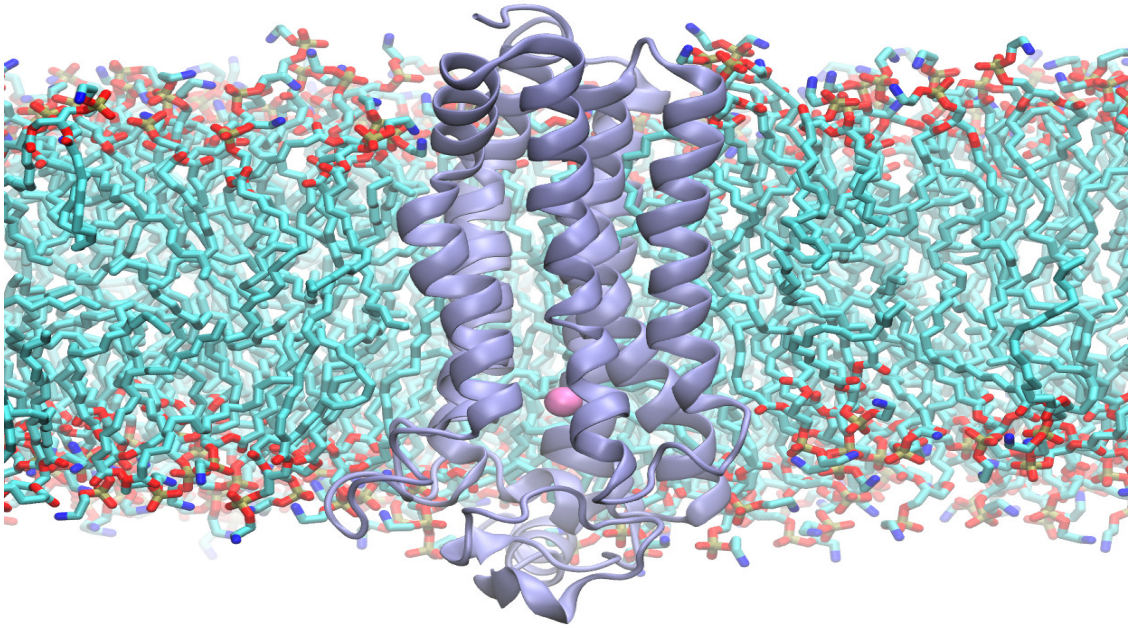
CHALMERS
UNIVERSITY OF TECHNOLOGY



Understanding the Mechanism of PAQR-2 through Modeling and Simulations

Master's Thesis in Nanotechnology

HILDA SANDSTRÖM

Department of Microtechnology and Nanoscience
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Understanding the Mechanism of PAQR-2 through Modeling and Simulations

HILDA SANDSTRÖM

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Microtechnology and Nanoscience
Chalmers University of Technology
Gothenburg, Sweden 2017

Understanding the Mechanism of PAQR-2 through Modeling and Simulations
HILDA SANDSTRÖM

Supervisor: Samuel Genheden, Chemistry and Molecular Biology, University of Gothenburg
Examiner: Elsebeth Schröder, Microtechnology and Nanoscience, Chalmers University of Technology

Department of Microtechnology and Nanoscience
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Typeset in LaTeX
Gothenburg, Sweden 2017

Understanding the Mechanism of PAQR-2 through Modeling and Simulations
HILDA SANDSTRÖM
Department of Microtechnology and Nanoscience
Chalmers University of Technology

Abstract

PAQR-2 is a transmembrane protein in *C. elegans* with seven transmembrane helices and a zinc site. As a homolog to the human ADIPOR receptors, which are involved in regulating membrane fluidity and are of importance when studying glucose toxicity in diabetes patients, the mechanism of PAQR-2 holds many undiscovered secrets to understanding membrane fluidity regulation. PAQR-2 has been shown to be vital for survival in cold temperatures and in the presence of high glucose levels. Other fluidity sensitizing proteins have been shown to change conformation in different membrane environments. In this study, molecular dynamics simulations of PAQR-2 were done with the purpose of observing the structural response of PAQR-2 to different membrane environments. Both simulations of only the transmembrane domain and of the full protein were made. In addition, two loss of function mutants (d282n and g533r) were simulated and compared with the wild-type. Furthermore, the IGLR-2 protein that has been shown to be vital for the function of PAQR-2 was simulated and docked with PAQR-2 yeilding a likely structure for the PAQR-2:IGLR-2 complex. Simulations of PAQR-2 in a thick and ordered DPPE membrane revealed an adaptation of the membrane thickness to accommodate PAQR-2, rather than a structural change within the protein itself. The g533r mutation introduced novel interaction sites between the helices. The d282n mutation resulted in a loss of hydrogen bonds of the residue, which sits close to the zinc site. Protein-protein docking and PMF calculations using umbrella sampling revealed four possible PAQR-2:IGLR-2 complexes. The interactions of the highest scoring complex were analyzed and classified as being primarily weak interactions. The full protein model of PAQR-2 which includes both the transmembrane domain and the cytosolic domain, shows most promise as a model of PAQR-2, as it captures dynamics surrounding the zinc site predicted by the homolog model. Moreover, the full protein model describes different behaviour between wild-type and the d282n mutant not found in the model of only the transmembrane domain. Further optimization of the full protein is required as an un-physical loss of secondary structure occurs in the cytosolic domain when simulated close to the membrane.

Keywords: *C. elegans*, Molecular dynamics, Simulations, Transmembrane proteins, Glucose toxicity, Cold sensitivity, Membrane fluidity

Acknowledgements

I would like to thank my supervisor Samuel Genheden for all his guidance and encouragement. I am also very grateful to Professor Leif Eriksson for creating the homology models for PAQR-2, and to his group for their company. Lastly, I would like to thank Professor Marc Pilon for suggesting the project in the first place.

Hilda Sandström, Gothenburg, September 2017

Contents

List of Figures **xi**

List of Tables **xv**

1 Introduction **1**
1.1 Aim . . . 2
1.2 Question Formulation . . . 3
1.3 Method . . . 3

2 The cell membrane **5**

3 Computational Methods **9**
3.1 Molecular Mechanics and Dynamics . . . 9
3.2 Implementation of Molecular Dynamics Simulations . . . 10
3.2.1 Long-range Interactions . . . 10
3.2.1.1 Cut-off for Long Range Interactions . . . 10
3.2.1.2 Particle-Mesh Ewald Summation . . . 10
3.2.2 Boundary Conditions . . . 10
3.2.3 Time Step . . . 11
3.2.4 Temperature and Pressure Control . . . 11
3.3 Force Field Models . . . 12
3.3.1 Amber ff99sb-ildn . . . 12
3.3.2 The TIP3P Water Model . . . 12
3.3.3 The Martini Force Field . . . 12
3.4 Protein-Protein Docking . . . 13
3.5 Umbrella Sampling and WHAM for Calculation of the Potential of Mean Force . . . 14
3.6 Delimitations . . . 15

4 Methodology **17**
4.1 System setup: Protein and Membrane . . . 17
4.1.1 Preparation of All-Atom Simulations . . . 17
4.1.2 Preparation of Coarse-Grained Simulations . . . 18
4.2 Simulations . . . 18
4.2.1 All-Atom Simulation Specifics . . . 18
4.2.2 Coarse-Grained Simulation Specifics . . . 19
4.3 Protein-Protein Docking with Memdock . . . 20

4.4 Calculating Potential of Mean Force 20

5 Results and Discussion **21**
5.1 Simulations of PAQR-2 21
5.1.1 Simulations of the Transmembrane Domain of PAQR-2 21
5.1.1.1 Structure Comparison 21
5.1.1.2 Protein Dynamics 24
5.1.1.3 Membrane Properties 25
5.1.1.4 Interactions of the d282n PAQR-2 Mutant 27
5.1.1.5 Interactions of the g533r PAQR-2 Mutant 27
5.1.1.6 Reproducibility and Convergence of Trajectories . . . 29
5.1.2 Full Protein Simulations 30
5.1.2.1 Structure Comparison 32
5.1.2.2 Secondary Structure Loss in Cytosolic Domain . . . 32
5.1.2.3 Protein Dynamics 34
5.1.2.4 Membrane Properties 35
5.1.2.5 The d282n Mutant 35
5.2 The PAQR-2:IGLR-2 Complex 38
5.2.1 Evaluation of the Memdock Results 38
5.2.2 The PAQR-2:IGLR-2 complex simulations 38
5.2.3 Interaction Sites between PAQR-2 and IGLR-2 39

6 Conclusion **41**
6.1 Future Outlook 41

Bibliography **43**

A Appendix I **I**
A.1 Hydrogen bonds of ASP282/ASN282 in TMM simulations I
A.2 Hydrogen Bonds of ASP282/ASN282 in Full Protein Simulations . . . III
A.3 Hydrogen Bonds of ARG533 IV
A.4 Distance between Center of Mass in PAQR-2:IGLR-2 complex V

List of Figures

1.1 (a) The transmembrane domain of PAQR-2 with ASP282 and GLY533 are colored red. (b) The transmembrane domain of IGLR-2. 2

1.2 (a) *C. elegans* carrying eggs. (b)A normal tail of *C. elegans* and withered tail tip phenotype of the PAQR-2 mutant 3

2.1 (a) A bilayer made out of phospholipids. (b) A DOPE lipid. 6

2.2 (a) The left panel shows the chemical structure of a DOPE lipid. (b) The right panel shows the chemical structure of a DPPE lipid. 7

3.1 (a) IGLR-2 represented using an all-atom approach. (b) IGLR-2 in a coarse-grained representation. 13

3.2 (a) Histograms of the distance between the center of mass of two proteins for 21 different simulations. (b) The position of the first protein (black) and the second protein at 0 (red), 5 (orange), 10 (yellow), 15 (green) and 20 (blue) Å added distance between their centers of mass. 14

4.1 (a) The system before the protein (green) has been inserted into the membrane (dark blue). The protein inserted in the membrane. The water solvent is coloured in a lighter blue. 19

5.1 (a)-(b) A superposition of the structure of the TMD of PAQR-2 after a TMM simulation of the wild-type in a DOPE membrane (red) for 1.5 μs. Helices 5 and 6 are bent when compared with the original structure (blue). (c)-(d) Same as in (a)-(b) but here the original structure (blue) is compared with the structure in the DPPE membrane (orange). The distance between GLY464 in the two structures is shown. 22

5.2 A contact comparison between the PAQR-2 TMD from the different sets of simulations. A contact similarity of 1 means identical structures. This is shown along the diagonal. 24

5.3 (a) Correlation chart between the first principal modes of the 1.5 μs TMM simulations of PAQR-2. (b) The figure shows the characteristic fluctuation profile of each of the first modes from the TMD simulations. 25

5.4 The average chain order parameter of the DPPE membrane (A) and DOPE membrane (B) from the TMM simulations of the wild-type. The white cutout is where PAQR-2 sits. A completely ordered membrane has an order parameter of 1 and a completely disordered membrane has an order parameter of 0. 26
5.5 The membrane thickness of a DPPE membrane (A) and DOPE membrane (B) from the TMM simulations of the wild-type. The white cutout is where PAQR-2 sits. 26
5.6 The final snapshot from the TMM simulations - highlighting the vicinity of ASP282/ASN282. All distances are labeled in Ångström. 28
5.7 (a) The distance between the zinc ion and carbonyl of residue 282 in the wild-type and d282n mutant. (b) The zinc site and residue 282 of wild-type PAQR-2 after a TMD simulation in a DOPE membrane. All distances are labeled in Ångström. 29
5.8 The structure close to residue 533 in the wild-type structure after simulation in DOPE (a), and of the g533r PAQR-2 mutant in a DPPE membrane (b) and DOPE membrane (c). 29
5.9 (a) The model of the entire PAQR-2. (b) PAQR-2 colored according to a hydrophobicity scale going from red (hydrophobic) to blue (hydrophilic). All distances are labeled in Ångström. 30
5.10 (a) A superposition of the structure of the TMD of PAQR-2 after a full protein model simulation in a DOPE membrane (red). Helices 5 and 6 are bent when compared with the crystal structure. GLY464 has moved 11.5 Å. (b) Same as in (a) but here the original structure is compared with the structure in the DPPE membrane (orange). GLY464 has moved 7.5 Å. 31
5.11 The figure shows the structure of the first 300 residues of PAQR-2 after various treatment. In (a) the starting structure is showed. (b) shows the structure after 100 ns simulation in water. (c) shows the structure after a 100 ns simulation when connected to the TMD of PAQR-2 in a DOPE membrane. (d) The same as in (c) but instead of DOPE a DPPE membrane was used. 33
5.12 The secondary structure as a function of residue from different structures of the first 300 residues of PAQR-2. The loss of helices in the structure from the full protein simulations, here called DOPE and DPPE, is visible. 34
5.13 (a) Correlation chart between the first principal modes of the TMD of the full protein simulations (b) The figure shows the characteristic fluctuation profile of each of the first modes of the TMD in the full protein simulations. 34
5.14 The membrane thickness of a DPPE membrane (A) and DOPE membrane (B) from the full protein simulations of the wild-type. The white cutout is where PAQR-2 sits. 35
5.15 The average chain order parameter of the DPPE membrane (A) and DOPE membrane (B) from the full protein simulations of the wild-type. The white cutout is where PAQR-2 sits. 36

5.16 (a) The distance between the carbonyl of ASP282/ASN282 and the zinc ion as a function of time. (b) The zinc site at the end of a wild-type full model simulation. The distances are labeled in Ångström. . 36
5.17 Final snapshots of the structure in the vicinity of ASP282/ASN282. Distances are labeled in Ångström. 37
5.18 PAQR-2 (black) with the four selected docking results. The highest scoring (red), second (orange), third (green) and fourth (blue). (a) A top view with the PAQR-2 helices numbered from 1 to 7 in pink. (b) A side view. 39
5.19 The potential of mean force as a function of the distance between the center of mass of PAQR-2 and IGLR-2. The curves for docking structure one to four is plotted. The grey shaded area shows the estimated error. (a) The PMF curves in a DOPE membrane. (b) The PMF curves of a DPPE membrane. 40
5.20 (a) The PAQR-2:IGLR-2 complex colored using the hydrophobic scale, going from purple (hydrophobic) to blue (hydrophilic). (b) The PAQR-2:IGLR-2 complex where PAQR-2 is colored in blue and IGLR-2 is colored in beige. The residues making out the contact between the molecules have been colored orange. 40

6.1 View of the TMD of PAQR-2 from the side (a), above (b) and below (c). The identified pore volume is colored in red. 42

A.1 (a) Number of hydrogen bonds of ASP282 in the TMM simulation of the wild-type. (b) Number of hydrogen bonds of ASN282 in the TMM simulation of the d282n mutant. II
A.2 (a) Number of hydrogen bonds of ASP282 in the wild-type full protein simulations. (b) Number of hydrogen bonds of ASN282 in the d282n mutant full protein simulations. III
A.3 (a) Number of hydrogen bonds of the g533r mutant in the TMM simulations. Distance cut-off between donor and acceptor was set to 3.5 Å. (b) Length of the most present hydrogen bonds with GLU343 and GLN338 in the simulations. The main chain hydrogen bonds were primarily with ILE529 and LEU537. IV
A.4 (a) Overlay of the original PAQR-2:IGLR-2 complex (blue) and after a 100 ns simulation in a DPPE membrane. (b) same as in (a) but in a DOPE membrane. V
A.5 (a) Time evolution of distance between the center of mass of PAQR-2 and IGLR-2 in the PAQR-2:IGLR-2 complex. V

List of Tables

4.1 The histidine residues of PAQR-2 together with the position on which they are protonated. 17

5.1 The time average of the global RMSD of the protein backbone after a least square fit to the backbone of the TMD helices. 23
5.2 The percentage of the total fluctuations made out by the first principal component in the TMM simulations. 25
5.3 The percentage of the total fluctuations made out by the first principal component in the full protein simulations. 35

A.1 The hydrogen bond occupancy of the hydrogen bonds, of ASP282/ASN282 in the TMM simulations, that were present for more than 5 ns (0.3 %) in the simulation. I
A.2 The hydrogen bond occupancy of the hydrogen bonds of ASP282/ASN282 in the full protein simulations of the wild-type that were present for more than 5 ns (1 %) in the simulation. III
A.3 The hydrogen bond occupancy of the hydrogen bonds of the side chain of ARG533 that were present for more than 5 ns (0.3 %) in the simulation. IV

1
Introduction

In 2015 there were 415 million people with diabetes worldwide and 12 % of the global health expenditure was spent on diabetes [1]. Elevated glucose levels in diabetes patients are associated with a large increase in mortality rate in connection to cardiovascular disease [2]. However, the exact mechanism behind glucose toxicity remains unclear. The hormone adiponectin is known to be involved in glucose metabolism and interacts with two receptor proteins (ADIPORS) located in the cell membrane [3]. In addition, a study on obese monkeys showed that a decrease in adiponectin concentration was correlated to type 2 diabetes [4]. The receptor proteins ADIPOR1 and ADIPOR2 are enzymes known to mediate the production of sphingosine and free fatty acid though lipid metabolism of the lipid ceramide [5]. The crystal structure of the ADIPORS shows a transmembrane structure made out of seven helices, with a zinc binding site close to the intracellular interface [6]. The zinc binding site is believed to be important for the lipid metabolism activity of the receptors.

A homolog to the ADIPORS in the model organism *C. elegans* is the PAQR-2 protein, which has been studied by the group of Marc Pilon [7] [8]. The homologous genes which encode the proteins share the same ancestral gene. *C. elegans* is one of a handful of organisms that have been used as model system in molecular biology. Research on *C. elegans* began in the sixties, and has since led to many important scientific discoveries. There are a number of advantages to using *C. elegans* as a model system, these include a three day life cycle, its small size (1.5 mm), and its small and sequenced genome [9]. It is also transparent, hermaphroditic and possible to store in stocks using cryo-preservation. In the lab it is typically grown on agar plates with *E. coli*. It is inexpensive to maintain and easy to cultivate. In addition, 38 % of the protein encoding genes in *C. elegans* have human homologs [10], meaning that they share the same ancestral gene. Genes that are homologs through speciation typically retain the same function. Furthermore, 40 % of genes connected with diseases in humans have homologs in *C. elegans* [11]. Therefore, studies on *C. elegans* could be useful for understanding the function of its human relatives, the ADIPOR receptors.

Similar to the ADIPOR receptor, PAQR-2 has a seven transmembrane domain, with a zinc site close to the intracellular part of the membrane. The structure of the transmembrane domain of PAQR-2 is shown in the right side panel of Figure 1.1. Svensk et al. have isolated several loss of function mutations of PAQR-2 [7]. The PAQR-2 mutants showed impaired growth at low temperatures, a withered tail tip, an abnormal fatty acid composition and glucose intolerance. Their results suggest

(a) **(b)**

Figure 1.1: (a) The transmembrane domain of PAQR-2 with ASP282 and GLY533 are colored red. (b) The transmembrane domain of IGLR-2.

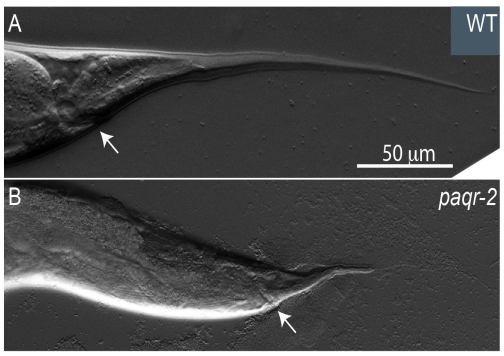
that like the ADIPORS, PAQR-2 involved with phospholipid metabolism. In Figure 1.2 *C. elegans* with wild-type and mutant PAQR-2 protein are shown. Svensk et al. have found evidence suggesting PAQR-2 regulates desaturase proteins, enzymes that can introduce double bonds in fatty acids. Furthermore, it has been found that there is an increased amount of unsaturated lipids in the membrane of *C. elegans* at low temperatures. This is important for *C. elegans* as it cannot regulate its body temperature, except through behavioural means. Therefore, without any fluidity regulation, its membrane will become more rigid at low temperatures. High glucose levels share the same effect [8].

In addition to their study of PAQR-2, Pilon's group have identified the protein IGLR-2 as a partner protein to PAQR-2. They have shown that the function of PAQR-2 affected by mutants in IGLR-2, and vice versa. IGLR-2 has a single transmembrane domain. It belongs to a family of proteins that are known to regulate activity of receptor proteins and voltage-gated protein channels. Through experiments they have shown that the proteins are located together in the membrane. The left panel of Figure 1.1 shows the structure of the transmembrane domain of IGLR-2.

1.1 Aim

This report presents a study with the aim to deepen the understanding of the mechanism of PAQR-2. This is achieved through an investigation of the structural response of PAQR-2 to membranes of varying fluidity using all-atom molecular dynamics simulations. In addition, protein-protein docking and coarse-grained as well as all-atom simulations of the PAQR-2:IGLR-2 complex in a membrane are performed to investigate their interaction.

(a)



(b)

Figure 1.2: (a) *C. elegans* carrying eggs. (b) A normal tail of *C. elegans* (upper) and withered tail tip phenotype of the PAQR-2 mutant (lower). Arrows indicate the location of the anus. Photographs courtesy of Professor Marc Pilon at the Department of Chemistry and Molecular Biology.

1.2 Question Formulation

- How does the wild type protein behave in a rigid or a more fluid membrane? Is there a change in the structure of the protein in the simulation. What parts of the protein are dynamic in the different membranes?
- How do the membrane properties, such as order parameter and thickness, change close to the protein in the different membrane models?
- What are the differences in structure and dynamics between wild type PAQR-2 and loss of function mutants?
- How do IGLR-2 and PAQR-2 interact?

1.3 Method

Molecular dynamics simulations have been a useful tool when investigating the interaction between proteins and lipid membranes [12]. The continuing improvement of algorithms and hardware have enabled longer simulations and larger systems to be studied. In particular, molecular dynamics simulations have been used as a tool to understand the mechanism of DesK, a protein that can activate fatty acid desaturase. A study by Cybulski et al. [13] found a mechanism driven by hydrophobic mismatch between the protein and the membrane. Another molecular dynamics study investigated the Mga2 protein in fungi [14]. The study investigated the response of the protein to different types of membranes. What they found was a a rotation based mechanism. Interestingly, most known proteins that regulate the fluidity of the membrane have a transmembrane domain of α-helices [15].

2

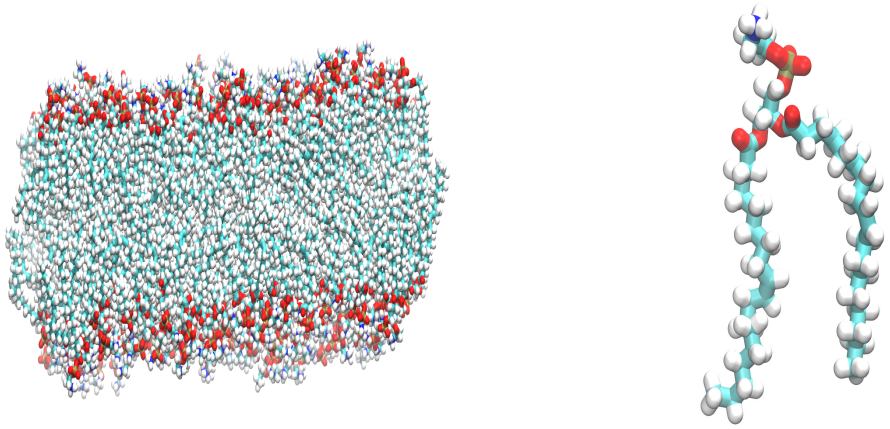
The cell membrane

The cell membrane consists of a self-assembled lipid bilayer and proteins. Its functions include sieving, protection and compartmentalization. It encapsulates the cell and its organelles. The lipid bilayer consists of two monolayers of phospholipids. There are a variety of types of phospholipids. What characterizes the phospholipids are the two hydrocarbon chains, which can vary in length and have zero or more double bonds, and the polar head group containing a phosphate group, with a substitution that also can vary. In short, the phospholipids have a polar end and a hydrophobic end. The hydrophobic effect drives the self-assembly of the lipids in water so that the head groups are at the interface between the membrane and the water and the hydrophobic ends are hidden in the inside of the membrane. Figure 2.1 shows the lipid bilayer and some different phospholipids.

The membrane contains a variety of lipids, including phospholipids and sterols, and proteins. The membrane composition is not random, but has domains. Both the domains and the proteins can diffuse in the membrane. The functionality of this organization is sensitive to changes in pressure and temperature [16]. The membrane must be able to adapt to proteins but also remain stable to act as a permeability barrier. The content of the membrane is also important, introducing sterols promotes order, while unsaturated lipids that typically have more conformational freedom promote a disordered membrane.

The state of the membrane can be characterized by a number of properties. Firstly, the order parameter is a measure of how structured the lipids are in the membrane. To evaluate the order parameter one imagines comparing the directions of the bonds in the fatty acid chains with the direction of the membrane normal. Another important property of the membrane is its thickness. The thickness can vary with varying length of the fatty acid tails of the lipid. The thickness also depends on the order of the membrane. More straight conformations produce a thicker membrane. Parameters that affect the thickness and order parameter of the membrane are temperature and degree of saturation.

Membranes undergo phase transitions. A self-assembled bilayer of one type of lipid has a main transition which occurs at a certain temperature which is referred to as the transition temperature. At temperatures above the transition temperature the lipids are in a liquid disordered phase, and below it they are in a solid ordered phase. What is more, this transition is accompanied with a lowering of the thickness as the temperature increases.

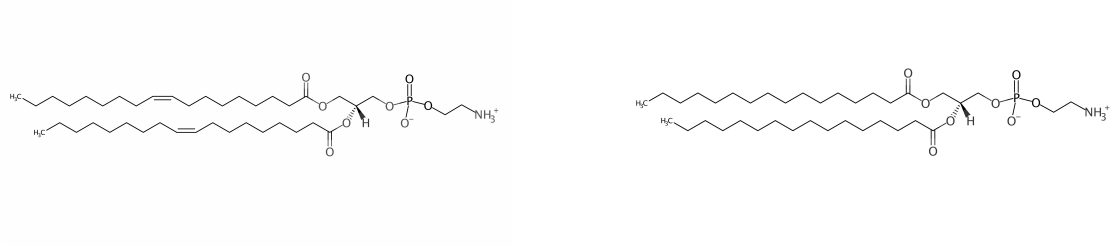


(a) **(b)**

Figure 2.1: (a) A bilayer made out of phospholipids. (b) A DOPE lipid.

The change in thickness of the bilayer can be important for the many integral membrane proteins that reside in the membrane. Typically the protein and membrane are arranged with a hydrophobic matching. When the membrane thickness changes, the protein may undergo conformational changes to reduce the hydrophobic mismatch created by the change. A known example of a protein that appears in different states in a thick or thin membranes is the transmembrane OmpA in some bacteria [16]. In addition, the transmembrane protein DesK in bacteria, which is known to regulate membrane fluidity, to take different conformations in thick and thin membranes [13]. This is believed to be a part of its mechanism as a temperature sensor.

C. elegans has 471 documented genes involved in lipid metabolics [17]. It has the ability to produce fatty acids by its own, and it also has uptake of fatty acids through its diet. Among the the resulting lipids are the phosphatidyl ethanolamines (PE) and phosphatidyl cholines (PC). Of the two, PE is the most common glycerophospholipid in *C. elegans* [18]. The two lipids chosen to represent the two model systems used in the study presented in this report are the PE lipids 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) and 1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine (DPPE). DPPE has a main transition temperature of 63 °C [19], so it is in its solid and ordered phase at room temperature. DOPE on the other hand has a transition temperature of -16 °C and will therefore be in its liquid phase at room temperature. Figure 2.2 shows the structure of the DOPE and DPPE lipids. The membrane of *C. elegans* does not contain sterols unlike other organisms, in which the sterols contribute to the structure of the membrane [17].



(a) **(b)**

Figure 2.2: (a) The left panel shows the chemical structure of a DOPE lipid. (b) The right panel shows the chemical structure of a DPPE lipid.

3

Computational Methods

3.1 Molecular Mechanics and Dynamics

In molecular dynamics simulations for classical systems Newton's equations of motion for the system are integrated to produce a trajectory of the system over time. The aim can be to study the interaction or structures of different molecules or to calculate thermodynamic properties of the system. In a system comprised of N particles with a position $\mathbf{q}_i$ and momentum $\mathbf{p}_i$. The equations of motion are

$$\frac{d\mathbf{q}_i}{dt} = \mathbf{p}_i \tag{3.1}$$

$$\frac{d\mathbf{p}_i}{dt} = \mathbf{F}(\mathbf{q}_i), \tag{3.2}$$

where the force on particle i arises from the potential V

$$\mathbf{F}(\mathbf{q}_i) = -\frac{\partial V}{\partial \mathbf{q}_i}. \tag{3.3}$$

In molecular systems the potential can include different types of interactions. For instance, there are bonded and non-bonded interactions. Bonded interactions arise between atoms in the same molecule and are functions of the interacting atom species as well as bond lengths and angles and torsions. In widely used force fields such as AMBER and CHARMM the potential energy function terms take on a similar form [20]. The bond and angle contributions are modelled as harmonic potentials, which is an appropriate approximation close to equilibrium. The torsion contribution is modelled as a Fourier sum. The non-bonded interactions are both electrostatic, described by a Coulomb potential, and dispersion and repulsion interactions. The Lennard-Jones potential is typically used to represent both the repulsion and dispersion contributions

$$\mathrm{V} = \sum_{bonds} k_b(l-l_0)^2 + \sum_{angles} k_a(\theta-\theta_0)^2 + \sum_{torsions} \tfrac{V_n}{2}[1+cos(n\phi-\gamma)] + \sum_{\substack{nonbonded\\pairs}} [k_e\tfrac{q_iq_j}{r_{ij}} + \epsilon[\tfrac{\rho^{12}}{r_{ij}^{12}} - 2\tfrac{\rho_6}{r_{ij}^6}]] \tag{3.4}$$

Here l refers to the bond length and θ to the bond angle. l_0 and θ_0 are the equilibrium values of the bond length and bond angle respectively. k_b and k_a are constants of the harmonic potentials. ϕ is a torsion angle and V_n is the maximum value of the torsion potential term. q_i refers to the charge of particle i and r_{ij} is the distance between particles i and j. k_e is Coulomb's constant. ϵ is the depth of the Lennard-Jones potential well and ρ is the distance where the potential well is at its minimum.

3.2 Implementation of Molecular Dynamics Simulations

The aim of this section is to provide sufficient information about the setup required for the molecular dynamics simulations, along with advantages and disadvantages with the different method options.

3.2.1 Long-range Interactions

3.2.1.1 Cut-off for Long Range Interactions

The non-bonded interactions are computationally demanding. Therefore, a cut-off is used for the non-bonded interactions. The cut-off radii for electrostatic potential, and the Lennard-Jones potential are parameters which need to be specified. Using a cut-off introduces the need to keep track of which particles interact at different times. This list of interacting particles is referred to as the neighbour list. The Verlet cut-off scheme [21] is an efficient way of calculating the non-bonded interactions because it clusters the particles and forms a neighbour list of clusters which is then used to find non-bonded interactions between particles.

3.2.1.2 Particle-Mesh Ewald Summation

Evaluating the electrostatic potential is a computationally expensive task. By utilizing the Particle-mesh Ewald (PME) summation the computing time is significantly improved. PME is an extension of Ewald summation, the principle of which is to divide the sum into short-range and long-range interactions. The short range interactions are calculated in real space and the long-range interactions in reciprocal space. In reciprocal space the long-range interaction term converges fast and can be truncated to save computation time. However, the long-range interactions scales as N^2. PME is faster than Ewald summation unless the system is very small. PME summation involves evaluating the charge density using fast Fourier transform (FFT). With PME the computing time scales as $N\log N$.

3.2.2 Boundary Conditions

In most molecular dynamics simulations periodic boundary conditions are used. They give the effect of a larger system using only a small unit cell. This less computationally expensive than actually simulating a larger system. However, a molecule should not be able to interact with itself or more than one copy of another molecule. This, together with the long-range interaction cut-off, introduces restraints on the box size. Each side of the box must be longer than the size of the protein in that direction added with two times the interaction cut-off length, R_c

$$L_i \geq 2R_c + l_i, \tag{3.5}$$

where L_i is the box length in direction i and l_i is the length of the protein in direction i.

3.2.3 Time Step

Choosing the right time step for a simulation requires consideration of which motions are included in the system. A long time step is desired for computational efficiency, but a too long time step will result in an accumulation of error in the trajectory [22]. A simulation of molecules needs a time step of 0.5 to 5 fs, depending on which types of movements are allowed in the simulation. If all types of movements, vibration, torsion, rotation and translation are considered, smaller time steps are needed. If bond vibrations are excluded, larger time steps are possible. Constraint dynamics is used to enable the use of larger time steps even when such movements are considered. LINCS [23] is the algorithm that per default is used by Gromacs when one wants to constrain bond lengths and angles. In addition, simulations in Gromacs can be implemented with virtual hydrogens, which means that the position of the hydrogens are calculated from the positions of the atom they are bonded to. This removes high frequency vibrations involving hydrogens. For some molecular units such as amine or hydroxide groups, the virtual site representation cannot be used because it would remove a rotational degree of freedom. Instead, the mass is repartitioned within the group by increasing the mass of the hydrogens and decreasing that of the the other atoms [24]. This will increase the moment of inertia, maintain equilibrium properties and not change the dynamics significantly [25].

3.2.4 Temperature and Pressure Control

The systems that molecular dynamics simulations model are often ones that are under constant pressure and or temperature, such as proteins found in the body. Different thermostats and barostats have been developed since the beginning of the 1980's to provide temperature and pressure control in simulations. Gromacs provides the option of different types of thermostats [25]. Two of them are weak coupling methods, the Berendsen thermostat and the velocity-rescaling thermostat. The Berendsen thermostat [26] is based on a physical model where the system is coupled to an external heat bath, which enables a fast and stable control of the temperature of the system. The advantage of the Berendsen thermostat is a fast exponential decay of the temperature towards the target value. The strength of the coupling is determined by the coupling time constant. A large value of the coupling constant produces a weak coupling to the bath. The disadvantage of the Berendsen thermostat is that it suppresses fluctuations in kinetic energy and therefore does not produce correct ensemble averages [25]. The velocity rescaling thermostat by Bussi, Donadio and Parinello [27], is a modified version of the Berendsen thermostat that produces a proper canonical ensemble. The velocity-rescaling thermostat has included a stochastic term that produces a correct kinetic energy distribution and maintains the advantage of the Berendsen thermostat of exponential decay of the temperature. Another type of thermostat available is one that provides a oscillatory relaxation, the Nosé-Hoover coupling thermostat. It has a much longer relaxation time than compared weak coupling methods like the Berendsen algorithm. It has the disadvantage of being non-ergodic.

Berendsen also devised a barostat method for scaling the box lengths and positions to maintain pressure. It is also a weak coupling method with the advantage of a fast relaxation time. It is necessary to specify the isothermal compressibility of the system as well as coupling time constants. For simulations from which one wants to calculate thermodynamical properties it is better to use the Parinello-Rahman pressure coupling [28],[29]. It is has an oscillatory relaxation and requires larger coupling coefficients than the weak coupling method. It also modifies the equation of motion of the particles.

3.3 Force Field Models

3.3.1 Amber ff99sb-ildn

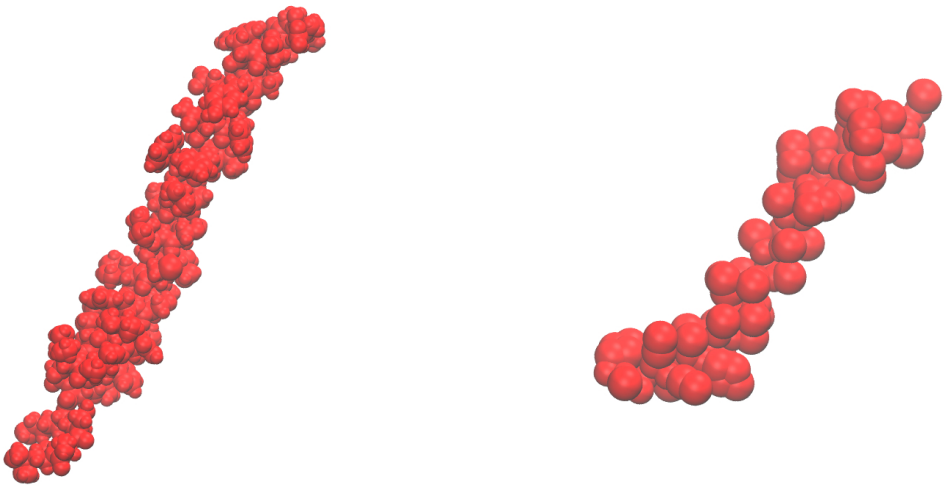
A force field includes the parameters that are needed to evaluate the different interaction energies from Equation 3.4. The Amber ff99sb-ildn force field [30] is one of a series of amber force fields. The parameters are found through a process that involves fitting the force field potentials to those of quantum mechanical calculations. Furthermore, crystal structures can be used for validation of the force field, after parametrization.

3.3.2 The TIP3P Water Model

There are a number of models of water that have been developed for use in molecular dynamics simulations. The TIP3P water model [31] was developed in 1983. It employs a rigid molecule geometry and has three charged sites, two on each of the hydrogen atoms and one on the oxygen. In the TIP3P model there is one van der Waals interaction point centered on the oxygen atom of the molecule [22]. When the flexibility in excluded as in the TIP3P model there are no vibrations that need to be considered.

3.3.3 The Martini Force Field

All-atom simulations include a detailed description of the system. But its high computational cost limits the time and length scales of the studied systems. To access longer time scales and reduce computational cost, force fields with less detailed descriptions have been developed. The Martini force field was developed in 2004 [32] and provides a coarse-grained description of molecules. In the Martini force field groups of four atoms are replaced with a bead representation. The bead is an effective interaction center. The Martini force field limits the type of interaction center to a few types, polar, nonpolar, apolar and charged. For computational efficiency all beads carry the same mass. The interactions between the beads are then calculated in a similar manner as for the all-atom force fields. The Lennard-Jones potential is used for non-bonded interaction and harmonic potentials are used for bonded interactions. The Martini force field was first developed for solvents and lipids, but has since been extended to include representations of amino acids [33].



(a) **(b)**

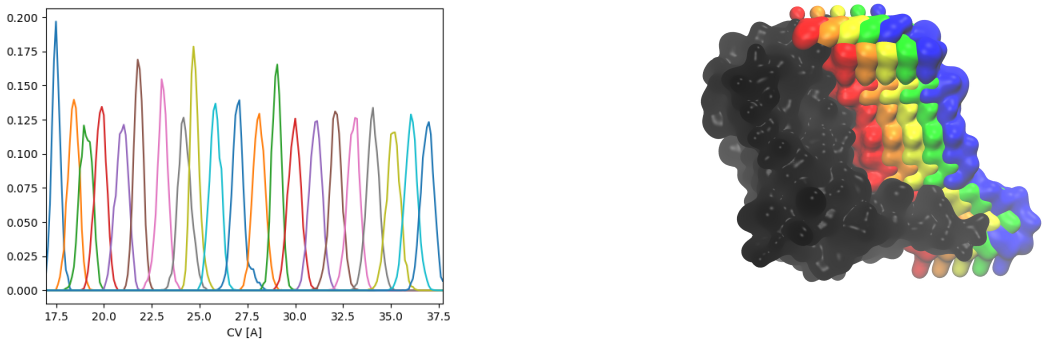
Figure 3.1: (a) IGLR-2 represented using an all-atom approach. (b) IGLR-2 in a coarse-grained representation.

The time step of the coarse-grained simulations can be in the range of 20 fs to 40 fs, which is a ten-fold increase compared to the all-atom simulations. In addition, the dynamics in the coarse grained simulations is faster than the all-atom case. Therefore, in coarse-grained simulations the effective time is 4 times the time of the simulation.

On the other hand, coarse-grained simulations suffer from limitations due to the approximations that make it so simple. The molecular resolution is limited as a result of the representation. The reduction in the number of particles affects the entropy, and therefore also the enthalpy of the system. However, free energy differences are reliable. Also, restraints need to be introduced to maintain the secondary structure of proteins. This means that the conformation of the protein will also be less likely to undergo conformational changes of any kind. The elastic network model [34] is one such tool which can be used to maintain the secondary structure. It introduces harmonic forces between non-bonded backbone beads.

3.4 Protein-Protein Docking

Docking two proteins entails finding the structure of the protein complex which is formed when the proteins interact. There are today a number of docking methods available and they can differ from each other. The process of docking two proteins involves generating a large population of possible structures for the complex. These solutions are then clustered and scored according to a scoring function that can vary depending on which program you are using. In general, the scoring function aims to classify the extent of the interaction between the proteins, be that of different



(a) **(b)**

Figure 3.2: (a) Histograms of the distance between the center of mass of two proteins for 21 different simulations. (b) The position of the first protein (black) and the second protein at 0 (red), 5 (orange), 10 (yellow), 15 (green) and 20 (blue) Å added distance between their centers of mass.

kinds (hydrogen bonding, solvation energies, electrostatics and others). To generate the population of solutions, one protein is fixated and the other is allowed to rotate and move around its partner. In addition, only some programs allow for flexible side chains of the proteins [35]. Memdock [36] is a docking program which was designed for the explicit use of docking of helical membrane proteins. It considers the structural constraints imposed by the membrane on the proteins and considers their orientation relative to the membrane. This limits the number of candidate structures.

3.5 Umbrella Sampling and WHAM for Calculation of the Potential of Mean Force

The potential of mean force can be calculated to measure the change in energy when increasing the separation of two proteins in the protein complex structure. Umbrella sampling is often used to improve sampling when calculating the potential of mean force. In short, umbrella sampling is used to increase sampling in regions that are energetically unfavourable. When using umbrella sampling a new biased potential is introduced. This bias is then accounted for when calculating the unbiased energy. In the case as for the protein complex it is convenient to perform many simulations with their own bias potential to get better sampling. WHAM [37] is an algorithm that can be used together with umbrella sampling to combine the results of the different simulations. WHAM stands for Weighted Histogram Analysis Method. Figure 3.2 shows an example of the resulting histograms from an umbrella sampling done with multiple simulations. The used WHAM software which has implemented the WHAM algorithm was [38] by Grossman.

3.6 Delimitations

The study presented in this report limits its investigation of the behaviour of PAQR-2 in two contrasting types of model membrane systems. However, there are a number of other ways to make a membrane more or less fluid (see Section 2). There are some additional general limitations when working with molecular dynamics simulations. The following part of this paragraph is a summary of the limitations presented in the manual for Gromacs [25]. Firstly, classical treatment of the molecules results in a need to correct energy terms when there are vibrations with small wavelengths. Furthermore, it is not possible to treat excited states of the molecule, they are always in the ground state. Any type of electron transport or electron transfer is not considered. However, this is not relevant for this study. Periodic boundary conditions are used to get the effect of simulating a larger system. However, if there exists long-range spatial correlations, the resulting system will be unnatural. Finally, the interactions between non-bonded particles are pair-additive. This excludes such interactions as those due to polarized atoms. The effect of this is a smaller screening effect in the system. This is noted to be somewhat mediated by the long-range cut-off of non-bonded interactions. Lastly, despite the more computationally efficient algorithms and the large capacity of supercomputers, the computational cost limits the duration and size of the simulations.

4
Methodology

4.1 System setup: Protein and Membrane

Below follows a presentation of the construction of the membrane-protein systems.

4.1.1 Preparation of All-Atom Simulations

The PAQR-2 protein structure was built using homology modeling. The transmembrane structure was based on the crystal structure of the human ADIPOR protein for residues 254-551 (PDB code: 3WXW) [6] and the respiratory complex for residues 551-581[39] (PDB code: 4HEA). A model of the first 300 residues was made using the structure of Csyb complexed with CoA-SH [40] (PDB code: 3WXY) [1]. The two models were joined by first fitting the structures using the McLachlan algorithm [41] implemented in the program ProFit [42]. The structures were then merged using Chimera [43]. The PAQR-2 mutants were made by substituting residues of the wild-type structure in Chimera. The IGLR-2 sequence was retrieved from the Entrez database [44]. TOPCONS [45] was used for prediction of transmembrane domain (TMD). The IGLR-2 structure was made in Chimera. OPM [46] was used to predict protein orientation in the membrane, and for confirmation of TMD. The force field used for the protein was Amber ff99sb-ildn [30]. Virtual hydrogens were used. The histidine residues were protonated manually at either the δ or ϵ position. The histidine residues, and their protonated state, are listed in Table 4.1. Residues of the types aspartic acid, arginine, glutamic acid and lysine were protonated according to pH 7. In addition harmonic restraints were used between the zinc ion and HIS511, HIS515 and HIS365. The equilibrium distance was set to 2 Å and the force constant to 10 kJ/Å^2

Protonated on	Histidine Residue
ϵ	110, 163, 183, 186, 193, 209, 214, 222, 227, 290, 309, 317, 336, 370, 458, 466, 475, 525, 527
δ	203, 230, 365, 511, 515, 535

Table 4.1: The histidine residues of PAQR-2 together with the position on which they are protonated.

[1] From previous work at the Department of Chemistry and Molecular Biology

The membrane and solvent structure was made using the CHARMM membrane builder [47] and minimized. Mass re-partitioning was used [24]. The force field for the lipids used was the Slipids force field [48],[49]. The TIP3P model [31] was used for the water.

4.1.2 Preparation of Coarse-Grained Simulations

The coarse-grained representations of the proteins were made from the all-atom structures using the Martinize script [50]. The protein was built into a membrane, and the system solvated, using the Insane script [51]. The martini force field version 2.2 was used [32][33].

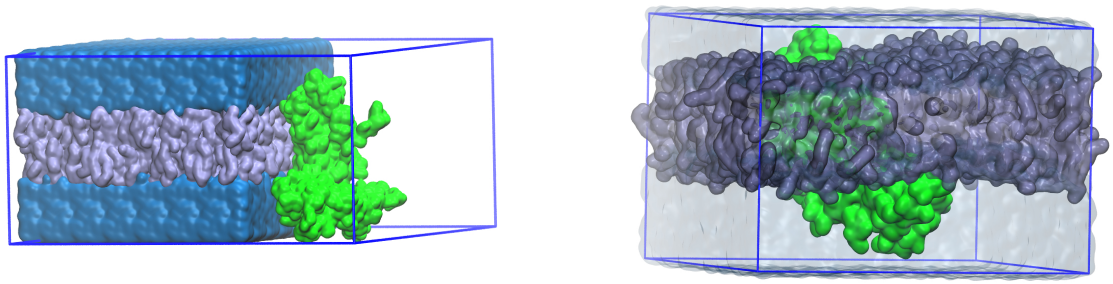
4.2 Simulations

All simulations were run using Gromacs version 5.1 [52] with periodic boundary conditions. Simulations were run using resources of the C3SE and PDC supercomputing facilities.

4.2.1 All-Atom Simulation Specifics

The protein (either PAQR-2, IGLR-2 or the PAQR-2:IGLR-2 complex) was inserted into the membrane using the protocol devised by Javanainen [53]. In this protocol the protein and membrane are placed next to each other in a box. Figure 4.1 shows a protein-membrane system before and after insertion. The protein was inserted into the membrane during a simulation where lateral pressure of 1000 bar was applied. Strong position restraints (100 kJ/mol/Å^2) were used to keep the system from deforming. During this simulation the pressure and temperature are both fixated using the Berendsen thermostat and barostat [26]. The reference pressure in the direction of the membrane normal was set to 1.013 bar. The reference temperature was set to 298 K and the coupling time constant was set to 0.5 ps. The pressure coupling constants were set to 30 ps for the transmembrane model (TMM) simulations and 40 ps for the full protein and PAQR-2:IGLR-2 complex. A time step of 2 fs was used for integration of the equations of motion. PME was used to treat long-range electrostatics. The Verlet-cut-off scheme [21] was used. The cut-off was set to 12 Å. A force-switch modifier was used. The neighbour list was updated every 20 fs. The simulation was 2 ns for the TMM simulations. It was extended to 3 ns for the full protein model and protein complex due to the increase in coupling coefficient. All bonds were constrained using LINCS [23]. The coordinates were saved every 10 ps.

Next, the membrane embedded protein system was simulated for 10 ns with position restraints of 10 kJ/mol/Å^2. The neighbour list was updated every 40 fs. The time step was set to 4 fs. The Berendsen thermostat and barostat were used. The temperature was set to 298 K with a coupling constant of 0.5 ps. The pressure was set to 1.013 bar with coupling constants of 10 ps. All bonds were constrained using the LINCS. The Verlet cut-off scheme was used. The cut-off was set to 12 Å. A force-switch modifier was used. PME was used to treat long-range electrostatics.



(a) **(b)**

Figure 4.1: (a) The system before the protein (green) has been inserted into the membrane (dark blue). The protein inserted in the membrane. The water solvent is coloured in a lighter blue.

The coordinates were saved every 10 ps.

The TMM simulations and the full protein model simulations had production runs of 500 ns. The production run of the all-atom model of the PAQR-2:IGLR-2 complex was 100 ns. Two independent production runs were made for the TMM. They were generated from the structure at the end of equilibration simulation. One of the production runs of the TMM was extended to 1500 ns. This was done for both the wild-type protein and the mutants. In all all-atom simulations the coordinates were saved every 20 ps. The time step used for the production runs was 4 fs. The pressure was fixed at 1.013 bar using the Parinello-Rahman barostat. The pressure coupling constant was set to 10 ps. The velocity-rescale thermostat was used to fixate the temperature at 298 K. The temperature coupling constant was set to 0.5 ps. Cutoff was applied to non-bonded interactions. The neighbour list was updated every 40 fs. The cut-off for both the van der Waals and Coulomb interactions was 12 Å. PME was used to treat the coulomb interaction. The cut-off was treated with force-switch modifier. All bonds were constrained using the LINCS.

4.2.2 Coarse-Grained Simulation Specifics

The production runs of the coarse-grained simulations were made using a 30 fs time step. The neighbour list was updated every 300 fs. The Verlet cut-off scheme was used for non-bonded interactions. Coulomb interactions were treated with reaction-field. The van der Waals interaction was treated with a potential-shift-verlet modifier. The cut-off for both the electrostatic and van der Waals interactions was set to 11 Å. The bonds were not constrained. The secondary structure was conserved using an elastic network [34]. The elastic bond strength was set to 5 kJ/(Å^2 mol), the lower bound was set to 5 Å, and the upper bound was set to 9 Å. The Parinello-Rahman barostat [28][29] was used with semiisotropic coupling. The pressure coupling constant was set to 10 ps and the reference pressure was set to 1.013 bar. The temperature was fixed at 298 K using the velocity-rescale thermostat. The temperature coupling constant was set to 0.5 ps. The coordinates were saved every

18 ps. The production run for the simulations for the calculation of the potential of mean force were run for 100 ns.

4.3 Protein-Protein Docking with Memdock

To investigate the interaction between IGLR-2 and PAQR-2 they were first docked and then simulated. PAQR-2 was fixated and IGLR-2 was flexible. The structures of PAQR-2 and IGLR-2 were uploaded to the Memdock web server [36]. Memdock returned the ten best scoring structures. After clustering the candidate structures and excluding structures where the TMD of IGLR-2 was in a bad agreement with the OPM prediction, four PAQR-2:IGLR-2 complex structures were chosen for further investigation.

4.4 Calculating Potential of Mean Force

The potential of mean force was calculated for four of the PAQR-2:IGLR-2 complexes returned from the docking step. Umbrella sampling and WHAM [38] was used. The biasing potential was chosen to be a harmonic potential. The domain sampled was a protein displacement of between 0 and 20 Å. 100 ns coarse-grained simulations were performed using the settings described above, see Section 4.2.2. The biasing potentials of the different simulations were spaced 1 Å apart, resulting in 21 simulations for each of the four complexes. The harmonic biasing potential has a force constant which determines the strength of the biasing potential. It was set to 20 kJ/(mol Å^2). In certain cases was increased to 40 kJ/(mol Å^2), or in one case 60 kJ/(mol Å^2), to ensure sufficient sampling. The potential of mean force was calculated the protein complex in both a DOPE and a DPPE membrane.

5

Results and Discussion

An analysis of the different models of PAQR-2 is presented below. A comparison between the structure and dynamics of PAQR-2 in different membranes is made. A second comparison is made between the wild-type and the d282n and g533r mutants. Lastly, analysis of the docking results as well as the interaction sites of the PAQR-2:IGLR-2 complex are discussed.

5.1 Simulations of PAQR-2

5.1.1 Simulations of the Transmembrane Domain of PAQR-2

The transmembrane model (TMM) simulations of PAQR-2 in a membrane required 180 lipids per leaflet. The dimensions of the initial protein structure were 66, 41 and 67 Å.

5.1.1.1 Structure Comparison

The protein structures were compared using both a contact based measure and a distance based measure.

Superpositions of the initial structure with the final snapshot of the wild-type and mutants are shown in Figure 5.1. There is no qualitative change in the overall arrangement of the helices of the protein. The bend between helices 5 and 6 is the most prominent change in structure. For instance, GLY464 moved 4.9 Å in the DOPE membrane and 7.6 Å in the DPPE membrane.

A contact comparison was made of all the 12 different simulations of PAQR-2. The contact similarity chart is plotted in Figure 5.2. The comparison and contact measure were calculated according to the method described in [54] [55]. Overall, the average protein structure in the TMM simulations of the wild-type and the PAQR-2 mutants show a contact similarity of around 85 %. There is no pattern indicating a difference in contacts between the proteins from the DOPE and DPPE respectively. The TMM simulation structures have more contact similarity with the original structure and with PAQR-2 in the PAQR-2:IGLR-2 complex simulations than the transmembrane domain (TMD) of the full protein model structures. The

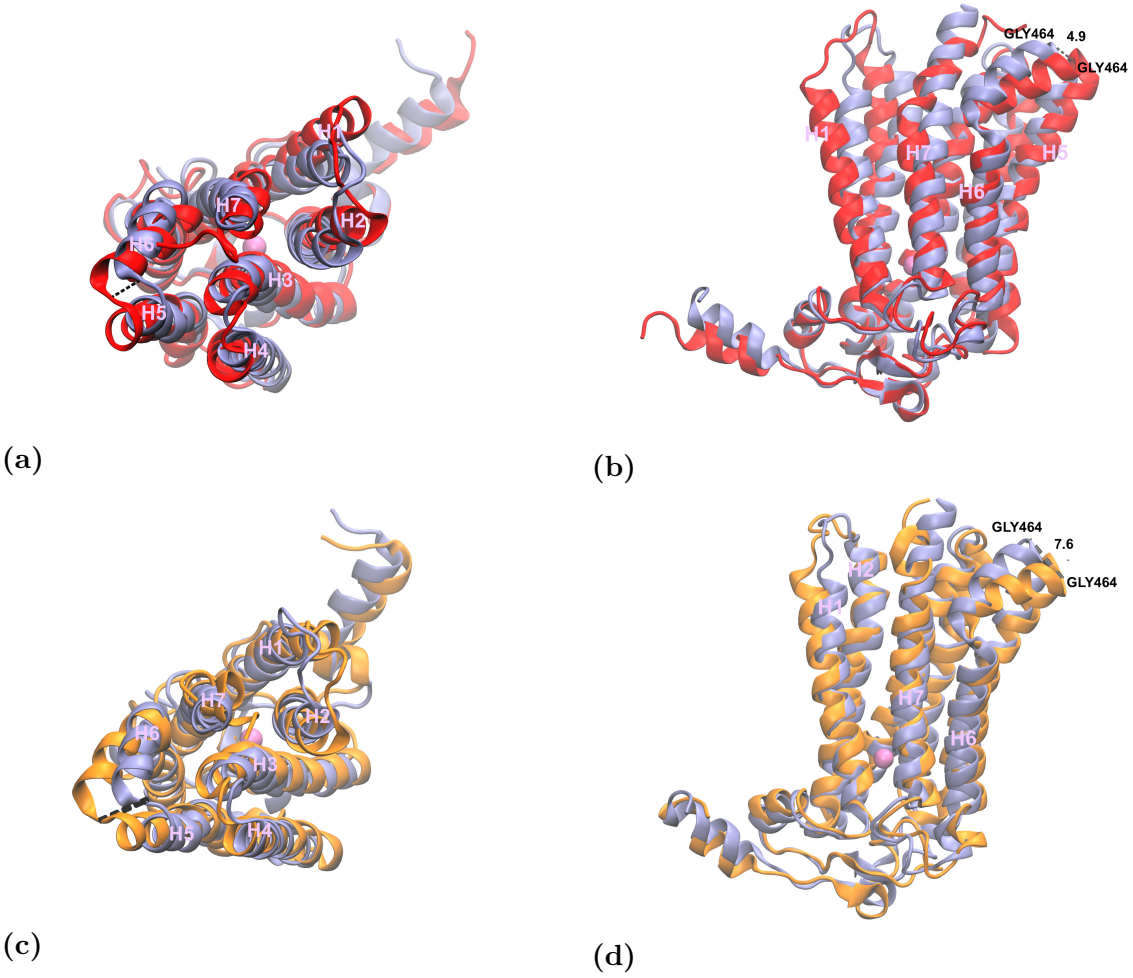


Figure 5.1: (a)-(b) A superposition of the structure of the TMD of PAQR-2 after a TMM simulation of the wild-type in a DOPE membrane (red) for 1.5 μs. Helices 5 and 6 are bent when compared with the original structure (blue). (c)-(d) Same as in (a)-(b) but here the original structure (blue) is compared with the structure in the DPPE membrane (orange). The distance between GLY464 in the two structures is shown.

g533r mutant has more contact similarity with itself in another membrane than any of the other set of TMM simulations. This indicates that the mutation alters the contact area within the structure.

The protein structures in the different simulations were also compared to the starting structure by calculating the global root mean square deviation (RMSD) for the protein backbone. The protein was fitted to the backbone of the TMD of the original structure. The average global RMSD for the different simulations is reported in Table 5.1. The global RMDS for the TMM model simulations are between 2.7 and 4.0 Å. The protein structures of the wild-type and the d282n mutant simulated in the DOPE membrane have a larger global RMSD than that in the DPPE membrane. For the g533r mutant the relationship is reversed.

Protein Trajectory	Time Averaged global RMSD (Å)
wt TMM DPPE	2.7 ± 0.4
wt TMM DOPE	3.3 ± 0.5
d282n TMM DPPE	2.7 ± 0.2
d282n TMM DOPE	3.2 ± 0.5
g533r TMM DPPE	4.0 ± 0.8
g533r TMM DOPE	3.2 ± 0.4
wt DPPE	5.7 ± 0.2
wt DOPE	4.8 ± 0.2
d282n DPPE	5.1 ± 0.3
d282n DOPE	4.8 ± 0.3
wt (complex) DPPE	3.5 ± 0.1
wt (complex) DOPE	3.3 ± 0.2

Table 5.1: The time average of the global RMSD of the protein backbone after a least square fit to the backbone of the TMD helices.

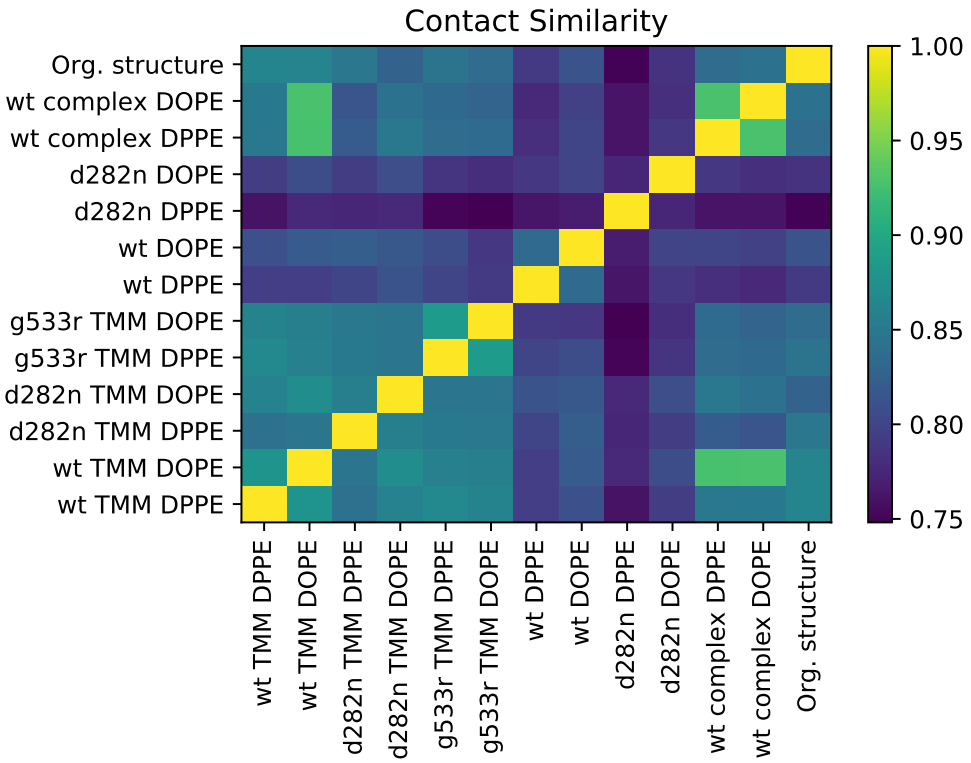
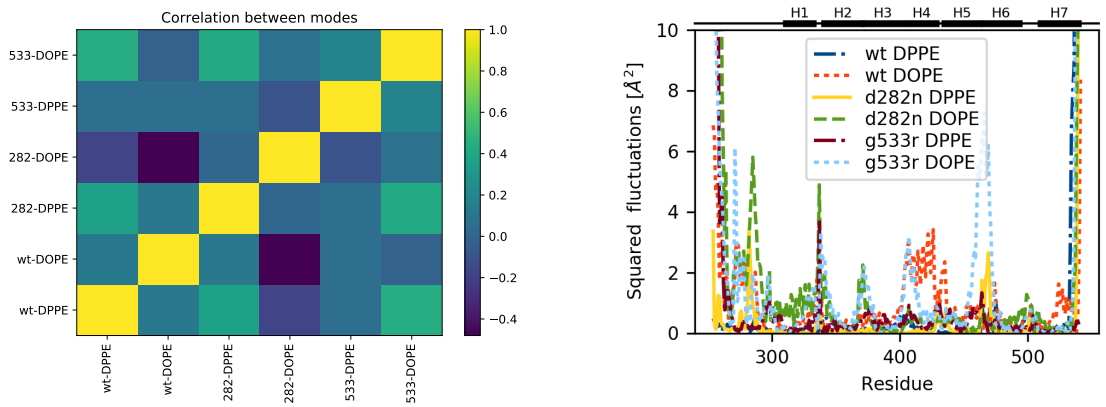


Figure 5.2: A contact comparison between the PAQR-2 TMD from the different sets of simulations. A contact similarity of 1 means identical structures. This is shown along the diagonal.

5.1.1.2 Protein Dynamics

To compare the protein dynamics in different membranes and between the wild-type and mutants, a principal component analysis (PCA) was performed using ProDy and NMWiz in VMD [56]. Using PCA the essential dynamics of the protein is summarized by different modes of movement. A mode describes a correlated set movements of the residues of the protein. The mode that made out the majority of the fluctuations during one simulation, the first principal component (PC), was compared with the first PC of the other simulations. The PCs made out between 26 and 51 % of the total fluctuations, see Table 5.2. In Figure 5.3 the first principal modes of the six 1.5 μs simulations are compared. The left panel (a) shows a correlation map and in (b) the right panel shows the squared fluctuation as a function of the residues for the first PCs. The fluctuations can primarily be observed at the residues close to either end of the membrane. The correlation map showed no strong correlation in movement between the simulations. Most notably, this analysis shows that the mutation promotes a larger fluctuation between helices 6 and 7 (light blue dotted peak). This is likely due to the type of mutation made at residue 533, where a glycine is replaced with an arginine. The arginine is more bulky which will likely cause the residues near it to become more perturbed when it moves.



(a) **(b)**

Figure 5.3: (a) Correlation chart between the first principal modes of the 1.5 μs TMM simulations of PAQR-2. (b) The figure shows the characteristic fluctuation profile of each of the first modes from the TMD simulations.

5.1.1.3 Membrane Properties

The membrane properties investigated here showed no apparent difference between the simulations of the PAQR-2 mutants and wild-type PAQR-2. The order parameter of the DOPE membrane was uniform throughout the membrane, see right panel of Figure 5.4. The average chain order parameter was 0.15. The membrane did not appear perturbed in any way by the protein. The same can be stated when studying the thickness plot of the DOPE membrane, see right panel in Figure 5.5. The thickness of the membrane was approximately 39 Å. The difference in the shape of the cutout of the protein (shown as white in Figures 5.4 and 5.5) in the different membranes suggests that the N-terminal part of PAQR-2 is more immersed in the DPPE membrane than the DOPE membrane.

Conversely, the DPPE membrane was more perturbed in the region surrounding the protein. The thickness around the protein was in the range between 35 and 40 Å, while the bulk thickness was above 47 Å, see the left side panel of Figure 5.5. The same perturbation could be observed in the plot of the order parameter in the DPPE membrane. It appears as if the DPPE lipids and not the protein, arranged

Protein Trajectory	Percentage (%)
wt TMM DPPE	45
wt TMM DOPE	37
d282n TMM DPPE	26
d282n TMM DOPE	43
g533r TMM DPPE	40
g533r TMM DOPE	51

Table 5.2: The percentage of the total fluctuations made out by the first principal component in the TMM simulations.

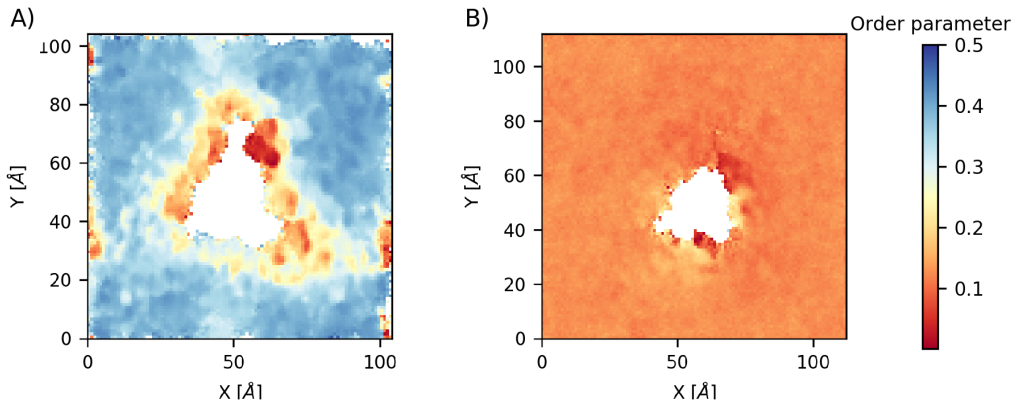


Figure 5.4: The average chain order parameter of the DPPE membrane (A) and DOPE membrane (B) from the TMM simulations of the wild-type. The white cutout is where PAQR-2 sits. A completely ordered membrane has an order parameter of 1 and a completely disordered membrane has an order parameter of 0.

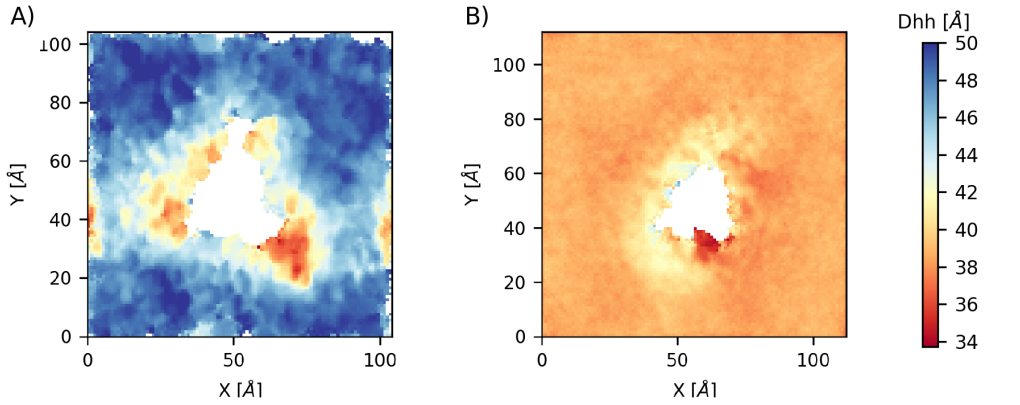


Figure 5.5: The membrane thickness of a DPPE membrane (A) and DOPE membrane (B) from the TMM simulations of the wild-type. The white cutout is where PAQR-2 sits.

themselves so to avoid a hydrophobic mismatch. This may explain why there was no detectable large-scale structural change in the protein when comparing the structure of the protein in the DOPE membrane and the DPPE membrane. This type of compression of lipids close to the protein has been observed elsewhere [57].

It is also possible that PAQR-2 diffuses to domains with a hydrophobic match. Hydrophobic mismatch can lead to a clustering of proteins and the creation of domains of protein clusters or a particular lipid composition [58]. One of the weaknesses of the uniform membrane model is that it cannot form lipid domains. Moving forward, simulations of PAQR-2 in membranes with more than one type of lipid investigate possibilities of domain formation surrounding PAQR-2.

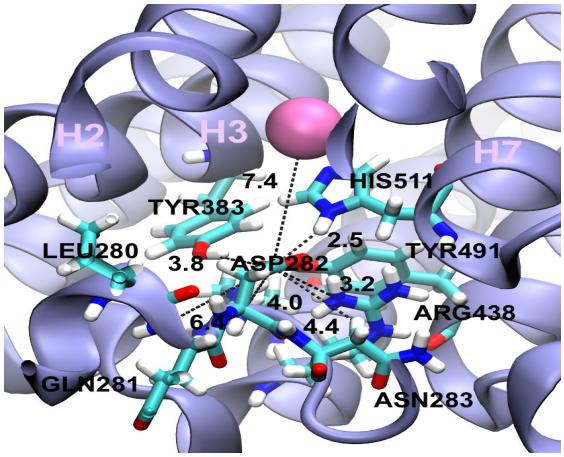
5.1.1.4 Interactions of the d282n PAQR-2 Mutant

In Figure 5.6 the final snapshots of the residues in vicinity of ASP282/ASN282 are shown (selected to be at most 2.5 Å from residue 282). In the initial structure ASP282 is elevated up into the bottom cavity of the TMD, see (a). However, it is still over 7 Å from the zinc ion, and likely does not take part in the coordination of the zinc site. When comparing the original interaction site with the final structures from the simulations of the wild-type protein in (b) and (c), the tyrosine residues are no longer close to ASP282. In (a) HIS511 which coordinates the zinc ion is close enough to participate as a proton donor in hydrogen bond between it and one of the oxygen of ASP282. The structures of the wild-type protein have conserved the close contact with ARG438, and the neighbouring residues GLN281 and ASN283. The TMD structure in the DPPE membrane is also still close to HIS511. New contacts have formed, including other arginine residues on helix 6 and 7. The close contacts of the d282n mutant do not overlap with the contacts of the original structure, apart from GLN281 and ASN283. Most contacts are between residue 282 and other residues in link between helix 1 and the cytosolic domain. In Appendix A the hydrogen bond occupancy is listed in Table A.1, together with a plot of the number of hydrogen bonds in the simulations over time in Figures A.1. There were many brief interactions but few that lasted long in either the wild-type or the d282n mutant simulations. The original structure of the wild-type only had one hydrogen bond with HIS511.

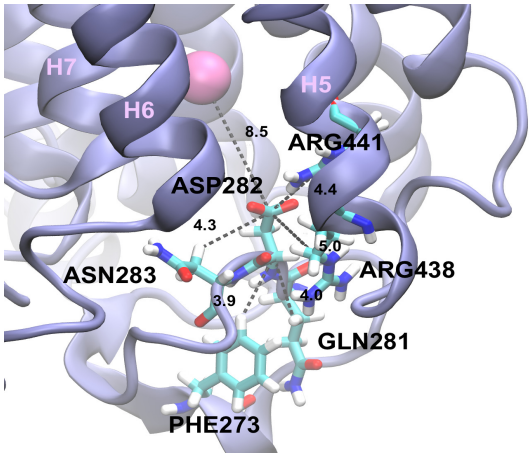
The location and type of substitution at residue 282 that renders PAQR-2 non-functional suggests that it is involved in some activity at the zinc site. When changing the residue from the aspartic acid to asparagnine, the residue no longer carries a negative charge at pH 7. Therefore, it will be less attracted towards the zinc site. When analyzing the distance between the carbon of the carbonyl group on the side chain of the residue and the zinc ion, there was no difference in the change in the distance over time, between the wild-type protein and the d282n protein. The distance over time is plotted in the left panel of Figure 5.7. In all simulations, the initial distance is 7 Å, and time average is 10 Å. In the right side panel residues within 3 Å from the zinc ion are shown together with residue 282.

5.1.1.5 Interactions of the g533r PAQR-2 Mutant

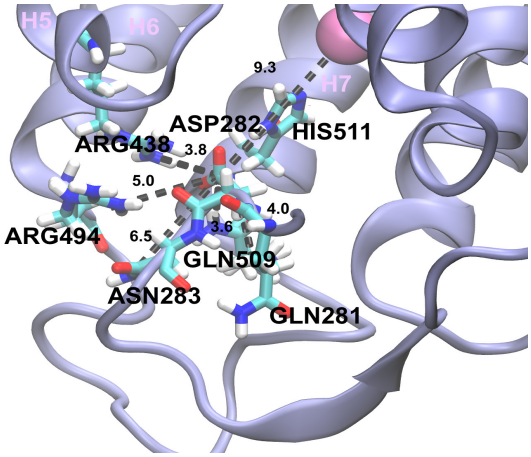
The mutation in residue 533 creates hydrogen bonds between helix 7 and helices 1 and 2. The interactions take place close to the extra-cellular part of the membrane. When glycine is replaced with an arginine residue, the residue becomes more bulky and positively charged. The conformation of residue 533 and residues in its immediate vicinity are shown in Figure 5.8. In one wild-type protein simulation (DOPE) there are only a pair of residues within 2.5 Å of it on helix 7, see (a). On the other hand, the arginine residue which replaced the glycine shares hydrogen bonds with GLU343 and GLN338 in the bend between helix 1 and 2. These interactions are likely to favour a close positioning of helix 1,2 and 7. If the mechanism of PAQR-2 requires some flexibility of the helical arrangement, to for example let other molecules enter the space between the helices, this would become more unlikely due to the



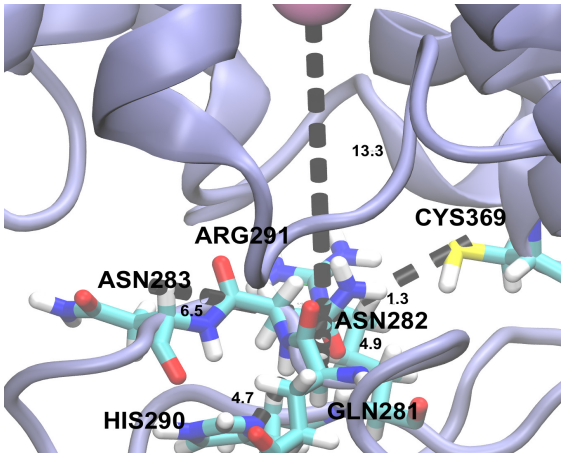
(a) Original PAQR-2 structure



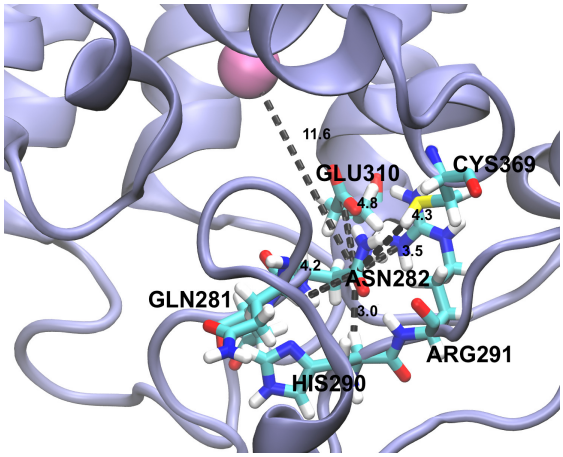
(b) wild-type in DOPE membrane



(c) wild-type in DPPE membrane

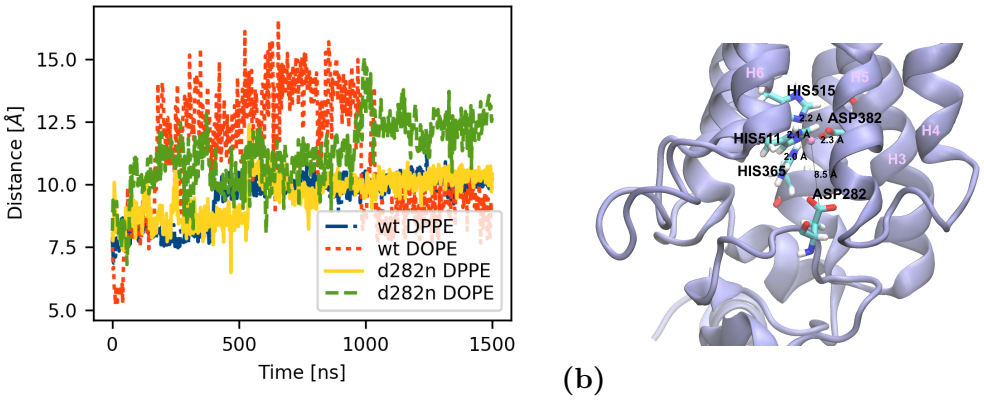


(d) d282n mutant in DOPE membrane



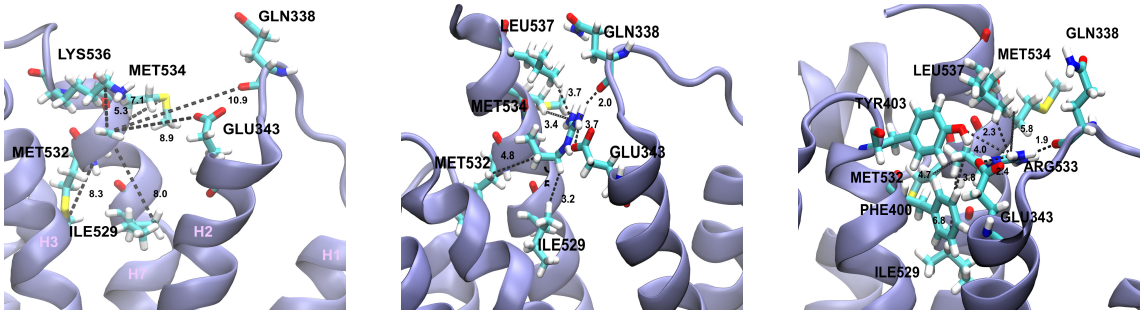
(e) d282n mutant in DPPE membrane

Figure 5.6: The final snapshot from the TMM simulations - highlighting the vicinity of ASP282/ASN282. All distances are labeled in Ångström.



(a) (b)

Figure 5.7: (a) The distance between the zinc ion and carbonyl of residue 282 in the wild-type and d282n mutant. (b) The zinc site and residue 282 of wild-type PAQR-2 after a TMD simulation in a DOPE membrane. All distances are labeled in Ångström.



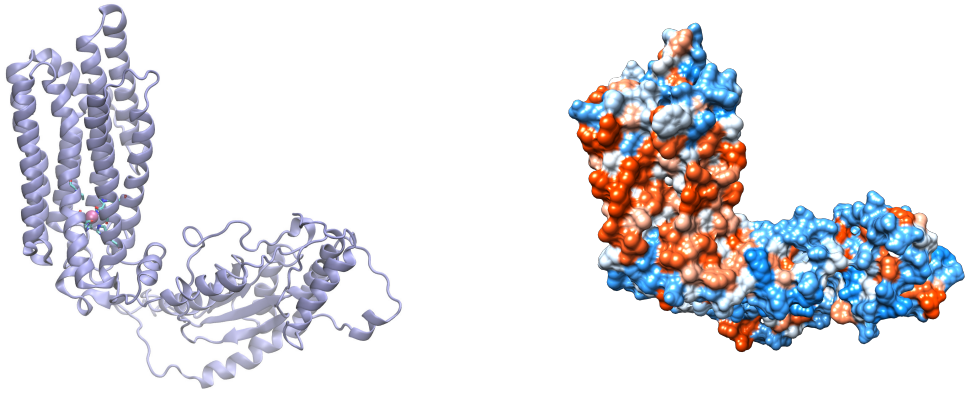
(a) (b) (c)

Figure 5.8: The structure close to residue 533 in the wild-type structure after simulation in DOPE (a), and of the g533r PAQR-2 mutant in a DPPE membrane (b) and DOPE membrane (c).

increased stability owing to the hydrogen bonds introduced with the mutation. In Section A.3 of the Appendix, the hydrogen bond occupancy is listed in Table A.3. In addition, the number of hydrogen bonds over time is plotted together with the length of the hydrogen bond as a function of time for the two most present hydrogen bonds in Figure A.3. It can be seen that there is a short interval in the beginning of the simulations were the residues move closer to create hydrogen bonds which then remain throughout the rest of the 1.5 μs.

5.1.1.6 Reproducibility and Convergence of Trajectories

A second set of 500 ns simulations were made to check for other behaviours not found in the first set. To check the reproducibility of the results, a comparison was made with the first set of TMM simulations. The membrane adaptation was visible also in the second set. A contact similarity analysis showed similarities above 0.8,



(a) (b)

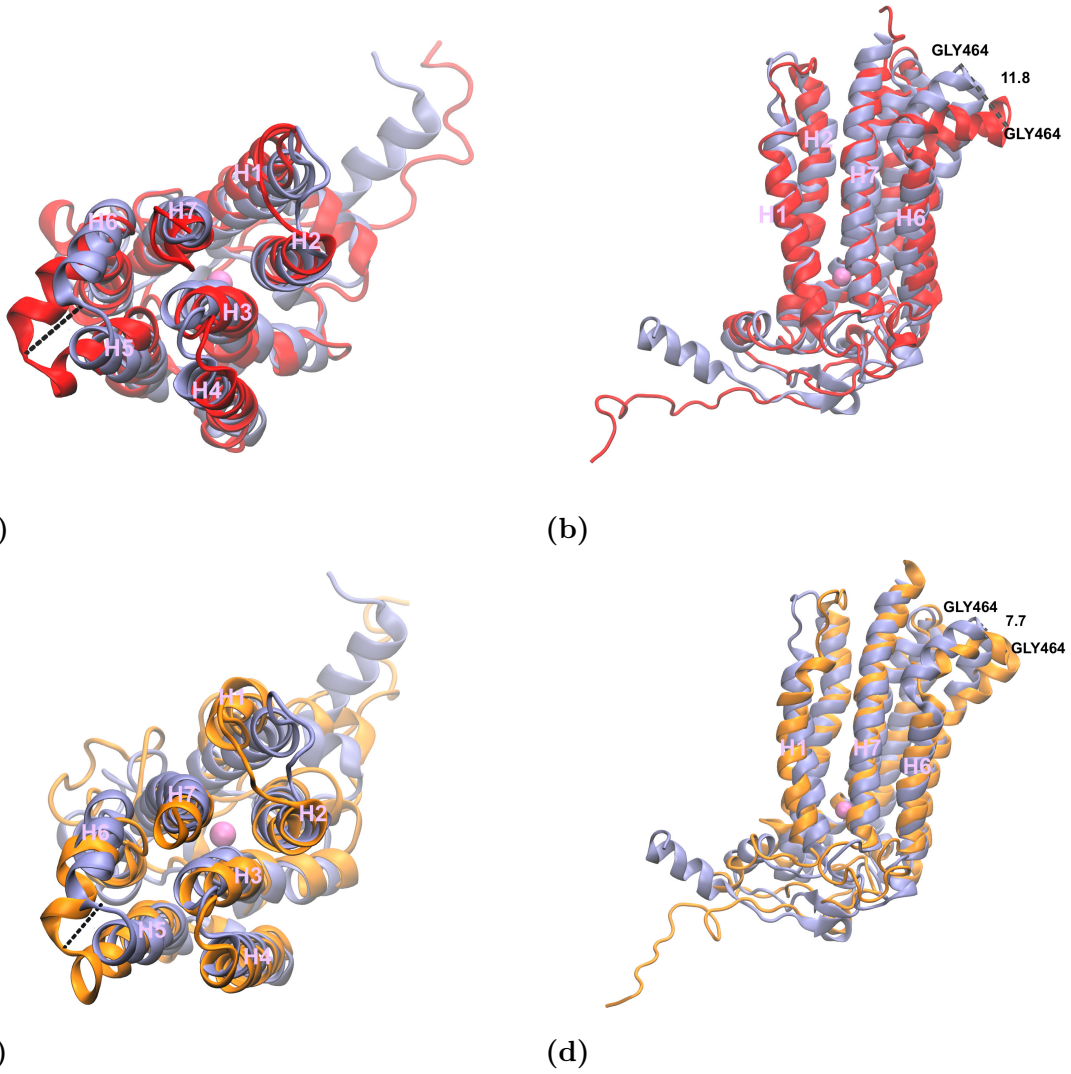
Figure 5.9: (a) The model of the entire PAQR-2. (b) PAQR-2 colored according to a hydrophobicity scale going from red (hydrophobic) to blue (hydrophilic). All distances are labeled in Ångström.

both when compared with the first set of simulations and in comparisons within the group. There was no significant similarity between independent runs of the same system either. This was also the case when comparing the protein dynamics. There was no significant difference between the two groups, and no significant similarity between two independent runs of the same system. More independent simulations are needed to confirm the results but were not carried out due to the time constraints of the project.

The first set of TMM simulations were extended twice, adding a total of 1000 ns to the original simulation time. The extension was made to try to observe slower motions. Extensions were motivated by results from the software Encore [59] that indicated that the protein was still exploring new conformations at 500 ns and 1000 ns. Using Encore one can study the rate of convergence of the trajectory. The upper limit for the duration of the simulation was set to 1500 ns so that the time spent on the simulations would not exceed what seemed reasonable within the project frame.

5.1.2 Full Protein Simulations

The full PAQR-2 protein model was created to investigate the structural effect of including the cytosolic domain as well as the C-terminal to the protein structure. The full protein can be seen in the right sde panel of Figure 5.9. As the d282n mutation sits in the intermediate part between the TMD and the cytosolic domain, simulations of the full protein model with the d282n mutation were also made. The dimensions of the full protein model were 93, 81 and 64 Å. This required 280 lipids per leaflet for both the DOPE and the DPPE membrane. In Figure 5.9, the full protein structure is shown, in (a) with secondary structure and in (b) coloured according to the hydrophobic scale. Figure 5.10 shows the TMD of PAQR-2 after structures from full protein simulation were fitted to the starting structure. The top of helices 5 and 6 have bent just as they did in the TMM simulations.



(a) **(b)**

(c) **(d)**

Figure 5.10: (a) A superposition of the structure of the TMD of PAQR-2 after a full protein model simulation in a DOPE membrane (red). Helices 5 and 6 are bent when compared with the crystal structure. GLY464 has moved 11.5 Å. (b) Same as in (a) but here the original structure is compared with the structure in the DPPE membrane (orange). GLY464 has moved 7.5 Å.

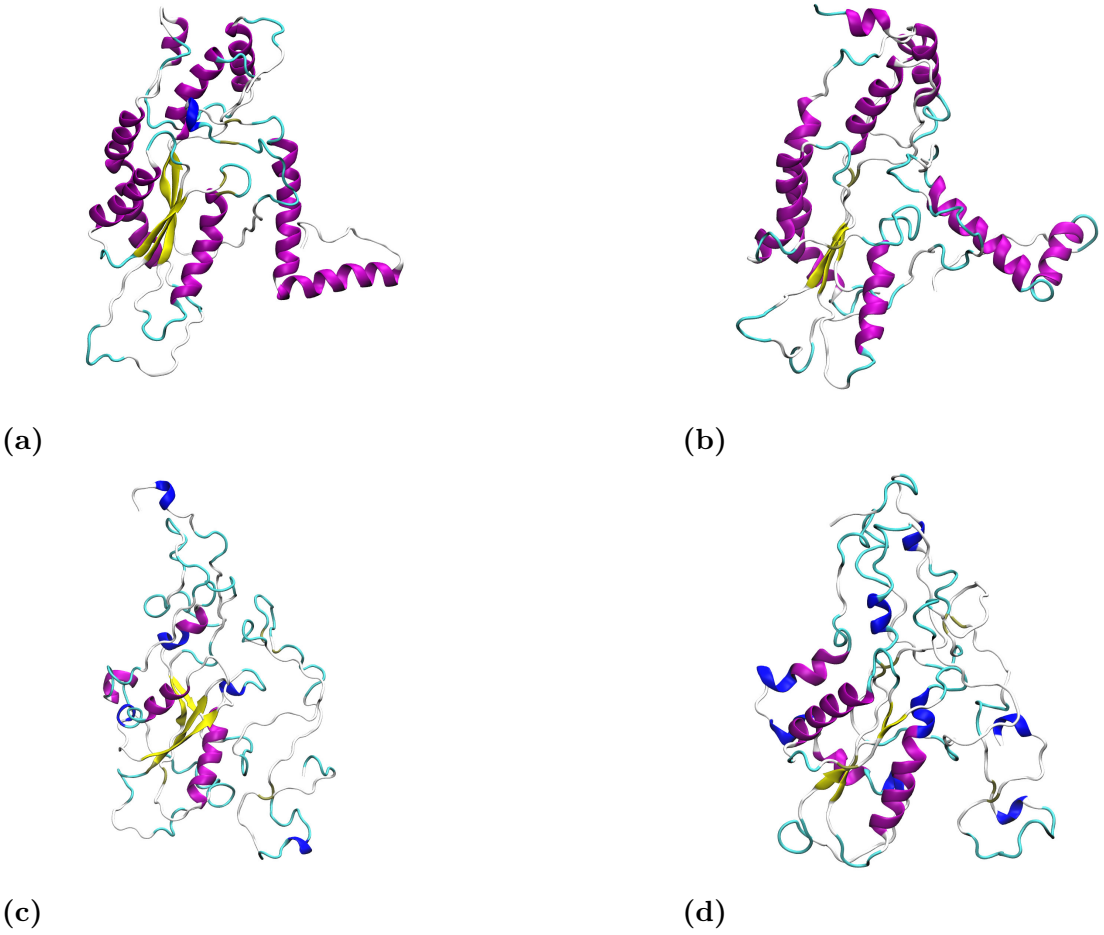
5.1.2.1 Structure Comparison

The contact similarity of the TMD of PAQR-2 in the full protein model simulations was shown in Figure 5.2. The TMD of PAQR-2 from the simulations of the full protein model show less similarity both within their group and with the original structure than the TMD simulations. The contact similarity was 75 to 80 %, both when comparing contacts for the TMD of the wild-type and the d282n mutant. This was also the case when comparing with the contacts of the TMM simulations. Some of their difference to the structures of the TMM simulations is likely due to the loss of the small helix at the N-terminal, see (a) and (b) of Figure 5.10. However, this does not explain why the TMD of the full protein model show less similarity within their group than the other groups. As can be seen in Table 5.1, the global RMSD was also higher for these simulations when compared with the original structure. Again, a likely contribution for this might be the loss of the helix before the N-terminal.

5.1.2.2 Secondary Structure Loss in Cytosolic Domain

Figure 5.11 shows the secondary structure of the cytosolic part of PAQR-2 in the beginning (a) and at the end of the simulations of PAQR-2 in the DOPE membrane (c) and DPPE membrane (d). There is a significant loss of the secondary structure of the cytosolic domain. This can more easily be seen in Figure 5.12 where the secondary structure of the cytosolic domain has been plotted as a function of residue. There it is clear that the α-helices at the beginning and end of the sequence of the protein, have disappeared in the simulations of PAQR-2 in the membranes.

To understand whether this was a problem of the protein model, or due to some effect of the membrane system, a simulation of the cytosolic part of the protein in a water box was performed. The structure at the end of this simulation can be seen in (b) of Figure 5.11. Its secondary structure is also shown in Figure 5.12. The secondary structure, and in particular the helices which disappeared in the membrane simulations remained at the end of the simulation of the cytosolic domain in the water box. Therefore, the model of the cytosolic domain is not necessarily wrong. However, it is possible that the way that the cytosolic domain is bent towards the membrane is not accurate. The cytosolic domain carries a large negative charge, as can be seen in the right side panel of Figure 5.9. Therefore, it may be more stable when accesing more of the solvent instead of the lipid head group which is a zwitter ion. Moving forward, a PMF calculation using the angle between the membrane and the cytosolic domain as the reaction coordinate could be made for optimization. It is also possible that the force field description at the protein-lipid-solvent interface does not produce proper behaviour. Typically, the force field parameteriztion is not done in systems with both membrane and proteins present. This can cause inaccuracies at the protein-lipid-solvent interface, as was observed for the Amber14sb/Slipids combination in [60].



(a) **(b)**

(c) **(d)**

Figure 5.11: The figure shows the structure of the first 300 residues of PAQR-2 after various treatment. In (a) the starting structure is showed. (b) shows the structure after 100 ns simulation in water. (c) shows the structure after a 100 ns simulation when connected to the TMD of PAQR-2 in a DOPE membrane. (d) The same as in (c) but instead of DOPE a DPPE membrane was used.

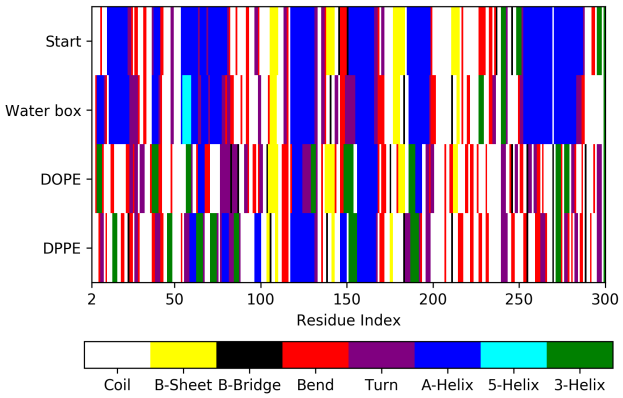
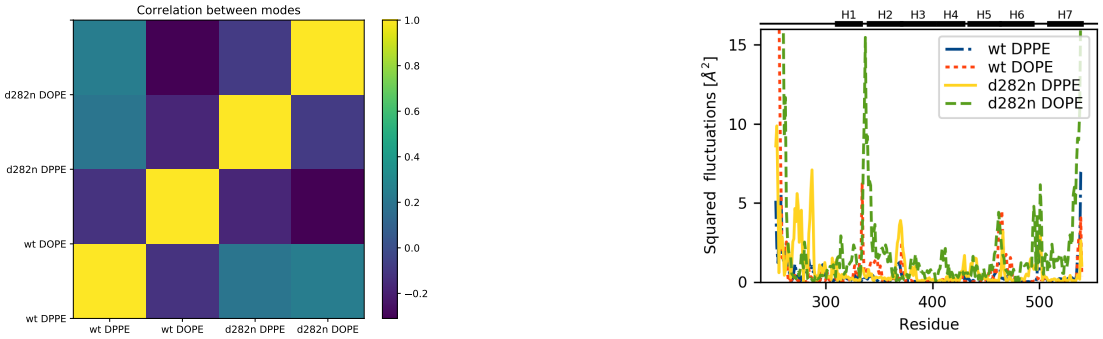


Figure 5.12: The secondary structure as a function of residue from different structures of the first 300 residues of PAQR-2. The loss of helices in the structure from the full protein simulations, here called DOPE and DPPE, is visible.



(a) **(b)**

Figure 5.13: (a) Correlation chart between the first principal modes of the TMD of the full protein simulations (b) The figure shows the characteristic fluctuation profile of each of the first modes of the TMD in the full protein simulations.

5.1.2.3 Protein Dynamics

A PCA was done for the TMD in the full protein simulations. No significant correlations between the PCs were identified, see the left side panel in Figure 5.13. The right side panel shows the fluctuation profile along the TMD sequence for the different full protein simulations. Both wild-type and d282n mutant profiles in DOPE membrane show large fluctuations in the bend between helix one and two. In addition, as for the TMM simulations, the largest fluctuations are coming from the ends of the TMD. The percentages of fluctuations made out by the PCs are listed in Table 5.3

Protein Trajectory	Percentage (%)
wt DPPE	29
wt DOPE	43
d282n DPPE	42
d282n DOPE	60

Table 5.3: The percentage of the total fluctuations made out by the first principal component in the full protein simulations.

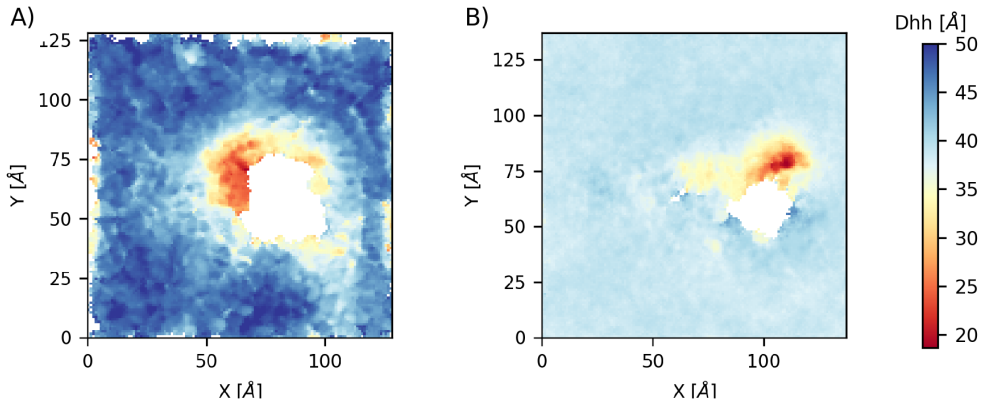


Figure 5.14: The membrane thickness of a DPPE membrane (A) and DOPE membrane (B) from the full protein simulations of the wild-type. The white cutout is where PAQR-2 sits.

5.1.2.4 Membrane Properties

The thickness of the DOPE and DPPE membrane are plotted in Figure 5.14. The values of the bulk thicknesses are similar to those seen during the TMM simulations, see Figure 5.5. However, the cytosolic domain creates a local area in the DOPE membrane where the thickness is significantly lowered, indicating that the cytosolic domain is pushing on the membrane. The DPPE membrane is also perturbed in the area around the protein. Also, the cutout for the protein is larger in the DPPE membrane than the DOPE membrane. This is likely because more of the cytosolic domain is inside the membrane.

The effect of the full protein on the membrane is also seen when plotting the order parameter, see Figure 5.15. While the bulk values again are similar to those seen in the TMM simulations, see Figure 5.4, the local area around the protein is more disordered.

5.1.2.5 The d282n Mutant

Differences between the wild-type protein and the PAQR-2 d282n mutant are observed when analyzing the zinc site of the simulations. Firstly, the distance between ASP282 and the zinc ion, in the simulations of wild-type protein, decrease with time. What is more, the fluctuations of the distance are also smaller for the wild-type. On

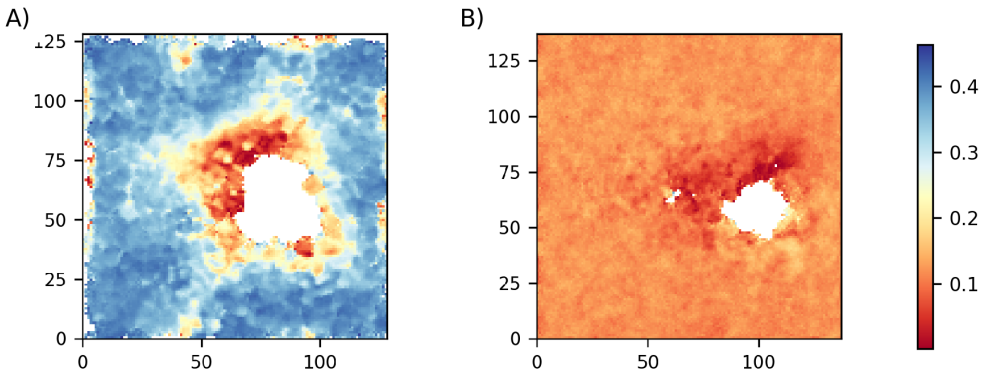
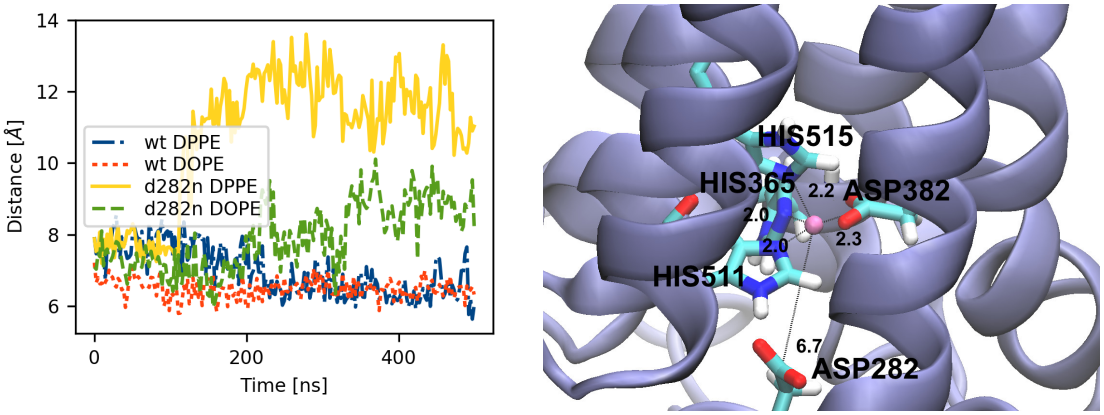


Figure 5.15: The average chain order parameter of the DPPE membrane (A) and DOPE membrane (B) from the full protein simulations of the wild-type. The white cutout is where PAQR-2 sits.

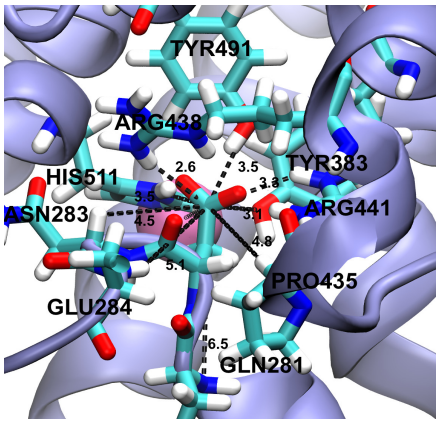


(a) **(b)**

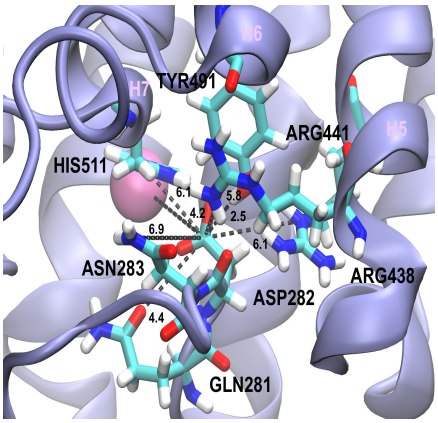
Figure 5.16: (a) The distance between the carbonyl of ASP282/ASN282 and the zinc ion as a function of time. (b) The zinc site at the end of a wild-type full model simulation. The distances are labeled in Ångström.

the other hand, the distance between ASN282 and the zinc ion in the d282n mutant increases over time, with larger fluctuations. This is not what was observed in the analysis of the zinc site for the TMM simulations. Even when comparing the results from only the first 500 ns of the TMM simulations there was no apparent trend in the time evolution of the distance between residue 282 and the zinc ion. Important to note however is the fact that while the distance in the wild-type protein appear to have reached its equilibrium value in both the DOPE and DPPE membrane, the large fluctuations in the PAQR-2 mutant in the DOPE membrane have not reached a plateau at the end of the simulation.

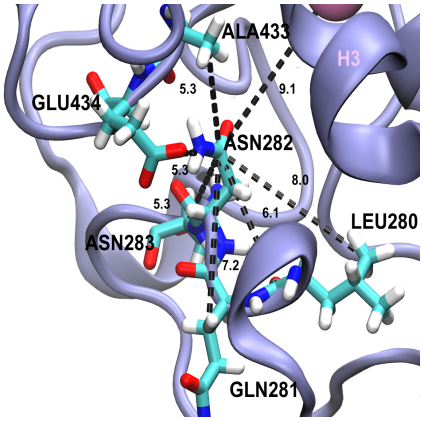
In Figure 5.17, final snapshots of the structure surrounding residue 282 are shown for the protein in different membranes, and for both the wild-type and d282n PAQR-2 mutant. Only residues that have some atom within 2.5 Å of residue 282 are shown.



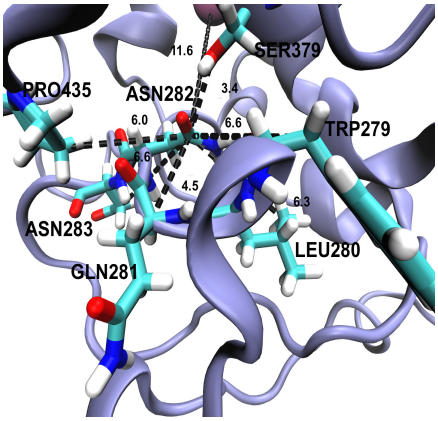
(a) wild-type PAQR-2 in DOPE membrane



(b) wild-type PAQR-2 in DPPE membrane



(c) d282n mutant in DOPE membrane



(d) d282n mutant in DPPE membrane.

Figure 5.17: Final snapshots of the structure in the vicinity of ASP282/ASN282. Distances are labeled in Ångström.

In (a) the wild-type protein in DOPE is shown to have conserved all but one (leucine 280) of the contacts found in the original structure which was based on the ADIPOR crystal structure. In addition, two new contacts with PRO435 and ARG441 are present. In this structure it would be possible for the ASP282 to make a hydrogen bond with residues such as HIS511, ARG441 or TYR491 and TYR383. Looking at the structure in (b), which is the structure of the PAQR-2 after a simulation in a DPPE membrane, one can locate the same contacts as for the structure in the DOPE membrane with the exception of PRO435 and TYR383. Conversely, the d282n PAQR-2 mutant structure only has conserved the interaction with LEU280 from the original structure in both membranes (in the DPPE membrane the PRO435 is a conserved as well). The new contacts that instead are present are with more hydrophobic residues. There is no longer a close contact to HIS511. The hydrogen bonds occupancy for ASP282/ASN282 is listen in Table A.2 in section A.2. The number of hydrogen bonds as a function of time is also plotted in Figure A.2. The number is more or less constant in the wild-type simulations while it decreases in the d282n mutant simulations.

5.2 The PAQR-2:IGLR-2 Complex

This section in devoted to answering the last of the aims of the study concerning how IGLR-2 and PAQR-2 interact.

Coarse-grained simulations and all-atom simulations of IGLR-2 were made to confirm the predicted orientation of IGLR-2 in the different membrane. The predicted TMD of IGLR-2 had 21 residues. After some tests it was found that when simulating the TMD with twelve extra residues on each side of the membrane the angles were in relatively good agreement with the OPM prediction of 38 ± 4 degrees tilt.

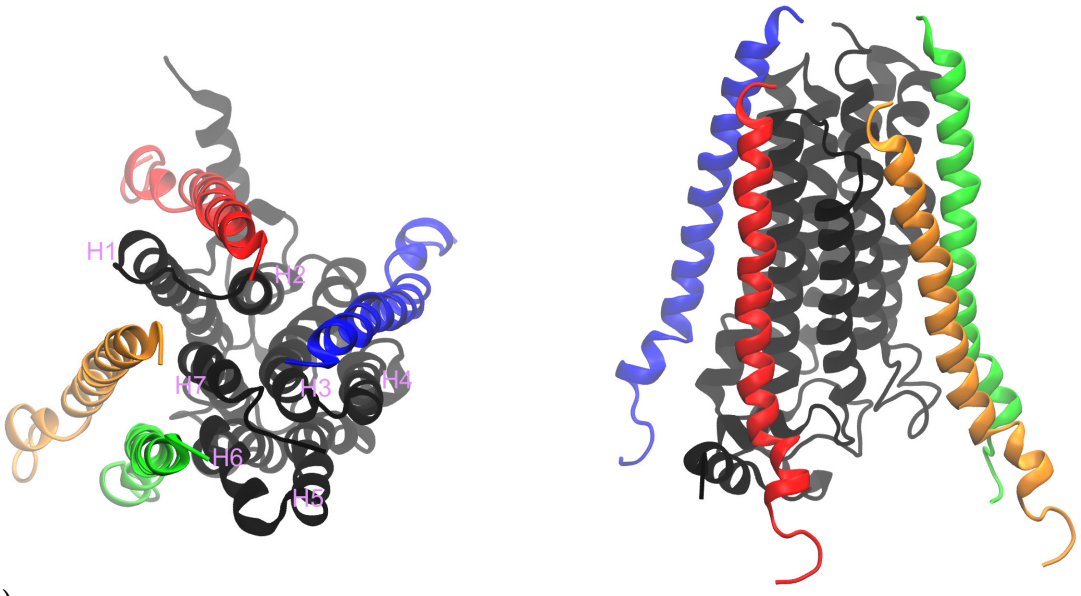
5.2.1 Evaluation of the Memdock Results

The protein-protein docking resulted in four possible candidate solutions. Although already scored by the docking program, the potential mean force calculation was performed for confirmation. Figure 5.18 shows PAQR-2 along with the four possible interaction sites of IGLR-2. The docking with the highest score in Memdock, was also the docking which had the most negative potential of mean force, see Figure 5.19. The PMF-curves in DOPE and DPPE differ both in their range and also in their shape. In the DOPE membrane (a), the PMF curves have a minimums between 15 and 20 Å. This is the characteristic shape of PMF curves. However, in the DPPE membrane the PMFs are at the most negative values when the proteins are as close as possible, given the experiment. The proteins are pushed together. The pressure in the DPPE membrane also affects the values of the PMF as the DPPE membrane is less able to adapt and arrange around the proteins when they are separated. As Memdock and the PMF calculation in the DOPE membrane both predicted complex with rank 1 as the most favourable, it was deemed the most likely candidate structure.

5.2.2 The PAQR-2:IGLR-2 complex simulations

The PAQR-2:IGLR-2 complex was simulated in a membrane with 240 lipids. The dimensions of the complex were 75 Å and 42 Å in the membrane plane, and 66 Å along the membrane normal.

The contact similarity between the PAQR-2 TMD structures from the PAQR-2:IGLR-2 simulations show a large contact similarity (>90 %) with eachother. This may be due to the fact that these structures were simulated for only 100 ns, compared with 500 ns for the full protein simulations and 1500 ns for the TMM simulations. However, the structures are more similar to each other than to the original structure. It could also mean that the IGLR-2 stabilizes the PAQR-2 structure. Their contact similarity with the structures from the TMD simulations are between 80 and 85 %.



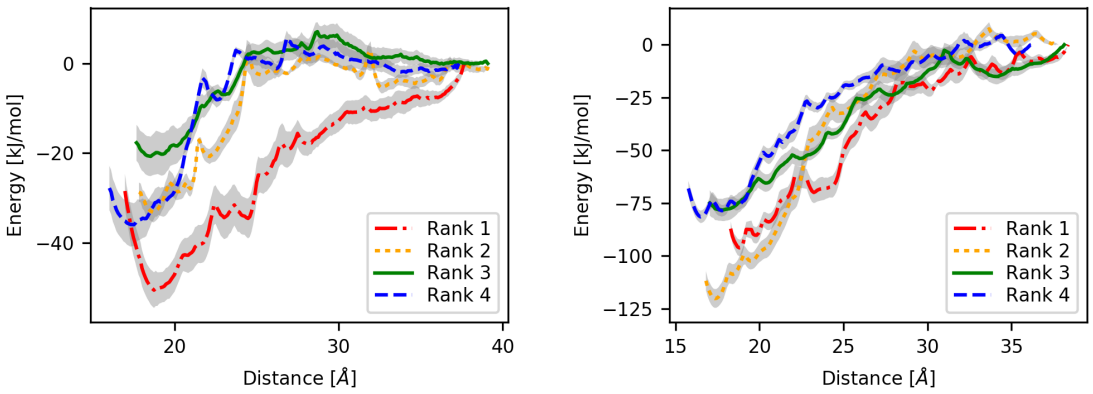
(a) (b)

Figure 5.18: PAQR-2 (black) with the four selected docking results. The highest scoring (red), second (orange), third (green) and fourth (blue). (a) A top view with the PAQR-2 helices numbered from 1 to 7 in pink. (b) A side view.

5.2.3 Interaction Sites between PAQR-2 and IGLR-2

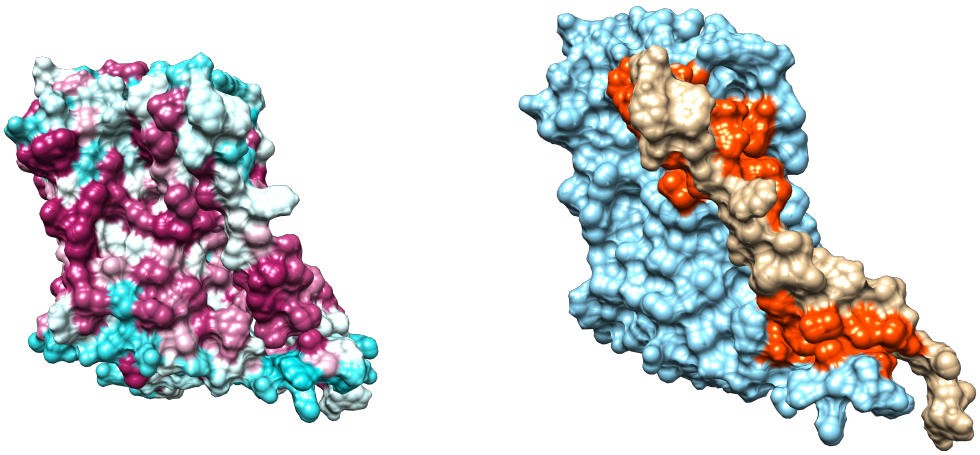
Further simulations of the highest scoring docking result were performed in a DOPE and DPPE membrane. Analysis of the interaction sites revealed that the contacts of interaction are some forms of weak interaction, either van der Waals or dipole - induced dipole. The interactions are primarily between IGLR-2 and helix 2 of PAQR-2. The analysis only showed a pair of hydrogen bonds, close to the N-terminal of IGLR-2. There were 80 more contacts detected in the complex that had been simulated in the DPPE membrane than in the DOPE membrane. The interaction sites on PAQR-2 are located on the first, second, and seventh helix. In addition there were contacts between IGLR-2 and the start of the cytosolic part of the protein. The contacts between the proteins can be seen in the left panel of Figure 5.20. The right panel shows the PAQR-2:IGLR-2 complex colored according to the hydrophobic scale.

One caveat of the docking result presented here is that it was done only using the TMD of IGLR-2 and PAQR-2. This was done as the study was primarily concerned with the TMD structures, and also for computational efficiency. The interaction between IGLR-2 and PAQR-2 close to the N-terminal implies that IGLR-2 would be in contact with the cytosolic domain, especially as it is arranged in the present model. However, as discussed above, the orientation of the cytosolic domain will probably be updated after an optimization. It would be interesting to see the arrangements of the proteins when both are modelled in full form.



(a) **(b)**

Figure 5.19: The potential of mean force as a function of the distance between the center of mass of PAQR-2 and IGLR-2. The curves for docking structure one to four is plotted. The grey shaded area shows the estimated error. (a) The PMF curves in a DOPE membrane. (b) The PMF curves of a DPPE membrane.



(a) **(b)**

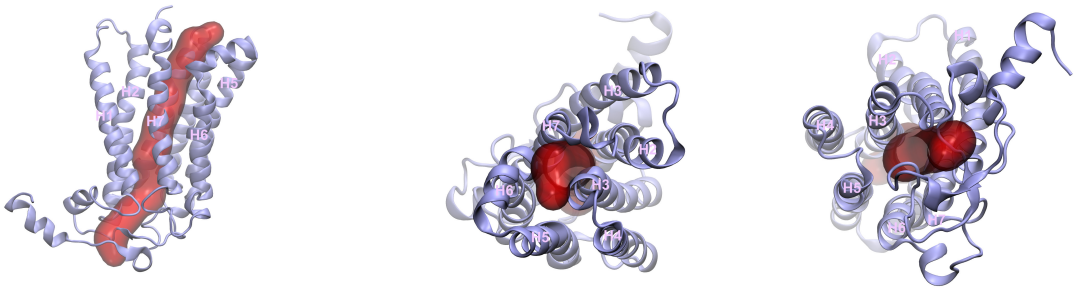
Figure 5.20: (a) The PAQR-2:IGLR-2 complex colored using the hydrophobic scale, going from purple (hydrophobic) to blue (hydrophilic). (b) The PAQR-2:IGLR-2 complex where PAQR-2 is colored in blue and IGLR-2 is colored in beige. The residues making out the contact between the molecules have been colored orange.

6
Conclusion

Simulations of two types of models of PAQR-2 were made in two types of membranes to study the response to a hydrophobic mismatch. In addition, simulations of two PAQR-2 mutants (d282n and g533r) were made. The structure of PAQR-2:IGLR-2 complex was found through protein-protein docking and PMF calculations. No evidence was found that suggests that the signalling mechanism of PAQR-2 is driven by a structural change of the PAQR-2 during simulations due to hydrophobic mismatch. Instead structural changes of the membrane compensated for the mismatch. Essential dynamics analysis revealed no strong correlation between the PCs of different simulations. Contact analysis showed that the contacts of the TMM model are more similar to the initial structure than the TMD of the full protein model. On the other hand close interactions of ASP282 are better preserved when using the full protein model than the transmembrane model. However, the full protein model suffers from the loss of structure of the cytosolic domain and therefore optimization of the model is needed. The g533r mutation introduces hydrogen bonds between residues on helix 7 and residues on the bend between helix 1 and 2, making a close conformation of the helices more stable. Finally, protein-protein docking of IGLR-2 and PAQR-2 revealed an interaction of the PAQR-2:IGLR:2 complex made out of primarily weak interactions.

6.1 Future Outlook

While some interesting observations could be made in the simulations, further simulations are needed to confirm the results presented here. In addition, the structure of the full protein model needs to be optimized so that the secondary structure of the cytosolic domain is stable in membrane simulations. When a functioning model is obtained, building a complete PAQR-2:IGLR-2 complex could provide additional insights to their interactions and mechanims. Furthermore, in the recent publication by Brooks et al. [3], a lipid molecule was docked to a channel in the ADIPOR protein and simulations were performed. Among other things, the study showed an interesting rearrangement of residues around the zinc site and lipid molecule which suggested a mechanism of the ADIPOR receptor as a hydrolase. While this type of docking was beyond the scope of the study presented here, some preliminary analysis of PAQR-2 revealed a possible pore structure. This can be seen in Figure 6.1. The analysis was done using PoreWalker [61] The next steps in the investigation of the mechanism of PAQR-2 could benefit from repeating the study of Brooks et al. on PAQR-2. The results of such a study would contribute to an interesting comparison.



(a) **(b)** **(c)**

Figure 6.1: View of the TMD of PAQR-2 from the side (a), above (b) and below (c). The identified pore volume is colored in red.

Bibliography

[1] I. D. Federation. [Online]. Available: `http://www.idf.org/about-diabetes/facts-figures`.

[2] M. Laakso and H. Cederberg, "Glucose control in diabetes: Which target level to aim for?: Review: Glucose control in diabetes", English, *Journal of Internal Medicine*, vol. 272, no. 1, pp. 1–12, 2012.

[3] I. Vasiliauskaité-Brooks, R. Sounier, P. Rochaix, G. Bellot, M. Fortier, L. Hoh Francois an De Colibus, C. Bechara, E. M. Saied, C. Arenz, C. Leyrat, and S. Granier, "Structural insights into adiponectin receptors suggest ceramidase activity", *NATURE*, vol. 000, pp. 1–4, 2017.

[4] K. Hotta, T. Funahashi, N. L. Bodkin, H. K. Ortmeyer, Y. Arita, B. C. Hansen, and Y. Matsuzawa, "Circulating concentrations of the adipocyte protein adiponectin are decreased in parallel with reduced insulin sensitivity during the progression to type 2 diabetes in rhesus monkeys", English, *Diabetes*, vol. 50, no. 5, pp. 1126–1133, 2001.

[5] K. E. Davis, M. R. Wade, N. Halberg, B. M. Barth, K. Sun, J. M. Rutkowski, R. A. Miller, Z. V. Wang, W. L. Holland, P. E. Scherer, H. H. Bui, J. T. Brozinick, V. M. Tenorio, M.-S. Kuo, B. T. Bikman, S. A. Summers, M. J. Birnbaum, and B. B. Zhang, "Receptor-mediated activation of ceramidase activity initiates the pleiotropic actions of adiponectin", *Nature Medicine*, vol. 17, no. 1, pp. 55–63, 2011.

[6] H. Tanabe, Y. Fujii, M. Okada-Iwabu, M. Iwabu, Y. Nakamura, T. Hosaka, K. Motoyama, M. Ikeda, M. Wakiyama, T. Terada, N. Ohsawa, M. Hato, S. Ogasawara, T. Hino, T. Murata, S. Iwata, K. Hirata, Y. Kawano, M. Yamamoto, T. Kimura-Someya, M. Shirouzu, T. Yamauchi, T. Kadowaki, and S. Yokoyama, "Crystal structures of the human adiponectin receptors", *Nature*, vol. 520, no. 7547, pp. 312–316, 2015.

[7] E. Svensk, M. Ståhlman, C. H. Andersson, M. Johansson, J. Borén, and M. Pilon, "Paqr-2 regulates fatty acid desaturation during cold adaptation in *C. elegans*", *PLoS Genetics*, vol. 9, no. 9, pp. 1–15, Sep. 2013.

[8] E. Svensk, R. Devkota, M. Ståhlman, P. Ranji, M. Rauthan, F. Magnusson, S. Hammarsten, M. Johansson, J. Borén, and M. Pilon, "Caenorhabditis elegans paqr-2 and iglr-2 protect against glucose toxicity by modulating membrane lipid composition", *PLoS Genetics*, vol. 12, no. 4, e1005982, 2016.

[9] D. L. Riddle, T. Blumenthal, B. J. Meyer, and J. R. Priess, *C. elegans ii*. Plymouth: Cold Spring Harbor Laboratory Press, 1998, vol. No. 33.

[10] NCBI Bookshelf - Free (e-book collection), *Wormbook: The online review of c. elegans biology*. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK19662/ (visited on 08/14/2017).

[11] E. Culetto and D. B. Sattelle, "A role for caenorhabditis elegans in understanding the function and interactions of human disease genes", *Human molecular genetics*, vol. 9, no. 6, pp. 869–877, 2000.

[12] E. Lindahl and M. S. Sansom, "Membrane proteins: Molecular dynamics simulations", English, *Current Opinion in Structural Biology*, vol. 18, no. 4, pp. 425–431, 2008.

[13] L. E. Cybulski, M. Martín, M. C. Mansilla, A. Fernéndez, and D. de Mendoza, "Membrane thickness cue for cold sensing in a bacterium", *Current Biology*, vol. 20, no. 17, pp. 1539–1544, 2010.

[14] R. Covino, S. Ballweg, C. Stordeur, J. B. Michaelis, K. Puth, F. Wernig, A. Bahrami, A. M. Ernst, G. Hummer, and R. Ernst, "A eukaryotic sensor for membrane lipid saturation", English, *Biophysical Journal*, vol. 112, no. 3, 508–509a, 2017.

[15] D. de Mendoza, "Temperature sensing by membranes", English, *Annual review of microbiology*, vol. 68, no. 1, pp. 101–116, 2014.

[16] O. G. Mouritsen and L. A. Bagatolli, *Life - as a matter of fat : Lipids in s membrane biophysics perspective*, 2nd ed. Springer, 2016.

[17] M. Witting and P. Schmitt-Kopplin, "The *Caenorhabditis elegans* lipidome: A primer for lipid analysis in *Caenorhabditis elegans*", *Archives of biochemistry and biophysics*, vol. 589, pp. 27–37, 2016.

[18] K. Satouchi, K. Hirano, M. Sakaguchi, H. Takehara, and F. Matsuura, "Phospholipids from the free-living nematode caenorhabditis elegans", *Lipids*, vol. 28, no. 9, p. 837, 1993.

[19] Avanti Polar Lipids, Inc., *Phase transition temperatures for glycerophospholipids*. [Online]. Available: https://avantilipids.com/tech-support/physical-properties/phase-transition-temps/ (visited on 08/30/2017).

[20] W. L. Jorgensen, J. Tirado-Rives, and B. J. Berne, "Potential energy functions for atomic-level simulations of water and organic and biomolecular systems", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 19, pp. 6665–6670, 2005.

[21] S. Páll and B. Hess, "A flexible algorithm for calculating pair interactions on simd architectures", *Computer Physics Communications*, vol. 184, no. 12, pp. 2641–2650, 2013.

[22] A. R. Leach, *Molecular modelling: Principles and applications*, 2nd ed. Pearson education limited, 2001.

[23] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "Lincs: A linear constraint solver for molecular simulations", *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[24] K. A. Feenstra, B. Hess, and H. J. C. Berendsen, "Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems", *Journal of Computational Chemistry*, vol. 20, no. 8, pp. 786–798, 1999.

[25] Gromacs, *Gromacs manual*, Visited 24-April-2017, 2016.3. [Online]. Available: http://manual.gromacs.org/documentation/2016.3/index.html#.

[26] H. Berendsen, J. Postma, W. van Gunsteren, A. DiNola, and J. Haak, "Molecular dynamics with coupling to an external bath", *Journal of Chemical Physics*, vol. 81, no. 8, pp. 3684–3690, 1984.

[27] D. Donadio, M. Parrinello, and G. Bussi, "Canonical sampling through velocity rescaling", *Journal of Chemical Physics*, vol. 126, no. 1, 014101-014101-7, 2007;2008;

[28] S. Nosé and M. L. Klein, "Constant pressure molecular dynamics for molecular systems", *Molecular Physics*, vol. 50, no. 5, pp. 1055–1076, 1983.

[29] M. Parrinello and A. Rahman, "Polymorphic transitions in single crystals: A new molecular dynamics method", *Journal of Applied Physics*, vol. 52, no. 12, pp. 7182–7190, 1981.

[30] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the amber ff99sb protein force field", *Proteins*, vol. 78, no. 8, 1950–NA, 2010.

[31] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water", English, *The Journal of Chemical Physics*, vol. 79, no. 2, p. 926, 1983.

[32] S. J. Marrink, A. H. de Vries, and A. E. Mark, "Coarse grained model for semiquantitative lipid simulations", *The Journal of Physical Chemistry B*, vol. 108, no. 2, pp. 750–760, 2004.

[33] S. J. Marrink and D. P. Tieleman, "Perspective on the martini model", *Chemical Society Reviews*, vol. 42, no. 16, pp. 6801–6822, 2013.

[34] X. Periole, M. Cavalli, S.-J. Marrink, and M. A. Ceruso, "Combining an elastic network with a coarse-grained molecular force field: Structure, dynamics, and intermolecular recognition", *Journal of Chemical Theory and Computation*, vol. 5, no. 9, pp. 2531–2543, 2009.

[35] A. A. Kaczor, J. Selent, F. Sanz, and M. Pastor, "Modeling complexes of transmembrane proteins: Systematic analysis of proteinprotein docking tools", *Molecular Informatics*, vol. 32, no. 8, p. 717, 2013.

[36] N. Hurwitz, D. Schneidman-Duhovny, and H. J. Wolfson, "Memdock: An α-helical membrane protein docking algorithm", *Bioinformatics (Oxford, England)*, vol. 32, no. 16, pp. 2444–2450, 2016.

[37] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman, "The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method", *Journal of Computational Chemistry*, vol. 13, no. 8, pp. 1011–1021, 1992.

[38] A. Grossfield, *Wham: An implementation of the weighted histogram analysis method*, version 2.0.9. [Online]. Available: `http://membrane.urmc.rochester.edu/content/wham/`.

[39] R. Baradaran, J. M. Berrisford, G. S. Minhas, and L. A. Sazanov, "Crystal structure of the entire respiratory complex i", *Nature*, vol. 494, no. 7438, p. 443, 2013.

[40] T. Mori, D. Yang, T. Matsui, M. Hashimoto, H. Morita, I. Fujii, and I. Abe, "Structural basis for the formation of acylalkylpyrones from two β-ketoacyl units by the fungal type iii polyketide synthase csyb", *The Journal of biological chemistry*, vol. 290, no. 8, pp. 5214–5225, 2015.

[41] A. D. McLachlan, "Rapid comparison of protein structures", *Acta Crystallographica Section A*, vol. 38, no. 6, pp. 871–873, Nov. 1982.

[42] A. C. R. Martin, *Profit*. [Online]. Available: `http://www.bioinf.org.uk/software/profit/`.

[43] E. F. Pettersen, T. D. Goddard, C. C. Huang, and G. S. Couch, "Ucsf chimera–a visualization system for exploratory research and analysis", *Journal of Computational Chemistry*, vol. 25, no. 13, pp. 1605–1612, 2004.

[44] Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information, *Entrez (internet)*, 2004. [Online]. Available: `https://www.ncbi.nlm.nih.gov/Class/MLACourse/Original8Hour/Entrez/`.

[45] K. D. Tsirigos, C. Peters, N. Shu, L. Käll, and A. Elofsson, "The topcons web server for consensus prediction of membrane protein topology and signal peptides", *Nucleic acids research*, vol. 43, no. W1, W401–W407, 2015.

[46] M. A. Lomize, A. L. Lomize, I. D. Pogozheva, and H. I. Mosberg, "Opm: Orientations of proteins in membranes database", *Bioinformatics*, vol. 22, no. 5, pp. 623–625, 2006.

[47] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Dávila-Contreras, Y. Qi, J. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im, "Charmm-gui membrane builder toward realistic biological membrane simulations", *Journal of Computational Chemistry*, vol. 35, no. 27, pp. 1997–2004, 2014.

[48] J. P. M. Jämbeck and A. P. Lyubartsev, "An extension and further validation of an all-atomistic force field for biological membranes", *Journal of chemical theory and computation*, vol. 8, no. 8, p. 2938, 2012.

[49] ——, "Derivation and systematic validation of a refined all-atom force field for phosphatidylcholine lipids", *The journal of physical chemistry. B*, vol. 116, no. 10, p. 3164, 2012.

[50] D. H. de Jong, G. Singh, W. F. w. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schafer, X. Periole, D. P. Tieleman, S. J. Marrink, and L. V. Schäfer, "Improved parameters for the martini coarse-grained protein force field", *Journal of Chemical Theory and Computation*, vol. 9, no. 1, pp. 687–697, 2013.

[51] T. A. Wassenaar, H. I. Ingolfsson, R. A. Boeckmann, D. P. Tieleman, and S. J. Marrink, "Computational lipidomics with insane: A versatile tool for generating custom membranes for molecular simulations", *Journal of Chemical Theory and Computation*, vol. 11, no. 5, pp. 2144–2155, 2015.

[52] M. Abraham, B. Hess, E. Lindahl, and D. van der Spoel, *Gromacs*, version 5.1. [Online]. Available: `http://www.gromacs.org`.

[53] M. Javanainen, "Universal method for embedding proteins into complex lipid bilayers for molecular dynamics simulations", *Journal of Chemical Theory and Computation*, vol. 10, pp. 2577–2582, 2014.

[54] I. Kufareva and R. Abagyan, "Methods of protein structure comparison", in *Homology Modeling: Methods and Protocols*, A. J. W. Orry and R. Abagyan, Eds. Totowa, NJ: Humana Press, 2012, pp. 231–257, ISBN: 978-1-61779-588-6. [Online]. Available: `https://doi.org/10.1007/978-1-61779-588-6_10`.

[55] B. Marsden and R. Abagyan, "Sad—a normalized structural alignment database: Improving sequence–structure alignments", *Bioinformatics*, vol. 20, no. 15, pp. 2333–2344, 2004.

[56] A. Bakan, L. M. Meireles, and I. Bahar, "Prody: Protein dynamics inferred from theory and experiments", *Bioinformatics (Oxford, England)*, vol. 27, no. 11, pp. 1575–1577, 2011.

[57] T. Kim, K. I. Lee, P. Morris, R. W. Pastor, O. S. Andersen, and W. Im, "Influence of hydrophobic mismatch on structures and dynamics of gramicidin a and lipid bilayers", *Biophysical Journal*, vol. 102, no. 7, pp. 1551–1560, 2012.

[58] D. Milovanovic, A. Honigmann, S. Koike, F. Gottfert, G. Pahler, M. Junius, S. Mullar, U. Diederichsen, A. Janshoff, H. Grubmuller, H. J. Risselada, C. Eggeling, S. W. Hell, G. v. d. Bogaart, and R. Jahn, "Hydrophobic mismatch sorts snare proteins into distinct membrane domains", *Nature Communications*, vol. 6, p. 5984, 2015.

[59] M. Tiberti, E. Papaleo, T. Bengtsen, W. Boomsma, and K. Lindorff-Larsen, "Encore: Software for quantitative ensemble comparison", *PLOS Computational Biology*, vol. 11, pp. 1–16, Oct. 2015.

[60] A. Sandoval-Perez, K. Pluhackova, and R. A. Böckmann, "Critical comparison of biomembrane force fields: Protein–lipid interactions at the membrane interface", *Journal of Chemical Theory and Computation*, vol. 13, no. 5, pp. 2310–2321, 2017.

[61] M. Pellegrini-Calace, T. Maiwald, and J. M. Thornton, "Porewalker: A novel tool for the identification and characterization of channels in transmembrane proteins from their three-dimensional structure", *PLOS Computational Biology*, vol. 5, pp. 1–16, Jul. 2009.

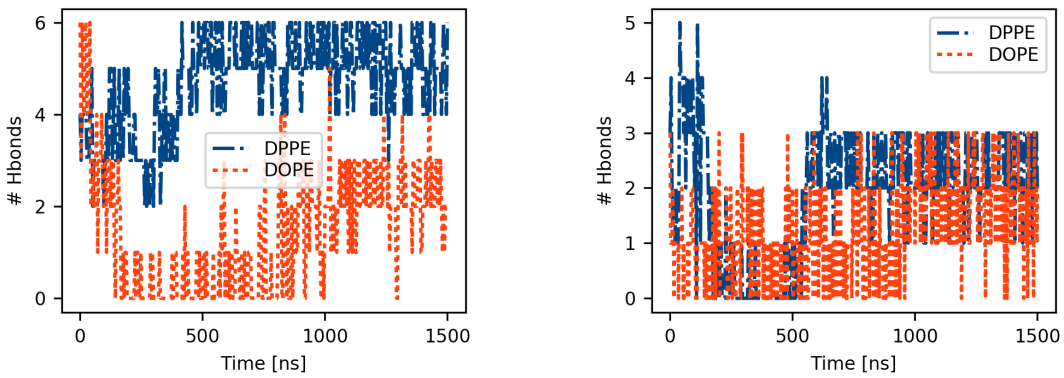
A

Appendix I

A.1 Hydrogen bonds of ASP282/ASN282 in TMM simulations

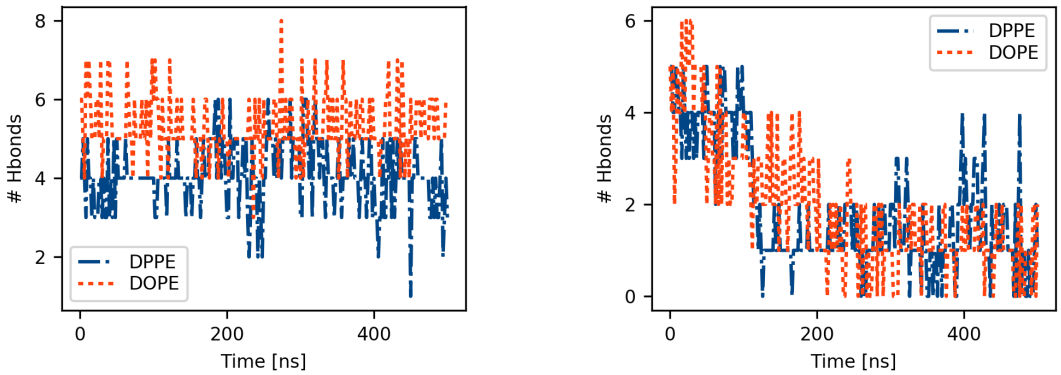
	Hydrogen bond occupancy (%)			
Residue	wt DPPE	wt DOPE	d282n DPPE	d282n DOPE
ARG441	-	31.5	-	-
ARG438	96.74	17.7	2.2	-
TYR383	1.90	2.70	-	-
HIS511	68.80	0.80	0.30	-
GLU310	-	3.10	10.7	1.40
ASN283	58.9	1.30	-	0.40
LEU280	2.40	1.10	-	-
ARG494	6.90	-	2.4	-
TYR491	6.50	-	0.80	-
HIS290	-	-	14.6	14.5
GLN509	-	-	0.50	-
GLU497	-	-	2.50	-
ARG291	-	-	29.3	10.31
GLY289	-	-	-	3.00
THR309	-	-	3.40	

Table A.1: The hydrogen bond occupancy of the hydrogen bonds, of ASP282/ASN282 in the TMM simulations, that were present for more than 5 ns (0.3 %) in the simulation.



(a) **(b)**

Figure A.1: (a) Number of hydrogen bonds of ASP282 in the TMM simulation of the wild-type. (b) Number of hydrogen bonds of ASN282 in the TMM simulation of the d282n mutant.



(a) **(b)**

Figure A.2: (a) Number of hydrogen bonds of ASP282 in the wild-type full protein simulations. (b) Number of hydrogen bonds of ASN282 in the d282n mutant full protein simulations.

A.2 Hydrogen Bonds of ASP282/ASN282 in Full Protein Simulations

	Hydrogen bond occupancy (%)			
Residue	wt DPPE	wt DOPE	d282n DPPE	d282n DOPE
TYR491	97.5	80.39	-	3.10
HIS511	61.4	90.3	-	-
ARG438	61.2	99.5	15.1	-
ASN283	1.20	-	-	-
TYR383	2.70	-	-	-
ARG441	47.8	89.6	5.40	3.50
TYR383	-	-	11.7	20.7
GLU310	-	-	29.9	34.8
TRP279	-	-	2.80	-
PRO435	-	-	-	1.40
ALA433	-	-	-	5.3

Table A.2: The hydrogen bond occupancy of the hydrogen bonds of ASP282/ASN282 in the full protein simulations of the wild-type that were present for more than 5 ns (1 %) in the simulation.

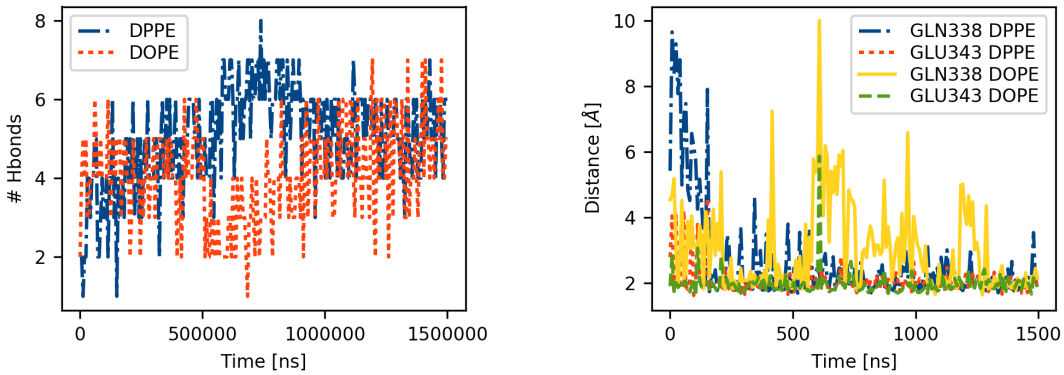
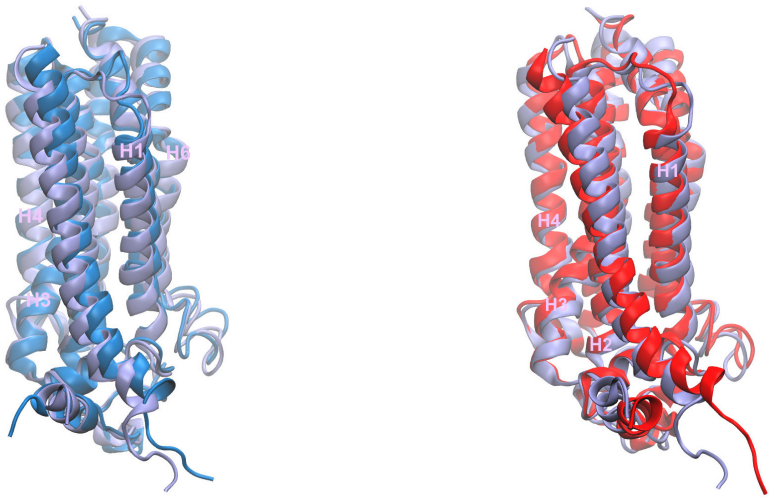


Figure A.3: (a) Number of hydrogen bonds of the g533r mutant in the TMM simulations. Distance cut-off between donor and acceptor was set to 3.5 Å. (b) Length of the most present hydrogen bonds with GLU343 and GLN338 in the simulations. The main chain hydrogen bonds were primarily with ILE529 and LEU537.

A.3 Hydrogen Bonds of ARG533

	Hydrogen bond occupancy (%)	
Residue	DPPE	DOPE
GLU343	55.4	70.2
GLN338	19.3	10.8
ASP337	7.00	-
LEU537	1.90	1.50
MET339	1.90	2.60
ILE529	1.50	0.40
LYS536	-	2.40
PHE400	-	0.40
TYR347	-	0.30

Table A.3: The hydrogen bond occupancy of the hydrogen bonds of the side chain of ARG533 that were present for more than 5 ns (0.3 %) in the simulation.



(a) **(b)**

Figure A.4: (a) Overlay of the original PAQR-2:IGLR-2 complex (blue) and after a 100 ns simulation in a DPPE membrane. (b) same as in (a) but in a DOPE membrane.

A.4 Distance between Center of Mass in PAQR-2:IGLR-2 complex

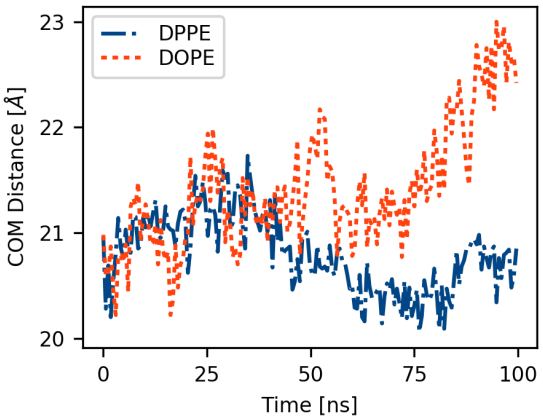


Figure A.5: (a) Time evolution of distance between the center of mass of PAQR-2 and IGLR-2 in the PAQR-2:IGLR-2 complex.