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

Valérie Biou[1], Ricardo Jorge Diogo Adaixo[2], Mohamed Chami[2], Pierre-Damien Coureux[3], Benoist Laurent[4], Yvette Ntsogo[1], Gisele Cardoso de Amorim[5], Nadia Izadi-Pruneyre[5], Christian Malosse[6], Julia Chamot-Rooke[6], Henning Stahlberg[2], Philippe Delepelaire[1]

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 $ExbB_5$-$ExbD_2$ 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

37 acquisition *via* the Has system from *S. marcescens*. $ExbB_{Sm}$ represents thus a new class of ExbB
38 protein and our results shed light on the specificity determinants between the ExbB-ExbD
39 complex and their associated TonB partners.
40
41
42 **Introduction**
43 Transport of nutrients across the Gram-negative outer membrane is either a diffusion-
44 facilitated or an active process, powered in that last case by the energy of the proton-motive
45 force (pmf) transmitted to specialized outer membrane transporters (the so-called TBDT's,
46 TonB dependent transporters). A complex of three cytoplasmic membrane proteins, TonB,
47 ExbB and ExbD, that together form the TonB complex (see [1] for a review) conveys the energy
48 of the pmf to the TBDT. In *Escherichia coli* K12, there is only one set of *tonB, exbB* and *exbD*
49 genes, whereas there are nine TBDT's [2], all energized by the same complex. ExbB and ExbD
50 respectively belong to the MotA-TolQ-ExbB and MotB-TolR-ExbD protein families, that are
51 involved in power generation and transmission in various processes (flagellar rotation for
52 MotAB, active transport of molecules across OM receptors for ExbBD and cell division for
53 TolQR), at the expense of pmf dissipation across the cytoplasmic membrane [3-6]. Those
54 proteins form complexes that associate with TonB (ExbBD), TolA (TolQR) or the flagellar rotor
55 (MotAB, stator of the flagellar motor). It has also been shown that ExbD does not accumulate
56 in the absence of ExbB, and that TonB does not accumulate either in the absence of the ExbBD
57 complex [7]. It is also known that the C-terminal domains of TonB and ExbD interact with each
58 other [8, 9]. ExbD TM has one essential strictly conserved aspartate residue [10], that is
59 thought to undergo cycles of protonation/deprotonation (coupled to pmf dissipation).
60 TBDT's comprise a 22-stranded ß-barrel anchored in the outer membrane. An N-terminal
61 domain folded inside the barrel and called the plug contains the main region of interaction
62 with TonB, the so-called TonB box located in the N-terminal periplasm-exposed part of TBDT.
63 TonB has its N-terminus localized in the cytoplasm, a single transmembrane segment,
64 followed by a Pro-Lys rich region long enough to span the periplasm and a structured C-
65 terminal domain interacting with the TBDT. Conformational changes elicited upon interaction
66 between TonB and the TonB box of the TBDT, and likely mediated via a rearrangement of the
67 plug domain, ultimately lead to entry of the substrate (initially bound to the extracellular side
68 of the receptor) into the periplasm by a yet unknown mechanism.

69 In *Serratia marcescens*, a close relative of *E. coli*, in which were identified about twenty
70 potential TBDT's, there are at least two TonB homologs: an ortholog of $TonB_{Ec}$ (having 48%
71 identity with $TonB_{Ec}$[11]) and HasB, a TonB paralog that is strictly dedicated to its cognate
72 outer membrane receptor HasR[12, 13]. The Has (heme acquisition system) system includes
73 the TBDT HasR, which recognizes both free heme and the high-affinity extracellular heme-
74 binding HasA hemophore. HasB is encoded within the *has* locus and displays low sequence
75 identity with either *E. coli* or *S. marcescens* TonB. Unlike $TonB_{Sm}$, HasB cannot complement
76 $TonB_{Ec}$ functions, nor it is able to drive heme entry *via* the HasR receptor in the presence of
77 $ExbBD_{Ec}$. A gain of function mutation in the TM domain of HasB was however isolated in *E.*
78 *coli*, restoring heme entry *via* HasR in the presence of $ExbBD_{Ec}$ [12]. HasB has the same
79 topology as TonB, and its C-terminal domain interacts specifically with HasR with a ca. 40-fold
80 higher affinity than the corresponding TonB domain [14]. Those observations prompted us to
81 identify the ExbBD complex that would function with HasB in *S. marcescens*. Hence, we first
82 identified the resident orthologous $ExbBD_{Sm}$ and found that this complex is active with both
83 HasB and TonB.
84 Recently several structures of *E. coli* ExbB and ExbBD have been solved, either by X-ray
85 crystallography or cryo-EM[15-17]. All samples studied included ExbB and ExbD, but not all
86 showed an ordered structure of ExbD. In these structures, $ExbB_{Ec}$[15] appeared as a pentamer,
87 each monomer of ExbB having three TM helices that extend into the cytoplasm, and with a
88 highly polarized charge distribution on the cytoplasmic side. The central pore is apolar, lined
89 by TM helices 2 and 3 of each monomer, creating a large hydrophobic cavity inside the
90 structure. Another study using X-ray crystallography, single-particle cryo-EM and electron
91 diffraction on two-dimensional crystals concluded that ExbB could undergo a pH-dependent
92 pentamer to hexamer transition, and that the hexameric ExbB could accommodate three ExbD
93 TM segments in its pore[16]. A more recent cryo-EM study of ExbBD reconstituted in
94 nanodiscs however confirmed the $ExbB_5ExbD_2$ stoichiometry consistent with previous DEER
95 results and identified the two TM helices of ExbD in the ExbB pore [16]. Several structures of
96 the flagellar MotAB motor complexes (related to ExbBD) from diverse bacteria
97 (*Campylobacter jejuni, Bacillus subtilis, Clostridium sporogenes, Vibrio mimicus, Shewanella*
98 *oneidensis,* PomAB from *Vibrio alginolyticus*) were recently published that also display a
99 $MotA_5MotB_2$ stoichiometry with some common topology elements [18, 19]. Similarly,
100 *Pseudomonas savastanoi* ExbBD exhibits the same stoichiometry.

101 Here we identify a new family of ExbB proteins with a long N-terminal extension whose
102 presence is strongly correlated to the presence of a *hasB* gene ortholog in the genome. We
103 characterized $ExbB_{Sm}$ alone and $ExbBD_{Sm}$ by single-particle cryo-EM and determined their
104 structures at 3.2 and 3.9Å, respectively. We show that in both cases, $ExbB_{Sm}$ behaves as a
105 stable pentamer; in the $ExbBD_{Sm}$ complex, we identify two TM helices of $ExbD_{Sm}$ in the pore
106 of ExbB pentamer, as it was shown for the *E. coli* case. Using NMR measurements, we also
107 show that the N-terminal periplasmic extension specific to $ExbB_{Sm}$ interacts with the globular
108 domain of HasB. Finally, *via* mutagenesis studies and bacterial growth assays we show that
109 the first transmembrane helix TM1 of ExbB contains specificity determinants for interaction
110 with HasB/TonB and that the periplasmic extension does play a role in heme acquisition *via*
111 the has system.
112
113 **Results**
114 **1. Orthologs of *E. coli exbB* and *exbD* in *Serratia marcescens* define a new ExbB family with**
115 **a N-terminal extension.**
116 Sequence analysis of strain Db11, a fully sequenced *S. marcescens* isolate (GenBank: *Serratia*
117 *marcescens* subsp. marcescens Db11, complete genome. ACCESSION HG326223,
118 https://www.ncbi.nlm.nih.gov/nuccore/HG326223.1), identified one putative operon
119 encoding orthologs of *E. coli* ExbB and ExbD (SMDB11_3479(ExbD) and SMDB11_3480(ExbB)).
120 In this case the *exbB*-like gene is close to and in the opposite orientation to a *metC* homolog,
121 as in *E. coli.* There is no conservation on the 3' end. Comparison of the coding sequences
122 indicated that the *exbD* gene encoded a 140 residue-long protein (71% identity with the 141
123 residue-long *E. coli* sequence). In contrast, the putative *exbB* gene encodes a much longer
124 protein than its *E. coli* counterpart (325 residues instead of 244 residues in *E. coli*, 73% identity
125 in the common part). The additional stretch of residues present in $ExbB_{sm}$ is located at the N-
126 terminus and corresponds to a putative signal peptide followed by a ~50 residues N-terminal
127 extension of the mature protein (Supplementary Figure S1).
128 A BLAST search on bacterial genomes indicated that such "extralong" ExbB's are found in
129 different Gammaproteobacteria, including Serratia, Yersinia, Pseudomonas, Erwinia/Dickeya,
130 and many genera in the Alphaproteobacteria class. Table SI lists representative species of
131 Alphaproteobacteria in which such long ExbB's are found. About 90% of the long ExbB-
132 containing species listed in this table also have a *hasB* gene ortholog.

133 An aminoacid sequence alignment of these ExbBs (Clustal omega
134 https://www.ebi.ac.uk/Tools/msa/clustalo/, Figure 1) also shows that the mature N-terminal
135 extensions are of variable length between a few and 150 residues with an average of 50
136 residues. These extensions are quite rich in Ala (24,6%) and Pro (14,5%) residues and therefore
137 likely to be unstructured but could be involved in protein-protein interactions as proline-rich
138 regions often are in signalling processes [20-22].
139 The Conserved Domain Database [23] contains currently two ExbB subfamilies in the cl00568
140 MotA_ExbB superfamily, TIGR02797 (containing *E. coli* ExbB, as well as longer ExbB's) and
141 TIGR02805 (containing *Haemophilus influenzae* ExbB, with a very short cytoplasmic domain
142 between TM1 and TM2); based on sequence data, it is then possible to subdivide TIGR02797
143 into 2 groups with "extralong" ExbB's in one subfamily, and *E. coli*-like sequences in the other.
144

145 **2. ExbBD$_{Sm}$ complements ExbBD$_{Ec}$ for iron acquisition and functionally associates with both**
146 **TonB and HasB for heme acquisition through HasR**

147 As a first step in the characterization of the identified ExbBD$_{sm}$, we tested whether this
148 complex would complement ExbBD$_{Ec}$. To this end, the plasmids pBADExbBD$_{Sm}$ and
149 pBADExbBD$_{Ec}$ were constructed, introducing the two genes, from *S. marcescens* and *E. coli*,
150 into pBAD24 vector under the control of p$_{araBAD}$ promoter. By using growth under iron
151 starvation conditions as a test, we could show that *exbBD$_{Sm}$* complements *exbBD$_{Ec}$*. The *E. coli*
152 C600*ΔexbBD* strain is indeed more sensitive than its wild-type counterpart to iron starvation
153 (induced for example by the Fe^{2+} chelator dipyridyl) as ExbBD are required to transduce pmf
154 to TonB for siderophore uptake. Its growth in the presence of dipyridyl is restored by the
155 expression of either pBADExbBD$_{Ec}$ or pBADExbBD$_{Sm}$ (Supplementary Figure S2).
156 HasB is a HasR-specific TonB ortholog, and as ExbBD$_{Ec}$ does not function with HasB (for heme
157 acquisition through HasA-HasR [12]), we then verified that ExbBD$_{Sm}$ would do so. To this end,
158 pBADExbBD$_{Sm}$ was introduced into strain C600*ΔhemAΔtonBΔexbBD* harbouring the
159 phasISRADEB/238 recombinant plasmid containing the whole has locus with its own
160 regulatory elements. This strain could grow in the presence of heme (down to a concentration
161 of 0.2µM) and arabinose, whereas the C600*ΔhemAΔtonBΔexbBD*(phasISRADEB/238 +
162 pBADExbBD$_{Ec}$) could not (see Supplementary Figure S3 bottom part). The strain
163 C600*ΔhemAΔexbBD*(phasISRADE/238 + pBADExbBD$_{Ec}$) could grow in the presence of heme
164 and arabinose, indicating that pBADExbBD$_{Ec}$ is functional (see Supplementary Figure S3 top

165 part). Likewise C600*ΔhemAΔexbBD*(phasISRADE/238 + pBADExbBD$_{Sm}$) could grow in the
166 presence of heme and arabinose. This indicates that ExbB$_{Sm}$ and ExbD$_{Sm}$ functionally associate
167 with either HasB or TonB$_{Ec}$ to allow HasA-mediated heme entry *via* HasR, whereas ExbBD$_{Ec}$
168 only functions with TonB leading to the heme entry *via* the HasA-HasR system (see
169 Supplementary Figure S3 top part). As already observed, a stronger iron starvation is
170 necessary when TonB is the link between the inner and outer membrane components [24].
171 Therefore, ExbBD$_{Sm}$ is the genuine *E. coli* ortholog, with a looser specificity.
172
173 **3. Characterization of ExbB$_{Sm}$ and ExbBD$_{Sm}$ and their structural analysis by cryo-EM show**
174 **ExbB$_5$ and ExbB$_5$D$_2$ stoichiometries**
175 To gain further insight into the possible differences between ExbBD$_{Sm}$ and ExbBD$_{Ec}$, we purified
176 both ExbB$_{Sm}$ and ExbBD$_{Sm}$ (as His-tagged proteins, see Materials and Methods section) in
177 LMNG micelles (Supplementary Figure S4). Both ExbB$_{Sm}$ and ExbBD$_{Sm}$ appear as homogeneous
178 oligomers at the last size exclusion chromatography (SEC) step. Mass spectrometry analysis of
179 intact ExbD$_{Sm}$ (16 131 Da) and ExbB$_{Sm}$ (29 557 Da) proteins, together with the determination
180 of the first amino-acid residues for ExbB (APAAN) showed that the predicted signal sequence
181 of ExbB$_{Sm}$ is absent from the mature sequence and has therefore been processed. It also
182 showed that ExbD initial Met residue was cleaved off (see Supplementary Figure S5). More
183 interestingly, mass spectrometry analysis of chloroform/methanol extracts of the purified
184 proteins also showed that both complexes contained native lipids (Supplementary Figure S6),
185 mostly phosphatidylglycerol (PG) and phosphatidylethanolamine (PE), and that ExbBD also
186 contained some cardiolipin (CL) (Supplementary Figure S7). Further analysis of aliphatic chain
187 composition of lipids shows evidence of specific composition (Supplementary Figure S8).
188 **3a ExbB structure**
189 As ExbB$_{Sm}$ presented specific features compared to ExbB$_{Ec}$, we set out to determine its
190 structure by single-particle cryo-EM. The 3D class average model clearly showed a pentameric
191 structure. It was refined and polished to obtain a resolution of 3.2Å. Maps were refined with
192 and without C5 symmetry and showed a 98% correlation. Therefore, the map used for model
193 building was the one with C5 symmetry.
194 The structure solved here and shown in Figure 2 A and B, is very similar to that of *E. coli* ExbB
195 5SV0 structure (that was co-purified with ExbD), with a 1.3Å rmsd over all Cα atoms (Table I).
196 This shows that ExbB is stable as a pentamer on its own. The periplasmic N-terminal extension

197 did not yield any visible electron density, precluding its structure determination. At the C-
198 terminus of each monomer, however, electron density was clearly defined, allowing structure
199 determination for an additional 10 residues in a helical conformation up to the last one (Gly
200 283, see Figure 2D, helix α8), before the tag. In the 5SV0 X-ray structure, a calcium ion is bound
201 to the five Glu 106 (of helix α2 in the TM region) on the cytoplasmic side. In $ExbB_{Sm}$, this
202 residue is replaced by an Asn, and we do not observe any density that could be assigned to a
203 metal ion.
204 The structure of the whole pentamer with one monomer coloured as a function of sequence
205 conservation in this sub-class of ExbB's (Supplementary figure 9) shows that the highest
206 residue conservation is inside the transmembrane channel, indicative of high functional
207 constraints, while the TM residues located at the membrane surface are more variable. The
208 same observation was made for those ExbB's without N-terminal signal sequence and
209 periplasmic extension [15].
210 Similarly to the $ExbB_{EC}$ structure, at the cytoplasmic junction with the membrane there are
211 "fenestrations", potentially leading to solvent or ion passage [15]. These vents are located at
212 the junction of adjacent monomers as shown in figure 2A. Additional density was also clearly
213 observed on the external surface, which we attributed to the phospholipids present in the
214 preparation. Each ExbB monomer appears to be associated with the equivalent of two PG
215 molecules, located at the inner leaflet of the cytoplasmic membrane (Figure 2A and C).
216 However, we were able to model only one PG molecule per monomer with confidence. Three
217 PE (phosphatidylethanolamine, the major *E. coli* phospholipid) molecules and one PG
218 molecule were identified as associated to the ExbBD pentamer in the *E. coli* complex after
219 reconstitution in nanodiscs [17]. In summary, our cryo-EM structure shows that ExbB from *S.*
220 *marcescens* is stable as a pentamer and associates strongly with specific phospholipids coming
221 from the inner leaflet of the membrane.
222
223 **3b $ExbBD_{Sm}$ structure**
224 We also solved the structure of the $ExbBD_{Sm}$ complex by cryo-EM (Figure 3 and Table I). As for
225 ExbB, the purified complex exhibited a symmetric peak on SEC (see Supplementary Figure S4).
226 Cryo-EM data were collected and processed as described in the Material and Methods section.
227 2D classes showed preferential orientations with 92% top views and only 8% side views. Due
228 to this very strong bias in particle orientation, the resolution achieved was 3.9 Å, precluding

229 reliable positioning of the side chains. As observed for ExbB alone, in the ExbBD complex, ExbB
230 behaves as a pentamer and no sign of other assemblies was found. Similarly, no density could
231 be attributed to the periplasmic extension. There are two clear densities inside the ExbB pore
232 that were assigned to the pair of ExbD TM segments (Figure 3 B and C). The charge distribution
233 is highly polarised on the cytoplasmic side: the region close to the membrane is positively
234 charged as expected and the distal part is negatively charged (Supplementary Figure S10, top
235 left). The central pore is apolar, lined by TM helices 2 and 3 of each monomer, creating a large
236 hydrophobic cavity inside the structure (Supplementary Figure S10, top right). In the refined
237 model, Asp 25 from ExbD monomer chain G faces Thr 218 from ExbB chain C, while Asp 25
238 from the ExbD chain F faces the interface between two ExbB monomers A and E
239 (Supplementary Figure S10, bottom). The estimation of pKa by using the program Propka [25]
240 server (http://server.poissonboltzmann.org/) shows that Asp 25 has a pKa of 7.3 for chain F
241 and 7.42 for chain G, both very peculiar for solvent-accessible acidic residues but that can be
242 found in buried active sites or membrane proteins [26]. This pKa should allow protonation and
243 deprotonation of ExbD Asp 25 at physiological pH.
244 As compared to the structure of $ExbB_{Sm}$, there is a slight «opening» towards the periplasmic
245 side (the distance of Ala 197 from one subunit to Leu 204 of the facing subunit varying from
246 25.5Å in ExbB to 29.8Å in ExbBD). This opening of the structure is limited to the periplasmic
247 part (Supplementary Figure S11). The RMSD between $ExbB_{Sm}$ alone and in complex with
248 $ExbD_{Sm}$ is 2Å over all $C\alpha$ of the pentamer, and the RMSD between $ExbB_{Sm}$ in complex with
249 $ExbD_{Sm}$ and $ExbB_{Ec}$ in complex with $ExbD_{Ec}$ (structure 6TYI chains A-E) is 1.7Å.
250 As compared to the 6TYI structure of *E. coli* ExbBD complex, there are some significant
251 differences that could be related to the specificity of function of $ExbB_{Sm}$. Both $ExbB_{Sm}$ and
252 $ExbBD_{Sm}$ inner pores are slightly wider than their *E. coli* counterparts at the periplasmic
253 entrance. Consequently, two channels cross the membrane region, (extending from the
254 periplasmic entrance to the Asp 25 residue from the ExbD TM deeply embedded in the ExbB
255 pore) that are clearly seen in the $ExbBD_{Sm}$ (as detected by the MoleOnline Server) while the *E.*
256 *coli* structure exhibits a single, much thinner channel as detected with the same parameters
257 of MoleOnline (Supplementary Figure S12). It is therefore possible that this structure
258 represents another state of the ExbBD motor complex, whereby the periplasmic side allows
259 solvent access to the Asp 25 residue of ExbD TM. In summary, we show that ExbBD from *S.*

260 *marcescens* has a 5:2 stoichiometry with a larger channel allowing solvent or proton transport
261 from the periplasmic side as compared with the *E. coli* structure.
262
263 **4. Specific functions of the N-terminal extension and of the first TM segment of $ExbB_{Sm}$**
264 **4a NMR shows an interaction between HasB and $ExbB_{Sm}$ periplasmic regions.**
265 The N-terminal extension of $ExbB_{Sm}$ contains almost 60% of Ala and Pro residues and is
266 predicted to be unstructured. In agreement with this, there was no observable electron
267 density for this region neither in the free form nor in the complex. Indeed, NMR analysis of a
268 44 residue-long synthetic peptide corresponding to this extension ($ExbB_{Sm}$ 1-44) did not show
269 any indication of secondary structure (unpublished results). Since this region is predicted to
270 be in the periplasm where the C-terminal domain of HasB, $HasB_{CTD}$ is also located, we
271 investigated the interaction of the $ExbB_{Sm}$ 1-44 synthetic peptide with $HasB_{CTD}$ by NMR. The
272 analysis of chemical shift perturbations (CSP) of amide resonances of $HasB_{CTD}$ upon addition
273 of the peptide, showed that the chemical environment of their corresponding residues was
274 modified as a consequence of their interaction (Supplementary Figure S13). The residues
275 showing the highest chemical shift perturbations are mapped on $HasB_{CTD}$ structure 2M2K in
276 Figure 4. They are mainly located on the helical face of $HasB_{CTD}$ forming a continuous surface
277 of interaction. An additional small pocket at the C-terminus of the protein is also involved in
278 this interaction.
279
280 **4b An $ExbB_{Sm}$ mutant devoid of the N-terminal extension is affected in its function**
281 To assess the potential function of the N-terminal periplasmic extension of $ExbB_{Sm}$, we
282 engineered a mutant lacking this extension (see Materials and Methods for details) and tested
283 its effect on the growth of a C600*ΔhemAΔexbBDΔtonB* strain harbouring plasmids
284 pAMhasISRADEB and pBADexbBD$_{Sm}$ or its derivatives. To get high sensitivity and numerical
285 data in our growth tests, we switched to growth in liquid medium in 48-well plates (see
286 Materials and Methods for details). We could then show that the lag between the start of the
287 experiment and the onset of growth is delayed by about 2 hours, for the mutant strain (in
288 which ExbB is devoid of its periplasmic extension), as compared to the WT strain. In addition,
289 the maximum OD reached after growth is slightly but significantly decreased (Supplementary
290 Figure S14, top). In this kind of experiment, the growth is related first to the expression of the
291 various has components under the control of their own regulatory elements, and to the

292 efficiency of the heme uptake process itself. Although we did not measure kinetics and extent
293 of the transcriptional activation of the has locus under these conditions, which is well beyond
294 the scope of our study, one possible interpretation of our results points to a likely role of the
295 N-terminal periplasmic extension in the activation of the transcription of the has locus. To
296 ensure that this difference was not actually due to a difference of expression of the two
297 constructs, we determined by Western blot their amounts in membrane preparations
298 (Supplementary Figure 14, bottom). As our anti-$ExbB_{Sm}$ antibody was not sensitive enough for
299 this measurement, we used His-tagged versions of ExbB. The experiment shows that the
300 amount of the ExbB variant deleted of its N-terminal extension is at least equal, if not slightly
301 higher than the amount of the wild-type protein, ruling out a possible decrease in protein
302 concentration.
303
304 **4c Swapping residues in TM1 between $ExbB_{Sm}$ and $ExbB_{Ec}$ strongly suggests that ExbB TM1**
305 **interacts with HasB/TonB.**
306 As mentioned above, HasB is a dedicated Has system TonB ortholog, that is functional with
307 $ExbBD_{Sm}$, but not with $ExbBD_{Ec}$. This clearly points to the looser specificity of $ExbBD_{Sm}$, as
308 compared to $ExbBD_{Ec}$, as $ExbBD_{Sm}$ is functional with both HasB and $TonB_{Ec}$. One reason for this
309 could be the higher stability of HasB itself in the cytoplasmic membrane, as it is known that
310 TonB can only accumulate in the cytoplasmic membrane when its ExbB and ExbD partners are
311 present. It is hypothesized that TonB is proteolytically degraded in the absence of ExbBD. We
312 therefore tested to what level HasB accumulates in the presence of either $ExbBD_{Ec}$ or $ExbBD_{Sm}$
313 and found that in the absence of either complexes, HasB was hardly detectable, whereas it
314 readily accumulated in their presence (Supplementary Figure S15). This observation therefore
315 contradicts the interpretation that HasB could be inactive in the presence of $ExbBD_{Ec}$ due to a
316 lack of accumulation, and also indicated that most likely HasB also interacts with $ExbBD_{Ec}$ so
317 as to be stabilized (i. e. withstanding proteolytic degradation), but in a non-functional manner.
318 Given the general framework put forward for the Mot complexes where the MotA (equivalent
319 of ExbB) rotation around MotB (the equivalent of ExbD) drives the rotation of the flagellum
320 basal ring [18, 19] we tested the possibility that ExbB residues exposed at the surface of the
321 protein might be involved in this functional interaction with HasB/TonB. The structure of ExbB
322 highlighted the conservation of the residues across the different bacterial genomes and
323 showed that the membrane-facing residues of the TM region of helix $\alpha 2$ and the cytoplasmic

324 residues were the least conserved as shown in the alignment and the structure
325 (Supplementary Figures S1 and S9). There was one conspicuous stretch of residues located in
326 the cytoplasmic leaflet of TM helix 1 (see Supplementary Figure S9): residues 76-88 of $ExbB_{Sm}$
327 -TILFAKGSELLRA-, corresponding to residues 39-51 of $ExbB_{Ec}$ -AIFFSKSVEFFNQ-, that is
328 the least conserved one between the two proteins. In this stretch, the conserved I, K and E
329 residues point toward the innermost part of the protein and the non-conserved residues point
330 to the outside and are available for protein-protein interaction (Supplementary Figure S16). It
331 was tempting to speculate that this stretch might correspond to an interaction region between
332 $ExbB_{Sm}$ and HasB. We therefore made a chimeric $ExbB_{Ec}$ protein with its 39-51 region replaced
333 by residues 76-88 from $ExbB_{Sm}$ named $ExbB_{Ec\text{-}Sm76\text{-}88}$ and asked whether this chimeric $ExbB_{Ec}$
334 protein would allow growth in the presence of HasB (see Supplementary Figure S3bis for the
335 constructions). As shown in Supplementary Figure S3, this chimeric protein $ExbB_{Ec\text{-}Sm76\text{-}88}$ was
336 active with both $TonB_{Ec}$ and HasB as it promotes growth with both hasISRADE and hasISRADEB
337 plasmids, strongly suggesting that this region is involved in specific interactions with HasB. We
338 also made the converse change where the 76-88 region of $ExbB_{Sm}$ was replaced by the 39-51
339 region from $ExbB_{Ec}$ ($ExbB_{Sm\text{-}Ec39\text{-}51}$). In our *in vivo* tests there was no major difference between
340 $ExbB_{Sm}$ and $ExbB_{Sm\text{-}Ec39\text{-}51}$ and $ExbB_{Sm\text{-}Ec39\text{-}51}$ was still active with HasB, although to a slightly
341 lower extent. To better locate the region responsible for specificity, we produced one
342 additional mutant where only residues 76-84 (TILFAKGSE) from $ExbB_{Sm}$ were exchanged
343 for residues 39-47 (AIFFSKSVE) from $ExbB_{Ec}$ yielding $ExbB_{Ec\text{-}Sm76\text{-}84}$. $ExbB_{Ec\text{-}Sm76\text{-}84}$ was still
344 active with HasB, showing that this 9-residue stretch is sufficient to transform ExbBEc into a
345 Sm-like phenotype (see Supplementary Figure S3 and S3bis for the visualisation of the
346 constructions). This set of experiments shows that intramembrane functional zones of ExbB
347 are crucial for the interaction between ExbB and HasB.
348 In summary, we showed using NMR and culture growth assays that $ExbB_{Sm}$ N-terminal
349 extension and a few residues in the α2 helix TM part are likely involved in specific interactions
350 with HasB.

351 **Discussion**

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

356 was subsequently exchanged for either C8E4 or C10E5. These shorter chain detergents might
357 have led to some destabilization of the pentameric structure and provoked the conversion to
358 the hexameric form. On the other hand, our data agree with those published by Lloubes,
359 Buchanan et al. [15, 17]. *In vivo*, a stoichiometry of 7ExbB, 2ExbD and 1TonB has been
360 determined in *E. coli* [27]. Although one cannot completely rule out some biases in this indirect
361 measurement, a physiologically relevant reason for an excess of ExbB would be to provide a
362 permanent scaffold to which TonB and/or ExbD could associate with, upon TonB complex
363 dissociation, that might occur during the catalytic cycle, as has already been proposed [28].
364 2) ExbB and ExbBD are co-purified with endogenous lipids
365 In their last study, Celia *et al.* reconstituted ExbBD from *E. coli* into nanodiscs with added
366 phospholipids from *E. coli*, and found mostly PE bound to ExbBD [17]. It is unclear whether
367 this phospholipid was a genuine tightly bound lipid co-purified or whether it came from the
368 lipids added during reconstitution in the nanodiscs. In our case, we did not add any lipid and
369 we found mostly PG and some PE in the ExbB case (and also CL for ExbBD) using mass
370 spectrometry. The cryo-EM structure shows that they are located in the inner leaflet, as in the
371 *E. coli* complex. PG and CL are negatively charged, whereas PE is zwitterionic. This might be
372 pertaining to the functioning of the complex, as the lipids are bound to a highly positively
373 charged interface, very close to the fenestrations seen in the structure and PE might therefore
374 be less tightly bound than PG or CL. Further analysis by mass spectrometry of the aliphatic
375 chains of the lipids present in our sample also revealed a specific composition in their length
376 and unsaturated nature.
377 One point also worth mentioning is the existence of "cavities" inside the ExbB tunnel,
378 potentially allowing the passage of proton/hydronium ion up to the Asp 25 residue of ExbD,
379 deeply embedded in the apolar medium of ExbB, and therefore with a pKa close to
380 physiological pH, allowing easy protonation/deprotonation steps of Asp 25 side chain. The
381 identified tunnels connect the cytoplasm to the periplasm *via* Asp 25, allowing us to propose
382 a possible trajectory for the proton transfer.
383 3) Motor model proposal
384 Regarding the coupling of the pmf dissipation with the mechanical work carried out by
385 MotAB/TolQR/ExbBD complexes, several models have been put forward that might be more
386 or less easy to reconcile with the structural data. A detailed mechanism was proposed for *C.*
387 *jejuni* MotAB complex, where cycles of protonation/deprotonation of the conserved Asp of

388 the MotB TM (equivalent to the conserved Asp 25 of the ExbD TM) are coupled to the rotation
389 in discrete steps of the MotA pentamer around MotB axis, provided that the periplasmic
390 domain of MotB is anchored to the peptidoglycan layer *via* its PG binding site [18]. Our ExbBD
391 structure indeed shows that ExbD monomers have two orientations relative to ExbB: one is
392 such that Asp 25 faces Thr 208 of ExbB and the other Asp 25 faces a hydrophobic region at the
393 interface of two ExbB monomers. These two states are compatible with
394 protonation/deprotonation cycles accompanied by a change in the environment upon
395 rotation.
396 Moreover, our NMR results show that ExbB interacts *via* its alanine and proline-rich
397 periplasmic N-terminal region with the HasB periplasmic domain, and growth studies of
398 mutant bacteria show that $ExbB_{Sm}$ α2 helix TM region likely carries specificity determinants of
399 the interaction with HasB; this last observation is in line with a previous TonB TM point mutant
400 partially suppressed by an ExbB α2 point mutant [29]. It also appears that HasB stabilisation
401 can occur without full functionality, as $ExbBD_{Ec}$ protects HasB from degradation, without
402 productively interacting with it. One possible interpretation is that stabilisation of HasB is
403 provided by its interaction with $ExbD_{Ec}$ periplasmic domain.
404 The MotA-MotB model also posits that MotA outer region interacts with the rotor of the
405 flagellum and thus that MotA rotation drives the rotation of the flagellum. A similar interaction
406 between the outermost TM of ExbB (α2) and the TM of HasB/TonB is supported by our
407 mutagenesis data, as we could swap the specificity of $ExbB_{Ec}$ towards HasB by exchanging a
408 short stretch of residues between $ExbB_{Sm}$ and $ExbB_{Ec}$. Moreover, the sequences of the TM
409 regions of HasB and TonB are different, apart from a conserved Histidine. Structure
410 comparison of the swapped regions shows that $ExbB_{Sm}$ has smaller residues that may be better
411 accommodated and less specific than the *E. coli* sequence (Supplementary Figure S16).
412 On the part of the periplasmic space close to the outer membrane, it is well known that the
413 C-terminal domain of HasB/TonB interacts with the "TonB/HasB box" of the outer membrane
414 receptor, and that this interaction is essential for the entry of the substrate bound to the
415 extracellular side of the TBDT into the periplasm [30]. AFM experiments also showed that by
416 exerting a pulling force on the C-terminal domain of TonB bound to the BtuB TBDT, a partial
417 unfolding of the TBDT plug occurs, potentially leading to the diffusion of the substrate in the
418 periplasm [31]. Kinetic measurements showed that the *in vivo* entry of a siderophore is a
419 rather slow process: during a cell division time a bacterial cell has to take up ca. 300000 iron

420 atoms from its surroundings. Under full induction, there are roughly 10-15000 FepA
421 siderophore receptors per cell and 1500 TonB complexes [27], meaning that each receptor
422 has to undergo 20 cycles during a generation, leading to a turnover time of the TonB complex
423 of ca 5-10 seconds. Several models have been proposed to account for the functioning of the
424 TonB complex, in conjunction with the entry of a substrate bound on the extracellular side of
425 a given receptor. Firstly, piston-like movements of TonB have been proposed, driving the
426 unfolding of the plug inside the receptor barrel. AFM experiments and steered molecular
427 dynamics also showed that monomeric TonB devoid of its TM segment was able to drive the
428 partial unfolding of the BtuB plug allowing thus vitamin B12 entry into the periplasm in a few
429 discrete steps [31]. It is also known that ExbD and TonB interact *via* their periplasmic parts
430 [32] and in particular residue 150 of TonB (positioned just upstream the C-terminal globular
431 domain) can make cross-links to the C-terminal domain of ExbD [33]. Secondly, the MotAB
432 model can be extrapolated to the ExbBD complex. In this model, ExbB would rotate around
433 ExbD, and this would drive the rotation of TonB, thanks to the specific interaction between
434 TonB TM and ExbB TM (in our case HasB TM and ExbB TMα2). It has also been hypothesized
435 that the C-terminal domain of TonB [34] and/or the C-terminal domain of ExbD have specific
436 binding sites for the peptidoglycan network [35]. This C-terminal domain can change
437 conformations along the catalytic cycle of protonation and deprotonation events at the
438 conserved Asp residue in ExbD TM. In addition, the periplasmic linker between the TM helix
439 and the C-terminal domain of TonB is rich in Pro and Lys residues and might adopt an extended
440 conformation of a sufficient length to span the periplasm [36]. We therefore propose a model
441 in which the force generated through rotation of TonB, driven by ExbB rotation, is not directly
442 transmitted to the TonB box of the receptor, but through the intermediary of the C-terminal
443 domain of ExbD, that might act as a wedge or a linchpin allowing force transmission and
444 converting the rotation into a pulling force exerted on the TonB/HasB box of the receptor.
445 Further studies are needed to test this model.
446

447 **Material and Methods**

448

449 **Strains and plasmid construction**

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

451 -Plasmid pBADExbBD$_{Sm}$ was constructed after amplification on the genomic DNA from strain

452 *S. marcescens* Db11 of a ca. 1.42kb fragment using primers ExbBDSm5' and ExbBDSm3'. The
453 PCR product was purified digested with *Eco*RI and *Sph*I, and ligated with pBAD24 digested with
454 *Eco*RI and *Sph*I. Correct clones were selected after sequencing of the insert.
455 -Plasmid pBADExbBD$_{Sm}$His6 was constructed by first amplifying on pBADExbBDSm a ca. 0.4kb
456 fragment with the following oligonucleotides SphIHisCtexbDSm and BglIIexbDSm; after
457 amplification, the fragment was purified, digested with *Bgl*II and *Sph*I, and ligated with
458 pBADExbBDSm digested with the same enzymes. The correct clones were selected after
459 sequencing of the insert. In biological tests, this plasmid was undistinguishable from its parent
460 plasmid pBADExbBDSm.
461 -Plasmid pBADExbBD$_{Ec}$ was constructed by amplification on genomic MG1655 DNA of a ca.
462 1.7kb fragment with the following oligonucleotides ExbBD5c, and ExbBD3c; the PCR product
463 was purified, digested with *Eco*RI and *Sph*I, and ligated with pBAD24 digested with the same
464 enzymes. Correct clones were selected after sequencing.
465 -Plasmid pBADExbB$_{Sm}$His6 was constructed as follows: a PCR fragment was amplified on
466 pBADExbBDSm with the oligos PBADFOR and ExbBHis6, digested with *Eco*RI and *Sph*I, and
467 ligated with pBAD24 digested with the same enzymes. Correct clones were selected by
468 sequencing.
469 Plasmid pBADExbBD$_{Ec\text{-}Sm76\text{-}88}$ was constructed as follows: Plasmid pBADExbBD$_{Ec}$ was amplified
470 with the two following 5' phosphorylated oligonucleotides: ExbBEcSm76-88.1 and
471 ExbBEcSm76-88.2. After digestion with *Dpn*I and self-ligation, correct clones were selected by
472 sequencing and the mutated fragments recloned in pBAD24.
473 Plasmid pBADExbBD$_{Ec\text{-}Sm76\text{-}84}$ was constructed in the same manner, with the following
474 oligonucleotides pairs: ExbBEcSm 76-84.1 and ExbBEcSm 76-84.2.
475 In the same manner, plasmid pBADExbBD$_{Sm\text{-}Ec39\text{-}51}$ was constructed by PCR on the
476 pBADExbBD$_{Sm}$ plasmid with the following couple of 5' phosphorylated oligonucleotides:
477 ExbBSmEc39-51.1 and ExbBSmEc39-51.2.
478 Plasmid pBADExbB$_{Smdelextss}$D was constructed similarly, by using as a template pBADExbBDSm,
479 with the two phosphorylated oligonucleotides ExbBdelextss1 and ExbBdelextss2.
480 To construct pBADExbB$_{Smdelextss}$His6, a ca. 0.6kb *Eco*RI-*Kpn*I fragment from pBADExbB$_{Smdelextss}$D
481 was exchanged for the corresponding fragment of pBADExbBHis6.
482 Plasmid pAMhasISRADEB, was constructed by digesting pAMhasISRADE and pSYC7 [37] by
483 *Kpn*I and *Hin*dIII, and purifying fragments of respectively ca. 9 and 7kbases, ligating them

484 together to obtain a plasmid with the whole *has* locus on a low copy number plasmid
485 (pAM238) under the control of its endogenous regulation signals.
486 All constructions were carried out in *E. coli* strain XL1-Blue.
487 Strains and other plasmids used in this study are shown in Table S2 and are from the laboratory
488 collection.
489
490 **Protein expression and purification**
491 The BL21DE3(pExbBD$_{Sm}$His6/pBAD24) or BL21DE3(pExbB$_{Sm}$His6/pBAD24) were grown at 37°C
492 either in TB or MDM medium, and induced with 0.2% arabinose at 1.5-2OD600nm (TB) or 5-
493 6OD600nm (MDM) and the incubation continued for 3hrs. The cells were harvested by
494 centrifugation (10000g 20minutes 4°C), washed once in 20mMTris-HCl pH8.0, flash-frozen in
495 liquid N_2 and kept at -80°C. Cells were broken in a Cell disruptor (Constant, UK) at 1kbar (10g
496 of cells in 40ml final of 20mM Tris pH 8.0 containing protease inhibitor (Roche, EDTA free), at
497 4°C. Benzonase was added and after ca. 15 minutes, the solution was centrifuged for 1hr at
498 100000g at 4°C. The pellet (crude membrane preparation) was resuspended in 20mM Tris pH
499 8.0 plus protease inhibitor cocktail (Roche EDTA free), flash-frozen in liquid N_2 and kept at -
500 80°C.
501 The crude membrane preparation was solubilized in 20mM Tris pH 8.0, 20mM Imidazole,
502 100mM NaCl, 10% Glycerol, 0.8% LMNG (10g of equivalent whole-cell pellet solubilized in
503 40ml), plus protease inhibitor cocktail (Roche EDTA-free) for 30 minutes at 15°C. After
504 centrifugation (1hr 100000xg), the supernatant was incubated with 2.5ml Ni-Agarose beads
505 (Thermo-Fischer His-Pure Ni-NTA #88222) preequilibrated with the same buffer except for the
506 detergent concentration (0.0015% LMNG). After three hours of incubation on a rotating wheel
507 at 4°C, the beads were washed three times with 25ml of pre-equilibration buffer and then
508 eluted with two times 25ml of the pre-equilibration buffer containing 200mM Imidazole. The
509 eluate was concentrated and washed on 100kDa cut-off centrifugal device, in pre-
510 equilibration buffer without NaCl and Imidazole. The resulting sample was loaded on a monoQ
511 HR10-100 column equilibrated with 20mM Tris pH 8.0, 10% glycerol, 0.0015%LMNG and
512 eluted with a gradient from 0 to 1M NaCl in the same buffer. Peak fractions were collected
513 and concentrated as before. The concentrated sample was then loaded on a Superose 6
514 increase column equilibrated with 20mM Tris pH8.0, 100mM NaCl, 0.0015%LMNG. The peaks
515 fractions were collected, concentrated and their concentration determined by using the

516 theoretical absorption coefficient of either ExbB5D2 (113790$M^{-1}cm^{-1}$), or ExbB5 (104850M-
517 $^{1}cm^{-1}$). They were then kept frozen at -80°C in aliquots until use. The SEC profiles as well as a
518 gel of a representative sample after purification are shown in supplementary data. Previous
519 attempts with DDM instead of LMNG yielded similar results, and LMNG was chosen, owing to
520 its very low CMC and its lower background in cryo-EM.
521
522 **Activity tests**
523 Three types of tests were used, either in liquid medium or on agar plates.
524 1. Growth tests in liquid medium: *E. coli* C600 and its *ΔexbBD* derivatives were transformed
525 with specified plasmids and their growth tested in liquid LB medium at 37°C, at various
526 dipyridyl concentrations to induce iron starvation. OD_{600nm} was measured after overnight
527 growth.
528 2. Growth tests on plates: the relevant plasmids were also transformed into *E. coli* C600
529 *ΔhemA*, a heme auxotroph strain (and derivatives thereof), and growth of the strains was
530 assayed as follows. Briefly, cells were first grown in LB medium (supplemented with delta-
531 aminolevulinic acid (50µg/ml) to bypass the effect of the *ΔhemA* mutation) at 37°C up to an
532 OD_{600nm} of 1, and then mixed with melted top agar (0.6%agar in LB), and poured onto LB agar
533 plates containing the appropriate antibiotics with arabinose at a concentration of 40µg/ml to
534 induce expression of the genes under the P_{araBAD} promoter[38]. Wells, punched with Pasteur
535 pipettes were filled with heme-albumin (at 5, 1 or 0.2µM), at two dipyridyl concentrations
536 (100 or 200µM) to induce iron starvation. Plates were incubated overnight at 37°C and scored
537 for growth around the wells. All experiments were performed in triplicate.
538 3. Growth curves in liquid medium: a few colonies of *E. coli* C600
539 *ΔhemAΔtonBΔexbBD*(pAMhasISRADEB + pBAD24 or derivatives thereof) were first inoculated
540 in 4ml of LB medium at 37°C with the corresponding antibiotics, and 100µM dipyridyl, 4µg/ml
541 arabinose but without delta-aminolevulinic acid. Once the culture reached an OD_{600nm} of ca.
542 1.2-1.5, it was diluted and inoculated in 48 well Greiner plates, in the same medium to which
543 was added 0.4µM He-BSA, as a heme source. The initial OD_{600nm} of the cultures was 0.001.
544 Each well contained 300µl of growth medium. Duplicates of each strain were made, and the
545 plate was incubated at 37°C with vigorous shaking (500rpm) in a Clariostar Plus Microplate
546 reader. OD_{600nm} was recorded every 30 minutes for 70 hours.
547

548 **N-terminal sequencing**

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

552

553 **Lipid extraction**

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

559

560 **Mass Spectrometry**

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

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

575

576 **Cryo-EM grid preparation and data acquisition**

577 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
578 holey carbon grids (Protochips Inc., USA) previously glow-discharged in air for 30s. Grids were
579 blotted for 2s at blot force 1 and vitrified in liquid ethane using a Vitrobot mark IV (FEI
580 company) operated at 10 °C and 100% relative humidity.

581 All data collection was performed with a Titan Krios (ThermoFisher Scientific) operated at 300
582 kV equipped with a K2 Summit direct electron detector (ThermoFisher Scientific) at the
583 European synchrotron research facility, ESRF (Grenoble, France). Movies were recorded in
584 electron counting mode with EPU software (ThermoFisher Scientific), aligned with Motioncor2
585 [39] and aligned images were processed with Gctf [40] using Scipion interface[41].

586 For ExbB data collection, 4567 movies were collected at a magnification of 165000x with a
587 nominal pixel size of 0.827Å using a defocus range from -1.5 to -2.5µm. Movies of 56 frames
588 were acquired using a dose rate of 8 electrons/$Å^2$/second over 7 seconds, yielding a
589 cumulative exposure of 55.95 electrons/$Å^2$.

590 For ExbBD data collection, 4043 movies were collected at a nominal magnification of 139000
591 with a pixel size of 1.067Å. Movies of 48 frames were acquired using a dose rate of 4.6
592 electrons/$Å^2$/second over 12 seconds yielding a cumulative exposure of 55.2 electrons/$Å^2$.

593

594 **Cryo-EM image processing and analysis**

595 **a) ExbB**

596 For ExbB, aligned movies were processed with Gctf [40] and only images with a resolution
597 higher than 4Å were kept; after visual inspection of the remaining images, processing was
598 carried out with Relion-3 [42]. Particles were extracted using a 2-fold binning, issued from a
599 manual picking and a 2-D classification of particles picked out from on 50 images. Automatic
600 extraction was performed using the selected 2D class averages. After several rounds of 2D and
601 3D classification, 161k particles were selected for 3D refinement. They were corrected for local
602 motion using Bayesian polishing option in Relion-3 and a post-refined map produced a 3.1Å
603 overall resolution with a 5-fold symmetry (Supplementary Figure S17).

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

608 glycerol starting structure and geometry were built using Phenix eLBOW [45] and was fit in
609 the map using Coot then refined along with the protein model in Phenix.real_space_refine.
610
611 **b) ExbBD**
612 Movies were drift-corrected and dose-weighted using MotionCorr2 [39]. Aligned dose
613 weighted averages were imported into Cryosparc2 [46] and contrast transfer function was
614 estimated using CTFFIND4.1 [47]. Micrographs with poor CTF estimation statistics or high drift
615 profiles were discarded. The remaining 3028 micrographs were used for automated particle
616 picking. Particles were extracted, Fourier cropped to 2 Å/px and 2D classified. The best 2D
617 classes were used as templates for automated particle picking resulting in 1.3 million particles.
618 After several rounds of classification, the best 600k particles were submitted to 3D
619 classification by means of multi-class ab-initio Reconstruction and Heterogeneous
620 Refinement. 158 k particles belonging to the best resolved classes were corrected for local
621 motion, re-extracted and used in Non-Uniform Refinement. The resulting refined map has a
622 nominal resolution of 4.56 Å.
623 Based on the previous map a soft mask lining the micelle was designed in UCSF Chimera [48]
624 and used to signal subtract the corresponding micelle density of particles in the refined map.
625 Localized refinement of the signal subtracted particles produced a map of the complex with
626 an estimated resolution of 3.96 Å judging by FSC at 0.143 criterion (Supplementary Figure
627 S18). Data have been deposited both at the EMDB and PDB databases (EMDB 10789 and
628 11806, PDB 6YE4 and 7AJQ, for $ExbB_{Sm}$ and $ExbBD_{Sm}$ respectively, see Table I for refinement
629 statistics).
630 Unless otherwise specified, the structure figures were made using UCSF Chimera software
631 [48].
632
633 **NMR experiments**
634 The C-terminal periplasmic domain of HasB, $HasB_{CTD}$ comprising residues 133-263, was
635 produced and purified as previously reported [14, 30].The peptide corresponding to the
636 periplasmic extension of $ExbB_{sm}$
637 (A_1PAANPAVTESVAPTTAPAPAAAAPESITPVNPAPTIQPPETRG$_{44}$- numbering with reference to
638 the mature protein) was synthetized by *Proteogenix*.

639 NMR experiments were acquired at 293 K on a 600 MHz Bruker Avance III spectrometer
640 equipped with a TCI cryoprobe. The spectra were processed with NMRpipe [49] and analyzed
641 with CcpNmr Analysis 2.4 software [50]. Proton chemical shifts were referenced to 2,2-
642 dimethyl-2-silapentane-5 sulfonate as 0 ppm. ^{15}N were referenced indirectly to DSS [51]
643 (Wishart et al., 1995). ^{1}H–^{15}N HSQC experiments were acquired on 0.15 mM $HasB_{CTD}$ in 50mM
644 sodium phosphate pH 7, 50 mM NaCl with or without the peptide. Aliquots from a solution of
645 peptide at 25mg/ml in the same buffer were added to the protein sample at 2:1 and 10:1
646 ratios. Chemical shift perturbations (CSPs) of backbone amide cross-peaks were quantified by
647 using the equation $CSP = [\Delta \delta H^2 + (\Delta \delta N \times 0.15)^2]^{1/2}$, where Δ δH and Δ δN are the observed ^{1}H
648 and ^{15}N chemical shift differences between the two experimental conditions. The ^{1}H and ^{15}N
649 resonance assignments were from Lefevre et al 2007 [52].
650
651 **Other biochemical methods**
652 SDS-PAGE and immunodetection with anti-HasB or anti-His6 antibodies (Abcam [HIS.H8]
653 (ab18184)) on whole cells or membrane preparations were carried out following standard
654 protocols. Secondary antibodies were coupled to alkaline phosphatase.
655
656 **BLAST search**
657 $ExbB_{Sm}$ was used as a search for BLAST for orthologs in complete bacterial genomes, focusing
658 on "long" ExbB's. Top hits were in *Serratia*, *Yersinia*, *Dickeya*, *Erwinia* and *Pseudomonas*. Those
659 genera were later excluded from successive BLAST searches to obtain orthologs in other
660 genera with higher p-value.
661
662 **Acknowledgments** This work was supported by grants from the ANR (HEMESTOCKEXCHANGE
663 ANR-12-BSV3-0022-01 and LABEX DYNAMO ANR-11-LABEX-0011-01). We thank Nathalie
664 Dautin for helpful suggestions and critical reading of the manuscript. We acknowledge
665 Emmanuel Frachon and Christophe Thomas from the Institut Pasteur platform "Production of
666 recombinant proteins" for large scale cultures used in this work, European Synchrotron
667 Radiation Facility for the provision of time on the Titan Krios at beamline CM01, the Institut
668 Pasteur Biological NMR Technological Platform for the use of the NMR spectrometers, the
669 staff at the ESRF CM01 facility in Grenoble for help in the data collection, the staff at the
670 Marseille proteomics facility and Djalila Djouadi and Diana Basto who initially contributed to
671 this project. We thank M. Nilges for interest in this work.
672
673

674 **Figure legends**
675 Figure 1:
676 Weblogo (http://weblogo.threeplusone.com/) representation of 131 « long » ExbB
677 aminoacid sequences; those with much shorter periplasmic extensions have been excluded
678 from the alignment. The orange boxes show the position of the three TM segments from the
679 *E. coli* 5SV0 structure.
680
681 Figure 2:
682 Representation of the pentameric structure of $ExbB_{Sm}$ solved by cryo-EM. A: the density map
683 is shown, each monomer with distinct shades of blue, view from the side. The grey densities
684 represent PG molecules. B: view from the cytoplasmic side. C: representation of one PG
685 density together with the fitting of the PG molecule inside the density, and with the stick
686 representation of the nearby ExbB residues using the Coot software [44]. D: superposition of
687 a monomer of $ExbB_{Ec}$ (dark grey) and $ExbB_{Sm}$ (red), represented as ribbon. The membrane
688 thickness is represented by the dotted lines and the transmembrane segments TM1 (part of
689 α2 helix), TM2 (part of α6 helix) and TM3 (part of α7 helix) are indicated.
690
691 Figure 3:
692 Structure of $ExbBD_{Sm}$ solved by cryo-EM. A: side view, the same color code as in Figure 2 for
693 the ExbB monomers is used. A: side view; B: view from the periplasmic space; the two ExbD
694 helices are seen respectively in red and magenta. C: clipped side view, with the helix of each
695 ExbD monomer represented (red and magenta). D and E, side and top views showing the
696 model built in the density maps.
697
698 Figure 4:
699 Interaction of $HasB_{CTD}$ with the periplasmic fragment of $ExbB_{Sm}$, as detected by NMR. $HasB_{CTD}$
700 residues shown on its structure (PDB code 2M2K) exhibiting the highest CSP in the presence
701 of the periplasmic peptide of $ExbB_{sm}$ are coloured in red (Top: surface representation;
702 bottom: cartoon representation). Right: Superposed ^{1}H–^{15}N HSQC spectra of 0.15 mM ^{15}N-
703 labelled $HasB_{CTD}$ in 50mM sodium phosphate, pH 7, 50 mM NaCl in the presence (red) or
704 absence (black) of the periplasmic peptide of $ExbB_{sm}$.
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719

720

721

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

722

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

724

725

726

727

728

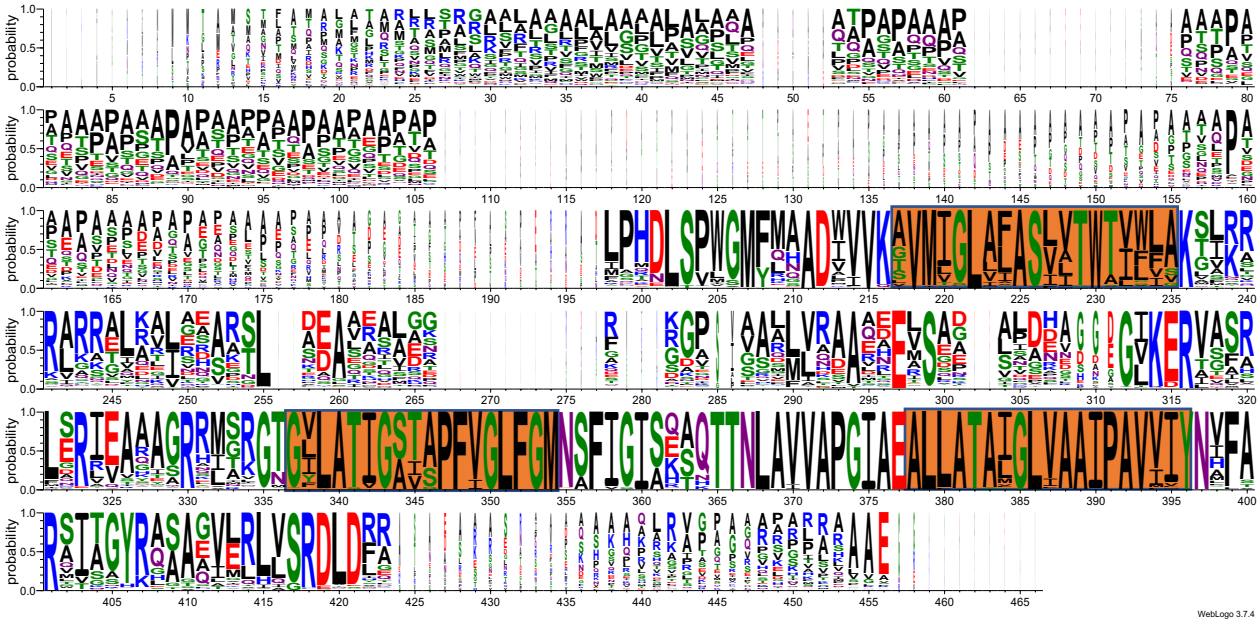
729

730

731

732

Figure 1



761 Figure 2
762
763

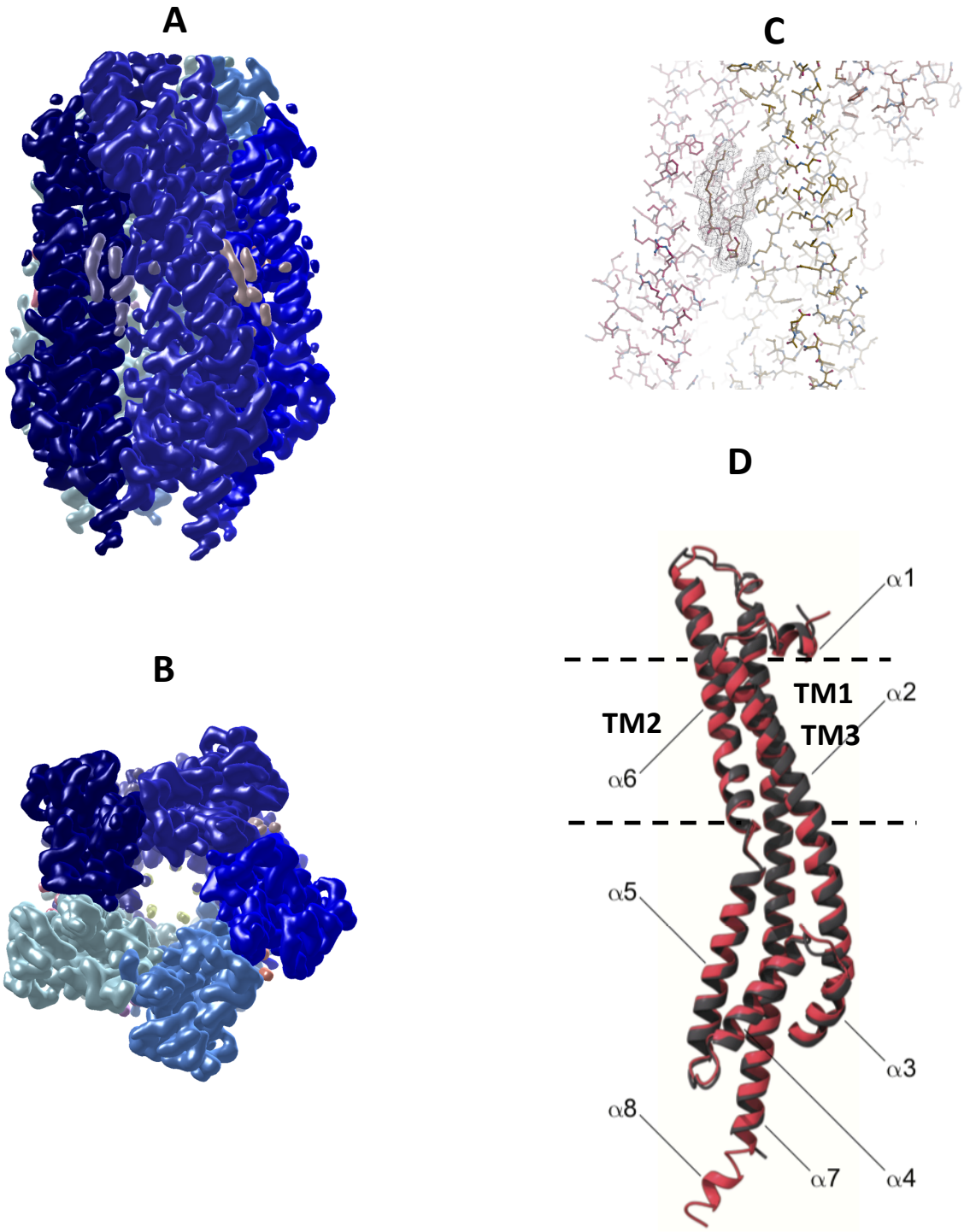


Figure 3

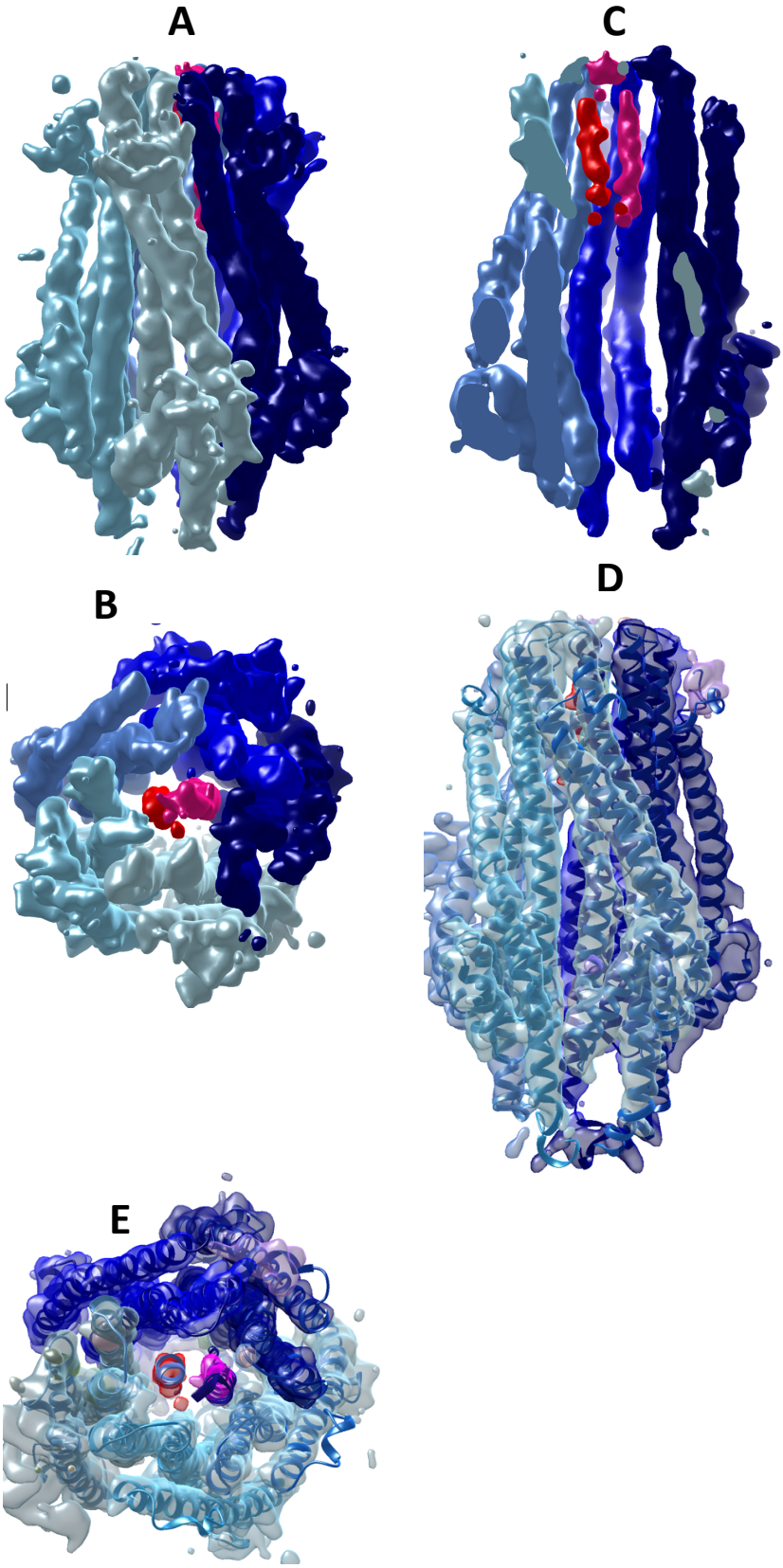
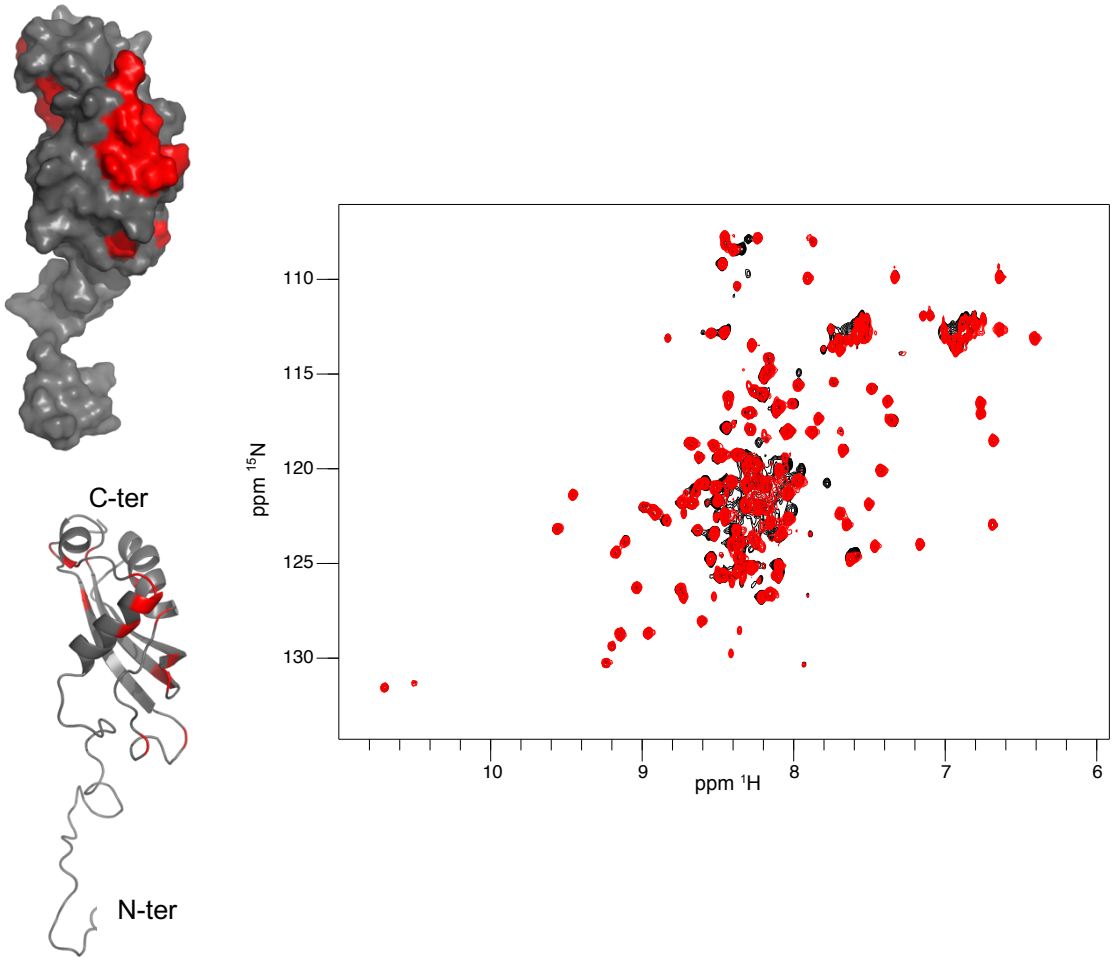


Figure 4



848 **References**
849

850 1. Noinaj, N., et al., *TonB-dependent transporters: regulation, structure, and function.*
851 Annu Rev Microbiol, 2010. **64**: p. 43-60.
852 2. Grinter, R. and T. Lithgow, *The structure of the bacterial iron-catecholate transporter*
853 *Fiu suggests that it imports substrates via a two-step mechanism.* J Biol Chem, 2019.
854 **294**(51): p. 19523-19534.
855 3. Cascales, E., R. Lloubes, and J.N. Sturgis, *The TolQ-TolR proteins energize TolA and*
856 *share homologies with the flagellar motor proteins MotA-MotB.* Mol Microbiol, 2001.
857 **42**(3): p. 795-807.
858 4. Braun, V., *The structurally related exbB and tolQ genes are interchangeable in*
859 *conferring tonB-dependent colicin, bacteriophage, and albomycin sensitivity.* J
860 Bacteriol, 1989. **171**(11): p. 6387-90.
861 5. Eick-Helmerich, K. and V. Braun, *Import of biopolymers into Escherichia coli:*
862 *nucleotide sequences of the exbB and exbD genes are homologous to those of the*
863 *tolQ and tolR genes, respectively.* J Bacteriol, 1989. **171**(9): p. 5117-26.
864 6. Braun, V. and C. Herrmann, *Evolutionary relationship of uptake systems for*
865 *biopolymers in Escherichia coli: cross-complementation between the TonB-ExbB-ExbD*
866 *and the TolA-TolQ-TolR proteins.* Mol Microbiol, 1993. **8**(2): p. 261-8.
867 7. Fischer, E., K. Gunter, and V. Braun, *Involvement of ExbB and TonB in transport across*
868 *the outer membrane of Escherichia coli: phenotypic complementation of exb mutants*
869 *by overexpressed tonB and physical stabilization of TonB by ExbB.* J Bacteriol, 1989.
870 **171**(9): p. 5127-34.
871 8. Ahmer, B.M., et al., *Characterization of the exbBD operon of Escherichia coli and the*
872 *role of ExbB and ExbD in TonB function and stability.* J Bacteriol, 1995. **177**(16): p.
873 4742-7.
874 9. Higgs, P.I., P.S. Myers, and K. Postle, *Interactions in the TonB-dependent energy*
875 *transduction complex: ExbB and ExbD form homomultimers.* J Bacteriol, 1998.
876 **180**(22): p. 6031-8.
877 10. Braun, V., et al., *Energy-coupled transport across the outer membrane of Escherichia*
878 *coli: ExbB binds ExbD and TonB in vitro, and leucine 132 in the periplasmic region and*
879 *aspartate 25 in the transmembrane region are important for ExbD activity.* J
880 Bacteriol, 1996. **178**(10): p. 2836-45.
881 11. Gaisser, S. and V. Braun, *The tonB gene of Serratia marcescens: sequence, activity*
882 *and partial complementation of Escherichia coli tonB mutants.* Mol Microbiol, 1991.
883 **5**(11): p. 2777-87.
884 12. Paquelin, A., et al., *Characterization of HasB, a Serratia marcescens TonB-like protein*
885 *specifically involved in the haemophore-dependent haem acquisition system.* Mol
886 Microbiol, 2001. **42**(4): p. 995-1005.
887 13. Benevides-Matos, N., C. Wandersman, and F. Biville, *HasB, the Serratia marcescens*
888 *TonB paralog, is specific to HasR.* J Bacteriol, 2008. **190**(1): p. 21-7.
889 14. Lefevre, J., et al., *Modulation by substrates of the interaction between the HasR outer*
890 *membrane receptor and its specific TonB-like protein, HasB.* J Mol Biol, 2008. **378**(4):
891 p. 840-51.
892 15. Celia, H., et al., *Structural insight into the role of the Ton complex in energy*
893 *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.

939 35. Boags, A.T., F. Samsudin, and S. Khalid, *Binding from Both Sides: TolR and Full-Length*
940 *OmpA Bind and Maintain the Local Structure of the E. coli Cell Wall.* Structure, 2019.
941 **27**(4): p. 713-724 e2.
942 36. Kohler, S.D., et al., *The proline-rich domain of TonB possesses an extended polyproline*
943 *II-like conformation of sufficient length to span the periplasm of Gram-negative*
944 *bacteria.* Protein Sci, 2010. **19**(4): p. 625-30.
945 37. Létoffé, S., J.M. Ghigo, and C. Wandersman, *Identification of two components of the*
946 *Serratia marcescens metalloprotease transporter: protease SM secretion in*
947 *Escherichia coli is TolC dependent.* J Bacteriol, 1993. **175**(22): p. 7321-8.
948 38. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by*
949 *vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-
950 30.
951 39. Zheng, S.Q., et al., *MotionCor2: anisotropic correction of beam-induced motion for*
952 *improved cryo-electron microscopy.* Nat Methods, 2017. **14**(4): p. 331-332.
953 40. Zhang, K., *Gctf: Real-time CTF determination and correction.* J Struct Biol, 2016.
954 **193**(1): p. 1-12.
955 41. Gomez-Blanco, J., et al., *Using Scipion for stream image processing at Cryo-EM*
956 *facilities.* J Struct Biol, 2018. **204**(3): p. 457-463.
957 42. Zivanov, J., et al., *New tools for automated high-resolution cryo-EM structure*
958 *determination in RELION-3.* Elife, 2018. **7**.
959 43. Afonine, P.V., et al., *Real-space refinement in PHENIX for cryo-EM and*
960 *crystallography.* Acta Crystallogr D Struct Biol, 2018. **74**(Pt 6): p. 531-544.
961 44. Emsley, P., et al., *Features and development of Coot.* Acta Crystallogr D Biol
962 Crystallogr, 2010. **66**(Pt 4): p. 486-501.
963 45. Moriarty, N.W., R.W. Grosse-Kunstleve, and P.D. Adams, *electronic Ligand Builder*
964 *and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint*
965 *generation.* Acta Crystallogr D Biol Crystallogr, 2009. **65**(Pt 10): p. 1074-80.
966 46. Punjani, A., et al., *cryoSPARC: algorithms for rapid unsupervised cryo-EM structure*
967 *determination.* Nat Methods, 2017. **14**(3): p. 290-296.
968 47. Rohou, A. and N. Grigorieff, *CTFFIND4: Fast and accurate defocus estimation from*
969 *electron micrographs.* J Struct Biol, 2015. **192**(2): p. 216-21.
970 48. Pettersen, E.F., et al., *UCSF Chimera--a visualization system for exploratory research*
971 *and analysis.* J Comput Chem, 2004. **25**(13): p. 1605-12.
972 49. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on*
973 *UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
974 50. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: development of a*
975 *software pipeline.* Proteins, 2005. **59**(4): p. 687-96.
976 51. Wishart, D.S., et al., *1H, 13C and 15N chemical shift referencing in biomolecular NMR.*
977 J Biomol NMR, 1995. **6**(2): p. 135-40.
978 52. Lefèvre, J., et al., *(1)H, (13)C and (15)N resonance assignments of the C-terminal*
979 *domain of HasB, a specific TonB like protein, from Serratia marcescens.* Biomol NMR
980 Assign, 2007. **1**(2): p. 197-9.
981