

**Studying transmembrane helix association, structure and stability using computational
and experimental techniques**

By

Gladys Díaz Vázquez

A dissertation submitted in partial fulfillment of
the requirement for the degree of

Doctor of Philosophy
(Biophysics)

at the

UNIVERSITY OF WISCONSIN-MADISON

2022

Date of final oral examination: 08/16/2022

The dissertation is approved by the following members of the Final Oral Committee:

Alessandro Senes, Professor, Biochemistry

Katherine A Henzler-Wildman, Professor, Biochemistry

Aaron A Hoskins, Associate Professor, Biochemistry

Qiang Cui, Professor, Chemistry

M Thomas Record Jr., Professor, Biochemistry

Table of Contents

Abstract.....	iii
Acknowledgments.....	v
Chapter 1: Thermodynamic studies of transmembrane protein association; experimental and molecular dynamics simulation approaches.....	1
1.1. Introduction:.....	1
1.2. Sedimentation equilibrium analytical ultracentrifugation (SEAUC).....	3
1.3. Use of Förster resonance energy transfer for the study of TM helix association. ..	7
1.4. Quantitative imaging Förster resonance energy transfer: A way to study transmembrane helix association in plasma membrane derived vesicles.....	11
1.5. Use of TOXCAT to obtain $\Delta\Delta G_{app}$ in an <i>in vivo</i> system.....	15
1.6. GALLEX, a tool to measure dimer stabilities in a biological membrane.....	17
1.7. Steric trapping.....	19
1.8. Use of single molecule photobleaching with total internal reflection fluorescence microscopy to measure oligomerization of transmembrane proteins.....	22
1.9. Free energy calculation of membrane helix association using Molecular Dynamics simulation techniques.....	26
1.10. Summary and concluding remarks: Experimental and molecular dynamics approaches to study thermodynamics of transmembrane protein oligomerization....	38
1.11. Outline of thesis Chapters 2 – 4: Studying transmembrane helix association, structure and stability using computational and experimental techniques.....	40
1.12. References.....	42
Chapter 2: Thermodynamic analysis of GAS _{right} dimerization supports a model in which stability is modulated by weak hydrogen bonding and van der Waals packing.....	49
Abbreviations.....	50
2.1. Abstract.....	51
2.2. Introduction.....	52
2.3. Results and Discussion.....	58
2.4. Conclusions.....	73
2.5. Experimental Procedures.....	77
2.6. Acknowledgments.....	83
2.7. Supplementary Figures and Tables.....	83
2.8. References.....	95
Chapter 3: The FtsLB subcomplex of the bacterial divisome is a tetramer with an uninterrupted FtsL helix linking the transmembrane and periplasmic regions.....	102
3.1. Abstract.....	103
3.2. Introduction.....	104
3.3. Results and discussion.....	112
3.4. Conclusions.....	141
3.5. Experimental Procedures.....	143
3.6. Author contributions.....	158
3.7. Acknowledgments.....	158
3.8. Supplementary Figures and Tables.....	160

3.9. References.....	192
Chapter 4: The coiled-coil domain of E. coli FtsLB is a structurally detuned element critical for modulating its activation in bacterial cell division.....	200
4.1. Abstract.....	201
4.2. Introduction.....	202
4.3. Results and Discussion.....	210
4.4. Conclusions.....	236
4.5. Experimental Procedures.....	239
4.6. Author contributions.....	247
4.7. Acknowledgments.....	247
4.8. Supplementary figures and tables.....	247
4.9. References.....	266
Chapter 5: Conclusions and Future Directions.....	274

Abstract

Membrane proteins carry out a diverse and essential set of functions, such as communication, transport, and enzymatic activity. Many of these membrane proteins must form homo- and hetero-oligomers to carry out their functions. A better understanding of the physical rules and properties that govern membrane protein association could give insight into their function and mechanism of action. In the research presented here a combination of computational and experimental techniques was used for the biophysical study of different membrane protein complexes.

In Chapter 2 I used FRET to study the oligomerization of transmembrane domains (TMDs) predicted to form GAS_{right} homodimers. I measured the thermodynamic stability of 9 constructs including the prototypical GAS_{right} homodimer glycophorin A and its monomeric mutant G83I. These TMDs associate with a wide range of thermodynamic stabilities, which compared quite favorably with their homodimerization propensities derived from the TOXCAT genetic reporter assay. Our findings suggested that TOXCAT is thermodynamically driven and further support the hypothesis that van der Waals and Ca-H hydrogen bonds are key players in the thermodynamic stability of GAS_{right} dimers. These results can be later combined with computational approaches to measure the relative contribution that different interactions have to the association of GAS_{right} helices.

In Chapter 3 and 4 we looked at the association between the membrane proteins FtsL and FtsB, which form an essential subcomplex of the bacterial divisome. Their specific mechanism of action is not fully understood and the experimental structure of

their complex has not been obtained. In our studies we combined computational modeling, molecular dynamics simulations, and experimental approaches to obtain a better understanding of the structure of FtsLB and its role in cell division. Our findings suggested FtsL and FtsB form a L₂:B₂ heterotetramer, with the more likely configuration being a four helix coiled coil in the TM region, and two L:B coiled coil dimers in the periplasmic region. Features identified in the predicted model, such as the transmembrane interface and the cluster of polar interfacial residues in the coiled coil, were found to be important for carrying out proper cell division.

Acknowledgments

I would like to thank my advisors Dr. Alessandro Senes and Dr. Qiang Cui, for giving me the opportunity to be part of their laboratories and learn from them about science and different interesting techniques that can be used to study biological systems. Thank you for your support, teachings and advice. I would also like to thank my committee members for their advice, which has helped me learn skills that are important to have when doing science, and have also helped me think about my research more thoroughly. I would also like to thank current and previous laboratory members for their help and support. Its been great to work in such a collaborative environment.

I would like to thank my friends and family, especially my mom Luz and my little brother Jan Luis. I am thankful for your support and presence in this journey. Thank you for cheering for me and believing in me. In addition I would like to especially thank my partner Samson.

Chapter 1: Thermodynamic studies of transmembrane protein association; experimental and molecular dynamics simulation approaches

1.1. Introduction:

Membrane protein interaction and oligomerization is a critical mechanism by which cells grow, divide, and sense their environment, yet the physical rules that govern this complicated dance remain poorly understood. A better understanding of the physical interactions that take place in the association of membrane protein helices could give insight into their structure, regulation and function of many biological processes. In order to obtain a better understanding of the physical rules that govern the association of transmembrane (TM) proteins, thermodynamic experiments of the process can be done. Thermodynamic studies give information about the stability of the system in specific conditions, and how changes to the system affect the stability of the membrane protein under study. In the simplest case of TM protein oligomerization, we consider two states: a monomeric state where the TM proteins are separated, and a dimeric state where they interact with each other (Figure 1.1). These two states exist in equilibrium and different methods can be used to obtain the equilibrium constant, and the free energy of association.

For a monomer-dimer equilibrium ($D \rightleftharpoons M^2$) the dissociation equilibrium constant can be defined as follows:

$$K_{eq} = \frac{\chi_M^2}{\chi_D}$$

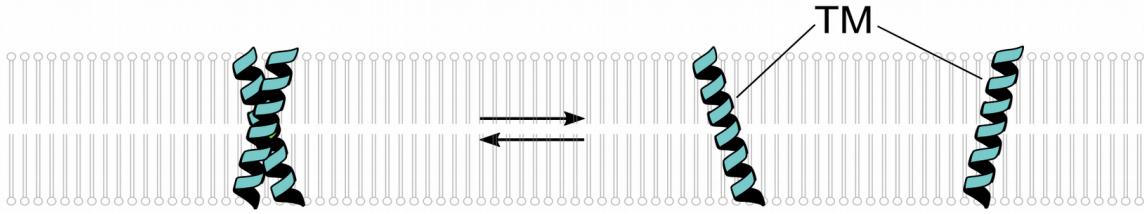


Figure 1.1: The two states observed in the dissociation of two membrane helices that are in a monomer-dimer equilibrium is shown. On the left we have the dimer state, and on the right the monomer state.

where K_{eq} represents the dissociation equilibrium constant, χ_M is the monomer mole fraction and χ_D is the dimer mole fraction. The equilibrium constant is related to the standard state free energy via the following equation:

$$\Delta G^\circ = -RT \ln(K_{eq})$$

where ΔG° is the standard state free energy change, R is the gas constant, and T is the temperature. Using these relations and experimental approaches where the relative monomer to dimer population can be varied and directly or indirectly measured, thermodynamic information can be extracted.

To date there exist a variety of *in vivo* and *in vitro* experiments to obtain thermodynamics of helical membrane protein association. In addition it is possible to use molecular dynamics (MD) simulations with enhanced sampling methods to get a better understanding of this process. One of the advantages of MD simulations is that an atomistic and molecular picture of the association process can be obtained, giving a better understanding on the possible contribution that different interactions have in the

association process and giving molecular level information that could be tested in the experiments.

In this chapter, a summary of different experimental and computational approaches that have been used to study thermodynamics of association of helical membrane proteins will be given. The discussion will start with an explanation of different experimental approaches and studies in which these approaches have been used. After that the use of MD simulations to study association of TM proteins will be discussed. For many of the studies that will be presented, the transmembrane domain (TMD) of the human erythrocyte sialoglycoprotein glycophorin A (GpA) was used as a model system of membrane protein oligomerization. GpA is a well known homodimer for which the NMR structure of the TMD was published in 1997 (MacKenzie et al., 1997). The thermodynamics of association of GpA has been extensively studied, and it has become a model system for studies done on the thermodynamics of dimerization of TM helices.

1.2. Sedimentation equilibrium analytical ultracentrifugation (SEAUC)

One of the techniques that has been used to study the reversible association of TM α -helices is analytical ultracentrifugation (Doura and Fleming, 2004; Fleming and Engelman, 2001; Fleming et al., 1997, 2004). In this approach the protein of interest is solubilized in detergent and centrifuged at high speeds. Over time a concentration gradient is generated within the sample based on the centrifugation rate, the buoyant force of each oligomeric state, and the equilibrium constant describing the transition of the protein between these states. Then absorbance vs radius information can be used to obtain molecular weight of the protein and possibly the equilibrium constant in the

corresponding environment. Several considerations need to be taken in order to use this approach. First, the detergent used to solubilize the membrane protein could contribute to the buoyant mass of the particle that is sedimenting in the ultracentrifuge. Due to this, the total buoyant molecular weight of the complex contains an uncertain contribution from the detergent. This effect must be mitigated by closely matching the density of the detergent micelles with that of the bulk solvent, generally limiting the detergents that can be used. Second, in order to estimate the equilibrium constant it is important to reach reversible association in the time scale of the experiment. A way to check for reversible equilibrium is by performing a set of SEAUC experiments at different initial concentrations and rotor speeds. If the resulting equilibrium constant is not the same across the different conditions, the sample did not equilibrate successfully.

In 1997 Fleming and coworkers used SEAUC to study the effect that specific point mutations have in the dimerization free energy of (GpA) in the detergent C₈E₅ (Fleming et al., 1997). They found that in these conditions, GpA had a dissociation constant of 240±50 nM. The mutations L75A and I76A both destabilized the dimer with dissociation constants of 1.7±0.2 μM and 4.2±0.9 μM respectively (Fleming et al., 1997). They found that the differences in association energetics could be accounted for by changes in van der Waals interactions between the helices (Fleming et al., 1997).

In another study Fleming and Engelman used alanine scanning mutagenesis and SEAUC on GpA to get a better understanding of how particular interactions between the helices impact the energetics of dimerization (Fleming and Engelman, 2001). They found that mutations to alanine at the interface of the dimer decreased association,

while mutations away from the interface seemed closer to WT or even led to slight stabilization of the dimer (Fleming and Engelman, 2001). Their findings could be rationalized by introduction of steric clashes or by removal of van der Waals interactions which suggested a detailed geometry and packing of the GpA dimer. In another study Doura and Fleming looked at the effect that double alanine substitutions have in the association energetics of GpA and found that most of the double mutants were not simply the sum of each individual mutation (Doura and Fleming, 2004). Their results could also be explained through changes in van der Waals interactions at the dimer interface (Doura and Fleming, 2004).

The use of SEAUC to obtain thermodynamic information on the association of GpA in detergent has increased our understanding of the interactions involved in this process and their possible contribution to the energetics of association. By doing experiments on wild-type GpA and different GpA mutants, they were able to look at the possible effects that different interactions might have in the associated state. The studies mentioned above suggest that van der Waals interactions seem to be key for the stability of the GpA dimer.

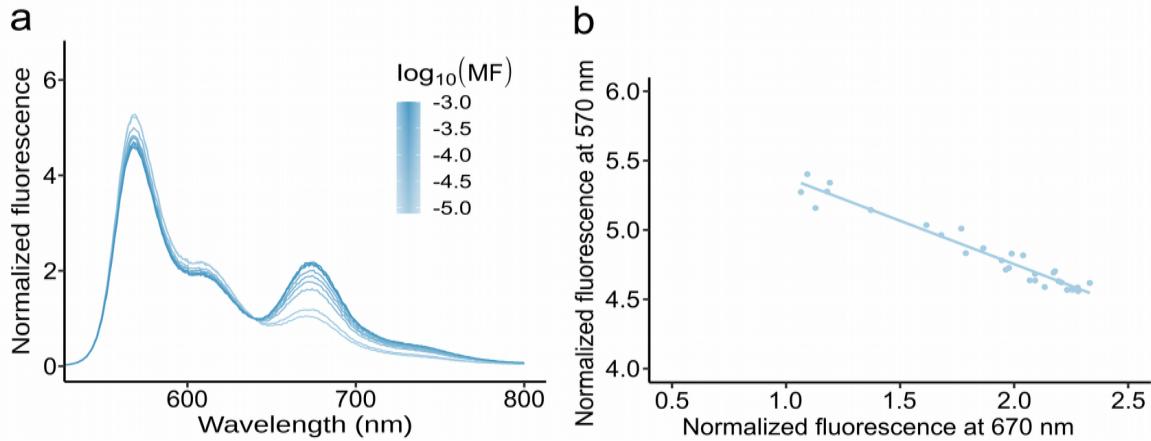


Figure 1.2: Panel a presents the normalized fluorescence intensity vs wavelength for the emission of Cy3-maleimide and Cy5-maleimide. Panel b is the donor vs acceptor normalized fluorescence intensity, where we can see the inverse linear relationship between the two.

Although SEAUC has been a useful experimental approach for the study of membrane protein association, it has limitations on the environment in which the process can be studied. As mentioned, reaching reversibility in the time scale of the experiments is important in order to obtain thermodynamic information. That limits the detergent environment in which the experiments can be performed. Additionally, the detergents must have a similar density to water, completely excluding many popular detergents used in other experiments (Burgess et al., 2008). The dimerization energetics of GpA have been examined in C14 betaine detergent micelles using SEAUC in addition to C₈E₅, where the free energy of dimerization was found to be more favorable in C₈E₅ than in C14 betaine (Fleming et al., 2004). However to date no SEAUC studies have been performed on GpA in more commonly used detergents like

dodecylmaltoside (DDM) (Burgess et al., 2008). In addition, although detergents represent a hydrophobic environment in which TM helices can be studied, the cell membrane is a lot more complex than a detergent micelle, with one of the main components being lipids. Other experimental approaches have been used to study this process in other membrane mimetics.

1.3. Use of Förster resonance energy transfer for the study of TM helix association

One of the main limitations of SEAUC is the environment in which the protein of interest can be studied. SEAUC thermodynamic studies are limited to folding of membrane proteins in a detergent in which reversibility can be obtained and in which the density is similar to water, which are relatively few. Förster resonance energy transfer (FRET) is another approach that has been widely used to study membrane protein association with the advantage that studies can be done not just in detergent but also in lipid bilayers (Adair and Engelman, 1994; Anbazhagan and Schneider, 2010; Khadria and Senes, 2015; Li et al., 2005a; Merzlyakov et al., 2006; You et al., 2005). FRET is the process by which two fluorescent molecules interact with each other via the nonradiative transfer of energy of one of these molecules called the donor, to the other molecule called the acceptor. For the transfer of energy to occur, the donor molecule needs to be excited with a wavelength that falls within its excitation spectra regime, and if the donor is in close proximity to the acceptor ($\sim <100$ nm), it can transfer its energy to the acceptor molecule as long as the emission spectra of the donor overlaps with the excitation spectra of the acceptor.

When using FRET for the study of TM protein association, a pair of fluorophores that follows the requirements mentioned above is selected and chemically added to the TM protein of interest or the construct where the TMD of interest is. Close proximity of two entities can be detected by looking at the emission spectra of the sample, where the nonradiative transfer of energy of the donor fluorophore, to the acceptor fluorophore would result in the decrease of the emission fluorescence of the donor, and the increase of the emission fluorescence of the acceptor (Figure 1.2). When studying thermodynamics of dimerization of helical membrane proteins with FRET care needs to be taken when choosing where in the construct the fluorophore is going to be added such that it does not affect the association energetics in a major way. In addition it is important to measure FRET in a wide range of concentrations that span the monomer to dimer transition so that the equilibrium constant can be extracted from the data. A key point is that for membrane proteins, the relevant measure of concentration is the amount of protein with respect to the hydrophobic environment, rather than the amount of protein relative to the volume of the full solution where it is (including water, buffer, salt etc.). Adding water to proteoliposomes will do nothing to the monomer-dimer population, but equilibrating the sample with more lipids will dilute the protein and drive the population towards the monomer state.

In 1994 Adair and Engelman published a study where they used resonance energy transfer to see if GpA forms dimers in DMPC lipid bilayers (Adair and Engelman, 1994). In their studies, they showed that GpA forms dimers in the lipid environment they used, supporting the possibility of GpA existing as a dimer in biological lipid bilayers and the

dimers previously observed in SDS not being just an effect of the detergent on it (Adair and Engelman, 1994; Lemmon et al., 1992). Since, many other studies have been published that used FRET to investigate TM helix oligomerization thermodynamics in different environments. In 1999 Fisher and coworkers used FRET to study the dimerization thermodynamics of GpA in various detergent micelles (Fisher et al., 1999). The detergents used were sodium dodecyl sulfate (SDS, a negatively charged detergent) and the zwitterionic detergents DDMAB and DPC (Fisher et al., 1999). They found the stability of the dimer to decrease in SDS when compared to the DDMAB ($K_D = 0.08 \pm 0.04 \mu\text{M}$ in DDMAB at 25°C) and DPC ($K_D = 0.16 \pm 0.08$ in DPC at 45°C) detergents (Fisher et al., 1999). This decrease in stability was shown to be independent of the secondary structure through far-UV circular dichroism, suggesting that helix formation and helix association are independent processes (Fisher et al., 1999).

There exist different types of lipid systems that could be used to look at association of TM proteins, and it is important to know if the system that is being used allow for the successful determination of thermodynamic quantities. In 2005 a study was published where You and coworkers used FRET to study the association of the fibroblast growth factor receptor 3 (FGFR3) TM domain in liposomes composed of POPC lipids (You et al., 2005). They performed FRET measurements in three different liposomal systems: multilamellar vesicles (MLVs), large unilamellar vesicles (LUV); and small unilamellar vesicles (SUVs). While the results obtained for FGFR3 in MLVs and LUVs were similar, they noticed a decrease in the amplitude of the fluorescence spectra for SUVs after several minutes, a behavior not observed in the other two lipid systems (You et al.,

2005). They mentioned that SUVs are known to be far from equilibrium and they believed the SUVs to be aggregating and coming out of solution (You et al., 2005). Although they did not know the exact mechanism for what they observed with the signal of the SUV sample, the results suggested the SUV liposomal system to not be appropriate for FRET measurements (You et al., 2005). One of the important findings that they obtained was that the determined FRET efficiencies depended solely on the protein-to-lipid ratio, which allows the use of FRET in liposomes for the thermodynamic studies of TM association in lipids membranes (You et al., 2005). The free energy of dimerization of the TM domain of FGFR3 was found to be approximately -3 kcal/mol in POPC bilayers (Li et al., 2005a).

Artificial lipid bilayers are an important environment for the study of TM helix association, but it is necessary to keep in mind that biological membranes are highly heterogeneous, composed of lipids with a wide variety of headgroups and hydrophobic tails. These different components can have a dramatic impact on the physical properties of the lipid bilayer and likely influence the events that happen in the cell. One key property of lipid bilayers is the thickness of the hydrophobic segment, which can influence TM helix properties. In 2010 Anbazhagan and Schneider used FRET to look at the dimerization of the TMD of GpA in different lipid environments by changing the thickness of the lipid bilayer (Anbazhagan and Schneider, 2010). They found that bilayer thickness affects the GpA TM monomer-dimer equilibrium, with hydrophobic matching giving place to more stable dimers (Anbazhagan and Schneider, 2010).

To this point we have discussed the use of FRET for the thermodynamic studies of TM homodimerization, but studying heterodimerization with this technique is also quite tractable. The only modification is to label one protein with the donor fluorophore and the other protein with the acceptor fluorophore. In 2006 Merzlyakov and coworkers used FRET to study heterodimerization of TM helices in liposomes (Merzlyakov et al., 2006). Duneau and coworkers published a study where they used FRET to determine the energetics of heterodimerization of the TM domain of the HER receptor family where they showed that these proteins can both homodimerize and heterodimerize with different affinities (Duneau et al., 2007). These relative affinities may be important for regulating their interactions and cell signaling processes.

As we can see from previous studies FRET has been a keystone experimental approach to examine membrane protein oligomerization. One of the key advantages of FRET is the fact that the monomer-dimer equilibrium can be studied not just in detergent, but also in lipid bilayers which more closely approximates biological membranes. FRET opened the door to researching thermodynamics of membrane helix association in different environments, but there remained the question of how to do thermodynamic experiments on this process in biological membranes. In the next section a method that uses FRET to study the energetics of association of transmembrane helices in cell membrane derived vesicles will be discussed.

1.4. Quantitative imaging Förster resonance energy transfer: A way to study transmembrane helix association in plasma membrane derived vesicles

Studying thermodynamics of association of membrane helices in biological membrane mimetics like detergents and artificial lipid bilayers has increased our

understanding of this process, but changing the mimetic can change the thermodynamic stability of the dimer. This begs the question: to what extent are thermodynamic stabilities obtained in detergent or artificial membranes biologically relevant? One of the differences between the biological environment and these mimetics is molecular crowding. How does the molecular crowding that exists in the biological environment affect the stability of the dimer? In order to address this question, scientists found a way to study TM helix association in cell-derived vesicles, introducing the method quantitative imaging Förster resonance energy transfer (QI-FRET) (Li et al., 2008).

In the QI-FRET method, TMDs are fused to donor and acceptor fluorescent proteins and transfected into cells. Vesicles derived from these cells are then analyzed with fluorescence microscopy. In order to determine the dissociation constant in this environment, three parameters need to be obtained from the microscopy: the donor concentration C_D , the acceptor concentration C_A , and the FRET efficiency E . In order to determine those three parameters, the donor-acceptor FRET pair needs to be chosen such that a donor scan (to get a fluorescence image of the donor only when it is excited), a FRET scan (to get the fluorescence of the acceptor when the donor is excited), and an acceptor scan (to get the fluorescence of the acceptor when it is excited) can be taken. Once an appropriate donor-acceptor pair has been chosen, the donor and acceptor bleed-through coefficients need to be determined, donor and acceptor intensities need to be calibrated by using purified proteins of known concentration, and the gauge factor needs to be determined. Once the necessary parameters are determined for a specific FRET pair, protein-protein interaction can be

studied in the cell-derived vesicles, by labeling the protein of interest and calculating C_A , C_D , and E , which requires the sensitized acceptor emission, and the intensity of the donor emission when the acceptor is not excited. With these parameters the dissociation constant can be determined. The QI-FRET method was validated by comparing the FRET efficiencies in cell-derived vesicles and in buffer where good agreement was found (Li et al., 2008).

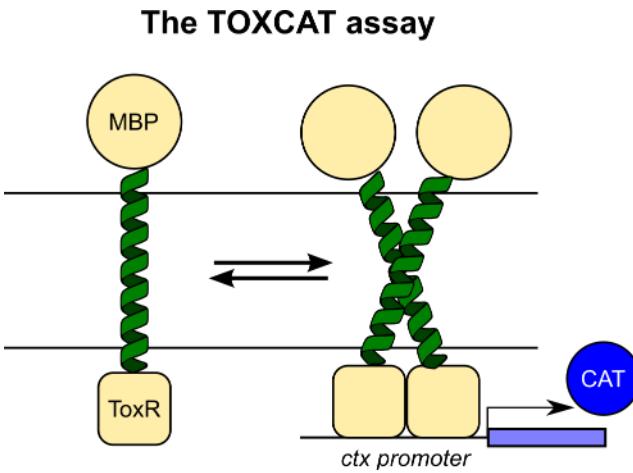


Figure 1.3: TOXCAT is a genetic reporter assay that gives information about TM homodimerization propensity in the inner membrane of *E. coli*. In this assay the TM of interest is fused to the transcription factor ToxR which upon dimerization results in the expression of the CAT enzyme.

In 2010 Chen and coworkers published a study where they used QI-FRET to determine the free energy of dimerization of the TMD of GpA (Chen et al., 2010). They found the free energy of dimerization for GpA in this environment to be -3.9 ± 0.2 kcal/mol (Chen et al., 2010).

The approaches discussed above are either *in vitro* or in the case of QI-FRET, *ex vivo*. However, it is possible to use genetic reporter assays to measure the relative stability of membrane protein oligomers in living cells. In the next two sections, studies that used TOXCAT and GALLEX to obtain thermodynamic information on dimerization of TM helices in the *E. coli* inner membrane will be discussed.

1.5. Use of TOXCAT to obtain $\Delta\Delta G_{app}$ in an *in vivo* system

To measure relative stabilities *in vivo*, it is necessary to couple the dimerization process to some sort of measurable output. This have been done by fusing the helical TMD of interest to a transcription factor, which upon association leads to a measurable change in expression of a reporter gene. TOXCAT is a genetic reporter assay that reports on TM helix homodimerization propensities in the *E. coli* inner membrane (Russ and Engelman, 1999). In the TOXCAT assay, the TM of interest is fused to the transcription factor ToxR and the periplasmic maltose binding protein MalE (Figure 1.3). The maltose binding protein is used to check that the protein is inserted in the inner membrane with the right direction (ToxR towards the cytoplasm), by transforming the plasmid into MalE-deficient *E. coli* and growing it in a media where maltose is the only source of carbon. The transcription factor ToxR binds to the DNA after it dimerizes, resulting in the expression of the enzyme chloramphenicol acetyltransferase (CAT) (Figure 1.3). This means that when the TM of interest dimerizes, ToxR dimerizes and activates transcription of CAT, whose expression is measured with an enzymatic assay which gives information about the homodimerization of the TMD. If this is done for different TM sequences comparison of their homodimerization propensities can be done to see for example how mutations in a sequence affects their homodimerization propensity in the *E. coli* inner membrane. But, there is a caveat: the homodimerization propensity of a TM can be affected by the concentration of that TM in the membrane (the amount of TM per lipid), and the concentration of the TM in the membrane can be

affected by the expression of the TM chimera. This means that the expression of the TMs under study should be taken into account when interpreting the results.

Even if expression is correctly accounted for, at first it was an open question whether TOXCAT was a thermodynamically driven process. In 2007 Duong and coworkers published a study where they used TOXCAT to determine the apparent change in free energy for a group of GpA point mutants that were previously studied using SEAUC (Duong et al., 2007). In order to obtain the $\Delta\Delta G_{app}$ with respect to wild type GpA, they came up with a theoretical approach that assumes that the mole fraction of dimeric fusion protein is very small in comparison to the mole fraction of monomeric fusion protein (Duong et al., 2007). By doing this they were able to simplify the expression for the association equilibrium constant such that it depends only on the dimeric mole fraction and the total mole fraction of the fusion protein. To obtain the dimeric mole fraction they used the CAT_{lysate} , and to obtain the total mole fraction of the fusion protein they used the western blot band intensity for the specific construct (Duong et al., 2007). But since these quantities relate to the specific mole fractions via a proportionality constant, the proportionality constants cancel out when calculating the $\Delta\Delta G_{app}$ with respect to wild type.

They compared the $\Delta\Delta G_{app}$ they obtained using TOXCAT with thermodynamic stabilities of the TMs that were previously studied by Fleming and coworkers (Doura and Fleming, 2004; Doura et al., 2004; Fleming and Engelman, 2001; Fleming et al., 1997) using SEAUC. They found a linear correlation with an $R = 0.74$ between the change in apparent free energy determined with TOXCAT, and the $\Delta\Delta G_{mut}$ obtained

with SEAUC (Duong et al., 2007). They also saw bigger discrepancies for certain mutants, suggesting that the association of the TMs in detergent vs TOXCAT is being affected differently. Their findings suggest TOXCAT to be giving homodimerization propensities that are thermodynamically driven, and that the main factors that affect relative dimerization between WT vs mutant (for many of the mutants), are present in both biological membranes and a detergent micelle.

1.6. GALLEX, a tool to measure dimer stabilities in a biological membrane

GALLEX is another genetic reporter assay that allows for the measurement of TM helix dimerization by fusing the TM helix of interest to the *E. coli* DNA binding domain protein LexA (Schneider and Engelman, 2003). Similar to TOXCAT (Russ and Engelman, 1999), the construct with the TM of interest fused to LexA, and the periplasmic maltose binding protein MalE, is expressed and inserted into the *E. coli* inner membrane. When dimerization of the TMD occurs, the transcription factor LexA homodimerizes and binds to the operator region on the DNA of *E. coli* resulting in the repression of the reporter gene lacZ and reduced expression of the reporter β -galactosidase (Finger et al., 2006). In the case of GALLEX, the expression of the construct with the TM of interest is controlled by a lac promoter, allowing for controlled expression of the chimera.

In 2006 Finger and coworkers published a study where they used the GALLEX system to obtain an apparent dissociation constant for GpA and several GpA mutants (Finger et al., 2006). They achieved this by varying expression levels of the chimeric protein with IPTG (which is used to induce expression through the lac promoter). They

then measured the activity of β -galactosidase at different IPTG concentrations (Finger et al., 2006). Expression of the chimera was analyzed via western blot, and they found that the protein concentration increased with IPTG concentration in the range 0.001 - 0.5 mM of IPTG (Finger et al., 2006). When comparing the activity of β -galactosidase with IPTG concentration they observed an inverse sigmoid with the linear region corresponding to the IPTG concentration range 0.001 - 0.1 mM (Finger et al., 2006). The fraction of monomeric protein was estimated from the β -galactosidase activity by assuming that the change in activity is linear and the maximal activity corresponds to 100% monomeric chimera. The fraction of monomer can then be related to the IPTG concentration. The concentration of IPTG at which the fraction of monomer is 0.5 was chosen as the apparent dissociation constant which can be used to obtain the apparent dimerization free energy (Finger et al., 2006). They obtained an apparent dimerization free energy for WT GpA of 7.51 kcal/mol, and a $\Delta\Delta G^{app}$ of the GpA mutants L75V, G79I, G83I, G79A, G83A, G79S, and G83S of 0.23 kcal/mol, 1.61 kcal/mol, 2.15 kcal/mol, 1.2 kcal/mol, 1.0 kcal/mol, 0.36 kcal/mol and 0.32 kcal/mol respectively (Finger et al., 2006).

The use of GALLEX for the study of TM helix association where the expression of the chimera is controlled by IPTG concentration allows for the determination of apparent dissociation constant (and relative apparent free energy) in the *E. coli* inner membrane. This tool adds to the study of how changes in TM sequence affects the relative stability of the TM domain in a more biologically relevant environment which could help scientist to better understand this process. Nonetheless, it does rely on some assumptions like linear proportionality of the signal and maximal activity of β -galactosidase

corresponding to 100% monomer, which may not be completely accurate. Additionally, with this approach absolute protein concentration could not be obtained, which is why the measured free energy is an apparent free energy instead of absolute. However, the ability to control and measure dimerization at a range of concentrations is a key advantage over TOXCAT.

1.7. Steric trapping

Scientists can do *in vitro* thermodynamic studies of TM helix association by changing the protein: detergent (or protein:lipid) mole fraction in order to drive the monomer-dimer equilibrium from one state to the other (Figure 1.1). However it can be challenging to drive the reaction from fully monomeric to fully dimeric for some proteins in certain environments. For example, GpA forms a very stable TM dimer in lipid bilayers, and diluting it enough in order to get a monomer-dimer equilibrium curve in lipid bilayers where the hydrophobic thickness of it better matches the proteins is challenging. In order to solve this problem such that thermodynamic stability of TMDs like GpA could be determined in that type of environment, Bowie and coworkers introduced a technique called steric trapping (Hong et al., 2010).

The steric trap technique utilizes the binding of monovalent streptavidin (mSA) to biotin to drive dissociation of the TM of interest. The TM of interest is fused to the protein Staphylococcal nuclease (SN) and a biotin acceptor peptide. The mSA has affinity for biotin, and can bind to one of the biotin tags. As mSA is a large protein, a second mSA cannot be accommodated on the opposite tag of the dimer unless the dimer dissociates into two monomers. When a second mSA binds to the biotin in the

other TM of the dimer, dissociation of the TM would occur, and the second binding event would reflect on the affinity of the TM domains for each other. This way, the dissociation of the TM domain can be controlled via mSA concentration or mSA affinity to biotin (which can be manipulated by doing mutations on mSA), rather than by dilution of the TM of interest. Detection of oligomerization is accomplished through a pyrene label, whose fluorescence is self-quenched in the dimer state.

Hong and coworkers published this method in 2010 where they determined the TM dimer stability of GpA wild type and some of its mutants in DM detergent, and in POPC lipid bilayers. They found the free energy of dissociation to be ~4-5 kcal/mol higher in POPC bilayers than in the DM detergent meaning that the GpA dimer is much more stable in POPC bilayers than in DM (Hong et al., 2010). When comparing the destabilization (with respect to wild type) that specific point mutations have in the stability of the dimer in the different environments, they observed I76A to be similar in DM and in POPC bilayers, but they found the destabilization of the dimer to be greater in POPC lipid bilayers than DM (and C₈E₅) (Fleming and Engelman, 2001) for G83I, V80A and T87A, which could be due to a more organized dimer in the POPC bilayer (Hong et al., 2010). The stability of the GpA dimer in POPC lipid bilayers is much higher than what was previously observed in plasma membrane derived vesicles (Chen et al., 2010). A possible reason for this is that the presence of other entities in the plasma membrane derived vesicles could be making the monomer state more stable (Hong et al., 2010).

Fortunately, steric trapping is amenable to a variety of membrane mimetics. In a later study Hong and Bowie examined the dimerization of the GpATM in lipid bilayers composed of *E. coli* phospholipid extracts. They found the dissociation constant of GpATM to be higher in the lipid bilayer composed of *E. coli* lipid extract than in the C_{16:0}C_{18:1c9}PC which means that the GpATM dimer is less stable in the *E. coli* lipid extract bilayer environment (Hong and Bowie, 2011). To decipher the reason for this destabilization Hong and Bowie looked at the GpATM stability in different lipid bilayer compositions and observed a destabilization of the dimer in the presence of negatively charged lipids (Hong and Bowie, 2011). They mentioned the presence of positively charged residues in the GpATM sequence in an area that could be at the interface of the membrane and the cytoplasm (Hong and Bowie, 2011). To test if electrostatics was playing a role in the destabilization of the dimer, they mutated four positively charged residues (RRKK) to QQNN (Hong and Bowie, 2011). The mutated GpATM resulted in the elimination of the destabilization that was previously observed for the wild type GpATM dimer in the presence of the negatively charged lipids (Hong and Bowie, 2011). In addition to looking at the effect of lipids in the stability of the GpATM dimer, they looked at the effect of inner membrane proteins and indeed found the presence of the inner membrane proteins to decrease the SNGpATM dimer stability (Hong and Bowie, 2011).

The introduction of the steric trapping technique open the doors to the study of highly stable dimers in lipid bilayers, expanding the dynamic range and the different environments and TMs that can be studied. As previously mentioned, the challenge with

these experiments is to be able to dilute the dimer such that the monomer state can be observed with sufficient signal. In the next session a technique that can also be used for the study of membrane protein oligomerization in a lipid bilayer at effectively “infinite” dilution will be discussed.

1.8. Use of single molecule photobleaching with total internal reflection fluorescence microscopy to measure oligomerization of transmembrane proteins

The techniques described above capture the bulk average of a population, which may ignore sample heterogeneity. In contrast, single molecule photobleaching with total internal reflection fluorescence microscopy (TIRF) uses the photobleaching event of a fluorophore attached to the protein of interest, to directly estimate the fraction of monomers and dimers in a sample. In this approach the protein of interest is purified, labeled with a fluorophore and reconstituted into lipid bilayers. The protein under study is then allowed to equilibrate in large multilamellar vesicles which are then extruded into liposomes creating a distribution of liposome sizes. At large dilutions, this extrusion step separates individual monomers and dimers into separate liposomes, freezing the equilibration. The liposomes are then bound to a treated glass slide which is then put in a TIRF microscope which uses the evanescent field to image very close to the glass slide. The number of photobleaching events in each region of interest can be quantified, and under ideal conditions (including complete labeling and perfect separation), the fraction of monomeric and dimeric species can be directly measured. In reality, fluorophore labeling is often incomplete, and larger liposomes or a higher protein to lipid ratio increases the chance of randomly capturing multiple proteins that were not interacting, giving the false impression of a larger subunit. In order to properly quantify

the monomer dimer population at the time of extrusion, different parameters need to be taken under consideration like the labeling efficiency, non-specific labeling and expected photobleaching probabilities for the system at different mole fractions which also depend on the liposome size distribution. Using a Poisson distribution to simulate the co-encapsulation of entities into liposomes of a known size distribution thermodynamic information for the monomer-dimer equilibrium can be obtained. The high sensitivity of this approach allows for the study of membrane protein dimerization with photobleaching probability distributions that can distinguish monomer dimer state at the point of MLVs extrusion in a dynamic range covering over five orders of magnitude ($\chi = 7.5 \times 10^{-10}$ to 7.5×10^{-5}), even going beyond maximal biological dilutions (Chadda et al., 2016).

In 2016 Chadda and coworkers published a study in which they used single molecule photobleaching analysis with TIRF microscopy to determine the dimerization free energy of a modified CIC Cl⁻/H⁺ antiporter and two additional mutants in lipid bilayers composed of a 2:1/POPE:POPG mixture (Chadda et al., 2016). The CIC Cl⁻/H⁺ antiporter was modified by mutating C85A/H234C to allow for labeling with Cy5-maleimide (Chadda et al., 2016). In addition they studied two monomerizing mutants (I422W and I201W/I422W) that they previously showed results in a functionally folded form of the protein (Robertson et al., 2010). They found the equilibrium constant for the modified CIC Cl⁻/H⁺ antiporter to be $\sim 10^8$ lipids/subunit yielding a dimerization free energy of -11.4 ± 0.1 kcal/mol, for the I422W mutant to be $\sim 10^6$ with a dimerization free

energy of -9.0 ± 0.3 kcal/mol, and for the I201W/I422W of $\sim 10^5$ resulting in a dimerization free energy of -7.3 ± 0.2 kcal/mol (Chadda et al., 2016).

This study heavily relied on previous analyses of extruded liposome size distributions to estimate the fraction of co-captured particles. However, the size distribution can dramatically change in different liposome compositions. In a later study, the same group published a work where they proposed a model free method to measure the dimerization free energy of CLC-ec1 in lipid bilayers. In this study they looked at the dependency that the fraction of dimer and the free energy have on the modeling by comparing the results with different liposome size distributions. The problem is that when the protein is highly diluted in the lipid bilayer the probability of encapsulating more than one state in a liposome is low, but as the protein concentration in the lipid bilayer increases this probability increases and it is important to take its contribution into account. In addition, the co-encapsulation of subunits is also dependent on the liposome size distribution and previously they observed that the data and model systematically deviated from each other at mole fractions greater than 10^{-6} , where they were seeing in the data more photobleaching events than what the model was predicting. The size distribution that they used for the model corresponded to a lipid composition that was a little different to the one they used. When they measured the liposome size distribution for the lipid composition they used, this systematic error was corrected and the WW mutant was found to be monomeric giving them a monomeric control (Chadda et al., 2018). For a dimeric control they used cross-linked R230C/L249C CLC-ec1 (Chadda et al., 2018). They found that the use of empirical

controls agrees with the results obtained when using a model as long as the liposome size distribution is correct (Chadda et al., 2018).

In a later study Chadda and coworkers used this technique to look at the effect that changing lipid composition has on the dimerization equilibrium of the CLC-ec1 protein (Chadda et al., 2021). They combined this experimental technique with coarse grained MD simulations to get a better understanding at the molecular level of how the lipids interact with the protein in the monomeric and dimeric state. From the simulations they saw that in the 2:1 POPE:POPG lipid composition there was a lipid defect near the protein in the monomer state where the lipid membrane get thinner and the bilayer structure get twisted (Chadda et al., 2021). The defects in the bilayer that occur to solvate the monomer state lead the membrane to adopt a high-energy conformation which gets eliminated when the protein dimerizes, behaving like an attractive force for dimerization (Chadda et al., 2021). They used the single molecule photobleaching technique to see if addition of shorter chain lipids to the lipid composition of the bilayer would lead to a more stable monomer. Their results show a significant destabilization of the dimer (in 20% DL 2:1 PE/PG) shifting the equilibrium towards the monomer state (Chadda et al., 2021). They estimated a lower limit of the $K_{D,Dimer} > 4.2 \times 10^6$ subunits/lipid (Chadda et al., 2021). The availability of the experimental approach discussed in this section allow the researchers to study TM dimerization in different lipid bilayer compositions with a dynamic range that expand several orders of magnitude where very stable dimers would fall.

As was mentioned in this section it is possible to also used MD simulations to study membrane proteins in more detail as Chadda and coworkers presented in their 2021 publication of CIC-ec1 dimerization in different lipid membrane compositions (Chadda et al., 2021). In the next section different MD simulation studies on transmembrane protein association will be discussed giving emphasis to free energy methods.

1.9. Free energy calculation of membrane helix association using Molecular Dynamics simulation techniques

When experimentally studying the association thermodynamics of TM helices, scientists can measure the thermodynamic stability of the system in different environments and on variants of the protein to better understand the forces involved. However, a molecular and atomistic picture of the process can not be clearly gained with the described experiments. One thing that can be done to get that information is to use MD simulations. While MD simulations are a powerful technique, there are many things to keep in mind when using and analyzing them for scientific insight. In this section different studies that have used MD simulations to look at the association of TM helices will be discussed. The focus is going to be on approaches that give the potential of mean force and research that works on getting a better understanding of the system. In addition work that has been done to improve the computational approach to better match the experimental results or improve the technique will be discussed.

When doing MD simulations, due to rough energy landscapes, one of the challenges is achieving enough sampling. What this means is that the system might have a hard time covering the different configurations that correspond to a state, and going from one state (example, monomer), to another state (for example, dimer). The more detailed a

simulation is, the more computationally expensive it is to run. In the case of simulations of TM helices, different force fields with different description levels exist. For example, simulations can be done using an all-atom force field like CHARMM36 (Klauda et al., 2010), but if this level of detail is not required, a coarse-grained force field like MARTINI (Marrink et al., 2004, 2007), can reduce the number of particles in the simulation and achieve higher sampling for the same amount of time. Another way to decrease computational expense is by using an implicit membrane model (like IMM1 (Lazaridis, 2003)) to represent the lipid bilayer, instead of modeling each separate lipid molecule. The choice of force field and model to use requires careful consideration of the questions that are being asked and the resources available. In the different studies that are going to be discussed in this session different models and force fields were used to study association of TM helices.

Another important consideration is that normal MD simulations can suffer from incomplete sampling in the time scale of the simulations. This could be due to high energy barriers in the free energy landscape, resulting in the system getting trapped in a local energy minimum. One way to deal with that problem is by using enhanced sampling methods such as umbrella sampling to help the system escape the energy barrier (Mori et al., 2016). Umbrella sampling is a free energy computational technique that applies a bias potential along a chosen reaction coordinate to help the system move from one thermodynamic state to another (Kästner, 2011). To obtain the potential of mean force, force integration or weighted histogram analysis method can be used (Kästner, 2011).

In 2006 Zhang and Lazaridis published a study where they presented a theoretical framework to calculate the free energy of association of TM helices in lipid bilayers and in detergent based on MD simulations performed using the implicit solvation system IMM1 (Zhang and Lazaridis, 2006). In this study they present the considerations for translational and rotational entropy in both environments and decomposed the standard association free energy into the effective energy, translational, rotational and conformational energy (Zhang and Lazaridis, 2006). They did the calculations for the extensively studied dimer GpA by running 1 ns MD simulations of the dimer and monomer in an implicit membrane, finding overall good agreement between their results and the available experimental data at the time (Zhang and Lazaridis, 2006). Importantly their framework not only correctly identified that GpA dimerization is stronger in bilayers than micelles, they also found the rotational entropy cost to be smaller in bilayers than in micelles, while the translational entropy cost is larger in bilayers than in micelles (Zhang and Lazaridis, 2006). This study helped in better understanding the contribution that different interactions have to the association of TM helices in different environments.

In another study published in 2012, Sengupta and Marrink used the MARTINI force field to perform coarse-grained MD simulations of GpA and two of its mutants (T87F and G79LG83LG86L) in a DPPC lipid membrane (Sengupta and Marrink, 2010). They did simulations where they started with the GpA monomers (WT sequence) separated and saw the formation of the dimer through the simulation trajectory, where the dimer structure seemed to agree with the NMR structure (Sengupta and Marrink, 2010). They

also simulated the self-assembly of the mutants mentioned above and saw the formation of dimers for both of the mutants but with varied structures and larger interhelical distance in comparison to WT (Sengupta and Marrink, 2010). In addition to the self-assembly simulations, they used umbrella sampling to obtain the potential of mean force for the different sequences. For the GpA WT PMF they observed two minimas at close distances one at 0.75 nm and another less deep minima at 0.9 nm (Sengupta and Marrink, 2010). In the GpA WT PMF profile they also saw a third deep minimum at 1.2 nm which corresponds to the regime where the TMs are just before lipid separation (Sengupta and Marrink, 2010). They converted their association free energy to a “detergent-like” standard state and compared the results with previous experimental results. They obtained an association free energy of -29.2 kJ/mol, which is comparable to the experimental association free energy estimate obtained in C₈E₅ of -30 kJ/mol (Fleming, 2002) and in C12 maltoside of -32 kJ/mol (Fisher et al., 2003). When comparing the GpA WT PMF to the mutants, they saw the mutants to have a PMF with a dimerization profile similar to the GpA WT but, the mutants had a destabilized dimer (Sengupta and Marrink, 2010). To get a deeper understanding of what could be the reason for this destabilization, they looked at different interaction energies like lipid-lipid, helix-helix and helix-lipid interactions. There they saw that the packing of the helices and the packing of the lipids around the helices is affected for the T87F mutant when compared to WT, contributing to the association decrease that is observed in the mutants (Sengupta and Marrink, 2010). This study gave insight into the TM association

process at the molecular level to see how specific interactions could be contributing to the process.

A different study done on gramicidin A (gA) was able to look at the effect of lipids in the association of the gA protein dimers. In this study Yoo and Cui looked at the association of gramicidin A dimers using coarse grained MD simulations in different lipid bilayers with the same headgroups but varying acyl tail lengths (DMPC, DPPC, and DSPC) (Yoo and Cui, 2013). To obtain the potential of mean force they used umbrella sampling where the reaction coordinate was the distance between the two gramicidin dimers (Yoo and Cui, 2013). They used WHAM and force integration to get the PMF, where the force integration allowed for the decomposition of the PMF into the contribution of different interactions that are related to the reaction coordinate (Yoo and Cui, 2013). Their results suggest that the association of the gA dimer is dependent on the membrane thickness or hydrophobic mismatch (Yoo and Cui, 2013). Being able to decompose the PMF into different contributions allows for a better understanding of the association process.

A question related to TM helix association is how are TM helices able to find each other before TM-TM interactions start to be relevant? Do helices must physically interact in order to begin dimerization, or are defects in the lipid bilayer somehow helping the TM helices to associate? In the previous section on the used of single molecule photobleaching analysis with TIRF microscopy it was discussed how Chadda and coworkers used a combination of MD simulations with single molecule photobleaching analysis to get a better understanding on the effect that changing lipid composition has

on the dimerization equilibrium of the CLC-ec1 protein. In the MD simulations they observed lipid defects around CLC-ec1 monomers that were absent in the dimer (Chadda et al., 2021). They combined these simulations with experiments where they added small quantities of short-chain lipids to their experiments, and found that they stabilized the monomer state of the protein, presumably by preferentially solvating this defect. In a study done by Kuznetsov and coworkers, all-atom MD simulations were used to get an atomistic picture of how the membrane promotes dimerization of transmembrane peptides (Kuznetsov et al., 2015). They tested three peptides, GpA TM, a poly-Ala and a poly-Leu peptide as monomers and dimers in a lipid bilayer composed of POPC lipids (Kuznetsov et al., 2015). To obtain the free energy profiles they used umbrella sampling with the interhelical distance between the center of mass as a reaction coordinate (Kuznetsov et al., 2015). Then the PMF profile was obtained by integrating the mean force (with respect to the distance) (Kuznetsov et al., 2015). The PMF profiles showed that GpA had the highest dimerization strength, with poly-Leu and poly-Ala both being relatively weak (Kuznetsov et al., 2015). However poly-Ala different to poly-Leu seemed to form dimers that are tightly packed but weak, while poly-Leu dimers seemed to be loose (Kuznetsov et al., 2015). In their results they observed that lipid perturbations seemed to be present to distances of about 20 Å from the helix axis with a pair of helices creating lipid perturbations of about 40 Å (Kuznetsov et al., 2015). It is possible that these perturbations guide TM domains towards one another even before they directly interact.

Umbrella sampling molecular dynamics (USMD) can be used to obtain the PMF for the association of TM helices where a series of windows that correspond to different coordinates along a chosen reaction path are generated, and a biased potential is imposed such that the system explores configurations around a point in space (Kästner, 2011). However, USMD can still suffer from incomplete sampling, especially at short TM interhelical distances where interactions between interfacial residues is likely to be strong. The strong interactions create an energetic barrier that is difficult to overcome on short time scales. One way to alleviate this issue is by using window-exchange USMD (WEUSMD), where the windows along the reaction coordinate are exchanged, which helps the system explore more conformational space. In a study published in 2012 by Park, Kim and Im, it was found that WEUSMD is more efficient than USMD for the study of TM helix association (Park et al., 2012). In the study by Park and coworkers (Park et al., 2012), they used WEUSMD to obtained the PMF for a pVNVV TM helix in an implicit membrane using the IMM1 model with a hydrophobic thickness of 23 Å. They found that the calculated association free energy using WEUSMD and USMD were very close to each other with an association free energy of -5.83 kcal/mol and -5.89 kcal/mol respectively (Park et al., 2012). However, the results obtained for USMD were more variable due to relatively worse sampling at short and intermediate interhelical distances (helix-helix distances less than ~14 Å) (Park et al., 2012).

Scientist have used enhanced sampling methods to try to address challenges related to incomplete sampling, but the problem of incomplete sampling for TM helix oligomerization can still be found. As discussed Park and coworkers previously showed

that WEUSMD has more efficient sampling than USMD at shorter interhelical distances for a specific TM (pVNVV) (Park et al., 2012). In another work Park and Im looked at the sampling power of WEUSMD on the association of the TM domain of the well known GpA (Park and Im, 2013). One of the reasons GpA is a good candidate for testing the sampling power of WEUSMD is the fact that the dimer has a very tightly packed interface, which could make it challenging for sampling conformational space at short interhelical distances (Park and Im, 2013). They used 1D WEUSMD with the distance between the helices as the collective variable, and 2D-WEUSMD with the distance and angle between the helices as the collective variables (Park and Im, 2013). They did the simulations in an implicit membrane using the IMM1 model with a hydrophobic thickness of 23 Å (Park and Im, 2013). They did 1D-WEUSMD starting from two different configurations, and found discrepancies between the PMFs at $r_{\text{HH}} < 8 \text{ \AA}$, suggesting sampling problems at short distances (Park and Im, 2013). They moved to 2D-WEUSMD and obtained better sampling than with 1D-WEUSMD, where they were able to obtain the minima at shorter distances, and observed a path along the angle (as the distance between the helices decreases), that can help bypass the barrier (Park and Im, 2013). Their results show that 1D-WEUSMD with the distance between the helices as a reaction coordinate suffers from incomplete sampling at shorter distances due to barriers along the crossing angle (Park and Im, 2013). The addition of a second reaction coordinate along the crossing angle helped to bypass the barrier and achieve better sampling (Park and Im, 2013). This makes 2D-WEUSMD (in comparison to 1D-WEUSMD) a better approach for determining PMF profiles for TM association of helices

that have tight packing (short interhelical distance) and suffer from hidden barriers along the crossing angle.

As can be seen different collective variables can be used when using umbrella sampling to obtain the potential of mean force. The chosen collective variable or variables is going to depend on the system and process under study. In 2017 a work by Domański and coworkers was published where they calculated the PMF for GpA dimerization, using an alternate choice of collective variable than interhelical distance or crossing angle (Domański et al., 2017). In this study they performed coarse-grained MD simulations of GpA dimerization in POPC lipids using the MARTINI force field (Domański et al., 2017). To determine the PMF they used replica exchange umbrella sampling (REUS) (Domański et al., 2017). The used collective variable for the GpA system was the distance root-mean-square displacement between two configurations with a cut off distance between 0.1 and 0.6 nm and only considering interhelical interaction for the TM region (Glu-72 to Ile-92) (Domański et al., 2017). They performed simulations with different starting configurations: ranging from unbound to bound and one with the windows spaced uniformly along the collective variable (Domański et al., 2017). They observed the PMF to be independent of initial configuration and to converge at $\sim 6 \mu\text{s}$ (Domański et al., 2017). However, it is important to note that the dimer state that the simulations converged on was inconsistent with available experimental data. Nonetheless, in this study they saw how convergence is achieved faster when using replica exchange which agrees with the results from Parks, Kim and Im, where they saw WEUSMD to have better sampling than USMD (Park et al., 2012);

and how the use of different initial conditions served as a test for convergence (Domański et al., 2017). The use of the D_{RMS} collective variable for the GpA system revealed a more complex PMF in comparison to previously obtained PMFs (Domański et al., 2017). In a later study Domański and coworkers determined the free energy landscape for the association of GpA using all-atom MD simulation with the CHARMM36 force field instead of the Martini force field (Domański et al., 2017, 2018). In this study they used umbrella sampling with the reaction coordinate defined as the D_{rms} and WHAM to unbias the umbrella sampling trajectories (Domański et al., 2018). While there were some problems with convergence when starting from different configurations, they mentioned the PMFs to have consistent features (Domański et al., 2018). Surprisingly in their results the monomer state of GpA was found to be more stable than the dimer state (Domański et al., 2018). They performed the simulations with other force fields (CHARMM36m (Huang et al., 2017) and AMBER Slipids model (Jämbeck and Lyubartsev, 2012b, 2012a)) and said that they observed the dissociation of the dimer into the monomer state which is very unexpected for this system as GpA have been observed to dimerize with high affinity (Domański et al., 2018). Because of the discrepancy with experiments they decided to do some minor adjustments to the protein-lipid non-bonded interactions which they hypothesized is where the mismatch might be (Domański et al., 2018). With the correction they observed better agreement with the experimental data (Domański et al., 2018). In their results the CHARMM36 force field identified the GpA native state as the more stable dimeric state, but rate the monomeric GpA configuration as the most energetically stable (Domański et al., 2018).

In another study Domański and coworkers studied the dimerization of the GpA-TM residues 69-97 in a POPC membrane using all-atom MD simulations with the CHARMM36 force field (Domański et al., 2020). Based on previous studies they added an adjustment of the protein-lipid interaction to have the dimer state more stable (Domański et al., 2018). They used REUS simulations to obtain the energetic information on the dimerization with WHAM to unbias it (Domański et al., 2020). They did two different initial conditions: one where the helices were separated, and another where they were close together. Interhelical D_{rms} was chosen as the collective variable (Domański et al., 2020). They observed a hidden barrier and proposed a new collective variable that is a combination of the D_{rms} and a crossing angle (Domański et al., 2020). In their results they observed two associated states, a non-native intermediate and the native dimer state where the two are separated by a free energy barrier (Domański et al., 2020).

As mentioned one of the challenges when studying TM helix association in a lipid bilayer is incomplete sampling which could lead to a calculated dimerization free energy that would be biased towards a specific state (native dimer state for example). This makes the choice of collective variable very important. In another study, Majumder and coworkers looked at the binding free energy of GpA using MD simulations at the coarse grained level (Martini v3) with enhanced sampling methods (umbrella sampling and REUS) (Majumder et al., 2022). In this study they looked at the binding energy of GpA using different collective variables (Majumder et al., 2022). They did one-dimensional US using the distance between the TM protein helices center of mass (COM) (D_{com}), and

D_{rmsd} which have been explored in previous studies, and they also did a two dimensional umbrella sampling using the xy-projection of the COM-COM distance where they saw better sampling (Majumder et al., 2022). They found that using the COM or RMSD suffers from incomplete sampling which leads to the computation of erroneous binding constant (Majumder et al., 2022). The challenges might be related to the fact that the GpA system seems to have multiple minima in the PMF that seem to contribute to the binding equilibria (Majumder et al., 2022).

In this section different free energy MD simulation studies done on the association of TM helices were discussed. Several considerations need to be taken into account when using this approach. First, the force field or model that will be used for the simulations is important both in terms of resources required and the scientific question being addressed. Second, it is important to pay attention to the sampling, especially when obtaining the PMF, such that the landscape sufficiently explores the configurational space. Third, when using the umbrella sampling technique it is important to choose collective variables that will help the system explore the different configurations that the system could take during the reaction. Another important point is to compare the MD simulation results with available experimental observations. If the simulations do not match what happens in the experiments the results would not be reliable. One of the main advantages of using MD simulations to study TM helix association is the molecular and atomistic level understanding that the approach give about the process under study. MD simulations have helped scientists better understand the different interactions that take place in the association of TM helices and interpret the experimental observations.

1.10. Summary and concluding remarks: Experimental and molecular dynamics approaches to study thermodynamics of transmembrane protein oligomerization

Membrane proteins are involved in many biological processes that are essential for the successful functioning of the cell. Membrane protein functions include transport into and out of the cell, signaling processes, cell to cell communication, and cell proliferation (Li et al., 2021). Incorrect functioning of these systems could lead to illnesses such as cancer and autoimmune diseases (Du and Lovly, 2018; Marinko et al., 2019). One of the important steps for the correct functioning of membrane proteins is successful folding of the protein. As previously mentioned membrane protein folding could include the oligomerization of independent helices in the membrane environment (Popot and Engelman, 1990). Oligomerization of membrane helices is especially important for single pass membrane proteins like integrins (Li et al., 2005b; Yin et al., 2006), and receptor tyrosine kinases (Bocharov et al., 2008, 2012; Mineev et al., 2010). A good understanding of the physical interactions that take place in the oligomerization of TM helices could give insight into their structure, how certain biological processes are regulated and their function. A way to obtain that understanding is by using experimental and computational approaches to determine the thermodynamic stability of these oligomers, and how changes to the environment and the protein sequence affect that stability.

In this chapter different experimental approaches that have been used for the thermodynamic study of TM protein association were discussed. In addition a discussion on the used of MD simulations to study TM helix association was presented. The experimental approaches discussed have advanced our understanding of the

association of TM helices by doing experiments in different environments including detergents, lipid bilayers with different lipid compositions, plasma membrane derived vesicles and the *E. coli* inner membrane. The ability to perform these experiments in different environments have taught us that changing the environment can have a big impact on the absolute stability of the associated state. By doing the experiments on different mutants of the TMDs, scientist have been able to identify which face of the helix might be involved in the interactions with the other protein, and which specific amino acids seem to contribute more to the associated state. In addition different experimental approaches allow for the study of proteins with different stability by expanding the dynamic range that can be covered.

The development of all these techniques have been essential in obtaining a better understanding of this process, but what about what is happening at the molecular level? What is specifically happening to the system when the amino acid sequence or environment is being changed? To obtain information about this process at the molecular and atomic level scientist can use MD simulations. One of the main advantages of MD simulations is that it can help to visualize what is actually causing the changes in stability and even what is the relative contribution that different interactions have to the process. Although MD simulations can be key in getting a better understanding of the process, it is important to be able to compare results from simulations with experimental observations. If the MD simulations results do not match experimental observations the simulations could be misleading, but the combination of MD simulations with experimental approaches can be extremely powerful. The

experiments can give a base to see if the MD simulation results are reliable, and the MD simulations can help guide the experiments and increase the understanding on what is specifically going on in the process under study.

1.11. Outline of thesis Chapters 2 – 4: Studying transmembrane helix association, structure and stability using computational and experimental techniques

In the following chapters we combined the use of computational and experimental techniques to get a better understanding of the biophysical properties of specific TM helices. In Chapter 2 a thermodynamic study on the association of TM sequences that form GAS_{right} dimers is presented. In this study the thermodynamic stability of a group of predicted GAS_{right} dimers was determined in detergent using FRET. The group of predicted GAS_{right} dimers that we investigate covered a wide range of thermodynamic stabilities. One of the challenges of this type of study is to be able to experimentally obtain the transition from monomer state to dimer state such that the data can be fitted and the equilibrium constant extracted. Being able to obtain data for TM helices that dimerized with different affinities, but are expected to have similar geometries, allowed us to overcome this challenge. The results were compared with previously obtained TOXCAT homodimerization propensities finding a striking agreement between the two. Using this relationship, we were able to predict the ΔG° in the detergent used for a larger set of TMs by using their TOXCAT homodimerization propensities. Our findings suggest that TOXCAT is a thermodynamically driven process and that the main drivers of relative stability in both environments for the sequences we used are similar. In

addition, the results support the previous suggestion that vdW and hydrogen bond interactions seem to be key players in the stability of GAS_{right} dimers.

In the next two chapters, we move to a group of single pass TM proteins that are part of the divisome. The divisome is a multiprotein complex involved in the division of bacteria. In our studies we focused on the divisome proteins FtsL and FtsB, which form a complex. In Chapter 3 a structural analysis of the divisome FtsLB subcomplex is presented. The experimental structure of the complex has not been solved. In this chapter a computational model for their possible structure is proposed. The model is supported by coevolutionary restraints and experimental data, which suggested FtsL forms a continuous helix spanning the TM and coiled coil domains, while these domains in FtsB are separated by a flexible linker. MD simulations of the model were performed to look at the dynamics and stability of the complex. The results point to FtsLB forming an L₂:B₂ heterotetramer. Although the FtsLB heterotetramer was modeled as a coiled-coil bundle from the TM region up to the periplasmic region, in the MD simulations water invaded the lower coiled-coil of the periplasmic region where there was a presence of interfacial polar residues. This suggested the possibility of this region opening into two LB dimers and the TM region forming a stable four-helix bundle (model presented in Chapter 4).

In Chapter 4, we focus on the polar cluster of amino acids that led to water infiltrating the model of the FtsLB complex during the MD simulations. In this study we used a combination of *in vivo* and *in vitro* experiments, along with computational modeling and MD simulations to get a better understanding on how amino acid changes in this region

affect the structure of the complex and the division of *E. coli*. The findings showed that mutation of these polar residues to “canonical” hydrophobic residues increases the thermal stability of the complex but causes cell division defects. In addition, a new tetrameric model for the FtsLB subcomplex is proposed based on our previous findings. In this new model the periplasmic region divides into two LB coiled coils instead of the four helices staying together. MD simulations were performed in the new model and suggest a more stable coiled coil region for this model compared with the previous one. The results suggest the stability of the complex needs to be properly tuned for correct cell division. We proposed the Y-model to be a more likely structure than the I-model based on our observations although further experimentation is needed to validate the model.

1.12. References

- Adair, B.D., and Engelman, D.M. (1994). Glycophorin A Helical Transmembrane Domains Dimerize in Phospholipid Bilayers: A Resonance Energy Transfer Study. *Biochemistry* 33, 5539–5544. <https://doi.org/10.1021/bi00184a024>.
- Anbazhagan, V., and Schneider, D. (2010). The membrane environment modulates self-association of the human GpA TM domain—Implications for membrane protein folding and transmembrane signaling. *Biochim. Biophys. Acta BBA - Biomembr.* 1798, 1899–1907. <https://doi.org/10.1016/j.bbamem.2010.06.027>.
- Bocharov, E.V., Mineev, K.S., Volynsky, P.E., Ermolyuk, Y.S., Tkach, E.N., Sobol, A.G., Chupin, V.V., Kirpichnikov, M.P., Efremov, R.G., and Arseniev, A.S. (2008). Spatial Structure of the Dimeric Transmembrane Domain of the Growth Factor Receptor ErbB2 Presumably Corresponding to the Receptor Active State. *J. Biol. Chem.* 283, 6950–6956. <https://doi.org/10.1074/jbc.M709202200>.
- Bocharov, E.V., Mineev, K.S., Goncharuk, M.V., and Arseniev, A.S. (2012). Structural and thermodynamic insight into the process of “weak” dimerization of the ErbB4 transmembrane domain by solution NMR. *Biochim. Biophys. Acta BBA - Biomembr.* 1818, 2158–2170. <https://doi.org/10.1016/j.bbamem.2012.05.001>.

Burgess, N.K., Stanley, A.M., and Fleming, K.G. (2008). Determination of Membrane Protein Molecular Weights and Association Equilibrium Constants Using Sedimentation Equilibrium and Sedimentation Velocity. In Methods in Cell Biology, (Elsevier), pp. 181–211.

Chadda, R., Krishnamani, V., Mersch, K., Wong, J., Brimberry, M., Chadda, A., Kolmakova-Partensky, L., Friedman, L.J., Gelles, J., and Robertson, J.L. (2016). The dimerization equilibrium of a CIC Cl⁻/H⁺ antiporter in lipid bilayers. *eLife* 5, e17438. <https://doi.org/10.7554/eLife.17438>.

Chadda, R., Cliff, L., Brimberry, M., and Robertson, J.L. (2018). A model-free method for measuring dimerization free energies of CLC-ec1 in lipid bilayers. *J. Gen. Physiol.* 150, 355–365. <https://doi.org/10.1085/jgp.201711893>.

Chadda, R., Bernhardt, N., Kelley, E.G., Teixeira, S.C., Griffith, K., Gil-Ley, A., Öztürk, T.N., Hughes, L.E., Forsythe, A., Krishnamani, V., et al. (2021). Membrane transporter dimerization driven by differential lipid solvation energetics of dissociated and associated states. *eLife* 10, e63288. <https://doi.org/10.7554/eLife.63288>.

Chen, L., Novicky, L., Merzlyakov, M., Hristov, T., and Hristova, K. (2010). Measuring the Energetics of Membrane Protein Dimerization in Mammalian Membranes. *J. Am. Chem. Soc.* 132, 3628–3635. <https://doi.org/10.1021/ja910692u>.

Domański, J., Hedger, G., Best, R.B., Stansfeld, P.J., and Sansom, M.S.P. (2017). Convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association. *J. Phys. Chem. B* 121, 3364–3375. <https://doi.org/10.1021/acs.jpcb.6b08445>.

Domański, J., Sansom, M.S.P., Stansfeld, P.J., and Best, R.B. (2018). Balancing Force Field Protein–Lipid Interactions To Capture Transmembrane Helix–Helix Association. *J. Chem. Theory Comput.* 14, 1706–1715. <https://doi.org/10.1021/acs.jctc.7b00983>.

Domański, J., Sansom, M.S.P., Stansfeld, P.J., and Best, R.B. (2020). Atomistic mechanism of transmembrane helix association. *PLOS Comput. Biol.* 16, e1007919. <https://doi.org/10.1371/journal.pcbi.1007919>.

Doura, A.K., and Fleming, K.G. (2004). Complex Interactions at the Helix–Helix Interface Stabilize the Glycophorin A Transmembrane Dimer. *J. Mol. Biol.* 343, 1487–1497. <https://doi.org/10.1016/j.jmb.2004.09.011>.

Doura, A.K., Kobus, F.J., Dubrovsky, L., Hibbard, E., and Fleming, K.G. (2004). Sequence Context Modulates the Stability of a GxxxG-mediated Transmembrane Helix–Helix Dimer. *J. Mol. Biol.* 341, 991–998. <https://doi.org/10.1016/j.jmb.2004.06.042>.

Du, Z., and Lovly, C.M. (2018). Mechanisms of receptor tyrosine kinase activation in cancer. *Mol. Cancer* 17, 58. <https://doi.org/10.1186/s12943-018-0782-4>.

- Duneau, J.-P., Vegh, A.P., and Sturgis, J.N. (2007). A Dimerization Hierarchy in the Transmembrane Domains of the HER Receptor Family. *Biochemistry* 46, 2010–2019. <https://doi.org/10.1021/bi061436f>.
- Duong, M.T., Jaszewski, T.M., Fleming, K.G., and MacKenzie, K.R. (2007). Changes in Apparent Free Energy of Helix–Helix Dimerization in a Biological Membrane Due to Point Mutations. *J. Mol. Biol.* 371, 422–434. <https://doi.org/10.1016/j.jmb.2007.05.026>.
- Finger, C., Volkmer, T., Prodöhl, A., Otzen, D.E., Engelman, D.M., and Schneider, D. (2006). The Stability of Transmembrane Helix Interactions Measured in a Biological Membrane. *J. Mol. Biol.* 358, 1221–1228. <https://doi.org/10.1016/j.jmb.2006.02.065>.
- Fisher, L.E., Engelman, D.M., and Sturgis, J.N. (1999). Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain 1 Edited by G. von Heijne. *J. Mol. Biol.* 293, 639–651. <https://doi.org/10.1006/jmbi.1999.3126>.
- Fisher, L.E., Engelman, D.M., and Sturgis, J.N. (2003). Effect of Detergents on the Association of the Glycophorin A Transmembrane Helix. *Biophys. J.* 85, 3097–3105. [https://doi.org/10.1016/S0006-3495\(03\)74728-6](https://doi.org/10.1016/S0006-3495(03)74728-6).
- Fleming, K.G. (2002). Standardizing the Free Energy Change of Transmembrane Helix–Helix Interactions. *J. Mol. Biol.* 323, 563–571. [https://doi.org/10.1016/S0022-2836\(02\)00920-8](https://doi.org/10.1016/S0022-2836(02)00920-8).
- Fleming, K.G., and Engelman, D.M. (2001). Specificity in transmembrane helix-helix interactions can define a hierarchy of stability for sequence variants. *Proc. Natl. Acad. Sci.* 98, 14340–14344. <https://doi.org/10.1073/pnas.251367498>.
- Fleming, K.G., Ackerman, A.L., and Engelman, D.M. (1997). The effect of point mutations on the free energy of transmembrane α -helix dimerization. *J. Mol. Biol.* 272, 266–275. <https://doi.org/10.1006/jmbi.1997.1236>.
- Fleming, K.G., Ren, C.-C., Doura, A.K., Eisley, M.E., Kobus, F.J., and Stanley, A.M. (2004). Thermodynamics of glycophorin A transmembrane helix dimerization in C14 betaine micelles. *Biophys. Chem.* 108, 43–49. <https://doi.org/10.1016/j.bpc.2003.10.008>.
- Hong, H., and Bowie, J.U. (2011). Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *J. Am. Chem. Soc.* 133, 11389–11398. <https://doi.org/10.1021/ja204524c>.
- Hong, H., Blois, T.M., Cao, Z., and Bowie, J.U. (2010). Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proc. Natl. Acad. Sci. U. S. A.* 107, 19802–19807. <https://doi.org/10.1073/pnas.1010348107>.

- Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., and MacKerell, A.D. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* *14*, 71–73.
<https://doi.org/10.1038/nmeth.4067>.
- Jämbeck, J.P.M., and Lyubartsev, A.P. (2012a). Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids. *J. Phys. Chem. B* *116*, 3164–3179. <https://doi.org/10.1021/jp212503e>.
- Jämbeck, J.P.M., and Lyubartsev, A.P. (2012b). An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *J. Chem. Theory Comput.* *8*, 2938–2948. <https://doi.org/10.1021/ct300342n>.
- Kästner, J. (2011). Umbrella sampling: Umbrella sampling. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* *1*, 932–942. <https://doi.org/10.1002/wcms.66>.
- Khadria, A.S., and Senes, A. (2015). Fluorophores, environments, and quantification techniques in the analysis of transmembrane helix interaction using FRET: Analysis Of Transmembrane Helix Interaction Using FRET. *Biopolymers* *104*, 247–264.
<https://doi.org/10.1002/bip.22667>.
- Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A.D., and Pastor, R.W. (2010). Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* *114*, 7830–7843. <https://doi.org/10.1021/jp101759q>.
- Kuznetsov, A.S., Polyansky, A.A., Fleck, M., Volynsky, P.E., and Efremov, R.G. (2015). Adaptable Lipid Matrix Promotes Protein–Protein Association in Membranes. *J. Chem. Theory Comput.* *11*, 4415–4426. <https://doi.org/10.1021/acs.jctc.5b00206>.
- Lazaridis, T. (2003). Effective energy function for proteins in lipid membranes. *Proteins Struct. Funct. Genet.* *52*, 176–192. <https://doi.org/10.1002/prot.10410>.
- Lemmon, M.A., Flanagan, J.M., Hunt, J.F., Adair, B.D., Bormann, B.J., Dempsey, C.E., and Engelman, D.M. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* *267*, 7683–7689.
[https://doi.org/10.1016/S0021-9258\(18\)42569-0](https://doi.org/10.1016/S0021-9258(18)42569-0).
- Li, E., You, M., and Hristova, K. (2005a). Sodium Dodecyl Sulfate–Polyacrylamide Gel Electrophoresis and Förster Resonance Energy Transfer Suggest Weak Interactions between Fibroblast Growth Factor Receptor 3 (FGFR3) Transmembrane Domains in the Absence of Extracellular Domains and Ligands. *Biochemistry* *44*, 352–360.
<https://doi.org/10.1021/bi048480k>.

- Li, E., Placone, J., Merzlyakov, M., and Hristova, K. (2008). Quantitative Measurements of Protein Interactions in a Crowded Cellular Environment. *Anal. Chem.* *80*, 5976–5985. <https://doi.org/10.1021/ac800616u>.
- Li, F., Egea, P.F., Vecchio, A.J., Asial, I., Gupta, M., Paulino, J., Bajaj, R., Dickinson, M.S., Ferguson-Miller, S., Monk, B.C., et al. (2021). Highlighting membrane protein structure and function: A celebration of the Protein Data Bank. *J. Biol. Chem.* *296*, 100557. <https://doi.org/10.1016/j.jbc.2021.100557>.
- Li, W., Metcalf, D.G., Gorelik, R., Li, R., Mitra, N., Nanda, V., Law, P.B., Lear, J.D., DeGrado, W.F., and Bennett, J.S. (2005b). A push-pull mechanism for regulating integrin function. *Proc. Natl. Acad. Sci.* *102*, 1424–1429. <https://doi.org/10.1073/pnas.0409334102>.
- MacKenzie, K.R., Prestegard, J.H., and Engelman, D.M. (1997). A Transmembrane Helix Dimer: Structure and Implications. *Science* *276*, 131–133. <https://doi.org/10.1126/science.276.5309.131>.
- Majumder, A., Kwon, S., and Straub, J.E. (2022). On Computing Equilibrium Binding Constants for Protein–Protein Association in Membranes. *J. Chem. Theory Comput.* *18*, 3961–3971. <https://doi.org/10.1021/acs.jctc.2c00106>.
- Marinko, J.T., Huang, H., Penn, W.D., Capra, J.A., Schlebach, J.P., and Sanders, C.R. (2019). Folding and Misfolding of Human Membrane Proteins in Health and Disease: From Single Molecules to Cellular Proteostasis. *Chem. Rev.* *119*, 5537–5606. <https://doi.org/10.1021/acs.chemrev.8b00532>.
- Marrink, S.J., de Vries, A.H., and Mark, A.E. (2004). Coarse Grained Model for Semiquantitative Lipid Simulations. *J. Phys. Chem. B* *108*, 750–760. <https://doi.org/10.1021/jp036508g>.
- Marrink, S.J., Risselada, H.J., Yefimov, S., Tieleman, D.P., and de Vries, A.H. (2007). The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B* *111*, 7812–7824. <https://doi.org/10.1021/jp071097f>.
- Merzlyakov, M., You, M., Li, E., and Hristova, K. (2006). Transmembrane Helix Heterodimerization in Lipid Bilayers: Probing the Energetics behind Autosomal Dominant Growth Disorders. *J. Mol. Biol.* *358*, 1–7. <https://doi.org/10.1016/j.jmb.2006.01.086>.
- Mineev, K.S., Bocharov, E.V., Pustovalova, Y.E., Bocharova, O.V., Chupin, V.V., and Arseniev, A.S. (2010). Spatial Structure of the Transmembrane Domain Heterodimer of ErbB1 and ErbB2 Receptor Tyrosine Kinases. *J. Mol. Biol.* *400*, 231–243. <https://doi.org/10.1016/j.jmb.2010.05.016>.

- Mori, T., Miyashita, N., Im, W., Feig, M., and Sugita, Y. (2016). Molecular dynamics simulations of biological membranes and membrane proteins using enhanced conformational sampling algorithms. *Biochim. Biophys. Acta BBA - Biomembr.* *1858*, 1635–1651. <https://doi.org/10.1016/j.bbamem.2015.12.032>.
- Park, S., and Im, W. (2013). Two Dimensional Window Exchange Umbrella Sampling for Transmembrane Helix Assembly. *J. Chem. Theory Comput.* *9*, 13–17. <https://doi.org/10.1021/ct3008556>.
- Park, S., Kim, T., and Im, W. (2012). Transmembrane Helix Assembly by Window Exchange Umbrella Sampling. *Phys. Rev. Lett.* *108*, 108102. <https://doi.org/10.1103/PhysRevLett.108.108102>.
- Popot, J.L., and Engelman, D.M. (1990). Membrane protein folding and oligomerization: the two-stage model. *Biochemistry* *29*, 4031–4037. <https://doi.org/10.1021/bi00469a001>.
- Robertson, J.L., Kolmakova-Partensky, L., and Miller, C. (2010). Design, function and structure of a monomeric CIC transporter. *Nature* *468*, 844–847. <https://doi.org/10.1038/nature09556>.
- Russ, W.P., and Engelman, D.M. (1999). TOXCAT: A measure of transmembrane helix association in a biological membrane. *Proc. Natl. Acad. Sci.* *96*, 863–868. <https://doi.org/10.1073/pnas.96.3.863>.
- Schneider, D., and Engelman, D.M. (2003). GALLEX, a Measurement of Heterologous Association of Transmembrane Helices in a Biological Membrane. *J. Biol. Chem.* *278*, 3105–3111. <https://doi.org/10.1074/jbc.M206287200>.
- Sengupta, D., and Marrink, S.J. (2010). Lipid-mediated interactions tune the association of glycophorin A helix and its disruptive mutants in membranes. *Phys. Chem. Chem. Phys.* *12*, 12987. <https://doi.org/10.1039/c0cp00101e>.
- Yin, H., Litvinov, R.I., Vilaire, G., Zhu, H., Li, W., Caputo, G.A., Moore, D.T., Lear, J.D., Weisel, J.W., DeGrado, W.F., et al. (2006). Activation of Platelet $\alpha IIb\beta 3$ by an Exogenous Peptide Corresponding to the Transmembrane Domain of αIIb . *J. Biol. Chem.* *281*, 36732–36741. <https://doi.org/10.1074/jbc.M605877200>.
- Yoo, J., and Cui, Q. (2013). Membrane-Mediated Protein-Protein Interactions and Connection to Elastic Models: A Coarse-Grained Simulation Analysis of Gramicidin A Association. *Biophys. J.* *104*, 128–138. <https://doi.org/10.1016/j.bpj.2012.11.3813>.
- You, M., Li, E., Wimley, W.C., and Hristova, K. (2005). Förster resonance energy transfer in liposomes: Measurements of transmembrane helix dimerization in the native bilayer environment. *Anal. Biochem.* *340*, 154–164. <https://doi.org/10.1016/j.ab.2005.01.035>.

Zhang, J., and Lazaridis, T. (2006). Calculating the Free Energy of Association of Transmembrane Helices. *Biophys. J.* 91, 1710–1723.
<https://doi.org/10.1529/biophysj.106.081224>.

Chapter 2: Thermodynamic analysis of GAS_{right} dimerization supports a model in which stability is modulated by weak hydrogen bonding and van der Waals packing

Gladys Díaz Vázquez^{1,2}, Qiang Cui³, and Alessandro Senes¹

¹Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706

²Biophysics Graduate Program, University of Wisconsin-Madison, Madison, WI 53706

³Department of Chemistry, Boston University

A preprint of this chapter is available on BioRxiv at:

<https://doi.org/10.1101/2022.07.11.499632>.

Abbreviations

TM: transmembrane

GpA: glycophorin A

FRET: Förster resonance energy transfer

SN: staphylococcal nuclease

TMD: transmembrane domain

DM: n-decyl- β -d-maltopyranoside

OD600: optical density at 600 nm

VDW: van der Waals

WT: wild type

2.1. Abstract

The GAS_{right} motif, best known as the fold of the glycophorin A transmembrane dimer, is one of the most common dimerization motifs found in membrane proteins. GAS_{right} is characterized by its hallmark GxxxG-like sequence motifs (GxxxG, AxxxG, GxxxS, and similar). Structurally, GAS_{right} displays a right-handed crossing angle and very short inter-helical distance. The helical backbones are in contact, favoring the formation of networks of weak hydrogen bonds between Cα–H carbon donors and carbonyl acceptors on opposing helices (Cα–H···O=C hydrogen bonds). To understand the factors that modulate the stability of GAS_{right}, we previously presented a computational and experimental structure-based analysis of 26 predicted dimers. We found that the contributions of van der Waals packing and Cα–H hydrogen bonding to stability, as inferred from the structural models, correlated well with experimental dimerization propensities obtained with the *in vivo* assay TOXCAT. Here we assess this model rigorously and quantitatively by measuring the free energy of dimerization of a representative subset of 7 of the 26 original TOXCAT dimers using FRET. The thermodynamic data show strikingly good agreement between the original propensities and their ΔG° of association in detergent, suggesting that TOXCAT is a thermodynamically driven process. From the correlation between TOXCAT and thermodynamic stability, we back-calculate the predicted free energy for all the 26 GAS_{right} dimers. We show that these energies correlate with the *in silico* ΔE scores of dimerization that were computed on basis of their predicted structure. These findings corroborate our original model with quantitative thermodynamic evidence, confirming

and strengthening the hypothesis that van der Waals and Ca–H hydrogen bond interactions are the key modulators of GAS_{right} stability.

2.2. Introduction

Membrane protein oligomerization is a fundamental process in the life of a cell. Oligomerization is especially important for bitopic proteins, i.e. the membrane proteins that contain a single transmembrane (TM) helix. The association of these TM helices can be optimized for stability in constitutive dimers, such as the case of glycophorin A (GpA) (MacKenzie et al., 1997). In other instances, stability is tuned appropriately to support dynamic association, which can be critical for regulating signal transduction or activation in important biological systems, such as integrins (Li et al., 2005; Yin et al., 2006) and receptor tyrosine kinases (Bocharov et al., 2008, 2012; Mineev et al., 2010), to name a few. Understanding the interplay between the forces involved in TM helix oligomerization could support the prediction of structure and stability, the identification of potential conformational changes, and the interpretation of the effect of mutations in these system, providing insight into biological function and regulation.

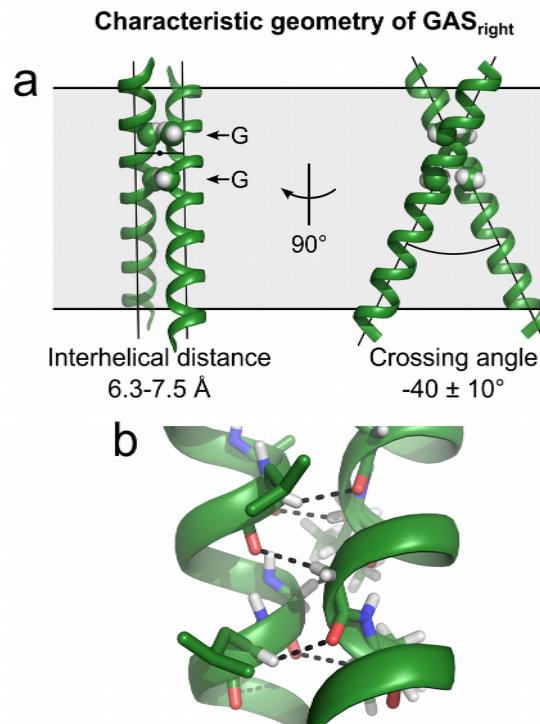


Figure 2.1: **Geometry of the GAS_{right} motif.** *a*, GAS_{right} is a right-handed helical dimer with a short inter-helical distance (6.3-7.5 Å) and a right-handed crossing angle of approximately -40° . The GxxxG sequence pattern at the crossing point (indicated by arrows) allows the backbones to come into contact. *b*, This contact enables formation of networks of 4 to 8 weak inter-helical H-bonds between Ca-H donors and carbonyl oxygen acceptors at the crossing point of the dimer.

In the present study, we focus on the energetics of association of an important structural dimerization motif known as GAS_{right} (Mueller et al., 2014), which is best known as the fold of the prototypical GpA dimer (MacKenzie et al., 1997). The GAS_{right} motif is one of the structural motifs most commonly observed in dimeric transmembrane complexes (Walters and DeGrado, 2006; Zhang et al., 2015a). It is characterized by the

presence of the small amino acids, Glycine, Alanine, and Serine (GAS) at the dimer interface, and a right-handed crossing angle of approximately -40° between the two helices. The small residues are separated by three amino acids and arranged on the same face of the helix to form GxxxG-like sequence motifs (GxxxG, GxxxA, SxxxG, etc.) (Brosig and Langosch, 1998; Russ and Engelman, 2000; Senes et al., 2000) (Figure 2.1).

They form a flat face that allows the helical backbones to come in close contact, promoting tight packing. The contact between the backbones at the specific geometry of GAS_{right} promotes the formation of networks of weak hydrogen bonds between $\text{Ca}-\text{H}$ carbon donors and carbonyl acceptors on opposite helices ($\text{Ca}-\text{H}\cdots\text{O}=\text{C}$) (Mueller et al., 2014). Carbon is a weak hydrogen bond donor, but $\text{Ca}-\text{H}$ groups are activated by the flanking electron-withdrawing amide groups in the peptide backbone, and thus the strength of $\text{Ca}-\text{H}$ hydrogen bonds has been estimated to be as much as one-half of that of N–H donors in vacuum (Scheiner et al., 2001; Vargas et al., 2000). Therefore, multiple $\text{Ca}-\text{H}$ hydrogen bonds occurring at the dimerization interface are expected to contribute significantly to the free energy of association in GAS_{right} dimers (Mueller et al., 2014; Senes et al., 2001).

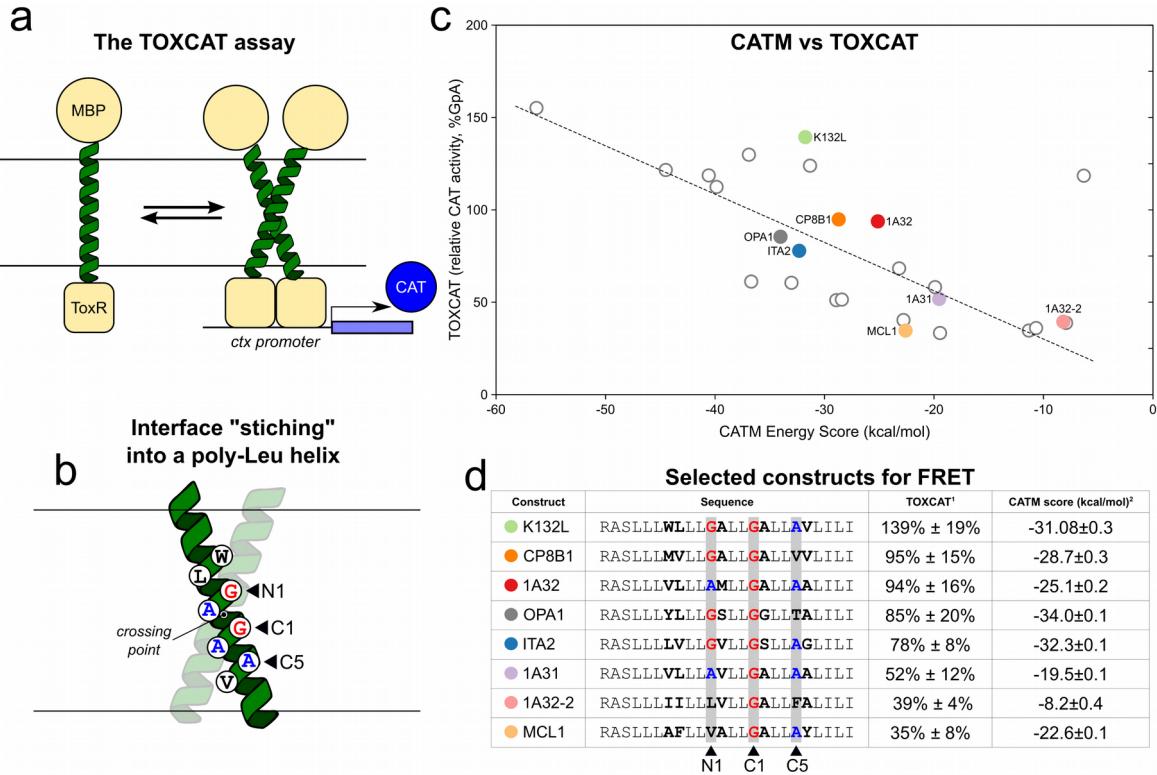


Figure 2.2: Construct selection. The selection of GAS_{right} constructs is a subset of sequences previously studied using TOXCAT (Anderson et al., 2017). *a*, TOXCAT is an *in vivo* assay. The TM domain under investigation is fused to the ToxR transcriptional activator. TM helix association results in the expression of a reporter gene in *E. coli* cells (CAT), which can be quantified. *b*, To reduce variability in TOXCAT, the eight amino acids that formed the interface as predicted by CATM were “stitched” into a standardized poly-Leu sequence (Anderson et al., 2017). *c*, Correlation between the energy score predicted with the program CATM and their dimerization propensity measured with TOXCAT assay for a series of 26 standardized GAS_{right} dimers in our previous study (Anderson et al., 2017). The

data suggest that two of the primary interactions that contribute to modulating the dimerization stability of these constructs are a combination of VDW interactions and weak $\text{Ca}-\text{H}$ hydrogen bonding. The 8 constructs selected for the present study are highlighted in color. *d*, List of the subset of eight GAS_{right} dimers analyzed in this study. The subset covers a range of TOXCAT homodimerization propensities and CATM energy scores.

To test the hypothesis that $\text{Ca}-\text{H}$ hydrogen bonds, along with van der Waals (VDW) packing, are major drivers of stability in GAS_{right} dimers, we previously used a combined computational experimental approach (Anderson et al., 2017). Specifically, we predicted the structure of a series of 26 GAS_{right} dimers using the program CATM (Mueller et al., 2014) and compared the energy score that was calculated with the experimental dimerization propensities obtained with TOXCAT (Russ and Engelman, 1999), a genetic assay that measures oligomerization in the *Escherichia coli* membrane (Figure 2.2a). To reduce some of the variability that is typical of a biological assay, we redesigned the constructs by “stitching” the 8 positions predicted by CATM to be at the helix-helix interface into a standardized helix of 21 amino acids consisting of a poly-Leu backbone (LLLxxLLxxLLxxLLxxLILI, where the x represents the variable interfacial amino acids (Anderson et al., 2017). Such standardization was important to reduce the differences in protein expression level. It also allowed us to focus on the forces that play a role at the interaction interface, isolating them from other variables that could contribute to dimerization stability, such as the length of the hydrophobic region of the TM helices and the position of the crossing point in the dimer, which were features shared by all

constructs (Figure 2.2b). The predicted structures were partially validated by mutating the critical interfacial glycine at C1 (the position at the interface that is in the closest contact with the opposing backbone) to a large isoleucine, a mutation that is expected to completely abolish dimerization by creating steric clashes (Anderson et al., 2017).

As shown in Figure 2.2c, the study identified a statistically significant relationship between experimental stability estimated with TOXCAT and the computational energy scores of the constructs (Anderson et al., 2017). However, neither CATM nor TOXCAT are rigorous methods for assessing the free energy of association of TM dimers. CATM is a program that can predict the structures of known GAS_{right} dimers with high accuracy, but its energy function is relatively simple, consisting of an unweighted sum of only three terms: VDW, hydrogen bonding, and an implicit solvation function. The ΔE score it produces is obtained by calculating the differential of these three terms between static monomeric and dimeric models. TOXCAT reports relative propensities of association in actual biological membranes, but it does it indirectly, through the stimulation of the expression of a reporter gene (Figure 2.2a). The assumption is that the signal is proportional to the amount of dimer present in the membrane, and thus the response is thermodynamically driven (Duong et al., 2007), but the assay could be affected by the variability of a living biological system and in particular by differential expression levels of the constructs. Despite these limitations, by sampling a large set of validated GAS_{right} constructs of different stabilities with these two methods, we were able to identify features that correlated statistically with the apparent stability of the dimers (Anderson et al., 2017). The data indicated that packing and Ca–H hydrogen bonds were major

factors that modulated the stability of these constructs and suggested that stability can be governed by inter-helical geometry, which in turn is enabled by the sequence (Anderson et al., 2017).

In this study, we take another step forward towards formulating a quantitative model of the energetics of GAS_{right}. We selected a representative subset of the original 26 GAS_{right} dimers (Anderson et al., 2017) and used Förster Resonance Energy Transfer (FRET) to determine their free energy of association in detergent *in vitro*. We found a strong correlation exists between the free energy of association in detergent and the dimerization propensity obtained previously with TOXCAT. The data also suggest that there is a striking correspondence between the $\Delta\Delta G^\circ$ of association measured *in vitro* and the same quantities in the membrane of *E. coli*. The study provides a thermodynamic foundation for the model that identifies a combination of VDW packing and weak hydrogen bonding as the primary drivers and modulators of stability of GAS_{right} dimers.

2.3. Results and Discussion

2.3.1. FRET construct design

While the majority of oligomerization studies of TM helices based on FRET have been performed on synthetic peptides (Anbazhagan and Schneider, 2010; Duneau et al., 2007; Fisher et al., 1999; Maeda et al., 2018; You et al., 2005), we opted for biological expression to increase throughput. Specifically, we used a chimeric construct that fuses a soluble Staphylococcal nuclease (SN) domain to the TM helices, a strategy used in previous thermodynamic association studies based on other methods (Bu and

Engelman, 1999; Doura and Fleming, 2004; Fleming, 2004; Fleming et al., 1997; Hong and Bowie, 2011; Hong et al., 2010; Lemmon et al., 1992a, 1992b). As a minor change, we added a 15-amino acid long flexible linker between the TM helix and SN moieties, to structurally decouple the orientation of the two domains and thus minimize any influence of the geometry of the TM region on the FRET efficiency. For labeling, we screened three surface positions, including an isosteric mutation (S3C), and two Ala positions that were previously found not to dramatically affect SN's stability when changed to cysteine (A60C and A112C) (Kim and Stites, 2008). Variant A112C yielded calculated labeling efficiencies of 70-80% (supplementary Table S2.7.1) when reacted with either Cy5-maleimide (acceptor), or Cy3-maleimide (donor) and was thus selected for the study (Figure 2.3a).

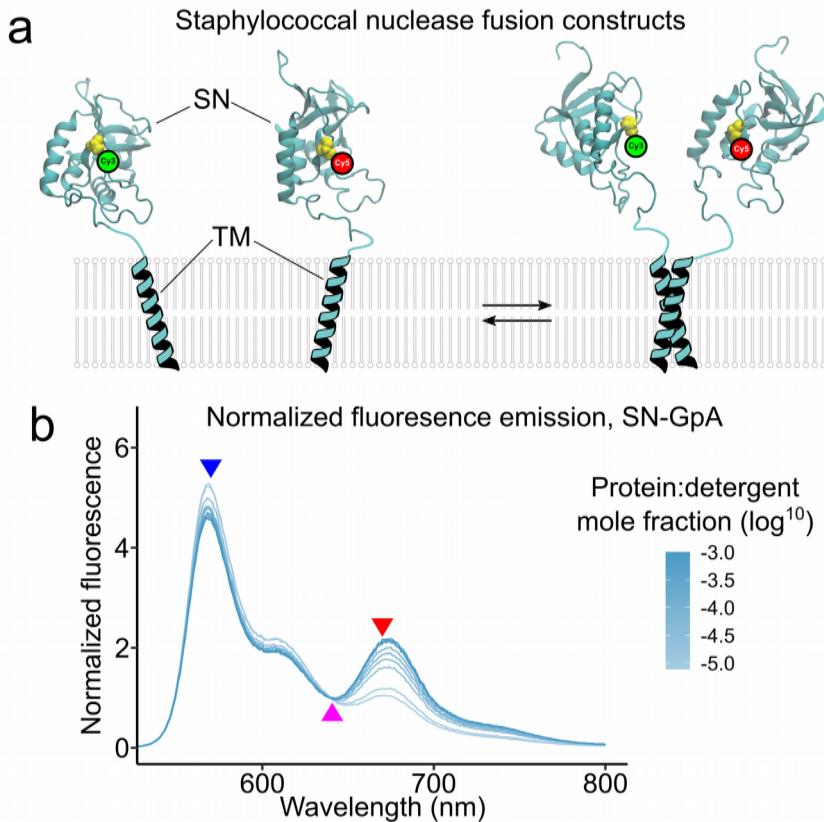


Figure 2.3: The FRET experiment. *a*, The reaction under study corresponds to a monomer-dimer equilibrium. The N-terminus of the TM helices of interest were fused to the C-terminus of the soluble staphylococcal nuclease (SN) through a 15-residues flexible linker. SN was labeled at Cys residue introduced at position 112 with either Cy3-maleimide (donor) or Cy5-maleimide (acceptor) for FRET experiments. *b*, Example of a series of fluorescence spectra obtained at different mole fractions (moles of protein/moles of detergent), driving the reaction from a monomeric to a dimeric state (construct SN-GpA). The fluorescence spectra were normalized using the isosbestic point indicated by the magenta arrow at 640 nm.

The donor emission at 570 nm (blue arrow) decreases and the acceptor emission at 670 nm (red arrow) increases as the protein mole fraction is increased. Original and normalized spectra of all constructs are shown in supplementary Figure S2.1.

From the 26 constructs from our previous TOXCAT study (Anderson et al., 2017) we excluded five that contained a cysteine in their sequence to avoid possible issues with non-specific labeling. The constructs that did not clone efficiently or expressed poorly were also eliminated. This yielded a final set of eight constructs that cover a wide range of homodimerization propensities in TOXCAT and energy scores predicted by CATM (Figure 2.2d). These constructs were expressed, purified in *n*-decyl- β -D-maltopyranoside (DM) detergent, and labeled.

2.3.2. Theoretical framework of the FRET experiments and global fitting analysis

FRET experiments were performed by keeping the amount of protein constant and changing the concentration of the DM detergent in solution, under the well-established assumption that the equilibrium of the dimerization reaction is governed not by the total volume but by the “hydrophobic volume”, given that the TM helices are constrained to reside in the interior of a detergent micelle (Fleming, 2002). Protein concentration is thus expressed as protein:detergent mole fraction, which was explored over a range of approximately 1:10⁻⁵ to 1:10⁻³ to drive the reaction from monomer to dimer states. The mole fraction was corrected by subtracting the critical micellar concentration from the detergent concentration, to account for the presence of monomeric (non-micellar) detergent.

An example series of donor-acceptor fluorescence spectra at increasing protein mole fraction is illustrated in Figure 2.3b. The spectra were first corrected by subtracting the contribution to emission of acceptor-only samples with the same concentration and then normalized using the isosbestic point (640 nm for our system), to correct for changes in signal that are not related to the transfer of energy between the fluorophores, such as small errors in the estimate of concentration (Hessels and Merkx, 2016). Original and normalized spectra of all constructs are shown in supplementary Figure S2.1.

The FRET series was validated by plotting donor emission at 570 nm against acceptor emission at 670 nm, which yielded the expected inverse linear relationship (supplementary Figure S2.2). Equilibrium constants of dissociation were obtained by fitting the normalized fluorescence intensity at 670 nm (NF_{670}) as a function of the protein: detergent mole fraction (χ_T) using equation 1 (derived in supplementary material Derivation S2.1):

$$NF_{670} = NF_D + (NF_M - NF_D) \frac{-K_D + \sqrt{K_D^2 + 8\chi_T K_D}}{4\chi_T} \quad (\text{equation 1})$$

Three parameters were derived from the fitting. Two of them were global parameters, NF_D and NF_M , i.e. the normalized fluorescence intensities at the limits in which the samples are fully dimeric (at infinite concentration) or monomeric (at infinite dilution), respectively. The third parameter was the desired dissociation equilibrium constant K_D , which was individually fitted to the data of each construct.

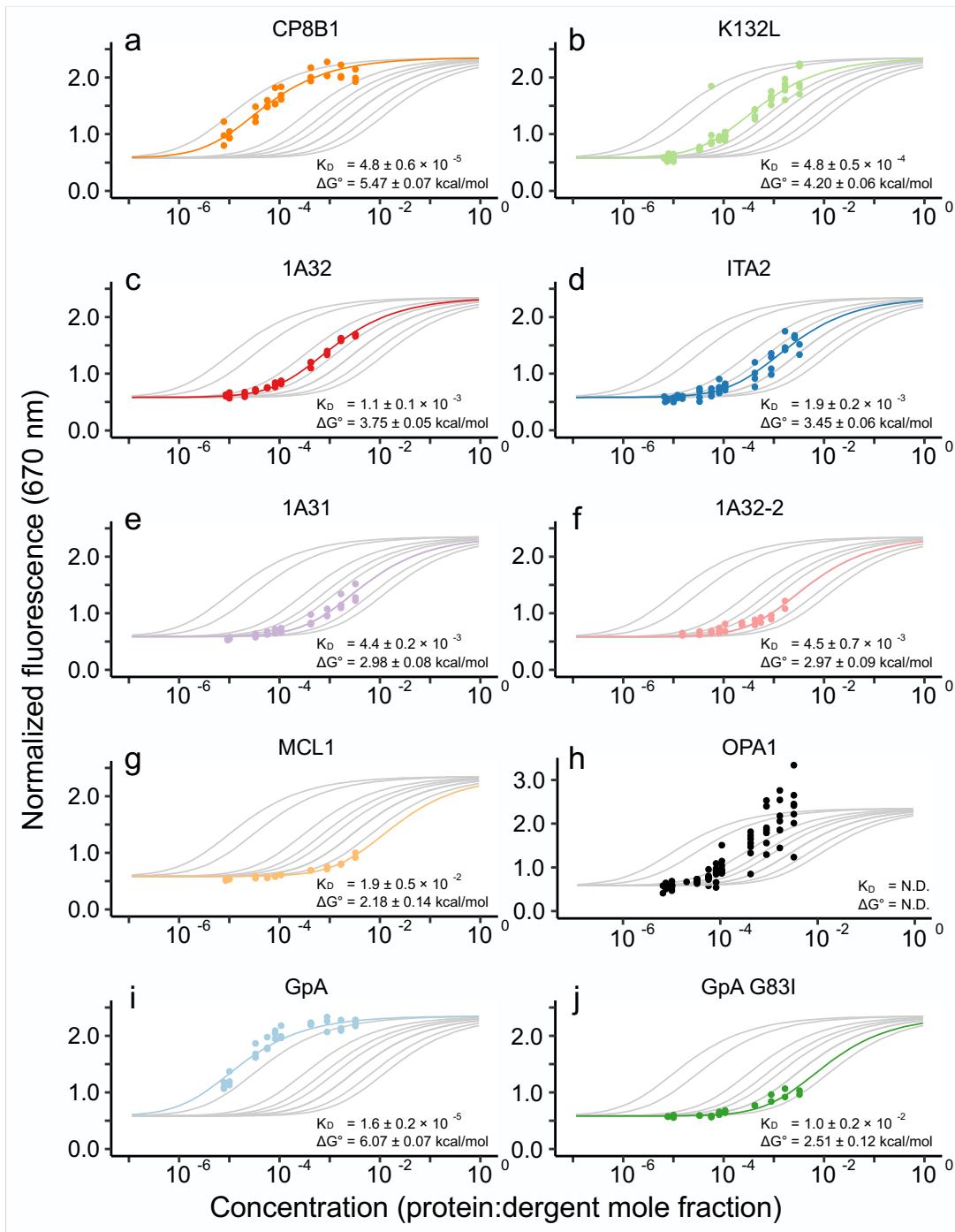


Figure 2.4: Thermodynamic analysis, global fitting of FRET data. The graphs show the fluorescence data as a function of concentration in FRET experiments for the eight selected GAS_{right} constructs from the 26 previously studied using TOXCAT (a-h). The prototypical GAS_{right} dimer GpA (i) and its monomeric mutant GpA G83I (j) were also included as controls. The gray lines correspond to the global fit of the FRET data, which yielded the dissociation constants (K_D). The construct OPA1 (h) was excluded from the global fitting because it showed a larger emission range and its stoichiometric analysis indicated it forms higher-order oligomers (supplementary Figure S2.3).

A major challenge presented by this type of FRET analysis is obtaining enough coverage of the association range (from mainly monomeric to mainly dimeric) so that all three parameters can be calculated with sufficient confidence. As illustrated in Figure 2.4, the weakest dimers (MCL1, 1A32-2, and 1A31) cover primarily the monomeric region, whereas the strongest dimers (CP8B1 and the control GpA) cover most of the transition and approach a fully dimeric state but lack the monomeric baseline. Fitting the NF_D and NF_M parameters globally solves this problem since these baselines are expected to be similar across all constructs.

2.3.3. The selected GAS_{right} dimers cover a wide range of thermodynamic stabilities in vitro

Figure 2.4a-h shows the normalized fluorescence data as a function of mole fraction for the eight selected GAS_{right} constructs. It also shows the fitted binding curves and the

calculated free energies of dissociation (with the exception of construct OPA1, panel *h*, which was not included in the fitting). Two additional controls were also measured: the highly stable GpA dimer (panel *i*) and its monomeric G83I variant (*j*). Inclusion of these two controls of known stability contributed to establishing the NF_D (wild type GpA) and NF_M (G83I variant) baselines in the global fitting analysis.

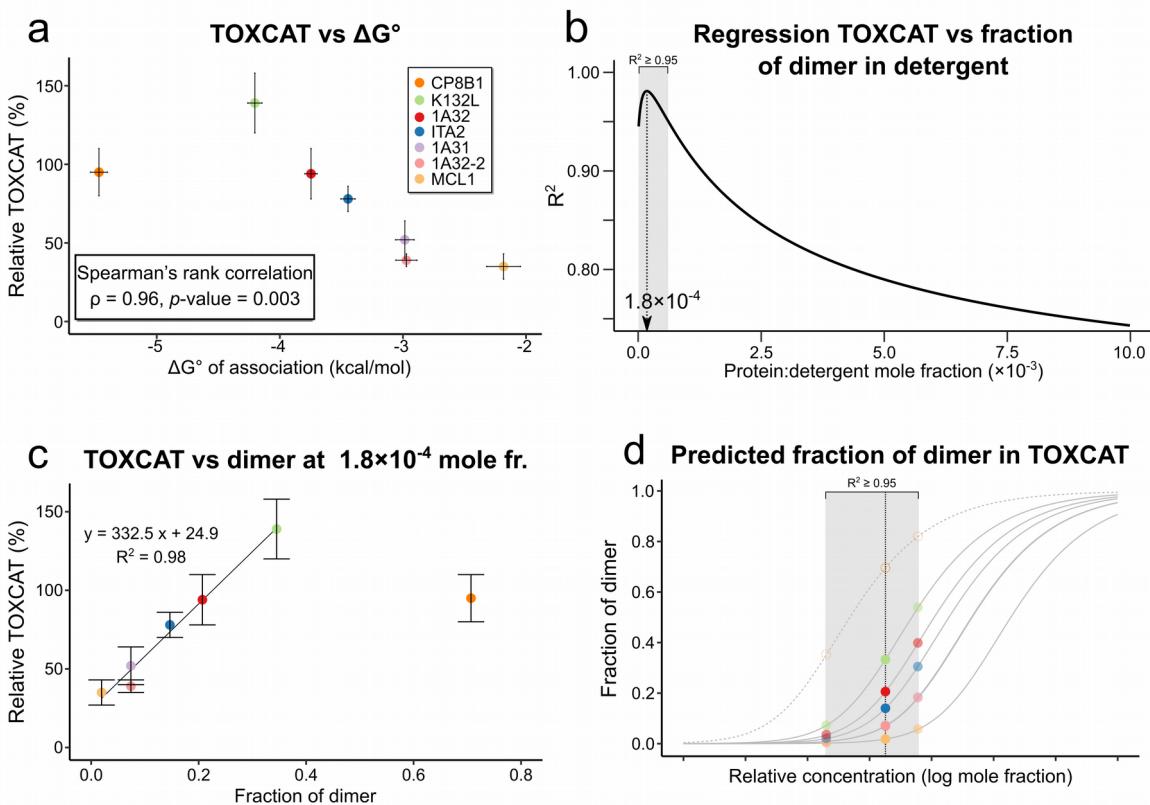


Figure 2.5: Comparison of the FRET results with the previous TOXCAT analysis. *a*, The thermodynamic stabilities of the constructs obtained here in detergent are compared to their previously determined TOXCAT homodimerization propensities (Anderson et al., 2017). The expected inverse non-linear relationship is observed. A clear outlier in the trend can be observed, corresponding to CP8B1. This outlier was included in the statistical analysis. *b*, Results of the linear regression analysis between TOXCAT and fraction of dimer calculated at different mole fractions. Here the regression coefficient R^2 is plotted (linear fits plotted in supplementary Figure S2.4). The coefficient is maximized at a 1.8×10^{-4} mole fraction. The region in which $R^2 \geq 0.95$ is shaded, corresponding to

concentrations between 2×10^{-5} and 6×10^{-4} mole fractions. The outlier CP8B1 was not included in this linear fit. *c*, Linear fit of TOXCAT vs fraction of dimer at 1.8×10^{-4} mole fraction. The outlier CP8B1 was not included in this linear fit. *d*, The predicted fraction of dimer in the biological membrane of *E. coli*, based on the linear relationship observed between TOXCAT and the fraction of dimer, under the assumption that CAT expression is directly proportional to fraction dimer. The data suggest the constructs are mostly monomeric in TOXCAT, ranging from 2% for the weakest dimer (MCL1) to 35% for the strongest (K132L) from the conditions of maximum correspondence with the free energy data. The region that covers the predicted fraction dimer for all concentrations that produced an $R^2 \geq 0.95$ is also indicated (shaded).

Overall, the constructs cover a wide range of thermodynamic stabilities with ΔG° values ranging approximately from 2 to 6 kcal/mol. There is good consistency among the baselines of the various samples. Specifically, the constructs that occur in a prevalent unassociated state at the highest dilutions (ITA2, K132L, G83I, 1A32-2, 1A32, MCL1, and 1A31) are all in good agreement with the globally fit monomeric baseline ($NF_M = 0.58 \pm 0.01$). Correspondingly, the two constructs that approach a fully associated state at their highest concentrations (GpA and CP8B1) are consistent with the calculated dimeric baseline ($NF_D = 2.35 \pm 0.03$). Finally, the data points of the construct that span a significant portion of the monomer/dimer transition (GpA, CP8B1, ITA2, K132L, and 1A32) fit the curves well. Only the construct OPA1 displayed a larger

emission range than the others, suggesting it may form higher-order oligomeric states. This was confirmed by stoichiometry analysis by varying donor-acceptor ratio (Adair and Engelman, 1994; Condon et al., 2018) (supplementary Figure S2.3). For this reason, OPA1 was not included in the global fitting and was no longer considered in the analysis.

The standard association free energy that we obtained for GpA was -6.07 ± 0.07 kcal/mol. This value is comparable with previous thermodynamic studies of GpA in other detergents by analytical ultracentrifugation, which reported free energies of -5.7 ± 0.3 and -9.0 ± 0.1 kcal/mol in C14 betaine (Fleming, 2004) and C₈E₅ (Fleming et al., 1997), respectively. As expected, the GpA G83I monomeric variant is much destabilized (-2.51 ± 0.12 kcal/mol). These results demonstrate that the global fit strategy across a set of constructs of varied stability enables the determination of the dissociation constant for a series of samples that individually do not cover a sufficient range from dissociated to associated states.

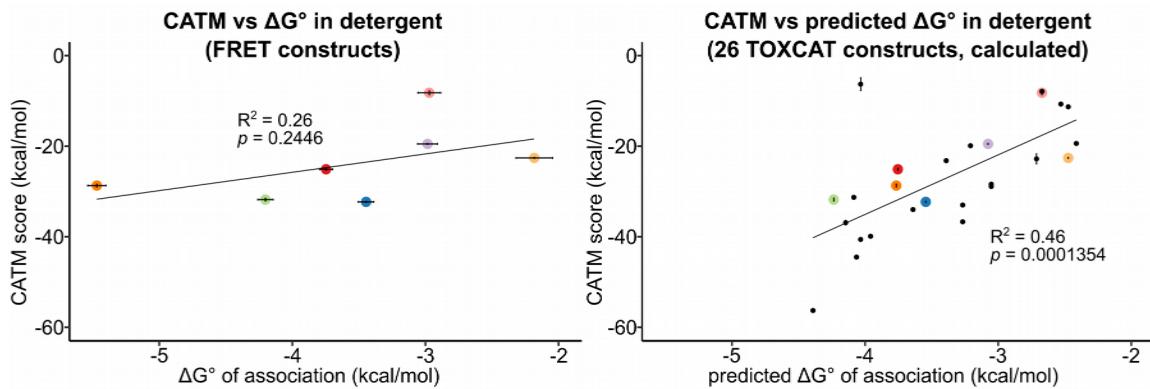


Figure 2.6: Comparison of experimental and computational energies of dimerization. *a*, The experimental ΔG° of association for the 7 constructs studied here do not result in a significant linear correlation with the previously determined CATM energy scores (Anderson et al., 2017). *b*, The comparison was expanded by predicting the free energy of association for all 26 original constructs from their previously determined TOXCAT data (Anderson et al., 2017). A significant correlation between this set and the CATM energy score was observed.

2.3.4. Strong correspondence of relative stabilities in detergent and the *E. coli* membrane

After obtaining the thermodynamic stability of the constructs in detergent *in vitro*, we compared it with their dimerization propensities in biological membranes obtained previously with TOXCAT (Anderson et al., 2017). As shown in Figure 2.5a, we observed the expected non-linear monotonic inverse relationship between the ΔG° of association and TOXCAT signal (Spearman's rank correlation, $\rho = 0.96$, p -value = 0.003). Only one notable outlier does not follow the proportionality, CP8B1, which is the most stable

among the constructs in this study but displayed only intermediate stability in the TOXCAT study. Among all constructs of our previous TOXCAT study, CP8B1 was the one that displayed the largest variation in its degree of expression (Fig S2 in Anderson et al. (Anderson et al., 2017)), therefore this discrepancy may be possibly due to significant differences in protein concentration of CP8B1 in the *E. coli* membrane compared to the other constructs.

The overall good correlation with the FRET data suggests that the response of TOXCAT is indeed governed by the thermodynamic stability of the dimers and that the expression of the CAT reporter gene is directly proportional to the amount of dimer that occurs in the membrane (Anderson et al., 2017). It also suggests that there is a close relationship between the stability of the constructs in detergent and the membrane. This relationship cannot be resolved directly from the TOXCAT data because of two unknown quantities. First, the free energies of dimerization of the constructs in the two environments are likely different (Chen et al., 2010). In addition, the concentration of constructs in the *E. coli* membrane is unknown. However, if we make the assumption that the expression level in TOXCAT is relatively similar across this set of standardized poly-Leu constructs (as assessed by Western blotting (Anderson et al., 2017)), we can ask whether the relative stability ($\Delta\Delta G^\circ$) of the series of constructs is conserved between the two environments.

We asked this question by checking if there is a correlation between the amount of dimer in detergent and their TOXCAT signal, which we assume is a proxy for the amount of dimer in the membrane. To do so, we performed a series of linear

regressions between TOXCAT signal and the dimer fraction in detergent calculated across a range of mole fractions that cover the entire transition from fully monomeric to fully dimeric, as shown in supplementary Figure S2.4 (the construct CP8B1 was excluded from this analysis since it is a clear outlier in the TOXCAT vs stability relationship).

We found that TOXCAT and dimer fraction correlate extremely well at certain concentration regimes. The regression coefficient plotted against protein:detergent mole fraction reaches a maximum of $R^2 = 0.98$ at a mole fraction of 1.8×10^{-4} (Figure 2.5b,c). At this mole fraction, the amount of dimer ranges from approximately 2% for the weakest dimer to 35% for the strongest (Figure 2.5d). It should be noted that the correlation remains very strong for a range of concentrations. As a reference, the range in which $R^2 \geq 0.95$ is between 2×10^{-5} and 6×10^{-4} mole fraction (shaded region in Figure 2.5b). We conclude that the monomer/dimer equilibria of the six constructs occur in the membrane at levels of dimerization that are close to the regime observed at 1.8×10^{-4} in detergent (dashed line in Figure 2.5d) and most likely falling on the left half of the binding curve (the shaded area in Figure 2.5d). In this region, most constructs go from all nearly monomeric to reaching halfway in the saturation curve for the most stable. These data, therefore, are in reasonable agreement with a previous suggestion that TOXCAT constructs exist in mostly monomeric state in the membrane (Duong et al., 2007).

These results indicate that there is a striking correspondence between relative stability ($\Delta\Delta G^\circ$) in detergent and the *E. coli* membrane for this set of constructs. They

confirm for the hypothesis that the TOXCAT process is governed by thermodynamic stability (Duong et al., 2007). Finally, the findings provide validation for the experimental data that was used to base the model that a combination of VDW packing and weak hydrogen bonding act as the primary drivers and modulators of stability of GAS_{right} dimers (Anderson et al., 2017).

2.3.5. The CATM energy score captures predicted stabilities in detergent

After examining the relationship with TOXCAT, we proceeded to compare the free energy of association *in vitro* with the energy score calculated with the structural prediction program CATM (Mueller et al., 2014). We first compared the experimental ΔG° of association in detergent of the subset of seven GAS_{right} constructs analyzed here with their CATM energy score (Figure 2.6a). The linear regression did not produce a significant correlation ($R^2=0.26$ $p\text{-value}=0.2446$), indicating that CATM does not capture well the energetics of association of this particular subset of constructs. We then took advantage of the derived relationship between TOXCAT signal and the ΔG° of association to back-calculate predicted ΔG° values in detergent for all 26 original constructs (supplementary Table S2.7.2) and assess the correlation with their CATM energy score. In this case, we found a reasonable correlation ($R^2=0.46$) with highly significant $p\text{-value} = 0.0001354$ (Figure 2.6b). The majority of the points form a clear trend dispersed around the line, with only two data points performing as clear outliers (TNR12, ROMO1). This discrepancy between Figure 2.6a and b is not unreasonable: although a clear correlation can be identified with a larger set of 26 constructs, the same

correlation may not become apparent with a smaller number of data points due to the noisy relationship and the loss of statistical power.

It should be noted that in Figure 2.6*b* there is a large difference in the scale of the two axes, with approximately 17 kcal/mol of CATM energy corresponding to just 2 kcal/mol of predicted ΔG° in DM. This difference is not straightforward to interpret physically. It appears that CATM somehow over-estimates the interactions, highlighting a need for recalibrating its energy function, which is not unexpected. The CATM function is simple, as an unweighted sum of only three terms, VDW interactions (CHARMM 22 (MacKerell et al., 1998)), hydrogen bonding (SCRWL 4 (Krivov et al., 2009)), and implicit solvation (IMM1 (Lazaridis, 2003)). The score ignores other potentially important terms, such as, for example, electrostatics or entropic contributions. With the availability of a sufficiently large data set of validated GAS_{right} TOXCAT constructs, it would be possible to derive (and rigorously test) an effective energy function to predict the relative stability of these dimers in detergent and potentially in the biological membrane. Nevertheless, it is notable that the simple energy function of CATM in its current form, appears to be a reasonable predictor of the relative stability of a large set of GAS_{right} dimers, confirming the model.

2.4. Conclusions

Understanding the physical basis of folding and association in membrane proteins remains challenging because of the technical difficulties posed by these systems, together with the complexity of a process occurring in an anisotropic and highly heterogeneous milieu (Hong, 2014). Many factors have been proposed to play a role in

membrane protein folding and association. Among them are lipid-specific effects (such as solvophobic exclusion, specific lipid binding, lateral pressure, and hydrophobic matching (Chadda et al., 2021; Cymer and Schneider, 2010; Cymer et al., 2012; Duneau et al., 2017; Langosch and Heringa, 1998; Marsh, 2008; Soubias et al., 2015; Yano et al., 2015)) and a variety of physical interactions (for example, aromatic and cation- π interactions (Hong et al., 2007; Johnson et al., 2007; Sal-Man et al., 2007)). However, two prominent candidates have emerged as the major determinants of stability in membrane proteins, namely the packing of apolar side chain and hydrogen bonding.

Packing is ubiquitous at helix-helix interfaces. These helices consist primarily of hydrophobic amino acids and their packing result in extensive and favorable VDW interactions (Hong, 2014; Joh et al., 2009; Langosch and Heringa, 1998; MacKenzie and Engelman, 1998). Structural informatics indicate that side chains pack more efficiently in membrane proteins (Adamian and Liang, 2001; Javadpour et al., 1999; Joh et al., 2009; Zhang et al., 2015b) whereas mutations that disrupt packing are generally strongly detrimental to stability (Baker and Urban, 2012; Duong et al., 2007; Guo et al., 2016; MacKenzie and Engelman, 1998), which clearly indicates that good packing is a necessary feature of stable membrane proteins and complexes. The question is to what extent packing can be a main driving force beyond a mere necessity. The favorable packing of apolar amino acids at helical interfaces in the associated state is counterbalanced by the favorable VDW interactions that the amino acids form with the lipids acyl chains in the unassociated state. Therefore, for packing to be an actual

driving force, the protein-protein interactions in the folded state need to overcome the contributions of protein-lipid interactions that are lost upon folding. Although this question remains open, the recent design of a stable oligomeric complex created solely around apolar packing offers proof-of-principle evidence that is consistent with VDW interactions as an actual driving force (Mravic et al., 2019).

The second main candidate that has emerged as an important force for membrane protein folding is hydrogen bonding. There is little doubt that “canonical” hydrogen bonds (i.e. those involving O-H and N-H donors) can, at least in some instances, promote folding when polar side chains are present at helix-helix interfaces (Choma et al., 2000; Gratkowski et al., 2001; Zhou et al., 2000, 2001). This is because the loss of hydrogen bonding when the helices are separated is not compensated by the interactions between the polar side chains with the apolar lipid environment, which are not as favorable. The same argument should apply to the weak $\text{Ca}-\text{H}$ hydrogen bonds, the signature trait of $\text{GAS}_{\text{right}}$ dimers, but measuring their contribution to association has been exceedingly difficult (Arbely and Arkin, 2004; Yohannan et al., 2004). A major technical challenge is the fact that donor ($\text{Ca}-\text{H}$) and acceptor groups (backbone carbonyl) are both part of the backbone, making a rational mutation strategy difficult to implement.

In this study, we advance these debates as they apply to the $\text{GAS}_{\text{right}}$ motif, with a thermodynamic analysis of association of eight representative candidates from the original pool of 26 constructs we previously studied with TOXCAT. Of these constructs, one formed higher oligomers whereas a second stood out as a clear outlier in the

apparent relationship between ΔG° of association and TOXCAT signal. For the remaining six constructs, we found that there is a striking correspondence between their relative stability ($\Delta\Delta G^\circ$) in detergent and their predicted stability in the biological membrane of *E. coli* calculated from TOXCAT (Figure 2.5c). Additionally, we found that when the free energy of association of the entire original pool of 26 constructs is back-calculated from TOXCAT using this apparent relationship, a significant correlation is observed with the ΔE calculated with CATM based on their predicted structural models (Figure 2.6b). Therefore, the present data provide quantitative thermodynamic validation that is consistent with our previous findings (Anderson et al., 2017) and supports them.

The GAS_{right} motif is a versatile dimerization motif found both in structural proteins that require maximal stability, as well as in dynamic proteins, such as receptors, in which stability needs to be finely modulated to support their function. Combined, the present and previous studies provide insight into how sequence, structure, and energetic factors modulate this stability. The analysis of the predicted structural models identified significant trends that suggested that geometry can affect stability (Anderson et al., 2017). Specifically, shorter interhelical distances and crossing angles near -40° tend to produce the most stable dimers. In turn, these ideal geometries are favored by the presence of GxxxG motifs (as opposed to other GxxxG-like motifs, such as AxxxG or GxxxA). In particular, dimerization appears to become most effective when a Gly residue is present at the interfacial position designated as N1, forming a GxxxG with the Gly at C1 (Figure 2.2b). None of these rules are stringent; it is possible for dimers containing GxxxG motifs to assume near-ideal geometry and, yet, their stability could be

detuned if the shape of the residues at the dimer interface is not conducive to good packing (Doura and Fleming, 2004). However, the analysis suggests that a geometric “sweet spot” exists, where weak hydrogen bonding and packing can be optimized, and which can be exploited by nature when maximal stability is required for function.

2.5. Experimental Procedures

2.5.1. Plasmid cloning

A gblock containing the soluble domain of Staphylococcal nuclease (SN) (amino acids: ALA1-GLN149) fused to the TM domain of GpA (amino acids: E72-R96) through a flexible linker (amino acids: TSGG[SGGG]₂SGGS) was inserted into a pet28a plasmid by restriction free cloning (van den Ent and Löwe, 2006). Mutations to the soluble domain SN to obtain S3C, A60C, and A112C were introduced using either QuickChange mutagenesis or double primer mutant method (Zheng, 2004). Restriction free cloning (van den Ent and Löwe, 2006) was used to replace the TM domain of GpA with the TM domain of the 8 GAS_{right} constructs obtained from the TOXCAT plasmid. All protein sequences are reported in supplementary Table S2.7.3.

2.5.2. Protein expression

Plasmids were transformed into BL21 (DE3) non-tunner cells and plated in LB agar plates containing 50 µg/ml of kanamycin and incubated at 37 °C overnight. Colonies were inoculated in ~3 mL of LB media containing 50 µg/ml of kanamycin and incubated in a shaker overnight at 37 °C. Overnight cultures were inoculated into ZYP-5052 autoinduction media (Studier, 2005) (1 ml of overnight culture per 500 ml of autoinduction media) containing 400 µg/ml of kanamycin and incubated in a shaker at

37 °C until the culture reached approximately 0.8 OD₆₀₀, at which point the temperature was lowered to 25 °C and cells were left to grow overnight. Cells were collected by centrifugation at 5,000 ×g for 10 min. Pellets were washed with cell wash buffer (50mM TrisHCL pH 7.9, 100mM NaCl), then centrifuged again at 5,000 ×g for 10 min. The pellets were flash frozen in liquid nitrogen and stored at -80 °C.

2.5.3. Purification of constructs

Cell pellets were mixed with lysis buffer (50 mM Tris HCl pH 7.9, 5 mM EDTA, 1 mM PMSF, 1 mg/ml Lysozyme and 5 mM BME) at a ratio of 10 ml of buffer per gram of cells, then lysed via sonication. The lysate was centrifuged at 10,000 ×g for 15 min. The membrane fraction was then isolated via ultracentrifugation of the supernatant at ~185,500 ×g for 30 min. The membrane pellet was then resuspended with approximately 5 ml/g of cells in solubilization buffer (50 mM Tris HCl pH 7.9, 1 M NaCl, 18 mM *n*-decyl-β-D-maltopyranoside and 10 mM TCEP), and left rotating overnight at 4 °C.

The membrane fraction was centrifuged at 5,000 ×g for 10 min. The supernatant was mixed with approximately 1 ml of Ni-NTA per gram of cells and equilibrated in solubilization buffer (50mM TrisHCl pH 7.9, 200mM NaCl, 5.4mM DM and 1mM TCEP). The sample was bound to the resin by rotating at 4 °C for at least 2 hours. The resin was loaded onto a gravity column and washed with wash buffer (50 mM Tris HCl pH 7.0, 200 mM NaCl, 5.4 mM DM, 1mM TCEP and 30 mM Imidazole). The purified protein was eluted with elution buffer (50 mM Tris HCl pH 7.0, 200 mM NaCl, 5.4 mM DM, 1mM TCEP and 200 mM imidazole). Samples were loaded on a 10DG desalting gravity

column to remove the imidazole. The desalted fractions containing protein were then ultracentrifuged at 100,000 ×g for 30 min at 4 °C to remove possible aggregates. The final protein concentration was determined via UV-Vis spectroscopy using calculated extinction coefficients (supplementary Table S2.7.3).

2.5.4. Labeling of chimeras with Cy3-maleimide or Cy5-maleimide

The constructs were labeled on-column with either Cy3-maleimide or Cy5-maleimide. First, the protein samples were mixed with Ni-NTA equilibrated with buffer (50mM TrisHCl pH 7.0, 200 mM NaCl, 5.4mM DM and 1mM TCEP) and bound in batch at 4 °C for at least 2 hours. The resin was loaded on a gravity flow column and the flow through was collected. The resin was then washed with nonreducing wash buffer (50 mM TrisHCl pH 7.0, 150 mM NaCl, and 5.4 mM DM). The resin was then mixed with labeling buffer (50 mM Tris HCl pH 7.0, 150 mM NaCl, 5.4 mM DM and 0.1 mM fluorophore) using a 15-fold excess of fluorophore to protein, and incubated for 15 min. Samples were quenched for 5 min by mixing 100 mM L-cysteine in a 100-fold excess with respect to fluorophore. The column was washed with 75 ml of labeling wash buffer (50 mM Tris HCl pH 7.0, 150 mM NaCl, 5.4 mM DM and 30 mM Imidazole) and the constructs were eluted using labeling elution buffer (50 mM Tris HCl pH 7.0, 150 mM NaCl, 5.4 mM DM and 200 mM Imidazole). Purity was assessed by SDS-PAGE.

The most concentrated fractions were collected and dialyzed twice in 200 ml of dialysis buffer (50 mM Tris HCl pH 7.0, 150 mM NaCl, and 5.4 mM DM) using a dialysis bag with a molecular weight cutoff of 12-14 kD. The samples were then ultracentrifuged at 100,000 ×g for 30 min at 4 °C to get rid of possible aggregates. The absorption

spectra of the samples was measured to calculate the protein concentration P using the following equation:

$$P = \frac{A_{280} - (A_{cy} CF)}{\epsilon_{280} d}$$

where A_{280} represents the absorbance at 280 nm, A_{cy} correspond to the absorbance at 553 nm for Cy3-maleimide or at 653 nm for Cy5-maleimide, ϵ_{280} is the molar absorptivity for the corresponding construct, d is the cuvette path length (1 cm), and CF is the correction factor to account for fluorophore absorbance at 280 nm (0.04 for Cy5 and 0.1 for Cy3). To calculate the labeling efficiency E_{cy} the following equation was used:

$$E_{cy} = \frac{A_{cy}}{[protein] \epsilon_{cy} d}$$

where ϵ_{cy} represents the molar absorptivity for Cy3 ($150,000 \text{ M}^{-1}\text{cm}^{-1}$) or Cy5 ($250,000 \text{ M}^{-1}\text{cm}^{-1}$). The labeling efficiencies for all constructs are reported in supplementary Table S2.7.4.

2.5.5. Sample preparation for FRET experiments in detergent

Samples at different mole fractions (moles of protein/moles of detergent), were prepared by mixing 1.3 μM of Cy5-maleimide labeled protein with 1.3 μM of Cy3-maleimide labeled protein for a constant protein concentration of 2.6 μM in a buffer containing 50 mM TrisHCl pH 7.0, 150 mM NaCl, and varying concentrations of DM detergent such that mole fraction concentrations from $\sim 1 \times 10^{-3}$ to $\sim 1 \times 10^{-5}$ were covered. In addition equivalent donor-only and acceptor-only samples were produced, each with a protein concentration of 1.3 μM . Samples were equilibrated at 4 °C in dark tubes.

Fluorescence emission spectra were collected at various time points to check sample equilibration.

2.5.6. FRET experiments

Emission spectra measurements were collected in a Tecan infinite M1000 pro plate reader. Samples were excited at 500 nm and emission spectra were collected from 505 nm to 800 nm every 1 nm step. The acceptor-only sample emission spectra were subtracted from the FRET samples emission spectra. Then the data was normalized by dividing every data point by the isosbestic point (640 nm) (supplementary Figure S2.1). Equilibrium constants of dissociation were obtained by fitting the normalized fluorescence intensity at 670 nm (NF_{670}) as a function of the protein: detergent mole fraction (χ_T) using equation 1 (derived in supplementary material Derivation S2.1):

$$NF_{670} = NF_D + (NF_M - NF_D) \frac{-K_D + \sqrt{K_D^2 + 8\chi_T K_D}}{4\chi_T} \quad (\text{equation 1})$$

Three parameters were derived from the fitting: NF_D , the normalized fluorescence intensity at the limit in which the samples are fully dimeric (i.e., at infinite concentration), NF_M , the normalized fluorescence baseline at the limit in which the samples are fully monomeric (at infinite dilution), and K_D , the desired dissociation equilibrium constant. The NF_M and NF_D parameters were fit as global variables, whereas each K_D , was fit individually to their respective construct's data.

The fraction of dimer ϕ_D as a function of protein: detergent mole fraction χ_T and dissociation constant K_D was calculated with the following equation:

$$\phi_D = \frac{\chi_D}{\chi_T} = 1 - \frac{-K_D + \sqrt{K_D^2 + 8K_D\chi_T}}{4\chi_T}$$

The dissociation free energy was determined from K_D using the following equation:

$$\Delta G^\circ = -RT \ln(K_D)$$

where R is the gas constant, and T is the temperature. Statistical analysis was performed using the R software (R Core Team, 2020), with the aid of the tidyverse (Wickham et al., 2019) and drc (Ritz et al., 2015) packages.

2.5.7. Determination of the oligomeric state

Stoichiometry experiments were performed as previously described (Adair and Engelman, 1994) to confirm the oligomeric state of the constructs. Briefly Cy3-maleimide and Cy5-maleimide labeled protein samples were mixed at different donor:acceptor ratios from 10:90 to 90:10 with 10% increments while keeping the total protein concentration constant. The same was done for Cy3-maleimide and unlabeled protein samples. All samples were prepared at a mole fraction of 4.8×10^{-4} . The relative donor quenching Q was obtained using the following equation:

$$Q = 1 - \frac{F}{F_0}$$

where F is the experimental quenched fluorescence for a specific donor:acceptor ratio, and F_0 is the experimental unquenched fluorescence for the same amount of donor. To estimate the oligomeric state, the relative donor quenching Q was plotted as a function of donor fraction P_D , and fit to the following equation:

$$Q(P_D) = k(1 - P_D^{n-1})$$

to obtain the oligomeric state n . The parameter k is defined by the following equation:

$$k = \left(1 - \frac{f_Q}{f_D}\right)$$

where f_Q represents the molar fluorescence of the quenched donor, and f_D represents the molar fluorescence of the donor in the absence of acceptor. Separate fits were performed for different integer values of n and the oligomeric state was evaluated by examining the residuals of the fits.

2.6. Acknowledgments

This work was supported by National Institutes of Health (NIH) grant R35-GM130339 to A.S. and National Science Foundation grants CHE-1710182 to A.S. and DMS-1661900 to Q.C.. G.D.V. acknowledge the support of a SciMed GRS Fellowship.

2.7. Supplementary Figures and Tables

Table S2.7.1: Initial labeling efficiencies of the constructs used for the screening of the labeling position on Staphylococcal nuclease (SN).

Construct	Cy5-maleimide (%)	Cy3-maleimide (%)
SN-A112C-GpA-TM	79.8	77.1
SN-A112C-GpA-G83I-TM	72.4	70.1
SN-A60C-GpA-TM	65.4	76.4

Table S2.7.2: The found relationship between TOXCAT and fraction of dimer was used to predict the standard free energy of the 26 constructs that we previously studied using TOXCAT and CATM (Predicted ΔG°). The relative free energy in the *E. coli* inner membrane was obtained with respect to poly-Leu GpA ($\Delta\Delta G^\circ$ poly-Leu GpA).

TM	Predicted ΔG°	$\Delta\Delta G^\circ$ poly-Leu GpA
ROMO1	-4.39	0.31
K132L	-4.24	0.15
CAD12	-4.15	0.06
GLPA	-4.08	0
SEM6B	-4.06	-0.02
STAB1	-4.03	-0.05
TNR12	-4.03	-0.05
F159B	-3.96	-0.13
CP8B1	-3.77	-0.32
1A32	-3.75	-0.33
OPA1	-3.64	-0.45
ITA2	-3.54	-0.54
SHSA7	-3.39	-0.69
EPCR	-3.27	-0.82
FCRL4	-3.27	-0.82
MGT5B	-3.21	-0.88
1A31	-3.08	-1.01
TNR1B	-3.05	-1.03
STBD1	-3.05	-1.03
MUC3A	-2.71	-1.37
1A32-2	-2.67	-1.41
COX14-2	-2.67	-1.41
1A31-2	-2.53	-1.56
TNR1B2	-2.47	-1.61
MCL1	-2.47	-1.61
ANPRC	-2.41	-1.67

Table S2.7.3: The amino acid sequence of each of the constructs used is presented, with the TM region in bold. The extinction coefficient used to determine the protein concentration for each of the constructs is also presented.

Plasmid name	Sequence	Molar extinction coefficient ($M^{-1}cm^{-1}$)
pSN-GpA	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LIIIFGVMA IVIGTILLISYGIRNPSEFKSSVDKLAAL EHHHHHH*	17,420
pSN-GpA-G83I	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LIIIFGVMA IVIGTILLISYGIRNPSEFKSSVDKLAAL EHHHHHH*	17,420
pSN-CP8B1	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIMVLLG ALLGALLVV LIGIRNPSEFKSSVDKLAAL	15930
pSN-K132L	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIWLLL G ALLGALLAV LIGIRNPSEFKSSVDKLAAL	21430
pSN-1A32	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIVLLL A MLLGALLAAL IIGIRNPSEFKSSVDKLAAL	15930
pSN-ITA2	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIVLLL G VLLGSLLAGL IIGIRNPSEFKSSVDKLAAL	15930
pSN-1A31	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIVLLL A VLLGALLAAL IIGIRNPSEFKSSVDKLAAL	15930
pSN-1A32-2	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLLLLLL VLLGALLFAL IIGIRNPSEFKSSVDKLAAL	15930
pSN-MCL1	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVAVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLIAFLLV ALLGALLAY LIGIRNPSEFKSSVDKLAAL	17420
pSN-OPA1	MATSTKKLHKEPATLIKAIDGDTVKLMLYKQQPMTFRLLLVDTPETKHPKGVEKYGPEAS AFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGMVNEALVRQGLAKVCYVYKPNN THEQLLRKSEAQAKKEKLNINWEDNADSGQTSGGGGGGGGGSGGSILGNEIT LLLYLLL G SLLGGLLTAL IIGIRNPSEFKSSVDKLAAL	17420

Table S2.7.4: Calculated labeling efficiencies for all of the different constructs.

Construct	Cy5 (%)	Cy3 (%)
SN-GpA	68	87
SN-G83I	77.4	79.9
SN-ITA2	74	80
SN-ITA2	89.6	94.7
SN-K132L	76	77
SN-K132L	73.5	76.6
SN-OPA1	71	85
SN-OPA1	71	82.5
SN-1A32-2	92.9	94.6
SN-1A32	90	97
SN-MCL1	60	68
SN-CP8B1	72	78
SN-1A31	72	85

Derivation S2.1: Derivation of the equation used for global fitting of the fluorescence data to obtain the dissociation constants.

To derive the equation to fit the normalized fluorescence intensity at 670 nm (NF_{670}) as a function of mole fraction to determine the equilibrium dissociation constant K_D , we started from the following relationships:

$$(1) \quad K_D = \frac{\chi_M^2}{\chi_D}$$

$$(2) \quad \chi_T = \chi_M + 2\chi_D$$

$$(3) \quad F_T = C_P \left[F_M \left(\frac{\chi_M}{\chi_T} \right) + 2 F_D \left(\frac{\chi_D}{\chi_T} \right) \right]$$

Where χ_M is the monomer mole fraction, χ_D is the dimer mole fraction, χ_T is the total protein mole fraction, F_T is the total fluorescence, F_D is the fluorescence intensity of a fully dimeric sample, F_M is the fluorescence intensity of a fully monomeric sample and C_P is the total protein concentration used in the sample.

Rearranging (1) for χ_D we get:

$$(4) \quad \chi_D = \frac{\chi_M^2}{K_D}$$

Using (4), we eliminate χ_D from (2):

$$(5) \quad \chi_T = \chi_M + \frac{2\chi_M^2}{K_D}$$

Rearranging (5) we get:

$$(6) \quad 2\chi_M^2 + K_D \chi_M - K_D \chi_T = 0$$

Solving the quadratic equation (6) for χ_M we get:

$$(7) \quad \chi_M = \frac{-K_D + \sqrt{K_D^2 + 8K_D \chi_T}}{4}$$

Rearranging (2) for χ_D we get:

$$(8) \quad \chi_D = \frac{\chi_T - \chi_M}{2}$$

Using (8), we eliminate χ_D from (3):

$$(9) \quad F_T = C_P \left[F_M \left(\frac{\chi_M}{\chi_T} \right) + F_D \left(\frac{\chi_T - \chi_M}{\chi_T} \right) \right]$$

Rearranging (9) we get:

$$(10) \quad F_T = C_P \left[(F_M - F_D) \frac{\chi_M}{\chi_T} + F_D \right]$$

Using (7), we eliminate χ_M from (10):

$$(11) \quad F_T = C_P \left[(F_M - F_D) \frac{-K_D + \sqrt{K_D^2 + 8\chi_T K_D}}{4\chi_T} + F_D \right]$$

At the isosbestic point (640 nm, in our case) the fluorescence of the donor and acceptor samples are equal by definition, therefore:

$$(12) \quad F_{M, iso} = F_{D, iso}$$

and thus, at this wavelength, the quadratic term drops and the total fluorescence in (11) simplifies to:

$$(13) \quad F_{T,iso} = C_P F_{D,iso}$$

Dividing the fluorescence of the wavelength of interest (11) by the fluorescence at the isosbestic point (13) cancels out the protein concentration C_P , resulting in the equation:

$$(14) \quad NF_{670} = \frac{F_{D,670}}{F_{D,iso}} + \left(\frac{F_{M,670} - F_{D,670}}{F_{D,iso}} \right) \frac{-K_D + \sqrt{K_D^2 + 8\chi_T K_D}}{4\chi_T}$$

where NF_{670} refers to the normalized fluorescence intensity at the wavelength of interest (670 nm, in our case).

If we now redefine the normalized fluorescence of the fully monomeric (NF_M) and dimeric samples (NF_D) as the ratio of the fluorescence at 670 nm and the fluorescence at the isosbestic point:

$$(15) \quad NF_M = \frac{F_{M,670}}{F_{D,iso}}$$

$$(16) \quad NF_D = \frac{F_{D,670}}{F_{D,iso}}$$

and substitute (15) and (16) in (14), we obtain the final form of the equation (equation 1):

$$(17) \quad NF_{670} = NF_D + (NF_M - NF_D) \frac{-K_D + \sqrt{K_D^2 + 8\chi_T K_D}}{4\chi_T}$$

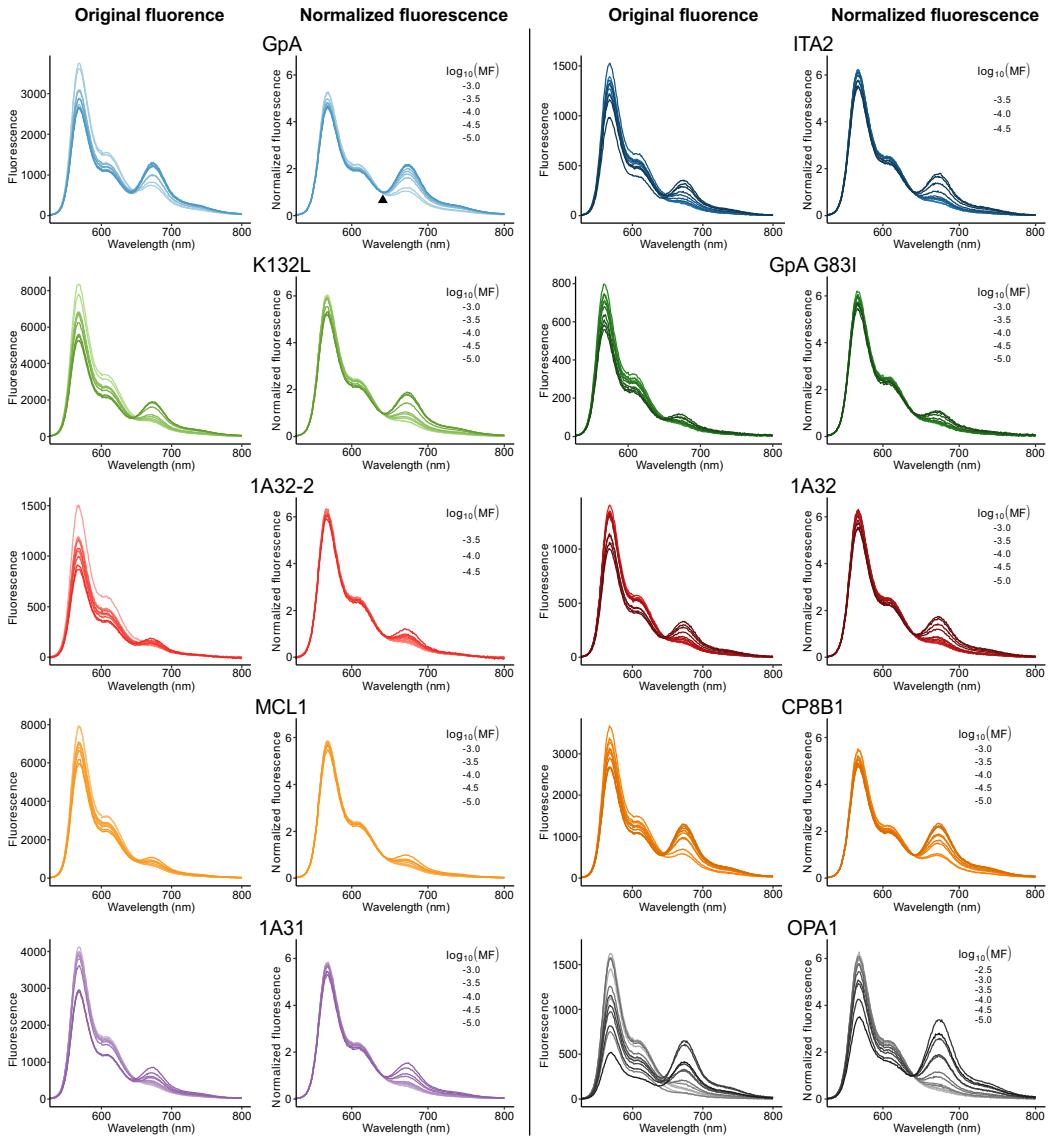


Figure S2.1: The original emission spectra of the different samples and their normalized version using the isosbestic point (640 nm for the FRET pair used). The normalization was introduced to correct for small changes in signal that are unrelated to the transfer of energy between the donor and acceptor fluorophores, such as small errors in the estimate of concentration.

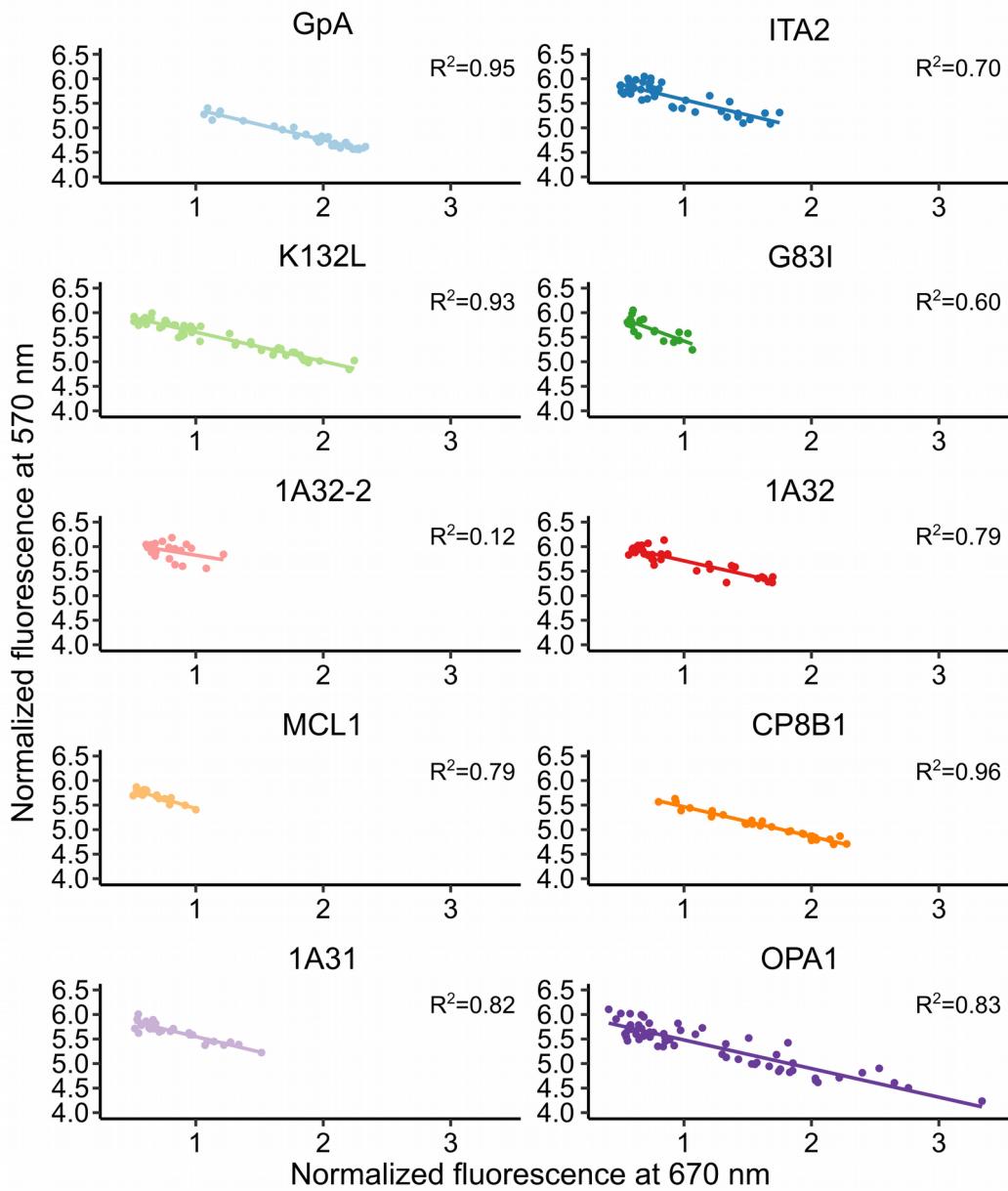


Figure S2.2: FRET data was validated by plotting acceptor emission at 670 nm vs donor quenching at 570 nm. The expected inverse relation is observed. OPA1 displayed a larger emission range when compared to the other constructs suggesting it may be forming higher order oligomers.

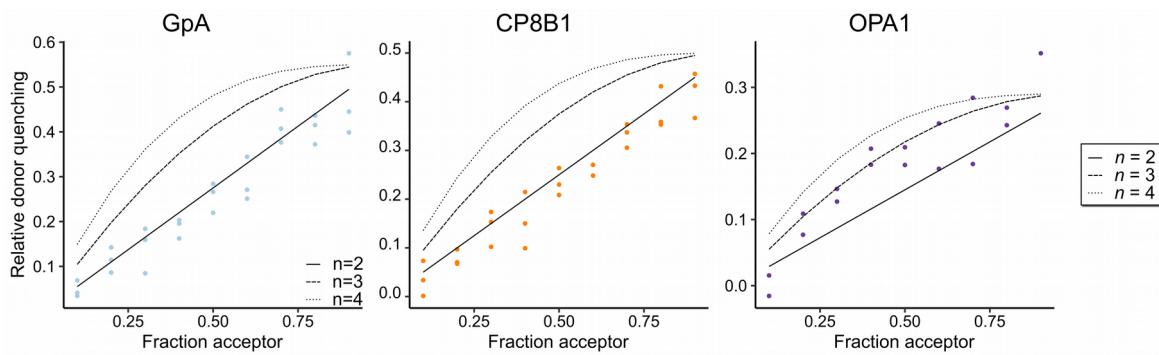


Figure S2.3: Stoichiometry experiments for the two outlier constructs CP8B1 and OPA1, and the control GpA. The experiments were carried out by varying the donor-labeled:acceptor-labeled construct ratio from 10:90 to 90:10 with 10% increments at a mole fraction of 4.8×10^{-4} . Both GpA and CP8B1 fit well to a linear model, indicating they associate as dimers ($n=2$). OPA1 fit better to a $n=3$ model indicating it is forming a higher-oligomer, likely a trimer.

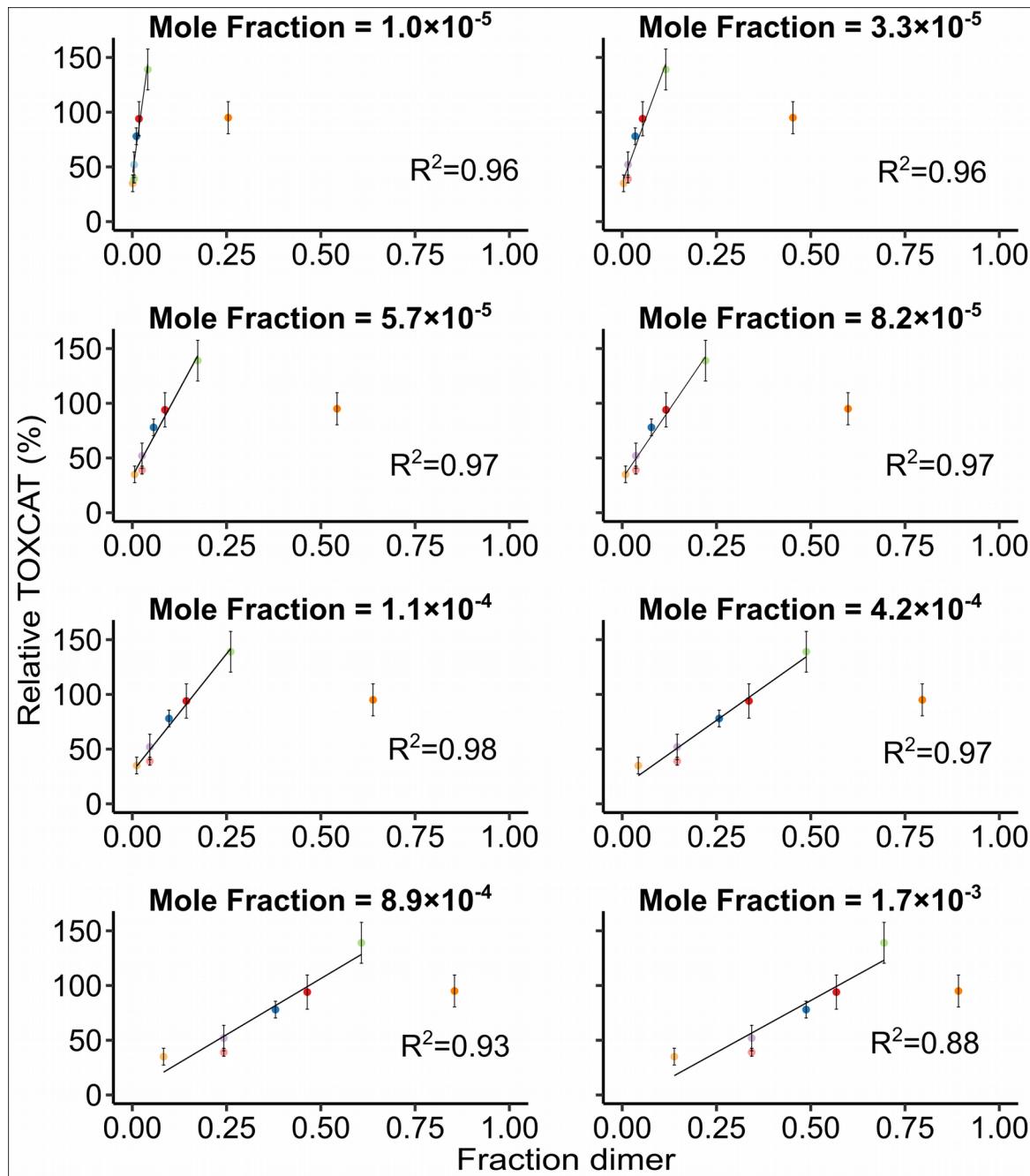


Figure S2.4: Comparison of the FRET results with the previous TOXCAT analysis, linear regression. Results of the linear regression analysis between TOXCAT and

fraction of dimer calculated at different protein: detergent mole fractions to identify the concentration regime at which the linear relationship between dimer fraction in detergent and TOXCAT (taken as proxy for dimer fraction in the membrane) is maximized. The clear outlier CP8B1 was not included in this linear fit.

2.8. References

- Adair, B.D., and Engelman, D.M. (1994). Glycophorin A Helical Transmembrane Domains Dimerize in Phospholipid Bilayers: A Resonance Energy Transfer Study. *Biochemistry* 33, 5539–5544. <https://doi.org/10.1021/bi00184a024>.
- Adamian, L., and Liang, J. (2001). Helix-helix packing and interfacial pairwise interactions of residues in membrane proteins. *J. Mol. Biol.* 311, 891–907. <https://doi.org/10.1006/jmbi.2001.4908>.
- Anbazhagan, V., and Schneider, D. (2010). The membrane environment modulates self-association of the human GpA TM domain—Implications for membrane protein folding and transmembrane signaling. *Biochim. Biophys. Acta BBA - Biomembr.* 1798, 1899–1907. <https://doi.org/10.1016/j.bbamem.2010.06.027>.
- Anderson, S.M., Mueller, B.K., Lange, E.J., and Senes, A. (2017). Combination of Ca–H Hydrogen Bonds and van der Waals Packing Modulates the Stability of GxxxG-Mediated Dimers in Membranes. *J. Am. Chem. Soc.* 139, 15774–15783. <https://doi.org/10.1021/jacs.7b07505>.
- Arbely, E., and Arkin, I.T. (2004). Experimental measurement of the strength of a C alpha-H...O bond in a lipid bilayer. *J. Am. Chem. Soc.* 126, 5362–5363. <https://doi.org/10.1021/ja049826h>.
- Baker, R.P., and Urban, S. (2012). Architectural and thermodynamic principles underlying intramembrane protease function. *Nat. Chem. Biol.* 8, 759–768. <https://doi.org/10.1038/nchembio.1021>.
- Bocharov, E.V., Mineev, K.S., Volynsky, P.E., Ermolyuk, Y.S., Tkach, E.N., Sobol, A.G., Chupin, V.V., Kirpichnikov, M.P., Efremov, R.G., and Arseniev, A.S. (2008). Spatial Structure of the Dimeric Transmembrane Domain of the Growth Factor Receptor ErbB2 Presumably Corresponding to the Receptor Active State. *J. Biol. Chem.* 283, 6950–6956. <https://doi.org/10.1074/jbc.M709202200>.
- Bocharov, E.V., Mineev, K.S., Goncharuk, M.V., and Arseniev, A.S. (2012). Structural and thermodynamic insight into the process of “weak” dimerization of the ErbB4

transmembrane domain by solution NMR. *Biochim. Biophys. Acta BBA - Biomembr.* 1818, 2158–2170. <https://doi.org/10.1016/j.bbamem.2012.05.001>.

Brosig, B., and Langosch, D. (1998). The dimerization motif of the glycophorin A transmembrane segment in membranes: Importance of glycine residues: Glycophorin A dimerization motif. *Protein Sci.* 7, 1052–1056. <https://doi.org/10.1002/pro.5560070423>.

Bu, Z., and Engelman, D.M. (1999). A method for determining transmembrane helix association and orientation in detergent micelles using small angle x-ray scattering. *Biophys. J.* 77, 1064–1073. [https://doi.org/10.1016/S0006-3495\(99\)76956-0](https://doi.org/10.1016/S0006-3495(99)76956-0).

Chadda, R., Bernhardt, N., Kelley, E.G., Teixeira, S.C., Griffith, K., Gil-Ley, A., Öztürk, T.N., Hughes, L.E., Forsythe, A., Krishnamani, V., et al. (2021). Membrane transporter dimerization driven by differential lipid solvation energetics of dissociated and associated states. *eLife* 10, e63288. <https://doi.org/10.7554/eLife.63288>.

Chen, L., Novicky, L., Merzlyakov, M., Hristov, T., and Hristova, K. (2010). Measuring the Energetics of Membrane Protein Dimerization in Mammalian Membranes. *J. Am. Chem. Soc.* 132, 3628–3635. <https://doi.org/10.1021/ja910692u>.

Choma, C., Gratkowski, H., Lear, J.D., and DeGrado, W.F. (2000). Asparagine-mediated self-association of a model transmembrane helix. *Nat. Struct. Biol.* 7, 161–166..

Condon, S.G.F., Mahbuba, D.-A., Armstrong, C.R., Diaz-Vazquez, G., Craven, S.J., LaPointe, L.M., Khadria, A.S., Chadda, R., Crooks, J.A., Rangarajan, N., et al. (2018). The FtsLB subcomplex of the bacterial divisome is a tetramer with an uninterrupted FtsL helix linking the transmembrane and periplasmic regions. *J. Biol. Chem.* 293, 1623–1641. <https://doi.org/10.1074/jbc.RA117.000426>.

Cymer, F., and Schneider, D. (2010). Transmembrane helix-helix interactions involved in ErbB receptor signaling. *Cell Adhes. Migr.* 4, 299–312. .

Cymer, F., Veerappan, A., and Schneider, D. (2012). Transmembrane helix-helix interactions are modulated by the sequence context and by lipid bilayer properties. *Biochim. Biophys. Acta* 1818, 963–973. <https://doi.org/10.1016/j.bbamem.2011.07.035>.

Doura, A.K., and Fleming, K.G. (2004). Complex Interactions at the Helix–Helix Interface Stabilize the Glycophorin A Transmembrane Dimer. *J. Mol. Biol.* 343, 1487–1497. <https://doi.org/10.1016/j.jmb.2004.09.011>.

Duneau, J.-P., Vegh, A.P., and Sturgis, J.N. (2007). A Dimerization Hierarchy in the Transmembrane Domains of the HER Receptor Family. *Biochemistry* 46, 2010–2019. <https://doi.org/10.1021/bi061436f>.

- Duneau, J.-P., Khao, J., and Sturgis, J.N. (2017). Lipid perturbation by membrane proteins and the lipophobic effect. *Biochim. Biophys. Acta Biomembr.* 1859, 126–134. <https://doi.org/10.1016/j.bbamem.2016.10.014>.
- Duong, M.T., Jaszewski, T.M., Fleming, K.G., and MacKenzie, K.R. (2007). Changes in Apparent Free Energy of Helix–Helix Dimerization in a Biological Membrane Due to Point Mutations. *J. Mol. Biol.* 371, 422–434. <https://doi.org/10.1016/j.jmb.2007.05.026>.
- van den Ent, F., and Löwe, J. (2006). RF cloning: A restriction-free method for inserting target genes into plasmids. *J. Biochem. Biophys. Methods* 67, 67–74. <https://doi.org/10.1016/j.jbbm.2005.12.008>.
- Fisher, L.E., Engelman, D.M., and Sturgis, J.N. (1999). Detergents modulate dimerization, but not helicity, of the glycophorin A transmembrane domain 1 Edited by G. von Heijne. *J. Mol. Biol.* 293, 639–651. <https://doi.org/10.1006/jmbi.1999.3126>.
- Fleming, K. (2004). Thermodynamics of glycophorin A transmembrane helix dimerization in C14 betaine micelles. *Biophys. Chem.* [https://doi.org/10.1016/S0301-4622\(03\)00296-5](https://doi.org/10.1016/S0301-4622(03)00296-5).
- Fleming, K.G. (2002). Standardizing the free energy change of transmembrane helix-helix interactions. *J. Mol. Biol.* 323, 563–571. .
- Fleming, K.G., Ackerman, A.L., and Engelman, D.M. (1997). The effect of point mutations on the free energy of transmembrane α -helix dimerization. *J. Mol. Biol.* 272, 266–275. <https://doi.org/10.1006/jmbi.1997.1236>.
- Gratkowski, H., Lear, J.D., and DeGrado, W.F. (2001). Polar side chains drive the association of model transmembrane peptides. *Proc. Natl. Acad. Sci. U. S. A.* 98, 880–885. <https://doi.org/10.1073/pnas.98.3.880>.
- Guo, R., Gaffney, K., Yang, Z., Kim, M., Sungsuwan, S., Huang, X., Hubbell, W.L., and Hong, H. (2016). Steric trapping reveals a cooperativity network in the intramembrane protease GlpG. *Nat. Chem. Biol.* 12, 353–360. <https://doi.org/10.1038/nchembio.2048>.
- Hessels, A.M., and Merkx, M. (2016). Simple Method for Proper Analysis of FRET Sensor Titration Data and Intracellular Imaging Experiments Based on Isosbestic Points. *ACS Sens.* 1, 498–502. <https://doi.org/10.1021/acssensors.6b00078>.
- Hong, H. (2014). Toward understanding driving forces in membrane protein folding. *Arch. Biochem. Biophys.* 564, 297–313. <https://doi.org/10.1016/j.abb.2014.07.031>.
- Hong, H., and Bowie, J.U. (2011). Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *J. Am. Chem. Soc.* 133, 11389–11398. <https://doi.org/10.1021/ja204524c>.

- Hong, H., Park, S., Jiménez, R.H.F., Rinehart, D., and Tamm, L.K. (2007). Role of aromatic side chains in the folding and thermodynamic stability of integral membrane proteins. *J. Am. Chem. Soc.* 129, 8320–8327. <https://doi.org/10.1021/ja068849o>.
- Hong, H., Blois, T.M., Cao, Z., and Bowie, J.U. (2010). Method to measure strong protein-protein interactions in lipid bilayers using a steric trap. *Proc. Natl. Acad. Sci. U. S. A.* 107, 19802–19807. <https://doi.org/10.1073/pnas.1010348107>.
- Javadpour, M.M., Eilers, M., Groesbeek, M., and Smith, S.O. (1999). Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys. J.* 77, 1609–1618. [https://doi.org/10.1016/S0006-3495\(99\)77009-8](https://doi.org/10.1016/S0006-3495(99)77009-8).
- Joh, N.H., Oberai, A., Yang, D., Whitelegge, J.P., and Bowie, J.U. (2009). Similar Energetic Contributions of Packing in the Core of Membrane and Water-Soluble Proteins. *J. Am. Chem. Soc.* 131, 10846–10847. <https://doi.org/10.1021/ja904711k>.
- Johnson, R.M., Hecht, K., and Deber, C.M. (2007). Aromatic and cation-pi interactions enhance helix-helix association in a membrane environment. *Biochemistry* 46, 9208–9214. <https://doi.org/10.1021/bi7008773>.
- Kim, Y.H., and Stites, W.E. (2008). Effects of Excluded Volume upon Protein Stability in Covalently Cross-Linked Proteins with Variable Linker Lengths. *Biochemistry* 47, 8804–8814. <https://doi.org/10.1021/bi800297j>.
- Krivov, G.G., Shapovalov, M.V., and Dunbrack, R.L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins* 77, 778–795. <https://doi.org/10.1002/prot.22488>.
- Langosch, D., and Heringa, J. (1998). Interaction of transmembrane helices by a knobs-into-holes packing characteristic of soluble coiled coils. *Proteins* 31, 150–159. [https://doi.org/10.1002/\(sici\)1097-0134\(19980501\)31:2<150::aid-prot5>3.0.co;2-q](https://doi.org/10.1002/(sici)1097-0134(19980501)31:2<150::aid-prot5>3.0.co;2-q).
- Lazaridis, T. (2003). Effective energy function for proteins in lipid membranes. *Proteins* 52, 176–192. <https://doi.org/10.1002/prot.10410>.
- Lemmon, M.A., Flanagan, J.M., Hunt, J.F., Adair, B.D., Bormann, B.J., Dempsey, C.E., and Engelman, D.M. (1992a). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J. Biol. Chem.* 267, 7683–7689. [https://doi.org/10.1016/S0021-9258\(18\)42569-0](https://doi.org/10.1016/S0021-9258(18)42569-0).
- Lemmon, M.A., Flanagan, J.M., Treutlein, H.R., Zhang, J., and Engelman, D.M. (1992b). Sequence specificity in the dimerization of transmembrane .alpha.-helices. *Biochemistry* 31, 12719–12725. <https://doi.org/10.1021/bi00166a002>.
- Li, W., Metcalf, D.G., Gorelik, R., Li, R., Mitra, N., Nanda, V., Law, P.B., Lear, J.D., DeGrado, W.F., and Bennett, J.S. (2005). A push-pull mechanism for regulating integrin

function. Proc. Natl. Acad. Sci. 102, 1424–1429.
<https://doi.org/10.1073/pnas.0409334102>.

MacKenzie, K.R., and Engelman, D.M. (1998). Structure-based prediction of the stability of transmembrane helix-helix interactions: the sequence dependence of glycophorin A dimerization. Proc. Natl. Acad. Sci. U. S. A. 95, 3583–3590. .

MacKenzie, K.R., Prestegard, J.H., and Engelman, D.M. (1997). A Transmembrane Helix Dimer: Structure and Implications. Science 276, 131–133. <https://doi.org/10.1126/science.276.5309.131>.

MacKerell, Bashford, D., Bellott, Dunbrack, Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins†. J. Phys. Chem. B 102, 3586–3616. <https://doi.org/10.1021/jp973084f>.

Maeda, R., Sato, T., Okamoto, K., Yanagawa, M., and Sako, Y. (2018). Lipid-Protein Interplay in Dimerization of Juxtamembrane Domains of Epidermal Growth Factor Receptor. Biophys. J. 114, 893–903. <https://doi.org/10.1016/j.bpj.2017.12.029>.

Marsh, D. (2008). Protein modulation of lipids, and vice-versa, in membranes. Biochim. Biophys. Acta 1778, 1545–1575. <https://doi.org/10.1016/j.bbamem.2008.01.015>.

Mineev, K.S., Bocharov, E.V., Pustovalova, Y.E., Bocharova, O.V., Chupin, V.V., and Arseniev, A.S. (2010). Spatial Structure of the Transmembrane Domain Heterodimer of ErbB1 and ErbB2 Receptor Tyrosine Kinases. J. Mol. Biol. 400, 231–243. <https://doi.org/10.1016/j.jmb.2010.05.016>.

Mravic, M., Thomaston, J.L., Tucker, M., Solomon, P.E., Liu, L., and DeGrado, W.F. (2019). Packing of apolar side chains enables accurate design of highly stable membrane proteins. Science 363, 1418–1423. <https://doi.org/10.1126/science.aav7541>.

Mueller, B.K., Subramaniam, S., and Senes, A. (2014). A frequent, GxxxG-mediated, transmembrane association motif is optimized for the formation of interhelical C α -H hydrogen bonds. Proc. Natl. Acad. Sci. 111, E888–E895. <https://doi.org/10.1073/pnas.1319944111>.

R Core Team (2020). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Ritz, C., Baty, F., Streibig, J.C., and Gerhard, D. (2015). Dose-Response Analysis Using R. PLOS ONE 10. .

Russ, W.P., and Engelman, D.M. (1999). TOXCAT: A measure of transmembrane helix association in a biological membrane. Proc. Natl. Acad. Sci. 96, 863–868. <https://doi.org/10.1073/pnas.96.3.863>.

- Russ, W.P., and Engelman, D.M. (2000). The GxxxG motif: A framework for transmembrane helix-helix association. *J. Mol. Biol.* 296, 911–919. <https://doi.org/10.1006/jmbi.1999.3489>.
- Sal-Man, N., Gerber, D., Bloch, I., and Shai, Y. (2007). Specificity in transmembrane helix-helix interactions mediated by aromatic residues. *J. Biol. Chem.* 282, 19753–19761. <https://doi.org/10.1074/jbc.M610368200>.
- Scheiner, S., Kar, T., and Gu, Y. (2001). Strength of the C α H...O hydrogen bond of amino acid residues. *J. Biol. Chem.* 276, 9832–9837. <https://doi.org/10.1074/jbc.M010770200>.
- Senes, A., Gerstein, M., and Engelman, D.M. (2000). Statistical analysis of amino acid patterns in transmembrane helices: the GxxxG motif occurs frequently and in association with β -branched residues at neighboring positions. *J. Mol. Biol.* 296, 921–936. <https://doi.org/10.1006/jmbi.1999.3488>.
- Senes, A., Ubarretxena-Belandia, I., and Engelman, D.M. (2001). The C α ---H...O hydrogen bond: a determinant of stability and specificity in transmembrane helix interactions. *Proc. Natl. Acad. Sci. U. S. A.* 98, 9056–9061. .
- Soubias, O., Teague, W.E., Hines, K.G., and Gawrisch, K. (2015). Rhodopsin/lipid hydrophobic matching-rhodopsin oligomerization and function. *Biophys. J.* 108, 1125–1132. <https://doi.org/10.1016/j.bpj.2015.01.006>.
- Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234. <https://doi.org/10.1016/j.pep.2005.01.016>.
- Vargas, R., Garza, J., Dixon, D.A., and Hay, B.P. (2000). How Strong Is the C α -H...OC Hydrogen Bond? *J Am Chem Soc* 122, 4750–4755. <https://doi.org/10.1021/ja993600a>.
- Walters, R.F.S., and DeGrado, W.F. (2006). Helix-packing motifs in membrane proteins. *Proc. Natl. Acad. Sci.* 103, 13658–13663. <https://doi.org/10.1073/pnas.0605878103>.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Gromelund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the tidyverse. *J. Open Source Softw.* 4, 1686. <https://doi.org/10.21105/joss.01686>.
- Yano, Y., Kondo, K., Kitani, R., Yamamoto, A., and Matsuzaki, K. (2015). Cholesterol-induced lipophobic interaction between transmembrane helices using ensemble and single-molecule fluorescence resonance energy transfer. *Biochemistry* 54, 1371–1379. <https://doi.org/10.1021/bi501528e>.
- Yin, H., Litvinov, R.I., Vilaire, G., Zhu, H., Li, W., Caputo, G.A., Moore, D.T., Lear, J.D., Weisel, J.W., DeGrado, W.F., et al. (2006). Activation of Platelet $\text{c}\text{llb}\beta 3$ by an

Exogenous Peptide Corresponding to the Transmembrane Domain of $\alpha 1\text{b}$. *J. Biol. Chem.* 281, 36732–36741. <https://doi.org/10.1074/jbc.M605877200>.

Yohannan, S., Faham, S., Yang, D., Grosfeld, D., Chamberlain, A.K., and Bowie, J.U. (2004). A C alpha-H...O hydrogen bond in a membrane protein is not stabilizing. *J. Am. Chem. Soc.* 126, 2284–2285. <https://doi.org/10.1021/ja0317574>.

You, M., Li, E., Wimley, W.C., and Hristova, K. (2005). Förster resonance energy transfer in liposomes: Measurements of transmembrane helix dimerization in the native bilayer environment. *Anal. Biochem.* 340, 154–164.
<https://doi.org/10.1016/j.ab.2005.01.035>.

Zhang, S.-Q., Kulp, D.W., Schramm, C.A., Mravic, M., Samish, I., and DeGrado, W.F. (2015a). The Membrane- and Soluble-Protein Helix-Helix Interactome: Similar Geometry via Different Interactions. *Structure* 23, 527–541.
<https://doi.org/10.1016/j.str.2015.01.009>.

Zhang, S.-Q., Kulp, D.W., Schramm, C.A., Mravic, M., Samish, I., and DeGrado, W.F. (2015b). The membrane- and soluble-protein helix-helix interactome: similar geometry via different interactions. *Struct. Lond. Engl.* 1993 23, 527–541.
<https://doi.org/10.1016/j.str.2015.01.009>.

Zheng, L. (2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* 32, e115–e115. <https://doi.org/10.1093/nar/gnh110>.

Zhou, F.X., Cocco, M.J., Russ, W.P., Brunger, A.T., and Engelman, D.M. (2000). Interhelical hydrogen bonding drives strong interactions in membrane proteins. *Nat. Struct. Biol.* 7, 154–160. .

Zhou, F.X., Merianos, H.J., Brunger, A.T., and Engelman, D.M. (2001). Polar residues drive association of polyleucine transmembrane helices. *Proc. Natl. Acad. Sci. U. S. A.* 98, 2250–2255. <https://doi.org/10.1073/pnas.041593698>.

Chapter 3: The FtsLB subcomplex of the bacterial divisome is a tetramer with an uninterrupted FtsL helix linking the transmembrane and periplasmic regions

This chapter was prepared for publication as:

Condon S.G.F.*¹, Mahbuba D.-A.*¹, Armstrong C.R., Diaz Vazquez G., Craven S.J., LaPointe L.M., Khadria A.S., Chadda R., Crooks J.A., Rangarajan N., Weibel D.B., Hoskins A.A., Robertston J.L., Cui Q., Senes A. 2018. "The FtsLB subcomplex of the bacterial divisome is a tetramer with an uninterrupted FtsL helix linking the transmembrane and periplasmic regions." *J. Biol. Chem.* 293, 1623–1641.

Statement of contribution

I performed the molecular dynamics simulations, worked on their analysis and contributed to writing the corresponding sections.

3.1. Abstract

In *Escherichia coli*, FtsLB plays a central role in the initiation of cell division, possibly transducing a signal that will eventually lead to the activation of peptidoglycan remodeling at the forming septum. The molecular mechanisms by which FtsLB operates in the divisome, however, are not understood. Here, we present a structural analysis of the FtsLB complex, performed with biophysical, computational, and *in vivo* methods, that establishes the organization of the transmembrane region and proximal coiled coil of the complex. FRET analysis *in vitro* is consistent with formation of a tetramer composed of two FtsL and two FtsB subunits. We predicted subunit contacts through co-evolutionary analysis and used them to compute a structural model of the complex. The transmembrane region of FtsLB is stabilized by hydrophobic packing and by a complex network of hydrogen bonds. The coiled coil domain probably terminates near the critical constriction control domain, which might correspond to a structural transition. The presence of strongly polar amino acids within the core of the tetrameric coiled coil suggests that the coil may split into two independent FtsQ-binding domains. The helix of FtsB is interrupted between the transmembrane and coiled coil regions by a flexible Gly-rich linker. Conversely, the data suggest that FtsL forms an uninterrupted helix across the two regions and that the integrity of this helix is indispensable for the function of the complex. The FtsL helix is thus a candidate for acting as a potential mechanical connection to communicate conformational changes between periplasmic, membrane, and cytoplasmic regions.

3.2. Introduction

The cell envelope of Gram-negative bacteria, such as *Escherichia coli*, consists of three layers: a cytoplasmic membrane, an outer membrane, and a rigid cell wall formed by a mesh of peptidoglycan located in the intervening periplasmic space. Cell division requires mechanisms for the separation of these three layers. The first step is the establishment of a division site and the segregation of the duplicated chromosome. This allows for the beginning of constriction, during which the coordinated activities of numerous peptidoglycan synthases and hydrolases remodel the cell wall. Finally, the process leads to membrane fusion and to the separation of two daughter cells. The large multiprotein complex that supports a majority of these functions is called the divisome.

At least 3 dozen proteins are known to participate in cell division in *E. coli*, but the essential components of the divisome consist of a core of 12 proteins. Their recruitment to mid-cell follows a hierarchical order of dependence ($\text{FtsZ} \rightarrow \text{FtsA-ZipA} \rightarrow \text{FtsE-FtsX} \rightarrow \text{FtsK} \rightarrow \text{FtsQ} \rightarrow \text{FtsL-FtsB} \rightarrow \text{FtsW} \rightarrow \text{FtsI} \rightarrow \text{FtsN}$; Figure 3.1a), which also reflects, in part, the timing and interactions that occur in the complex assembly (Buddelmeijer and Beckwith, 2004; Buddelmeijer et al., 2002; Chen and Beckwith, 2001; Ghigo et al., 1999; Goehring, 2005; Goehring et al., 2006; Hale and de Boer, 1999, 2002; Liu et al., 1999; Mercer and Weiss, 2002; Schmidt et al., 2004; Wang et al., 1998). The divisome assembles around the tubulin homolog FtsZ, which forms a ringlike structure at mid-cell (the Z-ring) (de Boer, 2010; Erickson et al., 2010; Lutkenhaus et al., 2012). FtsZ protofilaments treadmill around the circumference of the cell, providing a scaffold for the recruitment and movement of the components around the cell and ultimately leading to

incorporation of new peptidoglycan around the forming septum (Bisson-Filho et al., 2017; Yang et al., 2017). FtsZ is tethered to the plasma membrane by the cooperative action of ZipA, a single-pass transmembrane (TM) protein (Haney et al., 2001; Ortiz et al., 2016), and of FtsA, an actin homolog and peripheral membrane protein able to form protofilaments (Szwedziak et al., 2012). Other early components of the divisome include FtsEX, an ABC transporter-like complex that controls cell wall hydrolysis and possibly divisome assembly (Du et al., 2016; Yang et al., 2011), as well as FtsK, a DNA translocase important for chromosome segregation (Männik et al., 2017).

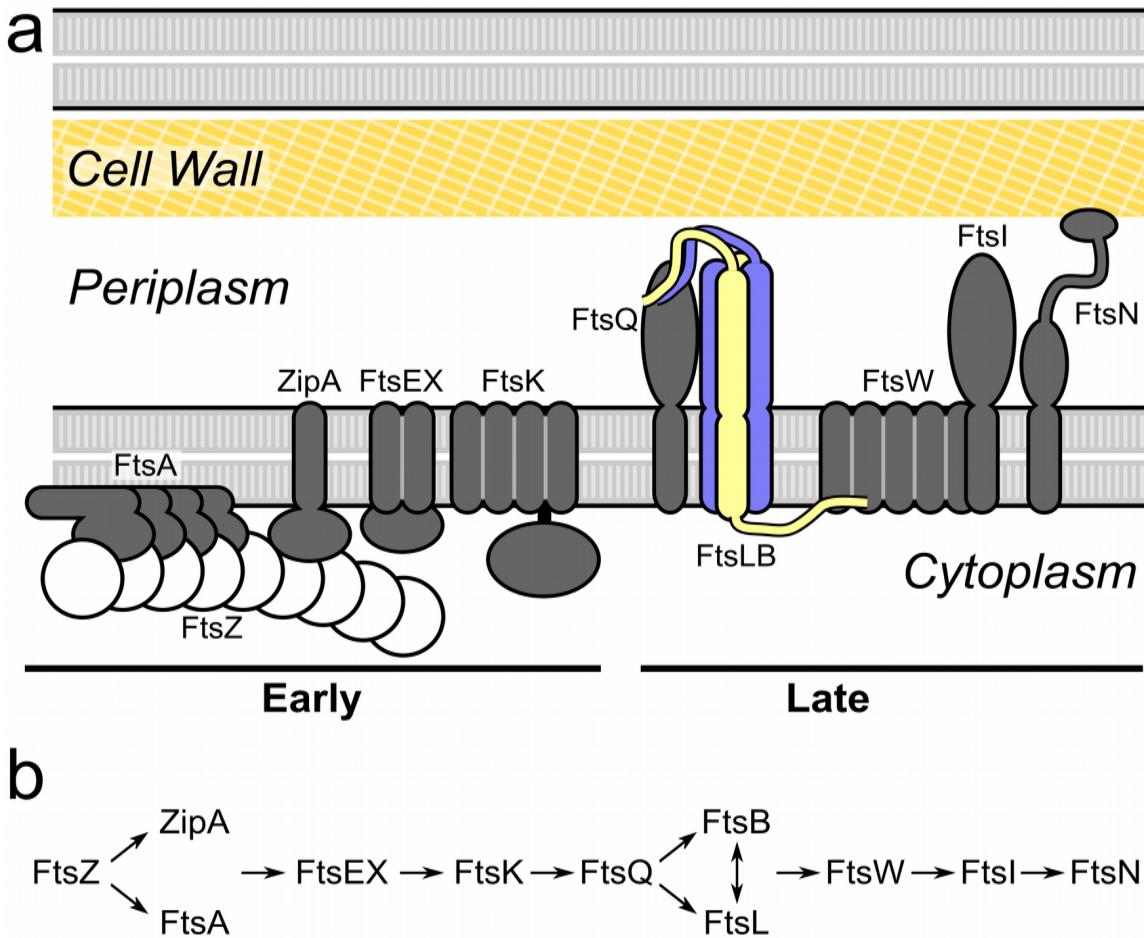


Figure 3.1: The essential proteins of the divisome. *a*, schematic representation of the divisome of *E. coli*. The complex assembles around a polymeric scaffold formed by FtsZ. With a few exceptions (FtsZ, FtsA, and FtsE), all essential components are integral membrane proteins. As reported in this work, the FtsLB complex is a heterotetramer formed by two FtsL subunits (yellow) and two FtsB subunits (blue). The complex forms an extended helical bundle that comprises the transmembrane helices and the periplasmic coiled coil. The C-terminal periplasmic tails of FtsLB mediate the binding to FtsQ. The cytoplasmic tail of FtsL

has been hypothesized to bind to FtsW. Description of the function of the various components is provided in the Introduction. *b*, the divisome of *E. coli* displays a characteristic hierarchy of recruitment at the division site, with the arrows indicating a dependence of a component on the one that precedes it in the sequence. The hierarchy also roughly corresponds to the order of recruitment, which can be subdivided into early and late components.

The late components of the divisome (FtsQ, FtsB, FtsL, FtsW, FtsI, and FtsN) localize approximately at the beginning of constriction (Aarsman et al., 2005). FtsQ, FtsB, and FtsL are three bitopic (single-pass) membrane proteins that form a complex whose exact function is not well understood (Buddelmeijer and Beckwith, 2004). The FtsQLB complex is required for the recruitment of FtsW, a large multispan membrane protein, and FtsI, a bitopic protein. FtsW and FtsI work in coordination to synthesize septal cell wall (Mercer and Weiss, 2002); FtsI is a penicillin-binding protein (PBP3) with transpeptidase activity (Vollmer and Bertsche, 2008), and FtsW is most likely its cognate glycosyltransferase (Cho et al., 2016; Emami et al., 2017; Meeske et al., 2016) (FtsW has also been proposed to be a flippase for peptidoglycan precursors (Mohammadi et al., 2011, 2014)). The last protein in the recruitment hierarchy is FtsN, a bitopic protein with an N-terminal TM domain, a predicted disordered region, and a C-terminal domain capable of recognizing septal peptidoglycan (Yang et al., 2004). The precise role of FtsN is not understood, but its accumulation at the septal ring represents a key event for triggering constriction. This activation may be mediated by some of the interactions that have been postulated for FtsN, which include FtsA, the peptidoglycan

synthase complex (FtsW, FtsI, and PBP1B), and the FtsQLB complex (Alexeeva et al., 2010; Busiek et al., 2012; Di Lallo, 2003; Karimova et al., 2005; Müller et al., 2007).

Because the divisome consists primarily of integral membrane proteins, its structural characterization has been lagging and limited to fragments of water-soluble domains (van den Ent et al., 2008; LaPointe et al., 2013; Massey et al., 2006; Mosyak, 2000; Yang et al., 2004). Here we focus on the structural organization of the FtsL-FtsB subcomplex (FtsLB), which has been implicated as a critical player in triggering constriction (Liu et al., 2015; Tsang and Bernhardt, 2015). Topologically, FtsL and FtsB are both small bitopic proteins with a nearly identical domain organization, which suggests that they may derive from a common evolutionary ancestor. Both proteins have short (or absent) N-terminal cytoplasmic tails, one TM domain, a juxtamembrane coiled coil, and C-terminal tails in the periplasm (Figure 3.1a). Peptides corresponding to the TM helices form a stable higher-order oligomer *in vitro*, with an equal number of FtsL and FtsB subunits, indicating that the TM region is an important contributor to the stability of the complex (Khadria and Senes, 2013). This is consistent with the observation that the TM domains are biologically important (Buddelmeijer and Beckwith, 2004; Gonzalez and Beckwith, 2009; Guzman et al., 1997). FtsL and FtsB associate *in vivo* even in the absence of FtsQ (Goehring et al., 2006; Robichon et al., 2008), although FtsQ is required for their recruitment to mid-cell (Buddelmeijer and Beckwith, 2004). The association with FtsQ to form the FtsQLB complex is primarily mediated by the C-terminal tails of FtsLB, which bind to the C-terminal end of FtsQ, as evidenced by truncation functional analysis (Gonzalez and Beckwith, 2009; Gonzalez et al., 2010) as

well as cross-linking performed *in vivo* (*van den Berg van Saparoea et al., 2013*). The periplasmic domains of the three proteins are sufficient to form soluble complexes with submicromolar binding affinities, as established for *E. coli* (Glas et al., 2015) and *Streptococcus pneumoniae* (Masson et al., 2009; Noirclerc-Savoye et al., 2004) proteins. Notably, these studies have been performed using solubilized FtsLB constructs that were fused to a stable heterodimeric coiled coil and thus forced into a dimeric state. However, FtsLB is likely to be a higher-order oligomer (such as a tetramer), as evidenced by biophysical analysis of their TM helices in isolation (Khadria and Senes, 2013). The tight binding affinity of these 1:1:1 FtsQLB soluble constructs therefore suggests that independent FtsQ-binding sites exist in FtsLB, each formed by one FtsL and one FtsB subunit.

The precise function of FtsLB has not yet been determined. It was originally hypothesized that FtsLB could have a structural role in stabilizing the divisome (Gonzalez and Beckwith, 2009; Gonzalez et al., 2010). The finding that FtsL is subject to active degradation unless it is stabilized by interaction with FtsB in *Bacillus subtilis* (Bramkamp et al., 2006; Daniel and Errington, 2000; Daniel et al., 2006) and *E. coli* (Gonzalez and Beckwith, 2009) suggested that FtsL levels may be rate-limiting for division (Bramkamp et al., 2006). More recent evidence indicates that FtsLB is an active participant in the decision making that controls the cell division process, playing a central role in triggering of septal peptidoglycan synthesis. This hypothesis is supported by observations that point mutations in a particular region of FtsL and FtsB alter the tightly regulated cell division process, allowing it to proceed even in situations in which

normally it would not occur (Liu et al., 2015; Tsang and Bernhardt, 2015). Tsang and Bernhardt (Tsang and Bernhardt, 2015) discovered that a single point mutation in FtsL (E88K) allows the cells to bypass the normally strict requirements for other division proteins, namely FtsK, ZipA, FtsN, and FtsA (although residual levels of FtsA expression appear to be still necessary). Independently, De Boer and colleagues (Liu et al., 2015) identified a series of mutants that bypass the need for FtsN. These mutants map to two short regions of the coiled coil of FtsL (residues 88–94) and FtsB (residues 55–59). These regions, which are located ~30 residues past the TM domain, were named the “constriction control domain” (CCD) (Liu et al., 2015).

The observed properties of these CCD mutants are consistent with a conformational change in FtsLB that has become deregulated. In other words, an OFF/ON structural transition in FtsLB may be part of the events that control the beginning of cell constriction. This transition may be triggered allosterically by FtsN itself, as suggested by the observation that the essential region of FtsN (E_{FtsN} , a short stretch of sequence that is separated by ~20 amino acids from the C-terminal side of the TM domain) is in a position that is topologically equivalent with the location of the CCD in FtsLB (Liu et al., 2015; Tsang and Bernhardt, 2015).

To elucidate the effect of the CCD mutants and the overall function of the FtsQLB complex in molecular detail, it is necessary to understand its structural organization. So far, two partial structures have been obtained. The structure of the periplasmic domain of FtsQ has been solved by X-ray crystallography, without the TM and cytoplasmic domains (van den Ent et al., 2008). Additionally, we previously determined the crystal

structure of the coiled coil domain of FtsB and also produced a computational model of the TM domain of the same protein, based on mutagenesis data (LaPointe et al., 2013). Both structures were obtained in the absence of FtsL and in a homodimeric form. In addition, computational models of the FtsQLB periplasmic region (both in the trimeric and hexameric state) have also been reported (Villanelo et al., 2011). These models were based on available experimental data regarding contacts between FtsQ and FtsLB, whereas the FtsLB component was predicted *ab initio*. Additionally, these models lacked the TM region.

Here, we present a structural analysis of the TM domain and periplasmic coiled coil of FtsLB, based on a combination of biophysical and computational methods with *in vivo* validation. The work includes *in vitro* FRET to determine the oligomerization state of FtsLB; a co-evolutionary analysis of sequence alignments to provide an extensive set of predicted contacts between positions in FtsL and FtsB; molecular modeling to compute a three-dimensional structural model of the complex; molecular dynamics in lipid bilayers to evaluate stability and dynamics of the model; and a functional analysis of mutants *in vivo* to experimentally test the model and investigate the biological relevance of the predicted structural features.

Our results confirm that FtsLB is a higher oligomer and point to an L₂B₂ tetramer. Overall, the phenotypic analysis is in good agreement with structural features identified computationally. The distribution of deleterious phenotypes among the TM region mutants is consistent with the predicted interface. We confirm that the helix of FtsB breaks between the TM and periplasmic regions, with the formation of a flexible linker,

as hypothesized previously (LaPointe et al., 2013). In contrast, we show that FtsL forms a continuous helix and that the integrity of this helix is crucial for function. We also found indications that the coiled coil might not be built for structural stability, at least not in the form of a canonical tetrameric helical bundle assembly, suggesting that the coil may either serve as a dynamic structural unit or else split into two independent domains.

3.3. Results and discussion

3.3.1. Co-evolutionary analysis identifies potential quaternary contacts

consistent with an extended helical bundle

To predict the positions that mediate the association of the FtsLB complex, we analyzed a paired alignment of FtsL and FtsB sequences from proteobacteria species using the EV-Couplings algorithm (Marks et al., 2011). EV-Couplings uses a maximum-entropy model of the evolutionary history of a protein (or protein complex) to infer potential tertiary (Marks et al., 2011) and quaternary (Hopf et al., 2014) contacts in its structure. It is based on the notion that, to maintain complementary interactions, amino acid changes influence the variation of positions that are in close proximity. The results of our EV-Couplings analysis are shown in Figure 3.2a. The figure displays the top 95 pairs of positions identified between FtsL and FtsB. The complete ranked list is provided in Table S3.8.1. FtsB-FtsB and FtsL-FtsL pairs are displayed in Figure S3.1.

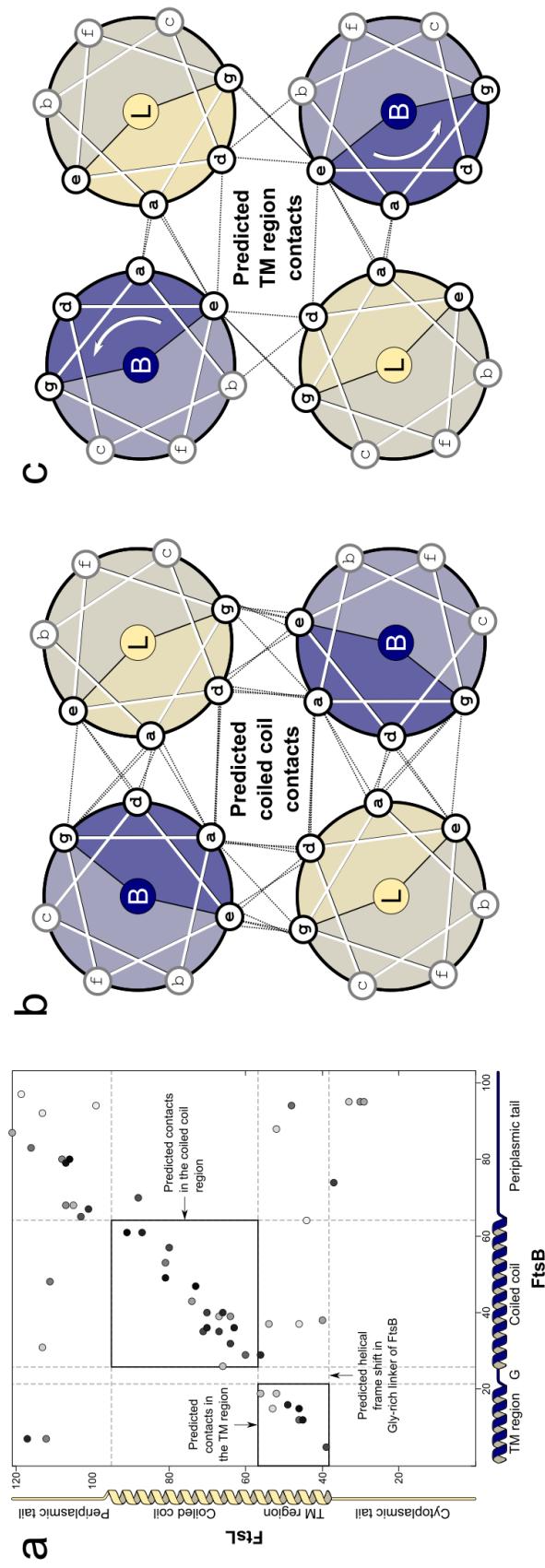


Figure 3.2: Co-evolutionary analysis identifies a self-consistent map of contacts in the FtsLB helical bundle. a, map of potential contacts between positions in the sequence in FtsB (x axis) and FtsL (y axis), inferred by co-evolutionary analysis. Displayed are the top 95 pairs identified by the analysis (the darkness of the shading reflects their ranking). The complete ranked list is provided in Table S3.8.1. The TM region and coiled coil are highlighted by boxes. Marked in the sequence of FtsB is also the position of a Gly-rich region between the TM and coil domains (G). b, predicted contacts in the coiled coil domain between FtsB (blue) and FtsL (yellow) displayed using a helical-wheel projection. The contacts are

in excellent agreement with the expected contacts for a canonical coiled coil. c, similar map for the TM region. The contacts are also consistent with the formation of a helical bundle. The interfacial positions of FtsL are on the same face of the interfacial position in the coiled coil, suggesting that FtsL forms a continuous helix. The positions in FtsB are rotated by ~60° with respect to those of the coiled coil (white arrow), indicating that a discontinuity is probably present between the two helical domains.

A notable diagonal pattern of co-evolving positions starts in the TM region of both proteins and continues in the coiled coil region, which is consistent with the contact map expected for an extended parallel helical bundle. As shown in the helical wheel diagram of Figure 3.2b, the co-evolving positions in the periplasmic region are also consistent with the expected geometry of a coiled coil (Buddelmeijer et al., 2002; Ghigo and Beckwith, 2000; Masson et al., 2009), occurring at positions of the “heptad repeat” (*abcdefg*) that are interfacial, at either the buried *a* and *d* positions or at the partially solvent exposed *e* and *g* positions. Moreover, these predicted contacts occur between pairs that are in proximity in a canonical coiled coil (e.g. *a-a*, *a-d*, and *g-e*, but not *a-e*). The pattern of connectivity remains in good agreement with a coiled coil contact map until positions 61 in FtsB and 91 in FtsL, after which it loses this periodicity. This range covers approximately five heptad repeats, which for FtsB is ~10–20 amino acids shorter than the consensus of the sequence-based predictors Paircoil (McDonnell et al., 2006), Multicoil (Wolf et al., 1997), Marcoil (Delorenzi and Speed, 2002), and Coils (Alva et al., 2016; Lupas et al., 1991) (inferring the length of the coiled coil of FtsL with

the same methods is difficult because it is poorly predicted (Ghigo and Beckwith, 2000)). The coiled coil predicted by the evolutionary analysis approximately reaches the CCD (55–59 in FtsB, 88–94 in FtsL, *highlighted in yellow* in Figure 3.2, *a* and *d*), suggesting that it is possible that this critical regulatory region occurs toward the end of the coil and thus in proximity of a structural transition.

3.3.2. A continuous FtsL helix and a discontinuous FtsB

As illustrated in the helical wheel diagram of Figure 3.2*c*, the TM region also displays a pattern of predicted contacts consistent with a bundle of parallel helices. The co-evolving positions are clustered toward the C-terminal side of the TM domains. Remarkably, the interfacial positions of the TM domain FtsB are in excellent agreement with those identified previously from their sensitivity to mutagenesis in a self-association assay (LaPointe et al., 2013).

Comparison of the contact maps (Figure 3.2, *b* and *c*) indicates that, in FtsL, the predicted interfaces of coiled coil and TM domains occur on the same face of the helix (involving the *a*, *d*, and *g* positions in the TM domain and the *a*, *d*, *e*, and *g* positions in the coil). This pattern indicates that FtsL forms a continuous helix across the membrane and the periplasmic region. A different outcome is observed for FtsB, for which the TM domain interface involves *a* and *e* but also *b* positions (*versus a*, *d*, *e*, and *g* positions in the coil) and thus appears rotated by ~60° with respect to the periplasmic region (as indicated by the arrow in Figure 3.2*c*). This rotation suggests the presence of a helical break. Rearrangements of the periodicity from the canonical heptad repeat are not uncommon in long coils. The 60° shift of the relative orientation of the interface between

the TM and coiled coil regions of FtsB would correspond to the insertion of four amino acids (*abcdabcde**f**g*), which is designated as a “stutter” (Brown et al., 1996; Hicks et al., 2002). Such mismatches influence the coil's local structure and can possibly facilitate conformational changes necessary for function (Schmidt et al., 2017; Strelkov et al., 2004). In this particular case, however, the presence of a conserved Gly-rich region (²²GKNG²⁵ in *E. coli* and ²²GKGG²⁵ in the consensus sequence of α-, β-, and γ-proteobacteria (LaPointe et al., 2013)) suggests that the juxtamembrane linker of FtsB is more likely to adopt a flexible and extended conformation instead of a local distortion of the helix.

3.3.3. FtsLB is a higher oligomer, probably an L2B2 tetrameric complex

To build a structural model for the FtsLB complex based on the evolutionary constraints, it was necessary to establish its oligomeric state. In previous work, we demonstrated that the isolated TM region of the FtsLB complex assembles to form a higher-order oligomer consisting of an equal number of FtsL and FtsB subunits, but we were unable to distinguish between heterotetrameric (L2B2), heterohexameric (L3B3), or even higher oligomeric forms (Khadria and Senes, 2013). Here, we assessed the stoichiometry of a construct that includes both periplasmic and TM domains by FRET analysis *in vitro*.

The FtsLB complex was overexpressed in *E. coli*, consisting of an N-terminally His-tagged FtsB and an N-terminally Strep-tagged FtsL. FtsL and FtsB were co-expressed. The proteins co-purify over sequential Ni-NTA and streptavidin columns in DM and DDM detergents, indicating that they form a stable complex with a density over Coomassie

staining compatible with a 1:1 stoichiometric ratio (Figure S3.2a). A C-terminal Cys residue was introduced in either FtsB or FtsL, and protein samples were separately labeled with either donor (Cyanine 3; Cy3) or acceptor (Cyanine 5; Cy5) fluorophores. Because the stoichiometric analysis is performed separately for the two proteins, to further confirm interaction between FtsL and FtsB in the experimental conditions, we mixed and equilibrated $\text{FtsL}_{\text{unlabeled}}\text{-}\text{FtsB}_{\text{Cy3}}$ and $\text{FtsL}_{\text{Cy5}}\text{-}\text{FtsB}_{\text{unlabeled}}$ samples, which produced a distinct FRET signal (Figure 3.2b).

We analyzed the stoichiometry of FtsLB by adopting a FRET method based on changing the relative fraction of donor- and acceptor-labeled molecules while the total protein concentration was maintained constant (Adair and Engelman, 1994). In these conditions, the theoretical variation of donor quenching (Q) depends on the oligomeric state of the complex, being proportional to $(1 - P_D^{n-1})$, where P_D is the relative fraction of donor-labeled protein, and n is the number of subunits in the complex (see Equation (4) under “Experimental procedures”). Due to the power law, no variation of donor quenching is expected for monomers as the donor fraction decreases, a linear increase of Q is expected for dimers, and a curved relationship is expected for trimers and higher oligomers. We performed two independent experiments, labeling only one of the two proteins with donor and acceptor fluorophores while leaving the other unlabeled. By labeling only either the FtsL or the FtsB moieties of the FtsLB complex, the procedure allowed us to calculate the number of subunits of FtsL and of FtsB individually.

As illustrated in Figure 3.3 (*left*), donor quenching increased linearly for FtsL-labeled samples when the acceptor fraction was increased. We used the sum of residuals

between the experimental data and the models to compare the fits. The residual of the two-subunit linear model (*dashed line*, residual (r) = 0.001) is 10-fold smaller than the residuals of the three-subunit model (*continuous line*, r = 0.01), indicating that two FtsL molecules are present in the FtsLB complex. The quenching data of FtsB (Figure 3.3, *right*) also display a progressive increase of donor quenching, establishing that at least two FtsB subunits are present in the FtsLB complex. However, the fit to two subunits (r = 0.0055) is only marginally better than the fit to three subunits (r = 0.0064); therefore, the data cannot differentiate between these two models, as in the case of FtsL. The fit, however, can reasonably rule out a four-subunit model (*dotted line*, r = 0.02).

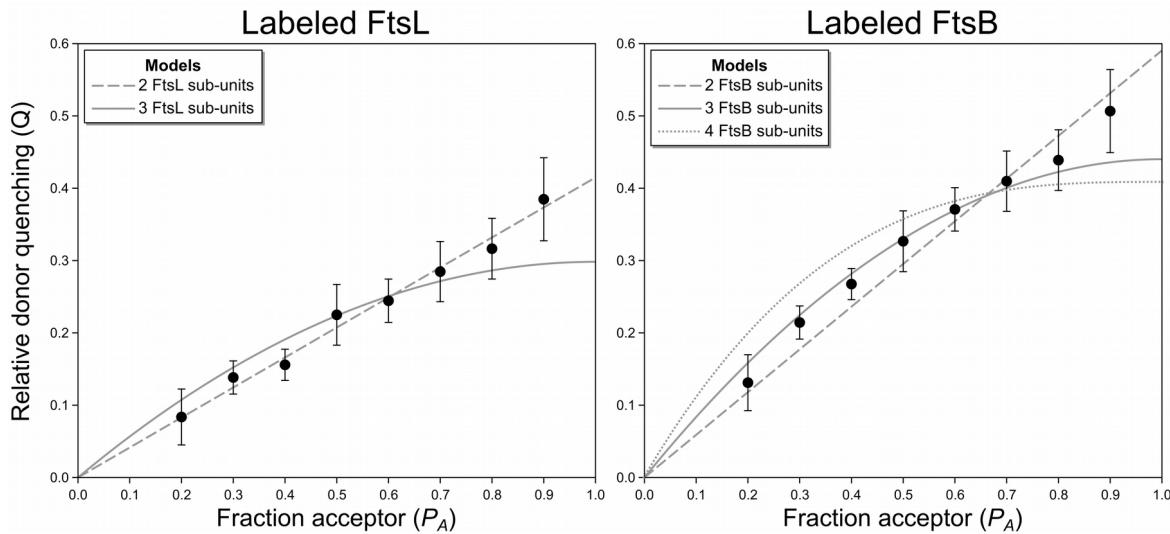


Figure 3.3: FRET analysis indicates FtsLB is a 2:2 tetramer. Analysis of the stoichiometry of Cy3- and Cy5-labeled FtsLB complexes was performed in DDM detergent. Two independent analyses were performed to count the subunits of FtsL (left) and FtsB (right) by labeling only one of the two proteins of the complex. Left, the FRET data of FtsL fit well to a two-subunit stoichiometry (dashed line, sum of residuals $r = 0.0010$) with respect to a three-subunit model (continuous line, $r = 0.0111$). Right, the fit for FtsB establishes that there are at least two subunits but does not differentiate between two- and three-subunit models ($r = 0.0055$ and 0.0064 , respectively). A four-subunit model can be excluded (dotted line, $r = 0.0224$). Considering prior evidence and likely symmetry, the most likely model for FtsLB is a 2:2 heterotetramer.

Overall, the analysis confirms that the FtsLB complex is a higher-order oligomer. Considering the entire body of available evidence, the most likely model is an L_2B_2

heterotetramer. Because FtsL fits well to a model containing two subunits, the possible states supported by the data are L_2B_2 or L_2B_3 stoichiometries. However, a 2:3 stoichiometry is in disagreement with previous FRET data that indicated that the FtsLB complex has an equivalent number of FtsL and FtsB subunits (Khadria and Senes, 2013). A 2:3 pentamer is also a less common oligomeric form in nature than a 2:2 heterotetramer (Levy et al., 2006). In addition, a 2:3 pentamer would be necessarily asymmetrical, whereas the co-evolutionary contacts are strongly consistent with a standard symmetrical organization (Figure 3.2, *b* and *c*). Finally, 2:2 would also be consistent with a previous estimate of the stoichiometry of the divisome based on ribosome profiling data (Egan and Vollmer, 2015).

3.3.4. Molecular modeling of the FtsLB helical bundle

We used a Monte Carlo docking procedure guided by the evolutionarily based distance constraints to create an all-atom computational model of the structure of the TM and coiled coil domains of the tetrameric complex (residues 1–61 of FtsB and 35–91 of FtsL). The model, which is illustrated in Figure 3.4a, consists of a four-helix bundle that spans the membrane and projects into the periplasmic region for ~55 Å. The structure is well packed, with the exception of a small void (~12 Å³) present at the level of the juxtamembrane linker region. A majority of the co-evolutionary distance constraints used to derive the model (*dashed lines* in the figure, listed in Table S3.8.1) are satisfied; of the 27 side chain pairs involved, 22 are in contact (minimal distance between heavy atoms (d_{\min}) < 5 Å), four are in proximity ($d_{\min} < 10$ Å), and only one

pair is separated by over 10 Å (FtsL Gln-66 with FtsB Ile-26). Because an L₃B₃ hexamer could not be entirely excluded, we also modeled this stoichiometry. The co-evolutionary restraints were well satisfied by the resulting model, but the model is significantly underpacked, with the presence of an incomplete pore spanning parts of the TM and coiled coil domains (Figure S3.3). For this reason, the L₃B₃ model was no longer pursued.

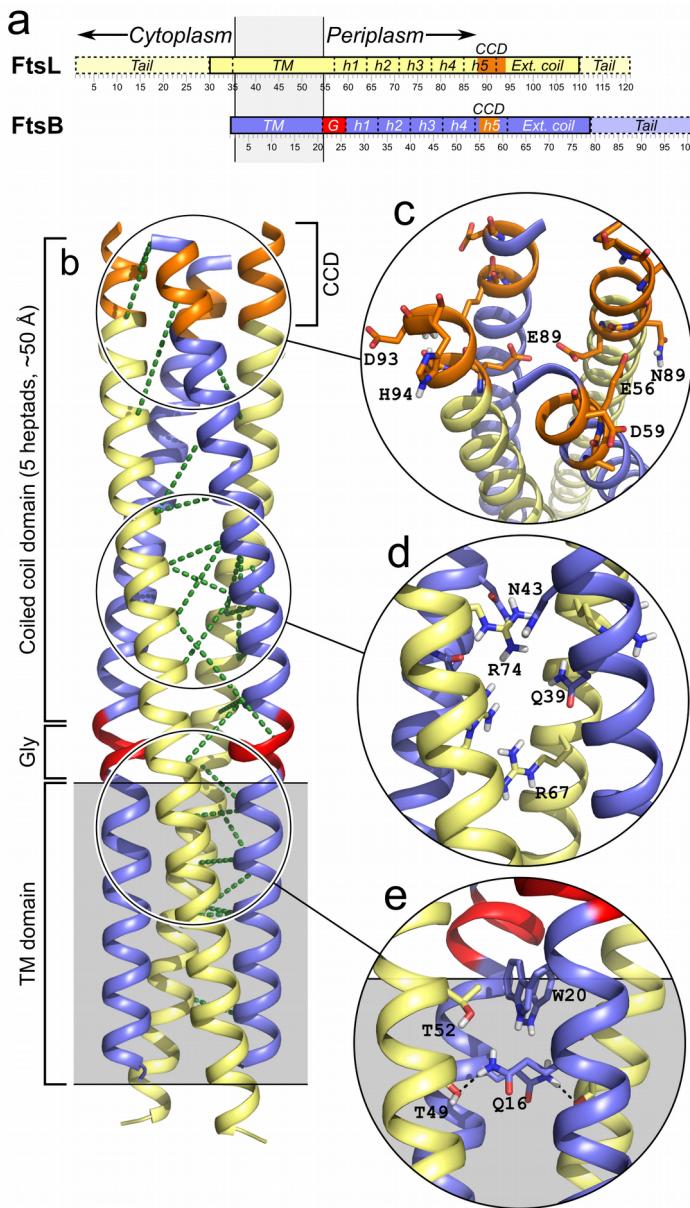


Figure 3.4: Computational model of FtsLB: A bundle with a helical break in FtsB.

a, schematic representation of the domains of FtsL and FtsB as discussed in this work. The five heptad repeats in the coiled coil are designated as h1, h2, etc. Ext. coil, extended coil region, as defined for the MD analysis. G (red), Gly-rich linker

in FtsB. The tails that are absent in the modeling are represented with dashed lines. The position of the CCD region is marked in orange. b, ribbon representation of the model of the FtsLB complex. The dashed lines represent the co-evolutionary constraints applied to obtain the model. FtsL (yellow) consists of continuous helices that traverse the membrane into the periplasmic region. The helix of FtsB (blue) is modeled in an unwound conformation in the juxtamembrane Gly-rich region (red). c, detail of the CCD region. If the coiled coil conformation persists beyond this region, the polar side chain of Glu-88 would be buried in the coil's core. d, the lower coiled coil also contains a cluster of very polar amino acids buried in the core (Gln-39 and Asn-41 from FtsB; Arg-67 and Arg-74 from FtsL). These amino acids are likely to be destabilizing if buried in a canonical coiled coil, as represented by the model. e, the TM region contains a cluster of amino acids able to form hydrogen bonds (Gln-16, Tyr-17, and Trp-20 from FtsB; Thr-49, Thr-52, and Thr-56 from FtsL), which may contribute to stabilize this domain.

The TM region of the L₂B₂ model forms a helical bundle characterized by a left-handed crossing angle (7° for FtsB and 10° for FtsL, tilt angle of the helices with respect to the membrane normal). The two TM helices of FtsB are in closer proximity to each other compared with the two helices of FtsL (interhelical distance of 13.6 and 15.2 Å, respectively). Remarkably, the conformation of FtsB is similar to a model of a FtsB dimer that we obtained previously (RMSD of 2.5 Å; Figure S3.4). This previous model

was obtained from the effect of point mutations on the homodimerization of the TM domain of FtsB in the absence of FtsL, as assayed with TOXCAT (LaPointe et al., 2013). The convergence of two completely independent sets of data, evolutionary information and experimental mutagenesis, to a similar model is a strong indication that the conformation and interface of the TM helices of FtsB are correctly predicted.

The C-terminal side of the TM region contains a number of side chains that can form hydrogen bonds, including the polar Gln-16. In the lowest-energy model, Gln-16 acts as a hydrogen bond donor to the side chain hydroxyl group of Thr-49 of FtsL (Figure 3.4b). Small changes in conformation would allow alternative interactions of Gln-16 with a number of other donor and acceptor groups. In particular, Gln-16 could form self-interactions with Gln-16 from the opposing chain as well as potential hydrogen bonds with Thr-52 of FtsL and Trp-20 of FtsB.

As expected, FtsL was modeled as a continuous helix across the TM and coiled coil domains. Also as expected, the change of orientation of the interface between the same two domains of FtsB required the introduction of a break in the α -helix in the juxtamembrane region (*red* in Figure 3.4). We opted to model the linker in the least perturbing conformation (an unwound helix), although the Gly-rich linker is likely to adopt a more extended conformation. The relative flexibility of the linker regions of FtsL and FtsB will be addressed later using molecular dynamics.

The periplasmic region (residues 27–61 in FtsB and 58–91 in FtsL) produced a well-packed canonical coiled coil. The coil is approximately five heptad repeats long (~35 amino acids). Interestingly, the domain is unusually rich in polar amino acids occurring

at *a* and *d* buried positions, with a total of 12. These polar amino acids are equally contributed by both proteins (three per subunit), although FtsB contains only neutral side chains (Gln-39, Asn-43, and Asn-50), whereas FtsL contains amino acids that are normally charged (Arg-67, Arg-74, and Glu-80). A region around the second and third heptad repeats of the coiled coil is particularly polar, where Arg-67 and Arg-74 from FtsL and Gln-39 and Asn-43 from FtsB occur in close proximity. The presence of so many buried hydrophilic side chains is interesting because they are likely to destabilize the coiled coil. In particular, the four Arg residues contributed by FtsL would be charged even if buried inside the protein core and thus very costly to desolvate (Harms et al., 2011). Another potentially charged side chain, Glu-80 in FtsL, is in the core in the modeled complex, but it is placed toward the end of the predicted coiled coil, near the CCD region, and thus it may be solvent-accessible.

3.3.5. Molecular dynamics suggest a stable TM region with an intricate network of hydrogen bonding

Two features, the presence of a flexible linker and a potentially destabilized coiled coil, raise questions about the dynamic properties of the FtsLB complex. To address them, we performed molecular dynamics (MD) simulations of the FtsLB model in explicit POPE bilayers. For this simulation, we extended the coiled coil conformation by ~20 amino acids beyond the region predicted by the co-evolutionary contacts, to avoid end effects and to test the coil boundaries. We will refer to this added region as the “extended coil” (residues 92–110 for FtsL and 62–79 for FtsB) and to the section predicted by the co-evolutionary analysis as the “lower coil” (residues 29–61 of FtsB

and 57–91 of FtsL). Three replica MD simulations were run for 260 ns (run 1; Figure 3.5) and for 200 ns (runs 2 and 3; Figure S3.5).

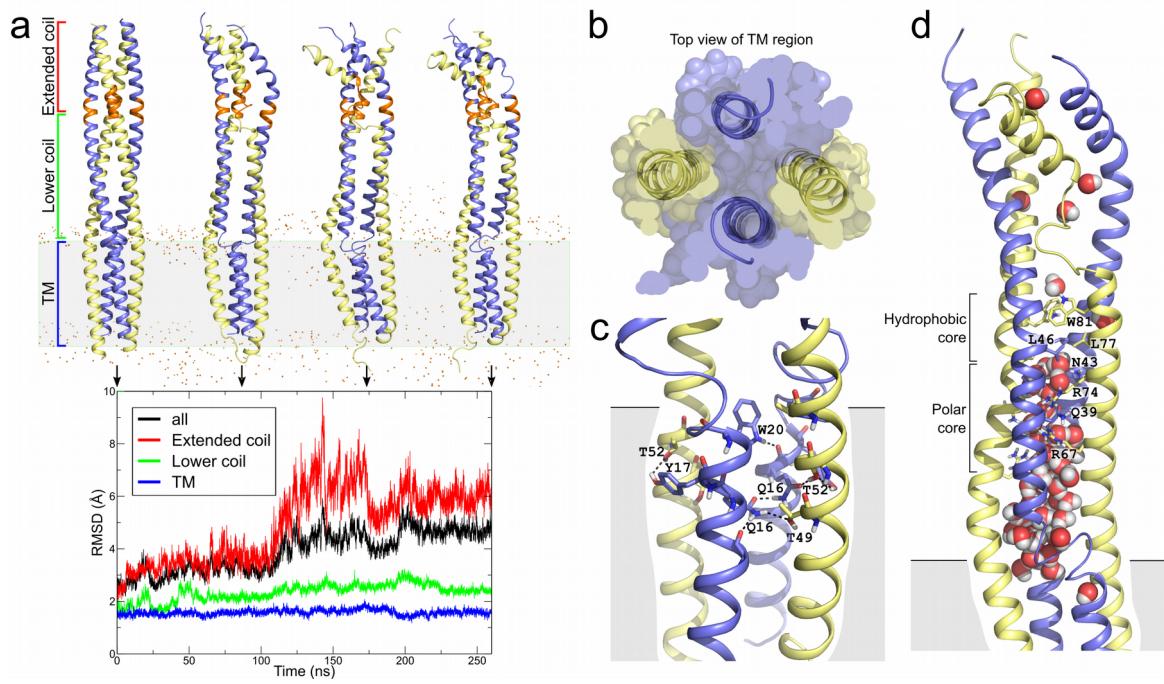


Figure 3.5: A stable TM region, a dynamic coiled coil. a, trajectory of the 260-ns run, with displayed conformations at 0, 87, 173, and 260 ns. The RMSD analysis is shown for the entire complex (black), the TM region (blue), the lower coil (green; 29–61 of FtsB and 57–91 of FtsL), and the extended coil (red, 62–79 of FtsB and 92–110 of FtsL). The CCD, which separates the lower from the extended coil, is in orange. As evident from the RMSD traces, the TM and lower coil are stable during the simulation, whereas the upper coil diverges, although it retains some helicity and interaction between pairs of FtsL and FtsB chains. The unfolding of the Gly-rich linker in the juxtamembrane region of FtsB is also evident. b, the TM region seen from above the membrane; the region rearranges during the simulation to bring the two FtsB helices (blue) closer to each other and

in contact, whereas the FtsL helices are now separated from each other. c, a configuration of the extended hydrogen bonding network present in the C-terminal side of the TM region, involving Gln-16, Tyr-17, and Trp-20 of FtsB and three threonine residues (Thr-49, -52, and -56) of FtsL. d, a view of the water that invades the core of the lower coiled coil, hydrating otherwise buried side chains or Arg-67, Arg-74, Gln-39, and Asn-43.

During the simulation, the RMSD from the initial structure increases to ~4, 5, and 6 Å in the three replica runs (*black traces* in Figure 3.5a and Figure S3.5). The majority of these changes are localized to the “extended coil” section (*red traces*), which quickly separates and partially unfolds. The lower coil remains relatively stable during the run (final RMSD around 2.3, 3.4, and 2.9 Å; *green traces*). The most stable region is the TM domain, which remains stable for the entire run across all three simulations, with average RMSDs around 1.6, 1.9, and 2.3 Å during the three runs (*blue traces*).

A slight rearrangement of the relative orientation of the TM helices is observed, which brings the two FtsB helices closer to each other by ~1–2 Å and, consequently, further separates the FtsL helices by a similar distance. This rearrangement gives the bundle a less “square” and more “rhomboid” configuration, one in which extensive packing occurs not only between FtsL and FtsB but also between the two FtsB helices, whereas the contacts between the two FtsL helices become reduced (Figure 3.5b). With the two helices of FtsB in closer proximity, their conformation becomes even more similar to our previous model of an FtsB dimer (RMSD ~2 Å; *Figure S3.4b*) (LaPointe et al., 2013).

A second important change that occurs in the TM region during the MD runs is a rearrangement and expansion of the hydrogen-bonding network. In the C-terminal section of FtsLB, there are 12 side chains in close proximity that are able to hydrogen-bond (three from each helix: Gln-16, Tyr-17, and Trp-20 in FtsB and three threonines, Thr-49, -52, and -56, in FtsL). With the exception of Trp-20, all of these side chains have both donor and acceptor groups. In addition, Gln-16 is also very flexible. Therefore, the network can rearrange in multiple configurations. Figure 3.5c illustrates one of the configurations observed, which displays a total of seven interhelical hydrogen bonds. The configuration of the hydrogen-bonding network varies over the simulations, but some interactions are predominant during the runs (Table S3.8.2). The most persistent interaction is between Tyr-17 of FtsB subunit A and the side chain of Thr-52 of FtsL subunit C (Tyr-17 of subunit B prefers to interact with the backbone carbonyl of Leu-48). Gln-16 interacts primarily with the side chain hydroxyl groups of Thr-49 and Thr-52 from FtsL and with Gln-16 from the opposed FtsB helix. Trp-20 primarily donates to the side chain hydroxyl group of FtsL Thr-56, but it can also donate to the carbonyl group of Gln-16 side chain. This extended network of hydrogen bonds is likely to contribute significantly to the association of the TM region, which can form a stable oligomer in isolation *in vitro* (Khadria and Senes, 2013).

As a control, we also performed a 160-ns simulation of a structural model that has comparably low energy but does not satisfy the evolutionary constraints in the TM region (“bad” model; *Figure S3.5d*). The TM region rearranges away from the initial model relatively quickly, reaching an RMSD of >3 Å in the first 60 ns. Its average (2.8 Å)

and maximum (3.5 Å) are higher than the RMSD of the TM region in the three replica runs of the “good” model (1.6, 1.9 and 2.3 Å, with maxima of 2.1, 2.4, and 2.7 Å, respectively).

3.3.6. A continuous FtsL helix and a dynamic coiled coil

Different sections of the coiled coil behave differently during the MD runs. These are notable in the helicity analysis presented in Figure S3.6. The “extended coil” (the region that extends beyond the pattern of co-evolutionary contacts consistent with a coiled coil (Figure 3.2)) unfolds partially as a bundle during the runs, although the region retains substantial helicity, and interactions still occur between pairs of FtsL and FtsB helices. The likely occurrence of a breakage of the helix between the “lower” and the “extended” coil regions was also hypothesized by a previous modeling analysis of the periplasmic region of the FtsLBQ complex (Villanelo et al., 2011). Interestingly, the location of the transition between these two regions roughly corresponds to the CCD.

The lower coil remains more stable during the run (final RMSD around 2.3, 3.4, and 2.9 Å). However, notable changes occur even in this region, providing further indication that the FtsLB complex may not be built to form a rigid, canonical, tetrameric coiled coil. Partial unfolding of the FtsL helix is observed in subunit C, and occasional unfolding is also notable in both FtsL subunits in the third replica run.

In the transition between the TM and coiled coil domains, the segment of 5–10 amino acids centered around the Gly-rich section of FtsB rapidly unfolds from the “unwound helix” conformation imposed by the modeling, unlinking the two domains and adopting a flexible and extended conformation (Figure 3.5a). Conversely, the

juxtapamembrane region of FtsL remains stable as a continuous helix throughout the duration of all three replica runs. The unfolding of the Gly-rich linker of FtsB creates an opening that allows water to access the core of the four-helix bundle (Figure 3.5d). Water molecules penetrate deeply up into the core of the coil, solvating the strongly polar side chains that would otherwise be buried in the structure (Arg-67 and Arg-74 from FtsL and Gln-39 and Asn-43 from FtsB, “polar core” in Figure 3.5d). In this region near the membrane, the coil becomes essentially separated by the water into two two-helix bundles. Water is excluded from the above layer, where the four-helix bundle becomes compact again around a hydrophobic core consisting of Leu-46 from FtsB and Leu-77 and Trp-81 from FtsL (“hydrophobic core” in Figure 3.5d).

With all caution in drawing conclusions from a theoretical model, it appears unlikely that a core so enriched in strongly polar amino acids would produce a very stable coiled coil. A possibility is that the stability of the coil is purposely “detuned” because the function of the FtsLB complex requires a weak coil, possibly to allow for a conformational change. A second hypothesis is that the region is actually designed to split into two separate two-helix coils. This split is most evident, as illustrated in the last frame of MD run number 3 (Figure S3.5). A “split” coiled coil would be consistent with the observation that a solubilized version of the periplasmic region of FtsLB forced into a heterodimeric form binds to FtsQ with high affinity (Glas et al., 2015; Masson et al., 2009; Noirclerc-Savoye et al., 2004), suggesting that association with FtsQ may be mediated by two independent binding domains of FtsLB. Such an organization would designate the TM region as the major factor driving tetramerization.

3.3.7. Functional analysis: The effect of mutations in the TM region is consistent with the predicted interface

To investigate how the model's structural features support the function of FtsLB, a series of rationally designed variants of the complex were tested *in vivo* for their ability to support cell division. Because FtsL and FtsB are essential proteins, the mutant proteins were introduced into strains in which a chromosomal copy of the wildtype protein is under the control of a repressible promoter (Gonzalez and Beckwith, 2009; Gonzalez et al., 2010). This allows for the depletion of the wildtype copy and the induction of the mutant version to reveal its phenotype. The expectation is that the most severe mutations will produce cells that elongate but are unable to divide, resulting in the formation of very long filaments, whereas less severe mutations will allow the cells to divide but will produce subpopulations of elongated cells. To assess each mutant, we measured the distribution of cell lengths, as exemplified in Figure 3.6a for the L15A mutant of FtsB. The data for each individual mutant are reported in Figure S3.7.

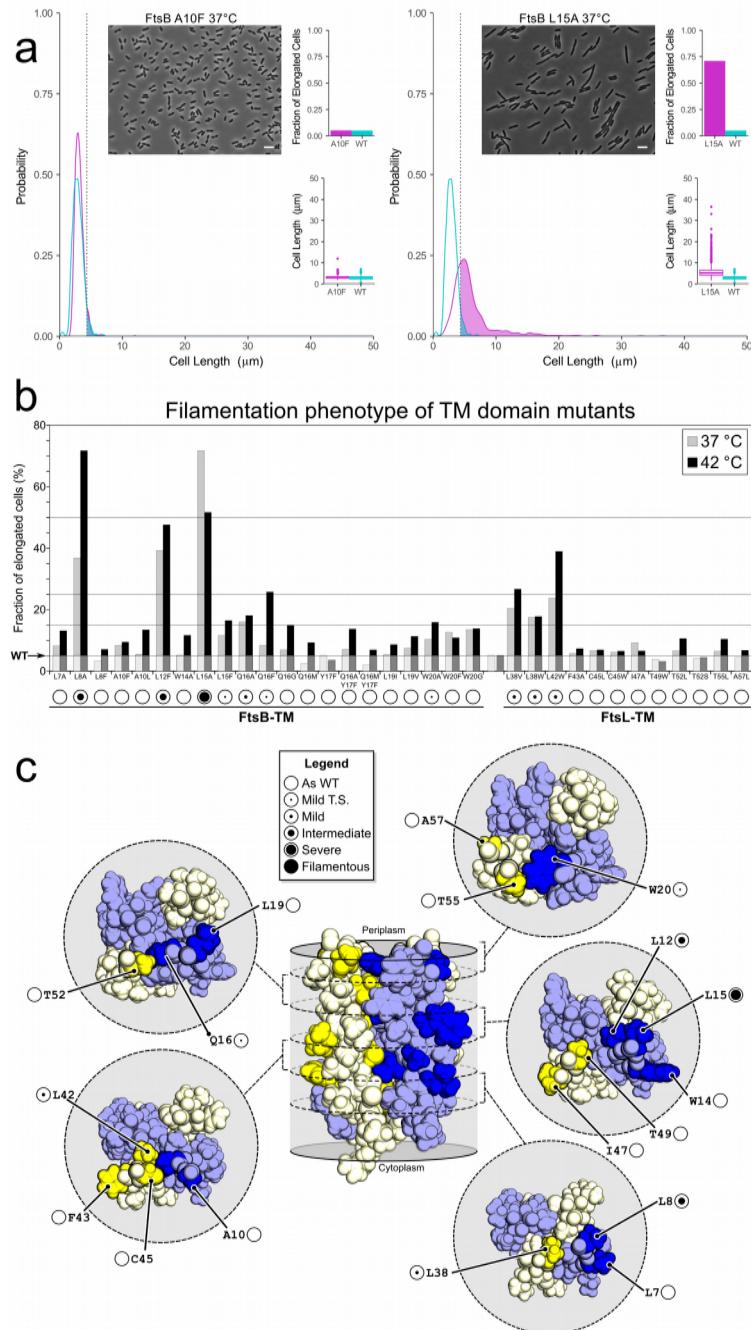


Figure 3.6: Mutations at the interface of the TM domain cause mild phenotypes in vivo. a, examples of in vivo analysis of point mutants with wildtype-like (A10F) and

defective (L15A) phenotypes. The distribution of cells lengths is compared between the wildtype (aqua) and the mutant (magenta). L15A displays 72% of cells that are longer than the 95th percentile in the wildtype distribution (colored areas, past the dotted line). As such, it is classified as a “severe” mutation. Graphs for each individual mutant are provided and explained in detail in Figure S3.7. Scale bar, 5 μm . b, phenotypes of TM domain mutants at 37 and 42 °C growth conditions. Classification is indicated using filled circles below. Filamentous, all cells are elongated at 37 °C; Severe, >50% of the cells are elongated; Intermediate, >25% of cells are elongated at 42 °C; Mild, >15% of cells are elongated at 42 °C; mild temperature-sensitive (Mild T.-S.), >15% of cells are elongated at 42 °C. c, location of the mutations within the structure of the TM domain. All mutations that display a cell division phenotype map within the interface of the helical bundle. All mutations on the outer surface of the bundle display As WT phenotypes.

Although none of the point mutations caused filamentous phenotypes, many produced elongated cells. To determine whether the difference between each mutant's and the wildtype's length distributions was significant, we used the Mann–Whitney *U* test (Mann and Whitney, 1947). However, as a consequence of the very large sample size of each experiment (500 to >1,000 cells), minuscule differences can be statistically significant even if they are not biologically relevant. To overcome this issue, we adopted an operative classification based on the fraction of elongated cells observed for each variant. We defined as “elongated” those cells that are longer than the 95th percentile in

the distribution of wildtype cell length (Figure 3.6a). Based on this threshold, we defined phenotypes as “mild temperature-sensitive” (*Mild T.S.*) when >15% (*i.e.* >3-fold compared with the wildtype) of cells were elongated only at the less permissive temperature of 42 °C; as “mild” when >15% of cells were elongated at 37 °C; as “intermediate” when >25% of cells were elongated (>5-fold); as “severe” when over 50% of the cells were elongated (>10 fold); and, finally, as “filamentous” when all cells were filamentous. The example of Figure 3.6a, which has 72% elongated cells at 37 °C, is therefore classified as “severe.”

The fractions of elongated cells for each variant in the TM region of FtsL and FtsB are reported in Figure 3.6b, with the relative classification reported at the *bottom*. The variants include a variety of drastic small-to-large amino acid changes (such as A10F in FtsB and L38W in FtsL), large-to-small changes (such as W14A and L8A in FtsB), and some conservative mutations (such as L19I in FtsB). We expected to observe division phenotypes when mutations affected the packing or hydrogen bonding at positions that mediate interaction between the helices. The results are in good agreement with this prediction. First, all positions that displayed impaired cell division phenotypes (Leu-8, Leu-12, Leu-15, Gln-16, and Trp-20 in FtsB; Leu-38 and Leu-42 in FtsL (Figure 3.6b)) occur at the helix-helix interfaces, as illustrated schematically in Figure 3.6c. In addition, all positions predicted to be away from the interface and exposed solely to lipids are classified as indistinguishable from wildtype (Leu-7, Ala-10, and Trp-14 in FtsB; Phe-43 and Ile-47 in FtsL). However, some of the TM region mutations that were predicted to be detrimental displayed little or no effect (*e.g.* Q16M, W20A, C45W, and T49W). This is

particularly noticeable for the C-terminal side of the TM bundle, a section characterized by an extensive hydrogen bonding network (Figure 3.5c). Mutations of Thr-49, Thr-52, and Thr-55 of FtsL and Tyr-17 of FtsB have no effect, and those of Gln-16 and Trp-20 of FtsB have little or no phenotype. Even a double mutation (Q16M/Y17F or Q16A/Y17F) that removes two hydrogen bonding groups is completely tolerated.

The finding is surprising because the C-terminal side of the TM bundle is the region with the highest number of co-evolutionary “connections,” suggesting that these positions would be structurally or functionally important. Moreover, Gln-16 and Trp-20 of FtsB, are nearly absolutely conserved in proteobacteria (LaPointe et al., 2013). A potential explanation is that the hydrogen bonding network may be sufficiently robust, extensive, and plastic to accommodate changes. Because of the large network, the removal of some donors and acceptors may not be sufficiently detrimental to destabilize the complex. Conversely, the N-terminal region, which is mediated exclusively by packing of hydrophobic residues, could be more sensitive to disruption of the complementary side chain packing. It is also possible that thermodynamic stability of the TM region is not strictly required for function, at least in the conditions tested.

To investigate whether there is an overall correspondence between the observed biological phenotypes and predictions based on the structural model, we calculated mutational energies *in silico* for all TM mutants and compared them with the fraction of elongated cells observed for each variant (using the 42 °C data to maximize the dynamic range of the experimental observations). We calculated the energies using a rigid-body model without backbone movement as well as with FoldX, a method

specifically calibrated to calculate the folding stability of protein mutants (Schymkowitz et al., 2005). The data are plotted in Figure S3.8. The energies produced by FoldX statistically correlate with the severity of the phenotypes by rank order Spearman correlation coefficient analysis ($r = 0.4631$, $n = 29$, $p < 0.01$) (Spearman, 1904). We conclude that perturbation of stability estimated on the basis of the structural model is a reasonably good predictor of the functional state of FtsLB.

3.3.8. The juxtamembrane and coiled coil of FtsB are tolerant to mutation

We applied a similar mutational approach to investigate the functional importance of the coiled coil and of the juxtamembrane region. The results are reported in Figure 3.7 (and Figure S3.7). We first tested whether the evolutionarily conserved Gly-rich juxtamembrane linker of FtsB ($^{22}\text{GKNG}^{25}$) is essential. A potential role for Gly is to provide structural flexibility, as suggested by our MD simulations. Alternatively, because of glycine's less restrictive Ramachandran distribution, this amino acid can also be required in a rigid structure to enable conformations of the backbone that are forbidden to all other amino acids.

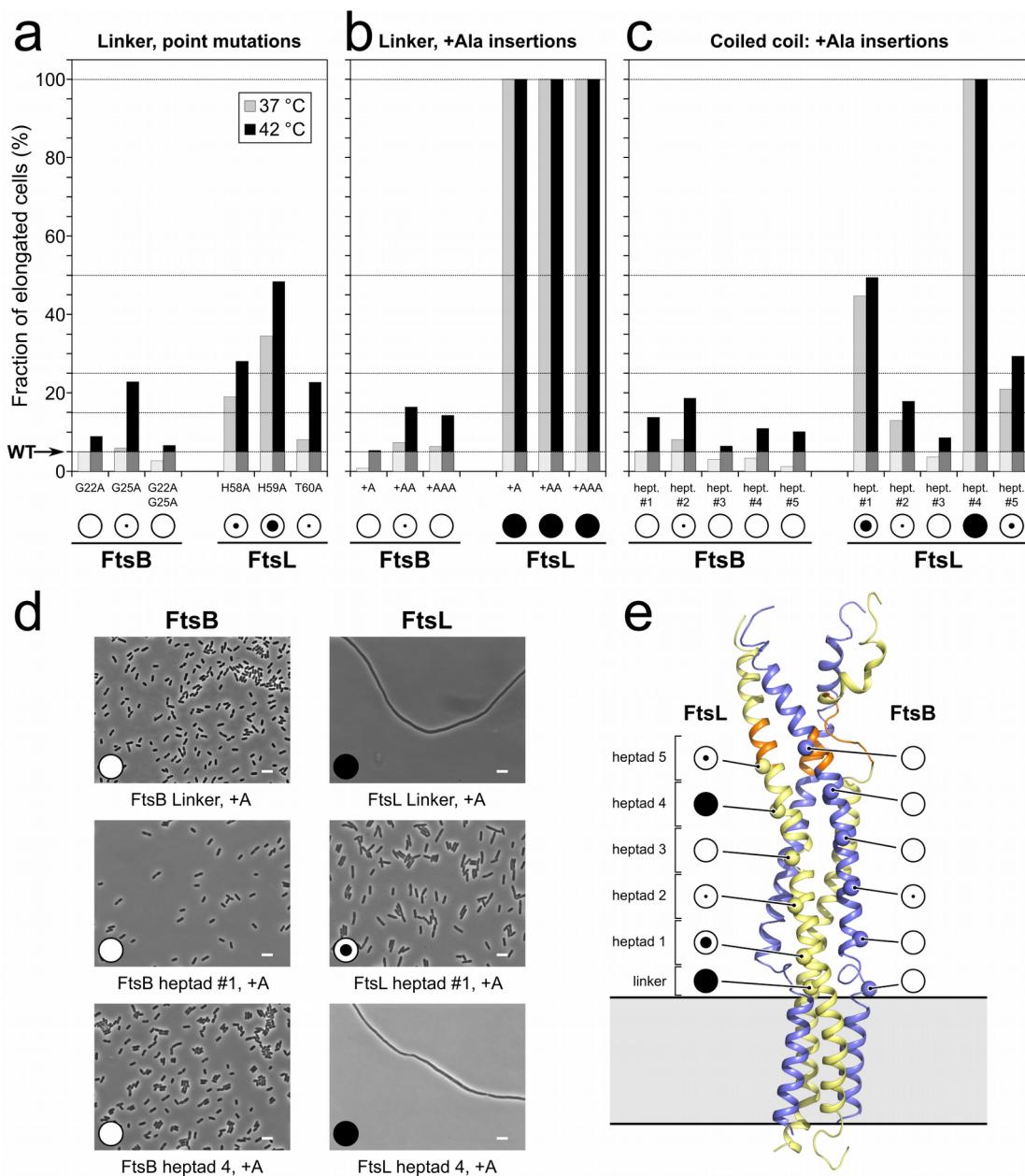


Figure 3.7: The juxta-membrane and coiled coil regions of FtsL are sensitive to mutation. a) Phenotypic analysis of juxta-membrane ("linker" region) point mutations of FtsL and FtsB, classified as detailed in Figure 3.6. Graphs for each

individual mutant are provided in Figure S3.7. b) Alanine insertion mutation (one to three extra residues) in the same region. c) Ala insertion mutations in the coiled coil region (operated between positions “b” and “c” in the heptad repeats). The evidence is consistent with a flexible linker, with the presence of a helical break, in FtsB, and with an uninterrupted helix that runs through the membrane and periplasmic regions for FtsL. Comparable expression levels of all variants displaying a phenotype were verified by Western blotting (Figure S3.9).

We tested individual and double Gly-to-Ala mutations at positions 22 and 25. They presented either no defects (G22A and G22A/G25A) or a mild TS phenotype (G25A) (Figure 3.7a). A possible explanation is that the structure of the linker retains sufficient flexibility even when Gly-22 and Gly-25 are substituted by Ala. To further test the flexibility of the linker, we inserted a series of Ala residues (up to three) between the TM domain and the coiled coil region (between positions Phe-21 and Gly-22), with the rationale that insertions should be better tolerated in a flexible region. The cells appear largely unaffected by the changes (only the two-Ala insertion mutation displays a mild TS phenotype; Figure 3.7b). The outcome is therefore consistent with a flexible FtsB linker and confirms the presence of a helical break between the membrane and periplasmic domains of FtsB. These observations are not consistent with the alternative hypothesis that Gly is required to enable a rigid backbone conformation.

We also applied this Ala insertion strategy to test the coiled coil region of FtsB. A single additional Ala residue was inserted between the *b* and *c* positions (*i.e.* in a position that is solvent-exposed) of each of the five heptad repeats. Most of these five

insertion mutants did not display any impairment, with the exception of a mild TS phenotype for the second heptad insertion after position Ala-37 (Figure 3.7c). Overall, the analysis suggests that the coiled coil of FtsB is likely to be quite plastic and able to tolerate insertion mutations all along its length.

3.3.9. Integrity of the juxtamembrane and coiled coil regions of FtsL is essential for function

Whereas the juxtamembrane and coiled coil domains of FtsB appeared tolerant to insertion, the opposite was observed for FtsL. We first introduced single, double, and triple Ala insertion mutations in the juxtamembrane region of FtsL (between positions Ala-57 and His-58). If the helix of FtsL is uninterrupted, the insertions should introduce strain into the structure because they would either produce a 100° rotation of the helix for each additional Ala or, more likely, introduce distortions into the helix. We observed that all insertion mutants, even the single Ala insertion, produced cells that were completely filamentous (Figure 3.7, *b* and cell images in *c*). These are the most dramatic phenotypes observed in the whole study. The results are consistent with the hypothesis that FtsL forms an uninterrupted helix that runs through the membrane and periplasmic regions.

We also tested the juxtamembrane linker of FtsL with a series of point substitutions (H58A, H59A, and T60A; Figure 3.7a). Interestingly, we observed some degree of functional disruption, particularly with the H59A mutation, which displays a 7-fold increase in the number of elongated cells compared with wildtype at 37 °C. In our model, His-59 is solvent-exposed and engaged only in minor interactions with the FtsB

helix. It is possible that this position is important for interactions with other components of the divisome, possibly with the periplasmic domain of FtsQ, which was shown to cross-link with this region of FtsL (van den Berg van Saparoea et al., 2013).

Finally, we scanned the entire coiled coil region of FtsL by Ala insertion (Figure 3.7c), as we did with FtsB. We observed a severe phenotype for the first heptad insertion, a mild phenotype for the second heptad insertion, and a wildtype phenotype for the third heptad insertion. A completely filamentous phenotype is then observed when Ala is inserted in the fourth heptad repeat (after position Arg-82). It is difficult to speculate what might cause such a dramatic defect, but we note that the mutation is in proximity to the CCD region of FtsL (residues 88–94, highlighted in *orange* in Figure 3.7e). Insertion at position 89, which is within the CCD and which would correspond to the fifth heptad, if the coiled coil persisted in this region, produced a mild phenotype (~4-fold increase in elongated cells). Overall, the periplasmic coiled coil of FtsL is sensitive to mutation, whereas the same domain of FtsB is tolerant, as schematically illustrated in Figure 3.7e. Interestingly, this outcome is the opposite of that observed for the mutagenesis of the TM region, where FtsB was the sensitive subunit, compared with the much milder phenotype displayed by FtsL (Figure 3.6).

3.4. Conclusions

The topology of the FtsLB complex, a helical bundle spanning the membrane and periplasmic space, with terminal tails available for binding other components, is widely conserved evolutionarily (Buddelmeijer and Beckwith, 2004; Masson et al., 2009). The widespread occurrence of FtsLB across a broad variety of bacterial species with

different cellular envelopes implies that this specific domain organization fulfills some important functional purpose, which is still unknown. To investigate this question, we have analyzed the structure-function relationship of the extended helical region of FtsLB, revealing a number of important features regarding the organization of the complex.

This study confirms that FtsLB is a higher-order oligomer and provides evidence that the complex consists of a heterotetramer. It identifies with a high degree of confidence the interface of the helical bundle region of the complex. It also produces a structural model of FtsLB validated experimentally through functional analysis. Further validation is also provided by a notable convergence of features in FtsB (the conformation of the TM helices and the presence of a flexible linker) that were also predicted in a prior model of FtsB based on a completely orthogonal set of experimental data (LaPointe et al., 2013).

The number of strongly polar, and even charged, amino acids that are predicted to be buried in the core of the coiled coil is a surprising and probably important finding, suggesting that the coil region may not be built for enhanced structural stability. A structural alternative to a monolithic four-helix coil could be the formation of two independent “dimeric” branches formed by one FtsL and one FtsB subunit. This possibility is consistent with previous observations that FtsLB fragments can bind to FtsQ with high affinity even when forced in a heterodimeric state (Glas et al., 2015; Masson et al., 2009; Noirclerc-Savoye et al., 2004).

We found that the integrity of the extended helix of FtsL is essential for the complex. A continuous helix is probably important for structural stability; given that the juxtamembrane linker of FtsB is disordered, a stable FtsL helix is probably necessary for nucleating the coiled coil. This hypothesis is in good agreement with the observation that amino acid insertions are most detrimental at the base of this domain. The helix of FtsL is also an interesting candidate for acting as a mechanical connection that could propagate a postulated allosteric conformational change (Liu et al., 2015; Tsang and Bernhardt, 2015) across the periplasmic, TM, and cytoplasmic regions. For example, the cytoplasmic region of FtsL is important for the recruitment of FtsW (Gonzalez et al., 2010); therefore, it is possible that FtsLB could control the activation of the FtsWI peptidoglycan synthase complex through direct interaction by coordinating a signal from the periplasm (*i.e.* interaction with FtsN). This study provides the structural groundwork necessary for investigating this and other hypotheses with coordinated biophysical and functional studies, which are necessary to finally clarify the precise role and the molecular mechanisms of the FtsLB complex in bacterial cell division.

3.5. Experimental Procedures

3.5.1. Co-evolutionary analysis

Sequences of FtsL and FtsB homologues were collected using the DELTA-BLAST algorithm on the RefSeq database (O'Leary et al., 2016). These sequences were filtered to include only proteobacterial species and then concatenated by matching the organism name in each record. The paired sequences were then aligned using the ClustalW algorithm with five guide tree and five hidden Markov model iterations (Li et

al., 2015). Columns in the alignment with a gap fraction higher than 0.3 were masked. This paired multiple sequence alignment was analyzed with the EV-Couplings algorithm (Marks et al., 2011) using default parameters and ignoring membrane topology, which would otherwise assume that sequential TM domains are anti-parallel. The top 95 (the number of unmasked columns in the multisequence alignment divided by 2) evolutionary constraints (*i.e.* the co-varying residues) predicted by EV-Fold between FtsL and FtsB were used for analysis and molecular modeling. The paired alignment is provided as a supplementary FASTA file. A total of 1291 sequences were used in the alignment. The number of effective sequences after reweighting for similarity is 883.1, corresponding ~4.6 sequences/residue. From this, ~40% of the top 95 contacts are estimated to be false positive pairs in the co-evolutionary analysis (Kamisetty et al., 2013).

3.5.2. Modeling the TM region of FtsLB

All modeling was performed using programs written in the Molecular Software Library (MSL) (Kulp et al., 2012). For the TM domain, ideal helices corresponding to residues 1–21 of FtsB and residues 35–58 of FtsL were generated. C_2 rotational symmetry (or C_3 , in the case of the hexamer) was preserved around the z axis for FtsB helices and FtsL helices. The geometry of the FtsL and FtsB bundles was defined by the following parameters: interhelical distance (d); rotation around the helical axis (ω); crossing angle (θ), and position of the crossing point (s). Additionally, the orientation of

the FtsL and FtsB bundles were changed by operating a rotation (Ω) and a translation (S) of each helix bundle about the z axis.

Starting from an initial random assignment of the parameters, the FtsLB TM helix complex was optimized using a Monte Carlo procedure that altered the interhelical geometry. The conformation of the side chains was periodically optimized with a 5% probability after each move. Side chain optimization was performed with a greedy trials algorithm using the backbone-dependent Energy-Based Conformer Library (bEBL) applied at the SL80 level (Subramaniam and Senes, 2012, 2014). Energies were calculated using the CHARMM 22 van der Waals function (MacKerell et al., 1998) and the hydrogen-bonding function of SCWRL 4 (Krivov et al., 2009) as implemented in MSL. Additionally, sigmoidal distance restraints were placed between the $C\alpha$ atoms of each pair of top co-evolving positions (Ovchinnikov et al., 2015), using the following form:

$$E_{\text{sigmoid}} = \frac{\omega}{1+e^{-a(r-r_0)}} + C \quad \text{Equation (1)}$$

Where E_{sigmoid} is the additional energy term, w is the weight, a is the slope, C is the intercept, r is the distance between the atoms, and r_0 is the distance cutoff. For these experiments, w was set to 10 kcal/mol, a was set to 0.5 \AA^{-1} , C was set to -2.5 kcal/mol, and r_0 was set to 10 Å.

For each pair of co-evolving residues, there are multiple pairs in the complex corresponding to the same residues on different chains. Only the restraints with the

lowest energies for each pair were added to the energy score, whereas the remaining ones were masked. The evolutionary constraints are listed in Table S3.8.1. Models were sorted by energy and clustered using a greedy algorithm and a C α RMSD threshold of 2.5.

As a negative control for the molecular dynamics simulations, a second model of the TM domain was produced by identifying a conformation with comparable energy of the initial model but with poor agreement with the co-evolutionary data (bad model). In this model, only one pair of co-evolving positions had a minimum heavy atom distance below 5 Å, and only two pairs had C α distances below 10 Å, whereas all pairs satisfy these conditions in the good model.

In silico mutational energies for the mutants in the transmembrane region were calculated in two ways. Rigid-backbone repacking was performed using MSL. Point mutations of interest were generated, and neighboring side chains were repacked using 100 rounds of a greedy trials algorithm. Mutated residues were repacked at the SL99 conformer level of the bEBL library; residues within 8, 16, and 20 Å were repacked at the SL95, SL90, and SL80 levels, respectively. van der Waals radii were scaled by 0.8. Energies of the mutants were subtracted from that of the wildtype to calculate their $\Delta\Delta E$. Mutations were also analyzed using the FoldX Suite (Schymkowitz et al., 2005). The models were first energy-minimized using the *RepairPDB* command with the membrane parameter set to true. Mutant $\Delta\Delta G$ calculations were performed using the *buildModel* command with the membrane parameter set to true.

3.5.3. Modeling the coiled coil domains of FtsL and FtsB

Supercoiled helices corresponding to residues 52–94 of FtsL and 21–63 of FtsB were generated by using a coiled coil generator based on a coiled coil parameterization described previously (North et al., 2001). The superhelical radius (r_1), superhelical pitch (P), helical rotation (Φ_1), and z-shift (s) of both FtsL and FtsB were freely altered, whereas the rise per residue (h) and helical radius (r_0) were kept constant. Additionally, the orientation of the FtsL and FtsB bundles were changed by operating a rotation (Ω) and a translation (S) of each helix bundle about the z axis. C_2 symmetry (C_3 for the hexamer) was preserved within the FtsL backbones and the FtsB backbones.

The coiled coils were optimized using a Monte Carlo procedure, changing their superhelical parameters starting from an initial assignment of random parameters. Side-chain conformational sampling was performed with a variable number of conformers with the positions that participate at a canonical coiled coil interface receiving higher sampling; the a and d positions were sampled at the SL80 conformer level of the bEBL library, e and g positions at the SL75 level, b and c at the SL70 level, and f at the SL60 level. Energies were calculated based on CHARMM 22 van der Waals and CHARMM 22 electrostatic terms. Additionally, sigmoidal restraints for each co-evolving pair in the coiled coil region were added, as described above.

3.5.4. Modeling the juxtamembrane regions of FtsL and FtsB

Top models of the TM and coiled coil domains were connected by aligning the helical residues 52–58 of FtsL, which were present in the models of both regions. The RMSD

between the C α atoms of these residues was minimized while keeping the main axis of both domains parallel to and centered on the z axis. The juxtamembrane regions of FtsL and FtsB were then replaced with loops corresponding to fragments from the PDB, as described previously (LaPointe et al., 2013). For FtsB, six-residue loops, corresponding to positions 21–26, with four flanking helical residues on each side, were used, with an additional sequence requirement that the fragment contain at least one glycine. For FtsL, 15-residue fragments with four flanking helical residues on each side were used with the requirement that the loop have helical secondary structure. Long helix fragments were used to better distribute minor deviation in alignment between the transmembrane and coiled coil helices. The connecting regions were optimized based on a greedy trials algorithm to minimize steric clashing, and the final model was minimized using BFGS constrained optimization using CHARMM (Brooks et al., 1983). Structural voids were analyzed using BetaCavityWeb (Kim et al., 2015). For the creation of the bad model, the TM model that did not satisfy the co-evolutionary constraints was aligned and connected to the same model of the coiled coil domain using an identical procedure.

3.5.5. All-atom molecular dynamics simulations

For the molecular dynamics simulations, the model's coiled coil region was extended, to avoid edge effects, to residues 110 (FtsL) and 79 (FtsB). The cytoplasmic side of FtsL was also extended to include residues 30–34, modeled in ideal α -helix. Four all-atom MD simulations (a 260-ns run, two 200-ns replica runs, and a 160-ns control run on the control bad model) were performed using the CHARMM 36 force field

and NAMD version 2.10 software (Klauda et al., 2010; Phillips et al., 2005). CHARMM-GUI membrane builder (Jo et al., 2008) was used to prepare systems composed of a POPE bilayer consisting of 301 lipids, the FtsLB tetramer, an ionic concentration of 0.150 m NaCl, and 59,034 TIP3P water molecules for hydration. The sizes of the boxes at the beginning of the simulation were $\sim 97 \times 97 \times 242 \text{ \AA}^3$ for runs 1, 2, and 3 and $97 \times 97 \times 245 \text{ \AA}^3$ for the control run. The simulations were initially minimized and equilibrated for 75 ps at an integration time of 1 fs/step and for 600 ps at an integration time of 2 fs/step. The integration time step for the production runs of each of the systems was 2.0 fs/step. The simulations were carried out in the NPT ensemble at a pressure of 1 atmosphere and a temperature of 310.15 K, using the Nose–Hoover Langevin piston and Langevin dynamics method. Particle mesh Ewald was used for electrostatic interactions, and a 12-Å cutoff was applied to Lennard–Jones interactions with a switching function from 10 to 12 Å. The RMSD analysis was performed using the RMSD trajectory tool in VMD (Humphrey et al., 1996). Hydrogen-bonding analysis was performed with an in-house script.

Helicity analysis was performed by measuring the backbone dihedrals and the distance between O_i and N_{i+4} for each residue in each selected frame of the simulation. If the O_i-N_{i+4} distance was between 2.0 and 4.2 Å and the backbone dihedral angles were within the favored α -helical region as defined in PROCHECK (Laskowski et al., 1996), the residue was classified as helical. If the O_i-N_{i+4} distance was within 5 Å and the backbone dihedral angles were within either the favored or

allowed α -helical region, the residue was classified as near-helical. Otherwise, the residue was classified as non-helical.

3.5.6. Cloning, expression, purification, and labeling of FtsLB constructs for FRET measurements

The His-tagged FtsB and Strep-tagged Cys-less (C41A and C45A) FtsL(35–121) (Table S3.8.3) were introduced into a modified pETDuet-1 vector at restriction sites Ncol/HindIII and NdeI/Xhol, respectively. For fluorophore labeling, cysteine mutations were introduced either to FtsB (S97C) or to Cys-less FtsL (I100C) via QuikChange mutagenesis (Novagene). All constructs were confirmed by DNA sequencing (Quintarabio).

The plasmids were transformed into BL21 (DE3) cells. Individual colonies were picked and grown overnight in 3 ml of LB broth containing 100 μ g/ml of ampicillin before being diluted 1:500 in 1 liter of ZYP-5052 autoinduction medium as described (Studier, 2005), and grown at 37 °C until reaching an A_{600} of 0.8, after which it was incubated for an additional 20 h at 25 °C. The cells were then lysed by sonication in 10 ml/g lysis buffer (50 mm NaCl, 50 mm HEPES, 10 mm TCEP, brought to pH 7.0 with NaOH) supplemented with 0.25 mg/ml lysozyme, 5 mm β -mercaptoethanol, and 1 mm phenylmethylsulfonyl fluoride. The inclusion body fraction was separated by centrifugation at 10,000 \times g for 20 min, followed by ultracentrifugation of the supernatant at 130,000 \times g for 30 min to isolate the cell membranes. The FtsLB complex was then extracted from the membrane fraction with lysis buffer supplemented with 18 mm *n*-decyl- β -D-maltopyranoside (DM; Anatrace) and 10 mm TCEP, rocking at 4

°C overnight. Solubilized protein was added to 3 ml of Ni-NTA-agarose resin (Qiagen) and rocked for 2 h at 4 °C before the resin was washed and the complex was labeled on-column with either Cy3 or Cy5 maleimide (Lumiprobe) for FtsB or FtsL. On-column fluorophore labeling was performed by running 3 column volumes of Ni wash buffer (300 mm NaCl, 25 mm HEPES, pH 8.0, 50 mm imidazole, 10 mm TCEP, brought to pH 7.0 with NaOH) supplemented with 5.4 mm DM, 3 column volumes Cy3 or Cy5 labeling buffer (100 µm Cy3 or Cy5 maleimide, 5.4 mm DM, 50 mm NaCl, 50 mm HEPES, pH 8.0, brought to pH 7.0 with NaOH), 6 column volumes of Ni wash buffer supplemented with 450 µm *n*-dodecyl-β-d-maltoside (DDM; Avanti Polar Lipids), and 2 column volumes of elution buffer (300 mm NaCl, 25 mm HEPES, pH 8.0, 450 µm DDM, 300 mm imidazole, brought to pH 7.0 with NaOH). For quantification, the elution fractions were dialyzed overnight at 4 °C against FRET buffer (300 mm NaCl, 25 mm HEPES pH 8.0, 450 µm DDM, brought to pH 7.0 with NaOH).

Labeling efficiency of each fraction was quantified by UV-visible spectroscopy, using absorbance at 280 nm (protein), 550 nm (Cy3), and 650 nm (Cy5). First, the concentration of the protein was calculated, taking into account the absorbance of the fluorophore:

$$[Protein] = \frac{A_{280} - (A_{Cy} \times CF_{Cy})}{\epsilon_{280}} \quad \text{Equation (2)}$$

Where A_{280} is the absorbance at 280 nm, A_{Cy} is the absorbance at 550 nm (Cy3-labeled samples) or 650 nm (Cy5-labeled samples), ϵ_{280} is the molar extinction

coefficient of FtsLB at 280 nm ($32,430 \text{ m}^{-1} \text{ cm}^{-1}$), and CF_{Cy} is the appropriate correction factor ($CF_{Cy3} = 0.11$ and $CF_{Cy5} = 0.05$) to subtract the contribution of the fluorophore to absorbance at 280 nm. Then the labeling efficiency P_{Cy} was calculated according to the following:

$$P_{Cy} = \frac{A_{Cy}}{[Protein] \times \epsilon_{Cy}} \quad \text{Equation (3)}$$

Where ϵ_{Cy} is the molar extinction coefficient of the fluorophore ($\epsilon_{Cy3} = 150,000 \text{ m}^{-1} \text{ cm}^{-1}$ at 550 nm, $\epsilon_{Cy5} = 250,000 \text{ m}^{-1} \text{ cm}^{-1}$ at 650 nm). Cysteine-less versions of the FtsL and FtsB constructs were also purified and labeled three times with the same protocol to determine background labeling, which was negligible.

A second round of purification could be performed by loading the Ni-NTA-purified FtsLB complex over a streptavidin column. The most concentrated Ni-NTA elution fractions were added to 3 ml of streptavidin resin and rocked for 2 h at 4 °C before being washed with 3 column volumes of streptavidin buffer (20 mM Tris, 300 mM NaCl, 0.1 mM EDTA, 0.5 mM TCEP, brought to pH 8.0 with NaOH) with 450 µM DDM, followed by elution with 3 column volumes of streptavidin buffer plus 2.5 mM desthiobiotin (initially dissolved in 1 ml of DMSO) and 450 µM DDM. Samples from each fraction of the nickel and streptavidin column were boiled in 4× SDS-PAGE loading buffer before being run on NuPage 4–12% BisTris protein gels at 150 V for 1 h (Thermo Fisher Scientific). Because the FtsLB complex obtained after Ni-NTA had sufficient

purity for FRET studies, this second chromatography was not routinely performed, but it was used initially to confirm the stability of the FtsLB complex.

3.5.7. FRET measurements

In all experiments, the FtsLB complex was diluted to 1.35 μm in FRET buffer, which contains 450 μm DDM, for a final protein/detergent molar ratio of 1:333. Experiments were performed by labeling either the FtsL moiety (FtsL-C41A/C45A/I100C + FtsB) or the FtsB moiety (FtsL-C41A/C45A + FtsB-S97C) in the protein complex. All FRET analysis was performed on an M1000 Tecan plate reader, with excitation set at 550 nm and emission recorded from 560 to 800 nm in 1-nm increments.

The formation of FtsL-FtsB oligomers *in vitro* was assessed by measuring FRET on a 50:50 mix of FtsL_{Cy3}-FtsB_{unlabeled} and FtsL_{unlabeled}-FtsB_{Cy5} complexes, which were compared with a 50:50 mix of FtsL_{Cy3}-FtsB_{unlabeled} and FtsL_{unlabeled}-FtsB_{unlabeled} (donor-only sample) complexes and a 50:50 mix of FtsL_{unlabeled}-FtsB_{Cy5} and FtsL_{unlabeled}-FtsB_{unlabeled} (acceptor-only sample) complexes.

The stoichiometric analysis of FtsL was performed by mixing FtsL_{Cy3}-FtsB_{unlabeled} (donor) and FtsL_{Cy5}-FtsB_{unlabeled} (acceptor) samples in different ratios, from 20:80 donor/acceptor up to 90:10, in 10% increments. To determine donor fluorescence in the absence of the acceptor, equivalent samples were produced by mixing FtsL_{Cy3}-FtsB_{unlabeled} (donor) and FtsL_{unlabeled}-FtsB_{unlabeled} (unlabeled) in the same molar ratios. The stoichiometric analysis of FtsB was performed similarly. Donor (Cy3)

fluorescence was recorded using its peak emission at 570 nm. Quenching (Q) of the donor fluorophore was calculated, as explained below, according to Equation (6), from donor/acceptor-labeled protein sample mixes (F) and donor/unlabeled protein sample mixes (F_0). Each experiment was independently replicated at least five times.

3.5.8. Fitting of experimental FRET data to obtain the number of subunits

Fitting for different oligomeric states was performed as described by Adair and Engelman (Adair and Engelman, 1994):

$$Q = \left(1 - \frac{f_Q}{f_D}\right) \left(1 - P_D^{n-1}\right) \quad \text{Equation (4)}$$

Where Q is related to the molar fluorescence of the quenched donor (f_Q) in the presence of the acceptor, the molar fluorescence of the donor f_D in the absence of acceptor, the number of subunits in the oligomer n , and the molar fraction of donor-labeled protein P_D .

P_D was calculated as follows:

$$P_D = \frac{[D]}{[D] + [A]} \quad \text{Equation (5)}$$

Where $[D]$ is the molar concentration of donor-labeled protein and $[A]$ is the molar concentration of acceptor-labeled protein.

Relative quenching was calculated as follows,

$$Q = 1 - \frac{F}{F_0} \quad \text{Equation (6)}$$

Where F is the experimentally measured quenched fluorescence for a certain donor/acceptor molar fraction and F_0 is the experimentally measured unquenched fluorescence of the same amount of donor, obtained in the absence of acceptor and in the presence of an equivalent amount of unlabeled protein. In Equation (4), the quantity related to the molar fluorescence can be treated as an overall unknown constant k .

$$\left(1 - \frac{f_Q}{f_D}\right) = k \quad \text{Equation (7)}$$

The parameter k was thus fit using a least square procedure to the experimental data Q as a function of donor fraction P_D according to Equation (8) for the different oligomeric states n .

$$Q(P_D) = k(1 - P_D^{n-1}) \quad \text{Equation (8)}$$

The relative quality of the fits to different oligomeric states was assessed by comparing the sum of the squared residuals between the experimental data and the models.

3.5.9. Bacterial strains, plasmids, and media for in vivo experiments

The phenotypic analysis was performed using depletion strain NB946 for FtsB (Buddelmeijer et al., 2002) and MDG277 for FtsL (Gonzalez and Beckwith, 2009). For all experiments described, bacterial cells were grown in LB medium supplemented with

100 µg/ml spectinomycin (Dot Scientific) and the appropriate carbon source. Medium was supplemented with 0.2% (w/v) l-arabinose (Sigma) or 0.2% (w/v) d-glucose (Sigma) to induce or repress, respectively, the expression of chromosomal copies of the wildtype genes regulated by the PBAD promoter. 20 µm isopropyl-β-d-1-thiogalactoside was added to the media to induce the expression of mutant genes regulated by the pTrc promoter in the plasmid. Point mutants and insertion mutants of FtsL and FtsB were constructed in plasmid pNG162 (Gonzalez and Beckwith, 2009) using either standard QuikChange mutagenesis or inverse PCR mutagenesis.

3.5.10. Depletion strain experiments

The protocol for the depletion strain experiment was adapted from Gonzalez and Beckwith (Gonzalez and Beckwith, 2009). In short, a mutated copy of FtsB or FtsL was transformed into its respective depletion strain. Strains were grown overnight at 37 °C on an LB plate supplemented with arabinose and spectinomycin. A single colony from the plate was grown overnight at 37 °C in 3 ml of LB medium supplemented with arabinose and spectinomycin. The overnight culture was then diluted 1:100 into fresh LB medium containing the same supplement and grown to an A_{600} of ~0.3. An aliquot of 1 ml of culture was washed twice with LB medium lacking any sugar and then diluted 1:100 into 3 ml of fresh LB medium supplemented with glucose and isopropyl-β-d-1-thiogalactoside to induce the expression of the mutated gene and the repression of the wildtype gene. The cells were then grown at 37 or 42 °C for 3.5 h before microscopy, the approximate time necessary to deplete the cells of the wildtype chromosomal copy (Gonzalez and Beckwith, 2009). Depletion strains provided with their respective

wildtype copy of the protein in the plasmid were tested as positive controls, and, similarly, depletion strains with no protein in the plasmid (empty vector) were tested as negative controls.

3.5.11. Microscopy and cell length measurement

10 µl of cell samples were mounted on a number 1.5, 24 × 50-mm (0.16–0.19-mm thickness) coverglass slide (Fisher). Cells were cushioned with a 3% (w/v) agarose gel pad to restrict the movement of the live cells. Cells were optically imaged using a Nikon Eclipse Ti inverted microscope equipped with crossed polarizers and a Photometrics CoolSNAP HQ2 CCD camera using a Nikon ×100 oil objective lens. Phase-contrast images of bacterial cells were recorded with a 70-ms exposure time using Nikon NIS Elements software. Multiple snapshots were collected for each experiment. All images were analyzed to measure the cell length in Oufti (Paintdakhi et al., 2016) using one single optimized parameter set.

3.5.12. Whole-cell lysate preparation and Western blotting

Expression level across all variants was assessed by Western blot analysis (Figure S3.9). 3.0 ml of cells were pelleted and resuspended in 500 µl of sonication buffer (25 mm Tris-HCl, 2 mm EDTA, pH 8.0). The cells were sonicated and centrifuged at 16,000 × g for 5 min before collecting the supernatant. Total protein concentration was determined by BCA assay (Pierce). 150 µl of lysates were mixed with 50 µl of 4× LDS sample buffer (Novex, Life Technologies) and boiled at 95 °C for 5 min. For each sample, the equivalent of 10 µg of total protein was separated by SDS-PAGE (Invitrogen) and transferred to polyvinylidene difluoride membrane (VWR). Horseradish

peroxidase-tagged anti-FLAG (M2) antibodies (Sigma; 1:1,000) were used for immunoblotting analysis.

3.6. Author contributions

S.G.C., D.-A.M., C.R.A., G.D.-V., and A.S. conceptualization; S.G.C., D.-A.M., and A.S. data curation; S.G.C. and A.S. software; S.G.C., D.-A.M., C.R.A., G.D.-V., and A.S. formal analysis; S.G.C., D.-A. M., C.R.A., G.D.-V., S.J.C., and A.S. investigation; S.G.C., D.-A.M., C.R.A., L.M.L., A.S.K., R.C., J.A.C., N.R., D.B.W., A.A.H., J.L.R., Q.C., and A.S. methodology; S.G.C., D.-A.M., C.R.A., G.D.-V., and A.S. writing—original draft; S.G.C., D.-A.M., C.R.A., G.D.-V., S.J.C., and A.S. writing—review and editing; C.R.A. validation; D.B.W., A.A.H., J.L.R., Q.C., and A.S. supervision; A.S. funding acquisition; A.S. project administration.

3.7. Acknowledgments

We thank Drs. Jon Beckwith and Mark Gonzalez for kindly providing plasmids and strains for the *in vivo* analysis and for guidance in performing the experiments. Computational resources from the Extreme Science and Engineering Discovery Environment (XSEDE), supported by National Science Foundation Grant OCI-1053575, are greatly appreciated. We thank Dr. Kai Cai for assistance during the biophysical studies and Dr. Elizabeth Caselle for critical reading of the manuscript.

This work was supported in part by National Institutes of Health (NIH) Grant R01-GM099752, National Science Foundation (NSF) Grant CHE-1415910 (to A. S.), and NSF Grant DMS-1661900 (to Q. C.). *The authors declare that they have no conflicts of interest with the contents of this article.* The content is solely the responsibility of the

authors and does not necessarily represent the official views of the National Institutes of Health.

3.8. Supplementary Figures and Tables

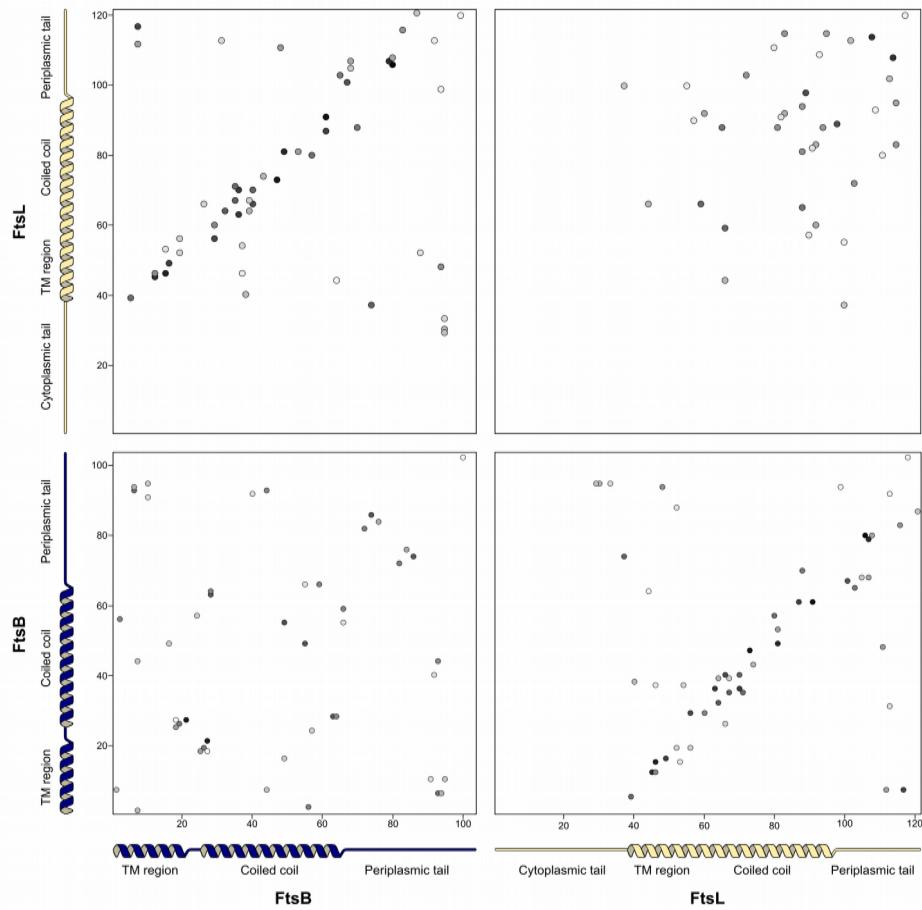


Figure S3.1: Co-evolutionary analysis of the paired alignment of FtsB and FtsL sequences. Map of potential contacts between positions in the sequence in FtsB and FtsL, inferred by co-evolutionary analysis. The top-left and bottom-right quadrants correspond to the FtsB-FtsL potential contacts reported in Figure 3.2. In addition, this figure reports the potential intra-subunit contacts, within FtsB (bottom-right) and within FtsL (top-right).

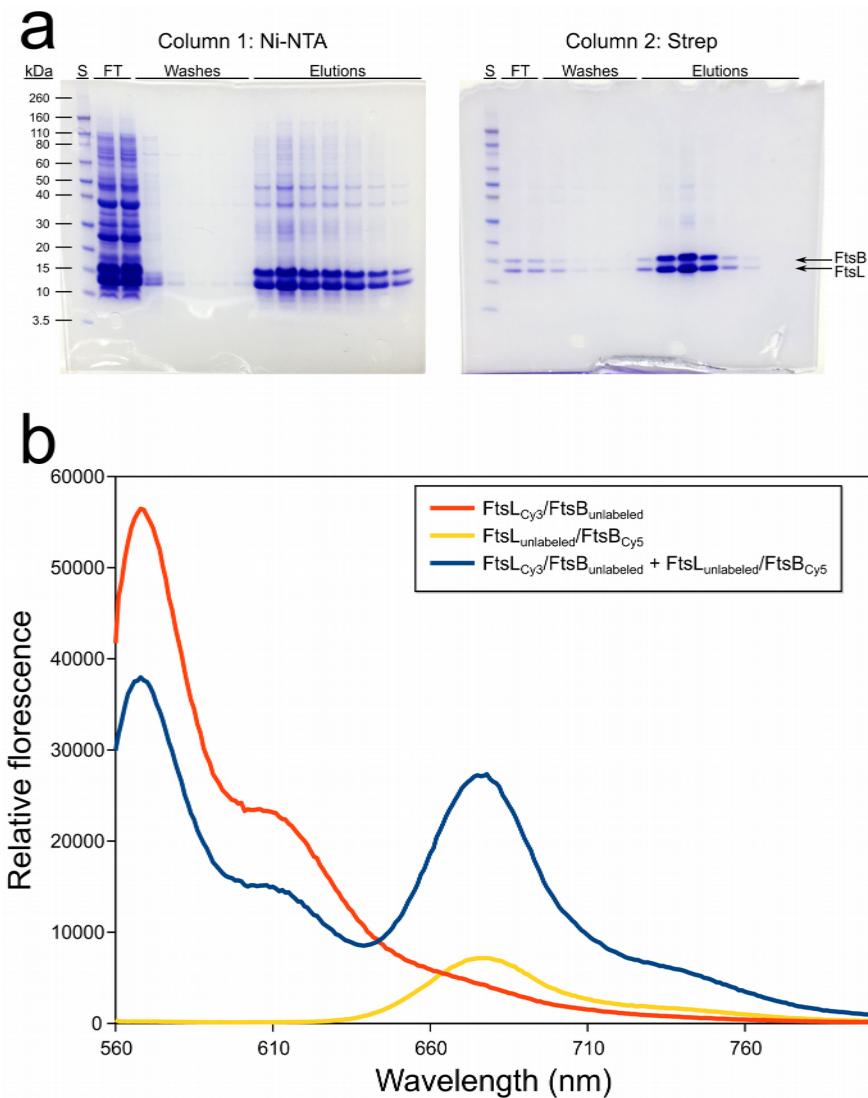


Figure S3.2: Determination of FtsL and FtsB association in vitro. a) Co-expressed strep-tagged FtsL and His-tagged FtsB proteins co-elute when run over consecutive Ni-NTA and streptavidin columns, demonstrating that FtsLB a stable complex in detergent. The density of the bands is consistent with a 1:1 stoichiometric ratio. S: standard ladder; FT: flow-through fractions. b) Acceptor-labeled FtsB ($\text{FtsL}_{\text{unlabeled}}/\text{FtsB}_{\text{Cy5}}$) and donor-labeled FtsL ($\text{FtsL}_{\text{Cy3}}/\text{FtsB}_{\text{unlabeled}}$)

were mixed in a 50:50 ratio and equilibrated and emission spectrum was recorded (blue line, excitation wavelength 550 nm). The mixed sample shows a substantial increase of acceptor emission around 680 nm, indicating proximity between FtsB_{Cy5} and FtsL_{Cy3}. The emission scan is compared to those of an acceptor-only sample (red, 50:50 mix of FtsL_{unlabeled} /FtsB_{Cy5} and FtsL_{unlabeled} /FtsB_{unlabeled}) and of a donor-only sample (yellow, 50:50 mix of FtsL_{Cy3} /FtsB_{unlabeled} and FtsL_{unlabeled} /FtsB_{unlabeled}). For these two control samples, mixing with FtsL_{unlabeled} /FtsB_{unlabeled} was performed to maintain identical total protein concentration and protein: detergent ratios. All other conditions are identical to the stoichiometric analysis of Figure 3.3.

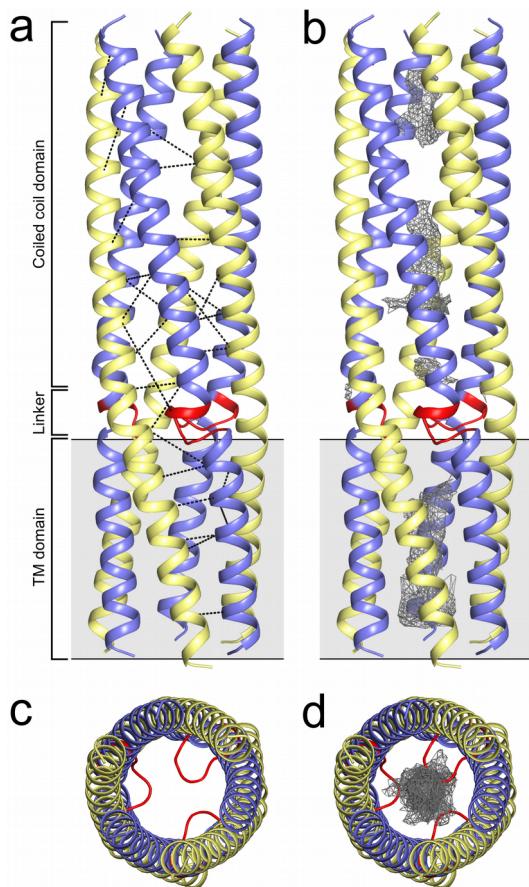
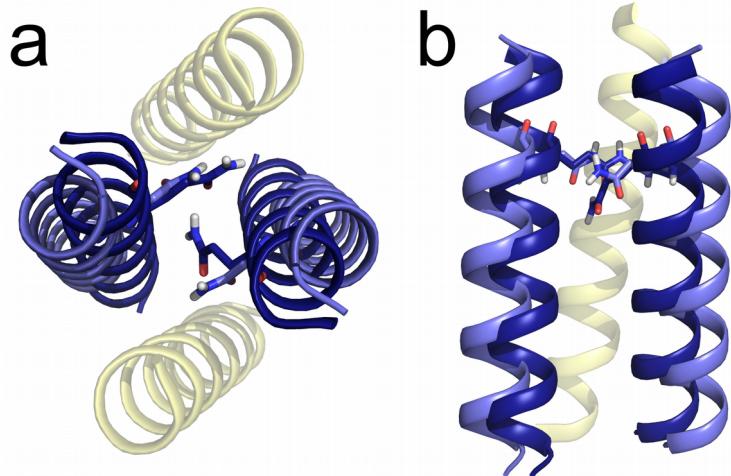
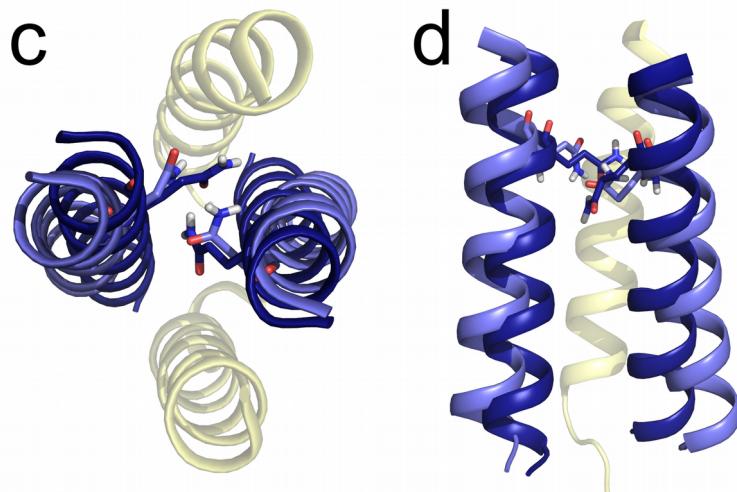


Figure S3.3: Hexameric model of the FtsLB complex. a) Model of the L3B3 hexamer with the co-evolutionary restraints (Table S3.8.1) shown as black dashes. Most of the evolutionary constraints are fit by this model. FtsB rendered in blue, FtsL rendered in yellow, and the juxtamembrane linker in FtsB highlighted in red. b) Solvent-accessible cavities rendered in gray wireframe. These cavities do not form a complete channel and are instead occluded by bulky side chains and the juxtamembrane linker of FtsB. c) and d) Top-down views of the hexameric model presented in a and b.

FtsB dimer vs FtsLB tetramer (initial model)



FtsB dimer vs FtsLB tetramer (MD model)



■ FtsB dimer – ■ FtsB and ■ FtsL in FtsLB

Figure S3.4: The configuration of the TM region of FtsB is similar to a previous model obtained in the absence of FtsL. Structural alignment of the L₂B₂ complex to a model of the FtsB homodimer. a) Top-down and b) side views of an alignment with the L₂B₂ model (blue, yellow) with the model of the homodimer (dark blue), which was previously generated using mutagenesis data based on a genetic

reporter assay of self-association (LaPointe et al., 2013). Residue Q16 is shown in sticks. The interface of FtsB is the same in both models, but the helices are farther apart in the L₂B₂ tetramer. c) and d) Top-down and side views of the FtsB homodimer model aligned with the last frame of the 260ns MD simulation. The FtsB helices have moved closer together, improving the RMSD of the alignment from 2.5 Å to ~2 Å.

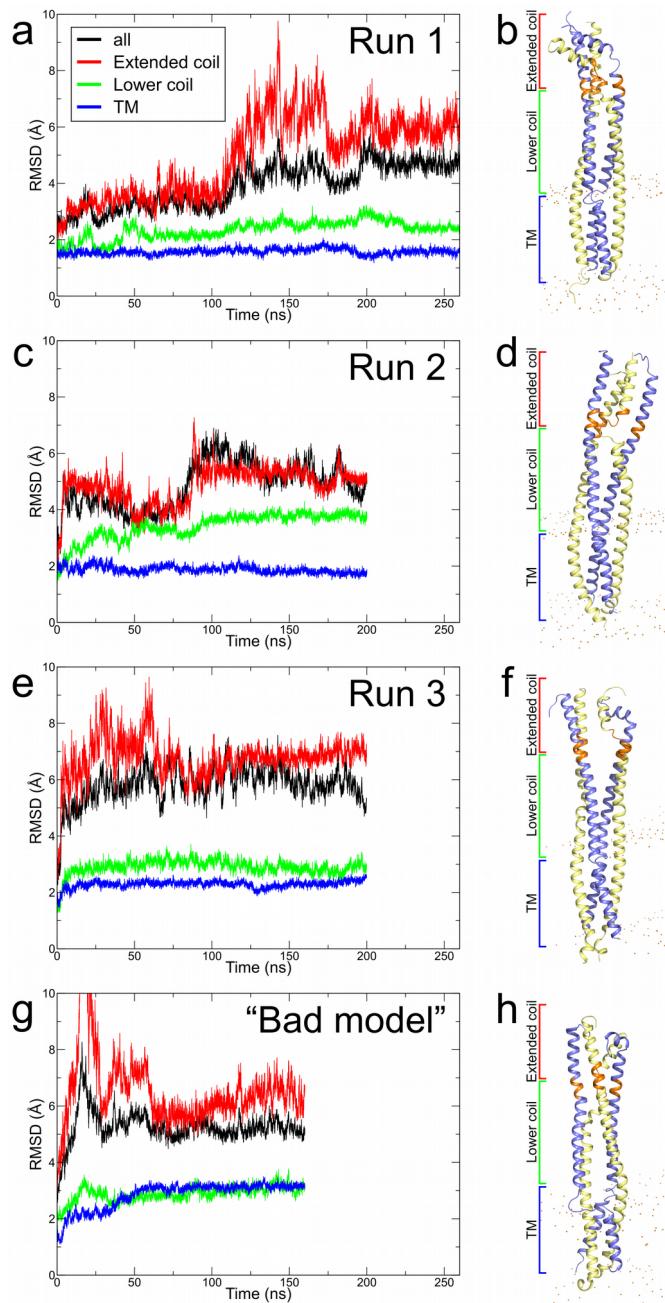


Figure S3.5: RMSD analysis of the three replica MD runs and of a “bad model”.

The figure is similar to Figure 3.5a (Run 1). The RMSD analysis is shown for the entire complex (black), the TM region (blue), the lower coil (green, 29-61 of FtsB

and 57-91 of FtsL) and the extended coil (red, 62-79 of FtsB and 92-110 of FtsL).

The CCD, which separates the lower from the extended coil is in orange. a) Trajectory of the 260 ns Run 1. b) Conformation of the last step of the trajectory. c and d) Trajectory of the 200 ns Run 2. e and f) Trajectory of the 200 ns Run 3. g and h) Trajectory of 160 ns run of a “bad model”, an alternate low energy model in which the conformation of the TM region does not satisfy most of the evolutionary constraints. It is notable how the RMSD of the TM region quickly rises to a higher value (approximately 3 Å, compared to the same variable in Runs 1-3, which is around 2 Å).

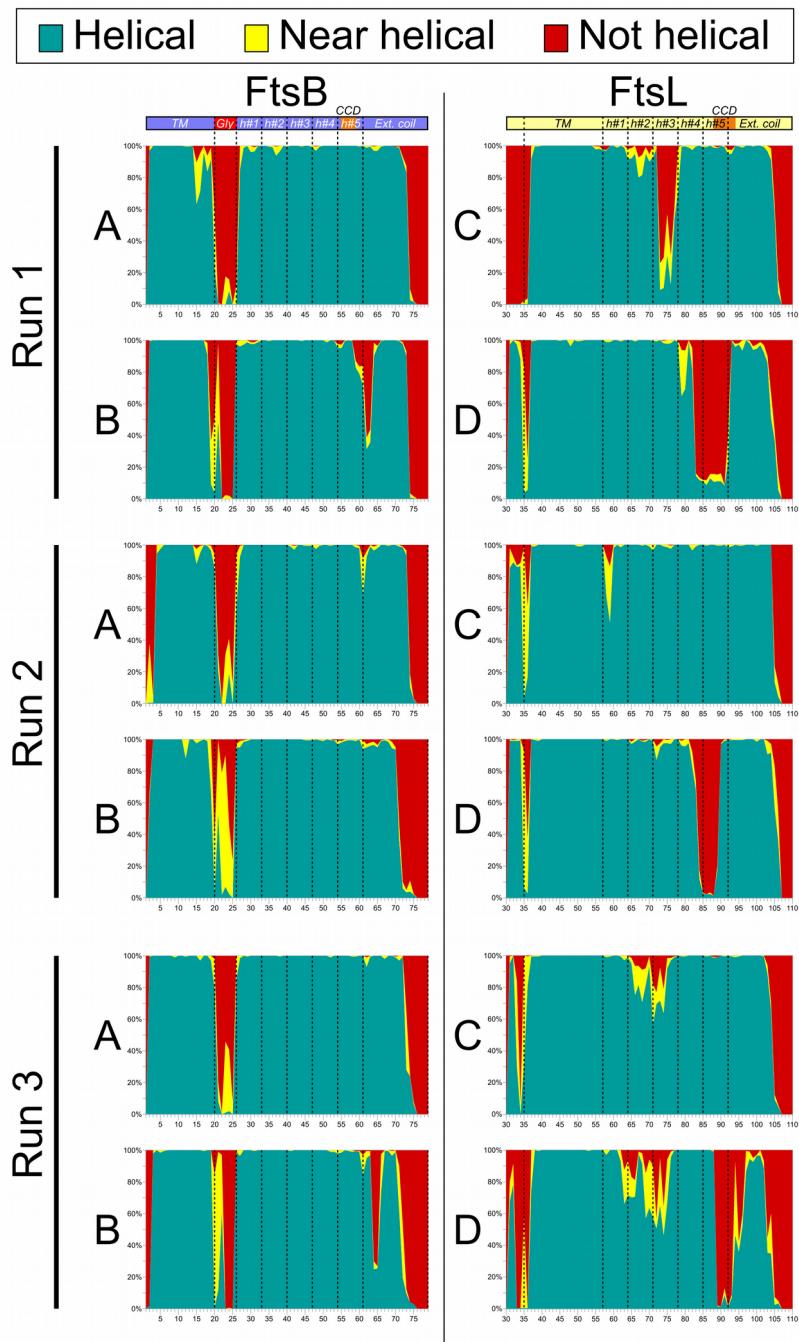


Figure S3.6: Analysis of helicity in the three replica MD runs. The graphs display the fraction of time each position adopted a helical conformation (green), nearly

helical (yellow) or non-helical (red). A and B are the two individual FtsB chains (left panels) and C and D are the two individual FtsL chains (right panels).

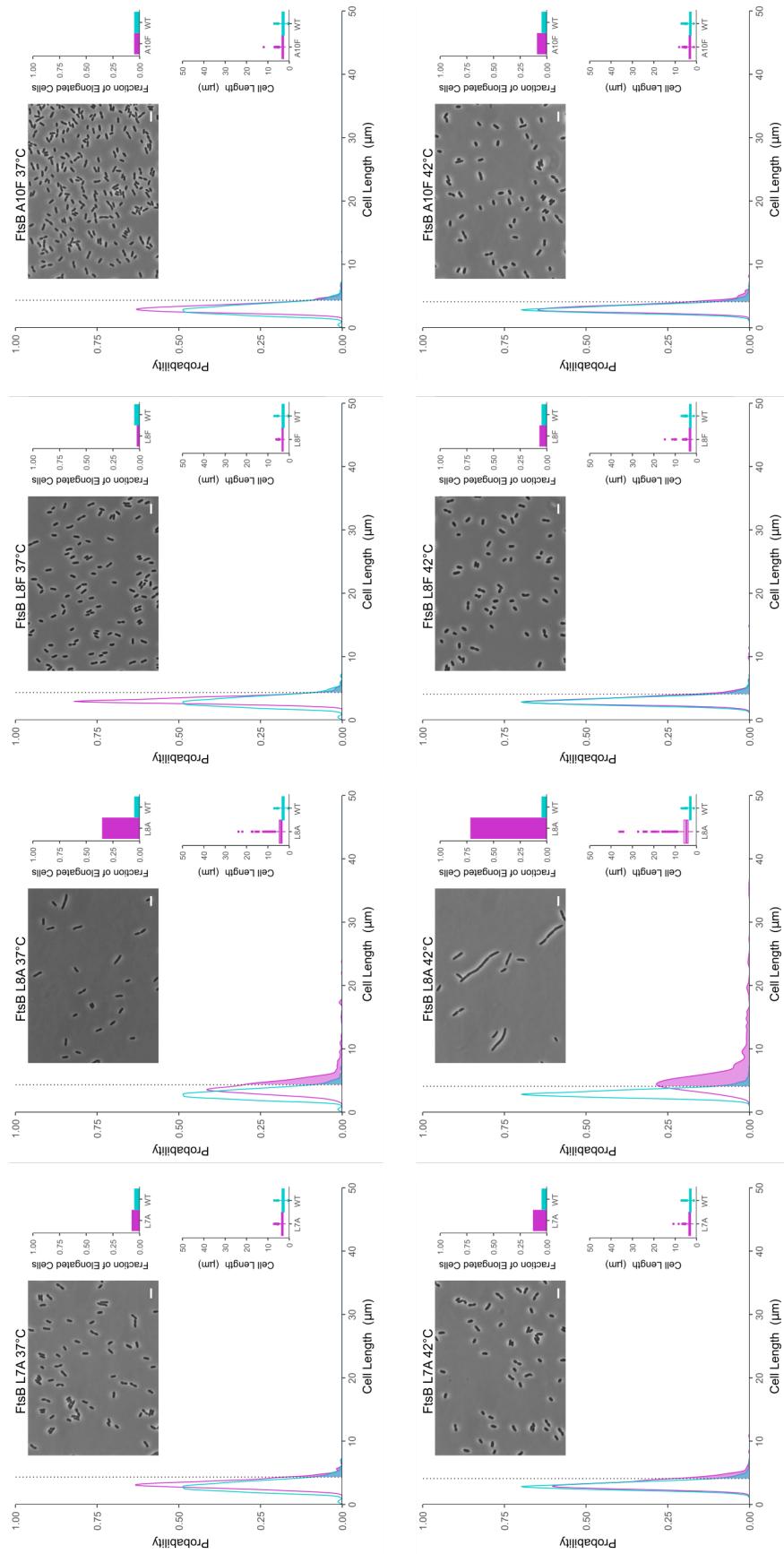


Figure S3.7: Full analysis of cell division phenotypes of point and alanine insertions mutants reported in Figure 3.6 and Fig 7. The distribution of cell lengths for each point substitution or insertion mutant was compared between wild type

(aqua) and mutant (magenta) cell population. Center: kernel density functions of mutant and wild-type cell lengths. The area above the 95th percentile of the wild-type distribution (dotted line) is shaded. Insert: representative photo of mutation. Scale bar corresponds to 5μm. Right, top: fraction of cells above the 95th percentile of wild-type cell length. Right, bottom: box and whisker plots of cell lengths for wild-type and mutant. Whiskers represent the interquartile range multiplied by 1.5.

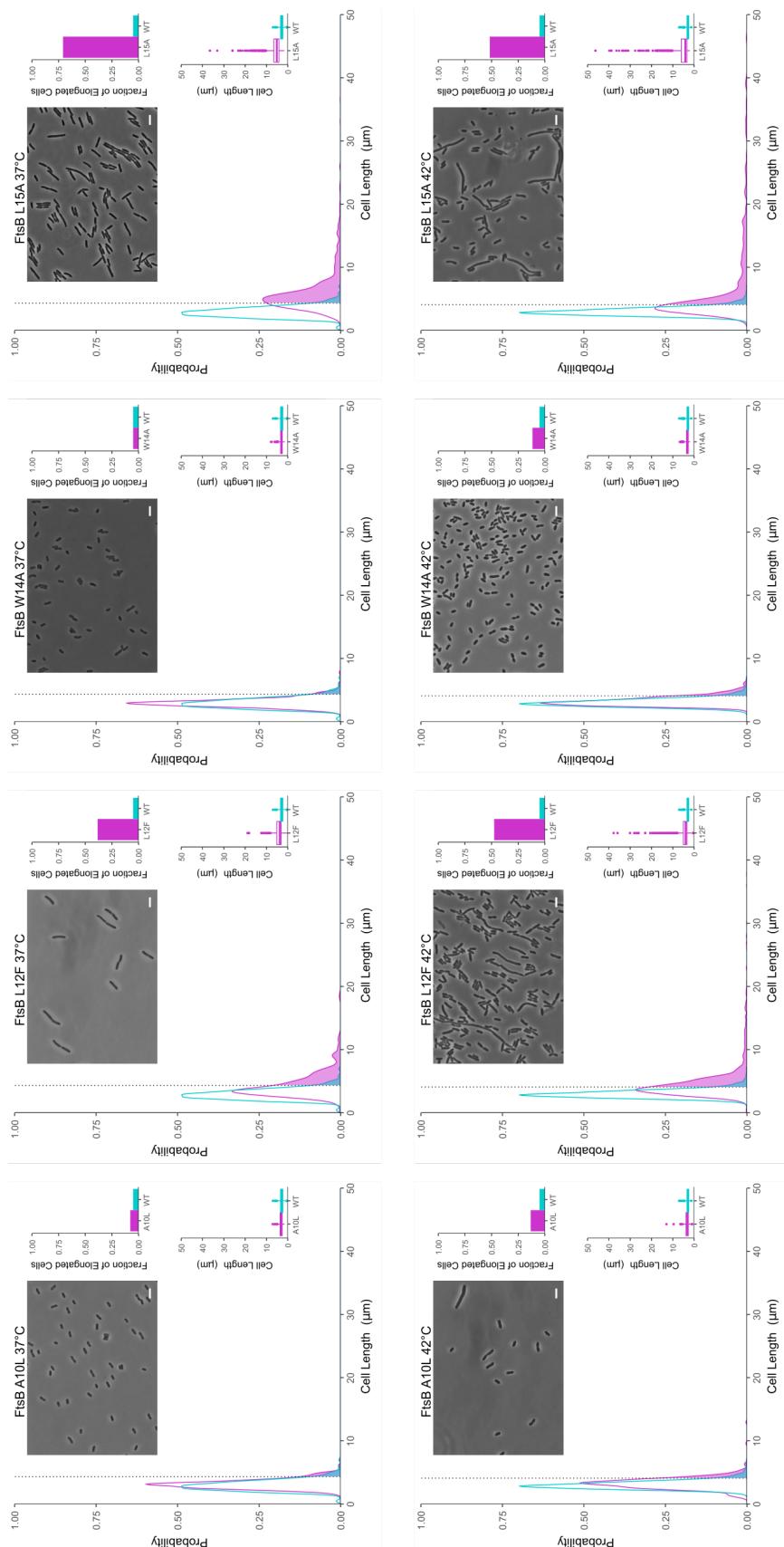


Figure S3.7, continued

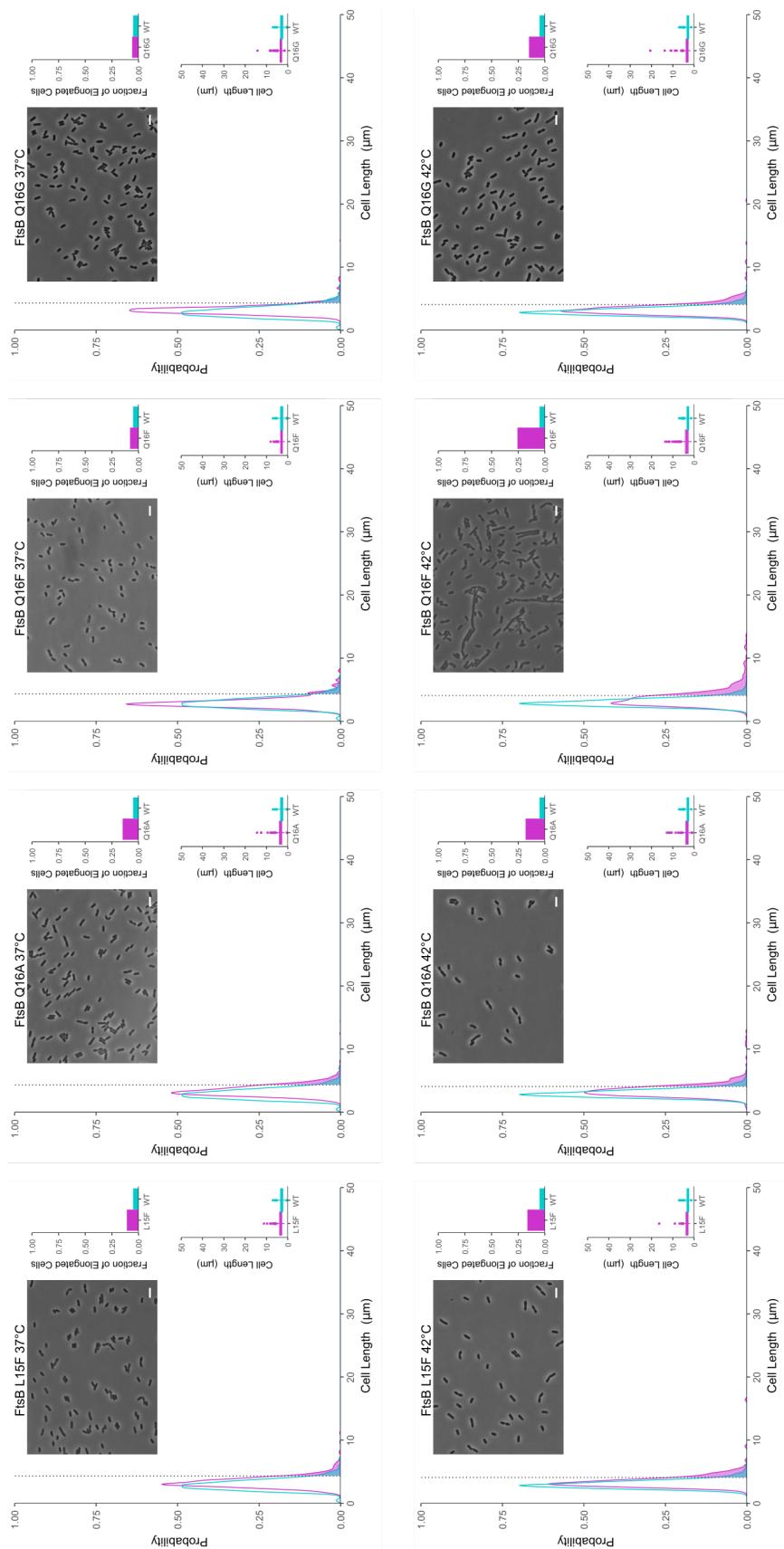


Figure S3.7, continued

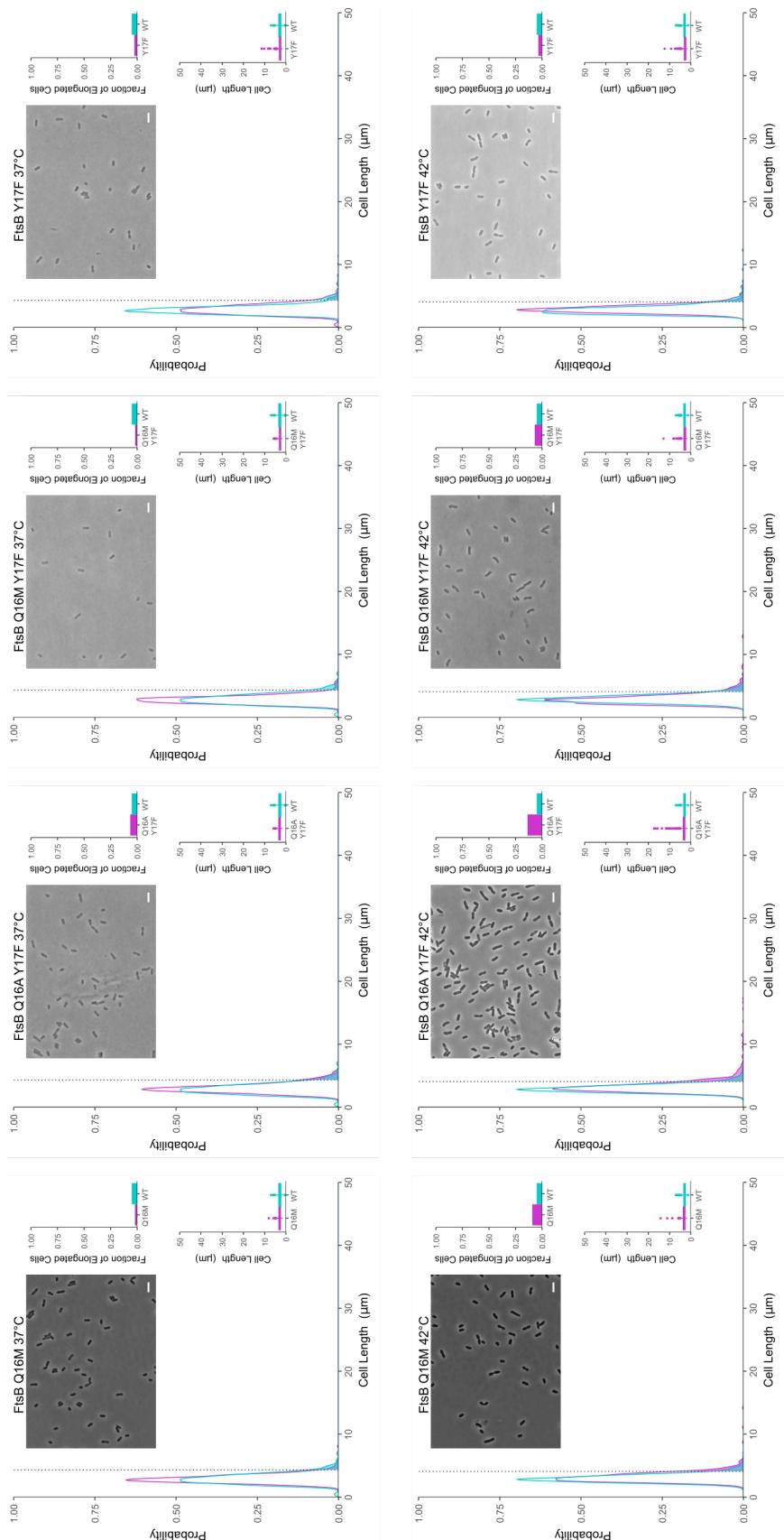


Figure S3.7, continued

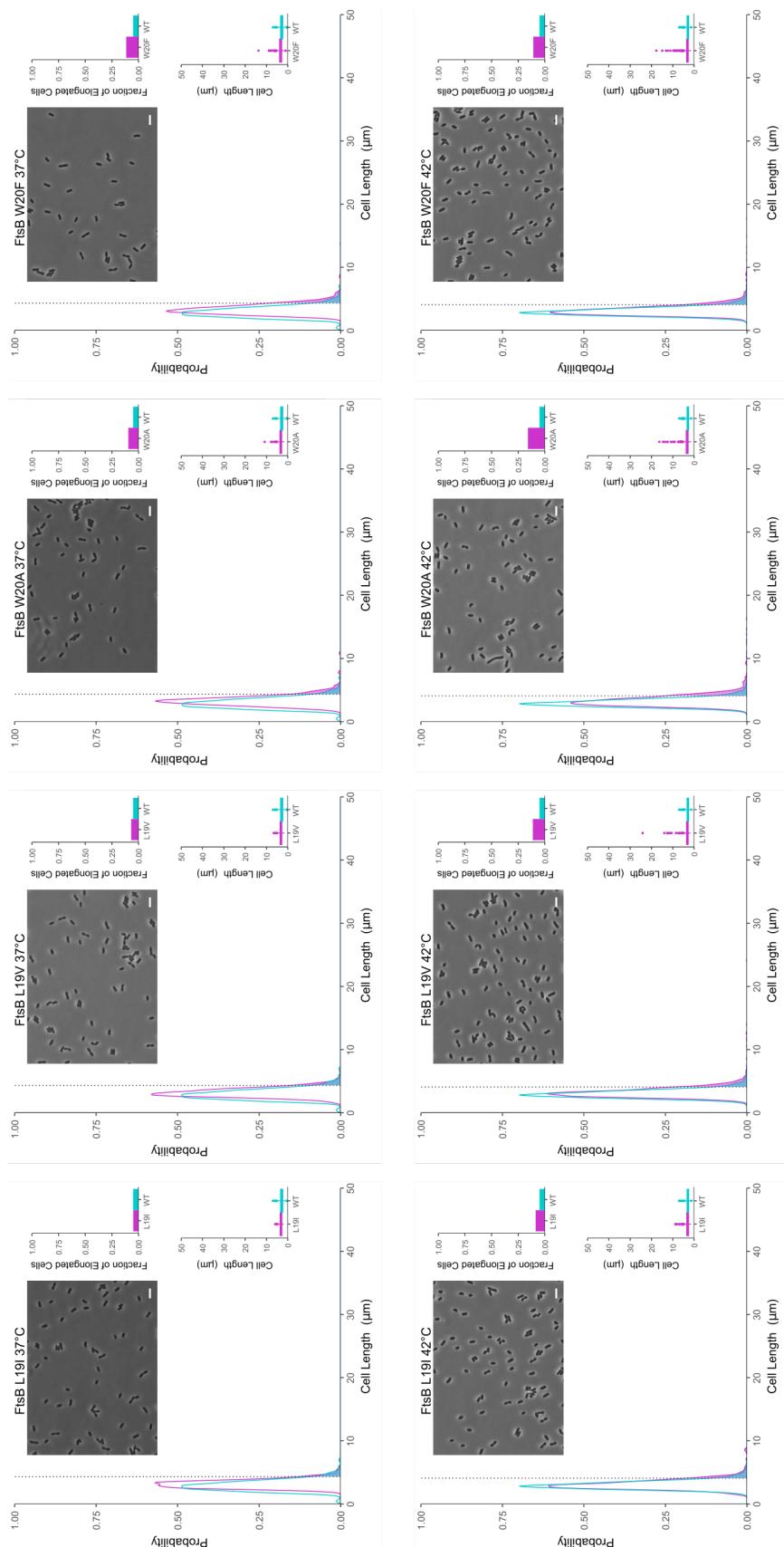


Figure S3.7, continued

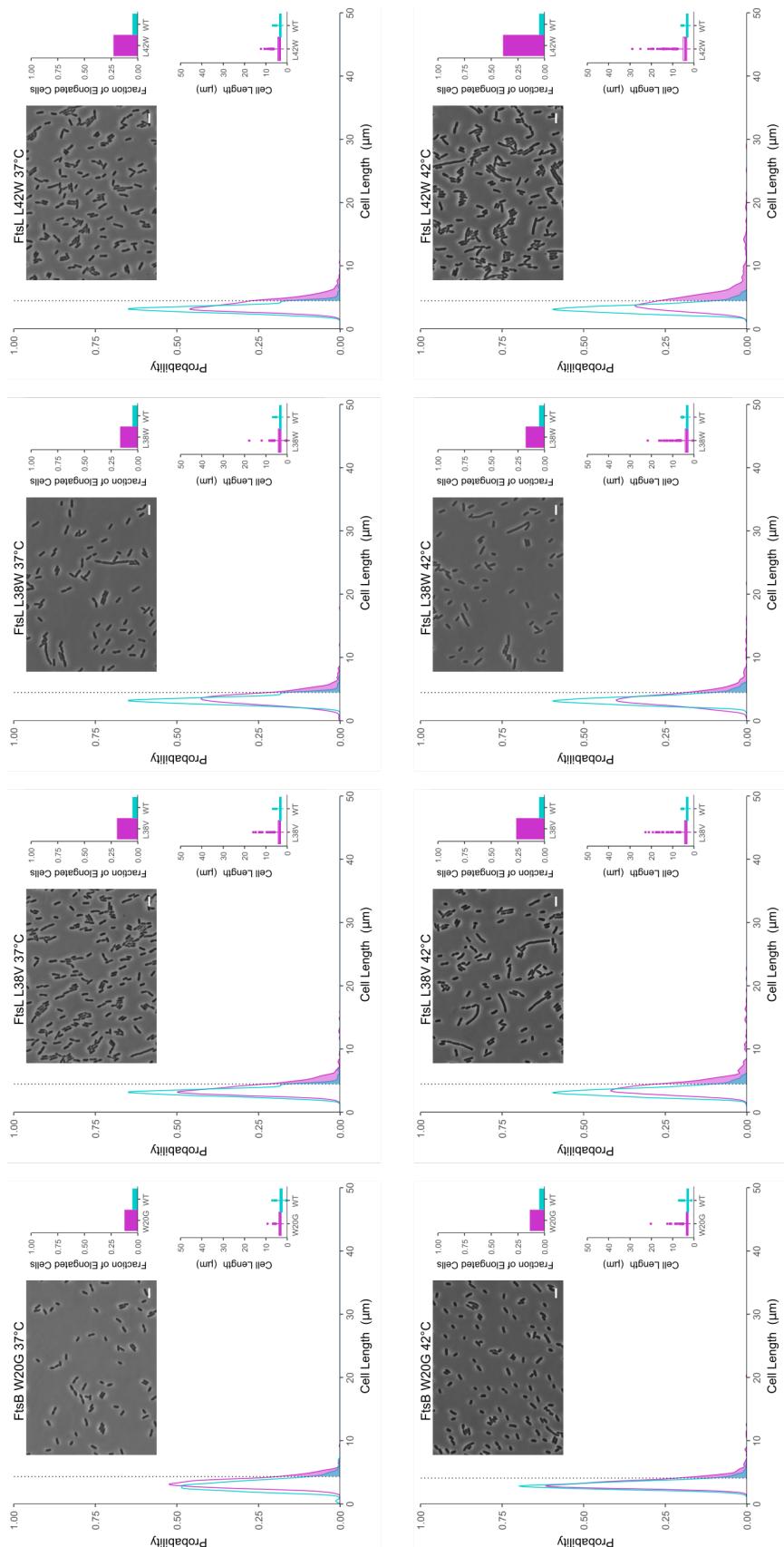


Figure S3.7, continued

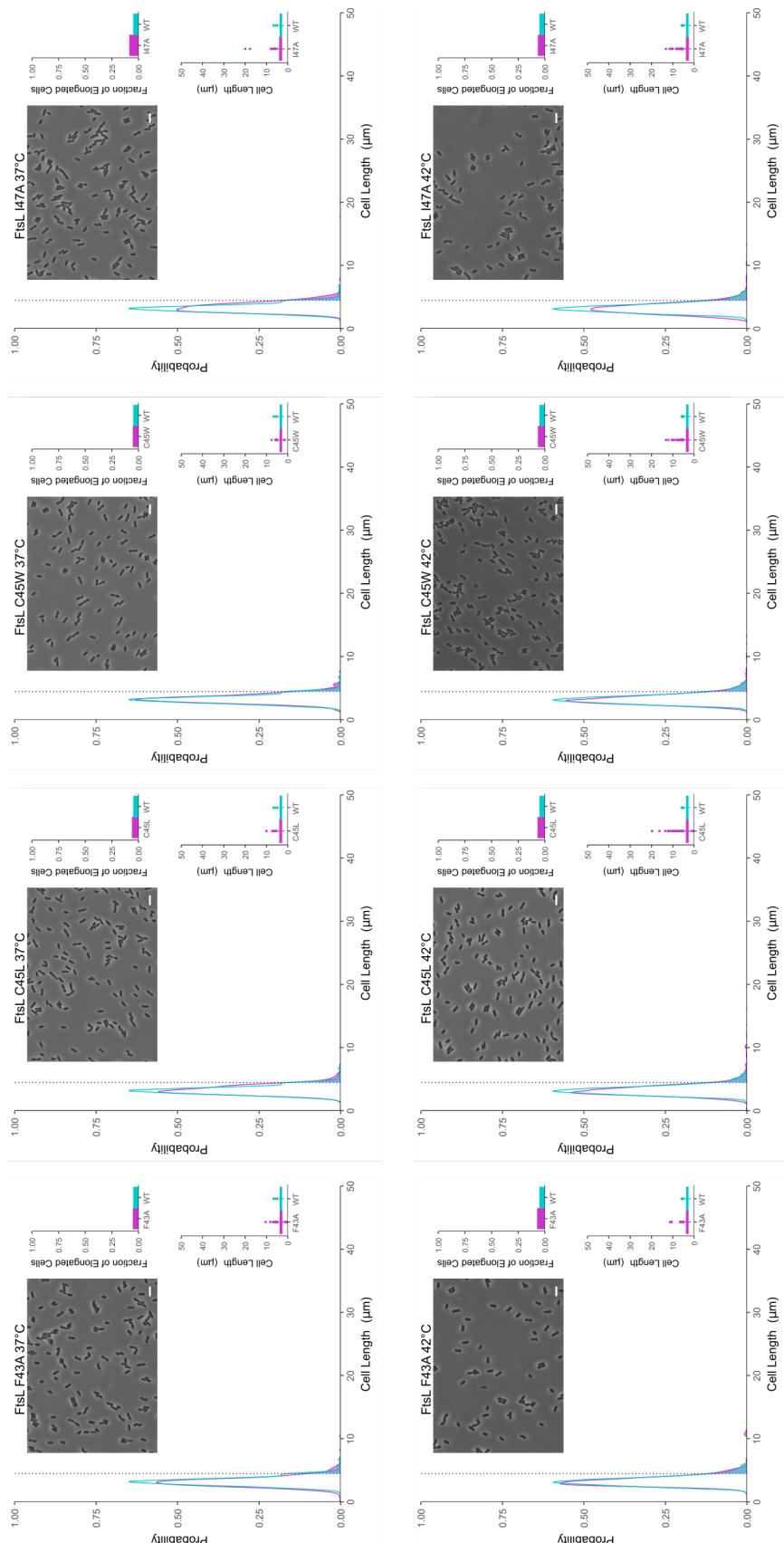


Figure S3.7, continued

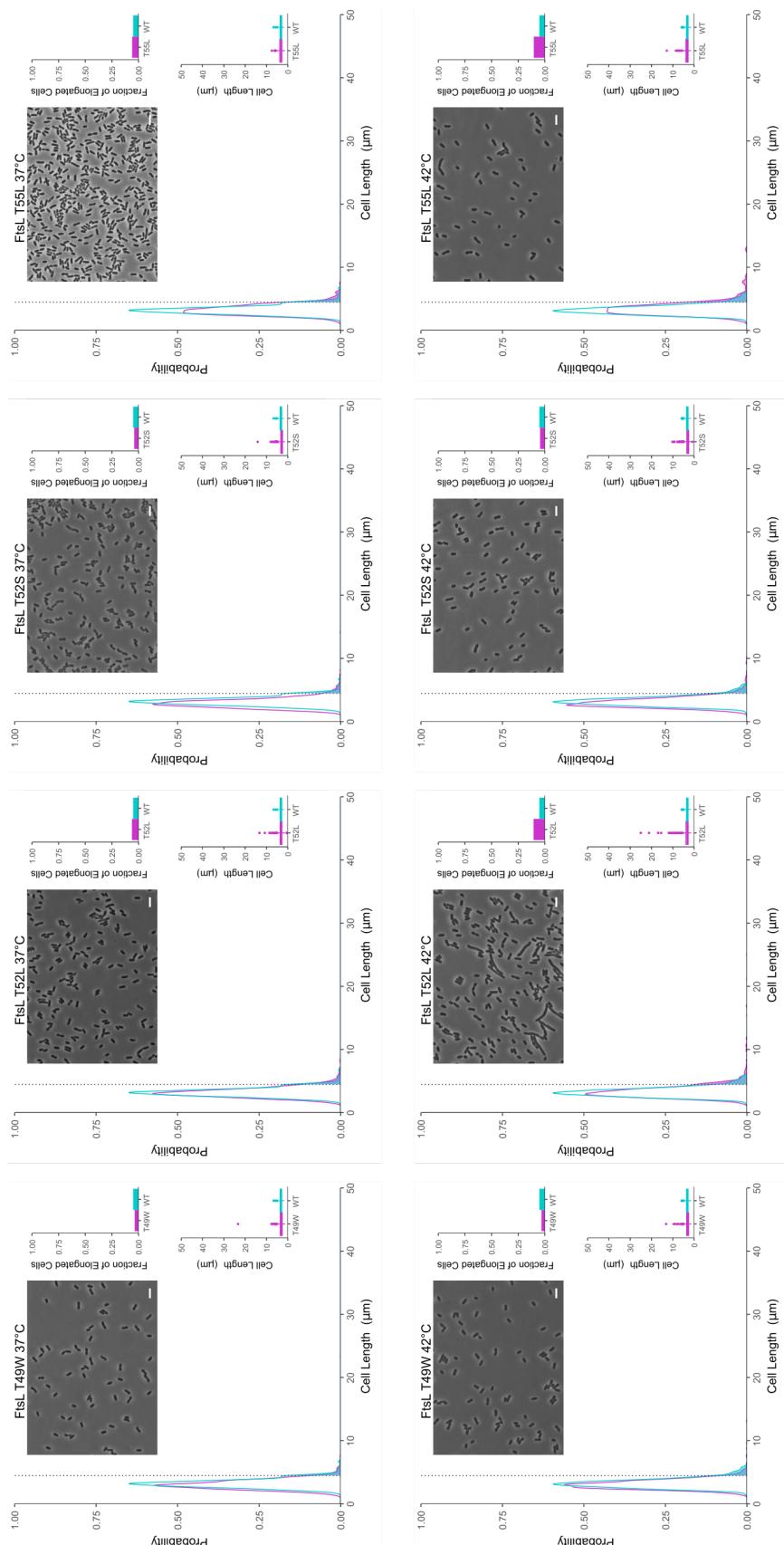


Figure S3.7, continued

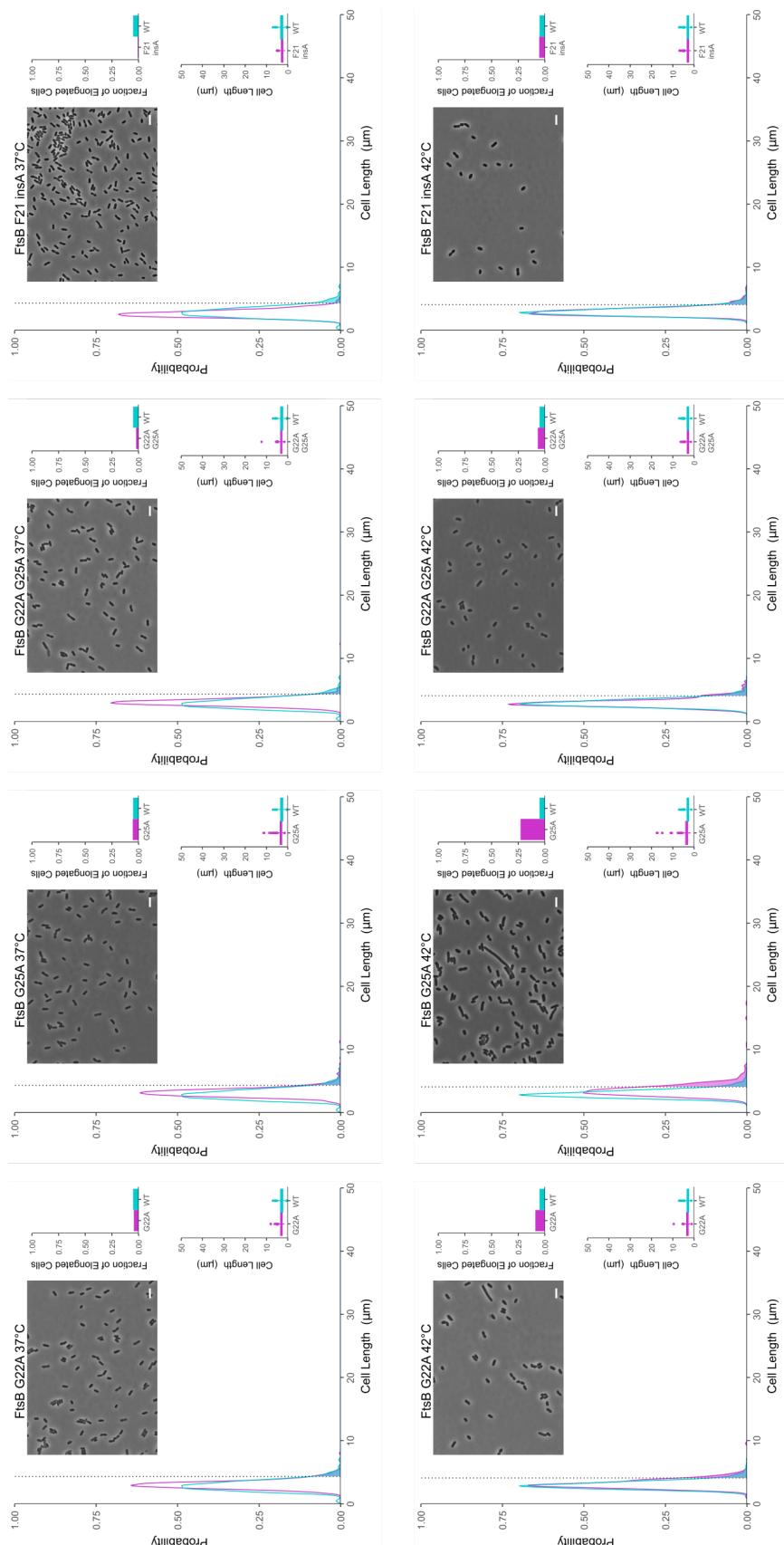


Figure S3.7, continued

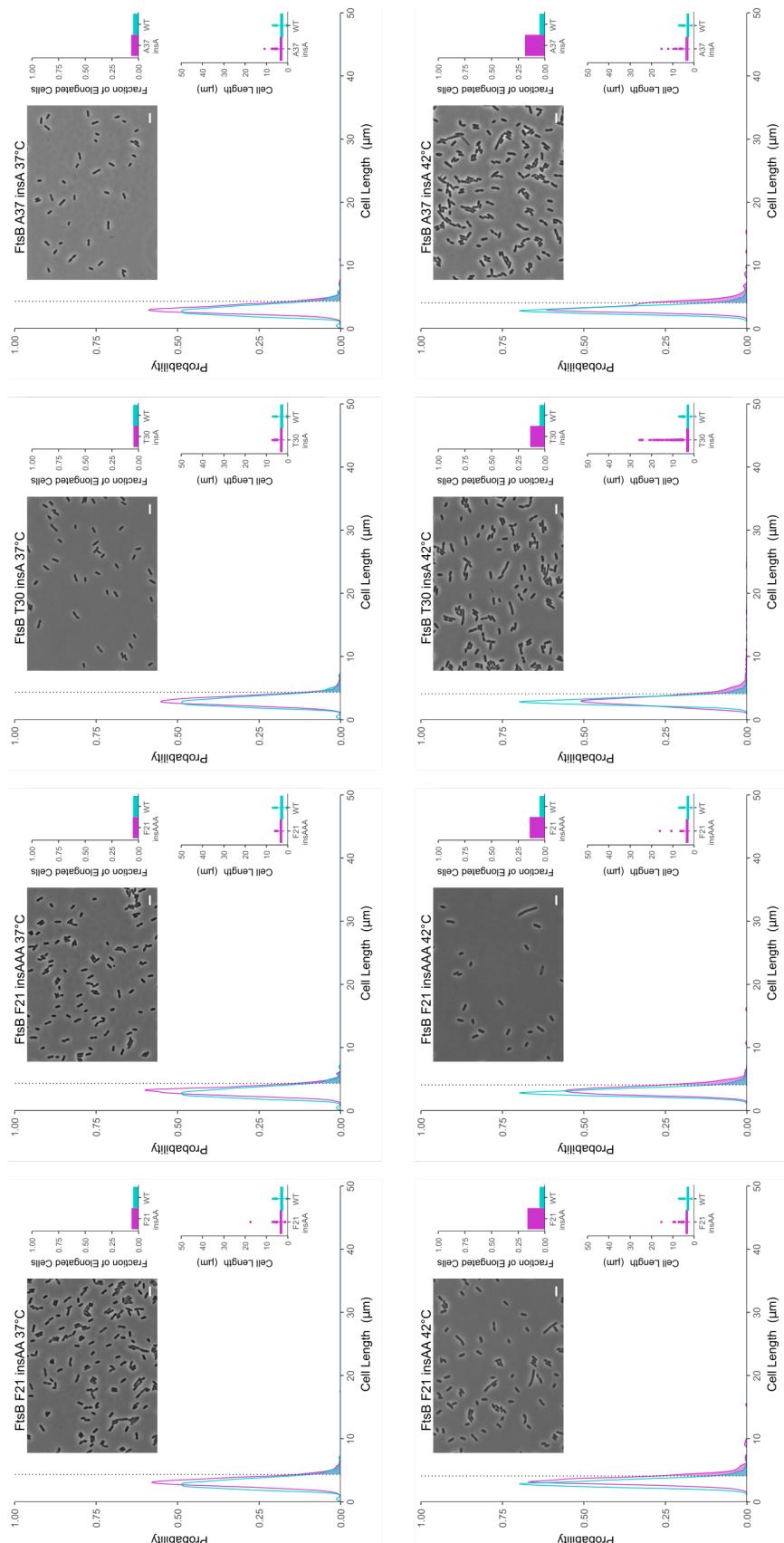


Figure S3.7, continued

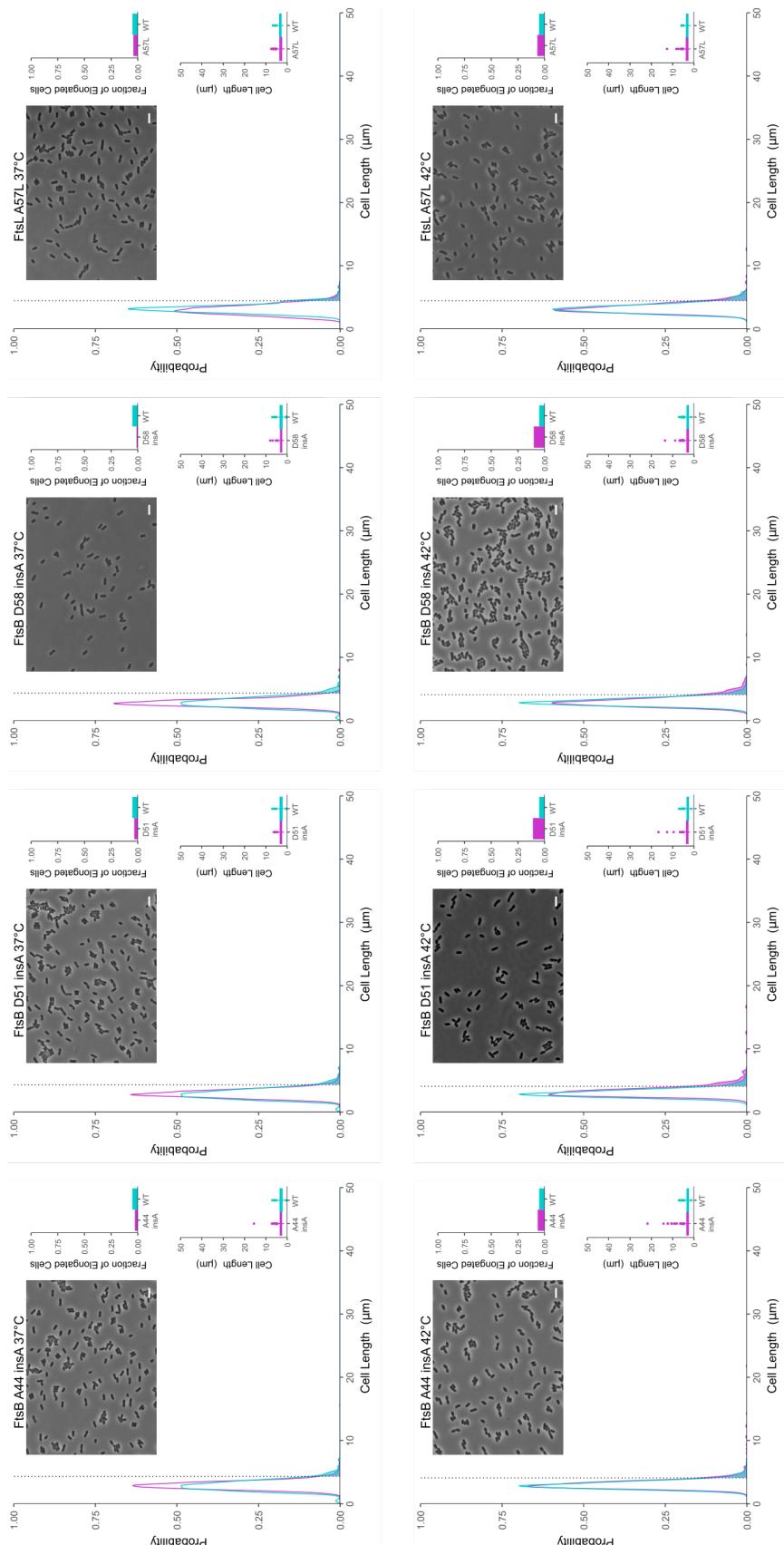


Figure S3.7, continued

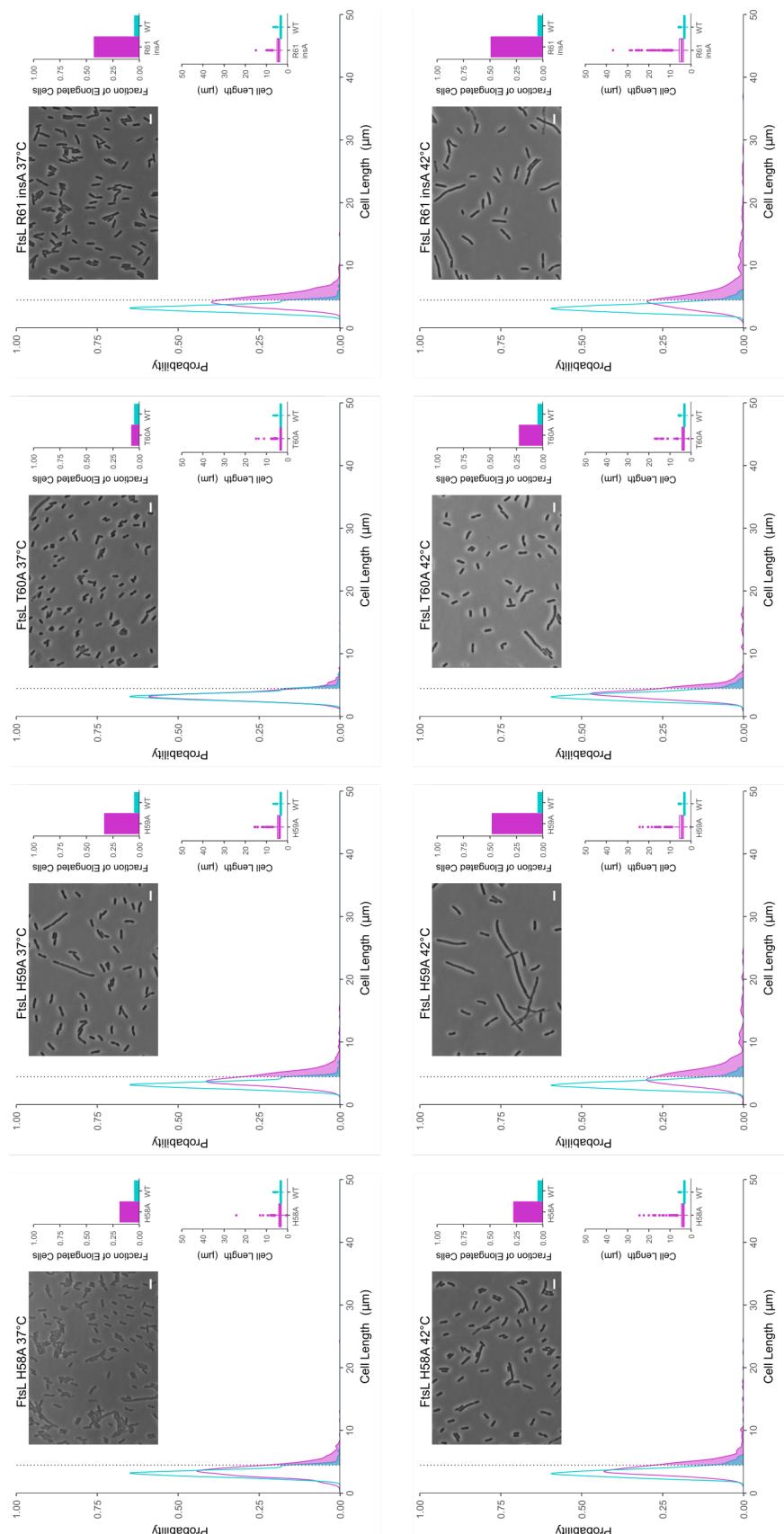


Figure S3.7, continued

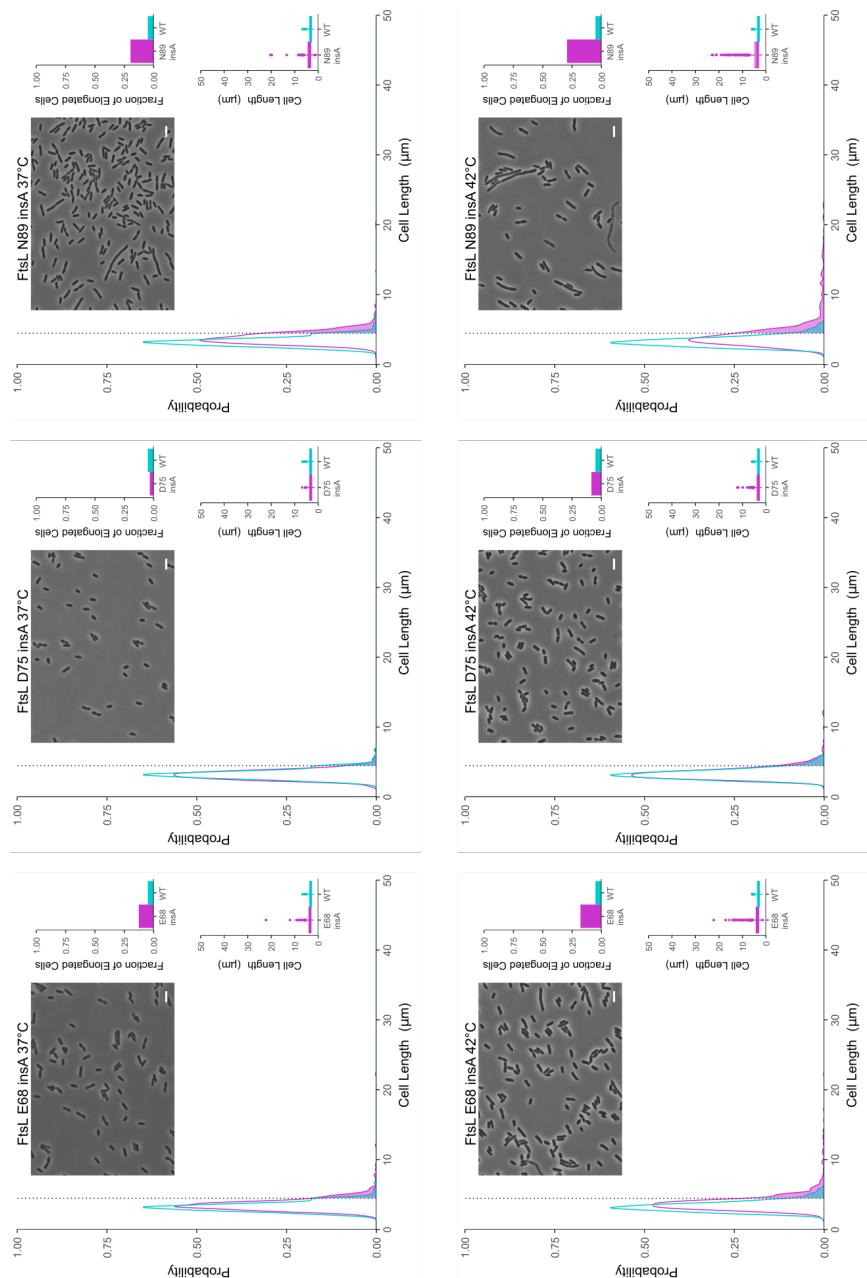


Figure S3.7, continued

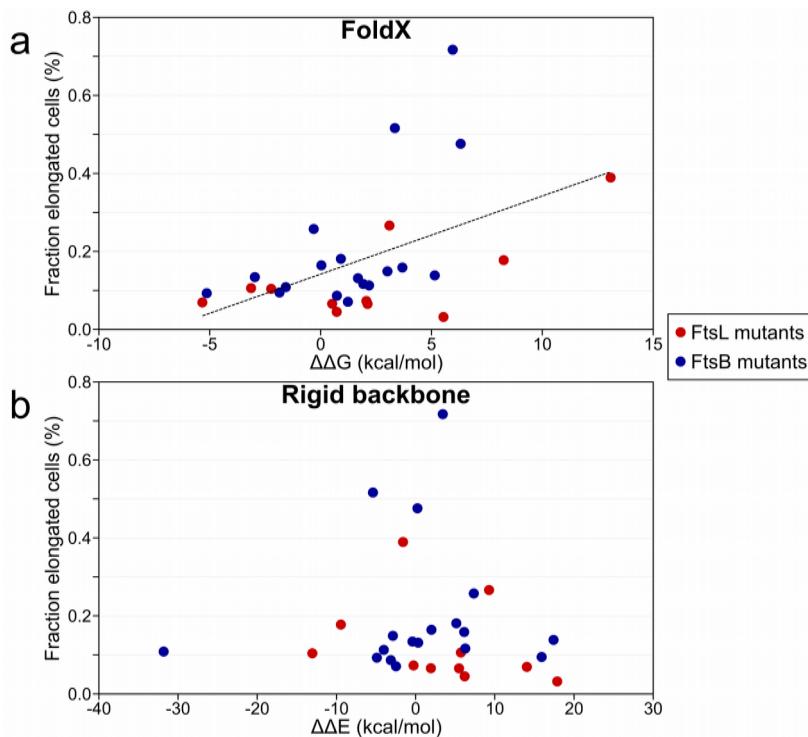


Figure S3.8: Comparison of the predicted mutational energy and the biological phenotypes of the mutants in the transmembrane region. Biological phenotype (expressed as the fraction of elongated cells, as defined in Figure 3.6) compared with the energy of destabilization of each mutant calculated in silico using a) FoldX or b) rigid body repacking in which the mutation was modeled without relaxing the backbone. The energies predicted by FoldX correlate statistically with the severity of the phenotypes by Rank Order Spearman Correlation Coefficient analysis, which does not assume a linear model ($r = 0.4631$, $N = 29$, $p < 0.01$), and by linear regression ($R^2 = 0.25$, $p < 0.01$, dashed line). Correlation with rigid backbone energies is not statistically significant.

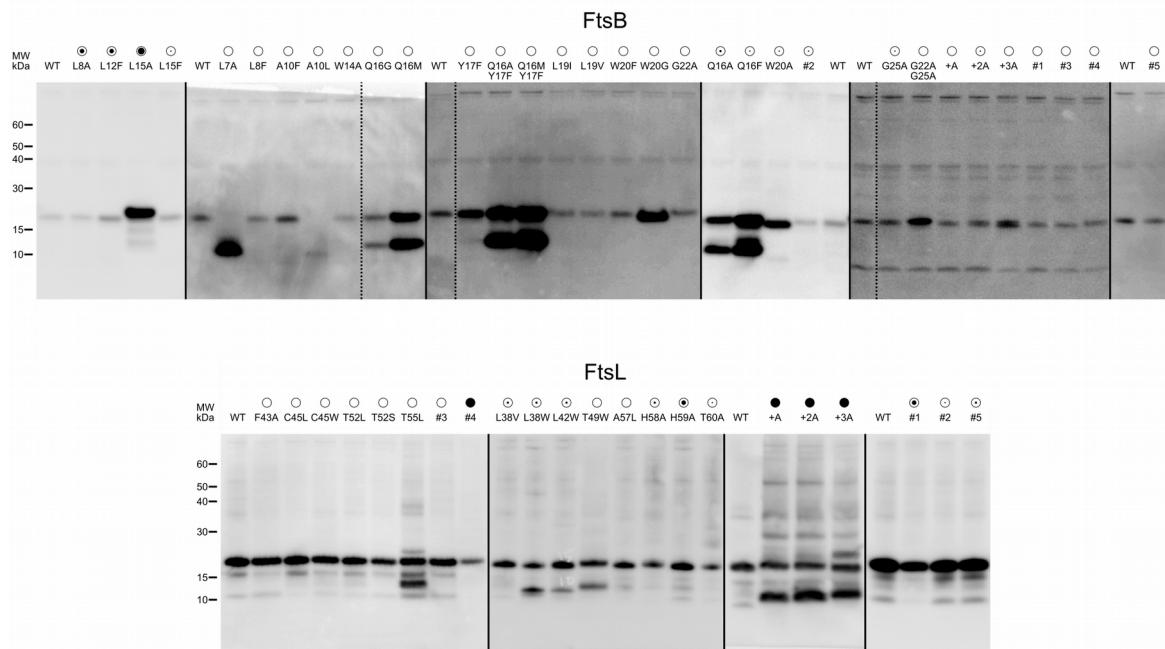


Figure S3.9: Expression level of FtsB and FtsL mutants assessed by Western Blot analysis. Image of all western blots of the FtsB and FtsL mutants tested in this work. The circles indicate the observed phenotypes, as defined in Figure 3.6. The FtsB wild type and its mutants display a notably fainter band with respect to the FtsL variants. Protein expression level of the FtsB and FtsL mutants with defective phenotypes are generally comparable and to the respective wild type. There are cases of FtsB mutants with increased protein level. These correspond to mutations in a specific region of the TM domain of FtsB, involving L15, Q16, Y17 or W20. These mutants with increased protein level correspond to interfacial positions that have a wild-type-like (Q16M, Y17F, Q16M/Y17F, Q16A/Y17F, W20A), mild temperature- ensitive (Q16F), mild (Q16A) phenotypes, or, in one case, a severe phenotype (L15A). It is possible that these apparent increases in

protein levels are due to reduced protein degradation, possibly because the mutations mask a site recognized by a membrane protease. Additionally, two FtsB mutants (L7A and A10L) present faint full-length bands and significant degradation. However, these mutants are viable and are classified as wild-type-like. Individual gels are separated by solid lines. Spliced lanes within the same gel are indicated by a dashed line.

FtsB		FtsL		Co-evo ranking		Structural model	
Position	Heptad ¹	Position	Heptad ¹	Rank ²	Score ³	Min distance (Å) ⁴	Cα distance (Å) ⁵
Other							
65		103		33	0.18404		
67		101		19	0.21003		
68		105		72	0.16013		
68		107		55	0.16661		
79		107		7	0.24894		
80		106		1	0.41824		
80		108		53	0.16726		
83		116		45	0.17608		
87		121		67	0.16224		
7		112		51	0.16889		
7		117		14	0.21928		
31		113		80	0.15693		
37		46		89	0.15313		
37		54		77	0.1583		
38		40		63	0.16335		
48		111		48	0.17405		
64		44		94	0.15196		
70		88		32	0.18635		
74		37		24	0.19954		
88		52		84	0.15532		
92		113		91	0.15285		
94		48		42	0.17769		
94		99		93	0.1522		
95		29		66	0.16249		
95		30		60	0.16504		
95		33		79	0.15708		

Table S3.8.1: Top 95 predicted co-evolutionary contacts between FtsL and FtsB

¹Residue assignment to a canonical coiled-coil heptad repeat, assuming a continuous transmembrane-coiled-coil helix.

²Pair ranking originated from the EV-Couplings algorithm.

³Pseudo-likelihood maximization (PLM) score computed for the pair of positions from the EV-Couplings algorithm.

FtsB		FtsL		Co-evo ranking		Structural model	
Position	Heptad ¹	Position	Heptad ¹	Rank ²	Score ³	Min distance (Å) ⁴	Cα distance (Å) ⁵
TM Domain							
T 5 e		P 39 a		28	0.18971	4.76	6.43
L 12 e		C 45 g		8	0.23962	3.61	7.56
L 12 e		I 46 a		41	0.17785	3.71	5.86
L 15 a		I 46 a		5	0.27006	4.01	8.11
L 15 a		V 53 a		86	0.15394	4.61	8.4
Q 16 b		T 49 d		12	0.23363	3.08	6.86
L 19 e		T 52 g		78	0.15754	3.58	8.4
L 19 e		T 56 d		75	0.15967	3.49	7.66
Coiled coil							
I 26 e		Q 66 g		76	0.15946	11.22**	16.52
Y 29 a		T 56 d		15	0.21667	7.97*	10.12
Y 29 a		T 60 a		39	0.17999	5.91*	8.69
V 32 d		T 64 e		22	0.20568	4.54	7.28
D 35 g		R 67 a		23	0.20543	4.03	8.08
D 35 g		V 71 e		27	0.19052	4.46	7.78
V 36 a		L 63 d		10	0.23707	5.38*	9.22
V 36 a		L 70 d		11	0.23583	3.78	7.7
Q 39 d		T 64 e		59	0.16512	9.17*	14.52
Q 39 d		R 67 a		74	0.15968	3.37	10.34
Q 40 e		Q 66 g		16	0.21574	3.54	9.58
Q 40 e		L 70 d		25	0.19574	3.48	6.94
N 43 a		R 74 a		62	0.16375	3.33	7.96
K 47 e		E 73 g		3	0.3737	3.88	9.46
R 49 g		W 81 a		6	0.25888	2.78	8.18
L 53 d		W 81 a		49	0.1721	3.52	10.29
I 57 a		E 80 g		29	0.1897	8.62*	13.43
N 61 e		E 87 g		13	0.22554	3.78	9.25
N 61 e		L 91 d		2	0.41222	4.06	6.67

Table S3.8.1, continued

⁴Minimum distance between heavy atoms of the residues in the computational model of the L₂B₂ tetramer. Asterisks denote outliers not well satisfied by the computational model. Single asterisk: distances above 5 Å; double asterisks: distances above 10 Å.

⁵Minimum distance between alpha carbons of the residues, which were used to compute sigmoidal distance restraints during the modeling procedure.

Chain	Position	Chain	Position²	Occupancy
A	Y17	C	T52	92.00%
A	W20	C	T56	81.00%
A	Q16	C	T49	81.00%
A	Q16	B	Q16	62.00%
B	Q16	B	W20	39.00%
B	Y17	D	L48b	31.00%
A	Q16	C	T52	26.00%
B	Q16	C	T49	18.00%
A	Y17	C	L48b	13.00%
A	Q16	A	Y17	11.00%
B	Y17	D	T52	10.00%
B	W20	A	Q16b	8.00%
B	Q16	D	T49	2.00%
A	Q16	A	V13b	1.00%

Table S3.8.2: Network of hydrogen bonds observed in TM domain of FtsLB during molecular dynamics simulation, run 1¹

¹The most frequently observed side-chain to side-chain or side-chain to backbone hydrogen bonding interactions in the molecular dynamics simulations are listed. Chains A and B correspond to FtsB while chains C and D correspond to FtsL. The percentage is cumulative of all possible combinations of donor-acceptor interactions between the two residues. Co-occurring interactions between different donors and acceptors within the same pair of residues are not double-counted.

²The letter “b” indicates that the hydrogen bonding involves a backbone group

Name	Sequence ¹
His ₆ -FtsB	<i>MGS</i> HHHHHHHHDYDIP <u>TSENLYFQGGSGMGKLTLLL</u> A <u>ILVWLQYSLWF</u> GKNGIHDYTRVNDDVAQQATNAKLKARNDQLFAEIDDLNGQEALEER ARNELSMTRPGETFYRLVPDASKRAQSAGQNNR
His ₆ -FtsB-S97C	<i>MGS</i> HHHHHHHHDYDIP <u>TSENLYFQGGSGMGKLTLLL</u> A <u>ILVWLQYSLWF</u> GKNGIHDYTRVNDDVAQQATNAKLKARNDQLFAEIDDLNGQEALEER ARNELSMTRPGETFYRLVPDASKRAQ C AGQNNR
Strep-FtsL ₃₅₋₁₂₁ -C41A-C45A (Cys-less FtsL)	<i>MSNWSHPQFEKDYD</i> <u>IPTSENLYFQGGSGFGKLPL</u> A LFIA I IILTAVTVVT TAHHTRLLTAQREQLVLERDALDIEWRNLILEENALGDHSRVERIATEK LQM Q HVDPSQENIVVQK
Strep-FtsL ₃₅₋₁₂₁ -C41A-C45A-I100C	<i>MSNWSHPQFEKDYD</i> <u>IPTSENLYFQGGSGFGKLPL</u> A LFIA I IILTAVTVVT TAHHTRLLTAQREQLVLERDALDIEWRNLILEENALGDHSRVER C ATEK LQM Q HVDPSQENIVVQK

Table S3.8.3: Sequences of FtsB and FtsL used for *in vitro* FRET experiments.

1The sequences of the added purification tags (His₆ or Strep) are in italics.

Mutations are highlighted in bold and underlined.

3.9. References

- Aarsman, M.E.G., Piette, A., Fraipont, C., Vinkenvleugel, T.M.F., Nguyen-Distèche, M., and den Blaauwen, T. (2005). Maturation of the *Escherichia coli* divisome occurs in two steps: Maturation of the divisome. *Mol. Microbiol.* 55, 1631–1645.
- Adair, B.D., and Engelman, D.M. (1994). Glycophorin A Helical Transmembrane Domains Dimerize in Phospholipid Bilayers: A Resonance Energy Transfer Study. *Biochemistry* 33, 5539–5544.
- Alexeeva, S., Gadella, T.W.J., Verheul, J., Verhoeven, G.S., and Den Blaauwen, T. (2010). Direct interactions of early and late assembling division proteins in *Escherichia coli* cells resolved by FRET: Bacterial division studied by spectral FRET. *Mol. Microbiol.* 77, 384–398.
- Alva, V., Nam, S.-Z., Söding, J., and Lupas, A.N. (2016). The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. *Nucleic Acids Res.* 44, W410–W415.
- van den Berg van Saparoea, H.B., Glas, M., Vernooij, I.G.W.H., Bitter, W., den Blaauwen, T., and Luijink, J. (2013). Fine-mapping the Contact Sites of the *Escherichia coli* Cell Division Proteins FtsB and FtsL on the FtsQ Protein. *J. Biol. Chem.* 288, 24340–24350.
- Bisson-Filho, A.W., Hsu, Y.-P., Squyres, G.R., Kuru, E., Wu, F., Jukes, C., Sun, Y., Dekker, C., Holden, S., VanNieuwenhze, M.S., et al. (2017). Treadmilling by FtsZ filaments drives peptidoglycan synthesis and bacterial cell division. *Science* 355, 739–743.
- de Boer, P.A. (2010). Advances in understanding *E. coli* cell fission. *Curr. Opin. Microbiol.* 13, 730–737.
- Bramkamp, M., Weston, L., Daniel, R.A., and Errington, J. (2006). Regulated intramembrane proteolysis of FtsL protein and the control of cell division in *Bacillus subtilis*: FtsL proteolysis in *Bacillus subtilis*. *Mol. Microbiol.* 62, 580–591.
- Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* 4, 187–217.
- Brown, J.H., Cohen, C., and Parry, D.A. (1996). Heptad breaks in alpha-helical coiled coils: stutters and stammers. *Proteins* 26, 134–145.
- Buddelmeijer, N., and Beckwith, J. (2004). A complex of the *Escherichia coli* cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region:

Assembly of a subcomplex of cell division proteins in *E. coli*. *Mol. Microbiol.* 52, 1315–1327.

Buddelmeijer, N., Judson, N., Boyd, D., Mekalanos, J.J., and Beckwith, J. (2002). YgbQ, a cell division protein in *Escherichia coli* and *Vibrio cholerae*, localizes in codependent fashion with FtsL to the division site. *Proc. Natl. Acad. Sci.* 99, 6316–6321.

Busiek, K.K., Eraso, J.M., Wang, Y., and Margolin, W. (2012). The Early Divisome Protein FtsA Interacts Directly through Its 1c Subdomain with the Cytoplasmic Domain of the Late Divisome Protein FtsN. *J. Bacteriol.* 194, 1989–2000.

Chen, J.C., and Beckwith, J. (2001). FtsQ, FtsL and FtsI require FtsK, but not FtsN, for co-localization with FtsZ during *Escherichia coli* cell division. *Mol. Microbiol.* 42, 395–413.

Cho, H., Wivagg, C.N., Kapoor, M., Barry, Z., Rohs, P.D.A., Suh, H., Marto, J.A., Garner, E.C., and Bernhardt, T.G. (2016). Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. *Nat. Microbiol.* 1, 16172.

Daniel, R.A., and Errington, J. (2000). Intrinsic instability of the essential cell division protein FtsL of *Bacillus subtilis* and a role for DivIB protein in FtsL turnover. *Mol. Microbiol.* 36, 278–289.

Daniel, R.A., Noirot-Gros, M.-F., Noirot, P., and Errington, J. (2006). Multiple Interactions between the Transmembrane Division Proteins of *Bacillus subtilis* and the Role of FtsL Instability in Divisome Assembly. *J. Bacteriol.* 188, 7396–7404.

Delorenzi, M., and Speed, T. (2002). An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. *Bioinformatics* 18, 617–625.

Di Lallo, G. (2003). Use of a two-hybrid assay to study the assembly of a complex multicomponent protein machinery: bacterial septosome differentiation. *Microbiology* 149, 3353–3359.

Du, S., Pichoff, S., and Lutkenhaus, J. (2016). FtsEX acts on FtsA to regulate divisome assembly and activity. *Proc. Natl. Acad. Sci.* 113, E5052–E5061.

Egan, A.J.F., and Vollmer, W. (2015). The stoichiometric divisome: a hypothesis. *Front. Microbiol.* 6.

Emami, K., Guyet, A., Kawai, Y., Devi, J., Wu, L.J., Allenby, N., Daniel, R.A., and Errington, J. (2017). RodA as the missing glycosyltransferase in *Bacillus subtilis* and antibiotic discovery for the peptidoglycan polymerase pathway. *Nat. Microbiol.* 2, 16253.

- van den Ent, F., Vinkenvleugel, T.M.F., Ind, A., West, P., Veprintsev, D., Nanninga, N., den Blaauwen, T., and Löwe, J. (2008). Structural and mutational analysis of the cell division protein FtsQ. *Mol. Microbiol.* **68**, 110–123.
- Erickson, H.P., Anderson, D.E., and Osawa, M. (2010). FtsZ in Bacterial Cytokinesis: Cytoskeleton and Force Generator All in One. *Microbiol. Mol. Biol. Rev.* **74**, 504–528.
- Ghigo, J.-M., and Beckwith, J. (2000). Cell Division in *Escherichia coli*: Role of FtsL Domains in Septal Localization, Function, and Oligomerization. *J. Bacteriol.* **182**, 116–129.
- Ghigo, J.-M., Weiss, D.S., Chen, J.C., Yarrow, J.C., and Beckwith, J. (1999). Localization of FtsL to the *Escherichia coli* septal ring. *Mol. Microbiol.* **31**, 725–737.
- Glas, M., van den Berg van Saparoea, H.B., McLaughlin, S.H., Roseboom, W., Liu, F., Koningstein, G.M., Fish, A., den Blaauwen, T., Heck, A.J.R., de Jong, L., et al. (2015). The Soluble Periplasmic Domains of *Escherichia coli* Cell Division Proteins FtsQ/FtsB/FtsL Form a Trimeric Complex with Submicromolar Affinity. *J. Biol. Chem.* **290**, 21498–21509.
- Goehring, N.W. (2005). Premature targeting of a cell division protein to midcell allows dissection of divisome assembly in *Escherichia coli*. *Genes Dev.* **19**, 127–137.
- Goehring, N.W., Gonzalez, M.D., and Beckwith, J. (2006). Premature targeting of cell division proteins to midcell reveals hierarchies of protein interactions involved in divisome assembly: Assembly of *Escherichia coli* division apparatus. *Mol. Microbiol.* **61**, 33–45.
- Gonzalez, M.D., and Beckwith, J. (2009). Divisome under Construction: Distinct Domains of the Small Membrane Protein FtsB Are Necessary for Interaction with Multiple Cell Division Proteins. *J. Bacteriol.* **191**, 2815–2825.
- Gonzalez, M.D., Akbay, E.A., Boyd, D., and Beckwith, J. (2010). Multiple Interaction Domains in FtsL, a Protein Component of the Widely Conserved Bacterial FtsLBQ Cell Division Complex. *J. Bacteriol.* **192**, 2757–2768.
- Guzman, L.M., Weiss, D.S., and Beckwith, J. (1997). Domain-swapping analysis of FtsI, FtsL, and FtsQ, bitopic membrane proteins essential for cell division in *Escherichia coli*. *J. Bacteriol.* **179**, 5094–5103.
- Hale, C.A., and de Boer, P.A. (1999). Recruitment of ZipA to the septal ring of *Escherichia coli* is dependent on FtsZ and independent of FtsA. *J. Bacteriol.* **181**, 167–176.
- Hale, C.A., and de Boer, P.A.J. (2002). ZipA Is Required for Recruitment of FtsK, FtsQ, FtsL, and FtsN to the Septal Ring in *Escherichia coli*. *J. Bacteriol.* **184**, 2552–2556.

- Haney, S.A., Glasfeld, E., Hale, C., Keeney, D., He, Z., and de Boer, P. (2001). Genetic Analysis of the *Escherichia coli* FtsZ·ZipA Interaction in the Yeast Two-hybrid System: CHARACTERIZATION OF FtsZ RESIDUES ESSENTIAL FOR THE INTERACTIONS WITH ZipA AND WITH FtsA. *J. Biol. Chem.* **276**, 11980–11987.
- Harms, M.J., Schlessman, J.L., Sue, G.R., and Garcia-Moreno E., B. (2011). Arginine residues at internal positions in a protein are always charged. *Proc. Natl. Acad. Sci.* **108**, 18954–18959.
- Hicks, M.R., Walshaw, J., and Woolfson, D.N. (2002). Investigating the Tolerance of Coiled-Coil Peptides to Nonheptad Sequence Inserts. *J. Struct. Biol.* **137**, 73–81.
- Hopf, T.A., Schärfe, C.P.I., Rodrigues, J.P.G.L.M., Green, A.G., Kohlbacher, O., Sander, C., Bonvin, A.M.J.J., and Marks, D.S. (2014). Sequence co-evolution gives 3D contacts and structures of protein complexes. *ELife* **3**.
- Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38.
- Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865.
- Kamisetty, H., Ovchinnikov, S., and Baker, D. (2013). Assessing the utility of coevolution-based residue-residue contact predictions in a sequence- and structure-rich era. *Proc. Natl. Acad. Sci.* **110**, 15674–15679.
- Karimova, G., Dautin, N., and Ladant, D. (2005). Interaction Network among *Escherichia coli* Membrane Proteins Involved in Cell Division as Revealed by Bacterial Two-Hybrid Analysis. *J. Bacteriol.* **187**, 2233–2243.
- Khadria, A.S., and Senes, A. (2013). The transmembrane domains of the bacterial cell division proteins FtsB and FtsL form a stable high-order oligomer. *Biochemistry* **52**, 7542–7550.
- Kim, J.-K., Cho, Y., Lee, M., Laskowski, R.A., Ryu, S.E., Sugihara, K., and Kim, D.-S. (2015). BetaCavityWeb: a webserver for molecular voids and channels. *Nucleic Acids Res.* **43**, W413–W418.
- Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A.D., and Pastor, R.W. (2010). Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **114**, 7830–7843.
- Krivov, G.G., Shapovalov, M.V., and Dunbrack, R.L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins Struct. Funct. Bioinforma.* **77**, 778–795.

- Kulp, D.W., Subramaniam, S., Donald, J.E., Hannigan, B.T., Mueller, B.K., Grigoryan, G., and Senes, A. (2012). Structural informatics, modeling, and design with an open-source Molecular Software Library (MSL). *J. Comput. Chem.* 33, 1645–1661.
- LaPointe, L.M., Taylor, K.C., Subramaniam, S., Khadria, A., Rayment, I., and Senes, A. (2013). Structural organization of FtsB, a transmembrane protein of the bacterial divisome. *Biochemistry* 52, 2574–2585.
- Laskowski, RomanA., Rullmann, J.AntoonC., MacArthur, MalcolmW., Kaptein, R., and Thornton, JanetM. (1996). AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8.
- Levy, E.D., Pereira-Leal, J.B., Chothia, C., and Teichmann, S.A. (2006). 3D Complex: A Structural Classification of Protein Complexes. *PLoS Comput. Biol.* 2, e155.
- Li, W., Cowley, A., Uludag, M., Gur, T., McWilliam, H., Squizzato, S., Park, Y.M., Buso, N., and Lopez, R. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 43, W580–W584.
- Liu, B., Persons, L., Lee, L., and de Boer, P.A.J. (2015). Roles for both FtsA and the FtsBLQ subcomplex in FtsN-stimulated cell constriction in *E. coli*: Cell fission regulation by FtsN, FtsA and FtsBLQ. *Mol. Microbiol.* 95, 945–970.
- Liu, Z., Mukherjee, A., and Lutkenhaus, J. (1999). Recruitment of ZipA to the division site by interaction with FtsZ. *Mol. Microbiol.* 31, 1853–1861.
- Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. *Science* 252, 1162–1164.
- Lutkenhaus, J., Pichoff, S., and Du, S. (2012). Bacterial cytokinesis: From Z ring to divisome. *Cytoskeleton* 69, 778–790.
- MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins [†]. *J. Phys. Chem. B* 102, 3586–3616.
- Mann, H.B., and Whitney, D.R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *Ann. Math. Stat.* 18, 50–60.
- Männik, J., Bailey, M.W., O'Neill, J.C., and Männik, J. (2017). Kinetics of large-scale chromosomal movement during asymmetric cell division in *Escherichia coli*. *PLOS Genet.* 13, e1006638.

- Marks, D.S., Colwell, L.J., Sheridan, R., Hopf, T.A., Pagnani, A., Zecchina, R., and Sander, C. (2011). Protein 3D Structure Computed from Evolutionary Sequence Variation. *PLoS ONE* 6, e28766.
- Massey, T.H., Mercogliano, C.P., Yates, J., Sherratt, D.J., and Löwe, J. (2006). Double-Stranded DNA Translocation: Structure and Mechanism of Hexameric FtsK. *Mol. Cell* 23, 457–469.
- Masson, S., Kern, T., Le Gouët, A., Giustini, C., Simorre, J.-P., Callow, P., Vernet, T., Gabel, F., and Zapun, A. (2009). Central Domain of DivIB Caps the C-terminal Regions of the FtsL/DivIC Coiled-coil Rod. *J. Biol. Chem.* 284, 27687–27700.
- McDonnell, A.V., Jiang, T., Keating, A.E., and Berger, B. (2006). Paircoil2: improved prediction of coiled coils from sequence. *Bioinformatics* 22, 356–358.
- Meeske, A.J., Riley, E.P., Robins, W.P., Uehara, T., Mekalanos, J.J., Kahne, D., Walker, S., Kruse, A.C., Bernhardt, T.G., and Rudner, D.Z. (2016). SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature* 537, 634–638.
- Mercer, K.L.N., and Weiss, D.S. (2002). The *Escherichia coli* Cell Division Protein FtsW Is Required To Recruit Its Cognate Transpeptidase, FtsI (PBP3), to the Division Site. *J. Bacteriol.* 184, 904–912.
- Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Distèche, M., de Kruijff, B., and Breukink, E. (2011). Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane: FtsW transports Lipid II across the membrane. *EMBO J.* 30, 1425–1432.
- Mohammadi, T., Sijbrandi, R., Lutters, M., Verheul, J., Martin, N.I., den Blaauwen, T., de Kruijff, B., and Breukink, E. (2014). Specificity of the Transport of Lipid II by FtsW in *Escherichia coli*. *J. Biol. Chem.* 289, 14707–14718.
- Mosyak, L. (2000). The bacterial cell-division protein ZipA and its interaction with an FtsZ fragment revealed by X-ray crystallography. *EMBO J.* 19, 3179–3191.
- Müller, P., Ewers, C., Bertsche, U., Anstett, M., Kallis, T., Breukink, E., Fraipont, C., Terrak, M., Nguyen-Distèche, M., and Vollmer, W. (2007). The Essential Cell Division Protein FtsN Interacts with the Murein (Peptidoglycan) Synthase PBP1B in *Escherichia coli*. *J. Biol. Chem.* 282, 36394–36402.
- Noirclerc-Savoye, M., Le Gouët, A., Morlot, C., Dideberg, O., Vernet, T., and Zapun, A. (2004). In vitro reconstitution of a trimeric complex of DivIB, DivIC and FtsL, and their transient co-localization at the division site in *Streptococcus pneumoniae*: Trimeric complex of DivIB, DivIC and FtsL. *Mol. Microbiol.* 55, 413–424.

- North, B., Summa, C.M., Ghirlanda, G., and DeGrado, W.F. (2001). D(n)-symmetrical tertiary templates for the design of tubular proteins. *J. Mol. Biol.* **311**, 1081–1090.
- O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* **44**, D733–D745.
- Ortiz, C., Natale, P., Cueto, L., and Vicente, M. (2016). The keepers of the ring: regulators of FtsZ assembly. *FEMS Microbiol. Rev.* **40**, 57–67.
- Ovchinnikov, S., Kinch, L., Park, H., Liao, Y., Pei, J., Kim, D.E., Kamisetty, H., Grishin, N.V., and Baker, D. (2015). Large-scale determination of previously unsolved protein structures using evolutionary information. *ELife* **4**, e09248.
- Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., and Jacobs-Wagner, C. (2016). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis: Oufti: image analysis software. *Mol. Microbiol.* **99**, 767–777.
- Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802.
- Robichon, C., King, G.F., Goehring, N.W., and Beckwith, J. (2008). Artificial Septal Targeting of *Bacillus subtilis* Cell Division Proteins in *Escherichia coli*: an Interspecies Approach to the Study of Protein-Protein Interactions in Multiprotein Complexes. *J. Bacteriol.* **190**, 6048–6059.
- Schmidt, K.L., Peterson, N.D., Kustusch, R.J., Wissel, M.C., Graham, B., Phillips, G.J., and Weiss, D.S. (2004). A Predicted ABC Transporter, FtsEX, Is Needed for Cell Division in *Escherichia coli*. *J. Bacteriol.* **186**, 785–793.
- Schmidt, N.W., Grigoryan, G., and DeGrado, W.F. (2017). The accommodation index measures the perturbation associated with insertions and deletions in coiled-coils: Application to understand signaling in histidine kinases: Accommodation Index for Insertions in Coiled-Coils. *Protein Sci.* **26**, 414–435.
- Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005). The FoldX web server: an online force field. *Nucleic Acids Res.* **33**, W382–W388.
- Spearman, C. (1904). The Proof and Measurement of Association between Two Things. *Am. J. Psychol.* **15**, 72.

- Strelkov, S.V., Schumacher, J., Burkhard, P., Aebi, U., and Herrmann, H. (2004). Crystal Structure of the Human Lamin A Coil 2B Dimer: Implications for the Head-to-tail Association of Nuclear Lamins. *J. Mol. Biol.* 343, 1067–1080.
- Studier, F.W. (2005). Protein production by auto-induction in high-density shaking cultures. *Protein Expr. Purif.* 41, 207–234.
- Subramaniam, S., and Senes, A. (2012). An energy-based conformer library for side chain optimization: Improved prediction and adjustable sampling. *Proteins Struct. Funct. Bioinforma.* 80, 2218–2234.
- Subramaniam, S., and Senes, A. (2014). Backbone dependency further improves side chain prediction efficiency in the Energy-based Conformer Library (bEBL): A Backbone-Dependent Energy-Based Library. *Proteins Struct. Funct. Bioinforma.* 82, 3177–3187.
- Szwedziak, P., Wang, Q., Freund, S.M., and Löwe, J. (2012). FtsA forms actin-like protofilaments: FtsA forms actin-like protofilaments. *EMBO J.* 31, 2249–2260.
- Tsang, M.-J., and Bernhardt, T.G. (2015). A role for the FtsQLB complex in cytokinetic ring activation revealed by an *fts L* allele that accelerates division: FtsL and divisome function. *Mol. Microbiol.* 95, 925–944.
- Villanelo, F., Ordenes, A., Brunet, J., Lagos, R., and Monasterio, O. (2011). A model for the *Escherichia coli* FtsB/FtsL/FtsQ cell division complex. *BMC Struct. Biol.* 11, 28.
- Vollmer, W., and Bertsche, U. (2008). Murein (peptidoglycan) structure, architecture and biosynthesis in *Escherichia coli*. *Biochim. Biophys. Acta BBA - Biomembr.* 1778, 1714–1734.
- Wang, L., Khattar, M.K., Donachie, W.D., and Lutkenhaus, J. (1998). FtsI and FtsW are localized to the septum in *Escherichia coli*. *J. Bacteriol.* 180, 2810–2816.
- Wolf, E., Kim, P.S., and Berger, B. (1997). MultiCoil: A program for predicting two-and three-stranded coiled coils: MultiCoil. *Protein Sci.* 6, 1179–1189.
- Yang, D.C., Peters, N.T., Parzych, K.R., Uehara, T., Markovski, M., and Bernhardt, T.G. (2011). An ATP-binding cassette transporter-like complex governs cell-wall hydrolysis at the bacterial cytokinetic ring. *Proc. Natl. Acad. Sci.* 108, E1052–E1060.
- Yang, J.-C., Van Den Ent, F., Neuhaus, D., Brevier, J., and Löwe, J. (2004). Solution structure and domain architecture of the divisome protein FtsN: Solution structure of FtsN. *Mol. Microbiol.* 52, 651–660.
- Yang, X., Lyu, Z., Miguel, A., McQuillen, R., Huang, K.C., and Xiao, J. (2017). GTPase activity–coupled treadmilling of the bacterial tubulin FtsZ organizes septal cell wall synthesis. *Science* 355, 744–747.

Chapter 4: The coiled-coil domain of *E. coli* FtsLB is a structurally detuned element critical for modulating its activation in bacterial cell division

This chapter was prepared for publication as:

Craven S.J^{1,2}, Condon S.G.F.^{1,2}, Díaz Vázquez G.^{1,3}, Cui Q.⁴, Senes A.¹. 2022. "The coiled-coil domain of *Escherichia coli* FtsLB is a structurally detuned element critical for modulating its activation in bacterial cell division." J. Biol. Chem. 298(1), 101460.

¹Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706

²Integrated Program in Biochemistry, University of Wisconsin-Madison, Madison, WI 53706

³Biophysics Graduate Program, University of Wisconsin-Madison, Madison, WI 53706

⁴Department of Chemistry, Boston University

Statement of contribution

I performed the molecular dynamics simulations, and worked on their analysis.

4.1. Abstract

The FtsLB complex is a key regulator of bacterial cell division, existing in either an *off* state or an *on* state, which supports the activation of septal peptidoglycan synthesis. In *Escherichia coli*, residues known to be critical for this activation are located in a region near the C-terminal end of the periplasmic coiled-coil domain of FtsLB, raising questions about the precise role of this conserved domain in the activation mechanism. Here, we investigate an unusual cluster of polar amino acids found within the core of the FtsLB coiled coil. We hypothesized that these amino acids likely reduce the structural stability of the domain and thus may be important for governing conformational changes. We found that mutating these positions to hydrophobic residues increased the thermal stability of FtsLB but caused cell division defects, suggesting that the coiled-coil domain is a “detuned” structural element. In addition, we identified suppressor mutations within the polar cluster, indicating that the precise identity of the polar amino acids is important for fine-tuning the structural balance between the *off* and *on* states. We propose a revised structural model of the tetrameric FtsLB (named the “Y-model”) in which the periplasmic domain splits into a pair of coiled-coil branches. In this configuration, the hydrophilic terminal moieties of the polar amino acids remain more favorably exposed to water than in the original four-helix bundle model (“I-model”). We propose that a shift in this architecture, dependent on its marginal stability, is involved in activating the FtsLB complex and triggering septal cell wall reconstruction.

4.2. Introduction

Cell division in bacteria is a complex process involving intricate coordination between numerous cellular components. Central to this coordination is the divisome, a multiprotein complex that in the Gram-negative bacterium *Escherichia coli* consists of a number of essential proteins (FtsZ, FtsA, ZipA, FtsE, FtsX, FtsK, FtsQ, FtsL, FtsB, FtsW, FtsI, and FtsN; Figure 4.1) as well as a suite of nonessential, conditionally essential, and redundant proteins (den Blaauwen et al., 2017; Du and Lutkenhaus, 2017). These proteins mediate the various functions necessary for cell division, including establishing the site of division, coordinating invagination of the inner and outer membranes, and remodeling the cell wall at midcell into a septum to compartmentalize the nascent daughter cells. If any of these functions is abrogated through deletion or mutation of essential proteins, the bacteria can continue to elongate and replicate their DNA, but they will be unable to divide and form distinct daughter cells. This will result in the formation of long filaments and eventual cell lysis and death.

Remodeling of the cell wall during division involves degradation of old peptidoglycan (PG) at the division site and synthesis of new material, leading to the formation of a septum that eventually splits into the poles of the nascent daughter cells. Numerous nonessential and redundant enzymes (e.g., periplasmic hydrolases) participate in PG reconstruction (Booth and Lewis, 2019; Egan and Vollmer, 2013; Egan et al., 2020), but the major synthetic activity is performed by the essential FtsWI complex (Cho et al., 2016; Fraipont et al., 2011; Mercer and Weiss, 2002). FtsW – a large multipass membrane protein – is a PG glycosyltransferase that polymerizes novel glycan strands from lipid II precursors (Meeske et al., 2016; Taguchi et al., 2019) (FtsW was also

previously reported to have lipid II flippase activity (Leclercq et al., 2017; Mohammadi et al., 2011, 2014)). FtsI is the cognate transpeptidase of FtsW and is responsible for crosslinking the glycan polymers to form a network of PG strands (Botta and Park, 1981; Mercer and Weiss, 2002). Forming the septum requires other PG synthases (e.g., either of the bifunctional glycosyltransferase/transpeptidases PBP1a or PBP1b (Nakagawà et al., 1979)), but functional redundancy between these proteins means that no individual component is essential apart from FtsWI.

The mere presence of FtsWI at midcell is not, however, sufficient for completion of cytokinesis. Instead, the complex must be switched on, and this activation (along with the consequent triggering of cell wall reconstruction) is a tightly regulated step of cell division (Du and Lutkenhaus, 2017; McCausland et al., 2021; Yang et al., 2021). In current models, activation begins with the midcell localization of FtsN (Gerding et al., 2009; Ursinus et al., 2004; Yang et al., 2004) (Figure 4.1c), which communicates with FtsWI through a cytoplasmic route involving FtsA (Liu et al., 2015; Park et al., 2021; Pichoff et al., 2015) and through a periplasmic route involving the FtsLB complex (Li et al., 2021; Liu et al., 2015; Marmont and Bernhardt, 2020; Park et al., 2020; Tsang and Bernhardt, 2015). In this work, we focus on the latter periplasmic activation route.

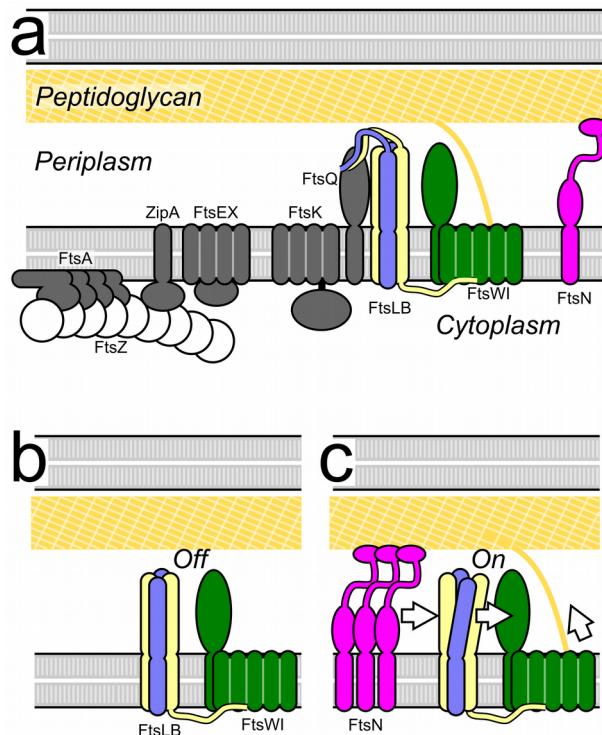


Figure 4.1: The current model for activation of cell division. *a*, Schematic representation of the essential components of the *E. coli* divisome. The polymeric ring formed by FtsZ establishes the site of division in coordination with other early components. The peptidoglycan synthase complex FtsWI is the main actor for the reconstruction of the cell wall, leading to the formation of a septum and, eventually, the poles of the nascent daughter cells. Its activation is tightly regulated by FtsN and FtsLB. *b*, In the current model, FtsWI is in complex with FtsLB but is initially inactive. *c*, Accumulation of FtsN at midcell somehow triggers a conformational change in the FtsLB complex, which in turn triggers peptidoglycan synthesis likely by a direct interaction with FtsWI.

FtsL and FtsB are both single-pass membrane proteins with a short (FtsL) or minimal (FtsB) N-terminal tail in the cytoplasm and a larger C-terminal domain in the periplasm. They form a heterotetrameric complex consisting of two FtsL and two FtsB subunits arranged into a long helical bundle formed by its transmembrane and periplasmic coiled-coil domains (Condon et al., 2018; Khadria and Senes, 2013) (Figure 4.2a). The N-terminal, cytoplasmic tail of FtsL is involved with recruiting FtsWI to the division site (Gonzalez et al., 2010), whereas the C-terminal, periplasmic tail of FtsB binds with high affinity to another divisome protein FtsQ (Choi et al., 2018; Glas et al., 2015; Kureisaite-Ciziene et al., 2018) and is needed for FtsLB's own recruitment to midcell (Ghigo and Beckwith, 2000; Ghigo et al., 1999; Goehring et al., 2006; Gonzalez et al., 2010). The structure of FtsLB has not been solved experimentally aside from fragments (Choi et al., 2018; Kureisaite-Ciziene et al., 2018; LaPointe et al., 2013), but computational structural models of its helical bundle region are available. Originally, the Monasterio group proposed models of the soluble coiled-coil region in complex with the periplasmic domain of FtsQ, in either trimeric or hexameric configurations (Villanelo et al., 2011). More recently, using a set of amino acid contacts postulated by evolutionary coupling, we derived a model that includes both the transmembrane region and the coiled coil (Condon et al., 2018). In this model, the transmembrane and coiled-coil domains of FtsL form a continuous helix, whereas FtsB contains a potentially flexible Gly-rich linker that breaks the helix between these two regions (LaPointe et al., 2013).

In current models, FtsLB regulates FtsWI septal PG synthesis activity by transitioning from an *off* state to an *on* state in response to a signal from FtsN (Figure

4.1c). This idea was initially proposed following the identification of a series of gain-of-function mutations within both FtsL and FtsB that enable survival in the absence of the normally essential FtsN (Liu et al., 2015; Tsang and Bernhardt, 2015). Along with subsequent work (Marmont and Bernhardt, 2020; Park et al., 2020), this led to the identification of two related regions at the C-terminal end of the FtsLB coiled coil that are central to its regulation of FtsWI. The first region, named the Constriction Control Domain (CCD, approximately residues 88-94 in FtsL and 55-59 in FtsB), houses the aforementioned $\Delta ftsN$ -suppressing mutations (Liu et al., 2015; Tsang and Bernhardt, 2015). The second region neighbors the CCD on the opposite helical face of FtsL specifically and is designated as Activation of FtsWI (AWI; positions 82-84, 86-87, and 90). Dominant-negative mutations in FtsL indicate that the AWI region directly interacts with and activates FtsWI (Marmont and Bernhardt, 2020; Park et al., 2020), suggesting that the *off/on* transition in FtsLB may involve conformational changes that make the AWI region available to interact with and activate FtsWI (Park et al., 2020). Normally, either direct or indirect interactions with FtsN are required to trigger such a change; however, the gain-of-function CCD mutations may induce similar structural rearrangements of FtsLB, thereby mimicking the signal from FtsN and bypassing its requirement to trigger septal PG reconstruction. Whatever conformational changes in FtsLB are required for this activation, they will likely depend on the extended helical topology that is at the core of the complex (Buddelmeijer and Beckwith, 2004; Condon et al., 2018; Ghigo and Beckwith, 2000; Masson et al., 2009).

One intriguing structural feature of FtsLB is the unusual presence of a cluster of strongly polar amino acids buried at interfacial positions of the coiled coil (i.e., at positions designated as “*a*” and “*d*” in the “*abcdefg*” heptad repeat). This “polar cluster” (as it will be referred to from now on) consists of two arginine residues (R67 and R74) in FtsL and of a glutamine residue (Q39) and two asparagine residues (N43 and N50) in FtsB (Figure 4.2*b*). Since canonical coiled coils contain primarily hydrophobic residues at the interfacial “*a*” and “*d*” positions (Woolfson, 2005), the polar cluster is likely to decrease the stability of the coiled coil (Condon et al., 2018). This suggests that these nonideal residues play some critical role in modulating the stability and dynamics of the FtsLB complex, which likely has important functional consequences – a hypothesis that we address in this present article.

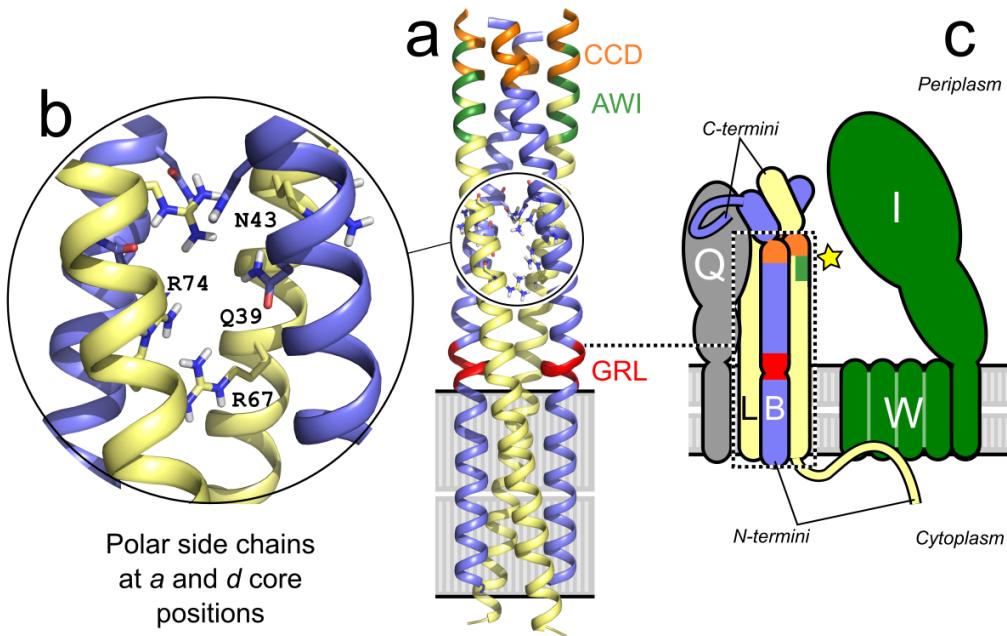


Figure 4.2: An unusual cluster of strongly polar amino acids in the core of the computational model (I-model) of FtsLB likely tunes its stability and conformation dynamics. *a*, Original computational model of the tetrameric helical bundle of FtsLB (Condon et al., 2018), referred to here as the I-model. The model is formed by two FtsL (yellow) and two FtsB (blue) units. FtsL forms a continuous helix across the membrane and periplasmic domains, whereas the helix of FtsB is interrupted in the juxta-membrane region by an unwound Gly-rich linker (GRL, red). Highlighted in orange and green are the CCD and AWI regions, respectively, which occur near a predicted hinge at the C-terminal end of the coiled coil. These regions are critical for FtsLB activation. *b*, The coiled coil contains a cluster of polar amino acids at core “*a*” and “*d*” positions, including two Arg residues in FtsL (positions 67 and 74) and Gln-39 and Asn-43 in FtsB. The

polar cluster is likely to be a destabilizing feature of the coiled coil. *c*, Schematic representation of FtsLB and its interactions with FtsQ and the FtsWI complex. The region corresponding to the model is enclosed in the dotted box. The interaction with FtsQ is largely mediated by the C-terminal tail of FtsB. The interaction with FtsW is mediated by the N-terminal tail of FtsL. The star indicates the putative activating contact between the AWI region of FtsL and FtsI.

We hypothesize that the coiled coil of FtsLB is detuned for stability in order to support the dynamics necessary for structural transitions that occur during the *off* to *on* switch at the heart of FtsWI regulation. Here, we show that hydrophobic mutations introduced in the polar cluster of FtsLB stabilize the coil but lead to cell division defects *in vivo*. We also show that the identity of the interfacial residues is important for the observed division phenotypes, indicating those residues may play a more nuanced role than simply destabilizing the hydrophobic coiled-coil interface and are likely to participate in the balance of forces that regulate the *off/on* transition of the complex. Additionally, because the presence of polar residues at “*a*” and “*d*” positions statistically favors the formation of two-stranded coiled coils (Gonzalez et al., 1996; Harbury et al., 1993; Wagschal et al., 1999), we investigated an alternative model of FtsLB in which the coiled-coil region splits into a pair of two-stranded coiled-coil domains (the Y-model), as opposed to the monolithic, four-stranded coil that we originally proposed (Condon et al., 2018) (the I-model). This revised model fits the available evidence as well as the original I-model, while displaying better behavior, and thus, we propose it as the more likely candidate for the structural organization of FtsLB.

4.3. Results and Discussion

4.3.1. The polar cluster of FtsLB is an unusual feature for coiled-coil structures

Although “a” and “d” positions of coiled coils tend to be mainly hydrophobic, polar amino acids can also occur there. To determine if the coiled coil of FtsLB is unusually rich in interfacial polar amino acids in comparison with other coiled coils, we performed a structural analysis of the 2,662 crystal structures available in the CC+ database (Testa et al., 2009). We found that polar amino acids such as Asp, Glu, His, Arg, Lys, Gln, and Asn, occur with a frequency of 17.6% at “a” or “d” positions, corresponding on average to approximately one polar amino acid every three heptad repeats (supplementary Table S4.8.2). Notably, the propensity to accommodate polar amino acids decreases as the number of helices forming the coiled-coil assembly increases. The frequency of polar amino acids is highest at 18.8% in coiled coils formed by two helices, and it decreases to 14.1% and 11.7% for three-stranded and four-stranded coiled coils, respectively. In contrast, 30% of the “a” and “d” positions in *E. coli* FtsLB are polar, confirming that the level of enrichment is quite high in comparison with the average coiled coil and nearly three times the average frequency for coiled coils that assume a tetrameric configuration, suggesting that the coiled coil of FtsLB is not designed for maximal stability.

4.3.2. The polar cluster is evolutionarily conserved

To investigate if the polar cluster is a conserved feature of FtsLB, we analyzed a multisequence alignment containing 2900 pairs of FtsB-FtsL sequences from diverse proteobacterial taxa. In *E. coli*, the coiled-coil region is predicted to extend for

approximately five heptad repeats, starting after the transmembrane helices and ending approximately at the CCD region (Condon et al., 2018). Three polar amino acids are found in FtsB at positions 2d (Gln-39, the “d” position in the second heptad repeat), 3a (Asn-43), and 4a (Asn-50). FtsL contains two arginine residues at positions 2a and 3a (Arg-67 and Arg-74, respectively). Additionally, FtsL contributes Glu-88 (5a), one of the critical amino acids in the CCD region (Liu et al., 2015; Tsang and Bernhardt, 2015) at the margin of the coiled coil.

The pattern of polar/nonpolar amino acids found at “a” and “d” positions in the alignment is summarized graphically in Figure 4.3. We found that this feature is highly conserved, even if the specific sequence is not. As highlighted by the main “paths” in the graph, five of the six polar positions in *E. coli* have a strong tendency to be polar in all species. The only position at which a polar amino acid is not particularly conserved corresponds to Arg-67 (2a) of FtsL. However, a polar amino acid occurs overall frequently across any of the “a” and “d” positions of the first two heptad repeats (71% of the sequences). In fact, the class of sequences to which *E. coli* belongs (i.e., those sequence in which both FtsL and FtsB contribute three polar amino acids each) is the most common (39%). Moreover, 84% of the sequences contain at least five polar amino acids (supplementary Figure S4.1). The analysis therefore suggests that, in proteobacteria, the coiled coil of FtsLB is far from being tuned for maximum stability and is thus designed to be dynamic or metastable for functional reasons.

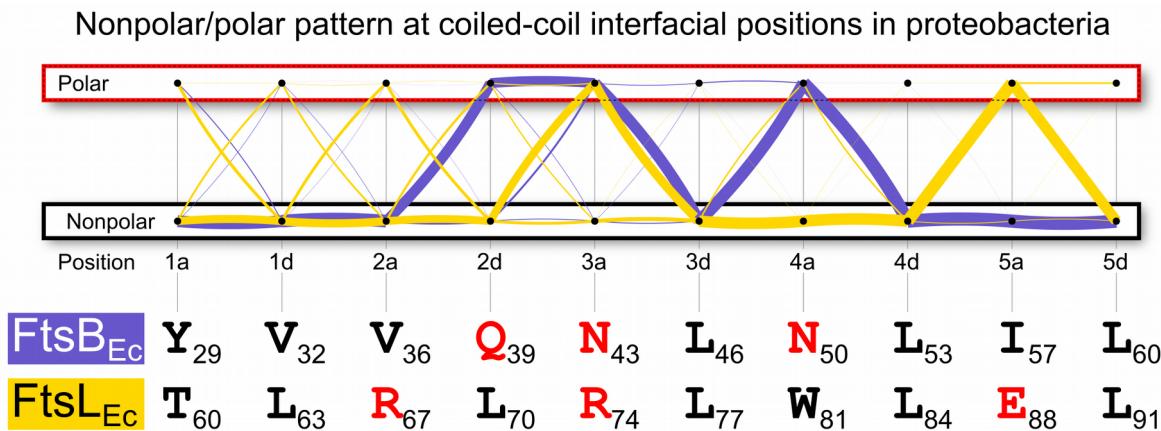


Figure 4.3: **Conservation of the polar cluster in the core “a” and “d” positions of the coiled coil of FtsLB in proteobacterial species.** The yellow (FtsL) and blue (FtsB) lines follow the pattern of polar (Asp, Glu, His, Arg, Lys, Gln, and Asn) and nonpolar amino acids in alignments of 2900 paired FtsB/FtsL sequences. The thickness of the lines is proportional to the number of sequences that follows a certain (polar/nonpolar) binary pattern. The graph evidences that polarity at positions corresponding to Gln-39, Asn-43, and Asn-50 is a highly conserved feature in FtsB. Positions Arg-74 and Glu-88 are also most frequently polar in FtsL. Polarity at the position corresponding to Arg-67 is not conserved in FtsL. However, there is overall a 71% probability that at least one position in the first two heptad repeats of FtsL contains another polar residue.

4.3.3. The polar cluster tunes the propensity of FtsLB to transition to an activated state

In order to test whether the conserved polar cluster is critical for function, we expressed mutant variants at these positions *in vivo* and examined whether these changes led to cell division defects. In our first round of experiments, we tested the effect of “idealizing” the coiled coil by converting each of the polar amino acids to a canonical hydrophobic residue (Ile or Leu, at “*a*” and “*d*” positions, respectively). We also mutated Trp-81 of FtsL to a small hydrophobic residue, since bulky aromatic residues tend to be excluded from natural coiled-coil interfaces (Woolfson, 2005). The effect of individual point mutations was assessed in complementation experiments in which we measured the length distribution of samples of at least 500 cells. As done previously to assess the severity of mutations that cause elongation defects (Condon et al., 2018), we measured the proportion of cells with lengths exceeding the 95th percentile of the length distribution observed for the WT (dashed vertical line in the four example length distributions displayed in Figure 4.4a-d. The fraction of long cells is reported for further mutants in the histograms in Figure 4.4e-h, grouped for FtsB nonideal-to-ideal, FtsL nonideal-to-ideal, FtsL charged-to-charged, and double substitutions at positions R67 and R74 in FtsL. Pictures, cell distributions, and metrics for all mutations are reported in supplementary Figure S4.2.

Three of the five polar-to-nonpolar individual point mutations tested (FtsB Q39L and N50I; FtsL R67I) resulted in notable defective phenotypes, displaying a fraction between 12-37% of elongated cells (Figure 4.4e-f). The remaining two polar-to-nonpolar mutations (FtsL R74I and FtsB N43I) were similar to WT. Mutation of the bulky Trp residue in FtsL (W81I) also resulted in a defective phenotype. To exclude that the

division defects were due to changes in protein expression levels, we performed western blot analyses (supplementary Figure S4.3), which indicated that each mutant was expressed to similar levels as WT, except for FtsB N43I, which showed increased expression. These observations support the hypothesis that the polar cluster plays a critical role in FtsLB function.

Arginine is one of the most polar amino acids, essentially never occurring in its neutral form, and thus it is one of the most destabilizing amino acids when buried within a protein core. For this reason, we further investigated the effects of mutating both Arg-67 and Arg-74 in FtsL (Figure 4.4h). The double charged-to-hydrophobic mutation (R67I+R74I) resulted in impaired division in a manner similar to the R67I single mutant (30% vs 33% elongated cells). We then asked whether the identity of these positions is important (Figure 4.4g). We first inverted the charge with Glu substitutions. The R67E mutation produced a division-defective phenotype (33% elongated cells) similar to that of R67I. Like the Ile substitution at the same position, R74E did not result in elongation, but somewhat surprisingly, this substitution resulted in a larger fraction of smaller cells (left-shifted peak compared to WT, Figure 4.4d). We then preserved the positive charges at the 67 and 74 positions of FtsL with Lys substitutions. While R67K produced WT-like cells, R74K had a similar phenotype to R74E, with a larger fraction of small cells (supplementary Figure S4.2c).

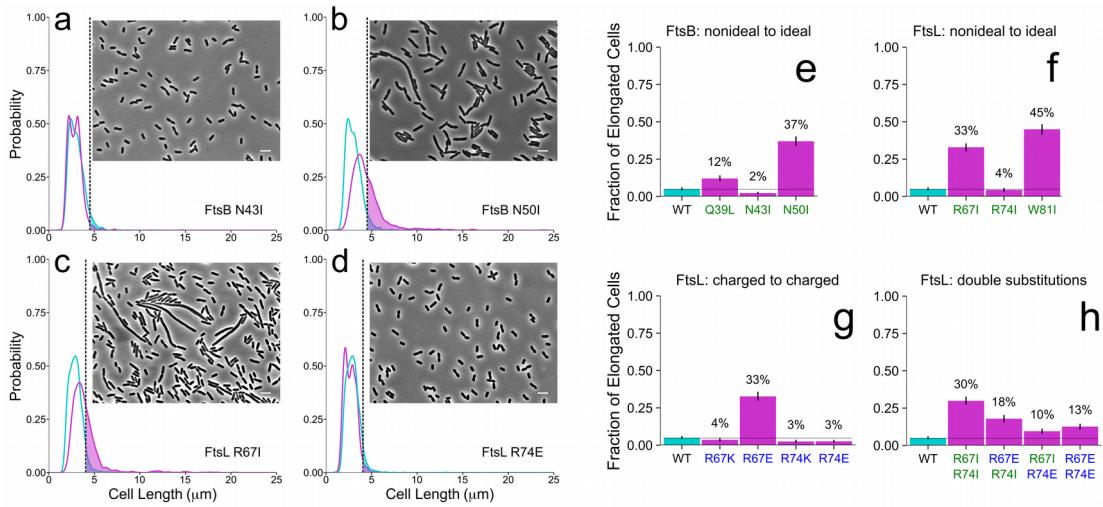


Figure 4.4: Mutations in the polar cluster cause both elongation and small cell phenotypes in vivo. Panels *a-d*: Phase-contrast images of representative mutants (5 μm scale bar) and cell length distributions of mutant cells (magenta) compared to wild type (cyan). The shaded areas to the right of the dotted lines represent the elongated cells (i.e., those that are longer than the 95th percentile in the WT distribution). The four examples include (*a*) a wild-type-like mutant (FtsB N43I), (*b-c*) two mutations that cause frequent elongation (FtsB N50I at 37% elongated and FtsL R67I at 33% elongated), and (*d*) a mutant that displays a higher fraction of shorter cells (FtsL R74E). Panels *e-h*: histograms of percent of elongated cells for all mutations at the nonideal coiled-coil positions (nonpolar mutations in green; charged mutations in blue). These include (*e*) FtsB and (*f*) FtsL nonideal-to-ideal mutations, (*g*) FtsL charged-to-charged mutations, and (*h*) FtsL double mutations at positions R67 and R74. Nonideal-to-ideal mutations tend to cause notable elongation phenotypes. A remarkable exception is position R74,

in which mutations appear to cause a small cell phenotype (likely due to deregulated division). R74E, in particular, suppresses the elongated phenotypes of R67I and R67E, indicating that these positions are important for governing the fine balance between the *on* and *off* conformations of FtsLB. Experiments were performed at 37 °C. Error bars represent the 95% confidence interval of the fraction of elongated cells estimated from 1000 replicates of bootstrap resampling.

The fact that charge reversal mutations at position R67 and R74 result in opposite phenotypes (elongation vs small cells, respectively) is interesting. Gain-of-function mutations in FtsLB that cause early cell division have been identified before (Liu et al., 2015; Tsang and Bernhardt, 2015), predominantly within the CCD. The reduced cell length of R74 mutants is particularly notable since a similar phenotype was never observed among a total of 55 mutations in the transmembrane region and coiled-coil positions of FtsLB that were assessed in our previous analysis with identical conditions and methodology (Condon et al., 2018). For this reason, we combined both charge reversal mutations (R67E+R74E) to observe their interplay. Interestingly, R67E+R74E resulted in a shorter cell length distribution (13% elongated cells compared to 33% for R67E alone, Figure 4.4d), supporting the hypothesis that R74E can suppress the R67E elongation defect by inducing early cell division. This is also consistent with the phenotypes we observed in the R67I+R74E and R74E+W81I double mutations (10% and 15% elongated cells, respectively), where the elongated phenotypes of R67I and W81I alone (33% and 45% elongated cells, respectively) are also suppressed, resulting in a decreased fraction of elongated cells.

In order to determine if FtsL R74E is unique or if other mutations at that position can suppress elongation defects, we tested further combinations. Ultimately, we found that the extent of observed suppression varies with the identity of the amino acid at position 74. The charged-to-nonpolar mutation R74I (which was WT-like alone and did not suppress the elongated phenotype of R67I) also appears to reduce the severity of the elongated phenotype when combined with R67E (33% vs 18% elongated cells for R67E and R67E+R74I, respectively). On the other hand, R74K did not noticeably suppress the elongation defect when combined with R67I, R67E, or W81I (supplementary Figure S4.2d), suggesting that a negative charge at position 74 (or perhaps merely the lack of a positive charge) is needed for the suppression effect.

Overall, the data suggest that the polar cluster of the FtsLB coiled coil is important for fine-tuning the propensity of the complex to transition to an activated state. In particular, the Arg residues at the 67 and 74 “a” positions in FtsL somehow play opposing roles in this regulation, and their specific residue identity is important.

4.3.4. Hydrophobic substitutions *in vitro* increase thermal stability

If the polar cluster plays a role in tuning FtsLB’s propensity to transition to an activated state, then polar-to-nonpolar substitutions likely stabilize the coiled-coil domain but may be detrimental for the complex to undergo conformational changes. To directly assess whether the polar cluster does in fact govern the structural stability of FtsLB, we investigated the thermal stability of the complex *in vitro* by monitoring secondary structure content using circular dichroism (CD). In this experiment, we

simultaneously mutated all four centrally located polar residues in both FtsB (Q39L and N43I) and FtsL (R67I and R74I).

We used a version of FtsL (FtsL_{35-121}) lacking the unstructured N-terminal tail, which is not necessary for the assembly of the FtsLB complex (Condon et al., 2018). Both FtsB and FtsL constructs retained the N-terminal purification tags (His and Strep tags, respectively), which do not cleave efficiently. To avoid interference from reducing agents in the sensitive UV region, both native cysteine residues of FtsL were replaced with alanine (C41A and C45A). We have shown previously that this Cys-less construct is stable and functional (Condon et al., 2018). The resulting constructs, His-FtsB/Strep- FtsL_{35-121} -C41A-C45A and His-FtsB-Q39L-N43I/Strep- FtsL_{35-121} -C41A-C45A-R67I-R74I, are termed “WT” and “4x-mutant”, respectively.

Figure 4.5a shows the CD spectra of the constructs solubilized in n-dodecyl-β-D-maltopyranoside (DDM) at low temperature (4 °C). Both constructs have the expected spectral signature of helical proteins with high secondary structure content. Little difference is noticeable between the two constructs, suggesting that the four mutations do not cause major changes in helical content at low temperature. We then investigated the difference in stability between WT and 4x-mutant comparing the CD signal at increased temperatures, monitoring ellipticity at one of the helical minima (224 nm, Figure 4.5b and supplementary Figure S4.4).

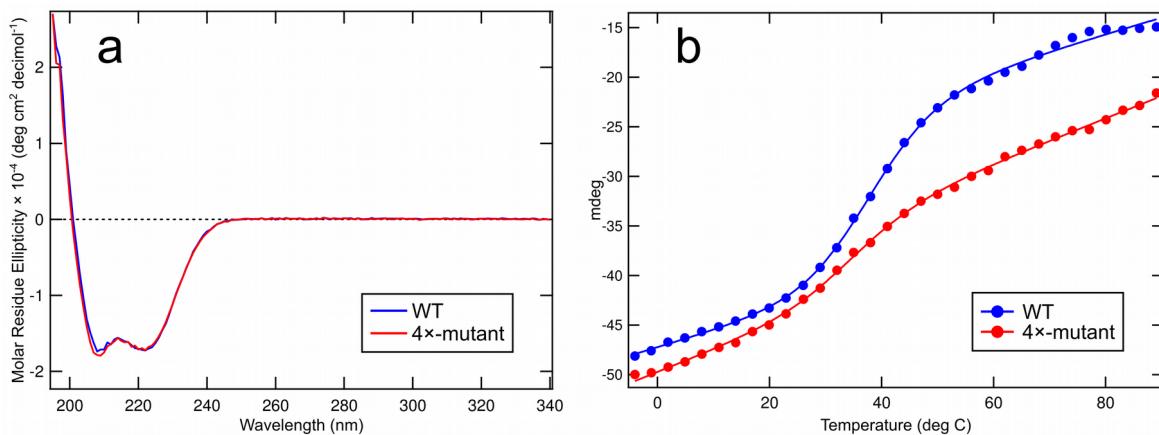


Figure 4.5: Conversion of the polar cluster to idealized hydrophobic residues increases thermal stability but affects folding cooperativity. a, Far-UV CD spectra of WT FtsLB (blue) compared to the 4x-mutant (red) at 4 °C. b, Representative temperature melting curves comparing WT FtsLB (blue) to the 4x-mutant (red). The WT melting curve displays a sigmoidal transition centered around 40 °C. In contrast, the 4x-mutant displays a curve that is shifted to the right but lacks a significant transition, possibly indicating a loss of cooperativity of unfolding. CD scans were monitored at 224 nm from 4 to 89 °C. Replica melting curves are included in supplementary Figure S4.4.

The WT construct (blue dots) displays a sigmoidal melting curve with a transition centered around 40 °C. Given the high thermal stability of transmembrane helices, the transition is most likely attributable to loss of helicity in the coiled-coil region. The relatively early unfolding transition confirms that the coiled coil of FtsLB is structured but not optimized for stability. The melting curve of the 4x-mutant is markedly different (red dots). The construct retains a higher degree of secondary structure at higher

temperatures, indicating that, as expected, the replacement of unfavorable polar side chains for canonical hydrophobic residues stabilized the coiled coil. However, the 4 \times -mutant's curve does not display the same degree of cooperativity as the WT, but rather, it shows a nearly linear loss of ellipticity lacking a clear transition. A sharp transition is consistent with a two-state unfolding process typical of well folded proteins; therefore, it is likely that the 4 \times -mutant version of the coiled coil of FtsLB can access alternative and potentially misfolded conformations. This notion becomes particularly interesting in light of the fact that FtsLB can be modeled in two alternative conformations, as discussed in the next section.

4.3.5. An alternative structural organization for FtsLB: the Y-model

Experimental evidence *in vitro* indicates that the FtsLB complex is a 2:2 FtsL:FtsB tetramer (Condon et al., 2018; Khadria and Senes, 2013; LaPointe et al., 2013). We previously modeled the complex in the simplest configuration consistent with this tetrameric state (i.e., a monolithic, four-helix bundle that extended across the transmembrane and periplasmic regions (Condon et al., 2018), Figure 4.2a). We noted, however, that it is unlikely that a tetrameric coiled coil would be stable with the number of polar residues present in its core. Consistently, the coiled-coil configuration did not appear stable during molecular dynamics (MD) simulations, where we observed a marked tendency of this region to open and recruit water within its interior.

The previously discussed analysis of the structural databases indicates that a polar coiled coil would be most likely to assume a two-stranded configuration. Experimental studies on model coiled coils also indicate that polar residues in core positions can

influence the number of helical strands in coiled coils, with a two-stranded coiled coil being better suited to accommodate polar residues (Gonzalez et al., 1996; Harbury et al., 1993). In particular, the four Arg residues contributed by the FtsL chains (two from each monomer) are particularly costly to bury, since Arg is essentially always protonated even in a hydrophobic environment. Indeed, a systematic study of all twenty amino acids in a model coiled coil found that Arg is the most destabilizing residue at “*a*” positions and also that, when present, this amino acid strongly favors a two-stranded over a three-stranded configuration (Wagschal et al., 1999). This finding is in good agreement with our analysis of the structural database, which shows that Arg is relatively frequent at “*a*” positions in two-stranded coiled coils (4.5%) but much more rare in four-stranded coils (0.5%, supplementary Table S4.8.2). This preference can be explained by the solvent accessibility of the side chains: in two-stranded coiled coils, the narrower interface allows the polar moieties at the end of the side chains to access the surrounding solvent and remain partially water-exposed. As the size of the helical bundle increases, these groups become increasingly more buried, resulting in a higher desolvation cost.

Following the lead suggested by the MD analysis of our original FtsLB model (Condon et al., 2018), we report a revised model of FtsLB in which the periplasmic region splits into a pair of two-stranded coiled-coil domains, each containing one FtsL and one FtsB chain (Figure 4.6a). This model (which we named the “Y-model” from its shape) is based on the same set of side chain contacts between FtsB and FtsL inferred by the co-evolutionary analysis we used to derive the original monolithic model (Condon

et al., 2018) (named here the “I-model”). The transmembrane region is modeled in the same four-helix-bundle configuration of the original I-model. As in the I-model, the helix of FtsL in the coiled-coil region is continuous with the transmembrane helix, as indicated by experimental analysis (Condon et al., 2018). In this configuration, the FtsL helix remains on the outward face of the “Y”, whereas the helices of FtsB occupy the inward face, in close proximity to each other and thus buried within the overall arrangement. With respect to the transmembrane domain, the helix of FtsB is rotated so that its interfacial “a” and “d” positions face the corresponding positions of FtsL, a rotation that is readily enabled by the structurally malleable, Gly-rich linker that occurs between the transmembrane and coiled-coil helices of FtsB (Condon et al., 2018; LaPointe et al., 2013).

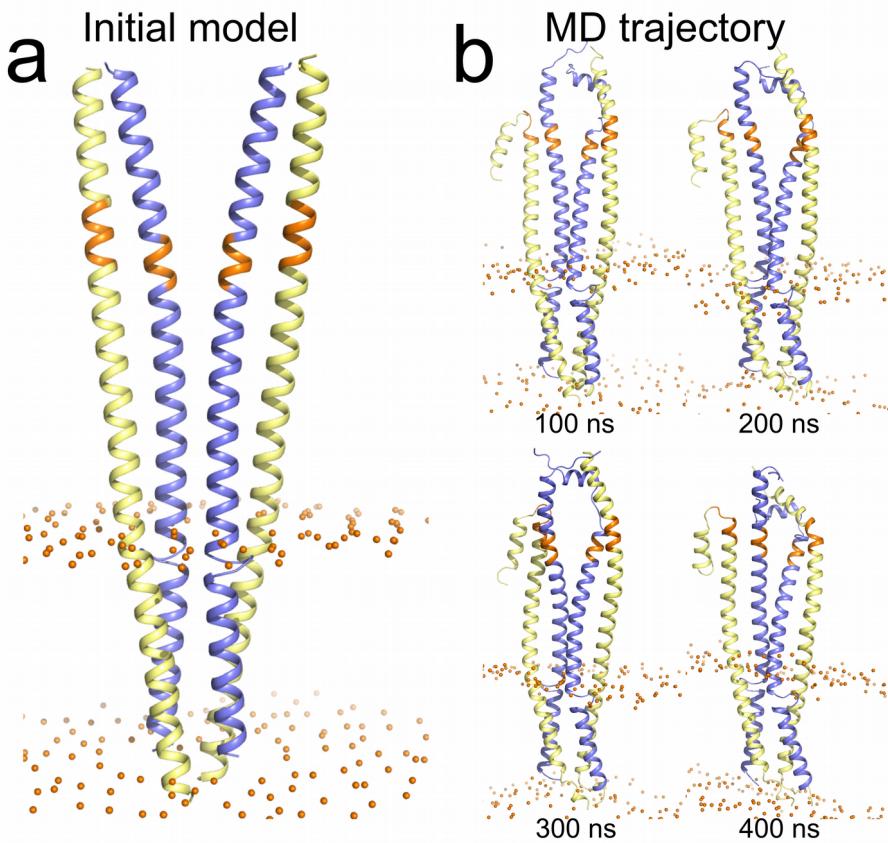


Figure 4.6: The Y-model of the FtsLB complex. *a*, Initial model. The model has the same transmembrane region of the original I-model. The coiled-coil region was modeled as a pair of two-stranded coiled-coil domains, each containing one FtsL and one FtsB helix. The modeling is based on the same set of side chain contacts between FtsB and FtsL inferred by the co-evolutionary analysis used to derive the I-model. Colored in orange is the CCD region. Spheres: lipid headgroup phosphate P atoms. *b*, Four frames of the trajectory of a 400 ns molecular dynamic run of the Y-model. The transmembrane and coiled-coil region remain relatively stable during the entire trajectory. The helices tend to break and

occasionally unfold in correspondence of a predicted hinge near the CCD region (orange). The RMSD analysis of all three replica run is reported in Figure S4.6 and in supplementary Table S4.8.4.

4.3.6. *AlphaFold2 modeling of FtsLB supports the structural features of the Y-model*

With the outstanding performance of the program AlphaFold2 (Jumper et al., 2021) at the recent CASP14 structural prediction competition, we decided to compare prediction of the FtsLB complex obtained with this method to our models. AlphaFold2 produced five models for FtsLB, which are ranked by their confidence score (supplementary Figure S4.5 and Table S4.8.1). Although the program was given a 2:2 FtsL:FtsB stoichiometry as the input, only the fifth ranked model produced a tetramer. Interestingly, the fifth model assumed a configuration that roughly resembles the Y-model, with its tetrameric assembly being mediated by the transmembrane region and two separated two-stranded coiled-coil domains (Figure S4.5a). However, rank model 5 and the Y-model align poorly (Ca RMSD of 11.11 Å). It should be noted that rank model 5 is a low-confidence model that is very loosely packed in its transmembrane region, and thus it is unlikely to be a good candidate for the structure of FtsLB.

The top four ranked AlphaFold2 models form two separated FtsB-FtsL heterodimers (Figure S4.5a). This is contrary to our experimental evidence, which indicates that FtsLB forms a heterotetramer (Condon et al., 2018; Khadria and Senes, 2013). In spite of this difference, however, ranked models 1 to 4 are structurally very similar to each

half of the tetrameric Y-model ($\text{C}\alpha$ RMSD values of 2.26-2.35 Å, Table S4.8.1), whereas their alignment with the I-model is less optimal ($\text{C}\alpha$ RMSD values of 3.26-3.47 Å). The AlphaFold2 models also display the same distinctive structural features that we identified, with identical interfaces, an unwound loop in the Gly-rich linker, and FtsL in a continuous helical configuration through the transmembrane and coiled-coil domains. The superimposition of AlphaFold2 ranked model 1 against the Y-model in supplementary Fig. S5c clearly illustrates that the two models are in excellent agreement. In summary, aside from the stoichiometry of the complex, the AlphaFold2 results provide a useful, independent validation of the structural organization of the Y-model.

4.3.7. The Y-model remains more stable in comparison to the I-model during MD simulations

To assess the stability of the Y-model and its dynamic properties, we performed MD simulations in explicit 1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) bilayers in conditions analogous to the previous MD simulations of the I-model (Condon et al., 2018). As done previously, the coiled coil was extended in the initial configuration by approximately 20 amino acids beyond its likely boundaries (the CCD region) to avoid end-effects. We will refer to this C-terminal region (residues 92-110 for FtsL and 62-79 for FtsB) as the post-CCD region. Three replica MD simulations were run for 400 ns each. The three trajectories are illustrated in Figure 4.6b and, in more detail, in supplementary Figure S4.6. Overall, we observed that the Y-model remained stable during the simulation time.

Similar to our previous simulations (Condon et al., 2018), the transmembrane region underwent only minor rearrangements, with average RMSDs of 2.4, 1.6, and 1.8 Å in the three replica runs (Figure S4.6, red traces, and supplementary Table S4.8.4). We found that the reconfigured coiled-coil region of the Y-model was very stable across all three replica runs with 1.4-1.7 Å average RMSDs (green and blue traces). For comparison, the average RMSDs for the same region of the I-model were 2.3-3.4 Å (Condon et al., 2018). The fact that the coiled coil of the Y-model remained well structured is in stark contrast with the simulation of the I-model, during which the four-stranded coiled coil opened, allowing water molecules into the core and in contact with the polar cluster.

As hypothesized, the terminal polar moieties of the long side chains (the amide group of FtsB Gln-39 and the guanidinium group of FtsL Arg-67 and Arg-74) are partially solvent exposed in the two-stranded coiled coils of the Y-model, whereas their nonpolar CH₂ groups contribute to hydrophobic packing at the interface. Solvent accessible surface area calculations indicate that on average the amide group of FtsB Gln-39 remains 20-40% solvent accessible, and the guanidinium groups of FtsL Arg-67 and 74 remain 40-60% accessible. The only polar group that remained nearly completely buried in the coiled coil was the shorter Asn-43 of FtsB, but its hydrogen bonding potential is satisfied by interacting with the backbone carbonyl group of Leu-70 and often with the side chain of Glu-73 on the opposed FtsL chain. In fact, the four side chains of the polar cluster tend to interact with each other in a shared network of hydrogen bonds (Figure 4.7b). Overall, the organization of the polar cluster appears clearly more favorable in

the Y-model in comparison with the I-model, in which these positions were unfavorably buried in a four-helix coiled-coil conformation.

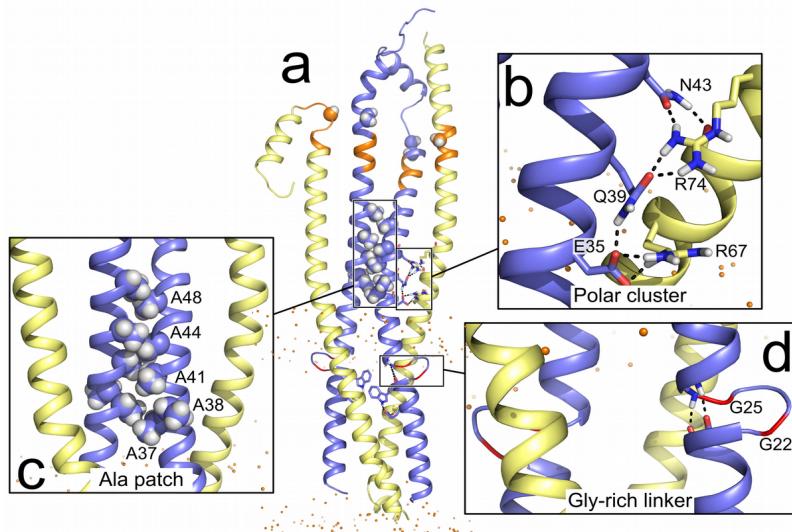


Figure 4.7: Features of the Y-model. *a*, MD frame of the Y-model highlighting three features of interest. *b*, The amide group of Gln-39 and the guanidinium group of Arg-67 and Arg-74 are partially solvent exposed whereas only the shorter Asn-43 is mainly buried. The polar cluster side chains tend to interact with each other in a shared network of hydrogen bonds. *c*, A patch of Ala residues on the opposite face of the FtsL/FtsB coiled-coil interfaces mediates the interaction of the two FtsB helices. The interaction is stable through the entire replica MD runs. *d*, The FtsB Gly-rich linker forms a short four-amino acid loop. The helical termini of the transmembrane and coiled-coil helices remain in contact in a configuration that resembles a continuous helix. There is hydrogen bond formation between the carbonyl groups of transmembrane residues 19-20 and the N-H groups of residues 26-27 of the coiled coil.

The coiled coils displayed a tendency to break near the CCD region, in correspondence to a likely hinge that contains Gly residues in both FtsB (positions 62 and 63) and FtsL (position 92). The terminal segments beyond this section are highly dynamic in our simulations, as indicated by their high RMSD traces (Figure S4.6, orange and purple traces). This is primarily because of the lever arm effect, since most of the RMSD arises from propagation of the unfolding of the hinge region, whereas the helical portion past the hinge stays mostly helical. The post-CCD segment of FtsB is essential for binding to FtsQ (Glas et al., 2015; Gonzalez and Beckwith, 2009) and is known to be structured in the FtsQLB complex (Choi et al., 2018; Kureisaite-Ciziene et al., 2018). This segment forms an 11-residue helix that starts right after the predicted hinge (position 64) and then associates with the C-terminal β -sheet of FtsQ by β -strand addition. FtsQ was not included in our simulations, and thus, it is not surprising that the C-terminal peptides unfolded in the simulation, although significant helical content was generally retained. The structure of the C-terminal segment of FtsL is not known, although presumably this terminal tail is also structured in the presence of FtsQ (or other components). In our simulations, this segment also tends to extend at the predicted hinge but otherwise retains significant helical content, similar to the corresponding segment of FtsB.

4.3.8. An Ala-rich patch in FtsB mediates intercoil contact in the Y-model

An interesting outcome of the MD simulations is a small but significant rearrangement of the two branches of the “Y”. The two coiled-coil domains are separated by approximately 7 Å in the initial model, but this gap closes in the

simulations. After this closure, the domains remain in permanent contact during all runs. Their contact is mediated by a mildly hydrophobic patch in FtsB consisting of five Ala residues that are clustered on the solvent-exposed, back side of the helix (positions 37, 38, 41, 44, and 48, Figure 4.7c). This interaction results in an additional helix-helix interface with an ~1,000 Å² contact area spanning approximately three helical turns and a right-handed crossing angle of approximately -35°.

The interaction forms rapidly within a few nanoseconds in runs 1 and 2. Once the contact is established, it persists for the entirety of the 400 ns runs, and thus, the two branches of the “Y” become locked into proximity. Only in replica run 3, the branches initially splay apart forming a more open “Y”, and it takes almost 40 ns for the branches to come in contact. In this run, when the contact is formed it is not symmetrical, involving Ala residue 41, 44, and 48 on one chain packing with residues 37, 41 and 44 in the other chain. Once this interaction is established, it also persists for the entirety of the 400 ns simulation, but it does not revert to the nearly symmetrical conformation observed in the other runs. Because of the asymmetry, one coil of FtsB appears to be pulled away from the transmembrane region, thus stretching the Gly-rich loop and breaking the contact between the termini of the transmembrane and coiled-coil helices.

To test whether the Ala patch is an essential feature of FtsLB, we drastically mutated the Ala residues and measured the resulting cell length phenotypes *in vivo*. All five Ala residues were mutated at once to combinations of Asp and Glu (AAAAA→EEEEDD and AAAAA→DDDEEE), according to the hypothesis that the replacement of hydrophobic Ala with negatively charged amino acids should destabilize

the interaction interface. As shown in Figure 4.8, the mutations did not cause elongation phenotypes, but their distributions were enriched in smaller cells compared to wild type (2.74 and 2.69 μm median length for the EEEDD and DDDEE mutants, respectively, compared to 2.93 μm for the wild type), in a fashion that is similar to the FtsL R74E mutant discussed previously.

Perhaps surprisingly, our experiments indicate that the disruption of the Ala patch feature does not cause major loss of function. If the Y-model is an accurate representation of the FtsLB complex, the data indicate that the added stability deriving from the interaction of the two coiled-coil branches is not a strong requirement. However, the observed enrichment in small cells is consistent with the hypothesis that fine-tuning the stability of the coiled coil participates in governing the balance between the *on* and *off* states of the complex. Specifically, it suggests that destabilization of the coiled coil tends to favor unregulated (early) cell division. It remains to be confirmed whether the hypothesized destabilization arises from the loss of the inter-branch hydrophobic interface or from direct destabilization of the coiled coil because of the presence of a patch of negative residues.

FtsB: Ala-patch mutants

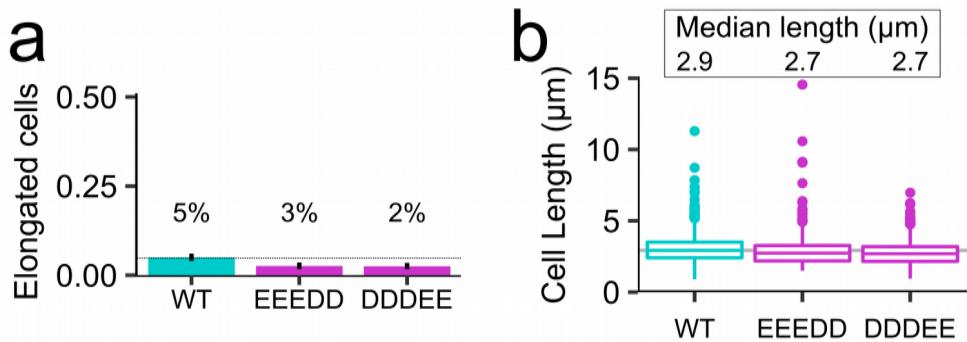


Figure 4.8: Disruption of the Ala-patch of FtsB results in smaller cells. A mildly hydrophobic patch of Ala residues is predicted to mediate the interaction of the two FtsB helices in the Y-model. All five Ala positions (37, 38, 41, 44, and 48) were simultaneously replaced with negatively charged residues, with a combination of three Glu and two Asp residues (EEEEDD) and vice versa (DDDEEE). *a*, Fraction of elongated cells for the mutants. Error bars represent the 95% confidence interval, estimated from 1000 replicates of bootstrap resampling. *b*, Median cell length distributions. The disruption of the potential interaction interface does not yield a defective division phenotype; however, a shift of the distribution with an increase of small cells suggests that the disruption of the Ala-patch induces some level of deregulated triggering of early cell division. Experiments were performed at 37 °C. Cell distributions are plotted in supplementary Figure S4.2.

4.3.9. The AWI positions of FtsL are surface-accessible in an initial model of the FtsQLB complex

The CCD and AWI positions are located in a region near the end of the predicted coiled coil in FtsLB, preceding a likely hinge created by three Gly residues in close proximity (positions 62-63 in FtsB and 92 in FtsL). CCD mutants rescue a $\Delta ftsN$ phenotype, suggesting that they induce changes in the FtsLB complex that mimic the activation signal given by FtsN (Liu et al., 2015; Tsang and Bernhardt, 2015). This process likely involves a conformational change that makes the AWI positions of FtsL available for interacting with FtsI, leading to activation of septal PG synthesis (Marmont and Bernhardt, 2020; Park et al., 2020).

In the Y-model, the AWI positions (82-84, 86-87, and 90) of FtsL are completely solvent exposed. Given that FtsLB exists as a complex with FtsQ, we assembled a preliminary model of the FtsQLB complex by aligning the region of FtsB that is in common between the Y-model and the cocrystal structure of the C-terminal FtsB fragment in complex with FtsQ (PDB code 6H9O (Kureisaite-Ciziene et al., 2018)). This fragment starts with the post-CCD helix (residues 64-74), following the putative hinge at the end of the coiled coil (62-63). Because of the flexible hinge, the orientation of the post-CCD helix is uncertain. However, the N-terminus of the periplasmic domain of FtsQ needs to be near the lipid bilayer for proper placement of its transmembrane helix. We found that by modeling the post-CCD helix as a nearly straight continuation of the coiled-coil helix, this constraint is satisfied and FtsQ is oriented in a position that does not collide with the FtsLB complex.

In this preliminary structural model of the FtsQLB complex (illustrated in Fig. 4.9*a* and *b*, with FtsQ depicted in gray), the surface of the coiled-coil region of FtsB is almost completely occluded. The helices of FtsB are flanked on one side by FtsL and on the opposite side by the helix of FtsB from the opposite coiled coil, and their remaining exposed face is then occluded by FtsQ. Conversely, the helix of FtsL packs only with its partner FtsB helix, whereas the opposite helical face remains exposed to solvent even in the presence of FtsQ. Interestingly, all amino acids of the AWI region of FtsL (in green) occur in this face, and thus remain solvent exposed. This configuration supports the AWI region as a good candidate surface for interaction with FtsWI. Conversely, the coiled coil of FtsB, which is nearly completely buried in the Y-model, would require major structural rearrangements to become available.

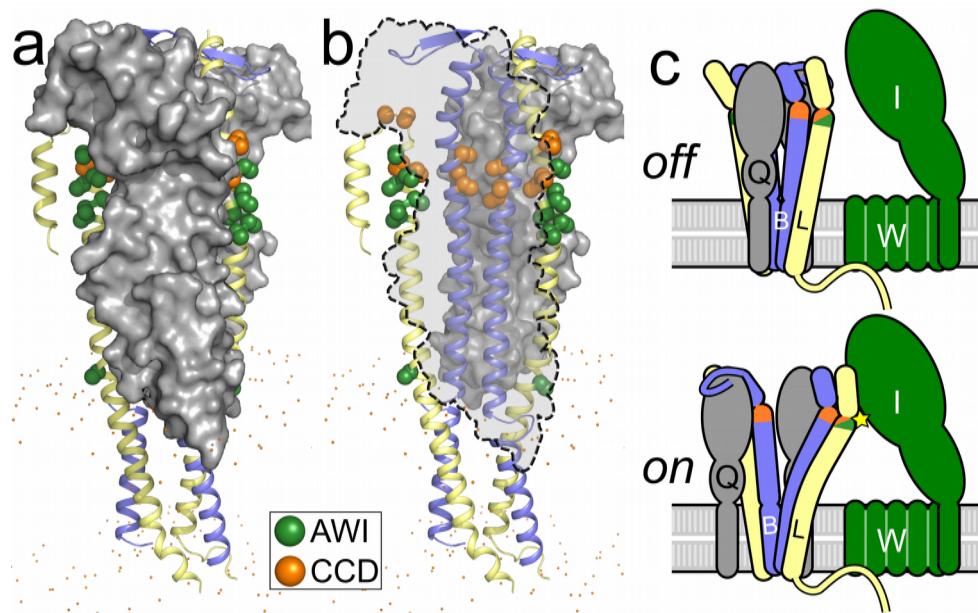


Figure 4.9: A structural shift in the coiled coil of FtsLB may be involved in its activation. *a*, Preliminary model of the FtsQLB complex, obtained by aligning the region of FtsB that is in common between the Y-model and the crystal structure of FtsQ (Protein Data Bank code: 6H9O). In this model, the four helices of the coiled coil of FtsLB form a flat bundle sandwiched between the two periplasmic domains of FtsQ (grey). *b*, See-through representation of the same model, with one of the two FtsQ subunits removed and represented by a dashed outline. The positions of the CCD and AWI domains of FtsB and FtsL are highlighted as orange and green spheres, respectively. FtsB is nearly completely buried. FtsL is more solvent exposed, and its AWI positions occur on a solvent-facing surface. *c*, Schematic representation of a potential model for FtsLB activation. A structural shift involving the coiled coil of FtsLB (potentially involving the separation of the two branches)

exposes the AWI region of FtsL to a hypothesized protein-protein interaction with FtsI (star). This transition may be favored by destabilization of the coiled-coil region.

4.4. Conclusions

In this article, we address the hypothesis that the coiled coil of FtsLB is a critical functional element. We explore the role of an unusual cluster of polar amino acids that occur at core positions within this domain and investigate whether the stability of the coiled coil is marginal by design. We propose that this characteristic may be critical for enabling the hypothesized FtsLB *off/on* transition that leads to the activation of the PG synthase machinery. Our evidence indicates that mutations affecting the conserved polar cluster lead to appreciable division phenotypes, indicating that this feature is functionally important. The fact that mutations to hydrophobic residues increase thermal stability *in vitro* but often lead to division defects *in vivo* suggests that the coiled-coil domain is indeed likely to be a “detuned” functional element, sacrificing some of the stability of the complex in order to enable proper function.

In addition to affecting the plasticity of the coiled coil, the polar cluster likely contributes to the shape of the periplasmic domain by causing it to branch out, as proposed in the revised Y-model of the complex. The branched Y-model satisfies all the known constraints (i.e., oligomeric state, inferred amino acid contacts, and mutagenesis data) as well as the original I-model, but it appears to be structurally more stable when assessed by MD, unlike the I-model in which the coiled coil was rapidly infiltrated by water during its simulation (Condon et al., 2018). The structural arrangement of the Y-

model is also in better agreement with predictions from AlphaFold2, after accounting for discrepancies between their overall stoichiometries. For these reasons, we propose the Y-model as the more likely configuration of the FtsLB complex.

This hypothesis that the polar cluster shapes the coiled coil into independent branches offers an interpretation for our observation *in vitro* that the “idealized” coiled coil (the “4×-mutant” FtsB_{Q39L,N43I}/FtsL_{R67I,R74I}) unfolds at higher temperatures but displays a broad noncooperative transition – an indication that the complex may assume multiple alternative configurations. In the “idealized” mutant, the coiled coil would not pay a steep penalty (i.e., the burial of polar side chains) to continue in the same four-stranded configuration of the transmembrane region, thus it is possible that a competition could occur between the I- and Y-configurations.

One of the most interesting outcomes of this analysis is the discovery of suppressor mutations that occur within the polar cluster. Specifically, changes in FtsL at position Arg-74 can suppress the phenotype of mutations at position Arg-67. The observations suggest a likely structural interplay between the two positions, where mutations cancel each other’s opposing tendencies to make division less likely (R67I and R67E, which display moderate elongation) or more likely (R74E, which is enriched in smaller cells), suggesting that these positions fine-tune the interactions that govern the balance between the *on* and *off* states of the complex. In addition, we noted that disruption of the mildly hydrophobic Ala-patch (EEEEDD and DDDEEE mutants) that mediates the contact between the two branches of the coiled coil in the Y-model MD simulations also results in a smaller than normal cell distribution. This finding suggests that separation of

the branches, which would be favored in the EEEEDD and DDDEEE mutants, facilitates the *on* transition. It is unclear if the R74E mutation operates in a similar manner, although in the model, substitution of R74 disrupts an intricate network of hydrogen bonding that connects the entire polar cluster, as illustrated in Figure 4.7*b*.

We propose a model in which the stability of the FtsLB coiled coil needs to be perfectly tuned. If the coiled coil is too stable, the complex does not respond appropriately to the FtsN-derived stimulus that normally causes activation of the complex. Conversely, if the coiled coil is further destabilized, FtsLB can become activated in an unregulated manner. Functional and metastable coiled-coil domains are known to be critical for other bacterial signaling processes such as those involving two-component systems (Schmidt et al., 2017), in which the coiled coil functionally connects the membrane-bound sensory domain with the response kinase domain in the cytoplasm. Specific examples include the *Staphylococcus aureus* antibiotic sensor NsaS (Bhate et al., 2018), the *Bacillus subtilis* thermosensor DesK (Albanesi et al., 2009; Fernández et al., 2019), and the *Bordetella pertussis* virulence factor BvgS (Lesne et al., 2016), each of which contains a coiled-coil domain with nonideal residues that enable conformational changes needed to regulate signaling activity. For these systems, the degree of hydrophobicity of the coiled-coil interface impacts the rigidity of the sensor and can even directly modulate the balance between kinase and phosphatase activity (Fernández et al., 2019; Lesne et al., 2016). The presence of nonideal residues at the interface of the FtsLB coiled coil suggests that a similar “tuning” of stability may be needed to support conformational rearrangements.

Further structural and biophysical characterization is necessary to validate this model of FtsLB and to fully understand how the structural organization of the coiled coil and its conformational changes operate in triggering septal PG reconstruction.

4.5. Experimental Procedures

4.5.1. Plasmid cloning

For the *in vivo* complementation experiments, mutant variants of FtsB or FtsL were cloned via standard QuikChange mutagenesis or inverse PCR into pMDG7 (Gonzalez and Beckwith, 2009) (flag3-FtsB) or pMDG29 (Gonzalez et al., 2010) (flag3-FtsL), respectively. For the CD experiments, the His-tagged FtsB and Strep-tagged Cys-less (C41A and C45A) FtsL_{35–121} were ligated into a modified pETDuet-1 vector at restriction sites Ncol/HindIII and Ndel/Xhol, respectively. Point mutations were introduced using standard QuikChange mutagenesis. All constructs were confirmed by DNA sequencing (Quintara Biosciences). A complete plasmid inventory is included in supplementary Table S4.8.3.

4.5.2. Bacterial strains, plasmids, and media for *in vivo* experiments

The phenotypic analyses were performed using depletion strains NB946 (Buddelmeijer et al., 2002) for FtsB and MDG277 (Gonzalez and Beckwith, 2009) for FtsL (both obtained from Jon Beckwith *et al.*) in which the WT copy of the protein of interest is under control of a repressible P_{BAD} promoter within the chromosome. These strains were transformed with plasmids containing either WT protein (positive control), empty vector (negative control), or a mutant version of the protein to test for defects in cell division as evidenced by an increase in cell length. For all experiments described,

bacterial cells were grown in LB medium supplemented with 100 µg/mL spectinomycin (Dot Scientific) and the appropriate carbon source. Medium was supplemented with 0.2% (w/v) L-arabinose (Sigma) or 0.2% (w/v) D-glucose (Sigma) to induce or repress, respectively, the expression of chromosomal copies of the WT genes regulated by the P_{BAD} promoter. About 20 µM isopropyl-β-D-1-thiogalactoside (IPTG) was added to the medium to induce the expression of mutant genes regulated by the P_{trc} promoter in the plasmid.

4.5.3. Depletion strain experiments

The protocol for the depletion strain experiments was adapted from Gonzalez and Beckwith (Gonzalez and Beckwith, 2009). In short, a mutated copy of FtsB or FtsL was transformed into its respective depletion strain. Strains were grown overnight at 37 °C on an LB plate supplemented with arabinose and spectinomycin. A single colony from the plate was grown overnight at 37 °C in 3 mL of LB medium supplemented with arabinose and spectinomycin. The overnight culture was then diluted 1:100 into fresh LB medium containing the same supplements and grown to an OD₆₀₀ of ~0.3. An aliquot of 1 mL of culture was washed twice with LB medium lacking any sugar and then diluted 1:100 into 3 mL of fresh LB medium supplemented with glucose, IPTG, and spectinomycin to induce expression of the mutated gene in the plasmid and to repress the WT gene in the chromosome. The cells were then grown at 37 °C for 3.5 hr, the approximate time necessary to deplete the cells of the WT chromosomal copy (Gonzalez and Beckwith, 2009). The cells were then placed on ice to stop growth before imaging. Depletion strains provided with a WT copy of their respective protein in the

plasmid were tested as positive controls, and, similarly, depletion strains with no protein in the plasmid (empty vector) were tested as negative controls.

4.5.4. Microscopy and cell length measurements

About 10 μL of cell samples were mounted on a number 1.5, 24 X 50 mm (0.16 – 0.19 mm thickness) cover glass slide (Fisher or VWR). Cells were cushioned with a 3% (w/v) agarose gel pad to restrict the movement of the live cells. Cells were optically imaged using a Nikon Eclipse Ti inverted microscope equipped with crossed polarizers and a Photometrics CoolSNAP HQ2 CCD camera using a Nikon X100 oil objective lens. Phase-contrast images of bacterial cells were recorded with a 50 ms exposure time using Nikon NIS Elements software. Multiple snapshots were collected for each experiment. All images were analyzed to measure the cell length in Oufti (Paintdakhi et al., 2016) using one single optimized parameter set and manual verification. Confidence intervals of the fraction of elongated cells for each mutant were computed from 1000 replicates of bootstrap resampling (Efron, 1979).

4.5.5. Western blots

Expression level across all variants was assessed by Western blot analysis (supplementary Figure S4.3). About 3.0 mL of cells were pelleted and resuspended in 300 μL of lysis buffer (50 mM HEPES pH 8.0, 50 mM NaCl) with 5 mM β -Mercaptoethanol (βME). The cells were sonicated and centrifuged at 21,000g for 10 min before collecting the supernatant. Total protein concentration was determined by bicinchoninic acid assay (Pierce). About 120 μL of lysates were mixed with 40 μL of 4× LDS sample buffer (Novex, Life Technologies) with βME and boiled at 98 °C for 3 min.

For each FtsL or FtsB sample, the equivalent of 7 µg or 15 µg, respectively, of total protein was separated by SDS-PAGE (Invitrogen) and transferred to polyvinylidene difluoride membrane (VWR). Horseradish peroxidase-tagged anti-FLAG (M2) antibodies (Sigma; 1:1,000) were used for immunoblotting analysis.

4.5.6. Protein expression and purification for CD

Plasmids were transformed into BL21(DE3) cells (NEB) and plated overnight at 37 °C on LB agar with 100 µg/mL ampicillin. Cells were washed off the plates with 1 mL LB broth and inoculated into 1 L of ZYP-5052 autoinduction medium as described (Studier, 2005) and grown at 37 °C until reaching an OD₆₀₀ of ~0.8, after which they were incubated overnight at 22 °C. Following expression, cells were pelleted, resuspended in cell wash buffer (100 mM NaCl, 10 mM HEPES pH 8.0), pelleted again, flash frozen, and stored at -80 °C for future use. The cells were then lysed by sonication in 10 mL/g lysis buffer (50 mM NaCl, 50 mM HEPES pH 8.0) supplemented with 0.5 mg/mL lysozyme, 5 mM β-Mercaptoethanol (βME), 1 mM phenylmethylsulfonyl fluoride, 1 mM EDTA, and a protease inhibitor cocktail providing (final concentrations) 8 µM leupeptin (Peptides International), 11.2 µM E-64 (Peptides International), 0.32 µM aprotinin (ProSpec), and 0.32 mM 4-(2-aminoethyl)benzenesulfonyl fluoride (Gold BioTechnology). The inclusion body fraction was separated by centrifugation at 10,000g for 20 min, followed by ultracentrifugation of the supernatant at 180,000g for 30 min to isolate the cell membranes. The FtsLB complex was then extracted from the membrane fraction with lysis buffer supplemented with 18 mM n-decyl-β-D-maltopyranoside (DM; Anatrace) and 5 mM βME, rocking at room temperature overnight. Nonresuspended

debris was separated from the solubilized protein via centrifugation at 10,000*g* for 20 min. The supernatant was added to ~3 mL of Ni-NTA-agarose resin (Qiagen) and rocked for 2 hr at 4 °C for batch binding before performing gravity-flow purification. Purification was performed by running 10 column volumes of Ni wash buffer (300 mM NaCl, 25 mM HEPES pH 8.0, 50 mM imidazole, 1 mM βME) supplemented with 510 μM n-dodecyl-β-D-maltopyranoside (DDM; Avanti Polar Lipids) and 10 column volumes of elution buffer (300 mM NaCl, 25 mM HEPES pH 8.0, 300 mM imidazole, 1 mM βME) also supplemented with 510 μM DDM. Protein purity was assessed via SDS-PAGE (Invitrogen).

4.5.7. CD experiments

Purified FtsLB protein was dialyzed twice at room temperature for at least 2 hr into 1 L CD buffer (10 mM phosphate buffer pH 7.4, 100 mM NaF) supplemented with 170 μM DDM (1× critical micelle concentration to prevent detergent exchange), then overnight at 4 °C in 1 L CD buffer supplemented with 510 μM DDM. Samples were kept at 4 °C or on ice from this point forward. Protein concentration was determined against the final dialysis buffer using A_{280} and an extinction coefficient of 32,430 M⁻¹cm⁻¹ for the FtsLB complex (calculated via ExPASy). Protein was diluted to ~14 μM, then filtered with 0.22 μm (13 mm diameter) PVDF syringe filters (CELLTREAT) before redetermining the final protein concentration. Samples were degassed in a vacuum chamber for at least 30 min, then centrifuged for 20 min at 21,000*g*. The final dialysis buffer was also filtered and degassed in the same manner to use as a blank in the CD experiments. CD spectra were obtained using an Aviv model 420 CD spectrometer and quartz cuvettes with a 0.1

cm pathlength. All spectra were recorded in 1 nm increments, with either a 10 s or 20 s averaging time, and after a 5 min equilibration time upon reaching a 0.3 °C deadband. The spectra were baseline corrected by buffer subtraction. For the CD-monitored thermal melting experiments, the samples were heated at 3 °C intervals with a 20 s equilibration time. Because the transitions were not reversible, detailed thermodynamic analyses were not carried out, and the curves were only fitted to sigmoidal transitions to calculate their temperature midpoints (T_m).

4.5.8. Bioinformatic analysis

Homologues of FtsB and FtsL were collected using the DELTA-BLAST algorithm (Boratyn et al., 2012) on the RefSeq database (O’Leary et al., 2016). FtsB-FtsL pairs were selected by the NCBI taxonomic identifier. In the case of multiple sequences per taxa, the one with the lowest E-value to the query *E. coli* FtsB or FtsL sequence was selected. Proteobacterial sequences were identified via the NCBI taxonomy database (Federhen, 2012). Sequences were aligned using the MAFFT algorithm (Katoh and Standley, 2013). Statistical analyses were performed in R (R Core Team, 2019) with the aid of the following packages: tidyverse (Wickham, 2017), tidymodels (Kuhn and Wickham, 2020), Biostrings (Pagès et al., 2019), zoo (Zeileis and Grothendieck, 2005), taxize (Chamberlain et al., 2019), rentrez (Winter, 2017), and tidygraph (Pedersen, 2019).

4.5.9. Molecular modeling

Modeling of the rearranged FtsLB complex was performed as described previously (Condon et al., 2018). Briefly, the FtsLB heterodimer was modeled using a Monte Carlo

procedure to model supercoiled helical bundles (Grigoryan and DeGrado, 2011). The superhelical radius (r_1), superhelical pitch (P), helical rotation (Φ_1), and z-shift (s) of both FtsL₅₂₋₉₄ and FtsB₂₁₋₆₃ were freely altered, whereas the rise per residue (h) and helical radius (r_0) were kept constant. Energies were calculated based on CHARMM 22 van der Waals and CHARMM 22 electrostatic terms with additional sigmoidal distance restraints for each pair of evolutionary couplings in the coiled-coil region (Condon et al., 2018). The heterodimeric FtsLB coiled coil was then aligned with one half of the previously modeled heterotetrameric transmembrane domain using residues 52-58 of FtsL, which were present in both models. Both domains were kept parallel to the Z-axis. The juxtamembrane regions of FtsL and FtsB were then replaced with loops corresponding to fragments from the PDB. For FtsB, six-residue loops (corresponding to positions 21–26) with four flanking helical residues on each side were used, with an additional sequence requirement that the fragment contain at least one glycine. For FtsL, 15-residue fragments with four flanking helical residues on each side were used with the requirement that the loop have helical secondary structure. This arrangement was made C₂-symmetric to generate the Y-model. Finally, the side chains were repacked using a greedy trials algorithm, and the model was minimized using BFGS constrained optimization in CHARMM (Brooks et al., 1983).

4.5.10. *Protein structure prediction using AlphaFold2*

The FtsLB complex was predicted with AlphaFold using the AlphaFold2_advanced notebook from Colabfold (Jumper et al., 2021; Mirdita et al., 2021), which allows for predictions of protein complexes. The full sequences of FtsB and FtsL were used as

input, with two chains for each protein to model a 2:2 heterotetramer. Mmseqs2 (Steinegger and Söding, 2017) was chosen for MSA generation, and the use_turbo option was enabled.

4.5.11. All-atom molecular dynamic simulations

For the MD simulations, the model's coiled-coil region was extended to avoid end-effects to residues 110 (FtsL) and 79 (FtsB). The cytoplasmic side of FtsL was also extended to include residues 30-34, modeled in ideal α -helix. Three 400 ns all-atom MD simulations were performed using the CHARMM36m force field (Huang et al., 2017) and NAMD 2.10 software (Klauda et al., 2010; Phillips et al., 2005). CHARMM-GUI membrane builder (Jo et al., 2008) was used to prepare systems composed of a POPE bilayer consisting of 301 lipids, the FtsLB tetramer, an ionic concentration of 0.150 M NaCl, and 59,163 TIP3P water molecules for hydration. The size of the boxes at the beginning of the simulation were approximately $97 \times 97 \times 242 \text{ \AA}^3$. The simulations were initially minimized and equilibrated for 75 ps at an integration time of 1 fs/step and for 600 ps at an integration time of 2 fs/step. The integration time step for the production runs of each of the systems was 2.0 fs/step. The simulations were carried out in the NPT ensemble at a pressure of 1 atmosphere and a temperature of 310.15 K, using the Nose-Hoover Langevin piston and Langevin dynamics method. Particle Mesh Ewald was used for electrostatic interactions, and a 12 \AA cutoff was applied to Lennard-Jones interactions with a switching function from 10 to 12 \AA . The RMSD analysis was performed using the RMSD trajectory tool in VMD (Humphrey et al., 1996). Hydrogen bonding analysis was performed with an in-house script.

4.6. Author contributions

S. J. C., S. G. F. C., and A. S. conceptualization; S. G. F. C., and G. D. V. software; S. J. C., S. G. F. C., G. D. V., and A. S. formal analysis; S. J. C., S. G. F. C., and G. D. V. investigation; S. J. C., S. G. F. C., and A. S. writing-original draft; G. D. V. and Q. C. writing-review & editing; Q. C. and A. S. supervision; Q. C. and A. S. funding acquisition.

4.7. Acknowledgments

This work was supported in part by National Institutes of Health (NIH) Grant R01-GM099752 and R35-GM130339 to A.S. and NSF Grant CHE-1829555 to Q.C. S.J.C. was supported in part by a Dr. James Chieh-Hsia Mao Wisconsin Distinguished Fellowship. S.G.F.C. was supported in part by the Arthur B. Michael Departmental Fellowship and the William R. & Dorothy E. Sullivan Wisconsin Distinguished Graduate Fellowship. G.D.V. was supported in part by a SciMed GRS Fellowship.

4.8. Supplementary figures and tables

Table S4.8.1: Alignment scores ($C\alpha$ RMSD) of the five AlphaFold2 dimeric FtsLB models against half of the Y- and I-models (FtsB_A 1-60, FtsL_C 40-91)

Alphafold 2 Model	pLDDT¹	Y-model $C\alpha$ RMSD (Å)	I-model $C\alpha$ RMSD (Å)
Rank_1	79.78	2.33	3.47
Rank_2	78.12	2.26	3.40
Rank_3	77.63	2.27	3.36
Rank_4	76.35	2.35	3.26
Rank_5	61.94	2.71	2.31

¹pLDDT score (predicted Local Distance Difference Test): Alphafold2's overall confidence metric of the model (100 = most confident, 0 = least confident)

Table S4.8.2: Frequency of polar amino acids at a and d positions in a database of 2,662 crystal structures of coiled coils from the CC+ database.

Two-stranded parallel coiled coils			
Amino acid	a positions ¹	d positions ¹	a+d positions ¹
D	0.97%	1.13%	1.05%
E	3.34%	4.26%	3.80%
H	1.79%	1.46%	1.63%
K	4.72%	2.79%	3.76%
N	3.22%	1.72%	2.47%
Q	3.45%	2.57%	3.01%
R	4.47%	1.59%	3.03%
Total polar	21.97%	15.53%	18.76%

Three-stranded parallel coiled coils			
Amino acid	a positions ¹	d positions ¹	a+d positions ¹
D	0.38%	0.76%	0.57%
E	2.05%	1.39%	1.72%
H	1.15%	1.46%	1.31%
K	0.83%	0.95%	0.89%
N	1.86%	4.50%	3.19%
Q	5.00%	5.26%	5.13%
R	1.09%	1.46%	1.27%
Total polar	12.38%	15.77%	14.09%

Four-stranded parallel coiled coils			
Amino acid	a positions ¹	d positions ¹	a+d positions ¹
D	0.16%	1.20%	0.69%
E	2.86%	2.07%	2.46%
H	1.14%	1.59%	1.37%
K	1.14%	0.48%	0.81%
N	1.39%	1.75%	1.57%
Q	2.45%	5.18%	3.83%
R	0.49%	1.43%	0.97%
Total polar	9.63%	13.71%	11.69%

¹Pseudocount-Adjusted frequency

Table S4.8.2, continued

All parallel coiled coils			
Amino acid	<i>a</i> positions ¹	<i>d</i> positions ¹	<i>a+d</i> positions ¹
D	0.84%	1.08%	0.96%
E	3.14%	3.77%	3.46%
H	1.69%	1.45%	1.57%
K	4.00%	2.38%	3.19%
N	2.93%	1.99%	2.46%
Q	3.50%	3.10%	3.30%
R	3.77%	1.54%	2.65%
Total polar	19.87%	15.32%	17.60%

¹Pseudocount-Adjusted frequency

Table S4.8.3: Strains and plasmids used in this work.

Strain/ plasmid	Description	Parent vector	Source
BL21(DE 3)	chemically competent <i>E. coli</i> for protein overexpression	-	New England BioLabs (C2527)
NB946	FtsB depletion strain	-	Buddelmeijer et al., 2002
MDG277	FtsL depletion strain	-	Gonzalez & Beckwith, 2009
pSJC020	His-FtsB Strep-FtsL35- 121 C41A/C45A (Cys-less FtsL)	pETDuet-1	Condon et al., 2018
pSJC309	His-FtsB Q39L/N43I Strep-FtsL35-121 C41A/C45A/R67I/R74I	pETDuet-1	This paper
pNG162	IPTG-inducible, low-copy- number vector (empty)	pAM238	Goehring et al., 2006
pMDG7	flag3-FtsB	pNG162	Gonzalez & Beckwith, 2009
pSJC187	flag3-FtsB Q39L	pMDG7	This paper
pSJC188	flag3-FtsB N43I	pMDG7	This paper
pSJC208	flag3-FtsL N50I	pMDG7	This paper
pSJC287	flag3-FtsB A37D/A38D/A41D/A44E/A 48E	pMDG7	This paper
pSJC288	flag3-FtsB A37E/A38E/A41E/A44D/A 48D	pMDG7	This paper
pMDG29	flag3-FtsL	pNG162	Gonzalez et al., 2010
pSJC183	flag3-FtsL R67I	pMDG29	This paper

Table S4.8.3, continued

Strain/ plasmid	Description	Parent vector	Source
pSJC185	flag3-FtsL R74I	pMDG29	This paper
pSJC190	flag3-FtsL R67I/R74I	pMDG29	This paper
pSJC194	flag3-FtsL R67E	pMDG29	This paper
pSJC201	flag3-FtsL R74E	pMDG29	This paper
pSJC220	flag3-FtsL W81I	pMDG29	This paper
pSJC254	flag3-FtsL R67E/R74E	pMDG29	This paper
pSJC304	flag3-FtsL R67I/R74E	pMDG29	This paper
pSJC324	flag3-FtsL R67K	pMDG29	This paper
pSJC325	flag3-FtsL R74K	pMDG29	This paper
pSJC326	flag3-FtsL R67E/R74I	pMDG29	This paper
pSJC334	flag3-FtsL R74E/W81I	pMDG29	This paper
pSJC335	flag3-FtsL R74K/W81I	pMDG29	This paper
pSJC336	flag3-FtsL R67I/R74K	pMDG29	This paper
pSJC337	flag3-FtsL R67E/R74K	pMDG29	This paper

Table S4.8.4: RMSD analysis of the three replica molecular dynamic runs of the FtsLB complex in the Y-model configuration

	RMSD average ± S.D. (Å)	Min RMSD (Å)	Max RMSD (Å)
Run 1			
Protein	8.3 ± 0.9	3.15	10.03
TM region	2.4 ± 0.4	1.41	2.98
Coiled coil domain, chain A and B	1.7 ± 0.4	0.85	2.94
Coiled coil domain, chain C and D	1.5 ± 0.3	0.82	2.43
Coiled coil, both domains	3.5 ± 0.6	1.55	5.16
Post-CCD, chain A and B	14.4 ± 1.8	2.46	17.00
Post-CCD, chain C and D	5.1 ± 0.7	2.74	6.76
Run 2			
Protein	7.7 ± 1.4	2.44	9.46
TM region	1.6 ± 0.3	0.92	2.05
Coiled coil domain, chain A and B	1.6 ± 0.3	0.89	2.35
Coiled coil domain, chain C and D	1.7 ± 0.3	0.88	2.57
Coiled coil, both domains	4.5 ± 0.7	1.63	6.14
Post-CCD, chain A and B	5.2 ± 1.8	1.27	11.69
Post-CCD, chain C and D	5.8 ± 1.5	1.99	7.88
Run 3			
Protein	7.5 ± 1.5	2.61	14.41
TM region	1.8 ± 0.2	1.24	2.32
Coiled coil domain, chain A and B	1.6 ± 0.3	0.89	2.35
Coiled coil domain, chain C and D	1.7 ± 0.3	0.88	2.57
Coiled coil, both domains	5.4 ± 1.0	2.06	10.65
Post-CCD, chain A and B	5.2 ± 1.8	1.27	11.69
Post-CCD, chain C and D	5.8 ± 1.5	1.99	7.88

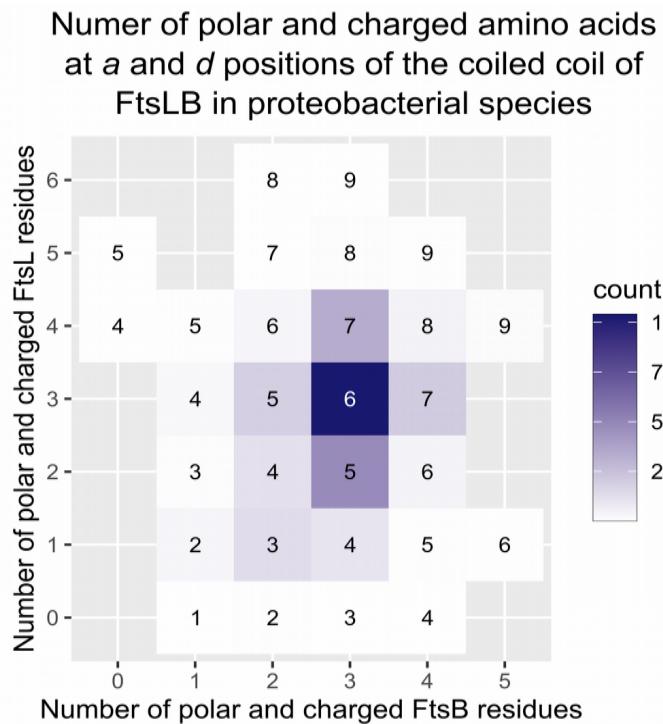


Figure S4.1: Conservation of the polar cluster in the core “a” and “d” positions of the coiled coil of FtsLB in proteobacterial species. Number of sequences with a given number of polar/charged amino acids (Asp, Glu, His, Arg, Lys, Gln, and Asn) contributed by FtsB (X axis) and FtsL (Y axis) at “a” and “d” positions in the five heptad repeats of the coiled coil (their sum is reported in the box). Data from an alignment of 2900 paired FtsB/FtsL proteobacterial sequences. The most frequent combination corresponds, by a large margin, to three polar residues contributed by both FtsB and FtsL, for a total of six polar residues. Although less frequent, combinations with a total of five or seven polar residues also occur relatively often. Overall, 84% of the sequences contain at least 5 polar residues.

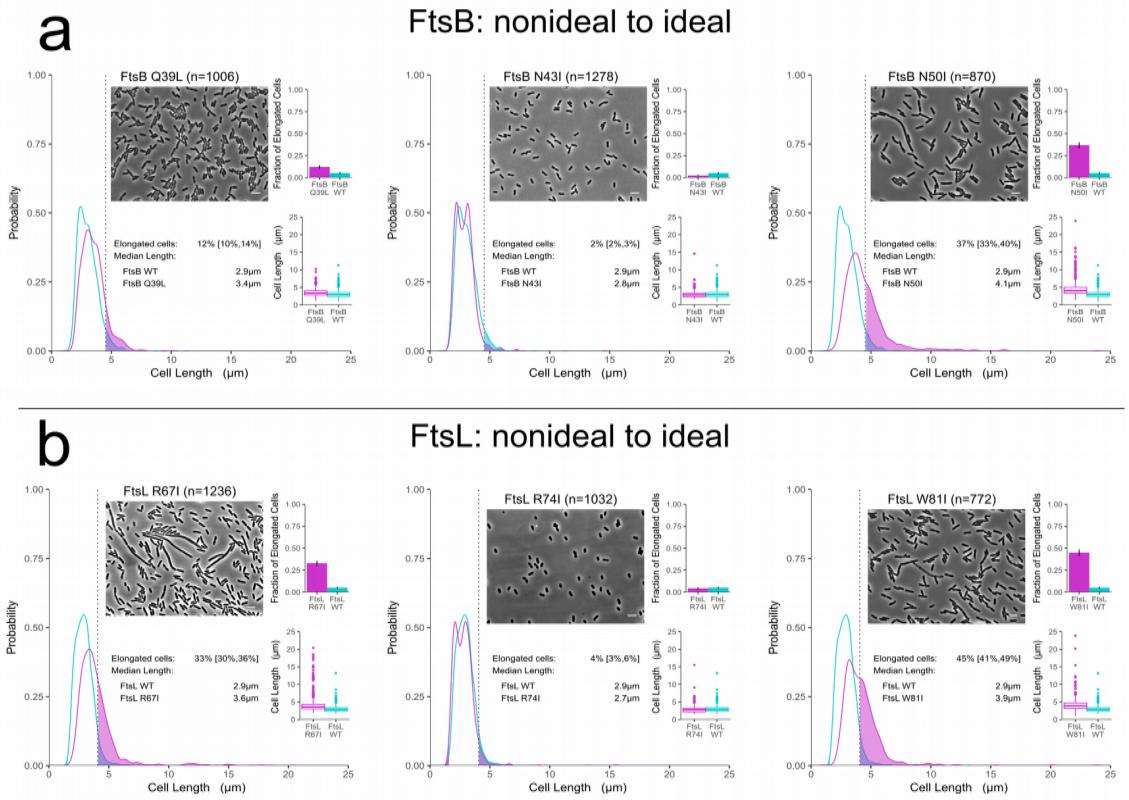


Figure S4.2: Cell length distribution of mutants. Phase-contrast images and cell length distributions of mutant cells (magenta) compared to wild-type (cyan). White scale bar = 5 μm . The shaded areas to the right of the dotted lines represent the fraction of cells that are longer than the 95th percentile in the WT distribution. The fraction is plotted as a histogram in the top inset. Error bars and bracketed values represent the 95% confidence interval for the fraction of elongated cells, estimated by 1000 bootstrap replications of the samples. The bottom inset is a box and whisker plot of the same distribution of the main panel.

a, Nonideal to ideal mutations in FtsB. *b*, Nonideal to ideal mutations in FtsL. Next pages: *c*, Charge variations in FtsL at R67 and R74 positions. *d*, Double

substitutions in FtsL at R67, R74, and W81 positions. e, Ala patch substitutions in FtsB in which all five Ala positions (37, 38, 41, 44, and 48) were simultaneously replaced with negatively charged residues, with a combination of three Glu and two Asp residues (EEEDD) and vice versa (DDDEE). Experiments were performed at 37 °C.

(continued on next page)

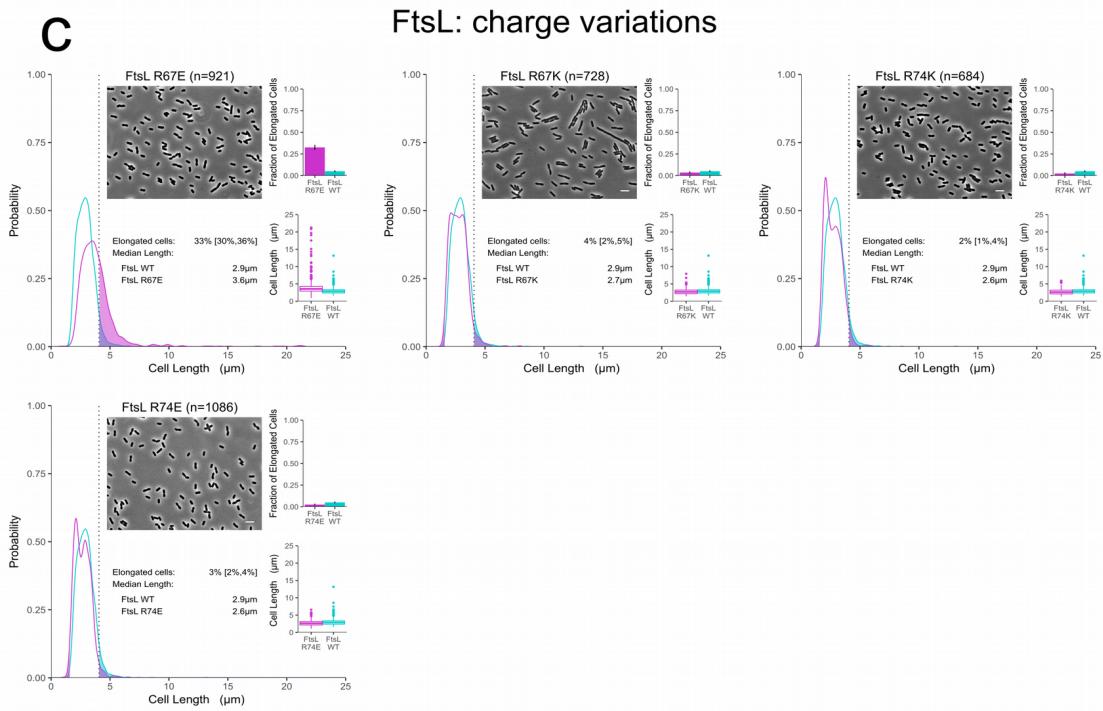


Figure S4.2: (continued from previous previous page and continues on next page)

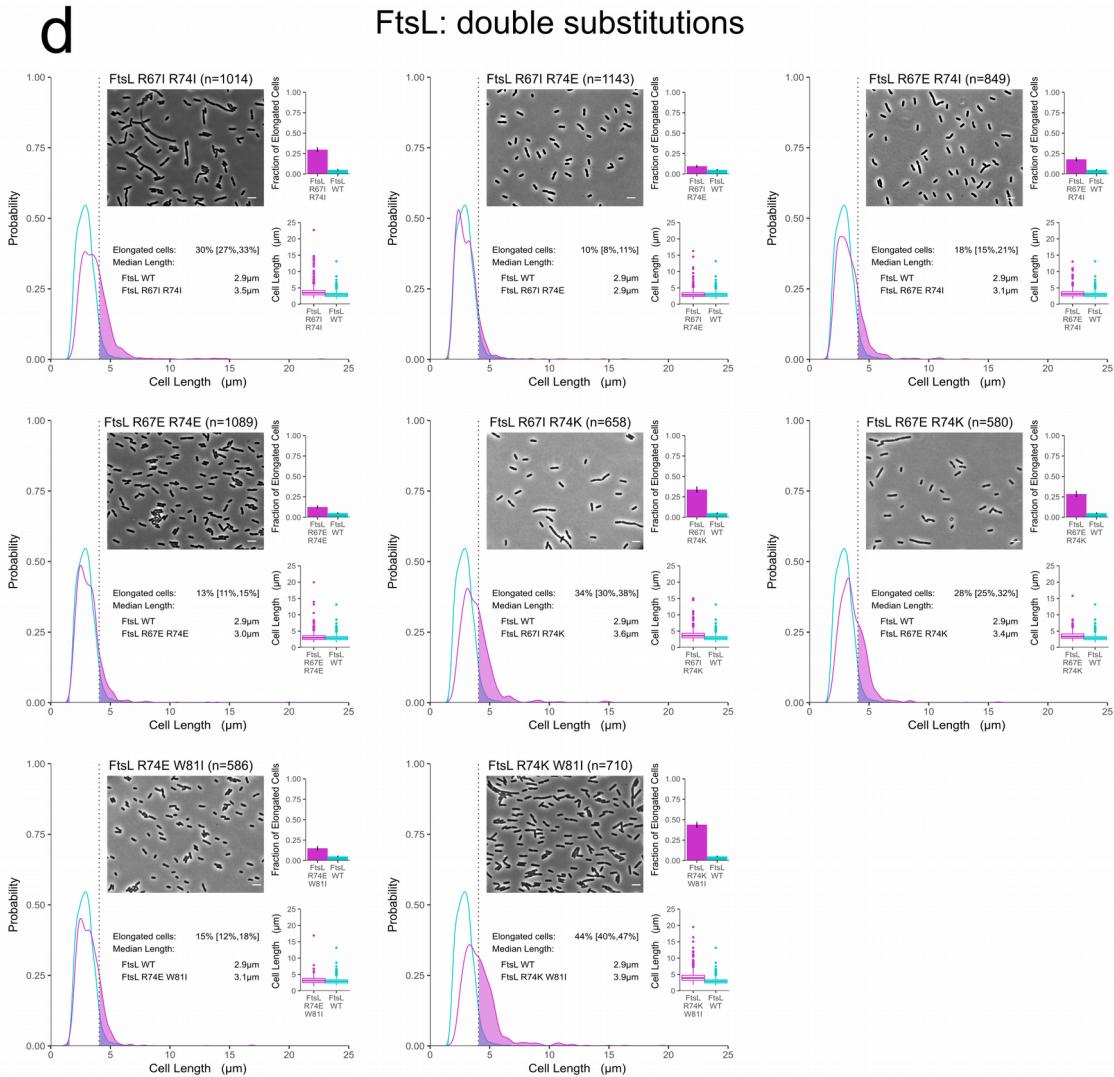
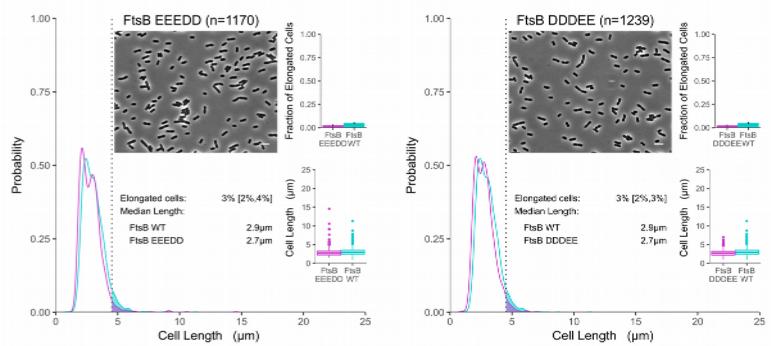


Figure S4.2: (continued from previous previous page and continues on next page)

e**FtsB: Ala patch substitutions****Figure S4.2: (continued from previous previous page).**

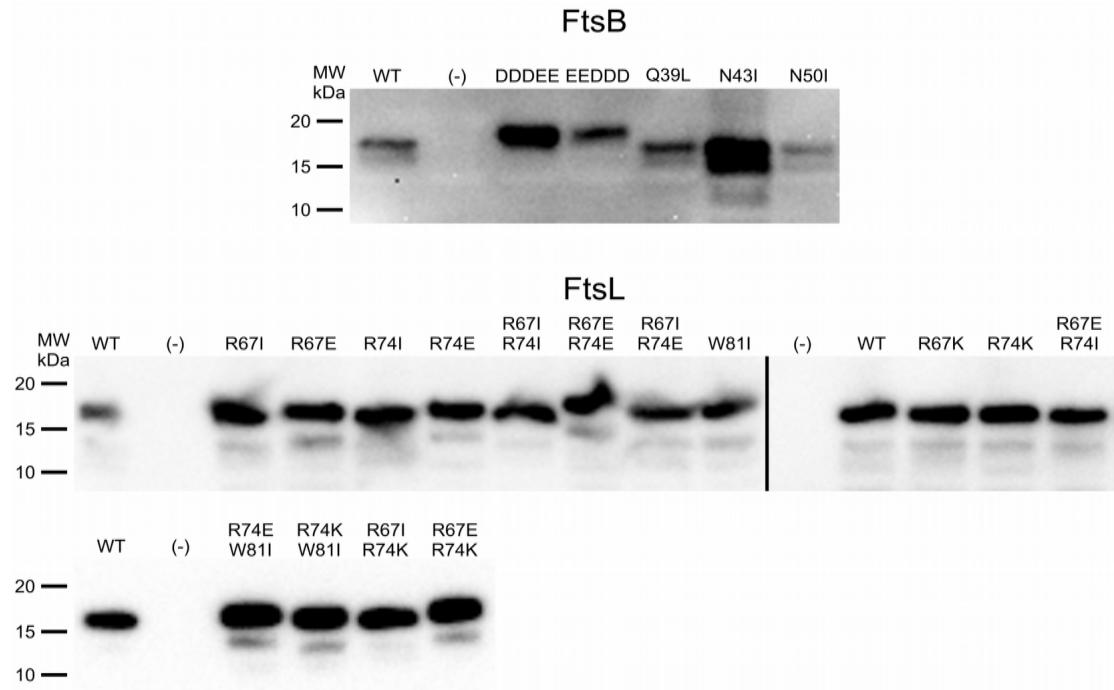


Figure S4.3: Expression level of FtsB and FtsL mutants assessed by Western Blot analysis. Image of all western blots of the FtsB and FtsL mutants tested in this work. Around twice as much whole cell lysate (normalized to total protein) was loaded for FtsB samples as for FtsL samples. Protein expression level of the FtsB and FtsL mutants with defective phenotypes are generally comparable to the respective wild type (WT). Negative controls (-) show no detectable signal for either protein. DDDEE is FtsB A37D/A38D/A41D/A44E/A48E, and EEDDD is FtsB A37E/A38E/A41E/A44D/A48D. Both show a slightly increased molecular weight, which may be due to the increased number of negatively charged residues in these mutants. There are cases of FtsB mutants with increased protein level

(N43I, in particular), though it is unclear why. Individual gels are separated by solid lines.

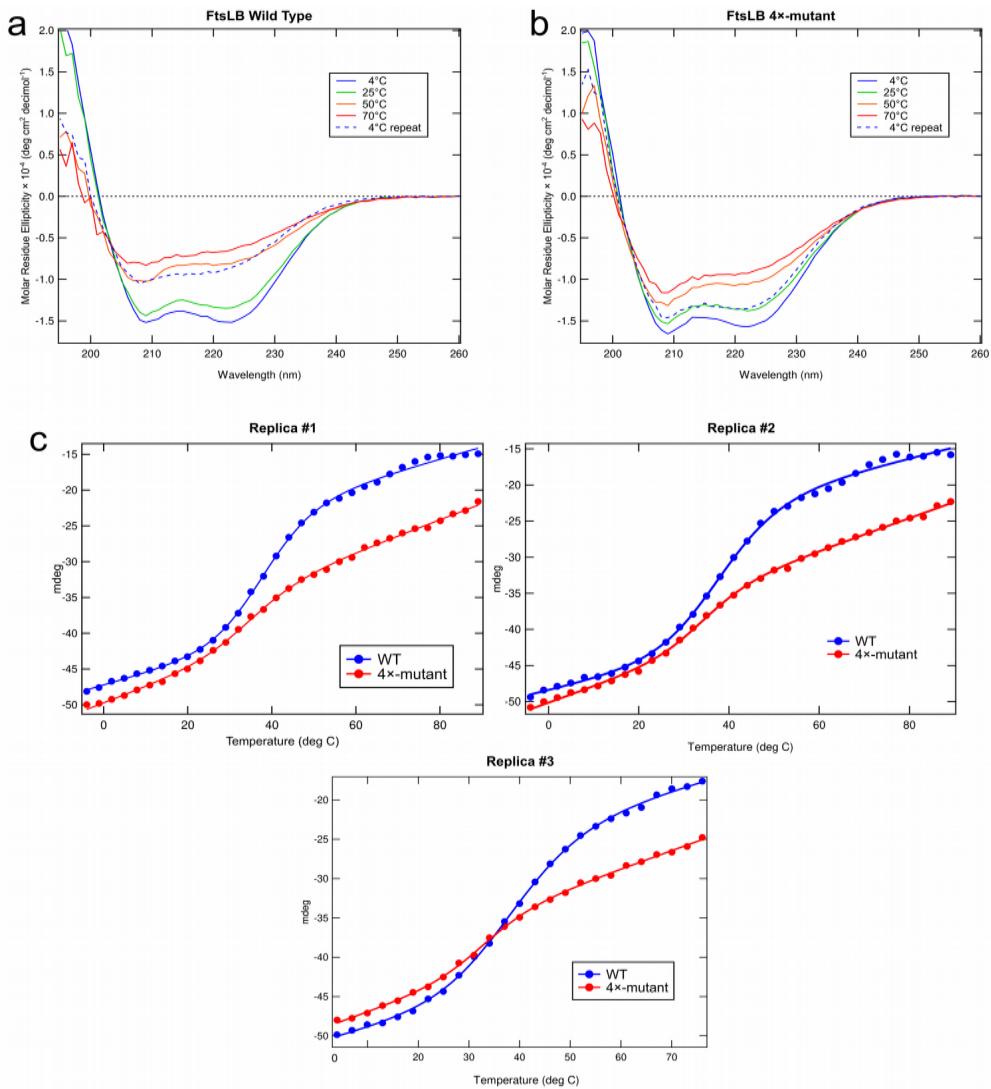


Figure S4.4: *a*, Far-UV CD spectra of WT FtsLB at 4, 25, 50, and 70 °C along with a 4 °C repeat after cooling the sample. *b*, Same analysis of the 4x-mutant. *c*, Replica experiments of temperature melting curves comparing WT FtsLB (blue) to the 4x-mutant (red). CD scans were monitored at 224 nm from 4 °C to 89 °C.

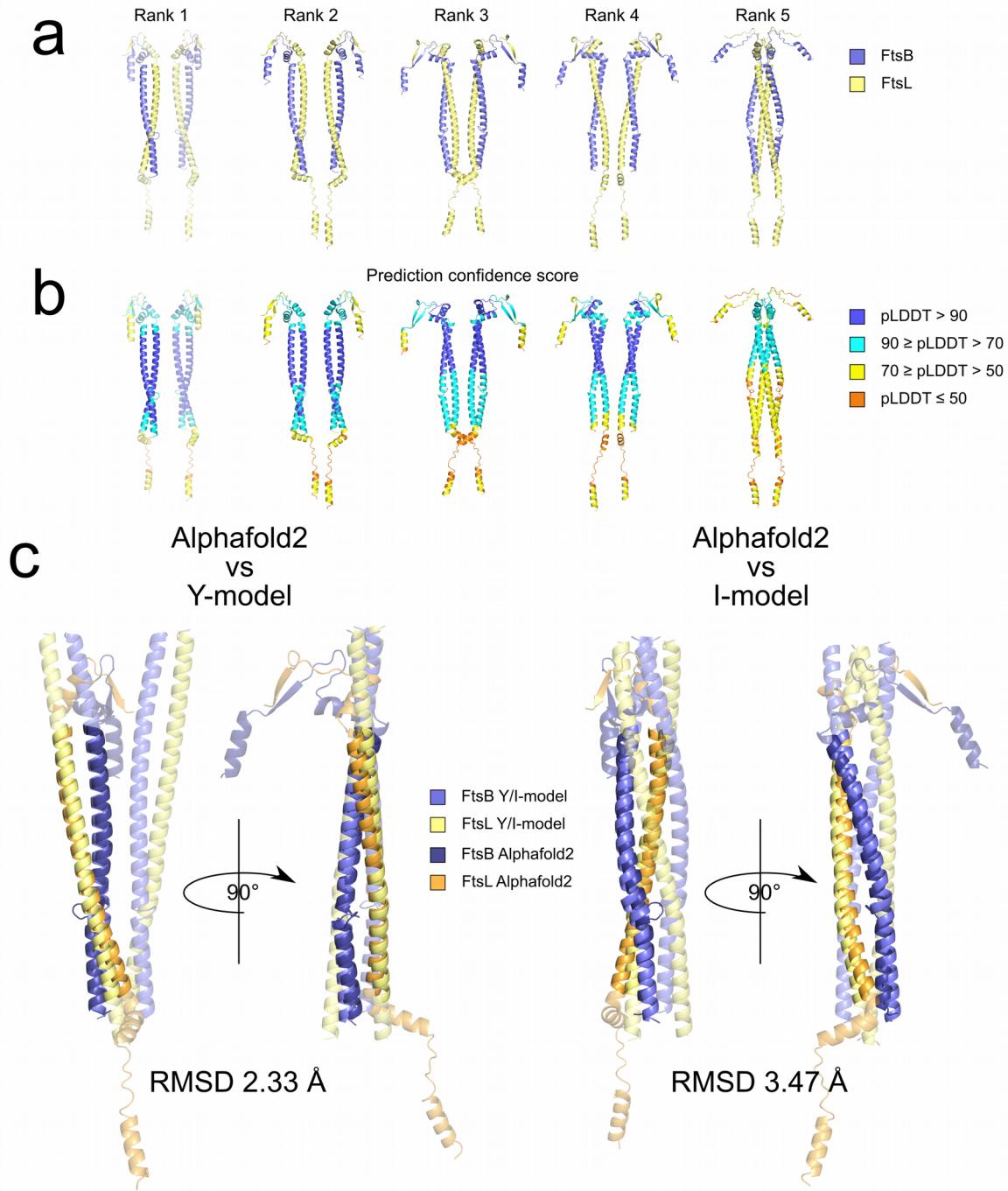


Figure S4.5: Alphafold2 models of FtsLB. *a*, Alphafold2 produced five ranked models of FtsLB, numbered 1-5 (best to worst). In spite of four subunits being

provided, rank models 1-4 display separated dimeric units with little or no interaction between them. Rank model 5 is organized in a tetrameric unit. However, the interactions between the pair of dimers are loose and significantly underpacked. *b*, The AlphaFold2 models have high confidence in the coiled-coil region, as color coded in the figure (100: highest confidence; 0: lowest confidence). *c*, Alignments of one dimer from the rank 1 model to half of the Y-model (left) and the I-model (right). Regions not used in the alignment are transparent. The alignment with the Y-model is excellent, with a C α RMSD of 2.33 Å, while the I-model alignment is less optimal, with an RMSD of 3.47 Å.

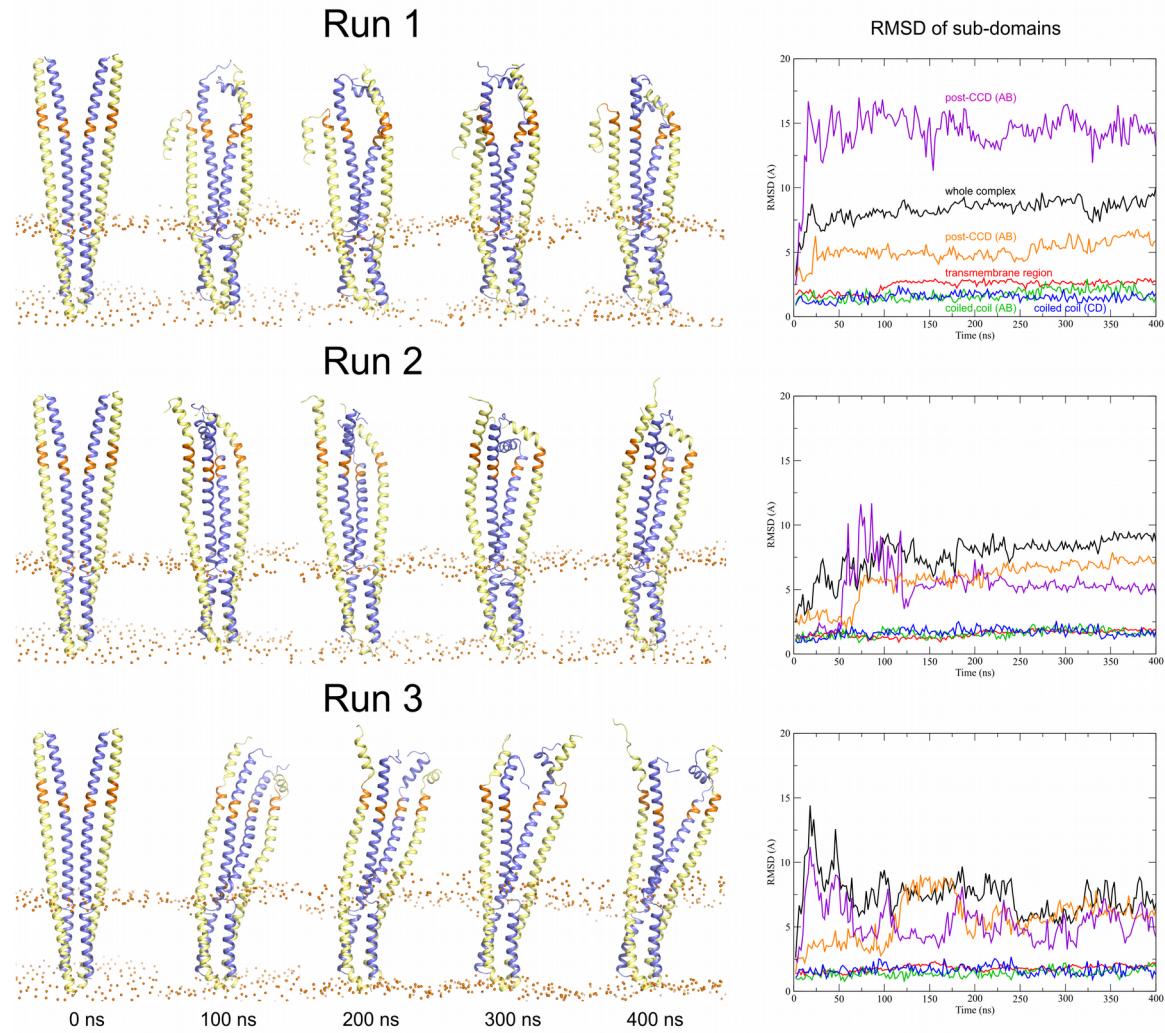


Figure S4.6: Trajectories of the three 400 ns replica MD runs of the FtsLB Y-model. The graph illustrates the fluctuations of the RMSD of the entire complex (black) and the individual subdomains: red: transmembrane region; green: coiled coil, chains A (FtsB) and B (FtsL); blue: coiled coil, chains C (FtsB) and D (FtsL); magenta: post-CCD region, chains A and B; orange: post-CCD region, chains C and D. The RMSD indicates that the transmembrane domain and coiled-coil

domains remain relatively stable during the entirety of the simulations. The RMSD analysis is summarized in supplementary Table S4.8.4.

4.9. References

- Albanesi, D., Martín, M., Trajtenberg, F., Mansilla, M.C., Haouz, A., Alzari, P.M., de Mendoza, D., and Buschiazzo, A. (2009). Structural plasticity and catalysis regulation of a thermosensor histidine kinase. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 16185–16190. <https://doi.org/10.1073/pnas.0906699106>.
- Bhate, M.P., Lemmin, T., Kuenze, G., Mensa, B., Ganguly, S., Peters, J.M., Schmidt, N., Pelton, J.G., Gross, C.A., Meiler, J., et al. (2018). Structure and Function of the Transmembrane Domain of NsaS, an Antibiotic Sensing Histidine Kinase in *Staphylococcus aureus*. *J. Am. Chem. Soc.* **140**, 7471–7485. <https://doi.org/10.1021/jacs.7b09670>.
- den Blaauwen, T., Hamoen, L.W., and Levin, P.A. (2017). The divisome at 25: the road ahead. *Curr. Opin. Microbiol.* **36**, 85–94. <https://doi.org/10.1016/j.mib.2017.01.007>.
- Booth, S., and Lewis, R.J. (2019). Structural basis for the coordination of cell division with the synthesis of the bacterial cell envelope. *Protein Sci. Publ. Protein Soc.* **28**, 2042–2054. <https://doi.org/10.1002/pro.3722>.
- Boratyn, G.M., Schäffer, A.A., Agarwala, R., Altschul, S.F., Lipman, D.J., and Madden, T.L. (2012). Domain enhanced lookup time accelerated BLAST. *Biol. Direct* **7**, 12. <https://doi.org/10.1186/1745-6150-7-12>.
- Botta, G.A., and Park, J.T. (1981). Evidence for involvement of penicillin-binding protein 3 in murein synthesis during septation but not during cell elongation. *J. Bacteriol.* **145**, 333–340. .
- Brooks, B.R., Brucolieri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., and Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.* **4**, 187–217. <https://doi.org/10.1002/jcc.540040211>.
- Buddelmeijer, N., and Beckwith, J. (2004). A complex of the *Escherichia coli* cell division proteins FtsL, FtsB and FtsQ forms independently of its localization to the septal region. *Mol. Microbiol.* **52**, 1315–1327. <https://doi.org/10.1111/j.1365-2958.2004.04044.x>.
- Buddelmeijer, N., Judson, N., Boyd, D., Mekalanos, J.J., and Beckwith, J. (2002). YgbQ, a cell division protein in *Escherichia coli* and *Vibrio cholerae*, localizes in codependent

fashion with FtsL to the division site. Proc. Natl. Acad. Sci. U. S. A. 99, 6316–6321. <https://doi.org/10.1073/pnas.092128499>.

Chamberlain, S., Szoecs, E., Foster, Z., Arendsee, Z., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., O'Donnell, J., Oksanen, J., et al. (2019). taxize: Taxonomic information from around the web.

Cho, H., Wivagg, C.N., Kapoor, M., Barry, Z., Rohs, P.D.A., Suh, H., Marto, J.A., Garner, E.C., and Bernhardt, T.G. (2016). Bacterial cell wall biogenesis is mediated by SEDS and PBP polymerase families functioning semi-autonomously. Nat. Microbiol. 1, 16172. <https://doi.org/10.1038/nmicrobiol.2016.172>.

Choi, Y., Kim, J., Yoon, H.-J., Jin, K.S., Ryu, S., and Lee, H.H. (2018). Structural Insights into the FtsQ/FtsB/FtsL Complex, a Key Component of the Divisome. Sci. Rep. 8, 18061. <https://doi.org/10.1038/s41598-018-36001-2>.

Condon, S.G.F., Mahbuba, D.-A., Armstrong, C.R., Diaz-Vazquez, G., Craven, S.J., LaPointe, L.M., Khadria, A.S., Chadda, R., Crooks, J.A., Rangarajan, N., et al. (2018). The FtsLB subcomplex of the bacterial divisome is a tetramer with an uninterrupted FtsL helix linking the transmembrane and periplasmic regions. J. Biol. Chem. 293, 1623–1641. <https://doi.org/10.1074/jbc.RA117.000426>.

Du, S., and Lutkenhaus, J. (2017). Assembly and activation of the Escherichia coli divisome. Mol. Microbiol. 105, 177–187. <https://doi.org/10.1111/mmi.13696>.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. Ann. Stat. 7. <https://doi.org/10.1214/aos/1176344552>.

Egan, A.J.F., and Vollmer, W. (2013). The physiology of bacterial cell division. Ann. N. Y. Acad. Sci. 1277, 8–28. <https://doi.org/10.1111/j.1749-6632.2012.06818.x>.

Egan, A.J.F., Errington, J., and Vollmer, W. (2020). Regulation of peptidoglycan synthesis and remodelling. Nat. Rev. Microbiol. 18, 446–460. <https://doi.org/10.1038/s41579-020-0366-3>.

Federhen, S. (2012). The NCBI Taxonomy database. Nucleic Acids Res. 40, D136–143. <https://doi.org/10.1093/nar/gkr1178>.

Fernández, P., Porrini, L., Albanesi, D., Abriata, L.A., Dal Peraro, M., de Mendoza, D., and Mansilla, M.C. (2019). Transmembrane Prolines Mediate Signal Sensing and Decoding in *Bacillus subtilis* DesK Histidine Kinase. MBio 10. <https://doi.org/10.1128/mBio.02564-19>.

Frapont, C., Alexeeva, S., Wolf, B., van der Ploeg, R., Schloesser, M., den Blaauwen, T., and Nguyen-Distèche, M. (2011). The integral membrane FtsW protein and

peptidoglycan synthase PBP3 form a subcomplex in *Escherichia coli*. *Microbiol. Read. Engl.* 157, 251–259. <https://doi.org/10.1099/mic.0.040071-0>.

Gerding, M.A., Liu, B., Bendezú, F.O., Hale, C.A., Bernhardt, T.G., and de Boer, P.A.J. (2009). Self-enhanced accumulation of FtsN at Division Sites and Roles for Other Proteins with a SPOR domain (DamX, DedD, and RlpA) in *Escherichia coli* cell constriction. *J. Bacteriol.* 191, 7383–7401. <https://doi.org/10.1128/JB.00811-09>.

Ghigo, J.M., and Beckwith, J. (2000). Cell division in *Escherichia coli*: role of FtsL domains in septal localization, function, and oligomerization. *J. Bacteriol.* 182, 116–129. <https://doi.org/10.1128/jb.182.1.116-129.2000>.

Ghigo, J.M., Weiss, D.S., Chen, J.C., Yarrow, J.C., and Beckwith, J. (1999). Localization of FtsL to the *Escherichia coli* septal ring. *Mol. Microbiol.* 31, 725–737. <https://doi.org/10.1046/j.1365-2958.1999.01213.x>.

Glas, M., van den Berg van Saparoea, H.B., McLaughlin, S.H., Roseboom, W., Liu, F., Koningstein, G.M., Fish, A., den Blaauwen, T., Heck, A.J.R., de Jong, L., et al. (2015). The Soluble Periplasmic Domains of *Escherichia coli* Cell Division Proteins FtsQ/FtsB/FtsL Form a Trimeric Complex with Submicromolar Affinity. *J. Biol. Chem.* 290, 21498–21509. <https://doi.org/10.1074/jbc.M115.654756>.

Goehring, N.W., Gonzalez, M.D., and Beckwith, J. (2006). Premature targeting of cell division proteins to midcell reveals hierarchies of protein interactions involved in divisome assembly. *Mol. Microbiol.* 61, 33–45. <https://doi.org/10.1111/j.1365-2958.2006.05206.x>.

Gonzalez, M.D., and Beckwith, J. (2009). Divisome under construction: distinct domains of the small membrane protein FtsB are necessary for interaction with multiple cell division proteins. *J. Bacteriol.* 191, 2815–2825. <https://doi.org/10.1128/JB.01597-08>.

Gonzalez, L., Woolfson, D.N., and Alber, T. (1996). Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* 3, 1011–1018. <https://doi.org/10.1038/nsb1296-1011>.

Gonzalez, M.D., Akbay, E.A., Boyd, D., and Beckwith, J. (2010). Multiple interaction domains in FtsL, a protein component of the widely conserved bacterial FtsLBQ cell division complex. *J. Bacteriol.* 192, 2757–2768. <https://doi.org/10.1128/JB.01609-09>.

Grigoryan, G., and DeGrado, W.F. (2011). Probing Designability via a Generalized Model of Helical Bundle Geometry. *J. Mol. Biol.* 405, 1079–1100. <https://doi.org/10.1016/j.jmb.2010.08.058>.

Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* 262, 1401–1407. .

Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmüller, H., and MacKerell, A.D. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* 14, 71–73.
<https://doi.org/10.1038/nmeth.4067>.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graph.* 14, 33–38, 27–28. .

Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29, 1859–1865.
<https://doi.org/10.1002/jcc.20945>.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589.
<https://doi.org/10.1038/s41586-021-03819-2>.

Katoh, K., and Standley, D.M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* 30, 772–780.
<https://doi.org/10.1093/molbev/mst010>.

Khadria, A.S., and Senes, A. (2013). The transmembrane domains of the bacterial cell division proteins FtsB and FtsL form a stable high-order oligomer. *Biochemistry* 52, 7542–7550. <https://doi.org/10.1021/bi4009837>.

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A.D., and Pastor, R.W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* 114, 7830–7843. <https://doi.org/10.1021/jp101759q>.

Kuhn, M., and Wickham, H. (2020). Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles.

Kureisaite-Ciziene, D., Varadajan, A., McLaughlin, S.H., Glas, M., Montón Silva, A., Luirink, R., Mueller, C., den Blaauwen, T., Grossmann, T.N., Luirink, J., et al. (2018). Structural Analysis of the Interaction between the Bacterial Cell Division Proteins FtsQ and FtsB. *MBio* 9. <https://doi.org/10.1128/mBio.01346-18>.

LaPointe, L.M., Taylor, K.C., Subramaniam, S., Khadria, A., Rayment, I., and Senes, A. (2013). Structural organization of FtsB, a transmembrane protein of the bacterial divisome. *Biochemistry* 52, 2574–2585. <https://doi.org/10.1021/bi400222r>.

Leclercq, S., Derouaux, A., Olatunji, S., Fraipont, C., Egan, A.J.F., Vollmer, W., Breukink, E., and Terrak, M. (2017). Interplay between Penicillin-binding proteins and SEDS proteins promotes bacterial cell wall synthesis. *Sci. Rep.* 7, 43306. <https://doi.org/10.1038/srep43306>.

Lesne, E., Krammer, E.-M., Dupre, E., Locht, C., Lensink, M.F., Antoine, R., and Jacob-Dubuisson, F. (2016). Balance between Coiled-Coil Stability and Dynamics Regulates Activity of BvgS Sensor Kinase in *Bordetella*. *MBio* 7, e02089. <https://doi.org/10.1128/mBio.02089-15>.

Li, Y., Gong, H., Zhan, R., Ouyang, S., Park, K.-T., Lutkenhaus, J., and Du, S. (2021). Genetic analysis of the septal peptidoglycan synthase FtsWI complex supports a conserved activation mechanism for SEDS-bPBP complexes. *PLoS Genet.* 17, e1009366. <https://doi.org/10.1371/journal.pgen.1009366>.

Liu, B., Persons, L., Lee, L., and de Boer, P.A.J. (2015). Roles for both FtsA and the FtsBLQ subcomplex in FtsN-stimulated cell constriction in *Escherichia coli*. *Mol. Microbiol.* 95, 945–970. <https://doi.org/10.1111/mmi.12906>.

Marmont, L.S., and Bernhardt, T.G. (2020). A conserved subcomplex within the bacterial cytokinetic ring activates cell wall synthesis by the FtsW-FtsI synthase. *Proc. Natl. Acad. Sci.* 117, 23879–23885. <https://doi.org/10.1073/pnas.2004598117>.

Masson, S., Kern, T., Le Gouët, A., Giustini, C., Simorre, J.-P., Callow, P., Vernet, T., Gabel, F., and Zapun, A. (2009). Central domain of DivIB caps the C-terminal regions of the FtsL/DivIC coiled-coil rod. *J. Biol. Chem.* 284, 27687–27700. <https://doi.org/10.1074/jbc.M109.019471>.

McCausland, J.W., Yang, X., Squyres, G.R., Lyu, Z., Bruce, K.E., Lamanna, M.M., Söderström, B., Garner, E.C., Winkler, M.E., Xiao, J., et al. (2021). Treadmilling FtsZ polymers drive the directional movement of sPG-synthesis enzymes via a Brownian ratchet mechanism. *Nat. Commun.* 12, 609. <https://doi.org/10.1038/s41467-020-20873-y>.

Meeske, A.J., Riley, E.P., Robins, W.P., Uehara, T., Mekalanos, J.J., Kahne, D., Walker, S., Kruse, A.C., Bernhardt, T.G., and Rudner, D.Z. (2016). SEDS proteins are a widespread family of bacterial cell wall polymerases. *Nature* 537, 634–638. <https://doi.org/10.1038/nature19331>.

Mercer, K.L.N., and Weiss, D.S. (2002). The *Escherichia coli* cell division protein FtsW is required to recruit its cognate transpeptidase, FtsI (PBP3), to the division site. *J. Bacteriol.* 184, 904–912. <https://doi.org/10.1128/jb.184.4.904-912.2002>.

Mirdita, M., Ovchinnikov, S., and Steinegger, M. (2021). ColabFold - Making protein folding accessible to all (Bioinformatics).

Mohammadi, T., van Dam, V., Sijbrandi, R., Vernet, T., Zapun, A., Bouhss, A., Diepeveen-de Bruin, M., Nguyen-Distèche, M., de Kruijff, B., and Breukink, E. (2011). Identification of FtsW as a transporter of lipid-linked cell wall precursors across the membrane. *EMBO J.* 30, 1425–1432. <https://doi.org/10.1038/emboj.2011.61>.

Mohammadi, T., Sijbrandi, R., Lutters, M., Verheul, J., Martin, N.I., den Blaauwen, T., de Kruijff, B., and Breukink, E. (2014). Specificity of the transport of lipid II by FtsW in *Escherichia coli*. *J. Biol. Chem.* 289, 14707–14718.
<https://doi.org/10.1074/jbc.M114.557371>.

Nakagawà, J., Tamaki, S., and Matsuhashi, M. (1979). Purified Penicillin Binding Proteins 1Bs from *Escherichia coli* Membrane Showing Activities of Both Peptidoglycan Polymerase and Peptidoglycan Crosslinking Enzyme. *Agric. Biol. Chem.* 43, 1379–1380. <https://doi.org/10.1080/00021369.1979.10863634>.

O’Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res.* 44, D733–D745. <https://doi.org/10.1093/nar/gkv1189>.

Pagès, H., Aboyoun, P., Gentleman, R., and DebRoy, S. (2019). Biostrings: Efficient manipulation of biological strings.

Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., and Jacobs-Wagner, C. (2016). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Mol. Microbiol.* 99, 767–777.
<https://doi.org/10.1111/mmi.13264>.

Park, K.-T., Du, S., and Lutkenhaus, J. (2020). Essential Role for FtsL in Activation of Septal Peptidoglycan Synthesis. *MBio* 11. <https://doi.org/10.1128/mBio.03012-20>.

Park, K.-T., Pichoff, S., Du, S., and Lutkenhaus, J. (2021). FtsA acts through FtsW to promote cell wall synthesis during cell division in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* 118, e2107210118. <https://doi.org/10.1073/pnas.2107210118>.

Pedersen, T.L. (2019). tidygraph: A Tidy API for Graph Manipulation.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J. Comput. Chem.* 26, 1781–1802. <https://doi.org/10.1002/jcc.20289>.

Pichoff, S., Du, S., and Lutkenhaus, J. (2015). The bypass of ZipA by overexpression of FtsN requires a previously unknown conserved FtsN motif essential for FtsA-FtsN interaction supporting a model in which FtsA monomers recruit late cell division proteins to the Z ring. *Mol. Microbiol.* 95, 971–987. <https://doi.org/10.1111/mmi.12907>.

R Core Team (2019). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Schmidt, N.W., Grigoryan, G., and DeGrado, W.F. (2017). The accommodation index measures the perturbation associated with insertions and deletions in coiled-coils:

Application to understand signaling in histidine kinases. *Protein Sci. Publ. Protein Soc.* 26, 414–435. <https://doi.org/10.1002/pro.3095>.

Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. *Nat. Biotechnol.* 35, 1026–1028. <https://doi.org/10.1038/nbt.3988>.

Studier, F.W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41, 207–234..

Taguchi, A., Welsh, M.A., Marmont, L.S., Lee, W., Sjodt, M., Kruse, A.C., Kahne, D., Bernhardt, T.G., and Walker, S. (2019). FtsW is a peptidoglycan polymerase that is functional only in complex with its cognate penicillin-binding protein. *Nat. Microbiol.* 4, 587–594. <https://doi.org/10.1038/s41564-018-0345-x>.

Testa, O.D., Moutevelis, E., and Woolfson, D.N. (2009). CC+: a relational database of coiled-coil structures. *Nucleic Acids Res.* 37, D315-322.
<https://doi.org/10.1093/nar/gkn675>.

Tsang, M.-J., and Bernhardt, T.G. (2015). A role for the FtsQLB complex in cytokinetic ring activation revealed by an ftsL allele that accelerates division. *Mol. Microbiol.* 95, 925–944. <https://doi.org/10.1111/mmi.12905>.

Ursinus, A., van den Ent, F., Brechtel, S., de Pedro, M., Höltje, J.-V., Löwe, J., and Vollmer, W. (2004). Murein (peptidoglycan) binding property of the essential cell division protein FtsN from *Escherichia coli*. *J. Bacteriol.* 186, 6728–6737. <https://doi.org/10.1128/JB.186.20.6728-6737.2004>.

Villanelo, F., Ordenes, A., Brunet, J., Lagos, R., and Monasterio, O. (2011). A model for the *Escherichia coli* FtsB/FtsL/FtsQ cell division complex. *BMC Struct. Biol.* 11, 28. <https://doi.org/10.1186/1472-6807-11-28>.

Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R.S. (1999). The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci. Publ. Protein Soc.* 8, 2312–2329. <https://doi.org/10.1110/ps.8.11.2312>.

Wickham, H. (2017). tidyverse: Easily Install and Load the “Tidyverse.”

Winter, D.J. (2017). rentrez: an R package for the NCBI eUtils API. *R J.* 9, 520–526. .

Woolfson, D.N. (2005). The design of coiled-coil structures and assemblies. *Adv. Protein Chem.* 70, 79–112. [https://doi.org/10.1016/S0065-3233\(05\)70004-8](https://doi.org/10.1016/S0065-3233(05)70004-8).

Yang, J.-C., Van Den Ent, F., Neuhaus, D., Brevier, J., and Löwe, J. (2004). Solution structure and domain architecture of the divisome protein FtsN. *Mol. Microbiol.* 52, 651–660. <https://doi.org/10.1111/j.1365-2958.2004.03991.x>.

Yang, X., McQuillen, R., Lyu, Z., Phillips-Mason, P., De La Cruz, A., McCausland, J.W., Liang, H., DeMeester, K.E., Santiago, C.C., Grimes, C.L., et al. (2021). A two-track model for the spatiotemporal coordination of bacterial septal cell wall synthesis revealed by single-molecule imaging of FtsW. *Nat. Microbiol.* 6, 584–593. <https://doi.org/10.1038/s41564-020-00853-0>.

Zeileis, A., and Grothendieck, G. (2005). zoo: S3 Infrastructure for Regular and Irregular Time Series. *J. Stat. Softw.* 14, 1–27. <https://doi.org/10.18637/jss.v014.i06>.

Chapter 5: Conclusions and Future Directions

In the works presented here, a combination of computational and experimental techniques was used for the biophysical study of transmembrane helices. As a graduate student I was involved in two main projects. One of these projects was focused on thermodynamic studies of the structural GAS_{right} motif, while the other was related to the bacterial divisome.

In Chapter 2, we presented an experimental thermodynamic study of a group of predicted GAS_{right} dimers. Using FRET we successfully determined the free energy of association of seven GAS_{right} helices plus the two controls GpA and its monomerizing mutant G83I in the detergent n-decyl-β-D-maltopyranoside (DM). In our results we observed a wide range of stabilities for the group of helices under study. In addition we observed a striking correlation between previously obtained TOXCAT homodimerization propensities and association free energies for six of the helices (excluding the outlier CP8B1). We used the strongly linear relationship between TOXCAT and fraction of dimer at a specific concentration to predict the association free energy for the other GAS_{right} dimers that were previously studied with TOXCAT and the computational approach CATM. We saw a statistically significant correlation between the predicted free energies and the CATM energy scores. These findings suggest TOXCAT is a thermodynamically driven process, and that Ca-H hydrogen bonds and vdW interactions are key contributors to the association of the GAS_{right} dimers.

It is important to consider the fact that the Ca-H hydrogen bond present in this motif occurs between the backbones of the helices, making it challenging to use a mutational strategy to understand their contribution to the association energetics. The use of experimental and computational techniques to better understand the contribution of this interaction is going to be essential. The thermodynamic experiments presented in Chapter 2 can be later combined with computational approaches to figure out the relative contribution that different interactions have to the association of $\text{GAS}_{\text{right}}$ helices, including the Ca-H hydrogen bond. The used of MD simulations to look at how different $\text{GAS}_{\text{right}}$ dimers behave could give us insight into the atomic and molecular interactions that are present. We could use WEUSMD with integration of mean force to dissect the potential of mean force (PMF) into different energy contributions so that we can learn about the possible relative contributions that different interactions have to the process. As mentioned in Chapter 2, understanding the interplay between the physical interactions involved in the association of $\text{GAS}_{\text{right}}$ helices could support the prediction of structure and stability (which are extremely challenging to experimentally measure for these systems), the identification of potential conformational changes (which could lead to a better understanding of their functional role and mechanism), and the interpretation of the effect of mutations in these systems.

In chapters 3 and 4 we look at the FtsL-FtsB subcomplex of the divisome. The experimental structure of their complex has not been solved, but we have proposed a computational model of it that is backed up by *in vivo*, *in vitro* and computational studies. My main contribution to these studies has been performing all-atom MD

simulations to look at the dynamics and stability of the predicted models. The first predicted model which we referred to as the I-model was an L₂:B₂ tetramer forming a four helix bundle spanning the TM domain and the periplasmic coiled-coil region of the complex (Chapter 3). In the MD simulations we observed that the TM region stayed stable throughout the simulations, while water infiltrated the periplasmic coiled coil interface due to a cluster of conserved and buried polar residues. This observation suggested the possibility of a L₂:B₂ tetramer where the TM forms a four helix bundle, while the periplasmic region divides into two two-stranded LB coiled coils. In the new model (Y-model), the polar residues that were buried in the I-model were more exposed to the solvent (Chapter 4). In the Y-model MD simulations the water infiltration that was observed in the I-model simulations was not present. The MD simulations showed the periplasmic coiled coil region of the Y-model to be more stable throughout the simulation when compared to the I-model. The MD simulations help us to get a view of the possible conformational changes that the FtsLB subcomplex could take, which was critical for discriminating between the two models. The MD simulations were also critical to assess the dynamic stability of the TM region where we observed the two FtsB TMDs to get closer to each other throughout the simulations (Chapter 3).

It is important to keep in mind that the divisome is composed of many proteins and in our MD studies the FtsLB subcomplex is isolated which could affect the results. A future direction for this would be to add other components of the divisome to the complex. For example we could see how the addition of FtsQ, a divisome protein known to interact with the C terminus of FtsB, affects the dynamics of the complex especially the

periplasmic region. We could also look at how FtsN presence influences the interaction between the FtsLB subcomplex and FtsI. Because of the size of the system we need to consider the computational expense and might need to move to a less expensive MD simulation approach (for example by using coarse grained simulations instead of all-atom). The MD simulation results need to be interpreted with care, but they still can guide us into a better understanding of how FtsLB achieves its function and help interpret the results of mutagenesis experiments.