Structural and molecular determinants for the interaction of ExbB from *Serratia marcescens* and HasB, a TonB paralog.

Valérie Biou[1#], Ricardo Jorge Diogo Adaixo[2], Mohamed Chami[2], Pierre-Damien Coureux[3], Benoist Laurent[4], Véronique Yvette Ntsogo Enguéné [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

Running title: Interaction specificity between ExbB and TonB/HasB

For correspondence: # Valérie Biou, valerie.biou@ibpc.fr,

Philippe Delepelaire, philippe.delepelaire@ibpc.fr

* Present addresses:
Véronique Yvette Ntsogo Enguéné: Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK
Gisele Cardoso de Amorim: Instituto de Bioquímica Médica, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
Henning Stahlberg: Centre d'imagerie Dubochet UNIL-EPFL-UNIGE & Laboratoire de microscopie électronique biologique UNIL-EPFL, Lausanne, Switzerland

Abstract.

ExbB and ExbD are cytoplasmic membrane proteins that associate with TonB to convey the energy of the proton-motive force to outer membrane receptors in Gram-negative bacteria for iron uptake. The opportunistic pathogen *Serratia marcescens* (*Sm*) possesses both TonB and a heme-specific TonB paralog, HasB. $ExbB_{Sm}$ has a long periplasmic extension absent in other bacteria such as *E. coli (Ec)*. Long ExbB's are found in several genera of Alphaproteobacteria, most often in correlation with a *hasB* gene. We investigated specificity determinants of $ExbB_{Sm}$ and HasB. We determined the cryo-EM structures of $ExbB_{Sm}$ and of the ExbB-$ExbD_{Sm}$ complex from *S. marcescens*. $ExbB_{Sm}$ alone is a stable pentamer, and its complex includes two ExbD monomers. We showed that $ExbB_{Sm}$ extension interacts with HasB and is involved in heme acquisition and we identified key residues in the membrane domain of $ExbB_{Sm}$ and $ExbB_{Ec}$, essential for function and likely involved in the interaction with TonB/HasB. Our results shed light on the new class of inner membrane energy machinery formed by ExbB,ExbD and HasB.

Introduction

Transport of nutrients across the Gram-negative outer membrane is either a diffusion-facilitated or an active process. In the latter case, the process is powered by the energy of the proton-motive force (pmf) transmitted through the periplasm to specialized outer membrane (OM) 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. Those proteins are involved in power generation and transmission in various processes (MotAB drives flagellar rotation, the ExbB-ExbD complex referred to as ExbBD energises active transport of molecules across OM receptors and TolQR is involved in cell division), at the expense of pmf dissipation across the cytoplasmic membrane (3-6). They form complexes that associate respectively with the flagellar rotor, TonB and TolA. It has also been shown that ExbD does not accumulate in the absence of ExbB, and that TonB does not accumulate in the absence of the ExbBD complex (7). The C-terminal domains of TonB and ExbD reside in the periplasm and interact with each other (8, 9). ExbD TM (Trans-Membrane segment) has one strictly conserved aspartate residue (10), which is thought to undergo cycles of protonation/deprotonation (coupled to pmf dissipation) and is essential to its function.

TBDT's comprise a 22-stranded ß-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.

The N-terminus of TonB is localized in the cytoplasm, followed by a single transmembrane segment, and by a Pro-Lys rich region long enough to span the periplasmic space and ends with a structured C-terminal domain interacting with the TBDT TonB box. The substrate binds the extracellular side of the TBDT receptor, triggering conformational changes that are transmitted to the periplasmic side and allow the interaction between TonB and the TonB box of the TBDT, leading to the substrate entry into the periplasm by a yet unknown mechanism. A rearrangement of the plug domain has been proposed to occur creating a path for the substrate across the TBDT.

In *Serratia marcescens*, a close relative of *E. coli*, in which about twenty potential TBDT's were identified, there are at least two TonB homologs: an ortholog of $TonB_{Ec}$ (48% aminoacid identity with $TonB_{Ec}$(11)) and HasB, a TonB paralog that is strictly dedicated to its cognate outer membrane receptor HasR(12, 13). HasB has the same topology as TonB, and its C-terminal domain interacts specifically with HasR with a circa (ca.) 40-fold higher affinity than the corresponding TonB domain (14). 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. The Has system has been functionally reconstituted in *E. coli*. Unlike $TonB_{Sm}$, HasB could not complement $TonB_{Ec}$ functions (12), nor was it 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). This mutation corresponds to a 6 base-pair duplication leading to a longer TM segment for HasB by inserting AL into CLVLVLALHLLVAALLWP resulting in CLVLVLALALHLLVAALLWP.

Recently several structures of *E. coli* ExbB and ExbBD have been solved, either by X-ray crystallography or cryo-EM(14-16). All samples studied included ExbB and ExbD, but not all showed an ordered structure of ExbD. In these structures, $ExbB_{Ec}$(14) appeared as a pentamer. Each monomer of ExbB has 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 (15). A more recent cryo-EM study of ExbBD reconstituted in nanodiscs however confirmed the $ExbB_5ExbD_2$ stoichiometry with two TM helices of ExbD identified in the ExbB pore (16). This is consistent with previous DEER (Double Electron Electron Resonance) results (15). Similarly, *Pseudomonas savastanoi* ExbBD exhibits the same stoichiometry (17). Structures of the flagellar MotAB motor complexes (related to ExbBD) from several bacteria (*Campylobacter jejuni, Bacillus subtilis, Clostridium sporogenes, Vibrio mimicus, Shewanella oneidensis,* PomAB from *Vibrio alginolyticus*) were recently published. They all display a $MotA_5MotB_2$ stoichiometry and share some topology elements with ExbBD complex (17, 18). The discovery of HasB and its functional specificities prompted us to identify the ExbBD complex that would function with it in *S. marcescens*. In this study, we identified the orthologous $ExbBD_{Sm}$ and found that this complex is active with both HasB and TonB. We characterised 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 purified $ExbB_{Sm}$ alone and $ExbBD_{Sm}$ and determined their structures by single-particle cryo-EM at 3.1 and 3.96Å resolution, respectively. We show that in both cases, $ExbB_{Sm}$ behaves as a stable

pentamer; in the $ExbBD_{Sm}$ complex, we observe 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 C-terminal 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 of ExbB 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 an 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 *E. coli*, the *exbBD* operon is surrounded on the 5' side, and in the opposite direction, by *metC* (encoding a cystathionine lyase) and on the 3' side by *yhgA* (encoding an aldehyde reductase). In *S. marcescens* Db11, the *exbB*-like gene is close to and in the opposite orientation to a *metC* homolog, as in *E. coli.* At the 3' end of the operon, and in the opposite direction there are genes related to sucrose metabolism. 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

147 corresponds to a putative signal peptide followed by a ~50 residue N-terminal extension of the
148 mature protein (Figure 1A).
149 A BLAST search (19) on bacterial genomes indicated that such "extralong" ExbB's are found
150 in different Gammaproteobacteria, including *Serratia*, *Yersinia*, *Pseudomonas*,
151 *Erwinia*/*Dickeya*, and many genera in the Alphaproteobacteria class. Table I lists representative
152 species of Alphaproteobacteria in which such long ExbB's are found. Interestingly, about 90%
153 of the long ExbB-containing species listed in this table also have a *hasB* gene ortholog.
154 An alignment of these ExbB aminoacid sequences (represented as a Logo on Figure 1B) also
155 shows that the mature N-terminal extensions are of variable lengths between a few and 150
156 residues with an average of 50 residues. These extensions are quite rich in Ala (24,6%) and Pro
157 (14,5%) residues and therefore likely to be unstructured but could be involved in protein-protein
158 interactions as proline-rich regions often are in signalling processes (20-22).
159 The Conserved Domain Database (23) currently contains two ExbB subfamilies in the cI00568
160 MotA_ExbB superfamily, TIGR02797 (containing *E. coli* ExbB, as well as longer ExbB's) and
161 TIGR02805 (containing *Haemophilus influenzae* ExbB, with a very short cytoplasmic domain
162 between TM1 and TM2). Based on sequence data, TIGR02797 can be divided into 2 groups:
163 those with "extralong" ExbB's in one subfamily, and *E. coli*-like sequences in the other. Along
164 with the existence of HasB in the bacterial species, the identification of this N-terminal addition
165 prompted us to further characterise the ExbB-ExbD complex from *S. marcescens*.
166
167 **2. ExbBD$_{Sm}$ complements ExbBD$_{Ec}$ for iron acquisition and functionally associates with**
168 **both TonB and HasB for heme acquisition through HasR.**
169

As a first step in the characterization of the identified $ExbBD_{Sm}$, we tested whether this complex complements $ExbBD_{Ec}$. To this end, the plasmids $pBADExbBD_{Sm}$ and $pBADExbBD_{Ec}$ were constructed, introducing the two genes from *S. marcescens* or *E. coli*, into pBAD24 vector under the control of arabinose-inducible 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 Fe^{2+} chelator dipyridyl, DiP) as ExbBD are required to transduce pmf to TonB for siderophore uptake. Its growth in the presence of DiP is restored by the expression of either $pBADExbBD_{Ec}$ or $pBADExbBD_{Sm}$ (Figure 2).

We then tested the functionality of ExbBDSm in heme acquisition via the Has system reconstituted in a heme auxotroph *E. coli* strain. To avoid interference from the chromosomal *exbBD* operon and from *tonB*, we used the C600Δ*hemA*Δ*exbBD* and the C600Δ*hemA*Δ*exbBD*Δ*tonB* strains, transformed with various recombinant plasmid pairs, one bringing the complete Has locus, with or without the *hasB* gene, the other one with the *exbBD* operon or its derivatives under the control of the P_{araBAD} promoter. In this kind of experiment, the bacterial growth reflects both the expression of the various Has components under the control of their own regulatory elements (24), and the efficiency of the heme uptake process itself. Two kinds of tests were carried out, one on solid substrate in Petri dishes, allowing to see haloes of bacterial growth around wells punched in the agar and containing the heme source, the other one in liquid medium in microplates with an absorbance plate reader, allowing to record growth at regular intervals over extended periods. The latter type of tests allows a more precise and quantitative description of the phenotypes as it is realized in a homogeneous medium, which is not the case with Petri dishes. Six strains were constructed: one set deleted from TonB C600Δ*hemA*Δ*exbBD*Δ*tonB* (pAMHasISRADEB + pBAD24),

194 C600Δ*hemA*Δ*exbBD*Δ*tonB* (pAMHasISRADEB + pBADExbBD$_{Ec}$),
195 C600Δ*hemA*Δ*exbBD*Δ*tonB* (pAMHasISRADEB + pBADExbBD$_{Sm}$), and one set with TonB
196 C600Δ*hemA*Δ*exbBD* (pAMHasISRADE + pBAD24), C600Δ*hemA*Δ*exbBD* (pAMHasISRADE
197 + pBADExbBD$_{Ec}$), and C600Δ*hemA*Δ*exbBD* (pAMHasISRADE + pBADExbBD$_{Sm}$). The
198 ExbBD constructs for Ec and Sm are schematized on the first two lines of Figure 3A. The
199 results are reported in Figures 3B (Petri dishes, overnight observation), 3C (Petri dishes, 36h
200 observation) and 3D (microplates, recording over 66h). As expected, control strains (with
201 pBAD24) did not grow. ExbBD$_{Sm}$ was functional with both HasB and TonB (middle series of
202 holes in Figure 3C indicated with orange and blue dots, and orange and blue dots curves in
203 Figure 3D), although with quite dramatically different kinetics and yield, the onset of growth
204 occurring at 5-6hrs for the HasB-ExbBD$_{Sm}$ pair, and at 22-23hrs for the TonB-ExbBD$_{Sm}$ pair.
205 Moreover, the final OD of TonB-ExbBD$_{Sm}$ is half that of HasB-ExbBD$_{Sm}$. Under our
206 experimental setup, ExbBD$_{Ec}$ was also functional with both TonB and more surprisingly with
207 HasB (bottom series of holes in Figure 3C, and green and grey dots curves in Figure 3D), the
208 onset of growth occurred at around 15-16hrs for TonB, and 18hrs for HasB; however, the
209 maximal OD was lower for the HasB-ExbBD$_{Ec}$ pair (0.43 vs. 0.72). Finally, the ExbBD$_{Sm}$-
210 HasB pair seems less sensitive to iron starvation than all the other ones, since similar results
211 are obtained at 100 and 200µM Dipyridyl (Figure 3B), which is not the case for the other pairs,
212 as already observed in the TonB-ExbBD$_{Ec}$ case (25). These results might seem at odds with the
213 previous results obtained on HasB (12) which showed that HasB was nonfunctional with
214 ExbBD$_{Ec}$. However, the experimental setup was quite different, both in terms of plasmids used,
215 strains and conditions of observation. Therefore, ExbBD$_{Sm}$ is the *E. coli* ExbBD ortholog, able
216 to associate both with HasB and TonB$_{Ec}$.
217

3. Characterization of $ExbB_{Sm}$ and $ExbBD_{Sm}$: Specific function of the N-terminal extension of $ExbB_{Sm}$

3.a Purification and mass-spectrometry analyses of protein and associated lipids

To gain further insight into the possible differences between $ExbBD_{Sm}$ and $ExbBD_{Ec}$, we purified both $ExbB_{Sm}$ and $ExbBD_{Sm}$ (as C-terminal 6-His-tagged proteins on ExbB for ExbB purification, and C-terminal 6-His tag on ExbD for ExbBD purification, see Materials and Methods section) in LMNG (lauryl maltose neopentyl glycol) micelles (Figure 4, A and B). Both $ExbB_{Sm}$ and $ExbBD_{Sm}$ appear as homogeneous oligomers at the last size exclusion chromatography (SEC) step.

Mass spectrometry analysis showed that purified $ExbD_{Sm}$His6 has a mass of 16 131 Da close to its theoretical mass of 16261.7 Da. Purified $ExbB_{Sm}$ has a measured mass of 29 557 Da and its predicted mature sequence has a theoretical mass of 29556.8 Da. Together with the determination of the first amino-acid residue for ExbB, this confirmed 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 S1). More interestingly, mass spectrometry analysis of chloroform/methanol extracts of the purified proteins also showed that both complexes contained native lipids (Supplementary Figure S2), mostly phosphatidylglycerol (PG) and phosphatidylethanolamine (PE), and that ExbBD also contained some cardiolipin (CL) (Supplementary Figure S3). Further analysis of aliphatic chain composition of lipids shows evidence of specific composition with a majority of PG with 34 carbons and 2 unsaturations (Supplementary Figure S4). Considering that the $ExbBD_{Sm}$ complex is functional in *E. coli*, and that the phospholipid composition of *S. marcescens* is quite similar to that of *E. coli* (26), this affinity for lipids may disclose a specific recognition that may be important for the function.

3b NMR shows an interaction between $HasB_{CTD}$ and $ExbB_{Sm}$ periplasmic extension.

The presence in $ExbB_{Sm}$ of an N-terminal extension residing in the periplasm led us to investigate its function. This extension contains more than 50% of Ala and Pro residues and is predicted to be unstructured by disorder prediction servers such as IUPRED2 (27). 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 (our 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 potential 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 because of their interaction (Figure 5 A and Supplementary Figure S5). Perturbed residues are mainly located on the helical face of $HasB_{CTD}$ forming a continuous surface of interaction (Figure 5B and 5C) (R175, R178-K180, K184, Q192, T200, L201, Q204, H206, A232, A240, G246). Interestingly, this face is on the opposite side of the third beta strand of HasB that was previously shown to interact with HasR (28). In addition, the residues of a small pocket at the C-terminus of HasB (D255, R259) show also high CSP and might be involved in the interaction with ExbB or subject to a conformation change induced by this interaction.

3c 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 $ExbB_{Smdelextss}$ mutant lacking this extension (Figure 3A) (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 in liquid

266 culture. We observed that the lag between the start of the experiment and the onset of growth
267 is at least 2 hours longer for the $ExbB_{Smdelextss}$ mutant strain, as compared to the WT strain
268 (Figure 6A, compare orange and yellow curves) and a control strain (medium blue). In addition,
269 the maximum OD reached after growth is slightly but significantly decreased (Figure 6 A). To
270 ensure that this difference was not due to a difference of expression of the two constructs, we
271 compared their amounts in membrane preparations by immunodetection. As our anti-$ExbB_{Sm}$
272 antibody was not sensitive enough for this measurement, we used the His-tagged version of
273 $ExbB_{Sm}$ readily detectable with anti-His6 antibodies. Coomassie blue stained SDS-PAGE
274 showed that the 3 conditions have similar total amounts of protein (Figure 6B). The Western
275 blot shows that the amount of the $ExbB_{SmHis6}$ variant deleted of its N-terminal extension is at
276 least equal, if not slightly higher, to the amount of the wild-type protein (Figure 6C, lanes 2 and
277 3), ruling out a possible decrease in protein concentration. The N-terminal periplasmic
278 extension of $ExbB_{Sm}$, is therefore likely involved in the functioning of the Has system, whether
279 at the transcription activation of the Has locus, and/or at any later step.
280

4 $ExbB_{Sm}$ and $ExbBD_{Sm}$ structural analysis by cryo-EM show $ExbB_5$ and $ExbB_5D_2$ stoichiometries.

4a $ExbB_{Sm}$ structure

As $ExbB_{Sm}$ presented specific features compared to $ExbB_{Ec}$, we set out to determine its structure by single-particle cryo-EM, see Material and Methods for details. The 3D class average model clearly showed a pentameric structure. It was refined and polished to obtain a resolution of 3.1Å using the Fourier shell correlation (FSC) gold-standard criterion at 0.143. Maps were refined with and without C5 symmetry and showed a 98% correlation in density. Therefore, we chose to use the map based on C5 symmetry for model building.

The structure solved here and shown in Figure 7 A (side view) and B (cytoplasm view), has the same α-helix bundle topology and is very similar to that of *E. coli* ExbB 5SV0 structure (that was co-purified with ExbD but did not show any ordered density for ExbD (14)), with a 1.3Å root mean square deviation (rmsd) over all Cα atoms (Table II). This shows that $ExbB_{Sm}$ is stable as a pentamer on its own. The periplasmic N-terminal extension did not yield any visible density, precluding its structure determination. At the C-terminus of each monomer, however, density was clearly defined, allowing structure determination for an additional 10 residues in a helical conformation up to the last one (helix α8 finishing with Gly 283, see Figure 7C) before the His-tag that is present but disordered. In the 5SV0 X-ray structure, a calcium ion, present in the crystallisation solution, is bound to the five Glu 106 (of helix α2 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 (Supplementary Figure S6) 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 ExbB sequences lacking N-terminal signal sequence and periplasmic extension (14).

The cryo-EM density of $ExbB_{Sm}$ shows vents located at the interface of adjacent monomers and at the height of the cytoplasmic junction with the inner membrane leaflet (Figure 7A). These vents may allow solvent or ion passage. Additional density was also clearly observed on the external surface, and we attributed it 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 7A, D and E). 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 (16). The cryoEM density map, when displayed at a level that shows the detergent belt, shows density inside the ExbB membrane pore (supplementary figure S7A). This density is too noisy to allow model building (supplementary figure S7B). However, it corresponds to a region with positive electrostatic charge on the top and bottom of the pore and neutral or hydrophobic in the middle (supplementary figure S7C). This density could be due to the presence of lipid or detergent. Interestingly, it is located at a different height as compared to the detergent belt and external lipids. In summary, our cryo-EM structure shows that $ExbB_{Sm}$ is stable as a pentamer and associates strongly with specific phospholipids coming from the inner leaflet of the membrane.

4b $ExbBD_{Sm}$ structure

We also solved the structure of the $ExbBD_{Sm}$ complex by cryo-EM (Figure 8 and Table II). As for ExbB, the purified complex exhibited a symmetric peak on SEC (see Figure 4, B). Cryo-

329 EM data were collected and processed as described in the Material and Methods section. In
330 contrast to $ExbB_{Sm}$, 2D classes showed preferential orientations with 92% top views and only
331 8% side views. Due to this strong bias in particle orientation, the resolution achieved was 3.96
332 Å, precluding reliable positioning of the side chains. However, our $ExbB_{Sm}$ model fitted readily
333 into the density. As observed for $ExbB_{Sm}$ alone, in the $ExbBD_{Sm}$ complex, $ExbB_{Sm}$ behaves as
334 a pentamer, and no sign of other assemblies was found. Similarly, no density could be attributed
335 to the periplasmic extension. Two clear densities inside the $ExbB_{Sm}$ pore were assigned to the
336 pair of $ExbD_{Sm}$ TM segments (Figure 8 B and C). The charge distribution is highly polarised
337 on the cytoplasmic side: the region close to the membrane is positively charged and the distal
338 part is negatively charged as observed for $ExbB_{Ec}$ (Supplementary Figure S8A). The central
339 pore of $ExbB_{Sm}$ is apolar, lined by TM helices 2 and 3 of each monomer, creating a large
340 hydrophobic cavity inside the structure (Supplementary Figure S8B). The ExbD helices were
341 initially oriented similarly to $ExbD_{Ec}$ in structure 6TYI and remained stable during refinement.
342 Interestingly, the ExbD TM helices are at the same height as the density observed in the pore
343 of ExbB alone, and closer to the periplasm than the membrane bilayer. In the refined model,
344 Asp 25 from $ExbD_{Sm}$ monomer chain G faces Thr 218 from $ExbB_{Sm}$ chain C, while Asp 25
345 from the $ExbD_{Sm}$ chain F faces the interface between two $ExbB_{Sm}$ monomers A and E
346 (Supplementary Figure S8C and D). The estimation of pKa, by using the program Propka (29)
347 server, shows that Asp 25 has a pKa of 7.3 for chain F and 7.4 for chain G, both very peculiar
348 for solvent-accessible acidic residues but that can be found in buried active sites or membrane
349 proteins (30). This pKa should allow protonation and deprotonation of ExbD Asp 25 at
350 physiological pH.
351 As compared to the structure of $ExbB_{Sm}$, there is a slight «opening» towards the periplasmic
352 side (the distance of Ala 197 from one subunit to Leu 204 of the facing subunit varying from

353 25.5Å in $ExbB_{Sm}$ to 29.8Å in $ExbBD_{Sm}$, much greater than any error expected due to the
354 difference in resolution of the two structures). This opening of the structure is limited to the
355 periplasmic part (Supplementary Figure S9). The rmsd between $ExbB_{Sm}$ alone and in complex
356 with $ExbD_{Sm}$ is 2Å over all Cα of the pentamer, and the rmsd between $ExbB_{Sm}$ in complex with
357 $ExbD_{Sm}$ and $ExbB_{Ec}$ in complex with $ExbD_{Ec}$ (structure 6TYI chains A-E) is 1.7Å.
358 As compared to the 6TYI structure of *E. coli* ExbBD complex, we observe significant
359 differences that may be related to the specificity of function of $ExbB_{Sm}$. Both $ExbB_{Sm}$ and
360 $ExbBD_{Sm}$ inner pores are slightly wider than their *E. coli* counterparts at the periplasmic
361 entrance. Consequently, two channels cross the membrane region, (extending from the
362 periplasmic entrance to the Asp 25 residue from the ExbD TM deeply embedded in the ExbB
363 pore) that are clearly seen in the $ExbBD_{Sm}$ (as detected by the MoleOnline (31) Server, Figure
364 9, A and C; the average diameter is around 3Å) while the *E. coli* structure exhibits a single,
365 much thinner channel as detected with the same parameters of MoleOnline (Figure 9, B and D;
366 average diameter around 2Å). It is therefore possible that this structure represents another state
367 of the ExbBD motor complex, whereby the periplasmic side allows solvent access to the Asp
368 25 residue of ExbD TM. The different physico-chemical conditions (nanodiscs vs. detergent
369 micelles, 200 vs. 100mM NaCl, pH7,4 vs. pH8) might also influence such parameters.
370 In summary, we show that ExbBD from *S. marcescens* has a 5:2 stoichiometry with a larger
371 channel allowing solvent or proton transport from the periplasmic side as compared with the *E.*
372 *coli* structure.
373
374 **5. Swapping residues in TM1 between $ExbB_{Sm}$ and $ExbB_{Ec}$ strongly suggests that ExbB**
375 **TM1 interacts with HasB/TonB.**

As already mentioned, TBDT function requires a productive association between TonB or its orthologs/paralogs with ExbBD complexes. It is known that the periplasmic domains of ExbD and TonB interact, and there are both genetic and structural pieces of evidence for the interaction between the TM helix of TonB and the first TM helix of ExbB. Our experimental system gave us the opportunity to investigate the specificity determinants between HasB/TonB and ExbB. Previous work (12) lends support to the absence of interaction between HasB and $ExbBD_{Ec}$. However, our data rather favor a less efficient interaction between HasB and $ExbBD_{Ec}$ than with $ExbBD_{Sm}$. *E. coli* can grow with the HasB-$ExbBD_{Ec}$ pair, although more slowly and with a lower yield: the final OD was 0.78 for HasB-$ExbBD_{Sm}$ and 0.43 for HasB-$ExbBD_{Ec}$ (Figure 3 D orange and grey curves). As it is known that ExbBD is required to stabilize TonB, we first tested whether HasB was stabilized by ExbBD. We could show that, although HasB does not accumulate in the absence of ExbBD, it is readily present in comparable amounts with either $ExbBD_{Sm}$ or $ExbBD_{Ec}$ (Figure 10). This observation therefore indicates that most likely HasB also interacts with $ExbBD_{Ec}$ so as to be stabilized (i. e., withstanding proteolytic degradation), but in a less functional manner than with $ExbBD_{Sm}$.

A general framework put forward for the Mot complexes indicates that MotA (equivalent of ExbB) rotation around MotB (equivalent of ExbD) drives the rotation of the flagellum basal ring (17, 18). We tested the possibility that ExbB residues exposed at the surface of the protein might be involved in a functional interaction with HasB/TonB, as has been proposed in the case of the TonB-ExbBD complex from *P. savastanoi* (17). The superposition of residue conservation on the structure of ExbB showed that the membrane-facing residues of the TM region of helix α2 and the cytoplasmic residues were the least conserved (Figures 1 and S6). We noted one conspicuous stretch of residues located in the cytoplasmic leaflet of TM helix 1 (see Supplementary Figure S6):

residues 76-88 of $ExbB_{Sm}$ TILFAKGSELLRA corresponding to

residues 39-51 of $ExbB_{Ec}$ AIFFSKSVEFFNQ

(identical residues are underlined) 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 (Figure 11 E and F).

We speculated that this region of ExbB might establish interaction with HasB. We made a chimeric $ExbB_{Ec}$ protein, with its 39-51 region replaced by residues 76-88 from $ExbB_{Sm}$, named $ExbB_{Ec\text{-}Sm76\text{-}88}$, and asked how this chimeric $ExbB_{Ec}$ protein would allow growth in the presence of HasB (see Figure 3A for the constructions), as compared to $ExbB_{Ec}$. As shown in Figure 6A, this chimeric protein $ExbB_{Ec\text{-}Sm76\text{-}88}$ (green dots) was more active with HasB than $ExbB_{Ec}$ (grey), the onset of growth took 14hrs compared to 18hrs for $ExbBD_{Ec}$ and the maximum OD was 0.74 compared to 0.43 for $ExbBD_{Ec}$. We also made the inverse change where the 76-88 region of $ExbB_{Sm}$ was replaced by the 39-51 region from $ExbB_{Ec}$ ($ExbB_{Sm\text{-}Ec39\text{-}51}$, magenta). This mutant has a degraded behaviour as compared to $ExbB_{Sm}$, since the onset of growth occurs at 19hrs compared to 6hrs, and the final OD is 0.47, compared to 0.78. 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 $ExbB_{Ec\text{-}Sm76\text{-}84}$ (see Figure 3 A). $ExbB_{Ec\text{-}Sm76\text{-}84}$ was quite comparable to $ExbB_{Ec\text{-}Sm76\text{-}88}$ (see deep-blue and green curves on Figure 6A). This 9-residue stretch is therefore sufficient to alter $ExbB_{Ec}$ to be better adapted to HasB. This set of experiments shows that the intramembrane functional zones of ExbB are crucial for the growth of *E. coli* in our experimental set up and therefore likely govern the interaction between ExbB and HasB.

In summary, we showed using NMR and culture growth assays that $ExbB_{Sm}$ N-terminal extension (section 3b) and a few residues in the α2 helix TM region (section 5) are likely involved in specific interactions with HasB.

Discussion

1) $ExbBD_{Sm}$ 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 (15). 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. Our data agree with those published by Lloubes, Buchanan et al. (14, 16). *In vivo*, a stoichiometry of 7ExbB, 2ExbD and 1TonB has been determined in *E. coli* (32). Although one cannot completely rule out some bias 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 (33).

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 (16). It is unclear whether this phospholipid was a genuine tightly bound, co-purified lipid or whether it came from the lipids added during reconstitution in the nanodiscs. In our case, we did not add any lipid and our mass spectrometry identification of co-purified lipids found mostly PG and some PE in the

447 ExbB case (and CL for ExbBD). The cryo-EM structure shows that they are located in the inner
448 leaflet, as in the *E. coli* complex. PG and CL are negatively charged, whereas PE is zwitterionic.
449 This might be pertaining to the functioning of the complex, as the lipids are bound to a highly
450 positively charged interface, very close to the fenestrations seen in the structure and PE might
451 therefore be less tightly bound than PG or CL. Further analysis by mass spectrometry of the
452 aliphatic chains of the lipids present in our sample also revealed a specific composition in their
453 length and unsaturated nature.
454 One point also worth mentioning is the existence of "channels" inside the ExbBD tunnel,
455 potentially allowing the passage of proton/hydronium ion up to the Asp 25 residue of ExbD,
456 deeply embedded in the apolar medium of ExbB, and therefore with a pKa close to
457 physiological pH, allowing easy protonation/deprotonation steps of Asp 25 side chain. The
458 identified channels connect the cytoplasm to the periplasm *via* Asp 25, allowing us to propose
459 a possible trajectory for the proton transfer via those channels.
460 3) Motor model proposal
461 Regarding the coupling of the pmf dissipation with the mechanical work carried out by
462 MotAB/TolQR/ExbBD complexes, several models have been put forward that might be more
463 or less easy to reconcile with the structural data. A detailed mechanism was proposed for *C.*
464 *jejuni* MotAB complex, where cycles of protonation/deprotonation of the conserved Asp of the
465 MotB TM (equivalent to the conserved Asp 25 of the ExbD TM) are coupled to the rotation in
466 discrete steps of the MotA pentamer around MotB axis, provided that the periplasmic domain
467 of MotB is anchored to the peptidoglycan layer *via* its peptidoglycan binding site (18). Our
468 ExbBD structure indeed shows that ExbD monomers have two orientations relative to ExbB:
469 Asp 25 from one monomer faces Thr 208 of ExbB and Asp 25 from the second monomer faces
470 a hydrophobic region at the interface of two ExbB monomers. In the first state previously

protonated Asp25 could be deprotonated in contact with polar Thr208 while the second, deprotonated Asp 25, could pick a proton from the periplasm via the channel. Rotation of the the ExbB pentamer around the two ExbD chains would then lead to a new protonation/deprotonation cycles accompanied by a change in the environment upon rotation. How is this rotative energy conveyed to the TBDT?

The binding of an iron-loaded substrate on the extracellular binding site of a TBDT triggers a reaction cascade ultimately leading to the substrate entry into the periplasm. The inner membrane TonB(HasB) complex conveys the energy of the pmf to the TBDT most likely by a rearrangement of the plug inside the barrel to allow substrate access to the periplasm and its capture by periplasmic binding proteins. Specific interactions between the TonB(HasB) box of the receptor and the C-terminal domain of TonB(HasB) are essential in this process (28). A wealth of data has also accumulated documenting specific interactions between $TonB_{CTD}$ and $ExbD_{CTD}$, depending upon the energy state of the cell. It is also known that ExbD and TonB interact *via* their periplasmic parts (34) 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 (35). There have also been indications and suggestions that both $TonB_{CTD}$ and $ExbD_{CTD}$ could interact with the peptidoglycan sacculus, providing anchor points to allow force transmission. Molecular modelling works hypothesized that the C-terminal domain of TonB (36) and the C-terminal domain of ExbD (37) have specific binding sites for the peptidoglycan network. Atomic force microscopy 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 entry of the substrate in the periplasm (38). Our NMR results also show that ExbB interacts *via* its alanine and proline-rich periplasmic N-terminal region with the HasB periplasmic domain on the side opposite to that of the TonB box interaction, thus

rendering a tripartite interaction possible. Our *in vivo* growth results with the $ExbB_{Smdelextss}$ mutant devoid of the periplasmic extension point to a possible role of the N-terminal periplasmic extension in the activation of the transcription of the Has locus (figure 6).

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. Similarly, the MotAB rotating model can be extrapolated to the ExbBD complex. In this model, ExbB would rotate around ExbD, driving the rotation of TonB/HasB thanks to the specific interaction between TonB/HasB TM and ExbB TM α2. This kind of interaction is supported by our mutagenesis data, as we could strongly increase the efficiency of $ExbB_{Ec}$ with HasB by exchanging a short stretch of residues between $ExbB_{Sm}$ and $ExbB_{Ec}$. It is also in line with a previous TonB TM point mutant partially suppressed by an ExbB α2 point mutant (39). 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 (Figure 11, compare structures on figures 11C and 11D and helical wheels on figures 11E and 11F). Similar interactions have been proposed by Deme et al. on the TonBExbBD complex from *P. savastanoi* (17), where an extra cryo-EM density is seen outside the ExbB pentamer and was tentatively assigned to the TM domain of TonB. The residues from $ExbB_{Ec}$TM1 (S34 and A39) identified as co-evolving with TonB TM region (18) are only a little deeper in the bilayer than the residues we exchanged between $ExbB_{Ec}$TM1 (39-51) and $ExbB_{Sm}$TM1, that are closer to the cytoplasm. More specifically, whereas S34 is conserved between $ExbB_{Ec}$ and $ExbB_{Sm}$ (S71), A39 is replaced by T76. Moreover, the hydrophobic core of HasB TM is likely shorter than that of TonB as seen in the sequence alignment of HasB and TonB TM domains:

```
HasBSmTM 18-39 RRC---LVLVLALH-LLVAALLWPRR
TonBSmTM 10-34 RRISVPFVLSVGLHSALVAGLLYAS-
TonBEcTM  7-31 RRFPWPTLLSVCIHGAVVAGLLYTS-
consensus      RR      L   LH   VA LL
```

This difference in length might influence the orientation of HasBTM inside the bilayer and potentially explain the gain of function of HasB6 mutant (in which the hydrophobic core has two more residues (39)) to be better suited to $ExbB_{Ec}$ than HasB. The presence of a proline residue in the TonB TM may also change its shape and the interaction with ExbB.

Several models have been proposed to account for the functioning of the TonB complex, in conjunction with the entry of a substrate bound to the extracellular side of a given receptor. In a first model, piston-like movements of TonB drive the unfolding of the plug inside the receptor barrel. *In vivo* proteolysis studies indicated that the C-terminal domain of TonB can change conformations during the catalytic cycle of protonation and deprotonation of the conserved Asp residue in ExbD TM (40). 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 (41). We therefore propose a model in which the force generated through the rotation of HasB/TonB, driven by ExbB rotation around ExbD, is not directly transmitted to the TonB box of the receptor, but could be mediated by the C-terminal domain of ExbD, that might act as an anchor point 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, in particular concerning protein-peptidoglycan interactions, the force needed to trigger TBDT plug opening, and how to distinguish between a rotation and a piston movement. Finally, given the wide range of lag periods we observe in our growth curves using different combinations of ExbB TM1 mutants

and TonB/HasB, one may hypothesize that the membrane interaction between ExbB and TonB/HasB influences the rate of transcription activation of the Has locus.

It is likely that the *in vivo* entry of a siderophore is a rather slow process: during a cell division time a bacterial cell must 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 (32), meaning that each receptor has to undergoes 20 cycles during a generation estimated to 30 minutes, leading to a turnover time of the TonB complex of about 5-10 seconds. As compared to flagellar motor that can operate at extremely high speeds (several hundred of rotations per second), even though the basic mechanisms are likely to be conserved with the ExbBD/TonB-HasB complex, it is much slower, which likely points to different coupling mechanisms.

Material and Methods

Strains and plasmid construction

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

-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 *Eco*RI and *Sph*I, and ligated with pBAD24 digested with *Eco*RI and *Sph*I. Correct clones were selected after sequencing of the insert.

-Plasmid pBADExbBD$_{Sm}$His6 (encoding ExbB$_{Sm}$ and a C-terminally His-tagged version of ExbD$_{Sm}$) was constructed by first amplifying on pBADExbBDSm a ca. 0.4kb fragment with the following oligonucleotides SphIHisCtexbDSm and BglIIexbDSm; after amplification, the fragment was purified, digested with *Bgl*II and *Sph*I, and ligated with pBADExbBDSm digested

568 with the same enzymes. The correct clones were selected after sequencing of the insert. In
569 biological tests, this plasmid was undistinguishable from its parent plasmid pBADExbBDSm.
570 -Plasmid pBADExbBD$_{Ec}$ was constructed by amplification on genomic MG1655 DNA of a ca.
571 1.7kb fragment with the following oligonucleotides ExbBD5c, and ExbBD3c; the PCR product
572 was purified, digested with *Eco*RI and *Sph*I, and ligated with pBAD24 digested with the same
573 enzymes. Correct clones were selected after sequencing.
574 -Plasmid pBADExbB$_{Sm}$His6 (encoding a C-terminally His-tagged version of ExbB$_{Sm}$) was
575 constructed as follows: a PCR fragment was amplified on pBADExbBDSm with the oligos
576 PBADFOR and ExbBHis6, digested with *Eco*RI and *Sph*I, and ligated with pBAD24 digested
577 with the same enzymes. Correct clones were selected by sequencing.
578 Plasmid pBADExbBD$_{Ec\text{-}Sm76\text{-}88}$ (encoding a chimeric ExbB$_{Ec}$ protein with its 39-51 residues
579 exchanged for the 76-88 residues from ExbB$_{Sm}$ and ExbD$_{Ec}$) was constructed as follows:
580 Plasmid pBADExbBD$_{Ec}$ was amplified with the two following 5' phosphorylated
581 oligonucleotides: ExbBEcSm76-88.1 and ExbBEcSm76-88.2. After digestion with *Dpn*I and
582 self-ligation, correct clones were selected by sequencing and the mutated fragments recloned
583 in pBAD24.
584 Plasmid pBADExbBD$_{Ec\text{-}Sm76\text{-}84}$ (encoding a chimeric ExbB$_{Ec}$ protein with its 39-47 residues
585 exchanged for the 76-84 residues from ExbB$_{Sm}$ and ExbD$_{Ec}$) was constructed in the same
586 manner, with the following oligonucleotides pairs: ExbBEcSm 76-84.1 and ExbBEcSm 76-
587 84.2.
588 In the same manner, plasmid pBADExbBD$_{Sm\text{-}Ec39\text{-}51}$ (encoding a chimeric ExbB$_{Sm}$ protein with
589 its 76-88 residues exchanged for the residues 39-51 from ExbB$_{Ec}$ and ExbD$_{Sm}$) was constructed
590 by PCR on the pBADExbBD$_{Sm}$ plasmid with the following couple of 5' phosphorylated
591 oligonucleotides: ExbBSmEc39-51.1 and ExbBSmEc39-51.2.

592 Plasmid pBADExbBD$_{Smdelextss}$ (encoding ExbB$_{Sm}$ deleted of its periplasmic extension and
593 ExbD$_{Sm}$) was constructed similarly, by using as a template pBADExbBDSm, with the two
594 phosphorylated oligonucleotides ExbBdelextss1 and ExbBdelextss2.
595 To construct pBADExbB$_{Smdelextss}$His6, a ca. 0.6kb *Eco*RI-*Kpn*I fragment from
596 pBADExbBD$_{Smdelextss}$ was exchanged for the corresponding fragment of pBADExbBHis6.
597 Plasmid pAMHasISRADEB (encoding HasI, HasS, HasR, HasA, HasD, HasE and HasB), was
598 constructed by digesting pAMHasISRADE (encoding HasI, HasS, HasR, HasA, HasD and
599 HasE) and pSYC7 (encoding HasD, HasE and HasB) (42) by *Kpn*I and *Hin*dIII, and purifying
600 fragments of respectively ca. 9 and 7 kbases, ligating them together to obtain a plasmid with
601 the whole *has* locus on a low copy number plasmid (pAM238) under the control of its
602 endogenous regulation signals (a fur box upstream of *hasI* and *hasR*, the *hasI* and *hasS* genes
603 respectively encoding the has specific sigma and anti-sigma factors, and the hasS box upstream
604 of *hasS* and *hasR*).
605 All constructions were carried out in *E. coli* strain XL1-Blue.
606 Strains and other plasmids used in this study are shown in Table III and are from the laboratory
607 collection.
608

609 **Protein expression and purification**

610 The BL21DE3(pExbBD$_{Sm}$His6/pBAD24) or BL21DE3(pExbB$_{Sm}$His6/pBAD24) were grown
611 at 37°C either in TB or MDM medium, and induced with 0.2% arabinose at 1.5-2OD600nm
612 (TB) or 5-6OD600nm (MDM) and the incubation continued for 3hrs. The cells were harvested
613 by centrifugation (10000g 20minutes 4°C), washed once in 20mMTris-HCl pH8.0, flash-frozen
614 in liquid N_2 and kept at -80°C. Cells were broken in a Cell disruptor (Constant, UK) at 1kbar
615 (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 N_2 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-Fisher 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$M^{-1}cm^{-1}$), or ExbB5 (104850$M^{-1}cm^{-1}$). 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 Figure 4. 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.

640

641 **Activity tests**

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

643 1. Growth tests in liquid medium: *E. coli* C600 and its Δ*exbBD* derivatives were transformed
644 with specified plasmids and their growth tested in liquid LB medium at 37°C, at various
645 dipyridyl concentrations to induce iron starvation. OD_{600nm} was measured after overnight
646 growth.

647 2. Growth tests on plates: the relevant plasmids were also transformed into *E. coli* C600 Δ*hemA*,
648 a heme auxotroph strain (and derivatives thereof), and growth of the strains was assayed as
649 follows. Briefly, cells were first grown in LB medium (supplemented with delta-aminolevulinic
650 acid (50µg/ml) to bypass the effect of the Δ*hemA* mutation) at 37°C up to an OD_{600nm} of 1, and
651 then mixed with melted top agar (0.6%agar in LB), and poured onto LB agar plates containing
652 the appropriate antibiotics with arabinose at a concentration of 40µg/ml to induce expression
653 of the genes under the P_{araBAD} promoter(43). Wells, punched with Pasteur pipettes were filled
654 with heme-albumin (at 5, 1 or 0.2µM), at two dipyridyl concentrations (100 or 200µM) to
655 induce iron starvation. Plates were incubated overnight at 37°C and scored for growth around
656 the wells. All experiments were performed in triplicate.

657 3.Growth curves in liquid medium: a few colonies of *E. coli* C600
658 Δ*hemA*Δ*tonB*Δ*exbBD*(pAMHasISRADEB + pBAD24 or derivatives thereof) were first
659 inoculated in 4ml of LB medium at 37°C with the corresponding antibiotics, and 100µM
660 dipyridyl, 4µg/ml arabinose but without delta-aminolevulinic acid. Once the culture reached an
661 OD_{600nm} of ca. 1.2-1.5, it was diluted and inoculated in 48 well Greiner plates, in the same
662 medium to which was added 0.4µM He-BSA, as a heme source. The initial OD_{600nm} of the
663 cultures was 0.001. Each well contained 300µl 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. OD_{600nm} was recorded every 30 minutes for 60-70 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 of 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 ExbB, the ExbBD complex or 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 software 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 MotionCorr2 (44) and aligned images were processed with Gctf (45) using Scipion interface(46).

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 -2.5µm. Movies of 56 frames were acquired using a dose rate of 8 electrons/$Å^2$/second over 7 seconds, yielding a cumulative exposure of 55.95 electrons/$Å^2$.

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

Cryo-EM image processing and analysis

a) ExbB

For ExbB, aligned movies were processed with Gctf (45) 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 (47) (Supplementary Figure S10). Particles were extracted using a 2-fold binning, issued from a manual picking and a 2-D classification of particles picked out from 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.

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 (48) and graphically adjusted with Coot (49). Lipid phosphatidyl glycerol starting structure and geometry were built using Phenix eLBOW (50) and was fit in the map using Coot then refined along with the protein model in Phenix.real_space_refine.

b) ExbBD

The data processing is summarised in Supplementary Figure S11. Movies were drift-corrected and dose-weighted using MotionCorr2 (44). Aligned dose weighted averages were imported

into Cryosparc2 (51) and contrast transfer function was estimated using CTFFIND4.1 (52). 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 (53) 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. 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 II for refinement statistics).

Unless otherwise specified, the structure figures were made using UCSF Chimera software (53).

NMR experiments

The C-terminal periplasmic domain of HasB, $HasB_{CTD}$ comprising residues 133-263, was produced and purified as previously reported (28, 54). The peptide corresponding to the periplasmic extension of $ExbB_{sm}$ (A_1PAANPAVTESVAPTTAPAPAAAAPESITPVNPAPTIQPPETRG$_{44}$- numbering with reference to the mature protein) was synthetized 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 (55) and analyzed with CcpNmr Analysis 2.4 software (56). Proton chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5 sulfonate as 0 ppm. ^{15}N were referenced indirectly to DSS (57) (Wishart et al., 1995). ^{1}H–^{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 $CSP = [\Delta \delta H^2 + (\Delta \delta N \times 0.15)^2]^{1/2}$, where $\Delta \delta H$ and $\Delta \delta N$ are the observed ^{1}H and ^{15}N chemical shift differences between the two experimental conditions. The ^{1}H and ^{15}N resonance assignments were from Lefevre et al 2007 (58).

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.

Author contribution

779 PD and VB conceived the study and wrote the manuscript. all authors contributed to the
780 manuscript and approved it. PD produced protein samples and performed microbiology
781 experiments. MC performed electron microscopy sample screening and data collection. RJDA,
782 PDC, VYNE and VB processed cryo-EM data. VB built structural models. VB and PD
783 interpreted models. BL installed and tested programs. CM and JCR performed mass spec
784 analyses. NIP and GCA performed NMR analyses. HS provided advice with data collection
785 and processing.
786
787 **Data availability**
788 Protein structures were deposited to the protein data bank with ID 6YE4 and 7AJQ. Electron
789 microscopy maps were deposited to the EMDB with ID 10789 and 11806.
790
791 **Acknowledgments** This work was supported by grants from the ANR
792 (HEMESTOCKEXCHANGE ANR-12-BSV3-0022-01 and LABEX DYNAMO ANR-11-
793 LABEX-0011-01). We thank Nathalie Dautin for helpful suggestions and critical reading of the
794 manuscript, as well as Andrew Thompson. We acknowledge Emmanuel Frachon and
795 Christophe Thomas from the Institut Pasteur platform "Production of recombinant proteins" for
796 large scale cultures used in this work, European Synchrotron Radiation Facility for the
797 provision of time on the Titan Krios at beamline CM01, the Institut Pasteur Biological NMR
798 Technological Platform for the use of the NMR spectrometers, the staff at the ESRF CM01
799 facility in Grenoble for help in the data collection, the staff at the Marseille proteomics facility
800 and Djalila Djouadi and Diana Basto who initially contributed to this project. We thank M.
801 Nilges for interest in this work and B. Miroux for constant support.
802

803
804

References

1. Noinaj N, Guillier M, Barnard TJ, Buchanan SK. 2010. TonB-dependent transporters: regulation, structure, and function. Annu Rev Microbiol 64:43-60.
2. Grinter R, Lithgow T. 2019. The structure of the bacterial iron-catecholate transporter Fiu suggests that it imports substrates via a two-step mechanism. J Biol Chem 294:19523-19534.
3. Cascales E, Lloubes R, Sturgis JN. 2001. The TolQ-TolR proteins energize TolA and share homologies with the flagellar motor proteins MotA-MotB. Mol Microbiol 42:795-807.
4. Braun V. 1989. The structurally related exbB and tolQ genes are interchangeable in conferring tonB-dependent colicin, bacteriophage, and albomycin sensitivity. J Bacteriol 171:6387-90.
5. Eick-Helmerich K, Braun V. 1989. 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 171:5117-26.
6. Braun V, Herrmann C. 1993. 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 8:261-8.
7. Fischer E, Gunter K, Braun V. 1989. 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 171:5127-34.
8. Ahmer BM, Thomas MG, Larsen RA, Postle K. 1995. Characterization of the exbBD operon of Escherichia coli and the role of ExbB and ExbD in TonB function and stability. J Bacteriol 177:4742-7.
9. Higgs PI, Myers PS, Postle K. 1998. Interactions in the TonB-dependent energy transduction complex: ExbB and ExbD form homomultimers. J Bacteriol 180:6031-8.
10. Braun V, Gaisser S, Herrmann C, Kampfenkel K, Killmann H, Traub I. 1996. 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 178:2836-45.
11. Gaisser S, Braun V. 1991. The tonB gene of Serratia marcescens: sequence, activity and partial complementation of Escherichia coli tonB mutants. Mol Microbiol 5:2777-87.
12. Paquelin A, Ghigo JM, Bertin S, Wandersman C. 2001. Characterization of HasB, a Serratia marcescens TonB-like protein specifically involved in the haemophore-dependent haem acquisition system. Mol Microbiol 42:995-1005.
13. Benevides-Matos N, Wandersman C, Biville F. 2008. HasB, the Serratia marcescens TonB paralog, is specific to HasR. J Bacteriol 190:21-7.
14. Celia H, Noinaj N, Zakharov SD, Bordignon E, Botos I, Santamaria M, Barnard TJ, Cramer WA, Lloubes R, Buchanan SK. 2016. Structural insight into the role of the Ton complex in energy transduction. Nature 538:60-65.
15. Maki-Yonekura S, Matsuoka R, Yamashita Y, Shimizu H, Tanaka M, Iwabuki F, Yonekura K. 2018. Hexameric and pentameric complexes of the ExbBD energizer in the Ton system. Elife 7.

16. Celia H, Botos I, Ni X, Fox T, De Val N, Lloubes R, Jiang J, Buchanan SK. 2019. Cryo-EM structure of the bacterial Ton motor subcomplex ExbB–ExbD provides information on structure and stoichiometry. Communications Biology 2:358.
17. Deme JC, Johnson S, Vickery O, Muellbauer A, Monkhouse H, Griffiths T, James RH, Berks BC, Coulton JW, Stansfeld PJ, Lea SM. 2020. Structures of the stator complex that drives rotation of the bacterial flagellum. Nat Microbiol 5:1553-1564.
18. Santiveri M, Roa-Eguiara A, Kuhne C, Wadhwa N, Hu H, Berg HC, Erhardt M, Taylor NMI. 2020. Structure and Function of Stator Units of the Bacterial Flagellar Motor. Cell 183:244-257 e16.
19. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-10.
20. Ihsan MZ, Ahmad SJ, Shah ZH, Rehman HM, Aslam Z, Ahuja I, Bones AM, Ahmad JN. 2017. Gene Mining for Proline Based Signaling Proteins in Cell Wall of Arabidopsis thaliana. Front Plant Sci 8:233.
21. Mishra AK, Choi J, Moon E, Baek KH. 2018. Tryptophan-Rich and Proline-Rich Antimicrobial Peptides. Molecules 23.
22. Ball LJ, Kühne R, Schneider-Mergener J, Oschkinat H. 2005. Recognition of proline-rich motifs by protein-protein-interaction domains. Angew Chem Int Ed Engl 44:2852-69.
23. Marchler-Bauer A, Anderson JB, DeWeese-Scott C, Fedorova ND, Geer LY, He S, Hurwitz DI, Jackson JD, Jacobs AR, Lanczycki CJ, Liebert CA, Liu C, Madej T, Marchler GH, Mazumder R, Nikolskaya AN, Panchenko AR, Rao BS, Shoemaker BA, Simonyan V, Song JS, Thiessen PA, Vasudevan S, Wang Y, Yamashita RA, Yin JJ, Bryant SH. 2003. CDD: a curated Entrez database of conserved domain alignments. Nucleic Acids Res 31:383-7.
24. Biville F, Cwerman H, Letoffe S, Rossi MS, Drouet V, Ghigo JM, Wandersman C. 2004. Haemophore-mediated signalling in Serratia marcescens: a new mode of regulation for an extra cytoplasmic function (ECF) sigma factor involved in haem acquisition. Mol Microbiol 53:1267-77.
25. Letoffe S, Delepelaire P, Wandersman C. 2004. Free and hemophore-bound heme acquisitions through the outer membrane receptor HasR have different requirements for the TonB-ExbB-ExbD complex. J Bacteriol 186:4067-74.
26. Aluyi HS, Boote V, Drucker DB, Wilson JM, Ling YH. 1992. Analysis of polar lipids from some representative enterobacteria, Plesiomonas and Acinetobacter by fast atom bombardment-mass spectrometry. J Appl Bacteriol 73:426-32.
27. Erdős G, Dosztányi Z. 2020. Analyzing Protein Disorder with IUPred2A. Curr Protoc Bioinformatics 70:e99.
28. de Amorim GC, Prochnicka-Chalufour A, Delepelaire P, Lefevre J, Simenel C, Wandersman C, Delepierre M, Izadi-Pruneyre N. 2013. The structure of HasB reveals a new class of TonB protein fold. PLoS One 8:e58964.
29. Li H, Robertson AD, Jensen JH. 2005. Very fast empirical prediction and rationalization of protein pKa values. Proteins 61:704-21.
30. Langsetmo K, Fuchs JA, Woodward C. 1991. The conserved, buried aspartic acid in oxidized Escherichia coli thioredoxin has a pKa of 7.5. Its titration produces a related shift in global stability. Biochemistry 30:7603-9.

31. Berka K, Hanák O, Sehnal D, Banás P, Navrátilová V, Jaiswal D, Ionescu CM, Svobodová Vareková R, Koca J, Otyepka M. 2012. MOLEonline 2.0: interactive web-based analysis of biomacromolecular channels. Nucleic Acids Res 40:W222-7.
32. Higgs PI, Larsen RA, Postle K. 2002. Quantification of known components of the Escherichia coli TonB energy transduction system: TonB, ExbB, ExbD and FepA. Mol Microbiol 44:271-81.
33. Ollis AA, Manning M, Held KG, Postle K. 2009. Cytoplasmic membrane protonmotive force energizes periplasmic interactions between ExbD and TonB. Mol Microbiol 73:466-81.
34. Garcia-Herrero A, Peacock RS, Howard SP, Vogel HJ. 2007. The solution structure of the periplasmic domain of the TonB system ExbD protein reveals an unexpected structural homology with siderophore-binding proteins. Mol Microbiol 66:872-89.
35. Ollis AA, Postle K. 2011. The same periplasmic ExbD residues mediate in vivo interactions between ExbD homodimers and ExbD-TonB heterodimers. J Bacteriol 193:6852-63.
36. Klebba PE. 2016. ROSET Model of TonB Action in Gram-Negative Bacterial Iron Acquisition. J Bacteriol 198:1013-21.
37. Boags AT, Samsudin F, Khalid S. 2019. Binding from Both Sides: TolR and Full-Length OmpA Bind and Maintain the Local Structure of the E. coli Cell Wall. Structure 27:713-724 e2.
38. Hickman SJ, Cooper REM, Bellucci L, Paci E, Brockwell DJ. 2017. Gating of TonB-dependent transporters by substrate-specific forced remodelling. Nat Commun 8:14804.
39. Larsen RA, Thomas MG, Wood GE, Postle K. 1994. Partial suppression of an Escherichia coli TonB transmembrane domain mutation (delta V17) by a missense mutation in ExbB. Mol Microbiol 13:627-40.
40. Ollis AA, Kumar A, Postle K. 2012. The ExbD periplasmic domain contains distinct functional regions for two stages in TonB energization. J Bacteriol 194:3069-77.
41. Kohler SD, Weber A, Howard SP, Welte W, Drescher M. 2010. 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 19:625-30.
42. Létoffé S, Ghigo JM, Wandersman C. 1993. Identification of two components of the Serratia marcescens metalloprotease transporter: protease SM secretion in Escherichia coli is TolC dependent. J Bacteriol 175:7321-8.
43. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121-30.
44. Zheng SQ, Palovcak E, Armache JP, Verba KA, Cheng Y, Agard DA. 2017. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. Nat Methods 14:331-332.
45. Zhang K. 2016. Gctf: Real-time CTF determination and correction. J Struct Biol 193:1-12.
46. Gomez-Blanco J, de la Rosa-Trevin JM, Marabini R, Del Cano L, Jimenez A, Martinez M, Melero R, Majtner T, Maluenda D, Mota J, Rancel Y, Ramirez-Aportela E, Vilas JL, Carroni M, Fleischmann S, Lindahl E, Ashton AW, Basham M, Clare DK, Savage K, Siebert CA, Sharov GG, Sorzano COS, Conesa P, Carazo JM. 2018.

Using Scipion for stream image processing at Cryo-EM facilities. J Struct Biol 204:457-463.

47. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife 7.

48. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. 2018. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol 74:531-544.

49. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.

50. Moriarty NW, Grosse-Kunstleve RW, Adams PD. 2009. electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation. Acta Crystallogr D Biol Crystallogr 65:1074-80.

51. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. 2017. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods 14:290-296.

52. Rohou A, Grigorieff N. 2015. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol 192:216-21.

53. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605-12.

54. Lefevre J, Delepelaire P, Delepierre M, Izadi-Pruneyre N. 2008. Modulation by substrates of the interaction between the HasR outer membrane receptor and its specific TonB-like protein, HasB. J Mol Biol 378:840-51.

55. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. 1995. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6:277-93.

56. Vranken WF, Boucher W, Stevens TJ, Fogh RH, Pajon A, Llinas M, Ulrich EL, Markley JL, Ionides J, Laue ED. 2005. The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins 59:687-96.

57. Wishart DS, Bigam CG, Yao J, Abildgaard F, Dyson HJ, Oldfield E, Markley JL, Sykes BD. 1995. 1H, 13C and 15N chemical shift referencing in biomolecular NMR. J Biomol NMR 6:135-40.

58. Lefèvre J, Simenel C, Delepelaire P, Delepierre M, Izadi-Pruneyre N. 2007. (1)H, (13)C and (15)N resonance assignments of the C-terminal domain of HasB, a specific TonB like protein, from Serratia marcescens. Biomol NMR Assign 1:197-9.

979 **Tables**

Afipia broomeae	Hartmannibacter diazotrophicus	Phreatobacter stygius
Agrobacterium rhizogenes	Hoeflea olei	Phyllobacterium zundukense
Aliihoeflea sp.	Hyphomicrobium nitrativorans	Proteobacteria bacterium
Aminobacter aminovorans	Insolitispirillum peregrinum	Pseudaminobacter manganicus
Amorphus coralli	Kaistia granuli	Pseudochrobactrum asaccharolyticum
Ancylobacter aquaticus	Ketogulonicigenium robustum	Pseudolabrys taiwanensis
Aquamicrobium aerolatum	Labrys okinawensis	Pseudorhizobium pelagicum
Aurantimonas manganoxydans	Magnetospirillum gryphiswaldense	Pseudorhodoplanes sinuspersici
Aureimonas flava	Mangrovicella endophytica	Reyranella massiliensis
Azorhizobium caulinodans	Mesorhizobium loti	Rhizobium tropici
Azospirillum brasilense	Methylobrevis pamukkalensis	Rhodoligotrophos sp.
bacterium A52C2	Methylocapsa acidiphila	Rhodomicrobium udaipurense
Bartonella apis	Methyloceanibacter stevinii	Rhodopseudomonas pseudopalustris
Beijerinckiaceae bacterium	Methylocella silvestris	Rhodospirillum centenum
Blastochloris sp.	Methyloligella halotolerans	Rhodovarius sp.
Bosea vaviloviae	Methylopila sp.	Roseococcus sp.
Bradyrhizobium yuanmingense	Methylovirgula ligni	Roseospirillum parvum
Brucella sp.	Microvirga aerophila	Sinirhodobacter sp.
Chelatococcus asaccharovorans	Neorhizobium alkalisoli	Sinorhizobium sp.
Ciceribacter lividus	Nitratireductor pacificus	Skermanella aerolata
Devosia psychrophila	Nitrobacter hamburgensis	Sphingobium chlorophenolicum
Dongia mobilis	Nitrospirillum amazonense	Sphingomonas sp.
Enterovirga rhinocerotis	Niveispirillum cyanobacteriorum	Sphingopyxis sp.
Erythrobacter	Ochrobactrum thiophenivorans	Sphingorhabdus sp.
Falsochrobactrum ovis	Paenirhodobacter enshiensis	Starkeya novella
Geminicoccus roseus	Paracoccus aminovorans	Tardiphaga robiniae
Granulibacter bethesdensis	Paramesorhizobium deserti	Telmatospirillum siberiense
Haematobacter massiliensis	Parvibaculum lavamentivorans	Variibacter gotjawalensis
Hansschlegelia zhihuaiae	Phaeospirillum fulvum	Xanthobacter autotrophicus

980
981

982 Table I: a list of selected species of Alphaproteobacteria where “extralong” ExbB’s are found.

983 All those species also contain HasB orthologs, except for the ones in red.

984

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–/Å2)	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)	3.1	3.96
FSC threshold	0.143	0.143
Map resolution range (Å)	3.1-5.5	3.96-8
Symmetry imposed	C5	C1
Initial model used	Phyre2 homology model based on 5SV0	Our ExbB structure
Model composition		
Non-hydrogen atoms	9145	9256
Protein residues	1180	1229
Ligands	PGT : 5	0
B factors (Å2)		
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 II: Data collection, processing, and refinement statistics for $ExbB_{Sm}$ and $ExbBD_{Sm}$.

Strains :
E. coli K12 C600
E. coli K12 C600Δ*hemA::Km*^{r}Δ*exbBD*
E. coli K12 C600Δ*hemA::Km*^{r}Δ*exbBD*Δ*tonB*
E. coli K12 XL1-Blue
E. coli K12 JP313
E. coli K12 JP313Δ*exbBD*Δ*tonB*
E. coli K12 C600Δ*exbBD*
E. coli BL21DE3
S. marcescens Db11

Plasmids :
pBAD24: lab collection
pBAD33: lab collection
pAM238: lab collection
pAMHasISRADE: lab collection
pAMHasISRADEB: this work
pBADExbBD$_{Ec}$: this work
pBADExbBD$_{Sm}$: this work
pBADExbBD$_{Ec\text{-}Sm76\text{-}88}$: this work
pBADExbBD$_{Ec\text{-}Sm76\text{-}84}$: this work
pBADExbBD$_{Sm\text{-}Ec39\text{-}51}$: this work
pBADExbBD$_{delextss}$: this work
pBADHasB: lab collection
pBADExbBD$_{Sm}$His6: this work
pBADExbB$_{Sm}$His6: this work
pBADExbB$_{Smdelextss}$His6: this work

Oligonucleotides:

ExbBDSm5’	5’-GGAGGAATTCACCATGAAAACGGCTGGCAAGAAT -3’
ExbBDSm3’	5’-AAGCTTGCATGCCTATTTGGCGGCGCCTTCCA-3’
SphIHisCtexbDSm	5’-TTGCATGCCTAATGGTGATGGTGATGGTGTTTGGCGCGCCTTC-3’
BglIIexbDSm	5’-GCGCTTAAATGAAGATCTGGACGACAGCGG-3’
ExbBD5c	5’-CAGGAGGAATTCACCATGGGTAATAATTTAATGCAGACGGA-3’
ExbBD3c	5’-AAGCTTGCATGCTTACTTCGCTTTGGCGGTTTCTT-3’
PBADFOR	5’-CTGACGCTTTTTATCGCAAC-3’
ExbBHis6	5’-AAGCTTGCATGCCTAATGGTGATGGTGATGGTGCCCCGCCCGCAGTTG-3’
ExbBEcSm76-88.1	5’P-GGCAGTGAACTGCTGCGCGCCAAGCGTCGCCTTAAGCGCGAG-3’
ExbBEcSm76-88.2	5’P-TTTAGCGAACAAAATGGTCCAGGTGACTACGGAGGCCAAAAT-3’
ExbBEcSm76-84.1	5’P-CTAAAGGCAGTGAATTCTTCAATCAGAAGCGTCG-3’
ExbBEcSm76-84.2	5’P-CGAACAAAATGGTCCAGGTGACTACGGAGGCCA-3’
ExbBSmEc39-51.1	5’P-GCGTAGAGTTCTTCAATCAGAAGCGCCGTCTGCGTCGCGA-3’
ExbBSmEc39-51.2	5’P-TCTTACTGAAGAAGATTGCCCAGGTCACGATAGACGCCAG-3’
ExbBdelextss1	5’P- GAAACCCGCGGCATGGACCTGTCCATTTGGGG-3’
ExbBdelextss2	5’P- TGCCTGCGCGCTGCCGGCCAGCCCCACAA-3’

Table III: list of strains, plasmids and oligonucleotide sequences used in this work

Figure Legends

Figure 1: ExbB sequence analysis. A: alignment of $ExbB_{Ec}$ (top) with $ExbB_{Sm}$ (bottom), made with DNA Strider v1.4x-2b, CEA. In red is shown the signal sequence from $ExbB_{Sm}$, and in blue the periplasmic extension, not present in $ExbB_{Ec}$. The consensus sequence is shown on the middle line, where + indicates similar amino-acid residues. B: Weblogo (24) representation of 131 « long » ExbB aminoacid sequences aligned with Clustal omega (25); 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: $ExbBD_{Sm}$ complements $ExbBD_{Ec}$. overnight growth in LB broth of *E. coli* C600Δ*exbBD*(pBAD24) (blue), C600Δ*exbBD*(pBAD24$ExbBD_{Ec}$) (orange), and C600Δ*exbBD*(pBAD24$ExbBD_{Sm}$) (grey), without (0) or with 100 and 200µM iron chelator di-pyridyl (DiP). One representative experiment is shown. The vertical axis represents mDO_{600nm} absorbance units.

Figure 3: Role of HasB and TonB on bacterial growth. A: representation of the constructions used in Figure 3 B, 3C, 3D and elsewhere in this work. ss refers to the signal sequence, ext to the periplasmic extension of $ExbB_{Sm}$, and tm1, tm2 and tm3 to the first, second and third ExbB transmembrane segments, respectively. B and C: growth around wells of *E. coli* C600Δ*hemA*Δ*tonB*Δ*exbBD* harbouring pAMHasISRADEB and specified recombinant plasmids (B and C left parts) or *E. coli* C600Δ*hemA*Δ*exbBD* harbouring pAMHasISRADE and specified recombinant plasmids (B and C right parts); the pictures were taken after overnight growth at 37°C (B), or after 36hrs growth (C). Arabinose concentration was 40µg/ml. Dipyridyl (DiP) iron chelator concentration was either 100 or 200µM. Heme-BSA concentrations inside the wells were 10, 2 and 0,4µM. D: growth curves in microplates of the same strains as in B/C; DiP concentration was 100µM, arabinose concentration 4µg/ml, Heme-BSA 1µM. Colored dots in B and C refer to the equivalent growth conditions in D.

Figure 4: Homogeneity assessment of ExbB and ExbBD. Size exclusion chromatography profiles of representative purification of respectively $ExbB_{Sm}$His6 (A) and $ExbBD_{Sm}$His6 (B) on a Superose 6 10/300 column. The 280nm absorbance is plotted as a function of elution volume. A Coomassie-stained gel of the pooled peak fractions is shown on the right of each profile together with a molecular weight ladder on the right (respectively 15, 25, 35, 40, 55, 70, 100, 130 and 170kDa). The faint band present in the Exb B_{Sm}His6 sample above 100kDa was identified as AcrB using mass spectrometry.

Figure 5:
Interaction of $HasB_{CTD}$ with the periplasmic fragment of $ExbB_{Sm}$, as detected by NMR. A: Superposed $^{1}H–^{15}N$ HSQC spectra of 0.15 mM ^{15}N-labelled $HasB_{CTD}$ in 50mM sodium phosphate, pH 7, 50 mM NaCl in the presence (red) or absence (black) of the periplasmic peptide of $ExbB_{Sm}$ residues 1-44. B and C: $HasB_{CTD}$ residues (PDB code 2M2K) exhibiting the highest CSP in the presence of the periplasmic peptide of $ExbB_{sm}$ are coloured red (B: surface representation; C: cartoon representation). The residues showing the highest CSP are indicated.

Figure 6: ExbB TM1 interacts with HasB. A: bacterial growth curves of *E. coli* C600Δ*hemA*Δ*exbBD*Δ*tonB* harbouring pAMHasISRADEB plasmid together with either pBAD24 (medium-blue), pBADExbBD$_{Sm}$ (orange) or pBADExbBD$_{Smdelextss}$ (yellow), pBADExbBD$_{Ec}$ (grey), pBADExbBD$_{EcSm76-88}$ (green), pBADExbBD$_{EcSm76-84}$ (dark-blue), pBADExbBD$_{SmEc39-51}$ (magenta) in the presence of 1µM He-BSA as heme source, 4µg/ml arabinose to induce ExbBD expression and 100µM dipyridyl to induce iron starvation (see Materials and Methods for further details). The optical path was ca. 3mm long, and each curve represents the mean of four replicates of the same culture, recorded every 30 minutes for 66 hours. Bottom: Coomassie-blue stained gel (B) and immunodetection (C) with anti-His6 antibody of membrane preparations of *E. coli* C600 harbouring either pBAD24 plasmid (1), pBADExbB$_{Sm}$ (2) or pBADExbB$_{Smdelextss}$ (3). The equivalent of 0.6 OD$_{600nm}$ was loaded in B, and of 0.3 OD$_{600nm}$ in C.

Figure 7: Representation of the pentameric structure of ExbB$_{Sm}$ solved by cryo-EM. A: the cryo-EM density map is shown, each monomer with distinct colours, view from the side. The yellow/grey regions represent non-protein density. The yellow regions were modeled as PG molecules. B: view from the cytoplasmic side. C: 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 α2 helix), TM2 (part of α6 helix) and TM3 (part of α7 helix) are indicated. D: close-up representation of one PG density; E: fitting of the PG molecule inside the density; nearby ExbB residues are depicted as sticks using the Coot software (49).

Figure 8: Structure of ExbBD$_{Sm}$ solved by cryo-EM. The same color code as in Figure 7 for is used the ExbB monomers and the two ExbD monomers are colored yellow and gold. A: side view; B: view from the periplasmic space. C: clipped side view, with the helix of each ExbD monomer represented (yellow and gold). For B and C, the insets show the clipping planes.

Figure 9: Visualisation of "tunnels" inside ExbBD structures from *S. marcescens* (A and C) and *E. coli* 6TYI structure (B and D), as calculated by the MoleOnline server with the same parameters. The Sm complex is colored as in Figure 8 and the Ec complex is coloured in shades of grey for ExbB and orange and red for ExbD. The tunnels are the green volumes running through the structure, viewed from the periplasmic space (A and B), and from the side (C and D). The average tunnel diameter is 3Å for *S. marcescens* and 2Å for *E. coli.*

Figure 10: ExbBD co-expression is necessary for HasB stabilisation. Immunodetection of HasB in whole cells of *E. coli* JP313Δ*exbBD*Δ*tonB*(pHasB33) also harbouring pBAD24 (control), pBAD24ExbBD$_{Ec}$, pBAD24ExbBD$_{Sm}$ or pBAD24ExbBD$_{Ec-Sm76-88}$, in the presence of either glucose (1mg/ml) or arabinose (40µg/ml), indicated by the + signs. The equivalent of 0.2OD$_{600nm}$ was loaded in each lane.

Figure 11: ExbB α2 helix: comparison between Sm and Ec. ExbB-ExbD structure with residues 76-88 coloured black and showing the aminoacid side chains. A, general view of the Sm pentamer; B, an enlarged view of A; C and D, two perpendicular close-up views of the exchanged region 76-88 in Sm (C) and 39-51 in Ec (D). E and F: Helical wheel representations of the swapped regions between ExbB$_{Sm}$ and ExbB$_{Ec}$ TM1. E: ExbBSm 76-88; F: ExbBEc 39-51. The red boxes show the helical face in interaction with the membrane.

1132

A

```
  1 ------------------------------------------------------------------------------MGNNLMQTDLSVWGMYQHADIVVKCVMIGLILASVVTWAIFFSK  44
                                                                                 DLS+WGMYQHAD VVK VMIGL+LAS+VTW I F+K
  1 MKTAGKNLNQGSFGQGRAQWGKAFGRSLMASMVLVVGLAGSAQAAPAANPAVTESVAPTTAPAPAAAAPESITPVNPAPTIQPPETRGMDLSIWGMYQHADAVVKAVMIGLVLASIVTWTILFAK  125
                      20                  40                  60                  80                 100                 120
              60                  80                 100                 120                 140                 160
 45 SVEFFNQKRRLKREQQLLAEARSLNQANDIAADFGSKSLSLHLLNEAQNELELSEGSDDNEGIKERTSFRLERRVAAVGRQMGRGNGYLATIGAISPFVGLFGTVWGIMNSFIGIAQTQTTNLAV  169
    E    KRRL+REQ  LAEARSL++A+++A +F  +S+S  LLN+AQNELELS  S+DN GIKERT FRLERRVAA  R MGRGNG+LATIGAISPFVGLFGTVWGIMNSFIGIA +QTTNLAV
126 GSELLRAKRRLRREQLALAEARSLDEASELAQNFSPESVSAVLLNDAQNELELSAESNDNNGIKERTGFRLERRVAAYSRNMGRGNGFLATIGAISPFVGLFGTVWGIMNSFIGIAHSQTTNLAV  250
              140                 160                 180                 200                 220                 240
          180                 200                 220                 240
170 VAPGIAEALLATAIGLVAAIPAVVIYNVFARQIGGFKAMLGDVAAQVLLLQSRDLDLEASAAAHPVRVAQKLRAG                                                    244
    VAPGIAEALLATA+GLVAAIPAVVIYN+FAR I G +A +GDVAAQVLLLQ RDLDL A+A A   + A +LRAG
251 VAPGIAEALLATAMGLVAAIPAVVIYNIFARVISGHRAQVGDVAAQVLLLQGRDLDLAATAEAKRSQHAHQLRAG                                                    325
```

B

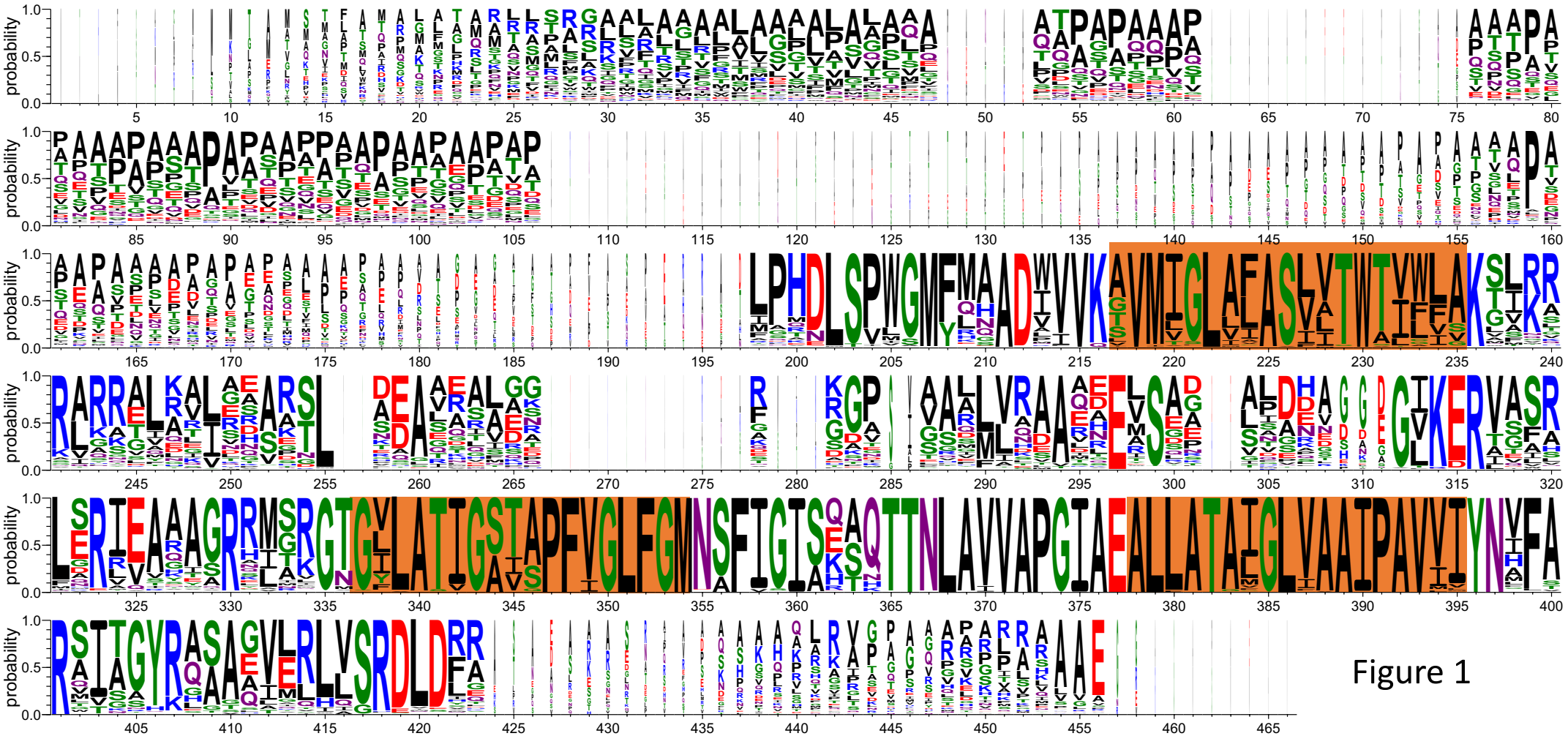


Figure 1

WebLogo 3.7.4

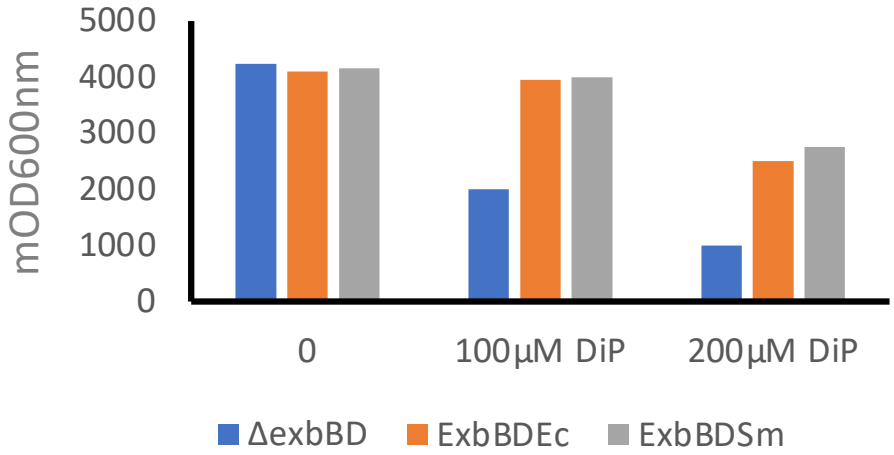
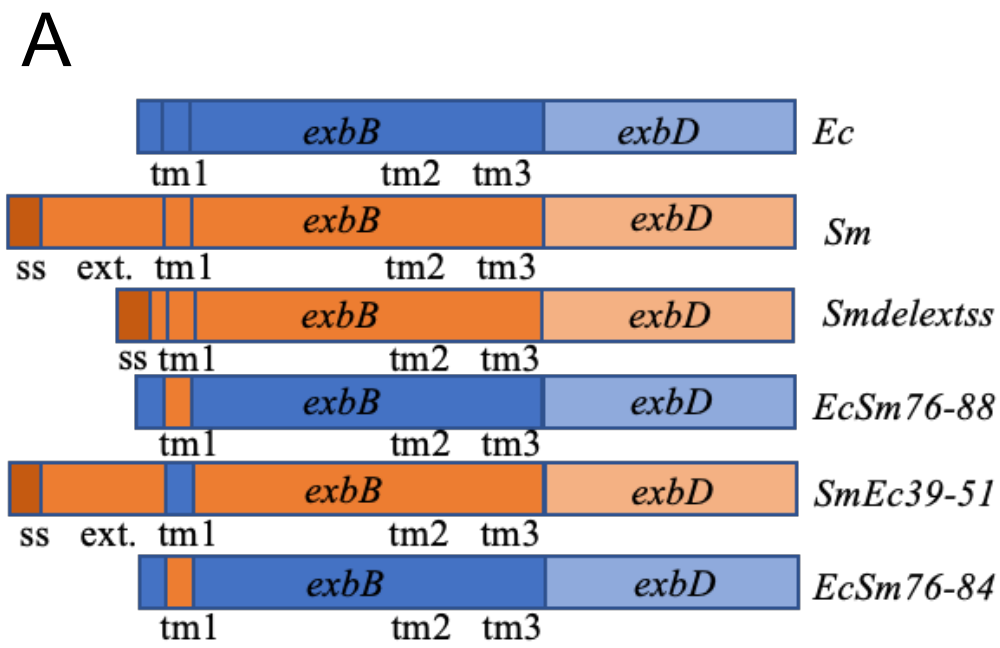
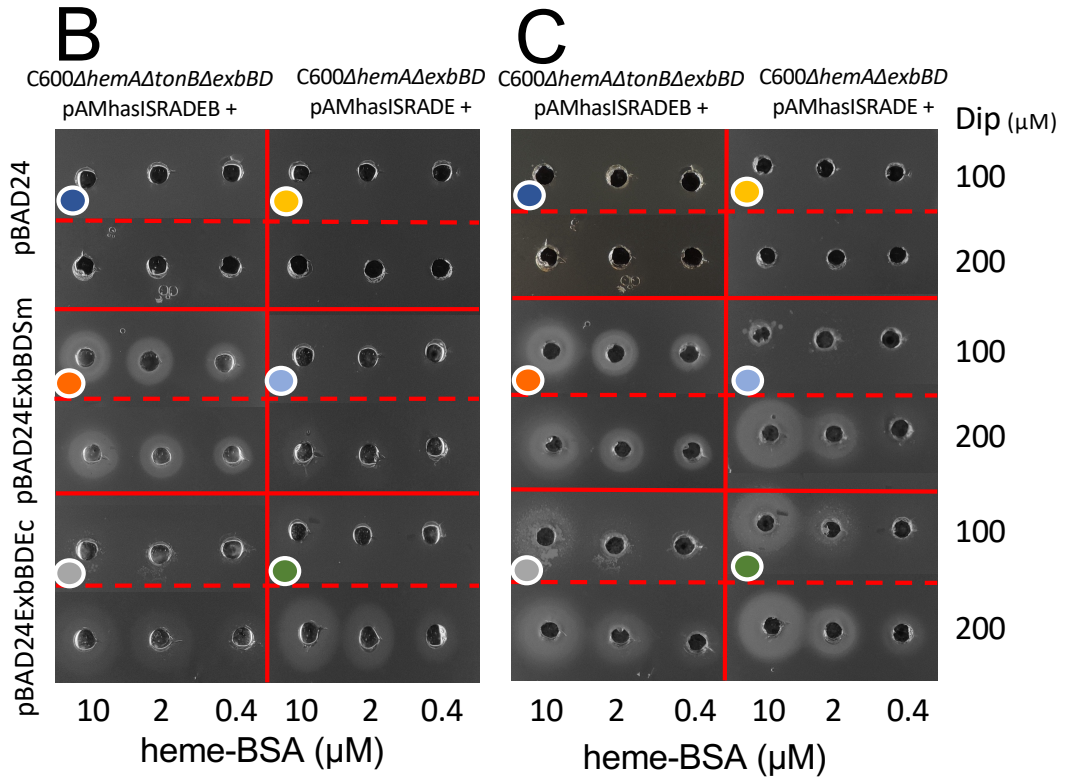

5000
4000
3000
2000
1000
0
mOD600nm
0
100µM DiP
200µM DiP
ΔexbBD
ExbBDEc
ExbBDSm

Figure 2


A
exbB
exbD
Ec
tm1
tm2
tm3
Sm
ss
ext.
Smdelextss
EcSm76-88
SmEc39-51
EcSm76-84


B
C
C600ΔhemAΔtonBΔexbBD
pAMhasISRADEB +
C600ΔhemAΔexbBD
pAMhasISRADE +
Dip (μM)
pBAD24
pBAD24ExbBDSm
pBAD24ExbBDEc
100
200
10
2
0.4
heme-BSA (μM)

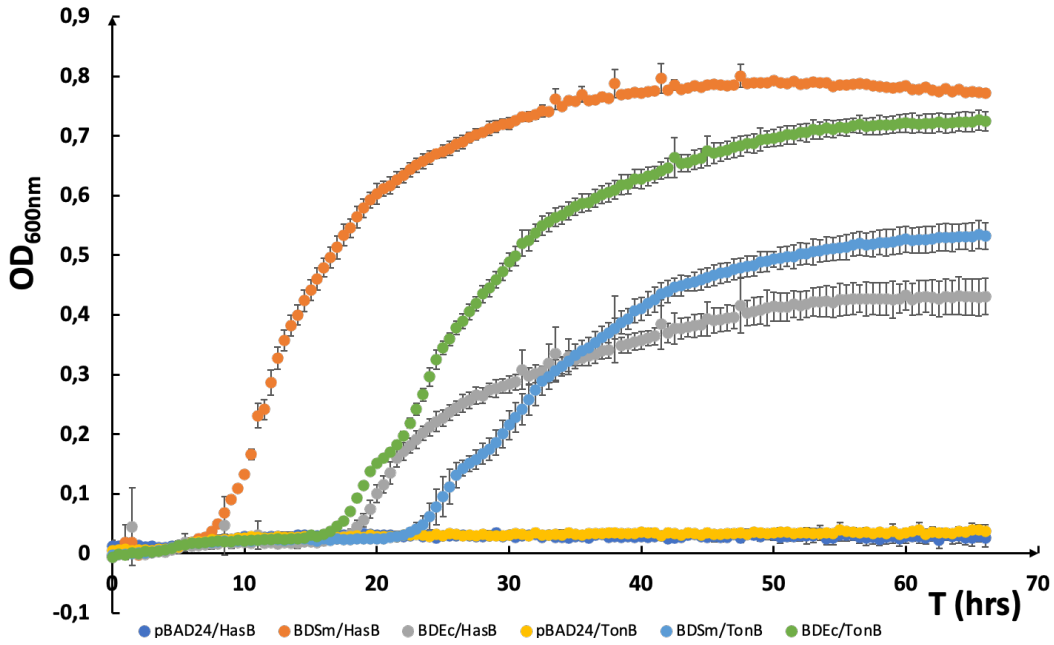

D
OD600nm
T (hrs)
pBAD24/HasB
BDSm/HasB
BDEc/HasB
pBAD24/TonB
BDSm/TonB
BDEc/TonB

Figure 3

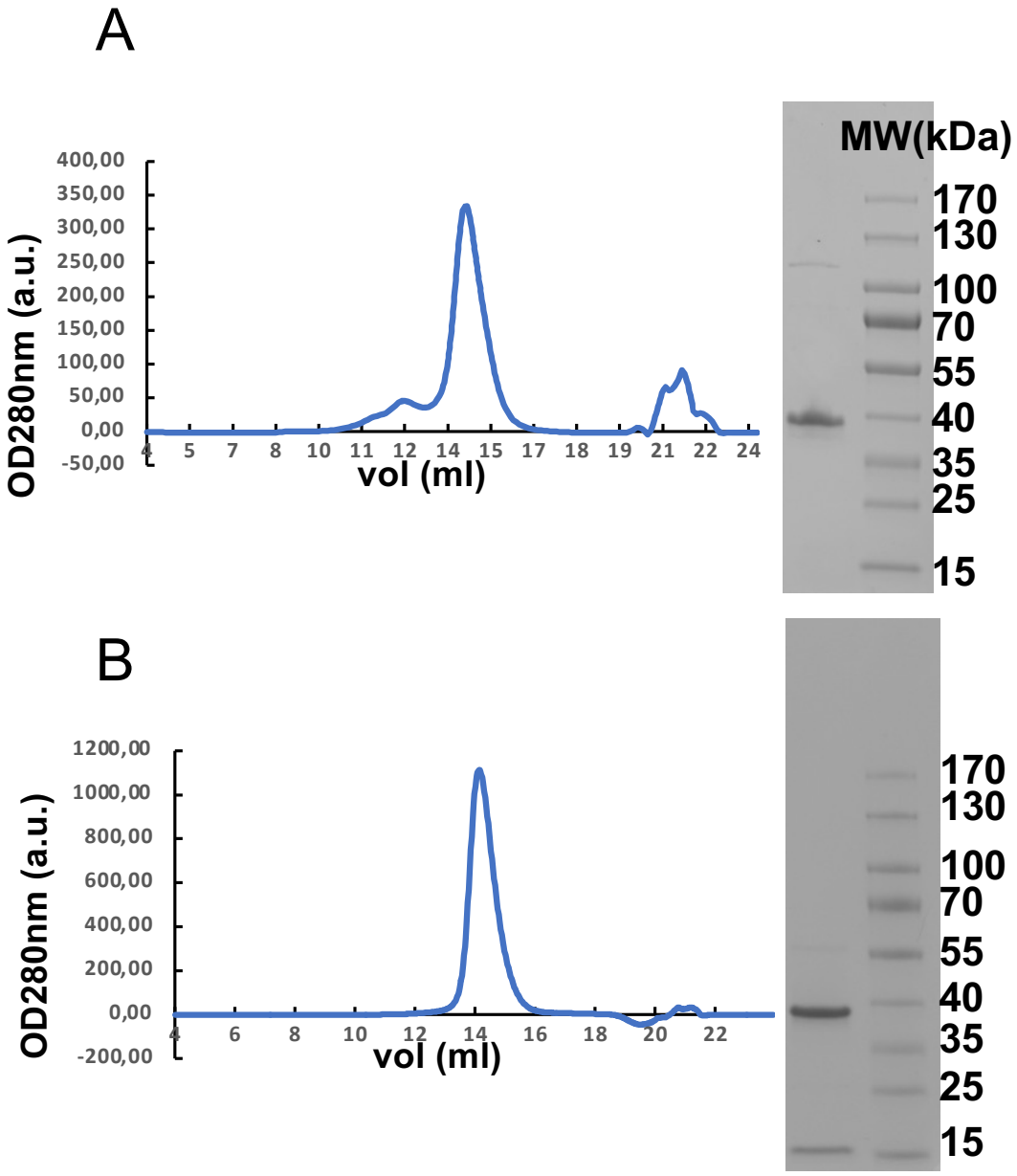

A
400,00
350,00
300,00
250,00
200,00
150,00
100,00
50,00
0,00
-50,00
OD280nm (a.u.)
4 5 7 8 10 11 12 14 15 17 18 19 21 22 24
vol (ml)
MW(kDa)
170
130
100
70
55
40
35
25
15
B
1200,00
1000,00
800,00
600,00
400,00
200,00
0,00
-200,00
OD280nm (a.u.)
4 6 8 10 12 14 16 18 20 22
vol (ml)
170
130
100
70
55
40
35
25
15

Figure 4

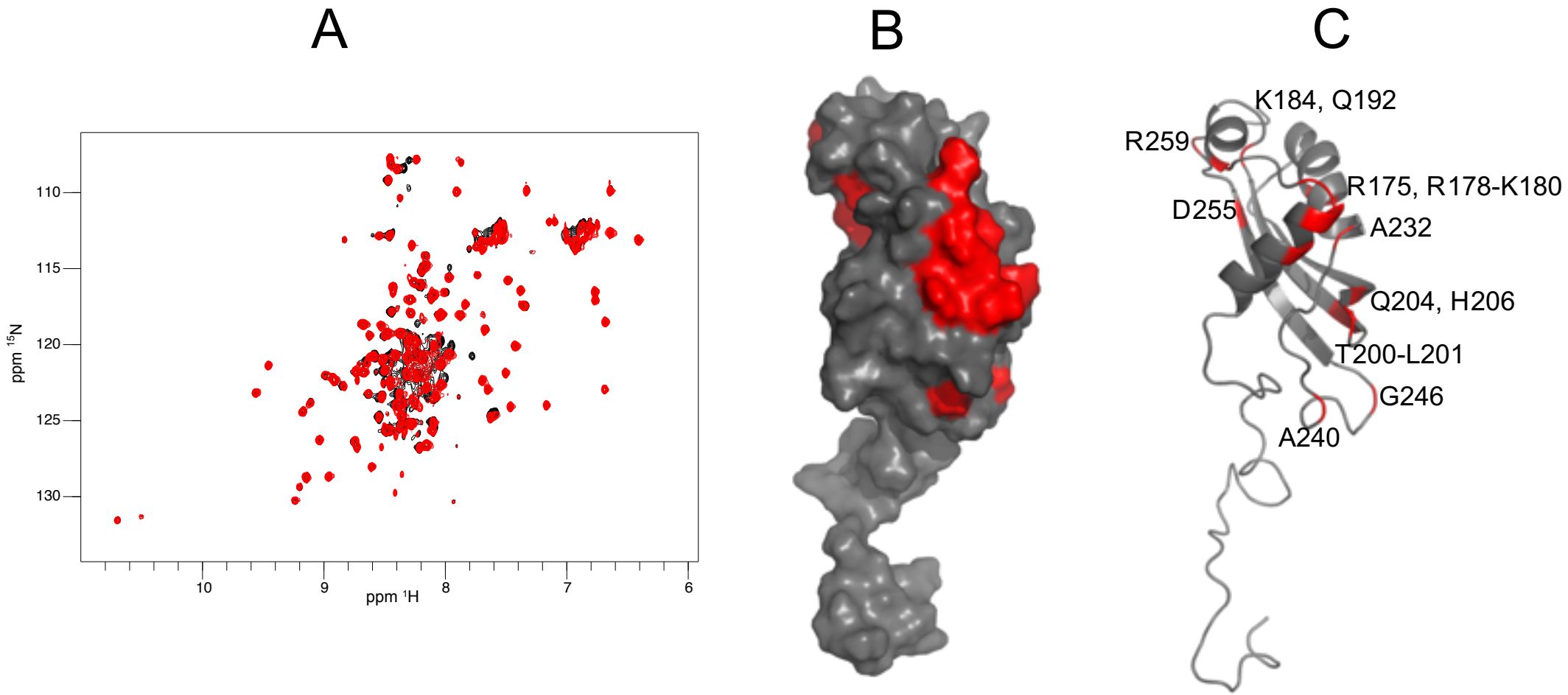

A
B
C
110
115
120
125
130
ppm 15N
10
9
8
7
6
ppm 1H
K184, Q192
R259
R175, R178-K180
D255
A232
Q204, H206
T200-L201
G246
A240

Figure 5

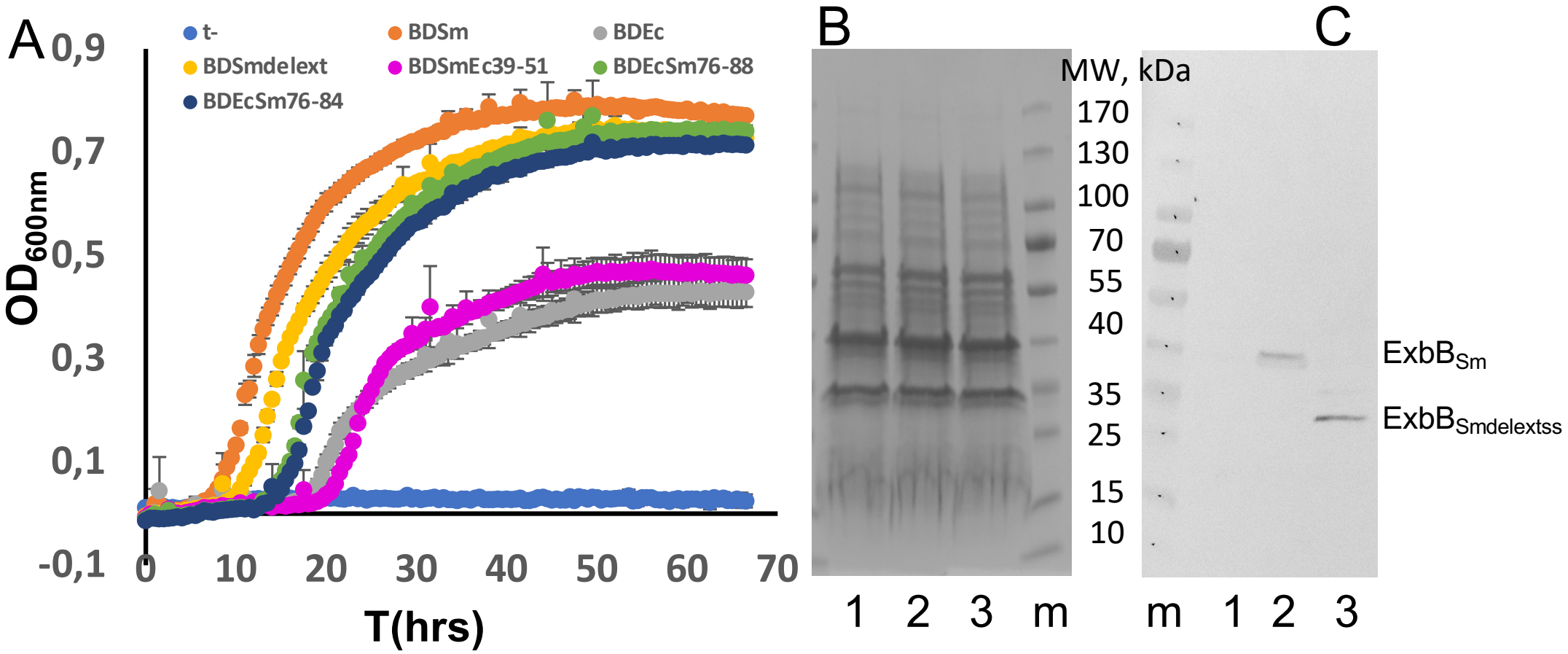

A
t-
BDSm
BDEc
BDSmdelext
BDSmEc39-51
BDEcSm76-88
BDEcSm76-84
OD600nm
0,9
0,7
0,5
0,3
0,1
-0,1
0
10
20
30
40
50
60
70
T(hrs)
B
MW, kDa
170
130
100
70
55
40
35
25
15
10
1
2
3
m
C
m
1
2
3
ExbBSm
ExbBSmdelextss

Figure 6

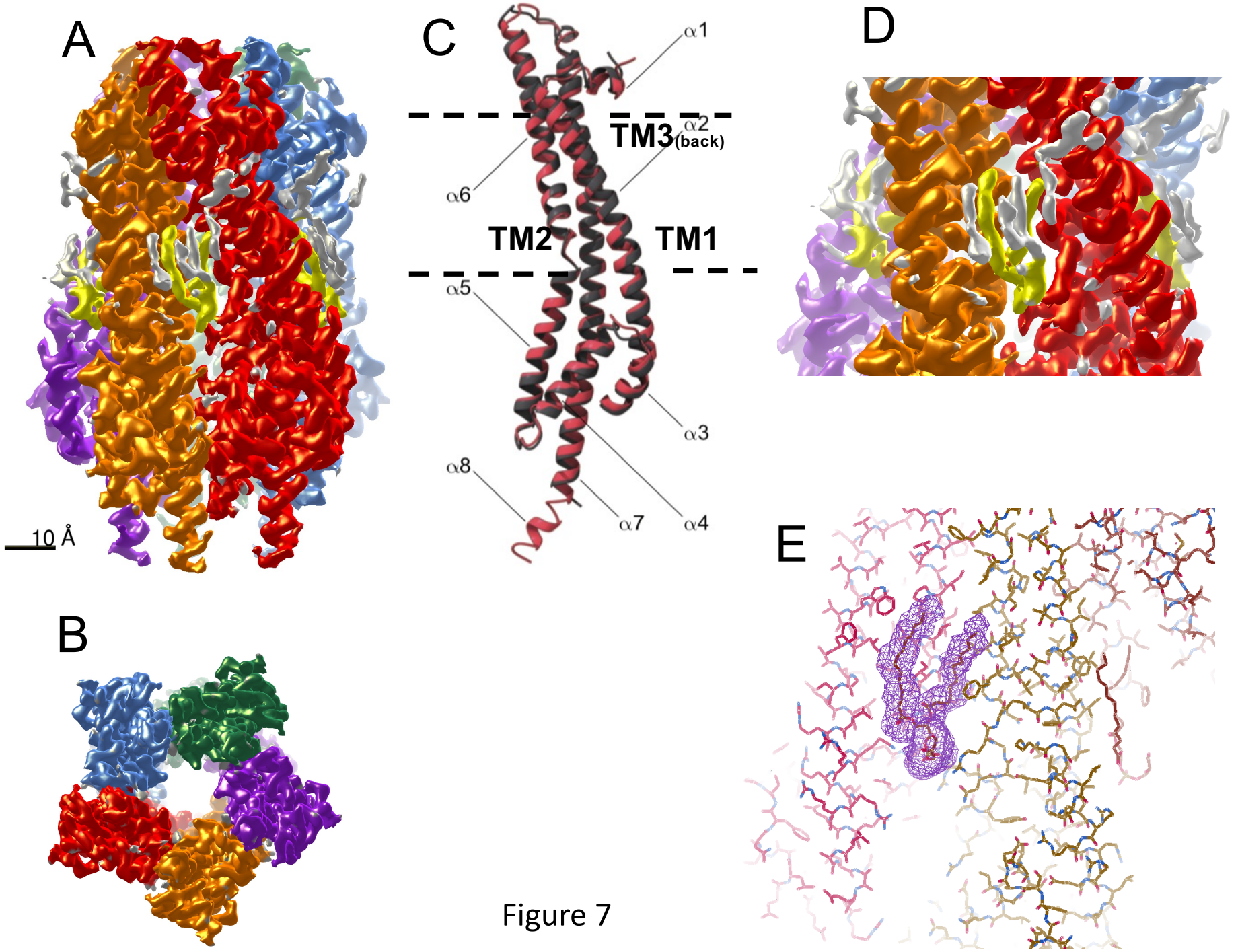


Figure 7

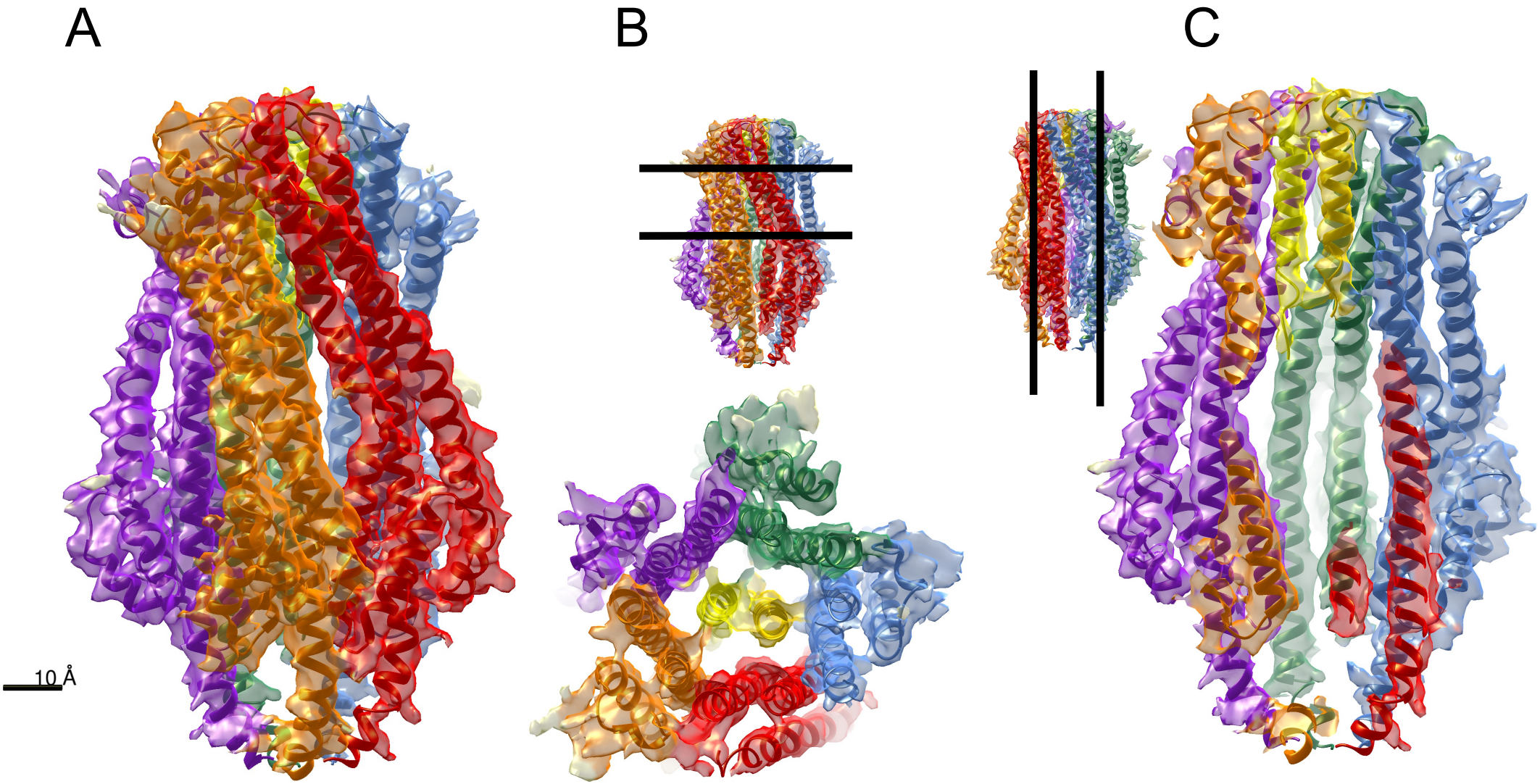


Figure 8

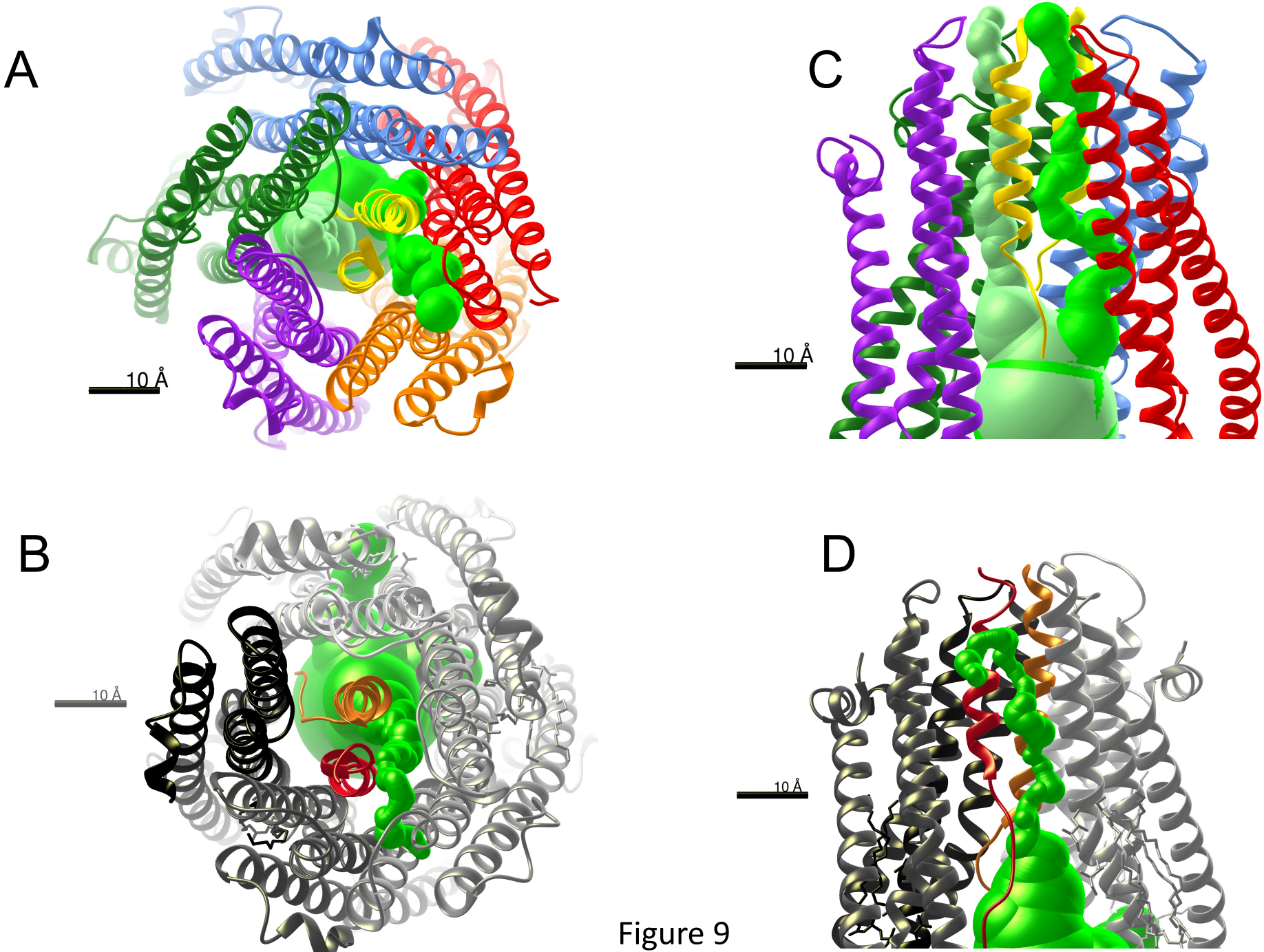


Figure 9

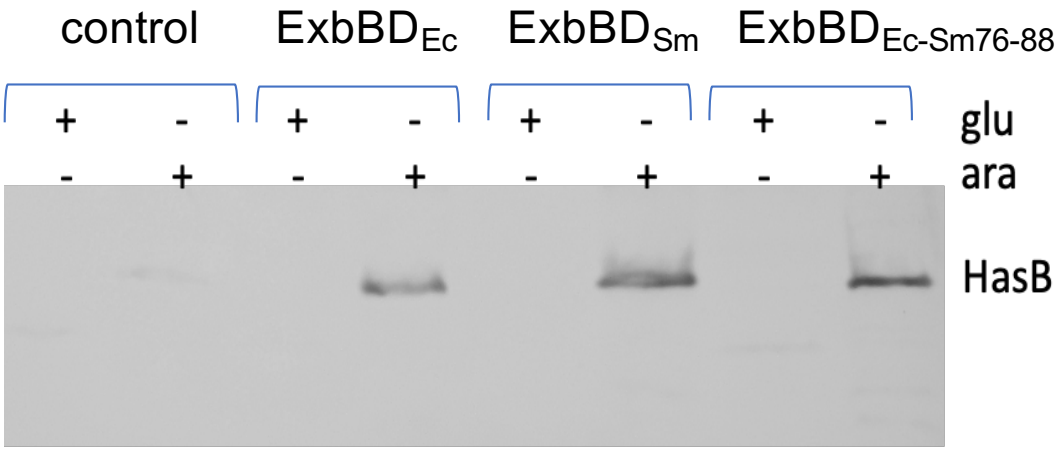


Figure 10

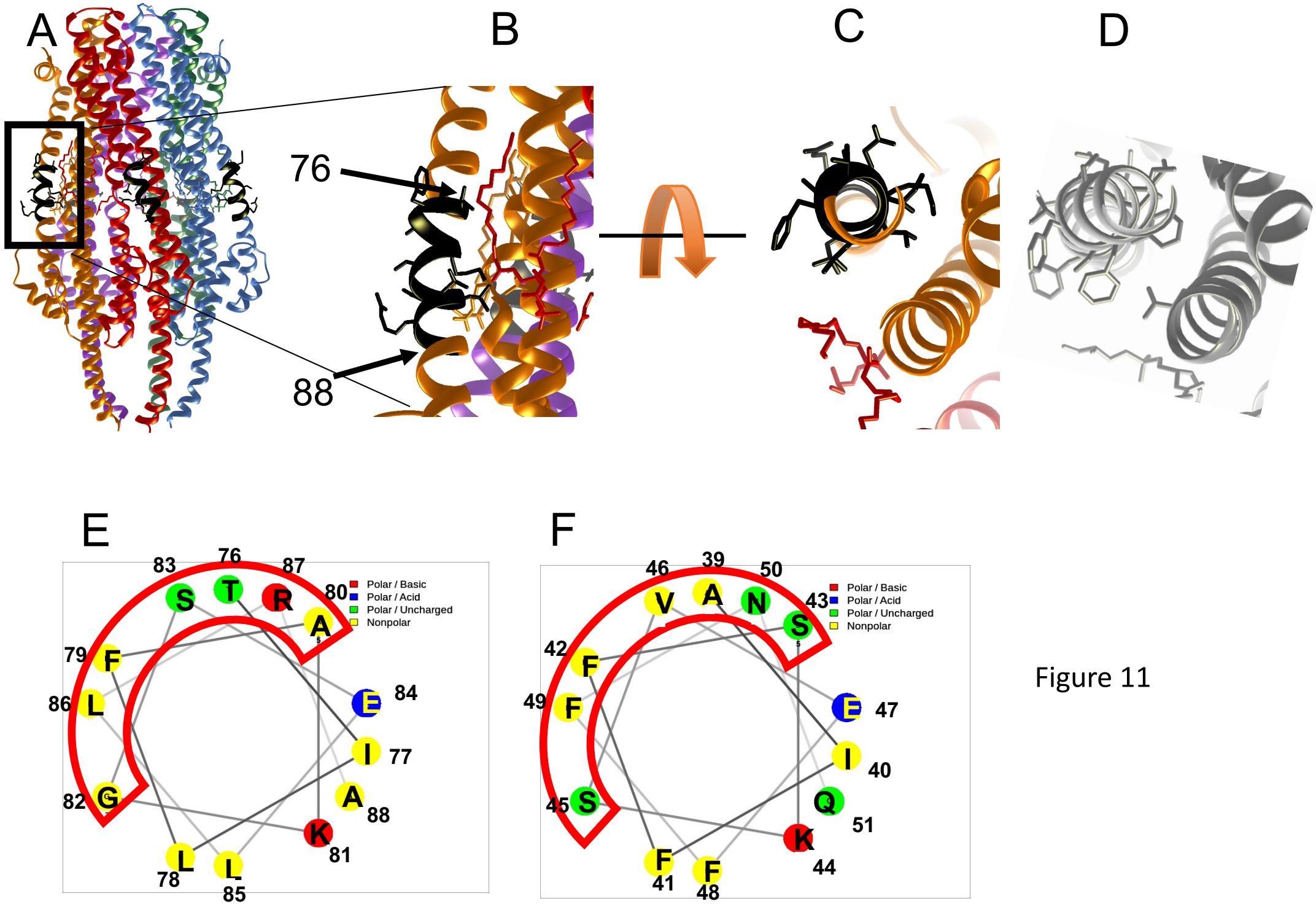

A
B
C
D
76
88
E
F
Polar / Basic
Polar / Acid
Polar / Uncharged
Nonpolar
83
76
87
80
79
86
84
77
82
88
81
78
85
46
39
50
43
42
49
47
40
45
51
44
41
48

Figure 11