Studies of IAPP from Type-II Diabetes and the M2 Proton Channel from Influenza with Novel Methods for Molecular Dynamics Simulations of Membrane Peptides

By

Michael J. Robertson

Under the Advisement of

Dr. Charles L. Brooks III

A senior honors thesis submitted in partial fulfillment of the requirements for the degree of
Honors Bachelor of Science in Biophysics
from the University of Michigan

Abstract

Membrane proteins comprise approximately 30% of proteins found in vivo, but remain a challenging topic for study with traditional Biophysical techniques. We were interested in studying the M2 Proton Channel from the Influenza Virus and Islet Amyloid Polypeptide (IAPP), the amyloid peptide found associated with Type-II Diabetes Mellitus. Studies of these two proteins are crucial for the understanding of the diseases they are associated with, but these studies are hindered by the proteins' dependance on interactions with membranes. We have employed new advances in the simulation of biomolecules with Molecular Dynamics to probe the behavior of these two proteins.

IAPP is the amyloid peptide known to form the fibrils associated with Type-II Diabetes, and is known to be toxic to pancreatic beta cells by some form of membrane disruption. In our work of IAPP, by combining our implicit anionic membrane model with all-atom replica exchange molecular dynamics, we produced results which suggested that anionic membranes increase the binding strength of the peptide to the membrane, as predicted by experiment. Our results also revealed structural differences for IAPP bound to anionic and uncharged membranes in the membrane-disrupting N-terminal helix of IAPP. Additionally, we were able to provide insight into how factors such as pH and C-terminal amidation, which are known to alter the disease-causing aspects of IAPP, affect the characteristics of the membrane bound state.

The M2 Proton Channel is a pH-gated channel that is responsible for the acidification of the Influenza viral interior, a necessary step in Influenza infection. For our simulation work of the M2-Proton Channel, by modeling the transmembrane portion of the channel in aqueous phase with pH Molecular Dynamics, we were able to calculate pKa values for the four pH sensing Histidine residues that compared favorably to NMR results. By repeating our simulations but imposing varying degrees of rotational symmetry, we showed that the Histidine-mediated pH gating interactions are well modeled as a dimer of dimers. Finally, by simulating only the four Histidine residues from the NMR structure of the M2 channel, our results suggested that the entire channel is necessary to produce its pH gating behavior.

Acknowledgements

I would like to thank Professor Charles L. Brooks III for accepting me as his student and providing me the opportunity to work on the projects contained in this thesis. His mentorship has been one of the most beneficial experiences in my development as a scientist.

Secondly, I would like to thank all of the other scientists in the Brooks lab, who have provided me with a great deal of suggestions and assistance over the years. Working in the Brooks lab has showed me the incredible power and beauty of computation and theory in Biophysics. Additionally, I would like to thank David Braun of the Brooks lab, for providing me with much needed computer support over the years.

I would also like to thank Professor Ramamoorthy and his group for providing me with the experimental data and ideas for simulation directions that have formed the basis for our studies of IAPP.

Finally, I would like to thank Jason Karslake and Josh Karslake for the assistance in proofreading this Thesis and providing suggestions for improvements.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background on IAPP

1.1.1 Diabetes Mellitus Type-II

Amyloid peptides are a particularly rich topic for research because, although the mechanism of disease is similar, they are implicated in a wide variety of oftentimes severe illnesses that affect many different parts of the body (Table 1). The protein we studied, IAPP, is associated with Diabetes Mellitus Type-II. Type-II diabetes, in addition to being one of the most prevalent diseases associated with amyloidosis, is a common chronic disease in the United States, affecting an estimated 9.6% of the adult population[1]. The pathology of Type-II Diabetes is characterized by the loss of the ability of pancreatic beta cells to produce sufficient insulin in response to elevated levels of blood glucose, due to insensitivity of cells to insulin. The resulting extended periods of hyperglycemia produce

Disease	Amyloid Protein	Official Abbreviation	Fibril Location
Alzheimer's Disease	Beta Amyloid	Aβ	Brain
Diabetes Mellitus Type 2	Islet Amyloid Polypeptide (Amylin)	IAPP	Pancreas Islets
Parkinson's Disease	Alpha-Synuclein	none	Brain
Transmissible Spongiform Encephalopathy	Prion	APrP	Brain
Huntington's Disease	Huntingtin	none	Brain

Table 1.1: List of some amyloid diseases, associated peptide found to comprise amyloid plaques, official peptide abbreviation, and location of amyloid plaque.

organ damage, particularly in the kidneys, cardiovascular system, nervous system, the retinas of the eyes, and the feet and legs[2]. This hyperglycemia will also damage the insulin-producing beta cells, further exacerbating the condition. Although the disease can be managed by careful monitoring of blood sugar levels coupled with diet and exercise, there will still be an increased risk of heart attack, stroke, and vision loss. Unlike Type-I Diabetes, Type-II has a more gradual onset and typically occurs in adults, a trait typical of many amyloid diseases. This slow onset of amyloid diseases which appears correlated to the formation of amyloid fibers is one of the main arguments suggesting that the fibers play a role in the pathology. In order to better understand these arguments, it is useful to discuss some of the fundamental research in the discovery of amyloid peptides.

1.1.2 The Discovery and Characterization of Amyloid Deposits

Amyloid deposits in tissue were given their name, which means 'starch-like', in 1854 by Rudolph Virchow, who demonstrated that the reaction of amyloid deposits to iodine staining and subsequent treatment with sulfuric acid was similar to that of cellulose[3]. It was later realized that if stained with congo red, these deposits would exhibit apple-green birefringence under polarized light[3], giving the first evidence for unique structural characteristics. In 1901, amyloid deposits were discovered in the Islets of Langerhans of the pancreas in association with diabetes mellitus, a phenomenon that was documented independently by Opie[4] and Weichselbaum *et al.*[5]. Around this time, several other diseases were shown to have associated Amyloid deposits, including Alzheimer's disease[6]. These plaques were found to be composed primarily of fibrils via electron microscopy by Cohen and Calkins[7] in 1959. X-ray diffraction of isolated Amyloid fibrils, first performed by Bonar *et al.* in 1969, finally elucidated a molecular level picture of amyloid plaques; the fibers were shown to be comprised of identical monomers in a beta pleated sheet conformation oriented perpendicular to the fibrilar axis[8] (Figure 1.1). This 'cross-beta sheet' fibril structure is the property used today to define an amyloid disease in a biophysical context, and shall be employed for the purpose of this thesis. Westermark P. *et al.* sequenced the purified protein extracted from pancreatic islet plaques in 1986, and showed it to be a novel peptide unlike insulin with an approximately 50% homology to the neuropeptide CGRP[9]. The Westermark group then showed in 1987 that this protein, which they determined to also be found in an aggregated form in diabetic cats, is a naturally secreted hormone peptide[10]. Based upon the local production of amyloid peptides but homologous fibril

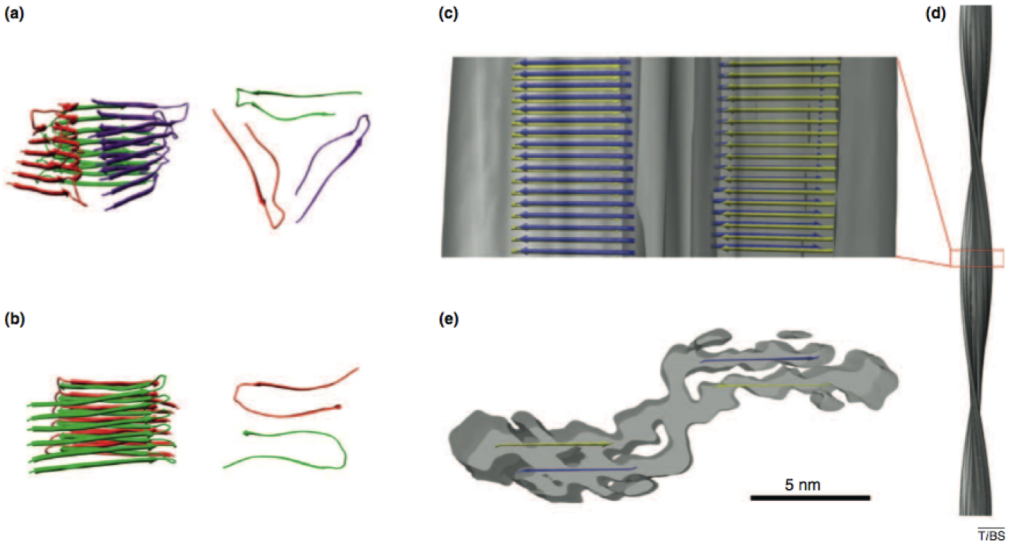

Figure 1.1: Proposed mature amyloid fibril structures. (a) and (b) represent two different models. (c), (d), and (e) represent Cryo-Electron Microscopy data of Aβ(1-40) with proposed beta sheet structure overlay.

structure, models were proposed for these diseases in which the amyloid fibrils played a key role in the molecular pathology. Preliminary experiments revealed that hIAPP could induce apoptosis in pancreatic beta cells[11] and that Aβ had toxic effects on neurons[12], supporting these models. However, further studies have shown that mature amyloid fibers are less toxic than originally suggested[13]. In light of these results, the focus has shifted in recent years to smaller oligomers of amyloid peptides as the toxic species responsible for the pathology. As such, the current challenge

in understanding amyloid diseases is discovering how amyloid peptides might fold and form these toxic small oligomers, and determining the mechanism of toxicity of these small oligomers. The first half of this challenge is what we studied in this work for the case of IAPP.

1.1.3 Role of IAPP in Type-II Diabetes

Even though research has implicated amyloid small oligomers as the toxic species, there is still a lack of consensus as to how these small oligomers kill cells. For IAPP, the large body of evidence points to some form of membrane disruption[14]. The method of permeabilization and/or destabilization of membranes is either by formation of a transmembrane oligomer pore or by larger-scale disruption involving small oligomers inserting into the membrane to induce excessive negative curvature strain. As our simulations will involve an implicit membrane model (the specifics of which are discussed in Section 1.3) it would be naive to attempt to explore the effect of IAPP upon membranes with our methods (at most we could simulate a barrel-stave model of a transmembrane pore, as even the toroidal pore model involves rearrangement of lipid beyond our membrane set-up). Rather, we want to explore how this peptide misfolds and oligomerizes, a topic that has also benefitted from important studies in recent years. Particularly important was the work of Knight *et al.*[15], who showed that the formation of IAPP fibers is catalyzed by the introduction of anionic, and not neutral lipid membranes to IAPP in vitro. This transition occurs over physiologically relevant concentration of anionic lipids (concentrations of 10% to 30% anionic lipid are observed in human beta cells); it has also been suggested that the concentration of anionic lipids in pancreatic β increases in the pre-diabetic condition[16]. The work of Knight *et al.* also demonstrated that IAPP, which is unstructured in solution, binds to anionic membranes and adopts an alpha-helical conformation. Together with their other result, this suggest that anionic membranes promote the folding of an alpha-helical structure of IAPP that acts as a crucial intermediate in the fibrillogenisis pathway. It has been shown by the Gafni, Ramamoorthy, and Steel groups that the N-terminal half of hIAPP, from residues Lys-1 to Ser-19, expressed as a fragment is capable of binding to and disrupting anionic membranes[17]. Coupled with the knowledge that the fragment from residues Ser-20 to Ala-29 is known to be amyloidgenic but weakly interacting with membranes, it is now hypothesized that the N-terminal half of the peptide is the primarily membrane binding portion of the peptide, whereas the C-terminal half participates in amyloid fibril formation.

1.1.4 Motivation for our Simulations of IAPP

Several recent works provided key pieces of motivation for our approach to studying Amyloid Peptides. Although NMR structures of monomers of the rat form of IAPP (rIAPP, Figure 1.3) [18] and a low pH, free-acid C-terminus human form IAPP (hIAPP, Figure 1.2)[19] have emerged in the past few years, only very recently has an NMR structure for hIAPP in physiologically relevant conditions (physiological pH and amidated C-terminus)[20] been published (Figure 1.2). This physiological hIAPP monomer was produced in SDS micelle, in an attempt to mimic the structure of an anionic membrane-bound state of hIAPP. In light of the evidence that the mechanism of hIAPP Beta-cell toxicity is by some form of membrane disruption, and that hIAPP oligomerization and fibrilization is 'catalyzed' by interactions with anionic (but not zwitterionic) cellular membranes, this NMR structure of physiological hIAPP in anionic membrane mimic provides the best known structural comparison for our simulations. We can use utilize it to examine how factors that are known to slow the rate of fibril formation, including lack of anionic membrane charge, low pH, and a

free acid C-terminus, influence both the structural characteristics and strength of membrane binding. The rat form of IAPP, being both non-amyloidgenic and non-toxic to pancreatic beta cells, will also be studied in our simulations to elucidate how its mutations give arise to its properties. From our studies, we hope to be able to gain insight into how membrane disruption and oligomerization might occur for physiological hIAPP but other IAPP variants or in other conditions.

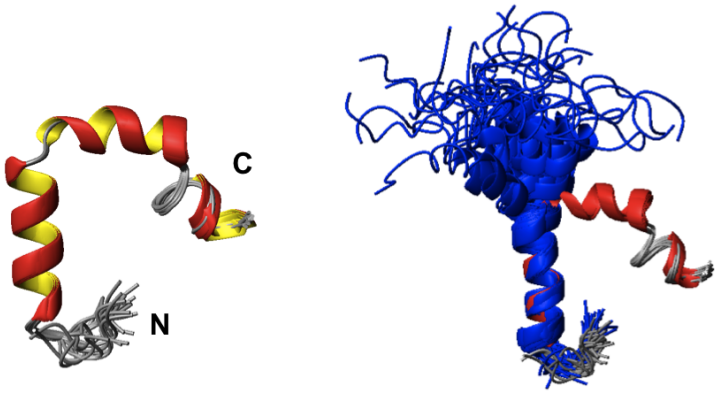

Figure 1.2: Structures of hIAPP from NMR. Above Left: hIAPP in anionic micelle at neutral pH with an amidated C-terminus, PDB:2L86. Above Right: In blue, cluster of structures from the NMR of hIAPP in anionic micelles at low pH with a free-acid C-terminus, PDB: 2KB8. Superimposed in red is the structure from the left, with an amidated C-terminus and low pH. Note the rigid kink at Histidine-18, the only titratable group, in the neutral pH structure.

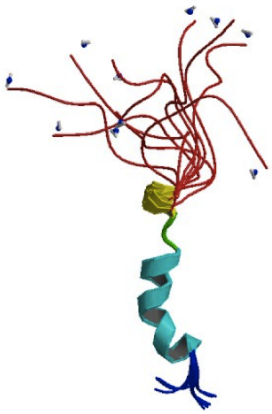

Figure 1.3: Structure of rIAPP from NMR. This structure was found in neutral pH with an amidated C-terminus in the presence of zwitterionic micelles.

Our approach is also motivated by several recent advancements in molecular dynamics techniques. The development of a Gouy-Chapman double layer potential to mimic membrane charge [21] that can be combined with the existing Generalized Born model for implicit membranes in CHARMM [22] has provided us a means to produce a fully implicit representation of an anionic lipid bilayer. This development provides the crucial means for measuring how membrane charge affects IAPP-membrane interactions. Our work is also largely possible due to all-atom replica exchange, a technique based in statistical mechanics that allows us to simulate a larger range of

peptide behavior in an equivalent amount of simulation time. All of these techniques are discussed more in depth in the Section 1.3.

1.2 M2 Proton Channel Background

1.2.1 Influenza A Infection

The Influenza A virus is a significant challenge in human health, with seasonal outbreaks causing 250,000-500,000 deaths a year and certain new strains resulting in pandemics with deaths in the millions[23]. A major factor in the virulency of the virus is its ease of infection, readily being transmitted through the air and by contact. Infection causes fatigue, sore throat, muscle pains, and fever. Although Influenza infection is not inherently life threatening, its effects increase the risk of complications of pre-existing conditions and contraction of a life-threatening infection. Because of this, Influenza fatalities are far more common in already weak individuals, such as children and the elderly. Vaccines exist for the virus; however, due to the rapid emergence of new strains of the virus, new vaccines are required each year. Antivirial adamantanes were once commonly prescribed for the treatment of Influenza, but the rapid rate of mutations has resulted in the virus, which had a 1% resistance in 1994 to amantadine, becoming almost completely resistant to these compounds today[24]. These drugs previously were able to bind to the M2 proton channel, preventing the release of the viral genetic material, but mutations have eliminated the binding target. Due to its previous success as a drug target, there is a great deal of interest in studying the M2. For this reason, it is important to understand it's specific role in the viral life cycle.

1.2.2 Role of the M2 Proton Channel in Influenza

The structure of the Influenza A virus is somewhat atypical in that it lacks a viral capsid composed of large numbers of a singular gene product in a repeating icosohedral geometry, as is typical of many viruses. Rather, Influenza is encapsulated by a 'viral envelope', a membrane derived from the lipids of the previously infected host cell[25]. This envelope possesses viral membrane proteins, and amongst these is the M2 proton channel (Figure 1.4). The infection begins when the target

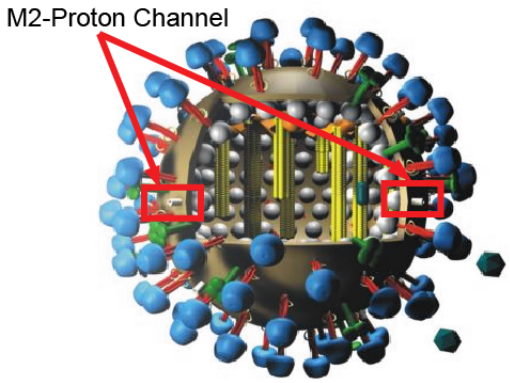

Figure 1.4: Structure of Influenza. The M2-Proton Channels are denoted with red arrows.

cell engulfs Influenza in an endosome and brings it inside of the cell towards the nucleus. As the

endosome diffuses deeper inside the cell, the pH on the interior of the endosome drops to 5.5-6[26]. This change in acidity causes the pH-gated M2 channels to open, allowing protons to flow into the viral interior[27]. This drop in pH inside of the virus will cause the RNP binding protein M1 to protonate and dissociate from the RNA/RNP complex. Concurrently, membrane proteins on the outside of the virus will trigger membrane fusion with the endosome, allowing the unbound RNA/RNP complex to exit the virus and enter the host cell, beginning the infection. There is a great interest in studying the mechanisms of proton conductance and pH gating of M2, both due to M2's previously mentioned use as a drug target and as a model system for a selective gated ion channel. In this work, we will primarily be examining the pH gating of the channel.

The M2 proton channel is a homotetrameric protein with 97 residues per chain. Each chain possesses a largely unstructured extravirial N-terminal region with interchain disulfide bonds, an alpha helical transmembrane domain of 25 residues, and a C-terminal region with a strongly amphipathic helix on the interior of the virus lying parallel to the membrane surface (See Figure 1.5 for an NMR structure of the transmembrane domain)[27]. If one expresses only the transmembrane portion of the protein, M2-TMP, the fragments are still able to form proton-selective ion channels in membranes, providing a good model system to work with[28]. Down the center of the transmembrane pore is a 'water wire' that conductions the protons through the channel. It has been shown with conductance studies of the full length protein[29][30] that lowering pH from 8.2 to 4.5 causes a 50 fold increase in proton conductance, with a 7 fold increase from a pH of 7.5 to 6.2, suggesting a 'pKa' for the channel opening of roughly 6. These results, combined with the NMR structures of the channel, suggest that the four Histidine-37 residues on the interior of the channel are responsible for mediating the pH sensing and conformational shifts in the channel. Additional mutation studies have suggested that the four Tryptophan-41 residues, one helical turn away from the Histidines, also play at least a steric role in occluding the pore in the closed conformation[31]. There have been several recent experimental works that have provided incredible detail into this gating mechanism that have served as the basis for our studies.

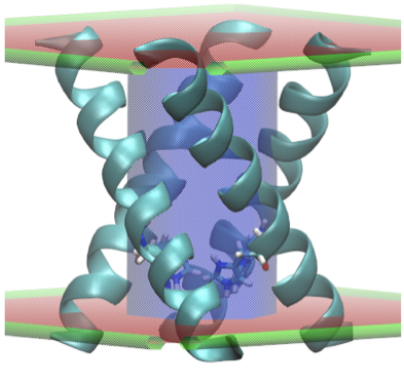

Figure 1.5: Solid state magic angle spinning NMR structure of M2-TMP. The planes represent a 25 Angstrom spacing to provide an idea of how M2-TMP sits in the hydrophobic core of the membrane. The blue cylinder provides a generous upper bound for the space water can occupy on the inside of the channel.

1.2.3 Motivation for Our Simulations of the M2 Proton Channel

In recent years, several NMR spetroscopists have taken advantage of the large difference in chemical shift between protonated and unprotonated Histidine to perform targeted measurements of the pH gating in M2. In particular, the Cross group performed solid-state magic angle spinning NMR on the transmembrane portion of the channel to produce pKa values for the four gating Histidines[32]. Their data suggested that the first two Histidines protonate at the same pKa with a value significantly greater than Histidine in solution, followed by a protonation event with a pKa near that of the channel opening pH and a final, much lower pKa protonation (Table 1.2). The

Residue	pKa	Error (+/-)
Histidine 1	<5	
Histidine 2	6.3	0.3
Histidine 3	8.2	0.2
Histidine 4	8.2	0.2

Table 1.2: Calculated pKas for the successive protonations of Histidine from the NMR work of the Cross group.

NMR spectrum for the channel at a pH where two of the four Histidines are protonated suggested that this state is comprised of a dimer of dimers, where an unprotonated Histidine and a protonated Histidine are forming a strong hydrogen bond. Based upon these results, they developed a model for the channel opening (Figure 1.6) where the first two protonation events form a stable dimer of dimers that is still a closed state, and that the channel does not open until the third proton is added to the cluster of Histidines.

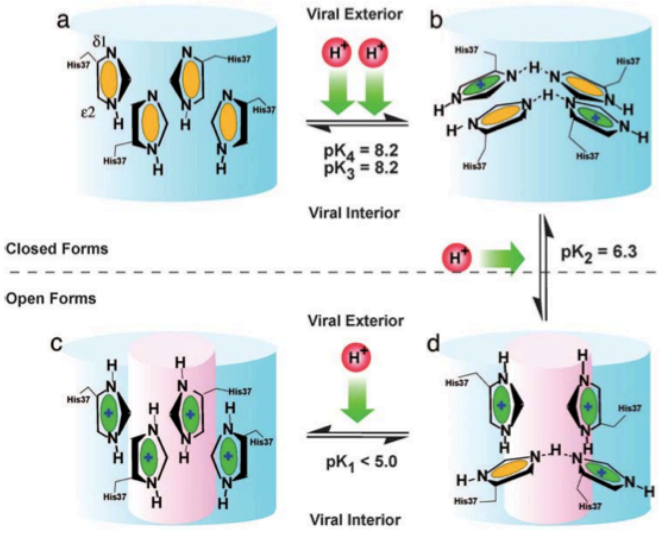

Figure 1.6: Model proposed by the Cross group for the pH gating of the M2 proton channel based upon their NMR results.

The Hong group also used solid state NMR to explore the gating interactions, focusing more on the fully unprotonated state (pH 8.5 in their tests) and the fully protonated state (pH 4.5)[33]. In their work, they showed that the open form of the channel has significantly increased sampling of Histidine rotamers, compared to fairly rigid Histidines in the closed form. They also showed that

in the completely unprotonated form, the ratio of Histidine with the epsilon nitrogen protonated compared to Histidine with the delta nitrogen protonated is 3:1.

In order to probe the pH gating of the channel computationally, we need a method where protonation state is a dynamic property that is allowed to change throughout the simulation. It is for this purpose that Constant pH Molecular Dynamics (CPHMD) have been developed in the Brooks lab, allowing one to not only explore the effects of pH on peptide structure and dynamics, but also to calculate how local environment affects pKa and tautomer fractions for protonatable residues. We will use this technique not only to attempt to recapitulate the pKas and tautomer fractions from the experimental works, but also to explore the effects of symmetry on these pKas and the channel. To do this, we will employ the IMAGES facility of CHARMM, which can be used to cause a peptide (in this case only some of the chains from the channel) to interact with mirror images of itself. In this way, we can impose symmetry on the channel. At this point, it is worth discussing some of the more specific details about the simulation techniques we will be employing to study the M2 Proton Channel and IAPP.

1.3 Underlying Physical Principles and Techniques

1.3.1 Molecular Dynamics

In an ideal world, we would model every molecule in our system using quantum mechanics. However, solving the Schrodinger equation for such a large system is entirely infeasable. Fortunately, classical mechanical calculations, referred to as Molecular Dynamics, can provide good approximations for the behavior of molecules. Many assumptions and theorems motivate this use of Molecular Dynamics, especially the Born-Oppenheimer approximation, which suggests that the total wavefunction of a molecule can be broken down into the product of the electronic and nuclear wavefunctions. This approximation is reasonable because of the large ratio in mass and thus velocity between the electrons and the atomic nuclei in a molecule; the nuclei can be considered stationary when calculating the motion of the electrons, and conversely the motion of the electrons is so rapid compared to that of the nuclei that they experience an averaged potential from the electronic motions. By assuming that the properties that require full calculation of electron motions are inconsequential for the study of biomolecule conformational dynamics, Molecular Dynamics was built around the sole consideration of nuclear coordinates and the net forces effecting them.

The most basic form for a Molecular Dynamics potential includes five terms; three terms describe 'bonded' interactions between an atom and any atoms it is chemically bonded to, and two 'non-bonded' interactions that can exist between any atoms. The first two bonded interactions, representing the energetic penalty of stretching or bending a chemical bond from some 'equilibrium' value, are modeled with a harmonic potential of the form derived from Hooke's Law. The final bonded interaction arises from the energetic cost of rotating a chemical bond, typically modeled as some cosine function to correctly place energetic minima and maxima with the corresponding *trans* and *gauche* positions. The total bonded potential is then given as:

$$V_{bonded}(x) = \sum_{bonds} \frac{1}{2} k_b (b - b_0)^2 + \sum_{angles} \frac{1}{2} k_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \frac{1}{2} k_\phi (1 + cos(n\phi - \delta))^2 \qquad (1.1)$$

where b and b_0 are the bond length and reference bond length; θ and θ_0 are the bond angle and reference bond angle; ϕ is the torsion angle; δ is a phase factor determining where the bond

torsion term passes through its minima; n is the multiplicity determining how many minima occur in one rotation; and k_b, k_θ, and k_ϕ are parameter constants. The values of the parameters used typically depend upon atom type and, for a given force field, were either determined experimentally or through quantum mechanical calculations. Typically, the choice of how to parameterize ones model is made to attempt to ensure the model produces accurate structural, thermodynamic, and/or spectroscopic data for the molecules simulated.

The non-bonded potential form is concerned with interactions that occur independent of chemical bonding. The first of these is the potential term from the interaction of partial charges, which can be simply modeled as a Coulombic potential. More difficult to model is the effect of Van der Waals forces; atoms need to both experience an attraction to one another due to instantaneously induced dipoles at longer ranges, as well as a strong repulsive affect at ranges short enough for electron clouds to be overlapping (Fermi repulsion). The most commonly used potential for approximating this phenomenon is the isotropic Lennard-Jones potential, giving us a complete form for the non-bonded potential of:

$$V_{nonbonded}(x) = \frac{1}{4\pi\epsilon\epsilon_0}\sum\frac{q_i q_j}{r_{ij}} + \sum\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right) \tag{1.2}$$

where ϵ and ϵ_0 are the permittivity of the vacuum and permittivity of the dielectric respectively; q_i and q_j are the charges of two interacting atoms; r_{ij} is the distance between interacting atoms for considering that given potential; and σ_{ij} is a parameter constant for the Lennard-Jones potential for the given forcefield.

By summing the contributions from the bonded and non-bonded potentials to produce a total potential for a given molecule, the force on a given atom can now be calculated with a simple derivative,

$$\mathbf{F}_i = -\frac{\partial V_{total}}{\partial \mathbf{r}_i} \tag{1.3}$$

depending upon the Molecular Dynamics system being used, this force is then iteratively solved with some form or extension of Newton's second law utilizing an algorithm with some small time step. This will allow for the acceleration of each individual atom to be calculated and thus the motion of the molecule to be approximated. One very common form of Newton's second law employed for this task is that found in the Langevin equation, which includes a retarding force caused by fluid viscosity and a noise term for the effects of Brownian motion:

$$m\frac{d^2\mathbf{r}}{dt^2} = -\frac{\partial V}{\partial \mathbf{r}} - m\beta\frac{d\mathbf{r}}{dt} + \mathbf{R}(t) \tag{1.4}$$

where m is the mass of the atom, β is a frictional coefficient, and $\mathbf{R}(t)$ the noise term. Utilizing this equation will estimate the effect of the random impacts of water molecules on the molecule being simulated, but as we will show later other tools are required to implicitly model the other effects of solvation. Additionally, due to the dependence of several of the terms upon temperature and the stochastic nature of the equation, Langevin Dynamics can provide a way to couple the simulated molecule to a heat bath when constant temperature is desired. The final step is to come up with a way to assign initial velocities, which is typically done with the Maxwell-Boltzman relation between average kinetic energy of particles in a system and the temperature,

$$T = \frac{1}{3Nk_B}\sum_i m < \mathbf{v}_i^2 > \tag{1.5}$$

where T is temperature, N is the number of atoms, k_B is the Boltzmann constant, m is the mass, and $< \mathbf{v}_i^2 >$ is the average velocity squared of the atom.

1.3.2 All-Atom Replica Exchange

It was first noticed by Levinthal in 1969 that due to the incredibly large number of degrees of freedom in even a peptide of modest length, it would take longer than the age of the universe for a peptide to reach its native folded state by sampling all possible states; even with a conservative estimate of possible conformations and incredibly rapid sampling rates[34]. This led to the hypothesis that protein folding is directed by an energy landscape dictated by intramolecular and solute-solvent interactions, with the native state corresponding to some deep free energy minima with steep 'walls'. It is with this hypothesis in mind that we apply our Molecular Dynamics force field to studying the conformational dynamics of peptides. One consequence of being able to study only relatively brief timescales with our model and computational resources is the possibility that by running a single simulation for a given temperature (particularly a temperature low enough to be physiological), one can become 'trapped' in a local energy minima that may not correspond to the native state. All-atom replica exchange was developed in response by Sugita and Okamoto[35] to provide a simulation protocol that would help ensure that a larger portion of the energy landscape is sampled and to prevent lasting kinetic trapping.

In all-atom replica exchange, multiple simulations or 'replicas' with identical experimental set-ups of the same molecule are run in parallel at various temperatures, typically exponentially spaced between two predetermined values. Although the experimental set up and the molecule must be the same, often times the conformation or orientation of the initial structures are varied from replica to replica to help direct sampling (especially when simulations are being run from experimental data that has provided several different possible structures, as is the case for our NMR data of IAPP). After a set number of dynamics steps, a protein database file (pdb) containing all atomic coordinates is saved for each replica and exchanges are attempted to swap the structures between two different temperature replicas if certain energetic criteria are met. The probability for a swap occurring is based upon the canonical Metropolis criterion,

$$p = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ e^{-\Delta} & \text{for } \Delta > 0 \end{cases} \tag{1.6}$$

for

$$\Delta = \left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j} \right) (E_j - E_i) \tag{1.7}$$

where k_B is the Boltzmann constant, T is the temperature of the replica, and E is the energy of that given replica.

Multiple studies have demonstrated the ability of Replica Exchange Molecular Dynamics simulations to produce informative structural and dynamical data for proteins. Perhaps most relevant is the work done by Khandogin and Brooks, which demonstrated a pH dependance for the folding of two fragments of the Aβ peptide[36].

1.3.3 Generalized Born with Simple Switching

An explicit method of modeling the effects of solvation in both aqueous phase and lipid bilayer would involve the simulation of a sufficient quantity of water and phospholipid molecules along

with the peptide. However, running such a simulation presents a large computational cost, which would severely limit the length of our simulations. The alternative is to implicitly produce the effects of solvation without including additional molecules in the simulation. One sucessful method of modeling implicit solvation is by solution to the Poisson-Boltzman equation (the details of this equation are discussed in section 1.3.4); however, as the equation is a non-linear differential equation, repeatedly calculating numerical solutions still proves computationally costly. Rather, it is often sufficient to consider the solvent as a continuum and only calculate the change in free energy in transferring the solute from vacuum to solvent. To do so we employ the method referred to as Generalized Born with simple SWitching (GBSW), an implicit solvation model that was extended by Im *et al*[22] to include a method for describing biological membranes.

The GBSW model has its roots in an early paper by Max Born[37], which first demonstrated a continuum electrostatic model of calculating solvation energy of an ion. The basic approach treats an ion in solution as a point charge q (where $q = Ze$) of radius a in a dielectric continuum slab with a dielectric constant of ϵ. By considering the energy of an electric field,

$$U = \frac{1}{2}\int \epsilon\epsilon_0 \mathbf{E}^2 d\mathbf{r} \tag{1.8}$$

where U is the energy, ϵ and ϵ_0 are the dielectric and vacuum permittivities respectively, and E is the electric field; and calculating the electric potential for the previously described ion in solution,

$$V(r) = \frac{q}{4\pi\epsilon\epsilon_0}\frac{1}{r} \tag{1.9}$$

where V(r) is the potential, q is the electric charge, ϵ and ϵ_0 are again the dielectric and vacuum permittivities, and r is the radial distance, we can calculate the change in electric energy lost in solvation. First, we can calculate the electric field from the electric potential,

$$\mathbf{E} = -\nabla V = -\nabla\frac{q}{4\pi\epsilon\epsilon_0 r} = \frac{q}{4\pi\epsilon\epsilon_0}\frac{\mathbf{r}}{r^3} \tag{1.10}$$

we can then plug this into equation 1.8 and integrate from r=a to infinity,

$$U = \frac{4\pi}{2}\int_a^\infty dr r^2 \epsilon\epsilon_0 \left(\frac{q}{4\pi\epsilon\epsilon_0}\frac{\mathbf{r}}{r^3}\right)^2 = \frac{4\pi}{2}\epsilon\epsilon_0\frac{q^2}{16\pi^2\epsilon^2\epsilon_0^2}\int_a^\infty dr\frac{1}{r^2} = -\frac{q^2}{8\pi\epsilon\epsilon_0}\frac{1}{r}|_a^\infty = \frac{q^2}{8\pi\epsilon\epsilon_0}\frac{1}{a} \tag{1.11}$$

to obtain an equation for the electric field energy in terms of the previously defined variables. The energy of solvation is then the difference between the energy in vacuum (or some low dielectric) and the energy in a high-dielectric environment,

$$U_{solvation} = U - U_0 = -\frac{q^2}{8\pi\epsilon_0}\left(1 - \frac{1}{\epsilon}\right)\frac{1}{a} \tag{1.12}$$

where U_0 is the field energy in vacuum and U is the field energy in dielectric. We can see from this equation that it is dependent upon the radius of the ion a, referred to in this context as a Born radius, which must be determined somehow (Born himself calculated the electrostatic solvation energy of ions based upon radii determined from crystal diffraction).

For the GBSW, the self-electrostatic solvation energy is expressed as

$$\Delta G^0_{elec,\alpha} = -\frac{1}{2}\left(\frac{1}{\epsilon_p} - \frac{1}{\epsilon_s}\right) q_\alpha^2 \left(\frac{1}{\eta_\alpha} - \frac{1}{4\pi}\int_{r>\eta_\alpha} d\mathbf{r}\frac{\mathcal{V}(\mathbf{r};\{\mathbf{r}_\alpha\})}{|\mathbf{r}-\mathbf{r}_\alpha|^4}\right) \tag{1.13}$$

where $\Delta G^0_{elec,\alpha}$ is the 'Coulomb field approximation' (comparable to the result we previously derived), ϵ_p is the dielectric constant in the interior of the solute, ϵ_s is the dielectric constant in the solvent, η_α is an arbitrarilly defined start point for integration to avoid the issue of the function being undefined at $|\mathbf{r} - \mathbf{r}_\alpha| = 0$, $\mathbf{r}_\alpha$ is the Born radius and $\mathcal{V}(\mathbf{r})$ is a solute volume function that is equal to 1 in the interior of the solute and 0 in the solvent. Since this term represents the solvent inaccessibility, we can define it in a way that includes the presence of the membrane as an infinite slab lying in the x-y plane that is solvent inaccessible (and thus has a low dielectric constant). It should be noted that this 'Coulomb field approximation' neglects the effect of solvent polarization at the dielectric boundary; the full form of $\Delta G_{elec,\alpha}$ utilized in the GBSW includes an additional term similar to that in equation 1.13 that serves to correct this omission.

It is also important to remember that the electrostatic free energy is not the sole component of the free energy of solvation. We must also consider the free energy that arises from the entropic penalty of the solute forming a cavity in the solvent. The equation for this utilized in the GBSW is of the form

$$\Delta G_{np} = \gamma S \tag{1.14}$$

where γ is a phenomenological surface tension coefficient and S is the solvent-acessable surface area. A similar procedure to that used for equation 1.13 is employed in order to define the solvent-acessable surface area S in the context of the infinite slab solvent-inacessable membrane.

Brooks and Im have demonstrated the ability of the combination of the GBSW implicit membrane with replica-exchange molecular dynamics to model the interfacial folding and insertion of membrane peptides[38]. Bu and Brooks have further extended the GB implicit membrane and AAREX techniques to model the assembly of transmembrane helix oligomers within good agreement of NMR experimental data[39].

1.3.4 Gouy-Chapman Theory

For simulating the behavior of monomeric peptides that are expected to fold as peripheral membrane proteins, it is necessary to consider the effects of membrane charge and/or potential on the folded structure. This is particularly important for our work as the NMR structures we used for starting structures and benchmarking were found in anionic sodium dodecyl sulfate (SDS) or zwitterionic dodecylphosphocholine (DPC) micelles. As previously discussed, our simulations will be utilizing the GBSW implicit solvent model of the membrane, so explicitly modeling these charged phospholipids or phospholipid-mimic molecules is not an option. Instead, we implemented a one-dimensional potential function to describe the electric potential as a function of distance away from the bilayer surface by utilizing Gouy-Chapman theory.

In order to properly motivate Gouy-Chapman theory, we must first discuss the fundamental principles of the behavior of electric charges in bulk solution. The main goal is to describe the electric potential that results from a charged surface solvated in a solution containing ions, where the combination of electric attraction and entropic effects will generate a 'cloud' of appropriately charged counter-ions in the vicinity of the charged surface, a phenomenon known as an 'electric double layer'. The first step is to come up with a means of relating the change in electric potential to some density of charge. By considering Gauss's law,

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon\epsilon_0} \tag{1.15}$$

where $\mathbf{E}$ is the electric field, ρ is the charge density, ϵ is the dielectric permittivity, and ϵ_0 is the vacuum permittivity, and using the appropriate definition of potential,

$$\mathbf{E} = -\nabla V \tag{1.16}$$

where V is the scalar electric potential, we can substitute equation 1.16 into equation 1.15 to yield the Poisson equation relating electric potential to charge density,

$$\nabla^2 V = -\frac{\rho}{\epsilon\epsilon_0} \tag{1.17}$$

In our situation, it is more convenient to utilize the one-dimensional form of the Poisson equation, given that the electric potential away from a membrane lying in the xy plane will be constant in both the x and y directions, and only vary as one moves away from the membrane in the z direction:

$$\frac{d^2V}{dz^2} = -\frac{\rho}{\epsilon\epsilon_0} \tag{1.18}$$

Now we must solve for the charge density, which will be related to the concentration of ions as a function of z displacement. This concentration, in turn, will be dependent upon the electric potential as a function of z displacement. By examining the Boltzmann distribution,

$$N_i = g_i e^{-E_i/k_B T} \tag{1.19}$$

where N_i is the number of particles occupying a state of energy E_i, g_i is the degeneracy of these states, T is the temperature, and k_B is the Boltzmann constant. By substituting the electric potential energy $q{\cdot}V(z)$ for E_i, replacing N_i with an expression of concentration as a function of z, $c(z)$, and adding a factor of c_0 to serve as an initial condition and provide the correct dimensions, we can produce an equation for concentration as a function of $V(z)$,

$$ec(z) = qc_0 e^{-qV(z)/k_B T} \tag{1.20}$$

By plugging this into equation 1.18, we produce the Poisson-Boltzmann equation, a non-linear differential equation:

$$\frac{d^2V}{dz^2} = \frac{qc_0}{\epsilon\epsilon_0} e^{-qV(z)/k_B T} \tag{1.21}$$

for all previously defined variables.

As we have already reduced the Poisson-Boltzmann equation to its one-dimensional form, we only need to identify the proper boundary conditions and find a solution for our set up. We have already established that the membrane will be defined as an infinite sheet of some charge (in our case negative, so as to reproduce the SDS anionic membrane mimics), so one of our boundary conditions will be the surface potential. T. Lazaridis[21] demonstrated a viable method of producing a Gouy Chapman 'electric double layer' solution to the Poisson-Boltzman equation for membranes that is adaptable to implicit models of solvent and membrane. By using a surface potential $V(0)$ of the form

$$sinh\left(\frac{qV(0)}{2k_B T}\right) = \frac{\sigma}{\sqrt{8\epsilon\epsilon_0 k_B T\rho}} \tag{1.22}$$

where q is the charge of an electrolyte, k_B is the Bolzmann constant, T is the temperature, σ is the surface charge density, ϵ_0 is the permittivity of vacuum, ϵ is the relative permittivity of

the dielectric, and ρ is the number of ions per volume; we can produce a solution to the Poisson Boltzmann equation for the potential as a function of distance from the membrane of the form

$$V(r) = \frac{2k_BT}{q} ln \frac{[1 + \alpha e^{(-\kappa r)}]}{[1 - \alpha e^{(-\kappa r)}]} \tag{1.23}$$

where α is defined as

$$\alpha = \frac{e^{(e_q V(0)/2k_BT)} - 1}{e^{(e_q V(0)/2k_BT)} + 1} \tag{1.24}$$

with e_q as the elementary charge; and κ, the Debye length, is defined to be

$$\kappa = \sqrt{\frac{2\rho O q^2}{\epsilon\epsilon_0 k_B T}} \tag{1.25}$$

where O is some offset from the nonpolar-polar interface (there is assumed to be some small distance in Angstroms between the charged plane and the nonpolar-polar interface to produce a more accurate picture of a charged headgroup on an anionic membrane molecule). The potential inside the membrane is taken to be a constant equal to that of the surface potential.

For implementation with the CHARMM program and the MMTSB package, the user specifies the anionic lipid mol fraction f_a, the area per lipid A, the valence of the electrolyte, and the offset for the plane of smeared charge. We can then calculate the surface charge density,

$$\sigma = \frac{f_a q}{A} \tag{1.26}$$

to plug into our potential functions. In his work, Lazaridis implemented the Gouy Chapman double layer with an implicit solvent/membrane model in the CHARMM forcefield. He used this model to correctly predict the location and orientation of several peptides known to bind to anionic membranes, including Melittin and Magainin-2, in good agreement with experimental data. In addition, calculated binding energies compared favorably to experimental thermodynamic data; together these results suggest that this approach can be successfully applied to study the interactions of peptides with anionic membranes.

1.3.5 pH Molecular Dynamics

In order to properly model the behavior of our peptide at different pH values, we need a method where protonation state is a dynamic property that is allowed to change throughout the simulation. It is for this purpose that Constant pH Molecular Dynamics were developed (CPHMD)[40]. In CPHMD, we utilize a coordinate, lambda, where a lambda value of 0 corresponds to the protonated state, and a lambda value of 1 corresponds to the unprotonated state. To impose this relation, we couple lambda to a coordinate theta with

$$\lambda_i = sin^2(\theta_i) \tag{1.27}$$

such that no matter what value theta has, lambda is constrained to be between 0 and 1. We then use this theta variable as a 'coordinate' for a fictitious atom that is given a 'mass' similar to that of

hydrogen. The kinetic energy for the fictitious theta atom is included in an extended Hamiltonian formalism,

$$H_{extended}(r,\lambda) = H_{hybrid}(r,\lambda) + \sum_i \frac{m_i}{2}\frac{d\theta_i}{dt}^2 + U^*_{biasing}(\lambda) \tag{1.28}$$

where H_{hybrid} is a Hamiltonian for the real atoms of the form

$$U_{int}(r) + U_{GBnonpolar}(r) + U_{Coulomb}(r,\lambda) + U_{GBelec}(r,\lambda) + U_{vdW}(r,\lambda) \tag{1.29}$$

with U_{int} and $U_{GBnonpolar}$ for the internal energy and non-polar Generalized Born terms that do not depend upon lambda; and $U_{Coulomb}$, U_{GBelec}, U_{vdW}, which are energy terms representing the Coulombic, electric portion of the Generalized Born term, and the van der Waals terms that are linearly interpolated from values for the protonated to the unprotonated state depending upon the current value of lambda.

The $U^*_{biasing}$ that appears in the extended Hamiltonian takes the form

$$\sum_i (-U_{model}(\lambda_i) + U_{pH}(\lambda_i) - U_{barrier}(\lambda_i) \tag{1.30}$$

where $U_{model}(\lambda)$ is a potential of mean force derived from a model compound (typically the titrating amino acid in solution) designed to flatten the free energy surface along the titration coordinate, $U_{pH}(\lambda)$ is a free energy of protonation dependent upon the simulation pH given by

$$U_{pH}(\lambda_i) = -log(10)k_BT\lambda_i(pKa_{ref} - pH) \tag{1.31}$$

where pKa_{ref} is a reference pKa from a model compound. Finally, the $U_{barrier}$ term in the biasing potential is a potential energy term designed to provide an energetic penalty to having a lambda value between 0 and 1, in an attempt to force lambda to occupy values near either extreme. This is because values of lambda between 0 and 1 are not physical and cannot be considered to correspond to any real state; when analyzing our simulations, we only consider some narrow range of values as valid physical states (usually, and in the case of our simulations, we considered $0 \leq \lambda \leq 0.1$ as the protonated state and $0.9 \leq \lambda \leq 1$ as the unprotonated state) and do not consider these 'mixed' states in calculating pKa. The formula for the pKa calculation is

$$S = \frac{1}{1 + 10^{n(pKa-pH)}} \tag{1.32}$$

where S is the fraction unprotonated, n is a hill coefficient, and pH is the simulation pH for that given window. This PHMD formalism has also been extended using a methodology similar to that employed for λ to include a variable 'x' corresponding to the tautomer form (i.e. for unprotonated Histidine whether the single proton is on the epsilon or delta nitrogen), allowing for one to also look at tautomer interconversion simultaneously with protonation state[41].

With all of these factors combined into equation 1.28, we are now able to couple protonation state to the dynamics of the protein and some pH of the simulation, letting us look at how the environment shifts pKa, how the structure and dynamics of the protein change as a function of pH, and other phenomena. This technique has been shown to recapitulate the pKas of model compounds, and predict pKa shifts and pH dependent behavior in several peptides[42][43].

Chapter 2

Simulation Methods

2.1 Full Length IAPP Simulations

Based upon the previously mentioned structures of hIAPP (both amidated and non-amidated) and rIAPP in anionic micelles, we designed our first simulations to see if we could predict or explain any differences in behavior between the peptides bound to membranes. Additionally, we wanted to test for possible differences in the monomeric forms when peripheral to anionic and uncharged membranes; and when modeled in a low pH environment.

We began by extracting structures for hIAPP both with a amidated C-terminus in SDS micelle (found by NMR) as provided by the Ramamoorthy group (PDB:2L86), and with an unmodified C-terminus in SDS micelle from the Alexandrescu group (PDB:2KB8), and converting both sets of structures to the all-hydrogen PARAM22 format for the CHARMM force field with the assistance of the MMTSB toolset. All simulations and minimizations were run with the Generalized Born with simple SWitching (GBSW) model of the membrane with a membrane hydrophobic core width of 25 Å, a 1.5 Åwidth of dielectric smoothing between the membrane and the solvent, and a 0.03 kcal/(mol*$\AA^2$) surface tension coefficient. Additionally, in both minimization and in simulation, the PATCH facility of CHARMM was utilized to establish a disulfide bond between Cys-2 and Cys-7, as well as to control whether the C-terminus is of the amidated (CT2) or unmodified (CTER) type. The first 12 structures for each NMR ensemble were minimized with 50-step steepest descent minimization followed by minimization with the Newton-Raphson minimization performed over either 500 steps or until the energy decrease between steps became less than $1.0*10^{-5}$ kcal/mol; whichever occurred first. We also set the GB salt concentration to be 100 mM, in parallel to the salt concentration used in the NMR experiments. For our simulations, we utilized the MMTSB[44]

Peptide	Sequence
hIAPP	KCNTA TCATQ RLANF LVHSS NNFGA ILSST NVGSN TY
rIAPP	KCNTA TCATQ RLANF LVRSS NNLGP VLPPT NVGSN TY
M2-TMP	SSDPL VVAAS IIGIL HLILW ILDRL

Table 2.1: Table of Amino Acid Sequences for Relevant Peptides

toolsets all-atom replica exchange (AAREX) facilities to perform replica exchange molecular dynamics with CHARMM version c36a3. All simulations utilized 12 replicas, each assigned one of the 12 NMR structures prepared, with temperatures exponentially spaced from 298 K to 450 K, and were run with a 2 femtosecond time step. Coordinates were recorded for analysis and replica

exchanges were attempted every 1 picosecond of the simulation, with all simulations being run for a total 20 nanoseconds.

A separate simulation was run for the amidated C-terminus structure with the CT2 patch and the unmodified structure with the CTER patch. Copies of these two simulations were also run with the addition of the Gouy-Chapman double layer, with a mol fraction of anionic lipids of 0.3, an area per lipid of 70 $\AA^2$, a valence of 1, and a double-layer offset from the hydrophobic core of 3 $\AA$. Simulations to mimic the effects of low pH were run for both amidated and non-amidated structures with the previously described procedure for the given NMR structures; however, the histidine residue was patched with the doubly-protonated HSP variant rather than the singly protonated HSE/HSD. These ensembles of structures underwent the same minimization procedures and were used to set up identical simulations as the regular pH form for both charged and uncharged membranes. This brought our total number of simulations up to 8; allowing us to completely test the various effects amidation, alteration of pH, and introducing a charged membrane has upon membrane-peripheral hIAPP structure.

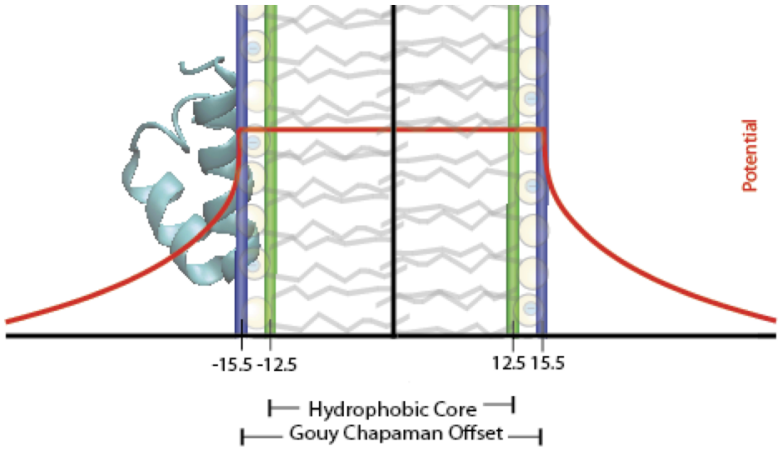

Figure 2.1: Plot depicting our membrane setup. The GBSW hydrophobic core lies between the green bars. The blue bars depict the start of the Gouy-Chapman double layer, with the red line approximating the potential due to the GC layer as a function of Z-position.

We also prepared rIAPP from the NMR structure in DPC micelles produced by the Ramamoorthy group (PDB:2KJ7) for the PARAM22 format for the CHARMM force field; with a disulfide bond patched between Cys-2 and Cys-7 and an amidated C-terminus. Only 10 structures were available in the pdb file, which were minimized with the same procedure for hIAPP. Simulations were run identical to those for hIAPP in an anionic membrane, with the 10 structures parsed out over 12 temperature windows; some structures being repeated. These steps were repeated for rIAPP prepared with the an uncharged membrane. Due to some unusual differences between the rIAPP data and our results for hIAPP in which rIAPP tended to form transmembrane states, which will be discussed in the results section, we repeated our simulations for rIAPP in anionic and uncharged membranes with identical procedures, with the exception of translating the starting NMR structures 12.5 Angstroms in the Z-direction, to place the peptide at the periphery of the membrane. The results for these four rIAPP simulations were used for comparison with respective hIAPP simulations.

2.2 Simulations of the M2 Proton Channel

For our work on the M2 proton channel, we first wanted to explore whether or not it would be feasible to employ our pH Molecular Dynamics in combination with our low dielectric, Generalized Born model of the membrane without additional adjustment of the PHMD parameters. To study this, we performed replica exchange molecular dynamics over 8 windows exponentially spaced from 298K to 400K of a single Histidine molecule inside of a GBSW implicit membrane. The GBSW was set up to provide a membrane hydrophobic core width of 25 Å, a 1.5 Åwidth of dielectric smoothing between the membrane and the solvent, and a 0.03 kcal/(mol*$\mathring{A}^2$) surface tension coefficient with a salt concentration of 120 mM. The Histidine was pinned in place with a 5 kcal/mol/$\mathring{A}^2$ harmonic constraint applied to the α-Carbon to prevent it from diffusing out of the membrane. Separate simulations were run with identical conditions at PHMD pH values spaced from 1 to 10 in increments of 1 pH unit. Our simulations were run with a 2 fs time step for 1 ns total simulation time, with coordinates (both real coordinates as well as λ and x for all titrating groups) saved every 500 steps. Langevin dynamics for all heavy atoms was used with a friction coefficient of 5.0 ps^{-1}. As we will discuss in the results section, this method produced an unreasonable pKa for the Histidine, suggesting that direct combination of the GBSW implicit membrane with PHMD would not succeed for our simulations.

In light of this, we ran simulations of the NMR structure of the closed state of the transmembrane portion of the M2 proton channel (PDB:1NYJ) with the GBSW implicit solvent model without the membrane (i.e., in a dielectric of 80 to mimic that of water). With this approach, we assumed that the 1 ns length of simulation that is required to obtain a calculation of pKa would be brief enough to maintain the integrity of the channel. The NMR structure of the channel was minimized with the all-hydrogen PARAM22 format for the CHARMM force field with 50-step steepest descent minimization followed by minimization with the Newton-Raphson minimization performed over either 500 steps or until the energy decrease between steps became less than $1.0*10^{-5}$ kcal/mol; whichever occurred first. For our simulations of the full transmembrane portion of the channel, we utilized identical simulation methods as those employed for just one Histidine (with the exception of the absence of the membrane). Each simulation was repeated 3 times so that we could obtain statistics on our calculated values.

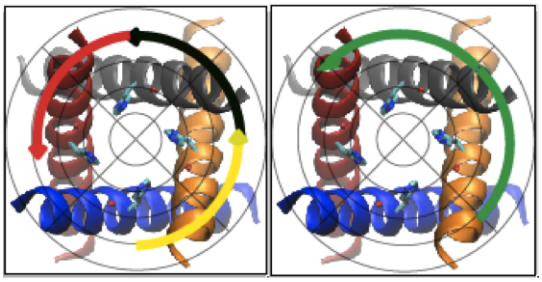

Figure 2.2: Above Left: Representation of the channel recreated from a monomer with three rotational images, at 90, 180, and 270 degrees. Here the yellow, black, and red helices are 'images' of the blue monomer. Above Right: Representation of the channel recreated from a dimer with one 180 degree rotational image. Here the red and black helices are 'images' of the blue and yellow dimer of helices

To explore the effects of symmetry, we extracted just one chain from the minimized M2-TMP structure to produced a pdb file containing just a monomer of the channel. Two adjacent chains

were also extracted from M2-TMP to produce a pdb file of a dimer. The identical procedure for the full tetramer was then performed for simulating these structures, with the addition of the use of the images facility to 'recreate' the full tetrameric channel. The images facility allows for a peptide to interact with an image of itself produced by some form of symmetry rotation. In the case of the dimer, this was one 180 degree rotational symmetry image; for the monomer, it was three rotational symmetry images of 90, 180, and 270 degrees (Figure 2.2). This approach, as previously mentioned, has been shown to faithfully reproduce the structure of M2 with the GBSW implicit membrane[39]. Although this reduces the ability to resolve pKas from separate Histidines, we will be able to gain insight into how forcing symmetry upon the channel affects calculated pKa values.

We also wanted to test if the pKa shifts of the gating Histidine result from the interactions of four Histidines with themselves, or whether the rest of the channel plays some role. Do do this we began by extracting just the coordinates for the four Histidine residues from the closed-state M2-TMP NMR structure. We then repeated our simulations for the full tetramer, dimer with one rotational image, and monomer with three rotational images with identical procedures, using just these Histidine residues pinned in place with 5 kcal/mol/$\mathring{A}^2$ harmonic restraints applied to the alpha and beta carbons. This allows us to calculated pKa values for these four Histidines as they interact with themselves and/or their images.

Chapter 3

Results

3.1 Simulations of IAPP

3.1.1 Membrane Charge Effects the Character and Strength of hIAPP Membrane Binding

Once all of our data was collected, our first interest was studying how the addition of membrane charge affected the binding of hIAPP to the membrane surface. For this purpose, we first extracted all of the saved coordinates from the last nanosecond of simulation from the 298K window of the replica exchange (1,000 sets of coordinates in total) for amidated hIAPP in both anionic and uncharged membranes. We then used the MMTSB toolset to cluster these structures based upon a mutual RMSD of 3 Angstroms (i.e., in order to be included in the cluster a given structure had to be within a 3 Angstrom rmsd of the 'center' of the cluster, see Table 3.1 for a breakdown of the clustering results). The structure closest to the center of each cluster was visualized to serve as a comparison point, which will be referred to as the 'representative structure' for a given cluster.

By comparison of the representative structures for each of our simulations, it quickly became apparent that the alteration of membrane charge causes a noticeable change in the orientation of the N-terminal helix with respect to the membrane. For those simulations with the membrane charge turned on, the N-terminal helix lies fairly flat against the membrane, such that most of the hydrophobic groups on the helix are inserted into the membrane. By contrast, the results of simulations with neutral membranes show a tendency for the N-terminal helix to drift away from the membrane surface (Figure 3.1). This procedure of clustering was repeated for the first nanosecond, and comparison of the representative structures reveals that this effect occurs quite rapidly in the simulation (Figure 3.2). To quantify this, we measured the tendency of each residue to be buried in the membrane in the last nanosecond of simulation. To do this we counted how many times each residue had coordinates within the 25 Angstrom wide membrane region in all of the structures from the last nanosecond that were used for the clustering. Both a count of how often each residue had at least one atom buried and a total number of atoms buried was taken to provide a more complete characterization of the orientation of IAPP and the depth of its penetration into the membrane (Figure 3.3). This characterization agrees with our qualitative observation, showing a tendency for residues closer to the N-terminus to be inserted in the membrane much less frequently when membrane charge is not present. This analysis was repeated for the first nanosecond of simulation (Figure 3.4), which confirms that these differences in membrane binding occur very rapidly (it

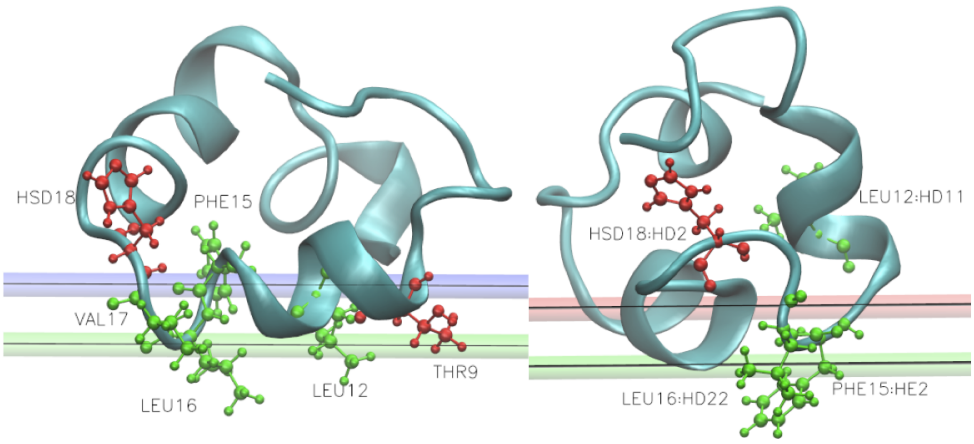

Figure 3.1: Above Left: Amidated hIAPP in 30% anionic membrane. The blue bar represents where the negative membrane charge is smeared (where the potential begins to drop off), the green bar represents the beginning of the hydrophobic core, the green amino acids are hydrophobic, and the red amino acids are hydrophilic. Above Right: Amidated hIAPP in an uncharged membrane. The same color scheme as before is adopted but with the red bar denoting where the membrane charge layer would be if it had not been turned off. Note the N-terminal helix for hIAPP in uncharged membrane only has Phe-15 and Leu-16, where the peptide is beginning to unfold, still interacting with the membrane; the rest of the helix is in solution .

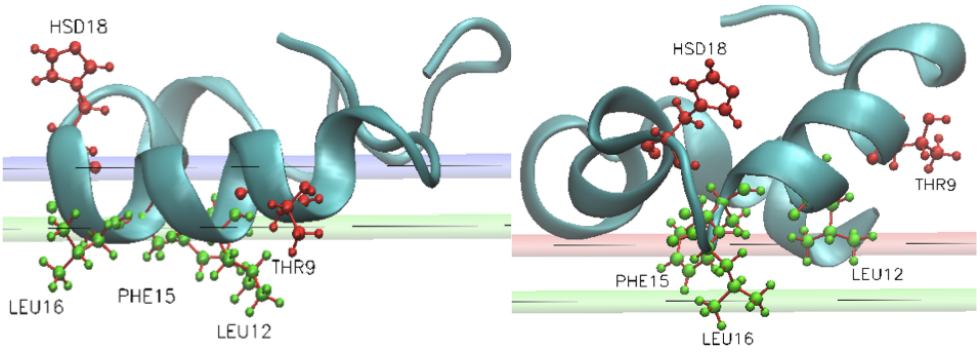

Figure 3.2: The same color scheme and representations were employed as those in Figure 3.1 Above Left: Amidated hIAPP in 30% anionic membrane. Above Right: Amidated hIAPP in an uncharged membrane. Note how the difference in the binding of the N-terminal helix is becoming apparent even after short timescales.

should be noted that although our protein was initially positioned inside the membrane parallel to its surface, Figure 3.4 suggests that IAPP diffused to the periphery of the membrane in roughly 0.1 ns, as estimated from all residues being buried in approximately 10% of this first nanosecond). It may be curious that, even if anionic membrane charge is present, the positively charged N-terminus does not seem to interact strongly enough with the membrane to have it or any adjacent residues buried. However, one must keep in mind that the disulfide bond between residues Cys-2 and Cys-7 creates strong steric restraints on conformations the N-terminus can possibly occupy.

It also became apparent from comparison of the structures visualized in VMD that there is an effect on the secondary structure of IAPP when bound to an uncharged membrane. To quantify this, we performed the DSSP algorithm for secondary structure identification on all of our structure files from the last nanosecond of simulation, and plotted the fractional occupancies of alpha-helical

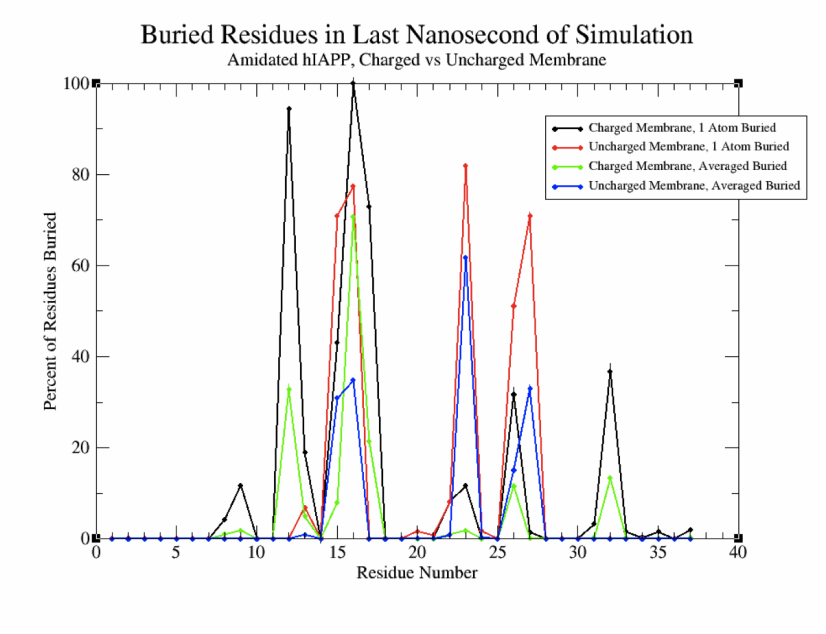

Figure 3.3: Plot of percentages of residues buried in the hydrophobic region in the last 1 ns of simulation time, for amidated hIAPP. The red and black lines, for charged and uncharged membranes respectively, correspond to a percentage of system coordinates over the last 1 ns where a given residue has at least 1 atom with a Z-coordinate of magnitude $\leq$12.5 $\AA$. The blue and green plots, again for charged and uncharged membranes respectively, correspond to a summation of the percent of atoms for each residue with Z-coordinates of magnitude $\leq$12.5 $\AA$ over the last 1 ns.

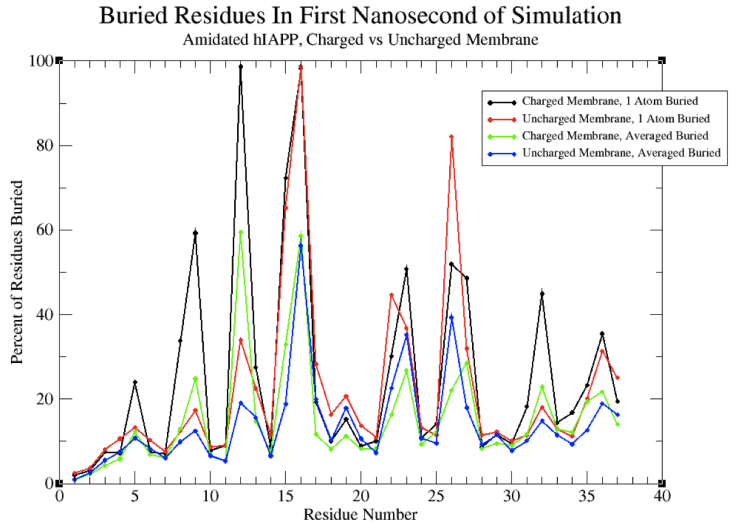

Figure 3.4: Plot of percentages of residues buried in the hydrophobic region in the first 1 ns of simulation time, for amidated hIAPP. The same representational scheme in Figure 3.3 has been repeated.

and 3-10 helical conformations for each residue (Figure 3.5). From the plot we observe a large decrease in alpha helical content from residues Asn-14 to Val-32 coupled with an increase in helical

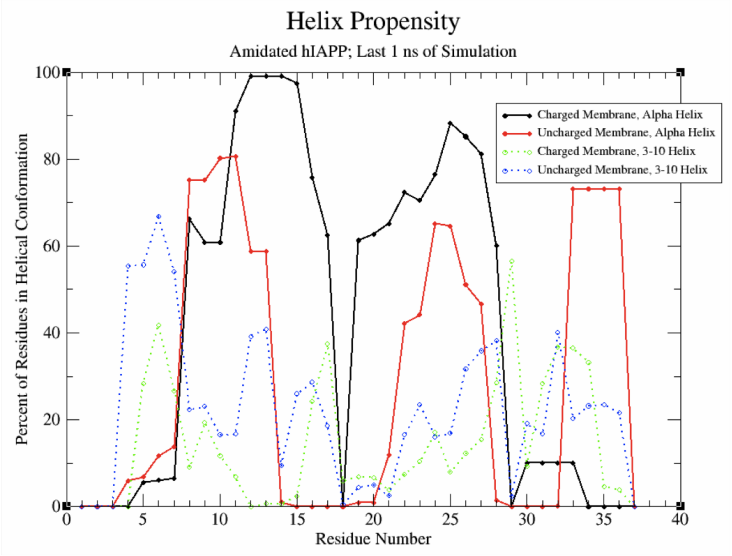

Figure 3.5: Plot of percent of residues in a helical conformation, as determined by the DSSP secondary structure protocol, for amidated hIAPP in charged and uncharged membranes. The solid black and red lines represent alpha-helical character for hIAPP in Charged and Uncharged membranes respectively. The dotted green and blue lines represent 3-10 helical character, again for hIAPP bound to charged and uncharged membranes.

content for residues Val-32 to Tyr-37 in uncharged membranes. For residues Phe-15 to Ser-20 the amount of 3-10 helical character is about the same for both membranes. This would seem to suggest that without the anionic membrane present, hIAPP begins to at least partially unfold starting from the His-18 at which the helical structure of the peptide is kinked.

	Number of Clusters	Size of Largest Cluster	RMSD of Representative Structure from Cluster Center
Charged Membrane, First Nanosecond	12	318	0.41
Charged Membrane, Last Nanosecond	4	532	0.80
Uncharged Membrane, First Nanosecond	8	752	0.48
Uncharged Membrane, Last Nanosecond	4	809	0.96

Table 3.1: Clustering results for the first and last nanosecond of simulation for amidated hIAPP in either charged or uncharged membranes. The representative structure is the structure closest to the center of the largest cluster.

These results are consistent with the NMR data , which suggest that in anionic micelle, hIAPP should have strong alpha helical character from residues Cys-7 to Val-17 and from Asn-21 to Ser-28, and a short 3-10 helical segment from Gly-33 to Asn-35. We also predict that the N-terminal helix lies peripheral to the anionic membrane, consistent with both the NMR results and the studies from the Miranker group. Although our results seem to display more helical character in the region from Val-17 to Asn-21 than the NMR data, it should be noted that the alpha-proton chemical shifts from the experiment only suggest that the Histidine is not in a helical conformation; the inference that this full region is not in an alpha helical conformation is based upon the lack of NOEs between Leu-16 and Ser-19 and between His-18 and Ser-20. It would be our suggestion from our simulations that there is some weaker alpha-helical character between Val-17 and Asn-22 that is not being captured with the NMR methods employed, which can explain the relative stiffness in the bend at His-18 compared to the other NMR structures (the structure of hIAPP in low pH,

for example, also shows N-terminal helical character with a break from residues Val-17 and Asn-21 but has a much greater flexibility in this region of the peptide).

Taken together, the lack of buried residues in the N-terminal 'membrane-binding' region of the peptide coupled with the unfolding of the helical region of the peptide in the absence of membrane charge suggests that the peptide binds more strongly to the anionic membrane which stabilizes the alpha helical conformation we see in the NMR structure. These two factors could also be contributors to the tendency of anionic and not neutral membranes to catalyze amyloid fiber formation. To probe this further, we wanted to explore another factor that effects both structure and fibril formation, pH.

3.1.2 Low pH Causes Local Unfolding around Histidine-18

The primary difference between the previously discussed NMR structures of hIAPP in anionic micelle at low pH without an amidated C-terminus and the structure of hIAPP in physiological conditions was in the structure of the peptide around His-18 and the C-terminus. The low pH, free carboxylic acid variant lacks the significant kink in the alpha helical structure at residue His-18 and the NMR data suggests the alpha helix is most stable at the N-terminus and decreases in stability moving towards the C-terminus. To compare this result with our data, we looked at the last nanosecond of our two simulations for amidated hIAPP in anionic micelle (started from the physiological structure) with His-18 having either both nitrogens protonated (low pH). We

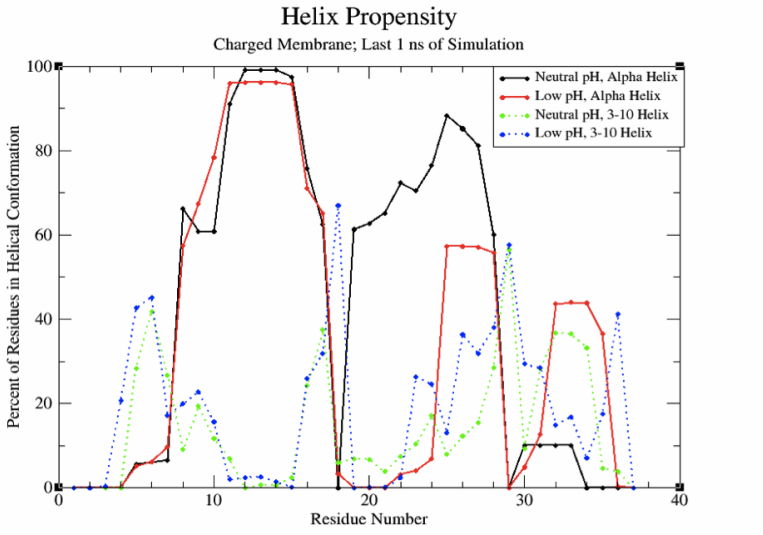

Figure 3.6: Plot of percent of residues in a helical conformation, as determined by the DSSP secondary structure protocol, for amidated hIAPP in charged membranes. The solid black and red lines represent alpha-helical character for hIAPP at neutral and low pH respectively. The dotted green and blue lines represent 3-10 helical character, again for hIAPP at neutral and low pH.

performed the DSSP algorithm for secondary structure identification on all of the structures selected and counted the frequency each residue occupied either alpha helical or 3-10 helical secondary structure, as done for the physiologically simulated peptide. These results were plotted for each residue as a fraction of occupancy of the last nanosecond of simulation time compared to the results

for the peptide in physiological conditions, i.e. neutral pH (Figure 3.6). Here, we see a large drop in alpha helical character in the low pH form from residues His-18 to Gly-24 compared to hIAPP at higher pH, consistent with the alpha-hydrogen chemical shifts for the region and the reduced strength of helical conformation from NOEs for the C-terminal helical part of the peptide. This unfolding of the peptide around His-18 sheds light on the increased flexibility of the hIAPP at this Histidine in low pH compared to physiological pH.

We were also interested in examining whether the change in pH affects how hIAPP binds to anionic micelles, so we repeated the method of buried residue calculations, which we used previously to explore the effects of membrane charge, on the structures from the last nanosecond of our simulations for hIAPP in low pH with an anionic membrane present. We then plotted these results against those found previously for the physiological pH results with the anionic membrane present. (Figure 3.7). This plot would suggest that the low pH form has perhaps slightly tighter binding,

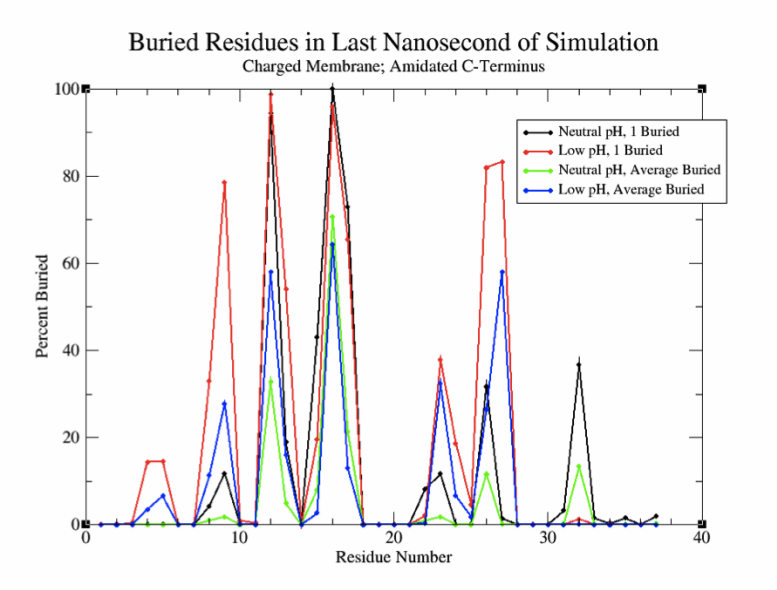

Figure 3.7: Plot of percentages of residues buried in the hydrophobic region in the last 1 ns of simulation time, for amidated hIAPP. The red and black lines, for neutral pH and low pH respectively, correspond to a percentage of system coordinates over the last 1 ns where a given residue has at least 1 atom with a Z-coordinate of magnitude $\leq$ 12.5 $\mathring{A}$. The blue and green plots, again for neutral pH and low pH respectively, correspond to a summation of the percent of atoms for each residue with Z-coordinates of magnitude $\leq$ 12.5 $\mathring{A}$ over the last 1 ns.

as evidenced by the increase in buried residues from Lys-1 to Arg-11. However, the main difference between the two peptides is actually a slight rotation of the N-terminal helix, which results in slightly different buried residues, primarily in residues Ala-13 and Phe-15. By examining the representative structure for hIAPP at low pH with the anionic membrane present, we can see that this is most likely due to the protonated Histidine being pulled towards the anionic surface, causing the N-terminal helix to rotate with it, which could potentially be another factor in the decrease in oligomer and fibril formation in low pH conditions.

3.1.3 The Free Acid C-terminus alters the Interactions of the C-terminal Helix with Anionic Membranes

We then wanted to check if the reduction of fibril formation for hIAPP with a free acid C-terminus is partially the result of the negatively charged free acid C-terminus affecting hIAPP-anionic membrane interactions. So, for our simulations of the free acid hIAPP we again performed the analysis of secondary structure and buried residues for the structures from the last nanosecond of the replica at 298K, with the plot of buried residues given in Figure 3.8. We noticed little difference at the secondary structure level between the amidated and the free acid C-terminus data; the free acid did have a kink in the helical structure around residue Cys-7, but this is most likely frustration due to the disulfide bond between Cys-7 and Cys-2 that is unrelated to the C-terminus at the other end of the peptide. The plot of buried residues reveals some differences in the C-terminal half of

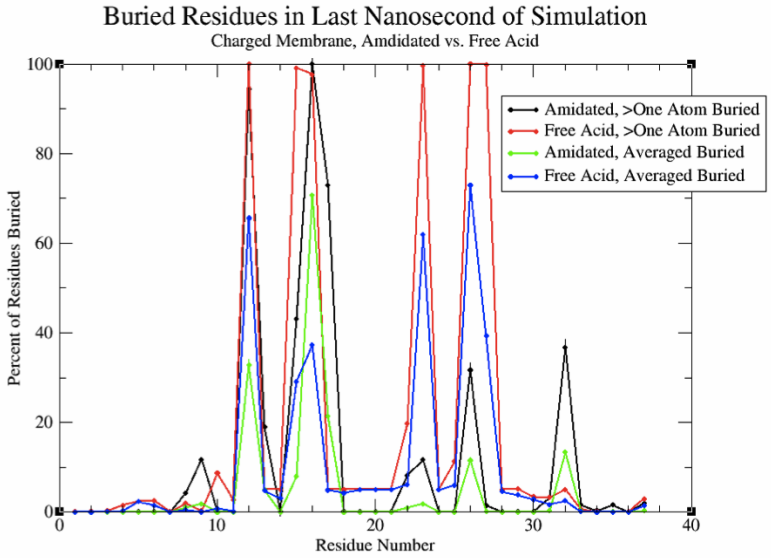

Figure 3.8: Plot of percentages of residues buried in the hydrophobic region in the last 1 ns of simulation time, for hIAPP. The red and black lines, for amidated and free acid C-terminus respectively, correspond to a percentage of system coordinates over the last 1 ns where a given residue has at least 1 atom with a Z-coordinate of magnitude $\leq$ 12.5 $\mathring{A}$. The blue and green plots, again for amidated and free acid termini respectively, correspond to a summation of the percent of atoms for each residue with Z-coordinates of magnitude $\leq$12.5 $\mathring{A}$ over the last 1 ns.

the peptide that are rather distinct. In particular, from residue Ser-28 to Tyr-37 in the free acid C-terminus, we see almost no buried residues, even at Val-32 where the Amidated peptide has a population of buried residues. This is most likely the result of the negatively charged free acid C-terminus being repelled by the anionic membrane surface. For the free acid structures, we also find that residues Phe-23, Ile-26, and Leu-27 are buried for the entirety of the last nanosecond, and residue Phe-23 and Ile-26 are buried quite deeply (greater than 60% of the residue buried). One might suggest that, as the C-terminal portion of the peptide is more flexible and interacts more weakly with the anionic membrane even with the amidated terminus, that it is possible that in the last nanosecond of our amidated hIAPP simulation we simply did not sample conformations where these hydrophobic residues were buried. However, comparisons to the buried fractions for these

residues in our simulations at low pH and without membrane charge show that, although these residues are buried with a higher frequency than the physiological hIAPP in anionic membrane, they still have different buried residues and depths of burial compared to the free acid C-terminus peptide. Based upon this result, I would suggest that when the free acid termini is repelled by the anionic membrane, the rest of the peptide undergoes conformational shifts that stabilize these hydrophobic residues being deeply buried in the hydrophobic core of the membrane. It is also worth noting that, if we examine Table 2.1, these buried residues are within 2 residues of the three 'beta-sheet blocking' proline substitutions in the non-amyloidgenic rat IAPP. This would suggest that some portion of the reduction in the rate of fibrillation for free-acid hIAPP is due to Phe-23, Ile-26, and Leu-27 being sequestered in the membrane where they are unable to participate in fibril formation.

3.1.4 Simulations of Rat IAPP Suggests Previously Unobserved Behavior

For our simulations of the non-amyloidgenic, non-toxic rat IAPP we noticed a distinct difference in the behavior of the peptide immediately. With our simulations of hIAPP, although our initial set up had the NMR structure lying parallel to the membrane surface in the membrane center (center of mass in the Z direction of about 0), the peptide rapidly diffused to the membrane surface, with only a very small minority of states forming a transmembrane structure. For rat IAPP, although the NMR structures were started in the same manner as hIAPP, clustering the structures for the last nanosecond of simulation revealed rIAPP had a very strong tendency to form a transmembrane alpha-helix in both anionic and zwitterionic membranes (Representative structures for both membranes are visualized in Figure 3.9). We can see from these figures that the

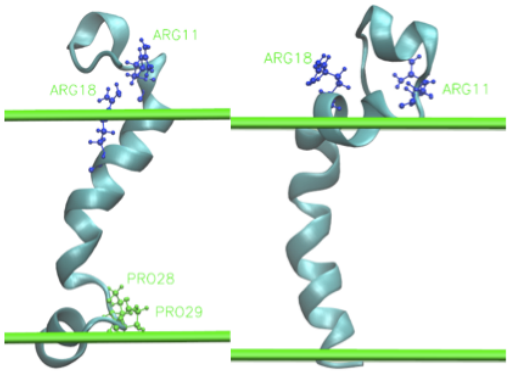

Figure 3.9: Two representative structures for rIAPP, the left in a charged membrane and the right in an uncharged membrane. The Arg-18 that is mutated from H-18 in hIAPP is depicted explicitly, as well as some of the Proline mutations.

transmembrane portion of the peptide is comprised of the C-terminal half of the peptide, where the majority of the residues that differ from rIAPP to hIAPP are located. Given that two of these mutations are from polar Serine in hIAPP to Proline in rIAPP, and the other mutations in the region are mostly from one hydrophobic group to another, it is not unexpected that rIAPP would be more stable as a transmembrane alpha helix.

To explore whether rIAPP has stable peripherally bound membrane states analogous to hIAPP, we repeated our simulations but with the initial NMR structures translated 12.5$\mathring{A}$ in the Z direction.

We then clustered the structures from the last nanosecond and performed both secondary structure analysis with the DSSP and buried residue analysis, with the procedures described previously for hIAPP. We plotted the helical content in the last nanosecond for rIAPP peripheral to an anionic membrane as well as the helical content for hIAPP in anionic membranes for comparison (Figure 3.10). We can see that for rIAPP peripheral to charged membranes, the substitution of Arg-18 for

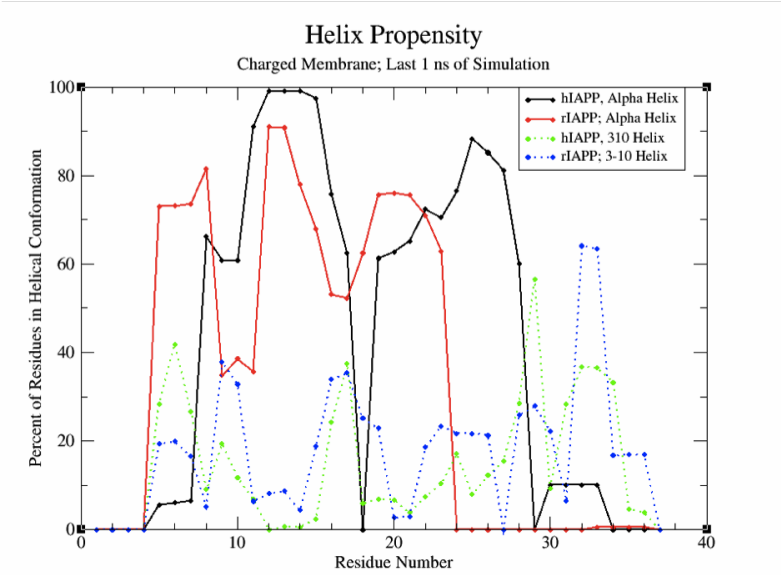

Figure 3.10: Plot of percent of residues in a helical conformation, as determined by the DSSP secondary structure protocol, for hIAPP and rIAPP in charged membranes, where the initial structures for rIAPP were started at the periphery of the membrane. The solid black and red lines represent alpha-helical character for hIAPP and rIAPP respectively. The dotted green and blue lines represent 3-10 helical character, again for hIAPP and rIAPP.

His-18 removes the break in helical structure at this position. rIAPP also has a complete lack of alpha-helical content in the Proline substituted region from Gly-24 to Thr-30, which is the part of hIAPP that is amyloidgenic and largely alpha helical when the peptide is bound to anionic membranes. Although this helical character from Ala-5 to Leu-23 is consistent with the NMR alpha-proton chemical shifts from the Ramamoorthy group, it should be noted that their data was taken in uncharged micelles.

Given that the structure of rIAPP was produced in uncharged membranes, and the role membrane charge plays in hIAPP fibrillation, we wanted to compare our results for rIAPP placed at the periphery of charged and uncharged membranes. For this purpose, we plotted the helical content data for the last nanosecond of our simulations of rIAPP with starting structures at the periphery of charged and uncharged membranes (Figure 3.11). It is apparent from this graph that rIAPP, when given a starting structure at the periphery of an uncharged membrane, almost completely unfolded in our simulations. This result is somewhat perplexing, as it seems to be in direct disagreement with the NMR structure taken in an uncharged micelle. It is possible that our model does not reproduce strong enough rIAPP-uncharged membrane interactions to stabilize the secondary structure from the NMR data, which would explain why the anionic membrane is able to produce the correct structure. We should also consider that our GBSW set-up is meant to mimic planar bilayers, and is not directly a model for the spherical micelles that were used in the experiment.

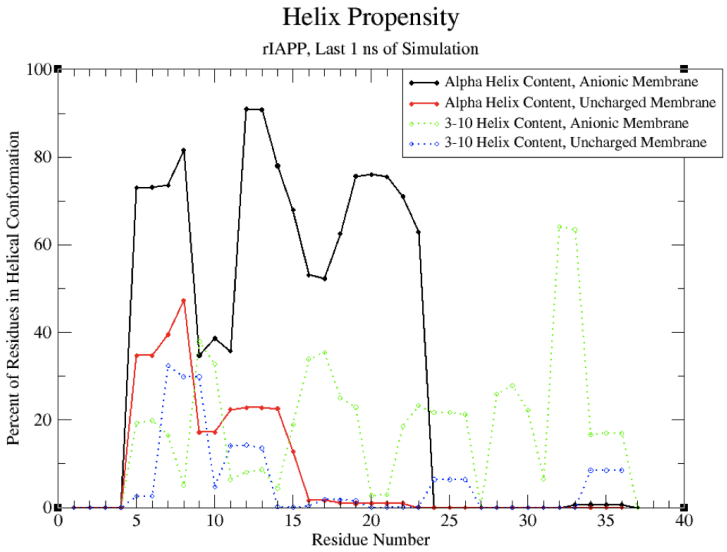

Figure 3.11: Plot of percent of residues in a helical conformation, as determined by the DSSP secondary structure protocol, for rIAPP in anionic and uncharged membranes based upon starting structures at the periphery of the membrane surface. The solid black and red lines represent alpha-helical character for rIAPP in charged and uncharged membranes respectively. The dotted green and blue lines represent 3-10 helical character, again for rIAPP in charged and uncharged membranes.

It is possible that effects of the micelle structure are reducing the energetic penalty of the burial of a portion of the peptide in the uncharged membrane, which could help stabilize alpha-helical secondary structure. This is consistent with our earlier rIAPP results that showed at least parts of the structure are stable as a transmembrane alpha helix. It should be noted as well that the GBSW membrane model is known to produce an energetic cost for the insertion of a peripheral peptide into the membrane that is high enough to prevent sampling of an insertion event in the timescales we simulated, which may have also hindered our ability to produce the correct rIAPP structure in uncharged membrane. Lastly, it should also be noted that, if the energy barrier for the insertion of the peptide across the membrane is too high for it to occur at physiological temperatures, then our transmembrane C-terminal helix is an artificial state that arises from our given initial conditions, and that it will not be observed in experiment. Further experimentation is needed to probe these possibilities.

3.2 Simulations of M2 Proton Channel

3.2.1 GBSW Test for Histidine

As previously noted, the first step in our simulations for the M2 Proton Channel was to examine whether the combination of our GBSW implicit membrane and CPHMD would produce reasonable calculations for pKa values. For our simulations of just Histidine, we extracted the lambda coordinates and found that even at a pH of 1, there was no protonated Histidine. This would suggest that our calculated pKa for Histidine would be much less than 1. This data was also poorly resolved, as at low pH we had a very high number of non-physiological mixed states, which can't be used to calculate fraction unprotonated. These two issues are not entirely unexpected, as in essence we

are exploring the free energy cost of transferring a proton from aqueous solution to vacuum. The poor resolution is then due to the free energy of protonation being much greater than the barrier energy we are using to bias our lambda values away from mixed states. This result would seem to suggest that the water molecules on the inside of the channel contribute polarization and/or hydrogen bonding that stabilizes the positive charges of the protonated Histidine.

3.2.2 PHMD of Tetrameric M2-TMP Recapitulates Experimenta pKa Shifts and Tautomer Fractions

Because of the difficulties with combining CPHMD with the GBSW implicit membrane model, we decided to instead model our peptide in aqueous phase. To take a look at our simulations for the full M2-TMP simulations, we extracted the lambda values from the replica at 298K for each pH and calculated the fraction of Histidine residues that were unprotonated. We then plotted this fraction unprotonated as a function of pH, and fit the curve for each unique Histidine to the Henderson-Hasselbalch equation (This is given in Figure 3.12 for the first trial). The pKa values that are calculated from this fitting for all three trials are given in Table 3.2. We can see in

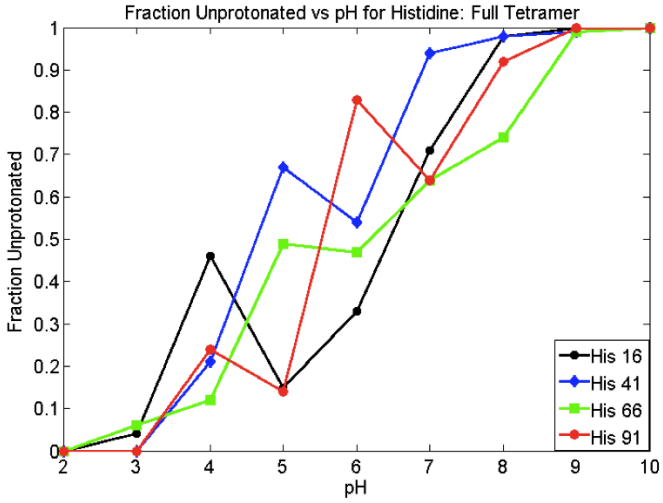

Figure 3.12: Plot of the fraction of unprotonated Histidine vs pH in the 1 ns of simulation for each Histidine in the channel.

Figure 3.12 the issues with this method, as none of the curves plotted for a given Histidine forms a smooth sigmoidal curve typical of titration. By looking at our calculated pKa values, we see that averaging over all trials for each Histidine residue produces pKas of about 5.5 for all Histidines, with negligible splitting or shifting of pKa values. Even if we attempt to naively correct these average values by averaging first the lowest pKa from each trial, then the next lowest pKa, and so on, we still see very little perturbance from the pKa of Histidine in solution. This result can be explained quite simply by considering that for our simulations of M2-TMP, where we have four Histidines on a homotetrameric protein with a starting structure that displays a high degree of four-fold rotational symmetry, there is no reason to expect that a given Histidine will consistently be the Histidine that has its pKa shifted down rather than up from one simulation to another. In order to get meaningful pKa calculations, we must instead take a page from the experimental work and consider an ensemble of four Histidines that can accept and release protons as the pH changes.

	Histidine 16	Histidine 41	Histidine 66	Histidine 91
Trial 1	4.0	5.6	6.5	5.5
Trial 2	4.9	5.9	4.1	6.4
Trial 3	6.4	4.8	6.0	5.6
Mean	5.1	5.4	5.5	5.8
Adjusted Mean	4.3	5.3	5.8	6.4

Table 3.2: Calculated pKas for the Four Histidines in our simulations of the full M2-TMP tetramer.

To implement this ensemble method, we plotted instead the sum of the fractions unprotonated for all Histidines as a function of pH (this is given in Figure 3.13 for the first set of trials). In order to fit this, we must now fit to the sum of four separate Henderson-Hasselbalch equations,

$$S = \sum_i \frac{1}{1 + 10^{n_i(pKa_i - pH)}} \tag{3.1}$$

where S is the fraction unprotonated and n_i and pKa_i are the hill coefficient and the pKa for the ith Histidine. The fit line for this equation is also plotted in Figure 3.13, and the pKa values from the fit are given in Table 3.3. We can see that this method yields a smooth titration curve that looks like the sum of sigmoidal curves.

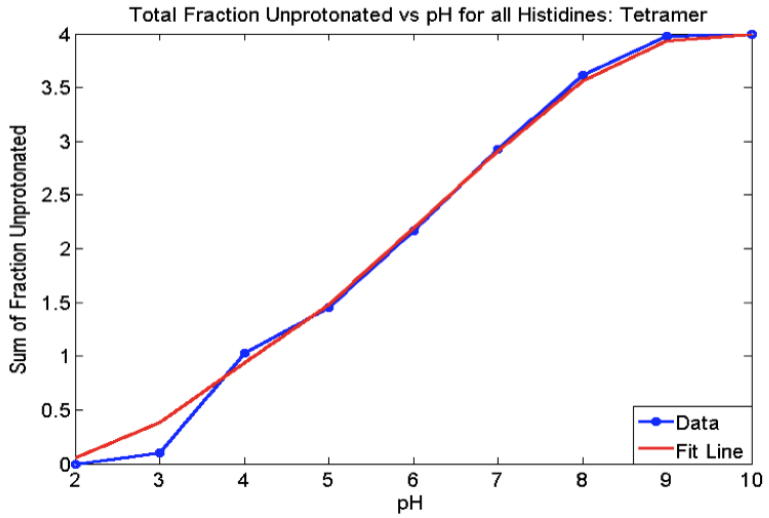

Figure 3.13: Plot of the sum of fraction of unprotonated Histidine vs pH in the 1 ns of simulation all Histidine in the channel.

It is also immediately apparent that these calculated pKa values are in much better agreement with the NMR data; averaged over all three trials, we have two higher pKas with similar values, one pKa near that of Histidine in solution, and one much lower pKa. The third protonation had a pKa of 5.2 $\pm$ 0.6, which is fairly consistent with the pKa of 6 for the opening of the channel from conductance studies. Although our approach does not produce the full magnitude of pKa shift for the first two Histidines within error bars (compare 7.1 to 8.2), we have reproduced the proper trend.

We also wanted to explore whether CPHMD of the full channel would be able to correctly produce the Tautomer fractions for the fully unprotonated channel in agreement with the NMR

	Histidine 1	Histidine 2	Histidine 3	Histidine 4
Trial 1	3.9	5.8	6.7	6.9
Trial 2	3.0	4.8	6.9	7.3
Trial 3	3.2	5.1	6.4	7.9
Mean	3.4	5.2	6.7	7.3
NMR Values	<5	6.3	8.2	8.2

Table 3.3: Calculated pKas for the Four Histidines in our simulations of the full M2-TMP tetramer, using the ensemble method. Note we have numbered our Histidines from 1 to 4 to reflect that rather than probing individual Histidines we are looking at successive protonations for the Histidine ensemble.

work of Hong *et al.*. We extracted x values from the simulations for the full M2-TMP tetramer at pH of 9, where we found no protonated Histidine present (8.5 in the NMR experiment) and calculated fraction of Histidine with its single proton on the epsilon nitrogen. We found a fraction of 73 ± 9% of epsilon-protonated Histidine, which is within error bars of the experimental value of 70%.

3.2.3 Imposed Twofold Symmetry is Able to Model Key Gating Features

As we did for our simulation of the M2-TMP tetramer, we extracted lambda coordinates and calculated fractions unprotonated for our simulations of the dimer with one 180 degree rotational image (imposed dimer of dimers) and of the monomer with four fold symmetry applied (imposed homotetramer. The values for the fractions unprotonated from the simulations of the dimer (for which only 2 Histidines were explicitly represented) were fit to pKas using equation 3.1 summed for 2 residues. For the simulations for the monomer with images, it was sufficient to use equation 1.32, as only one Histidine was present. The results averaged over all three trials are provided in Table 3.4, with the results for the explicit tetramer included for comparison. It is apparent

	Histidine 1	Histidine 2	Histidine 3	Histidine 4
Explicit Tetramer	3.4 ± 0.5	5.2 ± 0.6	6.7 ± 0.3	7.3 ± 0.5
Dimer of Dimers Imposed	5.4 ± 0.7	7.1 ± 0.6		
Homotetramer Imposed	4.8 ± 0.2			
NMR Values	<5	6.3 ± 0.3	8.2 ± 0.2	8.2 ± 0.2

Table 3.4: Mean values for all trials of the pKas calculated for the Histidines in the simulations of the Explicit M2-TMP Tetramer, the M2-TMP Dimer with Two-Fold Symmetry, and the M2-TMP Monomer with Four-Fold Symmetry, with NMR data provided for comparison.

from the results in Table 3.4 that the monomer with fourfold rotational symmetry is a very poor approximate for the channel, with a mean pKa of 4.8 ± 0.2 that is too low to correspond to the opening of the channel. This is consistent with the NMR data of the Cross group, which already revealed some degree of cooperatively and asymmetric interactions play a role in the gating of the channel. The results for the dimer with twofold rotational symmetry were far more promising, with the pKa for the second protonation event, 5.4 ± 0.7 being within error bars of the approximate pKa for the channel opening of 6 as well as the pKa for the protonation event that opens the channel from the NMR data of 6.3 ± 0.3. The dimer of dimers model also produced an inflated pKa for the first protonation event that is predicted by the Cross group to produce a dimer of dimers structure, although we were not within error bars of the experimentally calculated model (0.3 pKa unit difference). Inability to predict the full pKa shift for the first two protonation events

was also an issue with the explicit tetramer, so this is most likely an issue with the method of modeling the system rather than the dimer with imposed twofold symmetry being a bad model for the pH gating effects of the channel. Our results would suggest therefore that the dimer or dimers model is an accurate representation for the pH gating of the channel, based upon its ability to produce these reasonable pKa values.

3.2.4 Interactions of Histidines Alone are not Able to Predict pKa Splitting

For our simulations of just Histidine in an explicit tetramer and a dimer with one rotational image, we extracted lambda and x values from all three trials with the identical procedure as utilized for the full M2-TMP. We then fit this data to pKas, employing the ensemble method of pKa fitting (equation 3.1) when appropriate. The calculated pKa values are given in Table 3.5, along with the NMR values for reference. In these simulations of only Histidine residues held in place with

	Histidine 1	Histidine 2	Histidine 3	Histidine 4
Explicit Tetramer	6.7 ± 0.5	8.2 ± 0.6	8.4 ± 0.5	8.8 ± 0.1
Dimer of Dimers Imposed	7.9 ± 0.1	8.5 ± 0.1		
NMR Values	<5	6.3 ± 0.3	8.2 ± 0.2	8.2 ± 0.2

Table 3.5: Mean values for all trials of the pKas calculated for the Histidines in the simulations of the extracted Histidine residues from the NMR data. Listed are the values for a Histidine Tetramer and the Histidine Dimer with Two-Fold Symmetry with NMR data provided for comparison.

harmonic restrains, we calculate a mean pKa for the first two protonation events of roughly 8.5, compared to 8.2 from NMR experiment. The average calculated pKa values for these two Histidines are actually higher than the values from the NMR data (albeit mostly within error bars) in contrast to our studies of the full M2-TMP. The pKas for the third and fourth protonation events, however, were greatly overestimated in these simulations, at 8.2 ± 0.6 and 6.7 ± 0.5, as calculated for the explicit tetramer of Histidine residues, compared to 6.3 ± 0.3 and <5 from experiment. Although some degree of splitting is observed, particularly in the case of the final protonation for the explicit tetramer, we see a large reduction in the magnitude of these pKa splittings compared to both the NMR data and our previous simulation results for the full M2-TMP. We would suggest based upon these findings that structural rearrangements as the channel shifts from its open to closed form and/or interactions with other residues (of particular interest is Trp-41, the other gating residue, which can offer a hydrogen bond) are partially responsible for the Histidine pKa behavior that is observed in the NMR experiment of the Cross group.

Chapter 4

Concluding Discussion

4.1 Novel Insights in IAPP Fibril Formation

This work suggests several new structural features that are necessary for hIAPP to induce toxicity and/or oligomerize. Perhaps most ubiquitous was that hIAPP at low pH, hIAPP bound to uncharged membranes, and rIAPP all lacked the distinct structural bend at His-18; the other preparations of hIAPP lost secondary structure in this region and the rat form has no break in the helix. These results, taken together with the experimental evidence that this structural kink corresponds approximately to the bend between beta sheets in the fibril structure, suggest that the kink at Histidine-18 is involved in nucleating the eventual formation of the beta sheet fibrils.

The other commonly observed possible mechanism affecting amyloid fibril formation noted in our simulations was the increase in the burial of residues in the 'amyloidgenic' region region of the peptide from Ser-20 to Ala-30 in the membrane. This phenomenon was observed to some degree for all of the hIAPP simulations performed in conditions known to slow the rate of fiber formation, including the low pH, free acid C-terminus, and uncharged membrane results. It would be our suggestion that, in these conditions, the conformations where the amyloidgenic residues are sequestered in the membrane are stabilized, preventing these residues from participating in IAPP-IAPP interactions that will lead to fiber formation.

These suggestions would, of course, benefit from experimental verification that they are authentic occurrences in hIAPP-membrane interactions. Still, there are further simulations we can run to probe these results. In particular, it is worth exploring hIAPP with some small number of residues mutated to the rIAPP equivalent, to shed light onto which mutations are causing the various fiber interrupting behaviors. Also of interest would be simulating the N-terminal and C-terminal halves of the peptide separately, to examine which aspects of the properties of hIAPP bound to anionic membranes are properties that require the full peptide.

4.2 Towards Complete Modeling of pH Gating in the M2 Proton Channel

Our results have shown that CPHMD is able to reproduce the pKas for the gating Histidines in the M2 Proton Channel and thus the pH gating behavior. However, as we simulated the protein in high dielectric aqueous phase, we could not possibly use this approach to produce high accuracy

structural clusters for the opening and closing of the channel at a given pH. To obtain a complete picture for the conformational rearrangements in M2-TPM pH gating (and perhaps improve our pKa calculations), we need some new method of modeling the effects of the membrane on the peptide backbone, while still allowing for water to screen the interactions in the channel pore. Fortunately, the recent developments of a version of CPHMD that can work with explicit solvent and a model of the GBSW implicit membrane with a high dielectric cylinder representing a water pore provide multiple options for surmounting this future obstacle.

Bibliography

[1] Cowie C.C., Rust K.F., Byrd-Holt D.D., Gregg E.W., Ford E.S., Geiss L.S., Bainbridge K.E., Fradkin J.E., Prevalence of Diabetes and High Risk for Diabetes using A1C Criteria in the U.S. Population in 1988-2006. *Diabetes Care* 2010, **33**, 562-568

[2] Larsen P.R., Kronenberg H.M., Melmed S., Polonsky K.S., *Williams Textbook of Endocrinology*; Saunders, 2007

[3] Sipe J. and Cohen A., Review: History of the Amyloid Fibril. *Journal of Structural Biology* 2000, **130**, 88-98.

[4] Opie E.L., On relation of chronic interstitial pancreatitis to the island of Langerhans and to diabetes mellitus. *J. Exp. Med.* 1901, **5**, 397-428

[5] Weichselbaum A. and Stangl E., Zur Kenntnis der feineren Veranderungen das Pankreas bei Diabetes mellitus. *Wien klin Wochenshr* 1901, **14**, 968-972

[6] Alzheimer A., Uber eine Eigenartige Erkrankung der Himrinde. *Allgemeine Zeitschrift fr Physchiatrie und Phychish-Genrichtliche Medizin.* 1907, **64**, 146-148

[7] Cohen A.S., Calkins E., Electron Microscopic Observations on a Fibrous Component in Amyloid of Diverse Origins. *Nature* 1959, **183**, 1202-1203

[8] Bonar L., Cohen A.S., Skinner M.M., Characterization of the Amyloid Fibril as a Cross-Beta Protein. *Exp. Biol. Med.* 1969, **131**, 1373-1375

[9] Westermark P., Wernstedt C., Wilander E., Sletten K., A Novel Peptide in the Calcitonin Gene Related Peptide Family as an Amyloid Fibril Protein in the Endocrine Pancreas *Biochemical and Biophysical Research Communications* 1986, **140**, 827-831

[10] Westermark P., Wernstedt C., Wilander E., Hayden D.W., O'Brien T.D., Johnson K.H., Amyloid Fibrils in Human Insulinoma and Islets of Langerhans of the Diabetic Cat are Derived from a Neuropeptide-like Protein also Present in Normal Islet Cells. *PNAS* 1987, **84**, 3881-3885

[11] Lorenzo A., Razzaboni B., Weir G., Yanker B., Pancreatic Islet Cell Toxicity of Amylin Associated with Type-II Diabetes Mellitus. *Nature* 1994, **368**, 756-760

[12] Yanker B., Duffy L., Kirschner D., Neurotrophic and Neurotoxic Effects of Amyloid Beta Protein: Reversal by Tachykinin Neuropeptides. *Science* 1990, **250**, 279-282

[13] Konarkowska B., Aitken J., Kistler J., Zhang S., Cooper G., The Aggregation Potential of Human Amylin Determines its Cytotoxicity Towards Islet Beta-Cells. *FEBS J.* 2006, **273**, 3614-3624

[14] Janson J., Ashley R., Harrison D., McIntyre S., Butler P., The Mechanism of Islet Amyloid Polypeptide Toxicity is Membrane Disruption by Intermediate-Sized Toxic Amyloid Particles. *Diabetes* 1999, **48**, 491-498

[15] Knight J.D., Miranker A.D., Phospholipid Catalysis of Diabetic Amyloid Assembly. *Journal of Molecular Biology* 2004, **341**, 1175-1187

[16] Clement L., Kim-Sohn K.A., Magnan C., Kassis N., Adnot P., Kergoat M., Pancreatic Beta-Cell Alpha2A Adrenoceptor and Phospholipid Changes in Hyperlipidemic Rats. *Lipids* 2002, **37**, 501-506

[17] Brender J.,Lee E.,Marchello C., Gafni A., Steel D., Ramamoorthy A., Amyloid Fiber Formation and Membrane Distription are Separate Processes Localized in Two Distinct Regions of IAPP, the Type-2-Diabetes-Related Peptide. *JACS* 2008, **130**, 6424-6429

[18] Nanga R.P., Brender J.R., Xu J., Hartman K., Subramanian V., Ramamoorthy A., Three-Dimensional NMR Structure of Rat Islet Amyloid Polypeptide in DPC Micelles. *J. Am. Chem. Soc.* 2009, **131**, 8252-8261

[19] Patil S., Shihau X., Sheftic S., Alexandrescu A., Dynamic Alpha-Helix Structure of Micelle-bound Human Amylin. *Journal of Biological Chemistry* 2009, **284**, 11982-11991

[20] Nanga R., Brender J., Vivekanandan S., Ramamoorthy A., Structure and Membrane Orientation of IAPP in its Natively Amidated Form at Physiological pH in a Membrane Environment. *Biochim.Biophys.Acta* 2011, **1808**, 2337-2342

[21] Lazardis T., Implicit Solvent Simulations of Peptide Interactions with Anionic Lipid Membranes. *Proteins: Structure, Function, and Bioinformatics.* 2005, **58**, 518-527

[22] Im W., Feig M., Brooks C.L. III, An Implicit Membrane Generalized Born Theory for the Study of Structure, Stability, and Interactions of Membrane Proteins. *Biophysical Journal* 2003, **85**, 2900-2918

[23] World Health Organization. Influenza (Seasonal). http://www.who.int/mediacentre/factsheets/fs211/en/ (Accessed April 1, 2012).

[24] Deyde V.M., Xu X., Bright R.A., Shaw M., Smith C.B., Zhang Y., Shu Y., Gubareva L.V., Cox N.J., Kilmov A.I., Surveillance of Resistance to Adamantanes among Influenza A(H3N2) and A(H1N1) Viruses Isolated Worldwide. *J. Infect. Dis.* 2007, **192**, 249-257

[25] Compans R.W., Meier-Ewert H., Palese P., Assembly of Lipid-Containing Viruses. *Journal of Supramolecular Structure* 1974, **2**, 496-511

[26] Murphy R.F., Powers S., Cantor C.R., Endosome pH Measured in Single Cells by Dual Fluorescence Flow Cytometry: Rapid Acidification of Insulin to pH 6. *J. Cell Bio.* 1984, **98**, 1757-1762

[27] Schnell J.R., Chou J.J., Structure and Mechanism of the M2 Proton Channel of Influenza A Virus. *Nature* 2008, **451**, 591-600

[28] Nishimura K., Kim S., Zhang L., Cross T.A., The CLosed State of a H+ Channel Helical Bundle Combining Precise Orientational and Distance Restrains from Solid State NMR. *Biochemistry* 2002, **41**, 13170-13177

[29] Wang C., Takeuchi K., Pinto L.H., Lamb R.A., The Ion Channel Activity of Influenza A Virus M2 Protein: Characterization of the Amantadine Block. *J. Virol.* 1993, **67**, 5585-5594

[30] Wang C., Lamb R.A., Pinto L.H., Activation of the M2 Ion Channel of Influenza Virus: A Role for the Transmembrane Domain Histidine Residue. *Biophys. J.* 1995, **69**, 1363-1371

[31] Tang Y., Zaitseva F., Lamb R.A., Pinto L.H., The Gate of the Influenza Virus M2 Proton Channel is Formed by a Single Tryptophan Residue. *Journal of Biological Chemistry*, 2002, **277**, 39880-39886

[32] Hu J., Fu R., Nishimura K., Zhang L., Zhou H.X., Busath D.D., Vijayvergiya V., Cross T.A., Histidines, Heart of the Hydrogen Ion Channel from Influenza A Virus: Toward and Understanding of Conductance and Proton Selectivity. *PNAS*, 2006, **103**, 6865-6870

[33] Hu F., Wenbin L., Hong M., Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science*, 2010, **22**, 505-508

[34] Levinthal C. How to Fold Graciously. *Mossbauer Spectroscopy in Biological Systems: Proceedings of a meeting held at Allerton House* 1969, **22-24**

[35] Sugita Y., Okamoto Y., Replica-Exchange Molecular Dynamics Methods for Protein Folding. *Chemical Physics Letters*, 1999 **314**, 141-151

[36] Khandogin J., Brooks C.L. III, Linking Folding with Aggregation in Alzheimer's Beta-Amyloid Peptides. *PNAS* 2007, **104**, 16880-16885

[37] Born M., Volumen und Hydratationswarme der Ionen. *Z. Phys.* 1920, **1**, 45-48

[38] Im W. and Brooks C.L. III, Interfacial folding and membrane insertion of designed peptides studied via molecular dynamics simulations. *Proc. Natl. Acad. Sci. USA* 2005, **102**, 6771-6776

[39] Bu L., Im W., Brooks C.L. III, Membrane assembly of simple helix homo-oligomers studied via molecular dynamics simulations. *Biophys. J.* 2009, **92**, 854-863.

[40] Lee M.S., Salsbury F.R., Brooks C.L. III, Constant-pH Molecular Dynamics Using Continuous Titration Coordinates. *Proteins* 2004, **56**, 738-752

[41] Khandogin J., Brooks C.L. III, Constant pH Molecular Dynamics with Proton Tautomerism. *Biophys. J.* 2005, **89**, 141-157

[42] Khandogin J., Chen J., Brooks C.L. III, Exploring Atomistic Details of pH-dependent Peptide Folding. *Proc. Natl. Acad. Sci.* 2006, **103**, 18546-18550

[43] Arthur E.J., Yesselman J.D., Brooks C.L. III, Predicting extreme pK(a) shifts in staphylococcal nuclease mutants with constant pH molecular dynamics. *Proteins* 2011, **79**, 3276-3286

[44] Feig M., Karanicolas J., Brooks C.L. III, MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model.* 2004, **22**, 377-395.