Running title: Functional and structural characterization of *Serratia marcescens* ExbB: determinants of the interaction with HasB/TonB

Valérie Biou¹, Ricardo Jorge Diogo Adaixo², Mohamed Chami², Pierre-Damien Coureux³, Benoist Laurent⁴, Yvette Ntsogo¹, Gisele Cardoso de Amorim⁵, Nadia Izadi-Pruneyre⁵, Christian Malosse⁶, Julia Chamot-Rooke⁶, Henning Stahlberg², Philippe Delepelaire¹

1: Laboratoire de Biologie Physico-Chimique des Protéines Membranaires, Université de Paris, UMR 7099 CNRS, F-75005, Paris, France.

Institut de Biologie Physico-Chimique, F-75005, Paris, France.

2: Center for Cellular Imaging and NanoAnalytics, Biozentrum, University of Basel, Mattenstrasse 26, CH-4058, Basel, Switzerland

3: Laboratoire de Biologie Structurale de la Cellule, BIOC, UMR7654 CNRS/Ecole polytechnique, Palaiseau France

4: Plateforme de Bioinformatique, Université de Paris, FRC 550 CNRS, F-75005, Paris, France. Institut de Biologie Physico-Chimique, F-75005, Paris, France.

5: Structural Bioinformatics Unit, Department of Structural Biology and Chemistry, C3BI, Institut Pasteur, CNRS UMR3528, CNRS USR3756, Paris, France.

6: Mass Spectrometry for Biology Unit, CNRS USR 2000, Institut Pasteur, 75015 Paris, France

For correspondence: valerie.biou@ibpc.fr, philippe.delepelaire@ibpc.fr

Summary.

ExbBD is part of a cytoplasmic membrane molecular motor driven by the proton-motive force. It belongs to the larger family of motors involved in nutriment import across the outer membrane of Gram-negative bacteria (ExbBD), flagellar rotation (MotAB) or late steps of cell division in Gram-negative bacteria (TolQR). ExbB and ExbD are integral membrane proteins with three (ExbB) or one (ExbD) transmembrane segment. Here we have solved by single-particle cryo-EM the structures of ExbB alone and of the ExbB-ExbD complex of the opportunistic pathogen *Serratia marcescens*. ExbB_{Sm} alone behaves as a stable pentamer, and the complex displays the ExbB₅-ExbD₂ stoichiometry. This is similar to what has been observed for ExbB-ExbD complexes from *Escherichia coli* and *Pseudomonas savastanoi* as well as MotAB complexes from various species. We identified residues located in the first TM of ExbB_{Sm} and ExbB_{Ec} that are likely involved in the interaction with TonB/HasB and that are essential for function. ExbB_{Sm} has a ca. 40 residues long periplasmic extension absent in *E. coli*. Such long ExbB's are found in some Gammaproteobacteria, and several genera of Alphaproteobacteria. We show that this extension interacts with HasB, a dedicated TonB paralog from the heme acquisition system (Has) from *S. marcescens*. We also show that it is involved in heme

acquisition *via* the Has system from *S. marcescens*. ExbB_{sm} represents thus a new class of ExbB protein and our results shed light on the specificity determinants between the ExbB-ExbD complex and their associated TonB partners.

Introduction

Transport of nutrients across the Gram-negative outer membrane is either a diffusion-facilitated or an active process, powered in that last case by the energy of the proton-motive force (pmf) transmitted to specialized outer membrane transporters (the so-called TBDT's, TonB dependent transporters). A complex of three cytoplasmic membrane proteins, TonB, ExbB and ExbD, that together form the TonB complex (see [1] for a review) conveys the energy of the pmf to the TBDT. In *Escherichia coli* K12, there is only one set of *tonB*, *exbB* and *exbD* genes, whereas there are nine TBDT's [2], all energized by the same complex. ExbB and ExbD respectively belong to the MotA-TolQ-ExbB and MotB-TolR-ExbD protein families, that are involved in power generation and transmission in various processes (flagellar rotation for MotAB, active transport of molecules across OM receptors for ExbBD and cell division for TolQR), at the expense of pmf dissipation across the cytoplasmic membrane [3-6]. Those proteins form complexes that associate with TonB (ExbBD), TolA (TolQR) or the flagellar rotor (MotAB, stator of the flagellar motor). It has also been shown that ExbD does not accumulate in the absence of ExbB, and that TonB does not accumulate either in the absence of the ExbBD complex [7]. It is also known that the C-terminal domains of TonB and ExbD interact with each other [8, 9]. ExbD TM has one essential strictly conserved aspartate residue [10], that is thought to undergo cycles of protonation/deprotonation (coupled to pmf dissipation).

TBDT's comprise a 22-stranded β -barrel anchored in the outer membrane. An N-terminal domain folded inside the barrel and called the plug contains the main region of interaction with TonB, the so-called TonB box located in the N-terminal periplasm-exposed part of TBDT. TonB has its N-terminus localized in the cytoplasm, a single transmembrane segment, followed by a Pro-Lys rich region long enough to span the periplasm and a structured C-terminal domain interacting with the TBDT. Conformational changes elicited upon interaction between TonB and the TonB box of the TBDT, and likely mediated via a rearrangement of the plug domain, ultimately lead to entry of the substrate (initially bound to the extracellular side of the receptor) into the periplasm by a yet unknown mechanism.

In *Serratia marcescens*, a close relative of *E. coli*, in which were identified about twenty potential TBDT's, there are at least two TonB homologs: an ortholog of TonB_{Ec} (having 48% identity with TonB_{Ec}[11]) and HasB, a TonB paralog that is strictly dedicated to its cognate outer membrane receptor HasR[12, 13]. The Has (heme acquisition system) system includes the TBDT HasR, which recognizes both free heme and the high-affinity extracellular heme-binding HasA hemophore. HasB is encoded within the *has* locus and displays low sequence identity with either *E. coli* or *S. marcescens* TonB. Unlike TonB_{Sm}, HasB cannot complement TonB_{Ec} functions, nor it is able to drive heme entry *via* the HasR receptor in the presence of ExbBD_{Ec}. A gain of function mutation in the TM domain of HasB was however isolated in *E. coli*, restoring heme entry *via* HasR in the presence of ExbBD_{Ec} [12]. HasB has the same topology as TonB, and its C-terminal domain interacts specifically with HasR with a ca. 40-fold higher affinity than the corresponding TonB domain [14]. Those observations prompted us to identify the ExbBD complex that would function with HasB in *S. marcescens*. Hence, we first identified the resident orthologous ExbBD_{Sm} and found that this complex is active with both HasB and TonB.

Recently several structures of *E. coli* ExbB and ExbBD have been solved, either by X-ray crystallography or cryo-EM[15-17]. All samples studied included ExbB and ExbD, but not all showed an ordered structure of ExbD. In these structures, ExbB_{Ec}[15] appeared as a pentamer, each monomer of ExbB having three TM helices that extend into the cytoplasm, and with a highly polarized charge distribution on the cytoplasmic side. The central pore is apolar, lined by TM helices 2 and 3 of each monomer, creating a large hydrophobic cavity inside the structure. Another study using X-ray crystallography, single-particle cryo-EM and electron diffraction on two-dimensional crystals concluded that ExbB could undergo a pH-dependent pentamer to hexamer transition, and that the hexameric ExbB could accommodate three ExbD TM segments in its pore[16]. A more recent cryo-EM study of ExbBD reconstituted in nanodiscs however confirmed the ExbB₅ExbD₂ stoichiometry consistent with previous DEER results and identified the two TM helices of ExbD in the ExbB pore [16]. Several structures of the flagellar MotAB motor complexes (related to ExbBD) from diverse bacteria (*Campylobacter jejuni*, *Bacillus subtilis*, *Clostridium sporogenes*, *Vibrio mimicus*, *Shewanella oneidensis*, PomAB from *Vibrio alginolyticus*) were recently published that also display a MotA₅MotB₂ stoichiometry with some common topology elements [18, 19]. Similarly, *Pseudomonas savastanoi* ExbBD exhibits the same stoichiometry.

Here we identify a new family of ExbB proteins with a long N-terminal extension whose presence is strongly correlated to the presence of a *hasB* gene ortholog in the genome. We characterized ExbB_{Sm} alone and ExbBD_{Sm} by single-particle cryo-EM and determined their structures at 3.2 and 3.9Å, respectively. We show that in both cases, ExbB_{Sm} behaves as a stable pentamer; in the ExbBD_{Sm} complex, we identify two TM helices of ExbD_{Sm} in the pore of ExbB pentamer, as it was shown for the *E. coli* case. Using NMR measurements, we also show that the N-terminal periplasmic extension specific to ExbB_{Sm} interacts with the globular domain of HasB. Finally, *via* mutagenesis studies and bacterial growth assays we show that the first transmembrane helix TM1 of ExbB contains specificity determinants for interaction with HasB/TonB and that the periplasmic extension does play a role in heme acquisition *via* the *has* system.

Results

1. Orthologs of *E. coli* *exbB* and *exbD* in *Serratia marcescens* define a new ExbB family with a N-terminal extension.

Sequence analysis of strain Db11, a fully sequenced *S. marcescens* isolate (GenBank: *Serratia marcescens* subsp. *marcescens* Db11, complete genome. ACCESSION HG326223, <https://www.ncbi.nlm.nih.gov/nuccore/HG326223.1>), identified one putative operon encoding orthologs of *E. coli* ExbB and ExbD (SMDB11_3479(ExbD) and SMDB11_3480(ExbB)). In this case the *exbB*-like gene is close to and in the opposite orientation to a *metC* homolog, as in *E. coli*. There is no conservation on the 3' end. Comparison of the coding sequences indicated that the *exbD* gene encoded a 140 residue-long protein (71% identity with the 141 residue-long *E. coli* sequence). In contrast, the putative *exbB* gene encodes a much longer protein than its *E. coli* counterpart (325 residues instead of 244 residues in *E. coli*, 73% identity in the common part). The additional stretch of residues present in ExbB_{Sm} is located at the N-terminus and corresponds to a putative signal peptide followed by a ~50 residues N-terminal extension of the mature protein (Supplementary Figure S1).

A BLAST search on bacterial genomes indicated that such “extralong” ExbB’s are found in different Gammaproteobacteria, including *Serratia*, *Yersinia*, *Pseudomonas*, *Erwinia*/Dickeya, and many genera in the Alphaproteobacteria class. Table SI lists representative species of Alphaproteobacteria in which such long ExbB’s are found. About 90% of the long ExbB-containing species listed in this table also have a *hasB* gene ortholog.

An aminoacid sequence alignment of these ExbBs (Clustal omega <https://www.ebi.ac.uk/Tools/msa/clustalo/>, Figure 1) also shows that the mature N-terminal extensions are of variable length between a few and 150 residues with an average of 50 residues. These extensions are quite rich in Ala (24,6%) and Pro (14,5%) residues and therefore likely to be unstructured but could be involved in protein-protein interactions as proline-rich regions often are in signalling processes [20-22].

The Conserved Domain Database [23] contains currently two ExbB subfamilies in the cl00568 MotA_ExbB superfamily, TIGR02797 (containing *E. coli* ExbB, as well as longer ExbB's) and TIGR02805 (containing *Haemophilus influenzae* ExbB, with a very short cytoplasmic domain between TM1 and TM2); based on sequence data, it is then possible to subdivide TIGR02797 into 2 groups with "extralong" ExbB's in one subfamily, and *E. coli*-like sequences in the other.

2. ExbBD_{Sm} complements ExbBD_{Ec} for iron acquisition and functionally associates with both TonB and HasB for heme acquisition through HasR

As a first step in the characterization of the identified ExbBD_{Sm}, we tested whether this complex would complement ExbBD_{Ec}. To this end, the plasmids pBADExbBD_{Sm} and pBADExbBD_{Ec} were constructed, introducing the two genes, from *S. marcescens* and *E. coli*, into pBAD24 vector under the control of p_{araBAD} promoter. By using growth under iron starvation conditions as a test, we could show that *exbBD_{Sm}* complements *exbBD_{Ec}*. The *E. coli* C600Δ*exbBD* strain is indeed more sensitive than its wild-type counterpart to iron starvation (induced for example by the Fe²⁺ chelator dipyrindyl) as ExbBD are required to transduce pmf to TonB for siderophore uptake. Its growth in the presence of dipyrindyl is restored by the expression of either pBADExbBD_{Ec} or pBADExbBD_{Sm} (Supplementary Figure S2).

HasB is a HasR-specific TonB ortholog, and as ExbBD_{Ec} does not function with HasB (for heme acquisition through HasA-HasR [12]), we then verified that ExbBD_{Sm} would do so. To this end, pBADExbBD_{Sm} was introduced into strain C600Δ*hemA*Δ*tonB*Δ*exbBD* harbouring the phasISRDEB/238 recombinant plasmid containing the whole *has* locus with its own regulatory elements. This strain could grow in the presence of heme (down to a concentration of 0.2μM) and arabinose, whereas the C600Δ*hemA*Δ*tonB*Δ*exbBD*(phasISRDEB/238 + pBADExbBD_{Ec}) could not (see Supplementary Figure S3 bottom part). The strain C600Δ*hemA*Δ*exbBD*(phasISRDEB/238 + pBADExbBD_{Ec}) could grow in the presence of heme and arabinose, indicating that pBADExbBD_{Ec} is functional (see Supplementary Figure S3 top

part). Likewise *C600ΔhemAΔexbBD*(phasISRADE/238 + pBADExbBD_{Sm}) could grow in the presence of heme and arabinose. This indicates that ExbB_{Sm} and ExbD_{Sm} functionally associate with either HasB or TonB_{Ec} to allow HasA-mediated heme entry *via* HasR, whereas ExbBD_{Ec} only functions with TonB leading to the heme entry *via* the HasA-HasR system (see Supplementary Figure S3 top part). As already observed, a stronger iron starvation is necessary when TonB is the link between the inner and outer membrane components [24]. Therefore, ExbBD_{Sm} is the genuine *E. coli* ortholog, with a looser specificity.

3. Characterization of ExbB_{Sm} and ExbBD_{Sm} and their structural analysis by cryo-EM show ExbB₅ and ExbB₅D₂ stoichiometries

To gain further insight into the possible differences between ExbBD_{Sm} and ExbBD_{Ec}, we purified both ExbB_{Sm} and ExbBD_{Sm} (as His-tagged proteins, see Materials and Methods section) in LMNG micelles (Supplementary Figure S4). Both ExbB_{Sm} and ExbBD_{Sm} appear as homogeneous oligomers at the last size exclusion chromatography (SEC) step. Mass spectrometry analysis of intact ExbD_{Sm} (16 131 Da) and ExbB_{Sm} (29 557 Da) proteins, together with the determination of the first amino-acid residues for ExbB (APAAN) showed that the predicted signal sequence of ExbB_{Sm} is absent from the mature sequence and has therefore been processed. It also showed that ExbD initial Met residue was cleaved off (see Supplementary Figure S5). More interestingly, mass spectrometry analysis of chloroform/methanol extracts of the purified proteins also showed that both complexes contained native lipids (Supplementary Figure S6), mostly phosphatidylglycerol (PG) and phosphatidylethanolamine (PE), and that ExbBD also contained some cardiolipin (CL) (Supplementary Figure S7). Further analysis of aliphatic chain composition of lipids shows evidence of specific composition (Supplementary Figure S8).

3a ExbB structure

As ExbB_{Sm} presented specific features compared to ExbB_{Ec}, we set out to determine its structure by single-particle cryo-EM. The 3D class average model clearly showed a pentameric structure. It was refined and polished to obtain a resolution of 3.2Å. Maps were refined with and without C5 symmetry and showed a 98% correlation. Therefore, the map used for model building was the one with C5 symmetry.

The structure solved here and shown in Figure 2 A and B, is very similar to that of *E. coli* ExbB 5SV0 structure (that was co-purified with ExbD), with a 1.3Å rmsd over all Cα atoms (Table I). This shows that ExbB is stable as a pentamer on its own. The periplasmic N-terminal extension

did not yield any visible electron density, precluding its structure determination. At the C-terminus of each monomer, however, electron density was clearly defined, allowing structure determination for an additional 10 residues in a helical conformation up to the last one (Gly 283, see Figure 2D, helix $\alpha 8$), before the tag. In the 5SV0 X-ray structure, a calcium ion is bound to the five Glu 106 (of helix $\alpha 2$ in the TM region) on the cytoplasmic side. In ExbB_{Sm}, this residue is replaced by an Asn, and we do not observe any density that could be assigned to a metal ion.

The structure of the whole pentamer with one monomer coloured as a function of sequence conservation in this sub-class of ExbB's (Supplementary figure 9) shows that the highest residue conservation is inside the transmembrane channel, indicative of high functional constraints, while the TM residues located at the membrane surface are more variable. The same observation was made for those ExbB's without N-terminal signal sequence and periplasmic extension [15].

Similarly to the ExbB_{EC} structure, at the cytoplasmic junction with the membrane there are “fenestrations”, potentially leading to solvent or ion passage [15]. These vents are located at the junction of adjacent monomers as shown in figure 2A. Additional density was also clearly observed on the external surface, which we attributed to the phospholipids present in the preparation. Each ExbB monomer appears to be associated with the equivalent of two PG molecules, located at the inner leaflet of the cytoplasmic membrane (Figure 2A and C). However, we were able to model only one PG molecule per monomer with confidence. Three PE (phosphatidylethanolamine, the major *E. coli* phospholipid) molecules and one PG molecule were identified as associated to the ExbBD pentamer in the *E. coli* complex after reconstitution in nanodiscs [17]. In summary, our cryo-EM structure shows that ExbB from *S. marcescens* is stable as a pentamer and associates strongly with specific phospholipids coming from the inner leaflet of the membrane.

3b ExbBD_{Sm} structure

We also solved the structure of the ExbBD_{Sm} complex by cryo-EM (Figure 3 and Table I). As for ExbB, the purified complex exhibited a symmetric peak on SEC (see Supplementary Figure S4). Cryo-EM data were collected and processed as described in the Material and Methods section. 2D classes showed preferential orientations with 92% top views and only 8% side views. Due to this very strong bias in particle orientation, the resolution achieved was 3.9 Å, precluding

reliable positioning of the side chains. As observed for ExbB alone, in the ExbBD complex, ExbB behaves as a pentamer and no sign of other assemblies was found. Similarly, no density could be attributed to the periplasmic extension. There are two clear densities inside the ExbB pore that were assigned to the pair of ExbD TM segments (Figure 3 B and C). The charge distribution is highly polarised on the cytoplasmic side: the region close to the membrane is positively charged as expected and the distal part is negatively charged (Supplementary Figure S10, top left). The central pore is apolar, lined by TM helices 2 and 3 of each monomer, creating a large hydrophobic cavity inside the structure (Supplementary Figure S10, top right). In the refined model, Asp 25 from ExbD monomer chain G faces Thr 218 from ExbB chain C, while Asp 25 from the ExbD chain F faces the interface between two ExbB monomers A and E (Supplementary Figure S10, bottom). The estimation of pKa by using the program Propka [25] server (<http://server.poissonboltzmann.org/>) shows that Asp 25 has a pKa of 7.3 for chain F and 7.42 for chain G, both very peculiar for solvent-accessible acidic residues but that can be found in buried active sites or membrane proteins [26]. This pKa should allow protonation and deprotonation of ExbD Asp 25 at physiological pH.

As compared to the structure of ExbB_{Sm}, there is a slight «opening» towards the periplasmic side (the distance of Ala 197 from one subunit to Leu 204 of the facing subunit varying from 25.5Å in ExbB to 29.8Å in ExbBD). This opening of the structure is limited to the periplasmic part (Supplementary Figure S11). The RMSD between ExbB_{Sm} alone and in complex with ExbD_{Sm} is 2Å over all Cα of the pentamer, and the RMSD between ExbB_{Sm} in complex with ExbD_{Sm} and ExbB_{Ec} in complex with ExbD_{Ec} (structure 6TYI chains A-E) is 1.7Å.

As compared to the 6TYI structure of *E. coli* ExbBD complex, there are some significant differences that could be related to the specificity of function of ExbB_{Sm}. Both ExbB_{Sm} and ExbBD_{Sm} inner pores are slightly wider than their *E. coli* counterparts at the periplasmic entrance. Consequently, two channels cross the membrane region, (extending from the periplasmic entrance to the Asp 25 residue from the ExbD TM deeply embedded in the ExbB pore) that are clearly seen in the ExbBD_{Sm} (as detected by the MoleOnline Server) while the *E. coli* structure exhibits a single, much thinner channel as detected with the same parameters of MoleOnline (Supplementary Figure S12). It is therefore possible that this structure represents another state of the ExbBD motor complex, whereby the periplasmic side allows solvent access to the Asp 25 residue of ExbD TM. In summary, we show that ExbBD from *S.*

marcescens has a 5:2 stoichiometry with a larger channel allowing solvent or proton transport from the periplasmic side as compared with the *E. coli* structure.

4. Specific functions of the N-terminal extension and of the first TM segment of ExbB_{Sm}

4a NMR shows an interaction between HasB and ExbB_{Sm} periplasmic regions.

The N-terminal extension of ExbB_{Sm} contains almost 60% of Ala and Pro residues and is predicted to be unstructured. In agreement with this, there was no observable electron density for this region neither in the free form nor in the complex. Indeed, NMR analysis of a 44 residue-long synthetic peptide corresponding to this extension (ExbB_{Sm} 1-44) did not show any indication of secondary structure (unpublished results). Since this region is predicted to be in the periplasm where the C-terminal domain of HasB, HasB_{CTD} is also located, we investigated the interaction of the ExbB_{Sm} 1-44 synthetic peptide with HasB_{CTD} by NMR. The analysis of chemical shift perturbations (CSP) of amide resonances of HasB_{CTD} upon addition of the peptide, showed that the chemical environment of their corresponding residues was modified as a consequence of their interaction (Supplementary Figure S13). The residues showing the highest chemical shift perturbations are mapped on HasB_{CTD} structure 2M2K in Figure 4. They are mainly located on the helical face of HasB_{CTD} forming a continuous surface of interaction. An additional small pocket at the C-terminus of the protein is also involved in this interaction.

4b An ExbB_{Sm} mutant devoid of the N-terminal extension is affected in its function

To assess the potential function of the N-terminal periplasmic extension of ExbB_{Sm}, we engineered a mutant lacking this extension (see Materials and Methods for details) and tested its effect on the growth of a C600Δ*hemA*Δ*exbBD*Δ*tonB* strain harbouring plasmids pAMhasISRADEB and pBADexbBD_{Sm} or its derivatives. To get high sensitivity and numerical data in our growth tests, we switched to growth in liquid medium in 48-well plates (see Materials and Methods for details). We could then show that the lag between the start of the experiment and the onset of growth is delayed by about 2 hours, for the mutant strain (in which ExbB is devoid of its periplasmic extension), as compared to the WT strain. In addition, the maximum OD reached after growth is slightly but significantly decreased (Supplementary Figure S14, top). In this kind of experiment, the growth is related first to the expression of the various has components under the control of their own regulatory elements, and to the

efficiency of the heme uptake process itself. Although we did not measure kinetics and extent of the transcriptional activation of the has locus under these conditions, which is well beyond the scope of our study, one possible interpretation of our results points to a likely role of the N-terminal periplasmic extension in the activation of the transcription of the has locus. To ensure that this difference was not actually due to a difference of expression of the two constructs, we determined by Western blot their amounts in membrane preparations (Supplementary Figure 14, bottom). As our anti-ExbB_{Sm} antibody was not sensitive enough for this measurement, we used His-tagged versions of ExbB. The experiment shows that the amount of the ExbB variant deleted of its N-terminal extension is at least equal, if not slightly higher than the amount of the wild-type protein, ruling out a possible decrease in protein concentration.

4c Swapping residues in TM1 between ExbB_{Sm} and ExbB_{Ec} strongly suggests that ExbB TM1 interacts with HasB/TonB.

As mentioned above, HasB is a dedicated Has system TonB ortholog, that is functional with ExbBD_{Sm}, but not with ExbBD_{Ec}. This clearly points to the looser specificity of ExbBD_{Sm}, as compared to ExbBD_{Ec}, as ExbBD_{Sm} is functional with both HasB and TonB_{Ec}. One reason for this could be the higher stability of HasB itself in the cytoplasmic membrane, as it is known that TonB can only accumulate in the cytoplasmic membrane when its ExbB and ExbD partners are present. It is hypothesized that TonB is proteolytically degraded in the absence of ExbBD. We therefore tested to what level HasB accumulates in the presence of either ExbBD_{Ec} or ExbBD_{Sm} and found that in the absence of either complexes, HasB was hardly detectable, whereas it readily accumulated in their presence (Supplementary Figure S15). This observation therefore contradicts the interpretation that HasB could be inactive in the presence of ExbBD_{Ec} due to a lack of accumulation, and also indicated that most likely HasB also interacts with ExbBD_{Ec} so as to be stabilized (i. e. withstanding proteolytic degradation), but in a non-functional manner. Given the general framework put forward for the Mot complexes where the MotA (equivalent of ExbB) rotation around MotB (the equivalent of ExbD) drives the rotation of the flagellum basal ring [18, 19] we tested the possibility that ExbB residues exposed at the surface of the protein might be involved in this functional interaction with HasB/TonB. The structure of ExbB highlighted the conservation of the residues across the different bacterial genomes and showed that the membrane-facing residues of the TM region of helix $\alpha 2$ and the cytoplasmic

residues were the least conserved as shown in the alignment and the structure (Supplementary Figures S1 and S9). There was one conspicuous stretch of residues located in the cytoplasmic leaflet of TM helix 1 (see Supplementary Figure S9): residues 76-88 of ExbB_{Sm} -TILFAKGSELLRA-, corresponding to residues 39-51 of ExbB_{Ec} -AIFFSKSVEFFNQ-, that is the least conserved one between the two proteins. In this stretch, the conserved I, K and E residues point toward the innermost part of the protein and the non-conserved residues point to the outside and are available for protein-protein interaction (Supplementary Figure S16). It was tempting to speculate that this stretch might correspond to an interaction region between ExbB_{Sm} and HasB. We therefore made a chimeric ExbB_{Ec} protein with its 39-51 region replaced by residues 76-88 from ExbB_{Sm} named ExbB_{Ec-Sm76-88} and asked whether this chimeric ExbB_{Ec} protein would allow growth in the presence of HasB (see Supplementary Figure S3bis for the constructions). As shown in Supplementary Figure S3, this chimeric protein ExbB_{Ec-Sm76-88} was active with both TonB_{Ec} and HasB as it promotes growth with both hasISRADE and hasISRADEB plasmids, strongly suggesting that this region is involved in specific interactions with HasB. We also made the converse change where the 76-88 region of ExbB_{Sm} was replaced by the 39-51 region from ExbB_{Ec} (ExbB_{Sm-Ec39-51}). In our *in vivo* tests there was no major difference between ExbB_{Sm} and ExbB_{Sm-Ec39-51} and ExbB_{Sm-Ec39-51} was still active with HasB, although to a slightly lower extent. To better locate the region responsible for specificity, we produced one additional mutant where only residues 76-84 (TILFAKGSE) from ExbB_{Sm} were exchanged for residues 39-47 (AIFFSKSVE) from ExbB_{Ec} yielding ExbB_{Ec-Sm76-84}. ExbB_{Ec-Sm76-84} was still active with HasB, showing that this 9-residue stretch is sufficient to transform ExbB_{Ec} into a Sm-like phenotype (see Supplementary Figure S3 and S3bis for the visualisation of the constructions). This set of experiments shows that intramembrane functional zones of ExbB are crucial for the interaction between ExbB and HasB.

In summary, we showed using NMR and culture growth assays that ExbB_{Sm} N-terminal extension and a few residues in the α 2 helix TM part are likely involved in specific interactions with HasB.

Discussion

1) ExbB_{Sm} is a 5:2 complex at pH 8. Our structures were solved at pH 8, and examination of the 2D class averages only showed pentameric structures and not hexameric ones, contrarily to what has been observed in the Maki-Yonekura study, where ExbB was mostly seen as a hexamer in the high pH regime[16]. In their study, the solubilizing detergent was DDM, that

was subsequently exchanged for either C8E4 or C10E5. These shorter chain detergents might have led to some destabilization of the pentameric structure and provoked the conversion to the hexameric form. On the other hand, our data agree with those published by Lloubes, Buchanan et al. [15, 17]. *In vivo*, a stoichiometry of 7ExbB, 2ExbD and 1TonB has been determined in *E. coli* [27]. Although one cannot completely rule out some biases in this indirect measurement, a physiologically relevant reason for an excess of ExbB would be to provide a permanent scaffold to which TonB and/or ExbD could associate with, upon TonB complex dissociation, that might occur during the catalytic cycle, as has already been proposed [28].

2) ExbB and ExbBD are co-purified with endogenous lipids

In their last study, Celia *et al.* reconstituted ExbBD from *E. coli* into nanodiscs with added phospholipids from *E. coli*, and found mostly PE bound to ExbBD [17]. It is unclear whether this phospholipid was a genuine tightly bound lipid co-purified or whether it came from the lipids added during reconstitution in the nanodiscs. In our case, we did not add any lipid and we found mostly PG and some PE in the ExbB case (and also CL for ExbBD) using mass spectrometry. The cryo-EM structure shows that they are located in the inner leaflet, as in the *E. coli* complex. PG and CL are negatively charged, whereas PE is zwitterionic. This might be pertaining to the functioning of the complex, as the lipids are bound to a highly positively charged interface, very close to the fenestrations seen in the structure and PE might therefore be less tightly bound than PG or CL. Further analysis by mass spectrometry of the aliphatic chains of the lipids present in our sample also revealed a specific composition in their length and unsaturated nature.

One point also worth mentioning is the existence of “cavities” inside the ExbB tunnel, potentially allowing the passage of proton/hydronium ion up to the Asp 25 residue of ExbD, deeply embedded in the apolar medium of ExbB, and therefore with a pKa close to physiological pH, allowing easy protonation/deprotonation steps of Asp 25 side chain. The identified tunnels connect the cytoplasm to the periplasm *via* Asp 25, allowing us to propose a possible trajectory for the proton transfer.

3) Motor model proposal

Regarding the coupling of the pmf dissipation with the mechanical work carried out by MotAB/TolQR/ExbBD complexes, several models have been put forward that might be more or less easy to reconcile with the structural data. A detailed mechanism was proposed for *C. jejuni* MotAB complex, where cycles of protonation/deprotonation of the conserved Asp of

the MotB TM (equivalent to the conserved Asp 25 of the ExbD TM) are coupled to the rotation in discrete steps of the MotA pentamer around MotB axis, provided that the periplasmic domain of MotB is anchored to the peptidoglycan layer *via* its PG binding site [18]. Our ExbBD structure indeed shows that ExbD monomers have two orientations relative to ExbB: one is such that Asp 25 faces Thr 208 of ExbB and the other Asp 25 faces a hydrophobic region at the interface of two ExbB monomers. These two states are compatible with protonation/deprotonation cycles accompanied by a change in the environment upon rotation.

Moreover, our NMR results show that ExbB interacts *via* its alanine and proline-rich periplasmic N-terminal region with the HasB periplasmic domain, and growth studies of mutant bacteria show that ExbB_{Sm} α 2 helix TM region likely carries specificity determinants of the interaction with HasB; this last observation is in line with a previous TonB TM point mutant partially suppressed by an ExbB α 2 point mutant [29]. It also appears that HasB stabilisation can occur without full functionality, as ExbBD_{Ec} protects HasB from degradation, without productively interacting with it. One possible interpretation is that stabilisation of HasB is provided by its interaction with ExbD_{Ec} periplasmic domain.

The MotA-MotB model also posits that MotA outer region interacts with the rotor of the flagellum and thus that MotA rotation drives the rotation of the flagellum. A similar interaction between the outermost TM of ExbB (α 2) and the TM of HasB/TonB is supported by our mutagenesis data, as we could swap the specificity of ExbB_{Ec} towards HasB by exchanging a short stretch of residues between ExbB_{Sm} and ExbB_{Ec}. Moreover, the sequences of the TM regions of HasB and TonB are different, apart from a conserved Histidine. Structure comparison of the swapped regions shows that ExbB_{Sm} has smaller residues that may be better accommodated and less specific than the *E. coli* sequence (Supplementary Figure S16).

On the part of the periplasmic space close to the outer membrane, it is well known that the C-terminal domain of HasB/TonB interacts with the “TonB/HasB box” of the outer membrane receptor, and that this interaction is essential for the entry of the substrate bound to the extracellular side of the TBDT into the periplasm [30]. AFM experiments also showed that by exerting a pulling force on the C-terminal domain of TonB bound to the BtuB TBDT, a partial unfolding of the TBDT plug occurs, potentially leading to the diffusion of the substrate in the periplasm [31]. Kinetic measurements showed that the *in vivo* entry of a siderophore is a rather slow process: during a cell division time a bacterial cell has to take up ca. 300000 iron

atoms from its surroundings. Under full induction, there are roughly 10-15000 FepA siderophore receptors per cell and 1500 TonB complexes [27], meaning that each receptor has to undergo 20 cycles during a generation, leading to a turnover time of the TonB complex of ca 5-10 seconds. Several models have been proposed to account for the functioning of the TonB complex, in conjunction with the entry of a substrate bound on the extracellular side of a given receptor. Firstly, piston-like movements of TonB have been proposed, driving the unfolding of the plug inside the receptor barrel. AFM experiments and steered molecular dynamics also showed that monomeric TonB devoid of its TM segment was able to drive the partial unfolding of the BtuB plug allowing thus vitamin B12 entry into the periplasm in a few discrete steps [31]. It is also known that ExbD and TonB interact *via* their periplasmic parts [32] and in particular residue 150 of TonB (positioned just upstream the C-terminal globular domain) can make cross-links to the C-terminal domain of ExbD [33]. Secondly, the MotAB model can be extrapolated to the ExbBD complex. In this model, ExbB would rotate around ExbD, and this would drive the rotation of TonB, thanks to the specific interaction between TonB TM and ExbB TM (in our case HasB TM and ExbB TM α 2). It has also been hypothesized that the C-terminal domain of TonB [34] and/or the C-terminal domain of ExbD have specific binding sites for the peptidoglycan network [35]. This C-terminal domain can change conformations along the catalytic cycle of protonation and deprotonation events at the conserved Asp residue in ExbD TM. In addition, the periplasmic linker between the TM helix and the C-terminal domain of TonB is rich in Pro and Lys residues and might adopt an extended conformation of a sufficient length to span the periplasm [36]. We therefore propose a model in which the force generated through rotation of TonB, driven by ExbB rotation, is not directly transmitted to the TonB box of the receptor, but through the intermediary of the C-terminal domain of ExbD, that might act as a wedge or a linchpin allowing force transmission and converting the rotation into a pulling force exerted on the TonB/HasB box of the receptor. Further studies are needed to test this model.

Material and Methods

Strains and plasmid construction

Strains, plasmids, and oligonucleotides used are shown Table S2.

-Plasmid pBADExbBD_{Sm} was constructed after amplification on the genomic DNA from strain

S. marcescens Db11 of a ca. 1.42kb fragment using primers ExbBDSm5' and ExbBDSm3'. The PCR product was purified digested with *EcoRI* and *SphI*, and ligated with pBAD24 digested with *EcoRI* and *SphI*. Correct clones were selected after sequencing of the insert.

-Plasmid pBADExbBDSmHis6 was constructed by first amplifying on pBADExbBDSm a ca. 0.4kb fragment with the following oligonucleotides *SphI*HisCtexbDsm and *BglII*exbDsm; after amplification, the fragment was purified, digested with *BglII* and *SphI*, and ligated with pBADExbBDSm digested with the same enzymes. The correct clones were selected after sequencing of the insert. In biological tests, this plasmid was undistinguishable from its parent plasmid pBADExbBDSm.

-Plasmid pBADExbBD_{Ec} was constructed by amplification on genomic MG1655 DNA of a ca. 1.7kb fragment with the following oligonucleotides ExbBD5c, and ExbBD3c; the PCR product was purified, digested with *EcoRI* and *SphI*, and ligated with pBAD24 digested with the same enzymes. Correct clones were selected after sequencing.

-Plasmid pBADExbB_{Sm}His6 was constructed as follows: a PCR fragment was amplified on pBADExbBDSm with the oligos PBADFOR and ExbBHis6, digested with *EcoRI* and *SphI*, and ligated with pBAD24 digested with the same enzymes. Correct clones were selected by sequencing.

Plasmid pBADExbBD_{Ec-Sm76-88} was constructed as follows: Plasmid pBADExbBD_{Ec} was amplified with the two following 5' phosphorylated oligonucleotides: ExbBEcSm76-88.1 and ExbBEcSm76-88.2. After digestion with *DpnI* and self-ligation, correct clones were selected by sequencing and the mutated fragments recloned in pBAD24.

Plasmid pBADExbBD_{Ec-Sm76-84} was constructed in the same manner, with the following oligonucleotides pairs: ExbBEcSm 76-84.1 and ExbBEcSm 76-84.2.

In the same manner, plasmid pBADExbBD_{Sm-Ec39-51} was constructed by PCR on the pBADExbBDSm plasmid with the following couple of 5' phosphorylated oligonucleotides: ExbBSmEc39-51.1 and ExbBSmEc39-51.2.

Plasmid pBADExbB_{Smdelextss}D was constructed similarly, by using as a template pBADExbBDSm, with the two phosphorylated oligonucleotides ExbBdelextss1 and ExbBdelextss2.

To construct pBADExbB_{Smdelextss}His6, a ca. 0.6kb *EcoRI*-*KpnI* fragment from pBADExbB_{Smdelextss}D was exchanged for the corresponding fragment of pBADExbBHis6.

Plasmid pAMhasISRDEB, was constructed by digesting pAMhasISRDE and pSYC7 [37] by *KpnI* and *HindIII*, and purifying fragments of respectively ca. 9 and 7kbases, ligating them

together to obtain a plasmid with the whole *has* locus on a low copy number plasmid (pAM238) under the control of its endogenous regulation signals.

All constructions were carried out in *E. coli* strain XL1-Blue.

Strains and other plasmids used in this study are shown in Table S2 and are from the laboratory collection.

Protein expression and purification

The BL21DE3(pExbBD_{sm}His6/pBAD24) or BL21DE3(pExbB_{sm}His6/pBAD24) were grown at 37°C either in TB or MDM medium, and induced with 0.2% arabinose at 1.5-2OD_{600nm} (TB) or 5-6OD_{600nm} (MDM) and the incubation continued for 3hrs. The cells were harvested by centrifugation (10000g 20minutes 4°C), washed once in 20mMTris-HCl pH8.0, flash-frozen in liquid N₂ and kept at -80°C. Cells were broken in a Cell disruptor (Constant, UK) at 1kbar (10g of cells in 40ml final of 20mM Tris pH 8.0 containing protease inhibitor (Roche, EDTA free), at 4°C. Benzonase was added and after ca. 15 minutes, the solution was centrifuged for 1hr at 100000g at 4°C. The pellet (crude membrane preparation) was resuspended in 20mM Tris pH 8.0 plus protease inhibitor cocktail (Roche EDTA free), flash-frozen in liquid N₂ and kept at -80°C.

The crude membrane preparation was solubilized in 20mM Tris pH 8.0, 20mM Imidazole, 100mM NaCl, 10% Glycerol, 0.8% LMNG (10g of equivalent whole-cell pellet solubilized in 40ml), plus protease inhibitor cocktail (Roche EDTA-free) for 30 minutes at 15°C. After centrifugation (1hr 100000xg), the supernatant was incubated with 2.5ml Ni-Agarose beads (Thermo-Fischer His-Pure Ni-NTA #88222) preequilibrated with the same buffer except for the detergent concentration (0.0015% LMNG). After three hours of incubation on a rotating wheel at 4°C, the beads were washed three times with 25ml of pre-equilibration buffer and then eluted with two times 25ml of the pre-equilibration buffer containing 200mM Imidazole. The eluate was concentrated and washed on 100kDa cut-off centrifugal device, in pre-equilibration buffer without NaCl and Imidazole. The resulting sample was loaded on a monoQ HR10-100 column equilibrated with 20mM Tris pH 8.0, 10% glycerol, 0.0015%LMNG and eluted with a gradient from 0 to 1M NaCl in the same buffer. Peak fractions were collected and concentrated as before. The concentrated sample was then loaded on a Superose 6 increase column equilibrated with 20mM Tris pH8.0, 100mM NaCl, 0.0015%LMNG. The peaks fractions were collected, concentrated and their concentration determined by using the

theoretical absorption coefficient of either ExbB5D2 ($113790\text{M}^{-1}\text{cm}^{-1}$), or ExbB5 ($104850\text{M}^{-1}\text{cm}^{-1}$). They were then kept frozen at -80°C in aliquots until use. The SEC profiles as well as a gel of a representative sample after purification are shown in supplementary data. Previous attempts with DDM instead of LMNG yielded similar results, and LMNG was chosen, owing to its very low CMC and its lower background in cryo-EM.

Activity tests

Three types of tests were used, either in liquid medium or on agar plates.

1. Growth tests in liquid medium: *E. coli* C600 and its ΔexbBD derivatives were transformed with specified plasmids and their growth tested in liquid LB medium at 37°C , at various dipyriddy concentrations to induce iron starvation. $\text{OD}_{600\text{nm}}$ was measured after overnight growth.

2. Growth tests on plates: the relevant plasmids were also transformed into *E. coli* C600 ΔhemA , a heme auxotroph strain (and derivatives thereof), and growth of the strains was assayed as follows. Briefly, cells were first grown in LB medium (supplemented with delta-aminolevulinic acid ($50\mu\text{g}/\text{ml}$) to bypass the effect of the ΔhemA mutation) at 37°C up to an $\text{OD}_{600\text{nm}}$ of 1, and then mixed with melted top agar (0.6% agar in LB), and poured onto LB agar plates containing the appropriate antibiotics with arabinose at a concentration of $40\mu\text{g}/\text{ml}$ to induce expression of the genes under the P_{araBAD} promoter[38]. Wells, punched with Pasteur pipettes were filled with heme-albumin (at 5, 1 or $0.2\mu\text{M}$), at two dipyriddy concentrations (100 or $200\mu\text{M}$) to induce iron starvation. Plates were incubated overnight at 37°C and scored for growth around the wells. All experiments were performed in triplicate.

3. Growth curves in liquid medium: a few colonies of *E. coli* C600 $\Delta\text{hemA}\Delta\text{tonB}\Delta\text{exbBD}$ (pAMhasISRDEB + pBAD24 or derivatives thereof) were first inoculated in 4ml of LB medium at 37°C with the corresponding antibiotics, and $100\mu\text{M}$ dipyriddy, $4\mu\text{g}/\text{ml}$ arabinose but without delta-aminolevulinic acid. Once the culture reached an $\text{OD}_{600\text{nm}}$ of ca. 1.2-1.5, it was diluted and inoculated in 48 well Greiner plates, in the same medium to which was added $0.4\mu\text{M}$ He-BSA, as a heme source. The initial $\text{OD}_{600\text{nm}}$ of the cultures was 0.001. Each well contained $300\mu\text{l}$ of growth medium. Duplicates of each strain were made, and the plate was incubated at 37°C with vigorous shaking (500rpm) in a Clariostar Plus Microplate reader. $\text{OD}_{600\text{nm}}$ was recorded every 30 minutes for 70 hours.

N-terminal sequencing

The N-terminus of ExbB_{Sm} was determined at the Plateforme Protéomique de l'institut de microbiologie de la Méditerranée (Marseille), after blotting on a PVDF membrane a purified sample of ExbBD_{His6} run on SDS-PAGE and determined to be APAAN.

Lipid extraction

Chloroform (0.20 mL) and methanol (0.40 mL) were sequentially added to a sample of the ExbB or ExbBD complex (0.1 mL). The sample was vortexed for 10 min at room temperature and chloroform (0.2 mL) and water (0.2 mL) were further added. The organic phase was collected and the extraction procedure was repeated on the remaining aqueous phase. Combined organic layers were evaporated to dryness under argon and stored at -20 °C.

Mass Spectrometry

Dry lipid extracts obtained from the ExbBD complex and from an *E. coli* lysate were solubilized in chloroform/methanol (50/50). Samples were then nano-electrosprayed using a TriVersa NanoMate (Advion Biosciences, Ithaca, USA) coupled to a Synapt G2-Si mass spectrometer (Waters Corporation, Manchester, UK). The instrument was calibrated in negative ion mode (for lipids) and positive ion mode (for proteins) from 50 m/z to 2000 m/z with NaI (50 mg/ml) with an accuracy of 0.8 ppm in resolution mode. The following settings were chosen: sampling cone 40 V, source offset 40 V, source temperature 80°C, trap gas flow 5 mL/min, helium cell gas flow 180 mL/min. MS/MS spectra were recorded using CID (Collision Induced Dissociation) with a normalized collision energy set up to 30.

To measure the mass of the intact proteins of the ExbBD and purified ExbB complexes, a desalting step was performed using micro Bio-Spin™ 6 (BIO-RAD) with 500mM ammonium acetate. Both samples were analyzed in denaturing conditions after a two-fold dilution with acetonitrile 4% formic acid. Mass spectra were acquired in positive ion mode. All molecular weights were measured after MaxEnt1 deconvolution into neutral species.

Cryo-EM grid preparation and data acquisition

3 μL of either purified ExbB_{Sm} or ExbBD_{Sm} complex at ca. 1 mg/mL was applied to C-Flat 1.2/1.3 holey carbon grids (Protochips Inc., USA) previously glow-discharged in air for 30s. Grids were blotted for 2s at blot force 1 and vitrified in liquid ethane using a Vitrobot mark IV (FEI company) operated at 10 °C and 100% relative humidity.

All data collection was performed with a Titan Krios (ThermoFisher Scientific) operated at 300 kV equipped with a K2 Summit direct electron detector (ThermoFisher Scientific) at the European synchrotron research facility, ESRF (Grenoble, France). Movies were recorded in electron counting mode with EPU software (ThermoFisher Scientific), aligned with Motioncor2 [39] and aligned images were processed with Gctf [40] using Scipion interface[41].

For ExbB data collection, 4567 movies were collected at a magnification of 165000x with a nominal pixel size of 0.827Å using a defocus range from -1.5 to -2.5 μm . Movies of 56 frames were acquired using a dose rate of 8 electrons/Å²/second over 7 seconds, yielding a cumulative exposure of 55.95 electrons/Å².

For ExbBD data collection, 4043 movies were collected at a nominal magnification of 139000 with a pixel size of 1.067Å. Movies of 48 frames were acquired using a dose rate of 4.6 electrons/Å²/second over 12 seconds yielding a cumulative exposure of 55.2 electrons/Å².

Cryo-EM image processing and analysis

a) ExbB

For ExbB, aligned movies were processed with Gctf [40] and only images with a resolution higher than 4Å were kept; after visual inspection of the remaining images, processing was carried out with Relion-3 [42]. Particles were extracted using a 2-fold binning, issued from a manual picking and a 2-D classification of particles picked out from on 50 images. Automatic extraction was performed using the selected 2D class averages. After several rounds of 2D and 3D classification, 161k particles were selected for 3D refinement. They were corrected for local motion using Bayesian polishing option in Relion-3 and a post-refined map produced a 3.1Å overall resolution with a 5-fold symmetry (Supplementary Figure S17).

A homology model was built with the *Serratia marcescens* ExbB sequence using Phyre2 server from 5SV0 monomer and docked into the refined map. Refinement was done with Phenix Real Space Refine option with secondary structure, “non-crystallographic symmetry” and Ramachandran restraints [43] and graphically adjusted with Coot [44]. Lipid phosphatidyl

glycerol starting structure and geometry were built using Phenix eLBOW [45] and was fit in the map using Coot then refined along with the protein model in Phenix.real_space_refine.

b) ExbBD

Movies were drift-corrected and dose-weighted using MotionCorr2 [39]. Aligned dose weighted averages were imported into Cryosparc2 [46] and contrast transfer function was estimated using CTFFIND4.1 [47]. Micrographs with poor CTF estimation statistics or high drift profiles were discarded. The remaining 3028 micrographs were used for automated particle picking. Particles were extracted, Fourier cropped to 2 Å/px and 2D classified. The best 2D classes were used as templates for automated particle picking resulting in 1.3 million particles. After several rounds of classification, the best 600k particles were submitted to 3D classification by means of multi-class ab-initio Reconstruction and Heterogeneous Refinement. 158 k particles belonging to the best resolved classes were corrected for local motion, re-extracted and used in Non-Uniform Refinement. The resulting refined map has a nominal resolution of 4.56 Å.

Based on the previous map a soft mask lining the micelle was designed in UCSF Chimera [48] and used to signal subtract the corresponding micelle density of particles in the refined map. Localized refinement of the signal subtracted particles produced a map of the complex with an estimated resolution of 3.96 Å judging by FSC at 0.143 criterion (Supplementary Figure S18). Data have been deposited both at the EMDB and PDB databases (EMDB 10789 and 11806, PDB 6YE4 and 7AJQ, for ExbB_{sm} and ExbBD_{sm} respectively, see Table I for refinement statistics).

Unless otherwise specified, the structure figures were made using UCSF Chimera software [48].

NMR experiments

The C-terminal periplasmic domain of HasB, HasB_{CTD} comprising residues 133-263, was produced and purified as previously reported [14, 30]. The peptide corresponding to the periplasmic extension of ExbB_{sm} (A₁PAANPAVTESVAPTTAPAPAAAAPESITPVNPAPTIIQPPETRG₄₄- numbering with reference to the mature protein) was synthesized by *Proteogenix*.

NMR experiments were acquired at 293 K on a 600 MHz Bruker Avance III spectrometer equipped with a TCI cryoprobe. The spectra were processed with NMRpipe [49] and analyzed with CcpNmr Analysis 2.4 software [50]. Proton chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5 sulfonate as 0 ppm. ^{15}N were referenced indirectly to DSS [51] (Wishart et al., 1995). ^1H – ^{15}N HSQC experiments were acquired on 0.15 mM HasB_{CTD} in 50mM sodium phosphate pH 7, 50 mM NaCl with or without the peptide. Aliquots from a solution of peptide at 25mg/ml in the same buffer were added to the protein sample at 2:1 and 10:1 ratios. Chemical shift perturbations (CSPs) of backbone amide cross-peaks were quantified by using the equation $\text{CSP} = [\Delta \delta\text{H}^2 + (\Delta \delta\text{N} \times 0.15)^2]^{1/2}$, where $\Delta \delta\text{H}$ and $\Delta \delta\text{N}$ are the observed ^1H and ^{15}N chemical shift differences between the two experimental conditions. The ^1H and ^{15}N resonance assignments were from Lefevre et al 2007 [52].

Other biochemical methods

SDS-PAGE and immunodetection with anti-HasB or anti-His6 antibodies (Abcam [HIS.H8] (ab18184)) on whole cells or membrane preparations were carried out following standard protocols. Secondary antibodies were coupled to alkaline phosphatase.

BLAST search

ExbB_{Sm} was used as a search for BLAST for orthologs in complete bacterial genomes, focusing on “long” ExbB’s. Top hits were in *Serratia*, *Yersinia*, *Dickeya*, *Erwinia* and *Pseudomonas*. Those genera were later excluded from successive BLAST searches to obtain orthologs in other genera with higher p-value.

Acknowledgments This work was supported by grants from the ANR (HEMESTOCKEXCHANGE ANR-12-BSV3-0022-01 and LABEX DYNAMO ANR-11-LABEX-0011-01). We thank Nathalie Dautin for helpful suggestions and critical reading of the manuscript. We acknowledge Emmanuel Frachon and Christophe Thomas from the Institut Pasteur platform “Production of recombinant proteins” for large scale cultures used in this work, European Synchrotron Radiation Facility for the provision of time on the Titan Krios at beamline CM01, the Institut Pasteur Biological NMR Technological Platform for the use of the NMR spectrometers, the staff at the ESRF CM01 facility in Grenoble for help in the data collection, the staff at the Marseille proteomics facility and Djalila Djouadi and Diana Basto who initially contributed to this project. We thank M. Nilges for interest in this work.

Figure legends

Figure 1:

Weblogo (<http://weblogo.threeplusone.com/>) representation of 131 « long » ExbB aminoacid sequences; those with much shorter periplasmic extensions have been excluded from the alignment. The orange boxes show the position of the three TM segments from the *E. coli* 5SV0 structure.

Figure 2:

Representation of the pentameric structure of ExbB_{Sm} solved by cryo-EM. A: the density map is shown, each monomer with distinct shades of blue, view from the side. The grey densities represent PG molecules. B: view from the cytoplasmic side. C: representation of one PG density together with the fitting of the PG molecule inside the density, and with the stick representation of the nearby ExbB residues using the Coot software [44]. D: superposition of a monomer of ExbB_{Ec} (dark grey) and ExbB_{Sm} (red), represented as ribbon. The membrane thickness is represented by the dotted lines and the transmembrane segments TM1 (part of α 2 helix), TM2 (part of α 6 helix) and TM3 (part of α 7 helix) are indicated.

Figure 3:

Structure of ExbB_{Sm} solved by cryo-EM. A: side view, the same color code as in Figure 2 for the ExbB monomers is used. A: side view; B: view from the periplasmic space; the two ExbD helices are seen respectively in red and magenta. C: clipped side view, with the helix of each ExbD monomer represented (red and magenta). D and E, side and top views showing the model built in the density maps.

Figure 4:

Interaction of HasB_{CTD} with the periplasmic fragment of ExbB_{Sm}, as detected by NMR. HasB_{CTD} residues shown on its structure (PDB code 2M2K) exhibiting the highest CSP in the presence of the periplasmic peptide of ExbB_{Sm} are coloured in red (Top: surface representation; bottom: cartoon representation). Right: Superposed ¹H–¹⁵N HSQC spectra of 0.15 mM ¹⁵N-labelled HasB_{CTD} in 50mM sodium phosphate, pH 7, 50 mM NaCl in the presence (red) or absence (black) of the periplasmic peptide of ExbB_{Sm}.

Complex	ExbB	ExbB-ExbD
Deposition ID	EMD-10789, PDB 6YE4	EMD-11806, PDB ID 7AJQ
Data collection and processing		
Microscope type, camera	ESRF Titan Krios, Gatan K2	ESRF Titan Krios, Gatan K2
Magnification	165000	139000
Voltage (kV)	300	300
Electron exposure (e-/Å ²)	56	55
Defocus range (μm)	-1.5 to -2.5	-1 to -3
Pixel size (Å)	0.87	1.067
Initial particle images (no.)	850000	1373000
Final particle images (no.)	157000	158000
Map resolution (Å)	2.97	4
FSC threshold	0.143	0.143
Map resolution range (Å)	3.2-5.5	3.97-8
Symmetry imposed	C5	C1
Initial model used	Pdb id. 5SV0	Our ExbB structure
Model composition		
Non-hydrogen atoms	9145	9256
Protein residues	1180	1229
Ligands	PGT : 5	0
B factors (Å ²)		
Protein min/max/mean	21.58/ 145.06/ 45.97	93.07/ 411.58/ 169.35
Ligand min/max/mean	45.43/ 46.91/ 46.12	
R.m.s. deviations		
Bond length (Å) (# >4σ)	0.026 (3)	0.003 (0)
Bond angles (°) (# > 4σ)	1.56 (8)	0.573 (4)
Validation		
Refined model CC	0.85	0.77
MolProbity score	1.1	2.03
Clashscore	3	15.3
Poor rotamers (%)	0	0
Ramachandran plot		
Favored (%)	99	95
Allowed (%)	1	5
Disallowed (%)	0	0

Table 1: Data collection, processing, and refinement statistics for ExbB_{Sm} and ExbBD_{Sm}.

Figure 1

WebLogo 3.7.4

761 Figure 2

762

763

Figure 3

Figure 4

References

1. Noinaj, N., et al., *TonB-dependent transporters: regulation, structure, and function*. Annu Rev Microbiol, 2010. **64**: p. 43-60.
2. Grinter, R. and T. Lithgow, *The structure of the bacterial iron-catecholate transporter Fiu suggests that it imports substrates via a two-step mechanism*. J Biol Chem, 2019. **294**(51): p. 19523-19534.
3. Cascales, E., R. Lloubes, and J.N. Sturgis, *The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB*. Mol Microbiol, 2001. **42**(3): p. 795-807.
4. Braun, V., *The structurally related exbB and tolQ genes are interchangeable in conferring tonB-dependent colicin, bacteriophage, and albomycin sensitivity*. J Bacteriol, 1989. **171**(11): p. 6387-90.
5. Eick-Helmerich, K. and V. Braun, *Import of biopolymers into Escherichia coli: nucleotide sequences of the exbB and exbD genes are homologous to those of the tolQ and tolR genes, respectively*. J Bacteriol, 1989. **171**(9): p. 5117-26.
6. Braun, V. and C. Herrmann, *Evolutionary relationship of uptake systems for biopolymers in Escherichia coli: cross-complementation between the TonB-ExbB-ExbD and the TolA-TolQ-TolR proteins*. Mol Microbiol, 1993. **8**(2): p. 261-8.
7. Fischer, E., K. Gunter, and V. Braun, *Involvement of ExbB and TonB in transport across the outer membrane of Escherichia coli: phenotypic complementation of exb mutants by overexpressed tonB and physical stabilization of TonB by ExbB*. J Bacteriol, 1989. **171**(9): p. 5127-34.
8. Ahmer, B.M., et al., *Characterization of the exbBD operon of Escherichia coli and the role of ExbB and ExbD in TonB function and stability*. J Bacteriol, 1995. **177**(16): p. 4742-7.
9. Higgs, P.I., P.S. Myers, and K. Postle, *Interactions in the TonB-dependent energy transduction complex: ExbB and ExbD form homomultimers*. J Bacteriol, 1998. **180**(22): p. 6031-8.
10. Braun, V., et al., *Energy-coupled transport across the outer membrane of Escherichia coli: ExbB binds ExbD and TonB in vitro, and leucine 132 in the periplasmic region and aspartate 25 in the transmembrane region are important for ExbD activity*. J Bacteriol, 1996. **178**(10): p. 2836-45.
11. Gaisser, S. and V. Braun, *The tonB gene of Serratia marcescens: sequence, activity and partial complementation of Escherichia coli tonB mutants*. Mol Microbiol, 1991. **5**(11): p. 2777-87.
12. Paquelin, A., et al., *Characterization of HasB, a Serratia marcescens TonB-like protein specifically involved in the haemophore-dependent haem acquisition system*. Mol Microbiol, 2001. **42**(4): p. 995-1005.
13. Benevides-Matos, N., C. Wandersman, and F. Biville, *HasB, the Serratia marcescens TonB paralog, is specific to HasR*. J Bacteriol, 2008. **190**(1): p. 21-7.
14. Lefevre, J., et al., *Modulation by substrates of the interaction between the HasR outer membrane receptor and its specific TonB-like protein, HasB*. J Mol Biol, 2008. **378**(4): p. 840-51.
15. Celia, H., et al., *Structural insight into the role of the Ton complex in energy transduction*. Nature, 2016. **538**(7623): p. 60-65.

- 894 16. Maki-Yonekura, S., et al., *Hexameric and pentameric complexes of the ExbBD*
895 *energizer in the Ton system*. Elife, 2018. **7**.
- 896 17. Celia, H., et al., *Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD*
897 *provides information on structure and stoichiometry*. Communications Biology, 2019.
898 **2**(1): p. 358.
- 899 18. Santiveri, M., et al., *Structure and Function of Stator Units of the Bacterial Flagellar*
900 *Motor*. Cell, 2020. **183**(1): p. 244-257 e16.
- 901 19. Deme, J.C., et al., *Structures of the stator complex that drives rotation of the bacterial*
902 *flagellum*. Nat Microbiol, 2020.
- 903 20. Ihsan, M.Z., et al., *Gene Mining for Proline Based Signaling Proteins in Cell Wall of*
904 *Arabidopsis thaliana*. Front Plant Sci, 2017. **8**: p. 233.
- 905 21. Mishra, A.K., et al., *Tryptophan-Rich and Proline-Rich Antimicrobial Peptides*.
906 Molecules, 2018. **23**(4).
- 907 22. Ball, L.J., et al., *Recognition of proline-rich motifs by protein-protein-interaction*
908 *domains*. Angew Chem Int Ed Engl, 2005. **44**(19): p. 2852-69.
- 909 23. Marchler-Bauer, A., et al., *CDD: a curated Entrez database of conserved domain*
910 *alignments*. Nucleic Acids Res, 2003. **31**(1): p. 383-7.
- 911 24. Letoffe, S., P. Delepelaire, and C. Wandersman, *Free and hemophore-bound heme*
912 *acquisitions through the outer membrane receptor HasR have different requirements*
913 *for the TonB-ExbB-ExbD complex*. J Bacteriol, 2004. **186**(13): p. 4067-74.
- 914 25. Li, H., A.D. Robertson, and J.H. Jensen, *Very fast empirical prediction and*
915 *rationalization of protein pKa values*. Proteins, 2005. **61**(4): p. 704-21.
- 916 26. Langsetmo, K., J.A. Fuchs, and C. Woodward, *The conserved, buried aspartic acid in*
917 *oxidized Escherichia coli thioredoxin has a pKa of 7.5. Its titration produces a related*
918 *shift in global stability*. Biochemistry, 1991. **30**(30): p. 7603-9.
- 919 27. Higgs, P.I., R.A. Larsen, and K. Postle, *Quantification of known components of the*
920 *Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA*. Mol
921 Microbiol, 2002. **44**(1): p. 271-81.
- 922 28. Ollis, A.A., et al., *Cytoplasmic membrane protonmotive force energizes periplasmic*
923 *interactions between ExbD and TonB*. Mol Microbiol, 2009. **73**(3): p. 466-81.
- 924 29. Larsen, R.A., et al., *Partial suppression of an Escherichia coli TonB transmembrane*
925 *domain mutation (delta V17) by a missense mutation in ExbB*. Mol Microbiol, 1994.
926 **13**(4): p. 627-40.
- 927 30. de Amorim, G.C., et al., *The structure of HasB reveals a new class of TonB protein*
928 *fold*. PLoS One, 2013. **8**(3): p. e58964.
- 929 31. Hickman, S.J., et al., *Gating of TonB-dependent transporters by substrate-specific*
930 *forced remodelling*. Nat Commun, 2017. **8**: p. 14804.
- 931 32. Garcia-Herrero, A., et al., *The solution structure of the periplasmic domain of the*
932 *TonB system ExbD protein reveals an unexpected structural homology with*
933 *siderophore-binding proteins*. Mol Microbiol, 2007. **66**(4): p. 872-89.
- 934 33. Ollis, A.A. and K. Postle, *The same periplasmic ExbD residues mediate in vivo*
935 *interactions between ExbD homodimers and ExbD-TonB heterodimers*. J Bacteriol,
936 2011. **193**(24): p. 6852-63.
- 937 34. Klebba, P.E., *ROSET Model of TonB Action in Gram-Negative Bacterial Iron*
938 *Acquisition*. J Bacteriol, 2016. **198**(7): p. 1013-21.

35. Boags, A.T., F. Samsudin, and S. Khalid, *Binding from Both Sides: TolR and Full-Length OmpA Bind and Maintain the Local Structure of the E. coli Cell Wall*. Structure, 2019. **27**(4): p. 713-724 e2.
36. Kohler, S.D., et al., *The proline-rich domain of TonB possesses an extended polyproline II-like conformation of sufficient length to span the periplasm of Gram-negative bacteria*. Protein Sci, 2010. **19**(4): p. 625-30.
37. Létoffé, S., J.M. Ghigo, and C. Wandersman, *Identification of two components of the Serratia marcescens metalloprotease transporter: protease SM secretion in Escherichia coli is TolC dependent*. J Bacteriol, 1993. **175**(22): p. 7321-8.
38. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter*. J Bacteriol, 1995. **177**(14): p. 4121-30.
39. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy*. Nat Methods, 2017. **14**(4): p. 331-332.
40. Zhang, K., *Gctf: Real-time CTF determination and correction*. J Struct Biol, 2016. **193**(1): p. 1-12.
41. Gomez-Blanco, J., et al., *Using Scipion for stream image processing at Cryo-EM facilities*. J Struct Biol, 2018. **204**(3): p. 457-463.
42. Zivanov, J., et al., *New tools for automated high-resolution cryo-EM structure determination in RELION-3*. Elife, 2018. **7**.
43. Afonine, P.V., et al., *Real-space refinement in PHENIX for cryo-EM and crystallography*. Acta Crystallogr D Struct Biol, 2018. **74**(Pt 6): p. 531-544.
44. Emsley, P., et al., *Features and development of Coot*. Acta Crystallogr D Biol Crystallogr, 2010. **66**(Pt 4): p. 486-501.
45. Moriarty, N.W., R.W. Grosse-Kunstleve, and P.D. Adams, *electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation*. Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 10): p. 1074-80.
46. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination*. Nat Methods, 2017. **14**(3): p. 290-296.
47. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from electron micrographs*. J Struct Biol, 2015. **192**(2): p. 216-21.
48. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research and analysis*. J Comput Chem, 2004. **25**(13): p. 1605-12.
49. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes*. J Biomol NMR, 1995. **6**(3): p. 277-93.
50. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a software pipeline*. Proteins, 2005. **59**(4): p. 687-96.
51. Wishart, D.S., et al., *¹H, ¹³C and ¹⁵N chemical shift referencing in biomolecular NMR*. J Biomol NMR, 1995. **6**(2): p. 135-40.
52. Lefèvre, J., et al., *(¹H, (¹³C and (¹⁵N resonance assignments of the C-terminal domain of HasB, a specific TonB like protein, from Serratia marcescens*. Biomol NMR Assign, 2007. **1**(2): p. 197-9.