Using protein design to explore the role of sidechain packing in membrane protein stability

Written Proposal for Preliminary Examination

Submitted by: Gilbert Loiseau
Advisor: Alessandro Senes

Oral Examination Date:
June 30, 2020
1:00 PM

University of Wisconsin-Madison
Integrated Program in Biochemistry

Abstract

Membrane proteins subunits associate to fold into biologically essential structures as a result of forces including hydrogen bonding, weak polar interactions, and van der Waals packing. After a membrane protein is inserted into the membrane, a combination of these forces stabilizes the folded state. Previous research has investigated the impact of hydrogen bonding and weak polar interactions on membrane protein association, however, the contribution of van der Waals packing to association in the folded state remains poorly understood. Van der Waals packing is a complex set of diverse interactions comprised of lipid-lipid packing, lipid-protein packing, and protein-protein packing, which collectively stabilize both the unfolded and folded states. Each of these interactions must be taken into account to understand the contribution of packing to membrane protein association in the folded state, but prior research has found difficulty isolating these interactions from each other and from other stabilizing forces. Literature suggests that protein-protein packing, or sidechain packing, plays a significant role in membrane protein stability, demonstrating that optimized sidechain packing alone can stabilize the folded state. However, the relative contribution of this force to the stable structure of a membrane protein has not yet been determined. My work aims to control for other stabilizing forces to study the extent at which sidechain packing can stabilize membrane protein structure. To investigate this question, I will use helical dimers which are a simple and tractable model system for the association of membrane protein subunits during folding. Using established computational design techniques, I will engineer dimeric sequences with an array of sidechain packing energies to determine how much packing is necessary for stability of the folded state. Using a high-throughput *in vivo* assay, sort-seq, I will screen the dimerization propensity of each construct, to identify correlations between computational and experimental stability. Then, I will make point mutants on my designed structures and measure changes in thermodynamic stability using *in vitro* FRET. This will allow me to confirm the correlation seen between sidechain packing and dimerization propensity in sort-seq and estimate the free energies of my designed constructs. This work will provide a better understanding of how sidechain packing can facilitate and stabilize membrane protein association and addresses a crucial knowledge gap in our understanding of membrane protein biology.

Specific Aims

Proper membrane protein folding is necessary for essential biological functions such as cell signaling and gene regulation. Misfolding of membrane proteins often leads to disease phenotypes including growth defects and cancer. A variety of forces are necessary for membrane proteins to fold properly including hydrogen bonding, weak polar interactions, and van der Waals packing. In order to fully understand how membrane proteins fold for proper function, it is necessary to elucidate the energetic contribution of each of these interactions to the folded state. This contribution, or the ability to drive the transition from the unfolded to the folded state, has been characterized and quantified for hydrogen bonding and weak polar interactions, but research is lacking on the contribution of van der Waals packing. Van der Waals packing in the membrane is a complex force composed of lipid-lipid packing, lipid-protein packing, and protein-protein packing. Of these three, protein-protein packing, or sidechain packing, is of particular interest because we can manipulate protein sequence, allowing us to modulate sidechain packing. Previous research has demonstrated that disruption of sidechain packing within the core of bacteriorhodopsin destabilizes the protein structure while membrane protein design has shown that optimized sidechain packing stabilizes a redesigned phospholamban structure. However, outside of these individual systems, the contribution of sidechain packing to the folded state of membrane proteins has not yet been determined. With my research, I propose to characterize and quantify the extent at which sidechain packing contributes to membrane protein association.

In order to understand the general rules that govern the influence of sidechain packing on membrane protein association, I must explore a wide array of sequences and structures that rely on sidechain packing. To do so, I will use a high-throughput approach combining large scale computational sequence design and *in vivo* characterization of transmembrane (TM) helix association. Using a set of structural geometries and simple energetic terms, my lab previously demonstrated that dimeric structures can be predicted from sequence. These TM helices were then expressed in *Escherichia coli* and assessed experimentally using a reporter gene assay, producing a correlation between our predicted energy scores and reporter gene output to assess the accuracy of our predictions. With my research, I will use the same algorithms and methods used in prediction to design homodimeric sequences that associate based on sidechain packing energy. I will design hundreds of sequences using only nonpolar amino acids, minimizing the impact of non van der Waals forces including hydrogen bonding and weak polar interactions. These sequences will be designed along the interface of a fixed backbone geometry, where having increased sequence similarity should standardize the energetics of the unfolded state of my designed sequences. Then, by making single point mutants at positions on these well-packed interfaces to amino acids that decrease sidechain packing, I will have generated thousands of homodimers with a range of sidechain packing energies. These sequences will be characterized for their association using "**sort-seq**", a high-throughput method that will measure the dimerization propensity of my designed sequences. In this system, dimerization of proteins results in a fluorescence output which is used to **sort** cells into bins based on fluorescence as a readout of association strength. These cells are then sent for next-generation **sequencing** to identify the dimerization propensity of each sequence. After determining the correlation of sidechain packing energy to dimerization propensity, I will take a subset of these sequences and quantify their association stability using *in vitro* FRET. By ensuring this subset is composed of well-packed sequences and their corresponding mutants that decrease sidechain packing, I can quantify differences in sidechain packing contribution between designed and mutant sequences. Overall, my research will characterize and quantify the extent at which sidechain packing is a driving force for membrane protein association, improving our fundamental understanding of membrane protein association and folding.

Aim 1: High-throughput screening of the effects of sidechain packing on dimerization
I will analyze the protein databank for helices in close contact within membrane proteins, searching for common geometries between helices. These geometries will be used to design sequences with a range of sidechain packing which will then be tested using *in vivo* sort-seq. This aim will determine the extent to which sidechain packing influences dimerization.

Aim 2: Thermodynamic quantification of sidechain packing on dimerization
I will use a subset of designed sequences and point mutants to quantify the stability of sequences with a range of sidechain packing energies, determining the thermodynamic free energy of a variety of dimers stabilized by sidechain packing. This aim will quantify the extent to which sidechain packing influences dimerization.

Significance

Membrane proteins comprise 25-30% of the proteins found within protein-coding genes of various organisms (Fagerberg et al., 2010). Proper membrane protein folding is critical for essential biological functions, including cell signaling, ion balance, and gene regulation. Misfolding of membrane proteins has been found to be involved in several human diseases such as Parkinson's, cystic fibrosis, and cancer (Sanders and Myers, 2004; Gregersen et al., 2006). To understand how protein misfolding plays a role in disease states and progression, it is necessary to investigate how these proteins fold. However, studying membrane protein folding is inherently a difficult challenge because of their hydrophobic nature. Membrane proteins are difficult to express in yields high enough for biophysical experiments, and purification and solubilization of these proteins often lead to aggregation or unfolding (Carpenter et al., 2008). To combat these challenges, much of the research studying membrane protein folding is focused on understanding the biophysical forces that govern the folding process. Investigation of the biophysical forces that govern protein folding will allow us to determine why specific mutations result in diseases caused by misfolding. In addition, this knowledge can be applied to design new therapeutics that specifically target proteins in these misfolded states. Understanding the forces that enable membrane proteins to fold will contribute to our knowledge of how these fundamental forces are involved in human health and disease.

Proper membrane protein folding is regulated by a distribution of stabilizing hydrogen bonds, weak polar interactions, and van der Waals forces between the unfolded and folded states. Previous research has measured the contributions of both hydrogen bonding and weak polar interactions in the membrane and determined that these forces can drive membrane protein folding (Zhou et al., 2001; Yano et al., 2002; Johnson et al., 2007), but research is lacking on the contribution of van der Waals packing. This force is particularly important due to the nature of van der Waals interactions: Even if hydrogen bonding or polar interactions play a significant stabilizing role, because van der Waals occurs between any nonbonded atoms in close contact, it is a necessary force that is always present within the folded state. This means that van der Waals packing is essential for folding, but the extent at which packing can be a driving force for membrane protein folding is unclear. Without understanding the extent at which van der Waals packing contributes to folding, we cannot complete our understanding of how these other forces contribute to membrane protein association and folding.

The contribution of van der Waals packing to membrane protein folding can be broken down into three distinct interactions: lipid-lipid packing, lipid-protein packing, and protein-protein packing. Protein-protein (or sidechain) packing, is a technically feasible starting point because of the ability to manipulate sequences and determine changes in stability due to mutation. Previous research has demonstrated that disruption of packing within the core of bacteriorhodopsin destabilizes protein structure (Faham et al., 2004; Joh et al., 2009). In addition, a recent study using membrane protein design has shown that optimized sidechain packing can stabilize the folded state of phospholamban (Mravic et al., 2019). Although it is known that sidechain packing plays a role in stabilizing membrane protein structure in these individual systems, the energetic contribution of sidechain packing to the folded state of membrane proteins more generally has not yet been determined. I propose to characterize and quantify the extent at which sidechain packing is a driving force for membrane protein association. To do so, I will investigate the role that sidechain packing plays in the association of homodimers, a simple and tractable model for studying membrane protein folding (Popot and Engelman, 1990). Using large scale computational design on common dimeric backbone geometries found within the PDB, I will study the extent at which sidechain packing can be considered a driving force for the general population of membrane protein structures. Overall, my research will give insight into the extent at which fundamental sidechain packing impacts membrane protein association and folding, something that has not yet been done despite the importance of van der Waals packing to membrane protein folding.

The contribution of van der Waals packing to stabilizing membrane protein structure remains a significant gap in our understanding of membrane protein folding (Hong et al., 2014). Determining how sidechain packing impacts the association of membrane protein subunits will give us a better understanding of how membrane proteins assemble to fold stabilized structures. By understanding how these forces contribute to stability, we will

be closer to obtaining a holistic view of how all forces involved combine to stabilize membrane protein structure. This proposed work will increase our knowledge of the fundamental rules of membrane protein folding, and add to our understanding of complex membrane protein mechanisms, such as oligomerization and conformational change necessary for essential biological processes including signal transduction and ion transport (Sanders and Myers, 2004; Gregersen et al., 2006). Eventually, this knowledge can be used to design new functional membrane protein structures, advancing the field of synthetic biology.

Innovation

Understanding the fundamental impact of sidechain packing on membrane protein association and folding has relied heavily on monitoring stability changes in individual protein systems where other forces likely play a stabilizing role (Faham et al., 2004; Joh et al., 2009; Zhang et al., 2011). Controlling for changes in stabilizing forces and packing interactions in the membrane is a difficult but necessary task to understand the influence of sidechain packing on membrane protein association. Using a technically innovative high-throughput approach will allow me to specifically assess sidechain packing in a large variety of unique structures and sequences. Overall, my research will reveal general rules for how sidechain packing contributes to association, determining if it is just a necessary force or if it is a potential driving force in membrane protein association and folding.

Approach

Aim 1: High-throughput determination of how sidechain packing affects dimerization

Rationale:

The goal of this aim is to assess the contribution of nonpolar sidechain interactions to TM helix dimerization.

Experimental Design:

Aim 1a: Computational design of homodimeric sequences

I aim to explore the effect of sidechain packing on homodimerization. Using common geometries from the PDB as structural templates for computational design will allow me to obtain a general understanding of how sidechain packing influences membrane protein association and folding. Each of these geometries will be standardized with a poly-leucine backbone to control for expression and insertion of our sequences (Zhou et al., 2001; Anderson et al., 2017). To specifically vary the sidechain packing contributing to association, only positions at the dimer interface will be mutated. With these geometries as a structural template, I will use well-known computational algorithms to filter and search sequence space for amino acid combinations that pack at the dimerization interface (Desmet et al., 1992; Koehl and Delarue, 1994; Hansmann and Okamoto, 1999). Using a minimalistic set of energy functions that measure van der Waals packing, hydrogen bonding, and membrane implicit solvation (IMM1) (MacKerell et al., 1998; Lazaridis, 2003; Krivov et al., 2009), I will measure the stability of well-packed sequences at each geometry that rely on stabilization by packing. Overall, computational design will result in hundreds of dimers with differences in sidechain packing energies (Fig. 1).

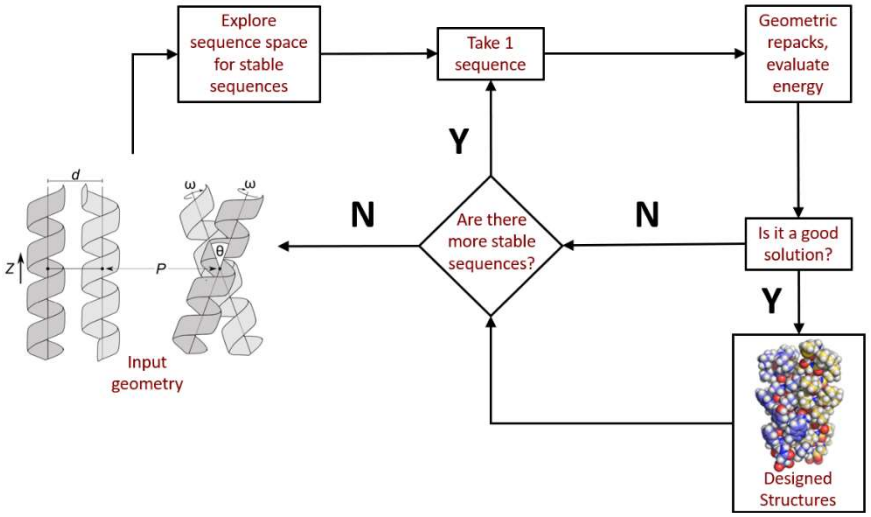

Fig. 1: Flowchart of design algorithm. A common natural geometry from the PDB is input as a structural template. Positions found at the interface of this geometry are identified and common computational methods search sequence space for well-packed homodimers. One of those sequences undergoes uses local geometric repacks to search for a geometry with the most stable energy. This energy is compared to the monomeric state and the design is added to a pool of successfully design sequences if it is more stable. The algorithm then repeats this process until there are no more sequences with good packing at the interface. This process can be repeated will other common geometries from the PDB, allowing me to design hundreds of sequences with an array of expected dimerization based on sidechain packing energies.

Aim 1b: *In vivo* characterization of designed sequences

The designed population of sequences will be evaluated for their experimental stability using sort-seq, a high-throughput method which combines fluorescence activated cell sorting (FACS) with next-generation sequencing (NGS) to evaluate dimerization propensity of TM domains (Fig. 2). An oligo pool library consisting of my designed TM sequences will be cloned into TOXGREEN plasmids (Armstrong and Senes, 2016). These TOXGREEN plasmids use the reporter gene sfGFP to quantify dimerization (as described in methods). Based on the expression of sfGFP, cells are sorted into different bins with different fluorescence thresholds. These plasmids can then be purified out of the cells in each bin and enumerated via NGS. Based on the counts present within each bin, the fluorescence profile, and thus dimerization propensity, will be reconstructed for each dimer. Saturating point mutations of each sequence to various hydrophobic amino acids will also be characterized on sort-seq, and a sensitivity to each mutation will be generated based on change in dimerization propensity. This mutagenesis and sensitivity profile will be used to confirm that the designed sequences associate via the predicted structure (Anderson (Thesis), 2019). Using this technique, I will be able to determine if there is a correlation between strength of sidechain packing and membrane protein association in hundreds of designed sequences at a time.

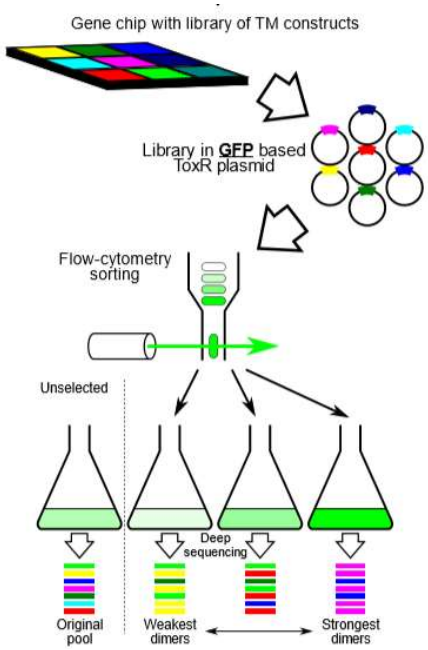

Fig. 2: TOXGREEN sort-seq schematic. Individual TM domains are synthesized on a chip using oligo pool technology and cloned into TOXGREEN plasmids. In the GFP-based assay, cells will fluoresce based on GFP expression correlating to the dimerization propensity of the expressed TM domain chimera. This GFP expression will be sorted by fluorescence-activated cell sorting (FACS) followed by next-generation sequencing of the bins.

Expected Results/Interpretation:

Previous research has demonstrated that understanding the structures of natural proteins can guide protein design (Korendovych et al., 2010; Joh et al., 2014; Huang et al., 2014; Lu et al., 2018). In particular, if a structure is commonly found in nature but not dependent on a specific sequence, then designing sequences for this structure is likely to be more successful (Zhou and Grigoryan, 2015). In my case, finding common geometries for TM helices in close contact is likely to yield successfully designed sequences for dimers. Much of the analysis for these geometries has been done (detailed in Preliminary Data below), and I am currently in the process of designing well-packed homodimers for these geometries. I expect that designed sidechain packing energy will correlate well to dimerization propensity found in sort-seq, resulting in a trend where sequences with weak sidechain packing energy results in an unstable dimer and strong packing energy results in a stable dimer. Overall, these results will give insight into the extent at which sidechain packing contributes to stable helix-helix association in membrane protein folding, and whether or not it can be considered a driving force for membrane protein association.

Potential Complications:

Protein design papers typically result in the structural determination of each of their designs, confirming the expected orientation of the design. However, my research aims to design hundreds of structures, making it difficult to confirm the structure of each design. To address this lack of structural data, I will use sort-seq to test point mutations at each position on my homodimers. If changes in dimerization are particularly sensitive at the interface, then I will have supporting information for the structures of my designs. In addition, if I decide that confirming the structure of some of my constructs is important to determine the accuracy of my designs, I will have a large population of proteins that I could test for expression, purification, and crystallization ability, which may make it possible to solve the structure of some of my designs. Another caveat is that I am assuming that if a geometry is more prevalent in the PDB analysis, then it may be more amenable to design (Zhou and Grigoryan, 2015). To determine if this is the case, I will also design sequences for geometries that are not as prevalent within the analysis. If these designed sequences are overall less energetically stable than our other designs and have lower dimerization propensity, then it will support our assumption that common geometries are more favorable for design. Finally, it is possible that our results show little correlation between our designed energies and dimerization propensity. We expect that small changes in sequence should have little effect on the monomer

energy of our TMs and that these changes would average out with a large number of sequences for analysis, but this would suggest that this is not always the case. I may have to use more sensitive methods to confirm differences between the energy of the monomeric and dimeric states of my designs such as molecular dynamics simulations, giving me a better understanding of why particular sequences do not follow a trend. These sequences can then be removed from the analysis, allowing me to determine if a trend between sidechain packing and dimerization exists with sequences where the changes in the monomer energy have little effect on dimerization.

Preliminary Results

Analysis of nonredundant membrane proteins determines common geometries of helix-helix association

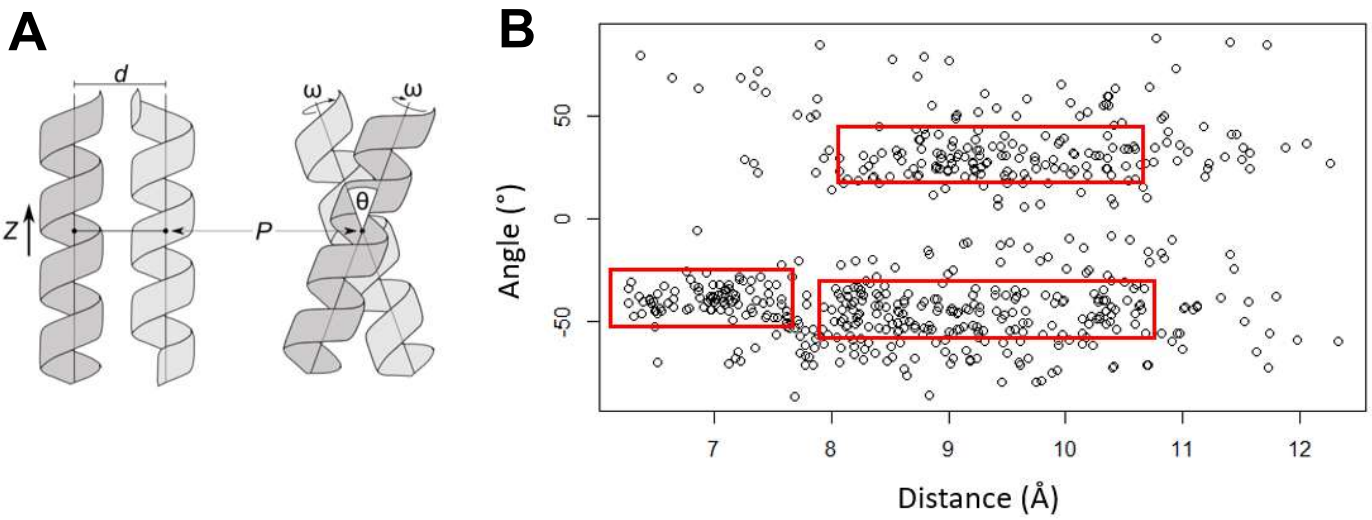

Fig. 3: Geometric analysis of membrane proteins in the PDB.
A) The geometric landscape of homodimerization is defined by four parameters: interhelical distance (d), crossing angle (θ), rotation of the helix around its axis (ω), and the vertical position (Z) relative to the crossing point between the helices (P). **B)** Each circle in this scatterplot represents one helical pair of the 835 pairs extracted from nonredundant membrane proteins found within the PDB. These pairs were plotted by their angle and distance, allowing me to determine areas with common geometries that may be amenable for sequence design. Boxed in red are three areas with particularly high density of common geometries. The leftmost box is a common geometry for sequences stabilized by weak hydrogen bonding, so we are likely not to use that area for sequence design. The other two boxes are currently being implemented as geometries for sequence design.

Before being able to design well-packed sequences, I needed to determine geometric orientations to place my TM backbones. I analyzed membrane proteins found within the PDB for helices in close contact within the lipid bilayer, aiming to find common natural geometries that will be used for protein design. To ensure that common geometries are not found as a result of analyzing multiples of the same protein, I analyzed the geometries of helical pairs from a nonredundant database of membrane proteins. Since I will be designing well-packed homodimeric sequences, I extracted the geometric parameters that define the homodimeric landscape (Fig. 3A). I then plotted these helical pairs as points on scatterplots with each of these parameters against one another, searching for high-density areas that represent geometric parameters that are common within membrane proteins. Figure 3B shows the scatterplot of distance vs angle, and areas of high density are boxed in red. In order to determine if any of these areas had common sequence motifs that led to association to potentially inform my computational design, I extracted the sequences from each pair and constructed sequence logos (data not shown). The box on the left was shown to contain the GAS_{right} motif, which is known to facilitate by the formation of weak hydrogen bonds (Anderson et al., 2017). The other two boxes did not have any common sequences motifs according to my sequence logos. To ensure that I do not have other stabilizing forces such as weak hydrogen bonds, the box containing GAS_{right} sequences was eliminated from consideration for protein design. My protein design algorithm will be run on the geometric space encompassed by the other two regions with a high density of common angles and distances. By using these common natural geometries to design well-packed sequences, I will be able to analyze the contribution of sidechain packing on membrane protein association in a general population of associating helices. My work will address the question of how fundamental sidechain packing impacts a general population of membrane protein structures, giving insight into the fundamental impact of packing on membrane protein association and folding.

Aim 2: Thermodynamic quantification of sidechain packing on dimerization

Rationale:

Although my previous aim will deduce the extent at which sidechain packing influences dimerization, it will not measure the biophysical strength of sidechain packing. Aim 2 will quantify the free energy of association of a subset of my designed sequences, allowing me to confirm the correlation seen in aim 1 with thermodynamic data. If there is good correlation between my data in both aims, I will be able to translate dimerization propensities to free energies of association as determined by the correlation in aim 2.

Experimental Design:

I will measure dimerization free energies of some of my designed sequences using a previously developed *in vitro* FRET strategy. I will engineer plasmids containing chimeric TM constructs of my designed helix and staphylococcal nuclease, a 16.9 kDa protein that saturation mutagenesis has shown to have no effect on dimerization of GpA *in vitro* (Lemmon et al., 1992; Fleming et al., 1997). These chimeric proteins will be expressed in *E. coli*, extracted from the membrane, and then purified in detergent micelles before resuspending them in phophatidylcholine lipid vesicles for *in vitro* study (Khadria & Senes, 2013). A position on SN will be mutated to cysteine, allowing for fluorescent labeling via maleimide crosslinking. The protein will be labeled by either Cy3 (donor) or Cy5 (acceptor), and the labeling efficiency will be calculated by comparing the amount of absorbance between total protein and label in a spectrophotometer. The protein concentrations for binding are then calculated for different mole fractions, or total protein concentration to lipid concentration ratios. The donor-labeled protein is mixed with equal amounts of acceptor-labeled protein and allowed to equilibrate before being measured on a fluorimeter. I will calculate the FRET efficiency by comparing donor fluorescence to fluorescence of donor and acceptor at different mole fractions. From this data, I will be able to calculate the free energy of association of my designed sequences. I will also study point mutations at the dimer interface of some of these sequences, destabilizing the dimer only by changes in sidechain packing. This will allow me to measure the difference in free energy of association between my wild type designs and point mutants to determine how sensitive dimerization is to small changes in sidechain packing.

Expectations/Interpretation:

I expect that FRET will allow me to determine the strength of association for a subset of my designed homodimers and reflect the correlation of packing with dimerization that I see in aim 1, confirming that changes in sidechain packing contributes to the free energy of association. I can then use the correlation found in aim 2 to translate the original dimerization propensities to free energy of association, resulting in estimated free energies for each of my designed constructs. Overall, I expect that I will be able to determine free energies of association for my constructs and the correlation will be very similar to aim 1.

Potential Complications:

In order to determine the change in association strength between my sequences with well designed sidechain packing and their mutants, the TM helices must associate within the detection range of mole fractions measurable by our FRET scheme. Current research in my lab is investigating the change in stability of membrane protein dimers by weak hydrogen bonding, where the measurable range of detecting dimerization is from 10^{-3} to 10^{-6} (protein/lipid mole fraction ratio). If my designed sequences do not dimerize within this range, it may be necessary to use alternate techniques to determine the free energy of association, such as sedimentation equilibrium analytical ultracentrifugation or single molecule photobleaching (Fleming et al., 1997; Chadda et al., 2016). In addition, biophysical measurements of association in lipid vesicles is a complicated process due to the ability for helices to be reconstituted in a flipped orientation in the bilayer. This would prevent dimerization from occurring if there is one flipped helix and one properly oriented helix, decreasing the total expected fluorescence and making interpretation of our results complex. A previous paper by Chadda et al. in 2016 accounts for this by assuming that flipping is random and dividing the protein concentration in half, resulting in good correlation between calculated and observed fluorescence. I will apply this method when calculating the free energy for my designed homodimers, allowing me to account for flipped of helices in lipid vesicles.

Methods

TOXCAT/TOXGREEN

TOXCAT is an *in vivo* reporter-based assay that measures helix-helix association (Fig. M1). The TM of interest is C-terminally fused to maltose binding protein (MBP) and N-terminally fused to ToxR, a dimerization dependent transcription factor from *vibrio cholera*. In TOXCAT, when the chimera dimerizes in the membrane, the ToxR domain dimerizes in the cytoplasm, upregulating the expression of the *ctx* promoter, engineered to express chloramphenicol acetyltransferase (CAT). The cell lysate is then assessed for CAT activity using acetyl-Coenzyme A (CoA) and chloramphenicol (CAM). Ellman's reagent reacts with the free CoA and ionizes to produce a yellow color that is detectable in a spectrophotometer at wavelength 412 nm. The propensity for dimerization is correlated to this detectable change in color.

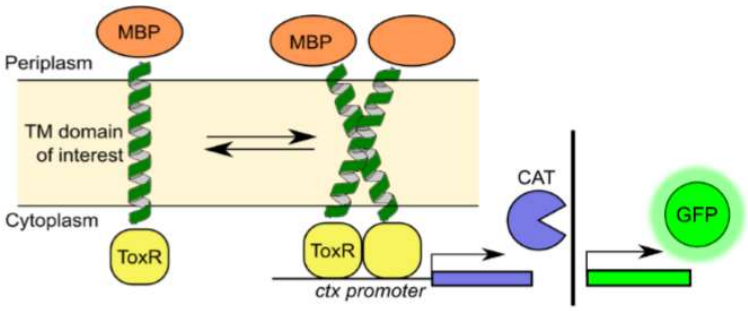

Fig. M1: TOXCAT/TOXGREEN.
TOXCAT is an *in vivo* assay that reports the stability of TM helices through the expression of chloramphenicol acetyltranferase (CAT). The TM domain is fused to maltose binding protein (MBP) to ensure proper insertion, and ToxR, a dimerization dependent transcription factor. When the helices dimerize, ToxR binds to the *ctx* promoter to induce expression of CAT. In TOXGREEN, the reporter gene is changed to sfGFP, allowing for faster and more high-throughput quantification of helix-helix association.

My lab developed a high-throughput variant of TOXCAT known as TOXGREEN, quantifying dimerization by protein expression of superfolder green fluorescent protein (sfGFP) rather than intensity of Ellman's reagent. TOXGREEN replaces the gene for CAT with the gene for sfGFP (Fig. M1). TOXGREEN was shown to replicate results of TOXCAT, demonstrating that it has comparable response, dynamic range, and sensitivity. However, because TOXGREEN relies on the fluorescence of GFP, the cells do not need to be lysed and reporter gene expression can be quantified with minimal sample processing. TOXGREEN has been successfully quantified using 96-well plates and fluorescence activated cell sorting, making it possible to rapidly assess the dimerization of many constructs at once. This technique has been implemented into the sort-seq procedure that will be used to measure the dimerization propensity of hundreds of my designed constructs at a time.

References

1. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlén, M. & Berglund, L. (2010). Prediction of the human membrane proteome. Proteomics 10, 1141–1149.
2. Sanders, C.R., and Myers, J.K. (2004). Disease-Related Misassembly of Membrane Proteins. Annu. Rev. Biophys. Biomol. Struct. 33, 25–51.
3. Gregersen, N., Bross, P., Vang, S., and Christensen, J.H. (2006). Protein Misfolding and Human Disease. Annu. Rev. Genom. Hum. Genet. 7, 103–124.
4. Carpenter, E.P., Beis, K., Cameron, A.D., and Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. Current Opinion in Structural Biology 18, 581–586.
5. 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.
6. Yano, Y., Takemoto, T., Kobayashi, S., Yasui, H., Sakurai, H., Ohashi, W., Niwa, M., Futaki, S., Sugiura, Y., and Matsuzaki, K. (2002). Topological Stability and Self-Association of a Completely Hydrophobic Model Transmembrane Helix in Lipid Bilayers. Biochemistry 41, 3073–3080.
7. Johnson, R.M., Hecht, K., and Deber, C.M. (2007). Aromatic and Cation−π Interactions Enhance Helix−Helix Association in a Membrane Environment †. Biochemistry 46, 9208–9214.
8. Faham, S., Yang, D., Bare, E., Yohannan, S., Whitelegge, J.P., and Bowie, J.U. (2004). Side-chain Contributions to Membrane Protein Structure and Stability. Journal of Molecular Biology *335*, 297–305.
9. 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.
10. 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.
11. Popot, J.L., and Engelman, D.M. (1990). Membrane protein folding and oligomerization: the two-stage model. Biochemistry 29, 4031–4037.
12. Hong, H. (2014). Toward understanding driving forces in membrane protein folding. Archives of Biochemistry and Biophysics 564, 297–313.
13. Zhang, Z., Norris, J., Schwartz, C., and Alexov, E. (2011). In Silico and In Vitro Investigations of the Mutability of Disease-Causing Missense Mutation Sites in Spermine Synthase. PLoS One 6.
14. Anderson, S.M., Mueller, B.K., Lange, E.J., and Senes, A. (2017). Combination of Cα–H Hydrogen Bonds and van der Waals Packing Modulates the Stability of GxxxG-Mediated Dimers in Membranes. J. Am. Chem. Soc. 139, 15774–15783.
15. Desmet, J., Maeyer, M.D., Hazes, B., and Lasters, I. (1992). The dead-end elimination theorem and its use in protein side-chain positioning. Nature 356, 539.
16. Koehl, P., and Delarue, M. (1994). Application of a Self-consistent Mean Field Theory to Predict Protein Side-chains Conformation and Estimate Their Conformational Entropy. Journal of Molecular Biology 239, 249–275.
17. Hansmann, U.H.E., and Okamoto, Y. (1999). New Monte Carlo algorithms for protein folding. Current Opinion in Structural Biology 9, 177–183.
18. 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.
19. Lazaridis, T. (2003). Effective energy function for proteins in lipid membranes. Proteins 52, 176–192.
20. Krivov, G.G., Shapovalov, M.V., and Dunbrack, R.L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. Proteins 77, 778–795.
21. Armstrong, C.R., and Senes, A. (2016). Screening for transmembrane association in divisome proteins using TOXGREEN, a high-throughput variant of the TOXCAT assay. Biochim Biophys Acta 1858, 2573–2583.
22. Anderson, S. M. (2019). Understanding the GAS_{right} motif: sequence, structure, and stability. Thesis. University of Wisconsin, Madison, WI.
23. Korendovych, I.V., Senes, A., Kim, Y.H., Lear, J.D., Fry, H.C., Therien, M.J., Blasie, J.K., Walker, F.A., and DeGrado, W.F. (2010). De Novo Design and Molecular Assembly of a Transmembrane Diporphyrin-Binding Protein Complex. J. Am. Chem. Soc. 132, 15516–15518.

24. Joh, N.H., Wang, T., Bhate, M.P., Acharya, R., Wu, Y., Grabe, M., Hong, M., Grigoryan, G., and DeGrado, W.F. (2014). De novo design of a transmembrane Zn2+-transporting four-helix bundle. Science 346, 1520–1524.
25. Huang, P.-S., Oberdorfer, G., Xu, C., Pei, X.Y., Nannenga, B.L., Rogers, J.M., DiMaio, F., Gonen, T., Luisi, B., and Baker, D. (2014). High thermodynamic stability of parametrical-ly designed helical bundles. Science 346, 481–485.
26. Lu, P., Min, D., DiMaio, F., Wei, K.Y., Vahey, M.D., Boyken, S.E., Chen, Z., Fallas, J.A., Ueda, G., Sheffler, W., et al. (2018). Accurate computational design of multipass transmembrane proteins. Science 359, 1042–1046.
27. 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.
28. 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. Journal of Molecular Biology 272, 266–275.
29. 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.
30. 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 ClC Cl−/H+ antiporter in lipid bilayers. ELife 5, e17438.