Chapter 1: Introduction

Table of Contents

[1.1 Introduction to single-pass membrane proteins 3](#_Toc161323742)

[1.2 Forces involved in membrane protein association 4](#_Toc161323743)

[1.3 GASright 5](#_Toc161323744)

[1.4 Studying membrane protein folding and structure 6](#_Toc161323745)

[1.4.1 Computational methods 6](#_Toc161323746)

[1.4.2 Experimental methods 6](#_Toc161323747)

[1.5 Overview of thesis 7](#_Toc161323748)

[1.6 References 8](#_Toc161323749)

****1.****1 ****Introduction to membrane proteins****

Membrane proteins (MP) comprise 25-30% of the proteins found within protein-coding genes of various organisms (Fagerberg et al., 2010). (go into the types of membrane proteins here: beta barrels, sheets, end with alpha helices). Proper membrane protein folding is critical for essential biological functions, including cell signaling, ion balance, and gene regulation (expand on this sentence with more details about signaling, ph balance, and the cascades involved in 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). (expand on how it is involved in each of these diseases, a sentence on each). To understand how protein misfolding plays a role in disease states and progression, it is necessary to investigate how these proteins fold (rewrite this sentence into something better, but basically yes).

However, studying membrane protein folding is inherently a difficult challenge because of their hydrophobic nature (more detail, maybe something about the membrane layer that houses these proteins is complex and difficult to completely account for computationally, like in large simulations like MD, 2-3 sentences). 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) (additional information about experiments). To combat these challenges, much of the research studying membrane protein folding is focused on understanding the biophysical forces that govern the folding process. (add in a bunch of the interesting membrane protein studies that you’ve read, grouping them by design, experiment, and both *or something like that*)

****1.2 Forces involved in membrane protein association****

* Define the forces involved
  + How did we get to these? When did they first originate?
* First studies to qualitatively identify these forces, first studies to quantify these forces, current work that utilizes these forces
* How do we quantify these?
  + What are the experiments that led to calculations to quantify? What main experiments?
* Reason why computation is valuable at turn of century???

(what do we gain from understanding biophysical forces? The ability to design and engineer novel proteins, exploring and predicting their structures, and training to do both *or something like that*) In addition, this knowledge can be applied to design new therapeutics that specifically target proteins in these misfolded states. (better ending sentence about how learning in depth about these forces will improve human life)

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.

****1.3 The importance of van der Waals packing****

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) (expand and draw out more from these studies, 1-2 sentences), but research is lacking on the contribution of van der Waals packing (expand on why).

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 (expand on this; could go into the other forces involved in the membrane?).

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 (say something about the complexities of this force in the membrane and contrast with soluble). 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. My research aims to characterize and quantify the extent at which sidechain packing is a driving force for membrane protein association.

****1.3 GASright****

* What is it? Why are we interested in it?
* Talk about it as a good system for understanding forces because …cite papers here
* Maybe look at Samantha’s and Gladys for guidance here

****1.4 Studying membrane protein folding and structure****

* General paragraph for why we need to understand, probably harping back on disease paragraph from earlier

****1.4.1 Computational methods****

* Talk about the importance of papers at the interfaces of all disciplines, needing to understand the computation
* Talk about the need to understand and utilize it as a tool to increase what we can learn and the rate at which we learn about these small systems, could even cite the lady who recently imaged the black hole because of her ability to store and process so much data
* Rosetta -> Alphafold -> all the new methods that are still coming up
* The use of AI in analyzing data and building new structures
* Thoughts about how to evaluate the field in the presence of AI?

****1.4.2 Experimental methods****

* **Talk about the experiments used in the past, what were they able to discover**
* **The establishment of experimental databanks like the PDB and memprotDB and others**
* **Go into the quantity increase in the last 15 years** 
  + **talk about the difference between number of structures in the PDB with number of structures predicted by Alphafold (and compare how long it took and maybe now how many papers are citing this powerful tool for their experiments)**
* **Combining computation expands the ability to think of more challenging and complex ways to tackle experiments**
* **The important methods for matching this high-throughput age of science:**
  + **FACS, gene chips, NGS, cloning strategies, expression strategies, machines**
  + **Other things that are able to match the throughput of computation**

****1.5 Overview of thesis****

****1.6 References****