Structural biology of energy-transducing proteins from *Escherichia coli*

By

Damien Biot-Pelletier

Department of Microbiology and Immunology

Faculty of Medicine, McGill University, Montreal, QC

July 2010

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

Abstract

The cell envelope of Gram-negative bacteria such as *Escherichia coli* is composed of an outer (OM) and cytoplasmic membrane (CM), which enclose a space known as the periplasm. For acquisition of essential nutrients, bacteria must transport molecules across both membranes. The import of certain scarce nutrients, such as iron and vitamin B_{12}, requires the input of energy; however, the OM is devoid of any source of energy. To drive the transport of scarce nutrients across the OM, bacteria use the proton motive force of the CM. The TonB–ExbB–ExbD protein complex located within the CM is responsible for harnessing the energy of this proton gradient and transducing it to OM receptors. While the existence and the role of this complex has been investigated many years ago by genetic means and by *in vivo* crosslinking experiments, structural and functional aspects remain largely unknown. To date, only partial structures of periplasmic domains of TonB and ExbD are known. Additionally, information on the structure of the complex as well as stoichiometry has not been deduced. We have therefore designed a strategy for the co-expression and co-purification of ExbB and ExbD in complex. Here we produced milligram amounts of pure ExbB–ExbD; size-exclusion chromatography revealed that our samples were homogenous and monodisperse. Observation of our samples by negative-staining electron microscopy revealed uniform particles from which 2D reconstructions were generated. These reconstructions suggested a pentameric and perhaps cylindrical arrangement for the complex. In parallel, crystallization trials have identified promising leads. Our optimization efforts have led to the successful detergent exchange of the protein

complex from dodecyl maltoside to decyl maltoside, to GNG-12 and to MNG-28; the latter two amphiphiles are proprietary and of undisclosed structure. We also adopted an NMR detergent quantitation method to standardize amounts of detergent in samples bound for crystallization. Preliminary attempts revealed that samples display lower detergent concentration when compared to protein-free buffer. Together, these efforts steer us towards a high-resolution structure of the ExB-ExbD complex, specifically to advance our understanding of the molecular mechanism of scarce nutrient import in Gram-negative bacteria.

Résumé

L'enveloppe cellulaire des bactéries à Gram négatif tel *Escherichia coli* est constituée d'une membrane externe (ME) et d'une membrane cytoplasmique (MC), entre lesquelles se trouve un espace nommé périplasme. L'acquisition de nutriments essentiels par les bactéries nécessite le transport de molécules à travers chacune des deux membranes. Le transport de certains nutriments rares, tel le fer et la vitamine B_{12}, requiert un apport en énergie. Cependant, la ME est dénuée de toute source d'énergie. Ainsi, pour alimenter en énergie le transport de nutriments rares à la ME, les bactéries font appel à la force proton motrice de la MC. Le complexe protéique TonB–ExbB–ExbD, situé à la MC, arnache l'énergie de ce gradient de protons et la transduit aux récepteurs de la ME. Bien que l'existence et le rôle de ce complexe aient été sondés il y a plusieurs années par des voies génétiques et des expériences de réticulation *in vivo*, plusieurs aspects de la structure et de la fonction du complexe restent inconnus. Jusqu'à maintenant, seule la structure des domaines périplasmiques de TonB et ExbD sont connus. Qui plus est, le structure du complexe et sa stoechiometrie n'ont pas été résolus. Nous avons donc élaborée une stratégie pour la co-expression et la co-purification d'ExbB et d'ExbD en complexe. Nous produisons plusieurs milligrammes d'ExbB–ExbD pur et la chromatographie par filtration sur gel révèle que nos échantillons sont homogènes et monodispersés. L'observation de nos échantillons par microscopie électronique en coloration négative révèle des particules de taille uniformes à partir desquelles des reconstructions 2D ont été générées. Ces reconstructions semblent suggérer un arrangement pentamérique et peut-être cylindrique. En outre, nos essais de

cristallisation nous ont permis d'identifier des pistes prometteuses. Nos efforts d'optimisation nous ont amenés à échanger avec succès notre protéine depuis le détergent original, le DDM, vers le DM, le GNG-12 et le MNG-28. Nous avons aussi adopté une méthode de quantification du détergent par RMN, avec pour objectif d'uniformiser la quantité de détergent présente dans les échantillons destinés à la cristallisaton. Les essais préliminaires de quantification révèlent que nos échantillons présentent une concentration de détergent réduite par rapport au tampon de purification. Ensemble, ces efforts nous rapprochent d'une structure à haute résolution du complexe ExbB–ExbD. Cette structure devrait parfaire notre compréhension des mécanismes moléculaires du transport des nutriments rares chez les bactéries Gram-négatives.

Acknowledgements

First, I must thank my supervisor, Dr James W. Coulton, for the opportunity to work on this challenging project as part of his research team. He has been a supportive and dedicated mentor, and for that I am very grateful. I must also thank him for his trust and for the many great opportunities he opened for me.

Next, I wish to thank members of the lab, past and present, for their scientific and technical support as well as for their good company. They are, in no particular order : Karron James, David Carter, Christopher Ng Thow Hing, Kristian Levey, Jacqueline Chung, Colins Vasquez, Maria Plesa, Justin Deme, Isabelle Racine-Miousse, Jaeseung Kim, Andrew Cottreau, Aleksandr Sverzhinky and Nathalie Croteau. I owe something special to each and every single one of them. Specifically, Maria Plesa and Kristian Levey introduced me to work with membrane proteins in general and with ExbB and ExbD in particular. Nathalie Croteau taught me the where and abouts of protein crystallization and I admire her sixth sense for crystals. Andrew Cottreau was a valuable EM partner and I thank him for picking up reconstructions after I left for the laboratory of our collaborator at UCLA. I also wish to acknowledge his input on the EM-related parts of my thesis, and I thank him for allowing me to display his reconstructions in the thesis. I wish to thank, also, Justin Deme and Jacqueline Chung for their valuable and patient revisions of my thesis.

I wish to thank our collaborator, Dr. Jeffrey Abramson at the University of California Los Angeles, who generously welcomed me his laboratory for three months and who taught me new approaches and techniques in crystallization. I remember his

lively personality with much pleasure. I must also thank the members of his lab, again in no particular order : Vincent Chaptal, Gabriel Mercado Besserer, Yazan Hadid, Aviv Paz, Rachna Ujwal, Jennifer Hsieh, Akira Watanabe, Le Du, Nir Arbel and Prakash Reddy.

I wish to thank Dr. Isabelle Rouiller and Dr. Mihnea Bostina from the Department of Anatomy and Cell Biology for teaching me the basics of electron microscopy. None of the EM work could have been done without their precious collaboration.

I need to thank my family, who provided me with the stimulating environment and education that have sparked my intellectual curiosity. In particular, I must acknowledge the early philosophical teachings of my father. The humanist ideal and rationalist principles that he taught me are what brought me into science. I must also thank my fiancée, Olivia Lafontaine, soon to become my wife, and who has been an enormous moral support. She is also an inspiring model of patience, discipline, perseverance, responsibility and hard work. Thanks to her, I am a better person, and a better scientist.

Lastly, I must acknowledge and thank the Natural Sciences and Engineering Research Council, McGill University and the Fonds de la recherche en santé du Québec for their financial support.

Table of Contents

page

List of Tables and Figures

page

Chapter 1 – Literature review and introduction to conducted research

1.1 Gram-negative bacteria and the organization of their cell envelope

Bacteria are classically divided into two large categories based on their ability to retain Gram stain. The difference between the Gram-positive and Gram-negative bacteria relies on the fundamental organization of their cell envelope (1). The envelope of bacterial cells is a key feature for their classification, but foremost, it is essential for survival and cell integrity. It is a line of defence that protects the cell from physical and chemical stresses, and also serves as a selective barrier that regulates critical processes of nutrient uptake. In *E. coli* and other Gram-negative bacteria, the cell envelope consists of three concentric components. Figure 1 outlines the general organization of the Gram-negative envelope. It is basically made of two membranes, outer and cytoplasmic, which together enclose a space known as the periplasm.

The outermost component of the Gram-negative cell envelope is the outer membrane (OM). Like all biological membranes, it is a double-layer of amphipathic lipids. One of its important features is the distinct chemical nature of each leaflet. The outer leaflet of the OM is made of glycolipids, mostly lipopolysaccharide (LPS). The structure of LPS (Figure 2) varies according to environmental conditions and from species to species, but a common core can be identified. The base of the molecule is lipid A, a six acyl chain phosphoglycoside. The lipid is derivatized by a complex oligosaccharide moiety,

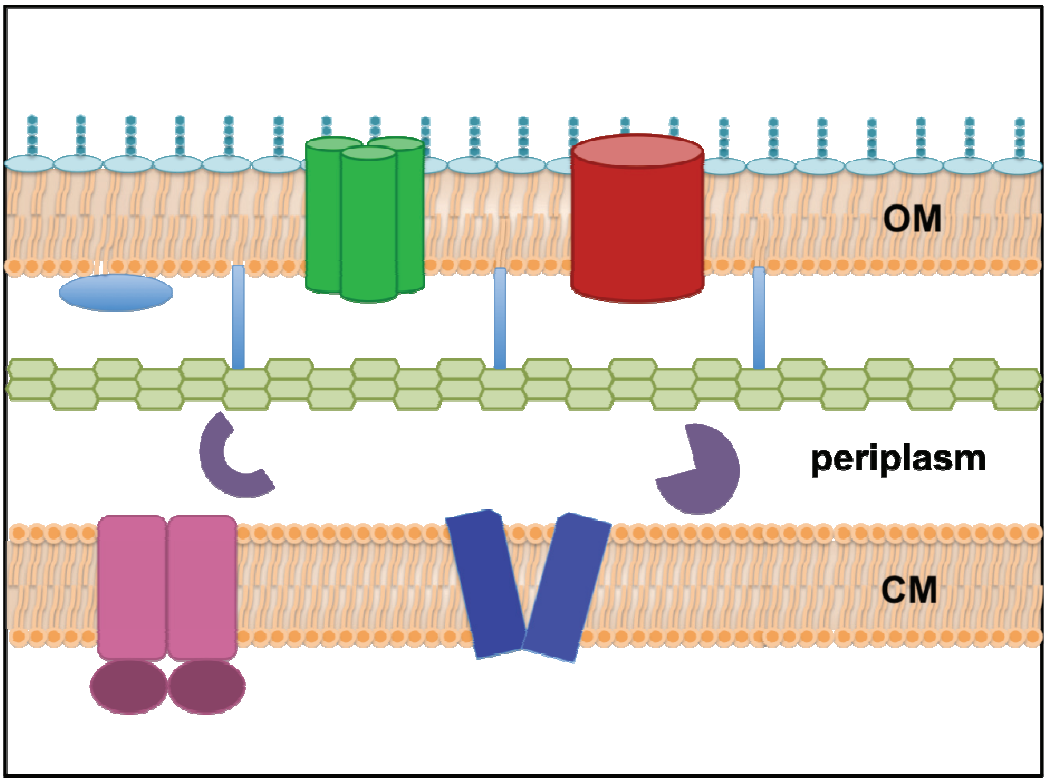

Figure 1. Schematic representation of the Gram-negative cell envelope. The envelope of Gram-negative bacteria consists of two membranes: an outer membrane (OM) and a cytoplasmic membrane (CM) which together enclose the periplasm. The outer leaflet of the OM is made of lipopolysaccharide, while its inner leaflet, similar to the CM, is predominantly made of phospholipids (orange balls and tails). Protein constituents of the OM are found as integral membrane proteins (green and red cylindrical structures) and membrane associated lipoproteins (blue ellipse with orange tail). In the periplasm, the peptidoglycan layer (kaki stack of hexagons) is anchored to the OM by Braun's lipoprotein (blue rectangles). Enzymes and proteins involved in transport are also found in the periplasm (purple circular structures). Integral membrane proteins, mostly involved in nutrient uptake are present in the CM (pink and blue shapes).

the variable part of which is known as the O-antigen (2). The large polar headgroup and coordinating divalent cations between its phosphate groups are what make the OM impermeable to hydrophobic molecule diffusion. Tight packing of the six acyl chains of LPS also contributes to the barrier quality of the Gram-negative envelope (2). In contrast, the inner leaflet, is made of a mixture of phospholipids (3).

The OM is the first obstacle that stands between the interior of the cell and its extracellular environment, excluding high molecular weight compounds and denying entry to bacteriophages and other harmful agents. Similarly, it retains components from inner compartments and prevents their dissemination into the environment. However, it is also a rather porous barrier containing porins and other transport proteins that permits free diffusion of water and low molecular weight compounds, as well as enabling controlled flow of sugars, vitamins and a plethora of other chemicals and nutrients (3). The rest of the protein component of the OM is made up of lipoproteins, most of which are of unknown function (4).

The periplasmic space, enclosed between the OM and the cytoplasmic membrane (CM), contains proteins that are involved in the shuttling of compounds travelling from the outside to the inside of the cell. Molecules and molecular assemblies bound for secretion transit by the periplasm and can be stored there (5,6). In addition, a number of enzymes - either involved in synthesis of periplasmic components or in reactions potentially harmful to the cell - are also found in the periplasm (3,4). A central feature of the periplasm is the presence of

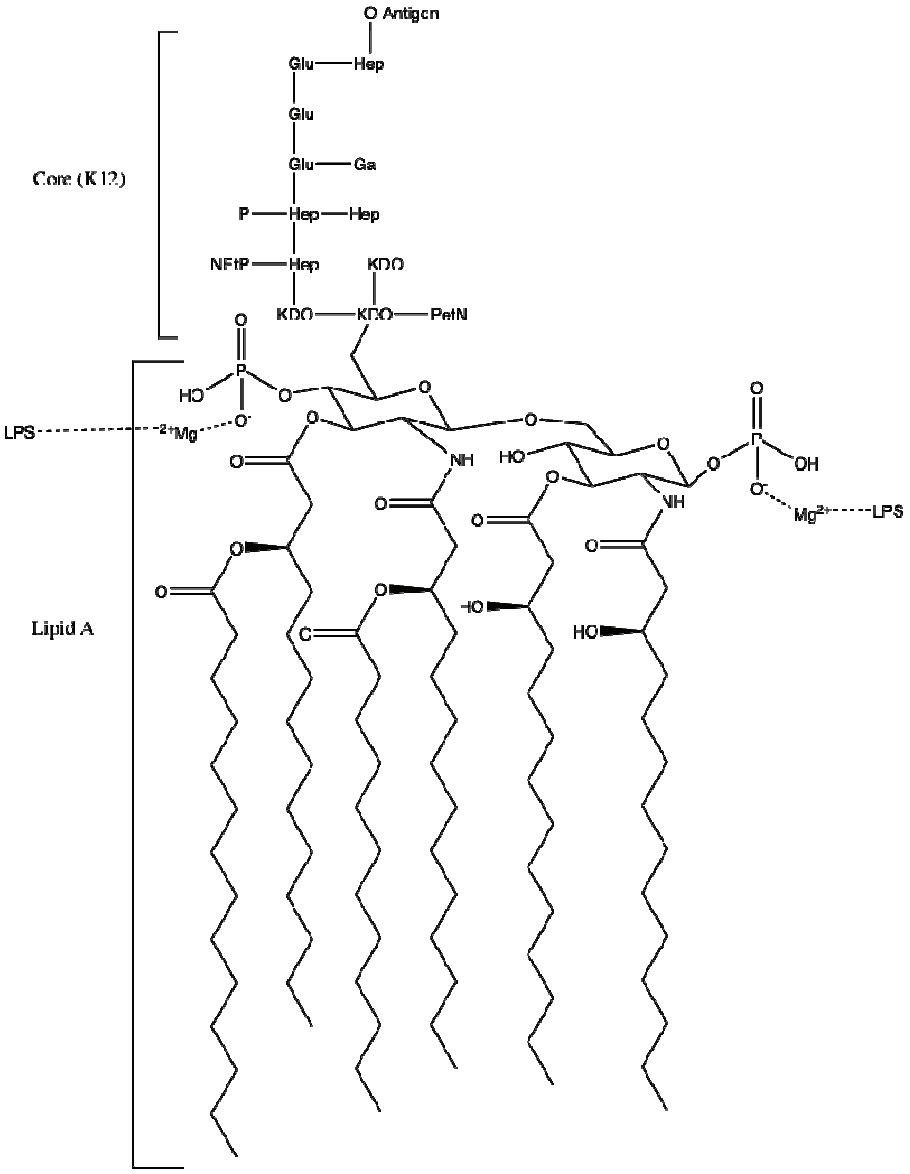

Figure 2. Core structure of lipopolysaccharide. The hydrophobic core of LPS, lipid A, is a tightly packed six-acyl chain phosphoglycoside. It is derivatized by a complex polysaccharide moiety. The *E. coli* K12 core of that moiety is presented here. The variable portion of the polysaccharide is known as O-antigen. Phosphate groups from neighboring LPS molecules coordinate divalent cations, contributing to the OM barrier. Abbreviations in the figure are the following: KDO, 3-deoxy-D-manno-oct-2-ulosonic acid; PetN, Phosphoethanolamine; Hep, l-glycero-D-manno-heptose; Glu, D-glucose; Gal, D-galactose.

the peptidoglycan layer. Peptidoglycan is a rigid polymer of N-acetylglucosamide-N-acetylmuramic acid disaccharides periodically crosslinked, in Gram-negatives, by a pentapeptide of sequence L-Ala-D-γ-Glu-diaminopimelate-D-Ala-D-Ala (7). It is anchored to the OM by Braun's lipoprotein, the most abundant protein in *E. coli* (8). The resistance of the peptidoglycan polymer to tensile forces contributes towards bacterial resistance to osmotic pressure and other mechanical stresses (9). This layer also plays a significant role in the determination of cell shape (4).

The CM is a phospholipid bilayer composed mainly of phosphatidylethanolamine and phosphatidylglycerol (10). Much like its outer counterpart, the CM is a major site of uptake and secretion performed by proteins embedded in the membrane. It is also central for cellular energetics as a consequence of the ion gradients that the cell maintains across it. In particular, the proton motive force, an H^+ gradient across the CM, directly or indirectly drives many energy-requiring cellular processes such as nutrient import and ATP synthesis. The CM also contains proteins involved in nutrient import, electron transport, electrophysiological homeostasis and lipid biosynthesis (4).

1.2 Nutrient import across the OM of Gram-negative bacteria

The transport of nutrients across the OM of Gram-negative bacteria occurs through two general processes. Water, ions, saccharides, amino acids and lipids are imported into the periplasm by diffusion through pores in the OM (Section 1.2.1). Scarce

nutrients such as vitamin B_{12} or iron require more elaborate uptake strategies that involve active transport (Sections 1.2.2, 1.3 and 1.4 and Figure 4).

1.2.1 Energy-independent transport

Porins are proteins that form the pores of the OM. Porins can themselves be divided into two groups based on the type of diffusion that they mediate. Non-specific or general porins are simple channels in the OM that allow the diffusion of any molecule that their central cavity can accommodate. Specific porins, as their name implies, allow the passage of only certain molecules based on precise chemical interactions. Energy-independent transporters of the FadL family, specialized in the import of hydrophobic compounds, are also found in the OM.

All porins share a common architecture. They all fold as hollow β-barrels and assemble into a characteristic homotrimeric organization. The hollow center of porins is constricted by loops that fold towards the interior of the barrel. Those loops constitute the inner lining of a porin and the amino acid side chains that decorate them form the structural basis for specificity and selectivity. The barrels of non-specific porins are made of 16 transmembrane β-strands, with a central constriction lined with polar moieties, both acidic and basic, allowing the passage of polar compounds of molecular weights of 600 Da or less. Three general porins are known in *E. coli*: OmpF, OmpC, and PhoE (11). All three show preferences of compounds based on charge and size. OmpF and OmpC have a preference for cations, with OmpF being able to accommodate slightly larger

substrates. PhoE, despite it being referred to as phosphoporin, has a general preference for all anions (11).

The prototype for specific porins is represented by protein LamB (or maltoporin), a protein involved in the import of maltodextrins and reported to be the receptor for phage λ (12). Similar to non-specfic porins, LamB is a homotrimer of β-barrels, but its crystal structure has unravelled structural determinants for specificity (13). LamB displays a slightly larger barrel than general porins, with 18 strands, and a very narrow constriction of only 5 Å in diameter. The hollow center of the barrel shows a continuous strip of aromatic residues, nicknamed the *greasy slide*, forming a π-stacked guiding path for polysaccharidic substrates from the extracellular to the periplasmic side of the porin. Similarily, the porin ScrY, specialized in the transport of sucrose and a wide range of other sugars including maltodextrins (11), adopts the same 18-stranded barrel fold, and assembles as trimers. While ScrY also displays a greasy slide, stacking interactions of aromatic residues with sucrose exclude fructofuranose rings and interacts strictly with glucopyranose moieties. ScrY possesses a seemingly disordered N-terminal domain that is absent from LamB (14).

FadL, a long chain fatty acid transporter, is the archetypal representative of a family of OM channels specialized in the transport of hydrophobic compounds. The high-resolution structure of this *E. coli* protein (15) shows an architecture distinct from that of OM porins. FadL is a monomeric 14-stranded β-barrel with a hatch domain occluding the center. It has been proposed that the fatty acid diffusion mechanism involves a conformational change where the hatch is displaced to allow passage of the

substrate (15,16). However, the presence of an unusual lateral opening in the FadL crystal structure with a protruding LDAO detergent molecule suggested a lateral diffusion model where the sole function of the protein is to get hydrophobic compounds through the LPS polar region and allow subsequent free diffusion in the OM bilayer. In adddition, the recently solved FadL structure of *Pseudomonas aeruginosa*, in conjunction with oleate uptake measurements in *E. coli* mutants of FadL with a constricted lateral opening, brought further evidence for this alternative mechanism (17).

1.2.2 Energy-dependent transport

Iron and vitamin B_{12} are essential nutrients involved in a number of metabolic processes such as electron transport and amino acid synthesis. They are considered scarce nutrients based on their low bioavailability. Vitamin B_{12} is synthesized only by a select number of microorganisms, while iron, in its non-toxic ferric form, is weakly soluble in oxic conditions and is thus found at attomolar concentrations in the environment (18). The context of infection mimics this situation; mammalian hosts sequester these two essential nutrients in protein-bound forms. Bacteria have thus evolved elaborate uptake systems to overcome this apparent paucity of critical organometallic nutrients (see Figure 3). As part of their strategy, bacteria and certain eukaryotic microorganisms secrete high-affinity iron chelators, called siderophores, into the environment and actively

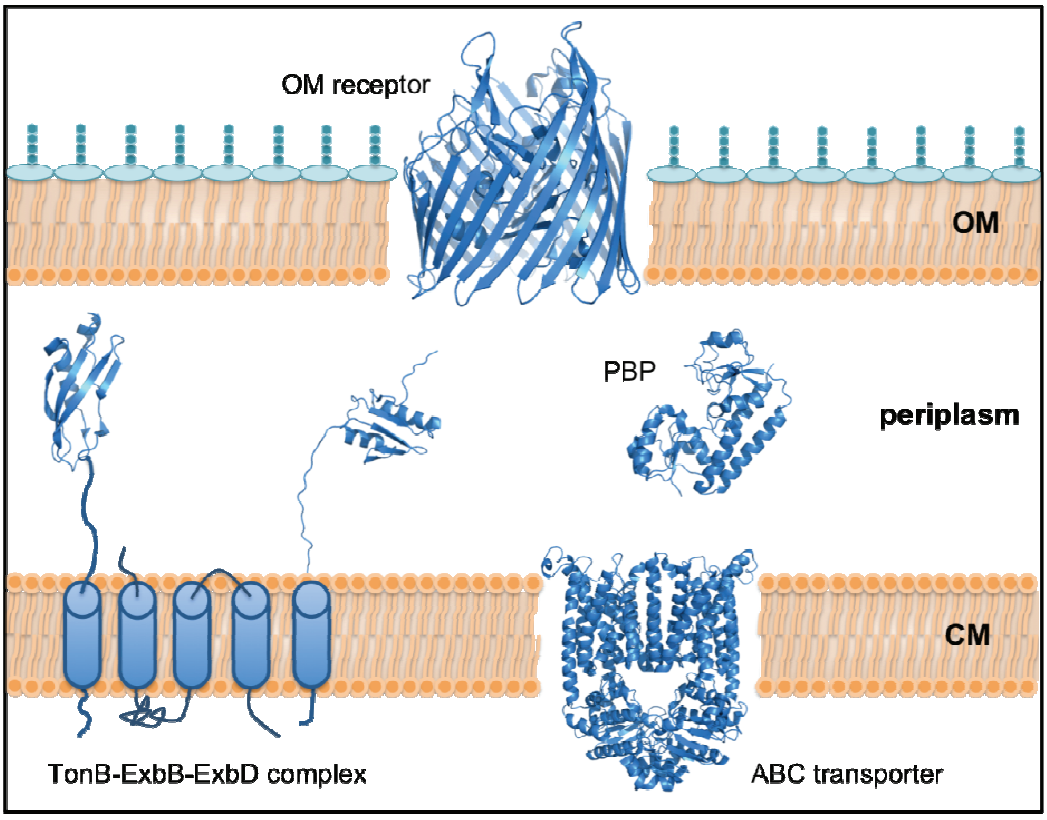

Figure 3. Components of a scarce nutrient uptake system. The import of scarce nutrients (e.g. vitamin B_{12}) in Gram-negative bacteria requires four main components: an OM receptor, the TonB–ExbB–ExbD complex, a PBP, and an ABC transporter. OM receptors (BtuB, PDB id 1NQE) are responsible for crossing the outer membrane, and are energized through direct contacts with the TonB–ExbB–ExbD complex (TonB PDB id 1XX3, ExbD PDB id 2PFU) which harnesses and transduces the proton motive force. After crossing the OM, nutrients are bound by a periplasmic binding protein (BtuF, PDB id 1N4D), and are shuttled to ATP-driven ABC transporters (BtuCD, PDB id 1L7V) to be uptaken into the cytoplasm. Cartoon cylinders and coils are used to represent helices and coils of unresolved structure, respectively.

uptake the iron-siderophore complexes. Heme and ferroproteins of mammalian hosts are other sources of iron exploited by pathogenic bacteria (19,20). Despite these strategies, the transport of iron and vitamin B_{12} across the bacterial cell envelope remains a thermodynamically unfavourable process. Indeed, because the concentration of scarce nutrients is typically higher inside cells than in their surrounding environment, the import of these nutrients is an uphill movement of molecules against a concentration gradient. Therefore, acquisition of iron and vitamin B_{12} by Gram-negative bacteria is an energy-dependant process mediated by specialized transport proteins termed OM receptors. Of note is the fact that many of these transporters are also exploited by bacteriophages and bacteriocins for their translocation inside of cells (21,22). Antibiotics are also known to target OM receptors (23)

Structures for ten OM receptors have been reported. Five of these structures are from *E. coli*: FhuA (23,24), FepA (25), FecA (26,27) BtuB (28), and Cir (29). Reported structures from other organisms are FpvA (30) and FptA (31) from *P. aeruginosa*, FauA (32) from *Bordetella pertussis*, HasR (33) from *Serratia marcescens* and ShuA (34) from *Shigella dysenteriae*. Like all other intrinsic OM proteins, OM receptors adopt the β-barrel fold. They are found as monomers and all share a common general architecture (Figure 4). Their structure can be divided in two large domains. A large barrel of 22 anti-parallel strands makes the channel portion of these proteins, while a cork domain occludes the center of the protein, preventing the free diffusion of molecules. Interestingly, while they mediate an energy-driven process, these receptors have no

source of energy readily available to them at the OM (35). Instead, energy is obtained from the proton motive force (PMF) of the cytoplasmic membrane and transduced by the TonB–ExbB–ExbD complex (Section 1.4). Another essential part of OM receptors is the Ton-box, a conserved 5 or 6 amino acid segment at their amino-terminus essential for interaction with the energy-transducing protein TonB (36).

The initial report of the FhuA structure in both its apo and holo forms (23) not only unravelled the general architecture of OM receptors, but also provided the paradigm for substrate binding and for the conformational changes that take place as a result of this interaction. In the holo structure, the hydroxamate siderophore ferrichrome is bound in a pocket surrounded by the extracellular apices of the protein. Binding site residues are provided by inter-strand loops that extend above the plane of the membrane bilayer. Other interacting positions are found on the extracellular face of the cork domain. Comparison of the ligand-free structure with the ferrichrome-FhuA complex shows very little variation in the β-barrel backbone structure. Changes are predominantly observed in the cork domain and, expectedly, in the binding pocket. Conformational changes upon ferrichrome binding are indicative of an induced-fit mechanism. On the periplasmic side of the cork domain, a major change in secondary structure is observed upon siderophore binding. A short hydrophobic helix termed the switch helix is comprised of residues 24 to 29, stabilized by a complementary pocket of the cork in the apo structure. The switch helix is completely destabilized in the holo form, and results in its unwinding. The unwound segment reorients 180^{o}

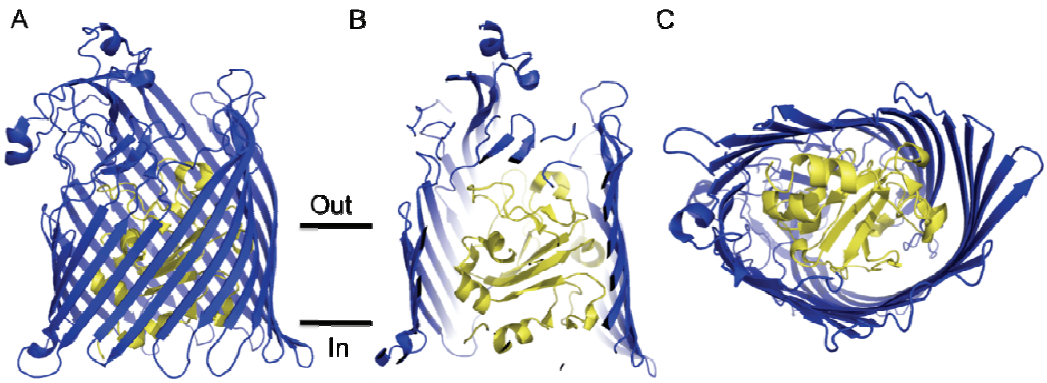

Figure 4. Overview of OM receptor structure. The structure of OM receptors is divided in two distinct domains : 22-stranded β-barrel domain (blue) and cork domain (yellow). Structure of the ferric hydroxamate receptor FhuA (PDB id 1QFG) is used as an example. Horizontal black bars indicate the approximate boundaries of the membrane, with In and Out indicating the periplasmic and extracellular sides, respectively. (A) View of the barrel in the plane of the membrane. (B) Cut-open view, in the plane of the membrane, revealing the structure of the cork domain. (C) View from the periplasmic side of the protein showing the occlusion of the hollow center of the barrel by the cork domain.

from its original axis, releasing the N-terminal Ton-box into the periplasm for interaction with TonB (23).

More recently, structures of FhuA (37) and BtuB (38), in complex with the C-terminal end of TonB, (Figure 5) have shed further light on the mechanism of energy-dependent OM transport. As in previous structures, complexation does not appear to significantly alter the overall fold and secondary structure of OM proteins. However, significant changes are observed at the N-terminal ends. As in holo-FhuA, the switch helix is unwound. In these structures, however, the Ton-box is ordered and visible: it adopts a β-strand secondary structure and participates in the formation of an inter-protein β-sheet with TonB. Loops of the periplasmic face of the cork domain also engage in interactions with TonB. Understanding the molecular interaction between OM receptors and TonB provides insight into import mechanism. The structures support a mechanism in which PMF-induced structural changes in TonB induce a pulling or a partial displacement of the cork that allows passage of the substrate (37,38).

Recently, other substrates for TonB-dependant OM transport that are different from the canonical vitamin B_{12} and iron-siderophore complexes have been identified in certain Gram-negative species. The first such example was identified in *Caulobacter crescentus*, which demonstrated transport of maltose and maltodextrins through MalA, a protein that shows more sequence similarity to TonB-dependant OM receptors than to the LamB maltoporin. It was demonstrated that MalA was strictly required for uptake of long maltodextrins (maltopentose and longer) and depends on TonB–ExbB–ExbD for its activity

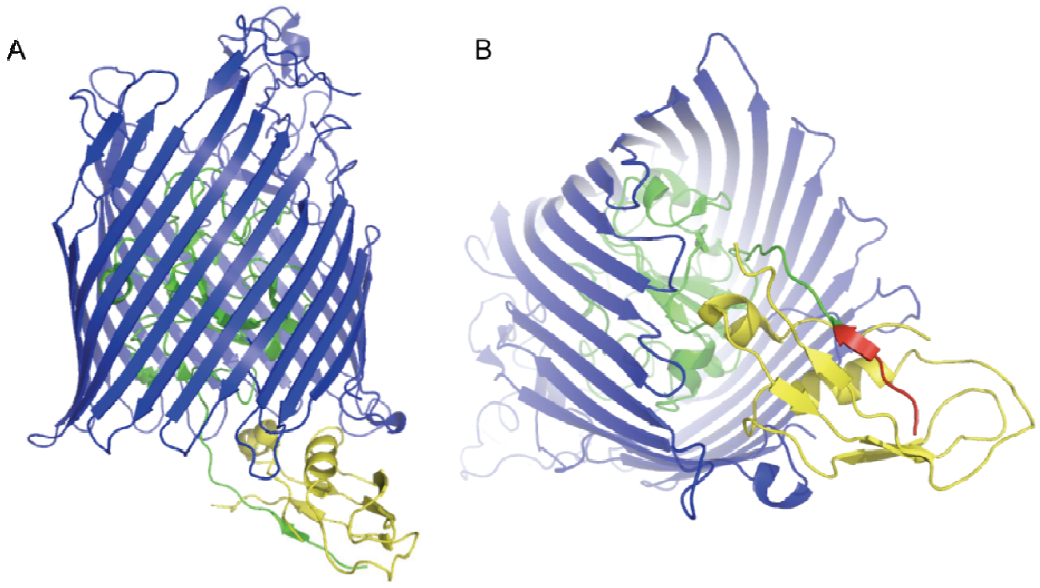

Figure 5. OM receptors interact with TonB to form a complex. Two views of the structure of FhuA in complex with TonB (PDB id 1GRX). The barrel domain of FhuA is in blue, the cork is in green and TonB is presented in yellow. (A) View of the complex in the plane of the membrane showing the cork (green segment) with the Ton-box extending into the periplasm for contact with TonB (yellow). (B) View from the periplasmic side revealing the TonB-FhuA site of interaction. The Ton-box is highlighted in red.

(39,40). The *Helicobacter pylori* OM receptor FrpB4, originally postulated to be an iron transporter, was demonstrated to be involved in nickel uptake, and was dependent upon the TonB complex for function (41). Studies in the phytopathogenic bacterium *Xanthomonas campestris* also suggest a role for one or several of its eight TonB homologues in sucrose acquisition (42). In addition, bioinformatic predictions suggest the existence of a wide number of TonB-dependant OM receptors throughout Gram-negative species (43). Together, these studies argue for a broader role for TonB-dependant uptake than originally expected.

1.3 Periplasmic binding proteins and influx ABC transporters

After they have crossed the OM, scarce nutrients are shuttled across the periplasm and then actively transported through the CM into the cytosol. These processes and the proteins that mediate them are discussed in this section.

Nutrients that have entered the periplasm either by diffusion or active transport are typically captured by specific high-affinity proteins termed periplasmic binding proteins (PBPs). At least 50 PBPs are known in Gram-negative bacteria. Lipoproteins that are homologous to PBPs are also found in the CM of Gram-positives (44). Over the last 26 years, crystal structures have been solved for more than a dozen PBPs. While they vary greatly in terms of size and sequence, PBPs show an overall structure conservation; they are all formed of two alpha-helical lobes separated by a groove in which the ligand is bound (44). PBPs have been classified into three groups based on the number of

polypeptide segments that connect their two lobes: group I has three interconnecting segments, group II has two, and group III has only one (44,45).

The function of PBPs is intimately linked to that of ATP Binding Cassette (ABC) transporters embedded in the CM. ABC transporters are one of the largest protein superfamilies known. In *E. coli* alone, 80 ABC transporters are described. As their name indicates, these transporters are nucleotide-binding proteins that use ATP hydrolysis to drive the movement of substrates across membranes, either in or out of cells. In Gram-negative bacteria, nutrients that reach the periplasm through OM-receptors are transported into the cytoplasm by specific ABC transporters. Vitamin B_{12} and iron-siderophore complexes have specific ABC transporters (46).

Structures for nine bacterial ABC transporters are available and all adopt a similar organization. ABC transporters have four domains: two nucleotide binding domains (NBDs) and two membrane-spanning domains (MSDs). Pairs of NBDs can be identical or different, leading to varying levels of symmetry in ABC transporters and to cooperative catalytic activity. These domains can also be separate polypeptide chains, or fused to form a single polypeptide. As their name indicates, the NBDs are responsible for binding, but also for hydrolysis of ATP. Hydrolysis of ATP by NBDs induces conformational changes in MSDs that allow release of substrates into the cytoplasm. The structure of the ABC transporter for vitamin B_{12}, formed by proteins BtuC and BtuD, was the first to be solved, both alone (47), and in complex with the PBP BtuF (48). It adopts a $BtuC_2D_2$ organization. BtuC forms a dimer of MSDs, each consisting of 10 transmembrane helices. Similar to NBDs, two MSDs can be identical or different, and

can exist as two polypeptides or as a single chain. At the interface between the two MSDs is the proposed B_{12} translocation pathway. BtuD, on the other hand, is the NBD and it also form dimers.

No structures exist for ABC transporters for siderophore uptake, such as FhuBC, although the NBDs from the ferric ion transporter FbpC from *Neisseria gonorrhea* were solved (49). Other structures of note include the maltose transporter of *E. coli* $MalFGK_2$, in complex with its cognate PBP (50), and the lipid transporter of *Salmonella typhimurium*, MsbA (51).

The role of PBPs is to bind their target ligand and to efficiently shuttle it across the periplasm. Interaction of PBPs with ABC transporters of the CM has been demonstrated *in vitro* (45). In addition, as mentioned in Section 1.4.1, recent studies (52,53) have shown *in vitro* interaction of periplasmic binding proteins FhuD and BtuF with TonB, suggesting a proposed ternary complex comprised of TonB, an OM receptor and a PBP. Together, these interaction studies with PBPs suggest a model of periplasmic shuttling of scarce nutrients where ligands are captured immediately after translocation through OM receptors and are later delivered to their specific ABC transporter at the CM.

1.4 The TonB–ExbB–ExbD complex

The PMF, a H^+ gradient actively maintained by cells across their CM (54), is used to drive many energy-dependant processes, such as nutrient import, ATP-synthesis and cell motility (54,55). Gram-negative bacteria have also evolved a strategy to exploit that

energy for the purpose of OM transport (56). Integral membrane proteins TonB, ExbB and ExbD, anchored in the CM, are proposed to form a complex that harnesses the PMF and transduces energy to OM receptors. Early pull-down experiments have probed the interactions that take place among the three components. His-tagged ExbB bound to Ni^{2+}-charged resin was shown to specifically retain both TonB and ExbD (57). Later, crosslinking studies have shown interactions between all components of the complex. These studies have also identified residues important for ExbD–TonB interactions. Indeed, aspartate 25 and alanine 92 in ExbD, and histidine 20 and alanine 150 in TonB have been shown to be critical for the *in vivo* crosslinking of those two proteins (58,59). Together, these proteins have been shown responsible for harnessing the PMF and transducing its energy to OM receptors (60).

The complex shares important sequence homology with two other PMF-dependant protein complexes of *E. coli*, TolAQR and MotAB. The TolAQR system is primarily involved in maintaining membrane stability and was reported to be exploited for colicin import and filamentous phage DNA insertion (61-63). Sequence homology between the Ton and Tol system has been confirmed by genetic studies establishing their functional redundancies. Complementation studies have demonstrated the ability of *tol* transformants to rescue transport phenotypes lost by *Δexb* mutations. Transport phenotypes of *Δtol* mutants were also shown to be complemented by tranformation of *exb* genes (64-66). The other system homologous to TonB–ExbB–ExbD is the MotAB stator complex of the bacterial flagellar motor (67,68). This homology has been used for modeling the transmembrane segments of TonB–ExbB–ExbD (67). Based on these

modeling outcomes and using sequence conservation patterns, putative proton pathways into TonB–ExbB–ExbD were proposed. This proposition is discussed in further detail in section 1.4.2.

1.4.1 TonB

TonB is the most studied and best understood component of the complex. It is a 26 kDa, bitopic membrane protein, 239 amino acids in length. The structure of TonB (Figures 6A and 7) can be divided into three domains. The first 33 amino acids encompass the protein's single transmembrane helical segment, placing the N-terminus in the cytoplasm (69), and the rest of the protein in the periplasm. Residues 33-155 form a long linker between the C-terminal domain and the transmembrane helix, which allows the protein to span the periplasm and contact OM receptors. Within that linker is a region between residues 66 and 100 which is known as the proline-rich region. The precise function of this highly entropic region is not understood, but its deletion does not impair TonB-dependant transport (70). Nevertheless, it appears to play a role in the formation of a complete TonB-FhuA complex (71).

Residues 155 to 239 form the C-terminal domain which has been identified as being involved in direct contacts with OM receptors. Extensive structural information for this domain shows the C-terminal portion of TonB with OM receptors FhuA and BtuB, revealing structural determinants critical to

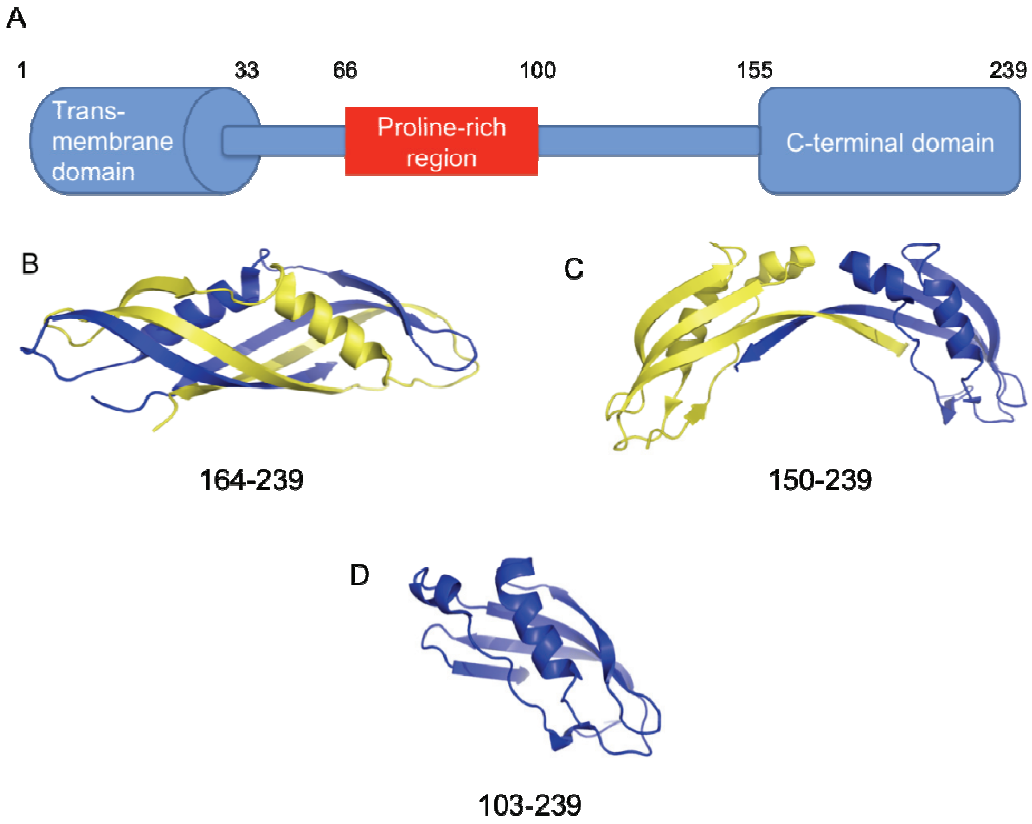

Figure 6. Overview of the structure of TonB. (A) Cartoon of domain structure of TonB highlighting its main domains. (B) Crystal structure covering residues 164 to 239 (PDB id 1QXX), showing an intimately intertwined TonB dimer. Monomers are coloured in blue for differentiation. (C) Longer dimeric structure of TonB covering residues 150-239 (PDB id 1U07), showing significant conformational changes compared to the shorter fragment in (B). (D) NMR solution structure of the 103-239 TonB fragment (PDB id 1XX3). Although monomeric, the NMR structure shows strong resemblance to (C), with the exception of a β-strand located at the dimer interface of (C) but folded into a β-hairpin in (D).

interaction (37,38). Prior to these reports, three structures for free periplasmic TonB had been solved (Figure 6BCD). The first crystal structure covered residues 164 to 239 (72) and revealed a dimeric organization that was confirmed by analytical ultracentrifugation measurements of TonB-FhuA complexes (73). The crystal structure of a longer construct was later reported (74). While the new construct showed a dimer interface inside the crystal, it adopted a markedly different fold and it stayed monomeric in solution. Shortly thereafter, the solution NMR structure of an even longer construct – covering residues 103-239 – was found to be monomeric (75). Furthermore, the ordered portion of the structure (residues 151-239) was in good agreement with the previous crystal structure (residues 164-239), with the exception of a folded β-hairpin that prevented dimer formation.

Recent phage display (PD) studies (52,53) identified FhuD-, BtuF-, and TonB-affinity selected peptides that suggest an interaction between TonB and periplasmic binding proteins (PBPs). Fluorescence, surface plasmon resonance (SPR) and dynamic light scattering (DLS) experiments further support that claim by providing direct physical evidence for these interactions. The implications of this finding for siderophore and vitamin transport were discussed in Section 1.3.

1.4.2 ExbB and ExbD, and the putative proton pathway

ExbB is a 26 kDa integral membrane protein composed of 244 amino acids. Its proposed topology (Figure 7) places its N-terminal end in the periplasm and its C-

terminus in the cytoplasm. It is proposed to have three transmembrane segments. Protection from proteolysis studies suggest that the residues involved in transmembrane helix formation are: 16-39 (helix I), 128-155 (helix II), and 162-194 (helix III) (76). Bioinformatic predictions closely match these results (77). Significant cytoplasmic domains are therefore predicted between helices I and II, as well as in the C-terminal portion that follows helix III. Structural information is not available to confirm those predictions.

The smallest protein in the complex, ExbD, has a molecular weight of 15.5 kDa and a length of 141 amino acids. It has a topology and overall structural organization similar to that of TonB. The amino-terminus is located in the cytoplasm and possesses a single transmembrane helix segment that covers residues 23 to 43, while the rest of the protein forms a large periplasmic domain (78). Bioinformatic predictions are in good agreement with this proposed topology (77). As for TonB, the structure for the periplasmic domain of ExbD was reported by NMR (Figures 7 and 8A) (79). The structure spans residues 44 to 141 and can be divided in three regions. Positions 44 to 63 are disordered and are proposed to form a flexible linker. They are followed by a well-folded segment spanning amino acids 64 to 133. The last segment, residues 133-141, is also observed as a disordered region. The central folded domain consists of a five-stranded β-sheet in a mixed parallel anti-parallel arrangement. Two α-helices flank the C- and N-terminal sides. The recently solved NMR structure of the TolR homologue of *Haemophilus influenzae*, albeit dimeric, shows the same overall secondary structure arrangement and the same fold (Figure 8B) (80). This fold also shows striking similarities

with the C-terminal lobes of the periplasmic binding proteins FhuD and CeuE (Figure 8C) (79). These observations raise interesting questions in regard of the observed *in vitro* interactions between FhuD and TonB (52), and considering the fact that proteins ExbD and TonB are found together as a complex. Monitoring of chemical shift perturbation by HSQC upon binding of C-terminal ExbD to C-terminal TonB or synthetic peptides of the proline-rich region suggest a weak interaction (79).

A series of mutagenesis studies have identified residues that are important for the activity of ExbB and ExbD. L132 replaced by Q in the periplasmic domain of ExbD, failed to restore the Ton-activity of *⊿exbD* mutants *of E. coli*. The sole charged residue in the membrane-spanning helix of ExbD, R25, displays the same phenotype when mutated to asparagine (57). An equivalent mutation in TolQ has a parallel effect (68). In ExbB, mutations V35E, V36D and A39E do not abolish TonB-dependant activities, but were shown to have a suppressive effect on TonB mutations H20Y, S16L and ΔV17, suggesting a role for these areas in TonB–ExbB interactions (81,82).

More recently, a molecular modeling study of the TonB–ExbB–ExbD complex used the transmembrane helix bundles of MotA-MotB as templates, and proposed two putative proton pathways within the complex. The first pathway involves helices II and III of ExbB and the sole transmembrane helix of ExbD. The pathway also relies on potential water molecules along the path to act as proton acceptors. The order of residues involved in the path are as follows: T181 –> S155 –> H_2O molecule adjacent to G151 –> T148 –> H_2O molecule adjacent to G144 –> D25 (ExbD). The alternative pathway that was proposed is similar to

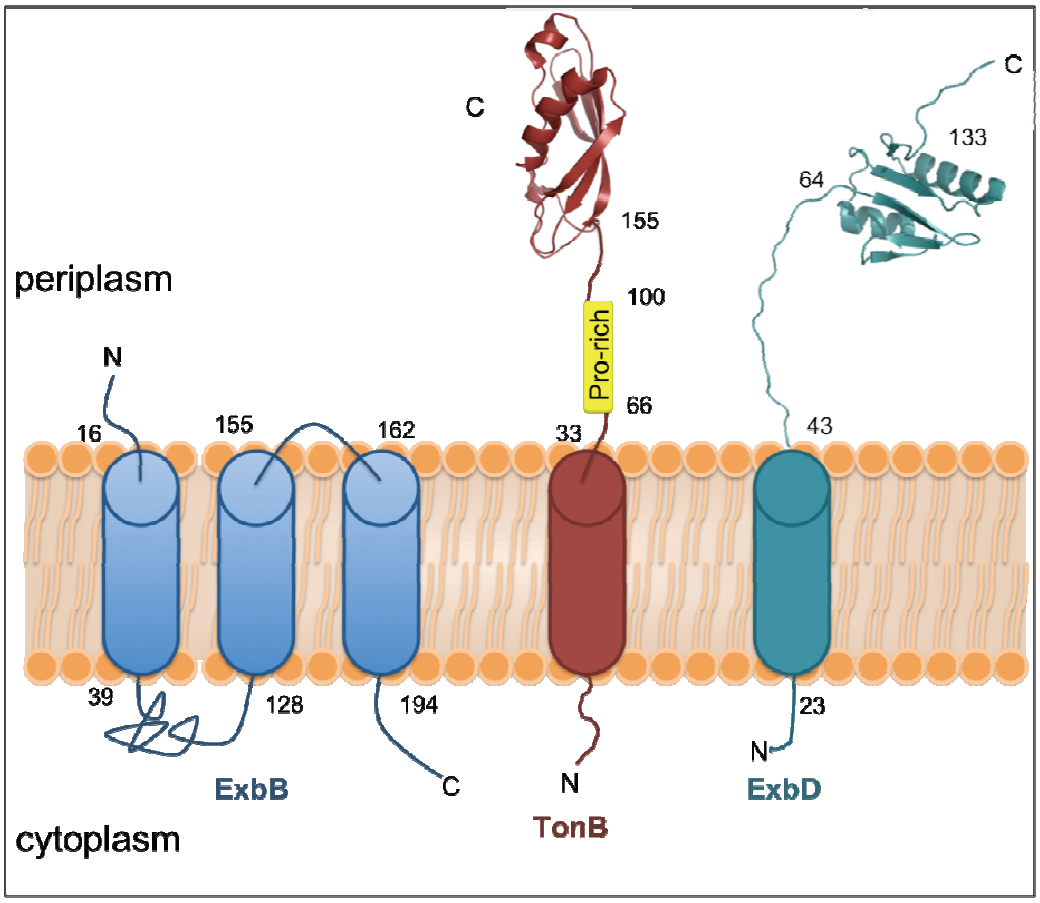

Figure 7. Structures and proposed topologies of components of the energy-transducing complex. The structure of ExbB (blue) is unknown. ExbB is proposed to have three transmembrane helices, with the N-terminal end in the periplasm and C-terminus in the cytoplasm. TonB (red) has a single transmembrane helix, with its amino-terminus in the cytoplasm. Structure of its C-terminal periplasmic domain is known (PDB id 1XX3). A 122 amino acid-long linker, containing the disordered proline-rich region, is found between the transmembrane helix and the resolved C-terminal domain. The proposed topology for ExbD (turquoise) is similar to that of TonB. Its periplasmic domain structure was also solved (PDB id 2PFU). Cartoon cylinders represent helices of unresolved structure. Numbers indicate sequence boundaries between depicted structural elements.

the first but also involves the transmembrane helix of TonB. The detailed second path is T181 (helix III of ExbB) –> H_2O molecule (adjacent to G184 in helix III of ExbB) –> H_20 in TonB –> S16 in TonB –> H_2O (adjacent to A188 in helix III of ExbB) –> T148 in helix II of ExbB –> H_2O (adjacent to G144 in helix 2 of ExbB) –> D25 in ExbD (67). A later mutagenesis study showed that threonines 148 and 181, when both mutated to alanine, abolish all ExbB activity. These two positions are conserved in the Exb/Tol/Mot systems and are part of both proposed proton pathways. The same study also identified glutamate 176, the sole charged residue in helix III of ExbB, to be critical for the protein's function (83).

1.4.3 Stoichiometry and mechanism of action

Genetic and structural studies have revealed many aspects of TonB–ExbB–ExbD function, but a complete and detailed understanding of its mechanism of action is still lacking. Two reasons account for our limited knowledge of the mechanism of the complex. One of these reasons is that the actual stoichiometry of the complex remains unknown. To establish the number of copies of each protein in the complex, Higgs *et al*. applied an immunoblot strategy, using purified ^{35}S-labelled proteins, to quantitate the abundance of each component of the complex in *E. coli* cells. It was concluded that the TonB:ExbB:ExbD ratio in *E. coli* was 1:7:2 (84). While this report reveals the relative abundance of each component of the complex in *E. coli*, it is not a definitive measurement of its actual stoichiometry.

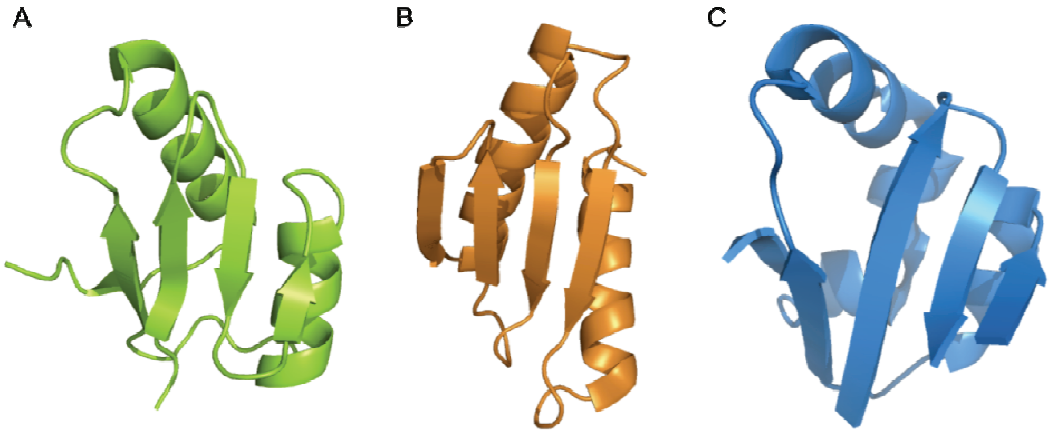

Figure 8. The core fold of periplasmic ExbD (green) shows structural homology to TolR (orange) and FhuD (blue). The 4-stranded, mixed parallel anti-parallel β-sheet of the periplasmic domain of ExbD (PDB id 2PFU), flanked by two alpha helices (A) is highly similar to the *H. influenzae* TolR structure (PDB id 2JWK) (B). It is also reminiscent of the C-terminal lobe of periplasmic binding protein FhuD (PDB id 1EFD) (C) and CeuE (not shown).

The second reason why the mechanism of action of the TonB–ExbB–ExbD complex remains elusive is that structural information on the complex is very fragmented. All reported structures concern periplasmic portions and have been solved for isolated components. Therefore, nothing is known about the structure of the transmembrane segments and the spatial arrangement they adopt. The biggest gap in our structural knowledge of the complex is the absence of any report on the structure of ExbB, the component proposed to be the most abundant within the complex.

Despite our incomplete knowledge of TonB–ExbB–ExbD structure and function, tentative mechanisms have been proposed. A first proposal relied on homology with MotAB and on the mechanisms of other proton-driven molecular machines of the CM, proposing a rotary mode of action (72). TonB-induced conformational changes of the cork-domain of OM receptors, believed to be generated by a rotary motion, form the basis of the model. The authors of that proposal acknowledged the inherently speculative nature of their proposition, and did not provide a detailed and exhaustive sequential model. However, the model did identify association and dissociation of TonB with OM receptors as key events, but did not determine which event requires energy.

A highly speculative and eminently more elaborate model was later proposed by Postle *et al*. In this proposition, the PMF, harnessed by ExbB and ExbD, is used to conformationally "energize" TonB. Energized TonB then extracts itself from the CM to contact nutrient-bound OM receptors and releases its energy by another conformational change, allowing the passage of the bound nutrient. TonB, now reverted

to its ground state, returns to the CM and re-associates with ExbB–ExbD (85). This model is questionable in many aspects. It underestimates the amount of energy required to stably extract a transmembrane helix from a lipid-bilayer. More importantly, it fails to explain how TonB is to retain its activated conformation once it has dissociated from ExbB–ExbD to freely shuttle through the periplasm. It does not explain what force prevents the activated and thus unstable TonB conformation from spontaneously dropping to the ground state if it is freely diffusing.

1.5 Structural biology of membrane proteins

Nearly 30% of the proteins encoded by any genome, including that of *E. coli*, are membrane proteins (86). Despite that fact, only 245 unique structures (87) out of the 65 000 found in the protein data bank are membrane proteins (as of May 29th 2010). This discrepancy is explained by the technical challenges involved in determining membrane protein structures. Indeed, membrane proteins present special difficulties in terms of their overexpression, solubilization, and crystallization. In this section, I aim to introduce membrane protein structural biology as well as some of the techniques and strategies that are specific to that field.

1.5.1 Membrane protein over-expression

The photosynthetic reaction center of the cyanobacterium *Rhodopseudomonas viridis* was the first membrane protein structure to be resolved, and was isolated from natural sources (88). However, for most proteins, low endogenous levels of expression

preclude such a strategy. Over-expression is therefore necessary. In addition, over-expressed recombinant proteins can be manipulated to generate mutants and variants of interest. The expression system of choice remains *E. coli* which can be easily grown to high cell densities at low cost, allowing the production of the milligram amounts of protein required for crystallographic studies. However, in heterologous expression, bacterial hosts fail to properly provide the post-translational modifications associated with many eukaryotic proteins. In addition, the membrane environment of an *E. coli* expression host differs from that of other organisms. Therefore, other systems have been explored which include the yeasts *Pischia pastoris* and *Saccharomyces cerevisiae*, Sf9 insect cells and human embryonic kidney (HEK) cell lines (89).

Over-expression of bacterial membrane proteins faces two specific bottlenecks. The first concerns the protein folding capacity of the expression system. *E. coli* membrane chaperones have demonstrated their importance for proper folding of membrane proteins (90): overwhelming their folding capacity during overexpression is a well-known concern (91). Membrane space to accommodate overexpressed proteins is the other specific factor in membrane protein expression. Indeed, membrane proliferation is known to promote the expression of membrane proteins (92). *E. coli* strains have been specifically designed and used to circumvent those problems (91). In addition, expressing protein entails the screening of media, induction time and temperature, and inducer concentration.

1.5.2 Detergents for structural biology of membrane proteins

Isolation of membrane proteins for structural studies and other applications requires extraction from the lipidic environment of the membrane and stabilization in solution by the use of detergents. Detergents, like membrane lipids, are amphipathic molecules composed of a polar or charged headgroup and of a hydrophobic tail of a certain length. Detergents, however, have a specific phase behaviour: at low concentrations, detergent molecules are found as monomers. When detergent concentration reaches a specific threshold known as the critical micellar concentration (CMC), monomers begin to self-associate to form structures known as micelles. The CMC is a characteristic property of each detergent. Above the CMC value, monomer concentration remains constant while the number of micelles increases with detergent concentration (Figure 9A). Depending on temperature, solution composition and other physical factors, detergents can form different phases that can be both beneficial or detrimental to membrane protein solubility, stability, and ability to crystallize. The ability of detergents to form micelles, however, is the central property exploited by membrane protein scientists. A micelle is a dynamic association of detergent molecules that exposes hydrophilic moieties to the solvent and packs hydrophobic tails away from it. The apolar interior of micelles, therefore, is the environment used to maintain stability of solubilized membrane proteins in an aqueous environment (93).

Identification of the best detergent for solubilization of a membrane protein is typically done by screening several classes of detergents and measuring the amount of target protein solubilized. Once in solution, the membrane protein is found as a protein

detergent complex (PDC), which may contain lipids in addition to protein and detergent (89). Consequently, the specific properties of PDCs will govern the membrane protein crystallization process. Four families of detergents dominate in their contribution towards crystallized membrane proteins. In order of importance, they are: the maltosides, the glucosides, the polyoxyethylene glycols, and the amine oxides (94).

1.5.3 Crystallography of membrane proteins

Crystallization of any protein exploits its solubility phase behaviour to generate out-of-solution and ordered proteinaceous structures represented as crystals. The phase diagram of protein solubility (Figure 9B) identifies four zones. The undersaturated zone corresponds to the conditions in which a protein is completely stable in solution, while the precipitation zone corresponds to conditions in which proteins come out of solution as an amorphous precipitate. Between these two zones are the nucleation and metastable zones. To obtain a crystal, a protein must reach the nucleation zone where nuclei form and then migrate to the metastable zone to allow crystal growth. Gaining control of the crystallization process of membrane proteins entails identifying and optimizing conditions that will yield crystals (95), in addition to accommodating specific properties of PDCs.

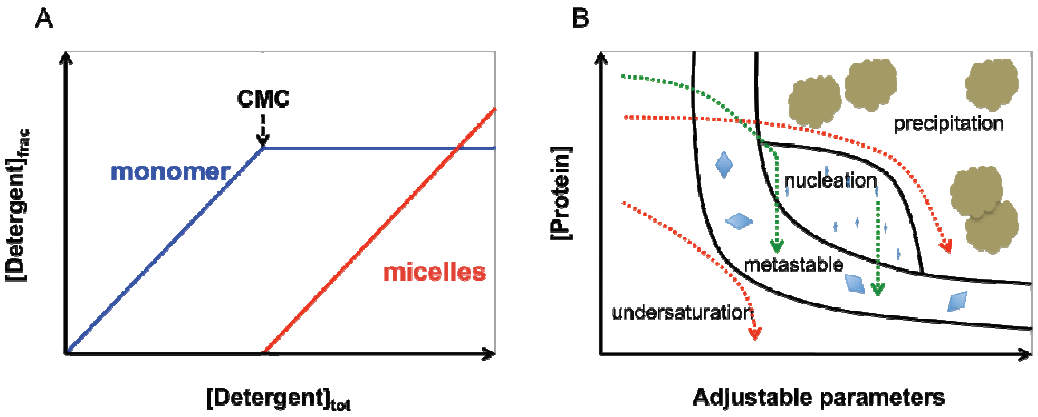

Figure 9. Detergents and proteins have unique phase behaviors that are exploited for solubilization and crystallization. (A) This panel illustrates the concept of critical micellar concentration (CMC). Below the CMC, the fraction of detergent monomer concentration ($[Detergent]_{frac}$) increases linearly as a direct function of total detergent concentration. At the CMC, detergent molecules begin assembling into micelles. As concentration increases above CMC, the monomer concentration remains stable while micelle concentration increases. (B) Phase diagram of protein crystallization contains four zones: the undersaturation, metastable, nucleation and precipitation zones. Crystal nuclei form in the nucleation zone, while crystal growth takes place in the metastable zone. The green dashed arrows are examples of condition paths that result in crystallization, while red dashed arrows indicate condition paths that will not yield crystals.

Membrane proteins are inherently hydrophobic. They often display few polar areas on their surface, which are usually segregated to intra- and extracellular apical portions. Therefore, membrane proteins have fewer regions available for crystal contacts than soluble proteins. This property makes it more difficult to obtain crystals of membrane proteins, but also hampers the formation of well-ordered, quality crystal packings that are necessary to obtain high-resolution X-ray diffraction data. Another obstacle that may prevent crystallization of membrane proteins is their propensity to aggregate. Once extracted from their native membrane environment, the extensive apolar surfaces of these proteins become available for non-specific hydrophobic interactions with other proteins, even when those surfaces are found within the micellar environment. The disordered and non-specific interactions that lead to aggregation are detrimental to crystallization. Assessing the monodispersity of PDCs by size-exclusion chromatography (SEC) or dynamic light scattering (DLS) is therefore useful for successful crystallization. The intrinsic properties of the protein and the choice of detergent are the main factors affecting aggregation.

Several strategies exist to tackle the specific challenges of membrane protein crystallization. The selection of a detergent suited for crystal growth is one of the most important steps. A detergent efficient at solubilizing a membrane protein might not be optimal for crystallization, and may require detergent exchange. Monitoring of detergent concentration in protein samples can be used to make sure that excessive detergent will not prevent crystallization. The optimal detergent should prevent, and

not promote aggregation. It must also adequately stabilize the membrane protein in the aqueous environment without covering surface regions required for the formation of crystal contacts. One strategy to promote crystal contacts involves adding a soluble element to the PDC, either through binding of a Fab fragment or by generation of recombinant fusion proteins that include a solvent exposed domain (e.g. lysozyme or GFP) (96,97).

Another approach to crystallizing membrane proteins consists of partitioning protein into a lipidic environment that is closer to the natural membrane protein environment, thereby facilitating close interactions and crystal growth. The use of lipidic cubic phases (or the related *sponge phases*) has been successfully applied to the crystallization of many membrane proteins, notably several G-protein-coupled receptors (98). A related method makes use of mixed lipid-detergent bilayer discs known as bicelles. Bicelles allow ease of use and flexibility of micelle-solubilized membrane proteins with the potential advantages of lipidic environments. So far, one novel structure, the mitochondrial voltage-gated channel VDAC1, was obtained from crystals grown in bicelles (98,99).

1.5.4 Electron-microscopy of membrane proteins

Electron microscopy (EM) has been used for the observation of cells and cellular structures for several decades. More recent technological developments have enabled the use of EM as a method to determine molecular structures. The small wavelength of

electron beams allows, at least in theory, the generation of molecular models at resolutions comparable to those reached by crystallographic means (100). However, technical limitations of EM results in reconstructions with resolutions between 10 Å and 20 Å. Nevertheless, EM has several advantages over crystallography that make it an interesting alternative. The generation of 2D crystals has been reported to improve EM outcomes, however crystallization is in no way a prerequisite to this technique. EM is therefore not limited by the possibility to crystallize a protein. It also does not require the levels of sample purity and homogeneity imposed by crystallography.

The central limitation faced by EM pertains to radiation damage. Generation of high-contrast, high-resolution images requires the use of electron beam intensities that may cause significant damage to biological samples. Low-dose beams can be used on biological samples, but at the cost of a poor signal-to-noise ratio. The classical means of overcoming this limitation is to fix and coat samples with a heavy metal stain that effectively shields samples from radiation damage, a method termed negative-staining. The downside of this technique is that it limits observation to areas of molecules that are accessible to the stain. The very nature of heavy-metal stains also denotes that resolution is limited to 20 Å, precluding the generation of high-resolution models. A strategy used to tackle the signal-to-noise limitation at low-dose is to average multiple particles, typically several thousands, to diminish noise (100). Another strategy is to enable the use of higher dose electron beams, for example by dropping the temperature of observed samples. Cryo-EM is a technique that freezes samples at temperatures of -150^{o}C or less, reducing radiation damage of about an order of

magnitude (101). The combined use of averaging and cryogenic temperatures has produced the highest resolutions of recent years.

1.6 Introduction to conducted research

Gram-negative bacteria are a leading cause of hospital-acquired infections, accounting for 70% of infections acquired in intensive care units in the United States. The *Enterobacteriaceae* family contributes to the majority of these infections (102). *E. coli* is a member of this family and is the leading cause of urinary tract infections worldwide. A major challenge facing the treatment of these infections is the increasing prevalence of antibiotic-resistant bacteria. For example, between 50% and 60% of *E. coli* isolates are resistant to β-lactams. Nosocomial pathogens of the Gram-negative genera *Klebsiella*, *Enterobacter* and *Serratia* are also reported as resistant to β-lactams, as well as to first generation cephalosporins (103). Early in the development of the first antibiotics, resistance was recognized as inevitable, but the massive use and misuse of antibiotics over the past decades have exerted a strong evolutionary pressure towards the selection of resistant strains (104). In parallel, recent years have seen a decline in the marketing of novel antimicrobial drugs due to concerns over their practical longevity, increasing costs and technical challenges that antimicrobial compound discovery represents (102). Consequently, the need for new antibiotics is pressing, and requires the exploitation of new bacterial targets.

Understanding processes that are critical for bacterial cell function is required to find novel ways to cure bacterial infections. Mechanisms of import of essential nutrients are potential targets. OM protein FhuA is already exploited as a receptor for antimicrobials such as albomycin and microcin peptides. The TonB–ExbB–ExbD complex, vital for the uptake of many essential nutrients, also appears as a prospective antibiotic target. In particular, structural information on constituent proteins within the complex and resolution of their stoichiometry would provide detailed insight into their molecular mechanism of action. More far reaching is the goal that structures of TonB, ExbB and ExbD as a complex be used for the structure-based drug design of antimicrobials specifically targeting Gram-negative bacteria.

To gain insight into the molecular structure of the TonB–ExbB–ExbD complex, we designed an over-expression and purification strategy for proteins ExbB and ExbD as a complex. ExbB–ExbD was successfully over-expressed and purified in milligram amounts. Investigation of the complex's behaviour in solution also identified conditions for optimal solubility, resulting in sample concentrations of more than 10 mg/ml. We adopted a nuclear magnetic resonance approach for the quantitation of detergent concentration in samples. Concentrated ExbB–ExbD samples were then used to evaluate the monodispersity of the complex by SEC and to perform crystallization trials. Several crystallization approaches were applied, generating promising leads that require optimization to yield well-diffracting crystals. Detergent-exchange was attempted to try to produce better diffracting crystals. In parallel, ExbB–ExbD samples were prepared for negative-staining EM. Grids were observed under EM revealing discrete and uniform

particles. High-resolution micrographs were collected and several thousands of particles were selected. Averaging of the selected particles allowed the generation of 2D reconstructions that suggest a pentameric organization for the complex.

Chapter 2 –Material and Methods

2.1 Bacterial strains and plasmids

ExbB–ExbD complex was expressed in the *Escherichia coli* M15 strain harbouring a pREP4 plasmid (Qiagen) which is a kanamycin resistance vector that constitutively expresses the lactose operator (*lacO*) for tighter control of *lac*-inducible constructs. The ExbB–ExbD complex was expressed from plasmid pExbBD (Figure 10), derived from the commercially available (Qiagen) pQE60 vector which carries an ampicillin resistance gene as a selection marker. Overexpression of recombinant proteins is under the control of the T5 promoter. The *exb* operon of *E. coli* was cloned into the NcoI/BamHI of the multiple cloning site (MCS) of pQE60, resulting in co-expression of wildtype ExbB in conjunction with ExbD with a hexahistidine tag appended at the C-terminal end. The construct also contains a Tobacco Etch Virus (TEV) protease site for cleavage of the His-tag from ExbD.

2.2 Protein expression

E.coli M15 cells harbouring pREP4 and pExbBD plasmids were grown in Terrific Broth (TB) supplemented with 10% (v/v) glycerol and 50 μg/ml ampicillin and 30 μg/ml kanamycin. An overnight pre-culture (1.5 ml) was used to inoculate 50 ml of TB. Cells were grown at 37^{o}C in a shaking waterbath incubator until they reached an A_{600} of 1.0 to 1.2. Overexpression of ExbB–ExbD was induced by addition of 0.5 mM isopropyl-β-D-1-thiogalactoside (IPTG). The culture was left

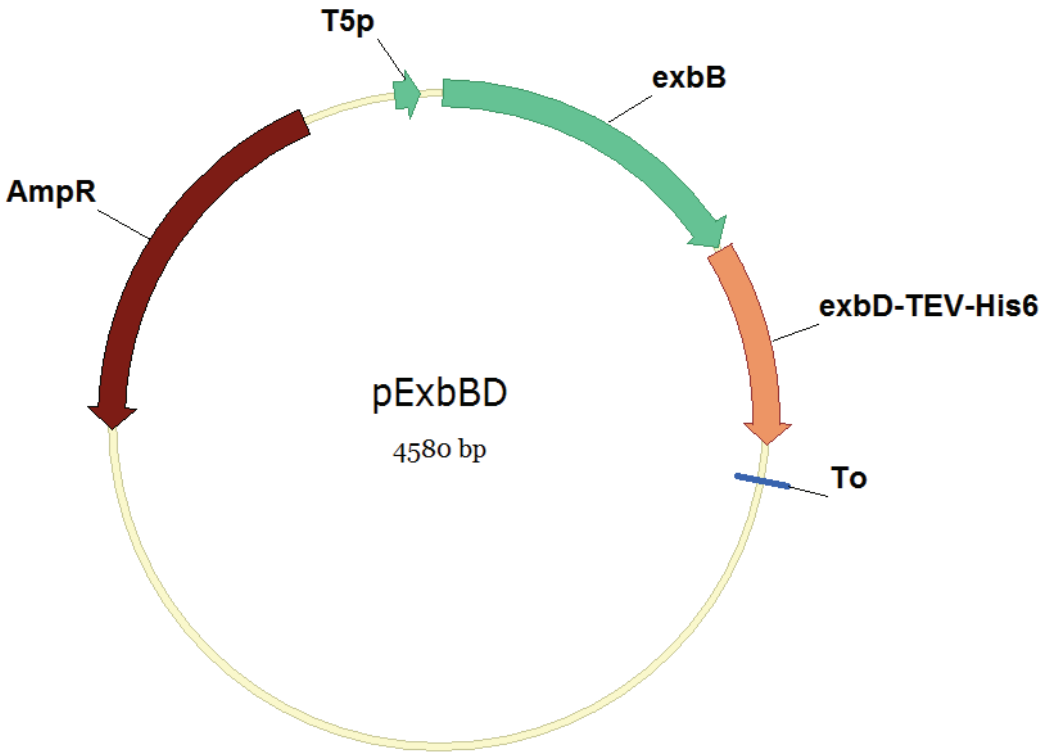

Figure 10. pExbBD plasmid. This construct was derived from the Qiagen pQE60 vector background. The *exb* operon is under the control of the T5 promoter (T5p) and attaches a TEV-cleavable C-terminal hexahistidine tag to ExbD. Transcription termination is directed by the lambda terminator (To). Induction by IPTG allows for coexpression of ExbB and ExbD-TEV-His_6. pExbBD also carries an ampicillin resistance gene for selection.

shaking at room temperature for 4 hours. The cells were then collected by centrifugation at 11 953 x g for 10 min in a Sorvall SS-34 rotor. Harvested cells were frozen at -20^{o}C or immediately lysed for expression tests (Section 2.3). Large scale expression was performed in the same conditions, but with a few differences. Overnight starter cultures of 25 ml were used to inoculate 1 L cultures in Fernbach flasks. Cells were harvested at 7268 x g for 15 minutes in a Sorvall H-6000 hanging bucket rotor.

2.3 Expression test

Cells harvested as described in Section 2.2 were resuspended in lysis buffer (50 mM Tris pH 8.0, 200 mM NaCl) supplemented with 5 mM $MgCl_2$, 100 µg/ml DNAse I, 100 µg/ml RNAse A, 10 µg/ml lysozyme and Complete protease inhibitor cocktail (Roche). The resulting cell suspension was then transferred to 1.5 ml microcentrifuge tubes and lysed by four 30 second rounds of sonication with a Sonic Dismembrator (Fisher model 300) at 50% power. Unlysed cells and large debris were removed by centrifugation at 13 000 rpm in an Eppendorf tabletop microcentrifuge. Total membranes were separated from soluble material by a one hour ultracentrifugation at 150 920 x g in a Beckmann MLA-130 rotor (Beckmann Optima Max tabletop ultracentrifuge) and resolubilized in 1% sodium dodecyl sulphate (SDS). Solubilized membranes and soluble material were analyzed by SDS-PAGE and analyzed by silver staining and Western blotting (Section 2.4).

2.4 Western blotting

Proteins separated by SDS-PAGE were transferred to an Immobilon-P Polyvinylidene fluoride (PVDF, Millipore) membrane for 20 minutes at 15 V with a Bio-Rad TransBlot SD semi-dry transfer apparatus. After transfer, the membrane was blocked for 1 h with 3% BSA in Tris-buffered saline (TBS, 50 mM Tris-HCl, 150 mM NaCl, pH 8.0). The membrane was rinsed quickly with TBS supplemented with 0.05% Tween20 (TBST) and incubated for 1 h with mouse anti-6xHis IgG1 monoclonal antibodies (Clontech) diluted 1/5000 in TBS. The membrane was then washed three time with TBST (two washes of 10 minutes and one of 5 minutes). Following washing, the membrane was incubated for 1 h with goat anti-mouse IgG alkaline-phosphatase conjugated antibody (Jackson ImmunoResearch) diluted 1/10 000 in TBS. Three additionnal washes were performed following incubation, as described above. The membrane was developed by incubation with Nitro blue tetrazolium chloride/5-Bromo-4-chloro-3-indolyl phosphate (NBT/BCIP) solution (Roche) diluted in developing buffer (100 mM $NaHCO_3$, 1 mM $MgCl_2$) for 10 min: the reaction was stopped with water.

2.5 Protein purification and TEV proteolysis

M15 pREP4 cells overexpressing recombinant ExbB–ExbD were resuspended in conditions identified in Section 3. Resuspended cells were lysed by two passages through an Emulsiflex homogenizer (Avestin). Unbroken cells and large debris were separated from the resulting lysate by low speed centrifugation (4300 x g in a Sorvall SS-34 rotor) for 10 minutes. Total membranes were collected by ultracentrifugation for 1

hour at 257 320 x g in a Beckmann Type 70.1 Ti rotor. Membranes were solubilized overnight at 4^{o}C in buffer containing 50 mM Tris pH 8.0, 200 mM NaCl and 1% (w/v) of detergent n-dodecyl-β-D-maltopyranoside (DDM) supplemented with protease inhibitor cocktail. Following overnight solubilization, total membrane sample was centrifuged for 45 minutes at 34 500 x g in a Sorvall SS-34 rotor to remove unsolubilized material prior to the first chromatographic step; supernatant containing DDM-solubilized proteins was retained.

The ExbB–ExbD complex present within the total membrane sample was captured by immobilized metal-ion chromatography (IMAC). DDM-solubilized proteins were applied to a Proffinity IMAC (Bio-Rad resin) column pre-equilibrated with equilibration buffer (50 mM Tris pH 8.0, 200 mM NaCl, 0.02% (w/v) DDM). Following sample loading, the column was washed with 10 column volumes (CV) of wash buffer (50 mM Tris pH 8.0, 200 mM NaCl, 35 mM imidazole, 0.02% (w/v) DDM). Bound material was eluted from the resin with elution buffer (25 mM Tris pH 8.0, 100 mM NaCl, 300 mM imidazole, and 0.02% (w/v) DDM).

Sample eluted from the IMAC resin was then subjected to cleavage by TEV protease to remove the C-terminal His-tag of ExbD. TEV protease was incubated with ExbB–ExbD complex in a 1:10 mass ratio for 72 hours at 4°C. Buffer used for cleavage consisted of 25 mM Tris pH 8.0, 100 mM NaCl, 0.02% (w/v) DDM and 10% (v/v) glycerol.

To further purify ExbB–ExbD preparations and to remove TEV protease, the reaction mixture was applied to an anion exchange chromatography (AEC) column. This step also allowed for buffer and detergent exchange to be performed (Sections 2.9 and

2.13 respectively). Poros HQ20 (Applied Biosystems) is the AEC chromatography medium. Sample was applied to the column pre-equilibrated with equilibration buffer containing 25 mM Tris pH 8.0, 100 mM NaCl, and 0.02% (w/v) DDM. Following sample loading, the column was washed with 5 CV of equilibration buffer, followed by 5 CV of 25 mM glycine pH 9.0, 100 mM NaCl, and 2 X CMC of the desired detergent. In the same buffer and detergent conditions, a gradient from 100 mM to 550 mM NaCl was formed over 20 CV, ending with a 5 CV step at 550 mM NaCl. After AEC, the protein was concentrated with a YM100 centrifugal filter concentrator (Millipore) according to manufacturer specification and applied to analytical size-exclusion chromatography (Section 2.6) or used in setting up crystallization trials (Section 2.10).

2.6 Size-exclusion chromatography

Size-exclusion chromatography (SEC) was performed on protein samples to assess their homogeneity. A 30 ml (10mm x 300mm ; Tricorn, GE) Toyopearl HW-55S (Tosoh) column was pre-equilibrated with 2 CVs of equilibration buffer containing 25 mM glycine pH 9.0, 100 mM NaCl, and 0.02% DDM. The sample, of a volume of 500 μl or less, was then injected onto the column. Proteins were separated by running equilibration buffer at a constant flow rate of 0.5 ml/min until all applied protein was eluted.

2.7 Negative-staining electron microscopy

Purified sample from IMAC was dialyzed extensively overnight through a 25 kDa nominal molecular weight cut-off (NMWCO) membrane in 100 μL dialysis buttons against dialysis buffer containing 25 mM Tris pH 8.0, 100 mM NaCl, and 0.02% DDM to remove imidazole. The dialyzed sample was then diluted to a protein concentration of 6-7 μg/ml with dialysis buffer. Five microlitres of the diluted sample containing purified ExbB–ExbD was applied to glow-charged carbon-coated electron microscopy (EM) copper grids and incubated for approximately one minute. The sample was blotted and rapidly replaced by 5 μl of 2% (w/v) uranyl acetate for staining and fixation. The grid was incubated with uranium salt for approximately 1 minute, then blotted to dryness. Grids were then ready for observation and could be conserved for extended periods of time. Micrographs were obtained with either a Tecnai T12 (135 000x to 165 000x magnification) or Tecnai F20 (50 000x magnification).

2.8 2D reconstruction of ExbB–ExbD particles

Particles were selected manually from unbinned micrographs using the *boxer* utility of the EMAN package (105). Initially, 2D reconstructions were built from 2000 particles by the maximum-likelihood (ML) multi-reference alignment method using the *xmipp* package (106). Initial reconstructions were used in *Signature* as search models for the automatic selection of an additional 45 321 particles (107). These reconstructions were then used through a second round of ML multi-reference alignment, generating reconstructions of higher definition.

2.9 Optimization of solubility

The effect of glycerol supplementation on ExbB–ExbD solubility was tested. Samples kept in 25 mM Tris pH 8.0, 100 mM NaCl, 10% glycerol, 0.02% DDM were concentrated using centrifugal concentrator units with 100 kDa NMWCO (YM100, Millipore) to the lowest possible volume and protein concentration was assessed using the bicinchronic acid (BCA) protein assay (Pierce).

The effect of buffer and pH on solubility was tested by dialyzing the sample against several buffer and pH conditions. Sample was dialyzed through a 25 kDa (NMWCO) membrane in 20 μl dialysis buttons against 10 ml of 100 mM NaCl, 10% glycerol, 0.02% DDM and 25 mM of either of six buffers (MES pH 5.5, PIPES pH 6.6, HEPES pH 7.5, Tris pH 8.0, bicine pH 8.5 or glycine pH 9.0). Buttons were observed after a one night dialysis for presence of a precipitate.

Selected buffer conditions were tested further by large-scale dialysis in 25 kDa NMWCO bags against a 100-fold volume excess of 100 mM NaCl, 10% glycerol, 0.02% DDM and 25 mM of either of three selected buffers (HEPES pH 7.5, bicine pH 8.5 or glycine pH 9.0). Concentration of ExbB–ExbD samples in those buffers followed. Centrifugal concentrator units with 100 kDa NMWCO (YM100, Millipore) were used. Frequent interruption of centrifugation and mixing of the sample was performed to prevent precipitation caused by concentration polarization. Concentration was stopped when samples displayed abundant precipitation or when very low volumes were

reached (150 μl or less). After concentration, protein concentration was measured in all three samples using the BCA kit.

2.10 Crystallization trials

2.10.1 Vapour-diffusion crystallization – low-throughput

ExbB–ExbD samples at a concentration of 10 mg/ml in 25 mM Gly pH 9.0, 100 mM NaCl, 10% glycerol, 0.02% DDM were mixed at a 1:1 ratio with mother liquor to form 1 μl x 1 μl drops on glass coverslips. The coverslips were inverted and applied to grease-sealed wells (24-well trays) containing 500 ml of mother liquor. Trays were incubated at 18^{o}C and drops were observed with an Olympus SZ30 microscope on a regular basis for the presence of crystals and other features of interest. The MbClass Suite I (NeXtal) incomplete factorial screen (96 conditions) as well as 23 basic screens (24 conditions each) were performed by this method.

2.10.2 Vapour-diffusion crystallization – medium-throughput

ExbB–ExbD samples at a concentration of 8, 10, 12 or 15 mg/ml in 25 mM Gly pH 9.0, 0.02% DDM were mixed at a 1:1 ratio with mother liquor using a Mosquito nanopipetting robot (TTP LabTech) to form 200 nl x 200 nl, 250 nl x 250 nl or 300 nl x 300 nl drops on 96-well tray adhesive lids. Lids were inverted and applied to 96-well tray with wells containing 100 ml of mother liquor. Trays were incubated at 4^{o}C or 20^{o}C and drops were observed with a Leica microscope on a regular basis for the presence of

crystals and other features of interest. The MemGold (Molecular Dimensions ; 96 conditions) and Abramson lab bicelle (UCLA ; 192 conditions) screens, in several detergent conditions, along with five basic screens (96 conditions each) were performed with this method.

2.10.3 Bicelle crystallization

ExbB–ExbD samples at a concentration of 12 mg/ml in 25 mM Gly pH 9.0, 0.02% DDM were mixed at a 1:7 ratio with a 35% 3:1 1,2-Diacyl-sn-Glycero-3-phosphocholine (DMPC):3-(cholamidopropyl) dimethylammonio-2-hydroxy-1-propanesulfonate (CHAPSO) bicelle mixture and incubated on ice for 30 minutes. The protein-bicelle mixture was then set-up as described in Section 2.10.2. Trays were incubated at 20^{o}C and observed with a Leica microscope on a regular basis for the presence of crystals and other features of interest. The Abramson lab bicelle screens (192 conditions) were performed with this method

2.10.4 Microbatch crystallization under oil – high-throughput

ExbB–ExbD sample at a concentration of 10 mg/ml in 25 mM Gly pH 9.0, 100mM NaCl, 10% glycerol, 0.02% DDM was sent to the Hauptman-Woodward high-throughput screening laboratory (Buffalo, New York) for mass crystallization using the microbatch-under-oil technique (108,109). A plate of 1536 wells, with each well containing a unique crystallization cocktail, was used for screening. All wells were imaged prior to addition of sample to control for the cocktail quality. Sample was then added to the well in a 1:1

ratio with the screening cocktail (200 nl protein and 200 nl cocktail). Crystallization experiments were incubated at 14^{o}C and each well was automatically imaged one day, one week, two weeks, three weeks, four weeks, five weeks and six weeks following sample addition. After imaging, digital images were made available and viewed on a computer screen.

2.11 Detergent exchange

On-column detergent exchange was performed by using AEC as described in Section 2.5. In short, sample was loaded and the complex was bound to a Poros HQ20 column pre-equilibrated with a buffer containing the original detergent. The column was washed with 5 CV of the original detergent-containing buffer, followed by an identical wash with a buffer containing 2 x CMC of the destination detergent. Gradient and elution were performed in 2 x CMC of the destination detergent. Detergents tested were n-decyl-β-D-decylmaltopyranoside (DM), n-octyl-β-D-glucopyranoside (OG), and experimental amphiphiles GNG-12 and MNG-28.

2.12 Detergent quantitation by nuclear magnetic resonnance

Detergent quantitation was performed following the methods developed by Maslennikov *et al.* (110,111). Shortly, 520 μL of samples to be detergent quantitated were mixed with 130 μL of a 5 mM DSS solution in D_2O. 1D proton spectrums were recorded at the Québec and Eastern Canada NMR Facility on a Varian INOVA 500 MHz

spectrometer. Comparison of integrated detergent peaks with integrated DSS standard peaks allowed calculation of detergent concentration.

Chapter 3- Results

3.1 Overexpression of ExbB–ExbD complex

Conditions for the overexpression of the ExbB–ExbD complex were tested following the protocol provided in Section 2.2.2. Optimal conditions for growth and expression were the following: Terrific Broth supplemented with 10% glycerol and a four-hour induction at an optical density of 1.0 to 1.2 with an IPTG concentration of 0.5 mM. Figure 11 shows the result of this expression strategy for the pExbBD construct and other plasmids.

The silver stained gel in Figure 11A shows that membranes of induced cells are enriched in two species when compared to non-induced cells. Indeed, two prominent bands appear in the membrane lane (red boxes), while the soluble protein content appears unchanged, strongly arguing for the expected overexpression of two membrane proteins. Both species migrate to distances that, when compared to the molecular weight ladder, agree well with the expected weights for ExbB (26 kDa) and ExbD (15 kDa) based on the molecular weight ladder. Confirming this observation, a similar construct lacking a TEV protease cleavage site in ExbD shows a highly similar pattern, but expectedly, with an ExbD band of a slightly lower molecular weight. Furthermore, another construct, carrying ExbB only, shows enrichment of only the higher molecular weight species. Western blotting with anti-His antibodies of the same samples reveals immunoreactive bands at the position of His-tagged ExbD, further confirming successful overexpression of the complex.

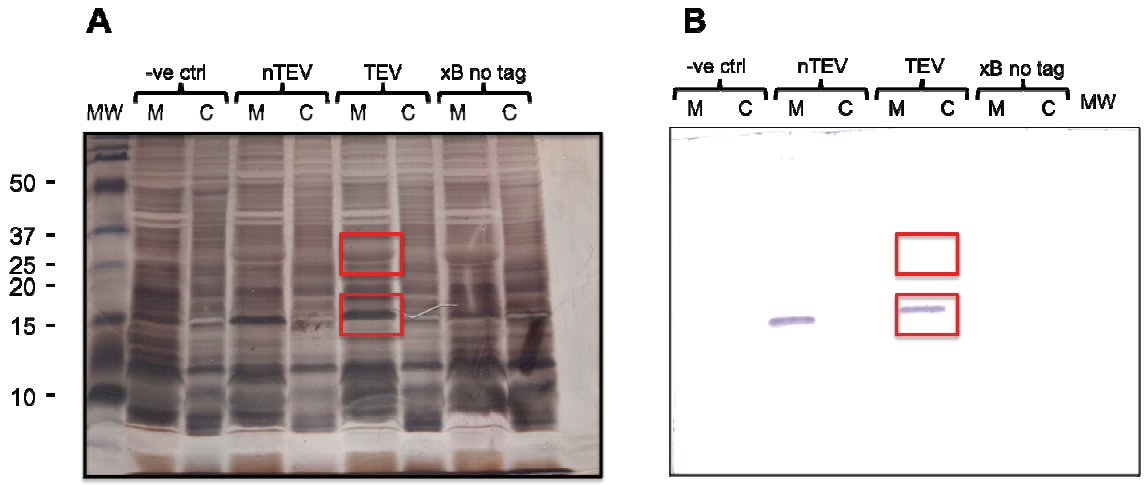

Figure 11. Overexpression of the ExbB–ExbD complex. M15 pREP4 cells overexpressing ExbB–ExbD were lysed by sonication. The lysate was cleared of unlysed cells and large debris by centrifugation. Total membranes were separated from soluble material by ultracentrifugation and resolubilized in SDS. Both solubilized membranes (M) and the soluble fraction (C) were analyzed by SDS-PAGE and Western Blot. One gel was silver stained (A) while the other was transferred to a PVDF membrane and immunoblotted with anti-His antibodies (B). The gel and Western blot show the expression test for three constructs. nTEV: ExbB–ExbD construct with uncleavable His-tag; TEV: ExbB–ExbD construct with cleavable His-tag (pExbBD), xB no tag: untagged ExbB construct. Negative control is uninduced cells carrying the TEV construct. Red boxes underline prominent bands of overexpressed ExbB (top) and ExbD (bottom). MW: molecular weight ladder.

3.2 Preparation of ExbB–ExbD samples

Successful expression of the ExbB–ExbD complex in *E. coli* membranes enabled its isolation by biochemical methods. Large-scale production of ExbB–ExbD was conducted as detailed in Section 2.2.2. Cells were lysed and total membranes were collected following the method in Section 2.5. Membranes were resolubilized in 1% DDM. The relative low cost and crystallization successes of this detergent for other membrane proteins explain the initial choice of DDM for membrane solubilization. In addition, membrane protein structural genomics initiatives have underlined the suitability of DDM as a reasonable first selection (112). Solubilized membranes, once cleared of unsolubilized material, are applied to an IMAC column. Figure 12A shows the chromatogram of that first chromatographic step. Two main signals are visible on this chromatogram. The flow-though (red arrow) consists of unbound material. The elution peak (red box), recovered by applying 300 mM imidazole to the column, contains material that has an affinity for the Ni^{2+} resin, such as His-tagged proteins. The silver stained gel in Figure 12C reveals the content of the IMAC elution peak. Two major species are observed, along with a number contaminants that non-specifically bound to the resin and/or to His-tagged protein. The two prominent bands correspond in all aspects to those seen in Figure 11A. The elution of the lower ExbD band is expected, as it was engineered to carry a His-tag. The upper band represented untagged ExbB that was pulled down by its interacting partner ExbD. This observation suggested the formation of a stable complex between the two proteins.

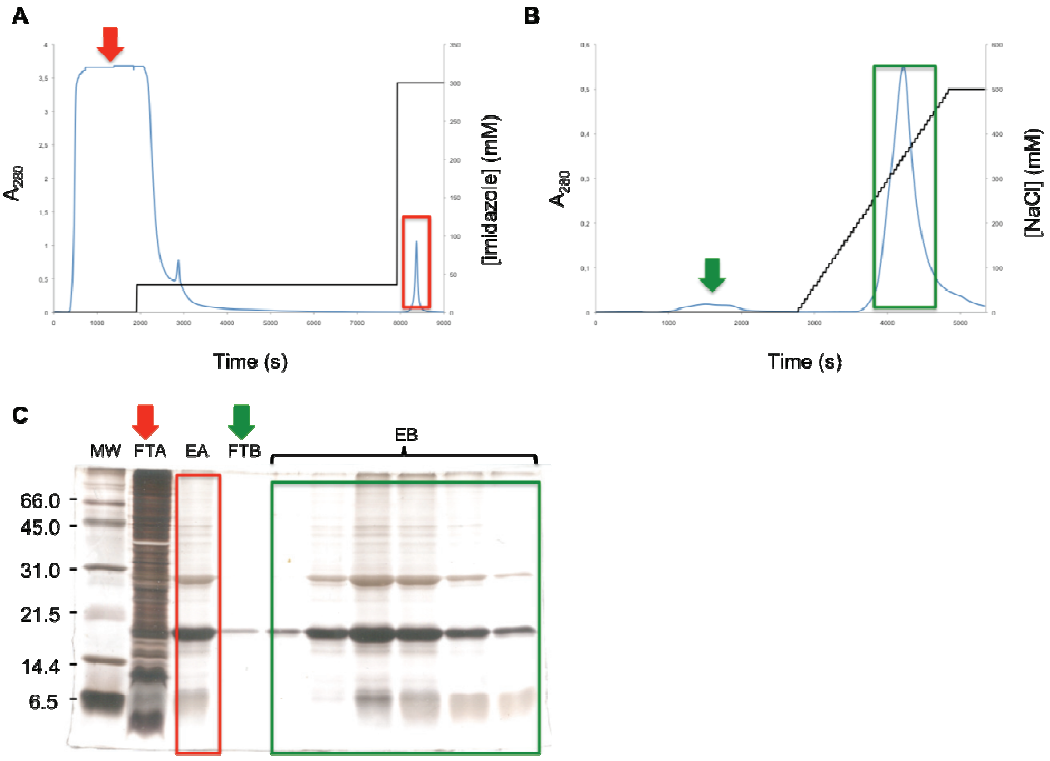

Figure 12. Purification of the ExbB–ExbD complex. Solubilized membranes containing ExbB–ExbD were applied to IMAC (chromatogram in A) to capture the His-tagged complex. Material not retained by the resin eluted in the flow-through (red arrows). Column was washed extensively with 35 mM imidazole. Captured material was eluted with high imidazole, and the elution peak (red boxes) was collected. After IMAC, the recovered material was further purified, cleared of TEV protease and buffer exchanged by AEC (chromatogram in B). Unbound material eluted in the flow-through (green arrows), while the bound material was further eluted by a salt gradient (green boxes). Material from those purification steps was submitted to SDS-PAGE and silver stained (C) to evaluate content and purity of the samples. MW: molecular weight ladder; FTA: flow-through from IMAC; EA: elution from IMAC; FTB: flow-though from AEC; EB: elution from AEC.

Anion exchange chromatography (AEC), was used as a second step in the purification of ExbB–ExbD. Figure 12B shows a typical AEC chromatogram for ExbB–ExbD. Unbound material is once again eluted in the flow-through while bound proteins elute following the application of a salt gradient. Figure 12C shows the content of both the AEC flow-through and elution. Differences in protein purity between IMAC and AEC samples may not be obvious. However, this step is used to remove TEV protease and to perform exchanges of buffer and detergent. In addition, this second chromatography can have an effect on the lipid and detergent content of the membrane protein and it may remove minor but critical material. The typical yield of this overexpression and purification strategy measured after AEC is more than 6 mg per litre of original cell culture.

After IMAC, ExbB–ExbD samples were typically submitted to TEV proteolysis for removal of the hexahistidine tag of ExbD. The cleavage reaction mixture, described in Section 2.5, was incubated for 72 hours at 4^{o}C. As mentioned in the previous paragraph, AEC was performed to remove TEV protease from the sample. To assess the success of cleavage and removal of TEV, SDS-PAGE and Western blotting of the samples was performed. The silver stained gel and Western blot in Figure 13 provide an example of a TEV cleaved ExbB–ExbD sample purified by AEC. The silver stain in Figure 13A shows the total content of the samples, both cleaved and uncleaved. The Western blot displays no His-reactive bands in the TEV cleaved sample lanes. Only the post-IMAC, uncleaved positive control displays an ExbD immunoreactive band. This

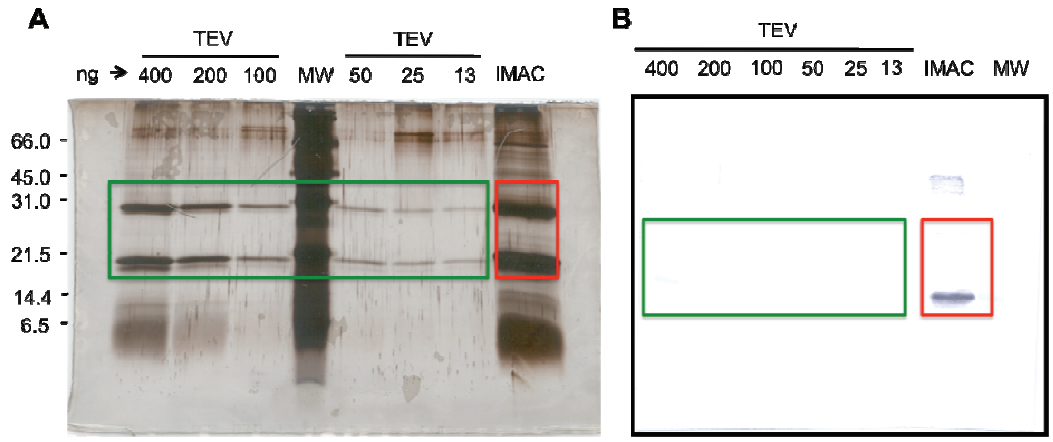

Figure 13. Cleavage of the His-tag using TEV protease. Material eluted from IMAC was pooled, supplemented with glycerol to a concentration of 10% and mixed with TEV protease. The amount of protease added corresponds to one tenth of the milligram amount of recovered ExbB–ExbD complex. The reaction mixture was incubated with stirring for 72 hours at 4°C. The sample was then submitted to AEC to remove TEV protease. Purity and the absence of protease was checked by running the sample on gel and performing silver staining and Western blotting. The presence of both residual tagged protein and contaminating TEV protease was revealed by blotting with anti-his antibodies, as both species carry His-tags. The silver stained gel in (A) shows a series of dilutions of a TEV cleaved and AEC-purified sample with 13 ng to 400 ng of protein loaded per lane (green box). A post-IMAC, pre-cleavage sample is run as a control (red box). The corresponding Western blot is shown in (B). The cleaved samples show no immunoreactive bands, indicating complete cleavage and removal of TEV protease. The control in lane IMAC shows an immunoreactive ExbD band.

demonstrates successful cleavage of the ExbD His-tag and the complete removal of His-tagged TEV protease from the sample.

Size-exclusion chromatography separates particles, such as proteins and protein detergent complexes, on the basis of molecular size. It therefore allows an assessment of protein sample homogeneity and the measurement of molecular weight. A single, monodisperse peak is indication of a homogeneous sample. Size-exclusion chromatography of ExbB–ExbD samples was performed as described in Section 2.6. Figure 14 shows the chromatogram of a TEV-cleaved sample purified by AEC. It displays a single, monodisperse peak indicative of a homogeneous sample. In addition, using the calibration curve prepared for this column (see upper-right panel), a molecular weight of 480 kDa was estimated for the ExbB–ExbD protein-detergent complex.

The purity and homogeneity of ExbB-ExbD samples made them amenable to techniques used in examining membrane protein structure. Samples obtained directly from IMAC were used for negative-staining electron microscopy (Section 3.3). TEV-cleaved, AEC-purified samples were used in crystallization trials after the solubility of the complex was optimized (Sections 3.4 and 3.5).

3.3 Negative-staining electron microscopy of ExbB–ExbD

Negatively stained preparations of ExbB–ExbD prepared as described in Section 2.7 were observed under the electron microscope. The observations revealed particles

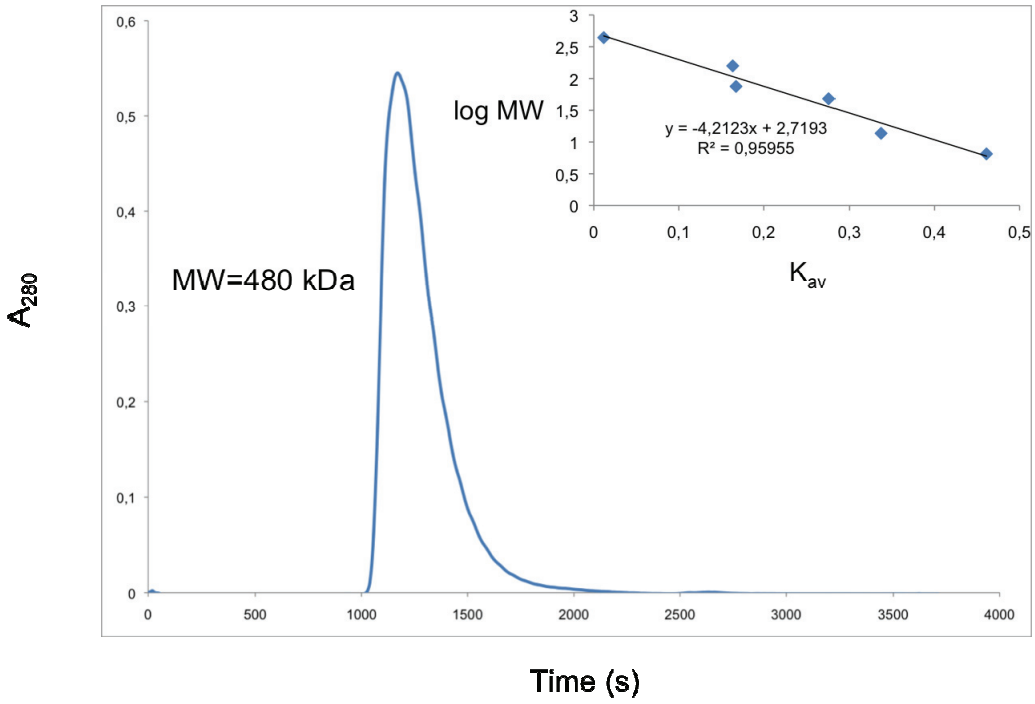

Figure 14. Monodispersity of the ExbB–ExbD complex by size-exclusion chromatography. The monodispersity of the complex was assessed by SEC on a Toyopearl HW-55S column. Chromatogram of a typical SEC run of the TEV-cleaved, AEC-purified complex is displayed. The presence of a single peak suggests monodispersity of the sample. The molecular weight (MW) of the complex was assessed by using the calibration curve prepared for this column (upper-right panel). The K_{av} is a corrected and normalized measure of a particle's elution volume from a size-exclusion column.

approximately 10 nm in diameter. The particles were well-defined and easily identifiable. They also displayed remarkable uniformity, confirming conclusions from SEC on the homogeneity of ExbB–ExbD samples. Based on this observation, high-resolution electron micrographs were collected to enable the generation of 2D reconstructions. Figure 15A shows a representative micrograph. From the collected micrographs, 2000 individual particles were selected and a sample of a few selected particles is given in Figure 15B. Preliminary 2D reconstructions were generated from those particles by the maximum-likelihood multi-reference alignment method using six class-averages. Figure 15C shows those preliminary reconstructions. Their poor definition precludes meaningful interpretation, but we surmised that the complex seemed to adopt a pentameric arrangement. Shortly thereafter, while I was visiting the laboratory of our collaborator – Dr. Jeffrey Abramson at the University of California Los Angeles – my colleague Andrew Cottreau used four of those preliminary reconstructions as search models for the automatic selection of 45 321 more particles. He used particles in another round of maximum-likelihood multi-reference alignment with 12 class-averages. Figure 15D shows his reconstructions. There is little variation from one class-average to another, meaning little heterogeneity among the particles. In practice, this is an indication that protein complexes show some level of alignment, preferentially presenting a certain face with respect to the grid.

Examination of the particle reconstructions in Figure 15D strongly suggests a pentameric organization for the complex, as was hypothesized from earlier

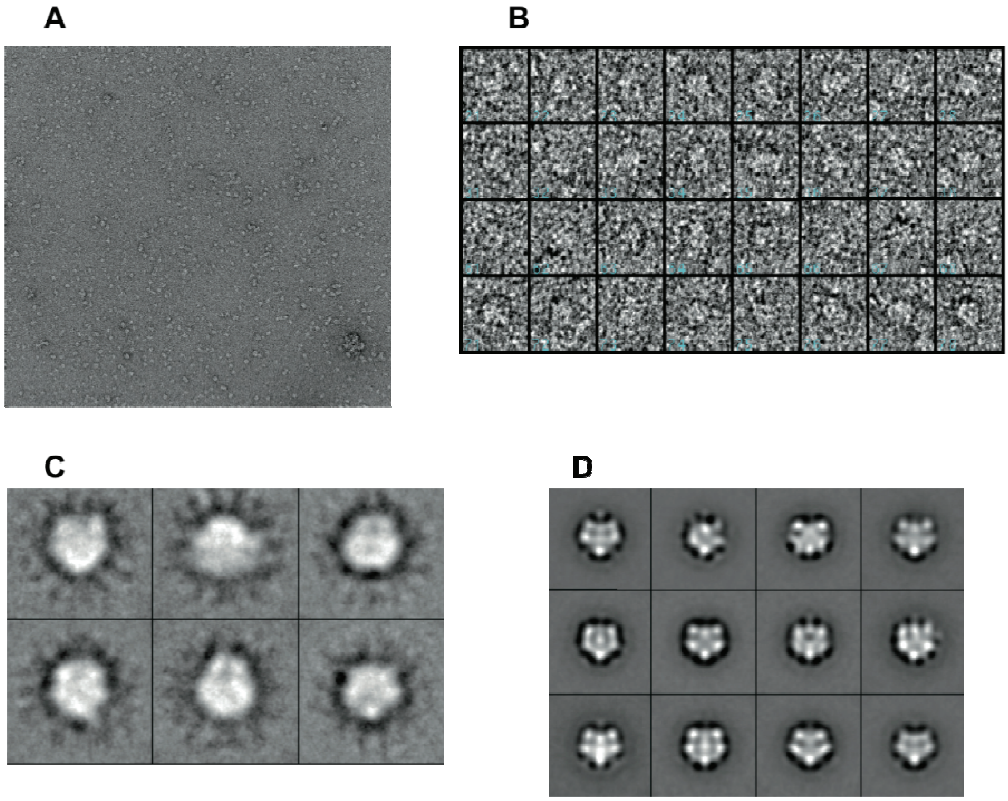

Figure 15. Electron microscopy of the ExbB–ExbD complex. Uncleaved, post-IMAC ExbB–ExbD samples were deposited onto glow-charged electron-microscopy grids. The grids were fixated and stained with uranyl acetate and then observed with a Tecnai F20 electron microscope. (A) Typical low-dose EM micrograph of negatively-stained ExbB–ExbD samples. The grid surface is littered with uniform particles. (B) In preparation for single-particle reconstruction, individual particles were selected from micrographs using the program boxer of the EMAN package. The panel shows a selection of boxed particles. (C) From 2000 manually selected particles, initial 2D reconstructions were generated by the maximum-likelihood multi-reference alignment method using six class-averages (D) Using four of my initial reconstructions as search models, my colleague Andrew Cottreau automatically selected 45 321 particles using *Signature*. He used those particles for another round of ML alignment to generate better defined reconstructions using 12 class-averages.

reconstructions. Five areas of high density are seen surrounding a central depression of somewhat lower density. This is especially obvious in the class-average situated at the second row, second column; other reconstructions are also indicative of this organization. This kind of organization also suggests a cylindrical, perhaps ring-like higher-order structure for the complex. The class-average at the second row, third column may be suggestive of such a structure. That being said, these 2D reconstructions are preliminary structural indications. However, their quality indicates that collection of more data should be pursued and that the generation of more informative 3D reconstructions is possible.

3.4 Optimizing solubility of ExbB–ExbD complex

Protein samples typically have concentrations of approximately 10 mg/ml when used for setting up crystallization trials. This usually requires that they be concentrated after the last purification step. Initial attempts at concentrating ExbB–ExbD samples in the original buffer (25 mM Tris pH 8.0, 100 mM NaCl, 0.02% DDM) proved unsuccessful. The protein precipitated at rather low concentrations and could not be kept stably at more than approximately 2 mg/ml. Additives and modifications to the composition of the buffer were used to increase protein solubility. The first attempt at increasing ExbB–ExbD solubility consisted in the simple addition of 10% glycerol to the original buffer solution. Concentration was attempted in these modified conditions and the maximum

Table I. Optimization of the solubility of ExbB–ExbD. Buffer conditions were screened to optimize the solubility of the complex. In (A), ExbB–ExbD was dialyzed against six buffer conditions. After dialysis, the samples were inspected and the observations are listed in the table. Samples dialyzed against the three buffers listed in table (B) were concentrated using a centrifugal filter concentrator and observed during the process. Concentration was stopped when a substantial precipitate was observed or when the samples reached a volume of 150 μl or less. Protein concentration of the samples was assayed.

A

Buffer	Observation
MES pH 5.5	Heavy precipitate
PIPES pH 6.6	Cloudy
HEPES pH 7.5	Clear
Tris pH 8.0	Clear
Bicine pH 8.5	Clear
Glycine pH 9	Clear

B

Buffer	Precipitate	Max conc. (mg/ml)
HEPES pH 7.5	Yes	2.77
Bicine pH 8.5	No	7.44
Glycine pH 9	No	6.35

solubility of the complex was increased to 4 mg/ml. While this was a clear improvement over initial attempts, it was insufficient for crystallization trials.

Screening of buffers and pHs was next attempted. Microvolumes of ExbB–ExbD sample were dialysed against six buffers covering a pH range going from 5.5 to 9.0. After overnight dialysis, the samples were observed and their appearance was noted. Table IA lists the buffers tested and the appearance of the samples after overnight dialysis. Acidic conditions lead to precipitation, while neutral and basic buffers are observed to remain clear. These simple observations clearly show that the solubility of the complex decreases at acidic pHs.

Following the observation that ExbB–ExbD was less soluble at low pH, further tests were focused on higher pH conditions. Specifically, HEPES pH 7.5, bicine pH 8.5, and glycine pH 9.0 were tested. Preparations of the complex were dialysed extensively against the three buffers and then concentrated. The concentration process was interrupted frequently to observe the samples and for mixing, which prevents concentration polarization – a phenomenon often observed with centrifugal concentrator filter devices. Concentration was stopped when samples displayed abundant precipitation or when they reached a very low volume. After concentration, protein content was assayed in all three samples. Table IB shows the results of this concentration test. A precipitate was observed rapidly in HEPES pH 7.5 when concentrated to just below 3 mg/ml. Bicine pH 8.5 and glycine pH 9.0 did not display precipitate and reached higher concentrations than was previously achieved. These results confirmed the higher solubility of the ExbB–ExbD complex at basic pH.

Because of compatibility issues with the Bicinchronic acid (BCA) protein assay, bicine was rejected and glycine was selected as the most desirable buffer. Larger amounts of ExbB–ExbD preparation were buffer exchanged to glycine pH 9.0 and concentrated to further test the protein concentrations that could be achieved in this condition. A concentration of 9.31 mg/ml – sufficient for crystallization trials – was measured. Subsequently, even higher concentrations were routinely achieved. Eventually, it was also determined that, at alkaline pHs, glycerol was not needed to obtain high concentrations of ExbB–ExbD.

Together, these solubility optimization tests provided the opportunity to initiate crystallization trials. They also provided information on the behavior of the protein in solution that can prove useful for crystallization and characterization of the complex (Section 4.1).

3.5 Crystallization of ExbB–ExbD complex

With pure, homogeneous and highly concentrated ExbB–ExbD samples in hand, crystallization trials were launched. Samples at 10 mg/ml in 25 mM glycine pH 9.0, 100 mM NaCl, 10% glycerol, 0.02% DDM were used to set-up an incomplete factorial screen of 96 conditions. The MbClass Suite I (Nextal) screen was used for initial trials. All observations will not be reported here, but the thought and experimental process followed for important leads will be outlined. Early after setting up the incomplete factorial screen, peculiar features akin to spherulites were observed in condition 16

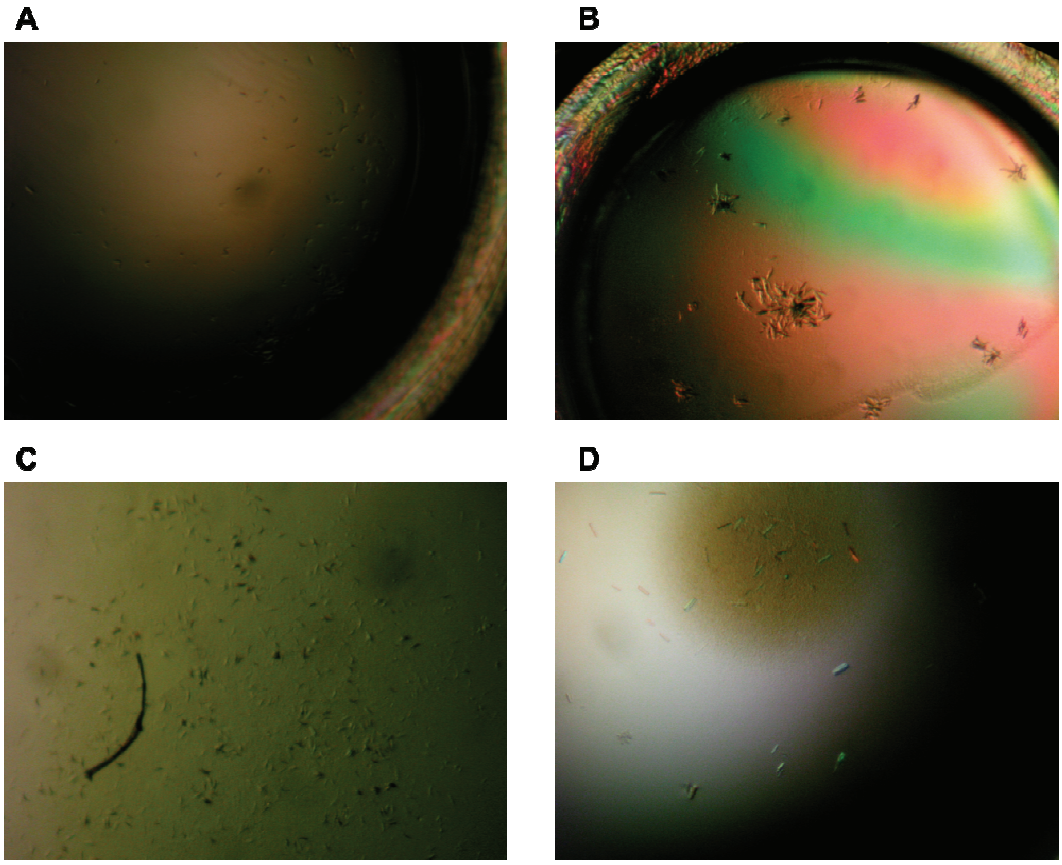

Figure 16. A crystallization lead: $MgSO_4$ and 2-methyl-2,4-pentanediol. This figure outlines progress in the pursuit of one ExbB–ExbD crystallization lead. In a basic screen of magnesium sulfate concentration *versus* pH, microcrystals such as those shown in (A) were observed after six weeks incubation. In parallel, also after six weeks incubation, the feather-like crystals in (B) were observed as part of an incomplete factorial screen. A basic screen around the condition in (B) testing $MgSO_4$ concentration *versus* MPD concentration was performed and the crystals were reproduced. Panel (C) shows the best crystals obtained from that basic screen. Another basic screen was performed, testing $MgSO_4$ concentration *versus* glycerol concentration, and rod crystals such as those in (D) were observed. For each panel, detailed conditions are:
(A) 1.4M $MgSO_4$, 0,1M HEPES pH 7.5, 0,1M Na/K tartrate (B) 0.6M $MgSO_4$, 0.1M HEPES pH 7.5, 4% MPD (C) 1.0M $MgSO_4$, 0.1M HEPES pH 7.5, 4% MPD (D) 0.2M $MgSO_4$, 0.1M HEPES pH 7.5, 4% MPD, 10% glycerol.

(0.4 M $MgSO_4$, 0.1 M Tris pH 8.5, 0.1 M Na/K tartrate). To investigate that condition, a basic screen testing $MgSO_4$ concentration *versus* pH was set-up. After six weeks, rod-shaped microcrystals were observed in several drops of that basic screen (Figure 16A). In parallel, condition 18 of the incomplete factorial screen (0.6 M $MgSO_4$, 0.1 M HEPES pH 7.5, 4% 2-methyl-2,4-pentanediol (MPD)) yielded feather-like microcrystals, also after six weeks (Figure 16B). As a follow-up on those two highly similar initial hits, a new basic screen was set-up, testing $MgSO_4$ concentration *versus* MPD concentration. Lawns of rod-shaped microcrystals were also obtained in many drops of that basic screen. Figure 16C provides an example. New, narrower trials testing the effect of 10% glycerol and screening $MgSO_4$ concentration *versus* MPD concentration as well as $MgSO_4$ concentration *versus* glycerol concentration were performed. Promising sharp rods were obtained, such as those displayed in Figure 16D. Our observations suggest that lower $MgSO_4$ concentrations are preferred to allow growth of fewer and larger crystals. It also appears that 4% MPD is the optimal precipitant concentration. Lastly, the use of glycerol as an additive appears to have a positive effect on the quality and number of crystals that are obtained. The crystals are still small and numerous, not yet permitting the collection of diffraction data, but the prospect to optimize this lead is valuable at this point.

Another similar lead was identified through several independent observations. Few rod-shaped microcrystals were first observed in 0.1 M $MgCl_2$, 0.1 M Tris pH 8.0,

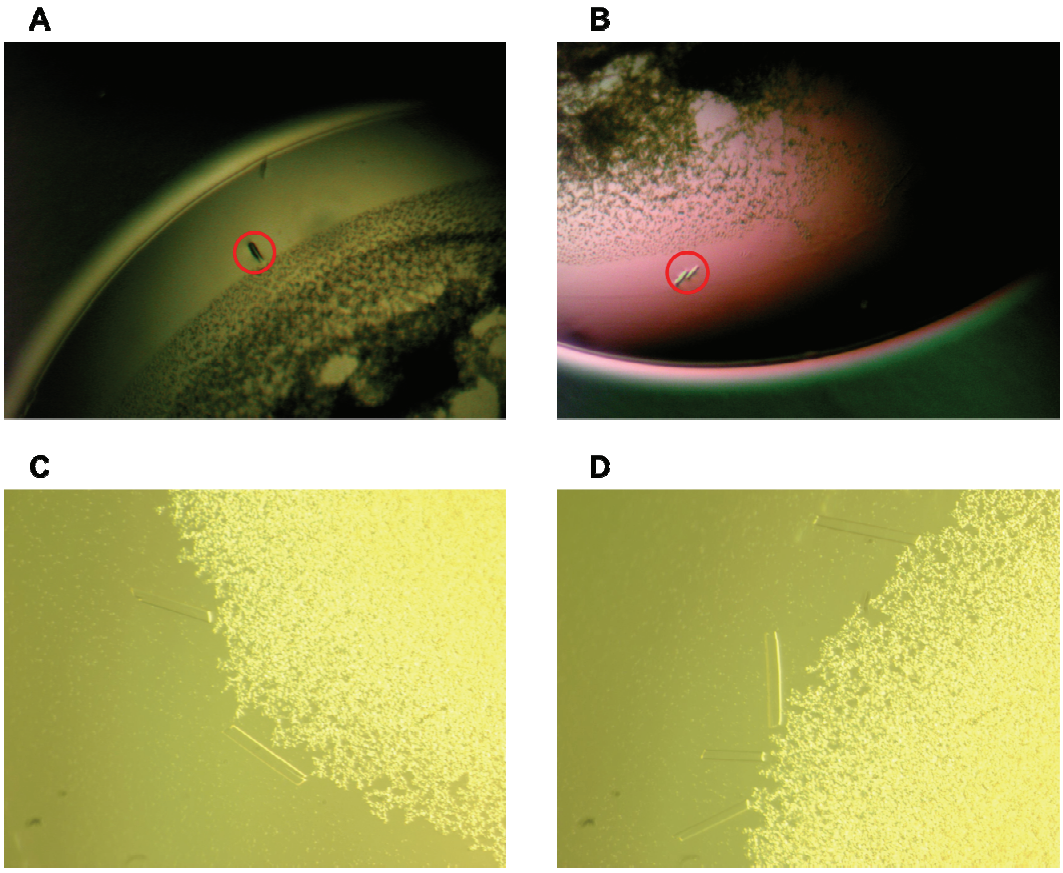

Figure 17. A crystallization lead: $MgCl_2$ and PEGs. This lead was identified as separate events. First, a basic screen of PEG 1500 concentration *versus* glycerol concentration identified a few microcrystals in 0.1 M $MgCl_2$, 0.1 M Tris pH 8.0, 25% PEG 1500 (panels A and B). After a two month incubation at 4^{o}C, the MemGold sparse matrix screen showed crystals in 0.1 M $MgCl_2$, 0.1 M Tris pH 8.5, 10% PEG 2000 (see panels C and D).

25% PEG 1500 (Figure 17AB). During an unrelated experiment, performing the MemGold sparse matrix screen at 4°C, similar microrods were observed in 0.1 M $MgCl_2$, 0.1 M Tris pH 8.5, 10% PEG 4000 after only 6 day incubation. Unfortunately, the crystals rapidly disappeared and no picture could be taken. Their disappearance, however, was encouraging in the sense that it supports the idea that they were protein crystals. Indeed, salt crystals tend to be hardier and more stable. In the same sparse matrix screen, at 4°C still, other crystals were observed after a two month incubation period (Figure 17CD). Those are by far the largest crystals that could be observed in ExbB–ExbD crystal trials so far. The crystals were large enough to be mounted on crystallographic loops and exposed to an X-ray beam from a rotating anode home source. Two crystals were tested: they displayed no salt diffraction, but also no protein diffraction. Poor diffraction from protein crystals indicates imperfect crystal packing and requires crystal optimization.

3.6 Detergent exchanges

As explained in Section 1.6.3, detergent can have a dramatic effect on the crystallization of membrane proteins. The exchange of samples to new detergents is among the tools available to the membrane protein crystallographer to optimize and facilitate the growth of crystals. With the objective of identifying new leads and optimizing existing crystallization leads, exchange of ExbB–ExbD samples from DDM to new detergents was performed following the on-column method detailed in Section 2.11. At this point, two detergents - n-decyl-β-D-decylmaltopyranoside (DM), n-octyl-β-

D-glucopyranoside (OG) – as well as two experimental amphiphiles – GNG-12 and MNG-28 provided by our collaborator Dr. Jeffrey Abramson at the University of California Los Angeles – were tested for their ability to efficiently stabilize the ExbB–ExbD complex in solution. Only OG failed to keep the complex in solution, causing the formation of a white and powdery precipitate only a few hours after exchange. The protein complex was stable in the three other detergents and sparse matrix crystal screens were set-up in each of them. No leads have been identified in those detergents after two months of observation.

3.7 Detergent quantitation

Because of the effect that detergents can have on crystallization, variations in their concentration should be avoided in samples destined to crystal trials. For the accurate reproduction of crystallization conditions, it is desirable to measure actual detergent concentration in samples. We wished to adopt a method of detergent quantitation, and we opted for the NMR strategy published by Maslennikov *et al.* (110,111). Samples containing 1 mM 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS) and 20% D_2O had their 1D proton spectrum recorded and intergrated DSS peaks were used as an internal standard to calculate detergent concentration from integrated detergent peaks. So far, the method was applied to quantitate detergent in two samples. First, the protein-free buffer (25 mM, 100 mM NaCl, 10% glycerol, 0.02% DDM) was quantitated. The measured DDM concentration was 0.015%, in close agreement with the theoretical or expected value of 0.02% (Figure 18).

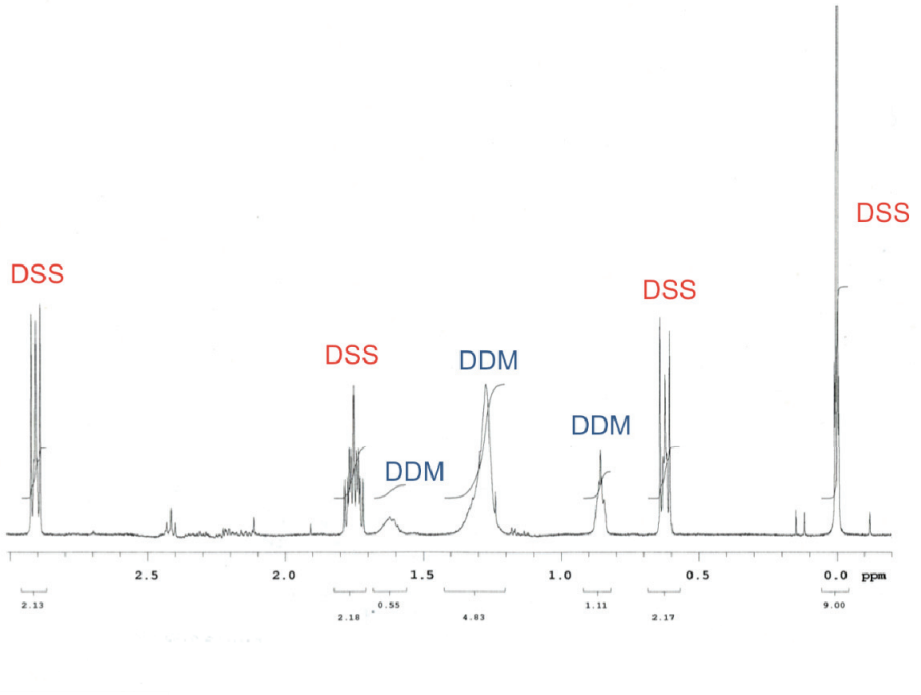

Figure 18. 1D proton spectrum of the ExbB–ExbD glycine buffer. The theoretical composition of the buffer was 25 mM glycine, 100 mM NaCl, 10% glycerol, 0.02% (or 0.39 mM) DDM. Sample was prepared as described in Section 2.12. Integration of DDM peaks (labelled in blue) was compared to that of DSS peaks (labelled in red) for quantitation of detergent content. The measured detergent concentration is 0.015%, or 0.29 mM.

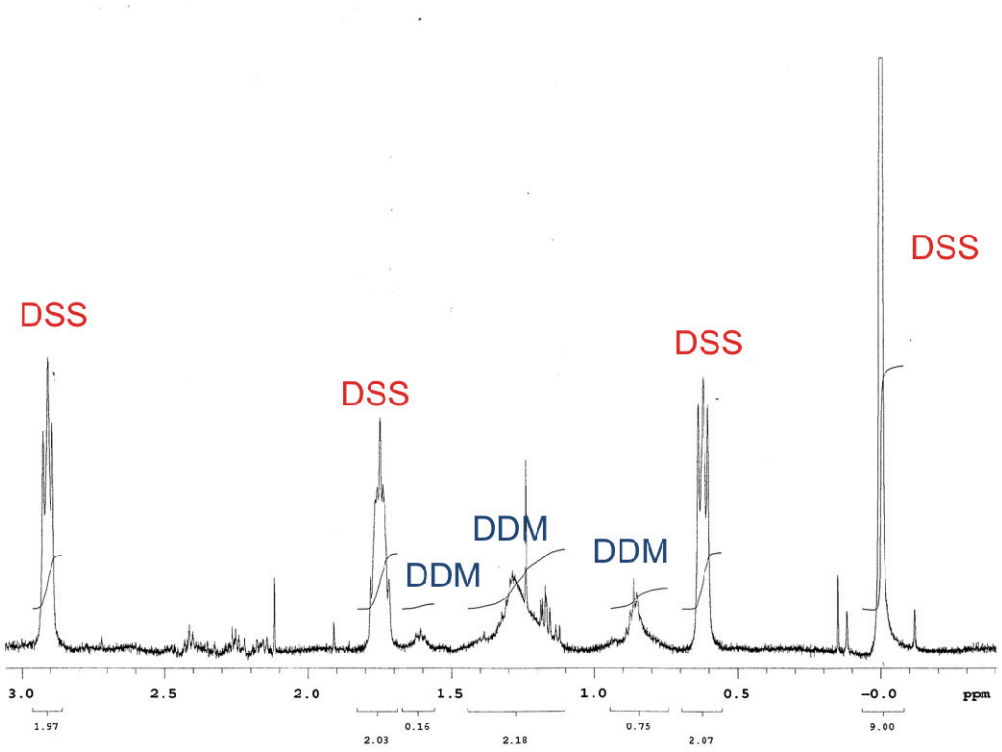

Figure 19. 1D proton spectrum of an ExbB–ExbD sample in glycine buffer. The theoretical composition of the buffer was 25 mM glycine, 250 mM NaCl, 10% glycerol, 0.02% (or 0.39 mM) DDM. Sample was prepared as described in Section 2.1.2. Integration of DDM peaks (labelled in blue) was compared to that of DSS peaks (labelled in red) for quantitation of detergent content. The measured detergent concentration is 0.007%, or 0.15 mM.

The same procedure was applied to a post-AEC sample. The expected buffer composition of the sample was 25 mM, 250 mM NaCl, 10% glycerol, 0.02% DDM with a protein concentration of 0.750 mg/ml. The measured DDM concentration was 0.007%, significantly lower than the anticipated 0.02% value (Figure 19). However, the protein appeared stable in solution, suggesting that only a minimal DDM concentration is necessary to keep ExbB–ExbD soluble.

This detergent quantification strategy has been successfully applied to ExbB–ExbD samples. It should be possible to use it to standardize detergent concentration in every preparation destined to crystallization, both for this project and future crystallization projects in the lab.

Chapter 4 – Discussion and Conclusion

4.1 Discussion

The role of proteins ExbB and ExbD proteins in harnessing the PMF and energizing OM transport has been firmly established. Their interaction with each other within a complex is also well documented (57,59,60). The stoichiometry of the complex, however, remains unknown. More importantly, the molecular determinants of their interaction and mechanism of action remains elusive because of limited structural information. The periplasmic domain of ExbD is the only portion of known structure, while ExbB remains structurally uncharacterized. Furthermore, the higher-order structure of the complex remains unknown. The objectives of this thesis research were to pursue structural determination by electron microscopy and macromolecular crystallography to gain insight into the structure of ExbB and ExbD as a complex.

The milligram amounts of protein necessary for crystallization require an efficient strategy for abundant overexpression of the target. Our optimized co-expression strategy in *E. coli* M15 pREP4 showed prominent overexpression of ExbB and ExbD readily from a silver stained gel of total membranes. This observation, confirmed by Western blotting, is not common with membrane proteins and prefigures abundant purification yields. Previous attempts in our lab at expressing each protein individually in different *E. coli* strains met with qualified success (not shown). Tight expression control in M15 pREP4 may explain much of the observed expression with this strategy. The

stabilizing effect of ExbB and ExbD on one another probably also plays a role. It may parallel the documented stabilizing effect of ExbB on TonB (113).

From total membranes that were enriched with ExbB–ExbD, immobilized metal-ion affinity chromatography allowed the recovery of abundant amounts of our purified target proteins. Silver-staining of IMAC elution fractions revealed two prominent species – ExbB and ExbD – and a number of minor contaminants. It should be noted that only ExbD was engineered with a His-tag at its C-terminus. ExbB was not affinity-tagged and does not display intrinsic affinity for divalent metal cations. The fact that ExbB eluted specifically and in large amounts with His-tagged ExbD from an IMAC column strongly argues for stable interaction between the two proteins. This data further complements previous pull-down studies in which tagged ExbB successfully retained native ExbD on nickel resin (57). In the subsequent AEC step of the purification protocol, both proteins co-eluted despite widely differing isoelectric points: this argues for the formation of a stable complex between ExbB and ExbD. This same anion exchange chromatography was used successfully for buffer and detergent exchange (see below). It follows the extended cleavage of the His-tag from ExbD by TEV protease. Cleavage at the engineered cut-site at the C-terminal end of ExbD was shown to be performed successfully, and to completion. Clearing of the protease, which is also His-tagged, from our samples was also shown to be complete. It is not known if cleavage of the tag will be beneficial, detrimental or even relevant to crystallization. Effect of His-tags on crystallization varies from protein to protein and is determined empirically based on crystallization success and later inspection of crystal packing. His-tags are often invisible

in crystal structures due to flexibility and disordered behavior. Examples of His-tags playing a critical role in the formation of crystal contacts are also known (23). Having the opportunity to keep or remove a tag means more options for the crystallographer and our cleavage strategy provides that option.

Homogeneity of protein samples is critical to all structural techniques. In electron microscopy, uniform, non-aggregated particles of identical nature are necessary to produce valid and meaningful reconstructions. Aggregated or heterogeneous samples will also impair the process of protein crystallization. The homogeneity of ExbB–ExbD samples was assessed by size-exclusion chromatography. Chromatograms showed a single, monodisperse peak indicative of a homogenous sample. It indicated that the samples were suitable for structural applications : electron microscopy and crystallization. Size-exclusion chromatography also allowed the determination of a molecular weight of 480 kDa for the protein-detergent complex in our samples. While the contribution of the detergent in this apparent molecular weight could not be determined with precision, it was strong evidence that the complex is a large multimer.

Samples recovered directly after IMAC and not submitted to TEV cleavage were used to prepare negative-staining EM grids. Micrographs of those grids showed uniform particles approximately 10 nm in diameter. The uniformity of these particles confirmed results from SEC, further arguing for a homogenous sample. The immediate consequence of that observation is that particles can be selected for single-particle reconstructions. From the more than 45 000 particles that were selected in total, 2D

reconstructions of the complex were generated by the joint efforts of myself and my colleague Andrew Cottreau. Those reconstructions argue for a pentameric arrangement of the complex. Five areas of higher density were seen surrounding a central depression, suggesting a hollow or cylindrical structure. The exact composition of each of the five peripheral centers of density is not known, therefore extrapolation from the pentameric arrangement cannot be used to make claims about the stoichiometry of ExbB–ExbD. To our knowledge, however, it is the first glimpse of structural information about proteins ExbB and ExbD as a complex. It is also strong incentive to pursue EM efforts to obtain 3D reconstructions at the highest possible resolution (Section 4.2).

Purified samples of ExbB–ExbD were shown to be monodisperse and homogeneous by SEC and EM. They therefore appeared suitable for crystallization trials. Unfortunately, solubility in the original conditions did not allow concentration of the complex at levels needed for crystallization. Use of 10% glycerol as an additive was shown to promote higher solubility, but highest improvements of the maximum achievable concentration were obtained by screening buffer and pH conditions. The ExbB–ExbD complex was most soluble at basic pH and precipitated readily in more acidic conditions. The immediate consequence of this finding was that it identified favourable conditions for protein concentration. Knowledge about the behavior of the complex in solution will prove useful in crystallization trials, as it provides tools to modulate solubility. Variation of the pH may be used deliberately to change solubility to promote and optimize crystallization. Information from this simple buffer screen may contain information about the complex itself. Indeed, we know that the solubility of a protein is

lowest at its isoelectric point. Because the stoichiometry of the complex is not known, a theoretical isoelectric point cannot be calculated. However, it can be calculated for the individual polypeptides. The values are 7.89 for ExbB and 4.70 for ExbD (114). The poor solubility of the complex at low pH would suggest that its pI is dominated by the ExbD component, potentially contradicting previous reports (84). This observation is also interesting considering that proteins ExbB and ExbD work closely with a protein gradient, the PMF. It can be hypothesized that the high proton concentration at low pH favours a conformation that decreases the solubility of the detergent-solubilized complex. Detergent is another factor that strongly influences the solubility of the complex, but its impact on the pH dependence of solubility is expected to be minimal because of the non-ionic nature of DDM.

The final step in the optimization of ExbB–ExbD preparations permitted the start of crystallization trials. Two similar leads were identified, and they have been optimized to some extent by setting up a number of basic screens. The first lead has identified the use 4% MPD as a precipitant in the presence of magnesium sulfate in HEPES pH 7.5 as conditions leading to rod-shaped crystals of relatively small size. Optimization screens suggest that the presence of $MgSO_4$ is necessary, but lower concentrations tend to lead to fewer and larger crystals. Screening of glycerol concentration also suggests that this additive promotes larger crystals. It should be noted that both MPD and glycerol are alcohols and we hypothesize that this family of compounds facilitates the crystallization of ExbB–ExbD. Our second lead uses polyethylene glycols (PEGs) as precipitants in the presence of magnesium chloride in Tris pH 8.0 to 8.5. This lead has not been

investigated as extensively as the previous, but it has produced the largest ExbB–ExbD crystals obtained to date. The size of those crystals was sufficient to mount them on crystal loops and attempt X-ray diffraction data collection on a homesource. No diffraction was observed from those crystals, indicating poor crystal packing. As with the previous lead, optimization will be necessary. Strategies to improve crystal quality are discussed in Section 4.2. An interesting feature of both leads is the requirement for a magnesium salt. Magnesium, or potentially other divalent cations, appear to be necessary for the crystallization of ExbB–ExbD. Further crystallization attempts will likely require magnesium ions.

One of the many strategies that can be adopted to optimize membrane protein crystallization is to exchange detergent. Exchange to several detergents was attempted on ExbB–ExbD samples. Exchange to DM and to experimental amphiphiles (GNG-12 and MNG-28) provided by a collaborator, was successful. The proteins can be kept at high concentrations in those detergents and crystallization screens were set-up in each of them. Crystals have not been observed in those detergents, but monitoring of the screens continues. The detergent DM, like DDM, belongs to the maltoside family of detergents. Among membrane proteins of known structure, maltosides are the crystallization detergents that are most frequently used. The second largest group of successfully used detergents are the glucosides (94). Exchange to OG, a common glucoside, was attempted: this led to precipitation of the protein. Other families of detergents may be tried. It is with such objective that exchange to experimental

amphiphiles was attempted. Other common families of detergents known to have been successfully used in membrane protein crystallization should also be tried (Section 4.2).

In another attempt at optimizing detergent conditions for crystallization, an NMR strategy for detergent quantitation was adopted. It was tested on the buffer used for crystallization as well as on a post-AEC protein sample. This test showed that the protein-free buffer has a composition that approximates its expected composition. Of note, however, is the much lower detergent content in the protein-containing sample. The cause of that reduction in detergent concentration is unclear, but it might be advantageous in minimizing amounts of detergent, while retaining solubility of the proteins. This strategy for detergent quantitation remains to be exploited fully and on a regular basis (Section 4.2).

4.2 Conclusion and future directions

The work presented here corresponds to the launch of structural studies of the ExbB–ExbD complex. A strategy for the overexpression of the complex in *E. coli* was optimized, leading to the successful and abundant production of ExbB–ExbD. It enabled the design of a purification protocol from which milligram amounts of pure and homogeneous protein complex preparation can be obtained. The samples generated by this approach were used in negative-staining electron microscopy. Under the EM microscope, uniform ExbB–ExbD particles were observed. From several tens of thousand of those particles, 2D reconstructions were generated that show a pentameric and perhaps channel-like structure for the complex. In parallel, the solubility of the complex

was optimized, identifying basic pHs as optimal for the concentration of our samples. The result was pure, homogeneous and concentrated samples suitable for crystallization trials. Initial trials have identified promising leads that require optimization, as diffraction-quality crystals have yet to be produced. In an attempt to optimize our crystals, exchange to several detergents was attempted. Exchange was successful into DM, into GNG-12 and into MNG-28, and crystal screens were set-up in those detergents. Crystals have yet to be observed in those detergents. A detergent quantitation strategy was also attempted to standardize and optimize detergent concentration in ExbB–ExbD samples.

Efforts to obtain structures of the ExbB–ExbD complex are well advanced but will require more experiments and much optimization. The promising 2D reconstructions generated by negative-staining EM mean that generation of 3D reconstructions can be attempted. More data will need to be collected. A tilt-series of images must be acquired to allow 3D reconstruction from negatively stained grids. Provided that these efforts are successful, higher-resolution reconstructions may be generated by cryo-EM strategies. Of particular interest is the prospect of using EM to determine the stoichiometry of ExbB–ExbD. Either ExbB and ExbD can be engineered to carry prominent fusions such as green fluorescent protein or glutathione-S-transferase. Preparations of the complex carrying such prominent fusions can be used in EM experiments like those performed for the native proteins. Comparison of reconstructions with and without fusions would then allow counting of the protruding fusions, therefore leading to determination of stoichiometry.

In parallel, crystallization trials will need to continue. Many avenues may be explored to optimize existing leads, and perhaps to identify new conditions for crystal growth. Detergent exchange has already been mentioned and attempted. Aside from maltosides and glucosides, the amine oxides and the polyoxoethylene glycols are two other detergent families frequently used with success in membrane protein crystallization (94). Lauryldimethylamine-oxide (LDAO) and tetraethylene glycol monooctyl ether (C_8E_4) are common representatives of the amine oxides and polyoxoethylene glycols, respectively. Exchange to these detergents should be attempted, and crystallization trials in these new conditions can be set-up. For all detergents, routine quantification by NMR will ensure optimal and reproducible crystallization conditions.

Aside from detergent exchange, existing leads may be optimized by modulating several factors. The chemical composition of crystallization conditions has been screened and needs to be screened further. Physical factors, in particular temperature, can be exploited to improve crystals. The largest ExbB–ExbD crystals obtained to date grew at 4^{o}C with PEGs and $MgCl_2$. Given the similarity of the MPD/$MgSO_4$ lead with the aforementioned condition, it is not unreasonable to propose that growing at 4^{o}C may promote the growth of better and larger crystals in that chemical environment as well.

Efforts to obtain a structure of the ExbB–ExbD complex have been advanced and are well underway. A 3D EM reconstruction is a prospect of the near future. On the other hand, optimization of existing crystal leads should generate well-diffracting

crystals suitable for data collection at a synchrotron. That data, in turn, will allow solving of a high-resolution structure for the complex.

References

1. Murray, P., Rosenthal, K., and Pfaller, M. (2009) *Medical Microbiology*, 6th ed., Mosby Elsevier, Philadelphia, PA
2. Ruiz, N., Kahne, D., and Silhavy, T. (2009) *Nat Rev Microbiol* **7**, 677-683
3. Beveridge, T. (1999) *J Bacteriol* **181**, 4725-4733
4. Silhavy, T., Kahne, D., and Walker, S. (2010) *Cold Spring Harb Perspect Biol* **2**, a000414
5. Fecycz, I., and Campbell, J. (1985) *Eur J Biochem* **146**, 35-42
6. Kessler, E., and Safrin, M. (1988) *J Bacteriol* **170**, 5241-5247
7. Meroueh, S., Bencze, K., Hesek, D., Lee, M., Fisher, J., Stemmler, T., and Mobashery, S. (2006) *Proc Natl Acad Sci U S A* **103**, 4404-4409
8. Ni, Y., Reye, J., and Chen, R. R. (2007) *Biotechnol Bioeng* **97**, 1347-1356
9. Dmitriev, B., Toukach, F., and Ehlers, S. (2005) *Trends Microbiol* **13**, 569-574
10. Raetz, C., and Dowhan, W. (1990) *J Biol Chem* **265**, 1235-1238
11. Nikaido, H. (2003) *Microbiol Mol Biol Rev* **67**, 593-656
12. Randall-Hazelbauer, L., and Shwartz, M. (1973) *J. Bacteriol.* **116**, 1436-1446
13. Schirmer, T., Keller, T., Wang, Y., and Rosenbusch, J. (1995) *Science* **267**, 512-514
14. Forst, D., Welte, W., Wacker, T., and Diederichs, K. (1998) *Nat Struct Biol* **5**, 37-46
15. van den Berg, B., Black, P., Clemons, W. J., and Rapoport, T. (2004) *Science* **304**, 1506-1509
16. van den Berg, B. (2005) *Curr Opin Struct Biol* **15**, 401-407
17. Hearn, E., Patel, D., Lepore, B., Indic, M., and van den Berg, B. (2009) *Nature* **458**, 367-370
18. Harris, W. R. (2002) *Molecular and Cellular Iron Transport*, Marcel Dekker, Inc., New York, NY
19. Prinz, T., Meyer, M., Pettersson, A., and Tommassen, J. (1999) *J Bacteriol* **181**, 4417-4419
20. Wandersman, C., and Stojiljkovic, I. (2000) *Curr Opin Microbiol* **3**, 215-220
21. Cascales, E., Buchanan, S., Duché, D., Kleanthous, C., Lloubès, R., Postle, K., Riley, M., Slatin, S., and Cavard, D. (2007) *Microbiol Mol Biol Rev* **71**, 158-229
22. Killmann, H., and Braun, V. (1994) *FEMS Microbiol Lett* **119**, 71-76
23. Ferguson, A., Hofmann, E., Coulton, J., Diederichs, K., and Welte, W. (1998) *Science* **282**, 2215-2220
24. Locher, K., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J., and Moras, D. (1998) *Cell* **95**, 771-778
25. Buchanan, S., Smith, B., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J. (1999) *Nat Struct Biol* **6**, 56-63
26. Ferguson, A., Chakraborty, R., Smith, B., Esser, L., van der Helm, D., and Deisenhofer, J. (2002) *Science* **295**, 1715-1719
27. Yue, W., Grizot, S., and Buchanan, S. (2003) *J Mol Biol* **332**, 353-368
28. Chimento, D., Kadner, R., and Wiener, M. (2003) *J Mol Biol* **332**, 999-1014

29. Buchanan, S., Lukacik, P., Grizot, S., Ghirlando, R., Ali, M., Barnard, T., Jakes, K., Kienker, P., and Esser, L. (2007) *EMBO J* **26**, 2594-2604
30. Cobessi, D., Celia, H., Folschweiller, N., Schalk, I., Abdallah, M., and Pattus, F. (2005) *J Mol Biol* **347**, 121-134
31. Cobessi, D., Celia, H., and Pattus, F. (2005) *J Mol Biol* **352**, 893-904
32. Brillet, K., Meksem, A., Lauber, E., Reimmann, C., and Cobessi, D. (2009) *Acta Crystallogr D Biol Crystallogr* **65**, 326-331
33. Krieg, S., Huché, F., Diederichs, K., Izadi-Pruneyre, N., Lecroisey, A., Wandersman, C., Delepelaire, P., and Welte, W. (2009) *Proc Natl Acad Sci U S A* **106**, 1045-1050
34. Cobessi, D., Meksem, A., and Brillet, K. (2010) *Proteins* **78**, 286-294
35. Koch, A. (1998) *Crit Rev Microbiol* **24**, 23-59
36. Schramm, E., Mende, J., Braun, V., and Kamp, R. (1987) *J Bacteriol* **169**, 3350-3357
37. Pawelek, P., Croteau, N., Ng-Thow-Hing, C., Khursigara, C., Moiseeva, N., Allaire, M., and Coulton, J. (2006) *Science* **312**, 1399-1402
38. Shultis, D., Purdy, M., Banchs, C., and Wiener, M. (2006) *Science* **312**, 1396-1399
39. Lohmiller, S., Hantke, K., Patzer, S., and Braun, V. (2008) *Microbiology* **154**, 1748-1754
40. Neugebauer, H., Herrmann, C., Kammer, W., Schwarz, G., Nordheim, A., and Braun, V. (2005) *J Bacteriol* **187**, 8300-8311
41. Schauer, K., Gouget, B., Carriere, M., Labigne, A., and de Reuse, H. (2007) *Mol Microbiol* **63**, 1054-1068
42. Blanvillain, S., Meyer, D., Boulanger, A., Lautier, M., Guynet, C., Denance, N., Vasse, J., Lauber, E., and Arlat, M. (2007) *PLoS One* **2**, e224
43. Schauer, K., Rodionov, D., and de Reuse, H. (2008) *Trends Biochem Sci* **33**, 330-338
44. Quiocho, F., and Ledvina, P. (1996) *Mol Microbiol* **20**, 17-25
45. Borths, E., Locher, K., Lee, A., and Rees, D. (2002) *Proc Natl Acad Sci U S A* **99**, 16642-16647
46. Davidson, A., and Chen, J. (2004) *Annu Rev Biochem* **73**, 241-268
47. Locher, K., Lee, A., and Rees, D. (2002) *Science* **296**, 1091-1098
48. Hvorup, R., Goetz, B., Niederer, M., Hollenstein, K., Perozo, E., and Locher, K. (2007) *Science* **317**, 1387-1390
49. Newstead, S., Fowler, P. W., Bilton, P., Carpenter, E. P., Sadler, P. J., Campopiano, D. J., Sansom, M. S., and Iwata, S. (2009) *Structure* **17**, 1213-1222
50. Oldham, M. L., Khare, D., Quiocho, F. A., Davidson, A. L., and Chen, J. (2007) *Nature* **450**, 515-521
51. Ward, A., Reyes, C. L., Yu, J., Roth, C. B., and Chang, G. (2007) *Proc Natl Acad Sci U S A* **104**, 19005-19010
52. Carter, D., Miousse, I., Gagnon, J., Martinez, E., Clements, A., Lee, J., Hancock, M., Gagnon, H., Pawelek, P., and Coulton, J. (2006) *J Biol Chem* **281**, 35413-35424

53. James, K., Hancock, M., Gagnon, J., and Coulton, J. (2009) *Biochemistry* **48**, 9212-9220
54. Jormakka, M., Byrne, B., and Iwata, S. (2003) *FEBS Lett* **545**, 25-30
55. Terashima, H., Kojima, S., and Homma, M. (2008) *Int Rev Cell Mol Biol* **270**, 39-85
56. Bradbeer, C. (1993) *J Bacteriol* **175**, 3146-3150
57. Braun, V., Gaisser, S., Herrmann, C., Kampfenkel, K., Killmann, H., and Traub, I. (1996) *J Bacteriol* **178**, 2836-2845
58. Higgs, P. I., Myers, P. S., and Postle, K. (1998) *J Bacteriol* **180**, 6031-6038
59. Ollis, A. A., Manning, M., Held, K. G., and Postle, K. (2009) *Mol Microbiol* **73**, 466-481
60. Braun, V. (1995) *FEMS Microbiol Rev* **16**, 295-307
61. Lazzaroni, J., Germon, P., Ray, M., and Vianney, A. (1999) *FEMS Microbiol Lett* **177**, 191-197
62. Lazzaroni, J., Dubuisson, J., and Vianney, A. (2002) *Biochimie* **84**, 391-397
63. Eick-Helmerich, K., and Braun, V. (1989) *J Bacteriol* **171**, 5117-5126
64. Braun, V., and Herrmann, C. (1993) *Mol Microbiol* **8**, 261-268
65. Brinkman, K., and Larsen, R. (2008) *J Bacteriol* **190**, 421-427
66. Braun, V. (1989) *J Bacteriol* **171**, 6387-6390
67. Zhai, Y., Heijne, W., and Saier, M. J. (2003) *Biochim Biophys Acta* **1614**, 201-210
68. Cascales, E., Lloubès, R., and Sturgis, J. (2001) *Mol Microbiol* **42**, 795-807
69. Postle, K. (1990) *Mol Microbiol* **4**, 2019-2025
70. Larsen, R., Wood, G., and Postle, K. (1994) *Mol Microbiol* **12**, 857
71. Khursigara, C., De Crescenzo, G., Pawelek, P., and Coulton, J. (2005) *Protein Sci* **14**, 1266-1273
72. Chang, C., Mooser, A., Plückthun, A., and Wlodawer, A. (2001) *J Biol Chem* **276**, 27535-27540
73. Khursigara, C., De Crescenzo, G., Pawelek, P., and Coulton, J. (2004) *J Biol Chem* **279**, 7405-7412
74. Ködding, J., Killig, F., Polzer, P., Howard, S., Diederichs, K., and Welte, W. (2005) *J Biol Chem* **280**, 3022-3028
75. Peacock, R., Weljie, A., Howard, S., Price, F., and Vogel, H. (2005) *J. Mol. Biol.* **345**, 13
76. Kampfenkel, K., and Braun, V. (1993) *J Biol Chem* **268**, 6050-6057
77. Nugent, T., and Jones, D. (2009) *BMC Bioinformatics* **10**, 159
78. Kampfenkel, K., and Braun, V. (1992) *J Bacteriol* **174**, 5485-5487
79. Garcia-Herrero, A., Peacock, R., Howard, S., and Vogel, H. (2007) *Mol Microbiol* **66**, 872-889
80. Parsons, L. M., Grishaev, A., and Bax, A. (2008) *Biochemistry* **47**, 3131-3142
81. Larsen, R., Thomas, M., Wood, G., and Postle, K. (1994) *Mol Microbiol* **13**, 627-640
82. Larsen, R., and Postle, K. (2001) *J Biol Chem* **276**, 8111-8117
83. Braun, V., and Herrmann, C. (2004) *J Bacteriol* **186**, 4402-4406
84. Higgs, P., Larsen, R., and Postle, K. (2002) *Mol Microbiol* **44**, 271-281
85. Postle, K., and Larsen, R. (2007) *Biometals* **20**, 453-465

86. Stevens, T., and Arkin, I. (2000) *Proteins* **39**, 417-420
87. White, S. Membrane Proteins of Known Structure.
88. Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985) *Nature* **318**, 618-624
89. Newby, Z., O'Connell, J. r., Gruswitz, F., Hays, F., Harries, W., Harwood, I., Ho, J., Lee, J., Savage, D., Miercke, L., and Stroud, R. (2009) *Nat Protoc* **4**, 619-637
90. Nagamori, S., Smirnova, I. N., and Kaback, H. R. (2004) *J Cell Biol* **165**, 53-62
91. Wagner, S., Bader, M. L., Drew, D., and de Gier, J. W. (2006) *Trends Biotechnol* **24**, 364-371
92. Arechaga, I., Miroux, B., Karrasch, S., Huijbregts, R., de Kruijff, B., Runswick, M. J., and Walker, J. E. (2000) *FEBS Lett* **482**, 215-219
93. Garavito, R. M., and Ferguson-Miller, S. (2001) *J Biol Chem* **276**, 32403-32406
94. Newstead, S., Ferrandon, S., and Iwata, S. (2008) *Protein Sci* **17**, 466-472
95. Chayen, N., and Saridakis, E. (2008) *Nat Methods* **5**, 147-153
96. Zhou, Y., Morais-Cabral, J. H., Kaufman, A., and MacKinnon, R. (2001) *Nature* **414**, 43-48
97. Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G., Thian, F. S., Kobilka, T. S., Choi, H. J., Yao, X. J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) *Science* **318**, 1266-1273
98. Johansson, L. C., Wohri, A. B., Katona, G., Engstrom, S., and Neutze, R. (2009) *Curr Opin Struct Biol* **19**, 372-378
99. Ujwal, R., Cascio, D., Chaptal, V., Ping, P., and Abramson, J. (2009) *Channels (Austin)* **3**, 167-170
100. Subramaniam, S., and Milne, J. L. (2004) *Annu Rev Biophys Biomol Struct* **33**, 141-155
101. Fujiyoshi, Y. (1998) *Adv Biophys* **35**, 25-80
102. Peleg, A. Y., and Hooper, D. C. (2010) *N Engl J Med* **362**, 1804-1813
103. French, G. L. (2005) *Adv Drug Deliv Rev* **57**, 1514-1527
104. English, B. K., and Gaur, A. H. (2010) *Adv Exp Med Biol* **659**, 73-82
105. Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007) *J Struct Biol* **157**, 38-46
106. Sorzanoa, C. O. S., c, Marabinia, R., b, Velázquez-Muriela, J., Bilbao-Castroa, J. R., Scheresa, S. H. W., Carazoa, J. M., b, and Pascual-Montanoa, A. *Journal of Structural Biology* **148**
107. Chen, J. Z., and Grigorieff, N. (2007) *J Struct Biol* **157**, 168-173
108. Luft, J., Collins, R., Fehrman, N., Lauricella, A., Veatch, C., and DeTitta, G. (2003) *J Struct Biol* **142**, 170-179
109. Koszelak-Rosenblum, M., Krol, A., Mozumdar, N., Wunsch, K., Ferin, A., Cook, E., Veatch, C., Nagel, R., Luft, J., Detitta, G., and Malkowski, M. (2009) *Protein Sci* **18**, 1828-1839
110. Maslennikov, I., Kefala, G., Johnson, C., Riek, R., Choe, S., and Kwiatkowski, W. (2007) *BMC Struct Biol* **7**, 74
111. Maslennikov, I., Krupa, M., Dickson, C., Esquivies, L., Blain, K., Kefala, G., Choe, S., and Kwiatkowski, W. (2009) *J Struct Funct Genomics* **10**, 25-35

112. Drew, D., Slotboom, D. J., Friso, G., Reda, T., Genevaux, P., Rapp, M., Meindl-Beinker, N. M., Lambert, W., Lerch, M., Daley, D. O., Van Wijk, K. J., Hirst, J., Kunji, E., and De Gier, J. W. (2005) *Protein Sci* **14**, 2011-2017
113. Fischer, E., Gunter, K., and Braun, V. (1989) *J Bacteriol* **171**, 5127-5134
114. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. in *The Proteomics Protocols Handbook* (Walker, J. M. ed.), Humana Press, Totowa, NJ. pp 571-607