shown as sticks. **(b)** Interactions between the donor sugar moiety and BcsA's D₂₄₆xD, FFC₃₁₈GS, T₃₄₁ED₃₄₃ and QxxR₃₈₂W₃₈₃ motifs. **(c)** Comparison of substrate- and product-bound states. Aligning the substrate's pyrophosphate group with the position of UDP in the product-bound state (Fig. 42) positions the donor's C1 carbon within approximately 2.9 Å of the acceptor's hydroxyl group. The substrate is shown as gray and black sticks for its UDP and glucose moieties, respectively.

must rotate either clock- or counter clockwise to be in

register with the preceding glucose units (Carpita, 2011; Yang et al., 2015). This

alternating rotation is most likely driven by the formation of intramolecular hydrogen

bonds (Carpita, 2011) and could be facilitated by UDP release after glycosyl transfer,

which minimizes steric restrictions at the active site (**Fig. 42**).

3.2.6 The donor glucose binds to a hydrophilic pocket underneath the acceptor

To stabilize the substrate-bound state of BcsA, we synthesized and employed a non-hydrolyzable phosphonate substrate analogue (Martin et al., 1990), (UDP-CH₂-Glc), in which a methylene bridge connects the donor glucose with UDP's β-phosphate.

Additionally, we also capped the cellulose polymer with galactose (see Methods), which cannot be extended due to an axial instead of an equatorial hydroxyl group at its C4 position.

The donor glucose inserts underneath the acceptor into a conserved hydrophilic pocket formed by BcsA's TED, HAKAG and FFCGS motifs (**Fig. 43a**), and the gating loop cooperates with the QxxRW motif to stabilize the substrate. In particular, the gating loop's Phe503 forms cation- π interactions with Arg382 of the QxxRW motif (Dougherty, 1996), which in turn forms a salt bridge with the nucleotide's β -phosphate as well as a hydrogen bond with the donor's C6 hydroxyl (**Fig. 43a and b**). Further, the donor's C3 hydroxyl interacts with the backbone carbonyl of Cys318 of the FFCGS motif and its ring oxygen is in hydrogen bond distance to the N ϵ of Trp383 of the QxxRW motif (**Fig. 43b**). Thus, the recognition of the donor's C3 and C6 hydroxyls and its ring oxygen appears particularly important for substrate selectivity.

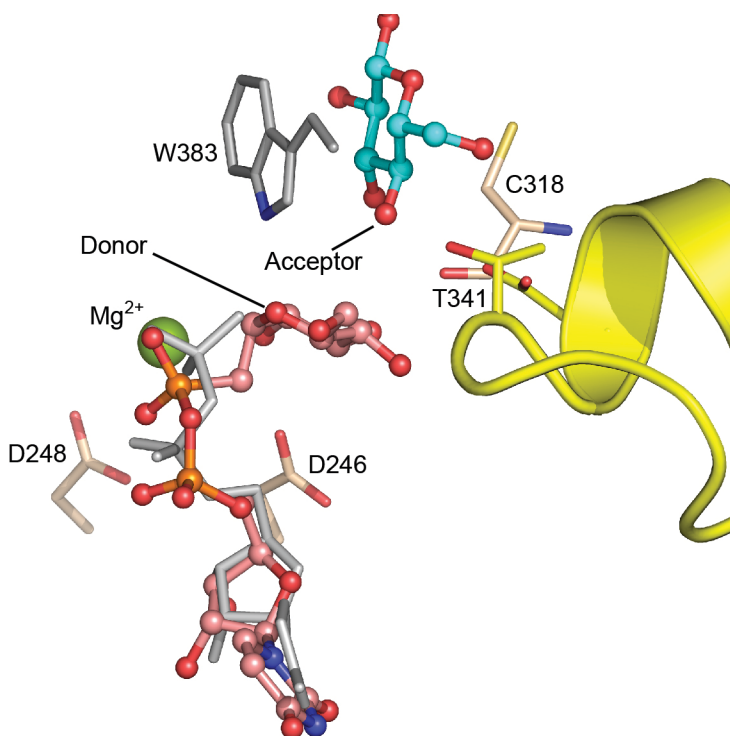


Figure 44 Comparison of the UDP conformation in the substrate and UDP-bound states of BcsA. The substrate-bound BcsA structure was superimposed with pdb entry 4P00 by secondary structure matching in Coot. The substrate is shown as "balls and sticks" in violet for the carbon atoms and the UDP molecule from pdb entry 4P00 is shown as gray sticks. BcsA's finger helix is shown as a yellow cartoon and the cellulose polymer is shown as cyan and red "balls and sticks" as observed in pdb entry 4P00. Magnesium is shown as a green sphere.

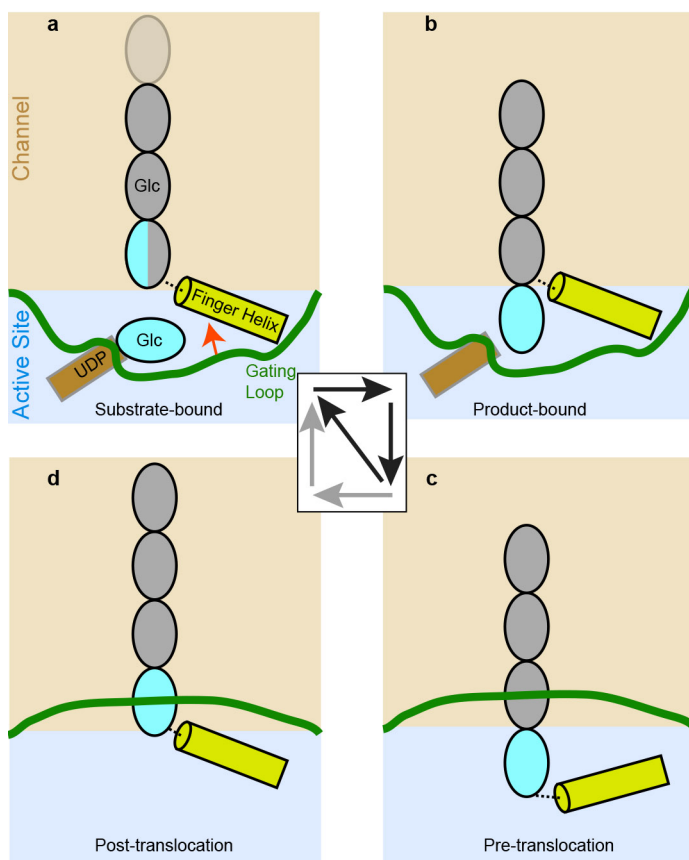


Figure 45 Model of cellulose biosynthesis. Cellulose biosynthesis might start with substrate binding to BcsA when the polymer's terminal glucose unit sits at the acceptor site at the entrance to BcsA's TM channel (a). At this time, the gating loop stabilizes the substrate and the finger helix in the 'up' position (red arrow). Glycosyl transfer generates the product-bound state (b) and retraction of the gating loop and UDP release allows the finger helix to reset to the 'down' position to contact again the polymer's terminal glucose unit (c). Substrate binding to this pre-translocation state and insertion of the gating loop could induce the upward movement of the finger helix and polymer translocation (a) or spontaneous translocation might precede substrate binding via a post-translocation state (d).

In the substrate-bound state, BcsA's finger helix is in

the 'up' conformation and positions Asp343 of the TED motif within 2.5 Å of the acceptor's C4 hydroxyl group, consistent with its likely function as general base during catalysis (**Fig. 43b**). However, the distance between the acceptor and the donor's C1 carbon is about 4.2 Å (assuming glucose instead of galactose as the polymer's terminal sugar), which is likely too far for a direct transfer. We note that the pyrophosphate group of UDP-CH₂-Glc is less deeply inserted into the active site compared to UDP in the product-bound state (**Fig. 44**), perhaps due to the substrate's methylene-bridge and/or the capping of the glucan with galactose. Repositioning UDP-CH₂-Glc according to the UDP conformation in the product-bound state places the donor's C1 carbon within approximately 2.9 Å of the acceptor's C4 hydroxyl, a suitable distance for glycosyl transfer (Vocadlo et al., 2001) (**Fig. 43c**). This distance is likely also maintained when the

acceptor is in the opposite orientation (as is the case for every other glucose unit), due to repositioning of the terminal glucose unit at the active site.

3.3 Discussion

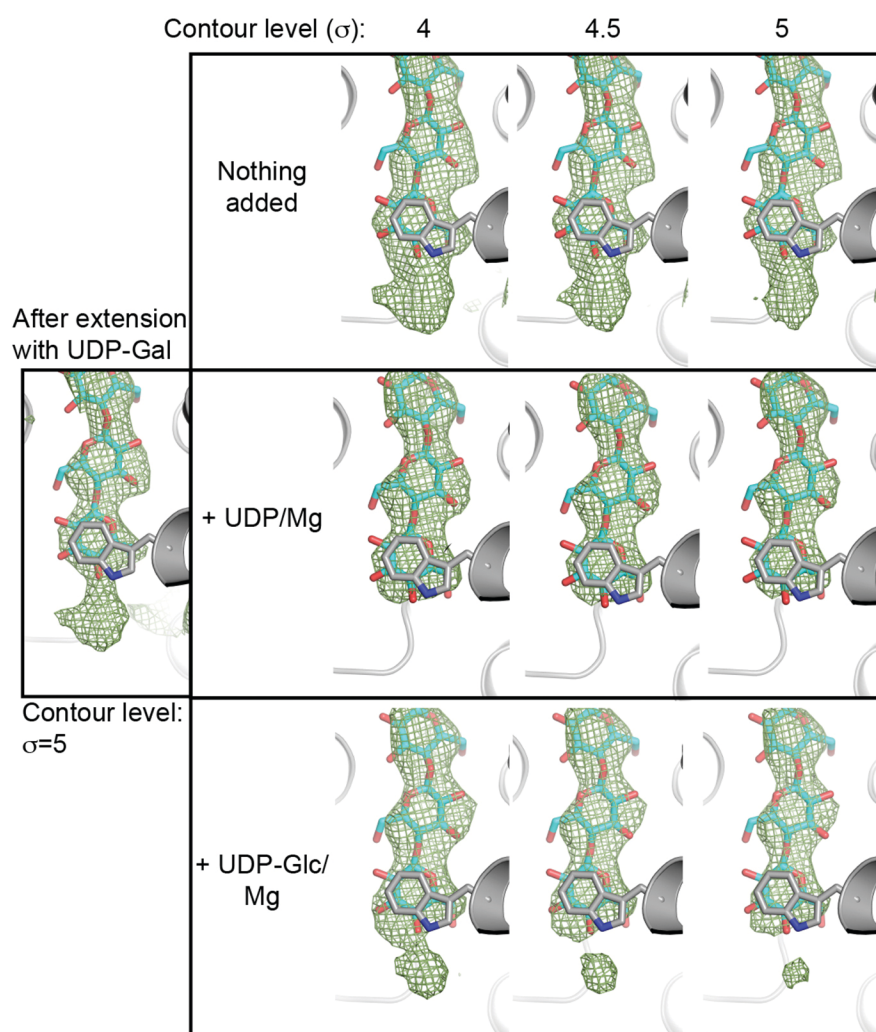


Figure 46 UDP-Glc induced polymer translocation. The nascent cellulose polymer was extended with a chain-terminating galactose residue upon soaking BcsA-B crystals with UDP-Gal. Following dilution of the substrate as described in Fig. 34, crystals were incubated for 150 min either in the absence of a nucleotide or in the presence UDP/Mg²⁺ or UDP-Glc/Mg²⁺, respectively. The unbiased SigmaA-weighted Fo-Fc difference electron density of the nascent polymer (green mesh) is shown at three different contour levels, indicating that UDP-Glc also induces polymer translocation.

3.3.1 Implications

for cellulose biosynthesis and membrane translocation

Upon substrate binding, BcsA's gating loop inserts into the catalytic pocket, thereby positioning the donor glucose for transfer and perhaps also stabilizing the UDP leaving group (Fig. 45). In this state, the acceptor glucose

rests next to Trp383 at the entrance to the TM channel and interacts with the TED motif at the N terminus of the finger helix. After glycosyl transfer, the newly added glucose unit aligns with the polymer and extends into the catalytic pocket next to UDP's pyrophosphate group. In this product-bound state, the gating loop remains inserted into the active site and the finger helix continues to point 'up' as observed in the substrate-bound state. Next, BcsA's gating loop retracts to release UDP and the finger helix resets to the 'down' position to interact again with the polymer's terminal glucose unit. Binding of a new substrate molecule to this pre-translocation state could elicit the translocation of the extended polymer (by an upward movement of the finger helix) through re-insertion of the gating loop into the active site. *In crystallo* translocation experiments with a galactose-capped polymer and UDP-Glc as substrate confirmed that the polymer can indeed be translocated when UDP-Glc/Mg²⁺ binds to the active site (**Fig. 46**). An alternative, perhaps slower, pathway could be the translocation of the polymer prior to substrate binding, facilitated either by random cycles of gating loop insertion and retraction or favorable interactions of the polymer in the extracellular milieu.

How could the opening of the gating loop and resetting of the finger helix be coordinated? The retraction of the gating loop after glycosyl transfer could be facilitated by the rotation of the newly added sugar into the plane of the polymer. Alternatively, it is possible that the Mg²⁺ coordination changes after glycosyl transfer, which in turn could affect the stability of UDP and the gating loop at the active site, as proposed for non-processive galactosyl transferases (Qasba et al., 2008).

BcsA's finger helix is capped at its N terminus by the TED motif, which is invariant among cellulose synthases. Within the motif, Thr341 and Asp343 form hydrogen bonds with the polymer's terminal sugar unit, which may enable the finger helix to exert force on the polymer during the upward movement. Both residues are well-suited for this task: as a β -branched amino acid, the side chain hydroxyl of Thr is sterically restricted, and Asp343 is further rigidified by interactions with its backbone amide proton and that of the following residues (**Fig. 47**).

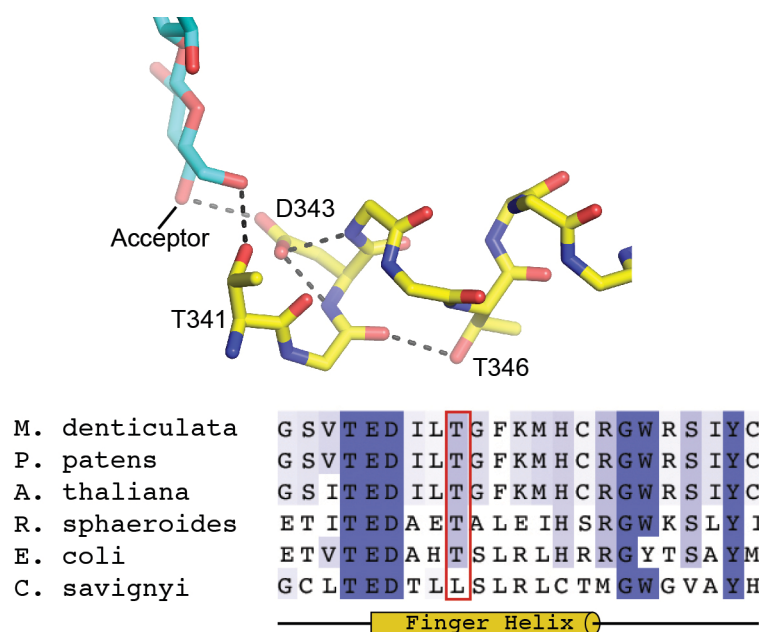


Figure 47 Stabilization of BcsA's finger helix by conserved residues. Top panel: Stick representation of BcsA's finger helix and nascent cellulose polymer shown in yellow and cyan for their carbon atoms. The finger helix is shown as a poly-glycine helix except for the labeled residues. Bottom panel: The finger helix's "TEDxxT" motif is conserved among pro- and eukaryotic cellulose synthases. Finger helix sequences are aligned for *Micrasterias denticulata* Cesa, *Physcomitrella patens* Cesa5, *Arabidopsis thaliana* Cesa8, *Rhodobacter sphaeroides* and *Escherichia coli* BcsA, and *Ciona savignyi* Cesa. The conserved threonine following the TED motif is indicated with a red box. Of note, the threonine residue is absent from the *Ciona* Cesa sequence, however, this protein contains a serine residue at the following position, which could perform a similar function.

How might the finger helix move downward without retro-translocating the glucan? N-terminal capping of α -helices with Asp residues has been shown to significantly stabilize the helical conformation in a pH dependent manner (Forood et al., 1993) (Dirr et al., 2005). During catalysis, Asp343 abstracts a

proton from the acceptor's C4 hydroxyl group (Lairson et al., 2008), thereby likely altering its interaction with the amide protons at the N terminus of the finger helix. Thus, we speculate that the destabilization of the finger helix during glycosyl transfer enables it to refold in the 'down' position, after UDP release and subsequent deprotonation of Asp343. This notion is supported by the position of another conserved residue within the finger helix. Thr346 sits at the membrane distal side of the finger helix just three residues C-terminal of Asp343, and its side chain hydroxyl provides an alternative hydrogen bond partner for the backbone carbonyl of the preceding Glu342 (**Fig. 47**). Threonine residues in α -helices, in particular in hydrophobic environments, often induce helical kinks, which could facilitate the repositioning of the finger helix (Cao and Bowie, 2012; Scharnagl et al., 2014).

On its own, the finger helix is insufficient for cellulose translocation, which requires substrate or UDP binding and gating loop insertion into the active site. Thus, we conclude that the free energy of substrate binding energizes cellulose translocation. Additional thermodynamic driving force for translocation may be generated by the base catalyst itself. The post-translocation state is likely energetically favorable due to a strong interaction between Asp343 and the acceptor's C4 hydroxyl. This interaction is broken upon protonation of its side chain during catalysis but re-established after polymer translocation.

Acknowledgements

We thank Tom Rapoport for critical comments on the manuscript and Justin Acheson for advice on reducing BcsA-B complexes *in crystallo*. Diffraction data were collected at the

Argonne National Laboratory's Advanced Photon Source (APS) beam lines 23-ID-D (GM/CA-), 22-ID (SER-) and 24-ID-C (NE-CAT). GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). The NE-CAT beam lines are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P41 GM103403). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). Data for this research was also in part collected at the APS SER-CAT beam line, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by ANL under Contract No. DE-AC02-06CH11357. J.L.W.M. is supported by a National Science Foundation Graduate Research Fellowship, Grant No. DGE-1315231. M.F. thanks the Austrian Science Fund (FWF) (J3293-B21) for an Erwin Schrödinger postdoctoral fellowship. This research was primarily supported by the National Institutes of Health, Grant 1R01GM101001, awarded to J.Z. S.G.W. thanks the Natural Sciences and Engineering Research Council of Canada for financial support.

Author contributions: J.T.M. and J.L.W.M. purified and crystallized BcsA-B and performed all crystal soaking experiments. J.T.M. cloned and analyzed all BcsA cysteine mutants. J.T.M. and J.L.W.M. collected and processed diffraction data and built and refined the BcsA-B models. M.F. synthesized the fluorinated and phosphonate UDP-Glc analogues and J.R. and H-M.C. synthesized the UDP-thio-galactose analogues. J.T.M., J.L.W.M. and J.Z. analyzed the data. J.Z. and J.L.W.M. wrote the paper and all authors edited to text.

Author information: The authors declare no competing financial interests. Structure factors and coordinates have been deposited at the Protein Data Bank under entry codes 5EJ1, 5EJY, and 5EJZ. Correspondence and requests for Materials should be sent to: jochen_zimmer@virginia.edu.

Table 2
Chapter 3 crystallographic data collection and refinement statistics.

	Product-bound	Substrate-bound	Pre-translocation
Data collection			
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	67.3, 216.8, 221.1	68.0 216.8 220.8	67.4 218.2 220.8
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å)	39.43–2.95 (3.01–2.94)*	35.05–2.90 (2.96–2.90)	29.48–3.4 (3.52–3.4)
<i>R</i> _{pim}	0.060 (0.512)	0.094 (0.514)	0.056 (0.638)
CC _{1/2} [^]	0.995 (0.636)	0.986 (0.577)	0.997 (0.787)
Mean <i>I</i> / <i>sI</i>	10.3 (1.6)	5.2 (1.2)	9.6 (1.3)
Completeness (%)	98.6 (82.4)	99.3 (90.0)	99.7 (100.0)
Redundancy	5.0 (4.2)	10.4 (9.1)	6.6 (6.9)
Refinement			
Resolution (Å)	34.35–2.94	34.92–2.95	29.48–3.4
No. reflections			
Total	68,776	70,259	86075
<i>R</i> _{free}	3,429	3,329	4367
<i>R</i> _{work} / <i>R</i> _{free} (%)	20.6/23.4	20.7/24.2	22.78/26.8
No. atoms			
Protein	10,673	10,725	10,618
β -1,4 glucan	198	188	199
c-di-GMP	92	92	92
UDP	25		
UDP-CH ₂ -Glc		36	
Mg ²⁺	2	2	
Lipids	89	90	69
<i>B</i> -factors			
Chain A	82.8	87.41	155.0
Chain B	72.2	79.4	147.0
Chain D	91.3	106.9	211.0
β -1,4 glucan	80.4	87.7	164.3
c-di-GMP	68.8	71.4	144.7
UDP	97.3		
UDP-CH ₂ -Glc		79.4	
Lipids	85.6	122.4	156.0
R.m.s deviations			
Bond lengths (Å)	0.003	0.003	0.004
Bond angles (°)	0.861	0.905	1.052

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

[^] Correlation between intensities from random half-data sets.

3.4 Methods

3.4.1 *In crystallo* cellulose synthesis

Rhodobacter sphaeroides BcsA–B was purified and crystallized as previously described (Morgan et al., 2014) by the bicelle crystallization method with the exception that gel filtration and crystallization were carried out in buffers lacking MgCl_2 . After the crystals grew to full size (about 2 weeks), cryo-protection was initiated by 3 successive 2 μL additions of cryo solution (well solution containing 20% glycerol) to the crystal mother liquor without added MgCl_2 , each addition separated by 10 minutes.

After the third addition of cryo solution, the polymer was elongated by adding 0.8 μL of 100 mM UDP-activated sugar (glucose or 2-fluoro glucose) in the absence of MgCl_2 to the ~ 8 μL crystallization drop for a final UDP-sugar concentration near 10 mM. The crystals were incubated with UDP-sugar for 2-3 hours at 30°C. After incubation, 6 μL of the crystallization solution was replaced with an equal volume of fresh cryo solution and this process was repeated twice to dilute the UDP-sugar concentration approximately 65-fold. Subsequently, crystals were then looped and flash-cooled in liquid N_2 at various time points.

3.4.2 *In crystallo* cellulose translocation

For *in crystallo* translocation experiments using wild-type BcsA–B, 0.8 μL of a solution containing 100 mM UDP and 250 mM MgCl_2 was added to the ~ 8 μL crystallization drop (see above) for a final concentration of ~ 10 mM UDP and 25 mM MgCl_2 . Crystals were looped and flash-cooled in liquid N_2 at various time points.

For translocation experiments using the BcsA-2Cmutant, cellulose was extended as described above and after completion of cryo-protection, sodium tetrathionate was added to a final concentration of ~1 mM or dithiobutylamine was added to a final concentration of ~100 mM, followed by incubation for 30 or 15 min, respectively. Then, UDP/MgCl₂ were added as described above, and crystals were harvested at various time points.

BcsA accepts UDP-Glc as well as UDP-Gal as substrates, however because galactose is the C4 epimer of glucose, elongation of the cellulose polymer with galactose is expected to stall after a single turn over. Thus, for translocation experiments using UDP-6-thio-Gal as substrate, polymer extension was performed as described above for other substrates with the exception that the cryo solution contained 25 mM MgCl₂, and the 100 mM UDP-6-thio-Gal solution contained 84 mM dithiothreitol (DTT).

3.4.3 Data collection

Diffraction data for wild type BcsA-B and its double cysteine mutant were collected and processed as previously described (Morgan et al., 2014). Diffraction data for UDP-6-thio-Gal were collected at 6.5 keV at NE-CAT to high redundancy. Phases were obtained by molecular replacement using a search model composed of pdb 4P02 with all ligands and residues 332-350 (finger helix) and 499-510 (gating loop) of BcsA omitted. All ligands except for the final 4 glucose units of the cellulose polymer were subsequently added, and the models were refined in Phenix_refine (Adams et al., 2010). Ramachandran analyses of the product-bound, substrate-bound and pre-translocation state structures identify 95.8/3.9/0.3%, 97.6/2.4/0.0% and 96.8/3.2/0.0% residues in the

preferred/allowed/outlier regions, respectively. Figures were prepared using PyMol (PyMol, DeLano Scientific) and crystallographic software is supported by SBGrid (Morin et al., 2013).

3.4.4 UDP-CH₂-Glc soak to generate the donor bound state

BcsA-B was crystallized in the presence of 1 mM UDP-Gal, which was added to the protein/bicelle solution prior to mixing with the crystallization well solution. Fully-grown crystals were cryo-protected as described above at 24°C. The crystals were then incubated with cryo-solution containing 1 mM UDP-CH₂-Glc and 10 mM MgCl₂ for 20 minutes, harvested, and flash-cooled in liquid N₂.

3.4.5 Finger helix cross-linking and activity assays

BcsA cysteine mutants were generated from the constructs described earlier (Morgan et al., 2013) by using the QuikChange mutagenesis technique, and the mutant BcsA-B complex was expressed and prepared as inverted membrane vesicles (IMVs) or purified and reconstituted into proteoliposomes (PLs) as previously described (Omadjela et al., 2013). BcsAB-containing PLs were diluted to 125 nM (or IMVs to 12% v/v) with 125 mM NaCl and 25 mM sodium phosphate, pH 7.2, and incubated with increasing concentrations of DTT or sodium tetrathionate each for 15 minutes at 37°C, prior to initiating cellulose biosynthesis.

To initiate cellulose biosynthesis, c-di-GMP, UDP-Glc, UDP-[³H]-Glc, and MgCl₂ were added, giving final concentrations of 100 nM BcsA-B, 20 mM sodium phosphate, pH 7.2,

100 mM NaCl, 20 mM MgCl₂, 5 mM UDP-Glc, 0.25 μ Ci UDP-[³H]-Glc, and 20 μ M c-di-GMP. Reactions were carried out in 25 μ L aliquots at 24°C for 15 minutes, terminated by adding 2% SDS, and the tritium-labeled cellulose product was quantified by scintillation counting as previously described (Omadjela et al., 2013). Reactions with IMVs containing BcsA Cys or I340 mutants were incubated at 24°C for 3 hrs or 30 min, respectively. All experiments were performed at least in triplicate and error bars represent the deviations from the means.

3.4.6 UDP-sugar analogue synthesis

3.4.6.1 General

Solvents and chemicals were of analytical grade and were purchased from Sigma-Aldrich. Inorganic pyrophosphatase from baker's yeast and uridine-5'-diphosphoglucose pyrophosphorylase (EC.2.7.7.9) from bovine liver were purchased from Sigma Aldrich. Analytical thin-layer chromatography (t.l.c.) was performed on MerckSilica Gel 60F254 (0.2mm thickness on aluminum. T.l.c. plates were visualized using UV light (254 or 365 nm) and by immersion in 10% ammonium molybdate in 2 M H₂SO₄, 5% H₂SO₄. Size exclusion chromatography was performed using Bio-Rad Bio-Gel P-2 Gel (column size: 70 cm x 1.5 cm) at 4°C eluting with deionised water. ¹H and ¹³C NMR spectra were recorded on a BrukerAvance 400inv Fourier Transform spectrometer. All spectra were recorded using an internal lock (deuterium) and are referenced internally to a residual solvent peak. ¹H and ¹³C chemical shifts are quoted in parts per million (ppm) downfield of tetramethylsilane, and chemical shifts (δ) are rounded to the nearest 0.1 ppm unless increased precision was required to distinguish resonances. Coupling constants (J) are

given in Hertz (Hz) and are quoted to the nearest 0.1 Hz. Proton and carbon spectral assignments were made based on COSY, TOCSY, APT, ^1H - ^{13}C -HSQC, and ^1H - ^{13}C -HMBC experiments as required. Mass spectra were acquired from sample dissolved in aqueous methanol on a Waters/Micromass instrument (electrospray ionization) and recorded using an ion-trap.

Synthesis of Uridine-5'-phosphoric- α -(D-glucopyranosyl)methylphosphonic anhydride:

The reaction mixture included glucose-1-phosphonate (Beaton, 2009) (8.8 mg, 32 μmol), UTP (18.6 mg, 33.6 μmol), 25 mM TRIS buffer (pH = 7.4, 6 mL reaction volume), 4 mM MgCl_2 , inorganic pyrophosphatase (5 U) and uridine-5'-diphosphoglucose pyrophosphorylase (15 U). The reaction was incubated at 37°C until TLC showed full consumption of starting materials (TLC eluent: EtOAc:H₂O:CH₃OH:NH₃ = 4:2:2:0.1). Subsequently the mixture was filtered through a 30 kDa cut off spin filter and freeze dried. The residue was dissolved in a minimum amount of water (200 μL), loaded on a Biogel P2 column and was eluted with water (flow: 13.5 mL/h) at 4°C. Fractions containing pure product as judged by TLC were freeze dried to give uridine-5'-phosphoric- α -(D-glucopyranosyl)methylphosphonic anhydride (Yield: 8.5 mg, 45%). Impure fraction were further purified by another P2 column to give more uridine-5'-phosphoric- α -(D-glucopyranosyl)methylphosphonic anhydride (Yield: 7.8 mg, 42%). ^1H -NMR (D₂O at 10°C) δ = 7.83 (d, J=8.12 Hz, 1H, H-6''), 5.86-5.75 (m, 1H, H-1'), 5.80 (d, J=8.16 Hz, 1H, H-5''), 4.32-4.18 (m, 3H, H-1, H-2', H-3'), 4.16-3.98 (m, 3H, H-4', H-5'a, H-5'b), 3.68 (dd, J=2.10, 12.07 Hz, 1H, H-6a), 3.60-3.46 (m, 2H, H-5, H-6b), 3.45

(pseudo t, $J=9.38$ Hz, 1H, H-3), 3.23 (pseudo t, $J=9.38$ Hz, 1H, H-4), 2.09 (ddd, $J=11.29$, 16.05, 16.17 Hz, 1H, H-1''a), 1.96 (ddd, $J=3.42$, 15.75, 20.51 Hz, 1H, H-1''b); ^{13}C -NMR (D_2O at 10°C) δ = 166.6 (C-4''), 152.1 (C-2''), 142.0 (C-6''), 102.9 (C-5''), 88.7 (C-1'), 83.48 (d, $J=9.08$ Hz, C-4'), 74.1 (C-3'), 73.4 (C-3), 72.9 (C-5), 72.46 (d, $J=4.94$ Hz, C-1), 71.36 (d, $J=12.63$ Hz, C-2), 70.3 (C-4), 69.8 (C-2'), 64.93 (d, $J=4.60$ Hz, C-5'), 60.9 (C-6), 24.35 (1C, d, $J=140.19$ Hz, C-1''). ^{31}P -NMR (D_2O) δ = 15.3 (d, $J=26.73$ Hz, 1P, CH_2P), -11.4 (d, $J=26.73$ Hz, 1P), LRMS (ESI negative mode): 563.0 $[\text{M-H}]^-$.

Synthesis of UDP-2-deoxy-2-fluoro- α -D-glucopyranose:

See extended Data Fig. 7b for reaction scheme. The reaction mixture included disodium 2-deoxy-2-fluoro- α -D-glucopyranose (Withers, 1986) (3.1 mg, 10 μmol), UTP (5.8 mg, 10 μmol), 25 mM TRIS buffer (pH = 7.4, 2 mL reaction volume), 4 mM MgCl_2 , inorganic pyrophosphatase (1.5 U) and uridine-5'-diphosphoglucose pyrophosphorylase (3.5 U). The reaction was incubated at 37°C until TLC showed full consumption of starting materials (~3 hours) (TLC eluent: $\text{EtOAc}:\text{H}_2\text{O}:\text{CH}_3\text{OH}:\text{NH}_3 = 4:2:2:0.1$). Subsequently the mixture was filtered through a 30 kDa cut off spin filter and freeze dried. The residue was dissolved in a minimum amount of water (100 μL), loaded on a Biogel P2 column and was eluted with water (flow: 13.5 mL/h) at 4°C . Fractions containing pure product as judged by TLC were freeze dried to give UDP-2-deoxy-2-fluoro- α -D-glucopyranose (Yield: 2.9 mg, 49%). All spectroscopic data were in accordance with previous literature (Stick, 2002). Previously not reported: ^{19}F -NMR

(D₂O) δ = -200.5 (1F, dd, J= 49.1, 12.9 Hz); ³¹P-NMR (D₂O) δ = -11.9 (d, J= 20.0 Hz, 1P), -13.7 (d, J= 20.0 Hz, 1P), LRMS (ESI negative mode): 567.0 [M-H]⁻.

Synthesis of Octyl UDP-6-thio- α -D-galactopyranose disulfide:

3-Methoxy-2-pyridyl 2,3,4-tri-*O*-acetyl-6-thioacetyl- β -D-galactopyranoside (**3**)

To a mixture of **1** (1.82 g, 4.5 mmol) in anhydrous dichloromethane (DCM) (25 mL), anhydrous ethyl acetate (2.5 mL) and acetic anhydride (0.2 mL), stirred at rt under N₂, was added TiBr₄ (3.3 g, 2.0 equiv) as one portion. After stirring overnight, more TiBr₄ (0.5 g) was added, and the mixture was stirred for another 6 h. The reaction mixture was diluted with DCM (150 mL), washed with cold water (100 mL) and cold brine (100 mL), dried over MgSO₄, filtered and evaporated, co-evaporated with dry toluene twice to afford a residue. To the residue were added dry toluene (15 mL) and silver 3-methoxy-2-pyridoxide (Lou, 1997), and the suspension was stirred vigorously for 1 h at 110 °C under N₂. After completion, the reaction mixture was filtered through Celite, and washed with ethyl acetate (200 mL). The combined filtrate was washed with saturated NaHCO₃ (100 mL) and brine (100 mL), dried over MgSO₄, filtered and concentrated. The residue was purified by flash column chromatography (2:1 petroleum ether-ethyl acetate) to afford the product as a white foam **3** (1.016 g, 48 %). ¹H NMR (CDCl₃, 300 MHz): δ 7.70 (dd, 1 H, J 1.3 Hz, J 4.9 Hz, Ar-H), 7.10 (dd, 1 H, Ar-H), 6.93 (dd, 1 H, J 7.8 Hz, Ar-H), 6.18 (d, 1 H, J_{1,2} 8.3 Hz, H-1), 5.52 (dd, 1 H, J_{2,3} 10.3 Hz, H-2), 5.45 (brd, 1 H, H-4), 5.12 (dd, 1 H, J_{3,4} 3.4 Hz, H-3), 3.90 (brt, 1 H, H-5), 3.81 (s, 3 H, OCH₃), 3.14 (dd, 1 H, J_{5,6a}, 7.8 Hz, J_{6a,6b} 13.8 Hz, H-6a), 3.02 (dd, 1 H, J_{5,6b}, 6.5 Hz, H-6b), 2.28 (s, 3 H,

CH₃CO), 2.15 (s, 3 H, CH₃CO), 1.97 (s, 3 H, CH₃CO), 1.93 (s, 3 H, CH₃CO). ¹³C NMR (CDCl₃, 75 MHz): δ 194.5, 170.5, 170.2, 169.5, 151.8, 144.3, 136.9, 119.4, 119.1, 94.0, 72.8, 71.5, 68.5, 67.8, 56.1, 30.5, 28.3, 20.83, 20.78, 20.7, EIS-MS: Calcd for [C₂₀H₂₅NO₁₀S + H]⁺: 472.1. Found m/z: 472.5.

Octyl 3-Methoxy-2-pyridyl 2,3,4-tri-*O*-acetyl-6-thio-β-D-galactopyranoside

Disulfide (4)

A solution of **3** (0.424 g, 0.9 mmol) in dry DMF (10 mL) was degassed with N₂ for 15 minutes, and then hydrazine acetate (162 mg, 2 equiv) was added as one portion. The reaction mixture was stirred for 0.5 h under N₂, diluted with ethyl acetate (100 mL), washed with saturated NaHCO₃ (50 mL) and brine (50 mL), dried over MgSO₄, filtered and concentrated, dried under vacuum for 1 h to give a residue. To the residue in dry DCM (40 mL), stirred at rt under N₂, were added diethyl-*N*-(octylsulfanyl)hydrazodicarboxylate, prepared by reaction of diethylazodicarboxylate with octanethiol in methylene chloride (Mukaiyama and Takahashi, 1968), (0.92 g, 2.7 mmol, 3 equiv) and trimethylamine (30 μL). The reaction mixture was stirred for 0.5 h under the same conditions, diluted with DCM (100 mL), washed with 1M HCl (60 mL), saturated NaHCO₃ (60 mL) and brine (60 mL), dried over MgSO₄, filtered and evaporated. The resulting residue was purified by flash column chromatography (4:1 petroleum ether-ethyl acetate) to afford the product **4** as a syrup (0.428 g, 83 %). ¹H NMR (CDCl₃, 400 MHz): δ 7.68 (dd, 1 H, J 1.2 Hz, J 4.8 Hz, Ar-H), 7.08 (dd, 1 H, Ar-H), 6.91 (dd, 1 H, J 8.0 Hz, Ar-H), 6.23 (d, 1 H, J_{1,2} 8.4 Hz, H-1), 5.53 (m, 2 H, H-2 & H-4), 5.18 (dd, 1 H, J_{2,3} 10.4 Hz, J_{3,4} 2.4 Hz, H-3), 4.09 (brt, 1 H, H-5), 3.79 (s, 3 H, OCH₃), 2.88 (dd, 1 H, J_{5,6a} 7.2 Hz, J_{6a,6b} 14.0 Hz, H-6a), 2.71 (dd, 1 H, J_{5,6b} 6.4 Hz, H-6b), 2.63

(t, 1 H, J 7.4 Hz, CH₂S), 2.12 (s, 3 H, CH₃CO), 1.97 (s, 3 H, CH₃CO), 1.92 (s, 3 H, CH₃CO), 1.57 (m, 2 H, CH₂), 1.26 (m, 10 H, 5XCH₂), 0.84 (t, 3 H, J 6.4 Hz, CH₃). ¹³C NMR (CDCl₃, 100 MHz): δ 170.3, 170.2, 169.5, 151.9, 144.3, 136.9, 119.4, 119.1, 94.1, 72.8, 71.6, 70.0, 68.7, 68.4, 56.1, 39.2, 38.9, 31.9, 29.3, 29.2, 29.1, 28.5, 22.7, 20.84, 20.79, 20.73, 14.2; EIS-MS: Calcd for [C₂₆H₃₉NO₉S₂ + H]⁺: 574.3. Found m/z: 574.6.

Octyl 3-Methoxy-2-pyridyl 6-thio-β-D-galactopyranoside Disulfide (5)

Ammonia was bubbled into a cooled (0 °C) solution of **4** (0.4 g, 0.7 mmol) in dry methanol (50 mL) for 10 minutes, and then the reaction mixture was stirred for 48 h under the same conditions. After this time the solvent was evaporated under reduced pressure. The resulting residue was purified by flash column chromatography (50:1 & 20:1 DCM-MeOH) to afford the product as a white foam (0.285 g, 91 %). ¹H NMR (CD₃OD, 400 MHz): δ 7.73 (dd, 1 H, J 1.2 Hz, J 4.8 Hz, Ar-H), 7.33 (dd, 1 H, Ar-H), 7.02 (dd, 1 H, J 8.0 Hz, Ar-H), 5.96 (d, 1 H, J_{1,2} 8.0 Hz, H-1), 4.03 (brd, 1 H, H-4), 3.93 (m, 2 H, H-2 & H-5), 3.89 (s, 3 H, OCH₃), 3.71 (dd, 1 H, J_{3,4} 3.6 Hz, J_{2,3} 9.6 Hz, H-3), 3.03 (dd, 1 H, J_{5,6a} 6.8 Hz, J_{6a,6b} 13.6 Hz, H-6a), 2.98 (dd, 1 H, J_{5,6b} 6.8 Hz, H-6b), 2.74 (t, 1 H, J 7.2 Hz, CH₂S), 1.65 (m, 2 H, CH₂), 1.37 (m, 10 H, 5XCH₂), 0.95 (t, 3 H, J 6.8 Hz, CH₃). ¹³C NMR (CD₃OD, 100 MHz): δ 152.6, 144.5, 136.6, 119.3, 118.6, 96.4, 74.1, 73.9, 70.4, 69.3, 55.2, 39.1, 38.7, 31.8, 29.1 (2 C), 28.9, 28.2, 22.5, 13.3. EIS-MS: Calcd for [C₂₀H₃₃NO₆S₂ + H]⁺: 448.2. Found m/z: 448.5.

Octyl 6-thio-α-D-galactopyranosyl phosphate Disulfide (6)

The 3-methoxy-2-pyridyloxy glycoside **5** (0.105 g; 0.234 mol) was concentrated twice from toluene (2 x 10 mL) before being dissolved in dry *N,N*-dimethylformamide (2 mL) under an argon atmosphere. This solution was transferred by canula to a flask containing dry H₃PO₄ (0.125 g; 1.27 mmol; 5.4 eq.). The original flask was rinsed with an additional 2 mL of the same solvent, which was then also transferred to the phosphoric acid containing flask by canula. The reaction was stirred under Argon at room temperature for approximately one hour before neutralization by addition of a slight excess of sodium hydroxide in water (600 uL). The solution was then partially concentrated by rotary evaporator under reduced pressure and purified by tC18 SepPak (Waters) to give **6** (0.033 g, 31 %) and recovered **5** (67 mg, 64%). ¹H NMR (D₂O, 400 MHz): δ 5.51 (ddd, 1 H, J_{1,2} 3.1 Hz, J_{1,P} 7.6 Hz, H-1; in ³¹P decoupled ¹H NMR signal collapses to d, J_{1,2} 3.1 Hz), 4.37 (dd, 1 H, H-5), 4.17 (brd, 1 H, J ~2 Hz, H-4), 3.99 (dd, 1 H, J_{3,4} 3.2 Hz, J_{3,4} 10.4 Hz, H-3), 3.79 (dd, 1 H, J_{2,3} 10.4 Hz, J_{2,1} 3.1 Hz, H-2), 3.09 (dd, 1 H, J_{6a,6b} 13.6 Hz, J_{6a,5} 6.4 Hz, H-6a), 3.00 (dd, 1 H, J_{6b,5} 8.0 Hz, H-6b), 2.91-2.79 (m, 2 H, SCH₂), 1.81-1.70 (m, 2 H, SCH₂CH₂), 1.5-1.3 (m, 10 H), 0.91 (m, 3H, CH₃). ¹³C NMR (D₂O, 100 MHz): δ 94.0 (d, ³J_{1,P} = 5.0 Hz, C-1), 69.98 (C-5), 69.95 (C-3), 69.6 (C-4), 69.3 (d, ⁴J_{2,P} = 6.0 Hz C-2), 38.4 (C-6), 38.3 (S-CH₂), 31.2, 28.5, 28.4, 28.4, 27.7, 22.2, 13.6. ³¹P NMR (D₂O, 162 MHz) δ 2.5. ESI-HRMS: Calcd for [C₁₄H₂₈O₈PS₂]: 419.0963. Found m/z: 419.0959.

Octyl UDP-6-thio-α-D-galactopyranose Disulfide (7):

The glycosyl phosphate **6** (9.1 mg; 20.5 umol) was suspended in methanol (1.0 mL) and cooled to 4 °C. Cold Amberlyst IR 120H (~ 20 mg) was added and the mixture was

stirred for 5 minutes in the fridge, during which time the remaining sugar solids appeared to dissolve. The mixture was filtered through cotton in a Pasteur pipette into cold pyridine (2 mL), rinsing with cold methanol (2 mL). The solution was concentrated on a rotary evaporator under reduced pressure ($< 30\text{ }^{\circ}\text{C}$), redissolved in dry pyridine and concentrated again (2 x 4 mL). The galactosyl phosphate pyridinium salt was then dissolved in dry pyridine (3 mL) and transferred by canula to a flask containing UMP morpholidate (28 mg; ~ 2 eq; Sigma) that had previously been dissolved in dry pyridine and concentrated (3 x 4 mL). To this solution was next added 1H-tetrazole (5 mg) in freshly distilled pyridine. The mixture was stirred for ~ 72 hours after which time methanol (1 mL) was added and the solution was concentrated to dryness. The residue was dissolved in 20 mM NH_4Cl and applied to a tC18 SepPak (Waters) preconditioned with methanol and then water (2 x column volume). The SepPak was eluted with water, 10, 20, 30, 50, and 70% methanol in 20 mM NH_4Cl , and then 100% methanol. The desired product was obtained in the 70% methanol fractions. After concentration the UDP-sugar was further purified on a Biogel P2 column and was eluted with water (flow: 13.5 mL/h) at 4°C to obtain **7** (6 mg; 39%). ^1H NMR (D_2O , 400 MHz): δ 7.90 (d, 1 H, J 8.4 Hz, =CH), 5.90 (d, 1 H, =CH), 5.87 (d, 1 H, J 3.2 Hz, H-1), 5.52 (dd, 1 H, $J_{1,2}$ 3.6 Hz, $J_{\text{P,H}}$ 7.6 Hz, H-1'), 4.30~4.08 (m, 6 H, H-2, H-3, H-4, H-5a, H-5b, H-5'), 4.02 (d, 1 H, H-4'), 3.84 (dd, 1 H, $J_{3,4}$ 3.2 Hz, $J_{2,3}$ 10.4 Hz, H-3'), 3.70 (dt, 1 H, H-2'), 2.92 (dd, 1 H, $J_{5,6a}$ 6.4 Hz, $J_{6a,6b}$ 13.6 Hz, H-6a), 2.84 (dd, 1 H, $J_{5,6b}$ 7.2 Hz, H-6b), 2.67 (t, 1 H, J 7.2 Hz, CH_2S), 1.58 (m, 2 H, CH_2), 1.18 (m, 10 H, $5\times\text{CH}_2$), 0.77 (t, 3 H, J 6.6 Hz, CH_3). ^{13}C NMR (D_2O , 150 MHz): δ 165.9, 151.3, 141.2, 102.1, 95.3 (d, $J_{\text{P,C}}$ 7.5 Hz), 88.1, 82.6 (d, $J_{\text{P,C}}$ 9.0 Hz), 73.4, 70.1, 68.9, 68.87, 68.8, 67.8 (d, $J_{\text{P,C}}$ 6.8 Hz), 64.2 (d, $J_{\text{P,C}}$ 6.8 Hz),

37.6, 37.5, 30.6, 27.7, 27.0, 21.5, 20.6 (2 C), 12.9. ^{31}P NMR (D_2O , 162 MHz): δ -10.3 (d, J 20.0 Hz), -12.0 (dd, J 6.0 Hz), $^{31}\text{P}(\text{H})$ NMR (D_2O , 162 MHz): δ -10.3 (d, J 20.1 Hz), -12.0 (d). ESI-HRMS: Calcd for $[\text{C}_{23}\text{H}_{39}\text{N}_2\text{O}_{16}\text{P}_2\text{S}_2]^-$: 725.1216. Found m/z : 725.120

Chapter 4: Structure of a Type-1 secretion system ABC transporter

Jacob Lowell Whitten Morgan^{1*} and Jochen Zimmer*

Department of Molecular Physiology and Biological Physics
University of Virginia School of Medicine
480 Ray C. Hunt Dr.
Charlottesville, VA 22908

¹Lead Contact

In this chapter I will present a paper that is under review at *Structure* concerning the structure and mechanism of a Type 1–Secretion System ABC Transporter. I began working on this project in early 2014, determined the structure in early 2015, struggled with functional assays until mid 2016, and wrote the manuscript in mid 2016. I am responsible for crystallizing and experimentally phasing PrtD, determining the structure, conceiving and performing functional assays, generating figures, writing the manuscript, and corresponding with editors.

4.1 Introduction

Gram-negative bacteria have evolved a number of mechanisms for protein secretion across the cell envelope, formed by the inner membrane, the peptidoglycan-containing periplasm, and the outer membrane. One pathway, the type-1 secretion system (T1SS), is utilized by many bacteria to secrete such substrates as toxins, virulence factors, proteases, lipases, and scavenging proteins. In particular, numerous pathogenic bacteria utilize the T1SS to secrete virulence factors, including *E. coli* hemolysin, *B. pertussis* adenylate cyclase toxin, and *L. pneumophila* and *V. cholera* RtxA (Binet et al., 1997; Delepelaire, 2004; Linhartová et al., 2010; Fuche et al., 2015)

The T1SS apparatus is composed of an inner-membrane (IM) ATP-binding-cassette (ABC) transporter, a periplasmic membrane-fusion protein (MFP), and an outer-membrane (OM) TolC-like porin. These components assemble in the presence of substrate, which contains a C-terminal secretion signal, to form a contiguous conduit across the periplasm (Thanabalu et al., 1998). The substrate is then transported from the cytoplasm to the medium with no apparent periplasmic intermediate (Koronakis et al., 1991).

The essential components of ABC transporters include a TM domain (TMD) and a nucleotide-binding domain (NBD). In most bacterial ABC exporters, both domains are formed by a single polypeptide chain with the TMD in the N-terminal and the NBD in the C-terminal half. These subunits assemble into either hetero- or homo-dimers that form two ATP binding sites at the NBD interface. Conformational changes of the NBDs

occurring during ATP binding and hydrolysis are coupled to the TMD to promote transport of substrates (Rees et al., 2009; Beek et al., 2014; Locher, 2016).

T1SS ABC transporters can be further divided into three classes depending on the presence and functionality of an N-terminal peptidase domain (Kanonenberg et al., 2013). The C39 group contains a functional C39 peptidase domain, which is involved in proteolytic processing of the substrate during secretion. The exported substrates tend to be small (≤ 10 kDa), and there is controversy as to whether they should be considered T1SS ABC transporters because the secretion signal is located at the N- instead of the C-terminus of the secreted peptide (Thomas et al., 2014). The CLD group, which includes the hemolysin transporter, HlyB, contains a degenerate ‘C39-like’ domain (CLD) of which the catalytic residue is mutated. The CLD is involved in binding the substrate as well as regulating ATPase activity of the NBDs (Lecher et al., 2012; Reimann et al., 2016). The third group contains no N-terminal extension, and its prototypical members are PrtD from *Dickeya dadantii* (Dd, formerly *Erwinia chymosthera*) and HasD from *Serratia marcescens* (Sm) (Kanonenberg et al., 2013).

The secretion signal for T1SS substrates is located at the C-terminus and comprises the last ~60 residues (Delepelaire, 2004; Thomas et al., 2014). The fusion of the secretion signal onto the C-terminus of an otherwise non-secreted protein is often sufficient to target it for secretion (Park et al., 2012). However, the specific properties of this secretion signal that target it to the T1SS are not well understood. At the sequence level, T1SS substrates have a number of unique characteristics. First, all, except the SmHasD

substrate, contain a ‘repeats-in-toxin’ (RTX) domain made up of the sequence GGxGxDxUx, where U is a hydrophobic and x is any residue. These ‘GG repeats’ form a calcium-binding domain that is disordered in the absence of calcium, but folds at higher (micromolar) calcium concentrations (Chenal et al., 2008; 2010). Thus, calcium binding plays a dual role for substrate secretion. In the cytoplasm, the polypeptide is disordered because calcium concentrations are low (approximately 300 nM), upon secretion into the extracellular milieu, where calcium can reach millimolar concentrations, however, calcium promotes folding (Jones et al., 1999; Thomas et al., 2014). Furthermore, the calcium-induced folding of the RTX domain is likely to play a role in energizing secretion by preventing backsliding and thus promoting vectorial transport of the substrate (Bumba et al., 2016). Second, T1SS substrates tend to be quite acidic, suggesting that substrate transport could be electrogenic (Delepelaire, 2004). Indeed, studies with *E. coli* hemolysin revealed a reduced secretion efficiency upon disruption of the membrane potential (Koronakis et al., 1991) while studies with *B. pertussis* CyaA suggested impaired secretion with increased basicity of the substrate (Bumba et al., 2016), both observations are consistent with an electrical contribution to secretion. Finally, T1SS substrates are quite long as compared to microcins, another class of protein toxins secreted in an ABC-transporter-dependent manner. While microcins only reach about 10 kD (Yang et al., 2014), T1SS substrates range in size from the 20 kD HasA hemophore to the 900 kD LapA adhesion protein .

The large size of T1SS substrates suggest that the transporter operates in a distinct mechanism from peptide ABC transporters such as microcin exporters and the human

antigen transporters, TAP1/2 (Choudhury et al., 2014; Grossmann et al., 2014; Eggensperger and Tampe, 2015; Lin et al., 2015). These transporters appear to operate by the ‘alternating access’ mechanism whereby the peptide binds from the cytoplasmic side of the membrane in a pocket of the inward-open conformation of the transporter. ATP binding at the NBDs induces a conformational change to the outward-open state, which exposes the binding site to the opposite side of the membrane where the substrate diffuses away from the binding pocket. ATP hydrolysis then returns the transporter to the inward-open conformation (Rees et al., 2009; Beek et al., 2014; Locher, 2016). A necessary feature of this mechanism is that the transporter forms a substrate-binding pocket large enough for the entire substrate. For T1SS ABC transporters, however, this does not seem feasible as substrates can be more than 8,000 amino acids in length (Delepelaire, 2004; Linhartová et al., 2010). Additionally, several experiments indicate that T1SS substrates must be unfolded in order to be translocation competent (Debarbieux and Wandersman, 2001; Bakkes et al., 2010), so the length of an unfolded T1SS substrate would far exceed the length of the transporter itself. These properties suggest that T1SS ABC transporters operate by a distinct mechanism with the unfolded polypeptide being processively moved through a pore formed by the transporter's TMDs (Linhartová et al., 2010; Kanonenberg et al., 2013; Thomas et al., 2014).

The *Dickeya dadantii* protease secretion system has long served as a model system for studying T1SS (Létoffé et al., 1990). It is composed of PrtD, the ABC transporter; PrtE, the MFP; and PrtF, the OM porin. These components assemble to secrete a number of

proteases encoded in the same operon: PrtA, PrtB, PrtC, and PrtG (Delepelaire and Wandersman, 1989; Ghigo and Wandersman, 1994).

Here, we show that secretion of Dd-PrtG across the plasma membrane requires all three transporter components, PrtD, -E, and -F, and demonstrate that secretion of PrtG *in vivo* requires ATP hydrolysis by PrtD. Furthermore, the crystal structure of a PrtD homologue from *Aquifex aeolicus* in an ADP-bound state reveals a novel, pore-forming architecture of the TMDs, as well as structural features likely implicated in substrate recruitment.

4.2 Results

4.2.1 Substrate secretion is dependent on ATP hydrolysis

In order to study substrate secretion *in vivo*, we reconstituted Dd-PrtD, PrtE, PrtF, and PrtG into *E. coli* carrying S-, His-, Flag-, and Myc- tags, respectively. The transporter components, PrtD-F, are under an IPTG-inducible promoter while the substrate, PrtG, is under an arabinose-inducible promoter. Upon expression of PrtDEF and PrtG, we observe continuing secretion of PrtG into the media upon addition of arabinose (**DdPrtDEF, Fig. 48A**), as demonstrated previously (Létoffé et al., 1996). However, if we express only PrtE and PrtF without PrtD, we observe no secretion, demonstrating that translocation is dependent on PrtD (**DdPrtEF, Fig. 48A**). Additionally, expression of only PrtD and PrtE (with no PrtF) yields no secretion, confirming that native *E. coli* TolC, which is homologous to PrtF, cannot replace PrtF in a functional T1SS complex (Binet and Wandersman, 1995)

We next tested whether PrtG secretion requires ATPase activity of PrtD.

In ABC transporters, replacing the Glu of the Walker B motif with a Gln has been frequently utilized to eliminate ATPase activity, while preserving the ability to bind ATP (Lee et al., 2014; Lin et al., 2015). Upon co-expression of the corresponding Dd-PrtD-E492Q mutant with PrtE and PrtF, we observe no secretion into the media even after 18 hr of PrtG induction, (**Right panel, Fig. 48A**). This is in agreement with previous observations on the hemolysin secretion system (Koronakis et al., 1995).

Blotting the cell extract for PrtG confirms that it is expressed but not secreted. These results indicate that PrtG secretion requires ATP hydrolysis and suggest that PrtD does not act as an ATP-gated channel as observed in ATPase incompetent mutants of the CFTR chloride channel (Csanady et al., 2010).

4.2.2 Transport across the inner membrane requires the entire T1SS complex

An important question in T1SS function is whether PrtD can transport substrate across the inner membrane in the absence of other transporter components. Because many T1SS substrates are toxins, it seems likely that the bacterium would prevent transporting these substrates into its own periplasm. It is possible, however, that the appearance of substrate in the periplasm is the stimulus that initiates recruitment of the MFP and subsequent assembly of the entire T1SS apparatus. Previous experiments on hemolysin secretion suggested that HlyB, the ABC transporter, requires both HlyD, the MFP, as well as TolC for substrate transport across the IM (Koronakis et al., 1997). However, HlyB contains

the N-terminal CLD, which PrtD lacks, and this domain is essential for HlyB activity and plays an important role in substrate encounter (Lecher et al., 2012). Furthermore, HlyB and HlyD appear to form a complex before substrate induction, whereas PrtD and PrtE do not (Létoffé et al., 1996; Thanabalu et al., 1998). Therefore, PrtD-PrtE-PrtF may assemble by a different mechanism from HlyB-HlyD-TolC.

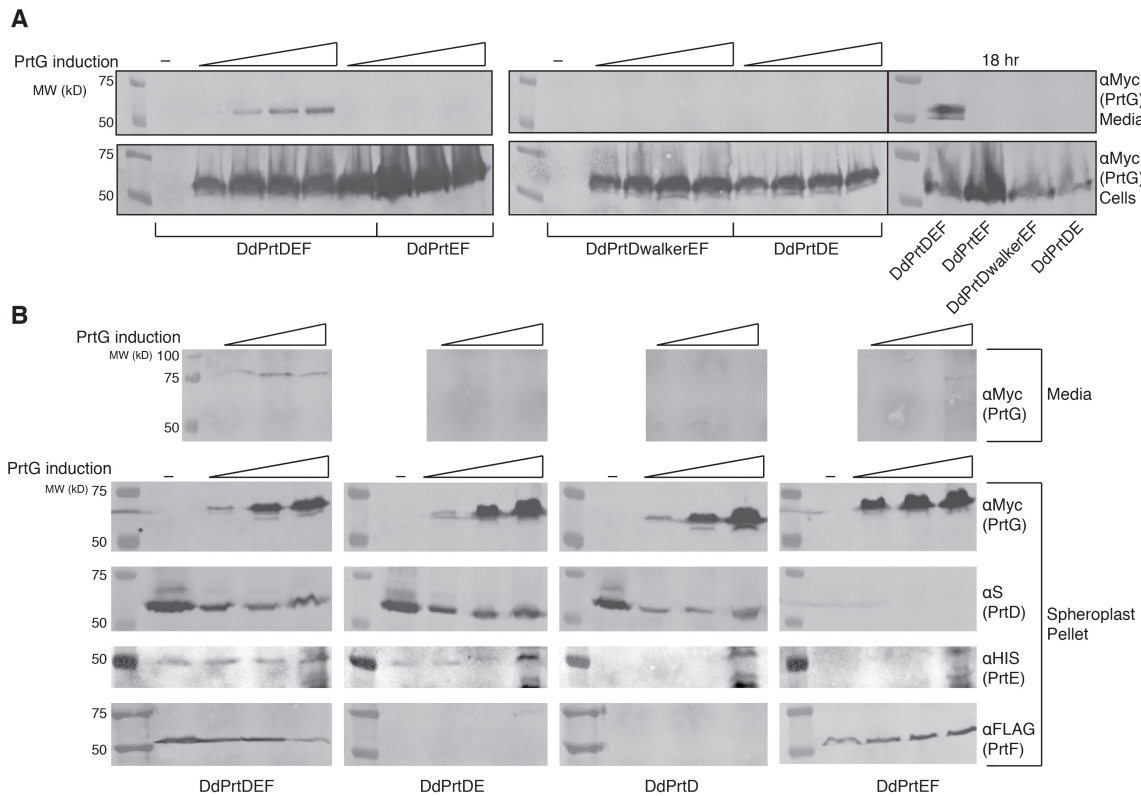


Figure 48 *In vivo* secretion of DdPrtG. (A) Secretion of DdPrtG. Time-course of PrtG expression in *E. coli* cells expressing the indicated Dd T1SS components. Increasing times represent 1, 2, 3, 4 hr post induction. Top, Western blot for myc-tagged PrtG in media. Bottom, Western blot for myc-tagged PrtG in cell pellet. (B) Secretion of DdPrtG in spheroplasts. *E. coli* cells expressing the indicated Dd T1SS components were spheroplasted before induction of PrtG for the indicated times. The presence of PrtG in the media and spheroplasts was detected by Western blotting. Spheroplasts were also analyzed for expression of all other T1SS components.

Spheroplasts are a convenient system to study IM protein secretion (Schlesinger, 1968; Fan et al., 2012) because, in the absence of an outer membrane, proteins that cross the IM are released into the media. To this end, we first expressed the T1SS transporter components of interest in *E. coli* and then generated spheroplasts by incubating the cells with EDTA and lysozyme in a high sucrose buffer. Following resuspension of the spheroplasts in fresh LB medium (supplemented with sucrose and 10 mM CaCl₂), expression of PrtG in the spheroplasts was induced with arabinose and the media was assayed for the presence of PrtG at different time intervals.

If PrtD, PrtE, and PrtF are all expressed (**DdPrtDEF, Fig. 48B**), PrtG is detected in the media at 1, 3, and 6 hours of induction (**DdPrtDEF top panel, Fig. 48B**), likely due to the presence of functional PrtDEF complex in the spheroplasts or remaining intact cells. Importantly, omission of PrtD results in no secretion of PrtG into the media (**DdPrtEF top panel, Fig. 48B**) despite comparable PrtG expression levels (**compare DdPrtDEF and DdPrtEF spheroplast pellet, Fig. 48B**), demonstrating that the spheroplasts are intact and that secretion into the media is specifically due to T1SS. Finally, expression of PrtD on its own or together with PrtE results in no secretion of PrtG into the media (**DdPrtD and DdPrtDE top panel, Fig. 48B**), indicating that neither PrtD on its own nor a complex of PrtD and PrtE is able to fully transport PrtG across the inner membrane.

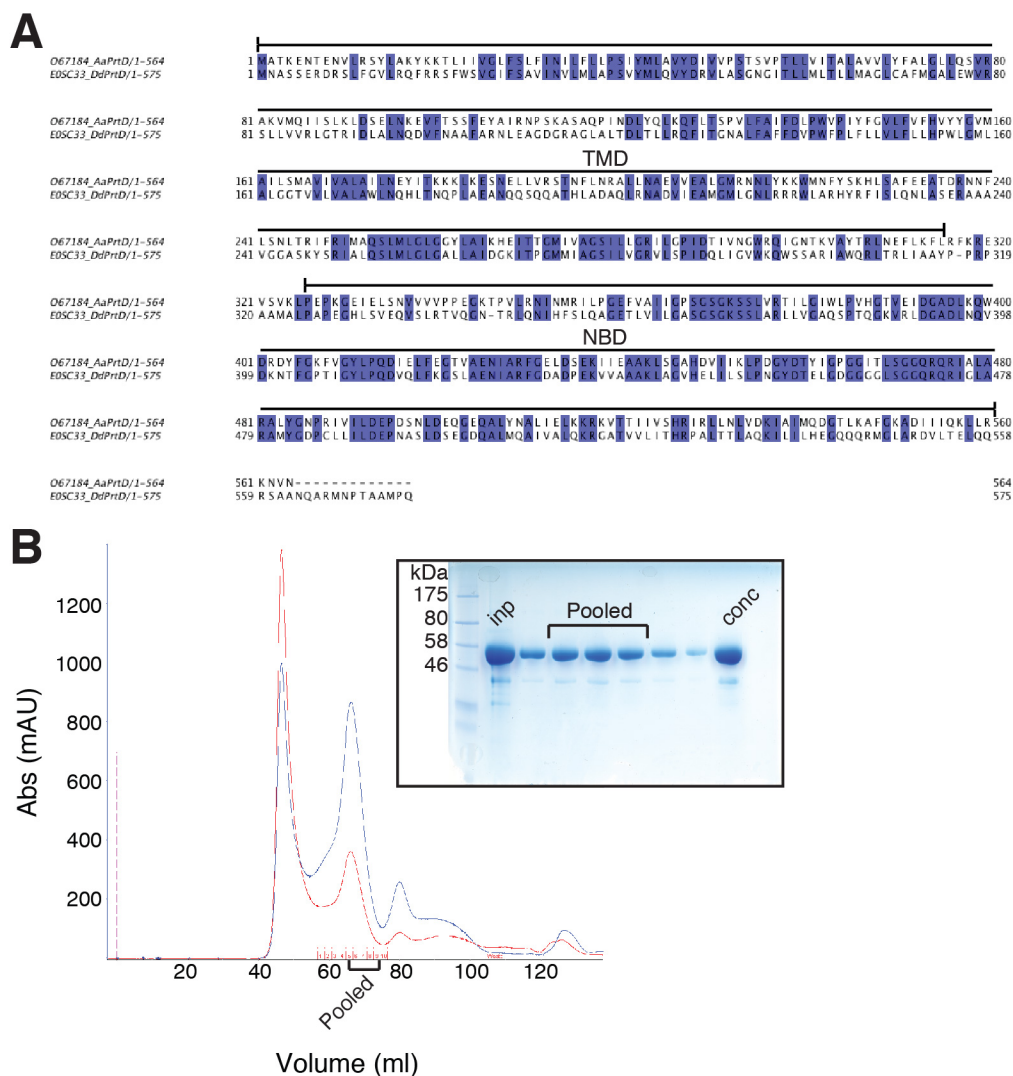


Figure 49 AaPrtD Properties. (A) Alignment of AaPrtD and DdPrtD. The coloring is according to identity and PrtD's NBD and TM regions are indicated. (B) Purification of AaPrtD. Gel filtration profile of AaPrtD solubilized in DDM. Inset shows Coomassie-stained gel of select fractions. Fractions pooled for crystallization are indicated both on profile and gel.

4.2.3 *Aquifex aeolicus* PrtD exhibits a typical ABC exporter fold

In order to gain structural insights into the transport mechanism of PrtD, we cloned a PrtD homologue from the hyperthermophilic bacterium *Aquifex aeolicus* (AaPrtD).

AaPrtD and DdPrtD exhibit 40% sequence identity comprising 37% in the TM region

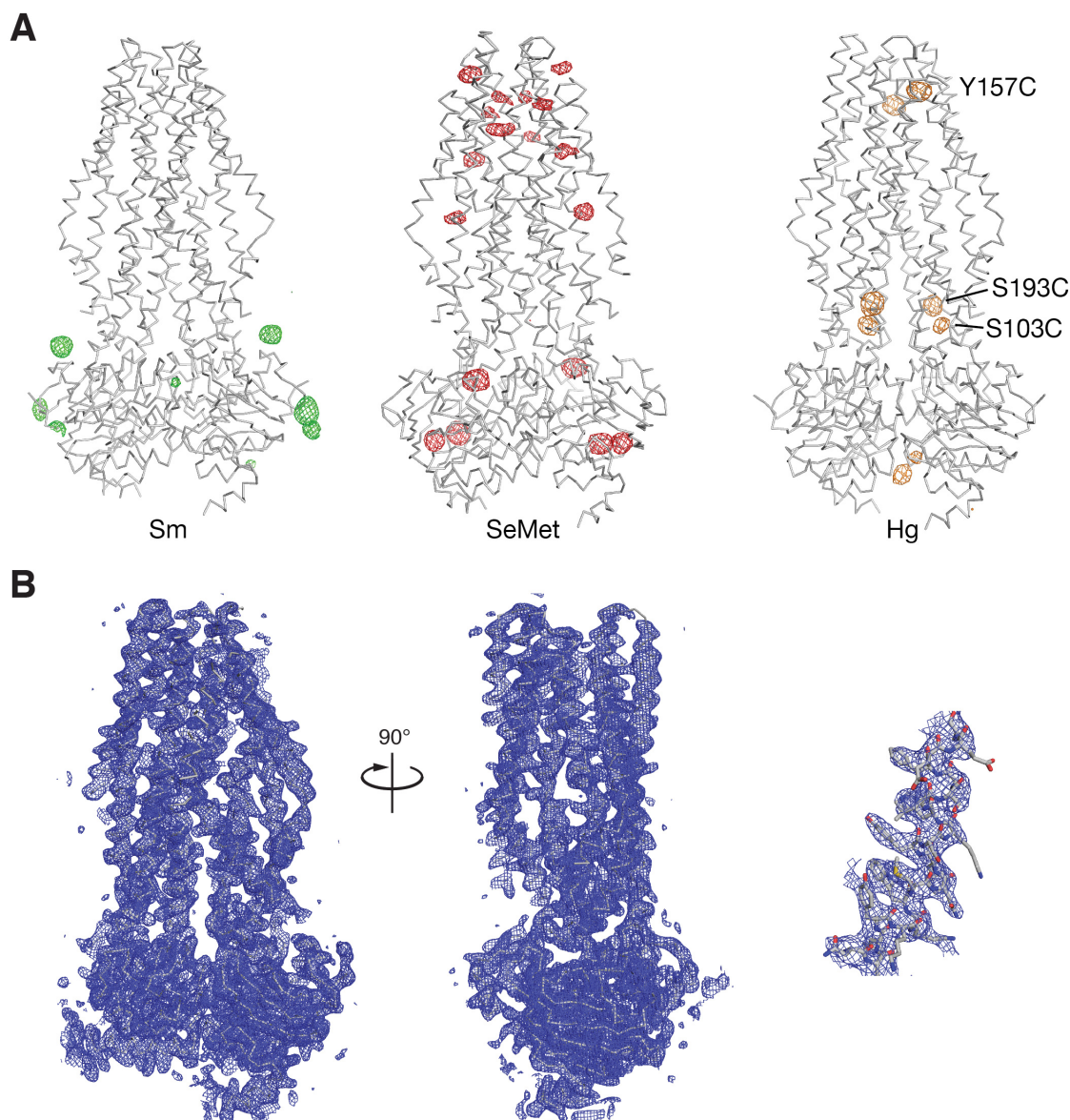


Figure 50 Phasing of AaPrtD. (A) Positions of anomalous scattering atoms used for phasing. The final AaPrtD structure is shown as a gray ribbon. Anomalous difference Fourier electron density maps for Sm-, Se-Met-, and Hg-derivatized crystals are shown in green, red, and orange and contoured at 5-, 4-, and 4- σ , respectively. Pb bound in approximately the same positions as Sm. (B) Final experimental electron density map following density modification and phase extension to 3.15 Å shown as a blue mesh contoured at 1 σ .

and 44% in the NBDs (**Fig. 49A**). AaPrtD can be expressed to high levels in *E. coli* and

purified in the detergent dodecyl- β -D-maltoside (DDM) via metal affinity and gel filtration chromatography (**Fig. 49B**). AaPrtD crystals were optimized in the presence of ADP to diffract to a maximum resolution of 3.15 Å (**Table 3**). Initial phases were determined experimentally using single-wavelength anomalous diffraction from samarium acetate-soaked crystals (**Fig. 50A**). Additional phase information was obtained from mercury chloride-soaked crystals of single-Cys AaPrtD mutants and from Se-Met-derivatized AaPrtD (**Fig. 50A**). Following density modification, the initial electron density was of high quality and allowed model building of almost the entire AaPrtD homodimer in the asymmetric unit (**Fig. 50B**).

The final AaPrtD structure, containing residues 11-315 and 326-559, reveals a fold typical of ABC exporters (**Fig. 51A**). It is comprised of a homodimer with each subunit containing six N-terminal TM helices and a C-terminal NBD (**Fig. 51B**). The NBDs closely interact with one ADP molecule bound at each of the active sites (**Fig. 51A and 53A**). The overall conformation could be described as ‘occluded’ with the TM domains closed on the periplasmic side despite the closure of the NBDs on the cytosolic side (**Fig. 51A**).

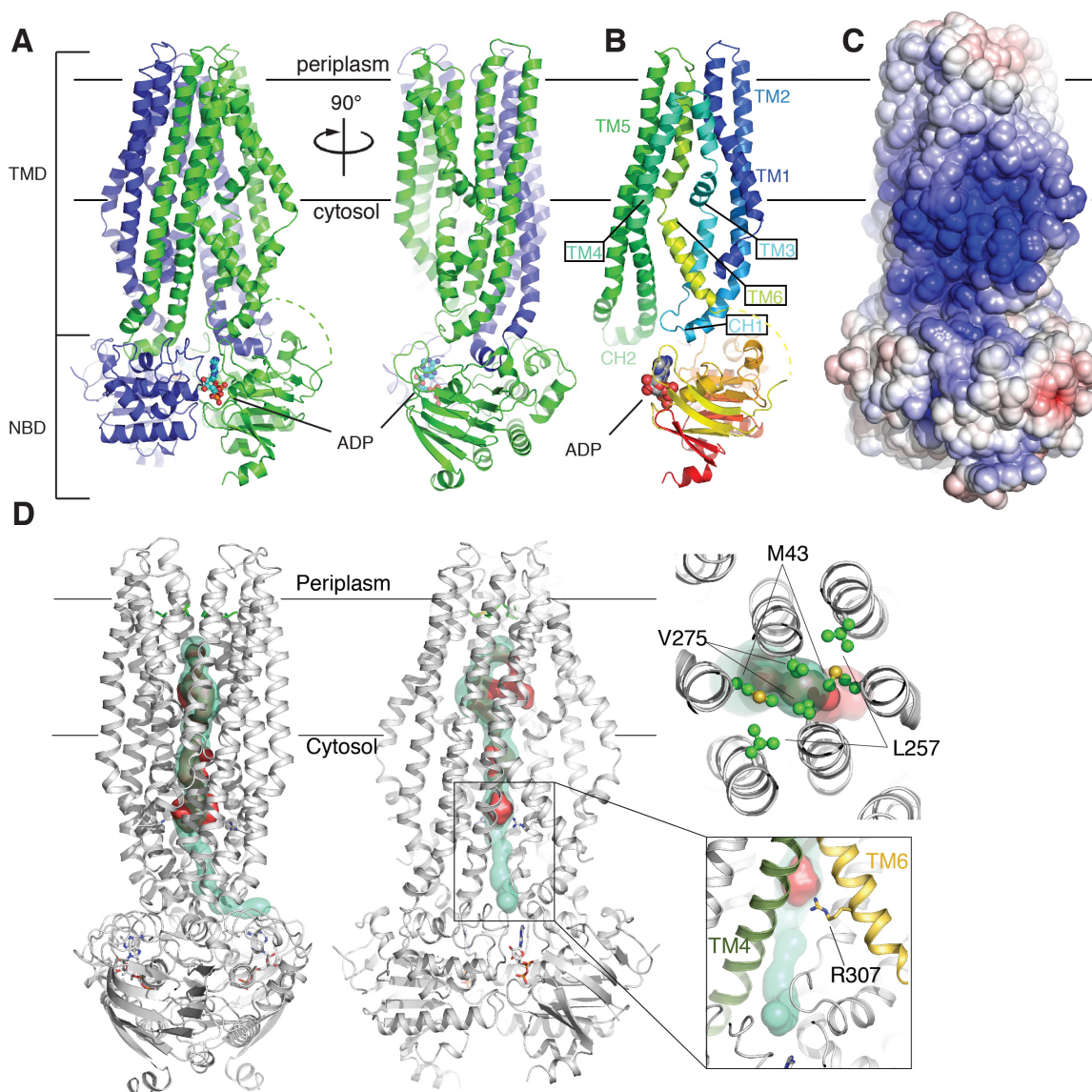


Figure 51 Structure of AaPrtD. (A) Overall architecture of AaPrtD homodimer. Subunits are colored green and blue, respectively, and ADP is shown as spheres. (B) Architecture of a PrtD monomer. The subunit is colored in a gradient from blue to red corresponding to the N- and C-termini, respectively. (C) Electrostatic properties. AaPrtD is shown as a solvent-accessible surface and colored according to electrostatic surface potential (Blue = +10 KT, Red = -10 KT). Electrostatics were calculated with APBS in Pymol using a Parse force field. (D) AaPrtD channel. AaPrtD is shown as a cartoon. The volume accessible to a 2.4 Å radius probe is shown as a red surface (calculated with HOLLOW), which is overlaid with a semi-transparent cyan surface representing a tunnel accessible to a 1.0 Å radius probe (calculated with CAVER). Right top panel: Periplasmic hydrophobic seal of PrtD. Right bottom panel: Putative substrate entry window formed by PrtD's TM helices 4 and 6.

AaPrtD shows a distinct architecture of its TM helices. Particularly striking features are the dramatic kinks in TM3 and TM6 near the cytosolic solvent-lipid interface (**Fig. 51B**).

These kinks, which point towards the dimer interface, create a highly-basic concave bowl on the PrtD surface formed by TM1, TM3, and TM6 (**Fig. 51B and C**). In addition to the sharp kink in helix 3, the coupling helix between TM2 and TM3 (CH1), which packs against the NBD of the same subunit, breaks and forms a short loop leading into TM3 (**Fig. 51B**). To our knowledge, the breaking of a coupling helix has not been observed before in an ABC exporter. Possible implications of this unique architecture of the TM region will be discussed later.

4.2.4 AaPrtD contains a narrow, occluded TM channel

AaPrtD's TM region forms a continuous channel approximately 40 Å long that spans almost the entire TM region (**Fig. 51D**). Based on accessibility to a 2.4 Å radius probe, the channel begins within the cytoplasmic section of AaPrtD about 20 Å from the cytosolic water-lipid interface, and extends about 3/4 of the way across the bilayer before it ends shortly before the periplasmic exit.

The channel is blocked on the cytoplasmic side by the conserved Arg307, right above CH1 and near where TM4 and TM6 separate (**Left panel, Fig. 51D**). On the periplasmic side, the channel is closed by a patch of conserved hydrophobic residues that point radially inwards (**Right panel, Fig. 51D**), thereby forming a seal similar to the hydrophobic 'pore ring' in the SecY protein translocation channel (Van den Berg et al., 2004). This 'gasket-like' hydrophobic ring may play a role in preventing leakage of protons and subsequent PMF collapse during protein translocation.

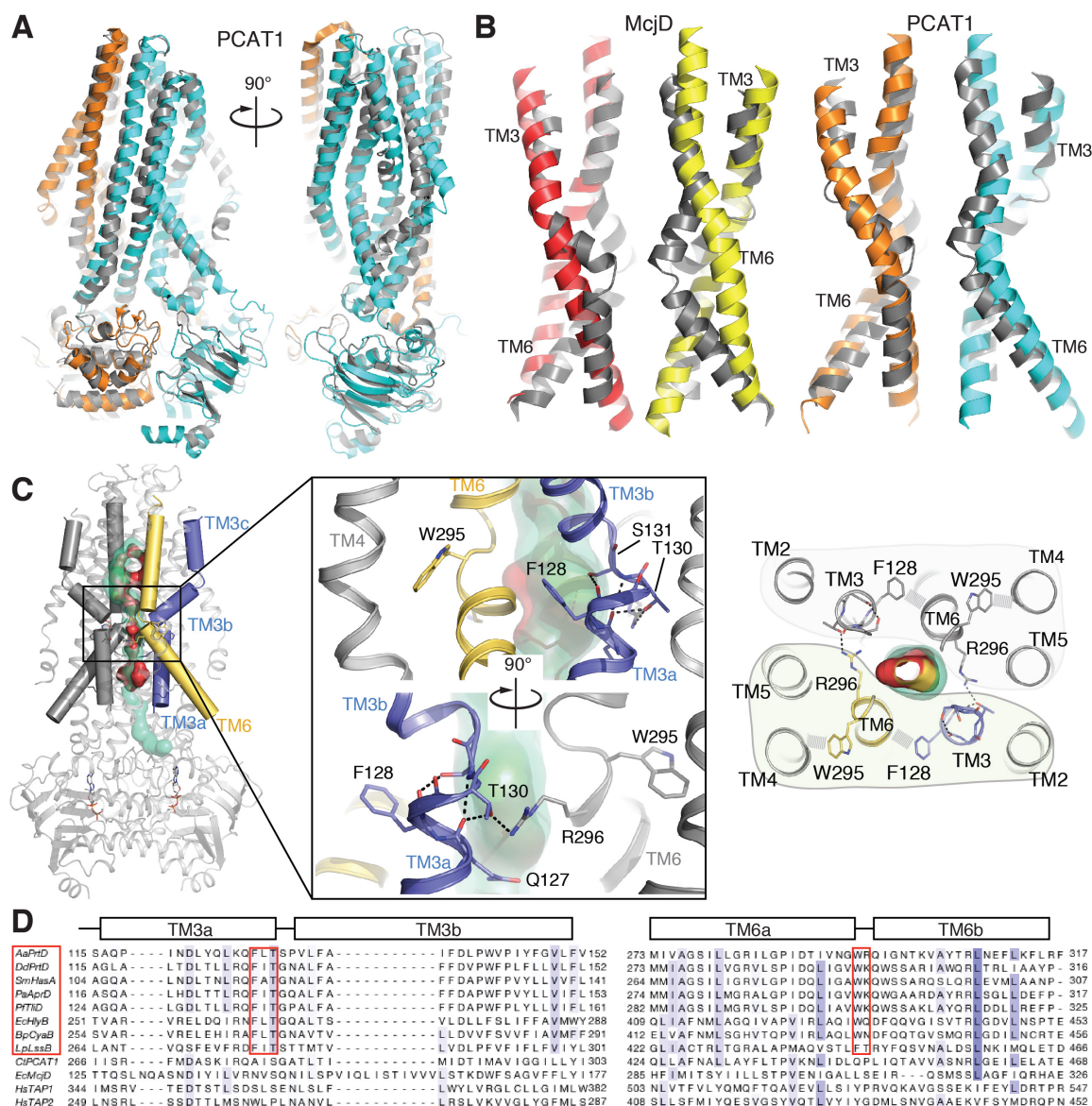


Figure 52 Comparison to peptide transporters. (A) Comparison between AaPrtD and PCAT1. PCAT1 (pdb:4S0F) is colored with one subunit orange and the other cyan. AaPrtD is colored dark grey. (B) Comparison of TM3 and 6. Left, McjD (pdb 4PL0) is shown as a dimer with one subunit colored yellow and the other red. Right, PCAT1 is colored as in (A). (C) Constriction of AaPrtD's pore by TM3 and 6. The TM channel is shown as in Figure 2D. AaPrtD TM3 and 6 are shown as a cylindrical cartoon with one subunit colored grey and the other colored blue and yellow, respectively. Key conserved residues stabilizing the kinks in TM3 and 6 are shown as sticks. Right panel: Interactions stabilizing the TM3 kink and connecting the PrtD protomers (shaded light grey and pale yellow, respectively) are shown. (D) Alignment of characterized TISS ABC transporters with peptide transporters. TISS transporters are indicated by a red box. The secondary structure is indicated above the alignment. The alignment sequences are colored according to identity.

T1SS substrates are most likely translocated in an unfolded, extended state. The PrtD pore dimensions observed in our structure are similar to the narrowest constrictions observed in the SecY channel containing a translocating polypeptide chain (Li et al., 2016). Using a smaller probe radius of 1.0 Å, the PrtD channel extends towards the NBDs and ends at a window towards the cytosol formed between TM4 and 6, right above the nucleotide-binding site (**Fig. 51D**). Although too narrow to accommodate an extended polypeptide chain in the current conformation, the pore might widen in a fully assembled and active T1SS.

4.2.5 Comparison to peptide transporters

The ADP-bound conformation of AaPrtD is most similar to the AMP-PNP-bound conformations of the bacteriocin transporters, PCAT1 and McjD (**Fig. 52A and 53B**), both of which are predicted to operate by ‘alternating access’ (Choudhury et al., 2014; Lin et al., 2015). In each case, the NBDs are closed around the bound nucleotides, yet the periplasmic exit of the transporter is also closed, which is in sharp contrast to the gaping periplasmic opening in the ADP-bound form of the multidrug extruder, Sav1866 (**Fig. 53C**) (Dawson and Locher, 2006).

The most striking difference between AaPrtD and the bacteriocin transporters is the sharp kinks in TM3 and TM6 (**Fig. 52B**). While these helices bow outward in McjD to create a spacious cavity, the inward kinks of AaPrtD form the most constricted region of the channel (**Left panel, Fig. 52B**). In PCAT1, TM3 and TM6 follow the same general path

as in AaPrtD, but the PCAT1 helices are continuous and the helical bends are far less pronounced (**Right panel, Fig. 52B**). TM3 in AaPrtD is especially distinct because it is broken up into three separate helices (TM3a-c), of which TM3a and 3b are connected by a 5 residue long loop (**Fig. 52C**). Hence, TM3 could not form a continuous helix unless the region either N-terminal (TM3a) or C-terminal to the break (TM3b) rotated by a quarter helical turn to accommodate the additional residue.

Each of the TM kinks of AaPrtD is stabilized by conserved residues. The TM3a-3b kink is enforced by an 'FxT' motif that is conserved in T1SS ABC transporters and not found in peptide transporters, such as PCAT1, McjD, and TAP1/2 (**Fig. 52D**). Similarly, the TM6 kink contains a conserved motif consisting of an aromatic residue (mainly W) followed immediately by a hydrophilic residue (R296 in AaPrtD) (**Fig. 52D**). The FxT motif stabilizes the TM3a-3b kink through multiple intra- and intermolecular interactions. First, F128 packs against TM6 of the same subunit while the side chain of the following T130 forms an H bond with the carbonyl oxygen of the preceding Q127. Q127, in turn, forms an H bond with the amide proton of S131. Second, R296 from TM6 of the opposite subunit reaches across to interact with the carbonyl of Q127 as well as the side chain hydroxyl of the FxT's T130. The conserved W295 of TM6 packs into a hydrophobic pocket of TM4 in the same subunit (**Fig. 52C**). Combined, the FxT motif stabilizes the TM3a-3b kink by anchoring residues 128-131 to TM6 of the same subunit while also tethering to TM6 of the opposite subunit through interactions with its R296 (**Fig. 52C**).

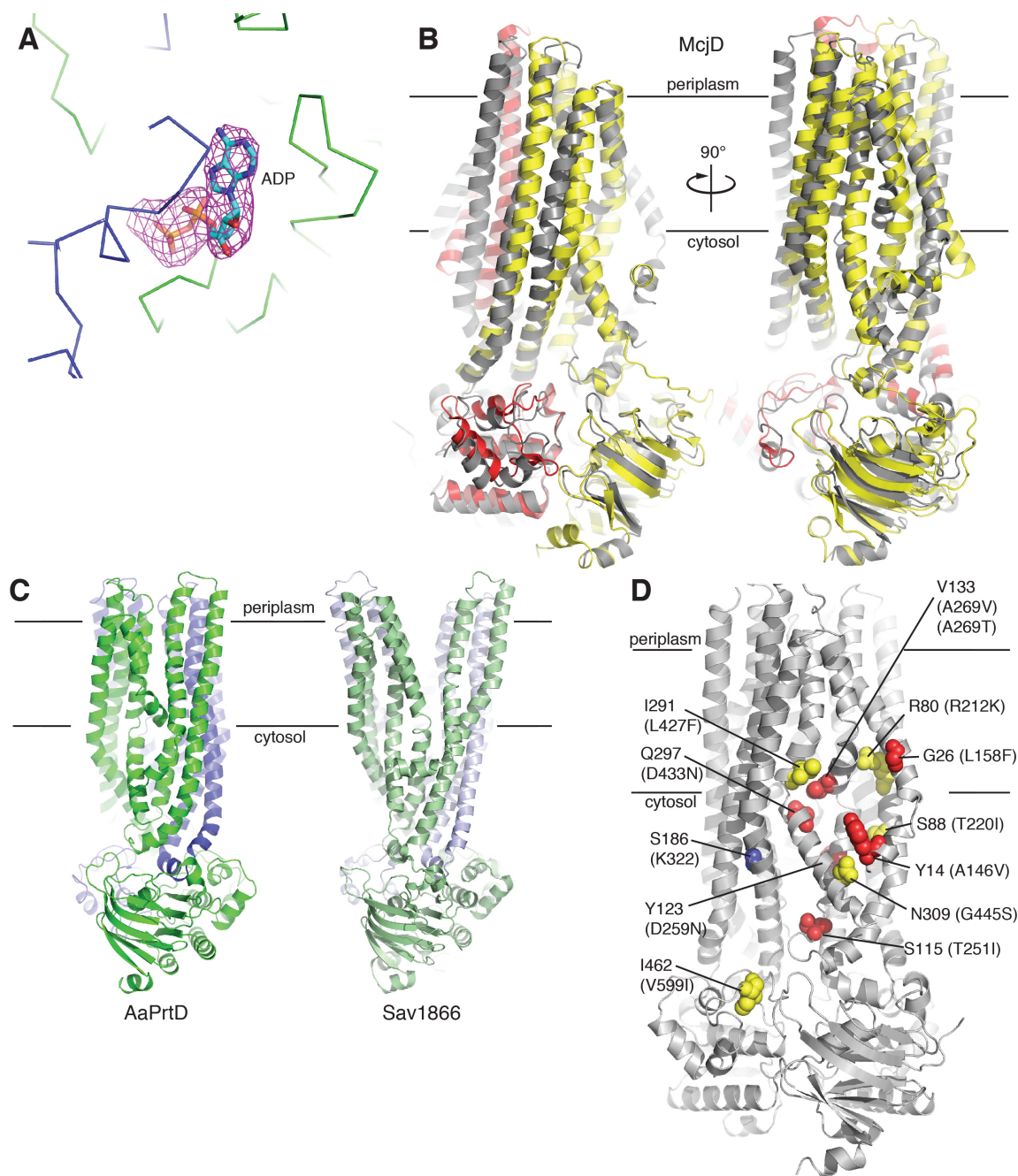


Figure 53 AaPrtD conformation. (A) Density for ADP. Unbiased $F_o - F_c$ difference electron density for ADP, contoured at 3.5σ , is shown as a magenta mesh. (B) Overlay of AaPrtD and McjD. McjD is shown as a yellow and red cartoon corresponding to different subunits of the homodimer. AaPrtD is shown as a dark gray cartoon. (C) Comparison of periplasmic openings of AaPrtD and Sav1866. AaPrtD is shown as in Figure 51A. Sav1866 is shown as a cartoon and colored light green and light blue corresponding to different subunits of the homodimer. (D) Location of residues corresponding to HlyB rescue mutants. AaPrtD is shown as a gray cartoon. HlyB and AaPrtD sequences were aligned as in Figure 3D. AaPrtD residues corresponding to HlyB mutants rescuing deletion of the N- or C-terminal half of the substrate's ~60 residue signal sequence are shown as spheres and colored yellow or red, respectively. Residue corresponding to substrate-crosslinked location is colored blue. AaPrtD residues are labeled, and corresponding HlyB mutations are shown in parentheses.

4.3 Discussion

The collective work on the PrtDEF system convey a model of substrate secretion whereby PrtD, PrtE, and PrtF diffuse independently of each other in the absence of substrate (Létoffé et al., 1996). PrtD displays a basal ATPase activity that is inhibited upon encountering substrate, perhaps by prevention of NBD closure in a manner similar to the inhibition of TAP1/2 by the viral peptide ICP47 (Delepelaire, 1994; Oldham et al., 2016). This inhibited conformation of PrtD binds PrtE, with the complex of PrtD and PrtE recruiting PrtF (Létoffé et al., 1996). Because ATP hydrolysis does not appear to be required to form an ABC transporter-MFP-TolC complex (Thanabalu et al., 1998; Masi and Wandersman, 2010) but is required for secretion, it seems likely that formation of the entire PrtDEF complex relieves the ATPase inhibition and allows substrate to be pushed across both membranes into the media. The stimulation of an ABC transporter's ATPase activity by an MFP has been previously demonstrated in the macrolide transport system, MacB–MacA–TolC (Tikhonova et al., 2007). Indeed, the MFP and TolC components seem to play more complex roles in T1SS than simple adapters. Cross-species complexes consisting of SmHasD-SmHasE-DdPrtF and DdPrtD-SmHasE-DdPrtF can secrete DdPrtC with the same efficiency as DdPrtD-E-F, indicating that DdPrtD is compatible with SmHasE and that SmHasE is compatible with DdPrtF. However, the DdPrtD-SmHasE-DdPrtF complex only secretes DdPrtC with 10% efficiency despite pairwise compatibility between each component in the complex (Binet and Wandersman, 1995).

In T1SSs, the ABC transporter-substrate interaction is complex. While the C-terminal secretion signal is required for secretion, experiments on the Has secretion system have demonstrated that expression of HasA substrates lacking the secretion signal can competitively inhibit secretion of WT HasA. Furthermore, mutations in specific regions of secretion-signal-deficient substrates can rescue the secretion of WT HasA, indicating that there are specific transporter-substrate interaction sites upstream of the secretion signal. Additionally, the secretion-signal-deficient substrate is sufficient to assemble the entire HasD-HasE-TolC complex despite being secretion incompetent (Masi and Wandersman, 2010). Experiments with the hemolysin secretion system also indicate that the MFP is also involved in binding substrate with specific regions of the cytoplasmic domain of HlyD being required for coupling the binding of substrate with the formation of a functional complex with TolC (Balakrishnan et al., 2001).

Several lines of evidence argue that substrate interacts with AaPrtD at the concave surface and near the cytosolic window where TM4 separates from TM6 (**Fig. 51**). First, random mutagenesis of the hemolysin transporter, HlyB, has revealed a number of suppressor mutations that rescue secretion of hemolysin carrying either a C-terminal truncation (**red residues, Fig. 53D**) (Zhang et al., 1993) or an internal deletion of the N-terminal half of the ~ 60 amino acid secretion signal (**yellow residues, Fig. 53D**) (Sheps et al., 1995). HlyB is able to secrete DdPrtB in small amounts (Delepelaire and Wandersman, 1990), indicating that HlyB likely recognizes its native substrate and DdPrtB through a similar mechanism. In the AaPrtD structure, all of the HlyB suppressor mutations are located near the cytosolic window and in the ‘concave bowl’ formed from

TM1, TM3, and TM6 (**Fig. 53D**). Furthermore, a study examining the effect of substrate on ATPase activity of purified HlyB found that interactions of the substrate's secretion signal with the N-terminal CLD of HlyB resulted in inhibition of ATP hydrolysis.

However, substrate regions upstream from the secretion signal also interact with HlyB and stimulate its ATPase activity (Reimann et al., 2016). The interaction interface of HlyB with this upstream region of substrate was determined to be at residue Lys322. The corresponding residue of PrtD, Ser186, is located on TM4 and sits right where TM4 and TM6 separate (**blue residue, Fig. 53D**). Additionally, this 'concave bowl' region as well as the TM4 and TM6 window region are highly basic (**Fig. 51C**), indicating that these regions could recruit acidic substrates through electrostatic interactions. Finally, the broken coupling helix leading into TM3 would be compatible with substrate entering through the cytosolic window between TM4 and TM6 (**Fig. 51B**). This region would need to be flexible in order to accommodate passage of polypeptide stretches with different sequences.

Currently we have no structural insights into an assembled PrtD-E-F complex. PrtF, the TolC homolog, most likely assembles into a trimer forming a large β -barrel porin capped on its periplasmic side by an α -helical cage (Koronakis et al., 2000). The recently determined structure of the soluble domain of the PrtE homologue, HlyD (Kim et al., 2016) confirmed its resemblance to AcrA, which mediates the interaction of the AcrB multi drug efflux pump with TolC. Similar to the architecture of the AcrB-AcrA-TolC complex (Du et al., 2014; Symmons et al., 2015), the PrtD-E-F complex may consist of a

periplasmic PrtE hexamer connecting the PrtD homodimer in the inner with a PrtF homotrimer in the outer membrane.

The most striking structural distinction between AaPrtD and the ‘alternating access’ peptide transporters are the sharp kinks in TM3 and TM6. The conservation of residues at these kinks as compared to peptide transporters suggests that the breaking of these helices plays a key role in the distinct mechanism of T1SS ABC transporters (**Fig. 52C and D**). Indeed, broken helices are directly implicated in interactions with and transport of unfolded polypeptides by translocating ATPases, such as ClpX or SecA (Erlandson et al., 2008; Martin et al., 2008; Bauer et al., 2014). Hence, it is tempting to speculate that AaPrtD interacts with the substrate at these kinks, yet we cannot exclude additional conformational changes upon complex formation with other T1SS components. An understanding of the mechanism by which these unique ABC transporters translocate unfolded polypeptides across the membrane will await structural analyses of an ABC transporter-MFP-TolC complex.

Author contributions

JLMW designed and performed experiments and wrote the manuscript. JZ designed experiments and wrote the manuscript.

Acknowledgements

We thank the beam line staff at APS GM/CA- and SER-CAT. Diffraction data were collected at the APS beam lines 23-ID-D (GM/CA-) and 22-ID (SER-). GM/CA@APS

has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006).

SER-CAT used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by

Argonne National Laboratory under Contract No. DE-AC02-06CH11357. J.L.W.M. was supported by a National Science Foundation Graduate Research Fellowship, Grant No.

DGE-1315231

Table 3

Chapter 4 crystallographic data collection and refinement statistics.

	Native	SmOAc-soaked	PbOAc-soaked	Se-Met
Data collection				
Space group	C 1 2 1	C 1 2 1	C 1 2 1	C 1 2 1
Cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> (Å)	118.6, 97.9, 179.8	116.8, 98.1, 177.8	117.0, 97.5, 178.3	117.6, 98.0, 179.2
<i>a</i> , <i>b</i> , <i>g</i> (°)	90, 100.5, 90	90, 103.2, 90	90, 102.2, 90	90, 102.8, 90
Resolution (Å)	34.3–3.12 (3.27–3.12)*	33.7–4.94 (5.52–4.94)	33.6–5.36 (5.99–5.36)	34.88–4.00 (4.47–4.00)
<i>R</i> _{pim}	0.061(0.509)	0.044 (0.195)	0.031 (0.114)	0.05 (0.199)
CC _{1/2} [^]	0.997 (0.647)	0.998 (0.940)	0.998 (0.971)	0.997 (0.955)
Mean <i>I</i> / <i>sI</i>	8.3 (1.3)	13.2 (4.6)	16.2 (6.5)	12.7 (5.2)
Completeness (%)	97.4 (84.0)	98.8 (97.0)	98.7 (96.7)	99.7 (99.6)
Redundancy	4.0 (3.2)	13.0 (12.7)	7.7 (7.7)	32.4 (32.6)
Refinement				
Resolution (Å)	33.5–3.15			
No. reflections				
Total	34,691			
<i>R</i> _{free}	1,729			
<i>R</i> _{work} / <i>R</i> _{free}	19.9/23.0			
No. atoms				
Protein	8,499			
ADP	52			
Mg ²⁺	2			
<i>B</i> -factors				
Chain A	144.4			
Chain B	147.4			
ADP	181.6			
Mg	210.8			
R.m.s deviations				
Bond lengths (Å)	0.003			
Bond angles (°)	0.736			

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

[^] Correlation between intensities from random half-data sets.

Table 3 (cont.)

Crystallographic data collection and refinement statistics.

	PrtD 103C Hg-soaked	PrtD 157C Hg-soaked	PrtD 193C Hg-soaked
Data collection			
Space group	C 1 2 1	C 1 2 1	C 1 2 1
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	119.0, 97.1, 174.9	121.0, 98.3, 176.7	117.4, 98.4, 178.6
<i>a</i> , <i>b</i> , <i>g</i> (°)	90, 102.8, 90	90, 102.8, 90	90, 102.5, 90
Resolution (Å)	34.3–4.92 (5.50–4.92)*	34.51–4.46 (4.99–4.46)	33.6–3.98 (4.45– 3.98)
<i>R</i> _{pim}	0.036(0.187)	0.023 (0.068)	0.025 (0.111)
CC _{1/2} [^]	0.999 (0.953)	0.999 (0.989)	0.999 (0.978)
Mean <i>I</i> / <i>sI</i>	12.9 (4.2)	19.5 (11.1)	18.8 (6.7)
Completeness (%)	98.0 (94.2)	98.4 (95.1)	99.3 (98.1)
Redundancy	7.5 (7.4)	13.1 (12.8)	6.6 (6.6)

4.4 Methods

4.4.1 Plasmid construction

Dickeya dadantii *prtD*, *prtE*, *prtF* and *prtG* genes were synthesized by Genewiz. *prtD* contained an upstream NdeI site and a downstream XhoI site. *prtF* contained an upstream EcoRI site, and engineered Flag tag on the C-terminus and a downstream HindIII site. *prtE* contained an upstream NcoI site and a downstream BamHI site. Following the BamHI site, the *prtE* gene contained a copy of the sequence of pACYCduet vector (Novagen) from 163 to 293 followed by an EcoRI site in order to generate an additional multiple cloning site in which to clone *prtF*. These components were cloned into a pACYCduet vector with the appropriate restriction sites giving PrtE an N-terminal His tag, PrtF a C-terminal Flag tag, and PrtD a C-terminal S tag. The *prtG* gene was synthesized with an upstream BglII site and a downstream HindIII site. Cloning *prtG* into pBAD/HisC vector (Invitrogen) generates an N-terminal His tag. A myc tag was inserted following residue 419 of the amino acid sequence using PCR overlap extension. The E492Q mutant PrtD was generated by site-directed mutagenesis.

Aquifex aeolicus *prtD* was amplified from genomic DNA (VF5 strain) with an upstream NdeI site and a downstream XhoI site. This construct was cloned into a pET30a vector (Novagen) to give a C-terminal His tag. S103C, Y157C, and S193C were generated by site-directed mutagenesis.

4.4.2 Whole cell secretion assays

E. coli C43 cells were transformed with pACYC vector encoding the transporter subunits and a pBAD vector encoding myc-tagged PrtG. The cells were inoculated into 100 ml of LB media supplemented with 100 µg/ml ampicillin and 25 µg/ml chloramphenicol and were grown at 37°C with shaking at 220 rpm. When the cells reached an OD₆₀₀ of ~0.6, 12 µg/ml of IPTG was added to induce expression of the transporter components. After 4.5 hr after IPTG induction, arabinose was added to induce PrtG expression. At the indicated time points, 10 ml of media were spun at 4000 rpm, 4°C, 10 min. The supernatant was then passed through a 0.22 µm filter. 1 ml of filtrate was precipitated by addition of 250 µl ~100% TCA. After incubating on ice for 30 min, the sample was spun at 14k rpm for 30 min. The supernatant was then discarded, and the pellet was resuspended in 20 µl 1.5 M Tris pH 8.8 and 20 µl reducing gel loading buffer by vigorous vortexing. The entire sample was separated on a 10% SDS PAGE gel and blotted for the α-myc tag on PrtG α-myc primary (Invitrogen) and IRDye800-conjugated α-mouse secondary (Rockland) antibodies. To determine the total protein expression levels, the cell pellet was resuspended in 100 µl loading buffer, 100 µl 20 mM Tris pH 7.5, 100 mM KCl. The sample was sonicated continuously using a microtip for 15 seconds. The sample was then placed on ice for 2-3 min and the sonication was repeated. 20 µl of sample was loaded onto a 10 % SDS PAGE gel and blotted for myc-tagged protein.

4.4.3 Spheroplast preparation and secretion assays

Cells were transformed and inoculated as with whole cell secretion assays. At an OD₆₀₀ of ~0.6, IPTG was added to 12 µg/ml to induce transporter components. After 4.5 hr with

IPTG, 10 ml of cell culture was diluted into 100 ml of fresh LB media containing 100 $\mu\text{g/ml}$ ampicillin, 25 $\mu\text{g/ml}$ chloramphenicol, and 12 $\mu\text{g/ml}$ IPTG. When the cells reached an OD_{600} of ~ 0.5 , 40 ml of culture was spun at 4000 rpm for 10 min, 4°C . The cell pellet was gently resuspended in 2.5 ml TES buffer (200 mM Tris pH 7.5, 500 mM sucrose, 1 mM EDTA) (Quan et al., 2013). Lysozyme was added to a final concentration of 2 mg/ml, and the cultures were incubated on ice. After 15 minutes, 20 mM MgCl_2 was added to stop the reaction, and the cells were spun at 4000 rpm, 4°C for 10 min (Koronakis et al., 1997). The supernatant was discarded, and the spheroplasts were gently resuspended in 40 ml of LB media supplemented with 250 mM sucrose, 100 $\mu\text{g/ml}$ ampicillin, 10 mM CaCl_2 . Based on the formation of round spheroplasts visible by light microscopy, conversion to spheroplasts was estimated to be $\geq 90\%$. 2% arabinose was added to induce PrtG. At the indicated time points, 10 ml culture was spun at 4000 rpm, 4°C for 10 min. The supernatant was passed through a $0.22 \mu\text{m}$ filter, then 2.5 ml $\sim 100\%$ TCA was added and the samples were incubated overnight on ice. The samples were spun at 4000 rpm, 4°C for ~ 2.5 hr. The supernatant was discarded, and the pellets were dried carefully. The pellets were then resuspended in $\sim 30 \mu\text{l}$ 1.5 M Tris pH 8.8 and $\sim 20 \mu\text{l}$ gel loading buffer. The entire sample was loaded onto a 10% SDS PAGE gel and blotted for myc-tagged protein.

In spheroplast experiments, the pellet was treated as the cell pellet above but was blotted for α -myc (Invitrogen), α -S (Millipore), α -FLAG (Sigma), and α -HIS (Qiagen) using an α -mouse secondary (Rockland).

4.4.4 AaPrtD purification

AaPrtD in pET 28a was transformed into *E. coli* C43 cells. The cells were inoculated into 3L of LB media containing 25 µg/ml kanamycin. The cells were grown at 37°C with 220 rpm shaking until the OD₆₀₀ reached ~0.6. At this time, the temperature was dropped to 30°C. After 20 min at 30°C, expression of AaPrtD was induced by the addition of 100 µg/ml IPTG. After 3 hrs of induction, the cells were harvested by spinning at 5500 rpm, 4°C for 10 min. The cell pellet from 3L culture was then resuspended in 50 ml Buffer A (20 mM Tris pH 7.5, 100 mM NaCl). The cells were then passed twice through a microfluidizer at >25 kPa pressure. ~50 mg PMSF (dissolved in 500 µl DMSO) was added between passes. The sample was then spun at 100,000g at 4°C for 1 hr to pellet the membranes. The supernatant was discarded and the membranes were flash frozen in liquid N₂ and stored at -80°C. The membranes from 3L culture were thawed in a 30°C water bath and resuspended in 50 ml Buffer B (50 mM sodium phosphate, pH 7.2, 100 mM NaCl, 20 mM imidazole, 2% Triton X-100). PMSF was added as before, and the sample was incubated at 4°C with rocking. After 1 hr, the sample was spun at 100,000g for 30 min. The supernatant was mixed with 5 ml Ni²⁺-NTA beads (Qiagen) and incubated with rocking at 4°C for 1 hr. The slurry was then added to a gravity flow column, the flow-through was collected and reapplied to the column. The beads were then washed with 110 ml W1 (Buffer A + 46 mM imidazole + 1mM dodecyl-b-D-maltoside (DDM)) followed by 40 ml W2 (Buffer A + 20 mM imidazole + 1mM DDM + 1.5 M NaCl). The protein was then eluted with 50 ml elution buffer (Buffer A + 270 mM imidazole + 1mM DDM). The eluted sample was concentrated to ~2 ml in a 100 kDa

cutoff spin concentrator, passed through a 0.22 μ M filter, and loaded onto a 120 ml S200 (16/60) column equilibrated with Buffer A + DDM (GF buffer).

Selenomethionine (Se-Met)-derivatized AaPrtD was produced by inhibition of the methionine-synthesis pathway as described (Morgan et al., 2013). Briefly, the cells were grown at 33°C in M9 media containing 0.2% glucose, 5% glycerol, 2mg/L thiamine, 0.1 mM CaCl_2 , and 2 mM MgSO_4 . When the OD_{600} reached 0.4, the temperature was reduced to 30°C. After 90 minutes, 100 mg each of Lys, Thr, and Phe while 50 mg each of Leu, Ile, and Val were added to inhibit methionine synthesis. Then, 60 mg Se-Met (Anatrace) was added. After 1 hr, IPTG was added at 500 mg/L final concentration. After 6 hours, cells were harvested and the Se-Met-derivatized protein was purified in the same way as native AaPrtD except that all buffers contained 5 mM β -mercaptoethanol, except for gel filtration buffer, which contained 0.5 mM TCEP.

4.4.5 AaPrtD structure determination

Fractions containing AaPrtD were collected and concentrated to ~10 mg/ml in an Amicon spin concentrator (100 kDa cutoff). The sample was then spun at 180,000g, 4°C for 20 min. Afterwards, MgCl_2 was added to 5 mM, and ADP was added to 2 mM. Initial crystals grew at 24°C in 16-21% PEG 1000, 100-400 mM $(\text{NH}_4)_2\text{SO}_4$, with 10 mM MES pH 6.5. Crystals, belonging to space group C2, appeared within 4 days and grew to a maximum size of ~0.5 x 0.5 x 0.5 μ M in ~2 weeks but diffracted poorly. Addition of 1x CMC octyl- β -D-glucoside (OG) to the protein immediately before setting up trays gave clusters of plate-like crystals that appeared after about 1 week, and grew to a maximum

size over 2 weeks. Crystals were cryoprotected by gradual replacement of the mother liquor with a solution containing 25% PEG 1000, 100 mM $(\text{NH}_4)_2\text{SO}_4$, 10 mM MES pH 6.5, 15% glycerol, 5 mM MgCl_2 , ADP, 20 mM OG. These crystals also belonged to space group C2 and diffracted to a maximum resolution of 3.15 Å. Diffraction data were collected at Advanced Photon Source beam lines, SER-CAT and GM/CA CAT. The data were processed in XDS (Kabsch, 2010). Initial phases were obtained using single-wavelength anomalous diffraction in ShelX (Sheldrick, 2010) on crystals derivatized by soaking with 10 mM samarium acetate for 30 min. These initial phases were used to calculate anomalous difference electron density maps for crystals soaked with lead acetate, for 3 single-cysteine derivatives (S103C, Y157C, and S193C) soaked with 10 mM mercury chloride for 10 min, as well as Se-Met derivatized PrtD. 10 samarium, 7 lead, 6 mercury, and 20 selenium positions were used to calculate new phases in MLPhare implemented in the CCP4 software package (Winn et al., 2011). Phases were extended to 3.15 Å using density modification with NCS averaging in Resolve (Terwilliger et al., 2008) as part of Phenix (Adams et al., 2010). The model was built in Coot (Emsley et al., 2010) and refined in Phenix_Refine (Adams et al., 2010) with TLS parameterization (Painter and Merritt, 2006a). All figures were generated using PYMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.). Internal PrtD cavities and tunnels were calculated using HOLLOW (Ho and Gruswitz, 2008) and CAVER (Kozlikova et al., 2014). The structure coordinates have been deposited in the protein data bank with the PDB code 5L22.

Chapter 5: Perspectives and conclusions

Over the past four years of graduate school, I have helped to take our molecular understanding of cellulose biosynthesis from a single snapshot of the inactive (without c-di-GMP) BcsA–B enzyme to a detailed molecular movie that illustrates the entire process of cellulose biosynthesis from c-di-GMP binding and allosteric activation by c-di-GMP to the conformational changes that couple synthesis with membrane translocation.

In Chapter 1, I described the first structure of the BcsA–B complex, which Jochen and I determined in the months leading up to my start of graduate school. This structure was hugely important in promoting our understanding of cellulose biosynthesis. It revealed how BcsA forms a channel for the translocating polysaccharide; the architecture of the GT domain, its position relative to the TM domain, and how the signature residues are positioned for catalysis; the position of the PilZ domain and its orientation relative to the the FT domain; and the architecture of BcsB as well as the basis for its interaction with BcsA. The initial BcsA–B structure gave us ideas about the activation by c-di-GMP as well as the coupling of synthesis and translocation. While many of these ideas turned out to be wrong, they still provided a foundation to build our development of more detailed questions and to interpret structural changes. On a more technical note, the initial BcsA–B structure, for which we determined phases experimentally, provided a model that I could use for molecular replacement, which made the later crystallography more manageable.

In Chapter 2, I presented the structure of BcsA–B bound to c-di-GMP. These crystals were obtained in bicelles, which is likely a more physiological environment for the protein than detergent. So, this structure served both as an important validation of the initial structure that was crystallized in detergent micelles while also revealing the mechanism by which c-di-GMP activates cellulose biosynthesis. Both the position of the ‘gating loop’ blocking the active site as well as the kinetic data showing that c-di-GMP-binding activates BcsA–B without changing K_M suggested that the c-di-GMP binding would result in movement of the gating loop. Indeed, the structure of c-di-GMP-bound BcsA–B in the absence of substrate clearly showed that the gating loop had moved from the front of the active site to a stable position above the active site, thus opening the active site. However, it was not immediately clear from the structure why the gating loop had moved, and it took re-examination of the initial structure in the absence of c-di-GMP to reveal that it was actually auto-inhibited by an interaction between R580, one of the signature c-di-GMP-binding residues of the PilZ domain, and the backbone of the gating loop. Upon encountering c-di-GMP, R580 rotates 180° in order to coordinate it. This rotation brings R580 away from the position where it can interact with the gating loop and thus releases the gating loop so that it can move to the more stable position above the active site. Soaking Mg^{2+} and UDP, which is a product of cellulose synthesis but is also a competitive inhibitor, into these crystals revealed another function of the gating loop beyond a moveable barrier. The loop clearly inserts into the active site where each of the conserved residues is involved in stabilizing UDP. Because UDP acts as a competitive inhibitor, its binding mode likely mimics substrate binding. Thus, it seems probable that gating loop insertion into the active site plays an important role in facilitating catalysis.

This work revealed for the first time the dynamic role of the gating loop in cellulose biosynthesis.

In Chapter 3, I present *in crystallo* experiments on BcsA–B, which is an even more convincing validation of the structure because the protein can undergo an entire cellulose biosynthesis cycle in the crystal. I discovered this remarkable feature of these crystals when I soaked UDP-Glc into the crystals in hopes of determining a substrate-bound state. I was surprised to see that the cellulose polymer was extended by a single glucose unit meaning that catalysis had actually occurred in the crystal. Intrigued by the idea of being able to experimentally separate synthesis from translocation, I extended the polymer by soaking the crystals with substrate and then exchanged the solution in order to dilute out the substrate and thus prevent additional extensions. With the polymer extended, I then performed experiments to try to find conditions that cause the polymer to translocate into the TM channel. I found that the pre-translocation state containing an extended polymer is quite stable in the absence of substrate. Extension of the polymer, dilution of substrate, and allowing the crystals to sit overnight did not support translocation. Instead, the ‘finger helix’, a short helix near the acceptor site which contains the catalytic base at its N-terminal tip, moved downwards to again establish an interaction with the newly-added sugar of the polymer. When these same crystals are subsequently soaked with UDP and Mg^{2+} , the terminal sugar disappears within 15 minutes. Our collaborators at University of British Columbia synthesized a UDP-thiogalactose derivative that allowed us to unambiguously identify the newly-added sugar, which then allowed us to confirm that the

disappearance of the newly-added sugar upon adding UDP is indeed due to its translocation into the BcsA channel.

In order to test whether movement of the finger helix is necessary for polymer translocation, my close friend and colleague, Josh McNamara, engineered a mutant BcsA with a Cys residue added on the finger helix and a Cys residue added to an adjacent helix. These Cys residues form a disulfide bond under oxidizing conditions and inhibit the steady-state activity of BcsA–B. These initial functional experiments somewhat suggest that immobilization of the finger helix prevents cellulose translocation, but the ability to separately observe cellulose synthesis and membrane translocation in the crystal allowed me to study the consequence of this disulfide mutant in a much more sophisticated system. I crystallized the double-Cys mutant and began by demonstrating that the formation of the disulfide bond does actually immobilize the finger helix. I then directly tested whether immobilization of the finger helix actually inhibits translocation. By extending the polymer then soaking UDP under oxidizing or reducing conditions, I demonstrated that the cellulose polymer only translocates when the disulfide bond is reduced. Thus, movement of the finger helix is necessary for cellulose polymer translocation. Again, it was not immediately clear why binding of UDP promotes finger helix movement and subsequent cellulose translocation. Upon close examination of the UDP-bound structure presented in Chapter 2, we noticed a close association between the gating loop and a rigid region upstream of the finger helix. This rigid region contains a conserved beta-branched residue (either Ile or Val), which we thought would make sense to transmit force from the gating loop to the finger helix. We tested this idea by mutating

this Ile residue, and we found that Ile and Val are interchangeable, whereas Leu, Ala, and Thr gave low levels of activity. These data support our idea that this residue is responsible for coupling insertion of the gating loop with upward movement of the finger helix.

These experiments along with Josh's soaking of crystals with a non-hydrolyzable UDP-Glc analogue give us structures of each state of the entire cellulose biosynthesis cycle. Some of the key structural features of cellulose synthase that our work revealed are the gating loop and the finger helix. Aside from an apparent rotation of the PilZ domain upon binding c-di-GMP, the motions of the gating loop and finger helix seem to be the only conformation changes necessary for the cellulose biosynthesis cycle. While the finger helix contains one of the signature residues of cellulose synthase, the importance of the gating loop was completely unexpected and appears to be a general feature of cellulose synthases (McNamara et al., 2015).

In Chapter 4, I present the structure of a T1SS ABC transporter. I began this project based on an interest in alternative biopolymer translocation systems. At the time, I thought that the T1SS ABC transporters were alone sufficient to transport a polypeptide across the membrane, which would likely make these ABC transporters the simplest polypeptide-transporting systems. I was fortunate to get crystals quickly after purifying the PrtD protein. However, it took another nine months to optimize the crystals to diffract to a useful resolution. Then, it took an additional three months to obtain the phase information that allowed me to determine the structure. The structure looks fairly similar

to other ABC exporters with the exception that it contains a channel instead of a large cavity and it contains highly-kinked helices that form a constriction in the channel near the cytoplasmic-membrane interface. Unable to envision a transport mechanism based on the structure, I tried to establish a functional assay to measure transport in a purified system composed of PrtD and the putative substrate. After almost a year with no results, I realized that this experimental approach has a fundamental flaw. It is designed to prove my hypothesis right instead of wrong. In other words, I would only learn whether isolated PrtD is sufficient to transport a polypeptide across the membrane is if it is indeed sufficient. So, I optimized an alternative approach, which was based on an experiment on the similar hemolysin secretion system (Koronakis et al., 1997). I expressed the components from a well-characterized homologous T1SS in *E.coli* and generated spheroplasts, which lack the OM. Then, I expressed the substrate and monitored its secretion into the media. Using this approach, I could optimize the experiment based on the complete and functional T1SS. I could then express only the PrtD component and test whether substrate is secreted into the media. Interestingly, neither isolated PrtD nor PrtD with PrtE (and no PrtF) could transport the substrate across the IM. So, isolated PrtD can bind substrate but cannot transport it until it forms a complex of all three components. These experiments directed our interpretation of the PrtD structure and gave us ideas about how it binds substrate.

Had I carried out this project with a better understanding of the literature and a more critical perspective, I probably would have approached the problem differently. It has been shown that complete ABC transporter–MFP–TolC complexes can be co-precipitated

by pulling on a substrate that contains a tightly-folded N-terminal tag such as GST (Létoffé et al., 1996). The tightly-folded domain gets stuck and inhibits the ABC transporter while stabilizing the whole T1SS complex. With the recent developments that allow near-atomic resolution structure determination by cryo electron microscopy, the structure of an entire assembled T1SS does not seem far-fetched. Perhaps focusing on this aspect from the beginning would have yielded more insight into how the T1SS works.

It is interesting to note that BcsA is not actually a transporter but is a genetic fusion between a passive TM channel and motor protein formed by the GT domain. This general transport scheme is reminiscent of co- and post-translational protein transport systems in which SecY forms a passive channel onto which the motor formed by the ribosome or SecA, respectively, docks to push the polypeptide through SecY (Park and Rapoport, 2012). While PrtD is an ABC transporter and is thus a fusion of a TM domain and ATPase domain, it is difficult to envision from the structure how the NBDs could be responsible for pushing the polypeptide substrate through the TM. The NBDs of PrtD are unremarkable, while the TM region contains structural features that differentiate it from canonical ABC transporters. Therefore, it seems likely that the TM region of PrtD is actually responsible for pushing the substrate, which would make its mechanism unique among the currently characterized biopolymer transport systems. However, we will need to wait for future experiments to reveal the transport mechanism of a T1SS.

Appendix 1: T1SS *in vitro* functional experiments

In this appendix, I describe in detail experiments that I performed in order to try to establish an *in vitro* transport assay with the T1SS ABC transporter. I began these experiments in mid 2015 and continued them until early 2016.

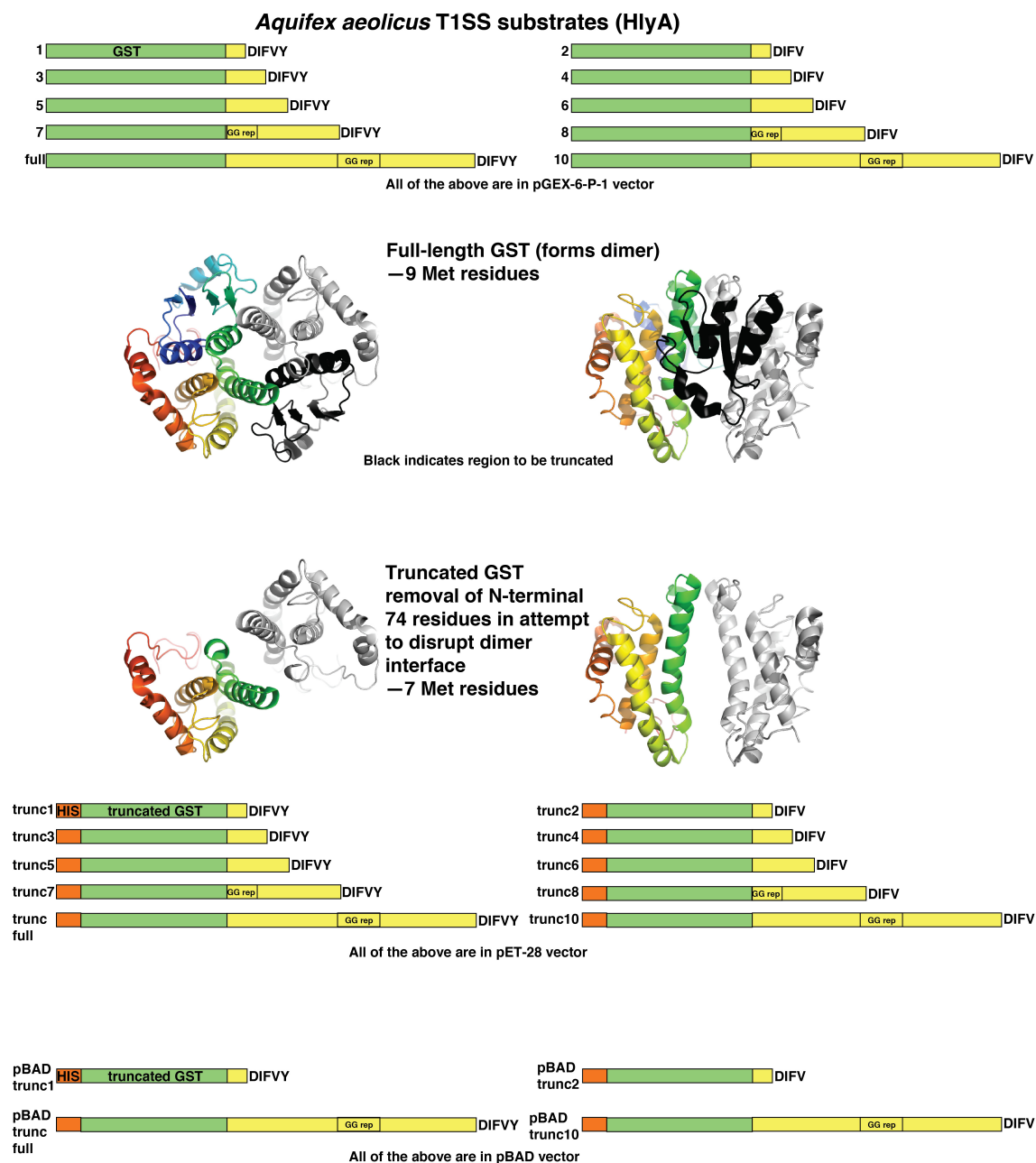
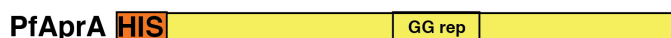


Figure 54 Cartoon of *Aquifex* T1SS substrate constructs.

Dd and *Pf* T1SS substrates

Pseudomonas fluorescens AprA (alkaline protease)



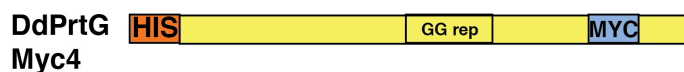
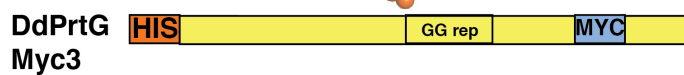
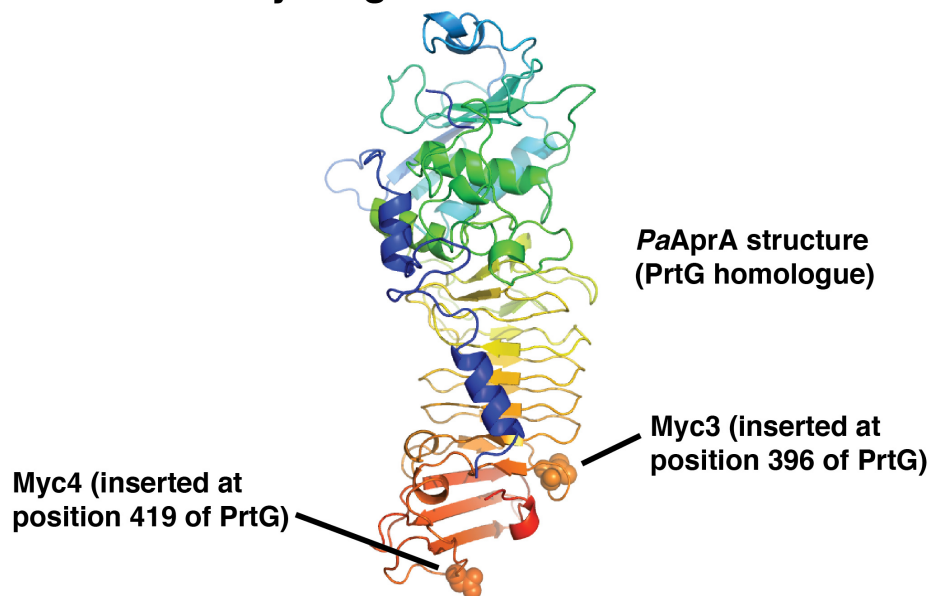
Available in both pet30a and pBAD (pBAD expression killed cells)

Dickeya dadantii PrtG (secreted protease)



Available in both pet30a and pBAD

Myc-tag insertions in DdPrtG



Both in pBAD vector

Figure 55 Cartoon of *Pf* and *Dd* T1SS substrate constructs.

App1.1 Expression, purification and reconstitution of AaPrtD

1. transform AaPrtD in pET 28a into E. coli C43 cells.
2. Inoculated cells into 3L of LB media (25 mg/L kanamycin).
3. Grow cells at 37°C with 220 rpm shaking until the OD₆₀₀ reaches ~0.6.
4. At this time, drop temperature to 30°C.
5. After 20 min at 30°C, induce expression of AaPrtD by adding 100 µg/ml IPTG
6. After 3 hrs of induction, the harvest cells by spinning at 5500 rpm, 4°C for 10 min
7. Resuspended cell pellet from 3L culture in 50 ml Buffer A (20 mM Tris pH 7.5, 100 mM NaCl)
8. Pass cells twice through a microfluidizer at >25 kPa pressure. Add ~50 mg PMSF (dissolved in 500 µl DMSO) between passes.
9. Spin sample in 45Ti at 42,000 rpm, 4°C for 1 hr to pellet the membranes.
10. Discard supernatant and flash freeze membranes in liquid N₂
11. Store at -80°C.
12. Thaw membranes from 3L culture in a 30°C water bath and resuspend in 50 ml Buffer B (50 mM sodium phosphate, pH 7.2, 100 mM NaCl, 20 mM imidazole, 2% Triton X-100).
13. Add PMSF as before
14. Incubate sample at 4°C with rocking for 1 hr.
15. Spin sample at 100,000g for 30 min
16. Mix supernatant with 5 ml Ni²⁺-NTA beads (Qiagen)
17. Incubate with rocking at 4°C for 1 hr
18. Add slurry to a gravity flow column
19. Collect the flow-through and reapply to the column
20. Wash beads with 110 ml W1 (Buffer A + 46 mM imidazole + 1mM DDM)

21. Wash beads with 40 ml W2 (Buffer A + 20 mM imidazole + 1mM DDM + 1.5 M NaCl)
22. Elut protein with 50 ml elution buffer (Buffer A + 270 mM imidazole + 1mM DDM).
23. Concentrate eluted sample was concentrated to ~2 ml volume in a 100 kDa cutoff spin concentrator
24. Pass concentrated sample through a 0.22 μ M filter
25. Load onto a 120 ml S200 (16/60) column equilibrated with Buffer A + DDM (GF buffer)
26. Collect fractions containing AaPrtD
27. Concentrate to OD₂₈₀ \approx 4 in an Amicon spin concentrator (100 kDa cutoff)
28. Reconstitute at 3 different lipid:protein (mass ratio) in 1:1 E coli total lipid (Ec TL):egg phosphatidylcholine (eggPC) at 4 mg/ml final lipid concentration
 - for 1 ml at 14:1 ratio: add (in order) 520 μ l GF buffer , 200 μ l Ec TL (10 mg/ml in 80 mM DDM), 200 μ l eggPC (10 mg/ml in 80 mM DDM), 80 μ l protein
 - for 1 ml at 30:1 ratio: add (in order) 564 μ l GF buffer , 200 μ l Ec TL (10 mg/ml in 80 mM DDM), 200 μ l eggPC (10 mg/ml in 80 mM DDM), 36 μ l protein
 - for 1 ml at 60:1 ratio: add (in order) 582 μ l GF buffer , 200 μ l Ec TL (10 mg/ml in 80 mM DDM), 200 μ l eggPC (10 mg/ml in 80 mM DDM), 18 μ l protein
29. Incubate mixtures on ice for 1 hr
30. Add scoop (~500 μ l) biobeads (suspended in H₂O) to fill tube
31. Rotate at 4°C for ~18 hr until mixture appears turbid
32. Carefully remove mixture from tube (try not to pipette biobeads) and transfer into new tube
33. spin at 14,000 rpm in 4°C for 30' to pellet proteoliposomes (PLs)
34. Discard supernatant
35. Resuspend PLs in 1ml 20 mM Hepes pH 7.5, 100 mM KCl
36. Repeat spin and resuspension to remove all detergent

37. In final resuspension, use 500 μ l to give a final lipid conc. of 8 mg/ml
38. Split into 20 μ l aliquots, snap freeze in liquid N₂ and store at -80°C

App1.2 *In vitro* transcription and translation of ^{35}S -Met labeled T1SS substrates

Uses E. coli T7 S30 Extract system for circular DNA (Promega)

—Use autoclaved water and freshly-autoclaved pipette tips and tubes to limit RNase contamination

—Upon receiving ^{35}S -methionine, split into ~10 μl aliquots and store at -80°C

—Carefully split (into autoclaved tubes) S30 premix (w/o amino acids) into 20 μl aliquots, and T7 S30 extract into 15 μl aliquots and store at -80°C to prevent multiple freeze-thaw cycles

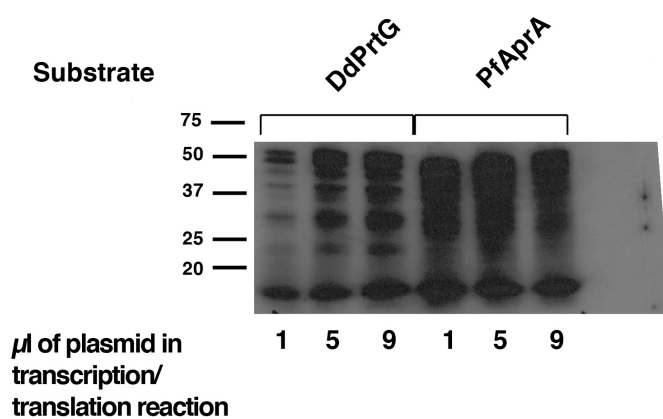
—use plasmid from miniprep (make sure construct has T7 promoter)

—it's best to optimize reaction by testing different plasmid concentrations

1. To tube, add in the following order: 20 μl S30 premix, 15 μl T7 S30 extract, 5 μl amino acid mixture (minus methionine), 1 μl ^{35}S -methionine
2. add between 1 and 9 μl of plasmid (stock plasmid conc. will probably be around 120 ng/ μl from miniprep)
3. add between 0 and 8 μl autoclaved water to bring final reaction volume to 50 μl
4. Place reactions at 37°C for 1 hr.
5. Place reactions on ice
6. load 1 μl onto gel
7. split remainder into 10 μl aliquots, freeze in liquid N_2 and store at -80°C
8. run gel
9. rock gel in ~100 ml of solution containing 40% ethanol and 3% glycerol for $\geq 30'$
10. dry gel for 2 hr at 80°C
11. expose gel in autoradiography cassette for 12–36 hr at -80°C
12. develop autoradiograph in biochemistry dept.

13. If denatured sample is desired, precipitate 50 μ l reaction with 13 μ l 100% TCA, incubate on ice for 30', spin at 4°C for 30', remove supernatant, and resuspend pellet in 25 μ l 8 M urea by vigorous vortexing (will likely take >24 hr of continuous vortexing)

Experiment from 12/22/2015



Transcription/translation reaction
loaded 2 μ l onto gel
exposed at -80°C for 8 hr

App1.3 Basic AaPrtD proteoliposome (PL) protease protection translocation experiments (variations will be described on specific experiments)

—Experiment is based on protease protection: if a radioactively-labeled polypeptide substrate is transported to the lumen of a vesicle, it will be protected from digestion by a robust, non-specific protease. Because the transport protein is an ATPase, the omission of ATP should result in no protection from the protease because the polypeptide substrate will not be transported to the lumen. Additionally, dissolving the vesicles with a detergent should also result in no protection because there will not be a lumen for the polypeptide substrate. Therefore, the ability to observe protection of a fragment of the radioactively-labeled substrate acts as an assay to monitor transport.

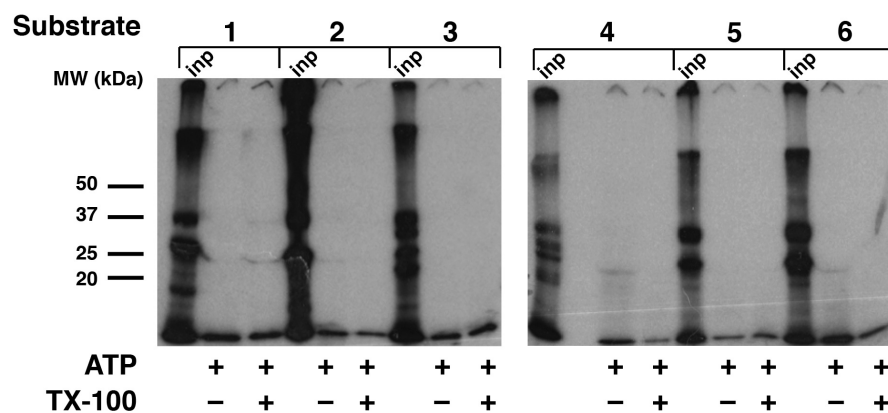
—50 μ l final reaction mixture contains 50 mM HEPES pH 7.5, 100 mM KCl, 2 mM MgCl_2 , 2 mM DTT, 2 mM ATP, 2mM MgCl_2 + PLs and substrate

—each experiment requires at least 3 reactions: one with ATP, one with no ATP or with an ATPase inhibitor such as vanadate (3 mM) or EDTA (3 mM), and one with ATP that will be dissolved with TX-100 as a negative control

1. Make master mix containing HEPES, KCl, MgCl_2 , and DTT
2. Aliquot into appropriate number of tubes
3. Add 2 μ l substrate
4. Add 1.5 μ l PLs
5. Lastly, add ATP to 2 mM
6. Incubate at the appropriate temperature for the appropriate time
7. Prepare fresh 10 mg/ml proteinase K (protK) solution dissolved in 50 mM Tris pH 8, 10 mM CaCl_2
8. Add TX-100 to 1% to negative control sample
9. Add protK to PL reaction to 1 mg/ml final conc.
10. incubate at room temp for 1 hr
11. Add 15 μ l TCA (~100% solution) and mix thoroughly by tapping tube
12. Incubate samples on ice for 30'
13. Spin samples at 14,000 rpm, 4°C, for 30'

14. Carefully remove supernatant (try to get all of it without getting pellet)
15. Resuspend pellet mixture of 20 μ l reducing SDS gel loading buffer and 20 μ l 1.5 M Tris pH 8.8
16. Shake vigorously until pellet is fully dissolved
17. Load 10-lane gel with 0.2 μ l substrate in one lane, ATP-containing sample in separate lane, ATP-omitted sample in separate lane, TX-100 containing sample in separate lane. Make sure to have at least 1 empty lane between radioactive samples
18. Run gel
19. Rock gel in \sim 100 ml of solution containing 40% ethanol and 3% glycerol for $\geq 30'$
20. Dry gel for 2 hr at 80°C
21. Expose gel in autoradiography cassette for 12–36 hr at -80°C
22. Develop autoradiograph in biochemistry dept.

Experiment from 4/1/2015

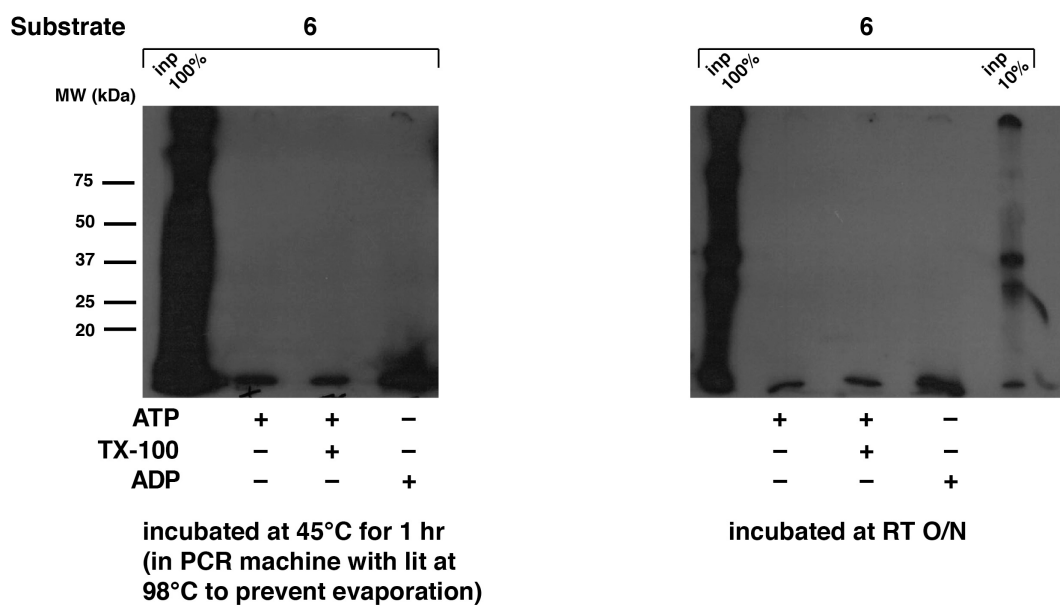


Incubated O/N at RT

used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

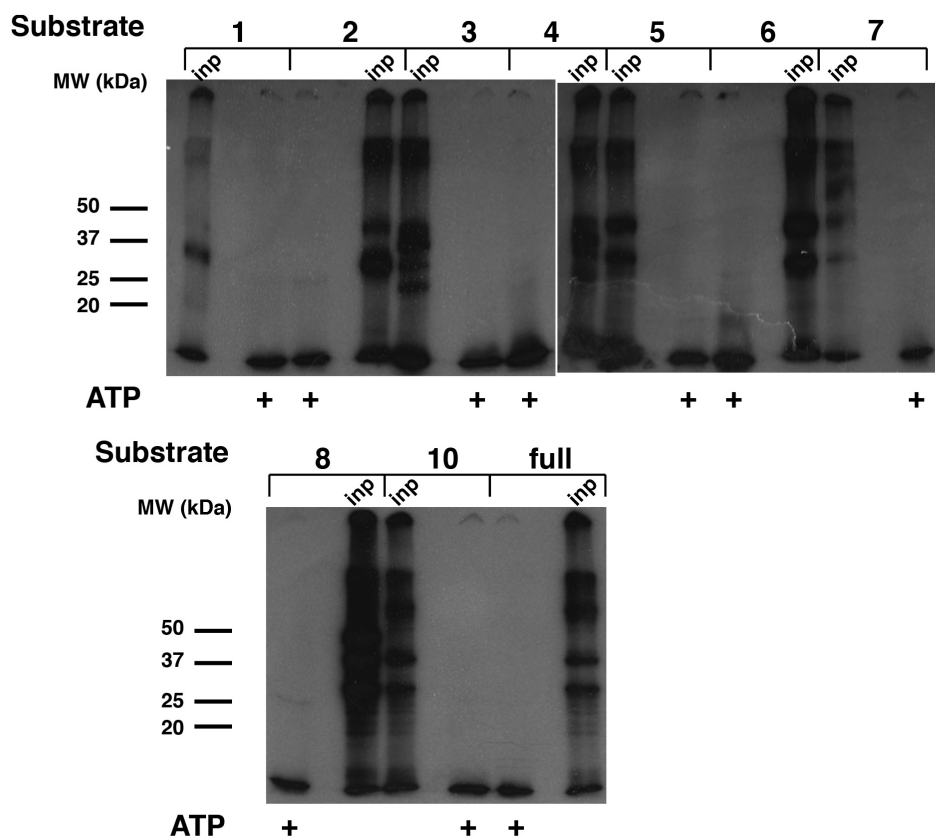
used 1.5 μ l PLs (14:1 lipid:protein ratio)

Experiment from 4/6/2015



used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)
 used 2 μ l PLs (14:1 lipid:protein ratio)
 added ATP as last component of reaction

Experiment from 4/9/2015



Incubated at RT O/N

used ATP regenerating system w/ 3mM PEP and 1 μ l LDH/PK mixture in each rxn

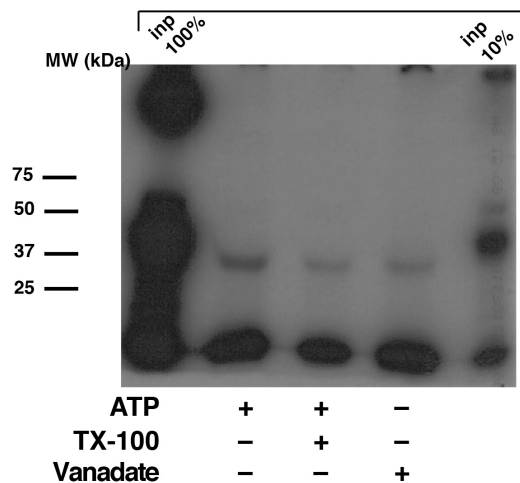
used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

used 2 μ l PLs (14:1 lipid:protein ratio)

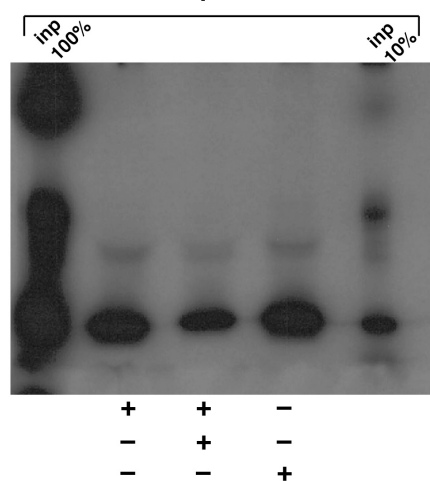
added ATP as last component of reaction

Experiment from 4/14/2015

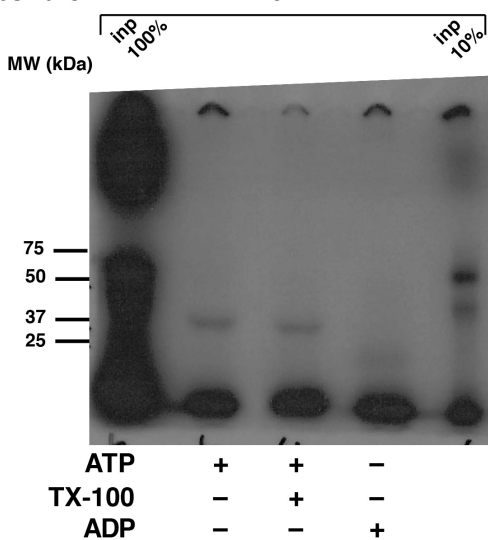
Substrate 2



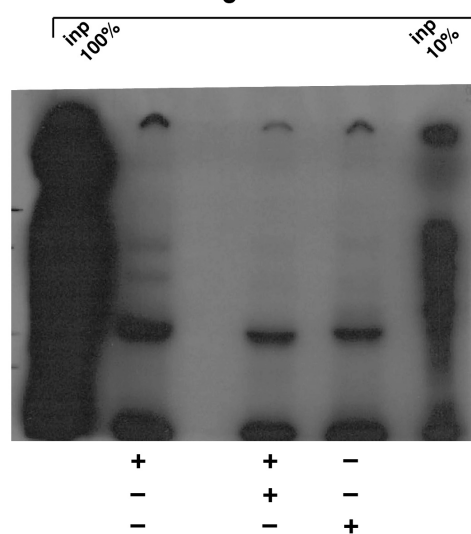
4



Substrate 6



8



Incubated at RT O/N

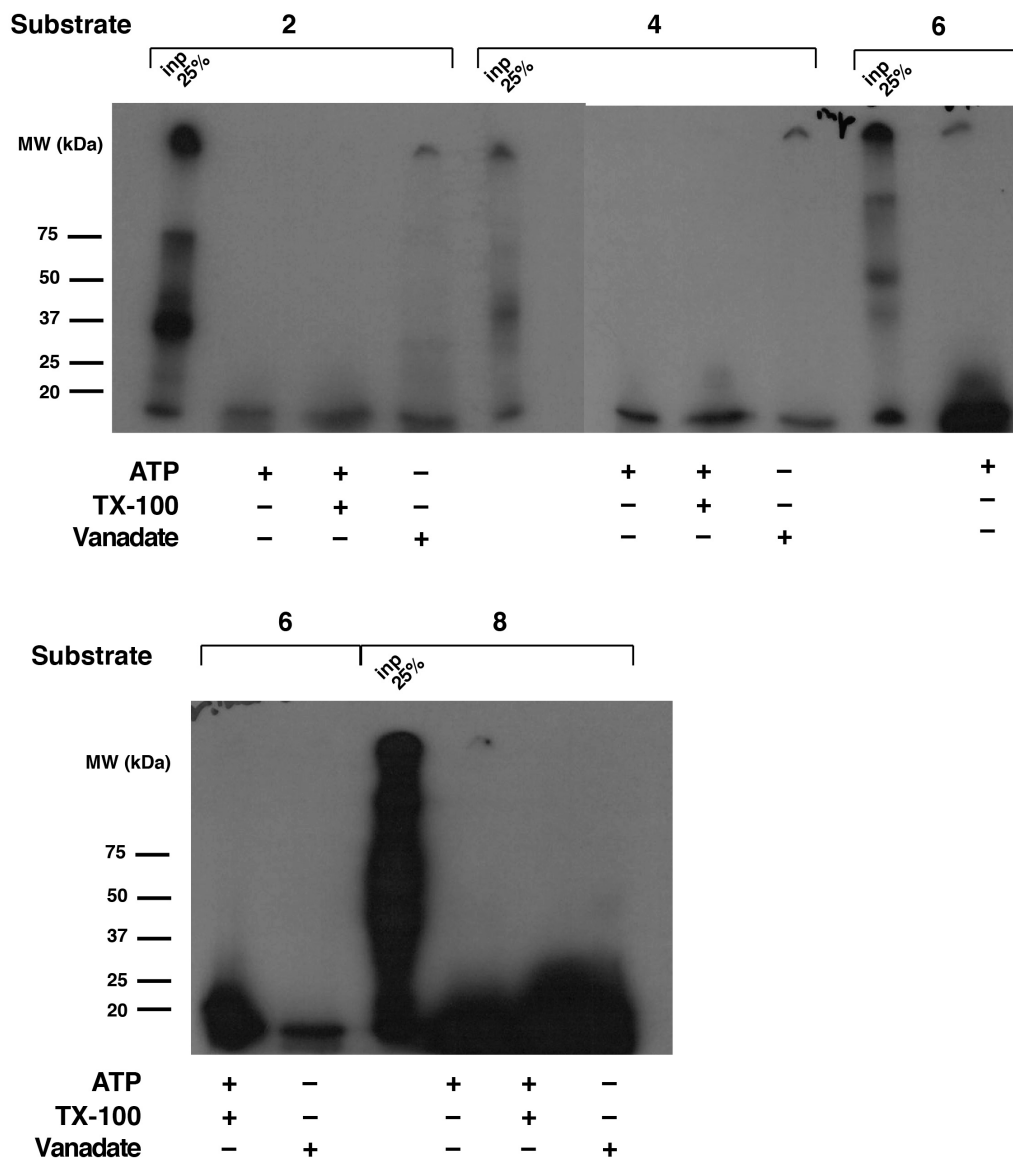
used ATP regenerating system w/ 3 mM PEP and 1 μ l LDH/PK per rxn

used 2 μ l substrate (directly from transcription/translation reaction (i.e. no TCA))

used 2 μ l PLs (14:1 lipid:protein ratio)

added ATP as last component of reaction

Experiment from 4/16/2015



Incubated at 55°C for 1 hr

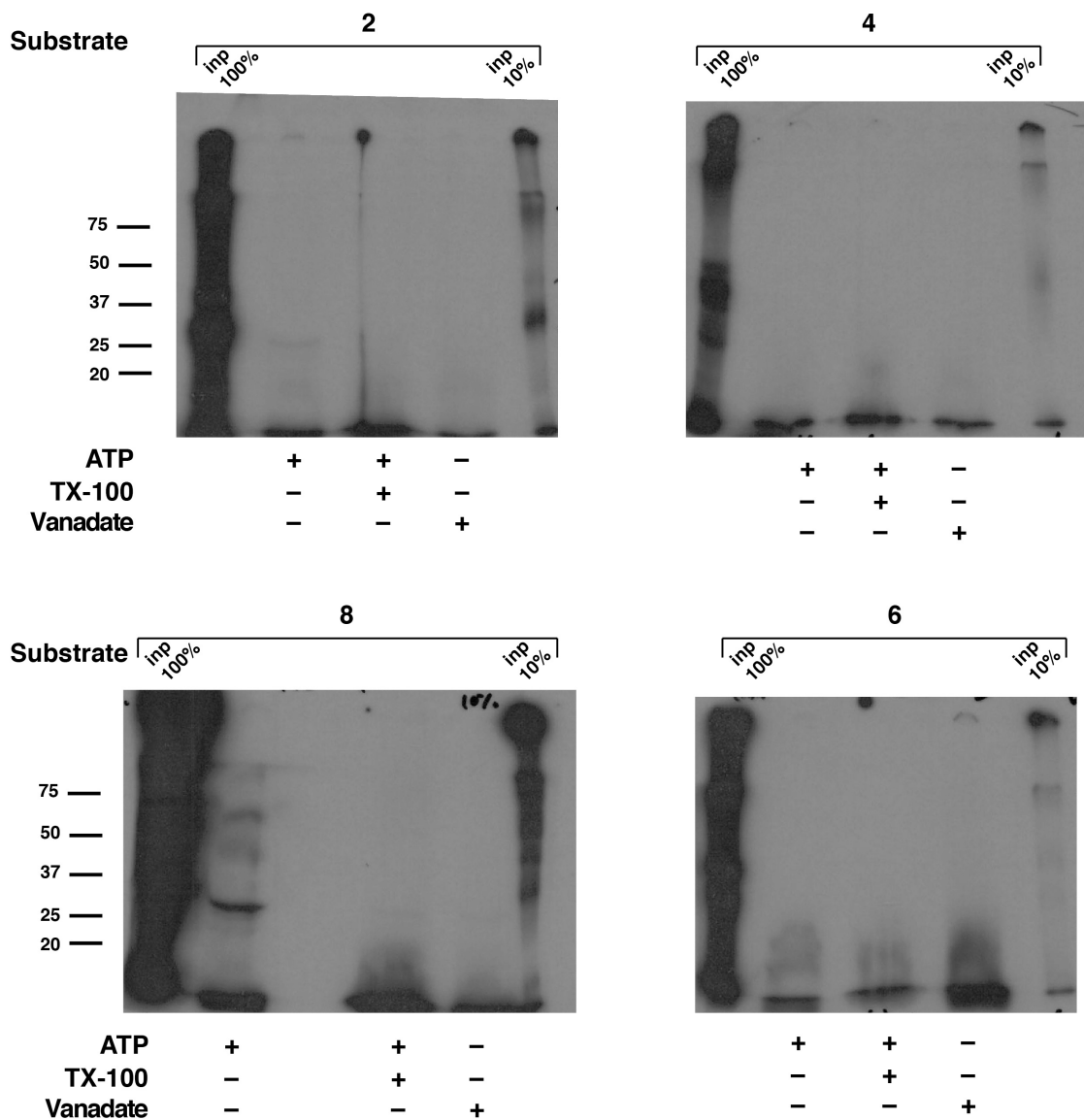
used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

used 1.65 μ l PLs (14:1 lipid:protein ratio)

Incubated with ATP for 1 min before adding substrate

as last component of reaction

Experiment from 4/16/2015



Incubated at RT O/N

used ATP regenerating system w/ 3 mM PEP and 1 μ l LDH/PK per rxn

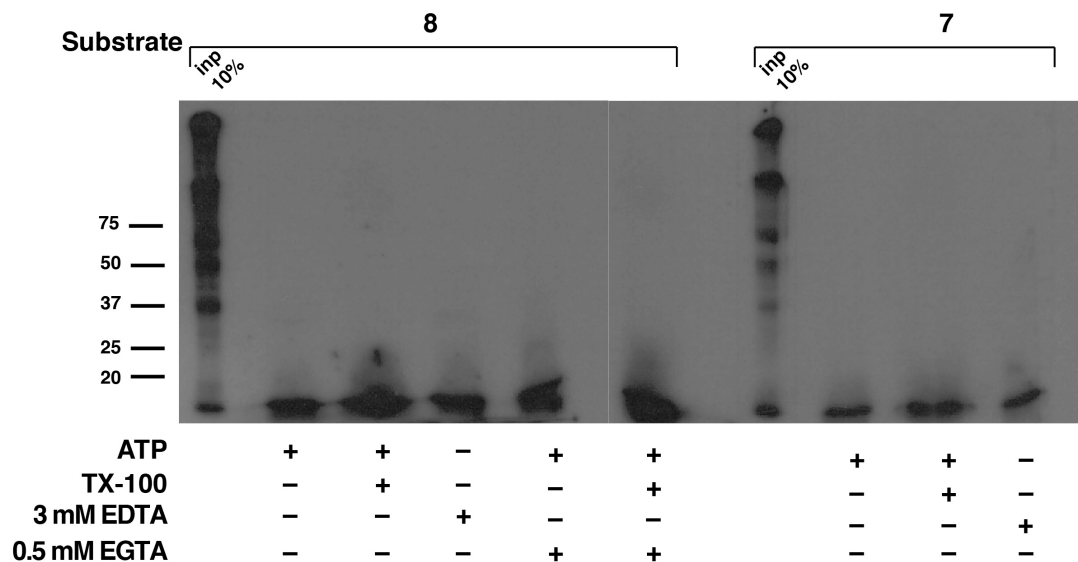
used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

used 1.65 μ l PLs (14:1 lipid:protein ratio)

Incubated with ATP for 1 min before adding substrate

as last component of reaction

Experiment from 4/23/2015



Incubated at RT O/N

used ATP regenerating system w/ 3 mM PEP and 1 μ l LDH/PK per rxn

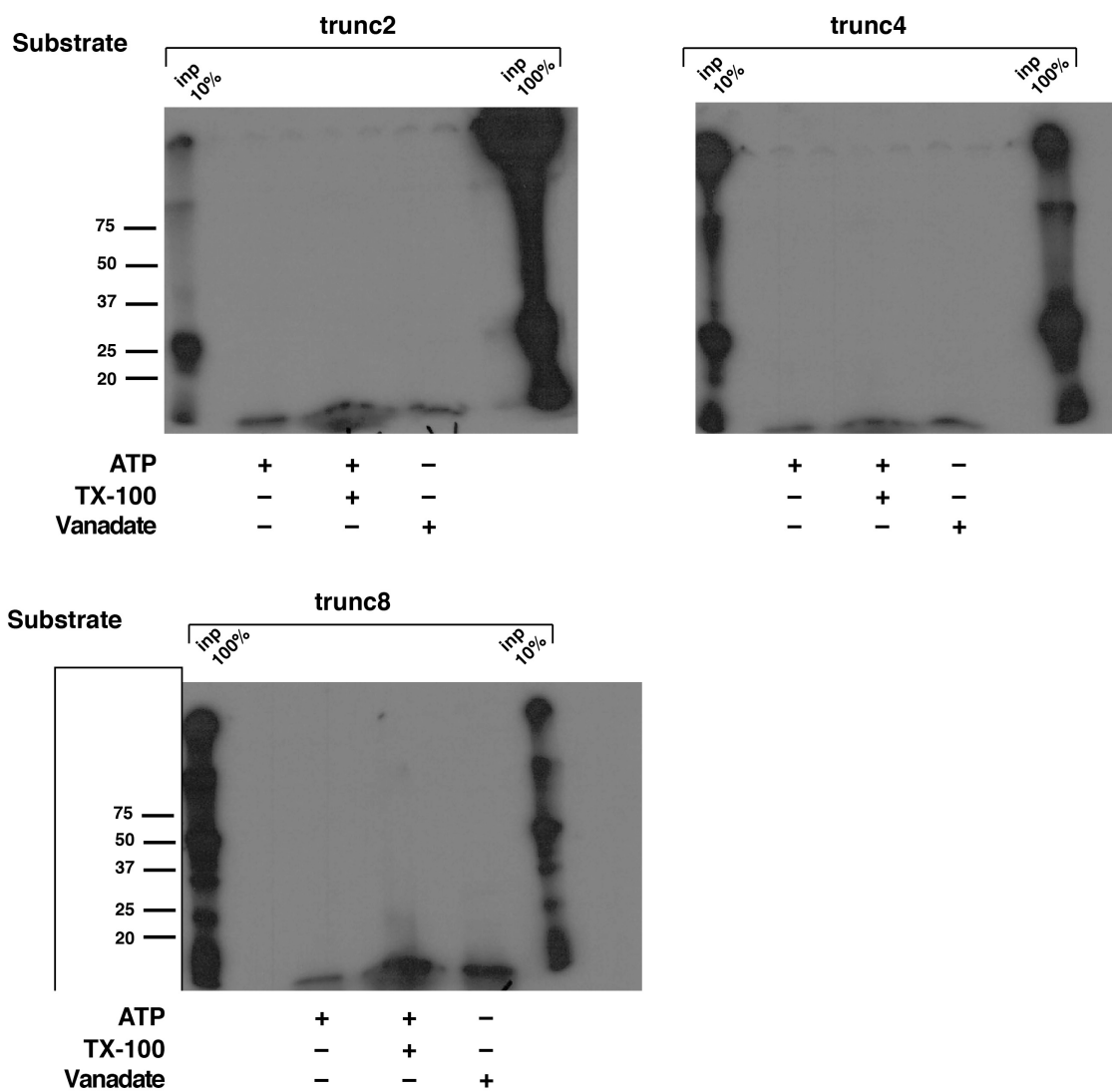
used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

used 1.65 μ l PLs (14:1 lipid:protein ratio)

Incubated with ATP for 1 min before adding substrate

as last component of reaction

Experiment from 5/5/2015



Incubated at RT O/N

used ATP regenerating system w/ 3 mM PEP and 1 μ l LDH/PK per rxn

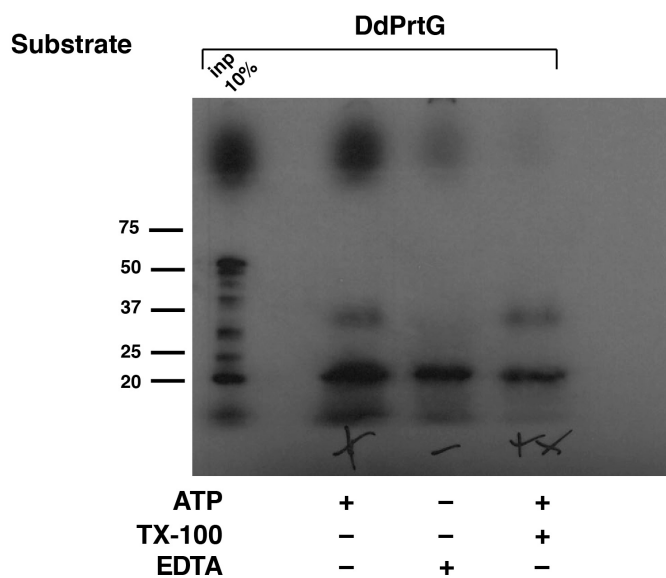
used 2 μ l substrate (TCA precipitated and resuspended in 25 μ l 8M urea)

used 1.5 μ l PLs (14:1 lipid:protein ratio)

Incubated with ATP for 1 min before adding substrate

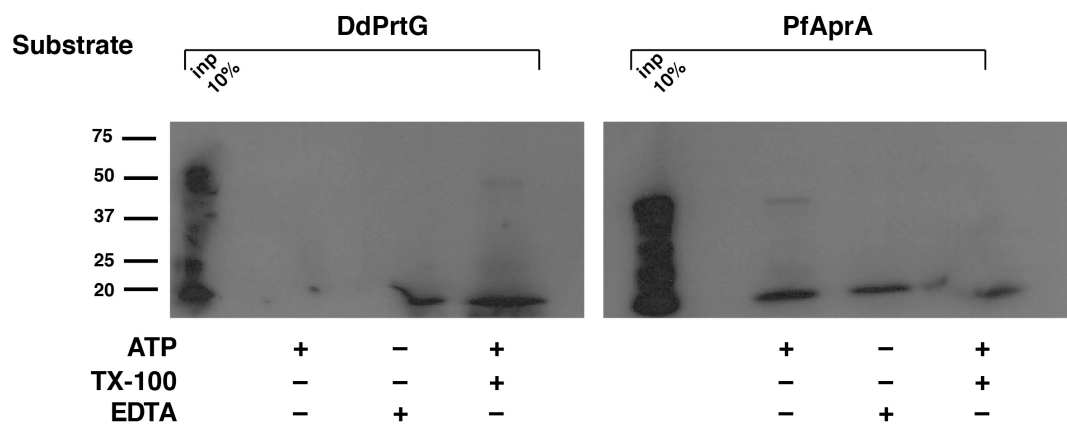
as last component of reaction

Experiment from 12/23/2015



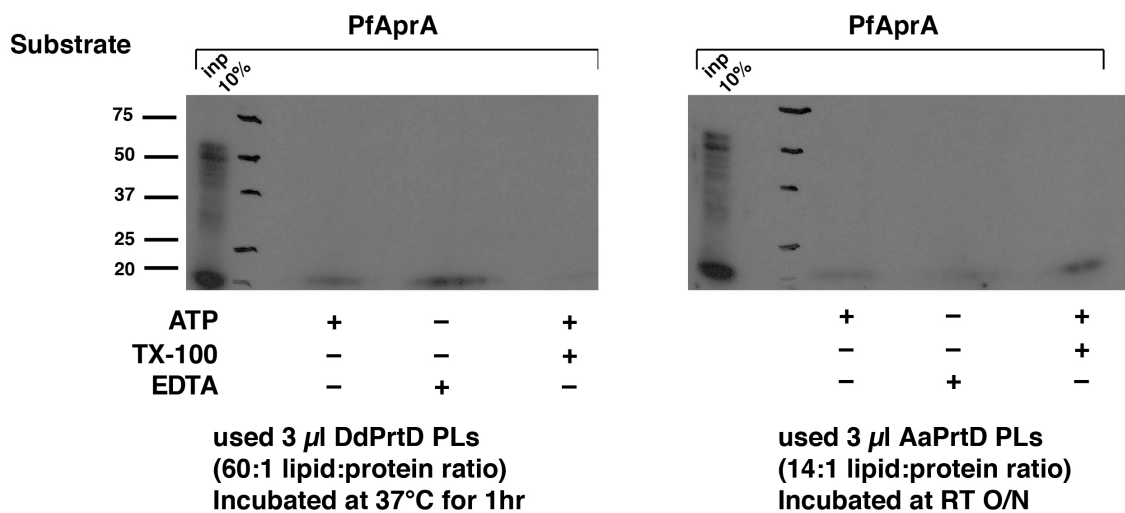
Incubated at 37°C for 1hr
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrtD PLs (60:1 lipid:protein ratio)
 added substrate as last component of reaction

Experiment from 12/28/2015



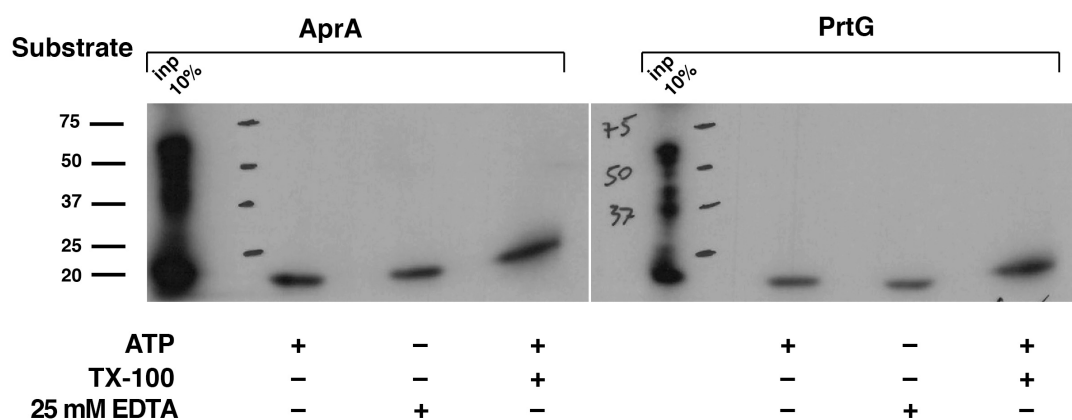
Incubated at 37°C for 1hr
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrtD PLs (60:1 lipid:protein ratio)
 added substrate as last component of reaction

Experiment from 12/30/2015



used 1.5 μ l substrate directly from transcription/translation rxn
(not TCA precipitated)
added substrate as last component of reaction

Experiment from 3/18/2015



used 3 μ l AaPrtD PLs (30:1 lipid:protein)

Incubated at 65°C for 25 min.

used 1.5 μ l substrate directly from transcription/translation rxn
(not TCA precipitated)

added substrate as last component of reaction

USED 2% TRITON TO DISSOLVE VESICLES

App1.4 Basic DdPrtD Inner-membrane vesicle (IMV) protease protection translocation experiments (variations will be described on specific experiments)

—Experiment is based on protease protection: if a radioactively-labeled polypeptide substrate is transported to the lumen of a vesicle, it will be protected from digestion by a robust, non-specific protease. Because the transport protein is an ATPase, the omission of ATP should result in no protection from the protease because the polypeptide substrate will not be transported to the lumen. Additionally, dissolving the vesicles with a detergent should also result in no protection because there will not be a lumen for the polypeptide substrate. Therefore, the ability to observe protection of a fragment of the radioactively-labeled substrate acts as an assay to monitor transport.

—20 μ l final reaction mixture contains 50 mM Tris pH 7.5, 100 mM NaCl, 5 mM MgCl_2 , 5 mM NADH or 25 mM D-lactate, 5 mM β -mercaptoethanol (β -me), 0.05 mg/ml creatine kinase (CK), 4mM EGTA, 25 mM phosphocreatine, 5 mM ATP, IMVs and substrate

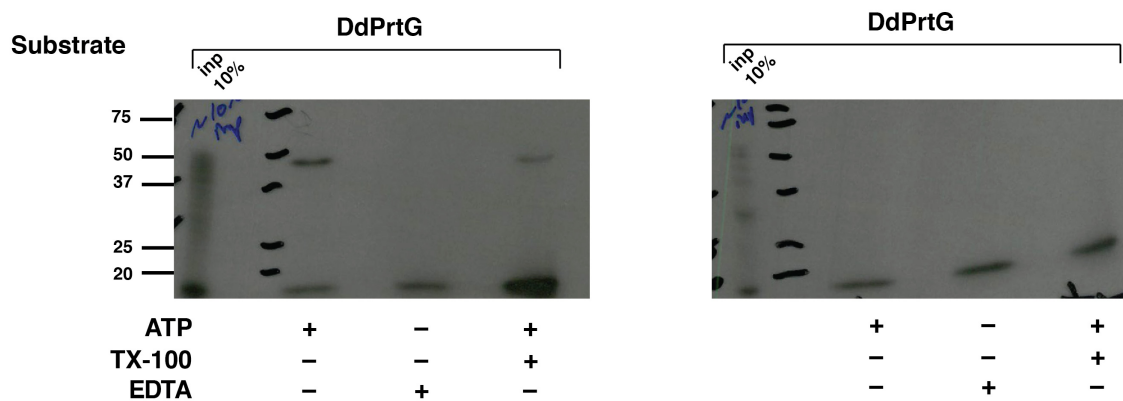
—Each experiment requires at least 3 reactions: one with ATP, one with no ATP or with a specific ATPase inhibitor such as vanadate (5 mM) or EDTA (25 mM), one with ATP that will be dissolved with TX-100 as a negative control

—IMV assay

1. Make master mix containing Tris, NaCl, β -me, EGTA, D-lactate, CK, and phosphocreatine.
2. Aliquot into appropriate number of tubes
3. Add 3 μ l IMVs to hopefully generate PMF and thus inhibit ATP synthase
4. Add ATP to 5 mM to two samples and EDTA to 25 mM to other
5. Lastly, add 1.5 μ l substrate
6. Incubate at 37°C for 1 hr
7. Prepare fresh 10 mg/ml proteinase K (protK) solution dissolved in 50 mM Tris pH 8, 10 mM CaCl_2
8. Add TX-100 to 1% to negative control sample
9. Add protK to PL reaction to 1 mg/ml final conc.
10. incubate at room temp for 1 hr
11. Add 15 μ l TCA (~100% solution) and mix thoroughly by tapping tube

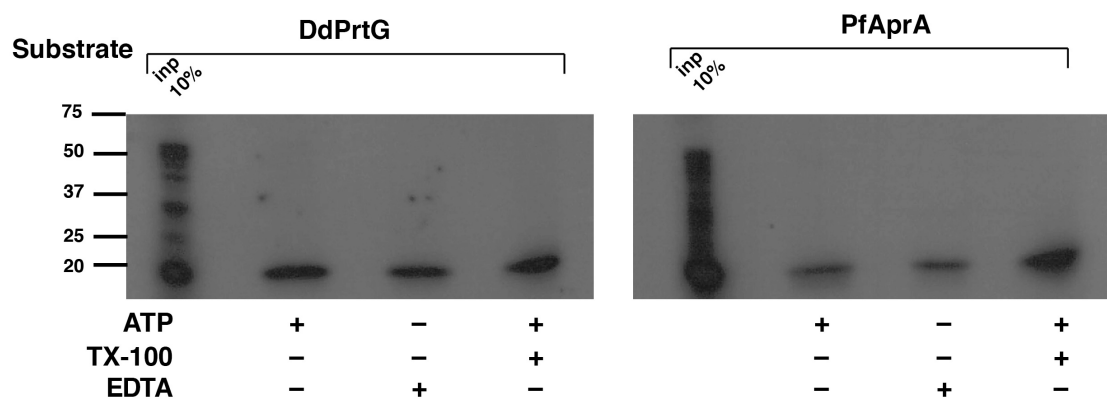
12. Incubate samples on ice for 30'
13. Spin samples at 14,000 rpm, 4°C, for 30'
14. Carefully remove supernatant (try to get all of it without getting pellet)
15. Resuspend pellet mixture of 20 μ l reducing SDS gel loading buffer and 20 μ l 1.5 M Tris pH 8.8
16. Shake vigorously until pellet is fully dissolved
17. Load 10-lane gel with 0.2 μ l substrate in one lane, ATP-containing sample in separate lane, ATP-omitted sample in separate lane, TX-100 containing sample in separate lane. Make sure to have at least 1 empty lane between radioactive samples
18. Run gel
19. Rock gel in ~100 ml of solution containing 40% ethanol and 3% glycerol for $\geq 30'$
20. Dry gel for 2 hr at 80°C
21. Expose gel in autoradiography cassette for 12–36 hr at –80°C
22. Develop autoradiograph in biochemistry dept.

Experiment from 1/18/2016



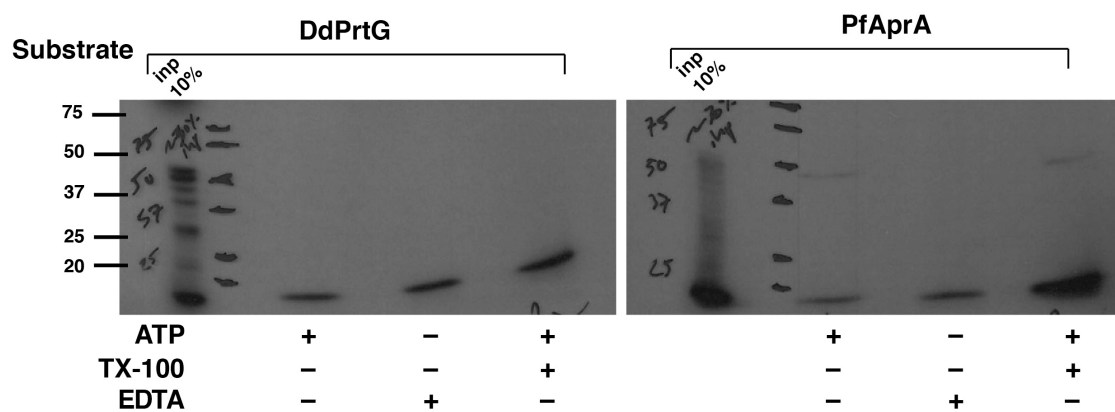
Incubated at 37°C for 1hr
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrtD IMVs
 added 5 mM NADH immediately before adding ATP
 added substrate as last component of reaction

Experiment from 1/24/2015



Incubated at 37°C for 1hr
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrtD IMVs
 added 5 mM NADH immediately before adding ATP
 added substrate as last component of reaction
 USED 2% TRITON TO DISSOLVE VESICLES

Experiment from 1/27/2015



Incubated at 37°C for 1hr

used ATP-regenerating system using 25 mM phosphocreatine and 0.05 mg/ml creatine kinase (CK)

used 1.5 μ l substrate directly from transcription/translation rxn (not TCA precipitated)

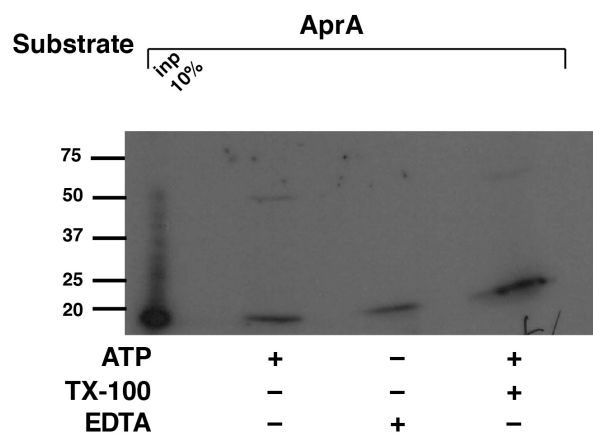
used 3 μ l DdPrtD IMVs

added 25 mM D-lactate immediately before adding ATP

added substrate as last component of reaction

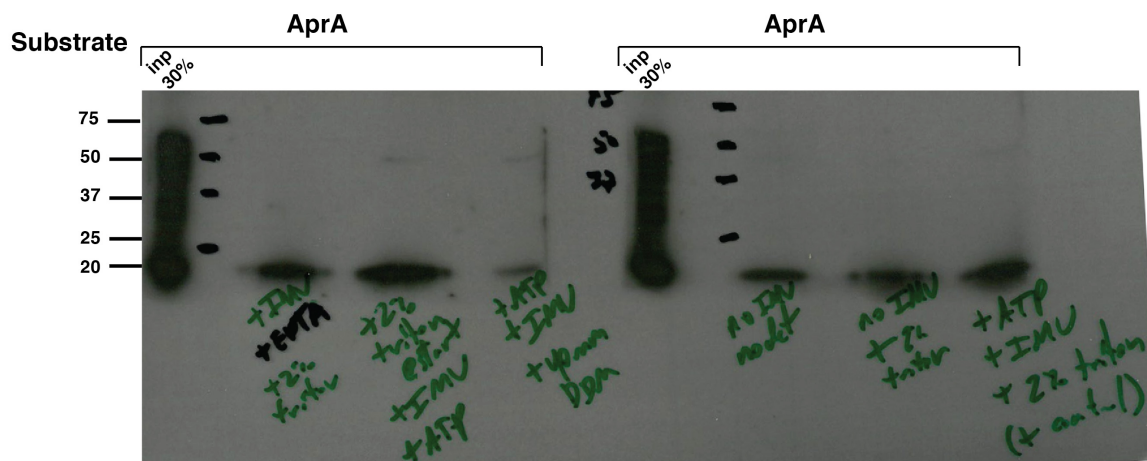
USED 2% TRITON TO DISSOLVE VESICLES

Experiment from 1/29/2015



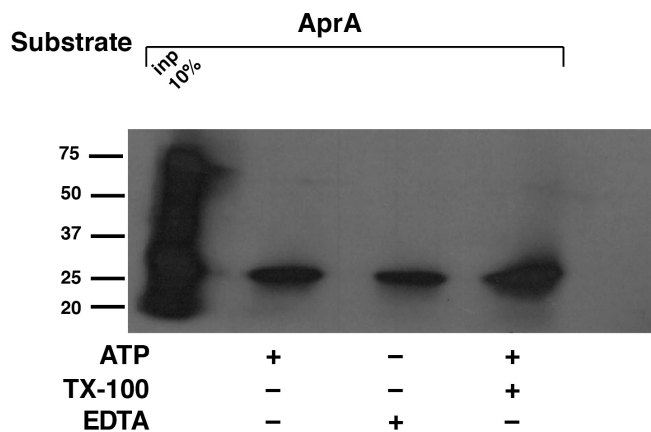
Incubated at 37°C for 1hr
 used ATP-regenerating system using 25 mM phosphocreatine
 and 0.05 mg/ml creatine kinase (CK)
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrD IMVs
 added 25 mM D-lactate immediately before adding ATP
 added substrate as last component of reaction
 USED 3% TRITON TO DISSOLVE VESICLES

Experiment from 2/3/2015



Incubated at 37°C for 1hr
 used ATP-regenerating system using 25 mM phosphocreatine
 and 0.05 mg/ml creatine kinase (CK)
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrd IMVs
 added 25 mM D-lactate immediately before adding ATP
 added substrate as last component of reaction
 USED 3% TRITON TO DISSOLVE VESICLES

Experiment from 2/9/2015



Used 40 μ l reaction volume
 Incubated at 37°C for 1hr
 used ATP-regenerating system using 25 mM phosphocreatine
 and 0.05 mg/ml creatine kinase (CK)
 used 1.5 μ l substrate directly from transcription/translation rxn
 (not TCA precipitated)
 used 3 μ l DdPrTD IMVs
 added 25 mM D-lactate immediately before adding ATP
 added substrate as last component of reaction
 added 25 mM EDTA to all samples before protK digestion
 USED 2% TRITON TO DISSOLVE VESICLES

Appendix 2: *Thermotoga maritima* growth protocol

In this appendix, I describe a method for growing *Thermotoga maritima* for the purpose of harvesting lipids and using them to make proteoliposomes that can withstand incubation at high temperatures.

1. In a 5 L bottle, combine:
 - 2.5 g yeast extract
 - 12.5 g tryptone
 - 37.5 g NaCl
 - 5 g MgSO₄
 - 1.25 g CaCl₂•2H₂O
 - 0.625 g NaHCO₃
 - 0.125 g KBr
 - 0.05 g Boric acid
 - 0.05 g potassium iodide
 - 0.0075 g Sodium tungstate
 - 0.005 g NiCl₂•6H₂O
2. Add 4.8L deionized water
3. Autoclave on cycle 2 with silicon cap loosely tightened
4. Prepare 10 mg/ml maltose solution (filter with 0.2 µm filter)
5. prepare 10 mg/ml KH₂PO₄ solution (filter with 0.2 µm filter)
6. prepare 10% Na₂S solution and adjust pH to 8 with sulfuric acid (do it in hood)
7. prepare 100 mg/ml stock of resazurin solution
8. Once media cools, add:
 - 5 ml KH₂PO₄ solution
 - 120 ml maltose solution
 - 10 ml resazurin solution
9. Spurge with nitrogen (use ~20 psi) by poking empty 21 G needle into silicon stopper and poking 21 G needle attached to nitrogen into silicon cap—spurge for 10 minutes and be careful to remove both needles at the same time

10. Make sure silicon cap is sealed tightly, media should change from yellow to pink to blue over the next 20–30 minutes as oxygen is consumed. (pink means oxygen is present, blue means that it is optimal anaerobic environment)
11. Once media turns blue, inoculate by adding 1% (V/V) of cells to media (poke needle into silicon cap) **DO NOT OPEN CAP** and be careful not to let too much air in during inoculation steps.
12. Place media in oven set at 80°C (no shaking) and allow to grow for 12–36 hours.
13. To harvest, spin media at 5500 rpm for 10 min (don't worry about oxygen anymore)
14. It is best to keep a few hundred milliliters of cells after growth (make sure to spurge the flask first) in order to use as a starter culture for future growths

References:

- Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* *301*, 610–615.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 213–221.
- Ahimou, F., Semmens, M.J., Haugstad, G., and Novak, P.J. (2007). Effect of protein, polysaccharide, and oxygen concentration profiles on biofilm cohesiveness. *Applied and Environmental Microbiology* *73*, 2905–2910.
- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008) *Molecular biology of the cell, 5th edition Garland Science, New York, NY*, 1616 pp
- Aloni, Y., Delmer, D.P., and Benziman, M. (1982). 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. U.S.A.* *79*, 6448–6452.
- Amikam, D., and Galperin, M.Y. (2006). PilZ domain is part of the bacterial c-di-GMP binding protein. *Bioinformatics* *22*, 3–6.
- Ardèvol, A., and Rovira, C. (2015). Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from ab Initio/Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J. Am. Chem. Soc.* *137*, 7528–7547.
- Asensio, J.L., Ardá, A., Cañada, F.J., and Jiménez-Barbero, J. (2013). Carbohydrate–Aromatic Interactions. *Acc. Chem. Res.* *46*, 946–954.
- Baker, P., Hill, P.J., Snarr, B.D., Alnabelseya, N., Pesttrak, M.J., Lee, M.J., Jennings, L.K., Tam, J., Melnyk, R.A., Parsek, M.R., et al. (2016). Exopolysaccharide biosynthetic glycoside hydrolases can be utilized to disrupt and prevent *Pseudomonas aeruginosa* biofilms. *Sci Adv* *2*, e1501632.
- Bakkes, P.J., Jenewein, S., Smits, S.H.J., Holland, I.B., and Schmitt, L. (2010). The Rate of Folding Dictates Substrate Secretion by the *Escherichia coli* Hemolysin Type 1 Secretion System. *Journal of Biological Chemistry* *285*, 40573–40580.
- Balakrishnan, L., Hughes, C., and Koronakis, V. (2001). Substrate-triggered recruitment of the TolC channel-tunnel during type I export of hemolysin by *Escherichia coli*. *Journal of Molecular Biology* *313*, 501–510.

- Balan, V. (2014). Current Challenges in Commercially Producing Biofuels from Lignocellulosic Biomass. *ISRN Biotechnology* 2014, 1–31.
- Balla, T. (2013). Phosphoinositides: Tiny Lipids With Giant Impact on Cell Regulation. *Physiol. Rev.* 93, 1019–1137.
- Bauer, B.W., Shemesh, T., Chen, Y., and Rapoport, T.A. (2014). A “Push and Slide” Mechanism Allows Sequence-Insensitive Translocation of Secretory Proteins by the SecA ATPase. *Cell* 157, 1416–1429.
- Beek, ter, J., Guskov, A., and Slotboom, D.J. (2014). Structural diversity of ABC transporters. *The Journal of General Physiology* 143, 419–435.
- Benach, J., Swaminathan, S.S., Tamayo, R., Handelman, S.K., Folta-Stogniew, E., Ramos, J.E., Forouhar, F., Neely, H., Seetharaman, J., Camilli, A., et al. (2007). The structural basis of cyclic diguanylate signal transduction by PilZ domains. *The EMBO Journal* 26, 5153–5166.
- Berg, J., Tymoczko, J., Stryer, L. (2002). *Biochemistry*, Sixth edition *WH Freeman*, NY, 1026 pp.
- Bi, Y., Hubbard, C., Purushotham, P., and Zimmer, J. (2015). Insights into the structure and function of membrane-integrated processive glycosyltransferases. *Current Opinion in Structural Biology* 34, 78–86.
- Binet, R., and Wandersman, C. (1995). Protein secretion by hybrid bacterial ABC-transporters: specific functions of the membrane ATPase and the membrane fusion protein. *The EMBO Journal* 14, 2298–2306.
- Binet, R., Létoffé, S., Ghigo, J.M., Delepelaire, P., and Wandersman, C. (1997). Protein secretion by Gram-negative bacterial ABC exporters--a review. *Gene* 192, 7–11.
- Bokranz, W., Wang, X., Tschäpe, H., and Römling, U. (2005). Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. *J. Med. Microbiol.* 54, 1171–1182.
- Brown, C., Leijon, F., and Bulone, V. (2012). Radiometric and spectrophotometric in vitro assays of glycosyltransferases involved in plant cell wall carbohydrate biosynthesis. *Nat Protoc* 7, 1634–1650.
- Buckstein, M.H., He, J., and Rubin, H. (2008). Characterization of nucleotide pools as a function of physiological state in *Escherichia coli*. *Journal of Bacteriology* 190, 718–726.
- Bumba, L., Masin, J., Macek, P., Wald, T., Motlova, L., Bibova, I., Klimova, N., Bednarova, L., Veverka, V., Kachala, M., et al. (2016). Calcium-Driven Folding of RTX Domain & β -Rolls Ratchets Translocation of RTX Proteins through Type I Secretion Ducts. *Molecular Cell* 62, 47–62.

- Buschiazzo, A., Ugalde, J.E., Guerin, M.E., Shepard, W., Ugalde, R.A., and Alzari, P.M. (2004). Crystal structure of glycogen synthase: homologous enzymes catalyze glycogen synthesis and degradation. *The EMBO Journal* 23, 3196–3205.
- Cantarel, B.L., Coutinho, P.M., Rancurel, C., Bernard, T., Lombard, V., and Henrissat, B. (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucl. Acids Res.* 37, D233–D238.
- Cao, Z., and Bowie, J.U. (2012). Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc. Natl. Acad. Sci. U.S.a.* 109, 8121–8126.
- Carpita, N.C. (2011). Update on mechanisms of plant cell wall biosynthesis: how plants make cellulose and other (1->4)- β -D-glycans. *Plant Physiology* 155, 171–184.
- Carroll, A., and Somerville, C. (2009). Cellulosic Biofuels. *Annu. Rev. Plant Biol.* 60, 165–182.
- Chan, P.H.W., Cheung, A.H., Okon, M., Chen, H.-M., Withers, S.G., and McIntosh, L.P. (2013). Investigating the structural dynamics of α -1,4-galactosyltransferase C from *Neisseria meningitidis* by nuclear magnetic resonance spectroscopy. *Biochemistry* 52, 320–332.
- Chang, A., Singh, S., Phillips, G.N., Jr, and Thorson, J.S. (2011). Glycosyltransferase structural biology and its role in the design of catalysts for glycosylation. *Current Opinion in Biotechnology* 22, 800–808.
- Chen, J. (2013). Molecular mechanism of the *Escherichia coli* maltose transporter. *Current Opinion in Structural Biology* 23, 492–498.
- Chenal, A., Guijarro, J.I., Raynal, B., Delepierre, M., and Ladant, D. (2008). RTX Calcium Binding Motifs Are Intrinsically Disordered in the Absence of Calcium: IMPLICATION FOR PROTEIN SECRETION. *Journal of Biological Chemistry* 284, 1781–1789.
- Chenal, A., Karst, J.C., Pérez, A.C.S., Wozniak, A.K., Baron, B., England, P., and Ladant, D. (2010). Calcium-Induced Folding and Stabilization of the Intrinsically Disordered RTX Domain of the CyaA Toxin. *Biophysj* 99, 3744–3753.
- Choudhury, H.G., Tong, Z., Mathavan, I., Li, Y., Iwata, S., Zirah, S., Rebuffat, S., van Veen, H.W., and Beis, K. (2014). Structure of an antibacterial peptide ATP-binding cassette transporter in a novel outward occluded state. *Proc. Natl. Acad. Sci. U.S.a.* 111, 9145–9150.
- Christen, M., Christen, B., Allan, M.G., Folcher, M., Jenö, P., Grzesiek, S., and Jenal, U. (2007). DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. *Proc. Natl. Acad. Sci. U.S.a.* 104, 4112–4117.

- Christopher, L.P., Yao, B., and Ji, Y. (2014). Lignin biodegradation with laccase-mediator systems. *Frontiers in Energy Research*.
- Cifuentes, C., Bulone, V., and Emons, A.M.C. (2010). Biosynthesis of Callose and Cellulose by Detergent Extracts of Tobacco Cell Membranes and Quantification of the Polymers Synthesized in vitro. *Journal of Integrative Plant Biology* 52, 221–233.
- Collaborative Computational Project, N.4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50, 760–763.
- Costerton, J.W., Stewart, P.S., and Greenberg, E.P. (1999). Bacterial biofilms: a common cause of persistent infections. *Science* 284, 1318–1322.
- Cotter, P.A., and Stibitz, S. (2007). c-di-GMP-mediated regulation of virulence and biofilm formation. *Current Opinion in Microbiology* 10, 17–23.
- Cowtan, K. (2010). Recent developments in classical density modification. *Acta Crystallogr. D Biol. Crystallogr.* 66, 470–478.
- Cowtan, K.D., and Zhang, K.Y. (1999). Density modification for macromolecular phase improvement. *Prog. Biophys. Mol. Biol.* 72, 245–270.
- Csanady, L., Vergani, P., and Gadsby, D.C. (2010). Strict coupling between CFTR's catalytic cycle and gating of its Cl⁻ ion pore revealed by distributions of open channel burst durations. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1241–1246.
- Cuthbertson, L., Kos, V., and Whitfield, C. (2010). ABC Transporters Involved in Export of Cell Surface Glycoconjugates. *Microbiology and Molecular Biology Reviews* 74, 341–362.
- Davies, J.C. (2002). *Pseudomonas aeruginosa* in cystic fibrosis: pathogenesis and persistence. *Paediatric Respiratory Reviews*.
- Dawson, R.J.P., and Locher, K.P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185.
- Debarbieux, L., and Wandersman, C. (2001). Folded HasA inhibits its own secretion through its ABC exporter. *The EMBO Journal* 20, 4657–4663.
- del Pozo, J.L., and Patel, R. (2007). The Challenge of Treating Biofilm-associated Bacterial Infections. *Clin Pharmacol Ther* 82, 204–209.
- Delepelaire, P. (1994). PrtD, the integral membrane ATP-binding cassette component of the *Erwinia chrysanthemi* metalloprotease secretion system, exhibits a secretion signal-regulated ATPase activity. *J. Biol. Chem.* 269, 27952–27957.
- Delepelaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochimica Et Biophysica Acta (BBA) - Molecular Cell Research* 1694, 149–161.

- Delepelaire, P., and Wandersman, C. (1989). Protease secretion by *Erwinia chrysanthemi*. Proteases B and C are synthesized and secreted as zymogens without a signal peptide. *J. Biol. Chem.* *264*, 9083–9089.
- Delepelaire, P., and Wandersman, C. (1990). Protein secretion in gram-negative bacteria. The extracellular metalloprotease B from *Erwinia chrysanthemi* contains a C-terminal secretion signal analogous to that of *Escherichia coli* alpha-hemolysin. *J. Biol. Chem.* *265*, 17118–17125.
- Deng, D., Xu, C., Sun, P., Wu, J., Yan, C., Hu, M., and Yan, N. (2014). Crystal structure of the human glucose transporter GLUT1. *Nature* *510*, 121–125.
- Dirr, H.W., Little, T., Kuhnert, D.C., and Sayed, Y. (2005). A conserved N-capping motif contributes significantly to the stabilization and dynamics of the C-terminal region of class Alpha glutathione S-transferases. *J. Biol. Chem.* *280*, 19480–19487.
- Dougherty, D.A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* *271*, 163–168.
- Du, D., Wang, Z., James, N.R., Voss, J.E., Klimont, E., Ohene-Agyei, T., Venter, H., Chiu, W., and Luisi, B.F. (2014). Structure of the AcrAB-TolC multidrug efflux pump. *Nature* *509*, 512–515.
- Eggensperger, S., and Tampe, R. (2015). The transporter associated with antigen processing: a key player in adaptive immunity. *Biol. Chem.* *396*, 1059–1072.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* *60*, 2126–2132.
- Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 486–501.
- Erlanson, K.J., Miller, S.B.M., Nam, Y., Osborne, A.R., Zimmer, J., and Rapoport, T.A. (2008). A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature* *455*, 984–987.
- Faham, S., and Bowie, J.U. (2002). Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. *Journal of Molecular Biology* *316*, 1–6.
- Faham, S., Boulting, G.L., Massey, E.A., Yohannan, S., Yang, D., and Bowie, J.U. (2005). Crystallization of bacteriorhodopsin from bicelle formulations at room temperature. *Protein Sci.* *14*, 836–840.
- Faham, S., Watanabe, A., Besserer, G.M., Cascio, D., Specht, A., Hirayama, B.A., Wright, E.M., and Abramson, J. (2008). The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na⁺/sugar symport. *Science* *321*, 810–814.

- Fan, E., Fiedler, S., Jacob-Dubuisson, F., and Müller, M. (2012). Two-partner secretion of gram-negative bacteria: a single β -barrel protein enables transport across the outer membrane. *Journal of Biological Chemistry* 287, 2591–2599.
- Flemming, H.C., Neu, T.R., and Wozniak, D.J. (2007). The EPS Matrix: The “House of Biofilm Cells.” *Journal of Bacteriology* 189, 7945–7947.
- Forood, B., Feliciano, E.J., and Nambiar, K.P. (1993). Stabilization of alpha-helical structures in short peptides via end capping. *Proc. Natl. Acad. Sci. U.S.A.* 90, 838–842.
- Forrest, L.R. (2013). (Pseudo-)Symmetrical Transport. *Science* 339, 399–401.
- Forrest, L.R., Krämer, R., and Ziegler, C. (2011). The structural basis of secondary active transport mechanisms. *BBA - Bioenergetics* 1807, 167–188.
- Franklin, M.J., Nivens, D.E., Weadge, J.T., and Howell, P.L. (2011). Biosynthesis of the *Pseudomonas aeruginosa* Extracellular Polysaccharides, Alginate, Pel, and Psl. *Front Microbiol* 2, 167.
- Friedman, L., and Kolter, R. (2004). Genes involved in matrix formation in *Pseudomonas aeruginosa* PA14 biofilms. *Mol. Microbiol.* 51, 675–690.
- Fuche, F., Vianney, A., Andrea, C., Doublet, P., and Gilbert, C. (2015). Functional Type 1 Secretion System Involved in *Legionella pneumophila* Virulence. *Journal of Bacteriology* 197, 563–571.
- Fujiwara, T., Komoda, K., Sakurai, N., Tajima, K., Tanaka, I., and Yao, M. (2013). The c-di-GMP recognition mechanism of the PilZ domain of bacterial cellulose synthase subunit A. *Biochem. Biophys. Res. Commun.* 431, 802–807.
- Gardner, K.H., and Blackwell, J. (1974). The hydrogen bonding in native cellulose. *Biochim. Biophys. Acta* 343, 232–237.
- Gentner, M., Allan, M.G., Zaehring, F., Schirmer, T., and Grzesiek, S. (2012). Oligomer Formation of the Bacterial Second Messenger c-di-GMP: Reaction Rates and Equilibrium Constants Indicate a Monomeric State at Physiological Concentrations. *J. Am. Chem. Soc.* 134, 1019–1029.
- Ghigo, J.M., and Wandersman, C. (1994). A carboxyl-terminal four-amino acid motif is required for secretion of the metalloprotease PrtG through the *Erwinia chrysanthemi* protease secretion pathway. *J. Biol. Chem.* 269, 8979–8985.
- Gloag, E.S., Turnbull, L., Huang, A., Vallotton, P., Wang, H., Nolan, L.M., Mililli, L., Hunt, C., Lu, J., Osvath, S.R., et al. (2013). Self-organization of bacterial biofilms is facilitated by extracellular DNA. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11541–11546.
- Gloster, T.M. (2014). Advances in understanding glycosyltransferases from a structural perspective. *Current Opinion in Structural Biology* 28, 131–141.

- Gray, M.C., Converse, A.O., and Wyman, C.E. (2003). Sugar monomer and oligomer solubility: data and predictions for application to biomass hydrolysis. *Appl. Biochem. Biotechnol.* *105 -108*, 179–193.
- Grossmann, N., Vakkasoglu, A.S., Hulpke, S., Abele, R., Gaudet, R., and Tampe, R. (2014). Mechanistic determinants of the directionality and energetics of active export by a heterodimeric ABC transporter. *Nature Communications* *5*, 1–10.
- Hay, I.D., Ur Rehman, Z., Moradali, M.F., Wang, Y., and Rehm, B.H.A. (2013). Microbial alginate production, modification and its applications. *Microb Biotechnol* *6*, 637–650.
- Hay, I.D., Wang, Y., Moradali, M.F., Rehman, Z.U., and Rehm, B.H.A. (2014). Genetics and regulation of bacterial alginate production. *Environ. Microbiol.* *16*, 2997–3011.
- Henrissat, B., Deleury, E., and Coutinho, P.M. (2002). Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? *Trends in Genetics* *18*, 437–440.
- Ho, B.K., and Gruswitz, F. (2008). HOLLOW: Generating Accurate Representations of Channel and Interior Surfaces in Molecular Structures. *BMC Struct Biol* *8*, 49.
- Hresko, R.C., Kraft, T.E., Quigley, A., Carpenter, E.P., and Hruz, P.W. (2016). Mammalian Glucose Transporter Activity Is Dependent upon Anionic and Conical Phospholipids. *Journal of Biological Chemistry* *291*, 17271–17282.
- Hubbard, C., McNamara, J.T., Azumaya, C., Patel, M.S., and Zimmer, J. (2012). The hyaluronan synthase catalyzes the synthesis and membrane translocation of hyaluronan. *Journal of Molecular Biology* *418*, 21–31.
- Islam, S.T., and Lam, J.S. (2014). Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway 1. *Can. J. Microbiol.* *60*, 697–716.
- Itoh, Y., Rice, J.D., Goller, C., Pannuri, A., Taylor, J., Meisner, J., Beveridge, T.J., Preston, J.F., and Romeo, T. (2008). Roles of pgaABCD genes in synthesis, modification, and export of the Escherichia coli biofilm adhesin poly-beta-1,6-N-acetyl-D-glucosamine. *Journal of Bacteriology* *190*, 3670–3680.
- Itoh, Y., Wang, X., Hinnebusch, B.J., Preston, J.F., and Romeo, T. (2005). Depolymerization of beta-1,6-N-acetyl-D-glucosamine disrupts the integrity of diverse bacterial biofilms. *Journal of Bacteriology* *187*, 382–387.
- Jackson, K.D., Starkey, M., Kremer, S., Parsek, M.R., and Wozniak, D.J. (2004). Identification of psl, a locus encoding a potential exopolysaccharide that is essential for Pseudomonas aeruginosa PAO1 biofilm formation. *Journal of Bacteriology* *186*, 4466–4475.
- Jardetzky, O. (1966). Simple allosteric model for membrane pumps. *Nature* *211*, 969–970.

- Jones, H.E., Holland, I.B., Baker, H.L., and Campbell, A.K. (1999). Slow changes in cytosolic free Ca^{2+} in *Escherichia coli* highlight two putative influx mechanisms in response to changes in extracellular calcium. *Cell Calcium* 25, 265–274.
- Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* 66, 125–132.
- Kanematsu, H., and Barry, D.M. (2015). Conditioning films. *Biofilm and Materials Science*.
- Kanonenberg, K., Schwarz, C.K.W., and Schmitt, L. (2013). Type I secretion systems - a story of appendices. *Res. Microbiol.* 164, 596–604.
- Karatan, E., and Watnick, P. (2009). Signals, Regulatory Networks, and Materials That Build and Break Bacterial Biofilms. *Microbiology and Molecular Biology Reviews* 73, 310–347.
- Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. *Science* 336, 1030–1033.
- Keegstra, K. (2010). Plant cell walls. *Plant Physiology* 154, 483–486.
- Keiski, C.-L., Harwich, M., Jain, S., Neculai, A.M., Yip, P., Robinson, H., Whitney, J.C., Riley, L., Burrows, L.L., Ohman, D.E., et al. (2010). AlgK is a TPR-containing protein and the periplasmic component of a novel exopolysaccharide secretin. *Structure* 18, 265–273.
- Kim, J.-S., Song, S., Lee, M., Lee, S., Lee, K., and Ha, N.-C. (2016). Crystal Structure of a Soluble Fragment of the Membrane Fusion Protein HlyD in a Type I Secretion System of Gram-Negative Bacteria. *Structure/Folding and Design* 24, 477–485.
- Kimura, S., Laosinchai, W., Itoh, T., Cui, X., Linder, C., and Brown, R. (1999). Immunogold labeling of rosette terminal cellulose-synthesizing complexes in the vascular plant *Vigna angularis*. *Plant Cell* 11, 2075–2086.
- Kimura, S., Ohshima, C., Hirose, E., Nishikawa, J., and Itoh, T. (2001). Cellulose in the house of the appendicularian *Oikopleura rufescens*. *Protoplasma* 216, 71–74.
- Klemm, D., Heublein, B., Fink, H.-P., and Bohn, A. (2005). Cellulose: Fascinating Biopolymer and Sustainable Raw Material. *Angew. Chem. Int. Ed.* 44, 3358–3393.
- Ko, J., Ryu, K.-S., Kim, H., Shin, J.-S., Lee, J.-O., Cheong, C., and Choi, B.-S. (2010). Structure of PP4397 Reveals the Molecular Basis for Different c-di-GMP Binding Modes by PilZ Domain Proteins. *Journal of Molecular Biology* 398, 97–110.
- Koo, H.M., Song, S.H., Pyun, Y.R., and Kim, Y.S. (1998). Evidence that a beta-1,4-endoglucanase secreted by *Acetobacter xylinum* plays an essential role for the formation of cellulose fiber. *Biosci. Biotechnol. Biochem.* 62, 2257–2259.

- Koronakis, E., Hughes, C., Milisav, I., and Koronakis, V. (1995). Protein exporter function and in vitro ATPase activity are correlated in ABC-domain mutants of HlyB. *Mol. Microbiol.* *16*, 87–96.
- Koronakis, V., Hughes, C., and Koronakis, E. (1991). Energetically distinct early and late stages of HlyB/HlyD-dependent secretion across both *Escherichia coli* membranes. *The EMBO Journal* *10*, 3263–3272.
- Koronakis, V., Li, J., Koronakis, E., and Stauffer, K. (1997). Structure of TolC, the outer membrane component of the bacterial type I efflux system, derived from two-dimensional crystals. *Mol. Microbiol.* *23*, 617–626.
- Koronakis, V., Sharff, A., Koronakis, E., Luisi, B., and Hughes, C. (2000). Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. *Nature* *405*, 914–919.
- Koyama, M., Helbert, W., Imai, T., Sugiyama, J., and Henrissat, B. (1997). Parallel-up structure evidences the molecular directionality during biosynthesis of bacterial cellulose. *Proc. Natl. Acad. Sci. U.S.A.* *94*, 9091–9095.
- Kozlikova, B., Sebestova, E., Sustr, V., Brezovsky, J., Strnad, O., Daniel, L., Bednar, D., Pavelka, A., Manak, M., Bezdeka, M., et al. (2014). CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics* *30*, 2684–2685.
- Kumari, M., Sunoj, R.B., and Balaji, P.V. (2012). Exploration of CH $\cdots\pi$ mediated stacking interactions in saccharide: aromatic residue complexes through conformational sampling. *Carbohydrate Research* *361*, 133–140.
- Lai-Kee-Him, J., Chanzy, H., Müller, M., Putaux, J.-L., Imai, T., and Bulone, V. (2002). In vitro versus in vivo cellulose microfibrils from plant primary wall synthases: structural differences. *J. Biol. Chem.* *277*, 36931–36939.
- Lairson, L.L., Henrissat, B., Davies, G.J., and Withers, S.G. (2008). Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* *77*, 521–555.
- Lecher, J., Schwarz, C.K.W., Stoldt, M., Smits, S.H.J., Willbold, D., and Schmitt, L. (2012). An RTX Transporter Tethers Its Unfolded Substrate during Secretion via a Unique N-Terminal Domain. *Structure/Folding and Design* *20*, 1778–1787.
- Lee, J.Y., Yang, J.G., Zhitnitsky, D., Lewinson, O., and Rees, D.C. (2014). Structural Basis for Heavy Metal Detoxification by an Atm1-Type ABC Exporter. *Science* *343*, 1133–1136.
- Lenardon, M.D., Munro, C.A., and Gow, N.A. (2010). Chitin synthesis and fungal pathogenesis. *Current Opinion in Microbiology* *13*, 416–423.

- Leslie, A.G.W. (2006). The integration of macromolecular diffraction data. *Acta Crystallogr. D Biol. Crystallogr.* *62*, 48–57.
- Levengood, M.R., Splain, R.A., and Kiessling, L.L. (2011). Monitoring Processivity and Length Control of a Carbohydrate Polymerase. *J. Am. Chem. Soc.* *133*, 12758–12766.
- Létoffé, S., Delepelaire, P., and Wandersman, C. (1990). Protease secretion by *Erwinia chrysanthemi*: the specific secretion functions are analogous to those of *Escherichia coli* alpha-haemolysin. *The EMBO Journal* *9*, 1375–1382.
- Létoffé, S., Delepelaire, P., and Wandersman, C. (1996). Protein secretion in gram-negative bacteria: assembly of the three components of ABC protein-mediated exporters is ordered and promoted by substrate binding. *The EMBO Journal* *15*, 5804–5811.
- Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* *531*, 395–399.
- Lin, D.Y.-W., Huang, S., and Chen, J. (2015). Crystal structures of a polypeptide processing and secretion transporter. *Nature* *523*, 425–430.
- Linhartová, I., Bumba, L., Masin, J., Basler, M., Osička, R., Kamanová, J., Procházková, K., Adkins, I., Hejnová-Holubová, J., Sadílková, L., et al. (2010). RTX proteins: a highly diverse family secreted by a common mechanism. *FEMS Microbiol Rev* *34*, 1076–1112.
- Lizak, C., Gerber, S., Numao, S., Aebi, M., and Locher, K.P. (2011). X-ray structure of a bacterial oligosaccharyltransferase. *Nature* *474*, 350–355.
- Lizak, C., Gerber, S., Zinne, D., Michaud, G., Schubert, M., Chen, F., Bucher, M., Darbre, T., Zenobi, R., Reymond, J.L., et al. (2014). A Catalytically Essential Motif in External Loop 5 of the Bacterial Oligosaccharyltransferase PglB. *Journal of Biological Chemistry* *289*, 735–746.
- Locher, K.P. (2016). Mechanistic diversity in ATP-binding cassette (ABC) transporters. *Nature Structural & Molecular Biology* *23*, 487–493.
- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P.M., and Henrissat, B. (2013). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucl. Acids Res.* *42*, D490–D495.
- Manat, G., Roure, S., Auger, R., Bouhss, A., Barreteau, H., Mengin-Lecreulx, D., and Touzé, T. (2014). Deciphering the Metabolism of Undecaprenyl-Phosphate: The Bacterial Cell-Wall Unit Carrier at the Membrane Frontier. *Microbial Drug Resistance* *20*, 199–214.
- Martin, A., Baker, T.A., and Sauer, R.T. (2008). Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nature Structural & Molecular Biology* *15*, 1147–1151.

- Martin, J.L., Johnson, L.N., and Withers, S.G. (1990). Comparison of the binding of glucose and glucose 1-phosphate derivatives to T-state glycogen phosphorylase b. *Biochemistry* *29*, 10745–10757.
- Masi, M., and Wandersman, C. (2010). Multiple Signals Direct the Assembly and Function of a Type 1 Secretion System. *Journal of Bacteriology* *192*, 3861–3869.
- Matsumoto, S., Shimada, A., and Nyirenda, J. (2013). Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation.
- Matthysse, A.G., Holmes, K.V., and Gurlitz, R.H. (1981). Elaboration of cellulose fibrils by *Agrobacterium tumefaciens* during attachment to carrot cells. *Journal of Bacteriology*.
- Mayer, R., Ross, P., Weinhouse, H., Amikam, D., Volman, G., Ohana, P., Calhoon, R.D., Wong, H.C., Emerick, A.W., and Benziman, M. (1991). Polypeptide composition of bacterial cyclic diguanylic acid-dependent cellulose synthase and the occurrence of immunologically crossreacting proteins in higher plants. *Proc. Natl. Acad. Sci. U.S.A.* *88*, 5472–5476.
- McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J Appl Crystallogr* *40*, 658–674.
- McCrate, O.A., Zhou, X., Reichhardt, C., and Cegelski, L. (2013). Sum of the parts: composition and architecture of the bacterial extracellular matrix. *Journal of Molecular Biology* *425*, 4286–4294.
- McFarlane, H.E., Döring, A., and Persson, S. (2014). The Cell Biology of Cellulose Synthesis. *Annu. Rev. Plant Biol.* *65*, 69–94.
- McNamara, J.T., Morgan, J.L.W., and Zimmer, J. (2015). A Molecular Description of Cellulose Biosynthesis. *Annu. Rev. Biochem.* *84*, 895–921.
- Merighi, M., Lee, V.T., Hyodo, M., Hayakawa, Y., and Lory, S. (2007). The second messenger bis-(3′-5′)-cyclic-GMP and its PilZ domain-containing receptor Alg44 are required for alginate biosynthesis in *Pseudomonas aeruginosa*. *Mol. Microbiol.* *65*, 876–895.
- Merzendorfer, H. (2006). Insect chitin synthases: a review. *J. Comp. Physiol. B, Biochem. Syst. Environ. Physiol.* *176*, 1–15.
- Merzendorfer, H., and Zimoch, L. (2003). Chitin metabolism in insects: structure, function and regulation of chitin synthases and chitinases. *J. Exp. Biol.* *206*, 4393–4412.
- Mohnen, D. (2008). Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* *11*, 266–277.
- Morgan, J.L.W., McNamara, J.T., and Zimmer, J. (2014). Mechanism of activation of

bacterial cellulose synthase by cyclic di-GMP. *Nature Structural & Molecular Biology* 21, 489–496.

Morgan, J.L.W., McNamara, J.T., Fischer, M., Rich, J., Chen, H.-M., Withers, S.G., and Zimmer, J. (2016). Observing cellulose biosynthesis and membrane translocation in crystallo. *Nature* 531, 329–334.

Morgan, J.L.W., Strumillo, J., and Zimmer, J. (2013). Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* 493, 181–186.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., and Sliz, P. (2013). Collaboration gets the most out of software. *eLife* 2, e01456.

Mukaiyama, T., and Takahashi, K. (1968). A convenient method for the preparation of unsymmetrical disulfides by the use of diethyl azodicarboxylate. *Tetrahedron Letters* 56, 5907–5908.

Newman, R.H., Hill, S.J., and Harris, P.J. (2013). Wide-Angle X-Ray Scattering and Solid-State Nuclear Magnetic Resonance Data Combined to Test Models for Cellulose Microfibrils in Mung Bean Cell Walls. *Plant Physiology* 163, 1558–1567.

Nishiyama, Y., Sugiyama, J., and Chanzy, H. (2003). Crystal structure and hydrogen bonding system in cellulose Ia from synchrotron X-ray and neutron fiber diffraction. *Journal of the American ...*

Nixon, B.T., Mansouri, K., Singh, A., Du, J., Davis, J.K., Lee, J.-G., Slabaugh, E., Vandavasi, V.G., O'Neill, H., Roberts, E.M., et al. (2016). Comparative Structural and Computational Analysis Supports Eighteen Cellulose Synthases in the Plant Cellulose Synthesis Complex. *Sci. Rep.* 1–14.

Notley, S.M., Pettersson, B., and Wågberg, L. (2004). Direct Measurement of Attractive van der Waals' Forces between Regenerated Cellulose Surfaces in an Aqueous Environment. *J. Am. Chem. Soc.* 126, 13930–13931.

Oldham, M.L., Hite, R.K., Steffen, A.M., Damko, E., Li, Z., Walz, T., and Chen, J. (2016). A mechanism of viral immune evasion revealed by cryo-EM analysis of the TAP transporter. *Nature* 529, 537–540.

Omadjela, O., Narahari, A., Strumillo, J., Mélida, H., Mazur, O., Bulone, V., and Zimmer, J. (2013). BcsA and BcsB form the catalytically active core of bacterial cellulose synthase sufficient for in vitro cellulose synthesis.

Painter, J., and Merritt, E.A. (2006a). Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr. D Biol. Crystallogr.* 62, 439–450.

Painter, J., and Merritt, E.A. (2006b). TLSMD web server for the generation of multi-group TLS models. *J. Appl. Cryst* 39, 109–111.

- Palmgren, M.G., and Nissen, P. (2011). P-Type ATPases. *Annu. Rev. Biophys.* *40*, 243–266.
- Park, E., and Rapoport, T.A. (2012). Mechanisms of Sec61/SecY-Mediated Protein Translocation Across Membranes. *Annu. Rev. Biophys.* *41*, 21–40.
- Park, Y., Moon, Y., Ryoo, J., Kim, N., Cho, H., and Ahn, J.H. (2012). Identification of the minimal region in lipase ABC transporter recognition domain of *Pseudomonas fluorescens* for secretion and fluorescence of green fluorescent protein. *Microb. Cell Fact.* *11*, 60.
- Parsek, M.R., and Fuqua, C. (2004). Biofilms 2003: Emerging Themes and Challenges in Studies of Surface-Associated Microbial Life. *Journal of Bacteriology*.
- Persson, K., Ly, H.D., Dieckelmann, M., Wakarchuk, W.W., Withers, S.G., and Strynadka, N.C. (2001). Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs. *Nat Struct Biol* *8*, 166–175.
- Pritt, B., O'Brien, L., and Winn, W. (2007). Mucoid *Pseudomonas* in cystic fibrosis. *Am. J. Clin. Pathol.* *128*, 32–34.
- Putney, J.W., and Tomita, T. (2012). Phospholipase C signaling and calcium influx. *Advances in Biological Regulation* *52*, 152–164.
- Qasba, P.K., Ramakrishnan, B., and Boeggeman, E. (2008). Structure and function of beta -1,4-galactosyltransferase. *Curr Drug Targets* *9*, 292–309.
- Quan, S., Hiniker, A., Collet, J.-F., and Bardwell, J.C.A. (2013). Isolation of bacteria envelope proteins. *Methods Mol. Biol.* *966*, 359–366.
- Ramakrishnan, B., Ramasamy, V., and Qasba, P.K. (2006). Structural snapshots of beta-1,4-galactosyltransferase-I along the kinetic pathway. *Journal of Molecular Biology* *357*, 1619–1633.
- Rees, D.C., Johnson, E., and Lewinson, O. (2009). ABC transporters: the power to change. *Nat Rev Mol Cell Biol* *10*, 218–227.
- Rehm, B. (2009). Alginate production: precursor biosynthesis, polymerization and secretion. *Alginates: Biology and Applications* *13*, 55–71.
- Reimann, S., Poschmann, G., and Ka, K. (2016). Interdomain regulation of the ATPas. *Biochemical Journal*.
- Ross, P., Weinhouse, H., Aloni, Y., Michaeli, D., Weinberger-Ohana, P., Mayer, R., Braun, S., de Vroom, E., van der Marel, G.A., van Boom, J.H., et al. (1987). Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. *Nature* *325*, 279–281.

- Römmling, U. (2002). Molecular biology of cellulose production in bacteria. *Res. Microbiol.* *153*, 205–212.
- Römmling, U., and Balsalobre, C. (2012). Biofilm infections, their resilience to therapy and innovative treatment strategies. *J. Intern. Med.* *272*, 541–561.
- Römmling, U., Galperin, M.Y., and Gomelsky, M. (2013). Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. *Microbiol. Mol. Biol. Rev.* *77*, 1–52.
- Römmling, U., Gomelsky, M., and Galperin, M.Y. (2005). C-di-GMP: the dawning of a novel bacterial signalling system. *Mol. Microbiol.* *57*, 629–639.
- Saraogi, I., and Shan, S.-O. (2014). Co-translational protein targeting to the bacterial membrane. *BBA - Molecular Cell Research* *1843*, 1433–1441.
- Saxena, I.M., and Brown, R.M. (1997). Identification of cellulose synthase (s) in higher plants: sequence analysis of processive β -glycosyltransferases with the common motif D, D, D35Q (R, Q) XRW'. *Cellulose* *4*, 33–49.
- Saxena, I.M., Brown, R.M., and Dandekar, T. (2001). Structure--function characterization of cellulose synthase: relationship to other glycosyltransferases. *Phytochemistry* *57*, 1135–1148.
- Saxena, I.M., Brown, R.M., Fevre, M., Geremia, R.A., and Henrissat, B. (1995). Multidomain architecture of beta-glycosyl transferases: implications for mechanism of action. *Journal of Bacteriology* *177*, 1419–1424.
- Saxena, I.M., Kudlicka, K., Okuda, K., and Brown, R.M. (1994). Characterization of genes in the cellulose-synthesizing operon (acs operon) of *Acetobacter xylinum*: implications for cellulose crystallization. *Journal of Bacteriology* *176*, 5735–5752.
- Scharnagl, C., Pester, O., Hornburg, P., Hornburg, D., Götz, A., and Langosch, D. (2014). Side-chain to main-chain hydrogen bonding controls the intrinsic backbone dynamics of the amyloid precursor protein transmembrane helix. *Biophys. J.* *106*, 1318–1326.
- Scheible, W.R., Eshed, R., Richmond, T., Delmer, D., and Somerville, C. (2001). Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in *Arabidopsis* *Ixr1* mutants. *Proc. Natl. Acad. Sci. U.S.a.* *98*, 10079–10084.
- Scheller, H.V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* *61*, 263–289.
- Schlesinger, M.J. (1968). Secretion of alkaline phosphatase subunits by spheroplasts of *Escherichia coli*. *Journal of Bacteriology* *96*, 727–733.
- Schmid, J., Sieber, V., and Rehm, B. (2015). Bacterial exopolysaccharides: biosynthesis pathways and engineering strategies. *Front Microbiol* *6*, 875.

- Schuman, B., Evans, S.V., and Fyles, T.M. (2013). Geometric Attributes of Retaining Glycosyltransferase Enzymes Favor an Orthogonal Mechanism. *PLoS ONE* 8, e71077.
- Schurr, M.J. (2013). Which bacterial biofilm exopolysaccharide is preferred, Psl or alginate? *Journal of Bacteriology* 195, 1623–1626.
- Serra, D.O., Richter, A.M., and Hengge, R. (2013). Cellulose as an architectural element in spatially structured *Escherichia coli* biofilms. *Journal of Bacteriology* 195, 5540–5554.
- Sheldrick, G.M. (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.* 66, 479–485.
- Sheps, J.A., Cheung, I., and Ling, V. (1995). Hemolysin transport in *Escherichia coli*. Point mutants in HlyB compensate for a deletion in the predicted amphiphilic helix region of the HlyA signal. *J. Biol. Chem.* 270, 14829–14834.
- Simm, R., Morr, M., Kader, A., Nimtz, M., and Römling, U. (2004). GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. *Mol. Microbiol.* 53, 1123–1134.
- Slabaugh, E., Davis, J.K., Haigler, C.H., Yingling, Y.G., and Zimmer, J. (2014). Cellulose synthases: new insights from crystallography and modeling. *Trends Plant Sci.* 19, 99–106.
- Soderberg, Timothy, "Organic Chemistry with a Biological Emphasis Volume I" (2016). *Chemistry Faculty*. Paper 1.
http://digitalcommons.morris.umn.edu/chem_facpubs/1
- Somerville, C. (2006). Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* 22, 53–78.
- Sondermann, H., Shikuma, N.J., and Yildiz, F.H. (2012). You've come a long way: c-di-GMP signaling. *Current Opinion in Microbiology* 15, 140–146.
- Stapper, A.P., Narasimhan, G., Ohman, D.E., Barakat, J., Hentzer, M., Molin, S., Kharazmi, A., Høiby, N., and Mathee, K. (2004). Alginate production affects *Pseudomonas aeruginosa* biofilm development and architecture, but is not essential for biofilm formation. *J. Med. Microbiol.* 53, 679–690.
- Steiner, S., Lori, C., Boehm, A., and Jenal, U. (2013). Allosteric activation of exopolysaccharide synthesis through cyclic di-GMP-stimulated protein-protein interaction. *The EMBO Journal* 32, 354–368.
- Stewart, A.G., Laming, E.M., Sobti, M., and Stock, D. (2014). Rotary ATPases—dynamic molecular machines. *Current Opinion in Structural Biology* 25, 40–48.
- Stewart, P.S., and Costerton, J.W. (2001). Antibiotic resistance of bacteria in biofilms. *The Lancet* 358, 135–138.

- Sutherland, I. (2001). Biofilm exopolysaccharides: a strong and sticky framework. *Microbiology (Reading, Engl.)* 147, 3–9.
- Symmons, M.F., Marshall, R.L., and Bavro, V.N. (2015). Architecture and roles of periplasmic adaptor proteins in tripartite efflux assemblies. *Front Microbiol* 6, 513.
- Tanford, C. (1982). Simple model for the chemical potential change of a transported ion in active transport.
- Terwilliger, T.C. (2004). Using prime-and-switch phasing to reduce model bias in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* 60, 2144–2149.
- Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* 64, 61–69.
- Tester, R.F., Karkalas, J., and Qi, X. (2004). Starch—composition, fine structure and architecture. *Journal of Cereal Science* 39, 151–165.
- Thanabalu, T., Koronakis, E., Hughes, C., and Koronakis, V. (1998). Substrate-induced assembly of a contiguous channel for protein export from E.coli: reversible bridging of an inner-membrane translocase to an outer membrane exit pore. *The EMBO Journal* 17, 6487–6496.
- Thomas, S., Holland, I.B., and Schmitt, L. (2014). The Type 1 secretion pathway — The hemolysin system and beyond. *BBA - Molecular Cell Research* 1843, 1629–1641.
- Thorens, B., and Mueckler, M. (2010). Glucose transporters in the 21st Century. *Am. J. Physiol. Endocrinol. Metab.* 298, E141–E145.
- Tikhonova, E.B., Devroy, V.K., Lau, S.Y., and Zgurskaya, H.I. (2007). Reconstitution of the Escherichia coli macrolide transporter: the periplasmic membrane fusion protein MacA stimulates the ATPase activity of MacB. *Mol. Microbiol.* 63.
- Tlapak-Simmons, V.L., Baron, C.A., and Weigel, P.H. (2004). Characterization of the purified hyaluronan synthase from Streptococcus equisimilis. *Biochemistry* 43, 9234–9242.
- Vagin, A., and Teplyakov, A. (2010). Molecular replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* 66, 22–25.
- Van den Berg, B., Clemons, W.M., Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C., and Rapoport, T.A. (2004). X-ray structure of a protein-conducting channel. *Nature* 427, 36–44.
- Venter, H., Shilling, R.A., Velamakanni, S., and Balakrishnan, L. (2003). An ABC transporter with a secondary-active multidrug translocator domain. *Nature*.

- Vocadlo, D.J., Davies, G.J., Laine, R., and Withers, S.G. (2001). Catalysis by hen egg-white lysozyme proceeds via a covalent intermediate. *Nature* *412*, 835–838.
- Vu, B., Chen, M., Crawford, R.J., and Ivanova, E.P. (2009). Bacterial Extracellular Polysaccharides Involved in Biofilm Formation. *Molecules* *14*, 2535–2554.
- Wang, X., Preston, J.F., and Romeo, T. (2004). The pgaABCD locus of *Escherichia coli* promotes the synthesis of a polysaccharide adhesin required for biofilm formation. *Journal of Bacteriology* *186*, 2724–2734.
- Whitney, J.C., and Howell, P.L. (2013). Synthase-dependent exopolysaccharide secretion in Gram-negative bacteria. *Trends in Microbiology* *21*, 63–72.
- Whitney, J.C., Hay, I.D., Li, C., Eckford, P.D.W., Robinson, H., Amaya, M.F., Wood, L.F., Ohman, D.E., Bear, C.E., Rehm, B.H., et al. (2011). Structural basis for alginate secretion across the bacterial outer membrane. *Proc. Natl. Acad. Sci. U.S.A.* *108*, 13083–13088.
- Wiling, J.N., Zaburdaev, V., De Volder, M., Losick, R., Brenner, M.P., and Weitz, D.A. (2013). Liquid transport facilitated by channels in *Bacillus subtilis* biofilms. *Proc. Natl. Acad. Sci. U.S.A.* *110*, 848–852.
- Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A., et al. (2011). Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* *67*, 235–242.
- Wright, E.M., Loo, D., Turk, E., and Hirayama, B.A. (1996). Sodium cotransporters. *Current Opinion in Cell Biology*.
- Wright, E.M., Loo, D.D.F., and Hirayama, B.A. (2011). Biology of human sodium glucose transporters. *Physiol. Rev.* *91*, 733–794.
- Xu, C., and Ng, D.T.W. (2015). Glycosylation-directed quality control of protein folding. *Nat Rev Mol Cell Biol* *16*, 742–752.
- Xu, K.D., Stewart, P.S., Xia, F., Huang, C.T., and McFeters, G.A. (1998). Spatial physiological heterogeneity in *Pseudomonas aeruginosa* biofilm is determined by oxygen availability. *Applied and Environmental Microbiology* *64*, 4035–4039.
- Yang, H., Zimmer, J., Yingling, Y.G., and Kubicki, J.D. (2015). How Cellulose Elongates--A QM/MM Study of the Molecular Mechanism of Cellulose Polymerization in Bacterial CESA. *J. Phys. Chem. B* *119*, 6525–6535.
- Yang, S.-C., Lin, C.-H., Sung, C.T., and Fang, J.-Y. (2014). Antibacterial activities of bacteriocins: application in foods and pharmaceuticals. *Front Microbiol* *5*, 241.
- Yu, H., Takeuchi, M., LeBarron, J., Kantharia, J., London, E., Bakker, H., Haltiwanger,

R.S., Li, H., and Takeuchi, H. (2015). Notch-modifying xylosyltransferase structures support an S_Ni-like retaining mechanism. *Nature Publishing Group* *11*, 847–856.

Zhang, F., Sheps, J.A., and Ling, V. (1993). Complementation of transport-deficient mutants of *Escherichia coli* alpha-hemolysin by second-site mutations in the transporter hemolysin B. *J. Biol. Chem.* *268*, 19889–19895.

Zhang, Z., Kim, S., Gaffney, B.L., and Jones, R.A. (2006). Polymorphism of the Signaling Molecule c-di-GMP. *J. Am. Chem. Soc.* *128*, 7015–7024.

Zhao, F.-Q., and Keating, A.F. (2007). Functional properties and genomics of glucose transporters. *Curr. Genomics* *8*, 113–128.

Zimmer, J., Nam, Y., and Rapoport, T.A. (2008). Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* *455*, 936–943.

Zogaj, X., Bokranz, W., Nimtz, M., and Römling, U. (2003). Production of cellulose and curli fimbriae by members of the family Enterobacteriaceae isolated from the human gastrointestinal tract. *Infect. Immun.* *71*, 4151–4158.