ELUCIDATING THE MECHANISM OF ACTION OF UNNATURAL AMINO ACID CONTAINING ANTIMICROBIAL PEPTIDES IN MEMBRANE ENVIRONMENTS

by

Tiffany D Clark

November 6, 2012

Directors of Dissertation: Rickey Hicks &
Libero Bartolotti

Interdisciplinary Doctoral Program in Biological Sciences
Major Department: Chemistry

Organism resistance continues to develop to the currently available antimicrobial compounds necessitating the development of innovative new therapeutic compounds with different specificities and mechanisms of action that provide acceptable therapeutic indices. Unnatural amino acid containing antimicrobial peptides could provide a novel avenue for the development of therapies with improved efficacy and pharmacokinetics over natural amino acid containing peptides which are prone to protease degradation.

Molecular dynamics (MD) simulations of antimicrobial peptides containing unnatural amino acids have been performed using explicit water and multiple model membrane types in all-atom simulations. The structural properties of peptides were investigated using both the canonical and isothermal-isobaric ensembles to further understand the mechanism through which the collections of AMPs exert their *in vitro* activity.

Simulations with micelle membrane models were conducted at 300 K to correlate with experimental circular dichroism (CD) data showing the secondary structure the peptides adopt in

the presence of an electrostatic membrane model. Analysis of the stabilized MD trajectory reflects peptide structural consistency with experimental data.

Simulations of the peptides with bilayer model membranes were conducted at the physiologically relevant 310 K to correlate with experimental cellular activity data which demonstrated the antimicrobial activity of the peptides without providing insight into the mechanism through which the activity was achieved. Long time scale simulations have noted distinct differences between bilayers in the presence of AMPs as compared to those without the peptide. Mixed bilayers with an anionic charge modeled bacterial membranes while a confluent zwitterionic bilayer modeled the mammalian membrane.

This research has demonstrated that force field parameters for unnatural amino acids can be derived from QM calculations. FF parameters derived from structures identified from a DFT approach have also been used to expand the AMBER ff03 force field. The FF parameters were able to able to model the interaction of the peptides which contain unnatural amino acids. The data is consistent with NMR data and further supported with CD spectroscopy.

ELUCIDATING THE MECHANISM OF ACTION OF UNNATURAL AMINO ACID CONTAINING ANTIMICROBIAL PEPTIDES IN MEMBRANE ENVIRONMENTS

A MOLECULAR DYNAMICS STUDY

A dissertation presented to the Faculty of the Department of Chemistry

East Carolina University

in Partial Fulfillment of Requirements for the Degree Doctor of Philosophy

By

Tiffany D. Clark

November 6, 2012

ELUCIDATING THE MECHANISM OF ACTION OF UNNATURAL AMINO ACID CONTAINING ANTIMICROBIAL PEPTIDES IN MEMBRANE ENVIRONMENTS

Tiffany D. Clark

APPROVED BY:

CO-DIRECTOR OF DISSERTATION:_____
/HEAD OF THE DEPARTMENT OF CHEMISTRY Rickey P. Hicks, PhD

CO-DIRECTOR OF DISSERTATION: _____
Libero J. Bartolotti, PhD

COMMITTEE MEMBER: _____
Donald R. Hoffman, PhD

COMMITTEE MEMBER: _____
Colin S. Burns, PhD

COMMITTEE MEMBER: _____
Anthony Kennedy, PhD

DIRECTOR OF THE INTERDISCIPLINARY DOCTORAL PROGRAM IN BIOLOGICAL SCIENCES

_____
Terry West, PhD

DEAN OF THE GRADUATE SCHOOL

_____
Paul J. Gemperline, PhD

For my Mom who taught me I could do anything I *wanted* and my Dad
who supported me through life's journey.

ACKNOWLEDGEMENT

I would first like to express my gratitude to the two gentlemen, Dr. Rickey P. Hicks and Dr. Libero Bartolotti, who have helped me gain the knowledge and experience on my way. I feel exceptionally lucky to have had two excellent mentors to help me navigate the research challenges I encountered. Dr. Hicks always provided the support and perspective when I was not able to find the best path forward. Dr. Bartolotti has been my computer guru, answered my questions at all hours of the day and night in addition to being an excellent role model.

I would like to thank my committee members, Dr. Donald Hoffman, Dr. Colin Burns, and Dr. Anthony Kennedy. I am fortunate to have had great scientists who have given their time, effort, and enthusiasm to help direct my research when I needed it most.

The Department of Chemistry at East Carolina University has been the best possible place for me to develop as a scientist. The faculty at ECU cares about their students and is always willing to teach anyone who wants to learn. Jeff Kinney at the Information Technology and Computing Systems center kept me from crashing my experiments (and everyone else's) and always answered my HPC questions with useful advice.

I would like to thank my friends and colleagues who have lightened the load, provided distraction when needed and have left me with memories I will never forget. Dr. Prasenjit Mukherjee has shared his vast scientific knowledge as well as his friendship which has helped me grow personally as well as professionally.

My education would not have been possible were not for the support, love, and encouragement of my family. They have always supported me in my educational endeavors. My sister has been my companion through the many late nights. My Mom, Myrna J. Clark, has

selflessly supported me throughout my education and my Dad, James A. Clark has supported me and tried to teach me how to balance work and play.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

A	Alanine
Å	Angstrom
ACP	Anticancer peptide
AFP	Antifungal peptide
Ala	Alanine
AMBER	Assisted Model Building with Energy Refinement
AMP	Antimicrobial peptide
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
β-Ala	Beta alanine
C	Cysteine
CACS	Center for Applied Computational Studies
CAPs	Cationic antimicrobial peptides
CD	Circular dichroism
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CG	Coarse-grain
CGMD	Coarse-grain molecular dynamics
Cmc	Critical micelle concentration
CPP	Cell-penetrating peptides
Cys	Cystine
D	Aspartic acid

DAB	Diaminobutyric acid
DFT	Density functional theory
DMPK	Drug metabolism and pharmacokinetics
DPC	Dodecylphosphocholine
DPR	Diaminopropionic acid
DWF	Debye-Waller Factor
E	Glutamic acid
EC	Electron correlation
F	Phenylalanine
FF	Force field
G	Glycine
GABA	Gamma-amino butyric acid
GAFF	General Amber Force Field
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
GROMACS	GROningen Machine for Chemical Simulations
GTO	Gaussian type orbitals
GUI	Graphical user interface
H	Histidine
HF	Hartree-Fock
His	histidine
I	Isoleucine

IC_{50}	Inhibitory concentration of 50% of the cells
Ile	Isoleucine
ITC	Isothermal titration calorimetry
LPS	Lipopolysaccharide
K	Lysine
LCAO	Linear Combination of Atomic Orbitals
LUV	Large unilamellar vesicle
Lys	Lysine
M	Methonine
MAP	Membrane active peptides
MD	Molecular Dynamics
MD_npt	Isobaric/isothermal ensemble
MD_nvt	Canonical ensemble
Met	Methionine
MIC	Minimum inhibitory concentration
MLP	Membrane-lytic peptides
MM	Molecular Mechanics
MOA	Mechanism of action
MSD	Mean-squared displacement
N	Asparagine
NAMD	NAnoscale Molecular Dynamics
NARMS	Nathonal Antimicrobial Resistance Monitoring System
NIH	National Institutes of Health

OIC	Octahydroindolecarboxylic acid
OPLS	Optimized Potentials for Liquid Simulations
ORN	Ornithine
P	Proline
PBSA	Poisson-Boltzmann Surface Area
PBCs	Periodic boundary conditions
PBS	Phosphate buffered saline
PBSA	Poisson-Boltzmann Surface Area
PDB	Protein Data Bank structural data file format
PES	Potential energy surface
PGTOS	Primitive Gaussian Type Orbitals
Phe	Phenylalanine
PME	Particle mesh Ewald
PMN	Polymorphonuclear leukocytes
POPC	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine
POPG	1-Palmitoyl-2-oleoyl-*sn*-glycero-3-[phospho-rac-(1-glycerol)]
Pro	Proline
Q	Glutamine
QM	Quantum mechanics
R	Arginine
RBC	Red blood cells (erythrocytes)
RI	Reliability index
RMS	Root mean square

RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
S	Serine
SA	Simulated annealing
SAR	Structure activity relationship
sAMP	synthetic antimicrobial peptide
SANDER	Simulated Annealing with NMR-Derived Energy Restraints
SCF	Self consistent field
Ser	Serine
SMAMPs	Synthetic mimics of antimicrobial peptides
SCF	Self-consistent field
SDS	Sodiumdodecylsulfate
SSH	Secure shell
STO	Slater-type orbitals
T	Threonine
Thr	Threonine
TIC	Tetrahydroisoquinolinecarboxylic acid
TIP3P	Three site water model
Trp	Tryptophan
Tyr	Tyrosine
V	Valine
Val	Valine
Vdw	Van der Waals

VMD	Visual Molecular Dynamics
W	Tryptophan
Y	Tyrosine
6-AHX	6-Aminohexanoic acid

CHAPTER 1

AN INTRODUCTION TO ANTIMICROBIAL PEPTIDES

1.1 INTRODUCTION

Antimicrobial peptides have evolved within species as they evolved over time. Man's search for ways to inhibit bacterial proliferation date back to the late 1860's with the work of Louis Pasteur, Robert Koch and their contemporaries.[1] The suggestion that microbes were responsible for infection and food spoilage slowly became an accepted theory within science. The discovery of chemical compounds having anti-microbial properties emerged with the use of carbolic acid (phenol) as a surgical disinfectant.[2] Since that time the quest for additional entities with antimicrobial properties has been the focus of much research. Alexander Fleming's identification of the antimicrobial enzyme lysozyme (1923) and a few years later, the serendipitous discovery of penicillin (1928)[3, 4] set the precedent for the identification of organisms producing chemicals with antimicrobial properties which offer defense to would be environmental invaders.[5] In 1975, scientists and pharmaceutical companies decided the war against infectious disease had been won and that further antibiotic development was unnecessary.[6] The results of this myopic decision without considerations for microbial resistance development led to many pharmaceutical companies abandoning further investigation into antibiotic drug development.[7] Some antibiotic resistance occurred shortly after the routine use of antibiotics began (1950's) however the continued expansion of the rate of resistance was not realized until the 1990's.[7] In 2000, the World Health Organization (WHO) realized the extent of antibiotic resistance developing worldwide,[7] and urged a more responsible utilization of antibiotics. Education regarding appropriate use of antibiotics continues today and is an ongoing process.[8] Almost since the time of penicillin's discovery drug resistance has been a problem.[5] Microbes are able to alter their cell walls to protect themselves from the antimicrobial drugs

which seek to disable or kill these organisms.[9] Emergence of numerous organisms with resistance to multiple antimicrobial therapies becomes an ever more serious problem considering the limited arsenal of compounds available and effective for the treatment of many infectious bacteria. Over the last 24 years, only 36 new antibacterial compounds were approved by the United States' Food and Drug Administration (FDA).[10] The lack of effective treatments has precipitated the writing of legislation to meet the "10 x'20 initiative outlined by the "Bad Bugs No Drugs"[10] campaign in 2003. The 10 x'20 initiative seeks to develop ten new antibacterial treatments by 2020. Through the establishment of a Limited Population Antibacterial Drug (LPAD) approval process to encourage development of compounds specifically targeting resistant bacteria.[11]

Infections acquired in hospitals have demonstrated increased resistance to current therapies and are of particular interest; the organisms include *Streptococcus* and *Enterococcus* (especially vancomycin-resistant *Enterococci* (VRE)), *Staphylococcus aureus* (methicillin-resistant *S. Aureus* (MRSA)), *Klebsiella pneumonia*, *Acinetobacter baumanii*, *Pseudomonas aeruginosa*, *Mycobacterium tuberculosis*, and *Enterobacter* species.[5, 8] The economic impact and the cost of human life has necessitated research efforts be invested in the development of effective treatments where current approved therapeutic compounds have failed. [8, 11]

Since the realization that many of the existing antibiotics are no longer able to counter these threats, the United States has even developed an agency, National Antimicrobial Resistance Monitoring System (NARMS): Enteric Bacteria, for following the rate of resistance development.[7] As is evident, there is a huge demand, despite little financial incentive, for further antimicrobial compound development with novel mechanisms of activity. It seems obvious that with the increased demand a financial incentive would coincide. The relatively short amount of

time needed to treat a patient with a successful antimicrobial compound has limited financial incentive for industrial development of new compounds. The need for the development of new antimicrobial agents cannot be understated.

1.2 MAGAININ

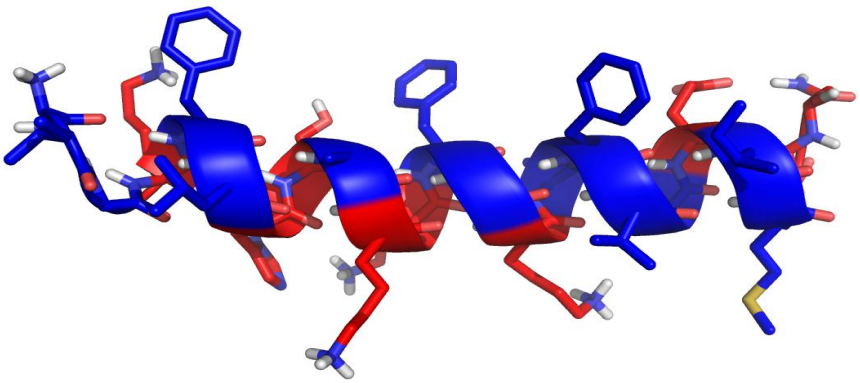


Figure 1.1 Amphipathic structure of magainin 2 PDBID: 2MAG. Hydrophilic residues have red carbon atoms. Hydrophobic residues have blue carbon atoms.

Magainins are a collection of naturally occurring cationic AMPs which were originally derived from the African tree frog, *Xenopus laevis*.[12-15] Magainin 2 is a peptide which has the following amino acid sequence: GIGKFLHSAKKFGKAFVGEIMNS. [12-15] Figure 1.1 shows the NMR derived alpha helical secondary structure as published in the protein data bank (PDB) with ID:2MAG, the structure of Magainin 2. Magainin has demonstrated broad antimicrobial activity making it of therapeutic interest. Magainin 2 was the electrostatic model upon which the unnatural amino acid containing peptides were modeled.[16, 17] The unnatural amino acid containing AMPs which resulted from the models included the unnatural amino acids beta alanine (β-Ala), gamma amino butyric acid (GABA), 6-amino hexanoic acid (6-AHX), ornithine (ORN), diamionbutanoic acid (DAB), diamionpropanoic acid (DPR), octahydroindole carboxylic acid (OIC), and tetrahydroisoquinoline carboxylic acid (TIC) (Figure 1.2).[16, 17]

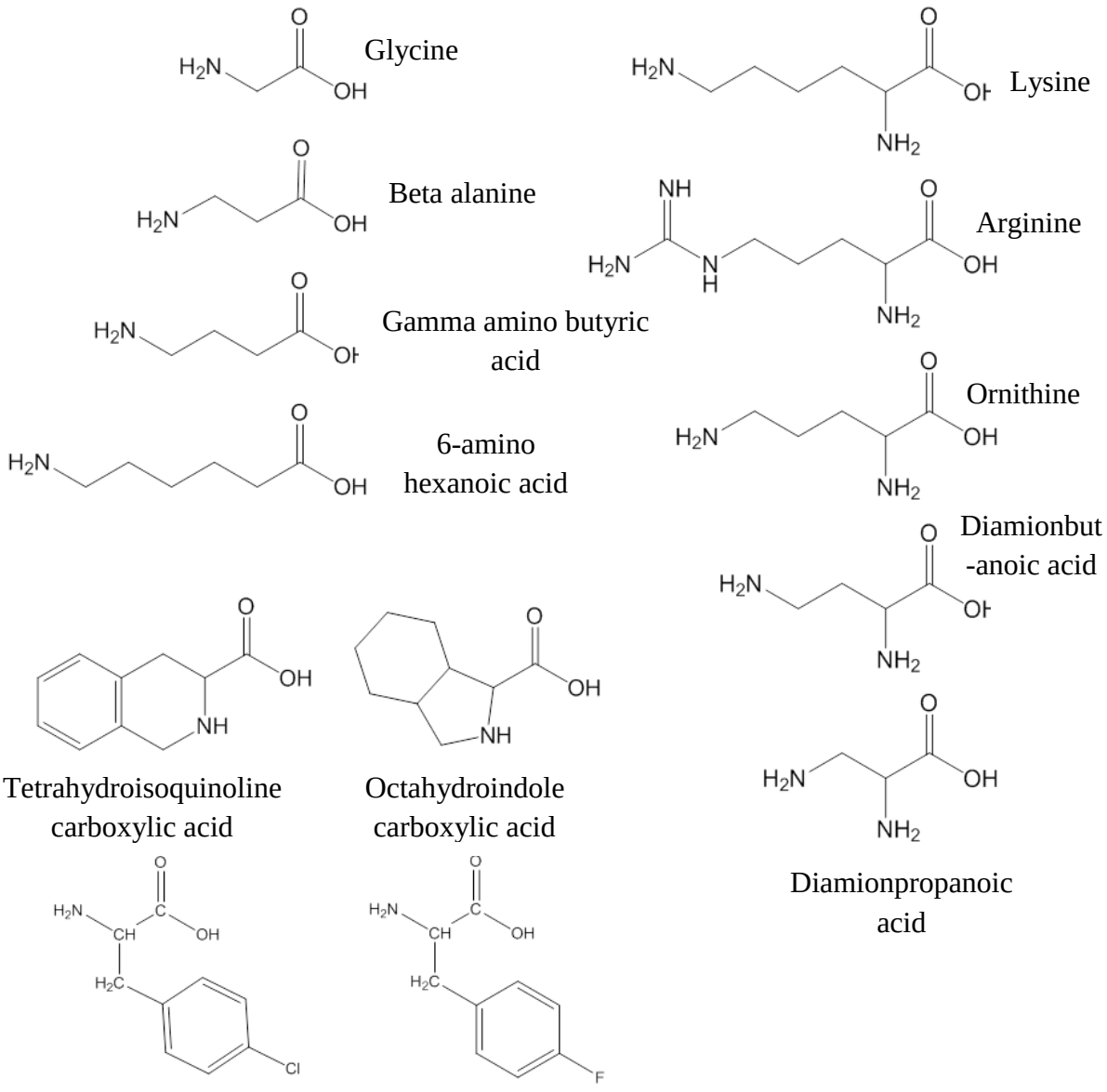


Figure 1.2 Amino acids incorporated into the AMP sequence

The skeleton of AMPs incorporating the unnatural amino acids is shown in Figure 1.3. Sterically hindered Tic-Oic dipeptides are separated by varying distances with spacer 1 which increases the number of carbon atoms in the peptide backbone. Gly, the smallest naturally occurring amino acid was used as reference for extension of the number of carbon atoms in the protein backbone. Extension to include a beta carbon in the backbone was accomplished by incorporating beta

alanine. Gamma amino butyric acid incorporated into the peptide sequence increased the flexibility of the backbone by adding gamma carbons into the backbone. The delta carbon was not synthesized but the epsilon carbon was generated by the inclusion of 6-Ahx.

1.3 ANTIMICROBIAL PEPTIDES

Antimicrobial peptides (AMPs) have evolved with species through the ages. According to Tomas Ganz, antimicrobial peptides are "polypeptides of fewer than 100 amino acids, found in host defense settings, and exhibiting antimicrobial activity at physiologic ambient conditions and peptide concentrations."[18] The presence of antimicrobial peptides appears to be ubiquitous in the biologic environment and wherever they have been sought, unique innate immune protection in the form of AMPs has been detected.

It wasn't until the early 1960's when H.I. Zeya and John K. Spitznagel characterized a cationic peptide found in rabbit and guinea pig neutrophil granules that the identification of polymorphonuclear leukocytes (PMN) (antimicrobial peptides was noted, though not fully understood) were isolated and proposed to have therapeutic potential.[19]

Concurrently, the identification of antimicrobial substances in insect hemolymph was made by Hans Boman and his confrere.[20] Francisco Garcia Olmedo and his colleagues were investigating antimicrobial peptides present in plants.[21] The discovery of human neutrophil antimicrobial peptides was not made until 1984 by Robert Lehrer, Michael Selsted, and Tomas Ganz.[22] They named their family of antimicrobial peptides "defensins" for the role they were assumed to play in protecting from microbial invaders. Robert Lehrer was responsible for initiating a search for the mechanism of action of the defensins. He identified the disruption of the cell membrane as a requirement for microbial death.[23] It was not until the early 1990's when

scientists became conversant amongst those who studied phagocytes, insects and plants and the potential for therapeutic development was realized.[18]

Due to their broad endemic nature, AMPs are good candidates as therapeutic agents. AMPs have numerous advantages over the traditional, small molecule, antimicrobial compounds. Attractive features of AMPs include reasonable host cell toxicity (relatively non-toxic), reduced resistance development from microbial invaders, and broad spectrum of activity. AMPs have demonstrated membrane disrupting capability beyond that of bacterial cells. Anticancer activity has also been identified for some of these peptides.[24] The membrane differences between normal and malignant cells make anticancer peptides (ACPs) a potential therapeutic option.[24]

The specificity of AMPs is a desirable therapeutic property. Avoiding damage to host cells while destroying alien cells decreases side effects resulting from offsite interactions which are associated with toxicity. Given the need for new, selective antimicrobial compounds with excellent toxicity profiles AMPs, and specifically antifungal peptides (AFP), are a novel collection of peptides which target the membrane surface. Antiviral and anti-parasitic AMPs have also been identified among the diverse class of peptides.[25] The dynamic activity demonstrated by naturally occurring AMPs makes them a good starting point for structure optimization and improving their therapeutic potential.

1.4 SECONDARY STRUCTURE

Anfinsen's theory outlining the correlation between the primary amino acid sequence, secondary structure, and function has led to the pursuit of secondary structure elucidation of many sequenced proteins. The secondary structure associated with an AMP is thought to directly correlate to its function. Presently naturally occurring AMPs are known to attain the

following three types of secondary structure/configuration: alpha helical, beta helix/sheet, or cyclic secondary structure. AMPs are known to have a random coil structure or a lack of secondary structure in solution, and only develop secondary structures (such as alpha helical, beta helix/sheet, or cyclic) in the presence of electrostatic charge environments such as membranes (Figure 1.4). [26,27]

AMPs are thought to require a minimum of 20 amino acids for membrane disruption

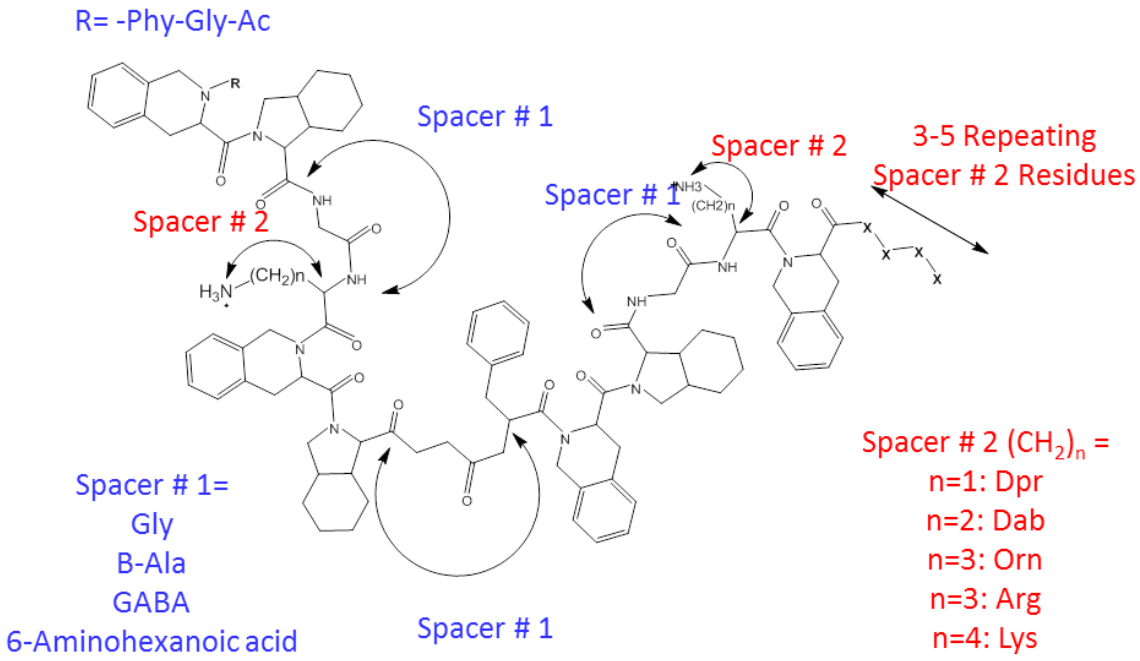


Figure 1.3 Line structure of the AMP designed with the unnatural amino acid, Tic-Oic dipeptides incorporated into the peptide sequence.

however; peptides with fewer amino acid residues have been shown to dimerize in order to disrupt the membrane lipid bilayer.[24] The peptides with the adopted secondary structure are thought to span the thickness of the bilayer in several of the proposed mechanisms. The theory of the requisite number of residues is thus dependent upon the secondary structure adopted by the peptide and the mechanism through which the membrane is disrupted.

Many attempts at AMP design have focused on arginine enrichment creating areas of cationic charge and cysteine incorporated into the peptide sequence for further oxidation. Additionally, inclusion of thiols for oxidation to form disulfide bonds sometimes creates cyclic peptides.[25] The cyclic peptides created through oxidation of cystein's thiol often create hydrophobic regions within the AMP which are able to interact with the hydrophobic region of the aliphatic lipid tails within the membrane core. Others attempts at AMP design focused on enrichment with tryptophan due to its preference for the interfacial region of lipid bilayers.[28] Movement of charged molecules across a lipid bilayer is challenging

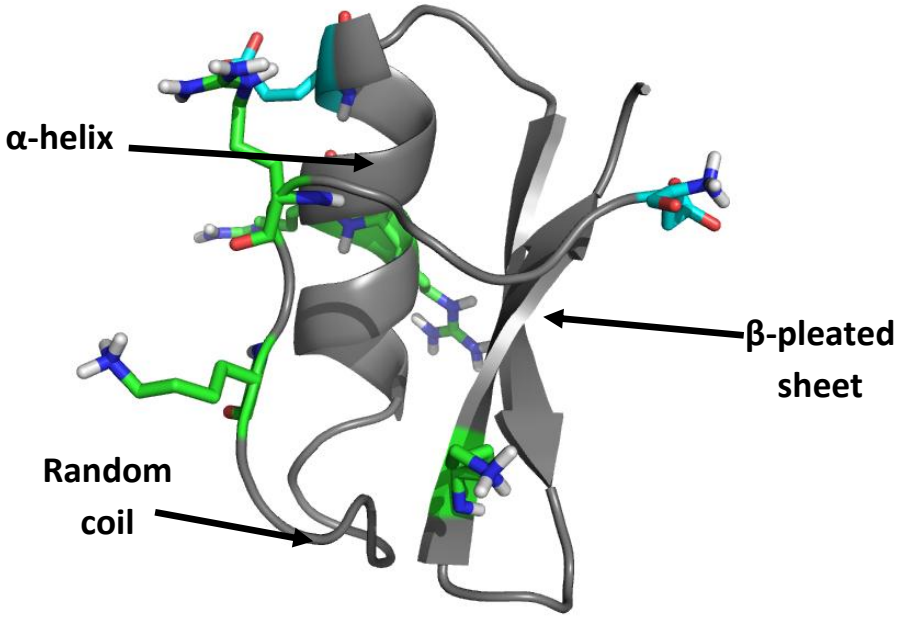


Figure 1.4 Common secondary structures associated with naturally occurring peptides.[27]

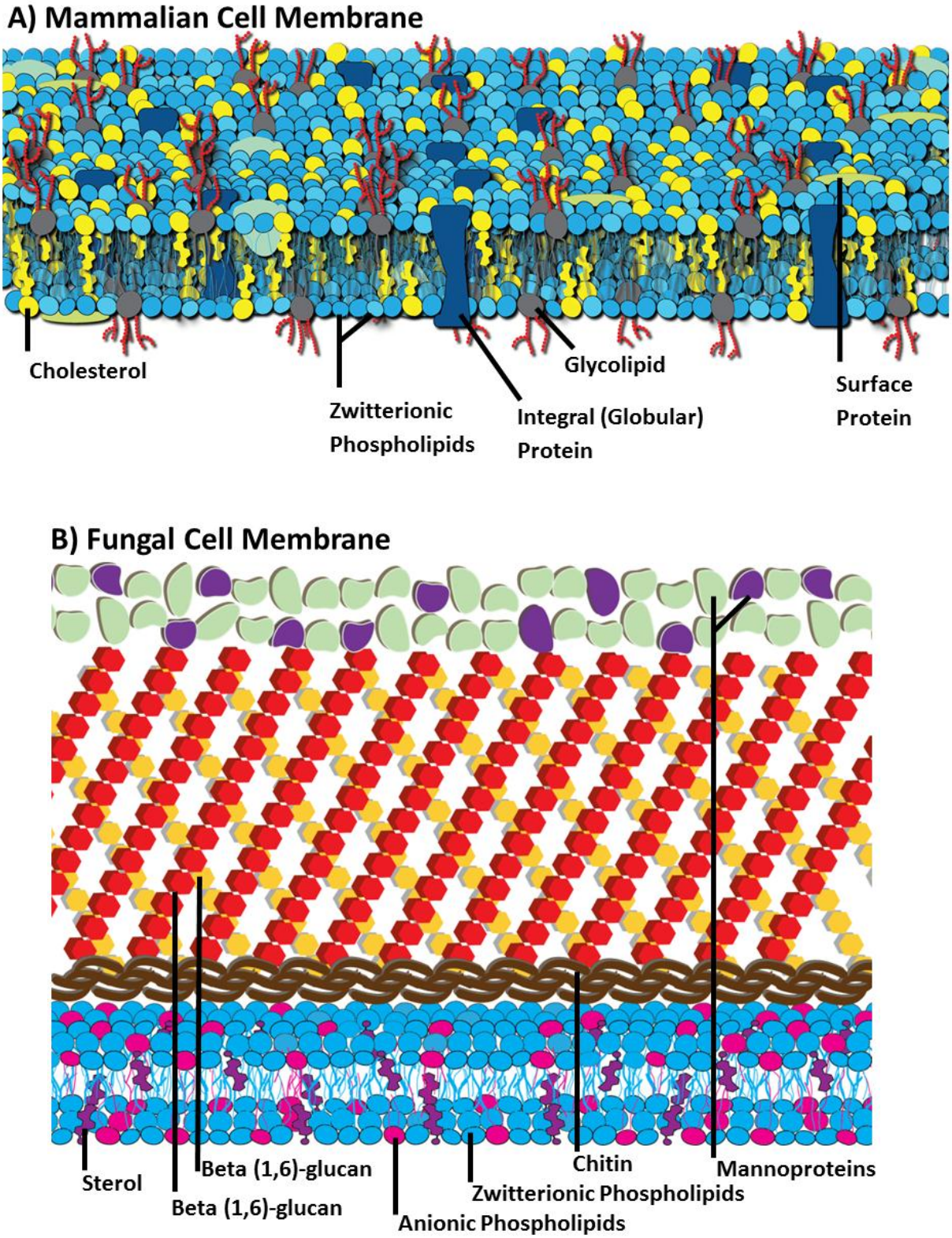


Figure 1.5 Eukaryotic membrane models A) Mammalian membrane model B) Fungal membrane model. [27]

due to the hydrophobic nature of the aliphatic lipid tails which must be traversed. Use of unnatural amino acids has enabled peptides to be more resistant to proteases and better tailor the amino acids to the target membrane.

It should be noted there is a distinct difference between AMPs and cell-penetrating peptides (CPP). While AMPs disrupt the membrane of a cell to the point of cellular death, CPPs are internalized by cells while avoiding the membrane destabilization and cell death.[29] CPPs are presently being investigated as drug delivery options for difficult to deliver peptide derived drugs. [29,30]

1.5 TARGETS OF AMPS

The structural and chemical characteristics associated with the membranes encapsulating each organism's cells is unique; not just to that organism but, within multicellular species membrane composition is unique to individual tissue types. The focus herein is on the components of the membrane which differentiate target cells from host cells. The structural features associated with eukaryotic cells which include mammalian and fungal organisms will be discussed. Additionally, the two most common membrane features, Gram negative vs. Gram positive, used in the differentiation of prokaryotic organisms will be discussed as well as how the mycobacteria differ from the bacteria which can be identified using the Gram stain technique.

The first generation of unnatural amino acid containing antimicrobial peptides was tested to identify selectivity for prokaryotic over eukaryotic cells. These two groups of cells with differing chemical compositions within the membrane are believed to be the result of electrostatic differences produced through variations in the membrane lipid composition.

1.6 EUKARYOTIC MEMBRANES

The Eukarya are distinguished from the prokaryotic organisms by the membrane bound nature of the cell nucleus. Both mammals and fungi are members of this taxon. Parasites, insects, reptiles, and amphibians are also members of the Eurarya. The membership of mammals to this group makes treating alien cells from fungi and parasites difficult to treat while maintaining low toxicity for the host cell due to the similarity in the chemical compositions of the membranes. The generic structural features which distinguish the mammalian cells from that of the fungi are discussed in the context of electrostatic modulating molecules which can be exploited for therapeutic selection.

1.6.1 MAMMALIAN MEMBRANES

Mammalian cells have a membrane charge closer to neutral as a result of the abundance of zwitterionic phospholipids (Figures 1.5 A and B). In addition to the zwitterionic phospholipids, eukaryotic cells frequently have bulky sterols incorporated into the membrane surface contributing polar but electrostatically neutral structure. Glycolipids which function as both an energy source and cellular markers are found in all eukaryotic cell membranes. Many of the *in vitro* toxicity tests focus on the effect the AMP has on erythrocytes or red blood cells

Hemolysis is associated with a lack of cellular specificity for foreign organisms over that of the mammalian host. AMPs with hemolytic properties would have limited utility due to toxicity.

1.6.2 FUNGAL MEMBRANES

The most common fungal infections present in the U.S. involve dermal surfaces exposed to prolonged periods of high humidity which create a hospitable environment for fungal infections such as athlete's foot, ring worm, jock itch and yeast infections. The community of immune-compromised patients including those with AIDS, undergoing cancer treatment, and organ transplant recipients is particularly at high risk for fungal infections which are commonly thought of as opportunistic. The aggrandized and diverse emergence of fungal infections has precipitated the need for new chemotherapeutic compounds. The penurious therapies approved for treatment of the eukaryotic, fungal infections have poor membrane permeating capabilities. The lack of cell penetration of antimycotic compounds necessitates high doses for effective antimycotic activity. The poor therapeutic profiles might be attributed to the membrane composition of mammalian and fungal cell walls (Figure 1.5 A and B) which share electrostatic similarity. Much like their bacterial counterparts, antifungal compounds have also experienced the drug resistant evolution to the therapeutic arsenal further handicapping fungal infection treatment.

The development of AFPs has great therapeutic potential. The last generation of

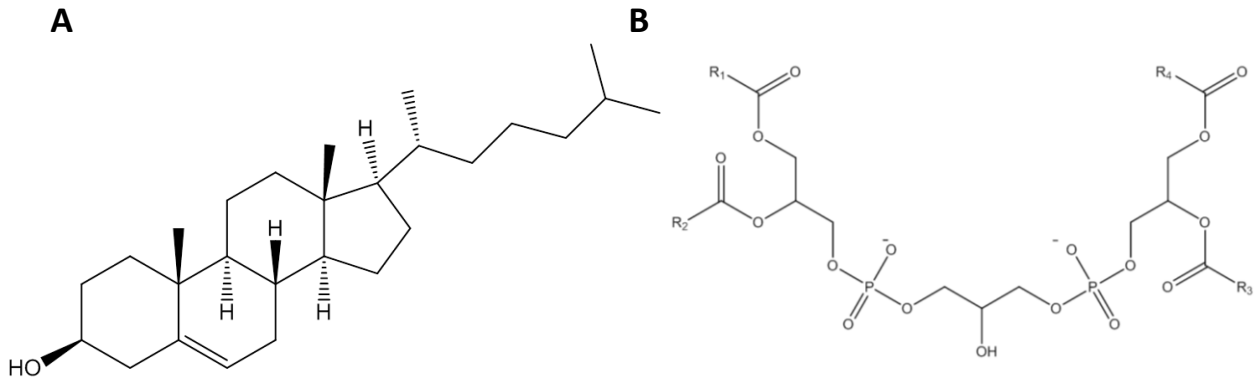


Figure 1.6 A) Cholesterol which is major components of eukaryotic cell membranes and **B)** Cardiolipin where R 1, R2, R3, and R4 represent aliphatic carbon tails

antifungal small molecule drugs is still fraught with many of the original compound limitations but AFPs administered as an adjuvant can act synergistically to improve the activity of the antimycotic by destabilizing the foreign cell membranes while avoiding host cells. Once the membranes have been destabilized, delivery of the antimycotic will have greater target delivery success enabling a lower dose and reduced side-effects.

1.7 PROKARYOTIC MEMBRANES

Prokaryotic cells differ from the eukaryotic due to the lack of a cell nucleus and other membrane bound organelles. Many species of bacteria (referring to the colloquial usage and not the genus) contribute and have demonstrated that they are an essential component of mammalian organisms and contribute to the health through a symbiotic relationship. The challenge is selecting for those organisms which are robust enough to overcome the normal flora resulting in an infection. In general the membrane surface of prokaryotic organisms has an anionic charge due to the large percentage of negatively charged phospholipids within the membrane.

The following description of the prokaryotic bacterial membranes both Gram positive and Gram negative has been published elsewhere.[26] The cartoons depicted in Figure 1.7 A & B give a simplified perspective of the membrane complexity which exists with the various membrane types. Within an organism the membrane composition varies which adds to the complexity of accurate model development. The membrane lipid composition varies among species and can even vary within a species due to the evolutionary adaptation of some enabling them to survive in environments the non-adapted species would not be able to survive.

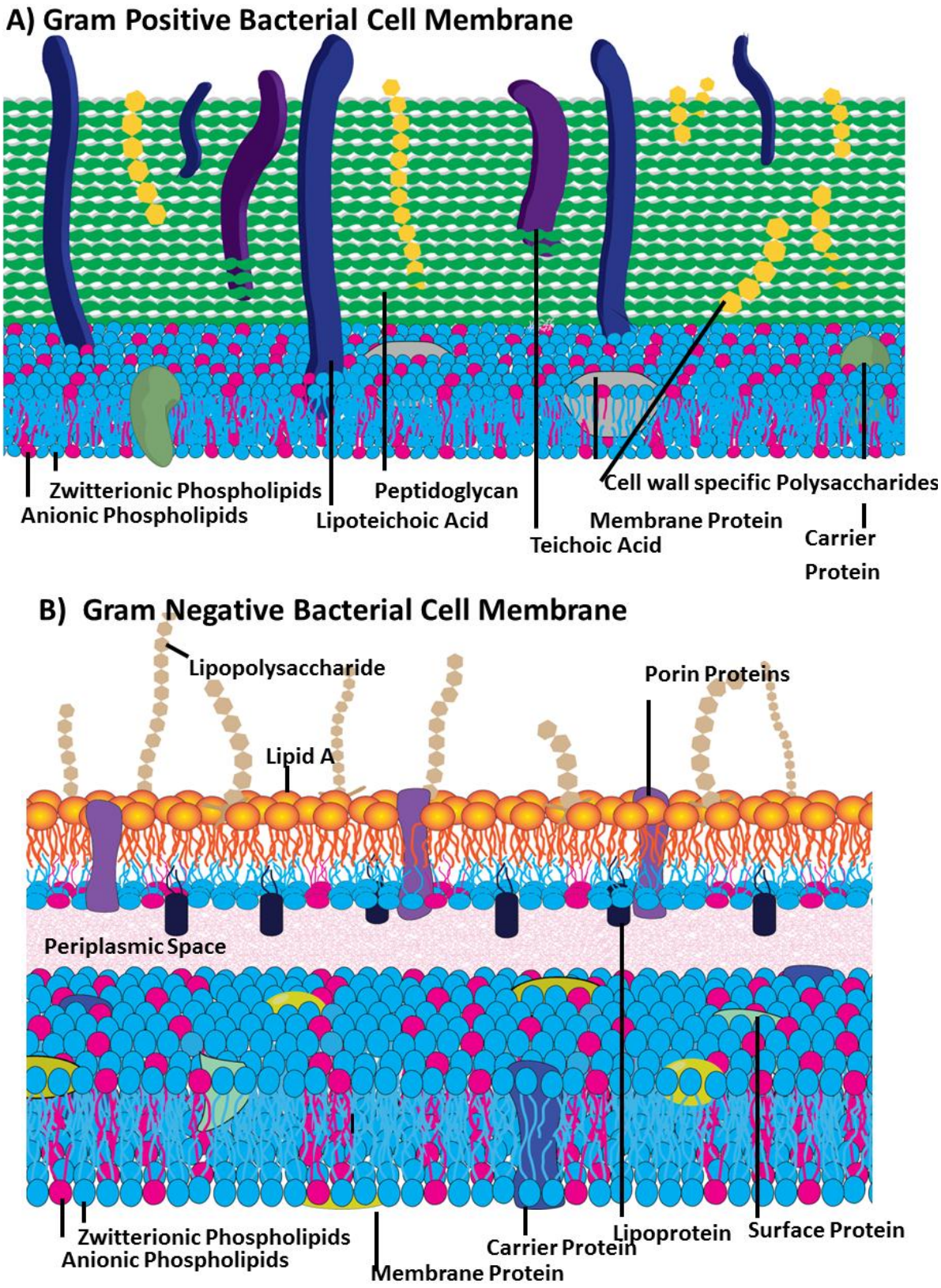


Figure 1.7 A) Model of prokaryotic gram positive cell membrane B) Model of prokaryotic gram negative cell membrane [26, 27]

1.6.1 Gram Positive Bacteria

Since the membrane is the target of AMPs and it is the disruption of the membrane which leads to the antimicrobial activity it is important to understand the specific molecules which are responsible for distinguishing one membrane type from another. Mammalian cell membranes have mostly hydrophobic cholesterol incorporated into the lipid bilayers. The inclusion of cholesterol with its small polar head group produced by the hydroxyl functional group decreases the flexibility associated with the membrane surface, see Figure 1.6 A. The lipid composition for several of the clinically relevant infectious organisms has been examined.[31] Extension of the membrane specificity would reason that the lipids, proteins, and other cellular components and the ratios of each contribute to the AMP selectivity observed between different organisms. One example of the unique membrane lipid composition responsible for AMP selectivity occurs with the Gram positive bacteria *Staphylococcus epidermidis*, which is composed primarily of POPG at 90% of the lipid composition.[26] The spore forming Gram positive bacteria *Bacillus subtilus* has a less anionic membrane with only 29% of the membrane comprised of POPG (Figure 1.9).[26]

1.7.2 Gram Negative Bacteria

Gram negative bacterial strains tend to have an inner and outer membrane with carbohydrate rings of peptidoglycan sandwiched between the membranes (Figure 1.7), thus reducing the amount of stain that is retained by the membrane when Gram stains are applied.[9] *Salmonella typhimurium*, *Pesudomonas cepacia* and *Escherichia coli* contain only 33, 18 and 6 percent POPG respectively reflecting the great variability even between the Gram negative bacteria.[26] *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*, which are members of the ESKAPE pathogens[10] contain 5 and 21 percent POPG respectively.[26] The zwitterionic lipid 1-Palmitoyl-2-Oleoyl-*sn*-Glycero-3-Phosphoethanolamine (POPE) has been identified as a major

component of Gram negative bacterial membranes.[26] *Escherichia coli*, *Klebsiella pneumonia* and *Pseudomonas aeruginosa* have membranes primarily composed of the anionic POPE lipid which contain 80, 82 and 60 percent POPE respectively.[32] The dual anionic charge associated with cardiolipin (Figure 16 B), identified as "Lipid A" in Figure 1.7 B offers four hydrophobic, aliphatic tails associated with the two phosphate ester moieties of the head groups.[33] Other structural features which differentiate prokaryotic and eukaryotic cell membranes include the sterols in the instance of fungi and cholesterol in the instance of mammals. The complex sugar coats composed of lipopolysaccharide and peptidoglycan challenge the disruption of the membrane yet provide unique structural characteristics.[9]

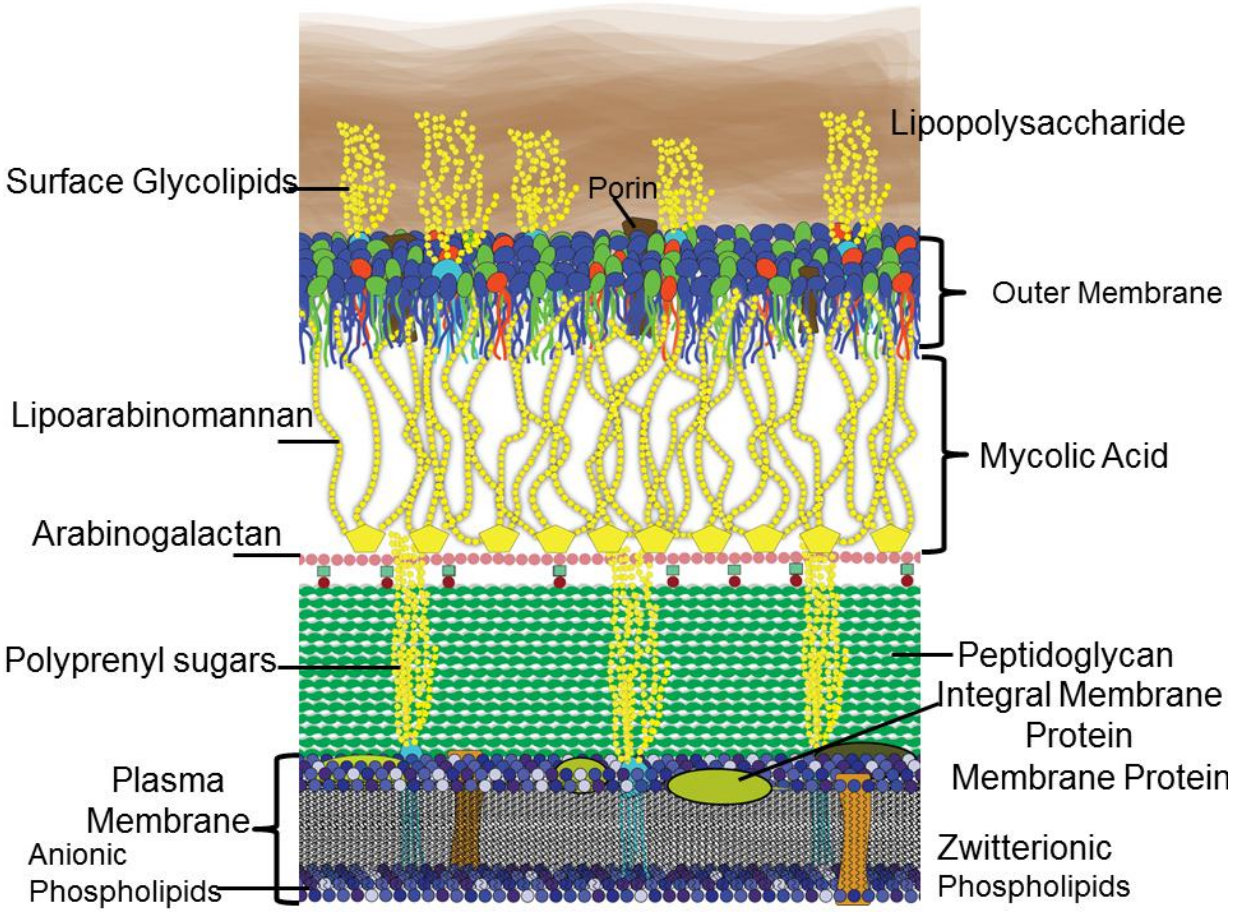


Figure 1.8 Model of prokaryotic mycobacteria cell membranes. [36-41]

1.7.3 Mycobacteria

Mycobacteria are prokaryotic organisms despite the original suspicion that they were a fungal species' (as evidenced by the "myco" root which means fungus). They are discussed here separately due to the complex membrane which is the target for therapeutic AMPs. Several clinically significant members of the mycobacteria include *Mycobacterium tuberculosis*, the causative agent of tuberculosis infections, *Mycobacterium leprae*, the causative agent of leprosy, and nontuberculosis mycobacterial infections caused by other mycobacterial species.[34] The expanding population of immunocompromised patients has led to the re-emergence with antibiotic resistance developing within this once well controlled collection of infections.[34] The mycobacteria are considered Gram-positive however they do not retain the requisite crystal violate stain on a consistent basis. A second staining technique, acid-alcohol staining, is needed to correctly differentiate mycobacteria from other prokaryotic organisms [5]. The World Health Organization estimates that as many as 32% of the world population is infected with *M. tuberculosis* infections which provides a glimpse of the significant need for the development of effective treatment options.[34, 35]

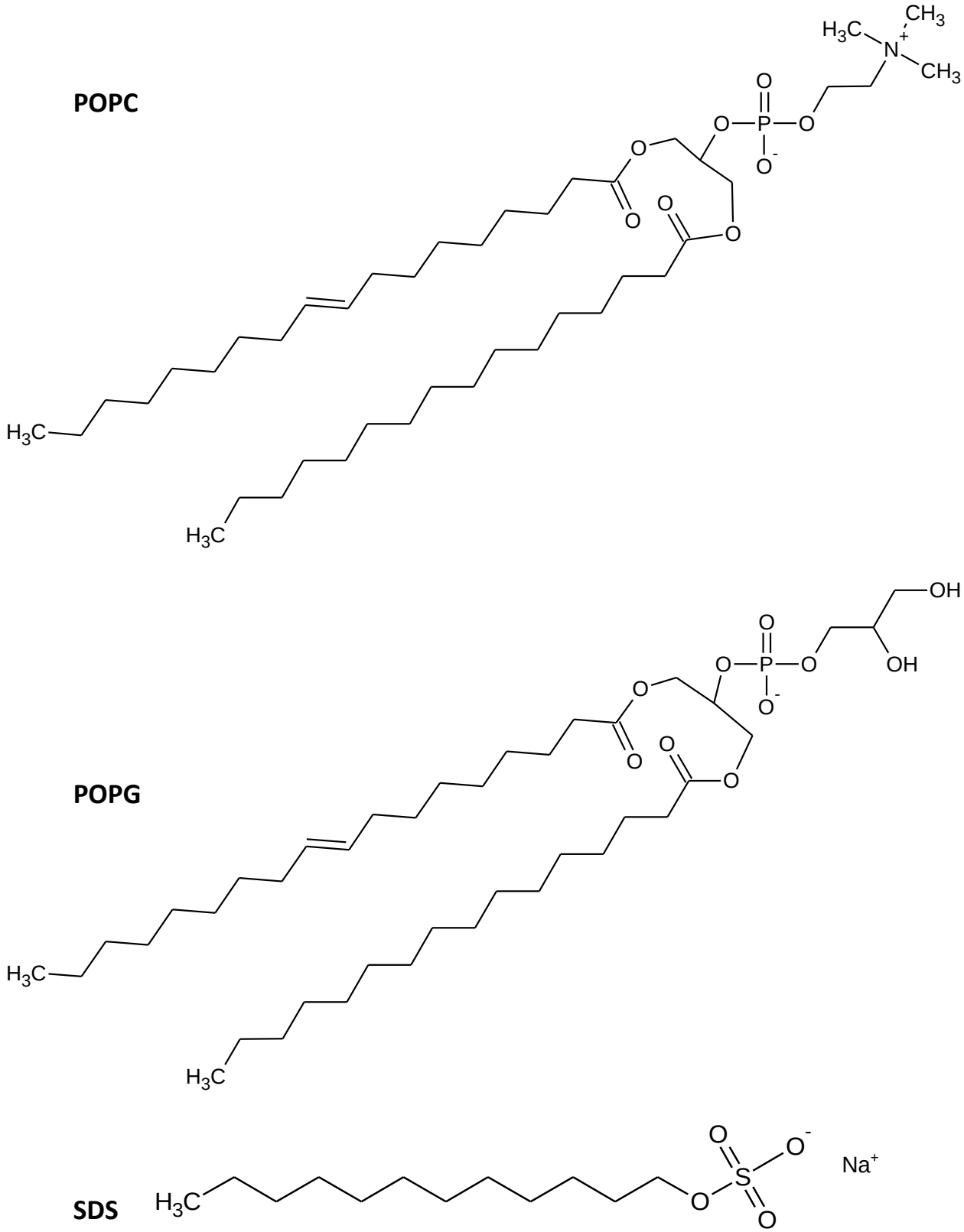


Figure 1.9 1-Palmitoyl-2-Oleoyl-*sn*-Glycero-3-Phosphocholine (POPC), 1-Palmitoyl-2-Oleoyl-*sn*-Glycero-3-[Phospho-rac-(1-glycerol)] (POPG) and sodium dodecylsulfate (SDS), the three lipid molecules used in the construction of the membrane models.

The complex membrane surrounding the mycobacterial cells has demonstrated immunomodulatory activities and has thus been the focus of antimicrobial compound development.[34, 35] The complex collection of lipids and carbohydrate molecules combined with the unique mycolic acid and unique polysaccharides such as arabinogalactan to create the dynamic species.

The cartoon depicted in Figure 1.8 shows a simplified perspective of a mycobacterium membrane; is an aggregation of information from several published sources sources.[36-41]

1.8 PROPOSED MECHANISMS OF ACTION

There have been three mechanisms through which AMPs have been hypothesized to exert their membrane disrupting activity. The first is the Barrel-stave mechanism where the peptides completely line a pore within the membrane. The second is there toroidal pore mechanism where the AMPs intermingle with the polar head groups of the membrane to create a membrane disruptive pore. The final mechanism is the carpet mechanism in which the AMPs have a detergent like effect creating micelles from the amphipathic membrane lipids after the bilayer structure has been disrupted. While the three mechanisms have been paired with many of the naturally occurring AMPs less is known about the mechanism of action through which the unnatural amino acid containing AMPs exert their activity. The common mechanism will be discussed in detail as a reference point for the ensuing investigation.

1.8.1 Barrel-Stave Mechanism

The variation in charge associated with bacterial membranes enables tailoring of the peptide charge for optimum electrostatic interaction. There are three proposed mechanisms through which antimicrobial peptides are speculated to exert membrane destabilization activity.[42]

Others distinguish the mechanism as one of two classes, those which function synergistically and those which have functional duality.[12] Identifying which mechanism or mechanisms correlate

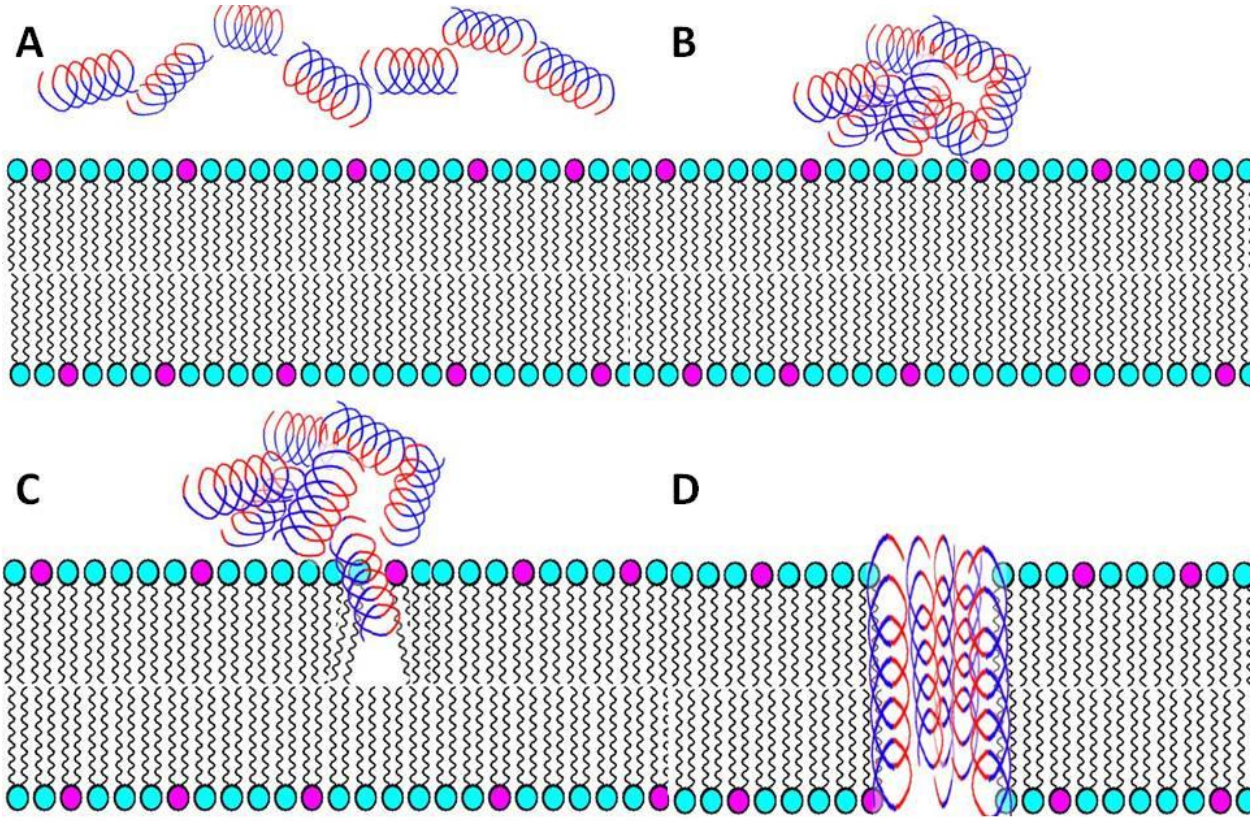


Figure 1.10 Barrel-Stave Mechanism. Hydrophilic residue regions are noted in red. Hydrophobic residue regions are noted in blue. Anionic POPG lipids are represented by magenta spheres and zwitterionic POPC lipids are represented with cyan colored spheres. Adapted from figures produced by Brogden.[34]

with individual AMP's function has been the subject of numerous studies.[12, 29, 43-52] The concentration of peptide relative to that of the lipids is critical in the membrane disruption mechanisms.[12] Aggregation of the peptide before the aggregate is attracted to the bilayer surface is a prerequisite. Once the critical peptide concentration has been achieved, aggregation can occur, however whether aggregation occurs prior to membrane attraction, i.e. in solution or at the surface of the membrane after the peptide has been attracted, has been debated. Despite the debate of when aggregation occurs, aggregation is a prerequisite for pore formation regardless of

the AMP type causing the disruption. Despite the diverse nature of AMP structure and selectivity they have been divided into three unique proposed mechanistic types: the barrel-stave, carpet, and toroidal pore.[53, 54]

The barrel-stave model was original introduced in the 1970s and has been considered the model mechanism for peptide-induced trans-membrane pore formation.[55] The barrel-stave model has been reported to be the natural mechanism of membrane disruption for several AMPs including alamethicin,[55] ceratotoxins,[56] peptaibols,[57] and distinctin.[56] Originally it was suggested that melittin and magainin function through a barrel-stave mode, although further research has demonstrated that they function through a toroidal pore mechanism. [42] In the barrel-stave model the peptide must be long enough to span the membrane.[45] Dimers or multimers form after binding of the AMP to the bacterial membrane. The formation of this assembly is a requisite step for pore formation through this mechanism.[58] Figure 1.10 is a cartoon depicting the mechanism of the barrel stave mechanism of bilayer disruption. Figure 1. 10A) the bilayer is surrounded by free peptide in solution. Figure 1. 10 B) electrostatic interactions between the peptides and bilayer cause aggregation. Figure 1. 10 C) the peptides aggregate on the bilayer surface leading to the disruption of the electrostatic interactions which hold the lipids together. Figure 1. 10 D) the peptides are fully inserted into the bilayer creating a pore lined with AMPs. 2-Oleoyl-1-palmitoyl-sn-glycero-3-phosphocholine (POPC) (Figure 1.9) is represented by cyan head groups while 1-Palmitoyl-2-Oleoyl-sn-Glycero-3-[Phospho-rac-(1-glycerol)] (POPG) (Figure 1.9) is represented by magenta head groups. The polar amino acids are represented by the red portion of the alpha helix cartoon while the hydrophobic portion is represented by the blue portion of the alpha helix cartoon. Figure is an adaptation of a figure from Brogden.[42]

1.7.2 Toroidal Pore Mechanism

The toroidal pore mechanism shares many common features with the barrel-stave pore model. They are similar in that they both require binding of the AMP to the bacterial membrane followed by aggregation of the AMP within the membrane which finally leads to the formation of pores.[58] This mechanism was originally used to describe magainin-induced pores.[59] The toroidal model is different from the barrel-stave model, in that the peptides are always in association with the lipid head groups even when they are perpendicularly inserted into the lipid bilayer. This association forces the bilayer to fold toward the hydrophobic center (from both the top and bottom of the bilayer) of the bilayer producing a pore which is lined by both the peptides

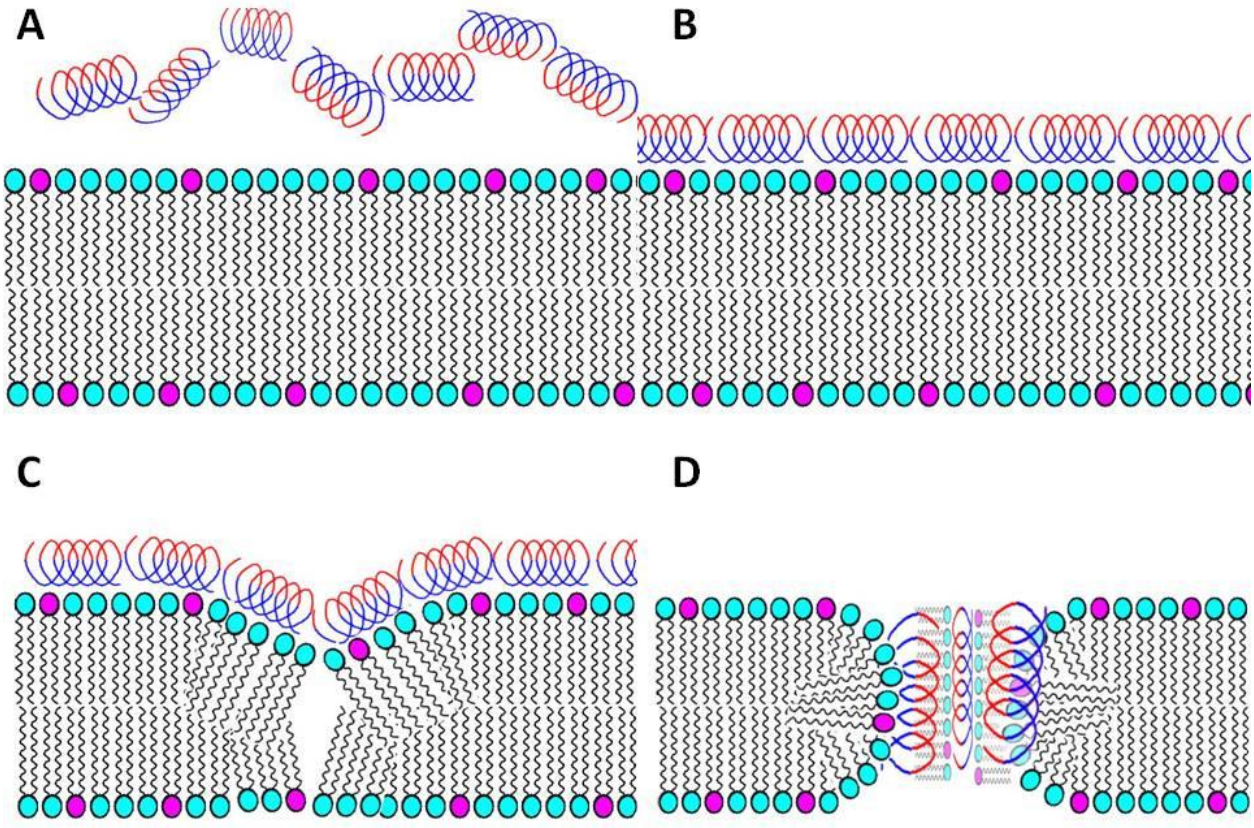


Figure 1.11 Torroidal Pore Mechanism. Hydrophilic is in red. Hydrophobic residues are in blue. Anionic POPG lipids are represented by magenta spheres and zwitterionic POPC lipids are represented with cyan colored spheres. Adapted from a figure created by Brogden.[34]

and the lipid head groups (see Figure 1.11).[55] Unlike the barrel-stave model, the peptides which act though the toroidal pore mechanism do not need to be long enough to span the entire thickness of the membrane.[45] This enables smaller peptides to destabilize the lipid bilayer by stabilizing the lipid pore rather than acting as a pore lining. The size of the pore created varies depending on the AMP which formed the pore and the lipid bilayer composition. In general, pores on the nanometer size scale are obtained using the toroidal pore mechanism.[52]

Figure 1.11 is a cartoon depicting the mechanism of the toroidal pore mechanism of bilayer disruption. Figure 1.11 A) the bilayer is surrounded by free peptide in solution. Figure 1.11 B) electrostatic interactions between the peptides and bilayer cause. Figure 1.11 C) the peptides aggregate on the bilayer surface leading to the disruption of the electrostatic interactions which hold the lipids together. In Figure 1.11 D) the peptides are fully inserted into the bilayer creating a pore lined with both AMPs and the polar head groups of the membrane. POPC (Figure 1.9) is represented by cyan head groups while POPG (Figure 1.9) is represented by magenta head groups. The polar amino acids are represented by the red portion of the alpha helix cartoon while the hydrophobic portion is represented by the blue portion of the alpha helix cartoon. These figures are an adaptation of a figure from Brogden.[42]

1.8.3 Carpet Mechanism

The third mechanism is the carpet model. In this model, peptides accumulate on the bilayer surface. The peptides orient parallel to the membrane surface through electrostatic attraction of the anionic phospholipid head groups which cover the surface of the membrane (Figure 1.12).[42] Through the binding of the peptide to the phospholipid membrane head groups a layer or 'carpet' is formed on the bilayer surface. Alignment of the peptide monomers on the surface occurs with the hydrophilic surface facing the phospholipid head groups.[46, 51, 54, 60] This

peptide binding continues until a threshold concentration is reached upon which permeation occurs in a detergent like manner.[61] The ability of an antimicrobial peptide to permeate a membrane is a concentration dependent process. At high peptide concentrations the peptides on the surface of the bilayer, through a detergent like activity, are able to form micelles.[42] When the critical concentration of peptide has been reached, transient, toroidal holes form in the membrane surface allowing small molecules to cross the bilayer prior to cellular collapse. Thus essentially when the use of the toroidal pore mechanism is engaged many times this constitutes the carpet mechanism.

A unique characteristic of the carpet mechanism is that the traditional requirement for the adoption of a specific secondary structure (such as an α-helix) upon binding to the membrane is not required by this mechanism.[54] Computational chemistry is able to provide useful tools and techniques which can complement data obtained from chemical experiments or it can be used for the investigation of systems which either cannot be studied using synthetic techniques or have produced inconclusive results. [12, 16, 17, 43, 45, 51, 54, 56, 57, 59, 62-64] During the course of this techniques will be utilized to provide the most rational explanation of the activity of our AMPs and their interaction with model eukaryotic and prokaryotic membrane systems. Energetic analysis of the interaction between the AMPs and the bilayers have been undertaken through MD simulation in which the atoms within a given system are allowed to interact over a short period of time by approximating the physical motion of the individual particles.

Understanding the mechanism through which each AMP exerts its *in vitro* activity can aid the design of the next generation of AMPs containing unnatural amino acids. Determining the secondary structure of the peptides containing unnatural amino acids can aid in the mechanism elucidation.

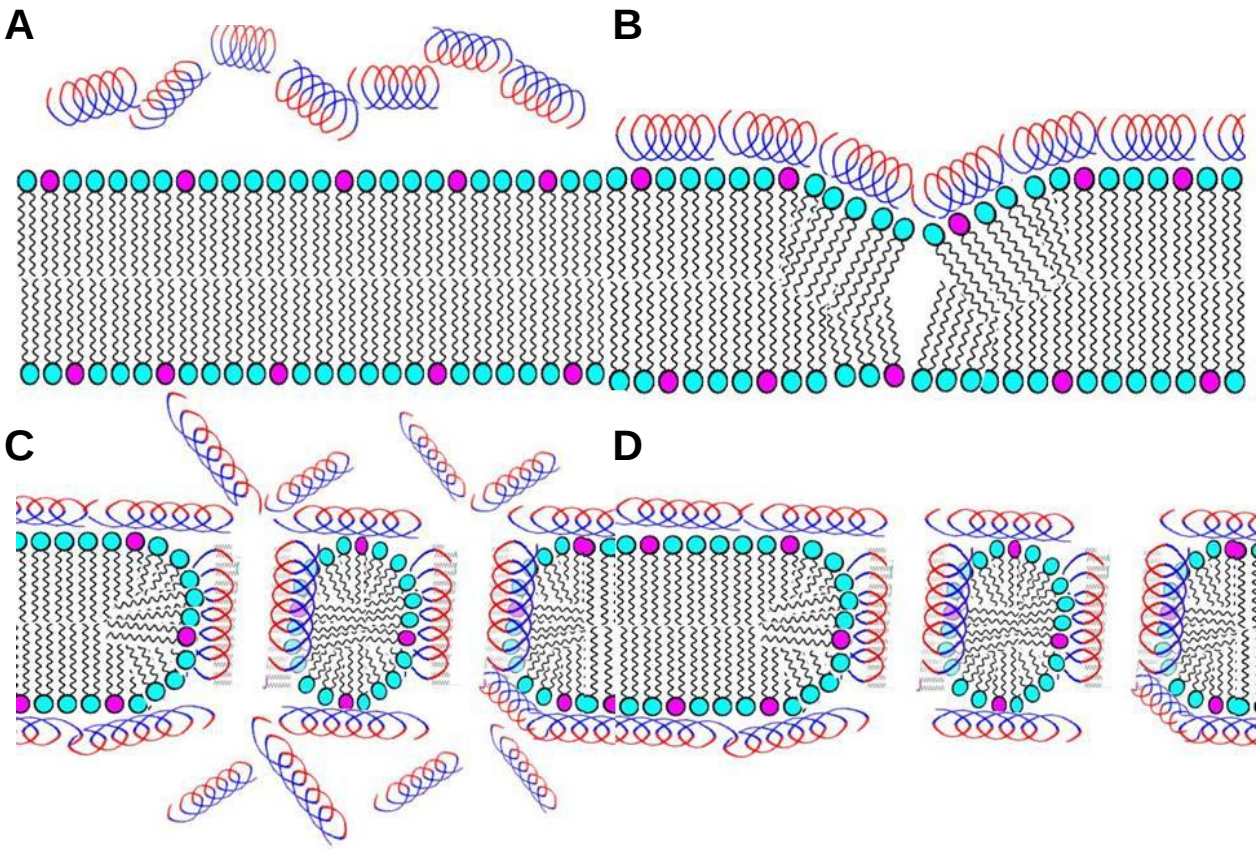


Figure 1.12 Carpet Mechanism. Hydrophilic residues are represented in red. Hydrophobic residues are shown in blue. Anionic POPG lipids are represented by magenta spheres and zwitterionic POPC lipids are represented with cyan colored spheres. Adapted from a figure by Brogden.[34]

1.9 BILAYER CURVATURE

Naturally occurring cells, and thus the membranes which exist as part of a larger spherical shape are influenced by the radius of the spherical cell. The curvature present in the membranes plays an important role in the membrane disruptive effects of AMPs.[65] MD simulations with BAR domains, which constitute a type of protein which drives the formation of tabulated and vesiculated membrane structures has demonstrated the ability to induce bilayer curvature when simulated using a coarse grain (CG) model.[66] The large number of atoms

required for the cooperative effect between the peptides and bilayer to occur makes the development and simulation of accurate model bilayer systems difficult to achieve with the requisite time scales at which the curvature can be observed to develop. This is a limitation of atomistic MD simulations present in all bilayer simulations. The main challenge any theoretical approach encounters is its ability to accurately replicate the system of interest. A detail as small as the curvature of the membrane can affect the ability of an AMP to form an interaction with the head groups of the bilayer lipids. There is evidence that the AMP concentration required for membrane thinning prior to pore formation can vary with the curvature of the membrane lipids. Differences between the curvature of the liposomes and micelles used in *in vitro* studies and the *in silico* studies have been another potential source of error.

Previous MD simulations with AMPs have primarily involved use of CG-MD[65, 67-70] The CG method minimizes the number of charge points which need to be considered at each step within the calculation and thus the offers a parsimonious approach for simulations with many atoms (100 000+ atoms). The speed gained with fewer calculations is achieved at the expense of atomic detail. Due to the loss of atomic perspective an all atom approach has been utilized in this research and will be discussed further in the next few chapters.

Environmental variables considered in the MD simulation are to model *in vivo* conditions as much as possible. Simulations have been conducted at 300 K for the micelle simulations to correlate with *in vitro* data[71] which was collected at room temperature. Bilayer simulations designed to model the environment of the *in vitro* studies have been conducted at 310 K, body temperature.

Previous MD simulations with AMPs have primarily involved use of CG-MD[65, 67-70] Use of CG-MD has been utilized because the small number of charge points, resulting from the

aggregation of numerous atoms into one charge point, increases the speed of the simulation. Having fewer charge points requires fewer calculations at each time step of the simulation.

1.10 MEMBRANE SPECIFICITY

The cell membrane's lipid composition varies for each organism and even each tissue type within an organism. The variation in the chemical composition between cell types can be a great target for therapeutic compound development. Tailoring an AMPs structure to interact with the invasive cell's membrane creates selectivity and specify.

1.11 COMPUTATIONAL THEORY

The goal of computational chemistry is to mathematically represent chemical reality. Molecules are composed of atoms which in computational models are represented as charged particles. In theory, including measures of all chemical properties (descriptors) associated with each atom should be able to reproduce experimentally derived results. Inclusion of such large numbers of descriptors is not computationally feasible with the current resources available. The truncation of descriptors may seem like an over simplification of the complex inter atomic interactions which occur between the atoms of a molecule and its surrounding environment but, it has demonstrated itself sufficient to be consistent with experimental data. The main focus is on the atomic charges because Coulomb interactions are the only physical force for chemical phenomena and van der Waals forces are responsible for handling the non-bonded interactions in the AMBER9 ff03 force field.[72] Considering the relative energy associated with the three dimensional arrangement of atoms, represented by collections of charged particles, the atomic interactions which occur during the reaction which result in the formation of unique products are

able to provide insight into other physical observable properties exhibited by the molecules.[72] The ultimate goal would be an exact solution of the Schrödinger Equation[72] (Equation 1.1) which at this time is feasible for only one and two particle systems, with the two particle system being separated into two pseudo one-particle problems and introducing a center of mass coordinate system.[72]

$$\mathbf{H}\boldsymbol{\Psi} = \frac{i d\boldsymbol{\Psi}}{dt} \qquad \text{Eqn. (1.1)}$$

The computational approaches will be discussed in further detail in Chapters 2 and 3.

1.12 STRUCTURE FUNCTION RELATIONSHIPS

In the 1960's Anfinsen proposed his "thermodynamic hypothesis" which states that a protein's native structure corresponds to the state with the lowest free energy of the protein in the water solvent system.[73] The correlation between the peptide sequence, the secondary structures and the resulting function suggests that determining the secondary structure can provide insight into the function of the peptide.[74] Use of the Generalized Born/Simulated Annealing is one approach that will identify low energy structures of the peptide systems. Using an implicit solvent specified through the use of a dielectric constant appropriate for water and a physiologically relevant ionic concentration the system was heated to 3000 K and cooled over 10 000 steps. Had an infinite number of steps been used for the cooling process, the global minimum energy structure, which has the lowest potential energy, would have been identified. Since an infinitely slow process is not feasible due to the time scale, an abbreviated cooling time has been used which resulted in a local minimum energy structure.

1.13 COMPUTATIONAL TIME SCALE AND THE RELATIONSHIP TO REALITY

1.13.1 Software

Many powerful software programs have been developed for the study of compounds at the molecular level yet widespread use and understanding is still limited within the chemistry community. Most prevalent in the literature covering bilayer simulations are utilization of GROMACS (GROningen MAchine for Chemical Simulations)[75], NAMD (NAnoscale Molecular Dynamics), and CHARMM (Chemistry at HARvard Macromolecular Mechanics)[76] software programs for MD simulation of bilayers. Despite the abundance of simulations with the above software simulations the bilayer simulations undertaken herein have been conducted with AMBER 9.[77] AMBER 9 was selected for the available expertise available at East Carolina University (ECU), in the Department of Chemistry. Additionally, the AMBER suite of programs offers a straight forward approach to add parameters to the force field for those molecules not previously included in the original force field parameter library.

Gaussian 03[78] is the 2003 release of the Gaussian series of electronic structure programs which utilizes the basic laws of quantum mechanics to predict energies, molecular structures, and vibrational frequencies of molecular systems. Calculation of these properties enables other data to be extrapolated. Understanding these components of a molecule under differing reaction conditions can enable accurate prediction of the most stable species of a compound when it is difficult or impossible to observe the structure experimentally due to short-lived intermediates and transition structures. In this study we will use it to calculate the minimum energy structures.[79]

The same level of theory, and basis sets were used in Gaussian for the parameterization of the amber force field. The high level QM calculations resulting from Gaussian were supported by x-ray structure data for each of the 20 standard amino acids. The same procedures were followed for each of the unnatural amino acids used in this study however structural data is not currently available.[79]

Materials Studio[80] is a comprehensive materials modeling and simulation software which has been well used in the research and development of pharmaceuticals. During the course of this research it was utilized for its graphical user environment which has allowed the development of starting structures for all calculations and simulations of peptide/model membrane systems.

Insight II is Accelrys software for the modeling of biological systems which has numerous functionalities, but here it was utilized for the construction of the lipid bilayers and the micelles. Manipulation of structure systems and bilayers with peptides can easily be accomplished using this program. While this software has plug-ins which could accomplish many of the calculations in this study; limited access prevents utilizing this as the major software. Discovery Studio is another Accelrys software which provides a molecular mechanics simulation environment for energy minimization and dynamics. It will be used to conduct pre-minimizations of the AMPs.

DMol3 is a density functional theory (DFT) program which is able to solve quantum mechanical equations efficiently for large systems.[81] DFT is based on the principle that the ground state electronic energy is determined completely by the electron density. It will be used to calculate the minimum energy of the structures of the lipid molecules which have more atoms

than the unnatural amino acids which compose the AMPs. DMol3 is able to efficiently optimize the larger structures.[81]

Pymol is a powerful software which enables molecular visualization of systems formatted for display of atomic coordinates.[82] Pymol has been used to visualize structures, take measurements, and create animations in addition to pictures.[82] Manipulation of lipid molecules in Pymol makes it possible to create the bilayer systems and enables easy adjustment of the distances between the polar head groups between the leaflets of the bilayers and the lipids within each leaflet.

Cygwin is a Linux-like environment for Windows which enables interface between the two environments.[83] All of the access to the high performance computing (HPC) cluster has been accessed through Cygwin software. SSH is an invaluable tool in the movement of large files between the HPC and personal computers by providing a secure environment for the transfer of files.

AMBER9 is a collection of about 50 programs including sander, xLEaP, antechamber, ptraj, mm-pbsa, which are used to study the molecular movements of atoms in different environments.[77] The different environments can include different water models, or an *in vaccuo* approach, implicit solvation or use of a nonpolar solvent.[77] Antechamber is an auxiliary program of AMBER9[77], for molecular mechanic (MM) studies which amongst other things, is able to find missing force field (FF) parameters and supply reasonable and similar substitutes.[77] This software has been utilized for all of the MD simulations undertaken during the course of this investigation.

While computer models are able to provide detailed representations; the model is only as accurate as the FF parameters ability to capture the actual physiological interactions.[84] It is

possible to include many details of the actual system in the model but, in the absence of crystallographic or NMR structural data, a model is an approximate representation of the system with error being inevitable.

Reaching the time scale at which AMPs interact with membranes is a challenge and requires large and powerful computational resources. The data can continuously be collected for a system however, as the simulation is allowed to progress, the limitations of the discipline encroach. Limited data storage space, time, and computer resources all limit the perspective which can be acquired. While the interaction of the AMPs with the membranes occurs on the second time scale presently that is beyond the capabilities of all but the most powerful computer systems. A few hundred nanoseconds of MD simulation have provided insight into the atomistic interactions of AMPs which contain unnatural amino acids and membrane models

1.13.2 Hardware Resources

All simulations conducted utilized East Carolina University's (ECU) Center for Applied Computational Studies (CACS) which presently provides computing access to more than twelve scientists across four educational institutes and government agencies. The CACS promotes the application of computational methods to describe and understand phenomena in the life sciences, geosciences, engineering and other disciplines. Computational methods provide the researcher and educator with a set of powerful tools to explore the vast unknowns and to expose the young minds to the wonders of science, mathematics and engineering. CACS's present high performance computational resources are comprised of an SGI Origin 350 with 32 processors and an SGI Altix 4700 with 128 processors. The SGI Origin 350 and SGI Altix 4700 were funded in part by a Major Research Instrumentation Grants from the National Science Foundation. While the researcher to processor ratio is reasonable for studying small molecules,

these systems are regulated mostly for 8 processor jobs. Large MD simulations, such as bilayer simulations, require larger computers with more powerful processors with excellent threading (communication between the processors) than those presently available at ECU.

1.14 EXPERIMENTAL DATA

Experimental data presented and discussed herein has been collected and published elsewhere.[16, 17, 62, 85- 88] The computational studies are supported by the experimental data in the form of circular dichroism (CD) spectroscopy measurements, isothermal titration calorimetry (ITC), Nuclear Magnetic Resonance (NMR), and calcein leakage data are discussed in the context of consistency with the computational models.

Chapter 2 will discuss the parameterization of the FF which is responsible for producing reasonable molecular geometries and interaction energies for the set of atoms, molecules, or peptides being modeled.[84] Chapter 3 is intended to assist in the development of an understanding of the structural and chemical contribution of each nonstandard and unnatural amino acid within the medicinal chemists' "amino acid tool box" to help take the guess work out of the design process and encourages the rationalization of the *in vitro* activity correlating to the structural features. Chapter 4 discusses MD simulations of AMPs which contain the Tic Oic dipeptide with SDS micelles. Chapter 5 discusses MD simulations of the same AMPs as were discussed in Chapter 4 in the presence of bilayers modeling the electrostatic properties of mammalian and prokaryotic membranes. Chapter 6 is a discussion of the observed evolution of a peptide containing the Tic Oic dipeptide when inserted into the bilayer surfaces of model mammalian and prokaryotic membranes.

1.15 REFERENCES

1. Williams, D. A., and Lemke, T. L. (2002) *Foye's principles of medicinal chemistry 5th edition.* Lippincott Williams & Wilkins, Baltimore, MD.

2. *Science, Technology and Medicine in Modern History : Devices and Designs : Medical Technologies in Historical Perspectives.* (2006) Palgrave Macmillan, Basingstoke, Hampshire, GBR.

3. American Chemical Society. (2012) *National historic chemical landmarks.* www.acswebcontent.acs.org/landmarks/landmarks/penicillin/discover.html ed., American Chemical Society, .

4. Yocum, R. R., Rasmussen, J. R., and Strominger, J. L. (1980) The mechanism of action of penicillin. penicillin acylates the active site of bacillus stearothermophilus D-alanine carboxypeptidase. *Journal of Biological Chemistry. 255*, 3977-3986.

5. Todar, K. (December 31, 2009) *The microbial world.* http://textbookofbacteriology.net/themicrobialworld/homepage.html ed., Department of Bacteriology University of Wisconsin-Madison .

6. Gillor, O., Kirkup, B. C., and Riley, M. A. (2004) Colicins and Microcins: The Next Generation Antimicrobials, in *Advances in Applied Microbiology* (Allen I. Laskin, Joan W. Bennett,and Geoffrey M.Gadd, Ed.) pp 129-146, Academic Press, .

7. Centers for Disease Control and Prevention. (2010) *National antimicrobial resistance monitoring system (NARMS): Enteric bacteria.* http://www.cdc.gov/NARMS/ ed .

8. Interagency Task Force on Antimicrobial Resistance. (2012) *2011 PROGRESS TOWARDS IMPLEMENTATION OF:A PUBLIC HEALTH ACTION PLAN TO COMBAT ANTIMICROBIAL RESISTANCE.*

9. Gillespie, S., and Bamford, K. (2012) *Medical Microbiology and Infection at a Glance (4th Edition).* John Wiley & Sons, Hoboken, NJ, USA.

10. Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B., and Bartlett, J. (2009) Bad bugs, no drugs: No ESKAPE! an update from the infectious diseases society of america. *Clin. Infect. Dis. 48*, 1-12.

11. Infectious Diseases Society of America. (2012) *Drugs for limited use in patients with serious or life-threatening bacterial infections.*

12. McCafferty, D. G., Cudic, P., Yu, M. K., Behenna, D. C., and Kruger, R. (1999) Synergy and duality in peptide antibiotic mechanisms. *Curr. Opin. Chem. Biol. 3*, 672-680.

13. Tachi, T., Epand, R. F., Epand, R. M., and Matsuzaki, K. (2002) *Position-dependent hydrophobicity of the antimicrobial magainin peptide affects the mode of Peptide−Lipid interactions and selective toxicity.* American Chemical Society .

14. Matsuzaki, K., Sugishita, K., Harada, M., Fujii, N., and Miyajima, K. (1997) Interactions of an antimicrobial peptide, magainin 2, with outer and inner membranes of gram-negative bacteria. *Biochim. Biophys. Acta. 1327*, 119-130 ER.

15. Unger, T., Oren, Z., and Shai, Y. (2001) The effect of cyclization of magainin 2 and melittin analogues on structure, function, and model membrane interactions: implication to their mode of action. *Biochemistry (N. Y.). 40*, 6388-6397.

16. Bhonsle, J. B., Venugopal, D., Huddler, D. P., Magill, A. J., and Hicks, R. P. (2007) Application of 3D-QSAR for identification of descriptors defining bioactivity of antimicrobial peptides. *J. Med. Chem. 50*, 6545-6553.

17. Hicks, R. P., Bhonsle, J. B., Venugopal, D., Koser, B. W., and Magill, A. J. (2007) De novo design of selective antibiotic peptides by incorporation of unnatural amino acids. *J. Med. Chem. 50*, 3026-3036.

18. Ganz, T. (2005) Defensins and other antimicrobial peptides: A historical perspective and an update. *Comb. Chem. High Throughput Screen. 8*, 209-217.

19. Zeya, H. I., and Spitznagel, J. K. (1963) Antibacterial and enzymic basic proteins from leukocyte lysosomes: Separation and identification. *Science. 142*, 1085-1087.

20. Boman, H. G., and Hultmark, D. (1987) Cell-free immunity in insects. *Annu. Rev. Microbiol. 41*, 103-126.

21. García-Olmedo, F., Molina, A., Alamillo, J. M., and Rodríguez-Palenzuéla, P. (1998) Plant defense peptides. *Peptide Science. 47*, 479-491.

22. Ganz, T., Selsted, M. E., Szklarek, D., Harwig, S. S., Daher, K., Bainton, D. F., and Lehrer, R. I. (1985) Defensins. natural peptide antibiotics of human neutrophils. *J. Clin. Invest. 76*, 1427-1435.

23. Lehrer, R. I. (2004) Primate defensins. *Nature Rev. Microbiol. 2*, 727-738 ER.

24. Hoskin, D. W., and Ramamoorthy, A. (2008) *Biochim. Biophys. Acta. 1778*, 357.

25. Hwang, P. M., and Vogel, H. J. (1998) *Cell. Biol. 76*, 235.

26. Hicks, R. P., Abercrombie, J. J., Wong, R. K., and Leung, K. P. (2012) Antimicrobial peptides containing unnatural amino acid exhibit potent bactericidal activity against ESKAPE pathogens. *Bioorg. Med. Chem.*

27. Hicks, R. P., and Clark, T. D. (In Press) Development of Synthetic Antimicrobial Peptides as Therapeutic Agents to Treat Drug Resistant Bacterial and Fungal Infections, in *Frontiers in Anti-Infective Drug Discovery*.

28. Chan, D. I., Prenner, E. J., and Vogel, H. J. (2006) Tryptophan- and arginine-rich antimicrobial peptides: Structures and mechanisms of action. *Biochimica et Biophysica Acta (BBA) - Biomembranes. 1758*, 1184-1202.

29. Trabulo, S., Cardoso, A. L., Mano, M., and De Lima, M. C. P. (2010) Cell-penetrating Peptides"Mechanisms of cellular uptake and generation of delivery systems. *Pharmaceuticals. 3*, 961-993.

30. Morris, M. C., Depollier, J., Mery, J., Heitz, F., and Divita, G. (2001) A peptide carrier for the delivery of biologically active proteins into mammalian cells. *Nat Biotech. 19*, 1173-1176.

31. Zhao, W., Róg, T., Gurtovenko, A. A., Vattulainen, I., and Karttunen, M. (2008) Role of phosphatidylglycerols in the stability of bacterial membranes. *Biochimie. 90*, 930-938.

32. Epand, R. M., and Epand, R. F. (2011) Bacterial membrane lipids in the action of antimicrobial agents. *Journal of Peptide Science. 17*, 298-305.

33. LeCocq, J., and Ballou, C. E. (1964) On the structure of cardiolipin*. *Biochemistry (N. Y.). 3*, 976-980.

34. Cosma, C., Sherman, D., and Ramakrishnan, L. (2003) The secret lives of the pathogenic mycobacteria. *Annu. Rev. Microbiol. 57*, 641-676.

35. Koul, A., Herget, T., Klebl, B., and Ullrich, A. (2004) Interplay between mycobacteria and host signalling pathways. *Nat Rev Micro. 2*, 189-202.

36. Abdallah, A. M., Gey, v. P., DiGiuseppe Champion, P. A., Cox, J., Luirink, J., Vandenbroucke-Grauls, C., Appelmelk, B. J., and Bitter, W. (2007) Type VII secretion mdash] mycobacteria show the way. *Nat Rev Micro. 5*, 883-891.

37. Röse, L. *Role of undecaprenyl phosphokinase in mycobacteria: Impact on biofilm formation, growth properties, persistence, and virulence.* http://edoc.hu-berlin.de/dissertationen/roese-lars-2004-06-09/HTML/chapter1.html ed.

38. Basso, L. A., Silva, Luiz Hildebrando Pereira da, Fett-Neto, A. G., Azevedo Junior, Walter Filgueira de, Moreira, Í. d. S., Palma, M. S., Calixto, J. B., Astolfi Filho, S., Santos, R. R. d., Soares, M. B. P., and Santos, D. S. (2005) The use of biodiversity as source of new chemical

entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases: A review. *Memórias do Instituto Oswaldo Cruz. 100*, 475-506.

39. van Ingen, J., Boeree, M. J., van Soolingen, D., and Mouton, J. W. (2012) Resistance mechanisms and drug susceptibility testing of nontuberculous mycobacteria. *Drug Resistance Updates. 15*, 149-161.

40. *Mycobacterium tuberculosis: Infection Control/Exposure control issues for oral healthcare workers course pages | DentalCare.com* http://www.dentalcare.com/en-US/dental-education/continuingeducation/ce316/ce316.aspx?ModuleName=coursecontent&PartID=3&SectionID=1 ed.

41. Zuber, B., Chami, M., Houssin, C., Dubochet, J., Griffiths, G., and Daffé, M. (August 15, 2008) Direct visualization of the outer membrane of mycobacteria and corynebacteria in their native state. *Journal of Bacteriology. 190*, 5672-5680.

42. Brogden, K. A. (2005) Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nat Rev Micro. 3*, 238-250.

43. Kandasamy, S. K., and Larson, R. G. (2004) Binding and insertion of α-helical anti-microbial peptides in POPC bilayers studied by molecular dynamics simulations. *Chem. Phys. Lipids. 132*, 113-132.

44. Shai, Y. (1999) Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by α-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1462*, 55-70.

45. Huang, H. W. (2006) Molecular mechanism of antimicrobial peptides: The origin of cooperativity. *Biochimica et Biophysica Acta (BBA) - Biomembranes. 1758*, 1292-1302.

46. Dennison, S. R., Whittaker, M., Harris, F., and Phoenix, D. A. (2006) Anticancer α-helical peptides and structure / function relationships underpinning their interactions with tumour cell membranes. *Curr. Protein Peptide Sci. 7*, 487-499.

47. Murzyn, K., and Pasenkiewicz-Gierula, M. (2003) *Construction of a toroidal model for the magainin pore.* Springer Berlin / Heidelberg.

48. Bocchinfuso, G., Palleschi, A., Orioni, B., Grande, G., Formaggio, F., Toniolo, C., Park, Y., Hahm, K., and Stella, L. (2009) *Different mechanisms of action of antimicrobial peptides: Insights from fluorescence spectroscopy experiments and molecular dynamics simulations.* John Wiley & Sons, Ltd.

49. Yang, L., Harroun, T. A., Weiss, T. M., Ding, L., and Huang, H. W. (2001) Barrel-stave model or toroidal model? A case study on melittin pores. *Biophys. J. 81*, 1475-1485 ER.

50. Straus, S. K., and Hancock, R. E. W. (2006) Mode of action of the new antibiotic for gram-positive pathogens daptomycin: Comparison with cationic antimicrobial peptides and lipopeptides. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1758*, 1215-1223.

51. Shai, Y., and Oren, Z. (2001) From "carpet" mechanism to de-novo designed diastereomeric cell-selective antimicrobial peptides. *Peptides. 22*, 1629-1641.

52. Sengupta, D., Leontiadou, H., Mark, A. E., and Marrink, S. (2008) Toroidal pores formed by antimicrobial peptides show significant disorder. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1778*, 2308-2317.

53. Yeaman, M. R., and Yount, N. Y. (2003) Mechanisms of antimicrobial peptide action and resistance. *Pharmacological Reviews. 55*, 27-55.

54. Shai, Y. (1999) Mechanism of the binding, insertion and destabilization of phospholipid bilayer membranes by alpha-helical antimicrobial and cell non-selective membrane-lytic peptides. *Biochim. Biophys. Acta. 1462*, 55-70.

55. Yang, L., Harroun, T. A., Weiss, T. M., Ding, L., and Huang, H. W. (2001) *Barrel-stave model or toroidal model? A case study on melittin pores.* Cell Press.

56. Campagna, S., Saint, N., Molle, G., and Aumelas, A. (2007) Structure and mechanism of action of the antimicrobial peptide piscidin. *Biochemistry (N. Y.). 46*, 1771-1778.

57. Monaco, V., Formaggio, F., Crisma, M., Toniolo, C., Hanson, P., Millhauser, G., George, C., Deschamps, J. R., and Flippen-Anderson, J. L. (1999) Determining the occurrence of a 310-helix and an α-helix in two different segments of a lipopeptaibol antibiotic using TOAC, a nitroxide spin-labeled Cα-tetrasubstituted α-aminoacid. *Bioorg. Med. Chem. 7*, 119-131.

58. Pálffy, R., Gardlík, R., Behuliak, M., Kadasi, L., Turna, J., and Celec, P. (2009) On the physiology and pathophysiology of antimicrobial peptides. *Mol Med. 15*, 51-59.

59. Huang, H. W. (2000) Action of antimicrobial peptides: Two-state model. *Biochemistry. 39*, 8347-8352.

60. Dennison, S. R., Wallace, J., Harris, F., and Phoenix, D. A. (2005) *Protein Pept. Lett. 12*, 31.

61. Jo, S., Kim, T., and Im, W. (2007) Automated builder and database of Protein/Membrane complexes for molecular dynamics simulations. *PLoS ONE. 2*, e880.

62. Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., Borschel, R., Mueller, A., Russell, A. L., Williams, B. C., and Hicks, R. P. (2010) Novel antimicrobial peptides that exhibit activity against select agents and other drug resistant bacteria. *Bioorg. Med. Chem. 18*, 5137-5147.

63. Hicks, R. P., Mones, E., Kim, H., Koser, B. W., Nichols, D. A., and Bhattacharjee, A. K. (2003) Comparison of the conformation and electrostatic surface properties of magainin peptides bound to sodium dodecyl sulfate and dodecylphosphocholine micelles. *Biopolymers. 68*, 459-470.

64. Ludtke, S. J. (1996) Membrane pores induced by magainin. *Biochemistry. 35*, 13723-13728.

65. Lee, M., Hung, W., Chen, F., and Huang, H. W. (2005) *Many-body effect of antimicrobial peptides: On the correlation between lipids spontaneous curvature and pore formation.* Cell Press.

66. Freddolino, P. L., Arkhipov, A., Shih, A. Y., Yin, Y., Chen, Z., and Schulten, K. (2008) Application of Residue-Based and Shape-Based Coarse Graining to Biomolecular Simulations. in *Coarse-Graining of Condensed Phase and Biomolecular Systems* (G. A. Voth, Ed.) pp 229, Chapman and Hall/CRC Press, Taylor and Francis Group.

67. Bond, P. J., Parton, D. L., Clark, J. F., and Sansom, M. S. P. (2008) *Coarse-grained simulations of the membrane-active antimicrobial peptide maculatin 1.1*. Cell Press.

68. Horn, J. N., Sengillo, J. D., Lin, D., Romo, T. D., and Grossfield, A. (2012) Characterization of a potent antimicrobial lipopeptide via coarse-grained molecular dynamics. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1818*, 212-218.

69. Sato, H., and Feix, J. B. (2006) Peptide–membrane interactions and mechanisms of membrane destruction by amphipathic α-helical antimicrobial peptides. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1758*, 1245-1256.

70. Tolokh, I. S., Vivcharuk, V., Tomberli, B., and Gray, C. G. (2009) Binding free energy and counterion release for adsorption of the antimicrobial peptide lactoferricin B on a POPG membrane. *Phys Rev E. 80,* 031911.

71. Russell, A. L. (2011) *Thermodynamic and spectroscopic investigations of novel antimicrobial peptides containing unnatural amino acids with model membrane systems.* Doctor of Philosophy ed., East Carolina University, Greenville, North Carolina.

72. Jensen, F. (2007) *Introduction to Computational Chemistry 2nd Edition.* John Wiley & Sons, Ltd.

73. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science. 181*, 223-230.

74. Ferrara, P., Apostolakis, J., and Caflisch, A. (2000) Computer simulations of protein folding by targeted molecular dynamics. *Proteins: Structure, Function, and Bioinformatics. 39,* 252-260.

75. Lindahl, E. e. a. (2001) *GROMACS 3.0: A package for molecular simulation and trajectory analysis.*

76. Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry. 29,* 1859-1865.

77. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER9.*

78. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004. (2003) *Gaussian 03 Gaussian inc, 340 Quinnipiac street, building 40, Wallingford, CT, 06492.*

79. Leach, A. (2001) *Molecular Modelling Principles and Applications 2nd Edition.*

80. Accelrys. *Materials studio.*

81. Delley, B. (19902003) *Accelrys DMol3*.

82. The PyMOL Molecular Graphics System. *The PyMOL molecular graphics system, version 1.2r3pre, schrödinger, LLC.* The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC ed, .

83. Red Hat, I. (1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008) *Cygwin API reference* .

84. Schlick, T. (2010) *Molecular Modeling and Simulation: An Interdisciplinary Guide, 2nd Edition.* Springer, New York.

85. Russell, A. L., Kennedy, A. M., Spuches, A. M., Venugopal, D., Bhonsle, J. B., and Hicks, R. P. (2010) Spectroscopic and thermodynamic evidence for antimicrobial peptide membrane selectivity. *Chem. Phys. Lipids. 163*, 488-497.

86. Russell, A. L., Williams, B. C., Spuches, A., Klapper, D., Srouji, A. H., and Hicks, R. P. (2012) The effect of the length and flexibility of the side chain of basic amino acids on the binding of antimicrobial peptides to zwitterionic and anionic membrane model systems. *Bioorg. Med. Chem. 20*, 1723-1739.

87. Russell, A. L., Kennedy, A. M., Spuches, A. M., Gibson, W. S., Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., and Hicks, R. P. (2011) Determining the effect of the incorporation of unnatural amino acids into antimicrobial peptides on the interactions with zwitterionic and anionic membrane model systems. *Chem. Phys. Lipids. 164*, 740-758.

88. Russell, A. L., Spuches, A. M., Williams, B. C., Venugopal, D., Klapper, D., Srouji, A. H., and Hicks, R. P. (2011) The effect of the placement and total charge of the basic amino acid clusters on antibacterial organism selectivity and potency. *Bioorg. Med. Chem. 19*, 7008-7022.

CHAPTER 2
FORCE FIELD DEVELOPMENT

ABSTRACT

Molecular dynamics (MD) simulations rely upon the force field parameters specified for each atom of a molecule within a given system. The force field (FF) parameters for the AMBER FF have been optimized for modeling the behavior of peptides including the 20 standard amino acids. The FF charge parameters for each of the nonstandard amino acids used in this study have been developed using high level quantum mechanical (QM) calculations using the same basis sets, 6-31G*, and level of theory i.e. Hartree-Fock (HF), as the standard amino acid parameters of the amber ff03 force field which has been applied to the other atoms in the simulation. The unnatural amino acids parameterized included beta alanine (β-Ala), gamma amino butyric acid (GABA), 6-amino hexanoic acid (6-AHX), ornithine (ORN), diamionbutanoic acid (DAB), diamionpropanoic acid (DPR), octahydroindole carboxylic acid (OIC), and tetrahydroisoquinoline carboxylic acid (TIC) (Figure 2.1). FF charge parameters were also calculated for the lipid molecules POPC, POPG, and SDS (Figure 2.2) using QM optimization followed by single point energy calculations using the density functional theory of the DMol3 program.

2.1 FORCE FIELD BACKGROUND

Providing a through explanation of the intricacies involved in FF development and their application in molecular MD is not possible in a few pages. The following pages seek to provide a superficial understanding of the theory which supports the *in silico,* or computationally derived data. The reader is directed to several extensive resources which are able to provide a more

complete explanation of the physical chemical and mathematical basis which supports *in silico* techniques.[1-3]

The success or failure of any MD simulation is based upon the parameters specified in the FF which determines how each atom and molecule behaves within the experimental system. While the placement and orientation of each molecule in the system play a role in the short term behavior of the system, any improper placement should be overcome if the FF parameters are able to accurately model the behavior of the atomic behavior. The mathematical functions describe the potential energy and are usually derived from experimental work, high-level quantum mechanical calculations, or both.[1-3] Due to the basis of the molecular mechanics (MM) of MD being based upon QM foundations the discussion will start with QM and its relationship to MM.

The potential energy of a protein system is a function of the energetic properties of the interatomic interactions and is expressed by the equations of motion, which is based on the atomic coordinates.[1-3] The coordinates may vary however; the parameters are an invariant variable during the course of the simulation. The coordinates are an invariant variable because they describe the geometric and energetic properties of the interparticle interactions however the type of atom at each coordinate does not change even though the coordinates do change.[3] Numerous FFs are available and have been specialized to more accurately describe the energy associated with unique systems such as proteins, small molecules, and metal atoms. All-atom FFs model individual atoms of a system, including hydrogen atoms while a united-atom FF identifies hydrogen attached to both carbon and hetero atoms as a single charge center.[1] For example the CH_2 groups in a united-atom force field would be treated as though there was a CH_2 atom type.[3] During the course of a MD simulation the forces applied to each atom are calculated

based upon the environmental conditions such as temperature.[3] Coarse-grained (CG) FFs can expedite simulations with large atom counts by merging multiple atoms together allowing longer time scales to be achieved through calculations.[4] Use of a CG force field which creates so called CG "sites" is known as coarse graining.[4] The MD simulations undertaken herein utilize an all atom FF. Atoms are considered as spheres with a specific radius and charge as opposed to a system consisting of individual sub-atomic particles (electrons, neutrons, and protons).[3] The result is that the bonding information must be explicitly provided as opposed to it being derived from first principles by solving the electronic Schrödinger equation, equation 2.1.[1]

$$\hat{H}\Psi_n = E_n\Psi_n \qquad \text{Eqn. 2.1}$$

Where $\hat{H}$ is the Hamiltonian operator, which is the sum of the kinetic and potential energy, E_k and E_p respectively.[1] E_n is the quantum states (eignevalues) which form a discrete set and correspond to the eigen functions, Ψ_n for the electrons and nuclei of a system.[1] The Schrödinger equation provides the energy states for a stationary quantum system.

The Born-Oppenheimer Approximation to the Schrödinger equation separates the motions of molecule movement into two levels, electrons and nuclei.[1] Electrons are treated as independent variables of $\hat{H}$ assuming the nuclei are fixed which is reasonable due to the large discrepancy in their masses.[1] The resulting eigenvalues represent the electronic energy levels of a molecule as a function of atomic coordinates which is known as Born-Oppenheimer energy surfaces (BOES).[1] The second level of the Born-Oppenheimer approximation enables investigation of the QM behavior of the nuclei using the electronic ground state as the potential energy of the Hamiltonian instead of the Coulombic potential.[1] The coupling between the nuclei and electronic motion is neglected.[3] The electronic portion is solved with nuclear positions as

parameters enabling the resulting potential energy surface (PES) to form the basis for solving the nuclear motion.[3] Solving such equations numerically requires significant computational resources necessitating a more efficient solution. The assumption that one electron is moving independently of the dynamics of all other electrons in a system means that the interactions between particles can be approximated by taking all interactions into account in an average which is called Hartree-Fock (HF) theory.[3] Using the HF model, electrons are described by an orbital, and the total wave function is given as a product of orbitals. Solving the Slater determinants for the molecular orbitals results in the lowest energy (of the molecule) with the restriction of the wave function being a single Slater determinant.[3] HF only accounts for the average electron-electron interactions neglecting the correlation between those electrons.[1, 3] Inclusion of correlation functions requires multi-determinant wave functions, since HF is the best single determinant wave function; they become computationally expensive but can improve the solutions of the Schrödinger equation.[3] The difference between the HF calculated energy and the lowest possible energy (global minima) are referred to as the Electron Correlation (EC) energy.[3]

Gaussian is a computational software which was originally developed by John Pople and initially released in 1970, it has evolved since.[5] It utilizes Gaussian orbitals to expedite calculations rather than using Slater-type orbitals.[5] The quantum chemical approach uses the Hartree-Fock (HF) method which approximates the ground state wave function for a multi-bodied system. Approximating solutions to the multi-bodied Hamiltonian approximates the true many-body wave function which is accomplished by solving a single Slater determinant of N spin-orbitals for each atom in a system.[1, 3]

QM should be the most reliable description of chemical processes due to its lack of experimental values involved in the calculation. Limitations arise from the approximations

required for small molecules since the Schrödinger equation can only be solved for hydrogen. Most molecules require the use of standard approximations further limiting the accuracy while requiring large computational resources.[1] The HF method was used in the optimization of the starting structure of the unnatural amino acids. Calculations used the 6-31G* basis set. The two types of basis functions, Slater Type Orbitals (STO) and Gaussian Type Orbitals (GTO) are used to improve the accuracy of the *ab initio* derived solution to the Schrödinger equation. The 6-31G* basis set is a STO, with a split valence basis where the core orbitals are a contraction of six primitive Gaussian Type Orbitals (PGTOs).[3] Split valence basis sets describe valence electrons using twice as many functions with exception of the core orbitals which are expected to vary minimally.[2] The 6-31G* basis set is not the most exhaustive set of functions but, it balances accuracy and computational expense. While larger basis sets, and more polarization functions are possible they do not always result in model improvement. It was selected to be consistent with the QM calculations done for the original development of force field parameters in the ff03 force field employed in the MD simulation.[6] The * denotes the use of polarization basis functions.[5]

Using a self-consistent field (SCF) calculation refines the wavefunction until self-consistency is reached.[1-3] SCF criterion for convergence was set at 'tight,' meaning full convergence was requested for each structure. A SCF uses an initial set of orbitals to generate a new set of orbitals and the procedure is repeated until the convergence criteria has been met. In this study the RMS convergence criteria has been set at 1.00×10^{-8}. Convergence criteria are a measure of the root mean square deviation (RMSD) from the starting structure. When the RMSD is no longer changing it is expected that a minimum energy structure has been identified. The electrostatic-potential derived atomic charges came from the Merz-Singh-Kollman (MK)

scheme in which atomic charges are fitted to reproduce the molecular electrostatic potential (MEP) at various points around the molecule.[2] MK fits the charges so the sum of all atomic charges equals the overall charge of the system.[7]

Structures were constructed using Cerius 2[7]to create the starting structures. Capping residues ACE, an acetyl group and NME, an N-methyl group were added to neutralize the backbone charges at the terminals. Gaussian 03[8] input files were created using the "antechamber" program.[5] Files were checked for atomic accuracy, appropriate charge, and the correct spin multiplicity. Spin multiplicity is given by 2S + 1, where S is the total electron spin for the molecule.[9] Paired electrons have a net zero spin and do not contribute.[5] One unpaired electron (of spin ½) results in a spin multiplicity of 2, the species is a doublet.[5] The amber program antechamber[5] calculated RESP charges based on the optimized, minimum energy structure identified by Gaussian03. Use of a "MAINCHAIN.txt" file, which specifies the head atom name, the amine nitrogen, the main chain atom, the alpha carbon, and the tail atom which is the carbonyl carbon.

The AMBER9[9] program was used to create the .prepin files which specifies the atomic coordinates, molecular charge, atomic charges as well as improper angles. Improper angles are "virtual" torsion angles which are specified in order to keep an atom and its neighbors in a specific configuration; this is the out of the plane bending which occurs. The .frcmod files, which defines any missing or undefined parameters including the atomic mass, bond distances, angles and dihedrals in addition to improper and non-bonded interactions if necessary. These parameter files were added to the FF parameter library of tLeap or xLeap, enabling peptide construction including the unnatural amino acids. tLeap and xLeap are both programs within the

AMBER 9 suite of programs.[9]

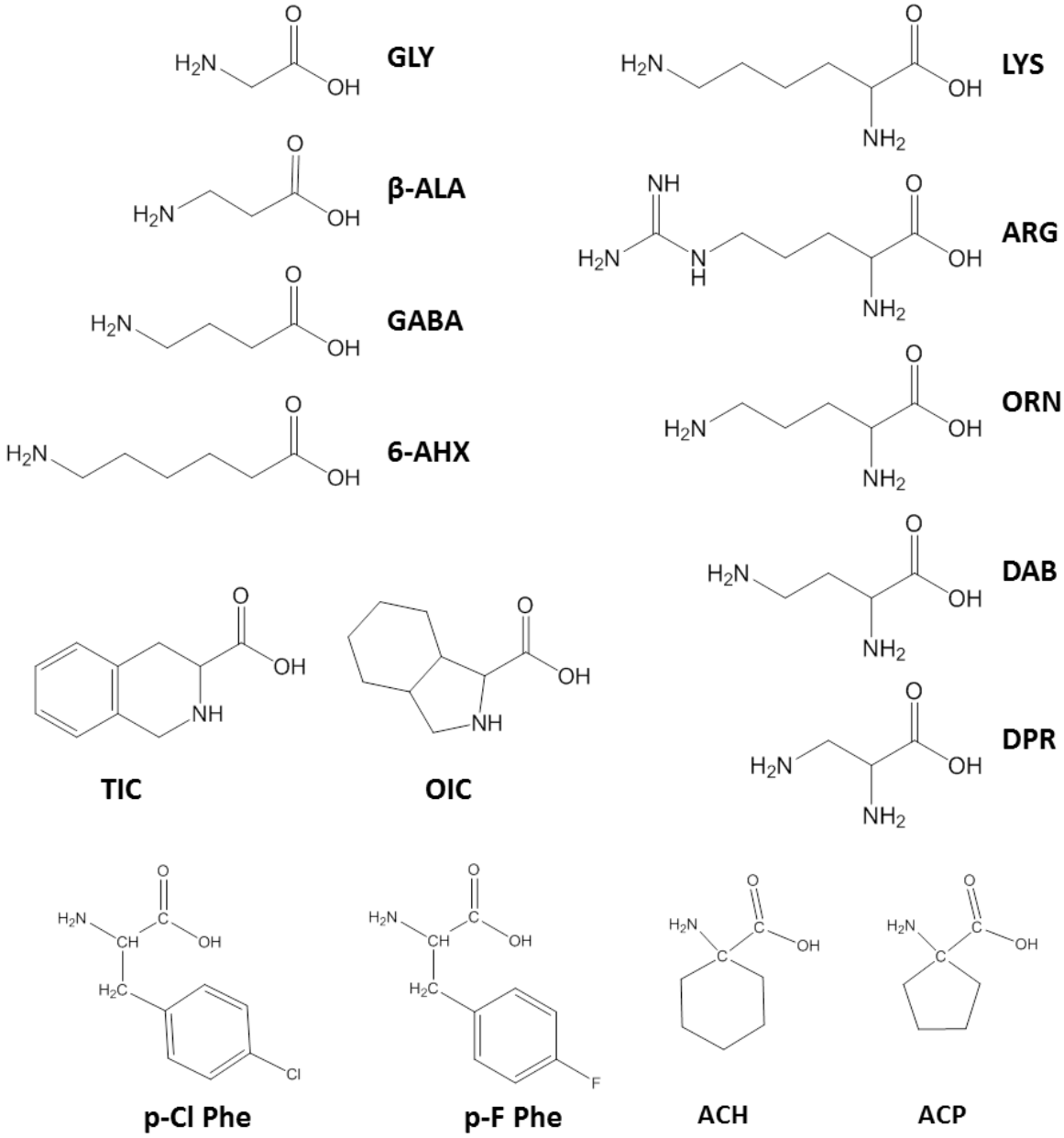


Figure 2.1 Amino acids for which force field parameters were developed.

2.2 DENSITY FUNCTIONAL THEORY

Density functional theory (DFT) is a QM modeling method which is used to understand the ground state electronic structure of many-body systems by calculating variational self-consistent solutions to DFT equations, expressed in numerical atomic orbital basis sets for many-body systems.[9] DFT is based on the proof by Hohenberg and Kohn which shows the ground state electronic energy can be determined completely by the electron density, ρ.[3, 10] DFT is an *ab initio* method. In principal DFT offers an exact energy value however, in practice, the approximation of the exchange correlation functional creates an inexact solution.[3] The wave function of the Schrödinger equation is represented by a single ground-state wave function where the electron density is represented as the sum of squares of orbital densities.[1] Using a DFT approach provides a good combination of accuracy and necessary computational resources enabling large systems to be studied more quickly than using traditional *ab initio* methods.[1]

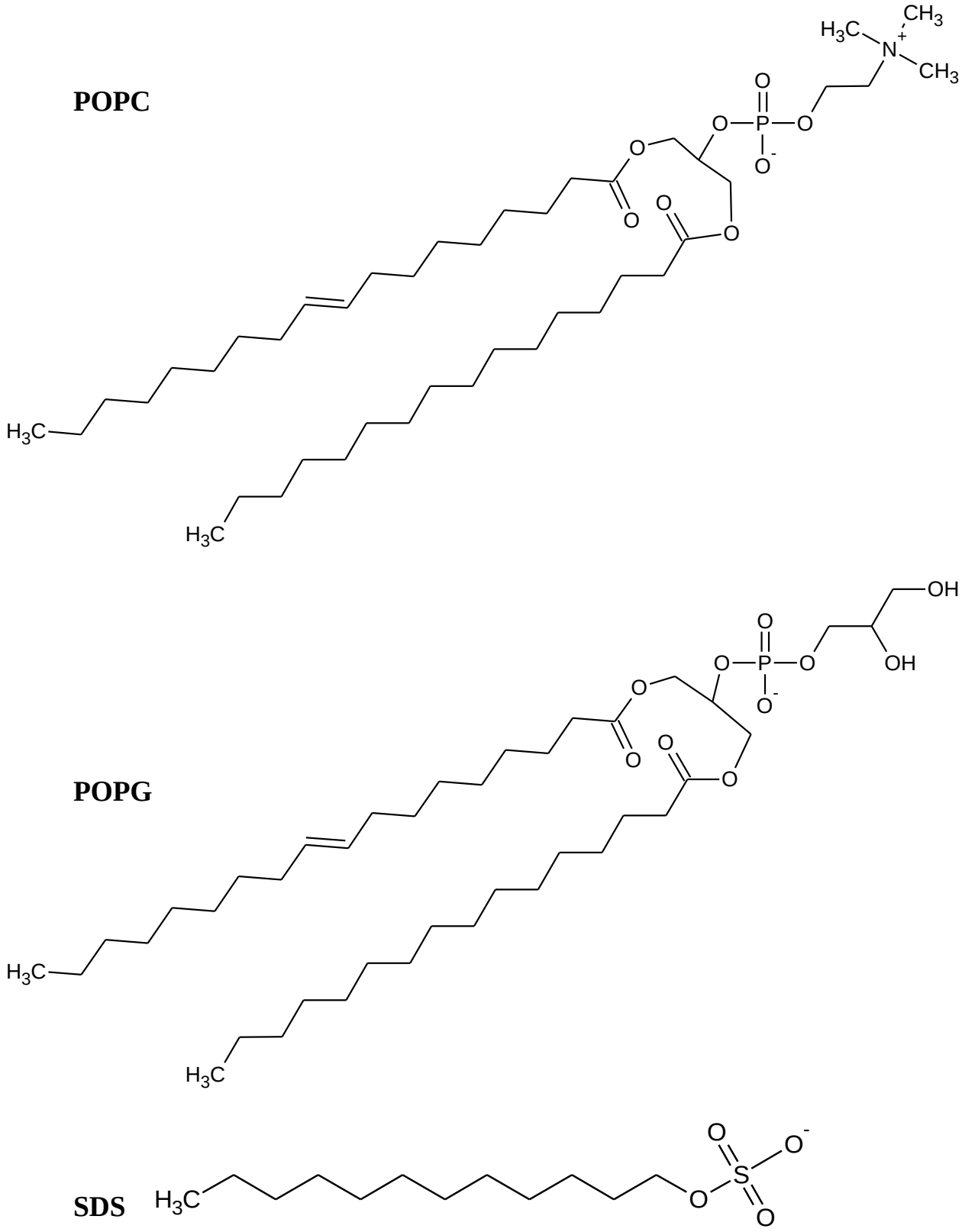


Figure 2.2 Lipid molecules parameterized using DFT. The R group in the top lipid is either a choline head group (POPC) or a glycerol head group (POPG). SDS is the lipid molecule used in the micelles.

Energy is minimized with respect to changes in ρ, subject to the constraints of charge and van der Waals interactions.[1] ρ is the electron density with respect to a specific point in space. In general it is easier to deal with electron density than the wave functions, however DFT is an iterative process (like HF) which makes it a better quantitative approach. When compared to the QM calculations it is still computationally expensive. An iterative process repeats itself until the desired target, in this instance a low energy structure with specified convergence criteria, has been achieved. There are no approximations which necessitate measurement of the exchange correlation energy (which is a function of ρ).[2] The equations are solved through approximations until convergence is achieved. When comparing HF and DFT, DFT will produce better result with comparable computational expense.[3] The lipid molecules used as component parts of the model membranes are shown in Figure 2.2.

2.3 PEPTIDE CONSTRUCTION

tLeAP and xLeAP from the AMBER9[3] suite of programs were utilized in the construction of the peptide sequences after successful library formation adding parameters for the non-parameterized residues. The peptide sequences listed in **Table 2.1** were constructed using the sequence command in tLeap. The MD simulations were conducted on the generic peptide sequence:

Ac-FX-Tic-Oic-**X**K-Tic-Oic-**X**F-Tic-Oic-**X**K-Tic-KKKK-$CONH_2$

where X is either the natural amino acid glycine, or the nonstandard amino acid β-alanine. After the peptides were constructed they had multiple bad contacts which necessitated a pre-minimization step using 10 000 steps of steepest descent optimization with Discovery Studio.[9]

After the pre-minimization step the atomic coordinates were appropriately spaced, and any overlap which was initially present had been removed.

Table 2.1: Representative Amino acid sequences for the AMP developed by Hicks and co-workers[11]

Peptide #	**amino acid sequence**
23	Ac-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-KKKK-$CONH_2$
29	Ac-Gaba-F-Tic-Oic-Gaba-K-Tic-Oic-Gaba-F-Tic-Oic-Gaba-K-Tic-KKKK-$CONH_2$
35	Ac-GF-F-Oic-GK-F-Oic-GF-F-Oic-GK-F-KKKK-$CONH_2$
36	Ac-βAla-F-Tic-Oic-βAla-K-Tic-Oic-βAla-F-Tic-Oic-βAla-K-Tic-KKKK-$CONH_2$
37	Ac-Ahx-F-Tic-Oic-Ahx-K-Tic-Oic-Ahx-F-Tic-Oic-Ahx-K-Tic-KKKK-$CONH_2$
38	Ac-F-Tic-Oic-K-Tic-Oic-F-Tic-Oic-K-Tic-KKKKKK-$CONH_2$
40	Ac-Gaba-F-Tic-Oic-Gaba-K-Tic-Oic-Gaba-F-Tic-Oic-Gaba-K-Tic-KKKKK-$CONH_2$
41	Ac-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-Orn-Orn-Orn-Orn-$CONH_2$
42	Ac-G-Fpa-Tic-Oic-GK-Tic-Oic-G-Fpa-Tic-Oic-GK-Tic-KKKK-$CONH_2$
43	Ac-GF-Tic-Oic-G-Orn-Tic-Oic-GF-Tic-Oic-G-Orn-Tic-Orn-Orn-Orn-Orn-$CONH_2$
45	Ac-GF-Tic-Oic-G-Dpr-Tic-Oic-GF-Tic-Oic-G-Dpr-Tic-Dpr-DprDpr-Dpr-$CONH_2$
46	Ac-βAla-Fpa-Tic-Oic-βAla-Dpr-Tic-Oic-βAla-Fpa-Tic-Oic-βAla-Dpr-Tic-Dpr-Dpr-Dpr-Dpr-$CONH_2$
50	Ac-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-βAla-KKKK-$CONH_2$
52	Ac-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-Ahx-KKKK-$CONH_2$
53	Ac-GF-Tic-Oic-G-Dab-Tic-Oic-GF-Tic-Oic-G-Dab-Tic-Dab-Dab-Dab-Dab-$CONH_2$
56	Ac-GF-Tic-Oic-GR-Tic-Oic-GF-Tic-Oic-GR-Tic-RRRR-$CONH_2$
61	Ac-KKKK-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-$CONH_2$
62	Ac-GF-Oic-Oic-GK-Oic-Oic-GF-Oic-Oic-GK-Tic-KKKK-$CONH_2$
64	Ac-GF-Tic-Oic-GK-Tic-Oic-GF-Tic-Oic-GK-Tic-KKK-$CONH_2$

Anfinsen's theory proposes a correlation of the secondary structure to the lowest free energy state of the primary protein sequence.[12, 13] Predicting the secondary structure of a peptide using *in silico* techniques has been of great interest due to the theorized correlation between the shape

and the function. Despite the recent success of structure prediction of peptides and proteins, which are restricted to the 20 standard amino acids, inclusion of nonstandard and unnatural amino acids in the peptide sequence adds several degrees of difficulty. Since the unnatural amino acids require synthetic means for their creation relatively less is known about their secondary structure and how that may correlate to observed antimicrobial activity. Given the importance associated with secondary structure, developing a technique which identifies low energy peptide conformations is expected to provide insight into the mechanism through which they act. Simulated annealing (SA) had been employed to further understand low energy secondary structures associated with peptides **23**, **29**, **36**, and **37**.

2.4 SIMULATED ANNEALING

MD utilizes temperature as a guide for the structural evolution of the molecules being simulated. Proteins with a well-defined secondary structure are known to lose structure at high temperatures. Simulated annealing uses the increase in molecule flexibility at high temperatures to determine the lowest energy structure. Simulated annealing uses temperatures as high as

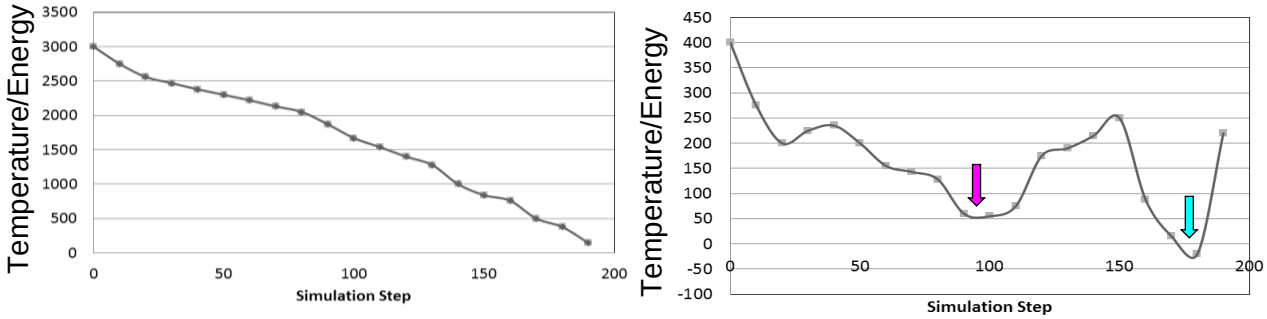


Figure 2.3 A) Exemplary plot reflecting the temperature and energy correlation to simulation step number and time if an infinitely slow cooling method is employed. B) Plot reflecting the temperature and energy correlation for actual simulated annealing where hundreds of thousands of steps can be employed in the cooling process

2000-3000 K to disrupt the intramolecular interactions responsible for the secondary structure of a peptide or protein. At high temperatures a wide area of the conformational space is sampled enabling the discovery of preferred, low energy atomic arrangements as the system is cooled to room temperature. If the cooling of the system occurred at an infinitely slow rate, the global minima should be achieved, see Figure 2.3 A. Due to the time restriction, SA seeks to model the results of an infinitely slow cooling process with a more reasonable time scale. SA however due to its nature does not ensure the global minima, or lowest possible energy structure, has been found for a molecule. Figure 2.3 B is a sample plot which shows the correlation of temperature and energy with the simulation step number and time. While the cooling process is able to find a minimum energy structure such as is noted by the pink arrow sometimes it gets "stuck" and has an energy barrier which is too great preventing identification of the global minima, noted in Figure 2.3 B with the cyan arrow. If an infinitely slow cooling process were feasible, the lowest energy structure would be identified. An infinitely slow cooling process would require an infinite amount of time, expediting the cooling process introduces more error while making good approximations within timescales enabling useful data collection despite the introduced errors.

2.5 MOLECULAR DYNAMICS

MD simulations can provide a quantitative and or a qualitative perspective correlating atomic movement to macroscopic observables, which in turn correlate to biological activity. Use of MD enables the potential effects of a change in the protein sequence; for example including nonstandard amino acids in the peptide sequence. Determining the effect of the new residues, such as whether the change will result in a large effect or if the change will be silent resulting in

negligible change in peptide behavior, needs to be determined. Observing the behavior of molecules at the atomic level can be correlated to the observable effects.

The trajectory resulting from this calculation specifies how a molecule moves over time based on the starting position and the velocity [14] based on Newton's equations of motion and is representative of what is observed in a physiologically relevant system. Newton formulated three laws pertaining to motion: 1. An object continues to move in a straight line at constant velocity unless acted upon by force, 2. Force is equal to the rate of change of the momentum of the object, and 3. Every action has an equal and opposite reaction.[2] MD is based on statistical mechanics and Newton's second law of motion which describes the relationship between force and momentum.

$$F = \frac{d\boldsymbol{p}}{dt} = \frac{d(m\mathbf{v})}{dt} = m\mathbf{a} \qquad \text{Eqn. (2.2)}$$

m**v** is the canonical momentum. The net force exerted on a particle is thus equal to the rate at which the momentum of the particle changes over time. Acceleration is $\mathbf{a} = \frac{dp}{dt}\frac{dv}{dt}$ the second law can be simplified as above (Equation 2.2).

The force field parameters can be applied to collections of atoms using various ensembles to model the experimental, *in vitro*, or *in vivo* conditions using constant-temperature and constant-pressure ensembles can be simulated. Starting with the acceleration, the calculation of the velocity and positioning by using integrals at any time. Newton's equation of motions is numerically solved using the finite difference methods to obtain a trajectory of the system. The trajectory is a set of configurations saved during a simulation, once the equation determining the

energy is solved the value can be compared with the starting energy value. The Assisted Model Building and Energy Refinement (AMBER) force field has been optimized for use with proteins and DNA.[2] The ff03 FF has charges derived from QM calculations that use a continuum dielectric to model the solvent polarization, and the ϕ and Ψ backbone torsions for proteins have been modified decreasing the preference for helical configurations for proteins. The gradient of the potential energy, U, function is determined by summing solving the potential energy equation with respect to time and the position of the atoms in the system. The potential energy is described by the following equation:

$$U(\mathbf{R}^{\mathrm{N}}) = U_{bonds} + U_{angles} + U_{dihedrals} + U_{van\ der\ Waals} + U_{electrostatics}$$

Eqn. (2.3)

$$U(\mathbf{R}^{\mathrm{N}}) = \sum_{bonds} K_r\left(r - r_{eq}\right)^2$$

$$+ \sum_{angles} K_\theta(\theta - \theta_{eq})^2$$

$$+ \sum_{dihedrals} \frac{V_n}{2}\left(1 + \cos[n\emptyset - \gamma]\right)$$

$$+ \sum_{i<j}^{atoms} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \qquad van\ der\ Waals$$

$$+ \sum_{i<j}^{atoms} \frac{q_i q_j}{\varepsilon\, R_{ij}{}_{ij}} \qquad electrostatics$$

$$+ \sum_{ij}^{atoms} \frac{q_i q_j}{f^{gb}(R_{ij})} + \sigma \boldsymbol{A} \qquad implicit\ solvation$$

Eqn. (2.4)

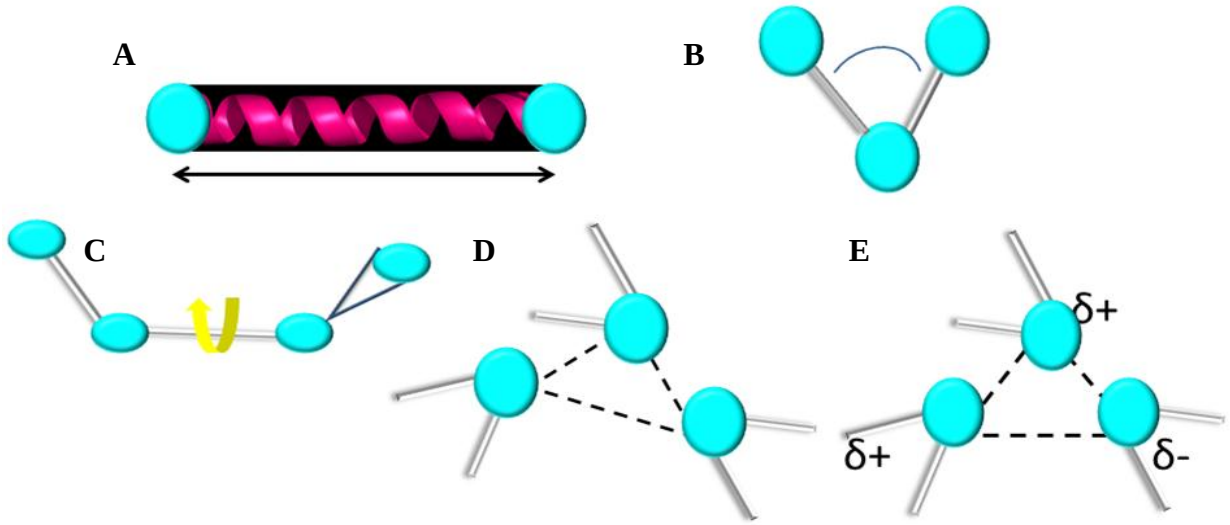


Figure 2.4 A. Internuclear Separation (bond Stretching) B. Angle Bending C. Bond rotation (torsion) D. Non-bonded Interactions (van der Waals) E) Non-bonded Interactions (electrostatic interactions. Adapted from the following reference.[2]

The parameters the FF is based upon are composed of the functional expression of the potential energy function of the molecular system (equations 2.3 and 2.4).[6] The potential energy is calculated where **R** is the function of positions of N particles, in this instance N is an atom.[1] r and r_{eq} are actual and ideal bond lengths, K_r, K_θ and V_n are force constants.[6] n is the torsional multiplicity and ϕ is a torsion and γ, a phase angle (Figure 2.4).[15] A and B are van der Waals coefficients, ϵ a dielectric constant, R_{ij} an interatomic distance and q are the atomic partial charges which are responsible for the long-range electrostatic interactions.[15]

The particle-mesh Ewald (PME) procedure is responsible for the long-range electrostatic interactions while the long range van der Waals interactions are estimated by a continuum

model.[15] The Lennard-Jones potential is a simple model which describes the interaction between a pair of neutral atoms or molecules.

$$V(r) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \qquad \text{Eqn. (2.5)}$$

Where ϵ is the depth of the potential well, σ is the distance at which the inter-particle potential is zero, and r is the distance between the particles.[6]

The MD simulation with AMBER9 enables the user to tailor many of the experimental parameters using a system of "flags" which specify the techniques applied to the simulations. The AMBER9 program has hundreds of thousands of setting variables specifying the parameters of each experimental simulation. A discussion of each variable is present in the AMBER9 user's manual.[2] Input files noting the parameters for the bilayer simulations have been included in **Appendix B**. Several of the parameters most frequently varied during the course of these experiments follow.

Periodic imaging geometries are a part of any simulation using periodic boundary conditions. AMBER 9 offers two box geometries, which the sub-programs LEaP and ptraj are able to support: rectangular parallelepiped and truncated octahedron.[6] Both have been explored during the course of this study. The repeating units are coupled to external pressures, velocities and temperatures which will each be specified as used.

Implicit solvation has been utilized in the preparation of minimum energy structures of peptide sequences which uses the generalized Born/surface area (GBSA) model which adds the following two terms to the "vacuum" potential function.[6] charge in a cavity with a specific

radius and the difference in energy between a vacuum and whatever the surrounding environment is as specified by the dielectric constant.[6]

2.6 ENERGY MINIMIZATION

Using the sequence command in tleap or xleap generates a peptide sequence which has all the appropriate atomic connections; however it may also include fictitious connections between atoms due to the software's limited experience with the peptides. Atoms are connected if there are within a reasonable bond distance and appear to have unsatisfied octets. Hydrogen atoms with their diverse naming conventions depending upon the software utilized can be problematic when loaded into tLeap or xLeap. For this reason, hydrogen atoms are removed from the peptide prior to loading it into the Leap program. Since the force field parameters for each of the amino acids of the peptide sequence are known to the program, the appropriate missing hydrogen atoms can be added to the structure. If the starting structure is too far removed from a reasonable minimum energy structure, such as having a sp3 hybridized carbon in close proximity to several other atoms which appears to be lacking a full octet, the software may not be able to recognize the true structure.

2.6.1 CONJUGANT GRADIENT VS STEEPEST DESCENT

Generally structures are optimized through the use of "minimization" steps prior to simulation. There are three classes of optimization methods which have commonly been employed and provide ways through which a minimum energy structure can be acquired. [3] The three methods of minimization are steepest descent, conjugate gradient, and Newton-Raphson methods have demonstrated utility in the identification of minimum energy structures.

Steepest descent is an iterative minimization method which attempts to improve the previous structure, based on the minimum energy calculated. The functions describing the molecule are calculated in a step wise fashion and at each step steepest descent moves the vector points down the energy scale rejecting any step which results in a higher energy value than the preceding step. Figure 2.5 A provides a depiction of the descent path which a molecule may follow relative to the energy. Steepest descent prevents movement up the gradient forcing a lower value ensuring a lower energy structure will be attained relative to the starting structure. The gradient and the direction of sequential energy sequential structures are orthogonal to one another.[3]

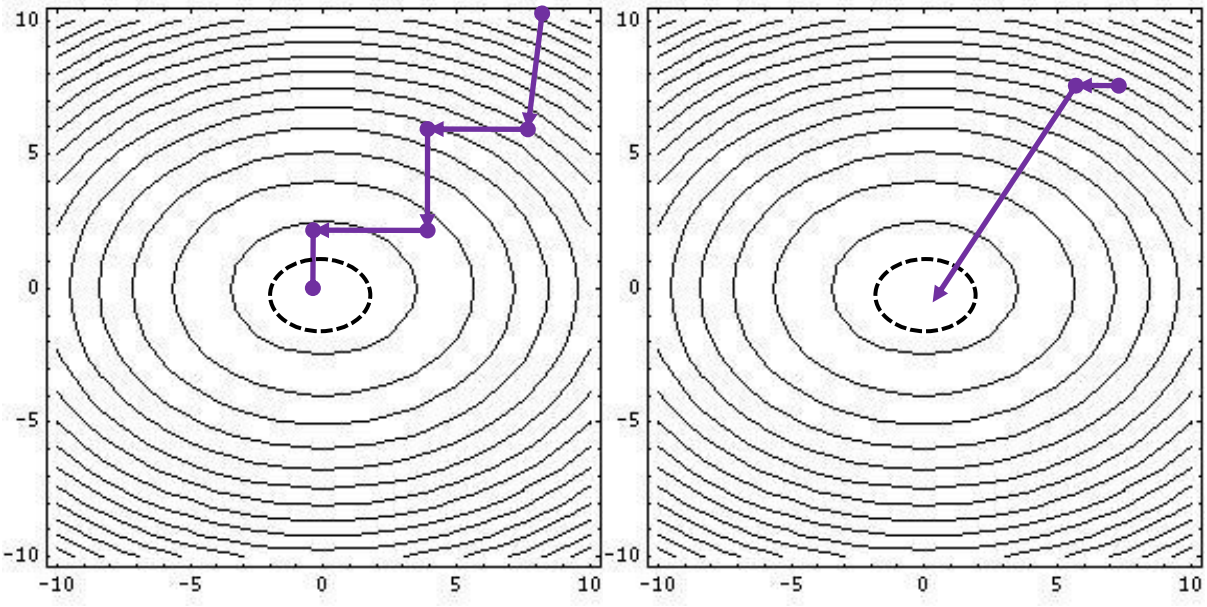


Figure 2.5 A) Steepest Descent minimization vs B) Conjugant Gradient Adapted from figures in in Frank Jensen's *Introduction to Computational Chemistry*.[3]

Conjugate Gradient (CG) tries to find a lower energy structure along a different gradient and is devoid of the oscillatory behavior of steepest descent.[2] CG moves orthogonally to the previous energy associated with a structure however the direction is conjugate. The search for

lower energy structures is not guaranteed to converge given the lack of direction associated with each step. Use of CG approach makes the starting structure very important because low energy structures are heavily dependent upon the structure which the minimization started.

CG will to converge to a minimum energy structure with the specified gradient; it just may not do so quickly. The only way to get efficient convergence is starting near the minimum (with a reasonable starting structure). Herein a combination of steepest descent and conjugate gradient has been used to improve the minimum energy structures. Use of steepest descent followed by conjugate gradient minimization will identify the best minimum energy structure with the smallest amount of CPU time.[2]

2.7 ENSEMBLES

2.7.1 ISOTHERMAL

Simulation using the isothermal ensemble creates a closed system which is simulated at constant temperature and volume.[2] Use of this ensemble is often referred to as NVT. The energy of the system is coupled to an external temperature bath. In this type of system the lower energy conformation is expected to be more prevalent than higher energy arrangements.[6]

2.7.2 ISOTHERMAL-ISOBARIC

Simulation using the isothermal-isobaric ensemble also creates a closed system which is connected to a pressure control in addition to an external temperature bath. [1, 3] Use of this ensemble is often referred to as NPT. The volume is correlated with the bath at a given pressure.[6] This correlates to states with low energy and volume is able to fluctuate.

2.7.3 GRAND CANONICAL ENSEMBLE

Use of the grand canonical ensemble occurs in a rigid cube which exchanges particles with an external reserve container which is correlated with a heat bath.[6] The energy is exchanged with an external temperature. The grand canonical ensemble is referred to as NμT. The number of molecules can vary by exchange with the reservoir at a specified chemical potential.[6]

2.7.4 ISOENERGETIC

Isoenergetic ensemble is a closed system contained in a rigid and insulated container commonly referred to as NVE.[3, 6] In this system energy is held constant and all states are equally likely thus limiting is relevance and correlation to experimental data.

2.8 MODEL OF WATER

All explicitly solvated simulations performed were solvated with the TIP3P[6] water model. Modeling the water dimer accurately has been a challenge for decades.[16] The TIP3P model, which is commonly used in CHARMM and Amber force fields) provides three sites (one

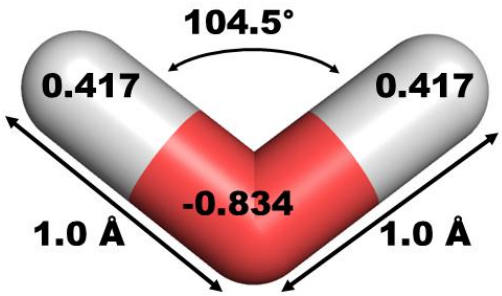


Figure 2.6 TIP3P Water model used for solvated simulations. Bond distances, angle and atomic charges.

water two hydrogen) to model the properties of liquid water.[16] The TIP3P model used with the CHARMM force field is slightly different because it includes well depth and radius for the hydrogen of the water molecules.[17] TIP3P is one of several simple models, and is considered to be rigid with fixed atom positions, which utilize pairwise potentials without use of three body terms or polarization effects and relies upon the Coulombic and Lennard-Jones expressions.[18] The TIP3P model of water has unique, slightly different geometries associated with each water molecule in the hydrogen charges of the Lennard-Jones parameters.[2] Figure 2.6 illustrates the geometry of a sample water molecule extracted from one of the solvated MD simulations. The TIP3P model was parameterized for simulations at 298 K and 1 atm; which models water at room temperature (~300 K) more accurately than physiologically relevant temperatures (310 K). Three point solvation models are very popular with molecular dynamics simulations due to the computational efficiency resulting from the mathematical simplicity of the parameter functions.[2] The selection of a water model can be influenced by many factors depending upon experimental constraints. Given the computational expense of larger models and the relatively smaller TIP3P water molecule offers a parsimonious approach. Considering the electrostatic interaction between two water molecules necessitates consideration of 9 site-site distances.[17] TIP3P water model includes the interaction calculated between oxygen and hydrogen of differing water molecules.[2] Point charges are assigned for each atom in addition to assignment of the Lennard-Jones parameter.[2] While this water model does not reproduce all of water's properties with good accuracy it provides an appropriate solvent environment for conducting simulations

2.9 PERIODIC BOUNDARY CONDITIONS

Periodic boundary conditions enable simulations to be performed using a relatively small sample of what may be a much larger system.[2] An example would be a bilayer simulation which is modeling a cell's membrane, which would be infinitely large. Through mirror images of the simulation being placed (Figure 2.7 & 2.8) on all sides of the original simulation setup a periodic array is created for the system.[2] Use of periodic boundary conditions enables fewer atoms to be simulated because the effects of the atoms in the surrounding images will be reflected in the original simulation box. An example is depicted in Figure 2.7 B. The central box represents the original simulation. If a particle moves out of the central box and into a surrounding box a corresponding particle moves in from the opposite side as is denoted with the black and red arrows. Depending upon the ensemble used during the course of the simulation, the number of particles in the simulation box remains constant through the constant replacement of particles

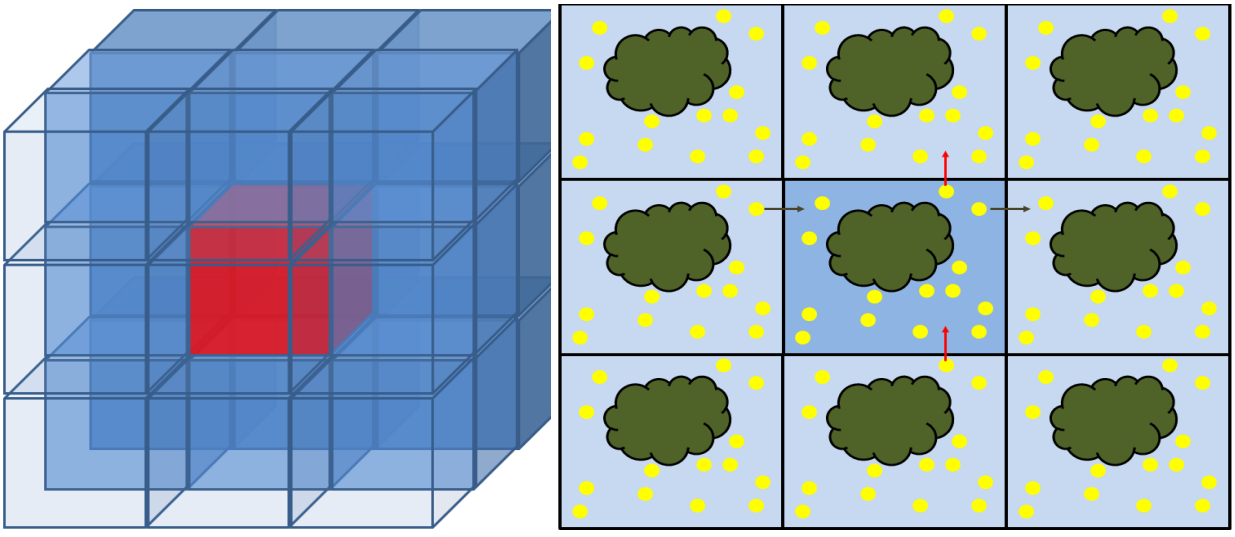


Figure 2.7 Periodic boundary conditions. A) The red cube in the center of the nonet of cubes represents the original simulation with the periodic images on all sides. B) The arrows reflect the effect of the movement of one atom out of the periodic box. Inspired by http://www.ccl.net/cca/documents/molecular-modeling/node9.html. Accessed 22 July 2010.

which move to surrounding boxes. The surrounding periodic boxes are considered to extend infinite distance from the original periodic box.

Multiple shapes of periodic systems exist besides a cubic shape including hexagonal, rhombic dodecahedron and an octahedral shape (a box with the corners cut off).[2] The diversion from the sharp corners of a cube is intended to decrease simulation time by decreasing the number of solvent molecules necessary to fill the periodic space. A smaller solvent box may be a more economic choice for simulations of small molecules or similarly solvent surrounded systems.

Use of a cubic box is especially useful when dealing with bilayer simulations which can be prepared in one of two ways for simulation. The first is to arrange the bilayer in a "sandwich" arrangement as outlined in Figure 2.7A. The central dark shaded box shows the polar head groups of the bilayer separated (by water solvent and appropriate ions) which is a variable enabling easy adjustment to accommodate peptides, small molecules, or complex systems simply by adjusting the distance between the polar head groups. Additionally, determining an appropriate size for the solvent box is relatively easy simply taking a few angstroms off the distance between the polar head groups to allow for equilibration of the system. Simulation box size is also straightforward in simply measuring the distance between the most terminal atoms of the hydrophobic region of the lipids and subtracting a few angstroms to encourage the hydrophobic tails to intermingle.

The second arrangement for a bilayer is the expected cross-section, Figure 2.7 B, with the hydrophobic tails of the lipids facing one another with the polar head groups at the surface. At first glance this appears to be the best starting point for a bilayer simulation but when considering the use of periodic boundaries the challenge of placing the solvent box (or two

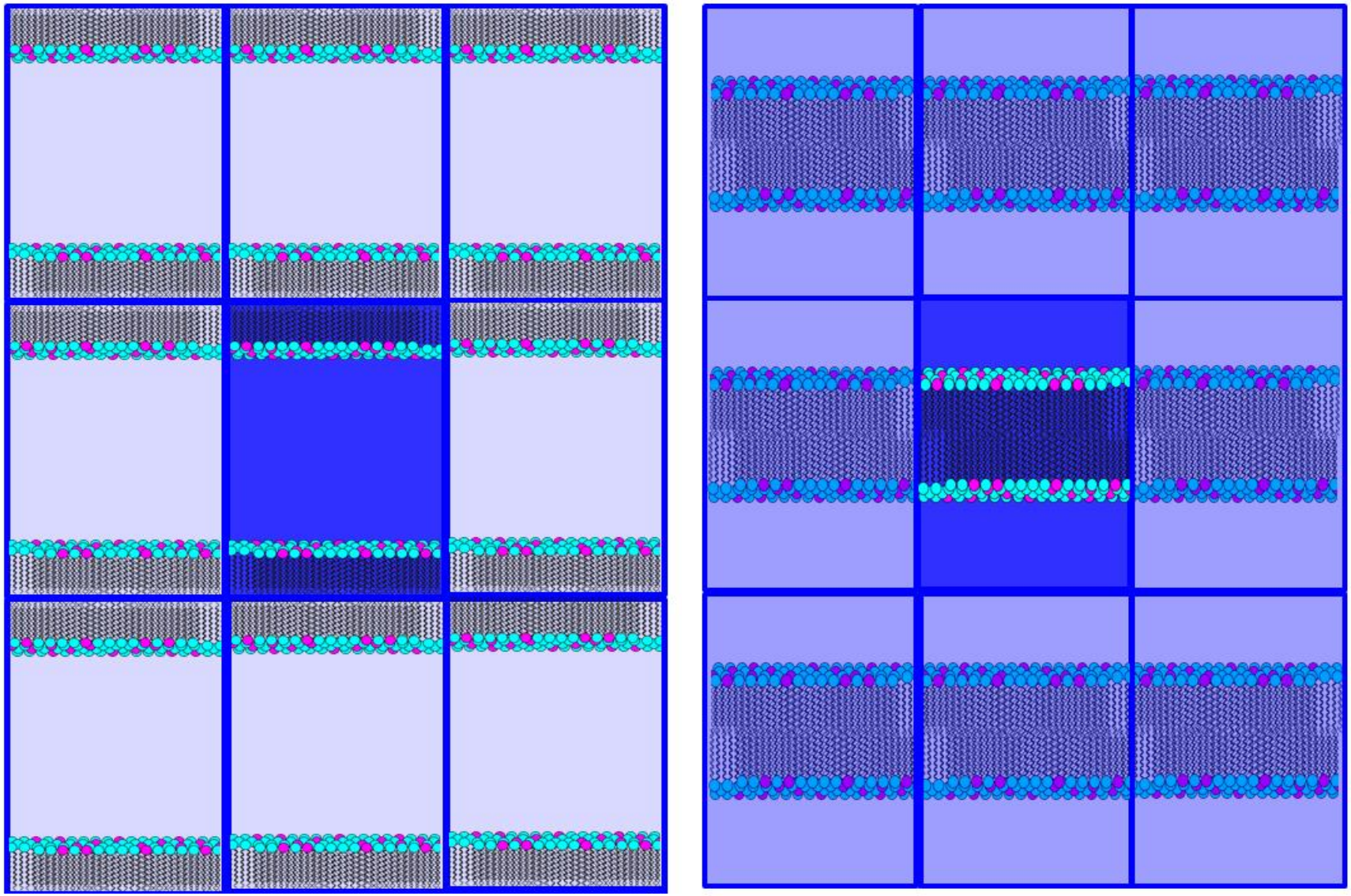


Figure 2.8 A) Sandwich configuration of membrane simulation B) bilayer membrane configuration

solvent boxes, one above the bilayer and one below, as the case may be) can present a challenge ensuring proper centering relative to the bilayer and then determining the final boundaries for the periodic system. Both of these approaches to periodic boundary simulations are technically correct however determining the appropriate placement of solvent the solvent box relative to the bilayer position is challenging and ensuring proper distribution of solvent around the bilayer is difficult to determine. For this reason use of the sandwich method will be utilized in the ensuing investigations.

2.10 THERMOSTATS

The velocities calculated for each atom at each step in an MD simulation depend upon the temperature. Due to the important role temperature plays in MD simulation and the type of simulation and the various ensembles make it is necessary to provide an external "water bath" coupling to maintain the desired temperature.

2.10.1 BERENDSEN

Use of the Berendsen temperature coupling system couples the simulation to an external heat bath with a specified temperature.[2] The external heat bath adds or removes energy as necessary enabling the velocity to be scaled proportionally to the external heat bath as the simulation progresses with respect to time.[19] The Berendsen thermostat is considered to be a weak coupling algorithm and does not generate a correct canonical ensemble. It may function better for large systems, hundreds to thousands of atoms. Use of the Nosé-Hoover thermostat is usually coupled to the Berendsen thermostat after equilibration to produce trajectories when using the canonical ensemble.[2]

2.10.2 ANDERSEN

Use of imaginary "collisions" which make the velocities random selection from the Maxwell-Boltzmann distribution,[2] and correspond to the specified temperature, at a given number of steps is known as Andersen temperature coupling.[2] The dynamics resulting from use of the Andersen temperature coupling are Newtonian which means the correlation functions are compounded and averaged over the trajectory.[6] The rate at which the collisions are allowed to occur is important is producing a simulation which models reality while not being too computationally expensive.[6]

2.10.3 LANGEVIN

"Langevin dynamics specify collision frequency which provides constant total energy classical dynamics."[20] This approach utilizes a Leapfrong integrator for the continued calculation the dynamics, and the kinetic energy is adjusted so it correlates with the harmonic oscillator case.[6] Langevin is only to be used with explicit solvent systems and models the viscous properties in a solvated system.[6]

2.11 SHAKE

SHAKE is a form of restraint applied during a MD simulation. Relative to the heavy atoms involved in a MD simulation the movement of hydrogen atoms is unlikely to affect the motion of the overall system. Here the motion associated with hydrogen atoms is removed enabling slightly longer time steps to be employed without introducing significant amounts of

error into the system.[6] The shake algorithm evolves as the coordinates change with respect to the original position of the hydrogen atoms.[21]

2.12 REFERENCES

1. Schlick, T. (2010) *Molecular Modeling and Simulation: An Interdisciplinary Guide, 2nd Edition.* Springer, New York.

2. Leach, A. (2001) *Molecular Modelling Principles and Applications 2nd Edition.*

3. Jensen, F. (2007) *Introduction to Computational Chemistry 2nd Edition.* John Wiley & Sons, Ltd.

4. Voth, G. A. (2009) *Gorse-graining of condensed phase and biomolecular systems.* CRC Press Taylor & Francis Group, Boca Raton, FL.

5. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004. (2003) *Gaussian 03, Gaussian inc, 340 Quinnipiac street, building 40, Wallingford, CT, 06492.*

6. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H.,

Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9 users' manual University of California, San francisco.* .

7. Singh, U. C., and Kollman, P. A. (1984) *An approach to computing electrostatic charges for molecules.* John Wiley & Sons, Inc.

8. *Cerius2, Version 4.10L, Conformational Search and Analysis Manual.* (2005) .

9. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9.*

10. Delley, B. (19902003) *Accelrys DMol3* .

11. Molecular Simulations Inc. (1999) *Discovery studio modeling environment, San diego:* .

12. Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., Borschel, R., Mueller, A., Russell, A. L., Williams, B. C., and Hicks, R. P. (2010) Novel antimicrobial peptides that exhibit activity against select agents and other drug resistant bacteria. *Bioorg. Med. Chem. 18*, 5137-5147.

13. Hicks, R. P., Bhonsle, J. B., Venugopal, D., Koser, B. W., and Magill, A. J. (2007; 2007) De novo design of selective antibiotic peptides by incorporation of unnatural amino acids. *J. Med. Chem. 50*, 3026-3036.

14. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science. 181*, 223-230.

15. Darden, T., Cisneros, A., Piquemal, J., and Elking, D. *Implementing advanced force fields for simulation of physical and biological processes.* http://www.niehs.nih.gov/research/atniehs/labs/lsb/ ed.

16. Jorgensen W L et, a. (1983) *Comparison of simple potential functions for simulating liquid water.*

17. Kukol, A. (2008) *Molecular modeling of proteins.* Humana Press, Totowa, NJ, USA.

18. MacKerell A D et, a. (1998) *All-atom empirical potential for molecular modeling and dynamics studies of proteins.*

19. Berendsen H J C, Postma J P M, van Gunsteren W F, DiNola A and Haak,J.R. (1984) *Molecular dynamics with coupling to an external bath.*

20. Andersen, H. C. (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys. 72,* 2384-2393.

21. Walker, R. (2004) *Amber workshop tutorials.* www.ambermd.org/tutorials ed., Ross Walker,.

CHAPTER 3
ELECTROSTATIC SURFACE POTENTIAL CALCULATIONS AND AMINO ACID ACTIVITY CORRELATIONS

ABSTRACT

Rational design has been applied to the development of small molecules for several decades and has provided a more succinct approach to identifying active molecules. Using a similar *in silico* approach to the development of synthetic antimicrobial peptides (AMPs) would optimize selectivity for the desired membrane type while minimizing the side effects by tailoring the peptide sequence to bind to the membrane of the targeted cell. The physical and structural properties associated with the unnatural amino acids used in the development of the first generation of magainin derived AMPs have been investigated and correlated with the observed *in vitro* activity. The electrostatic surface potential (ESP) was derived from the charges calculated for each of the unnatural and nonstandard amino acids. The ESP calculations were able to show the electron charge density varies with the number of carbon atoms separating the protonated amine of the side chain. The closer the charge is to the backbone the more diffused it becomes relative to the residues with the charge further from the backbone. The preferred regions of the Ramachandran plots of the 20 standard amino acids are consistent with other's experimental data in addition to the preferred torsion angles identified through spectroscopic means.[1] By extension of the reasonable angles produced for the standard amino acids, torsion angle calculations were also measured for the unnatural amino acids which varied the side chain and not the linker length within the peptide backbone. Additionally, the solvent accessible surface areas and volumes were calculated. For each residue the distances between polar atoms have been noted when relevant.

3.1 INTRODUCTION

The 20 standard amino acids encoded by the universal genetic code are the basis for most proteins, enzymes, and genetic material found in nature. Designing peptides composed entirely of the standard amino acids has some limitations including susceptibility to host enzymatic degradation. Inclusion of nonstandard and unnatural amino acids in the peptide sequence enhances their proteosomal degradation profiles and improves their therapeutic potential. The challenge associated with the inclusion of the nonstandard and unnatural amino acids is the difficulty of synthesizing them and the relatively small amount of data regarding the secondary structure associated with peptides containing such amino acids. While the 20 standard amino acid residues have been well characterized, the nonstandard and unnatural amino acids are not as well known. Nonstandard amino acids are not part of most protein sequences however they may be involved as intermediate structures in the metabolic pathways of the standard amino acids and include Gaba, β-Alanine, and ornithine. Amino acids which are not naturally encoded within the genetic code, the unnatural amino acids, require synthetic approaches for creation. The synthetically derived residues will be referred to as unnatural amino acids and include tetrahydroisoquinolinecarboxylic acid, octahydroindolecarboxylic acid, 6-aminohexanoic acid, diaminobutanoic acid, diamino propanoic acid, amino cyclohexane carboxylic acid, amino cyclopentane carboxylic acid, p-Cl phenylalanine, and p-F phenylalanine. Small molecules are continually optimized by manipulation of key structural motifs associated with improved potency, DMPK profiles, and efficacy. Each drug target has compounds which have been structurally sculpted for reasonable fit while using chemical intuition to minimize side effects resulting from activity against potential off-targets, metabolic by products etc. Developing an understanding of the structural and chemical contribution of each nonstandard and unnatural

amino acid within the medicinal chemists' "amino acid tool box" helps take the guess work out of the design process and encourages the rationalization of the *in vitro* activity correlating to the structural features.

3.2 FORCE FIELD PARAMETER DEVELOPMENT

The coordinates for the amino acids which did not have parameters included in the AMBER force field 03 were created using Materials Studio.[2] These included (a) β alanine (βAla), (b) GABA, (c) 6-Amino hexanoic acid (6-Ahx), (d)tetrahydroisoquinoline carboxylic Acid (Tic), with the carbonyl carbon in the axial position (e) octahydroindole carboxylic Acid (Oic), (f) ornithine (Orn), (g) diaminobutanoic acid (Dab), (h) diaminopropanoic acid (Dap), (i) aminocyclohexanoic acid (Ach), (j) aminocyclopentanoic acid (Acp), (k) p-Cl phenylalanine (pClPhe), and (l) p-F phenylalanine (pFPhe). Tripeptides, Ace-Ala-X-Ala-Nme, where Ace is the acetyl group used to neutralize the N-terminus and Nme is the N-methyl group used to cap and neutralize the C-terminus. These tripeptides were minimized using 5000 steps of steepest descent using Cerius 2[3] prior to *ab initio* calculations.

Antechamber was used to generate input files for GAUSSIAN03[4] which was utilized to calculate the minimum energy conformation of each amino acid structure. Hartree-Fock (HF) calculations using the 6-31G* basis set were used to optimize the structures and generate the electrostatic potential derived charges for numerous points around each molecule according to the Merz-Singh-Kollman scheme which fits the ESP to the point partial charges of the atoms within the molecule.[5]

This level of theory is consistent with that utilized for development of parameters of the standard amino acids which have been validated using experimental data.[6]

Force field parameters were developed for the following unnatural amino acids: (a) β-Ala, (b) Gaba, (c) 6-Ahx, (d) Tic, (e) Oic, (f) Orn, (g) Dab, (h) Dap, (i) Ach, (j) Acp, (k) pClPhe, and (l) pFPhe. Gaussian 03[4] was utilized to predict energies, molecular structures, and vibrational frequencies of the unnatural amino acid molecule systems. The resulting values were the basis for calculation of minimum energy structures upon which the force field parameters were derived.

Tripeptides were constructed from a library of residues, using tleap, which has been expanded to include the additional amino acids for which custom force field parameters were developed. Tripeptides were generated using the sequence command within the tleap program. Peptides were placed at the center of a solvent box which used the TIP3P model for water extending 10 angstroms from the peptide surface enabling the use of cubic periodic boundary conditions. The ionic concentration was adjusted using random replacement of water molecules, with sodium and chlorine ions, using the multi-scale modeling tools for structural biology (MMTSB) toolset.[7] The ionic strength of 150 mM NaCl was selected for its physiological relevance.[8] The peptides were minimized using a two-step process, first constraining the peptide atoms using a force constant of 15.0 kcal/(mol*angstrom) allowing the water to settle around the tripeptide followed by a second unrestrained minimization. Each minimization used 5000 steps of steepest descent with constant volume, no pressure scaling, and the Andersen temperature coupling.[9] Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method with a non-bonded cutoff truncated at 10.0 angstroms. Energy values were recorded every 25 steps. A second minimization was performed using 2500 steps of steepest descent followed by 500 steps of conjugant gradient which converged to a RMS of 7.3793×10^{-1}. Following minimization, the system was warmed to 310 K over a time period of 2.1 ns. 310 K

was the selected temperature due its physiological relevance. The non-bonded cutoff was set to 10 angstroms. A 1 fs time step was used while constraining hydrogen atom bond lengths using the SHAKE algorithm.[10-12] Anderson temperature scaling[9] was used to adjust the temperature of the system updating the temperature every 1 ps.

Peptides were subjected to constant pressure dynamics until density stabilization had been attained. The production run using constant volume dynamics continued until a 5 ns trajectory had been acquired with structure stabilization based upon a less than 2 angstrom deviation from the starting structure. Long-range electrostatic interactions were calculated using the PME method with a non-bonded cutoff set to 10 angstroms. A 1.5 fs time step was used while constraining the hydrogen atom bond lengths using the SHAKE algorithm.[10-12]

Simulation structures were recorded every 1.0 ps. The parameters selected for the simulation were designed to model physiological conditions as much as possible. All simulations were conducted at 310 K. Anderson temperature scaling was used to maintain the temperature of the system throughout the course of the production run.

3.3 CALCULATION OF THE ELECTROSTATIC SURFACE POTENTIAL

Electrostatic surface potential grid generation and analyses were performed with the Adaptive Poisson-Boltzman Solver (ABPS).[13] ABPS was written by Nathan Baker with J. Andrew McCammon and Michael Holst; and is supported by the National Biomedical Computation Resource, National Institutes of Health, National Partnership for Advanced Computational Infrastructure, and Washington University in St. Louis.[13] ABPS takes the charges generated either by a software such as PDB2PQR which are derived from a database of

values, primarily tailored to the standard amino acids or the .pqr file, which includes the atomic charge values can be derived from other means. The .pqr file, which included the atomic charges, was created from the minimized gas phase structure of the tripeptides prior to being imported into the ABPS[13] using a software plugin within PyMol.[14] The charge values from the .pqr file are distributed across the molecule based upon a two part grid shape with numerous customization options. Both the course and fine mesh length was -1, -1, -1. The grid center was at 0, 0, 0 with grid points every -1, -1, -1 units on the grid. The Linearized Poisson-Boltzmann Equation was used for the calculations. The dielectric constant for the protein was set at 2.0 while the surrounding solvent dielectric was set at 78.0. A dielectric constant value of 78.0 was selected to model the polarizability of a water solvated environment. The solvent radius was set to 1.4 angstroms. The temperature of the system was 310.0 K with a vacuum sphere density of 10 angstrom cube, which is the number of grid points per angstrom cubed square. The ionic concentration was set at 0.15 mM with a cation (+1) radius of 2.0 angstrom and an anion (-1) radius of 1.8 angstrom. No ions with +/- 2 charges were included in the model. The surface calculation method used harmonic average smoothing.

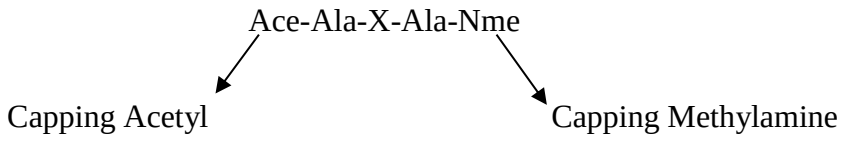
Electrostatic surface potential maps were created using the Solvent Accessible Surface Area (SASA) which gives an idea regarding a molecule's surface accessible to the surrounding solvent. The calculation is usually derived using a rolling ball technique in which a sphere with a specified radius is able to determine the surface area in square angstroms ($Å^2$). This can help predict the potential for the molecule to interact with the surrounding environment whether it is solvent or a bilayer. The SASA is also useful when performing implicit solvent simulations which are used to improve the prediction of protein secondary structure. Considering the volume of specific amino acids and the distribution of charge for each individual residue can help in the

rational design of therapeutic peptides. Tailoring the charge distribution of peptides can optimize for specific membrane charges (especially when targeting multiple drug resistant organisms).

Calculation of the SASA and volume calculations were performed with the UCSF Chimera package.[15] Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).[15]

3.4 BACKBONE CONFORMATION

The favored backbone conformations of standard amino acids have been well established.[1] However less is known about the nonstandard amino acids. To further investigate the preferred angles adopted by the nonstandard amino acids, capped tripeptides with nonstandard amino acids have been designed and are as shown below and utilized for modeling studies:



In the above generalized sequence representation the X represents either the standard or nonstandard amino acid. Standard amino acids containing tripeptides with a similar generalized sequence were also constructed and simulated as validation of the investigative approach. Tripeptides were minimized using 5000 steps of steepest descent with constant volume, no pressure scaling, and Andersen temperature coupling. Energy values were printed every 25

steps. A second minimization was performed using 2500 steps of steepest descent followed by 500 steps of conjugant gradient.

Intramolecular hydrogen bonds have been observed to form experimentally between the backbone and side chain heteroatoms.[16] Over the short time our simulations were conducted, formation of intramolecular hydrogen bonds were observed to form intermittently during the course of the solvated MD simulations. Prolonged warming steps were also associated with an increase in intramolecular hydrogen bond formation with the atoms of the peptide backbone.

3.5 LOG P

The Log P was calculated for both the standard and unnatural amino acids to give an idea of the solubility of the peptides when administered as a therapeutic compound and the potential for the residue to partition into the hydrophobic environment of the hydrophobic, aliphatic tail region of the lipid bilayers. The partition coefficient (P) is a measure of the concentration of a given compound which is dissolved in two immiscible phases at equilibrium. Log P is for neutral species only. The measurement is accomplished through the addition of buffers to the aqueous phase to maintain a particular pH at which the compound would be non-charged. Log D, which is able to take charged species into account through a pH dependent measure may be a more appropriate selection however the software for such measurements is not presently available.

Octanol and water are commonly used for such calculations and offer a measure of solubility in hydrophobic and highly polar solvents. Using the logarithm of the ratio of the concentrations of the compound dissolved in each of the solvents produces the log P value (equation 3.1).

$$\log P_{\frac{octanol}{water}} = \log\left(\frac{[solute]_{octanol}}{[solute]_{non-ionized\ water}}\right)$$ Eqn 3.1

The values listed in **Table 3.1** were derived from the Advanced Chemistry Development program in which Log P calculations calculate the octanol-water partition coefficient for neutral compounds under standard conditions at 25°C.[17] Calculations are provided with 95% confidence intervals or reliability index (RI) based upon a training set of >25 000 compounds in the training set and >18 000 log P values in the reference database.[17] While the training set is not specific for our compounds the fragments from which they are composed enable reasonable estimates for log P.[18] The Log P value is associated with the solvent type in which a molecule is expected to be soluble. Octanol and water, which are hydrophobic and highly polar solvents respectively, are commonly utilized for these measurements.

National Biomedical Computational Resource (NBCR) software Adaptive Poisson-Boltzmann Solver (APBS)[13] plugin within the PyMOL[14] software was used for the calculation of charges. The input files were derived from the previously calculated charges for each nonstandard amino acid (See Chapter 2 and **Appendix A**).

Table 3.1 Peptide sequences and the structural properties measured for each amino acid.

Peptide Sequence	Solvent Accessible Surface Area Å^3	Volume Å^3	Distance from Central Carbonyl Carbon to Ace Carbonyl Carbon	Distance from Central Carbonyl Carbon to Nme Carbonyl Carbon	Log P
Ala-**Lys**-Ala	324.1	302.4	3.7	3.3	-2.09 ±0.62
Ala-**Dab**-Ala	302.0	275.6	3.7	3.7	-2.09 ±0.64
Ala-**Orn**-Ala	308.6	291.0	3.3	3.8	-2.02 ±0.63
Ala-**Dpr**-Ala	281.9	262.2	3.8	3.7	-1.99 ±0.66
Ala-**Arg**-Ala	345.9	322.9	3.7	3.5	-2.84 ±0.68
Ala-**His**-Ala	307.7	292.1	3.7	3.7	-2.27 ±0.69
Ala-**Hip**-Ala	308.0	292.6	3.7	3.2	-2.27 ±0.69
Ala-**Tyr**-Ala	323.0	312.5	3.5	3.7	-0.63 ±0.67
Ala-**Leu**-Ala	304.7	291.0	3.6	3.8	-0.32 ±0.62
Ala-**Gly**-Ala	260.5	229.9	3.5	3.8	-2.08 ±0.61
Ala-**βAla**-Ala	275.2	244.0	4.6	3.7	-2.30 ±0.57
Ala-**GABA**-Ala	292.2	258.8	5.8	3.7	-2.08 ±0.57
Ala-**6Ahx**-Ala	308.4	273.4	7.1	3.7	-1.55 ±0.56
Ala-**Acp**-Ala	296.0	286.2	3.9	3.8	-1.16 ±0.61
Ala-**Ach**-Ala	310.6	305.5	3.8	3.3	-0.59 ±0.62
Ala-**Tic**-Ala	338.2	329.3	3.0	3.7	0.38 ±0.72
Ala-**Oic**-Ala	316.3	322.5	3.3	3.7	0.74 ±0.70
Ala-**Phe**-Ala	352.5	323.2	3.7	3.7	0.11 ±0.67
Ala-**Pro**-Ala	283.5	267.2	3.2	3.7	-1.37 ±0.70
Ala-**pClPhe**-Ala	356.3	248.9	3.3	3.8	0.70 ±0.67
Ala-**pFPhe**-Ala	336.3	326.4	3.2	3.7	0.16 ±0.71
Ala-**Trp**-Ala	324.4	324.9	3.7	3.7	0.03 ±0.67
Ala-**Gln**-Ala	312.7	292.3	3.7	3.6	-2.66 ±0.65
Ala-**Glu**-Ala	309.3	287.2	3.6	3.7	-2.26 ±0.78
Ala-**Ala**-Ala	271.6	245.3	3.7	3.7	-1.73 ±0.62
Ala-**Val**-Ala	277.7	296.5	3.1	3.6	-0.85 ±0.62
Ala-**Ile**-Ala	312.2	311.0	3.3	3.3	-0.32 ±0.62
Ala-**Ser**-Ala	284.9	264.7	3.4	3.6	-2.31±0.64
Ala-**Thr**-Ala	288.5	291.8	3.2	3.1	-1.96 ±0.64
Ala-**Cys**-Ala	295.6	276.4	3.3	3.1	-0.84 ±0.68
Ala-**Met**-Ala	304.0	315.3	3.3	3.3	-1.10 ±0.72
Ala-**Asn**-Ala	282.1	295.1	3.3	3.0	-2.57 ±0.71
Ala-**Asp**-Ala	281.0	286.0	3.0	3.4	-1.61 ±0.66

3.6 PHI/PSI ANGLE CALCULATIONS

The angles were calculated using the Cartesian coordinates (X,Y,Z) of each atom composing the relevant angle and using the vector between each atomic point to create vectors from which the angles were calculated. The torsion angle is the angle can be calculated by calculating the cross product of the vectors composing the angle. The sign of the torsion angle is determined by the relative orientation of the dot product of the same two vectors with reference to the middle vector. A python script was prepared for the calculation of the torsion angles. The scripts have been included in **Appendix C**. Protein databank (.PDB) files were created every step, 1 ps , for the 5 ns of the simulations. The first 2 ns was discarded as an equilibrating step and the coordinates for the central residue of the tripeptide were extracted. The second python script calculates the torsion angles first the phi angle followed by the psi angle and the two omega angles on either side of the central residue. The angles which were calculated were imported into Microsoft Excel 2010 enabling creation of the Ramachandran plots.

3.7 RAMACHANDRAN PLOTS

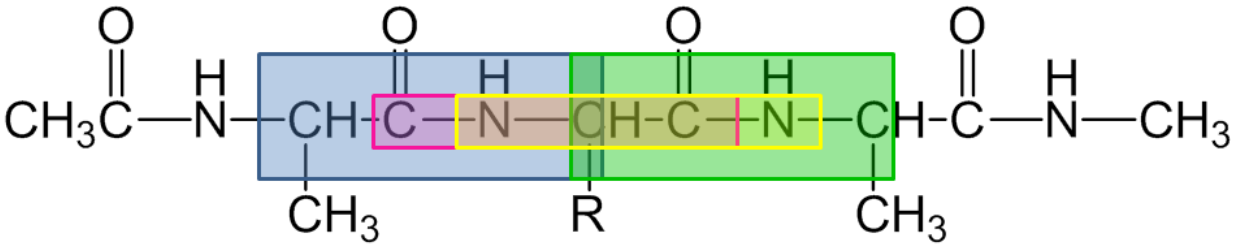


Figure 3.1 Atoms involved in the Omega 1 angle, outlined in blue. The Omega 2 angle is outlined in green. Phi angle, outlined in magenta and psi angle, outlined in yellow.

Ramachandran plots note the distribution of the phi and psi angles of an amino acid. Specifically the phi angle looks at the plane formed by the vectors resulting from the coordinates of the carbonyl carbon of the preceding residue, amino nitrogen, alpha carbon, and the carbonyl

carbon of the residue (Figure 3.1). The psi angle is formed from the plane formed by the vectors resulting from the coordinates of the amino nitrogen, the alpha carbon, the carbonyl carbon and the alpha carbon, the carbonyl carbon and the amino nitrogen of the following residue, The three points are used to define the plane between the atoms.

3.8 SPACER # 1 LINKERS

The outline of the AMP (Figure 1.15) includes four variable amino acid linkers which increase the number of carbon atoms in the backbone of the peptide. The inclusion of glycine (1 carbon), β-alanine (2 carbons), Gaba (4 carbons), and 6-Aminohexanoic acid (5 carbons) increases the flexibility of the peptide. The electrostatic properties of each peptide will be discussed in addition to the Ramachandran plots derived from averaging the final 2 ns of the 5 ns MD simulation.

3.8.1 GLYCINE

Glycine (Gly), is unique amongst the 20 standard amino acids because it lacks a chiral alpha carbon. This small nonpolar amino acid has relatively little charge due to the neutralization of the backbone associated charge (Figure 3.2 A). Gly is more flexible than the other amino acids due to the lack of steric hindrance because there are no side chains, only hydrogen atoms. The flexible nature of Gly is supported by the wide angle distribution observed in the Ramachandran plot (Figure 3.2 B). The normally disallowed regions of the Ramachandran plot are inhabited by the Gly residue, which has also been observed by others studying this small molecule.[1]

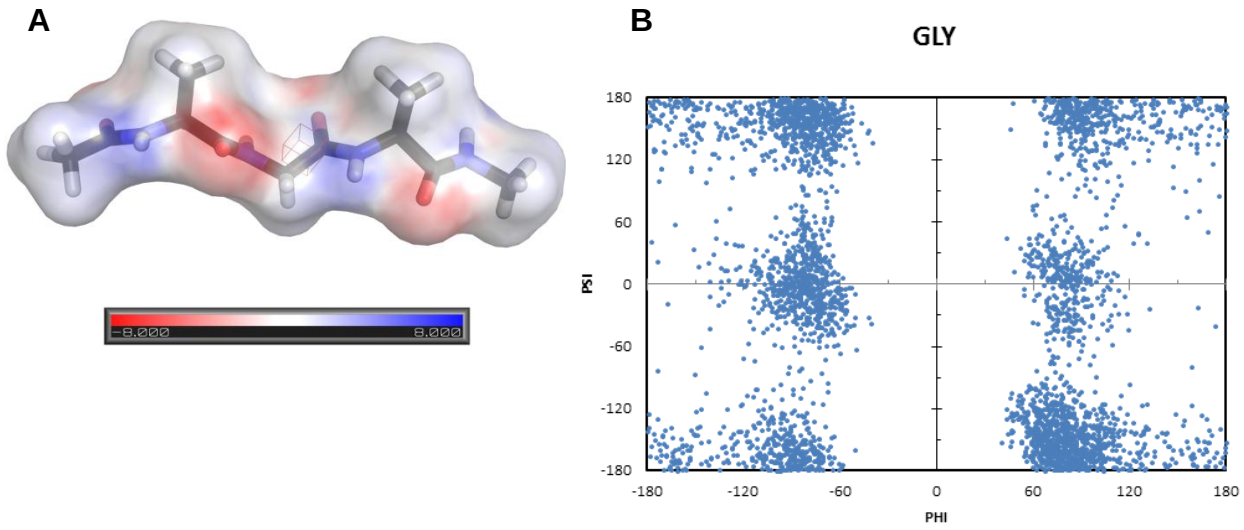


Figure 3.2 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala- Gly-Ala-Nme B. Ramachandran Plot of the phi/psi angle distribution for the central glycine residue of the tripeptide. The electro negative residues are shown with a darker red color and electro positive regions are shown with a blue surface.

3.8.2 β-Alanine

Beta-alanine, β-Ala, is a nonstandard, naturally occurring amino acid derived from the degradation of dihydrouracil and carnosine and is metabolized to acetic acid. This small nonpolar amino acid has relatively little charge due to the aliphatic nature of the alpha and beta carbon atoms incorporated into the backbone (Figure 3.3). The distance between the carbonyl oxygen atoms have increased from 3.5 angstrom observed in the Gly residue to 4.6 angstrom reflecting the inclusion of the beta carbon in the backbone. Addition of the beta carbon increases the backbone flexibility more so than what was observed in Gly. Due to the inclusion of the extra carbon in the backbone phi and psi angle calculations are not possible in the traditional sense and have been omitted.

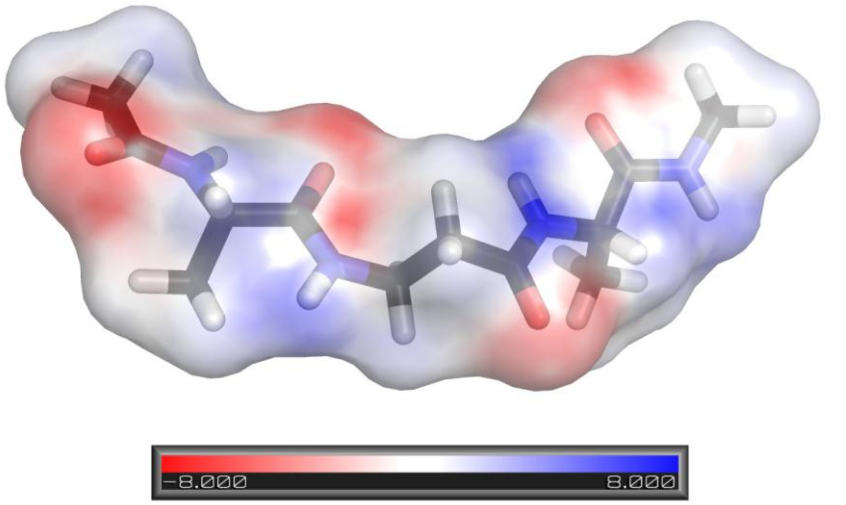


Figure 3.3 Electrostatic surface potential distribution in the tripeptide Ace-Ala-βAla-Ala-Nme

3.8.3 GAMMA AMINO BUTYRIC ACID

Gama aminobutyric acid (Gaba), is a nonstandard, naturally occurring amino acid which plays an important role as a neurotransmitter in the mammalian central nervous system. The charge distribution along the backbone is separated even farther than in the glycine or β-alanine. The white gap depicted in Figure 3.4 shows the spacing effect of the inclusion of four carbon atoms in the backbone. The distance between the carbonyl oxygen atoms has increased from 3.5 angstrom observed in the Gly, and 4.6 angstrom in β-Ala, to 5.8 angstrom in the Gaba residue. Addition of the beta, gamma, and delta carbon atoms increases the backbone flexibility more so than what was observed in either Gly or β-Ala. Due to the inclusion of the extra carbon atoms in the backbone phi and psi angle calculations are not possible in the traditional sense and have been omitted.

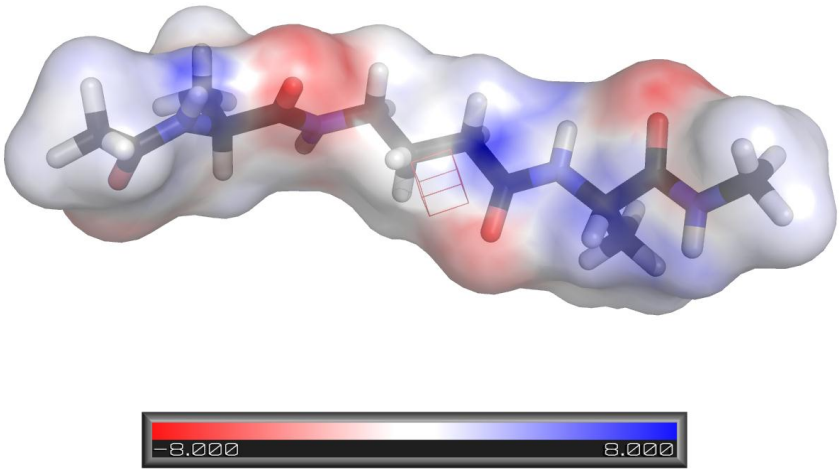


Figure 3.4 Electrostatic surface potential distribution for the tripeptide Ace-Ala-Gaba-Ala-Nme

3.8.4 6-AMINOHEXANOIC ACID

6-aminohexanoic acid (6-Ahx), is a nonstandard amino acid which can be synthesized from lysine. 6-Ahx has been used as a therapy for the treatment of bleeding disorders due to its ability to inhibit the proteolytic activity of plasmin. 6-Ahx is also a synthetic intermediate in the synthesis of nylon-6. The charge distribution along the backbone is separated even farther than in the glycine, β-alanine, or Gaba. The white gap depicted in Figure 3.5 shows the spacing effect of the inclusion of five carbon atoms in the backbone. The distance between the carbonyl oxygen atoms has increased from 3.5 angstrom observed in the Gly, to 4.6 angstrom in β-Ala, 5.8 angstrom in Gaba to 7.1 angstrom in the 6-Ahx residue. Addition of the beta, gamma, delta, and episilon carbon atoms increases the backbone flexibility more so than what was observed in either Gly, β-Ala, or Gaba. Due to the inclusion of the extra carbon atoms in the backbone phi and psi angle calculations are not possible in the traditional sense and have been omitted.

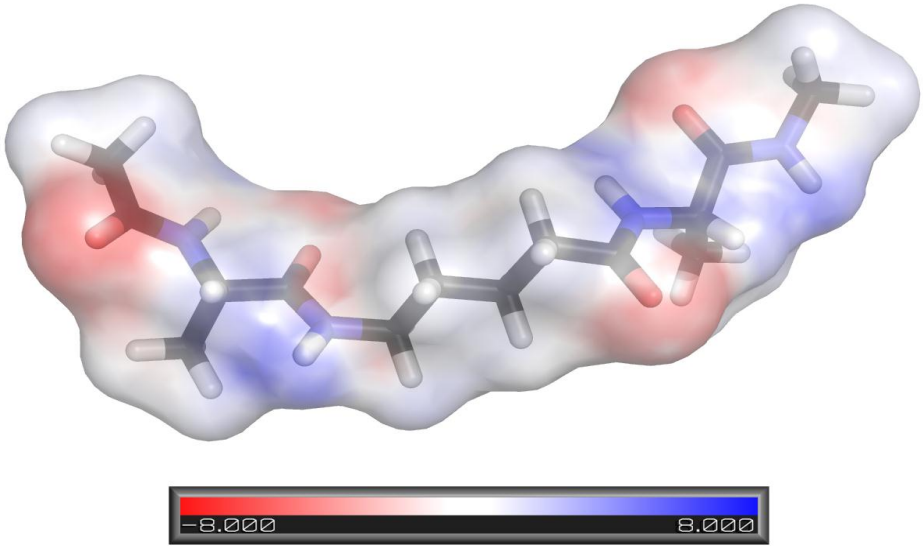


Figure 3.5 Electrostatic surface potential distribution for the tripeptide Ace-Ala-6-Ahx-Ala-Nme

3.9 SPACER # 2 LINKERS

The outline of the AMP (Figure 1.15) includes five variable amino acid linkers which increase the number of carbon atoms between the alpha carbon and the cationic charge of the side chain. The inclusion of diaminopropionic acid (1 carbon), diaminobutanoic acid (2 carbons), ornithine(3 carbons), arginine (3 carbons) and Lysine (4 carbons) into the peptide sequence changes the electronic charge distribution relative to the peptide backbone enabling the tailoring of the charge distribution relative to the desired target membrane charge type. The electrostatic properties of each tripeptide will be discussed in addition to the Ramachandran plots derived from averaging the final 2 ns of the 5 ns MD simulation.

3.9.1 DIAMINOPROPIONIC ACID

2,3-Diaminopropionic acid, Dpr, is an unnatural amino acid which is a synthetic amino acid. The charge distribution relative to the backbone was measured at the conclusion of the MD simulation. Dpr was found to have a distance of 2.6 Å between the alpha carbon and the

nitrogen atom of the side chain. The presence of the cationic charge close to the backbone creates a charge distribution along the backbone as evidenced by the blue shaded region in Figure 3.6 A. The figure shows the effect of the shorter carbon chain between the backbone and the side chain charge.

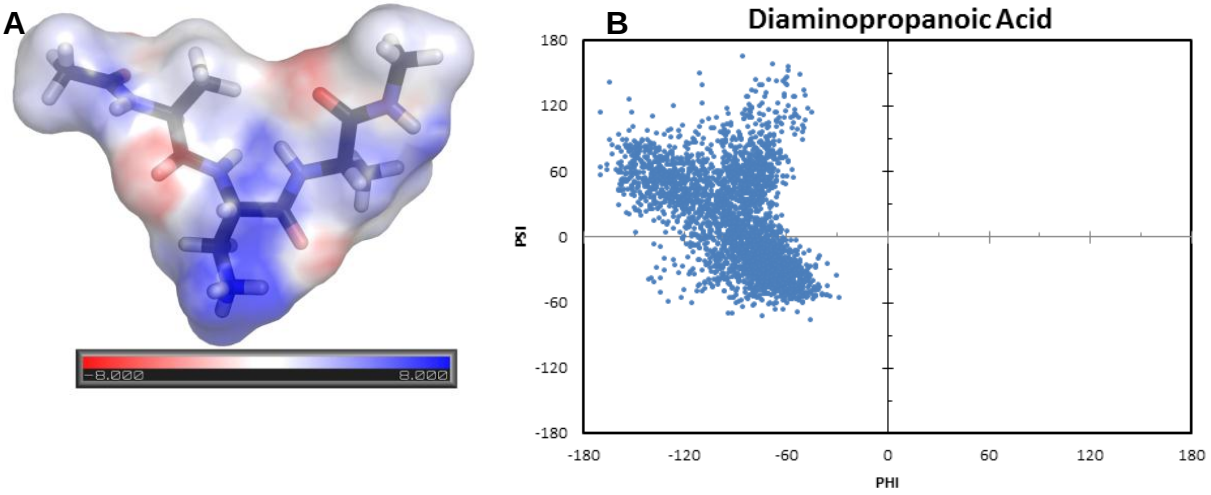


Figure 3.6 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Dpr-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Dap

Dpr is less flexible than residues which are able to locate the charge farther away from the charged atoms in the peptide backbone. The lack of flexibility results from the steric hindrance of the cationic protonated amine. The lack of flexibility of the Dpr residue is supported by the limited range of the phi and psi angle distribution observed in the Ramachandran plot (Figure 3.6 B). Another attribute to the observed distribution is the polar interactions between the cation and the carbonyl oxygen. The regions of the Ramachandran plot preferentially populated by Dpr suggest a right handed alpha helical shape or a β-sheet/turn secondary structure may be associated with the Dpr residue.

3.9.2 *Diaminobutanoic Acid*

2,3-Diaminobutanoic acid (Dab), is an unnatural amino acid which is a synthetic amino acid. The charge distribution relative to the backbone was measured at the conclusion of the MD simulation. Dab was found to have a distance of 3.4 angstrom between the alpha carbon and the nitrogen atom of the side chain. The presence of the cationic charge so close to the backbone creates a lack of electron density toward the backbone as evidenced by the blue shaded region in Figure 3.7 A which shows the effect of the shorter carbon chain between the backbone and the side chain charge.

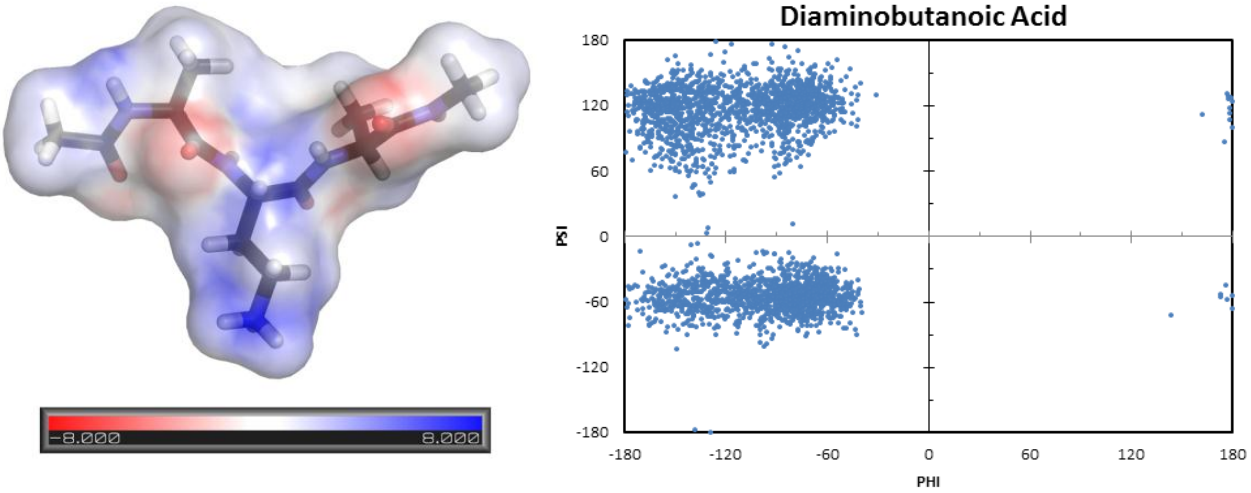


Figure 3.7 A. Electrostatic surface potential distribution of the tripeptide Ace-Ala-Dab-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Dab

Dab is less flexible than residues which are able to locate the charge farther away from the charged atoms in the peptide backbone. The charged atoms of the side chain have the charge distributed toward the backbone creating a slightly more diffuse charge region relative to the Dpr residue however; the charge is still more closely positioned relative to the other residues including Orn, Arg, and Lys. The lack of flexibility results from the steric hindrance of the cationic protonated amine. The lack of flexibility of the Dab residue is supported by the limited

range of the phi and psi angle distribution observed in the Ramachandran plot (Figure 3.7 B). The regions of the Ramachandran plot preferentially populated suggest a right handed alpha helical shape or a β-sheet/turn secondary structure may be associated with the Dab residue. The alpha helical preferred region is generally found in the lower left quadrant of the Ramachandran plot.

3.9.3 ARGININE

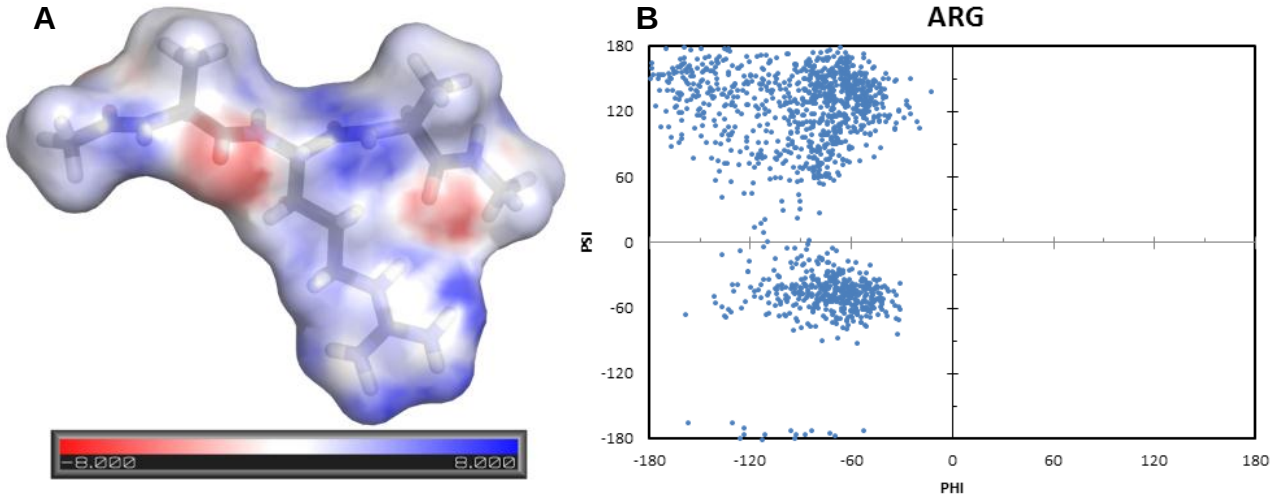


Figure 3.8 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Arg-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Arg

Arginine (Arg), is one of the 20 standard amino acids. The charge distribution relative to the backbone was measured at the conclusion of the MD simulation. Arg was found to have a distance of 4.8 angstrom between the alpha carbon and the carbon of the guanidinium group of the side chain. The presence of the cationic charge at such a distance from the backbone allows dispersion of the cationic charge to the backbone as evidenced by the blue shaded region in Figure 3.8 A which shows the dispersive effect of the guanidinium group, with its delocalized cation. Distribution of the charge across three carbon atoms back toward the backbone is reflective of the significant impact the cation has on the electron density of the Arg residue.

The three nitrogen atoms of the guanidinium group enable delocalization of the charge over a larger area which, based on the electrostatic surface potential map, shows the cationic charge extends toward the peptide backbone and creates a larger area of cationic charge. Arg is able to locate the charge of the guanidinium group farther away from the charged atoms in the peptide backbone than the Dpr and Dab residues. The length of the carbon linker between the alpha carbon and the side chain charge is sufficient to allow the side chain to fold back upon itself allowing the potential for intramolecular hydrogen bonding to occur. The flexibility of the Arg residue is supported by the broad range of the phi and psi angle distribution observed in the Ramachandran plot (Figure 3.8 B). The region of the Ramachandran plot preferentially populated by the Arg tripeptide suggest a right handed alpha helical shape or a β-sheet/turn secondary structure may be associated with the Arg residue. The regions populated on the Ramachandran plot are consistent with other groups findings.[1]

3.9.4 ORNATHINE

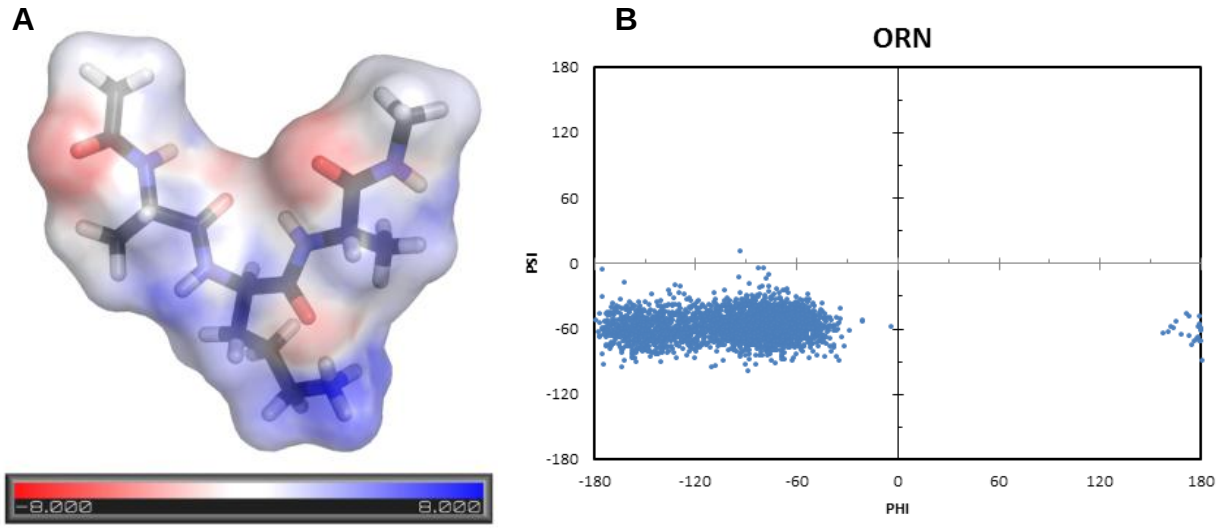


Figure 3.9 A. Electrostatic surface potential distribution of the tripeptide Ace-Ala-Orb-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Orn

Despite the similarity in carbon length between the alpha carbon and the charge centered atom between Arg and ornithine, Orn, the preferred secondary structure appears to be quite different. The distance between the alpha carbon and the protonated mamine of the Orn residue is 5.0 angstrom, similar to the 4.8 angstrom measurement associated with Arg. Both these residues have three carbon atoms linking the polar atom to the backbone. The charge distribution relative to the backbone was measured at the conclusion of the MD simulation. Much like the Arg residue the presence of the cationic charge three carbon atoms away from the backbone allows dispersion of the cationic charge to the backbone through intramolecular hydrogen bonding which is shown in Figure 3.9 A. The greater distance between the charge and the backbone increases the flexibility of the side chain, but not necessarily the backbone of the residue. The flexibility of the Orn residue side chain is not measured in the torsion angle calculations. The lack of flexibility in the backbone of this residue is supported by the narrow range of the phi and psi angle distribution observed in the Ramachandran plot (Figure 3.9 B). The region of the Ramachandran plot preferentially populated by the Orn tripeptide suggest a right handed alpha helical secondary structure may be associated with the Orn residue. Structural observations suggest hydrogen bonding could be a contributing factor.

3.9.5 LYSINE

Lysine, Lys, is a naturally occurring, standard amino acid. Here it is used as a reference reflecting the structural changes associated with the variation of the number of side chain carbon atoms separating the backbone from the cationic charge of the side chain. The distance between the alpha carbon and the protonated amine of the Lys residue is 6.4 Å. The charge distribution relative to the backbone was measured at the conclusion of the MD simulation. The shorter distances between the charge and the backbone allowed for the dispersion of the charge over a

greater area however lysine, with the large distance between the backbone and the hard charge of the protonated amine creates a separation which is not able to be distributed over a large area, see Figure 3.10 A. The presence of the cationic charge more than three carbon atoms away from the backbone allows flexibility in the side chain making intramolecular hydrogen bonding possible however it was not observed during the short simulation time scale studied. The preference of the Lys residue to adopt a right handed alpha helical secondary structure has been determined through crystallographic studies in addition to structure prediction using MD simulation.[1] The narrow range of the phi and psi angle distribution observed in the Ramachandran plot (Figure 3.10 B) is consistent with previous MD simulations and crystallographic data.[1] The large distance between the backbone and the protonated amine of the Lys residue make the formation of intramolecular hydrogen bonds possible and participation in cyclic five or six membered hydrogen bonds are expected.[19]

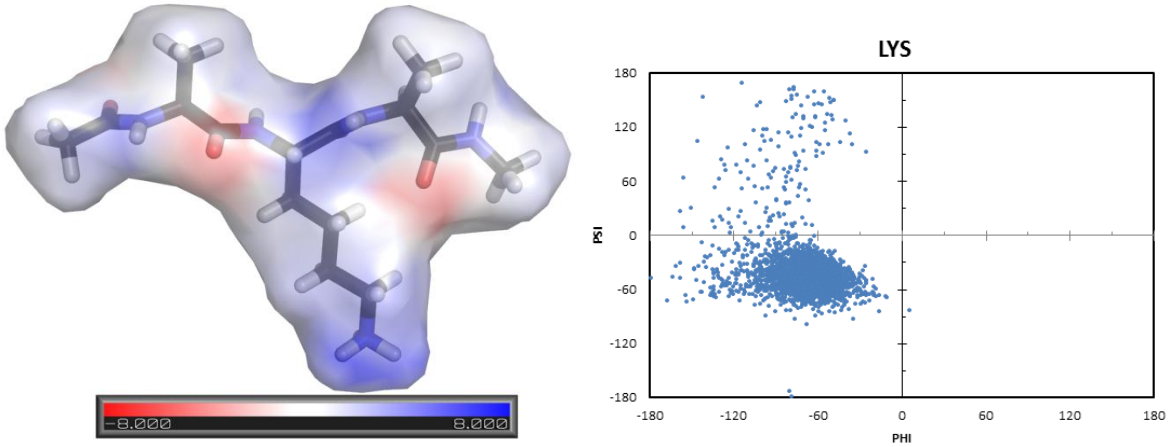


Figure 3.10 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Tic-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Tic

3.10 BICYCLIC UNNATURAL AMINO ACIDS

The outline of the AMP (Figure 1.15) includes three bicyclic amino acids placed in sequence together, Tic and Oic. A terminal single Tic residue is found just before the terminal four lysine residues. Inclusion of these sterically hindered residues in sequence is designed to force the secondary structure toward that of a β-sheet or turn. The structural properties of these bicyclic, unnatural amino acids have been studied using tripeptide MD simulations to examine the preferred angular arrangement of the backbone atoms. Similarities to naturally occurring amino acids phenylalanine and proline will be discussed as well. The electrostatic properties of each tripeptide will be discussed in addition to the Ramachandran plots derived from averaging the final 2 ns of the 5 ns MD simulation.

3.10.1 TIC

Tetrahydroisoquinolinecarboxylic acid (Tic) is a bicyclic, unnatural amino acid which has the side chain cyclizing via the amide nitrogen in an orientation similar to proline. The bicyclic nature of this residue creates a sterically bulky, hydrophobic side chain. This large nonpolar amino acid has relatively little charge due to aromatic ring fused to a nonaromatic ring in closest contact with the backbone. The lack of charge disassociation of the backbone atoms to those of the bicyclic rings is reflected in Figure 3.11 A. The side chain looping to the backbone severely limits the possible orientations possible. The lack of reflected in the observed in the Ramachandran plot (Figure 3.11 B). The preference, based on angle population appears to fall in both the right handed alpha helical region and the β-sheet/turn regions of the Ramachandran plot.

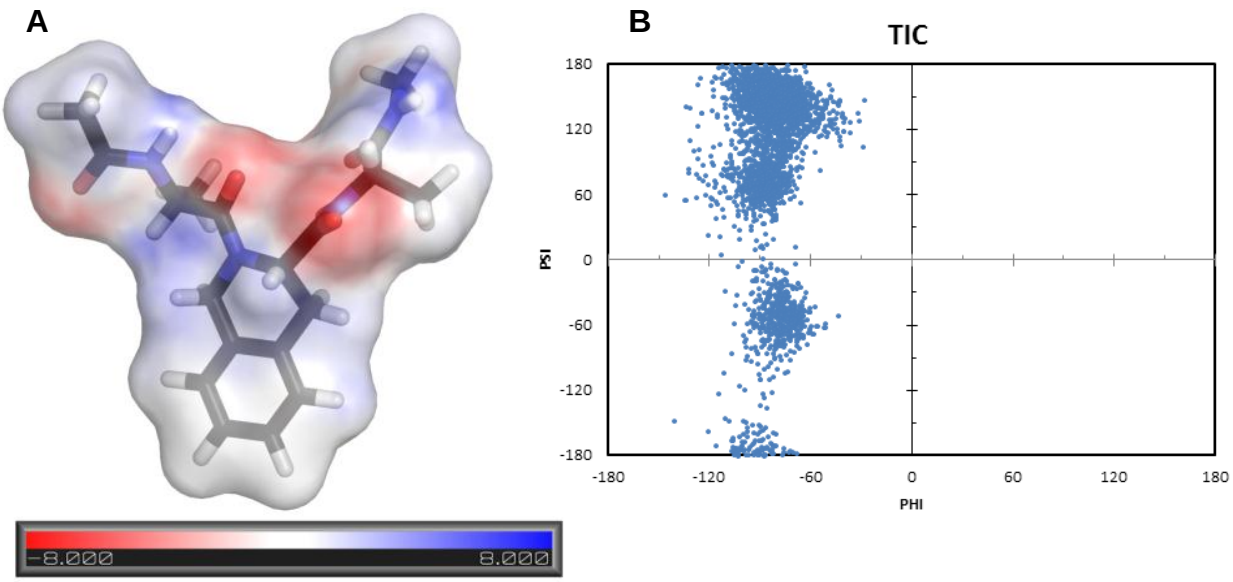


Figure 3.11 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Tic-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Tic

3.10.2 PHENYLALANINE

Phenylalanine (Phe) was the model for the design of unnatural amino acid Tic. The bulk of the hydrophobic aromatic ring was optimized while the addition of a carbon atom allowed the creation of the six membered ring hindering the peptide backbone. The nonpolar Phe is reflected in Figure 3.12 A. The charge is localized on the backbone atoms and the lack of charge density around the aromatic ring results in a bulky hydrophobic residue. The preferred torsion angles adopted by the Phe residue suggest a preference for a right handed alpha helical secondary structure, Figure 3.12 B. Comparison of the Phe residue to the unnatural Tic shows how the preferred secondary structure has been influenced. Despite the increased steric hindrance of Tic, the potential secondary structure expands beyond that observed for the monocyclic Phe. The MD simulation derived phi and psi angles are consistent with experimentally derived preferences for the Phe residue. [1]

3.10.3 OIC

Octahydroindolecarboxylic acid (Oic), is also a bicyclic, unnatural amino acid which has the side chain looping back to the amide nitrogen in an orientation similar to proline. The lack of aromaticity enables more flexibility in the side chain of the residue. The large nonpolar amino acid has no charge due to the aliphatic rings with the pyrrolidine ring providing the amido

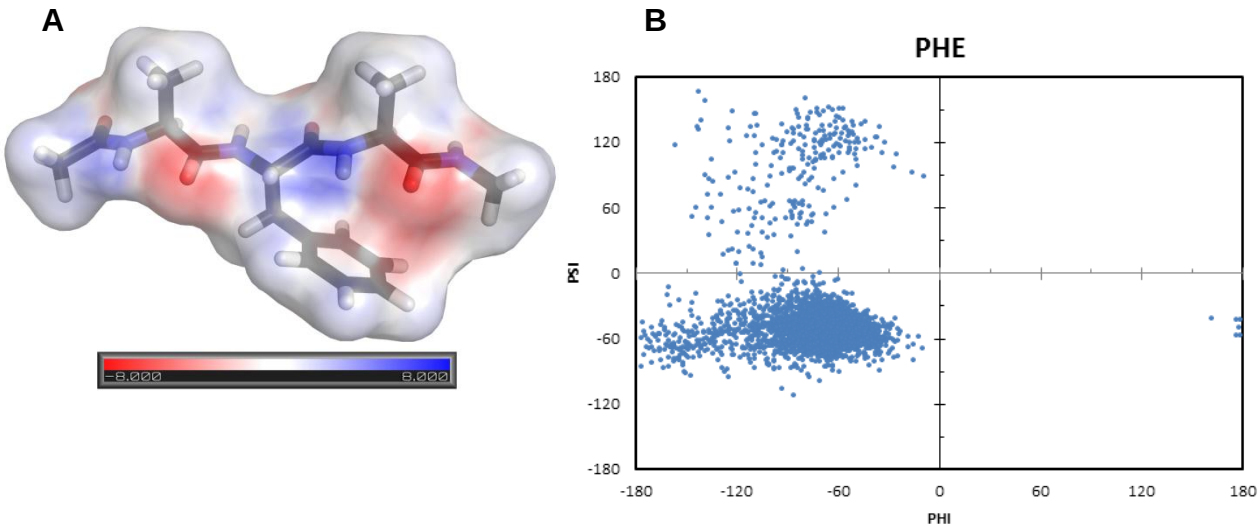


Figure 3.12 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Phe-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Phe

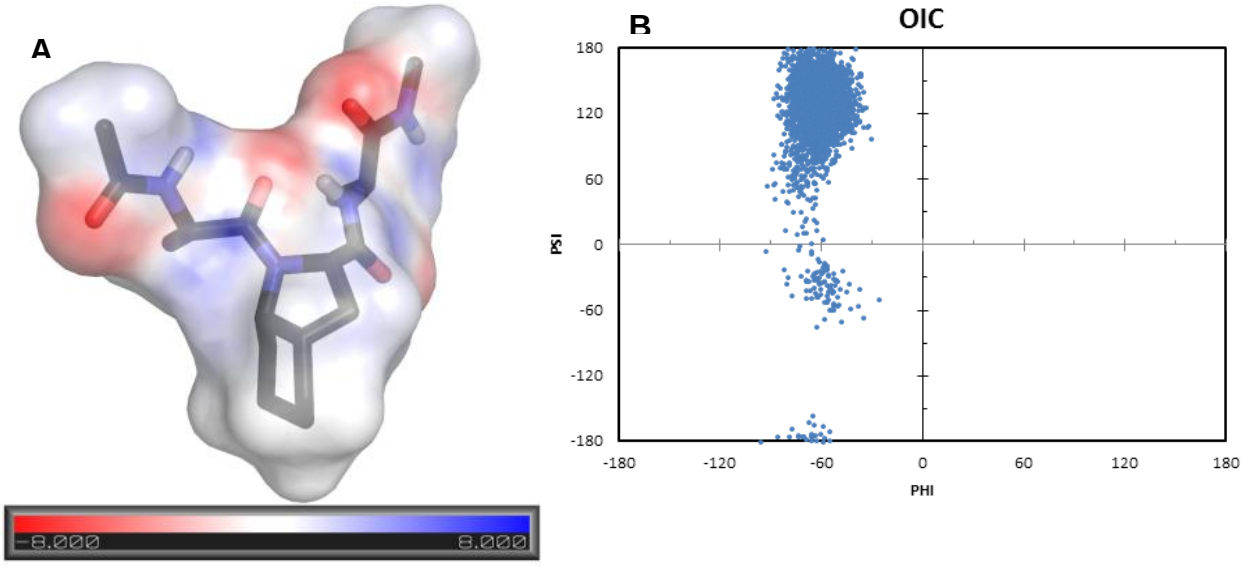


Figure 3.13 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Oic-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Oic

nitrogen and the alpha carbon of the backbone. The lack of charge disassociation of the backbone charge to the atoms of the bicyclic rings is reflected in Figure 3.13 A. The side chain cyclizing to the backbone severely limits the possible orientations possible. The lack of flexibility is reflected in the observed Ramachandran plot (Figure 3.13 B). The preference, based on angle population appears to fall the β-sheet/turn regions of the Ramachandran plot. The three chiral centers of the Oic residue warrant further investigation to elucidate the preferred structural combination for this residue.

3.10.4 PROLINE

Proline (Pro) is a naturally occurring amino acid which is commonly associated with β-sheets and turns within peptides and proteins. The nature of the pyrrolidine ring of Oic residue is structurally derived from Pro. The addition of the unsaturated six membered ring to the pyrrolidine ring of Oic residue is responsible for increasing the hydrophobic bulk of the residue relative to the smaller Pro residue. Despite the hydrophobic nature of the Pro residue, the close proximity of the side chain carbon atoms limit the hydrophobic nature of this residue, which was exaggerated in the Oic residue, see Figure 3.14 A. The effect observed of the addition of the unsaturated six membered ring to the pyrrolidine ring seems to be limiting the psi angle preference. Pro has two regions of preferred angular orientation, Figure 3.14 B relative to the single preferred region observed in Oic, Figure 3.13 B. The conclusion is that the addition of the unsaturated six membered ring to the pyrrolidine ring of Pro decreases the backbone flexibility forcing the conformation into a β-sheet/turn secondary structure.

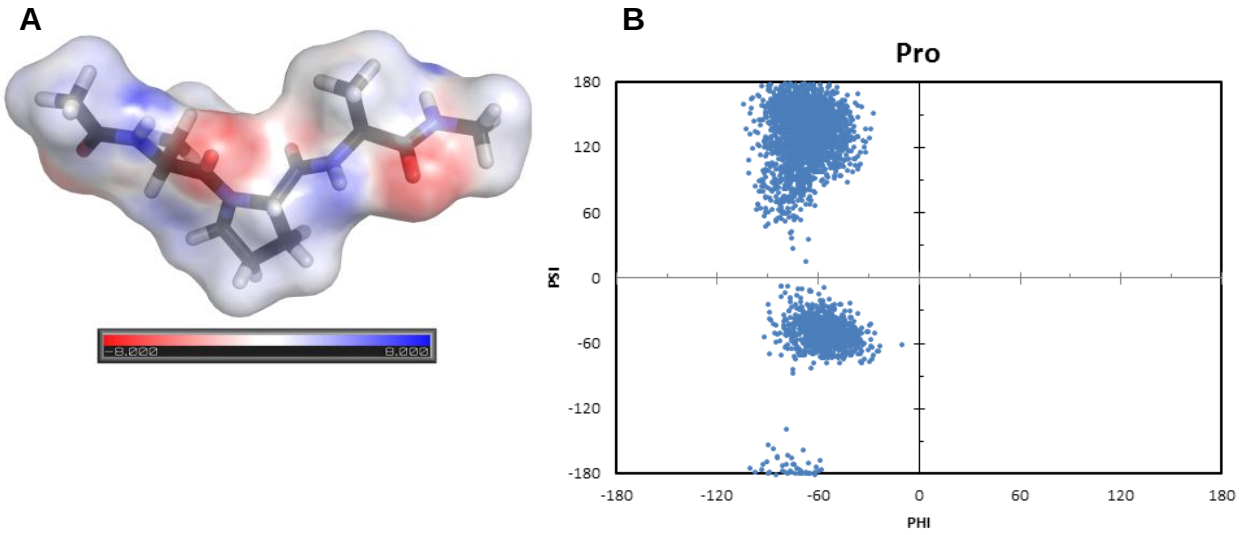


Figure 3.14 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Pro-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Pro

3.10.5 1-Aminocyclopentanecarboxylic Acid

1-aminocyclopentanecarboxylic acid (Acp) is a non-aromatic, sterically hindered unnatural amino acid. The electrostatic surface potential distribution shown in Figure 3.15 A shows the hydrophobic, electrically neutral nature of the Acp residue. The phi and psi angle measurements, shown in the Ramachandran plot in Figure 3.15 B, suggest this residue prefers a right handed alpha-helical secondary structure and thus could be a useful tool in influencing the peptide structure.

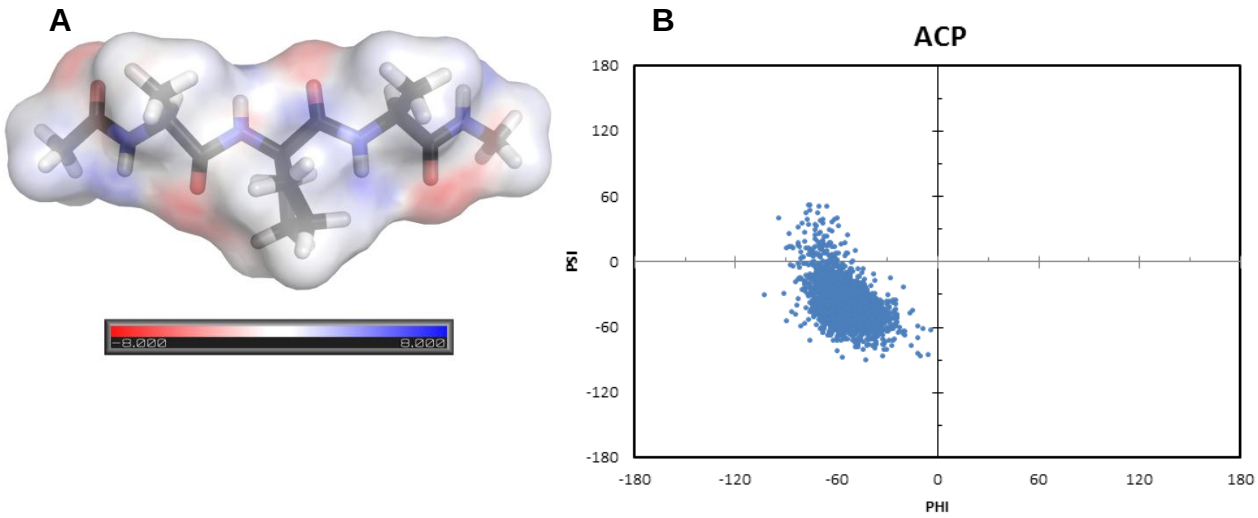


Figure 3.15 A. Electrostatic surface potential distribution B. Ramachandran plot phi/psi angle distribution for Acp

3.10.6 1-AMINOCYCLOHEXANECARBOXYLIC ACID

1-aminocyclohexanecarboxylic acid (Ach) is a non-aromatic, sterically hindered unnatural amino acid. The electrostatic surface potential distribution has shown in Figure 3.16 A the hydrophobic, electrically neutral nature of the Ach residue. The phi and psi angle measurements suggest the Ach residue prefers a right handed alpha-helical secondary structure

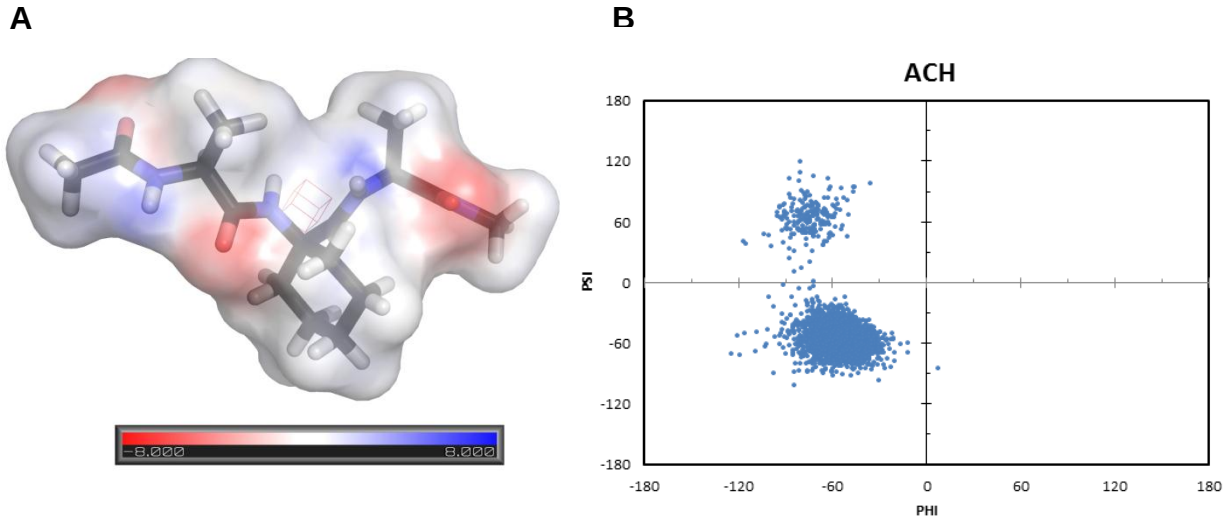


Figure 3.16 A. Electrostatic surface potential distribution B. Ramachandran plot phi/psi angle distribution for Ach

but, also has the flexibility to adopt beta-sheet/turn torsion angles. Due to the flexibility in the torsion angles adopted by the Ach residue it could be a useful tool in influencing the peptide structure and should be considered as a residue in the design of future unnatural amino acid containing AMPs.

3.10.7 p-FluoroPhenylalanine

Substitution of the phenylalanine at the 4 position created para-fluorophenylalanine (pFPhe). The electron withdrawing effects of the highly electronegative fluorine are not well reflected in the electrostatic surface potential shown in Figure 3.17 A. The Ramachandran plot suggests a preferred left handed alpha helical secondary structure (Figure 3.18 B), dramatically different than the substituted phenylalanine, Figure 3.12 B. The unusual secondary structure of

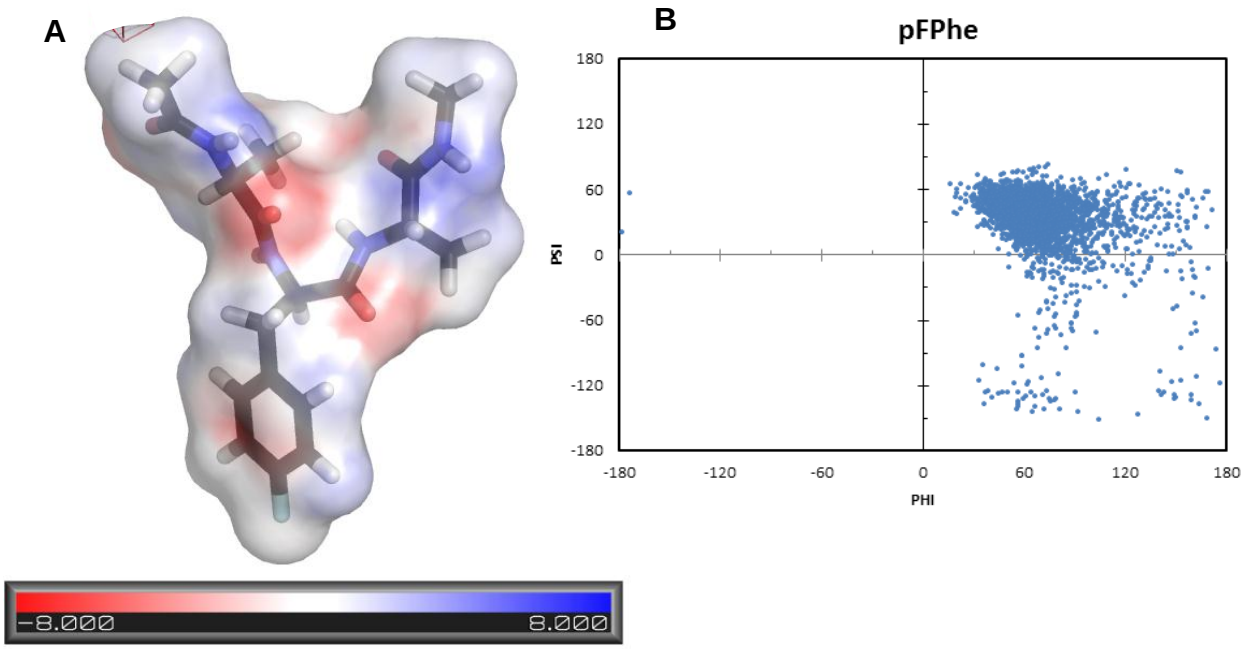


Figure 3.17 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-p-FPhe-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for p-F Phe

pFPhe is supported by an equally unusual CD spectrum.[20]

3.10.8 p-ChloroPhenylalanine

Substitution of the phenylalanine at the 4 position created para-chlorophenylalanine (pClPhe). The electron withdrawing effects of the highly electronegative chlorine are not well reflected in the electrostatic surface potential shown in Figure 3.18 A. The Ramachandran plot for pClPhe suggests a preferred left handed alpha helical secondary structure (Figure 3.18 B), dramatically different than the substituted phenylalanine, Figure 3.12 B. The pClPhe Ramachandran plot is consistent with the pFPhe plot (Figure 3.17 B) The unusual secondary structure of pClPhe is supported by an equally unusual CD spectrum.[20] The presence of angles populating the lower right hand quadrant of the Ramachandran plot is suspected of being transitional angles through which the preferred structure transitions. The lower right hand quadrant is usually considered a "forbidden" region however the standard amino acids have

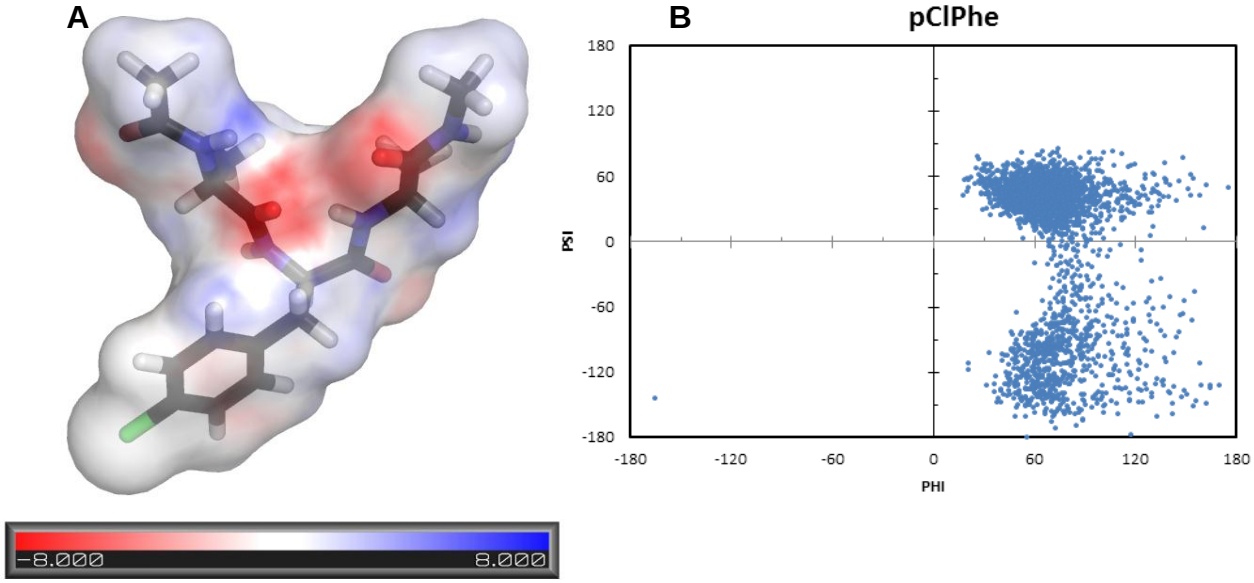


Figure 3.18 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-p Cl Phe-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for p-Cl Phe

demonstrated small populations which have populated the area.[1]

3.11 STANDARD AMINO ACIDS

Tripeptide simulations were also undertaken for all 20 of the standard amino acids to strengthen the validity of the experimental approach. The standard amino acids which were not part of the linkers or spacers are included.

3.11.1 VALINE

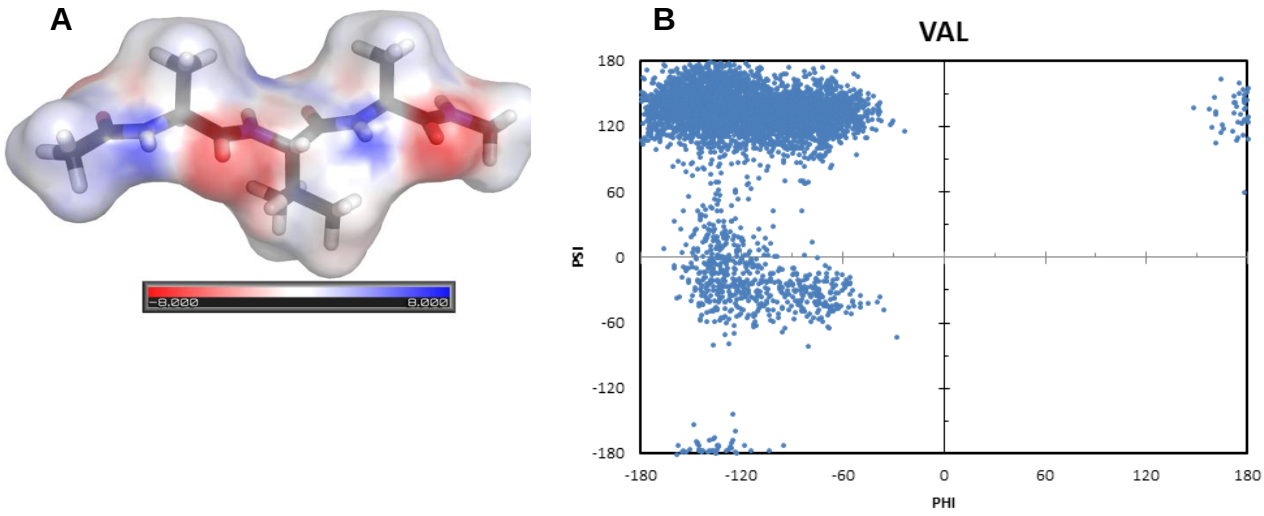


Figure 3.19 A. Electrostatic surface potential distribution of the tripeptide Ace-Ala-Val-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Val

3.11.2 LEUCINE

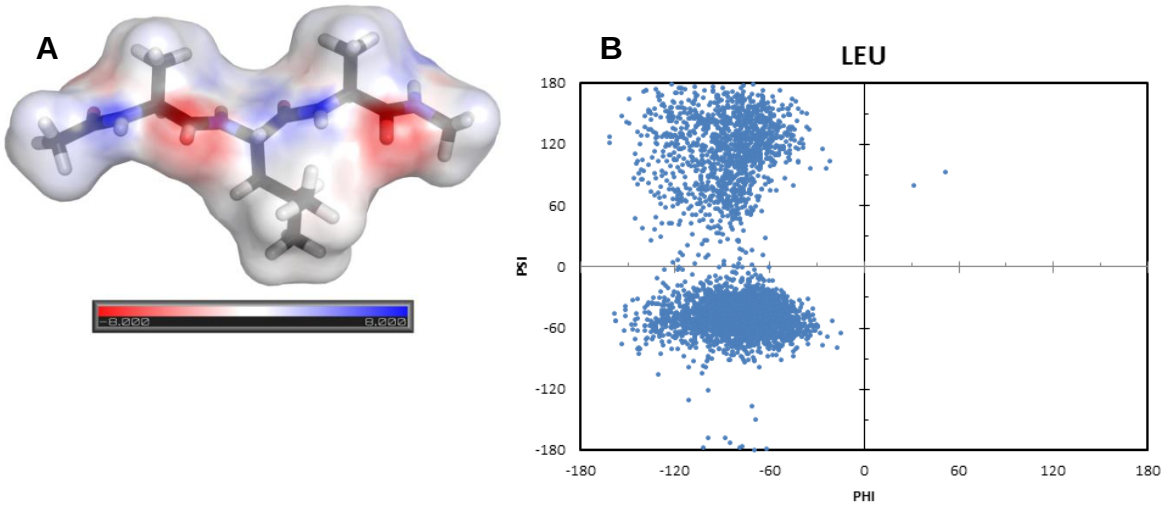


Figure 3.20 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Leu-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Leu

3.11.3 ISOLEUCINE

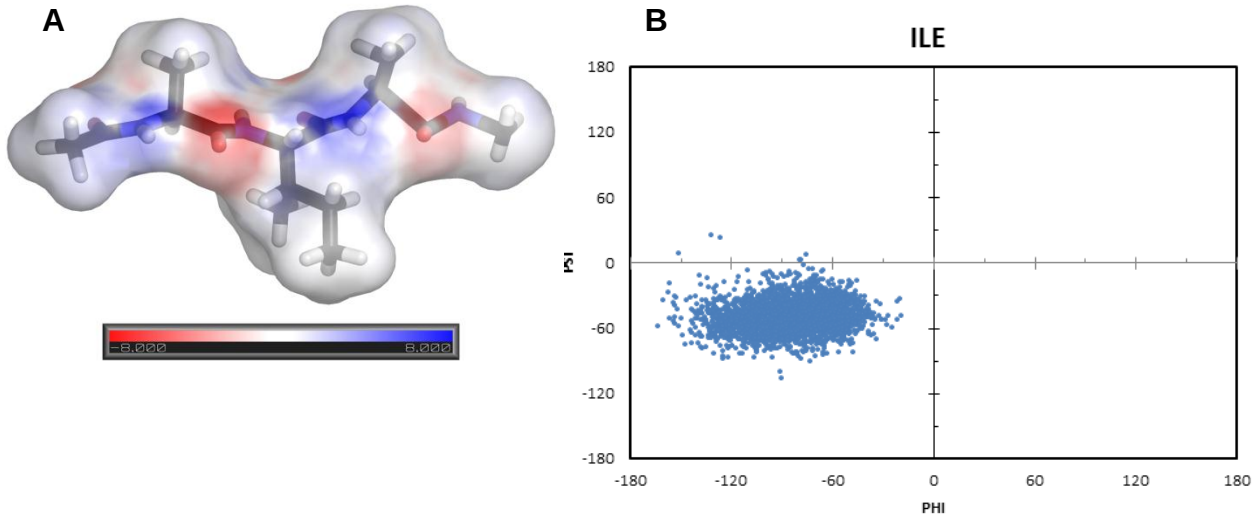


Figure 3.21 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Ile-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Ile

3.11.4 *TYROSINE*

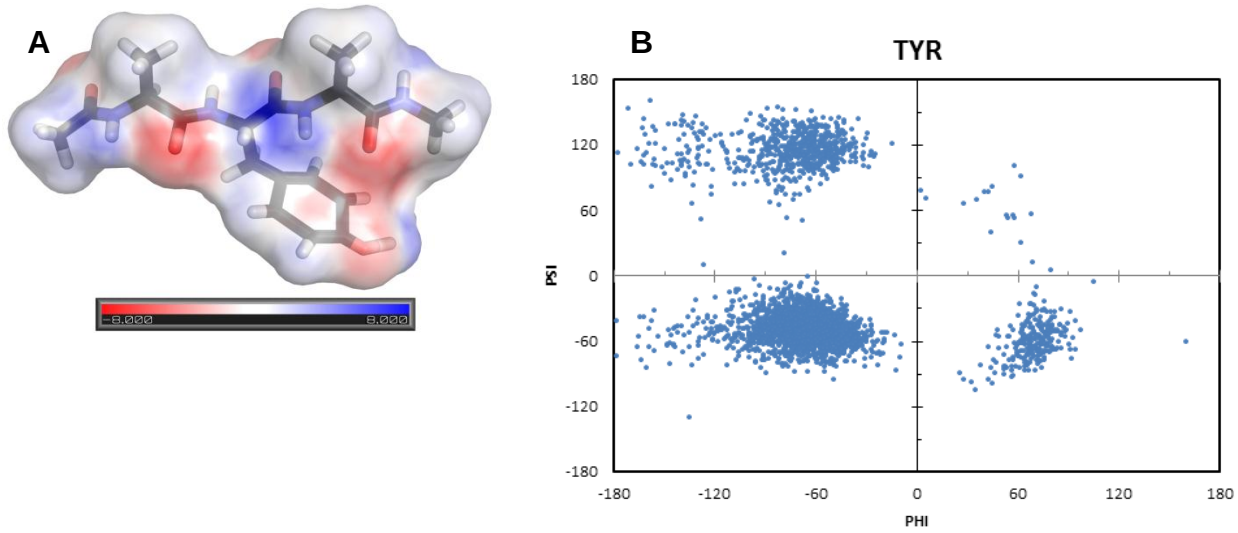


Figure 3.22 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Tyr-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Tyr

3.11.5 *TRYPTOPHAN*

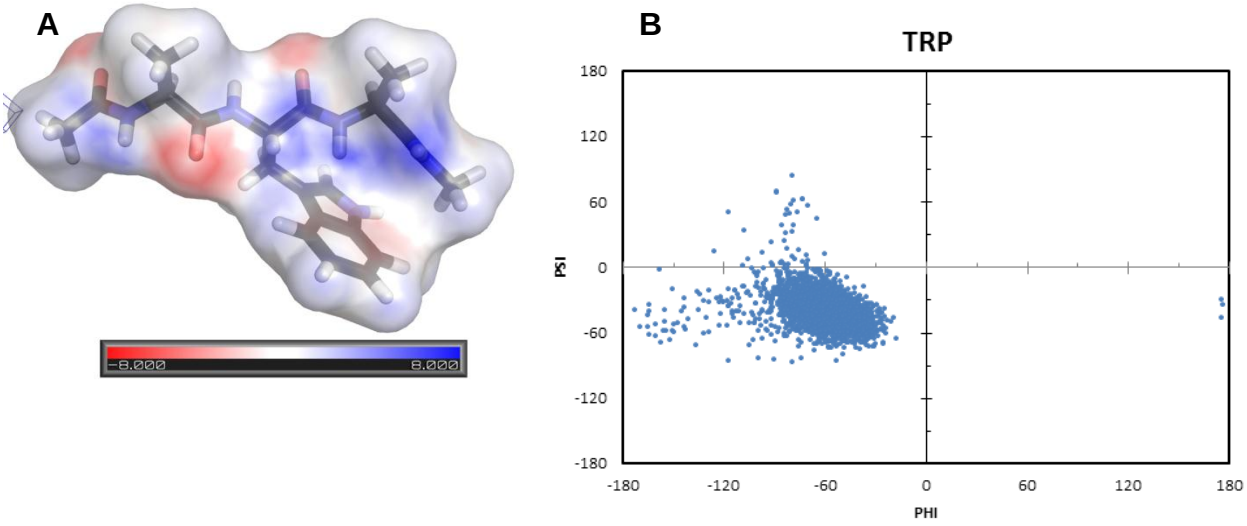


Figure 3.23 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Trp-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Trp

3.11.6 *Alanine*

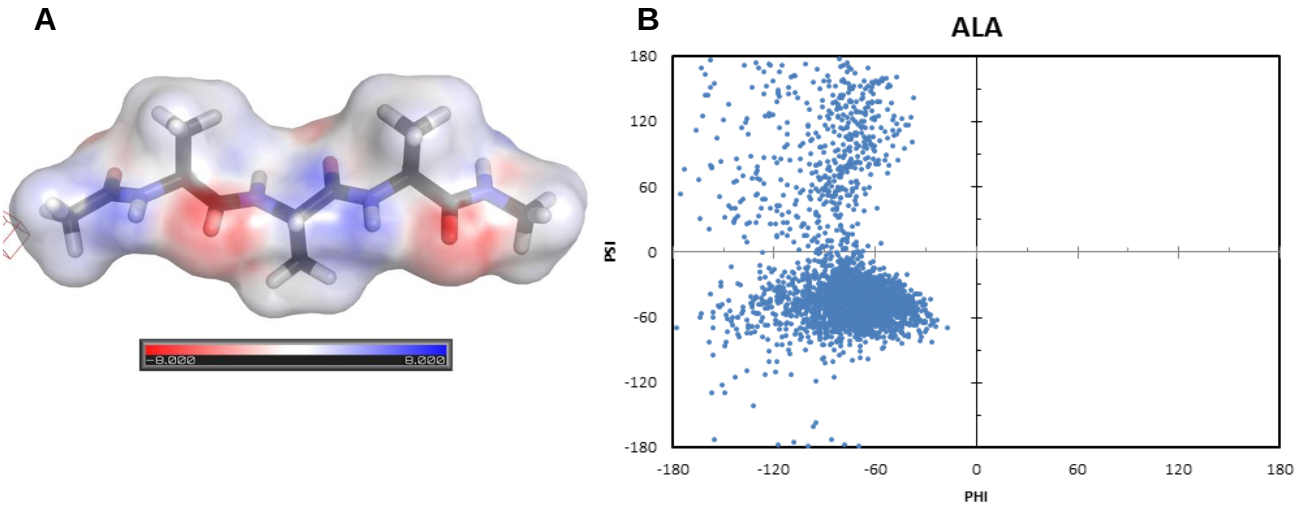


Figure 3.24 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Ala-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for alanine

3.11.7 *Serine*

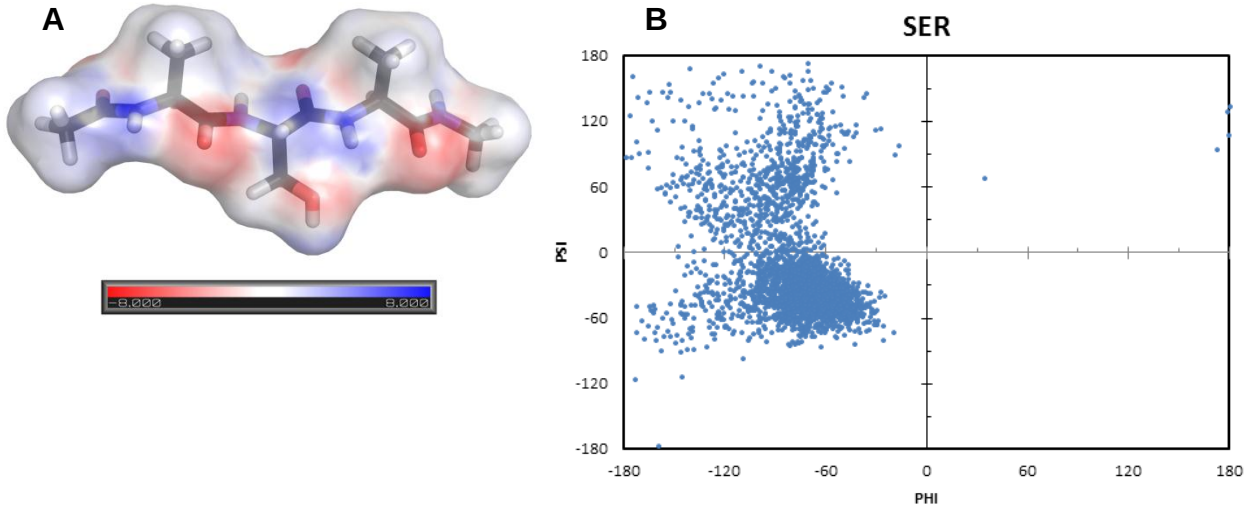


Figure 3.25 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Ser-Ala-Nme B. Ramachandran Plot phi/psi angle distribution for Ser

3.11.8 *THREONINE*

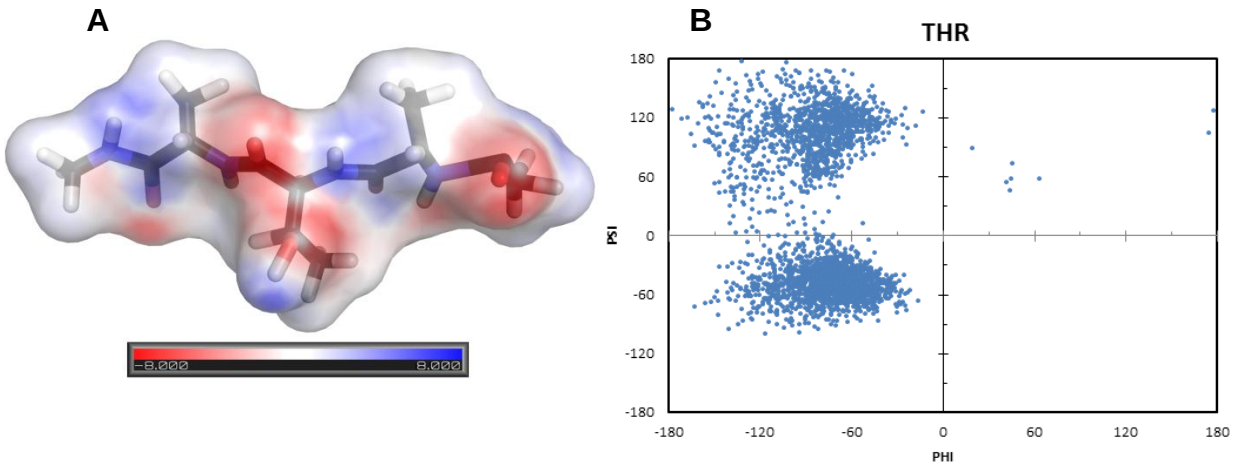


Figure 3.26 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Thr-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Thr

3.11.9 *CYSTEINE*

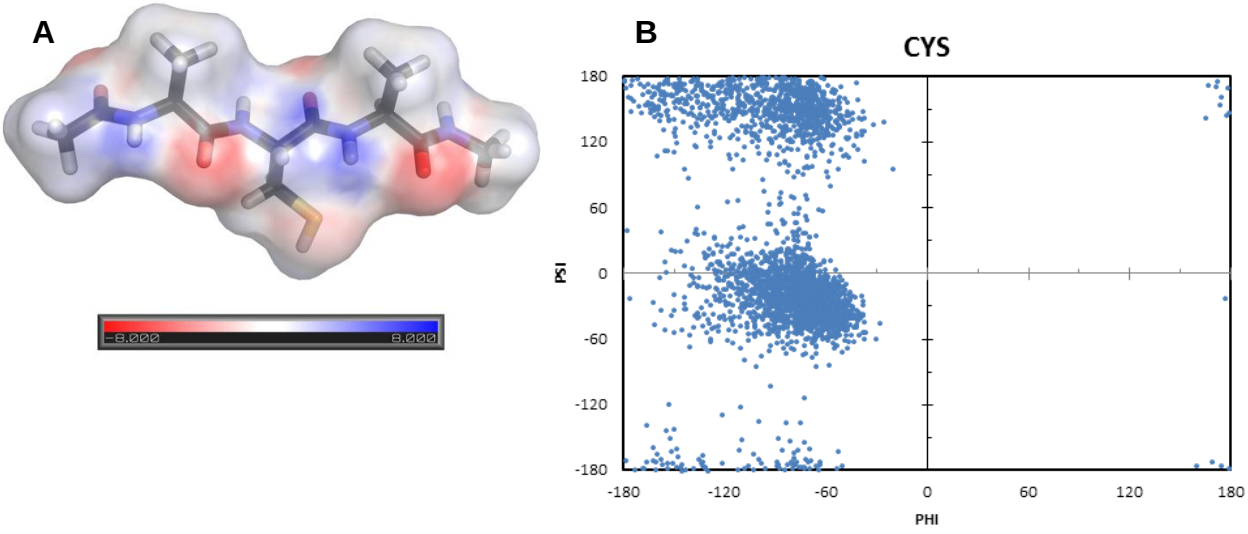


Figure 3.27 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Cys-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Cys

3.11.10 *METHIONINE*

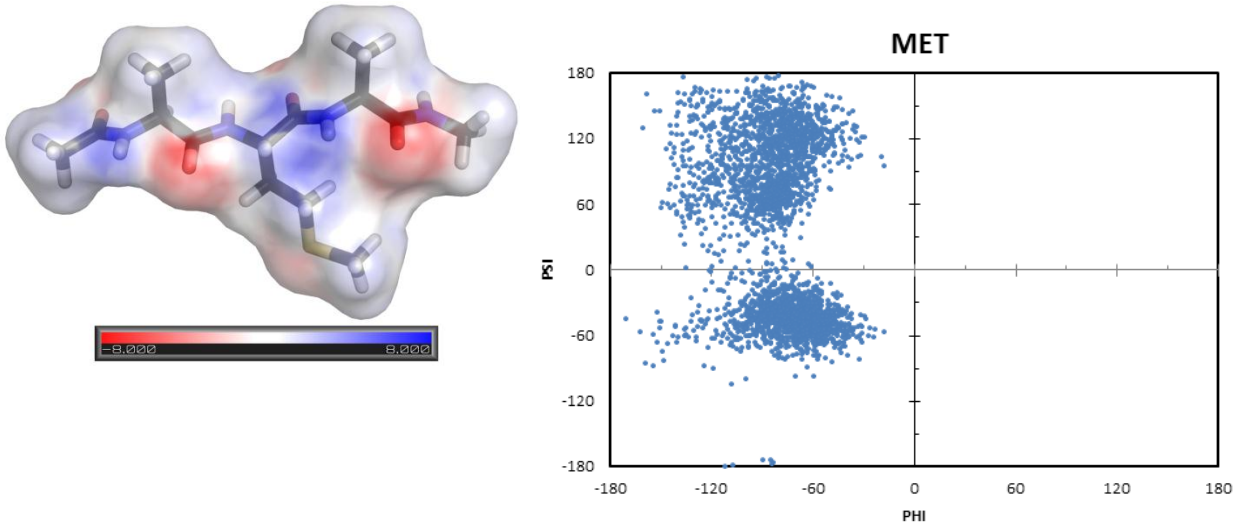


Figure 3.28 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Met-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Met

3.11.11 *ASPARAGINE*

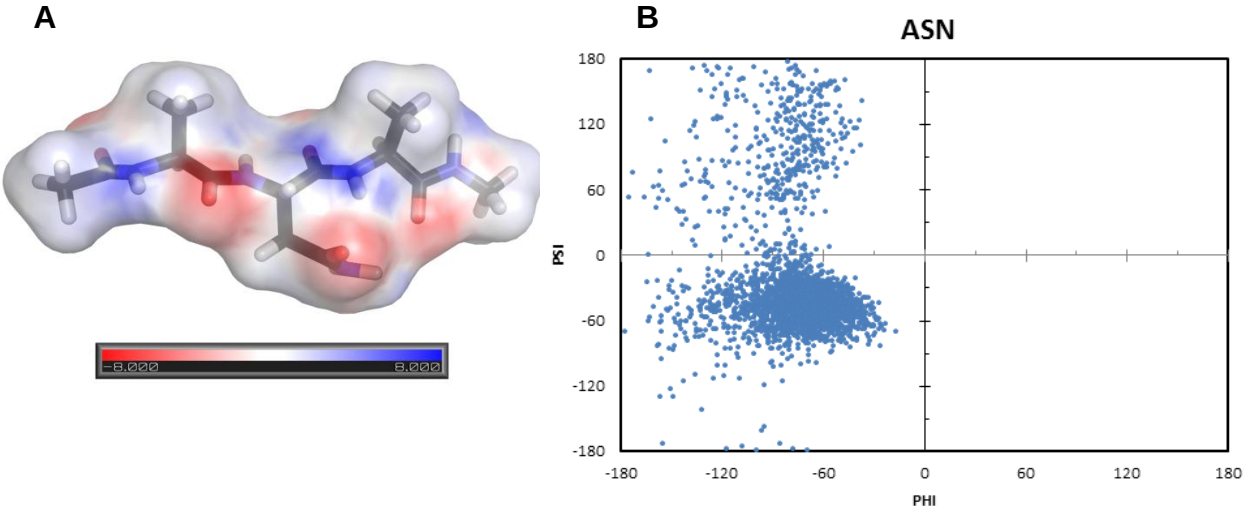


Figure 3.29 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Asn-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Asn

3.11.12 GLUTAMINE

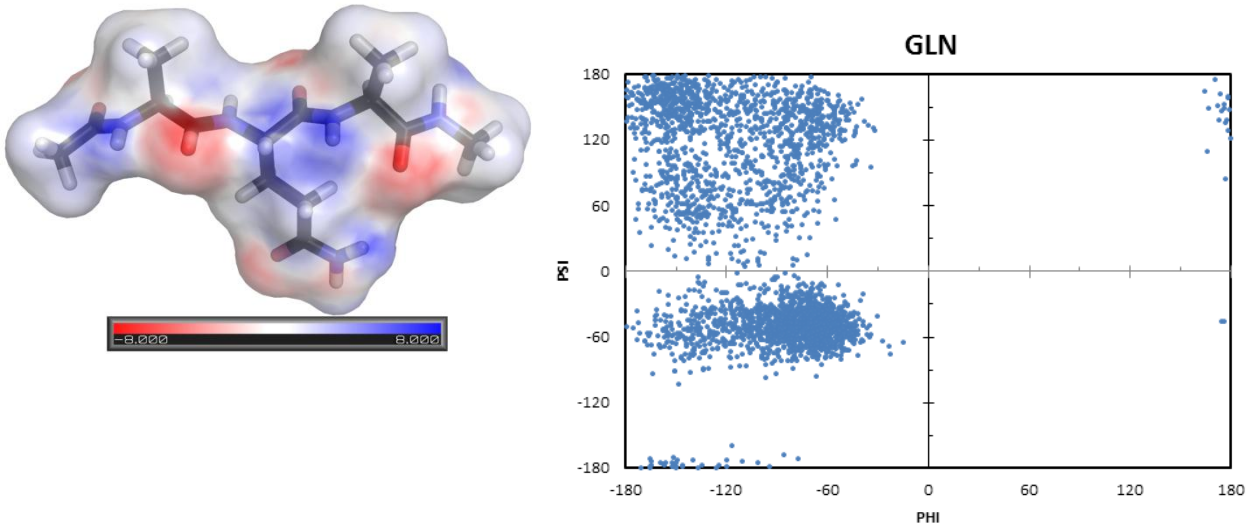


Figure 3.30 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Gln-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Gln

3.11.13 HISTIDINE

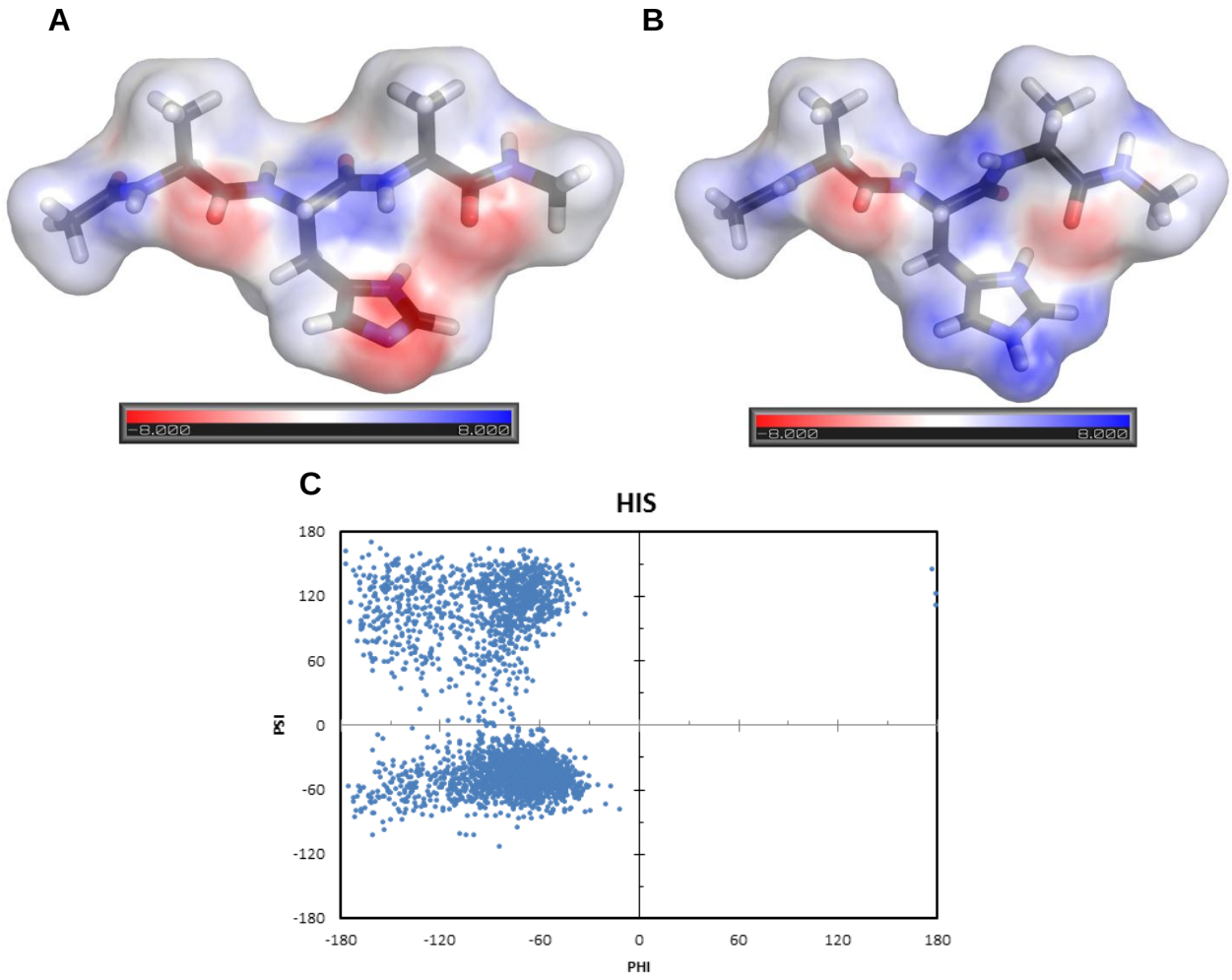


Figure 3.31 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-His-Ala-Nme, the unprotonated form of His B. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Hid-Ala-Nme, the protonated form of His C. Ramachandran plot phi/psi angle distribution for His

Within the AMBER program there are two neutral forms of histidine HisE and HisD in addition to one protonated for, HisP. The protonated and unprotonated forms have been simulated in the tripeptide arrangement. Figure 3.31 A shows the unprotonated form of histidine while Figure 3.31 B shows the protonated form. The effect of the additional proton is evident in the electrostatic surface potential map which reflects the anionic nature of the unprotonated form

of the His and the cationic charge associated with the protonated His residue. The angular phi and psi distribution for the two forms of this residue were not significantly different and the average structure is shown in Figure 3.31 C. The Ramachandran plot for His shows the preference for a right handed alpha helical secondary structure however also shows a less dense population present in the beta-sheet/turn region of the plot.

3.11.14 ASPARTIC ACID

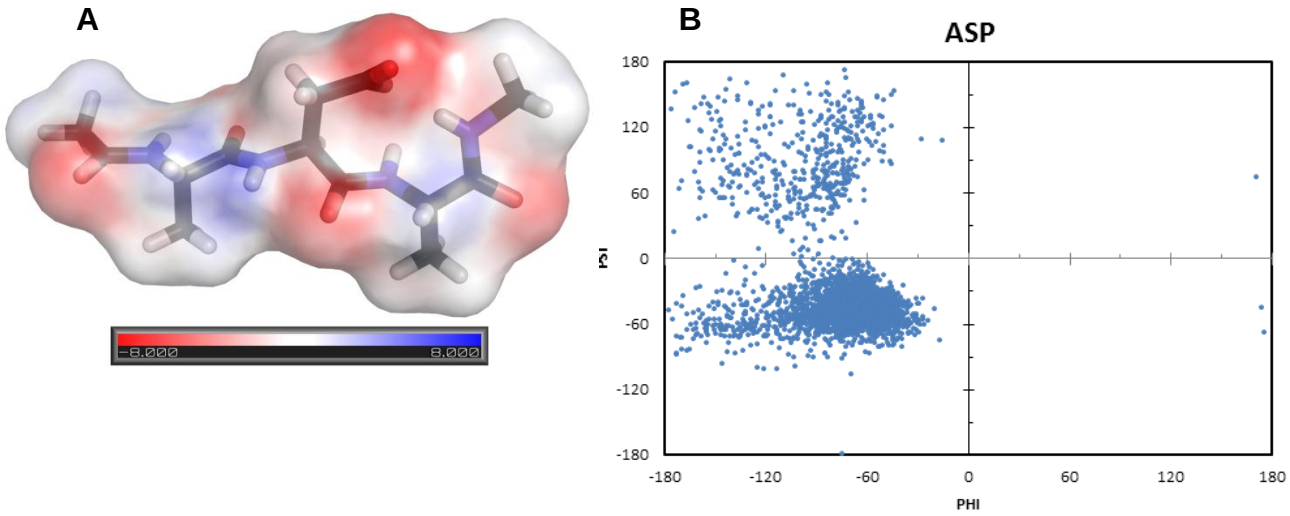


Figure 3.32 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Asp-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Asp

3.11.15 GLUTAMIC ACID

Electrostatic surface potential maps were created using the Solvent Accessible Surface Area (SASA) gives an idea regarding a molecule's surface accessible to the surrounding solvent. The calculation is usually derived using a rolling ball technique in which a sphere with a specified radius is able to determine the surface area in square angstroms ($Å^2$). This can help predict the potential for the molecule to interact with the surrounding environment whether it is solvent or a bilayer, the SASA is also useful when performing implicit solvent simulations which

is used to improve the prediction of protein secondary structure. Considering the volume of specific amino acids and the distribution of charge for each individual residue can help in the rational design of therapeutic peptides. Tailoring the charge distribution of peptides can optimize for specific membrane charges (especially when targeting multiple drug resistant organisms).

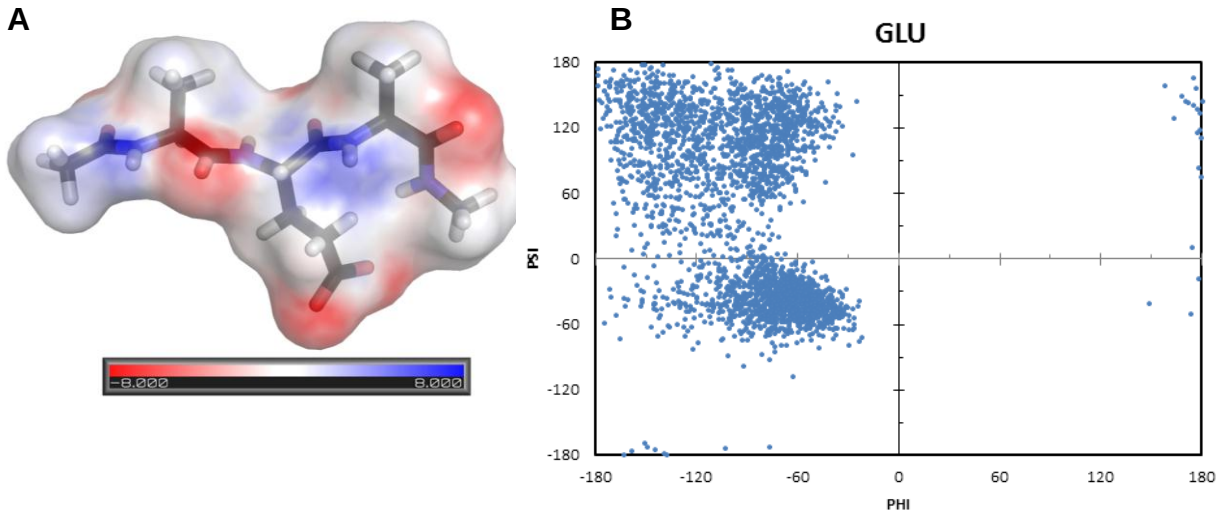


Figure 3.33 A. Electrostatic surface potential distribution for the tripeptide Ace-Ala-Glu-Ala-Nme B. Ramachandran plot phi/psi angle distribution for Glu

The tripeptide MD simulations for each of the 20 standard and unnatural amino acids have enabled a small sampling, 5 ns, of the conformational configurations of each system. Calculations of the torsion angles associated with each of the standard amino acids have been consistent with previously reported *in silico* experiments performed by other investigators.[1] Angle measurements derived from the experimentally derived structures and reported in the protein data bank (PDB) are also consistent with the calculated Ramachandran distributions for each of the standard amino acids. The Ramachandran plots are consistent with the experimental data for the 20 standard amino acids The consistency of the Ramachandran plots with literature values[1] validate the force field parameters because they were able to replicate the experimentally

derived angles and by extension the expanded force field has been applied to the unnatural amino acid enabling a sampling of the backbone torsion angles and comparison of the resulting Ramachandran plots with the known regions associated with the various secondary structures.[1]

Calculation of the solvent accessible surface area and volume calculations were performed with the UCSF Chimera package.[15] Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health (National Center for Research Resources grant 2P41RR001081, National Institute of General Medical Sciences grant 9P41GM103311).[15]

3.12 REFERENCES

1. Beck, D. A. C., Alonso, D. O. V., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences. 105*, 12259-12264.

2. Accelrys. *Materials studio.*

3. *Cerius2, Version 4.10L, Conformational Search and Analysis Manual.* (2005).

4. Gaussian 03, Revision C.02, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004. (2003) *Gaussian 03,Gaussian inc, 340 Quinnipiac street, Building 40, Wallingford, CT, 06492.*

5. Singh, U. C., and Kollman, P. A. (1984) *An approach to computing electrostatic charges for molecules.* John Wiley & Sons, Inc.

6. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V.,

Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9.*

7. Feig, M., Karanicolas, J., and Brooks III, C. L. (2004) MMTSB tool set: Enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model. 22*, 377-395.

8. Shimada, N., Ino, H., Maie, K., Nakayama, M., Kano, A., and Maruyama, A. (2011) Ureido-derivatized polymers based on both poly(allylurea) and poly(l-citrulline) exhibit UCST-type phase transition behavior under physiologically relevant conditions. *Biomacromolecules. 12*, 3418-3422.

9. Andrea, T. A., Swope, W. C., and Andersen, H. C. (1983) The role of long ranged forces in determining the structure and properties of liquid water. *J. Chem. Phys. 79*, 4576-4584.

10. Miyamoto, S., and Kollman, P. A. (1992) Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry. 13*, 952-962.

11. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) *J. Comput. Phys. 23*, 327.

13. Baker, N. A., Sept, D., Joseph, S., Holst, M. J., and McCammon, J. A. (2001) *Proc. Natl. Acad. Sci. 98*, 11037.

14. The PyMOL Molecular Graphics System. *The PyMOL molecular graphics system, version 1.2r3pre, schrödinger, LLC.* The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC ed.

15. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J Comput Chem. Oct., 25 (13)*, 1606.

16. Motta, A., Reches, M., Pappalardo, L., Andreotti, G., and Gazit, E. (2005) The preferred conformation of the tripeptide ala-phe-ala in water is an inverse -turn: Implications for protein folding and drug design. *Biochemistry (N. Y.). 44*, 14170-14178.

17. Advanced Chemistry Development, I. (2008) *ACD labs 12.0 ChemBasic LogP calculator.*

18. Petrauskas, A. A., and Kolovanov, E. A. (2000) *Perspectives in Drug Discovery and Design. 19*, 99.

19. Bilton, C., Allen, F. H., Shields, G. P., and Howard, J. A. K. (2000) Intramolecular hydrogen bonds: Common motifs, probabilities of formation and implications for supramolecular organization. *Acta Crystallographica Section B. 56*, 849-856.

20. Russell, A. L. (2011) *Thermodynamic and spectroscopic investigations of novel antimicrobial peptides containing unnatural amino acids with model membrane systems.* Doctor of Philosophy ed., East Carolina University, Greenville, North Carolina.

CHAPTER 4
MICELLE SIMULATIONS

ABSTRACT

Antimicrobial peptides are known to adopt specific secondary structure upon interaction with microbial cell surfaces and this unique phenomenon is dependent on the electrostatic charge distribution of the membrane surface.[1,2] SDS Micelles have been utilized as electrostatic models for membranes in numerous molecular modeling studies.[2-8] Experimental data is abundant for SDS micelles and their structures have been well established in numerous environments.[8] Molecular dynamics (MD) simulations were undertaken examining the effect of two different peptides, peptide **23** with a glycine linker and peptide **36** with the β-Alanine linker, on the SDS micelle. The exact mechanism of action of the AMPs has been elusive; making MD simulation of the system a good tool in understanding peptide-micelle interactions which contribute to the mechanism of action (MOA). The simulation data is consistent with DOSY NMR and circular dicroism (CD) structural data. Analysis of the Ramachandran plots for both peptide **23** and peptide **36** reflect the alpha helical and beta turn/sheet structural character associated with both peptides and corroborated by the CD and 2D NMR experiments. The orientation of the peptides relative to the micelle surface has offered insight into the orientations of the peptide, relative to the micelle surface, and provides a basis for further studies in various membrane environments.

4.1 MICELLE CONSTRUCTION

The SDS molecule previously parameterized (See Chapter 2 and **Appendix A**) was used to construct micelles with 50, 62, and 80 SDS units using a custom script[9] which evenly distributes the terminal carbon of the SDS molecules evenly around a sphere with a radius of 9 angstroms, with polar atoms pointing away from the center (Figure 4.1). The three sizes were selected to reflect the estimated experimental variability in the micelles used to conduct CD

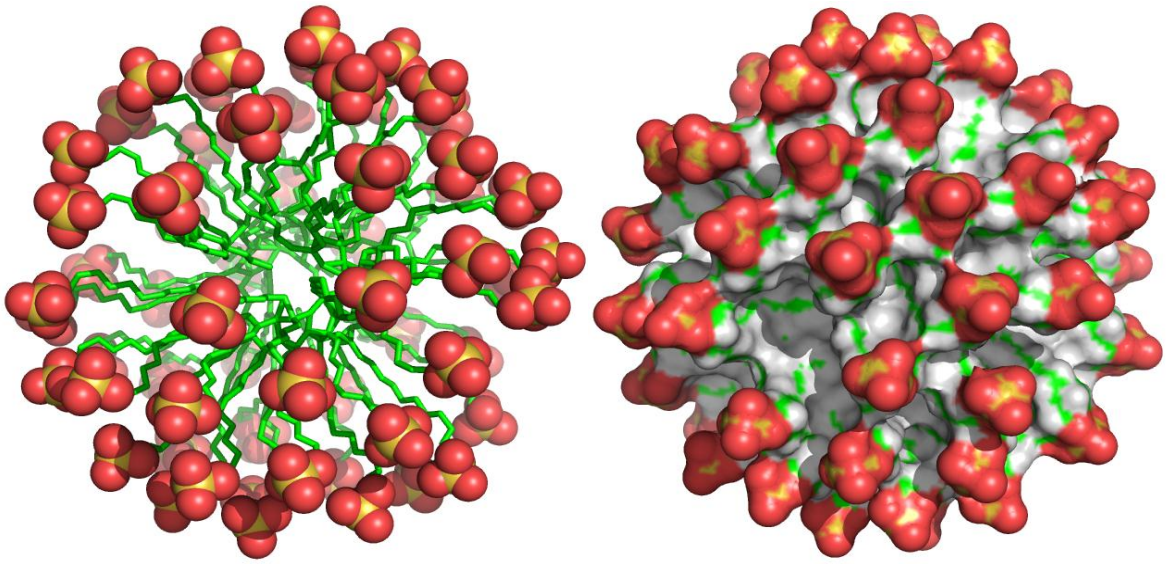


Figure 4.1 A) Starting 62 SDS micelle structure, sulfur is noted with yellow spheres and oxygen is noted with red spheres. B) The surface of the starting 62 SDS micelle structure

studies examining the secondary structure of AMPs.[10] Although three different aggregate micelles were constructed a parsimonious approach necessitated the selection of the most reasonable micelle for simulation. Micelles consisting of 62 individual SDS molecules were within the aggregation range associated with SDS micelles in the literature.[7, 11-13]

Water molecules were added to the micelle system in a two-step process. Each SDS molecule has a charge of -1. The micelle was neutralized by first randomly selecting 62 water molecules which were within 5 angstroms of each sulfur atom. The water molecules which fell into the previously specified criteria were randomly replaced using the MMTSB software to add 62 Na^+ ions to the system.[14] Following addition of sodium ions to neutralize the micelle a TIP3P, cubic water box extending 14 angstroms from the micelle surface was added.[15] Water molecules were randomly replaced using MMTSB software to bring the ionic strength to 0.15 M NaCl.[14] The ionic strength was selected due to its physiological relevance. The micelle was simulated using constant pressure dynamics for 300 ps prior to addition of any peptide. During this time density stabilization was observed however, the system did not stabilized based upon

the RMSD of the molecules relative to the starting minimized structure. This system was used in a preliminary equilibration simulation of the solvated, neutralized micelle alone (Figure 4.2).

The equilibration simulation was conducted in five steps: pre-minimization,

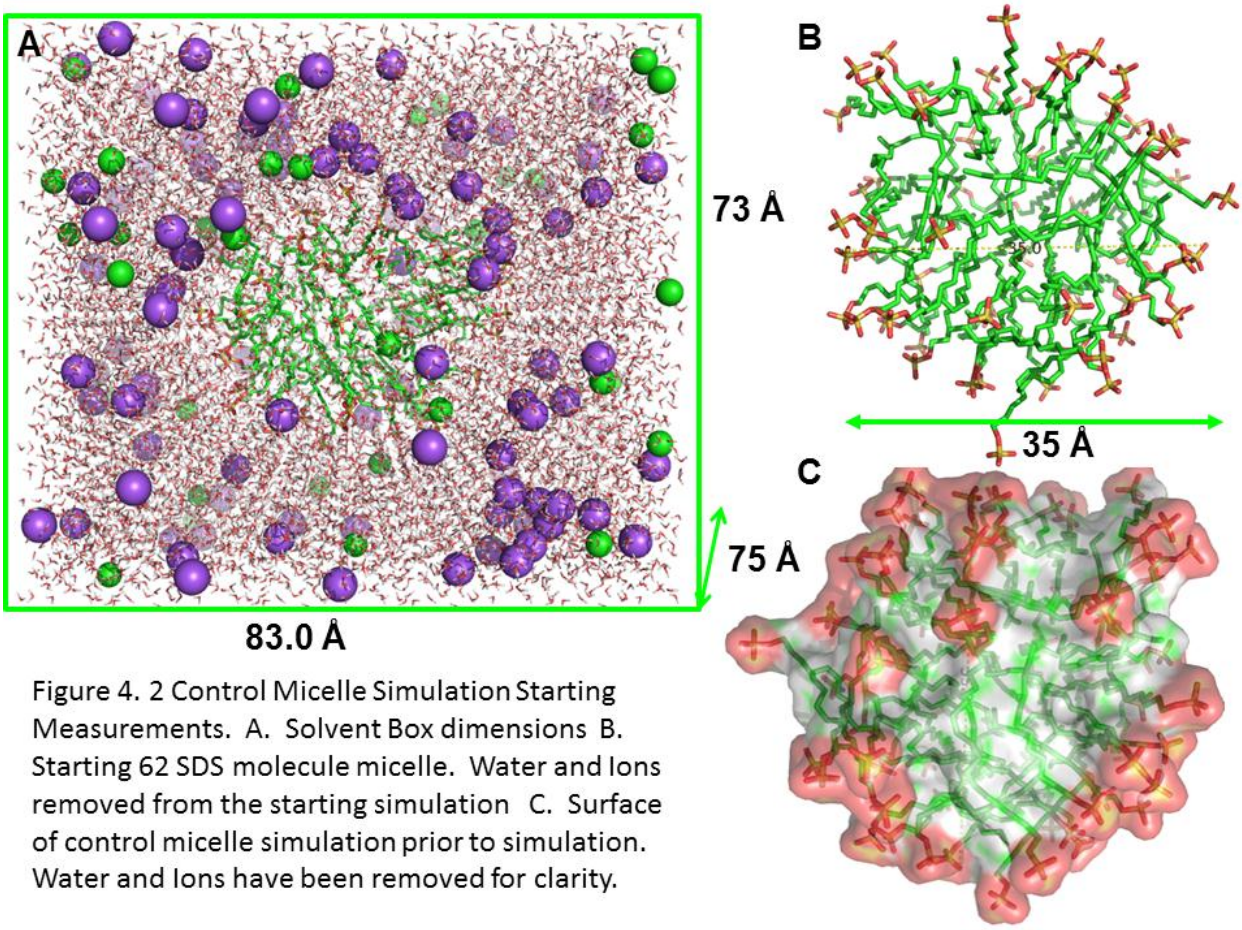


Figure 4. 2 Control Micelle Simulation Starting Measurements. A. Solvent Box dimensions B. Starting 62 SDS molecule micelle. Water and Ions removed from the starting simulation C. Surface of control micelle simulation prior to simulation. Water and Ions have been removed for clarity.

minimization, heating, constant pressure dynamics (equilibration phase), and constant volume dynamics (production phase). The micelle was minimized using a two-step process first constraining the micelle atoms allowing the water to settle around the surface of the micelle, followed by a second unrestrained minimization. Each minimization used 5000 steps of steepest descent with constant volume, no pressure scaling, and Andersen temperature coupling. Electrostatic interactions were truncated at 10.0 angstroms and were divided. Energy values

were printed every 25 steps. A second minimization was performed without restraints using 2500 steps of steepest descent followed by 500 steps of conjugant gradient.

Following minimization the system was warmed to 300K over 2.1 ns. The non-bonded cutoff was set to 10 angstroms. A 1 fs time step was used while constraining hydrogen atom bond lengths using the SHAKE algorithm.[16] Anderson temperature scaling was used to adjust the temperature of the system at each step. The step size was reduced to prevent SHAKE and other problems from developing.[17]

The micelle was subjected to constant pressure dynamics until density stabilization had been attained. The production run using constant volume dynamics continued until a 500 ps trajectory had been acquired. Micelle stabilization, based upon the RMSD plot was not attained during this short simulation period. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method with a non-bonded cutoff set to 12 angstroms. A 1.5 fs time step was used while constraining the hydrogen atom bond lengths using the SHAKE algorithm. Simulation structures were saved every 500 steps. The parameters selected for the simulation environment were designed to model *in vivo* conditions as much as possible. All simulations were conducted at 300 K. Anderson temperature scaling was used to maintain the temperature of the system throughout the course of the production run. A structural transition is observed for SDS molecules as the temperature is elevated making the temperature at which the simulation is conducted very important. Temperatures below 300 K are associated with a crystalline aggregate at temperatures above 310 K micelles are able to form a higher

temperatures.[5,18] The differences in temperature may be responsible for the observed experimental differences.

Much of the computational expense involved with MD simulations results from the number of atoms involved in a simulation. Reducing the number of atoms would increase the speed of the simulation. One way to accomplish the reduction of atoms involved in a simulation is to use an implicit model in place of the common explicit water models (ex. SPC, TIP3P, TIP4P). Due to the interest focusing on the solute of the system, in this case the micelle and peptides, reducing the water molecules to an implicit model seems an obvious solution to expedite data collection. Use of the Generalized Born simulation technique provides a description of solvent and counter ions as implicit model components by approximation of the linearized Poisson-Boltzmann equation.[19] The high-dielectric of the water solvent is accomplished using a continuum electrostatics model.[19] Solvation energies are combined with the total energy of the system and are included in the force terms driving the dynamics of the simulation.[19] Use of the Generalized Born model was attempted with the micelle alone and in the presence of peptide **23** and **36.** The Generalized Born model utilizes an implicit solvent system, reducing the total number of atoms involved in the simulation. Numerous attempts using the Generalized Born approach have demonstrated that with the current set of parameters, this technique may not be possible. Given the dynamic nature of micelles and the multi-bodied construct, applying the Generalized Born approach is of limited utility for the current system of interest.

Another technique for expediting md simulations is using an octahedral simulation box rather than the traditional cubic box. A octahedral box has the four corners sliced off creating a octahedral shaped face on each of the six sides of the traditional cubic box. Any reduction in the

number of atoms involved in the simulation will result in a decrease in the simulation time. Given the long time scale used to study the peptide micelle interaction and the large movements

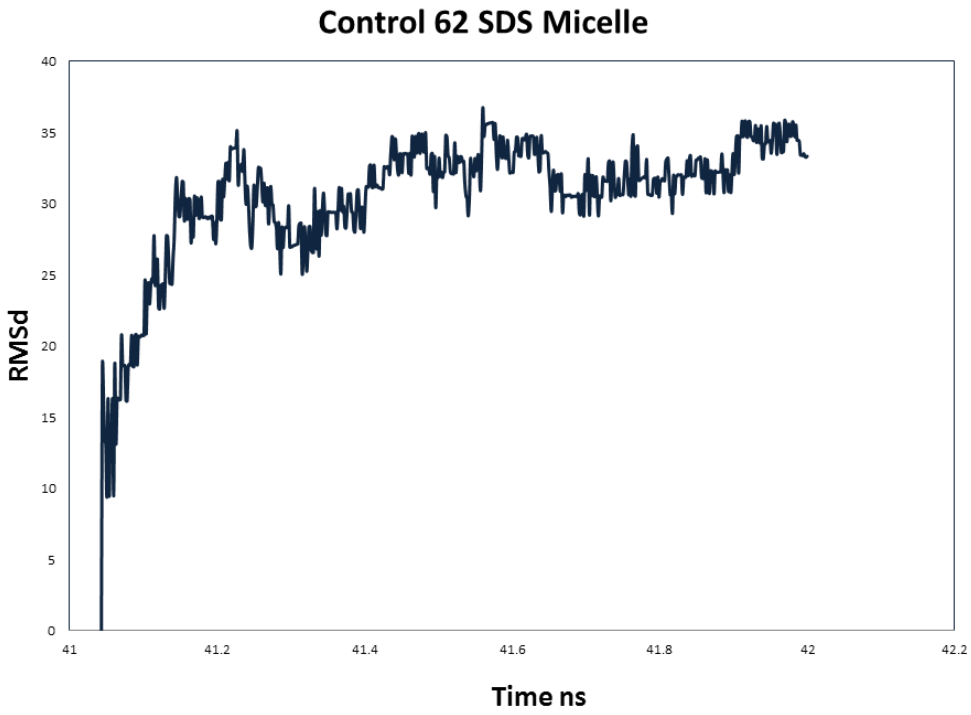


Figure 4.3 RMSD plot of the control 62 SDS molecule micelle. The final nanosecond of the simulation is shown.

associated with the micelle, use of the octahedral box with a comparable size box was not able to provide computational savings and the re-imaging of the system proved problematic. The necessary increased size of the octahedral box negated the savings expected from the smaller box and was thus abandoned.

4.2 CONTROL SIMULATION

The simulation of the control micelle in the solvated and ionic concentration of the system was adjusted. This system continued as a control simulation. The simulation was re-setup (ions and water were removed from the system) for consistency with the experimental

simulations with peptides **23** and **36**. The production run with the canonical ensemble (NVT) (see Chapter 2 for a discussion of the various ensembles) was continued for 42 ns. The resulting structure is depicted in Figure 4.3. The non-spherical nature of the micelle is consistent with other scientists' observations.[12,13] The exposure of the aliphatic, hydrophobic tails of the dodecylsulfate to the solvent surface[12,13] may be an unexpected occurrence but is well documented and consistent with other research group's simulation data associated with SDS micelles.[12,20]

The control micelle simulation was considered stable when the average slope of the RMSD plot, relative to the minimized structure, had a slope approaching 0. Stabilization is observed in this instance as opposed to equilibration, the more common measure of MD simulation completion, due to the dynamic nature of the micelle. The control micelle maintained the aggregate throughout the course of the trajectory with the individual SDS molecules continuing to oscillate relative to other SDS molecules of the micelle. The control simulation demonstrated the force field parameters developed for SDS were able to replicate the movement

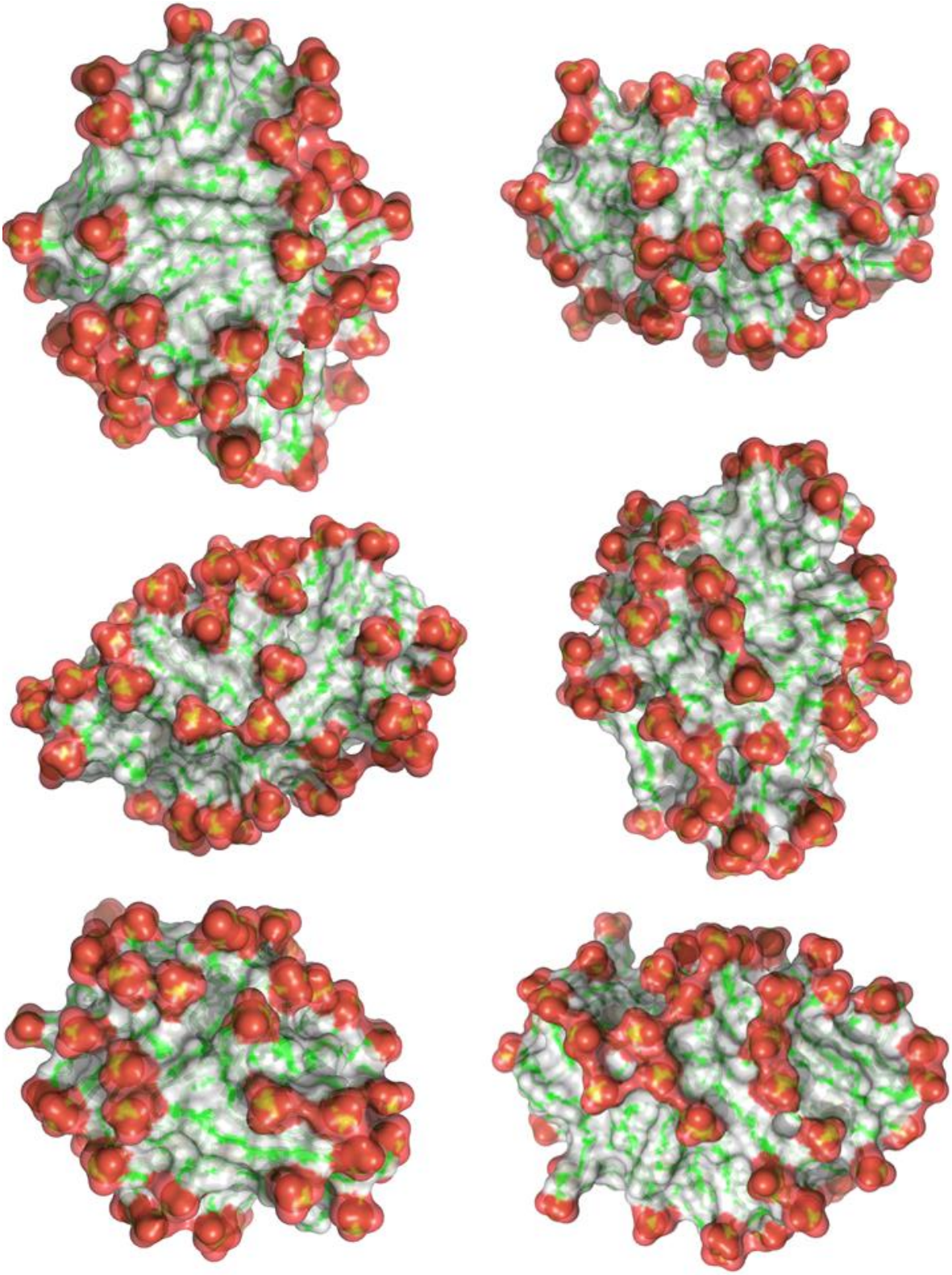


Figure 4.4 62 SDS micelle after 42 ns of md simulation. The 6 images (of each face of the PBC) are of the same micelle from the same time step noting the dynamic nature of the micelle and the non-spherical shape it adopts. Water and ion atoms have been removed for clarity.

other investigators had observed using other force fields and software.[12, 20]

4.3 PEPTIDE/MICELLE SIMULATIONS

There are several well studied techniques for the simulation of peptides with micelles. One starting setup places the peptide to be simulated at the center of the vacant micelle and is simulated with the expectation that the peptide will migrate to the micelle surface in its preferred orientation ending at the micelle surface.[4,6,7,13] The central placement of the peptide is expected to minimize placement bias allowing the peptide to move to its preferred orientation. Another approach to the peptide/micelle simulations is placing the peptide at the surface of the micelle, a sufficient distance away, enabling the peptide to be able to move and change conformation while still being close enough to provide potential for electrostatic interaction with the micelle.[3]

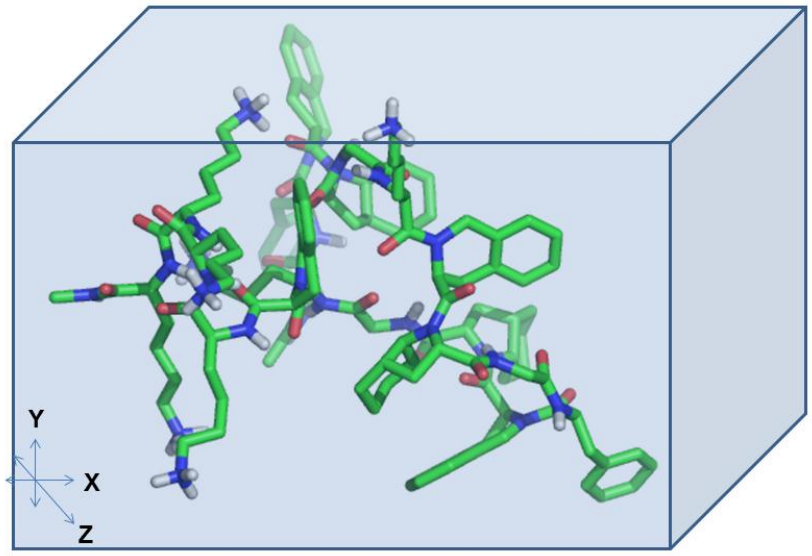


Figure 4.5 Peptide **23** reflecting the "faces" of the peptide used in preliminary simulations.

The second approach was utilized in this study with the hope that the simulation time required (and thus the computational expense) will be minimized. Preliminary studies of peptide/micelle interactions was undertaken to determine the effect of the peptide's "facial" placement relative to the micelle surface. Each of the four Y and Z axial faces of the peptide, as exemplified in Figure 4.5, was orientated toward the micelle surface. The peptides were placed within 7 angstroms of the micelle surface and allowed to progress through the five step of the simulation process. The resulting peptide orientations appear highly dependent upon the starting orientation, as would be expected. The polar residues do not seem to completely determine the final orientation of the peptide relative to the micelle surface, see Figure 4.6. Similar simulations were conducted for peptide **36** with similar results. The conclusion from the simulation is that while the starting orientation of the peptide relative to the micelle surface does have an impact upon the final orientation, it does not prevent re-orientation of the peptide to a preferred conformation throughout the course of the simulation.

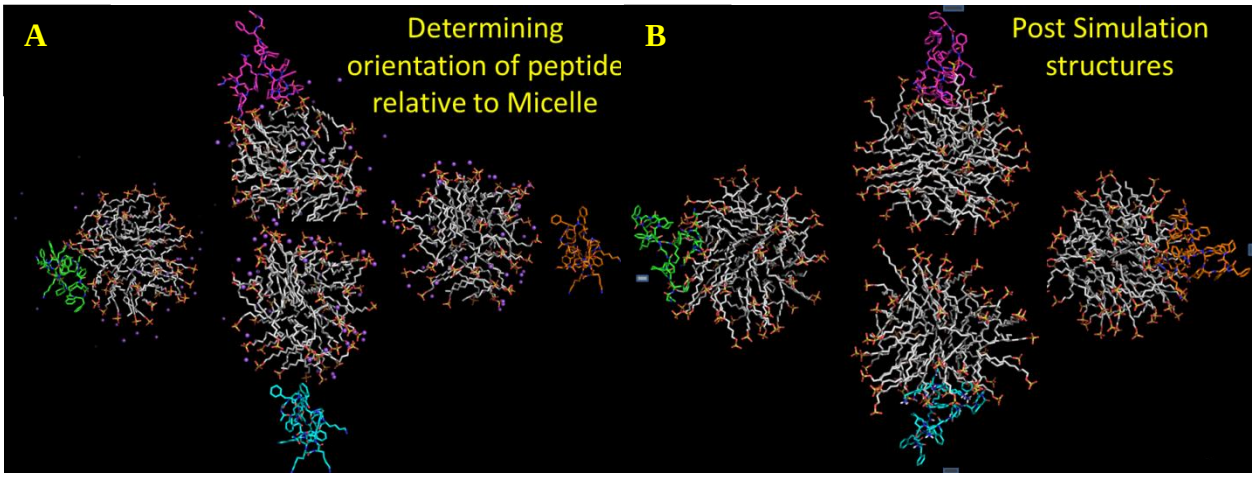


Figure 4.6 Peptide **23** Demonstrating the effect of orientation relative to the micelle surface A starting orientations of peptide **23**. B. Final orientation of peptide after molecular dynamics simulation.

4.3.1 PEPTIDE 23 WITH MICELLE

Peptide **23** was simulated with a 62 SDS molecule micelle for 25 ns. The same micelle as was used as the starting structure in the control micelle simulation was used as the starting structure for the peptide/micelle simulations. An initial box was placed extending 5 angstroms from the surface of the micelle. The system was prepared as previously described (Figure 4.7 A). Water molecules extended 18 angstroms from the micelle surface (Figure 4.7 B). The starting position of peptide relative to the surface was previously investigated using simulations with each of four "faces" of the peptide orientated toward the micelle surface (data not shown). While one of the cationic lysine residues was placed relatively close to the micelle surface (Figure 4.8) other nonpolar residues including Tic, Oic, and Phe were in comparable proximity with no portion of the peptide less than 6

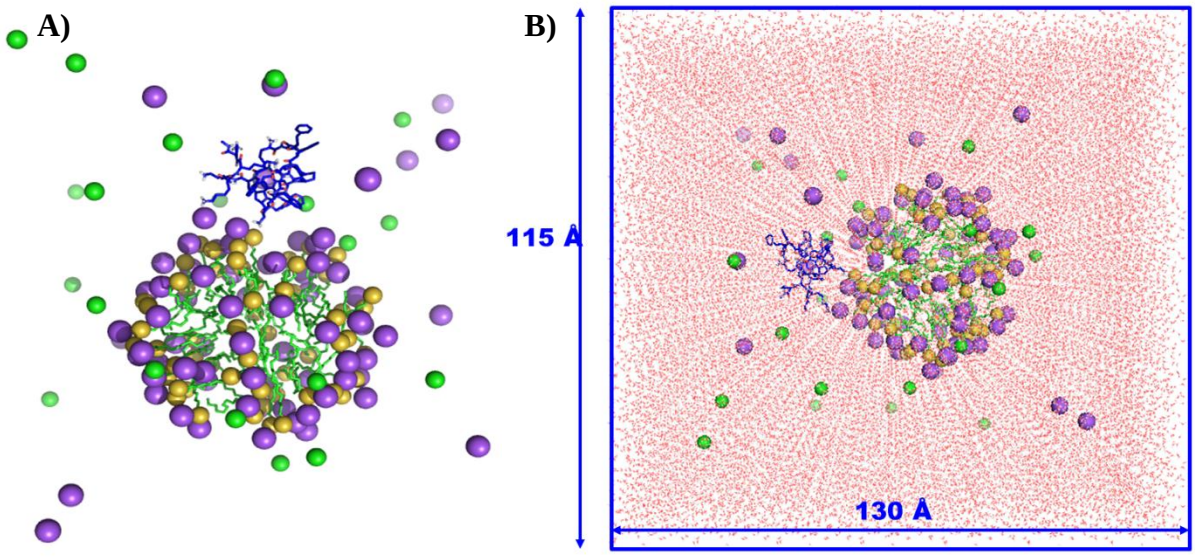


Figure 4.7 A) Starting sodium (purple) and chlorine (green) atoms. The sulfur atoms of SDS are noted with yellow spheres. Peptide **23** is shown in blue. B) The water box used in the periodic boundary simulation

angstroms away from the micelle surface.

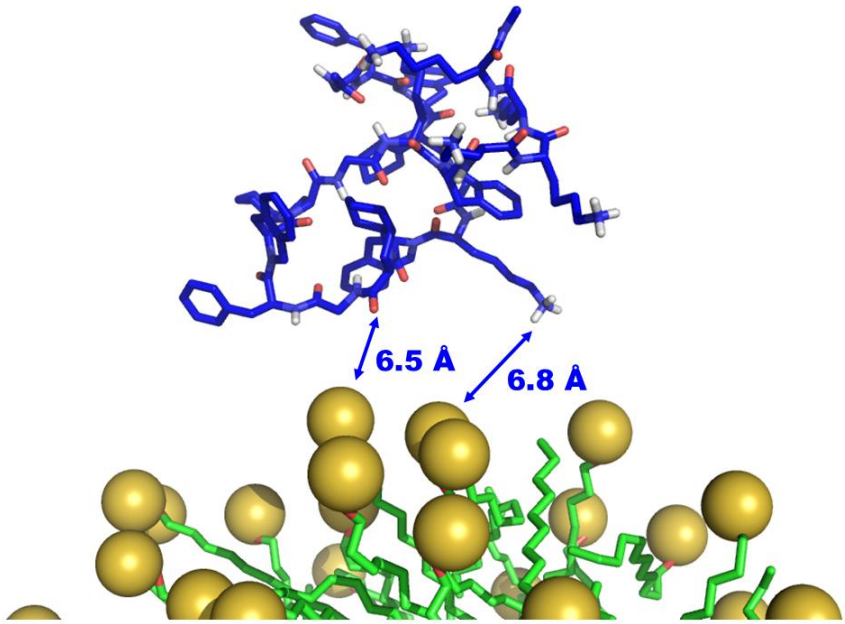


Figure 4.8 The sulfur atoms of SDS are noted with yellow spheres. Peptide **23** is shown in blue Distances noted.

Examining the electrostatic surface potential charge distribution of both peptide **23** and the 62 SDS micelle relative to one another (Figure 4.9) demonstrates the complementary placement of charged groups resulting in a favorable initial interaction profile. The highly electronegative micelle surface offers a powerful electrostatic pull relative to the electropositive peptide **23**. The net charge associated with peptide **23** is +6. The charge difference between the micelle and peptide **23** provides a powerful attractive force between them enabling the atomic interactions to be examined. The termini which would be in the ionized form at physiological pH has been artificially neutralized using an acetyl and methyl amine capping and neutralizing the N and C termini respectively.

The conclusion of the MD simulation produced the structures in Figure 4.10. Peptide **23** is shown with the N terminus interacting with the micelle likely through formation of

intermolecular backbone hydrogen bonds between the peptide and sulfate oxygen atoms of the

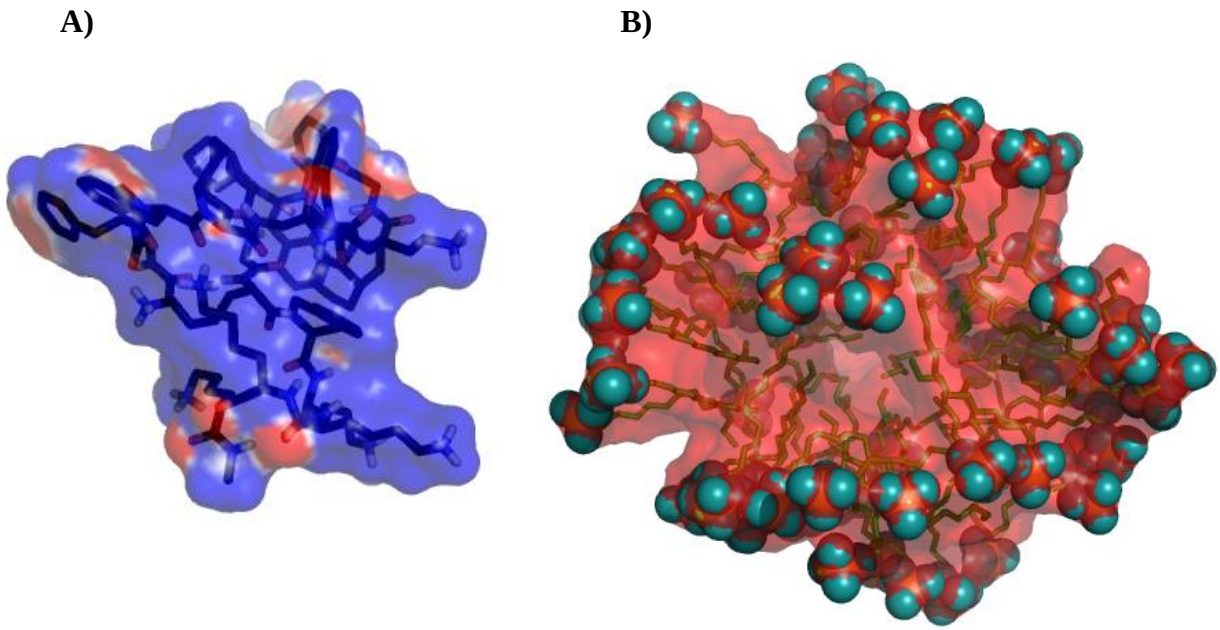


Figure 4.9 Electro positive surfaces of peptide **23** are shaded blue while electronegative surfaces are noted in red. The oxygen and sulfur atoms of SDS are noted with teal and yellow spheres respectively.

micelle. Four lysine residues at the C-terminus of peptide **23** are also interacting through a salt bridge interaction with the micelle. The shape of the micelle has departed from typically illustrated sphere. The polar head groups of the SDS molecules have been pushed away from the area of the micelle that the peptide is interacting with creating a trough in which the peptide resides.

Despite the prevalence of aliphatic carbon atoms of the SDS molecules exposed to the surface most of the aromatic rings of the peptide are found away from the SDS tails (Figure 4.10 B). The presence of the peptide on the micelle surface appears to disrupt the aggregation of the polar head groups (Figure 4.11). The SDS head groups would be expected to aggregate together

with the hydrophobic tails more tightly clustered in the center of the sphere enabling the aliphatic tails to be protected from the surrounding solvent.

Determining an endpoint for a MD simulation can be challenging and is influenced by numerous external factors. The most influential variable in this study has been the computational resources available. MD simulations with model bilayers significantly increase the number of atoms involved in the experiment. Any increase in the number of atoms will significantly increase the number of calculations required at each step of the simulation. The increase in the number of calculations needs to be met with a corresponding increase in

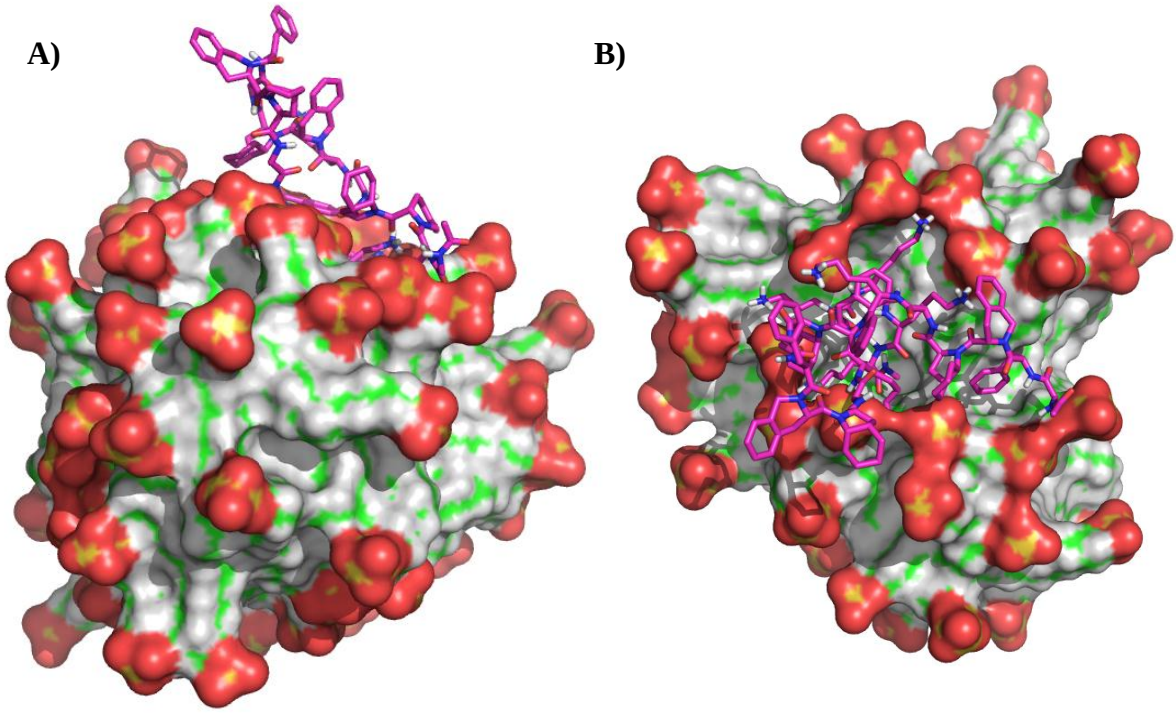


Figure 4.10 A. Final structure of 62 SDS molecules micelle with Peptide **23.** B. The top down perspective of peptide **23** with the 62 SDS micelle.

computing power in order to maintain the rate of data collection. The lack of a large computer cluster with good communication (threading) between the nodes of each processor have limited the rate at which these experiments have progressed.

Another consideration frequently employed in the analysis of MD simulations of peptide systems is the deviation which has occurred from a set point. The Root Mean-Square deviation (RMSD) measures the differences between values in the predicted model, in this instance the unit would be angstroms. The reference structure is the minimized system prior to the production run. The RMSD plot for peptide **23** with the micelle can be found in Figure 4.12.

The production run consisted of 25 ns of constant volume dynamics. Based on the RMSD plot (Figure 4.12) peptide **23** stabalized almost immediately. The RMSD is measured as the deviation occuring from the minimized structure of the system. The equilibrated state is evidenced by the average slope of the RMSD plots approaching 0, see Figure 4.12. Stabilization of the system does not mean the system is static, lipids in the micelle are expected to have continual molecule movement. The peptide achieves equilibration rather quickly, relative to the micelle. Visualization of the trajectory shows that peptide approaching the micelle surface with the C terminus becoming locked into the micelle. Hydrogen bonds are formed between the sulfonal oxygen and the amino groups of the lysine residues. Formation of these interactions appears to prevent the peptide from undergoing large structural changes.

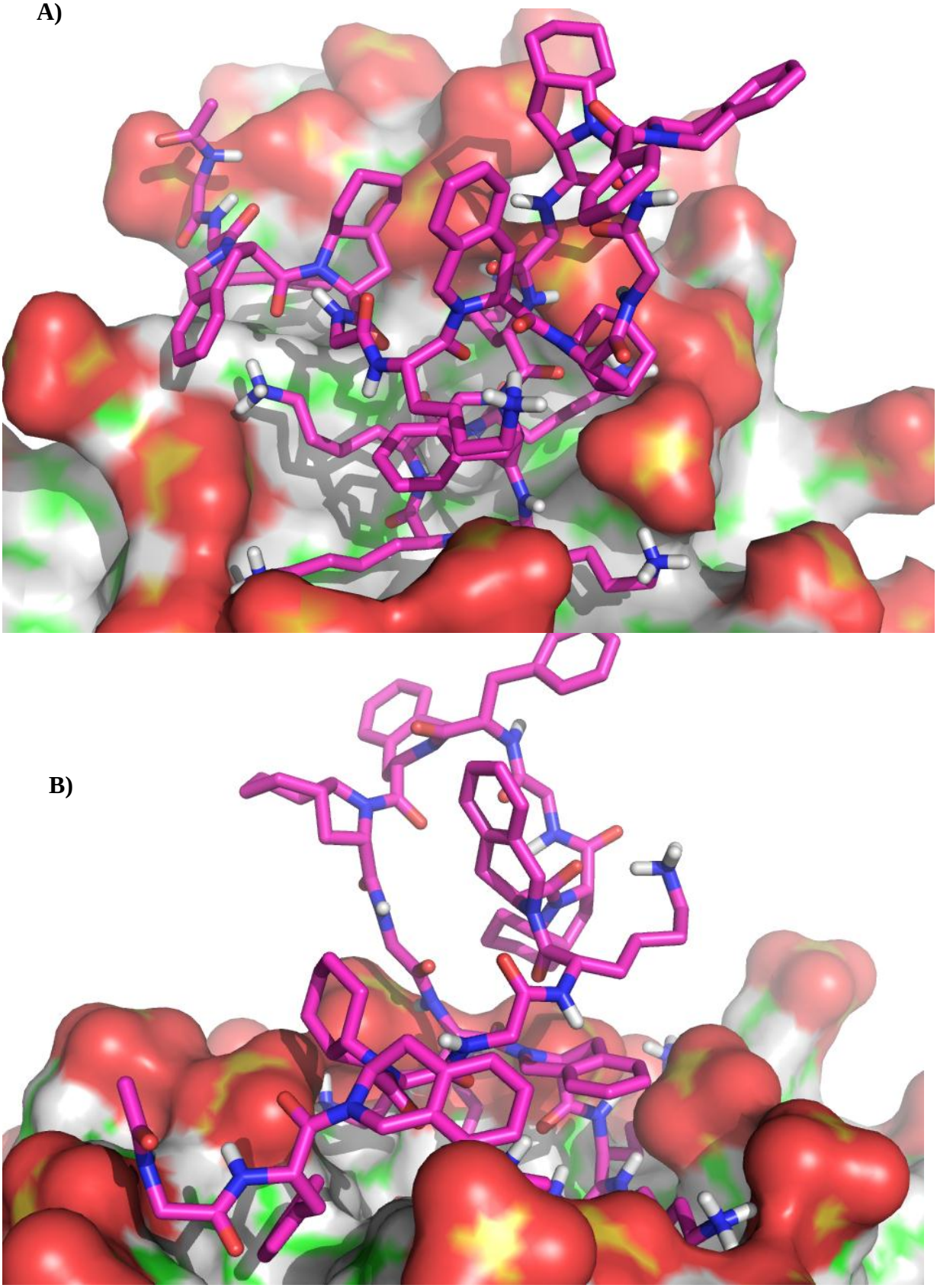


Figure 4.11 A & B End structure of 62 SDS molecules micelle with Peptide **23**

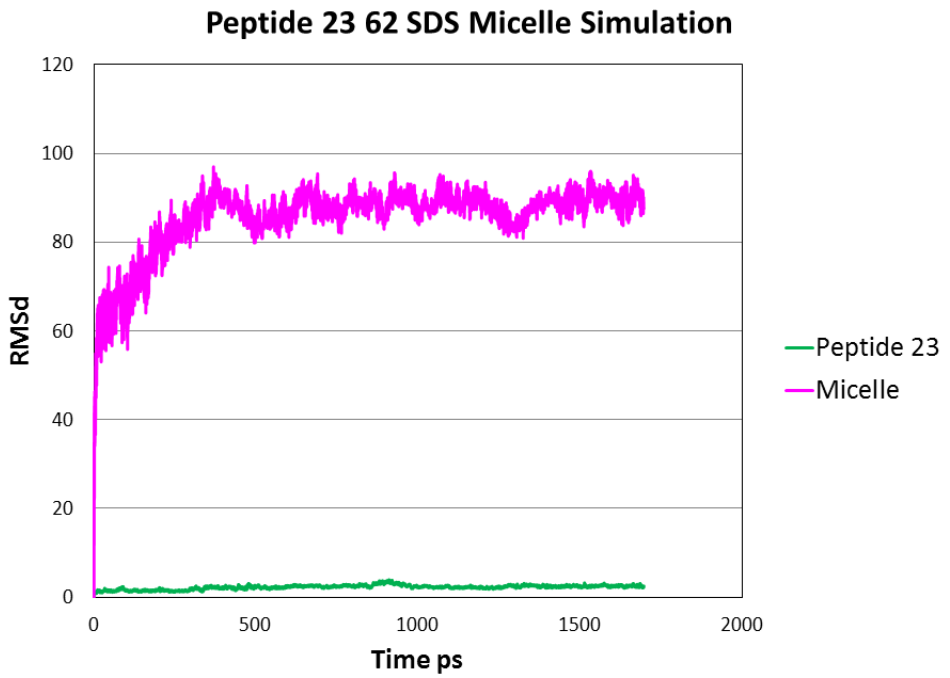


Figure 4.12 RMSD plot for Peptide **23** with the 62 SDS molecule micelle

4.3.2 PEPTIDE 36 WITH MICELLE

Peptide **36** was simulated with a 62 SDS molecule micelle for 25 ns. The same micelle as was used as the starting structure in the control micelle simulation was used as the starting structure for the peptide/micelle simulations. An initial box was placed extending 5 angstroms from the surface of the micelle. The box was solvated using the TIP3P water model. Water molecules within 5 angstroms of a sulfur atom were selected, while remaining water molecules were removed. The Multiscale Modeling Tools for Structural Biology (MMTSB)[14] software was utilized to randomly replace the water molecules with 62 sodium atoms to neutralize the surface charge of the micelle, remaining water molecules were removed. The coordinates were returned to tLeAP and a water box extending 18 angstroms from the micelle surface was added. The

ionic concentration of sodium and chlorine atoms was adjusted to 0.15 mM NaCl through random replacement of water molecules using MMTSB[14] software (Figure 4.13). Remaining water molecules were once again removed prior to adding the previous water box, extending 18 angstroms from the micelle surface (Figure 4.13). The starting position of peptide relative to the surface was previously investigated using simulations with each of four "faces" of the peptide orientated toward the micelle surface (data not shown). While one of the cationic lysine residues was placed relatively close to the micelle surface (Figure 4.14) other nonpolar residues including Tic, Oic, and Phe were in comparable proximity with no portion of the peptide less than 6 angstroms away from the micelle surface.

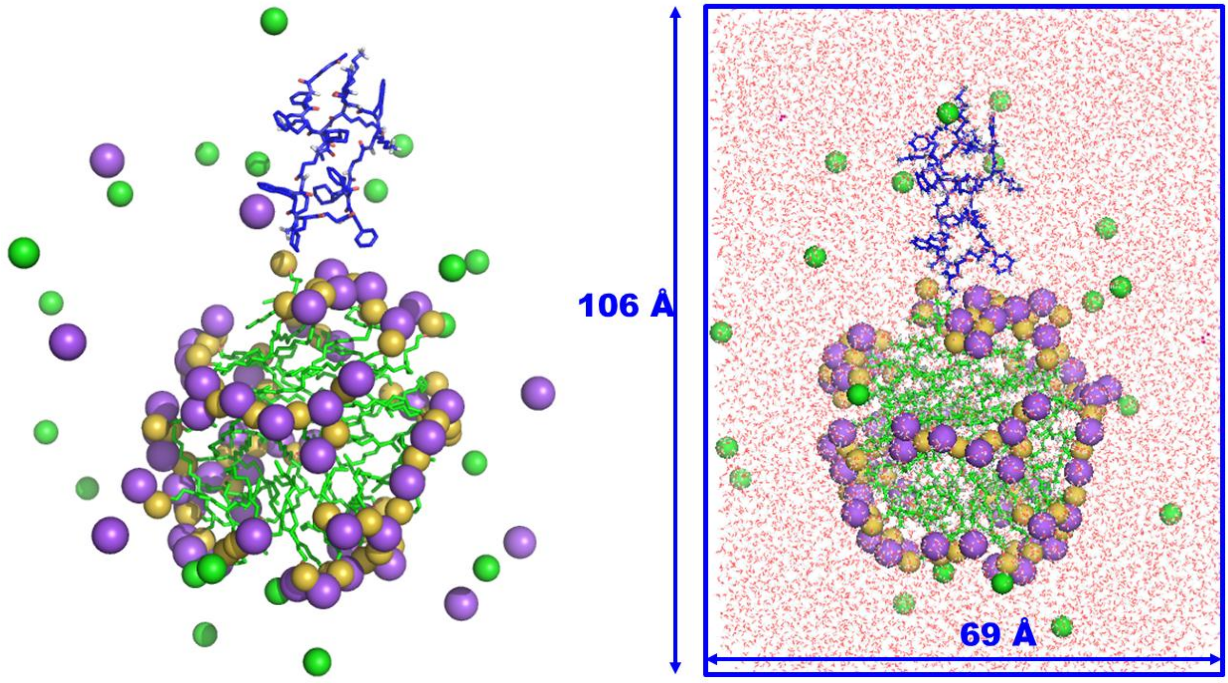


Figure 4.13 Electro positive surfaces are shaded blue while electronegative surfaces are noted in red. The oxygen and sulfur atoms of SDS are noted with red and yellow spheres respectively. The z direction is also 69 angstroms in length.

The electrostatic surface potential was calculated for peptide **36** and the micelle. The charge of peptide 36 is +6 with Gly substituted for β-Ala (Figure 4.15). Electronegative regions

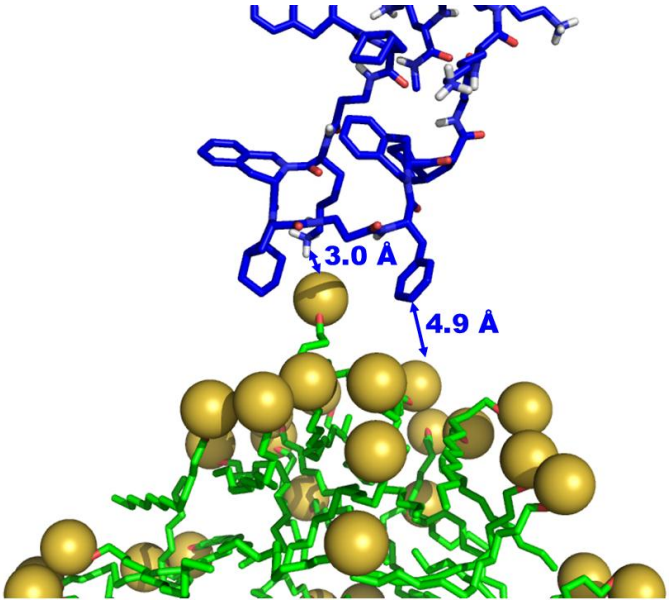


Figure 4.14 The sulfur atoms of SDS are noted with yellow spheres. Peptide **36** is shown in blue Distances noted

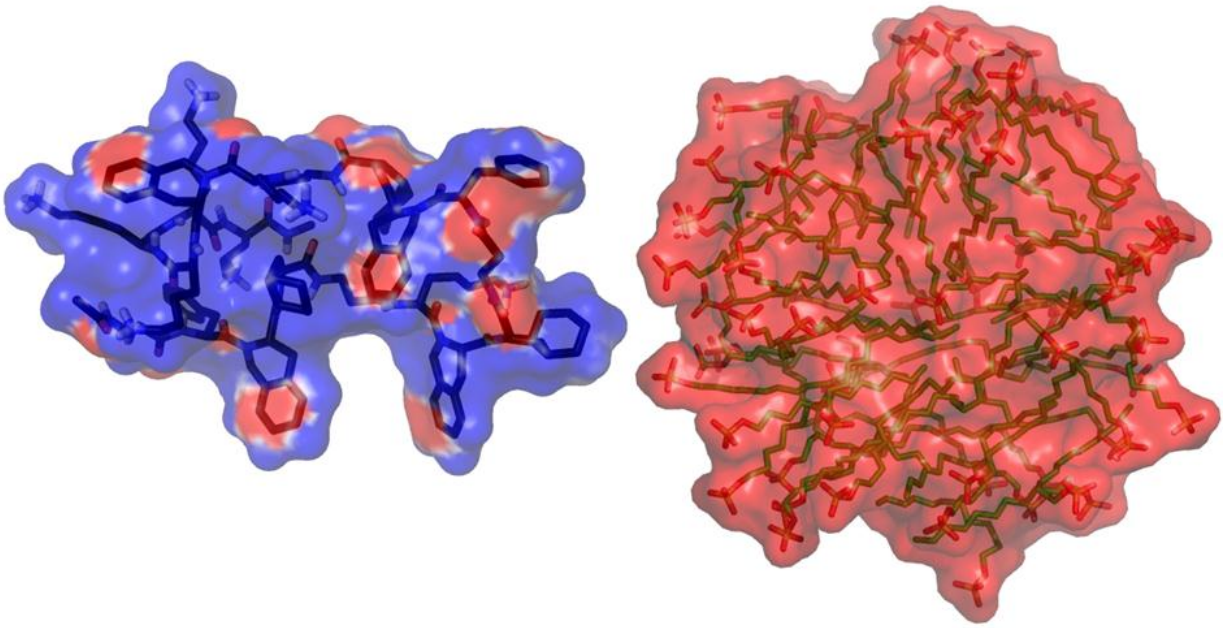


Figure 4.15 Electropositive surfaces of peptide **36** are shaded blue while electronegative surfaces are noted in red. The oxygen and sulfur atoms of SDS are noted with red and yellow spheres respectively.

are shown in red and electropositive regions are shown in blue.

The RMSD was calculated for the peptide and the micelle relative to the minimized structure prior to the production run (Figure 4.17). Both the peptide and the micelle stabilize within the 42 ns production run. The RMSD plot of peptide **36** in Figure 4.17 shows a slight deviation occurring around 30 ns into the simulation. The structure of peptide **36** after 15 ns is shown if Figure 4.18. The two termini are noted with yellow arrows. The peptide has adopted a "u" shape with the two termini 8.3 angstroms apart.

Comparing the structure of the peptide after the change at 30 ns (Figure 4.19) to the 15 ns structure (Figure 4.18) shows subtle movement of the termini away from one another. Peptide **36** was simulated with the 62 SDS molecule micelle until a 42 ns trajectory had been attained. The micelle is perturbed (Figure 4.19) from the starting structure (Figure 4.14).

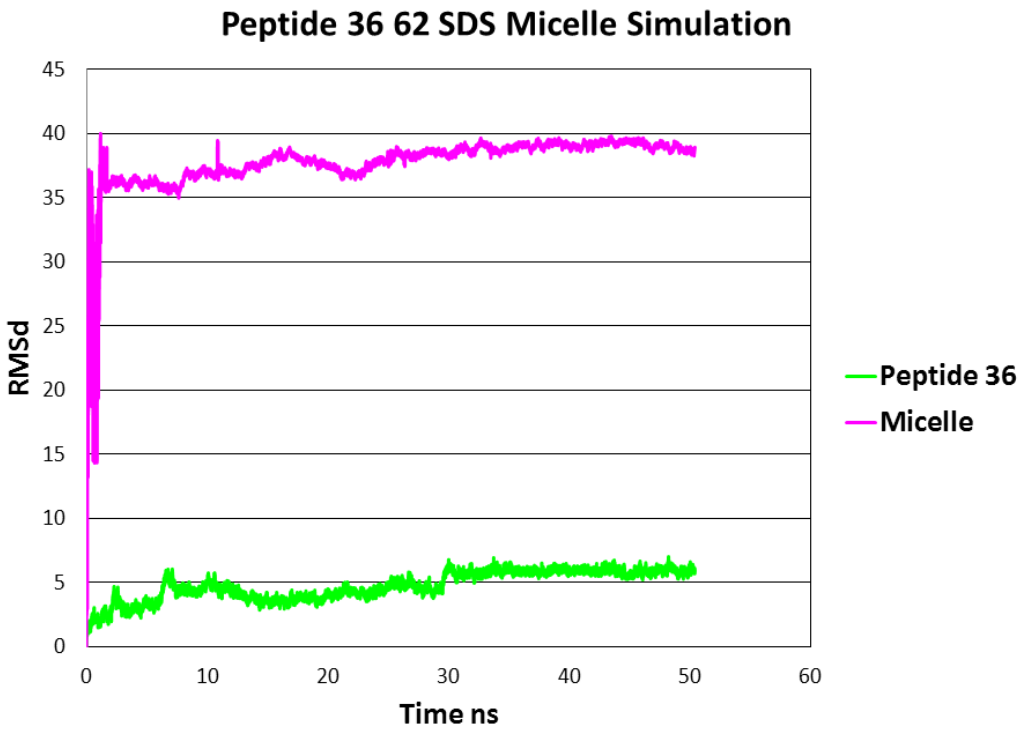


Figure 4.16 RMSD plot for Peptide **36** with the 62 SDS molecule micelle

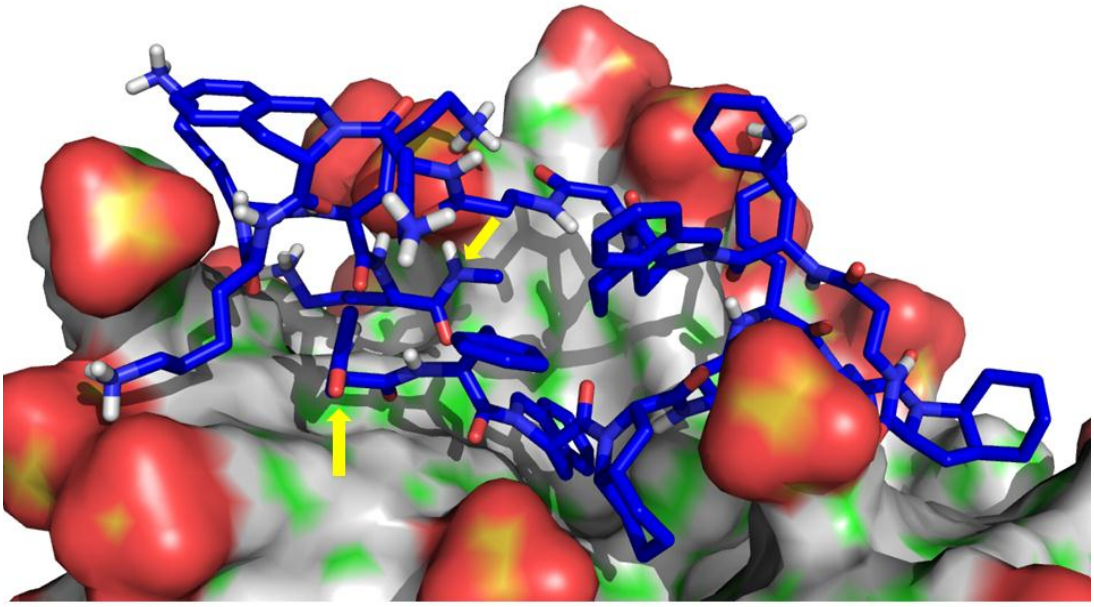


Figure 4.17 Peptide **36** with the 62 SDS molecule micelle. The N and C terminus are highlighted with yellow arrows.

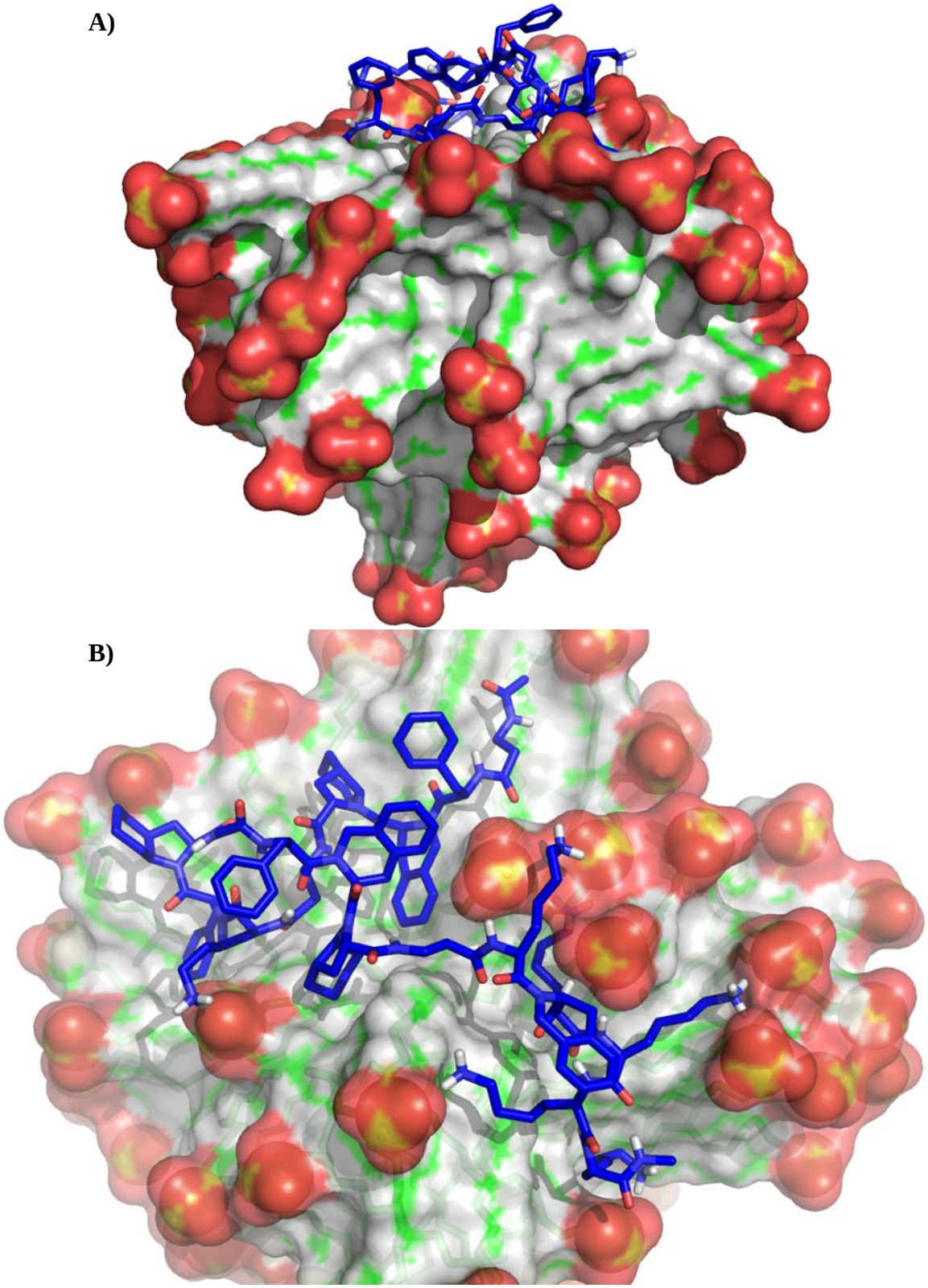


Figure 4.18 Peptide **36** at the conclusion of the 42 ns production md simulation. A. peptide **36** is shown to be well embedded into the micelle surface relative to the sulfate head groups of the SDS. B. Peptide **36** is shown to have moved the polar head groups of the micelle enabling hydrophobic rings to be protected from the surrounding solvent.

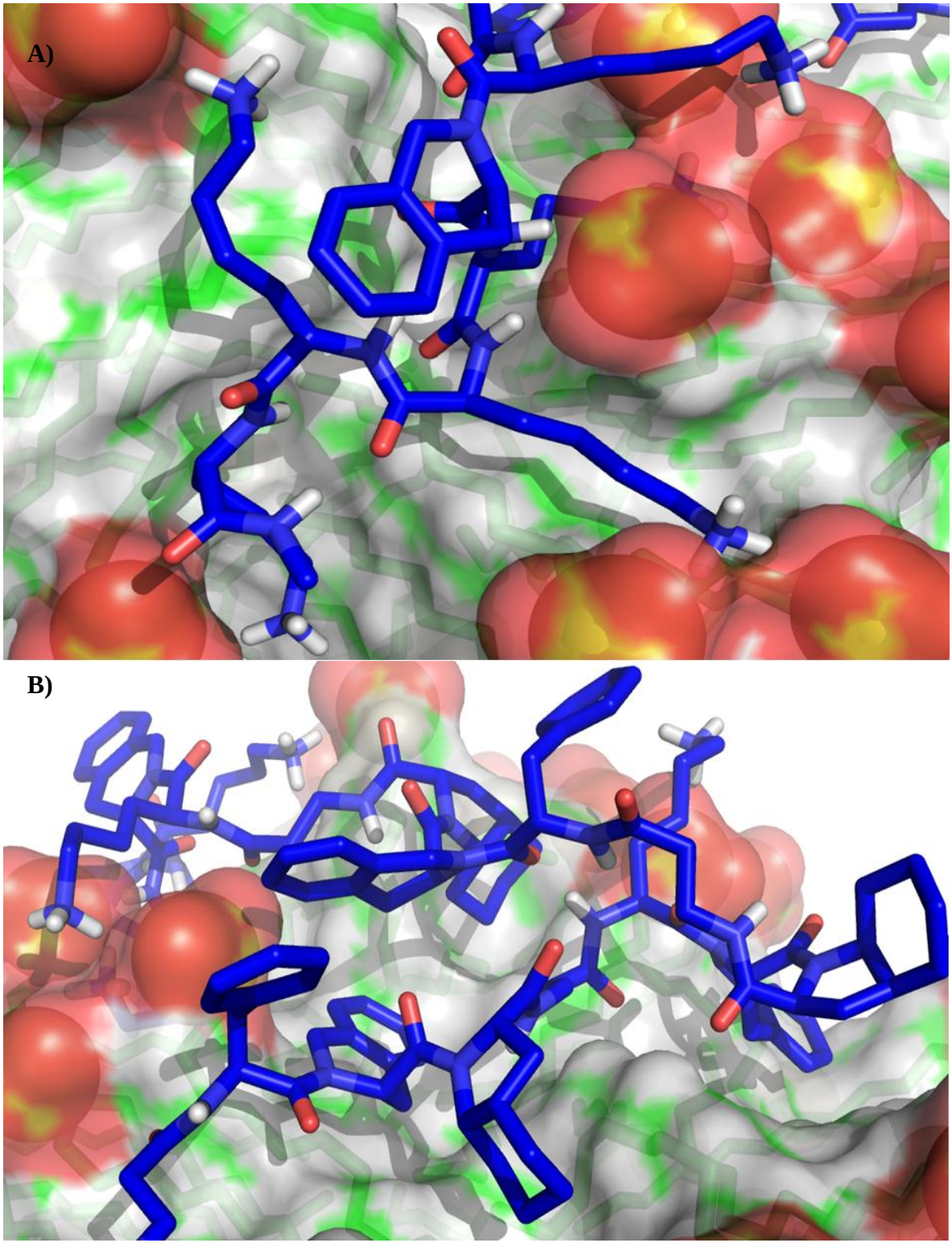


Figure 4.19 Peptide **36** at the conclusion of the 42 ns production md simulation. A. Peptide **36** is shown to be well embedded into the micelle surface relative to the sulfate head groups of the SDS. B. Peptide **36** is shown to have moved the polar head groups of the micelle enabling hydrophobic rings to be protected from the surrounding solvent.

4.4 DISCUSSION

Contrary to the idealized image expected for a micelle, with all the polar head groups exposed to the solvent surroundings and the hydrophobic, aliphatic tails all clustered safely beneath the solvent surface, a much more dynamic perspective is detected during experiments. The regions of exposed hydrophobic tails have been documented in the literature and supported with experimental data.[21]

The smaller molecules forming a micelle in this case result in a smaller number of atoms which should correlate to a shorter simulation time. The electrostatic model of a membrane formed by a micelle has limitations. The smaller aggregation size results in a smaller radius and an increased curvature modeling the membrane surface. A correlation between the spontaneous curvature of the electrostatic model and pore formation has been established.[22] The small size also limits the number of peptides which can be expected to interact with a single micelle. Due to the limited interaction possible between the peptide and the micelle the data does not directly correlate to in vitro experiments. Formation of peptide aggregates[22] which are requisite for bilayer disruption is not possible with such a small micelle. The small size is reflected by the lack of micelle disruption observed in experimental studies.[23] The micelle simulations have enabled the examination the surface interaction expected from the peptide suggesting appropriate peptide orientation for further study with larger model membrane systems.

The Ramachandran plots in Figure 4.20 reflect the average phi and psi angles for each of the noted residues over the last 2 ns of the production simulation from the micelle simulations. Calculations were performed using the same set of computer scripts mentioned in Chapter 3 and included in **Appendix B**. The preferred region for β-sheets is outlined in red while the allowed

region is outlined in green. Here no distinction is made between parallel and antiparallel secondary structures; they both fall within the noted region. The preferred region for right handed alpha helices is noted in magenta with the allowed region outlined in blue. The region outlined in purple is associated with left handed alpha helices. The region outlined in black has been noted to be a possible allowed region if a reduced radius hard sphere approach is utilized.[24]

The unique conformations adopted by each of the peptides during the course of the production run would suggest very different secondary structures. The averaged Ramachandran plots for the two peptides provide a different perspective. Peptide **23** has 8-9 residues found in the β-sheet preferred region and 6 residues in the right handed α-helical preferred region. The majority of the residues found outside the preferred region are glycine residues. Glycine residues due to their small size and lack of chirality have been found to adopt a diverse collection of angles.[25] The distribution of the peptides between the β-sheet and right handed α helical regions is consistent with the NMR data for peptide **23**.[26]

Ramachandran plots were also created for Peptide **36**; averaging the final 2 ns of the production run for each of the residues. Peptide **36** includes the unnatural amino acid β-alanine which prohibits the angle measurements due to the beta carbon's presence. Angle measurements for β-alanine have been excluded from the analysis. Eight residues were found in the β-sheet preferred region for peptide **36** with one lysine residue slightly falling outside of the allowed region. One residue was found in the right handed alpha helical allowed region with four

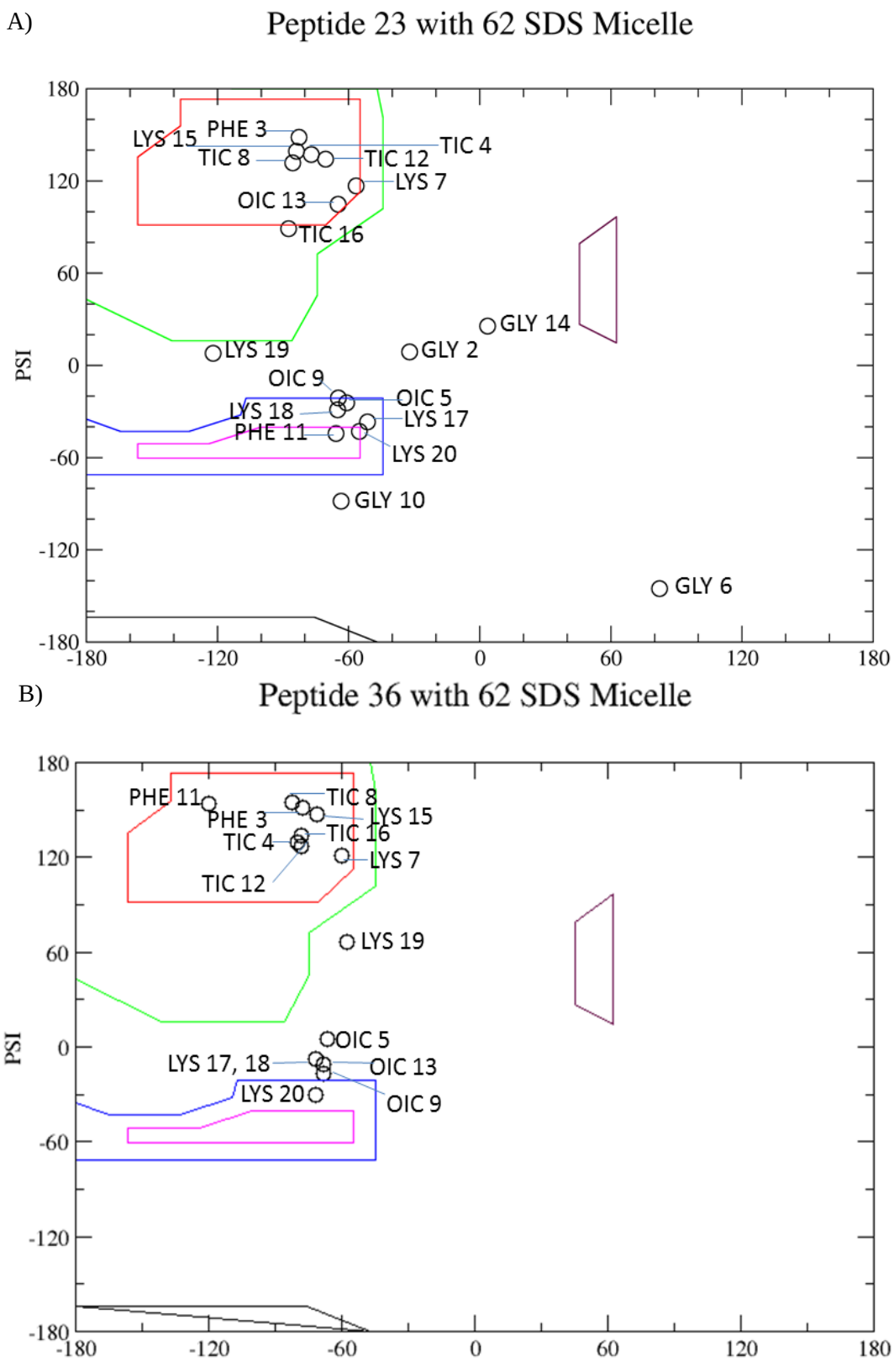


Figure 4.20 Ramachandran plots of the average residue angles from the last 2ns of the production run with the canonical ensemble. A) Peptide **23** B) Peptide **36**

residues very close to the right handed alpha helical allowed region.

The area of the micelles at the conclusion of the simulation was compared to the starting structure. Area and volume calculations were performed with the UCSF Chimera package. Chimera is developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS 9P41GM103311).[27] The starting micelle, at the end of the minimization, before peptides or the second solvent box and ions were added, had a volume of 2.270×10^4 Å^3 and an area of 8827 Å^2. The control micelle had a final volume of 2.250×10^4 Å^3 and an area of 8585 Å^2 which suggests that while the volume stays relatively constant the surface area changes. Variability in the area of a micelle is easily understood given the dynamic nature of this system and the speed which the individual SDS molecules move. The micelle simulated with peptide **23** had a final volume of 2.101×10^4 Å^3 and an area of 1.031×10^4 Å^2. The micelle simulated with peptide 36 had an area of 9910 Å^3. The micelle simulated with peptide **36** (Figure 4.16 A) was very deformed at the conclusion of the simulation inhibiting identification of a shape for which the volume could be calculated. The software builds surfaces using the solvent-accessible surface area. The addition of peptide to the micelle has the effect of increasing the area of the micelle.

The following discussion has been submitted for publication.[26] "Despite the lack of stabilization, animation of the trajectory has provided evidence of attraction between the SDS head groups and the Lys side chains of peptide **23**. At the conclusion of the simulation the N and C termini of compound **23** appear to bind to the polar surface of the micelle thus, allowing the positively charged Lys side chains to interact with the negatively polarized oxygen atoms of the SDS micelle. This data is consistent with NMR data.[26] The resulting binding orientation has

most of the aromatic and aliphatic rings of the peptide being exposed to solvent instead of being inserted into the hydrophobic core of the micelle. The micelle has adopted a less ridged configuration which enables some of the hydrophobic tails to be exposed to the solvent surface. This structure, with the larger hydrodynamic radius, is consistent with the observed decrease in the diffusion coefficient of the micelle after binding with peptide **23**. The polar head groups which are able to form hydrogen bonds with the bulk solvent water have deformed from the cartoon spherical shape commonly associated with a micelle. *In vitro* and *in silico* experiments support the dynamic nature of SDS micelles. Note the increased density of SDS head groups in the region of the micelle binding with the positively charged regions of the peptide. The side chains of the Lys residues of peptide **23** are embedded within the surface of the micelle (Figure 4.21). From this orientation it appears that the electrostatic interactions between the positively charged Lys side chains and the negative charged polar head group of the SDS micelle predominate during the binding process. The location of the aromatic side chains suggests that hydrophobic interactions between the peptide and the hydrophobic core of the SDS micelle do not play a major role in defining the binding process. This observation is consistent with the aromatic region of the 1D ^{1}H spectrum of peptide **23** bound to SDS micelles which shows similarity to the 1D ^{1}H spectrum of peptide **23** in the presence of buffer see Figure 4.21. Clearly, there are differences between the two spectra suggesting some interaction between the aromatic protons of peptide **23** (specifically the Phe residue located at the N-terminus and the Tic residue located before the Lys chain at the C-terminus) with the surface or core of the SDS micelle, however this interaction is not as pronounced as that observed in the presence of DPC micelles.

The NMR spectra for peptide **36** with the SDS micelles is shown in Figure 4.22. The peptide is inserted into the hydrophobic core of the micelle with the hydrophobic residues of the

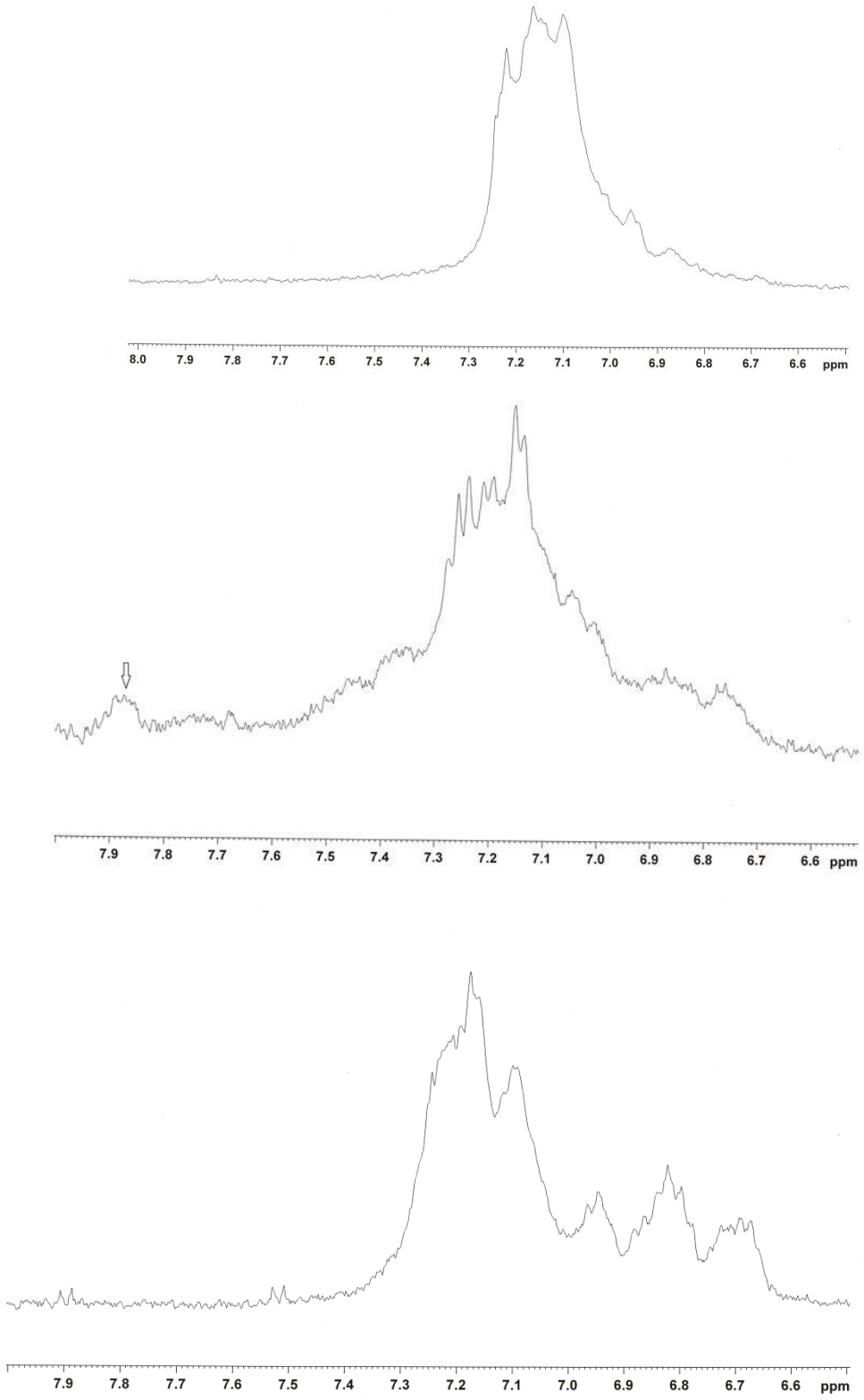


Figure 4.21 400 MHz ^{1}H spectra of peptide **23** in top) 150 mM sodium acetate buffer, pH=4.28 in D_2O, middle) 100 mM SDS micelle in 150 mM sodium acetate buffer, pH=4.28 in D_2O, bottom) 100 mM DPC micelle in 150 mM sodium acetate buffer, pH=4.28 in D_2O[26]

peptide predominately oriented toward the hydrophobic core of the micelle. The interaction of the aromatic side chains of peptide **36** are clearly indicated by the changes in the 1D ^{1}H spectrum see Figure 4.22.

The results of the MD simulation shown in Figure 4.23 indicates that the N-terminal Phe residue (shown in cyan) and the Tic residue preceding the four C terminal Lys residues (shown in yellow) are located very close to both the polar head groups and the hydrophobic core of the micelle."

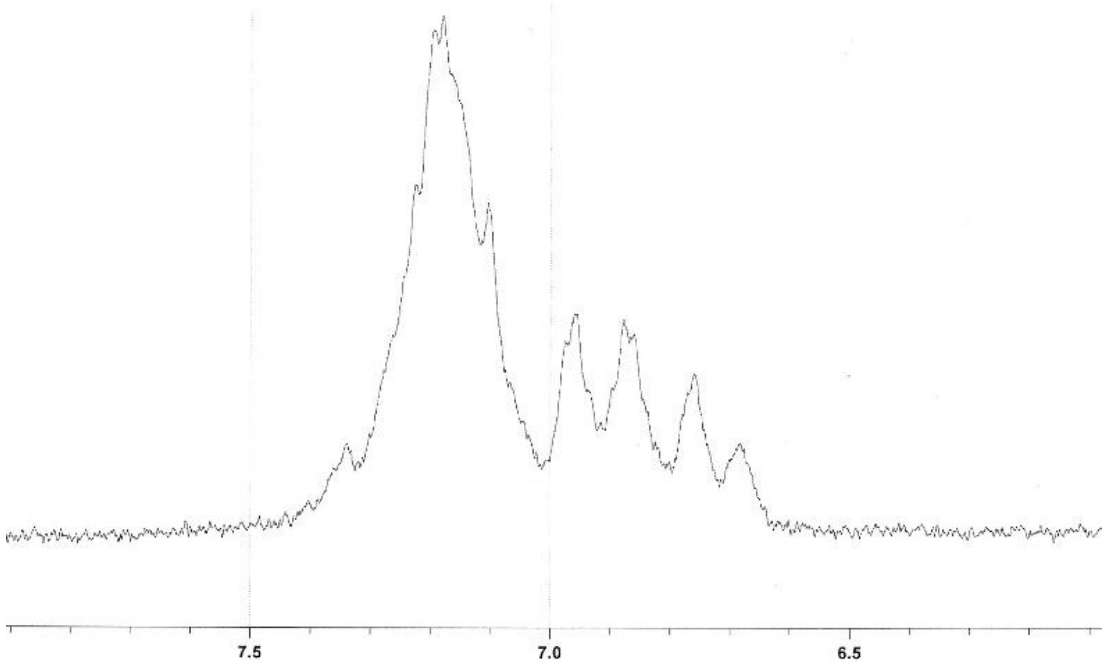


Figure 4.22 400 MHz ^{1}H spectra of peptide **36** in top) 100 mM SDS micelle in 150 mM sodium acetate buffer, pH=4.28 in D_2O, bottom) 100 mM DPC micelle in 150 mM sodium acetate buffer, pH=4.28 in D_2O[26]

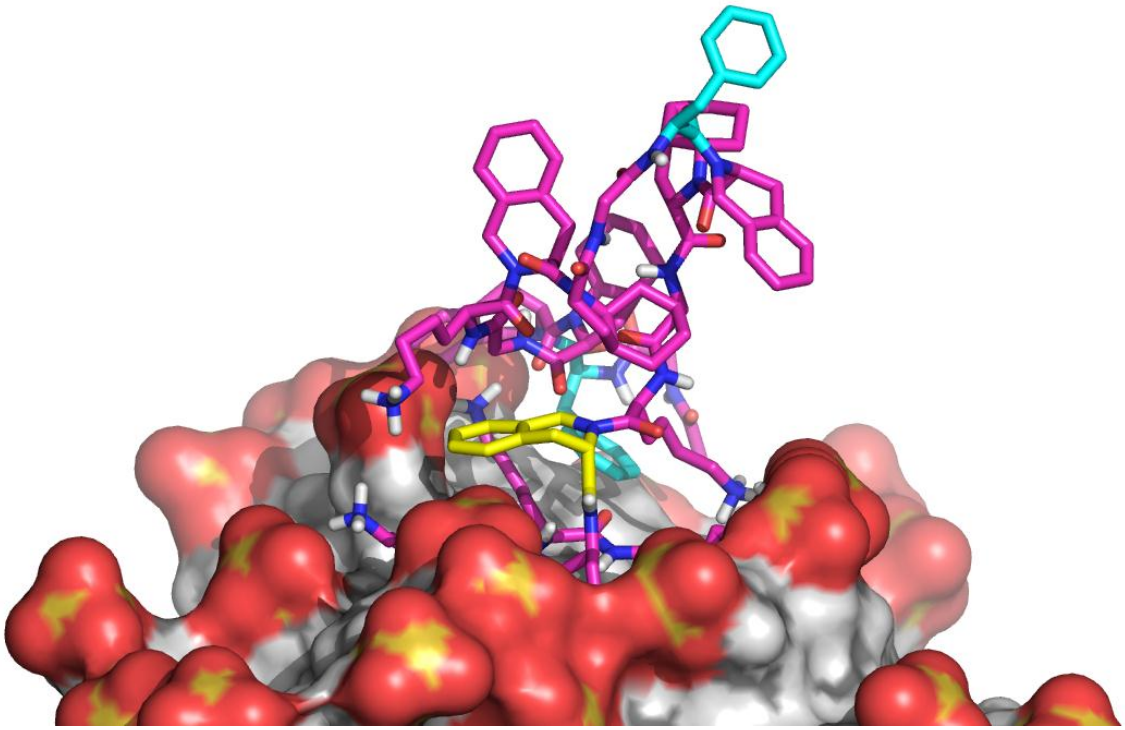


Figure 4.23 The results of the MD simulation for compound **23** indicates that the N-terminal Phe residue (shown in cyan) and the Tic residue preceding the four C-terminal Lys residues (shown in yellow) are located very close to both the polar head groups and the hydrophobic core of the micelle.[26]

4.5 CONCLUSION

The experimentally derived CD spectroscopy suggests that the peptides adopt a predominately alpha helical structure resulting from binding to SDS micelles. The differences observed during the course of the MD simulations reflect the differing secondary structures observed between peptide **23** and peptide **36**. Experimental data supports a different secondary structure when the same peptides are in the presence of a DPC micelle (zwitterionic) rather than the anionic SDS micelle. The resulting secondary structure in the presence of the zwitterionic membrane model suggests a mixture of both α-helical and beta-turn or sheet like structures being adopted. The Ramachandran plot derived from the average phi and psi angles, over the last 2 ns of the production run of peptide **23** and **36** provide support for the secondary structures with both

alpha helical and beta sheet/turn character as was suggested by the experimental NMR data.[26] The terminal structure resulting from the MD simulation peptide **23** adopts a structure which is a mixture of α-helical and β-turn/sheet characteristics. The terminal structure from the MD simulation of peptide **36** reflects a more extended structure interacting with the SDS micelle.

MD simulation data and experimental CD, DOSY NMR, and ^{1}H NMR is reflective of very different interactions (between the two peptides) with the SDS micelle. The electrostatic surface potential maps suggest this interaction is prominent in the binding mechanism of peptide **23** with the SDS micelle. Hydrophobic interactions provide a minor contribution to this interaction. In the simulation with peptide **23** many hydrophobic residues (excluding a phenylalanine and a TIC residue) are exposed to the surrounding solvent.

4.6 FURTHER RESEARCH

The *in silico* investigation into the AMP/micelle interactions can be continued to further understand the first step in peptide-lipid binding prefacing membrane disruption. Study of AMPs can be extended by including a mammalian micelle model using dodecylphosphocholine (DPC) as a zwitterionic mammalian membrane model (Figure 4.24). DPC micelles with AMPs have demonstrated unique experimental structural differences using CD spectroscopy as well as DOSY NMR experiments.[26] Observing the atomic level interactions and comparing AMP interactions with both SDS and DPC micelles can provide a more complete perspective of the initial steps involved in membrane disruption.

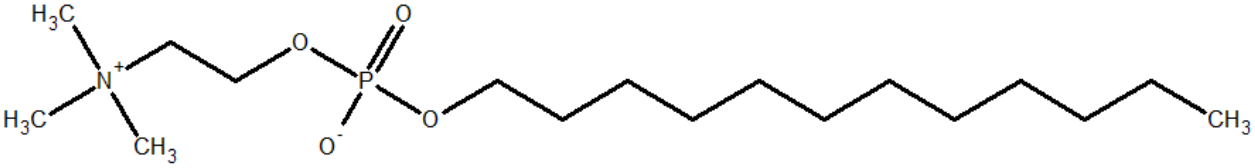


Figure 4.24 Structure of zwitterionic dodecylphosphocholine, DPC.

4.7 REFERENCES

1. Palazzesi, F., Calvaresi, M., and Zerbetto, F. *A molecular dynamics investigation of structure and dynamics of SDS and SDBS micelles.* The Royal Society of Chemistry.

2. Hicks, R. P., Mones, E., Kim, H., Koser, B. W., Nichols, D. A., and Bhattacharjee, A. K. (2003) Comparison of the conformation and electrostatic surface properties of magainin peptides bound to sodium dodecyl sulfate and dodecylphosphocholine micelles. *Biopolymers. 68*, 459-470.

3. Khandelia, H., Langham, A. A., and Kaznessis, Y. N. (2006) Driving engineering of novel antimicrobial peptides from simulations of peptide–micelle interactions. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1758*, 1224-1234.

4. Khandelia, H., and Kaznessis, Y. N. (2005) Molecular dynamics simulations of helical antimicrobial peptides in SDS micelles: What do point mutations achieve? *Peptides. 26*, 2037-2049.

5. Khandelia, H., and Kaznessis, Y. N. (2006) Molecular dynamics investigation of the influence of anionic and zwitterionic interfaces on antimicrobial peptides' structure: Implications for peptide toxicity and activity. *Peptides. 27*, 1192-1200.

6. Fuzo, C. A., Castro, J. R. M., and Degreve, L. (2008) Study of the antimicrobial peptide indolicidin and a mutant in micelle medium by molecular dynamics simulation. *GMR. 7*, 986-999.

7. Wymore, T., Gao, X. F., and Wong, T. C. (1999) Molecular dynamics simulation of the structure and dynamics of a dodecylphosphocholine micelle in aqueous solution. *J. Mol. Struct. 485-486*, 195-210.

8. Parker, W., and Song, P. S. (1992) *Protein structures in SDS micelle-protein complexes.* Cell Press.

9. Bartolotti, L. (2010) *Micelle script for InsightII.*

10. Russell, A. L., Kennedy, A. M., Spuches, A. M., Venugopal, D., Bhonsle, J. B., and Hicks, R. P. (2010) Spectroscopic and thermodynamic evidence for antimicrobial peptide membrane selectivity. *Chem. Phys. Lipids. 163*, 488-497.

11. Tieleman, D. P., van der Spoel, D., and Berendsen, H. J. C. (2000) Molecular dynamics simulations of dodecylphosphocholine micelles at three different aggregate sizes: Micellar structure and chain relaxation. *The Journal of Physical Chemistry B. 104*, 6380-6388.

12. Bruce, C. D., Berkowitz, M. L., Perera, L., and Forbes, M. D. E. (2002) Molecular dynamics simulation of sodium dodecyl sulfate micelle in water: micellar structural characteristics and counterion distribution. *The Journal of Physical Chemistry B. 106*, 3788-3793.

13. MacKerell, A. D. (1995) Molecular dynamics simulation analysis of a sodium dodecyl sulfate micelle in aqueous solution: Decreased fluidity of the micelle hydrocarbon interior. *J. Phys. Chem. 99*, 1846-1855.

14. Feig, M., Karanicolas, J., and Brooks III, C. L. (2004) MMTSB tool set: Enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model. 22*, 377-395.

15. Rakitin, A. R., and Pack, G. R. (2004) Molecular dynamics simulations of ionic interactions with dodecyl sulfate micelles. *The Journal of Physical Chemistry B. 108*, 2712-2716.

16. Miyamoto S and Kollman, P. A. (1992) *Settle-an analytical version of the shake and rattle algorithm for rigid water models.*

17. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER 9 users' manual University of California, San francisco.*

18. Sammalkorpi, M., Karttunen, M., and Haataja, M. (2007) Structural properties of ionic detergent aggregates: A large-scale molecular dynamics study of sodium dodecyl sulfate. *J Phys Chem B. 111,* 11722-11733.

19. Tsui, V., and - Case, D. A. (2000) *Theory and applications of the generalized born solvation model in macromolecular simulations.* John Wiley & Sons, Inc.

20. MacKerell, A. D. (2004) *Empirical force fields for biological macromolecules: Overview and issues.*

21. Tummala, N. R., and Striolo, A. (2008) Role of counter ion condensation in the self-assembly of SDS surfactants at the Water Graphite interface. *J Phys Chem B. 112,* 1987-2000.

22. Lee, M., Hung, W., Chen, F., and Huang, H. W. (2005) *Many-body effect of antimicrobial peptides: On the correlation between lipids spontaneous curvature and pore formation.* Cell Press.

23. Russell, A. L., Williams, B. C., Spuches, A., Klapper, D., Srouji, A. H., and Hicks, R. P. (2012) The effect of the length and flexibility of the side chain of basic amino acids on the binding of antimicrobial peptides to zwitterionic and anionic membrane model systems. *Bioorg. Med. Chem. 20,* 1723-1739.

24. Wong, M. (2006) *Classical potential energy methods.* www.hydra.vcp.monash.edu.au/modules/mod6/classpot.html ed.

25. Beck, D. A. C., Alonso, D. O. V., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences. 105,* 12259-12264.

26. Clark, T. D., Bartolotti, L., Russell, A. L., and Hicks, R. P. (2012) The application of NMR experiments and molecular dynamics simulations to investigate the binding interactions that occur between anionic and zwitterionic micelles and antimicrobial peptides containing unnatural amino acids. *Biopolymers. Submitted for publication*.

27. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) *J Comput Chem. Oct., 25 (13)*, 1606.

CHAPTER 5
BILAYER SIMULATIONS

ABSTRACT

A collection of molecular dynamics (MD) simulations were undertaken to study the behavior of unnatural amino acid containing antimicrobial peptides (AMPs) AMBER9 was used to simulate two model bilayers. The first were mixed bilayers with a 4:1 ratio of POPC:POPG creating an anionic membrane. The second was a confluent bilayer comprised of the zwitterionic POPC lipid. Peptide **23** which has a glycine residue linking the Tic Oic dipeptide units was added to each membrane environment. Over the course of the simulations membrane thinning was observed in the mixed bilayer in the presence of peptide **23**. Thinning was not observed in the simulation of peptide **23** with the control bilayer. Control simulations of each bilayer without peptides maintained the membrane thickness throughout the course of the simulation. The control simulations lend support to the selective membrane disruptive capabilities of peptide **23** in the presence of a mixed bilayer.

5.1 INTRODUCTION

The mechanism of membrane disruption for naturally occurring AMPs has been debated within the literature for over two decades.[1-6] The atomic events prefacing the disruption of the bilayers are poorly understood. The inclusion of unnatural amino acids increases the confusion with regard to how the membranes are disrupted due to the lack of accurate secondary structure information. Simulation of the AMPs seeks to understand the mechanism through which membranes become compromised. Other groups have used MD simulations to study the interaction of AMPs, which include unnatural amino acids with membranes[7]

Several online portals[8-10] are available to generate bilayers for use with other software programs, primarily the CHARMM (Chemistry at HARvard Macromolecular Mechanics)[11,12] and

GROMACS (GROningen MAchine for Chemical Simulations).[12,13] The large number of atoms composing bilayers requires substantial computational resources to simulate. The computational expense has necessitated the development of techniques to improve the efficiency of the simulation. Merging the hydrogen atoms bound to heavy atoms decreases the number of atoms which need to be considered and thus speeds up the simulation. Another technique used to speed the dynamics of a simulation is through the use of a coarse grain (CG) model. The CG model creates molecule "sites" for groups of atoms such as amino acids decreases the number of "atom sites" which need to be considered at each step of the simulation.[14]

Many of the published bilayer studies have started with the well cited bilayer constructed and simulated by the Tieleman group in Canada.[12,15] AMBER9[16] does not have the force field parameters for lipids which have been included in some of the later releases of the AMBER program such as the recently released AMBER12.[17]

In order to model *in vitro* data, simulated membrane systems have been constructed. A confluent POPC bilayer represents the eukaryotic cells while a mixed bilayer of POPC and POPG in a 4:1 ratio represents the electrostatic characteristics of prokaryotic cells.

Construction of each bilayer was conducted using Insight II[18] and pymol.[19] Within a 10 X 10 lipid unit membrane leaflet each lipid was placed 9.0 angstroms apart from neighboring lipids with a final composition of 200 lipid units (180 POPC:20 POPG). Anionic lipids were distributed within the bilayer as shown in Figure 5.1 where the • represents the zwitterionic POPC lipids and the circle with the x represents the anionic POPG lipids. The second leaflet was formed by copying the original leaflet and then rotating and translating to create a mirror image which completed the bilayer (considering boundary conditions). Counter ions were added to neutralize the overall charge from the solute and attain the desired ionic concentration using the

MMTSB's[20] software in the Cygwin environment.[21]

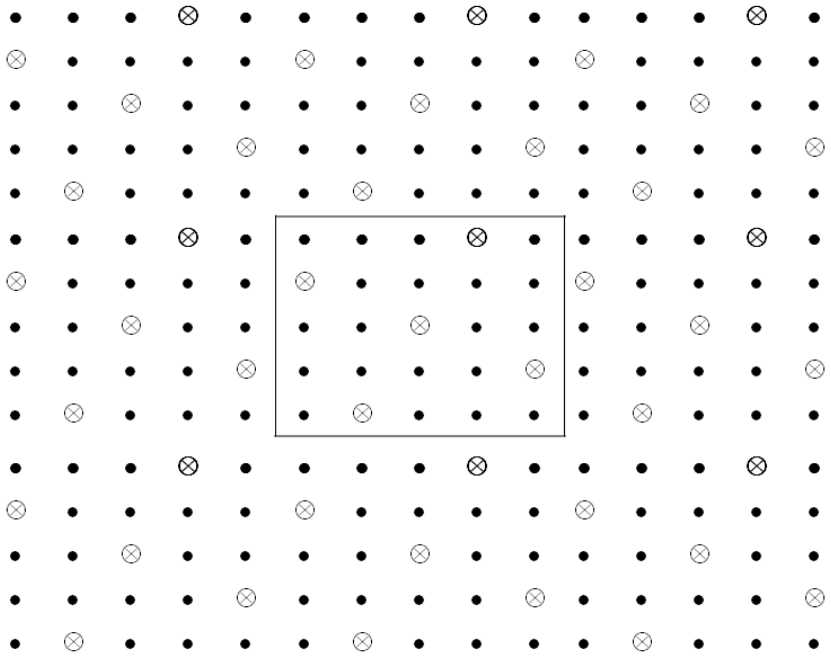


Figure 5.1 Lipid distribution where • represents POPC and the circle with the x represents POPG

5.2 MOLECULAR DYNAMICS SIMULATIONS OF AMPS IN MEMBRANE ENVIRONMENTS

Insight into the mechanism of AMP membrane disruption has been acquired from MD simulation using AMBER9.[16] Peptides **23** and **36** have been added to each bilayer system for simulation. While the exact number of AMPs involved in pore formation is unknown; previous experiments have suggested that 5-9 peptides are required to aggregate prior to membrane insertion.[22] In these simulations, a single peptide has been added to each bilayer to determine if simulations with peptides, which are partially composed of unnatural amino acids, are able to model the *in vitro*, *in vivo*, and spectroscopic activity observed in previous studies.[23-31]

All systems were solvated using the TIP3P water model. Point charges represent the electrostatic interactions. Bilayers were solvated though the generation of a water box with a volume slightly smaller than the volume between the two leaflets. This approach ensured water

molecules were only added to the hydrophilic region while maintaining the hydrophobic regions of the bilayers.

Simulations utilizing the canonical ensemble with anisotropic pressure and Andersen temperature coupling protocols have been employed.[32] The system have been minimized for 1000 steps of steepest descent before heating to 310 K over 20 ps and being subjected to an additional ns of constant pressure (isothermal-isobaric) MD. There is a critical point which exists for lipids between room temperature, 300 K and physiologically relevant temperature, 310 K. The result is that the comparison of MD simulations needs to be compared to experimental data collected under similar environmental conditions.

Long-range electrostatics have been calculated using the particle mesh Ewald (PME) method as recommended for membrane simulations, specifically those involving charged lipids.[33] The non-bonded cutoff has been set at 10 angstroms. FF parameters for each molecule in the system are outlined in Chapter 2. Once density stabilization had occurred the production run was started with the canonical ensemble. The simulations have been continued to the longest trajectory feasible given the computational resources. All MD simulations and analyses have been performed using AMBER9 unless otherwise noted. [16]

The formation of hydrogen bonds is an important intermolecular interaction which is suspected to play an important role in AMP/bilayer interactions. In Figure 5.2 a three site hydrogen bond system is outlined where D, represents the hydrogen bond donor and, A, represents the hydrogen bond acceptor. Hydrogen bonds were considered to be present when the Ha distance was 2.5 angstroms or less, the distance between the DA atoms was 3.5 angstroms or less and the D-H-A angle was 120° or greater or 90° or greater depending on the residue as outlined in Figure 5.2.

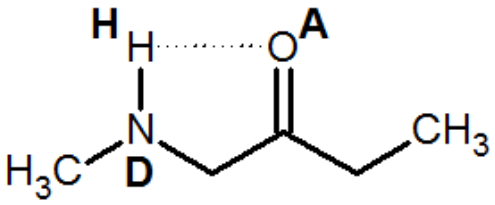


Figure 5.2 Criteria for hydrogen bond formation

5.3 SURFACE AREA OF THE BILAYER AND LIPID DENSITY

The literature values for the lipid area have a fair amount of variability, up to 25%.[34] The average area per lipid for POPC was 63.5 ± 0.3 angstroms squared. The average surface area per lipid or lipid density is often used as a measure of the force field's ability to model a system. Depending on the ensemble employed during the production phase of the MD simulation the variability of the surface area is a good metric to judge the correlation to real world systems. The consensus is that the use of the MD_NPT, isothermal-isobaric ensemble is most appropriate because it will allow the bilayer to expand and contract as the simulation progresses. The limiting factor is that the forces can become too great making the lipid density too high and the average surface area per lipid too low. This has been negated by many through the use of a surface area restraint which forces a constant surface area throughout the course of the simulation.

5.4 MIXED BILAYER 10 X 10 LIPIDS

A control mixed bilayer simulation was conducted as a reference for additional simulations. Figure 5.3 shows the starting simulation orientations for the control mixed bilayer simulation. Sodium ions are shown as purple spheres and chlorine ions are shown as the green spheres. Lipids in the mixed bilayer were separated 9 angstroms from one another. The lipid leaflets were placed 89.4 angstroms apart with X and Z dimensions creating a plane which measures 81 angstroms x 81 angstroms. The distance between the lipid leaflets was selected to be far enough apart that the peptide (even if it were to be fully extended, which has a length of approximately 35 angstroms, would not 'see' the other leaflet enabling the periodic conditions to be effectively replicated. Each leaflet was constructed with 10 lipids used in each dimension. The periodic boundary box had dimensions of 90, 136, 90 angstroms for the X, Y, and Z dimensions respectively. The increased size of the X and Z dimensions of the periodic boundary are to prevent overlap with the atoms in the adjoining periodic boxes. The smaller distance in the Y direction was to encourage the intermingling of the aliphatic carbon tails of the lipid molecules.

RMSD was used as an indicator of simulation completion. The RMSD is a measure of the deviation of the atomic coordinates relative to the minimized structure of the system. The RMSD plot shown in Figure 5.4 was created 20 ns into the simulation.

Animation of the trajectory did not identify any large changes throughout the course of the simulation. The re-imaged bilayer created at 20 ns into the simulation is shown in Figure 5.5 which shows the lipids intermingling of the aliphatic lipid tails. The thickness of the bilayer, the distance from the phosphate atom of the top leaflet to the phosphate of the bottom leaflet, is

virtually unchanged throughout the course of the simulation. Figure 5.6 shows the anionic lipid distribution from the top of the bilayer looking down the y axis shown the plane created with the x and z axes. The anionic lipids are shown with the purple surface and the zwitterionic lipid POPC is shown with a black surface.

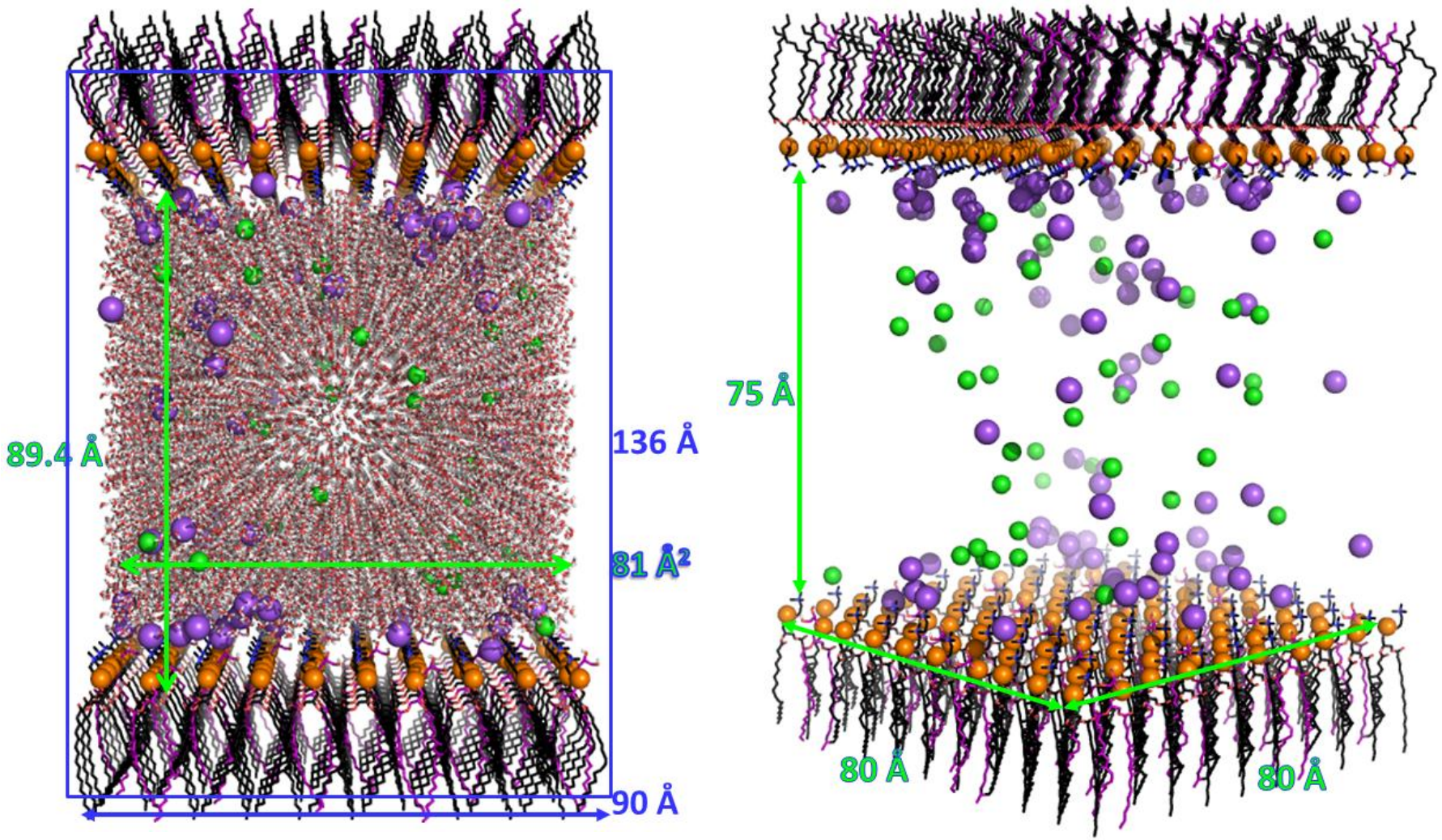


Figure 5.3 A) Water Box added to the control mixed bilayer simulations and periodic boundary box B) Control mixed bilayer at start of simulation

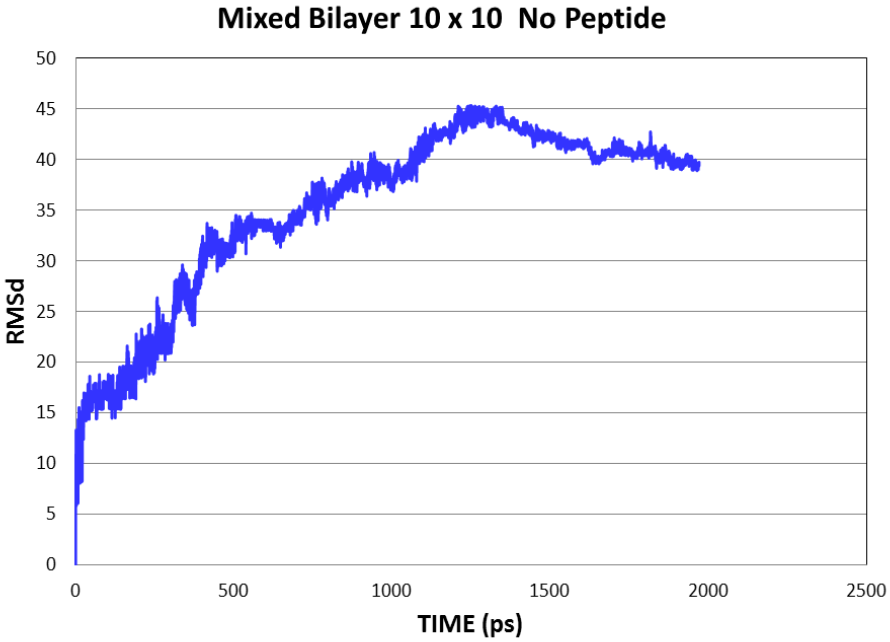


Figure 5.4 RMSD plot of the control mixed 10 x 10 bilayer at the end of 20 ns of MD simulation relative to the minimized bilayer structure.

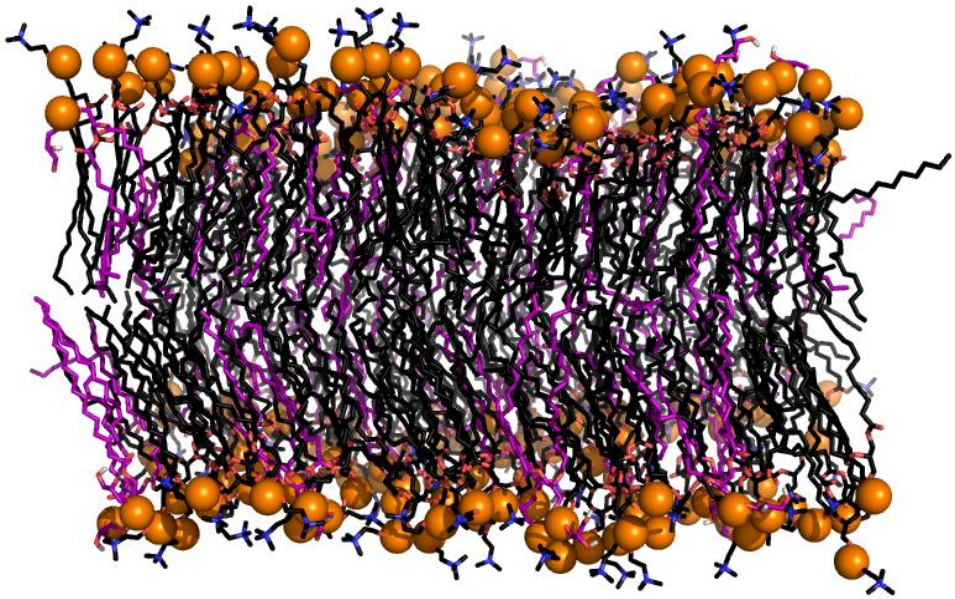


Figure 5.5 Re-imaged control mixed 10 x 10 lipid bilayer at the end of 20 ns of MD simulation

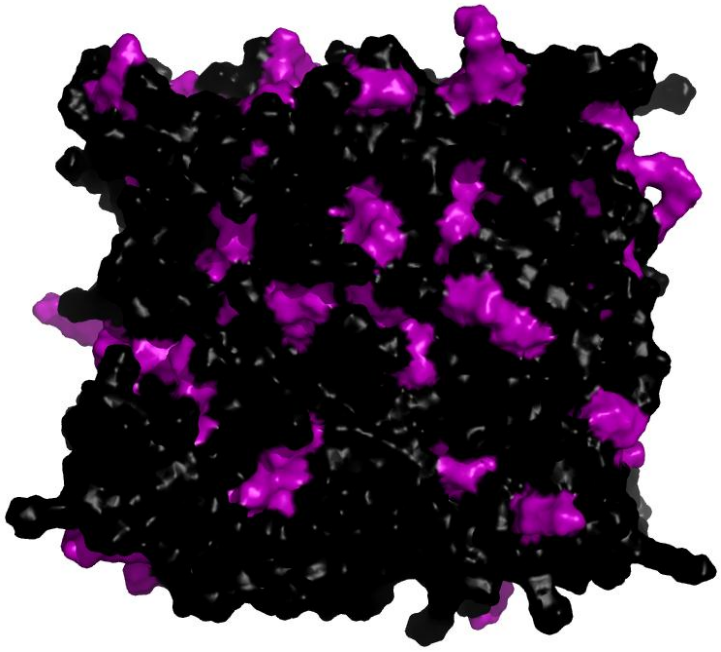


Figure 5.6 Control mixed 10 x 10 bilayer showing the x and z plane POPG lipid surfaces are shown in purple and POPC lipid surfaces are shown in black

5.5 CONTROL BILAYER 10 X 10 LIPIDS NO PEPTIDE

A control confluent bilayer simulation was conducted as reference for additional simulations. Figure 5.7 reflects the starting orientation of the control confluent bilayer simulation. Sodium ions are shown as purple spheres and chlorine ions are shown as the green spheres. Lipids in the mixed bilayer were separated by an average 8.7 angstroms from one another. The lipid leaflets were placed 68 angstroms apart with X and Z dimensions having 78 angstroms each. The distance between the lipid leaflets was selected to be far enough apart that the peptide (even if it were to be fully extended, which has a length of approximately 35 angstroms, would not 'see' the other leaflet enabling the periodic conditions to be effectively replicated. 10 lipids were used in each dimension. The periodic boundary box had dimensions of 90, 126, 90 angstroms for the X, Y, and Z dimensions respectively. The increased size of the

X and Z dimensions of the bilayer surface of the periodic boundary are to prevent overlap with the atoms in the adjoining periodic boxes. The thickness of the bilayer, the distance from the phosphate atom of the top leaflet to the phosphate of the bottom leaflet, is virtually unchanged throughout the course of the simulation. The smaller distance in the Y direction was to encourage the intermingling of the aliphatic carbon tails of the lipid molecules.

RMSD was used as an indicator of simulation completion. The RMSD is a measure of the deviation of the atomic coordinates relative to the minimized structure of the system. The RMSD plot shown in Figure 5.8 was created 15.8 ns into the simulation. The slope of the line is approaching zero suggesting the system has stabilized relative to the minimized structure.

Animation of the trajectory did not identify any large changes throughout the course of the simulation. The re-imaged bilayer created at 15.8 ns into the simulation is shown in Figure 5.9 which shows the intermingling of the aliphatic lipid tails. Figure 5.9 B shows the bilayer surface. Note the lack of bilayer thinning with the control bilayer. The distance between the surfaces of the two leaflets was constant throughout the course of the simulation. Zwitterionic lipid POPC is shown with a black surface.

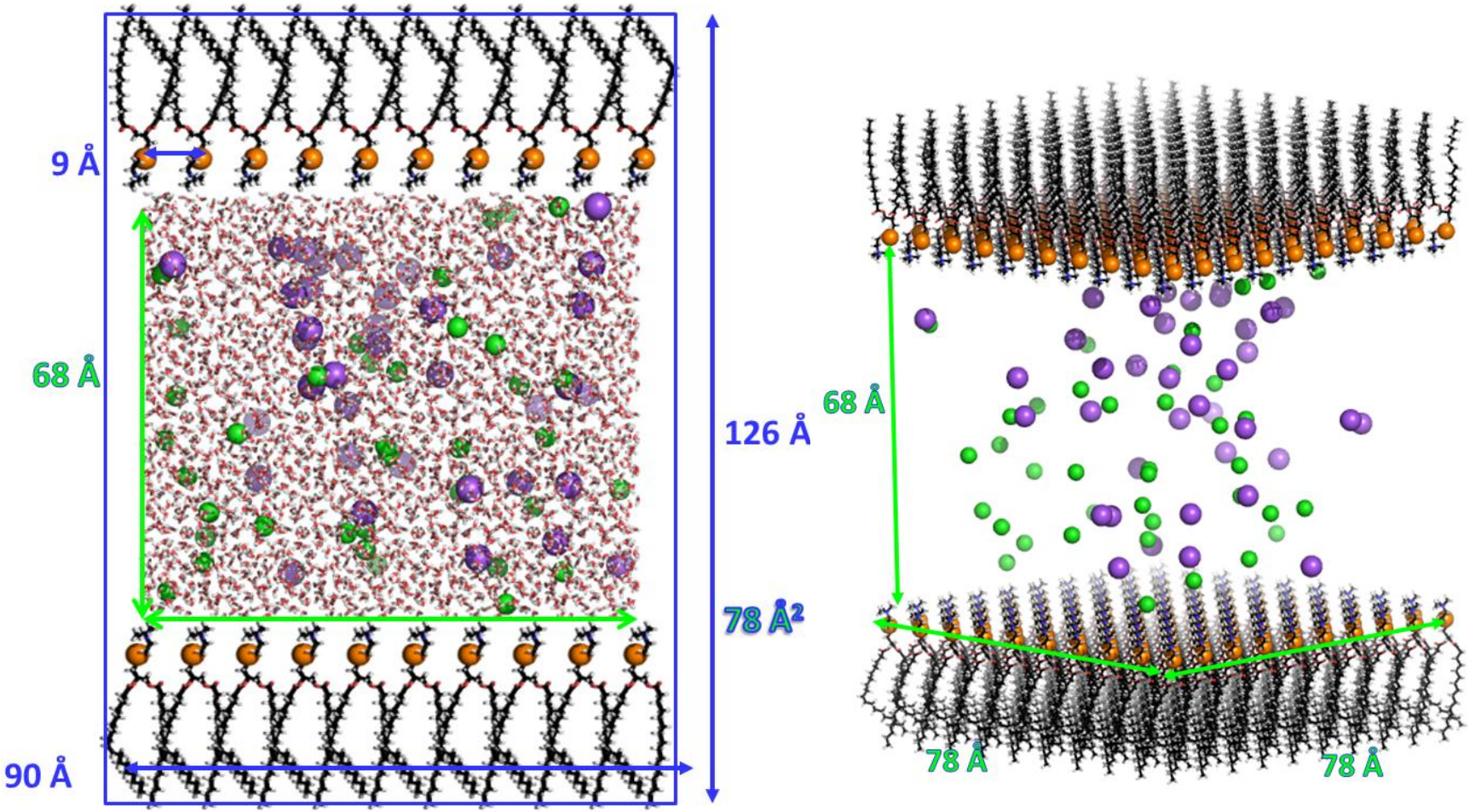


Figure 5.7 A) Water Box added to the control mixed bilayer simulations and periodic boundary box B) Control mixed bilayer at start of simulation

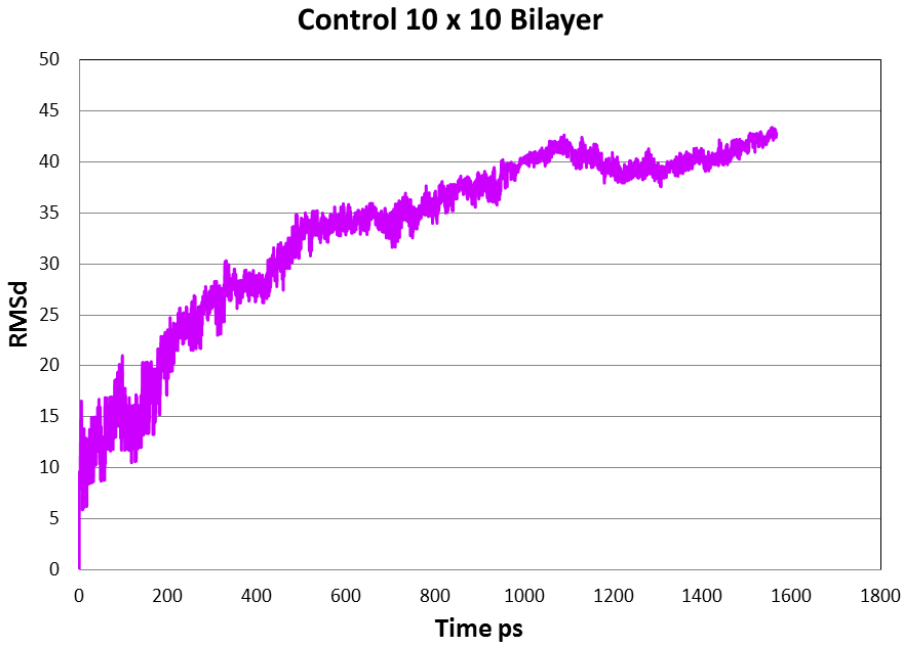


Figure 5.8 RMSD of the control bilayer simulation, no peptide.

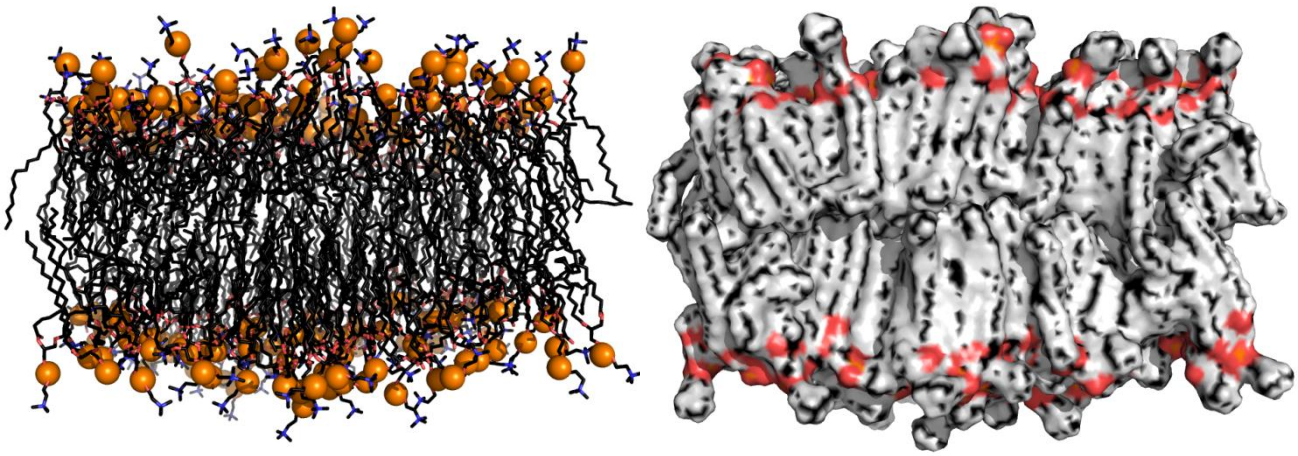


Figure 5.9 A) Control bilayer at the end of 1.5 ns trajectory. B) Control bilayer lipid solvent accessible surface at end of simulation without peptide

5.6 MIXED BILAYER 10 X 10 LIPIDS WITH PEPTIDE 23

The first mixed bilayer simulation was conducted with peptide **23**, which has glycine as the linker between the Tic Oic dipeptide units. Figure 5.10 A and B shows the starting simulation orientations for the mixed bilayer simulation with peptide **23**. Sodium ions are shown as purple spheres and chlorine ions are shown as the green spheres. Lipids in the mixed bilayer were separated 9 angstroms from one another. Anionic POPG lipids are shown with the carbon atoms colored purple while the zwitterionic lipids are shown with the carbon atoms colored black. The lipid leaflets were placed 78 angstroms apart with X and Z dimensions having 81 angstroms each. 10 lipids were used in each dimension. The water box added was slightly smaller to fill the space between the polar head groups of the lipids and measured 75 by 75 by 78 angstroms cubed. The periodic boundary box had dimensions of 90, 136, 90 angstroms for the X, Y, and Z dimensions respectively. The increased size of the X and Z dimensions of the periodic boundary are to prevent overlap with the atoms in the adjoining periodic boxes. The smaller distance in the Y direction was to encourage the intermingling of the aliphatic carbon tails of the lipid molecules. Placement of peptide **23** relative to the bilayer surface is shown in Figure 5.10 C. The peptide was placed in close proximity to the polar head groups of the bilayer yet far enough away to allow the water molecules to move in around the peptide between the peptide and the bilayer surface.

RMSD was used as an indicator of simulation completion. The RMSD is a measure of the deviation of the atomic coordinates relative to the minimized structure of the system. The RMSD plot shown in Figure 5.11 was created with the last 14 ns of a 241.5 ns trajectory. The RMSD for peptide **23** is shown in yellow while the bilayer is shown in purple**.** Peptide **23**

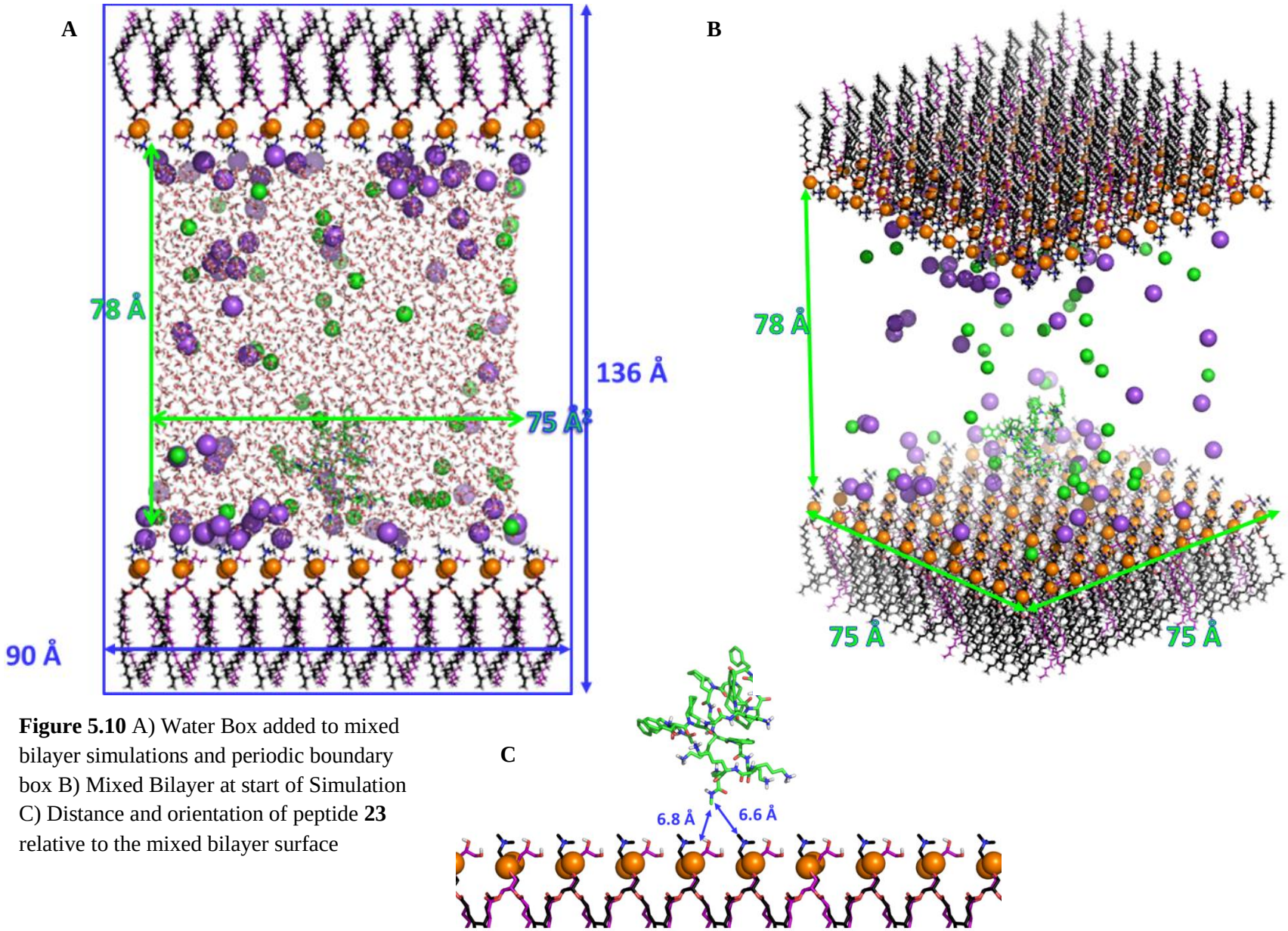


Figure 5.10 A) Water Box added to mixed bilayer simulations and periodic boundary box B) Mixed Bilayer at start of Simulation C) Distance and orientation of peptide **23** relative to the mixed bilayer surface

stabilized early in the trajectory, within the first 50 ns (data not shown). The mixed bilayer has not stabilized during the 241.5 ns trajectory. The large deviation observed near 234 ns mark of the simulation was determined to be an artifact of the re-imaging process which places a specified set of atoms at the center of the periodic box and adjusts the other molecules in relation to the specified atoms.

Animation of the trajectory reflected a very dynamic movement in the C terminal end of peptide **23** while the N terminal lysine residues appear to form interactions, probably hydrogen bonds, with the lipid head groups. Figure 5.12 shows the anionic lipids, POPG shown in purple and zwitterionic lipids POPC are shown in green. Figure 5.12 A shows the orientation of peptide

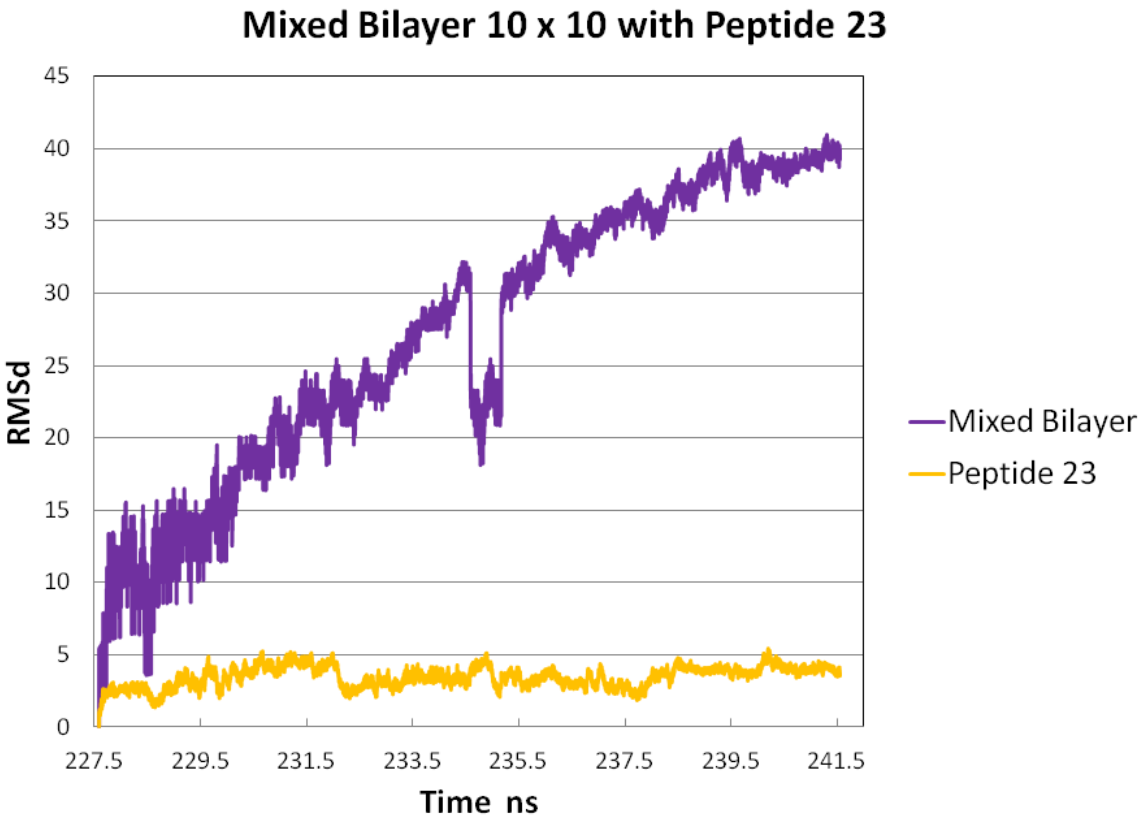


Figure 5.11 RMSD for the mixed bilayer with peptide **23**

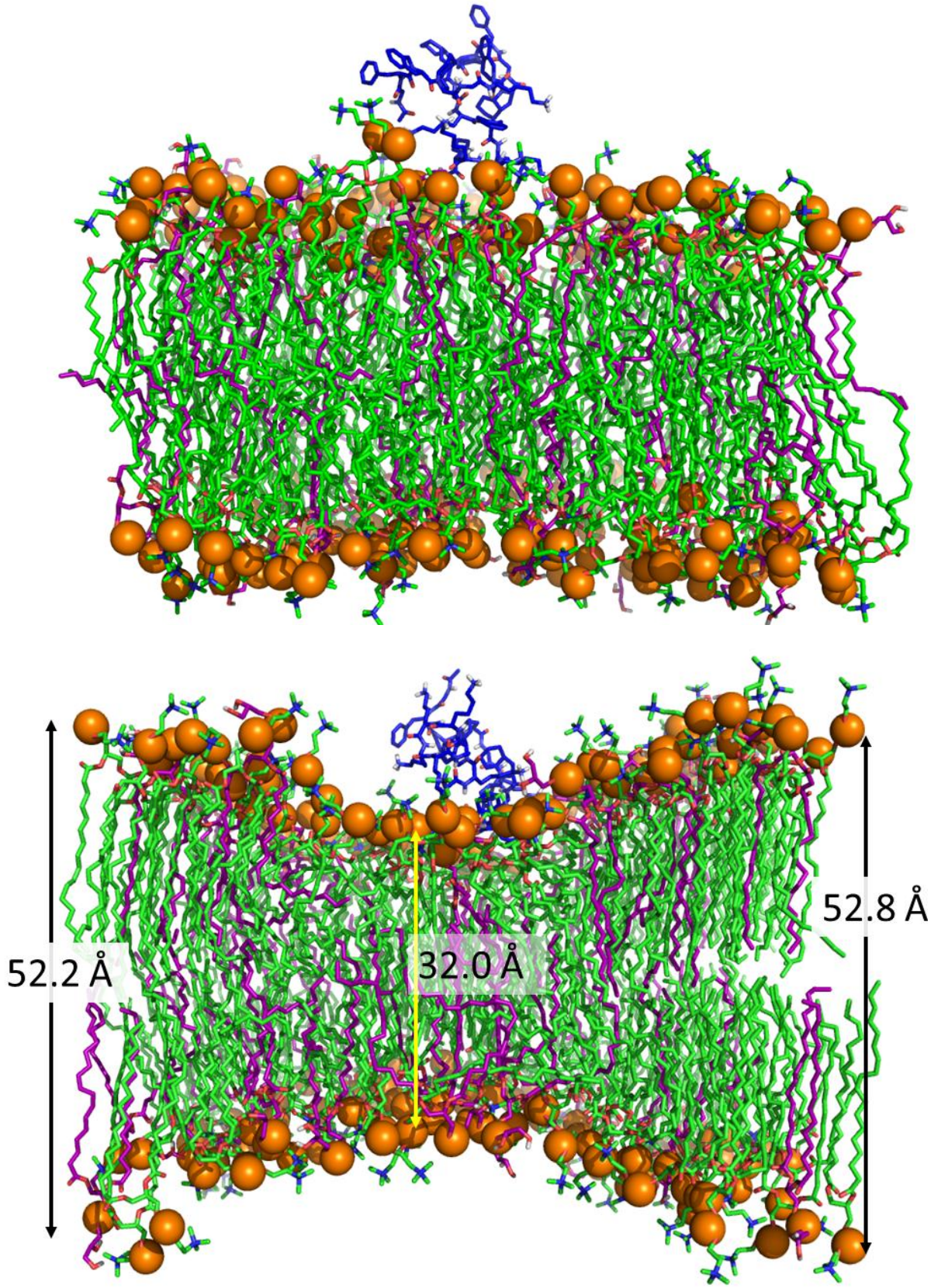


Figure 5.12 A) Mixed bilayer simulation with peptide **23** at start of production run B) Orientation of peptide **23** relative to bilayer at end of 241.5 ns of simulation

23 at the beginning of the production run; the starting distance between the phosphate atoms averages 39.73 angstroms. Figure 5.12 B shows the bilayer at the end of 241.5 ns of constant volume dynamics. A clear thinning of the bilayer is observed while the lipids farthest from the peptide have moved further apart. The thinning of the bilayer is reflects the powerful impact of peptide 23 on the bilayer. The time scale achieved during the course of this simulation is less than that required for membrane disruption which has been calculated to fall between milliseconds and seconds.[36] The other limitation of this simulation is the single peptide used. AMPs are known to require a critical concentration of peptide relative to the number of lipid molecules present. Aggregation of the lipids is required before membrane disruption thus complete membrane disruption is not expected.

Another interesting development over the course of the simulation is the lipid distribution with the bilayer surface. The simulation started with the POPG lipids evenly distributed throughout the bilayer relative to the location of the POPC lipids. Figure 5.13 A shows the distribution of the lipid types by showing the location of the phosphate atoms for each lipid. The POPC lipids are shown with orange spheres and the POPG lipids are shown with green spheres. Note the regions decreased lipid density, outlined in red. The location of the peptide is found in a region of low lipid density and more specifically mostly zwitterionic lipids. There are only two POPG lipids interacting with peptide **23**, noted with yellow arrows in Figure 5.13 B. The highly polar C terminus of peptide **23** with the 4 lysine residues in sequence seems the obvious choice for interaction with the anionic lipids.

Examining the lipid distribution using the solvent accessible provides an explicit comparison of the starting position of the anionic lipids relative to the zwitterionic lipids at the

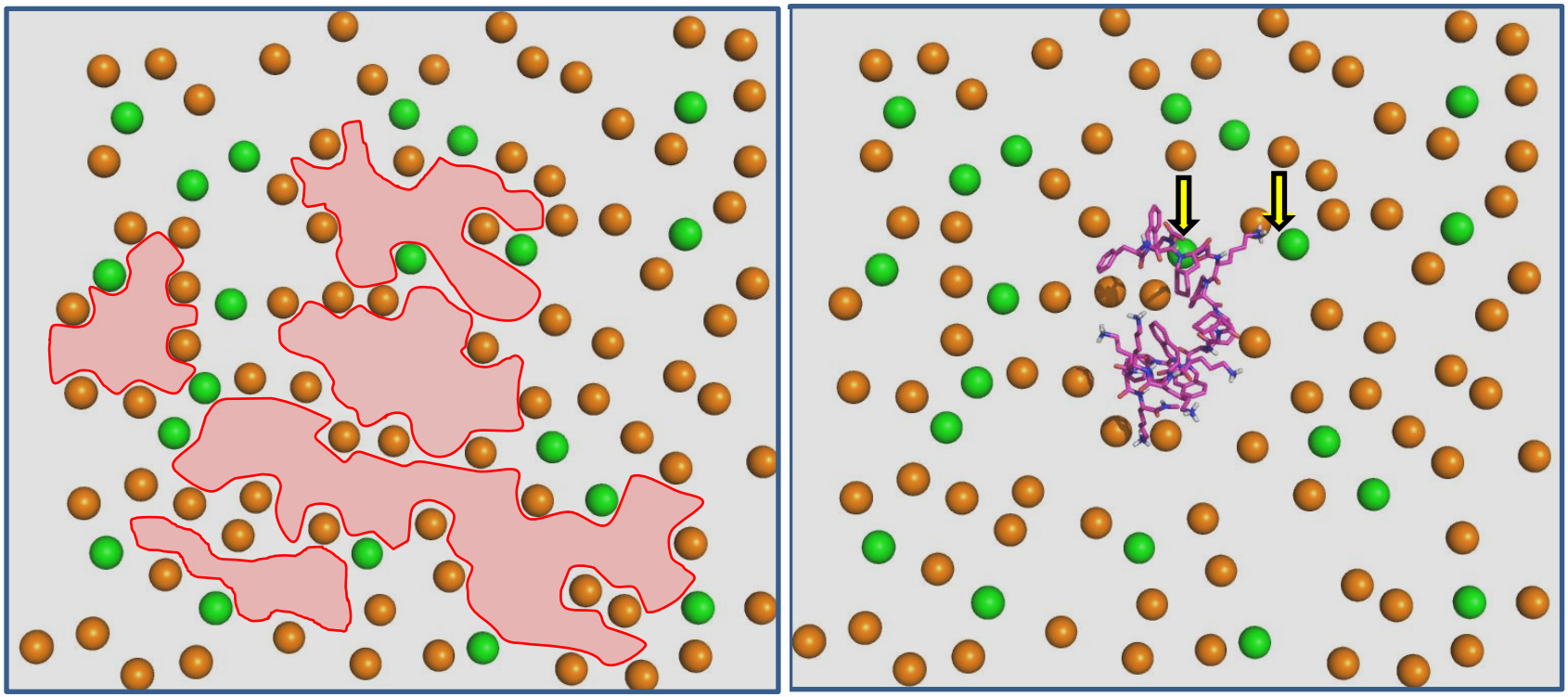


Figure 5.13 A) Distribution of anionic, POPG (Green) and zwitterionic, POPC lipids (orange). The red region highlights the area of low lipid density B) Position of peptide **23** relative to the anionic lipids of the mixed bilayer. The two POPG lipids interacting with peptide **23** are noted with yellow arrows.

start of the production run, Figure 5.14 A provides a good reference point for comparison at the 241.5 ns mark of the simulation, Figure 5.14 B.

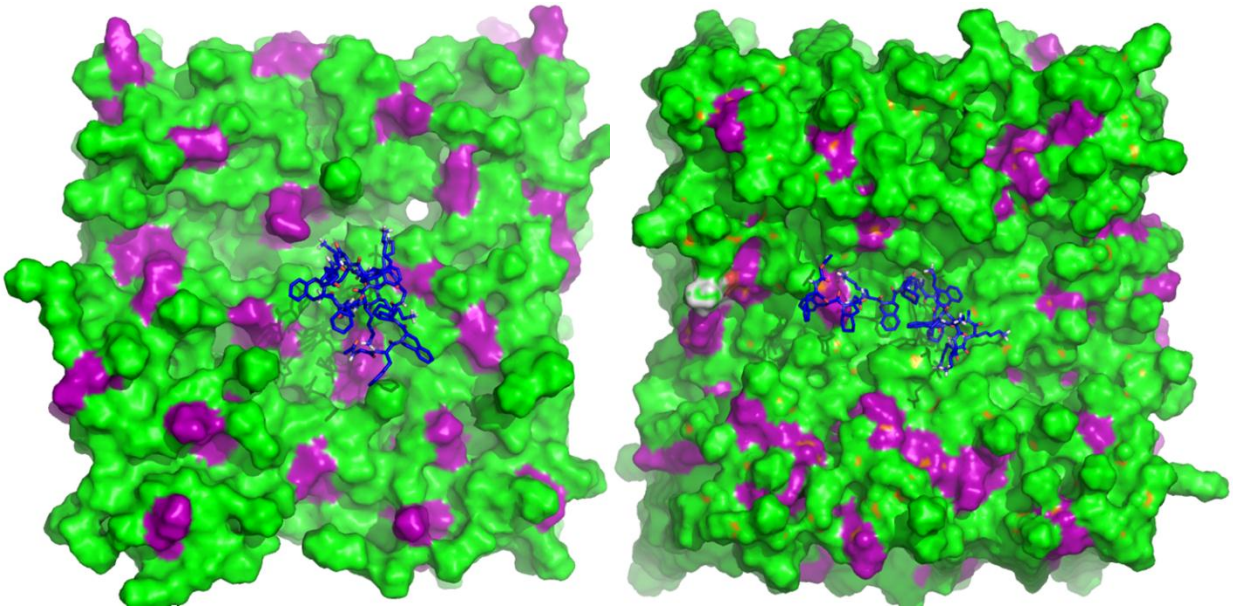


Figure 5.14 A) Surface of lipids at start of production run B) Distribution of the lipid surfaces 241.5 ns into the production simulation. Anionic POPG lipids are shown in purple, zwitterionic POPC lipids are shown in green. Peptide **23** is shown with blue carbon atoms.

The location of the peptide relative to the anionic lipids is relevant due to the mosaic associated with microbial membranes. The aggregation of minority lipids, POPG within the membrane surface correlates to the observed variation in the membrane charge distribution. The anionic lipid distribution at the conclusion of 241.5 ns is shown in Figure 5.14.

Figure 5.15 shows the Ramachandran plot for each residue from peptide **23** averaged over the last 2 ns of the production simulation. The glycine residues fall outside of the "allowed region" of the plot however; the small size and lack of chirality associated with the residue have been observed experimentally.[37] The residues cluster in the alpha helical region of the plot with

the exception of the Tic residues which are consistently in the beta sheet region of the plot. The preference for Tic has been previously observed during the course of the simulations (see Chapter 3).

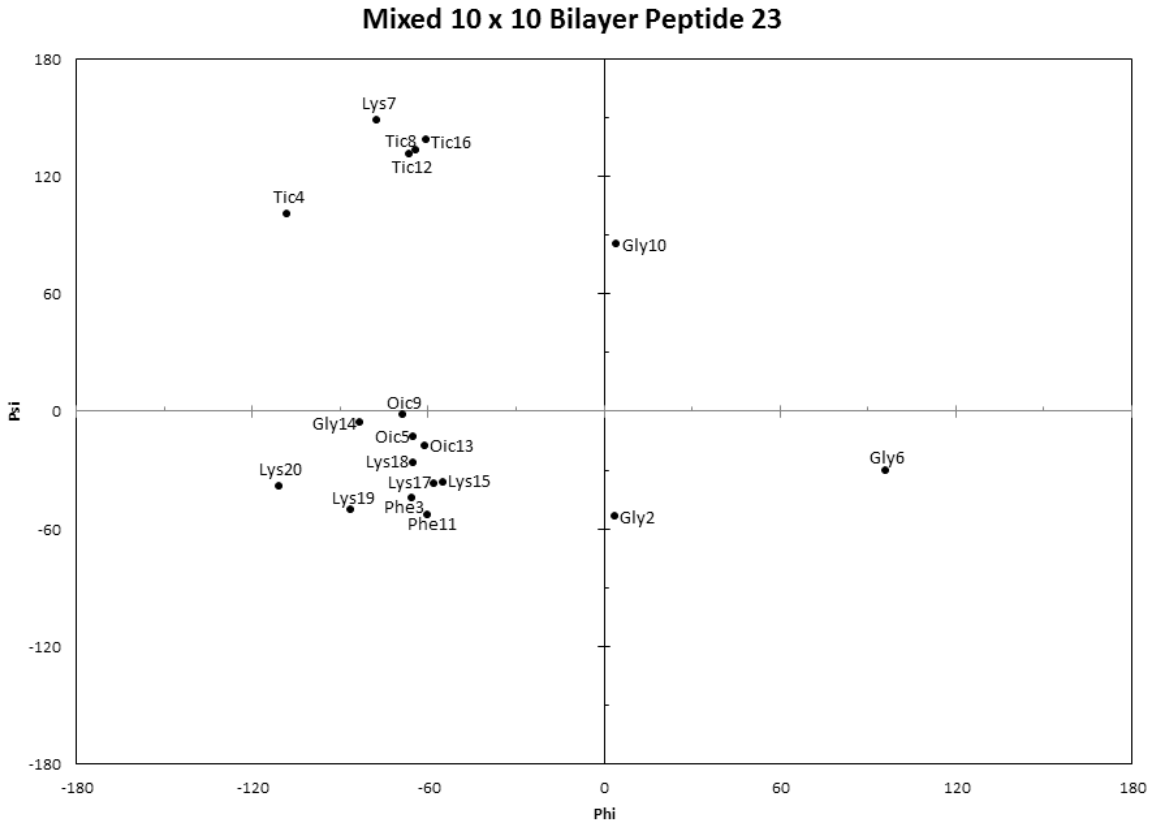


Figure 5.15 Ramachandran plot for peptide **23** in the presence of a mixed bilayer.

The lipid distribution at the conclusion of the MD simulation relative to peptide **23** suggests the peptide is able to influence the lipid distribution within the membrane. The two POPG lipids close enough for interaction with the peptide are likely interacting through hydrogen bonds. The first POPG is interacting with the side chain of a lysine residue and the other interaction is with a central lysine residue (see Figure 5.16). The lipid distribution effects

extend beyond the closest leaflet of the bilayer. The bottom leaflet, furthest away from the peptide shows a similar lipid distribution relative to the peptide's position (data now shown).

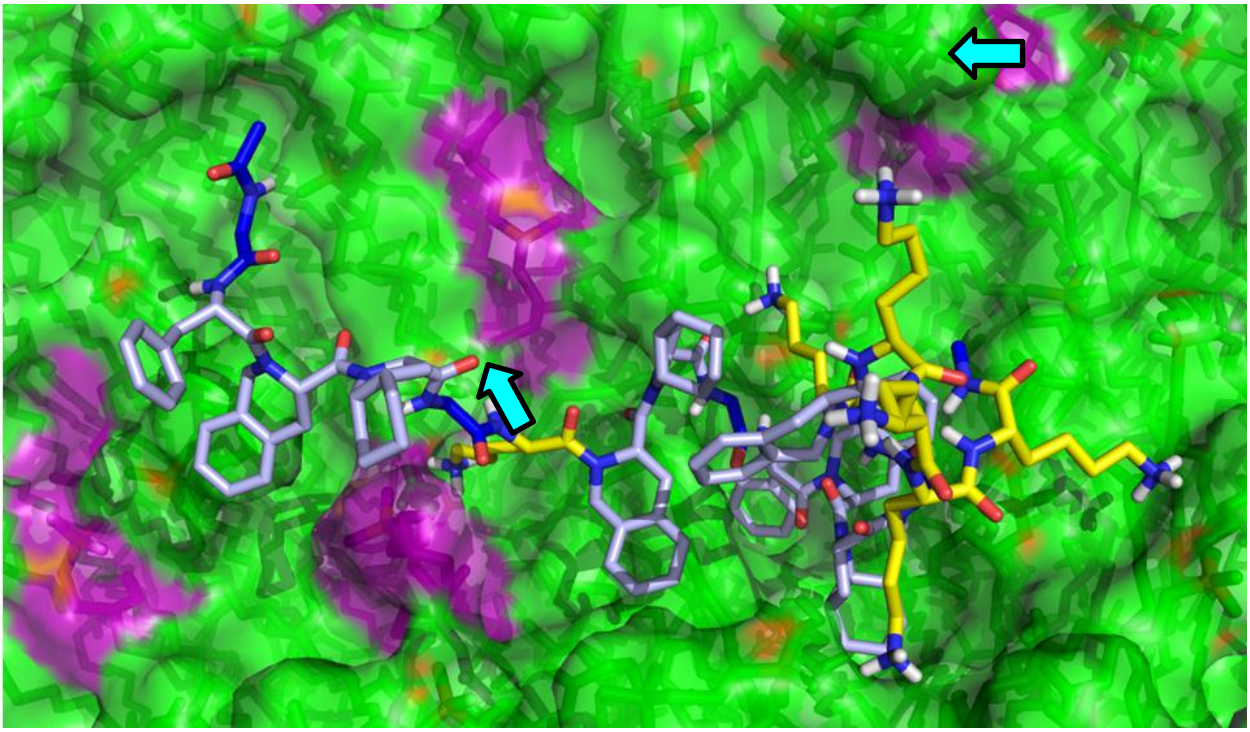


Figure 5.16 Surface of the mixed bilayer with peptide **23**. The two points of the peptide interacting with anionic lipids are pointed out with cyan arrows. Hydrophobic residues are shown in grey while the polar lysine residues are shown in yellow.

A closer look at peptide **23** relative to the bilayer surface is shown in Figure 5.16. Hydrophobic residues, Phe, Tic, and Oic are noted with grey carbon atoms while the cationic Lys residues are shown with yellow carbon atoms. The portions of the peptide interacting with anionic POPG lipids are pointed out with cyan arrows. The polar interactions appear to have held the peptide in close proximity to the bilayer surface enabling the hydrophobic rings of the terminal Phe, Tic, and Oic to move down into the hydrophobic region of the bilayer, see Figure 5.17.

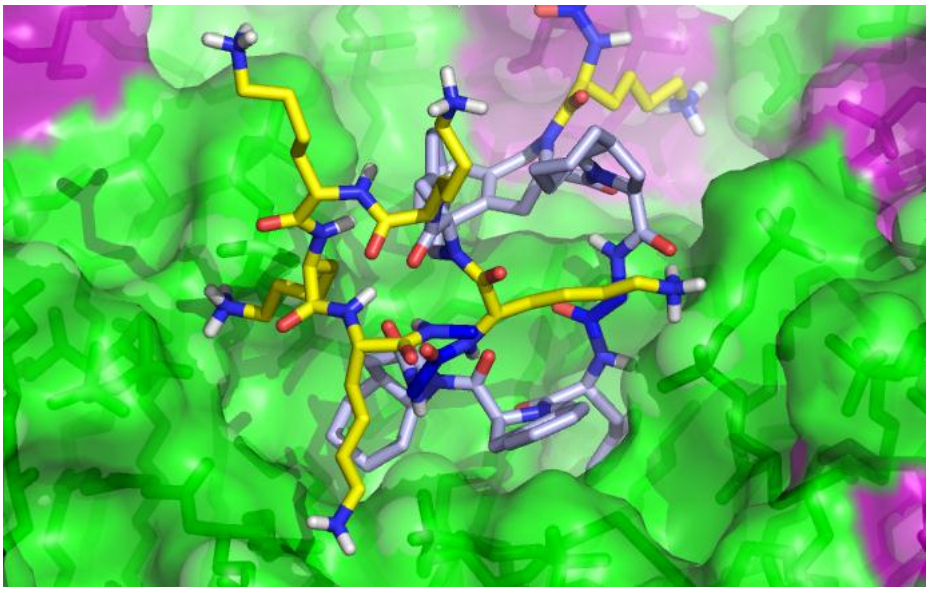


Figure 5.17 Surface of the mixed bilayer with peptide **23**. Hydrophobic residues are shown in grey while the polar lysine residues are shown in yellow.

5.7 CONTROL BILAYER 10 X 10 LIPIDS WITH PEPTIDE 23

The control bilayer simulation conducted with peptide **23** was conducted for comparison to control bilayer and the experimentally derived liposome data.[38] Figure 5.18 A and B shows the starting simulation orientations for the control bilayer simulation with peptide **23**. Sodium ions are shown as purple spheres and chlorine ions are shown as the green spheres. Lipids in the control bilayer were separated 9 angstroms from one another. Zwitterionic lipids are shown with the carbon atoms shaded black. The lipid leaflets were placed 78 angstroms apart with X and Z dimensions having 81 angstroms each. 10 lipids were used in each dimension. The water box added was slightly smaller to fill the space between the polar head groups of the lipids and measured 68 by 68 by 78 angstroms cubed. The periodic boundary box had dimensions of 90, 126, 90 angstroms for the X, Y, and Z dimensions respectively. The increased size of the X and Z dimensions of the periodic boundary are to prevent overlap with the atoms in the adjoining periodic boxes. The smaller distance in the Y direction was to encourage the intermingling of

the aliphatic carbon tails of the lipid molecules. Placement of peptide **23** relative to the bilayer surface is shown in Figure 5.18 C. The peptide was placed in close proximity to the zwitterionic head groups of the bilayer yet far enough away to allow the water molecules to move in around the peptide between the peptide and the bilayer surface.

RMSD was used as an indicator of simulation completion. The RMSD plot shown in Figure 5.19 was created with the last 20 ns of a 150 ns trajectory. The RMSD for peptide **23** is shown in yellow while the bilayer is shown in purple**.** The simulation was not continued due to the stabilization of both the peptide and the bilayer as calculated by an average slope near 0. Once the system has stabilized perturbation is unlikely within the timescale available.

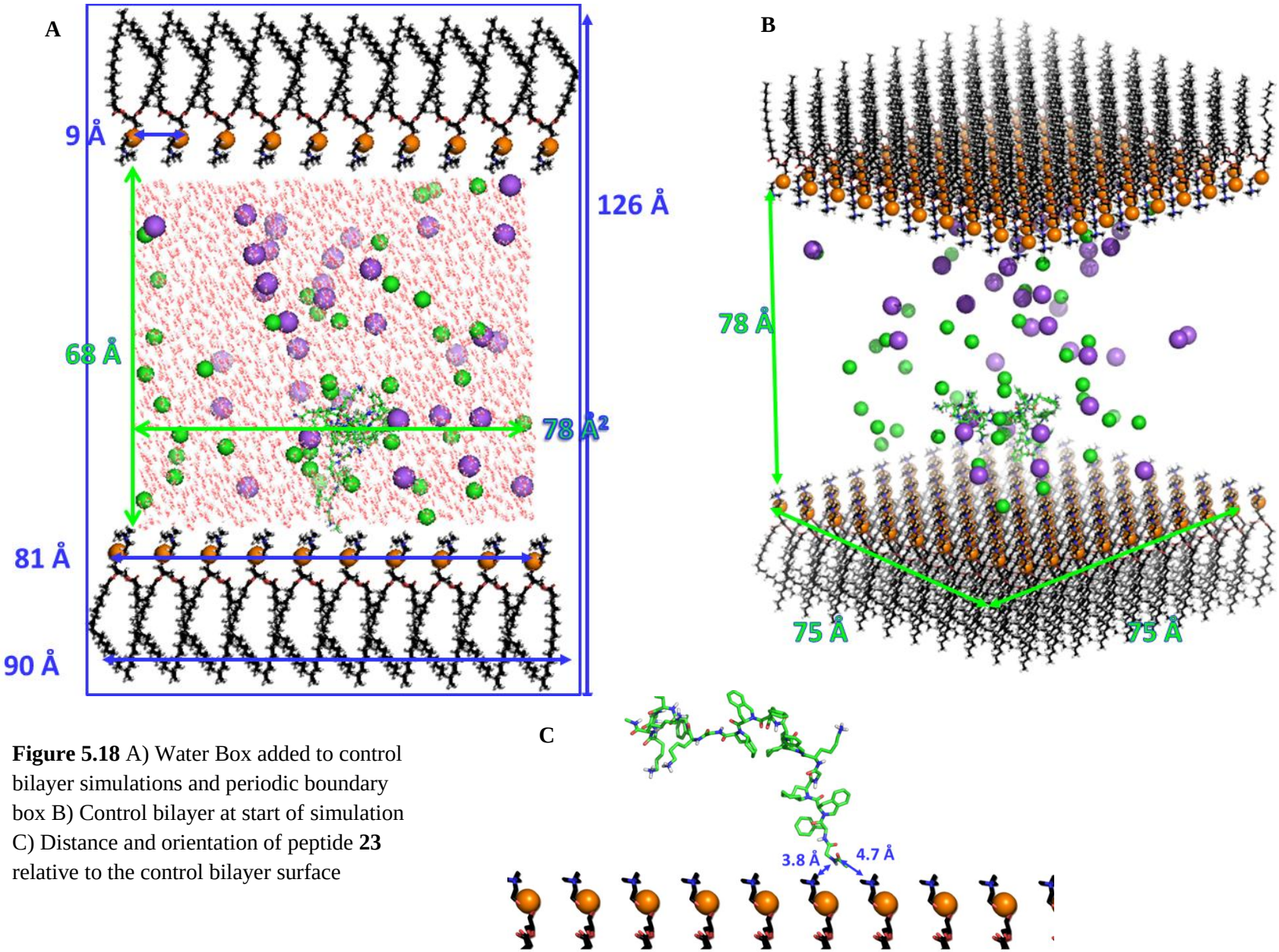


Figure 5.18 A) Water Box added to control bilayer simulations and periodic boundary box B) Control bilayer at start of simulation C) Distance and orientation of peptide **23** relative to the control bilayer surface

Figure 5.20 A shows the orientation of peptide **23** at the beginning of the production run; the starting distance between the phosphate atoms was approximately 41 angstroms. Figure 5.20 B shows the bilayer at the end of 150 ns of constant volume dynamics. The distance between bilayer thickness at the beginning of the simulation and the end is relatively unchanged, around 40 angstroms, given the expected dynamic nature of the bilayer. The control bilayer does not

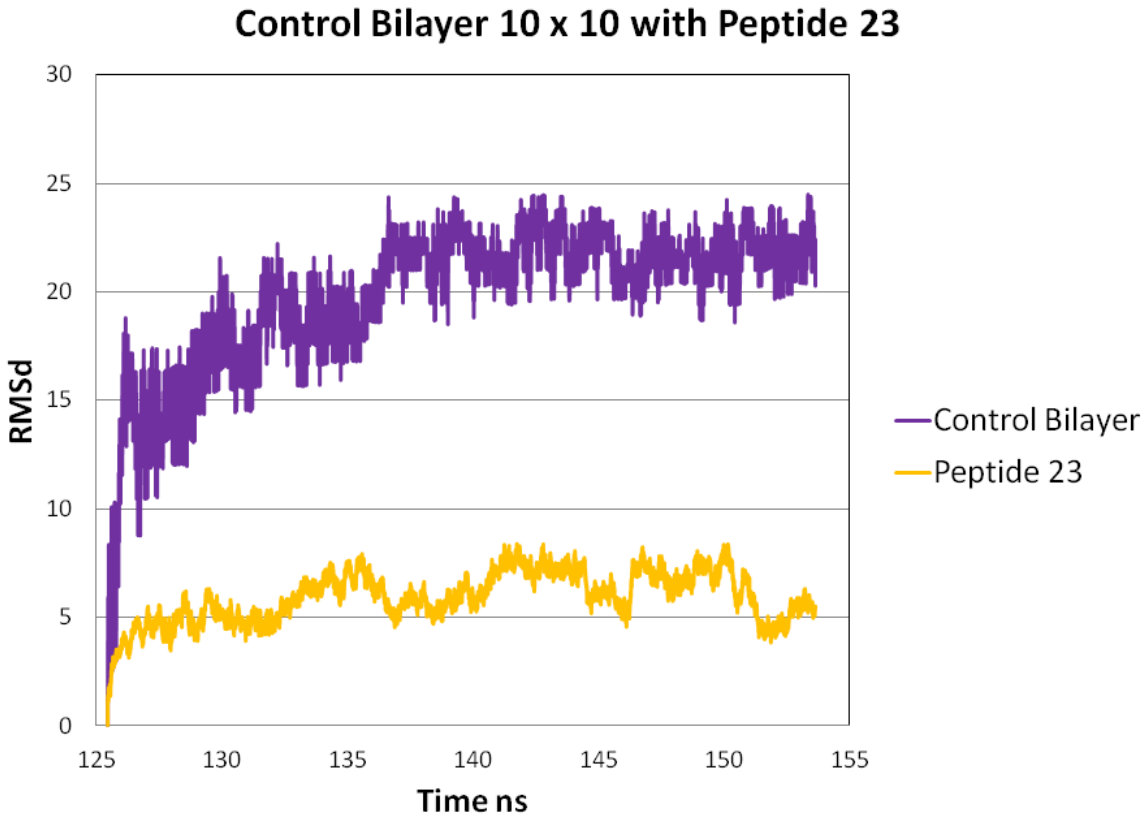


Figure 5.19 RMSD of the confluent POPC bilayer with peptide **23**

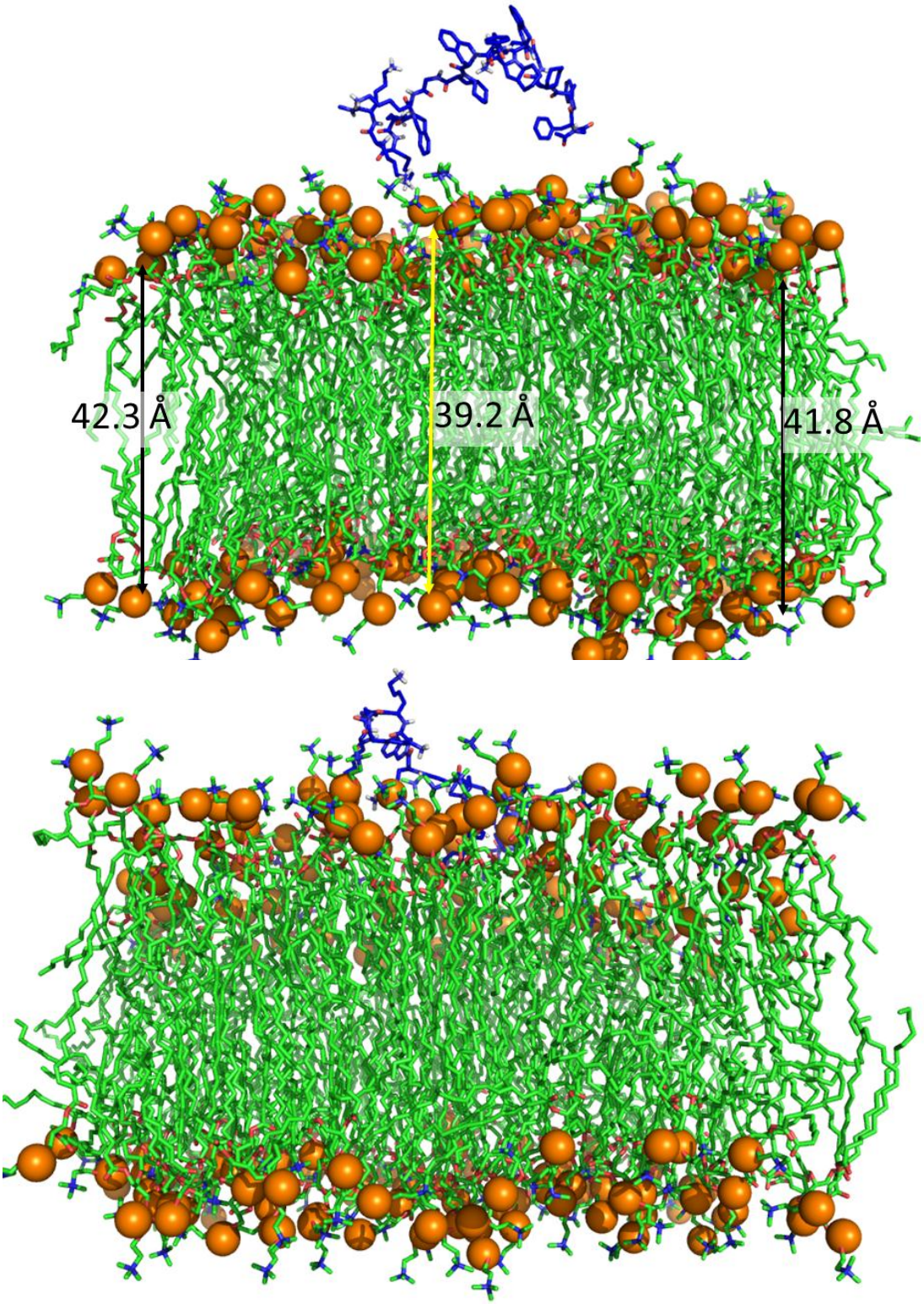


Figure 5.20 A) Control bilayer with peptide **23** at start of the production simulation B) Control bilayer after 150 ns of constant volume dynamics. POPC is shown with green carbon atoms peptide **23** is shown with blue carbon atoms.

display any evidence of membrane thinning as observed in the simulation with the mixed bilayer and peptide **23**. Given the peptide's interaction with the bilayer surface this is an interesting development (Figure 5.20 B). It is interesting to note the peptide's proximity to the lipids and even its ability to form hydrogen bonds with the choline head groups of POPC and the phosphate ester oxygen atoms (Figure 5.21) and yet, the lack of membrane thinning suggests a less disruptive interaction. The hydrophobic rings of the Phe, Tic, and Oic residues are shown with grey carbon atoms and the Lys residues are shown with yellow carbon atoms in Figure 5.21. The simulation started with the POPG lipids evenly distributed throughout the bilayer relative to the location of the POPC lipids.

Figure 5.22 shows the confluent POPC bilayer surface at the end of 150 ns of MD simulation. All but one of the lysine residues appears to be interacting with the polar head groups of the zwitterionic lipids. The peptide seems to be rather extended with regard to secondary structure. Magainin which was the electrostatic model for the AMPs with unnatural

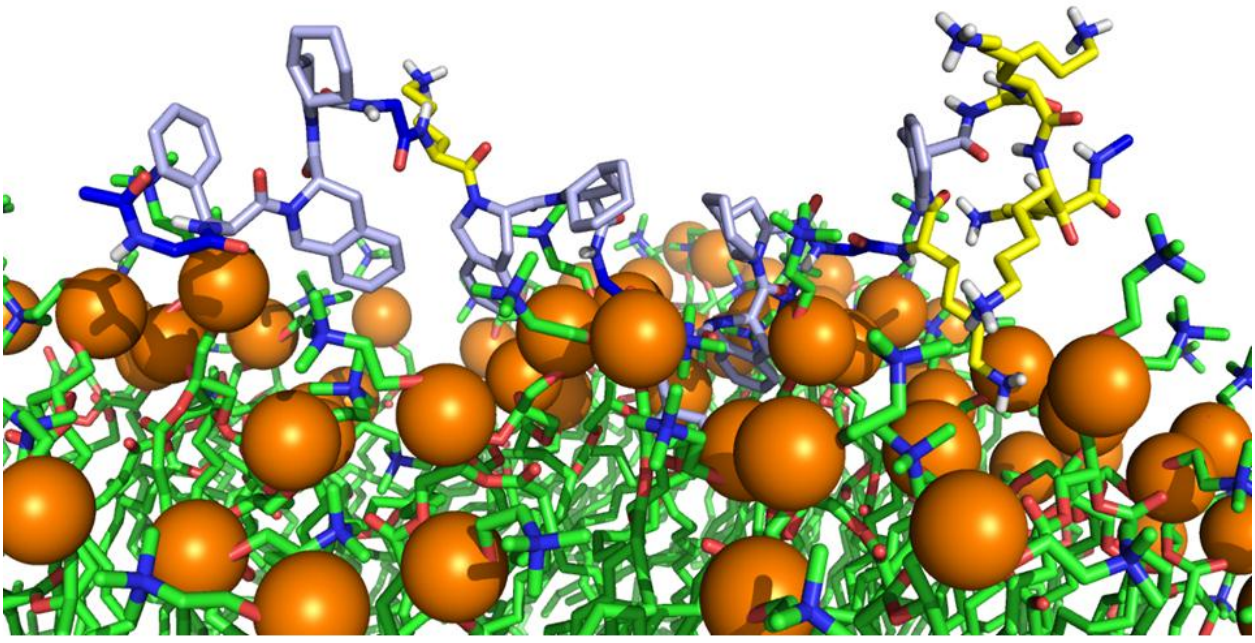


Figure 5.21 Peptide **23** with a confluent POPC bilayer. Hydrophobic rings are noted with grey carbon atoms and Lys residues are shown with yellow carbon atoms.

amino acids incorporated into the peptide sequence. Magainin is noted to adopt an amphipathic secondary structure with polar residues aligning themselves along one side of the alpha helical secondary structure and hydrophobic residues on the opposing side. It is interesting to note that three of the four Tic residues have aligned themselves toward the bilayer surface. The steric hindrance of the Tic Oic dipeptide discourages both residues from orienting toward the bilayer surface.

The speed with which the bilayer stabilized relative to time required for the bilayer with peptide **23** further suggests the disruptive influence of the peptide on the bilayer surface. This is consistent with the experimentally derived data. Peptide **23** has been observed to disrupt POPC

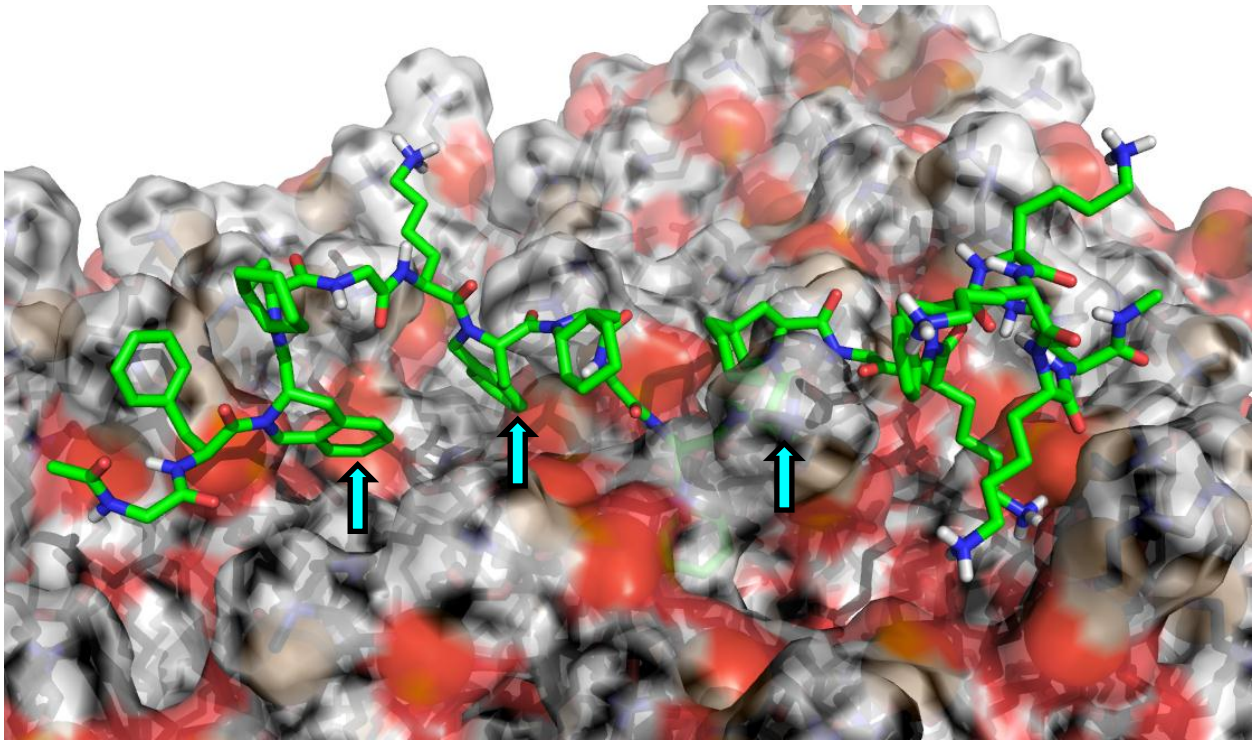


Figure 5.22 Peptide **23** with a confluent POPC bilayer surface displayed peptide **23** is shown with green carbon atoms and the POPC lipids are shown with black carbon atoms. Tic residues are identified with cyan arrows.

bilayers in a concentration dependent manner.

The Ramachandran plot shown in Figure 5.23 reflects the average phi & psi angles for peptide **23** over the last 2 ns of the production simulation with the control bilayer. Each residue was averaged over the trajectory. The glycine residues once again fall outside the "allowed regions" as identified by Ramachandran however, there is x-ray structural data to support the small achiral glycine residue being able to inhabit a larger region of the plot than other residues.[37] The Tic residues are positioned in the beta sheet region which is consistent with both the tripeptide simulations (Chapter 3) and the angles observed with peptide **23** in the presence of the mixed bilayer (Figure 5.15).

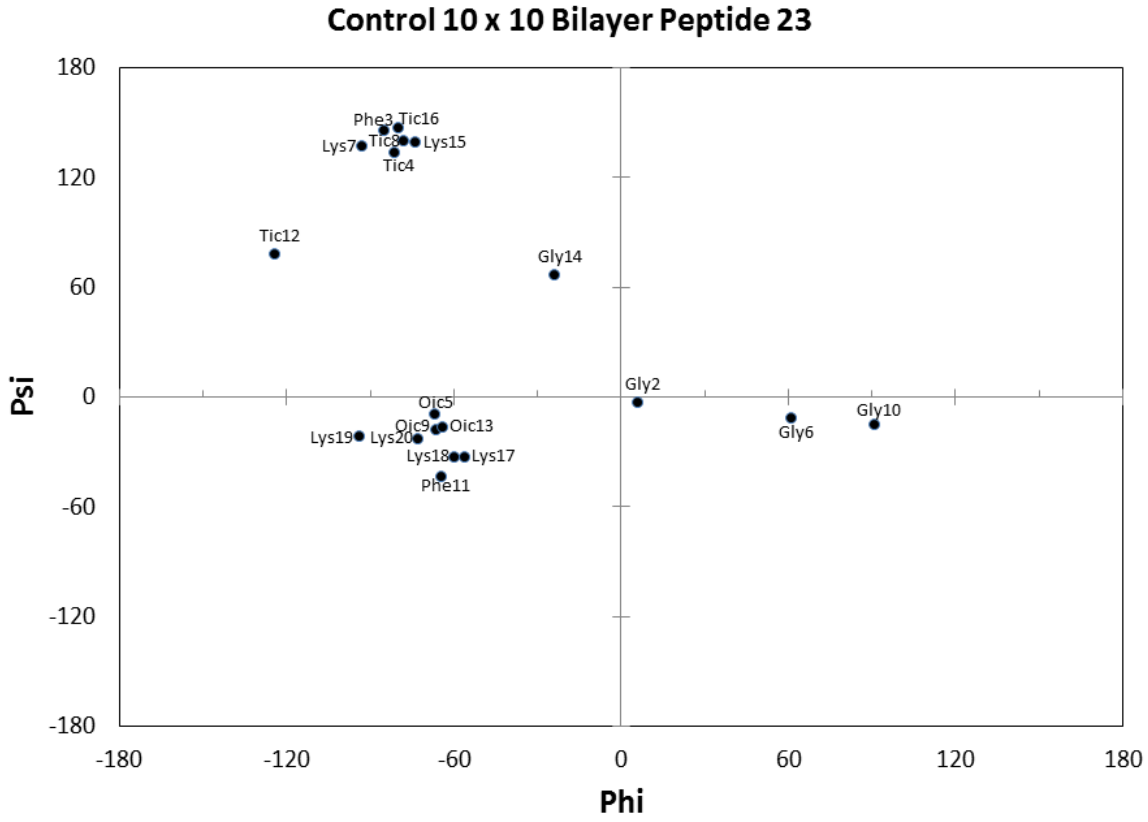


Figure 5.23 Ramachandran plot for peptide **23** in the presence of a control bilayer.

5.8 SUMMARY

MD simulations of mixed bilayers composed of a 4:1 ratio of POPC to POPG have been conducted in the presence of peptide **23** which has a glycine residue linking the Tic Oic dipeptides. Control simulations without peptides present were also conducted for reference and to ensure observed bilayer perturbations were a result of the peptides presence and not an artifact of the simulation.

The mixed bilayer in the presence of peptide **23** showed a membrane thinning effect after a 250 ns production simulation during which the RMSD of the bilayer did not stabilize but the peptide did relative to the start of the production simulation. When peptide **23** was simulated with the control, confluent POPC bilayer there were intermolecular interactions occurring however no membrane thinning was observed during the 150 ns trajectory during which the RMSD stabilized.

5.9 REFERENCES

1. Brogden, K. A. (2005) Antimicrobial peptides: Pore formers or metabolic inhibitors in bacteria? *Nat Rev Micro. 3*, 238-250.

2. Horn, J. N., Sengillo, J. D., Lin, D., Romo, T. D., and Grossfield, A. (2012) Characterization of a potent antimicrobial lipopeptide via coarse-grained molecular dynamics. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1818*, 212-218.

3. Cheng, J. T. J., Hale, J. D., Elliott, M., Hancock, R. E. W., and Straus, S. K. (2011) The importance of bacterial membrane composition in the structure and function of aurein 2.2 and selected variants. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1808*, 622-633.

4. Isaksson, J., Brandsdal, B. O., Engqvist, M., Flaten, G., Svendsen, J. S. M., and Stensen, W. (2011) A synthetic antimicrobial peptidomimetic (LTX 109): Stereochemical impact on membrane disruption. *J. Med. Chem. 54*, 5786-5795.

5. Langham, A., and Kaznessis, Y. N. (2010) *Molecular simulations of antimicrobial peptides.* Humana Press.

6. Mihajlovic, M., and Lazaridis, T. (2010) Antimicrobial peptides in toroidal and cylindrical pores. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1798*, 1485-1493.

7. Tsai, C., Hsu, N., Wang, C., Lu, C., Chang, Y., Tsai, H. G., and Ruaan, R. (2009) Coupling molecular dynamics simulations with experiments for the rational design of indolicidin-analogous antimicrobial peptides. *J. Mol. Biol. 392*, 837-854.

8. Jo, S., Lim, J. B., Klauda, J. B., and Im, W. (2009) *CHARMM-GUI membrane builder for mixed bilayers and its application to yeast membranes.* Cell Press.

9. Jo, S., Kim, T., Iyer, V. G., and Im, W. (2008) CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry. 29*, 1859-1865.

10. Jo, S., Kim, T., and Im, W. (2007) Automated builder and database of Protein/Membrane complexes for molecular dynamics simulations. *PLoS ONE. 2*, e880.

11. Brooks B R et, a. (1983) *CHARMM-a program for macromolecular energy, minimization, and dynamics calculations.*

12. Tieleman, D. P. (2006) *Membrane protein simulations with a united-atom lipid and all-atom protein model: Lipid protein interactions, side chain transfer free energies and model proteins.*

13. Zhao, W., Rog, T., Gurtovenko, A. A., Vattulainen, I., and Karttunen, M. (2007) *Atomic-scale structure and electrostatics of anionic palmitoyloleoylphosphatidylglycerol lipid bilayers with na+ counterions.* Cell Press.

14. Freddolino, P. L., Arkhipov, A., Shih, A. Y., Yin, Y., Chen, Z., and Schulten, K. (2008) Application of Residue-Based and Shape-Based Coarse Graining to Biomolecular Simulations. in *Coarse-Graining of Condensed Phase and Biomolecular Systems* (G. A. Voth, Ed.) pp 229, Chapman and Hall/CRC Press, Taylor and Francis Group.

15. Ash, W. L., Zlomislic, M. R., Oloo, E. O., and Tieleman, D. P. (2004) Computer simulations of membrane proteins. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1666*, 158-189.

16. Case, D. A., Darden, T. A., Cheatham, I., T.E., Simmerling, C. L., Wang, J., Duke, R. E., R. Luo, K. M., Merz, K. M., Pearlman, D. A., Crowley, M., Walker, R. C., Zhang, W., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Wong, K. F., Paesani, F., Wu, X., Brozell, S., Tsui, V., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Beroza, P., Mathews, D. H., Schafmeister, C., Ross, W. S., and Kollman, P. A. (2006) *AMBER9.*

17. Case, D. A., Darden, T. A., Cheatham III, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Hayik, S., Roitberg, A., Seabra, G., Swails, J., Goetz, A. W., Kolossváry, I., Wong, K. F., Paesani, J., Vanicek, R., Wolf, M., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Wang, J., Hsieh, G., Cui, D. R., Roe, D. H., Mathews, M. G., Seetin, M. G., Salomon-Ferrer, R., Sagui, C., Babin, V.,

Luchko, T., Gusarov, S., Kovalenko, A., and Kollman, P. A. (2012) *AMBER 12.* University of California, San Francisco.

18. Bartolotti, L. (2009) *Creation of lipids for insight II.*

19. The PyMOL Molecular Graphics System. *The PyMOL molecular graphics system, version 1.2r3pre, schrödinger, LLC.* The PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC ed.

20. Feig, M., Karanicolas, J., and Brooks III, C. L. (2004) MMTSB tool set: Enhanced sampling and multiscale modeling methods for applications in structural biology. *J. Mol. Graph. Model. 22*, 377-395.

21. Red Hat, I. (1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008) *Cygwin API reference* .

22. Toke, O., Cegelski, L., and Schaefer, J. (2006) Peptide antibiotics in action: Investigation of polypeptide chains in insoluble environments by rotational-echo double resonance. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1758*, 1314-1329.

23. Hicks, R. P., Bhonsle, J. B., Venugopal, D., Koser, B. W., and Magill, A. J. (2007; 2007) De novo design of selective antibiotic peptides by incorporation of unnatural amino acids. *J. Med. Chem. 50*, 3026-3036.

24. Bhonsle, J. B., Venugopal, D., Huddler, D. P., Magill, A. J., and Hicks, R. P. (2007) Application of 3D-QSAR for identification of descriptors defining bioactivity of antimicrobial peptides. *J. Med. Chem. 50*, 6545-6553.

25. Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., Borschel, R., Mueller, A., Russell, A. L., Williams, B. C., and Hicks, R. P. (2010) Novel antimicrobial peptides that exhibit activity against select agents and other drug resistant bacteria. *Bioorg. Med. Chem. 18*, 5137-5147.

26. Russell, A. L., Kennedy, A. M., Spuches, A. M., Venugopal, D., Bhonsle, J. B., and Hicks, R. P. (2010) Spectroscopic and thermodynamic evidence for antimicrobial peptide membrane selectivity. *Chem. Phys. Lipids. 163*, 488-497.

27. Russell, A. L., Spuches, A. M., Williams, B. C., Venugopal, D., Klapper, D., Srouji, A. H., and Hicks, R. P. (2011) The effect of the placement and total charge of the basic amino acid clusters on antibacterial organism selectivity and potency. *Bioorg. Med. Chem. 19*, 7008-7022.

28. Russell, A. L., Kennedy, A. M., Spuches, A. M., Gibson, W. S., Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., and Hicks, R. P. (2011) Determining the effect of the incorporation of unnatural amino acids into antimicrobial peptides on the interactions with zwitterionic and anionic membrane model systems. *Chem. Phys. Lipids. 164*, 740-758.

29. Russell, A. L., Williams, B. C., Spuches, A., Klapper, D., Srouji, A. H., and Hicks, R. P. (2012) The effect of the length and flexibility of the side chain of basic amino acids on the binding of antimicrobial peptides to zwitterionic and anionic membrane model systems. *Bioorg. Med. Chem. 20*, 1723-1739.

30. Hicks, R. (2012) *Circular dichroism of para substituted phenylalanine.*

31. Hicks, R. P., Abercrombie, J. J., Wong, R. K., and Leung, K. P. (2012) Antimicrobial peptides containing unnatural amino acid exhibit potent bactericidal activity against ESKAPE pathogens. *Bioorg. Med. Chem.*

32. Andersen, H. C. (1980) Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys. 72*, 2384-2393.

33. Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., and Pedersen, L. G. (1995) *J. Chem. Phys. 103*, 8577.

34. Feller, S. E., Huster, D., and Gawrisch, K. (1999) *J. Am. Chem. Soc. 121*, 8963.

35. Huang, H. W. (2006) Molecular mechanism of antimicrobial peptides: The origin of cooperativity. *Biochimica et Biophysica Acta (BBA) Biomembranes. 1758*, 1292-1302.

36. Som, A., Vemparala, S., Ivanov, I., and Tew, G. N. (2008) Synthetic mimics of antimicrobial peptides. *Peptide Science. 90*, 83-93.

37. Beck, D. A. C., Alonso, D. O. V., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences. 105*, 12259-12264.

38. Russell, A. L. (2011) *Thermodynamic and spectroscopic investigations of novel antimicrobial peptides containing unnatural amino acids with model membrane systems.* Doctor of Philosophy ed., East Carolina University, Greenville, North Carolina.

CHAPTER 6

SUMMARY AND FUTURE RESEARCH DIRECTIONS

Force field (FF) parameters have been developed for numerous unnatural amino acids including β-alanine, gamma aminobutyric acid, 6-aminohexanoic acid, tetrahydroisoquinoline carboxylic acid, with the octahydroindole carboxylic acid, ornithine , diaminobutanoic acid , diaminopropanoic acid , aminocyclohexanoic acid, aminocyclopentanoic acid, p-Cl phenylalanine, and p-F phenylalanine. In addition, a density functional approach (DFT) approach was used to create FF parameters for three model lipid molecules: sodium dodecyl sulfate (SDS), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), and 1-palmitoyl-2-oleoyl-sn-glycero-3-[phosphor-rac-(1-glycerol)] (sodium salt).

The structural features of the unnatural amino acids have been investigated using capped tri-peptides, Ala-X-Ala, where the X is the unnatural amino acid. The application is supported by the consistency with the simulation data resulting from the 20 naturally occurring amino acids which were also simulated.[1] Ramachandran plots have provided support for the mixed secondary structure observed in experimental CD studies. The electrostatic surface potential maps, which distribute the calculated charges from the previously determined FF parameters across the solvent accessible surface of each residue, shows the change in charge distribution (Chapter 2). Comparison of some the unnatural amino acids to the most structurally similar standard residues reflects the changes which result from inclusion of the unnatural amino acids in a peptide sequence. The structural information is meant to provide a "toolbox" for the medicinal chemist as they design AMPs with an emphasis on tailoring the peptide sequence, and charge distribution to the desired target membrane.

AMPs which include the unnatural amino acids have been simulated in three model membrane environments. The first was the smaller anionic SDS micelle which was able to electrostatically model microbial membrane surfaces. The second model membrane was created

using only the zwitterionic lipid POPC. The neutral overall charge resulting from a bilayer composed solely of POPC models the electrostatic charge distribution associated with mammalian cell surfaces. The final model membrane was composed of a 4:1 ratio of POPC/POPG which represents the net anionic surface associated with many bacterial species' membranes.

AMPs electrostatically model the parent peptide, the naturally occurring AMP magainin 2.[2-4] These peptides were developed based on structural features associated with pharmacological activity for select bacterial membranes. The activity of a particular AMP can be tailored to specific species based upon the lipid composition of the target membrane surface.

Two of the peptides have been studied using molecular dynamics (MD) simulations in model membrane environments. The first model was comprised of 62 SDS molecules. The micelle provided the necessary electrostatic charge requisite for AMP secondary structure to develop. Peptide **23** was the first peptide studied as a reference for the impact of increasing the number of carbon atoms in the peptide backbone. Peptide **23** links the bicyclic Tic Oic dipeptides with the standard amino acid glycine. The small size associated with the Gly residue enables increased flexibility relative to the other chiral standard amino acids. The mixed, alpha helical and beta turn/sheet character observed in CD studies is supported by the Ramachandran plot resulting from the average phi and psi angles over the last 3 ns of the MD simulation (Figure 4.20 A).[5]

The second peptide investigated with the micelle membrane model was peptide **36** which links the bicyclic Tic Oic dipeptides with the nonstandard amino acid β-alanine. Increasing the number of carbon atoms in the peptide backbone increases the flexibility of the peptide enabling

unique secondary structures to develop. The increase in backbone carbons caused the peptide to interact with micelle in a structurally unique way as compared to the interaction occurring with peptide **23** and the micelle. The presence of a beta carbon in the peptide backbone prohibited the calculation of the Ramachandran phi/psi distribution for the beta-alanine residues however the remaining peptides reflect the mixed alpha helical and beta turn/sheet character which is consistent with the CD data (Figure 4.20 B).

The second set of bilayer model membranes were simulated in an environment which was designed to be consistent with the *in vitro* data which demonstrated the pharmacologic activity of the unnatural amino acid containing AMPs.

MD simulations of mixed bilayers composed of a 4:1 ratio of POPC to POPG have been conducted in the presence of peptides **23**. Control simulations without peptides present were also conducted for reference. The mixed bilayer in the presence of peptide **23** showed a membrane thinning effect after a 250 ns production simulation during which the RMSD of the bilayer did not stabilize but the peptide did relative to the start of the production simulation (Figure 5.11).

The presence of peptide **23** appears to have an influence on the distribution of the anionic lipids within the membrane. Two lipids were observed to interact with peptide **23** during the course of the MD simulation. The remainder of peptide **23** was positioned in a region of low lipid density (Figure 5.18). The C terminal hydrophobic rings appear to be reaching the aliphatic tails of the lipid molecules.

The final model was also simulated with peptide **23** and was intended to model the electrostatic properties associated with mammalian membranes. The control bilayer consisted of

a confluent POPC bilayer. Despite the observed intermolecular interactions occurring between peptide **23** and the control bilayer, no membrane thinning was observed during the 150 ns trajectory during which the RMSD stabilized (Figure 5.19). The average phi and psi angles depicted in Figure 5.15 and Figure 5.23 suggest peptide **23** has an alpha helical preference as evidenced by the large number of residues populating the alpha helical preferred region of the

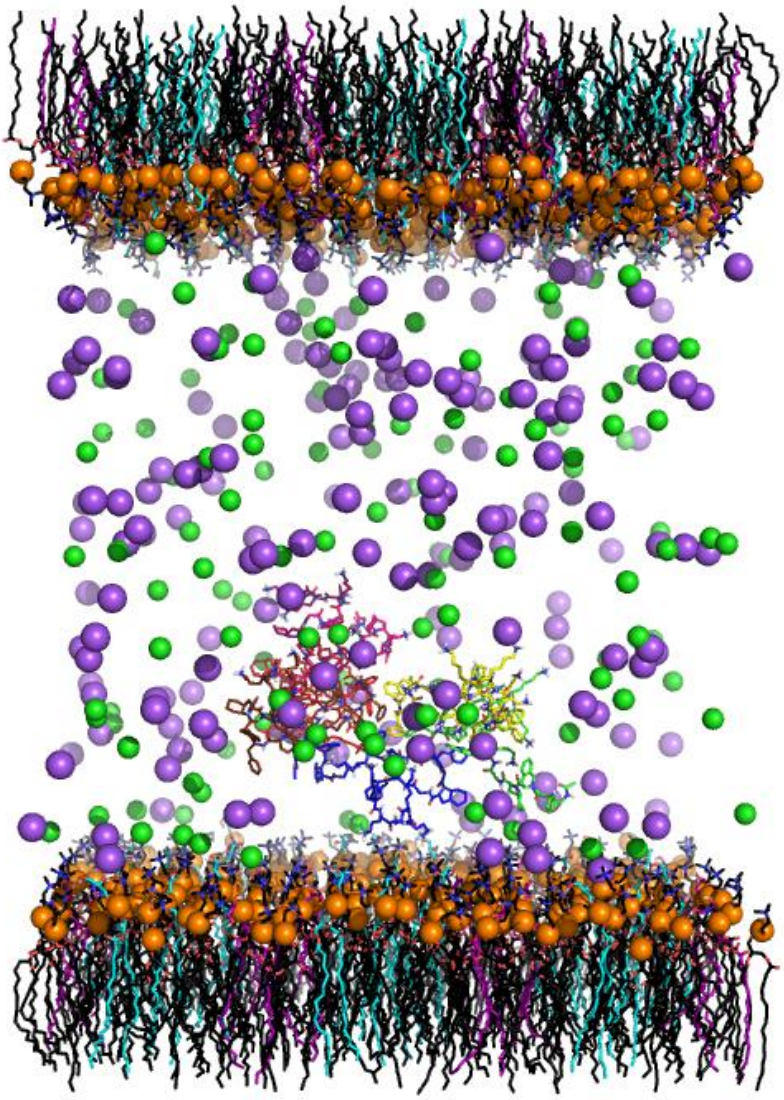


Figure 6.1 6 Peptide **23** with a mixed bilayer. POPC lipids are shown with black carbon atoms, phosphate atoms are noted with orange spheres. A racemic mixture of POPG is noted in cyan and purple carbon skeleton. Sodium atoms are noted with purple spheres and chloride atoms are noted with green spheres.

Ramachandran plots. The requisite electrostatic pull necessary for the secondary structure of the peptide has limited the utility of solvated peptide simulations. The random coil associated with AMPs in solution was also observed in the solvated MD simulation.

FUTURE RESEARCH

Future research efforts will continue the MD simulation of the mixed bilayer with peptide **23** and will also include simulation of peptide **36** in both the model bacterial and mammalian membranes. Comparison of the two peptides relative to each other can help decipher the ambiguous secondary structure suggested by the CD spectra. Analysis of the preferred angles associated with the peptide in the presence of each bilayer can help further our understanding of the mechanism through which membrane disruption occurs.

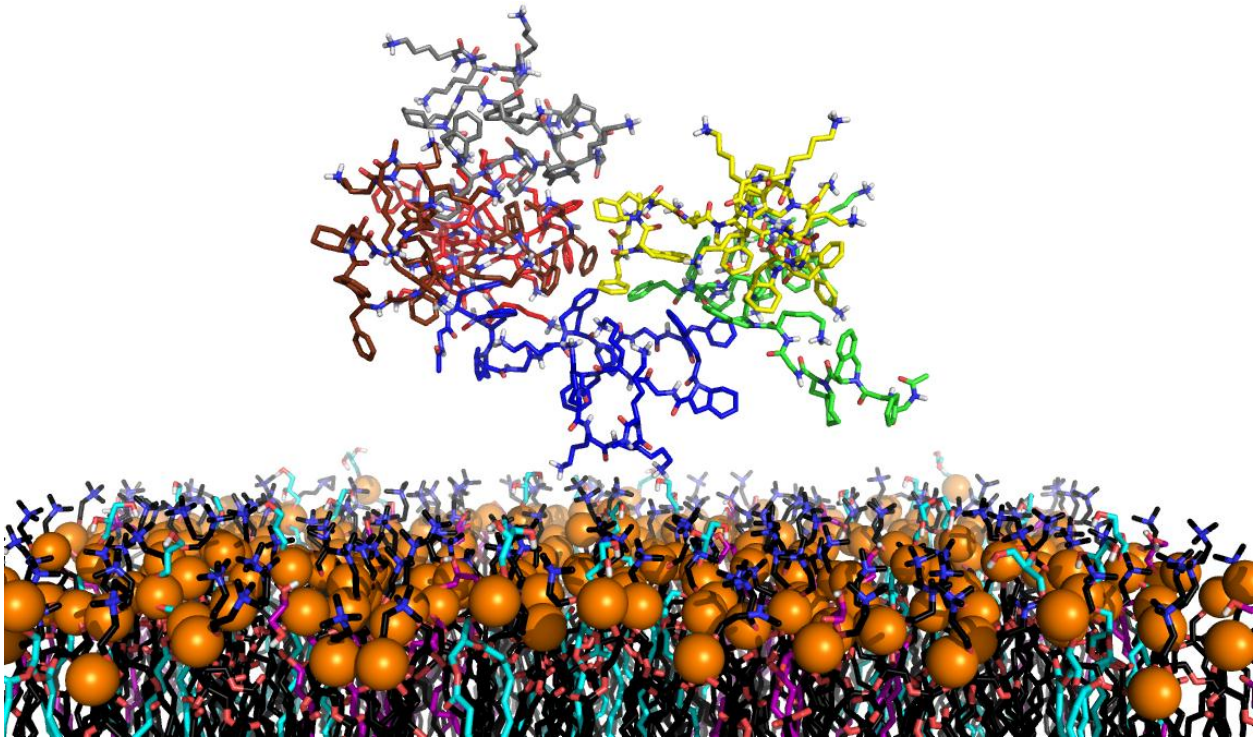


Figure 6.2 6 Peptide **23** with a mixed bilayer. POPC lipids are shown with black carbon atoms, phosphate atoms are noted with orange spheres. A racemic mixture of POPG is noted in cyan and purple carbon skeleton.

Larger computational resources would enable larger simulations to be conducted. MD simulations with a sufficient peptide to lipid ratio would provide valuable insight into the activity of this collection of AMPs. Simulations of numerous peptides, on a larger bilayer (Figure 6.1 & Figure 6.2), and a larger time scale would provide insight into the mechanism through which the membranes are disrupted.

REFERENCES

1. Beck, D. A. C., Alonso, D. O. V., Inoyama, D., and Daggett, V. (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proceedings of the National Academy of Sciences. 105*, 12259-12264.

2. Hicks, R. P., Bhonsle, J. B., Venugopal, D., Koser, B. W., and Magill, A. J. (2007; 2007) De novo design of selective antibiotic peptides by incorporation of unnatural amino acids. *J. Med. Chem. 50*, 3026-3036.

3. Bhonsle, J. B., Venugopal, D., Huddler, D. P., Magill, A. J., and Hicks, R. P. (2007; 2007) Application of 3D-QSAR for identification of descriptors defining bioactivity of antimicrobial peptides. *J. Med. Chem. 50*, 6545-6553.

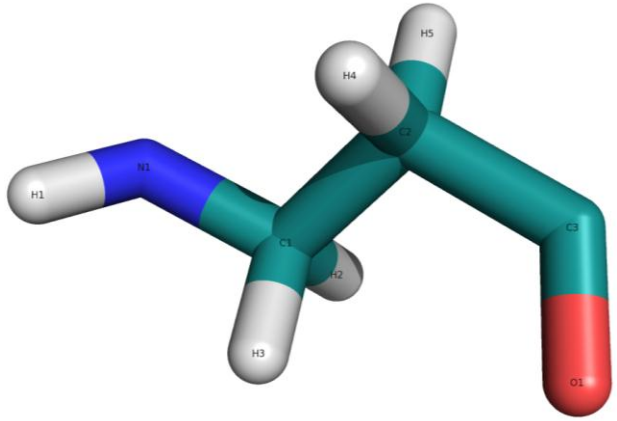
4. Venugopal, D., Klapper, D., Srouji, A. H., Bhonsle, J. B., Borschel, R., Mueller, A., Russell, A. L., Williams, B. C., and Hicks, R. P. (2010) Novel antimicrobial peptides that exhibit activity against select agents and other drug resistant bacteria. *Bioorg. Med. Chem. 18*, 5137-5147.

5. Russell, A. L., Kennedy, A. M., Spuches, A. M., Venugopal, D., Bhonsle, J. B., and Hicks, R. P. (2010) Spectroscopic and thermodynamic evidence for antimicrobial peptide membrane selectivity. *Chem. Phys. Lipids. 163*, 488-497.

APPENDIX A
FORCE FIELD PARAMETERS FOR NON-STANDARD AMINO ACIDS AND LIPID MOLECULES

UNNATURAL & NON-STANDARD AMINO ACIDS

B-ALANINE



.prepi file:

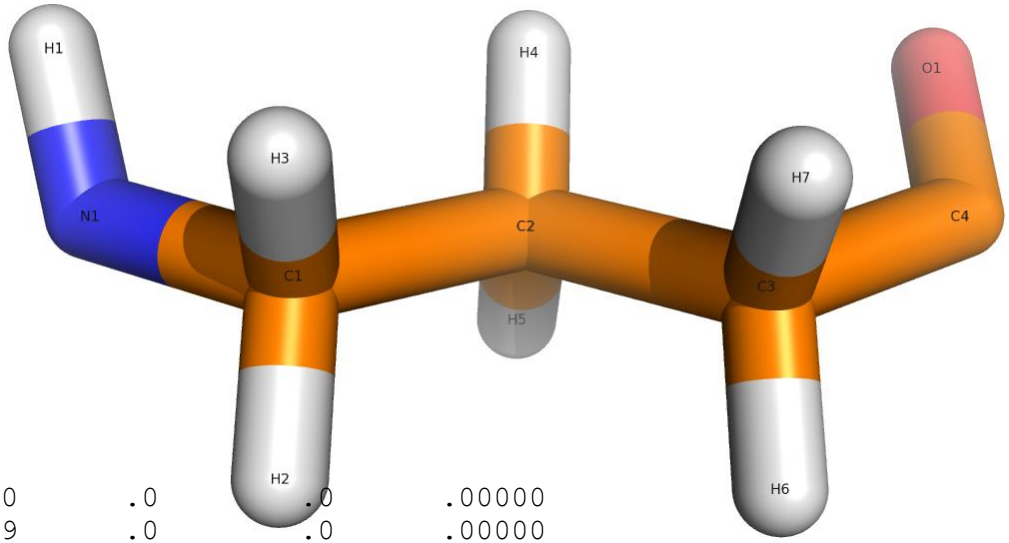
```
    0    0    2
Bala.prepi input
molecule.res
BAL   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0
.0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.75806
   5  H1    H     E    4   3   2     0.994   114.626  -155.677   0.38188
   6  C1    CT    M    4   3   2     1.450    17.741    97.543   0.30232
   7  H2    H1    E    6   4   3     1.082   108.840  -147.036   0.02196
   8  H3    H1    E    6   4   3     1.080   108.254    96.088   0.02196
   9  C2    CT    M    6   4   3     1.528   112.525   -25.145  -0.29408
  10  H4    HC    E    9   6   4     1.086   109.491    56.840   0.08374
  11  H5    HC    E    9   6   4     1.083   109.586   -60.944   0.08374
  12  C3    C     M    9   6   4     1.518   111.096   175.043   0.75195
  13  O1    O     E   12   9   6     1.203   121.588   -40.508  -0.59543
LOOP
IMPROPER
   -M   C1   N1   H1
   C2   +M   C3   O1
DONE
STOP
```

.frcmod file:

```
remark goes here
MASS
N  14.010        0.530               same as n
H  1.008         0.161               same as hn
CT 12.010        0.878               same as c3
H1 1.008         0.135               same as hc
HC 1.008         0.135               same as hc
C  12.010        0.616               same as c
O  16.000        0.434               same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-H1  337.30   1.092       same as c3-hc
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-CT   65.900     112.130   same as c3-c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
H1-CT-H1   39.400     108.350   same as hc-c3-hc
H1-CT-CT   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
HC-CT-HC   39.400     108.350   same as hc-c3-hc
HC-CT-C    47.200     109.680   same as c -c3-hc
```

```
DIHE
N -CT-CT-HC    1    0.156          0.000          3.000      same as X -c3-c3-X
N -CT-CT-C     1    0.156          0.000          3.000      same as X -c3-c3-X
H -N -CT-H1    1    0.000          0.000          2.000      same as X -c3-n -X
H -N -CT-CT    1    0.000          0.000          2.000      same as X -c3-n -X
CT-CT-C -O     1    0.000        180.000          2.000      same as X -c -c3-X
H1-CT-CT-HC    1    0.156          0.000          3.000      same as X -c3-c3-X
H1-CT-CT-C     1    0.156          0.000          3.000      same as X -c3-c3-X
HC-CT-C -O     1    0.000        180.000          2.000      same as X -c -c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  H1          1.4870  0.0157            same as hc
  HC          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

GABA



.prepi file:

```
    0    0    2

Gaba.prepi input
gaba.res
GAB   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0         .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0         .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1         .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.88538
   5  H1    H     E    4   3   2     0.998    98.122    -0.406   0.40641
   6  C1    CT    M    4   3   2     1.446    31.750  -131.165   0.41091
   7  H2    H1    E    6   4   3     1.082   108.978  -140.286  -0.01603
   8  H3    H1    E    6   4   3     1.086   107.764   104.877  -0.01603
   9  C2    CT    M    6   4   3     1.529   113.470   -16.683   0.07878
  10  H4    HC    E    9   6   4     1.083   109.841    57.387   0.00995
  11  H5    HC    E    9   6   4     1.087   109.508   -59.206   0.00995
  12  C3    CT    M    9   6   4     1.528   111.952   178.767  -0.42016
  13  H6    HC    E   12   9   6     1.087   110.743    61.542   0.07882
  14  H7    HC    E   12   9   6     1.088   109.911   -55.984   0.07882
  15  C4    C     M   12   9   6     1.521   112.328  -174.673   0.87128
  16  O1    O     E   15  12   9     1.201   122.132    31.524  -0.60732
LOOP
IMPROPER
   -M   C1   N1   H1
   C3   +M   C4   O1
DONE
STOP
```

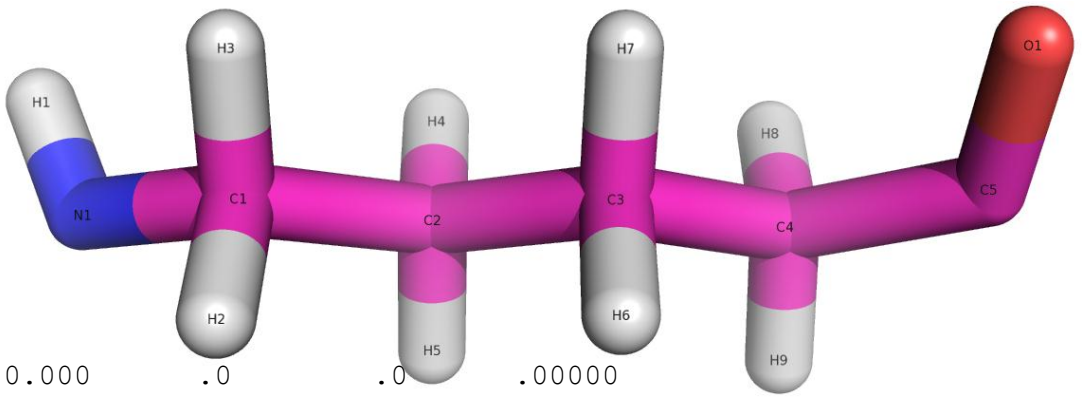
.frcmod file:

```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
H1 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-H1  337.30   1.092       same as c3-hc
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-CT   65.900     112.130   same as c3-c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-CT-CT   63.200     110.630   same as c3-c3-c3
H1-CT-H1   39.400     108.350   same as hc-c3-hc
H1-CT-CT   33.235     109.490   Calculated with empirical approach
CT-CT-C    63.800     110.530   same as c -c3-c3
HC-CT-HC   39.400     108.350   same as hc-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
HC-CT-C    47.200     109.680   same as c -c3-hc

DIHE
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
```

```
H -N -CT-H1    1    0.000          0.000           2.000      same as X -c3-n -X
H -N -CT-CT    1    0.000          0.000           2.000      same as X -c3-n -X
CT-CT-CT-HC    1    0.156          0.000           3.000      same as X -c3-c3-X
CT-CT-CT-C     1    0.156          0.000           3.000      same as X -c3-c3-X
H1-CT-CT-HC    1    0.156          0.000           3.000      same as X -c3-c3-X
H1-CT-CT-CT    1    0.156          0.000           3.000      same as X -c3-c3-X
CT-CT-C -O     1    0.000        180.000           2.000      same as X -c -c3-X
HC-CT-CT-HC    1    0.156          0.000           3.000      same as X -c3-c3-X
HC-CT-CT-C     1    0.156          0.000           3.000      same as X -c3-c3-X
HC-CT-C -O     1    0.000        180.000           2.000      same as X -c -c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  H1          1.4870  0.0157            same as hc
  HC          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

6-AMINOHEXANOIC ACID (6AH)

.prepi file:




```
    0    0    2

This is a remark line
6ahx.rf
6AH   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0       .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0       .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1       .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208   180.000  -0.66102
   5  H1    H     E    4   3   2     0.993   111.315  -138.505   0.36465
   6  C1    CT    M    4   3   2     1.450    22.498   108.272  -0.17774
   7  H2    H1    E    6   4   3     1.082   107.805  -135.812   0.12694
   8  H3    H1    E    6   4   3     1.084   107.618   108.460   0.12694
   9  C2    CT    M    6   4   3     1.528   113.419   -13.797   0.06894
  10  H4    HC    E    9   6   4     1.088   109.215    57.397   0.02278
  11  H5    HC    E    9   6   4     1.086   108.415   -58.817   0.02278
  12  C3    CT    M    9   6   4     1.529   112.090   179.213   0.19113
  13  H6    HC    E   12   9   6     1.087   109.494    56.776  -0.02482
  14  H7    HC    E   12   9   6     1.083   110.479   -60.333  -0.02482
  15  C4    CT    M   12   9   6     1.529   112.375   178.209  -0.42852
  16  H8    HC    E   15  12   9     1.088   109.770    56.833   0.08525
  17  H9    HC    E   15  12   9     1.086   110.411   -60.580   0.08525
  18  C5    C     M   15  12   9     1.519   112.250   175.180   0.82889
  19  O1    O     E   18  15  12     1.203   121.997   -40.631  -0.60663
LOOP
IMPROPER
   -M   C1   N1   H1
   C4   +M   C5   O1

DONE
STOP
```

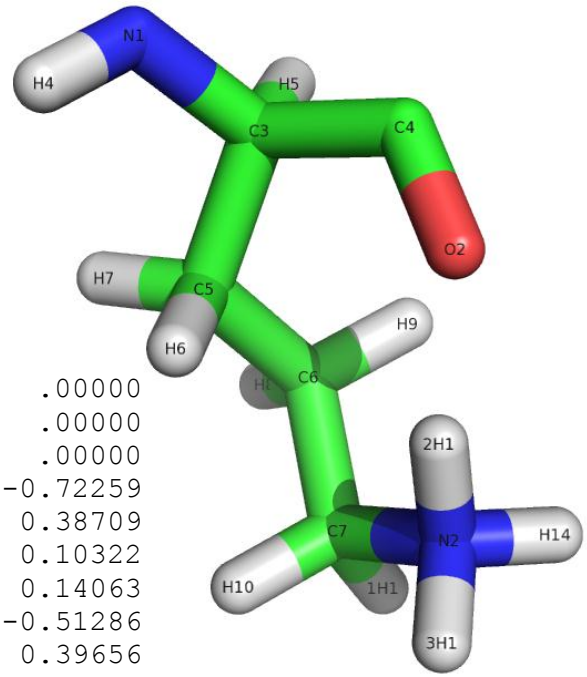
.frcmod file:

```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
H1 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-H1  337.30   1.092       same as c3-hc
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-CT   65.900     112.130   same as c3-c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-CT-CT   63.200     110.630   same as c3-c3-c3
H1-CT-H1   39.400     108.350   same as hc-c3-hc
H1-CT-CT   33.235     109.490   Calculated with empirical approach
HC-CT-HC   39.400     108.350   same as hc-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
CT-C -O    68.000     123.110   same as c3-c -o
HC-CT-C    47.200     109.680   same as c -c3-hc

DIHE
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
```

```
H -N -CT-H1    1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-CT    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-CT    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-CT-CT    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-C     1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O     1    0.000       180.000          2.000      same as X -c -c3-X
HC-CT-CT-C     1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-C -O     1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  H1          1.4870  0.0157            same as hc
  HC          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

ORNATHINE

.prepi file:




```
    0    0    2

This is a remark line
orn.res
ORN   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522    111.1       .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.72259
   5  H4    H     E    4   3   2     0.997    98.771   -60.443   0.38709
   6  C3    CT    M    4   3   2     1.438   134.505   152.660   0.10322
   7  C5    CT    3    6   4   3     1.537   110.731    61.104   0.14063
   8  N2    N3    3    7   6   4     1.501   110.900  -171.194  -0.51286
   9  H8    H     E    8   7   6     1.023   106.863   -53.870   0.39656
  10  H9    H     E    8   7   6     1.010   111.899    65.026   0.39656
  11  H10   H     E    8   7   6     1.010   111.981  -174.205   0.39656
  12  H6    HP    E    7   6   4     1.079   109.737   -52.128   0.09785
  13  H7    HP    E    7   6   4     1.080   111.584    70.203   0.09785
  14  H5    H1    E    6   4   3     1.084   106.529   -56.396   0.11861
  15  C4    C     M    6   4   3     1.532   113.557  -173.482   0.65633
  16  O2    O     E   15   6   4     1.216   118.260  -152.734  -0.55580


LOOP

IMPROPER
   -M   C3   N1   H4
   C3   +M   C4   O2

DONE
STOP
```

.frcmod file:

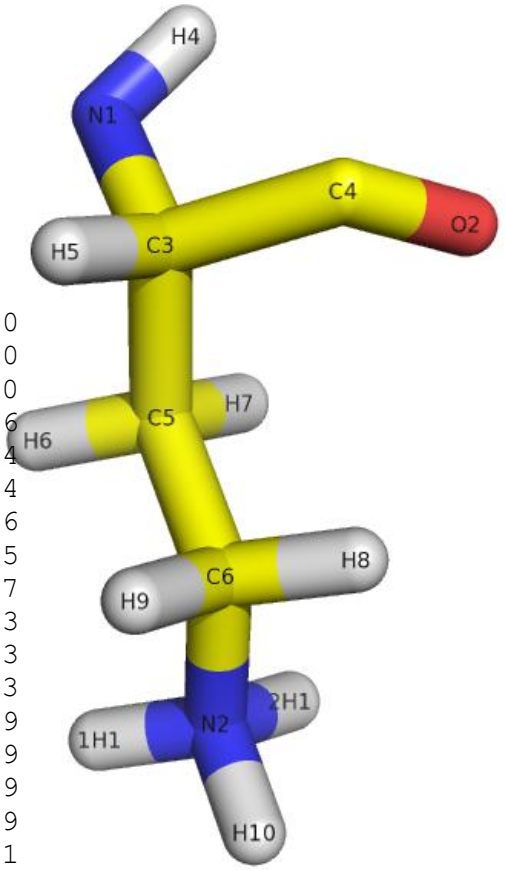
```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
N3 14.010        0.530              same as n4
HP 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
CT-N3  293.60   1.499       same as c3-n4
CT-HP  337.30   1.092       same as c3-hc
N3-H   369.00   1.033       same as hn-n4
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-N3   66.000     108.930   same as c3-c3-n4
CT-CT-HP   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
CT-N3-H    46.200     110.110   same as c3-n4-hn
N3-CT-HP   49.000     107.900   same as hc-c3-n4
H -N3-H    40.500     108.110   same as hn-n4-hn
```

```
HP-CT-HP   39.400     108.350   same as hc-c3-hc
H1-CT-C    47.200     109.680   same as c -c3-hc

DIHE
N -CT-CT-N3   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HP   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-H1   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-N3-H    1    0.156         0.000          3.000      same as X -c3-n4-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
N3-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
N3-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H -N3-CT-HP   1    0.156         0.000          3.000      same as X -c3-n4-X
HP-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HP-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  N3          1.8240  0.1700            same as n4
  HP          1.4870  0.0157            same as hc
  H1          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

DIAMINOBUTANOIC ACID

.prepi file:

```
    0    0    2
This is a remark line
dab.res
DAB   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.02446
   5  H4    H     E    4   3   2     1.012    92.662   -50.579   0.10134
   6  C3    CT    M    4   3   2     1.457   150.805   128.660   0.13954
   7  C5    CT    3    6   4   3     1.550   110.062    75.132  -0.10546
   8  C6    CT    3    7   6   4     1.540   113.939  -179.437   0.30085
   9  N2    N3    3    8   7   6     1.476   111.278   179.624   0.43497
  10  H10   H     E    9   8   7     1.010   109.825  -179.358   0.05783
  11  H11   H     E    9   8   7     1.010   109.967   -59.192   0.05783
  12  H12   H     E    9   8   7     1.010   109.681    61.017   0.05783
  13  H8    HP    E    8   7   6     1.092   110.719   -59.900  -0.05779
  14  H9    HP    E    8   7   6     1.091   110.059    58.261  -0.05779
  15  H6    HC    E    7   6   4     1.092   109.171   -57.294   0.05979
  16  H7    HC    E    7   6   4     1.091   108.859    57.841   0.05979
  17  H5    H1    E    6   4   3     1.093   108.411   -43.265   0.17641
  18  C4    C     M    6   4   3     1.529   108.530  -160.350  -0.29603
  19  O2    O     E   18   6   4     1.227   118.769   -82.936   0.09536
LOOP
IMPROPER
   -M   C3   N1   H4
   C3   +M   C4   O2

DONE
STOP
```

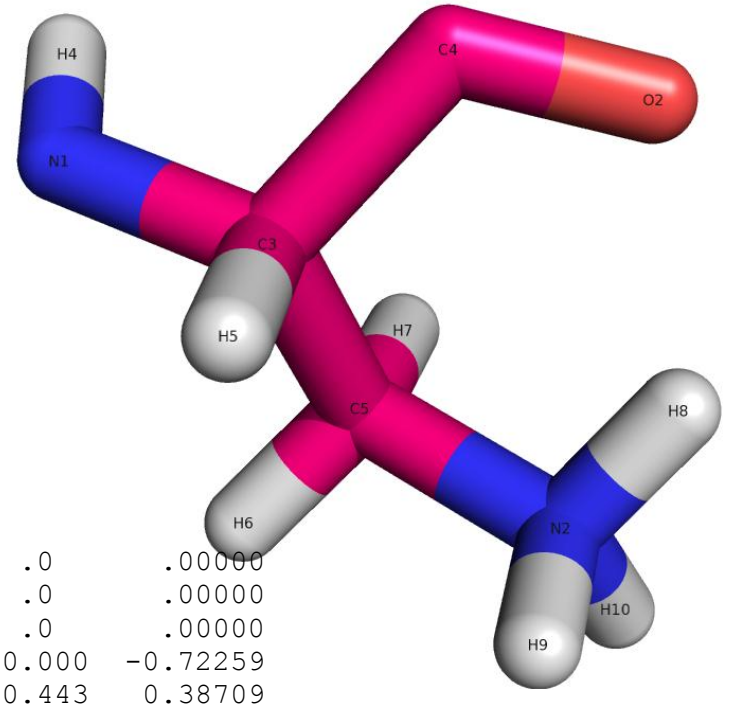


.frcmod file:

```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
N3 14.010        0.530              same as n4
HP 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
CT-HC  337.30   1.092       same as c3-hc
CT-N3  293.60   1.499       same as c3-n4
CT-HP  337.30   1.092       same as c3-hc
N3-H   369.00   1.033       same as hn-n4
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
CT-CT-N3   66.000     108.930   same as c3-c3-n4
```

```
CT-CT-HP   46.400     110.050   same as c3-c3-hc
CT-N3-H    46.200     110.110   same as c3-n4-hn
N3-CT-HP   49.000     107.900   same as hc-c3-n4
H -N3-H    40.500     108.110   same as hn-n4-hn
HP-CT-HP   39.400     108.350   same as hc-c3-hc
HC-CT-HC   39.400     108.350   same as hc-c3-hc
H1-CT-C    47.200     109.680   same as c -c3-hc
DIHE
N -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-H1   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CT-N3   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HP   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-CT-N3-H    1    0.156         0.000          3.000      same as X -c3-n4-X
CT-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
N3-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
H -N3-CT-HP   1    0.156         0.000          3.000      same as X -c3-n4-X
HP-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
IMPROPER
NONBON
  N           1.8240  0.1700          same as n
  H           0.6000  0.0157          same as hn
  CT          1.9080  0.1094          same as c3
  N3          1.8240  0.1700          same as n4
  HP          1.4870  0.0157          same as hc
  HC          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  C           1.9080  0.0860          same as c
  O           1.6612  0.2100          same as o
```

DIAMINOPROPIONIC ACID



.prepi file:

```
    0    0    2
This is a remark line
dpr.res
DPR   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.72259
   5  H4    H     E    4   3   2     0.997    98.771   -60.443   0.38709
   6  C3    CT    M    4   3   2     1.438   134.505   152.660   0.10322
   7  C5    CT    3    6   4   3     1.537   110.731    61.104   0.14063
   8  N2    N3    3    7   6   4     1.501   110.900  -171.194  -0.51286
   9  H8    H     E    8   7   6     1.023   106.863   -53.870   0.39656
  10  H9    H     E    8   7   6     1.010   111.899    65.026   0.39656
  11  H10   H     E    8   7   6     1.010   111.981  -174.205   0.39656
  12  H6    HP    E    7   6   4     1.079   109.737   -52.128   0.09785
  13  H7    HP    E    7   6   4     1.080   111.584    70.203   0.09785
  14  H5    H1    E    6   4   3     1.084   106.529   -56.396   0.11861
  15  C4    C     M    6   4   3     1.532   113.557  -173.482   0.65633
  16  O2    O     E   15   6   4     1.216   118.260  -152.734  -0.55580
LOOP
IMPROPER
   -M   C3   N1   H4
   C3   +M   C4   O2
```

```
DONE
STOP
```

.frcmod file:

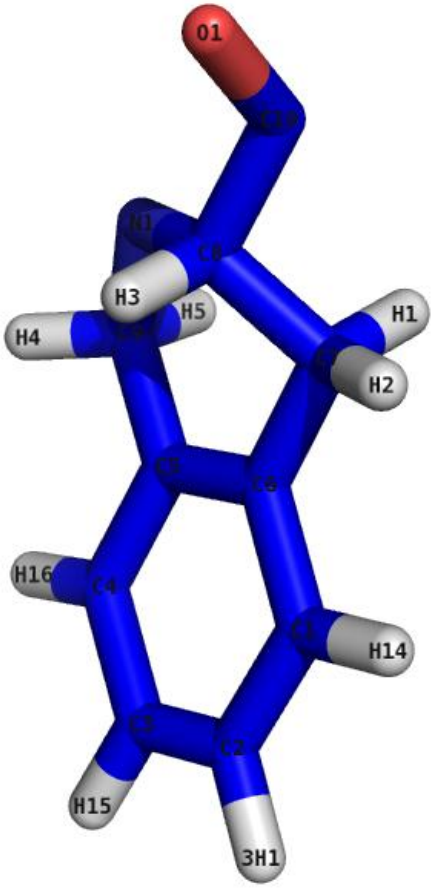
```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
N3 14.010        0.530              same as n4
HP 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
CT-N3  293.60   1.499       same as c3-n4
CT-HP  337.30   1.092       same as c3-hc
N3-H   369.00   1.033       same as hn-n4
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-N3   66.000     108.930   same as c3-c3-n4
CT-CT-HP   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
```

```
CT-N3-H    46.200     110.110   same as c3-n4-hn
N3-CT-HP   49.000     107.900   same as hc-c3-n4
H -N3-H    40.500     108.110   same as hn-n4-hn
HP-CT-HP   39.400     108.350   same as hc-c3-hc
H1-CT-C    47.200     109.680   same as c -c3-hc

DIHE
N -CT-CT-N3   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HP   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-H1   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-N3-H    1    0.156         0.000          3.000      same as X -c3-n4-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
N3-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
N3-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H -N3-CT-HP   1    0.156         0.000          3.000      same as X -c3-n4-X
HP-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HP-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER

NONBON
  N           1.8240  0.1700          same as n
  H           0.6000  0.0157          same as hn
  CT          1.9080  0.1094          same as c3
  N3          1.8240  0.1700          same as n4
  HP          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  C           1.9080  0.0860          same as c
  O           1.6612  0.2100          same as o
```

TIC

.prepi file:

```
    0    0    2
TIC.prepi input file
tic.res
TIC   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.22346
   5  C9    CT    3    4   3   2     1.517    30.684   135.785  -0.15007
   6  C5    CA    S    5   4   3     1.539   112.401   -20.564   0.06301
   7  C4    CA    B    6   5   4     1.538   120.657   149.531  -0.26369
   8  C3    CA    B    7   6   5     1.540   119.821  -178.904  -0.16552
   9  C2    CA    B    8   7   6     1.540   120.083     0.761  -0.17165
  10  C1    CA    B    9   8   7     1.541   119.865     0.637  -0.26564
  11  C6    CA    S   10   9   8     1.539   119.776    -0.963   0.09029
  12  C7    CT    B   11  10   9     1.534   121.009   176.302  -0.13781
  13  H1    HC    E   12  11  10     1.139   109.955   104.833   0.06498
  14  H2    HC    E   12  11  10     1.140   110.150   -16.240   0.06498
  15  H14   HA    E   10   9   8     1.140   120.068   178.964   0.19124
  16  H13   HA    E    9   8   7     1.140   120.106  -179.739   0.17933
  17  H15   HA    E    8   7   6     1.141   119.899  -179.321   0.17953
  18  H16   HA    E    7   6   5     1.141   120.046     1.253   0.20101
  19  H4    H1    E    5   4   3     1.140   108.987   100.053   0.13156
  20  H5    H1    E    5   4   3     1.139   108.870  -141.226   0.13156
  21  C8    CT    M    4   3   2     1.519    93.510   -19.738   0.16362
  22  H3    H1    E   21   4   3     1.141   108.900   -70.608   0.08010
  23  C10   C     M   21   4   3     1.540   111.624   168.341   0.40091
  24  O1    O     E   23  21   4     1.510   120.243   103.499  -0.56430
LOOP
   C6   C5
   C8   C7
```



```
IMPROPER
   -M   C8   N1   C9
   C6   C4   C5   C9
   C5   C3   C4  H16
   C4   C2   C3  H15
   C3   C1   C2  H13
   C6   C2   C1  H14
   C5   C1   C6   C7
   C8   +M  C10   O1

DONE
STOP
```

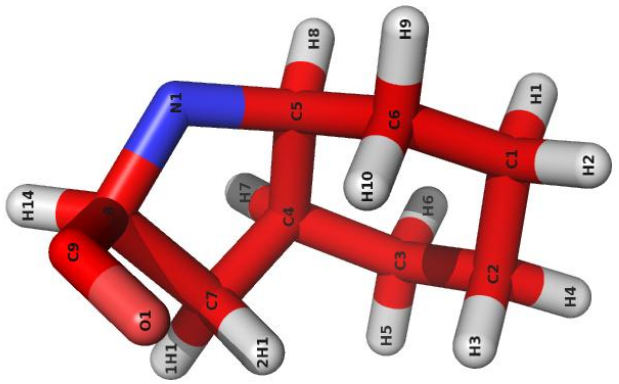
.frcmod file:


```
TIC.frcmod
MASS
N  14.010        0.530              same as n
CT 12.010        0.878              same as c3
CA 12.010        0.360              same as c2
HC 1.008         0.135              same as hc
HA 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -CT  330.60   1.460       same as c3-n
CT-CA  328.30   1.508       same as c2-c3
CT-H1  337.30   1.092       same as c3-hc
CA-CA  478.40   1.387       same as ca-ca
CA-HA  344.30   1.087       same as c2-hc
CT-HC  337.30   1.092       same as c3-hc
CT-CT  303.10   1.535       same as c3-c3
CT-C   328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o
```

```
ANGLE
N -CT-CA    66.341      111.760    Calculated with empirical approach
N -CT-H1    49.800      109.500    same as hc-c3-n
N -CT-CT    65.900      112.130    same as c3-c3-n
N -CT-C     66.700      111.560    same as c -c3-n
CT-N -CT    64.000      112.620    same as c3-n -c3
CT-CA-CA    64.300      123.420    same as c2-c2-c3
CA-CT-H1    47.000      110.490    same as c2-c3-hc
CA-CA-CA    67.200      119.970    same as ca-ca-ca
CA-CA-HA    50.300      119.700    same as c2-c2-hc
CA-CT-HC    47.000      110.490    same as c2-c3-hc
CA-CT-CT    63.700      110.960    same as c2-c3-c3
CT-CT-H1    46.400      110.050    same as c3-c3-hc
CT-CT-C     63.800      110.530    same as c -c3-c3
HC-CT-HC    39.400      108.350    same as hc-c3-hc
HC-CT-CT    46.400      110.050    same as c3-c3-hc
H1-CT-H1    39.400      108.350    same as hc-c3-hc
CT-C -O     68.000      123.110    same as c3-c -o
H1-CT-C     47.200      109.680    same as c -c3-hc

DIHE
N -CT-CA-CA    1    0.000          0.000          2.000      same as X -c2-c3-X
N -CT-CT-CA    1    0.156          0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC    1    0.156          0.000          3.000      same as X -c3-c3-X
N -CT-C -O     1    0.000        180.000          2.000      same as X -c -c3-X
CT-N -CT-CT    1    0.000          0.000          2.000      same as X -c3-n -X
CT-N -CT-H1    1    0.000          0.000          2.000      same as X -c3-n -X
CT-N -CT-C     1    0.000          0.000          2.000      same as X -c3-n -X
CT-CA-CA-CA    1    6.650        180.000          2.000      same as X -c2-c2-X
CT-CA-CA-HA    1    6.650        180.000          2.000      same as X -c2-c2-X
CT-CA-CA-CT    1    6.650        180.000          2.000      same as X -c2-c2-X
CA-CT-N -CT    1    0.000          0.000          2.000      same as X -c3-n -X
CA-CA-CA-CA    1    3.625        180.000          2.000      same as X -ca-ca-X
CA-CA-CA-HA    1    6.650        180.000          2.000      same as X -c2-c2-X
CA-CA-CT-HC    1    0.000          0.000          2.000      same as X -c2-c3-X
CA-CA-CT-CT    1    0.000          0.000          2.000      same as X -c2-c3-X
CA-CA-CT-H1    1    0.000          0.000          2.000      same as X -c2-c3-X
```

```
CA-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CA-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
HC-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
HA-CA-CA-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER
CA-CA-CA-CT         1.1         180.0          2.0        Using default value
CA-CA-CA-HA         1.1         180.0          2.0        Using default value

NONBON
  N           1.8240  0.1700          same as n
  CT          1.9080  0.1094          same as c3
  CA          1.9080  0.0860          same as ca
  HC          1.4870  0.0157          same as hc
  HA          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  C           1.9080  0.0860          same as c
  O           1.6612  0.2100          same as o
```

OIC



.prepi file:

```
    0    0    2

This is a remark line
oic.res
OIC   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2      0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1      1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0      1.522    111.1       .0      .00000
   4  N1    N     M    3   2   1      1.540   111.208  -180.000  -0.36092
   5  C5    CT    B    4   3   2      1.466   169.593   -80.218   0.07272
   6  C4    CT    3    5   4   3      1.537   101.904   139.335   0.06164
   7  C3    CT    3    6   5   4      1.533   114.181  -165.027  -0.12236
   8  C2    CT    3    7   6   5      1.532   113.578    45.134   0.02779
   9  C1    CT    3    8   7   6      1.530   110.937   -52.066   0.02579
  10  H1    HC    E    9   8   7      1.090   109.535   -62.841  -0.00620
  11  H2    HC    E    9   8   7      1.086   110.366  -179.864  -0.00620
  12  C6    CT    B    9   8   7      1.530   110.577    58.628  -0.08165
  13  H9    HC    E   12   9   8      1.087   110.739   179.813   0.03860
  14  H10   HC    E   12   9   8      1.085   109.477    62.635   0.03860
  15  H3    HC    E    8   7   6      1.087   110.439    69.153   0.01264
  16  H4    HC    E    8   7   6      1.087   109.305  -173.923   0.01264
  17  H5    HC    E    7   6   5      1.087   109.410   169.060   0.02786
  18  H6    HC    E    7   6   5      1.088   108.441   -75.479   0.02786
  19  H7    HC    E    6   5   4      1.087   106.955    75.198   0.00614
  20  C7    CT    B    6   5   4      1.534   102.764   -38.196   0.02406
  21  H11   HC    E   20   6   5      1.083   111.971   154.062   0.02667
  22  H12   HC    E   20   6   5      1.079   111.792   -84.941   0.02667
  23  H8    H1    E    5   4   3      1.082   109.960  -103.825   0.04973
  24  C8    CT    M    4   3   2      1.473    72.655    32.537  -0.03396
```

```
  25  H14   H1    E   24   4   3     1.081   108.152    61.857   0.05370
  26  C9    C     M   24   4   3     1.540   111.283   -57.274   0.62724
  27  O1    O     E   26  24   4     1.205   122.027   -97.747  -0.54906
LOOP
   C6   C5
   C8   C7
IMPROPER
   -M   C8   N1   C5
   C8   +M   C9   O1
DONE
STOP
```

.frcmod file:

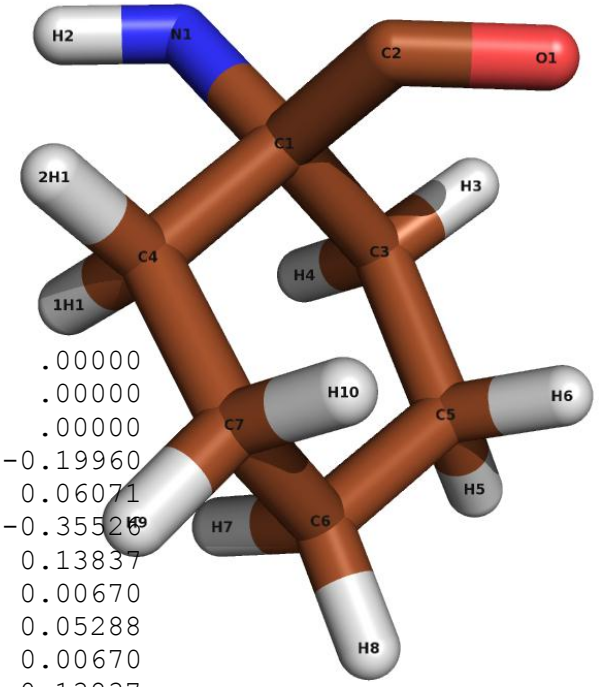
```
remark goes here
MASS
N  14.010        0.530              same as n
CT 12.010        0.878              same as c3
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
CT-HC  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-H1   49.800     109.500   same as hc-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
CT-N -CT   64.000     112.620   same as c3-n -c3
CT-CT-CT   63.200     110.630   same as c3-c3-c3
```

```
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-CT-H1   46.400     110.050   same as c3-c3-hc
HC-CT-HC   39.400     108.350   same as hc-c3-hc
CT-CT-C    63.800     110.530   same as c -c3-c3
CT-C -O    68.000     123.110   same as c3-c -o
H1-CT-C    47.200     109.680   same as c -c3-hc

DIHE
N -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
CT-N -CT-H1   1    0.000         0.000          2.000      same as X -c3-n -X
CT-N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
HC-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER


NONBON
  N           1.8240  0.1700          same as n
  CT          1.9080  0.1094          same as c3
  HC          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  C           1.9080  0.0860          same as c
  O           1.6612  0.2100          same as o
```

ACH

.prepi file:

```
    0    0    2

This is a remark line
ach.res
ACH   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.19960
   5  H2    H     E    4   3   2     1.109    95.742   175.896   0.06071
   6  C1    CT    M    4   3   2     1.508    24.553   -11.597  -0.35526
   7  C3    CT    3    6   4   3     1.547   108.577  -128.014   0.13837
   8  C5    CT    3    7   6   4     1.542   116.449  -147.120   0.00670
   9  C6    CT    3    8   7   6     1.540   109.799    51.627   0.05288
  10  C7    CT    3    9   8   7     1.541   109.452   -66.132   0.00670
  11  C4    CT    B   10   9   8     1.541   111.183    55.175   0.13837
  12  H11   HC    E   11  10   9     1.140    95.045    89.130   0.01269
  13  H12   HC    E   11  10   9     1.140   108.659  -157.919   0.01269
  14  H9    HC    E   10   9   8     1.139   109.318   176.720   0.00419
  15  H10   HC    E   10   9   8     1.136   108.748   -64.837   0.00419
  16  H7    HC    E    9   8   7     1.135   109.070    54.776   0.00145
  17  H8    HC    E    9   8   7     1.138   109.742   174.732   0.00145
  18  H5    HC    E    8   7   6     1.141   108.627   170.944   0.00419
  19  H6    HC    E    8   7   6     1.138   110.257   -69.117   0.00419
  20  H3    HC    E    7   6   4     1.140   107.182    86.273   0.01269
  21  H4    HC    E    7   6   4     1.140   114.401   -36.027   0.01269
  22  C2    C     M    6   4   3     1.540   108.193   -11.038   0.21708
  23  O1    O     E   22   6   4     1.510   119.737   -13.840  -0.13636
LOOP
   C4   C1
```



```
IMPROPER
   -M   C1   N1   H2
   C1   +M   C2   O1
DONE
STOP
```

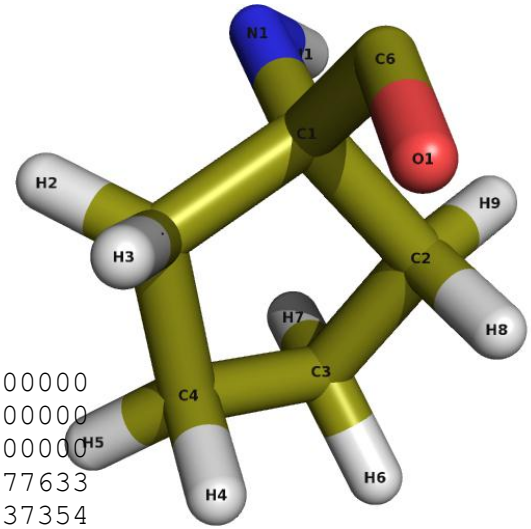
.frcmod file:

```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
HC 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-C   328.30   1.508       same as c -c3
CT-HC  337.30   1.092       same as c3-hc
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
CT-CT-C    63.800     110.530   same as c -c3-c3
HC-CT-HC   39.400     108.350   same as hc-c3-hc
```

```
DIHE
N -CT-CT-CT    1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O     1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT    1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C     1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CT-CT    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O     1    0.000       180.000          2.000      same as X -c -c3-X
CT-CT-CT-C     1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-C     1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-HC    1    0.156         0.000          3.000      same as X -c3-c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  HC          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

ACP

.prepi file:




```
    0    0    2
This is a remark line
acp.res
ACP   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522    111.1       .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208  -180.000  -0.77633
   5  H1    H     E    4   3   2     0.999    72.335    -8.888   0.37354
   6  C1    CT    M    4   3   2     1.449    42.298  -178.414   0.18994
   7  C2    CT    3    6   4   3     1.564   114.519  -121.832  -0.06626
   8  C3    CT    3    7   6   4     1.545   107.322  -129.327  -0.06953
   9  C4    CT    3    8   7   6     1.538   105.942   -16.899  -0.06953
  10  C5    CT    B    9   8   7     1.531   104.413    31.246  -0.06626
  11  H2    HC    E   10   9   8     1.082   110.067    83.446   0.04057
  12  H3    HC    E   10   9   8     1.083   113.455  -157.204   0.04057
  13  H4    HC    E    9   8   7     1.085   111.012   -88.035   0.02957
  14  H5    HC    E    9   8   7     1.084   112.273   152.937   0.02957
  15  H6    HC    E    8   7   6     1.085   112.336  -140.135   0.02957
  16  H7    HC    E    8   7   6     1.084   109.965   101.453   0.02957
  17  H8    HC    E    7   6   4     1.084   110.829   108.710   0.04057
  18  H9    HC    E    7   6   4     1.080   110.066    -7.497   0.04057
  19  C6    C     M    6   4   3     1.549   104.025     0.602   0.76234
  20  O1    O     E   19   6   4     1.202   121.519    14.503  -0.55848
LOOP
   C5   C1
IMPROPER
   -M   C1   N1   H1
   C1   +M   C6   O1
DONE
STOP
```

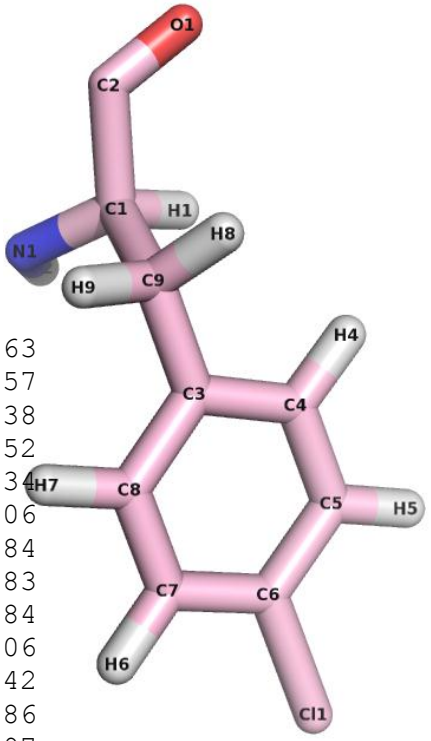
.frcmod file:

```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
HC 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-C   328.30   1.508       same as c -c3
CT-HC  337.30   1.092       same as c3-hc
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-C    66.700     111.560   same as c -c3-n
H -N -CT   46.000     116.780   same as c3-n -hn
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
CT-C -O    68.000     123.110   same as c3-c -o
CT-CT-C    63.800     110.530   same as c -c3-c3
HC-CT-HC   39.400     108.350   same as hc-c3-hc

DIHE
N -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
```

```
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X

IMPROPER

NONBON
  N           1.8240  0.1700            same as n
  H           0.6000  0.0157            same as hn
  CT          1.9080  0.1094            same as c3
  HC          1.4870  0.0157            same as hc
  C           1.9080  0.0860            same as c
  O           1.6612  0.2100            same as o
```

P-CL PHENYLALANINE

.mol2 file:




```
@<TRIPOS>MOLECULE
PCL
20   20    1
SMALL
USER_CHARGES

@<TRIPOS>ATOM
      1 C1          -1.4320   -0.5420    0.5430 Du        1 PCL          0.1063
      2 C2          -2.8510   -1.1830    0.3800 C         1 PCL          0.6257
      3 N1          -1.4270    0.8840    0.9730 N.3       1 PCL         -0.7238
      4 O1          -3.1080   -2.2450    0.9740 O.2       1 PCL         -0.6052
      5 C3           0.9530   -0.6060   -0.4830 Du        1 PCL          0.4134
      6 C4           1.7220   -1.5330    0.2520 Du        1 PCL         -0.3506
      7 C5           3.0980   -1.3470    0.4560 Du        1 PCL         -0.0284
      8 C6           3.7220   -0.2140   -0.0870 Du        1 PCL         -0.0583
      9 C7           2.9870    0.7270   -0.8200 Du        1 PCL         -0.0284
     10 C8           1.6100    0.5240   -1.0110 Du        1 PCL         -0.3506
     11 Cl1          5.4590    0.0230    0.1480 Cl        1 PCL         -0.1142
     12 C9          -0.5370   -0.8350   -0.7070 Du        1 PCL         -0.5686
     13 H1          -1.0100   -1.0870    1.3940 Du        1 PCL          0.0997
     14 H2          -1.0930    1.0500    1.9160 H         1 PCL          0.3997
     15 H4           1.2480   -2.4260    0.6640 Du        1 PCL          0.1989
     16 H5           3.6800   -2.0760    1.0180 Du        1 PCL          0.1387
     17 H6           3.4820    1.6000   -1.2430 Du        1 PCL          0.1387
     18 H7           1.0440    1.2520   -1.5950 Du        1 PCL          0.1989
     19 H8          -0.7070   -1.8900   -0.9660 Du        1 PCL          0.1842
     20 H9          -0.8900   -0.2340   -1.5530 Du        1 PCL          0.1842
@<TRIPOS>BOND
     1    1    2 1
     2    1    3 1
     3    1   12 1
     4    1   13 1
     5    2    4 2
```

```
     6    3   14 1
     7    5    6 1
     8    5   10 1
     9    5   12 1
    10    6    7 1
    11    6   15 1
    12    7    8 1
    13    7   16 1
    14    8    9 1
    15    8   11 1
    16    9   10 1
    17    9   17 1
    18   10   18 1
    19   12   19 1
    20   12   20 1
@<TRIPOS>SUBSTRUCTURE
     1 PCL         1 GROUP             0       ****    0 ROOT
```


.frcmod file:

```
remark goes here
MASS
CT 12.010        0.878              same as c3
C  12.010        0.616              same as c
N  14.010        0.530              same as n
O  16.000        0.434              same as o
CA 12.010        0.360              same as c2
Cl 35.450        1.910              same as cl
H1 1.008         0.135              same as hc
H  1.008         0.161              same as hn
HA 1.008         0.135              same as hc
HC 1.008         0.135              same as hc

BOND
CT-C   328.30   1.508       same as c -c3
CT-N   330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
```

```
C -O   648.00   1.214       same as c -o
C -N   478.20   1.345       same as c -n
N -H   410.20   1.009       same as hn-n
CA-CA  478.40   1.387       same as ca-ca
CA-CT  328.30   1.508       same as c2-c3
CA-HA  344.30   1.087       same as c2-hc
CA-Cl  328.80   1.722       same as c2-cl
CT-HC  337.30   1.092       same as c3-hc

ANGLE
CT-C -O    68.000     123.110   same as c3-c -o
CT-C -N    67.900     115.150   same as c3-c -n
CT-N -H    46.000     116.780   same as c3-n -hn
CT-N -C    63.900     121.350   same as c -n -c3
CT-CT-CA   63.700     110.960   same as c2-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
C -CT-N    66.700     111.560   same as c -c3-n
C -CT-CT   63.800     110.530   same as c -c3-c3
C -CT-H1   47.200     109.680   same as c -c3-hc
C -N -H    49.200     118.460   same as c -n -hn
N -CT-CT   65.900     112.130   same as c3-c3-n
N -CT-H1   49.800     109.500   same as hc-c3-n
N -C -O    75.800     122.030   same as n -c -o
CA-CA-CA   67.200     119.970   same as ca-ca-ca
CA-CA-HA   50.300     119.700   same as c2-c2-hc
CA-CT-HC   47.000     110.490   same as c2-c3-hc
CA-CA-CT   64.300     123.420   same as c2-c2-c3
CA-CA-Cl   58.400     122.790   same as c2-c2-cl
CT-CT-H1   46.400     110.050   same as c3-c3-hc
HC-CT-HC   39.400     108.350   same as hc-c3-hc
C -CT-HC   47.200     109.680   same as c -c3-hc
H1-CT-H1   39.400     108.350   same as hc-c3-hc

DIHE
CT-C -N -H    1    2.500       180.000          2.000      same as X -c -n -X
CT-C -N -CT   1    2.500       180.000          2.000      same as X -c -n -X
CT-N -C -O    1    2.500       180.000          2.000      same as X -c -n -X
```

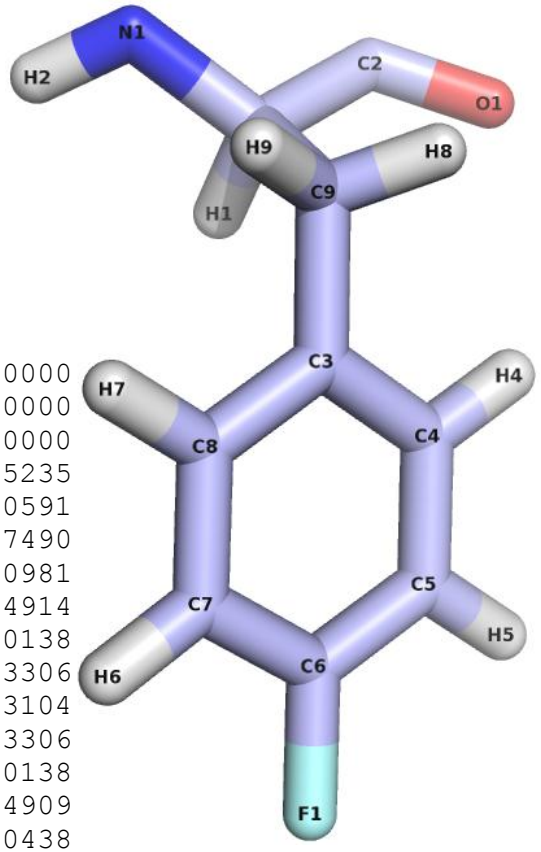
```
CT-CT-CA-CA   1    0.000          0.000          2.000      same as X -c2-c3-X
C -CT-N -H    1    0.000          0.000          2.000      same as X -c3-n -X
C -CT-N -C    1    0.000          0.000          2.000      same as X -c3-n -X
C -CT-CT-CA   1    0.156          0.000          3.000      same as X -c3-c3-X
C -CT-CT-HC   1    0.156          0.000          3.000      same as X -c3-c3-X
C -N -CT-H1   1    0.000          0.000          2.000      same as X -c3-n -X
N -CT-C -O    1    0.000        180.000          2.000      same as X -c -c3-X
N -CT-C -N    1    0.000        180.000          2.000      same as X -c -c3-X
N -CT-CT-CA   1    0.156          0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC   1    0.156          0.000          3.000      same as X -c3-c3-X
N -C -CT-HC   1    0.000        180.000          2.000      same as X -c -c3-X
O -C -CT-CT   1    0.000        180.000          2.000      same as X -c -c3-X
O -C -CT-H1   1    0.000        180.000          2.000      same as X -c -c3-X
O -C -N -H    1    2.500        180.000          2.000      same as X -c -n -X
N -C -CT-CT   1    0.000        180.000          2.000      same as X -c -c3-X
N -C -CT-H1   1    0.000        180.000          2.000      same as X -c -c3-X
CA-CA-CA-CA   1    3.625        180.000          2.000      same as X -ca-ca-X
CA-CA-CA-HA   1    6.650        180.000          2.000      same as X -c2-c2-X
CA-CT-CT-H1   1    0.156          0.000          3.000      same as X -c3-c3-X
CA-CA-CT-HC   1    0.000          0.000          2.000      same as X -c2-c3-X
CA-CA-CA-Cl   1    6.650        180.000          2.000      same as X -c2-c2-X
CA-CA-CA-CT   1    6.650        180.000          2.000      same as X -c2-c2-X
Cl-CA-CA-HA   1    6.650        180.000          2.000      same as X -c2-c2-X
CT-CT-N -H    1    0.000          0.000          2.000      same as X -c3-n -X
CT-CT-N -C    1    0.000          0.000          2.000      same as X -c3-n -X
CT-CA-CA-HA   1    6.650        180.000          2.000      same as X -c2-c2-X
H1-CT-N -H    1    0.000          0.000          2.000      same as X -c3-n -X
H1-CT-CT-HC   1    0.156          0.000          3.000      same as X -c3-c3-X
HA-CA-CA-HA   1    6.650        180.000          2.000      same as X -c2-c2-X
O -C -CT-HC   1    0.000        180.000          2.000      same as X -c -c3-X

IMPROPER
CT-N -C -O          1.1          180.0          2.0        Using default value
C -CT-N -H          1.1          180.0          2.0        Using default value
CA-CA-CA-CT         1.1          180.0          2.0        Using default value
CA-CA-CA-HA         1.1          180.0          2.0        Using default value
CA-CA-CA-Cl         1.1          180.0          2.0        Using default value
```

```
NONBON
  CT          1.9080  0.1094          same as c3
  C           1.9080  0.0860          same as c
  N           1.8240  0.1700          same as n
  O           1.6612  0.2100          same as o
  CA          1.9080  0.0860          same as ca
  Cl          1.9480  0.2650          same as cl
  H1          1.4870  0.0157          same as hc
  H           0.6000  0.0157          same as hn
  HA          1.4870  0.0157          same as hc
  HC          1.4870  0.0157          same as hc
```

P-F PHENYLALANINE

.prepi file:

```
    0    0    2

This is a remark line
pFPhe_4.res
FPH   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  N1    N     M    3   2   1     1.540   111.208   180.000  -0.85235
   5  H2    H     E    4   3   2     1.016   105.060    58.886   0.40591
   6  C1    CT    M    4   3   2     1.485   127.402  -159.092   0.17490
   7  C9    CT    3    6   4   3     1.575   114.015   120.885  -0.00981
   8  C3    CA    S    7   6   4     1.526   113.417   119.886  -0.04914
   9  C4    CA    B    8   7   6     1.412   120.830    81.583  -0.10138
  10  C5    CA    B    9   8   7     1.404   121.268  -178.417  -0.33306
  11  C6    CA    B   10   9   8     1.396   118.558     0.064   0.43104
  12  C7    CA    B   11  10   9     1.396   122.216    -0.410  -0.33306
  13  C8    CA    S   12  11  10     1.406   118.282     0.243  -0.10138
  14  H7    HA    E   13  12  11     1.092   118.784  -179.360   0.14909
  15  H6    HA    E   12  11  10     1.089   119.992  -179.202   0.20438
  16  F1    F     E   11  10   9     1.382   118.950   179.891  -0.24201
  17  H5    HA    E   10   9   8     1.090   121.529  -179.685   0.20438
  18  H4    HA    E    9   8   7     1.090   119.241     1.249   0.14909
  19  H8    HC    E    7   6   4     1.096   108.169  -117.992   0.06182
  20  H9    HC    E    7   6   4     1.096   108.303    -2.005   0.06182
  21  H1    H1    E    6   4   3     1.099   104.360  -121.546   0.05788
  22  C2    C     M    6   4   3     1.566   113.591    -9.092   0.65989
  23  O1    O     E   22   6   4     1.240   120.893   157.412  -0.53804
LOOP
   C8   C3
```



```
IMPROPER
   -M   C1   N1   H2
   C8   C4   C3   C9
   C3   C5   C4   H4
   C6   C4   C5   H5
   C7   C5   C6   F1
   C6   C8   C7   H6
   C7   C3   C8   H7
   C1   +M   C2   O1

DONE
STOP
```

.frcmod file:

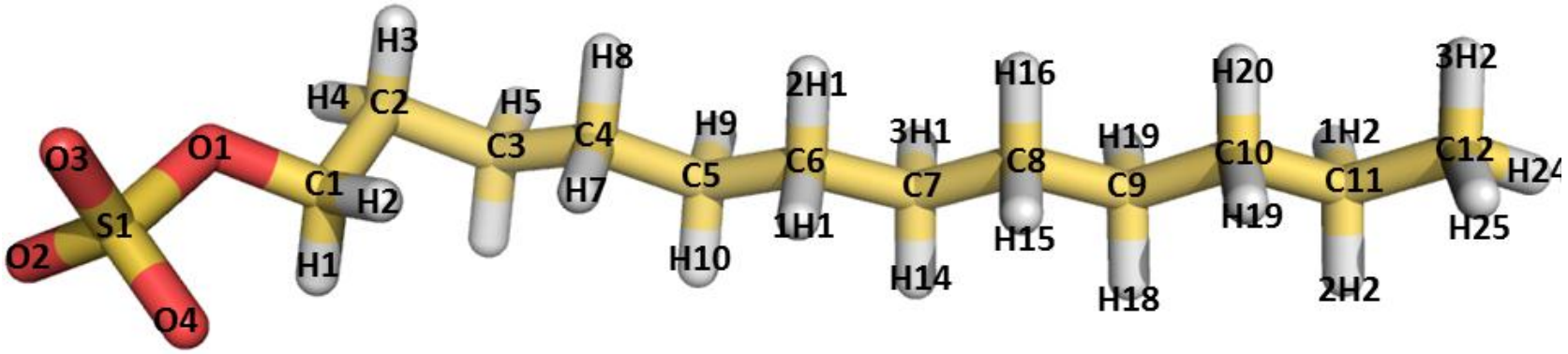
```
remark goes here
MASS
N  14.010        0.530              same as n
H  1.008         0.161              same as hn
CT 12.010        0.878              same as c3
CA 12.010        0.360              same as c2
HA 1.008         0.135              same as hc
F  19.000        0.320              same as f
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
C  12.010        0.616              same as c
O  16.000        0.434              same as o

BOND
N -H   410.20   1.009       same as hn-n
N -CT  330.60   1.460       same as c3-n
CT-CT  303.10   1.535       same as c3-c3
CT-H1  337.30   1.092       same as c3-hc
CT-C   328.30   1.508       same as c -c3
CT-CA  328.30   1.508       same as c2-c3
CT-HC  337.30   1.092       same as c3-hc
CA-CA  478.40   1.387       same as ca-ca
```

```
CA-HA  344.30   1.087       same as c2-hc
CA-F   368.70   1.340       same as c2-f
C -O   648.00   1.214       same as c -o

ANGLE
N -CT-CT    65.900      112.130   same as c3-c3-n
N -CT-H1    49.800      109.500   same as hc-c3-n
N -CT-C     66.700      111.560   same as c -c3-n
H -N -CT    46.000      116.780   same as c3-n -hn
CT-CT-CA    63.700      110.960   same as c2-c3-c3
CT-CT-HC    46.400      110.050   same as c3-c3-hc
CT-C -O     68.000      123.110   same as c3-c -o
CT-CT-H1    46.400      110.050   same as c3-c3-hc
CT-CT-C     63.800      110.530   same as c -c3-c3
CT-CA-CA    64.300      123.420   same as c2-c2-c3
CA-CT-HC    47.000      110.490   same as c2-c3-hc
CA-CA-CA    67.200      119.970   same as ca-ca-ca
CA-CA-HA    50.300      119.700   same as c2-c2-hc
CA-CA-F     67.800      124.010   same as c2-c2-f
HC-CT-HC    39.400      108.350   same as hc-c3-hc
H1-CT-C     47.200      109.680   same as c -c3-hc

DIHE
N -CT-CT-CA   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
N -CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
H -N -CT-CT   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-H1   1    0.000         0.000          2.000      same as X -c3-n -X
H -N -CT-C    1    0.000         0.000          2.000      same as X -c3-n -X
CT-CT-CA-CA   1    0.000         0.000          2.000      same as X -c2-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-CA-CA-CA   1    6.650       180.000          2.000      same as X -c2-c2-X
CT-CA-CA-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
CA-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CA-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
CA-CA-CA-CA   1    3.625       180.000          2.000      same as X -ca-ca-X
CA-CA-CA-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
```

```
CA-CA-CT-HC   1    0.000         0.000          2.000      same as X -c2-c3-X
CA-CA-CA-F    1    6.650       180.000          2.000      same as X -c2-c2-X
HA-CA-CA-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
HA-CA-CA-F    1    6.650       180.000          2.000      same as X -c2-c2-X
HC-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HC-CT-CT-C    1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X

IMPROPER
CA-CA-CA-CT          1.1         180.0          2.0        Using default value
CA-CA-CA-HA          1.1         180.0          2.0        Using default value
CA-CA-CA-F           1.1         180.0          2.0        Using default value

NONBON
  N           1.8240  0.1700          same as n
  H           0.6000  0.0157          same as hn
  CT          1.9080  0.1094          same as c3
  CA          1.9080  0.0860          same as ca
  HA          1.4870  0.0157          same as hc
  F           1.7500  0.0610          same as f
  HC          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  C           1.9080  0.0860          same as c
  O           1.6612  0.2100          same as o
```

LIPIDS

SDS

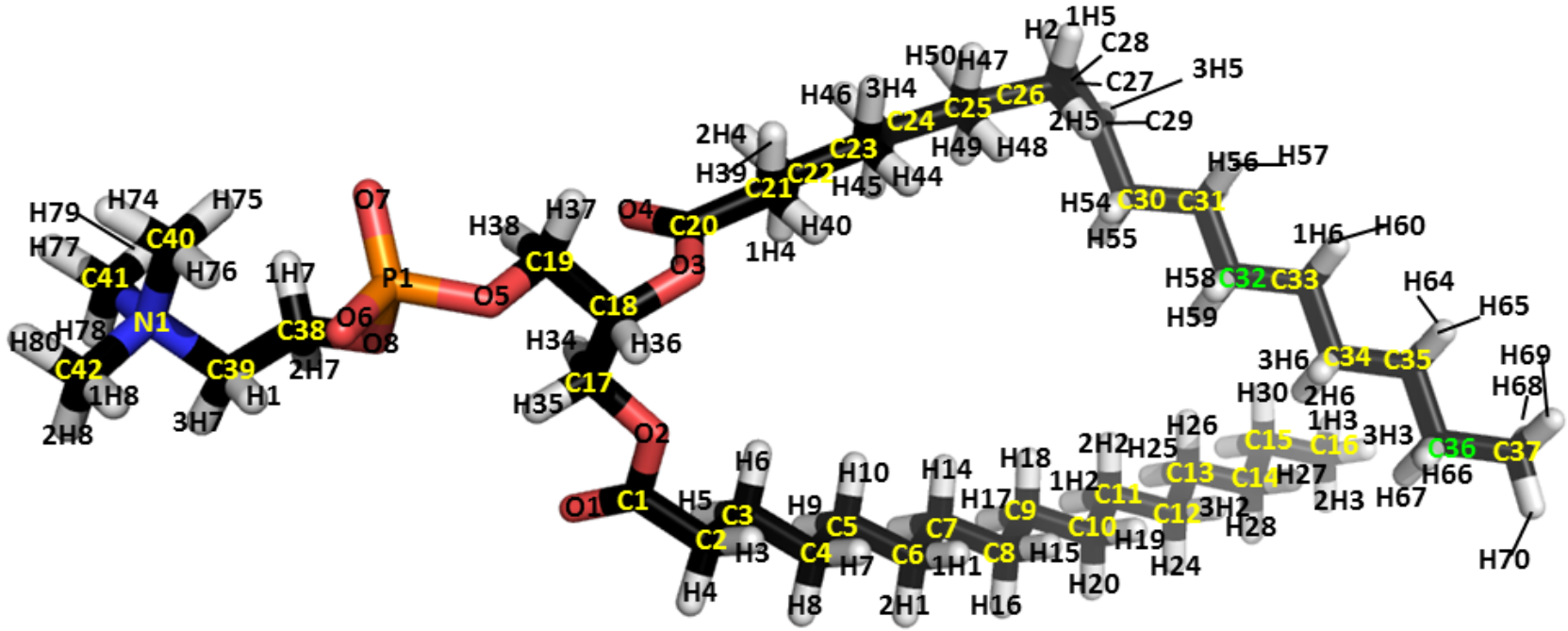


.prepi file:

```
    0    0    2
This is a remark line
sds.res
SDS   INT  0
CORRECT     OMIT DU   BEG
  0.0000
   1  DUMM  DU    M    0  -1  -2     0.000      .0        .0      .00000
   2  DUMM  DU    M    1   0  -1     1.449      .0        .0      .00000
   3  DUMM  DU    M    2   1   0     1.522   111.1        .0      .00000
   4  O2    o     M    3   2   1     1.540   111.208  -180.000  -0.67456
   5  S1    s6    M    4   3   2     1.438    19.047   132.668   1.29138
   6  O3    o     E    5   4   3     1.447   115.333   102.604  -0.67456
   7  O4    o     E    5   4   3     1.446   115.321  -122.068  -0.67456
   8  O1    os    M    5   4   3     1.633   101.805    -9.671  -0.48958
   9  C1    c3    M    8   5   4     1.406   116.952   176.279   0.21893
  10  H1    h1    E    9   8   5     1.088   106.699   136.387   0.01297
  11  H2    h1    E    9   8   5     1.081   110.255    19.894   0.01297
  12  C2    c3    M    9   8   5     1.524   110.389  -104.017  -0.07803
  13  H3    hc    E   12   9   8     1.083   108.194    61.303   0.04669
```

```
  14  H4    hc    E   12   9   8     1.087   107.771   -54.024   0.04669
  15  C3    c3    M   12   9   8     1.534   114.031  -175.268  -0.04926
  16  H5    hc    E   15  12   9     1.089   108.534   170.556  -0.00858
  17  H6    hc    E   15  12   9     1.091   109.647    55.328  -0.00858
  18  C4    c3    M   15  12   9     1.532   114.631   -67.980   0.02383
  19  H7    hc    E   18  15  12     1.087   109.980    62.375   0.02701
  20  H8    hc    E   18  15  12     1.087   109.091   -53.556   0.02701
  21  C5    c3    M   18  15  12     1.530   112.973  -175.607  -0.04733
  22  H9    hc    E   21  18  15     1.090   109.344    58.805   0.01570
  23  H10   hc    E   21  18  15     1.088   109.010   -56.948   0.01570
  24  C6    c3    M   21  18  15     1.532   112.921  -179.544  -0.06946
  25  H11   hc    E   24  21  18     1.089   109.288    61.243   0.01568
  26  H12   hc    E   24  21  18     1.088   108.585   -53.969   0.01568
  27  C7    c3    M   24  21  18     1.533   114.996  -175.511   0.02242
  28  H13   hc    E   27  24  21     1.089   108.580   172.050  -0.00530
  29  H14   hc    E   27  24  21     1.089   109.234    56.778  -0.00530
  30  C8    c3    M   27  24  21     1.531   114.871   -66.275   0.05251
  31  H15   hc    E   30  27  24     1.087   110.009    61.960  -0.00785
  32  H16   hc    E   30  27  24     1.090   109.149   -54.263  -0.00785
  33  C9    c3    M   30  27  24     1.530   112.992  -176.037  -0.06972
  34  H17   hc    E   33  30  27     1.089   109.301    58.076   0.00147
  35  H18   hc    E   33  30  27     1.089   109.279   -57.738   0.00147
  36  C10   c3    M   33  30  27     1.530   113.347  -179.826   0.03431
  37  H19   hc    E   36  33  30     1.089   109.290    58.034  -0.00750
  38  H20   hc    E   36  33  30     1.089   109.326   -57.806  -0.00750
  39  C11   c3    M   36  33  30     1.530   113.373  -179.881   0.14924
  40  H21   hc    E   39  36  33     1.088   109.309    58.010  -0.02747
  41  H22   hc    E   39  36  33     1.088   109.351   -57.957  -0.02747
  42  C12   c3    M   39  36  33     1.528   113.057  -179.996  -0.24073
  43  H23   hc    E   42  39  36     1.086   111.185   -59.899   0.04985
  44  H24   hc    E   42  39  36     1.087   111.385   179.975   0.04985
  45  H25   hc    E   42  39  36     1.087   111.144    59.958   0.04985
LOOP
IMPROPER
DONE
STOP
```

POPC



.mol2 file:

```
@<TRIPOS>MOLECULE
POC
  134   133     1     0     0
SMALL
resp
@<TRIPOS>ATOM
      1 C1          -4.4310   -2.8040    0.7390 C          1 POC       0.8311
      2 C2          -3.1620   -3.4380    1.3030 CT         1 POC      -0.3652
      3 C3          -1.9420   -3.1170    0.3910 CT         1 POC       0.1019
      4 C4          -0.6050   -3.7060    0.9010 CT         1 POC      -0.0136
      5 C5           0.5790   -3.3450   -0.0280 CT         1 POC      -0.0809
      6 C6           1.9660   -3.8310    0.4560 CT         1 POC       0.0519
```

7	C7	3.1040	-3.4010	-0.5000	CT	1 POC	0.0239
8	C8	4.5320	-3.7740	-0.0360	CT	1 POC	-0.0183
9	C9	5.6200	-3.2680	-1.0150	CT	1 POC	0.0599
10	C10	7.0780	-3.5320	-0.5690	CT	1 POC	-0.0053
11	C11	8.1180	-2.9390	-1.5500	CT	1 POC	-0.0029
12	C12	9.5920	-3.1180	-1.1190	CT	1 POC	0.0694
13	C13	10.5930	-2.4330	-2.0810	CT	1 POC	-0.0269
14	C14	12.0740	-2.5510	-1.6520	CT	1 POC	-0.0261
15	C15	13.0520	-1.8180	-2.6020	CT	1 POC	0.1280
16	C16	14.5320	-1.9180	-2.1680	CT	1 POC	-0.2429
17	O1	-5.1580	-3.3190	-0.1040	O	1 POC	-0.5977
18	O2	-4.6240	-1.5470	1.2600	OS	1 POC	-0.4534
19	C17	-5.7430	-0.7750	0.6930	CT	1 POC	-0.0010
20	C18	-5.9400	0.4770	1.5490	CT	1 POC	0.3483
21	C19	-7.2010	1.2800	1.1530	CT	1 POC	0.1009
22	O3	-4.7330	1.3380	1.6200	OS	1 POC	-0.6032
23	C20	-4.2420	1.9610	0.4990	C	1 POC	0.8540
24	C21	-3.0020	2.7810	0.8650	CT	1 POC	-0.2087
25	O4	-4.7340	1.8860	-0.6220	O	1 POC	-0.6060
26	C22	-1.9910	2.9400	-0.2940	CT	1 POC	0.0724
27	C23	-0.7370	3.7420	0.1280	CT	1 POC	0.0171
28	C24	0.3370	3.8610	-0.9800	CT	1 POC	0.0221
29	C25	1.6290	4.5640	-0.5010	CT	1 POC	-0.0242
30	C26	2.7470	4.6390	-1.5680	CT	1 POC	0.0170
31	C27	4.0760	5.2260	-1.0090	CT	1 POC	0.1055
32	C28	5.1910	5.3120	-2.0330	CM	1 POC	-0.2879
33	C29	6.2650	4.5020	-2.1530	CM	1 POC	-0.3128
34	C30	6.6150	3.2940	-1.3080	CT	1 POC	0.1620
35	C31	7.9030	3.4660	-0.4530	CT	1 POC	-0.0279
36	C32	8.2340	2.2100	0.3860	CT	1 POC	-0.0455
37	C33	9.4930	2.3470	1.2750	CT	1 POC	0.0548
38	C34	9.7810	1.0840	2.1220	CT	1 POC	-0.0161
39	C35	11.0360	1.1930	3.0200	CT	1 POC	-0.0253
40	C36	11.3090	-0.0780	3.8590	CT	1 POC	0.1697
41	C37	12.5650	0.0240	4.7520	CT	1 POC	-0.2488
42	O5	-8.3400	0.5530	1.6930	OS	1 POC	-0.5218
43	P1	-9.7590	0.5410	0.8620	P	1 POC	1.2943

44	O6	-10.8030	-0.1130	1.7400	O2	1 POC	-0.7816
45	O7	-10.0250	1.8320	0.1210	O2	1 POC	-0.7816
46	O8	-9.2960	-0.6210	-0.3210	OS	1 POC	-0.4951
47	C38	-10.1980	-0.7820	-1.4100	CT	1 POC	0.3475
48	C39	-11.5030	-1.5170	-0.9630	CT	1 POC	-0.0757
49	N1	-12.8580	-0.7730	-1.2030	N3	1 POC	0.1545
50	C40	-12.9240	0.5730	-0.4680	CT	1 POC	-0.4259
51	H1	-11.4440	-1.6720	0.1190	HP	1 POC	0.0991
52	H2	5.0970	6.1300	-2.7560	HA	1 POC	0.1490
53	H3	-2.9780	-3.0650	2.3200	HC	1 POC	0.0949
54	H4	-3.3230	-4.5230	1.3450	HC	1 POC	0.0949
55	H5	-2.1500	-3.4980	-0.6210	HC	1 POC	0.0149
56	H6	-1.8440	-2.0240	0.3040	HC	1 POC	0.0149
57	H7	-0.4030	-3.3280	1.9170	HC	1 POC	0.0203
58	H8	-0.6910	-4.8020	0.9860	HC	1 POC	0.0203
59	H9	0.3840	-3.7530	-1.0340	HC	1 POC	0.0142
60	H10	0.6130	-2.2480	-0.1460	HC	1 POC	0.0142
61	H11	2.1620	-3.4270	1.4640	HC	1 POC	-0.0134
62	H12	1.9600	-4.9300	0.5580	HC	1 POC	-0.0134
63	H13	2.9230	-3.8380	-1.4960	HC	1 POC	-0.0035
64	H14	3.0540	-2.3070	-0.6360	HC	1 POC	-0.0035
65	H15	4.7140	-3.3440	0.9640	HC	1 POC	-0.0010
66	H16	4.6140	-4.8680	0.0790	HC	1 POC	-0.0010
67	H17	5.4570	-3.7260	-2.0050	HC	1 POC	-0.0213
68	H18	5.4870	-2.1820	-1.1570	HC	1 POC	-0.0213
69	H19	7.2360	-3.0960	0.4320	HC	1 POC	-0.0072
70	H20	7.2440	-4.6180	-0.4620	HC	1 POC	-0.0072
71	H21	7.9750	-3.3900	-2.5470	HC	1 POC	-0.0039
72	H22	7.9120	-1.8620	-1.6720	HC	1 POC	-0.0039
73	H23	9.7260	-2.7040	-0.1050	HC	1 POC	-0.0197
74	H24	9.8280	-4.1930	-1.0450	HC	1 POC	-0.0197
75	H25	10.4730	-2.8570	-3.0930	HC	1 POC	-0.0007
76	H26	10.3290	-1.3640	-2.1660	HC	1 POC	-0.0007
77	H27	12.1910	-2.1450	-0.6330	HC	1 POC	0.0098
78	H28	12.3550	-3.6160	-1.5910	HC	1 POC	0.0098
79	H29	12.9440	-2.2260	-3.6210	HC	1 POC	-0.0165
80	H30	12.7640	-0.7550	-2.6660	HC	1 POC	-0.0165

81	H31	15.1890	-1.3820	-2.8700	HC	1 POC	0.0552
82	H32	14.8670	-2.9660	-2.1280	HC	1 POC	0.0552
83	H33	14.6850	-1.4820	-1.1680	HC	1 POC	0.0552
84	H34	-5.5130	-0.5320	-0.3480	H1	1 POC	0.0865
85	H35	-6.6590	-1.3750	0.7260	H1	1 POC	0.0865
86	H36	-6.0540	0.1730	2.5960	H1	1 POC	0.0907
87	H37	-7.1570	2.2730	1.6210	H1	1 POC	0.0740
88	H38	-7.2860	1.4040	0.0700	H1	1 POC	0.0740
89	H39	-3.3640	3.7710	1.1880	HC	1 POC	0.0646
90	H40	-2.5270	2.3270	1.7450	HC	1 POC	0.0646
91	H41	-1.6840	1.9400	-0.6410	HC	1 POC	-0.0073
92	H42	-2.4870	3.4240	-1.1470	HC	1 POC	-0.0073
93	H43	-1.0390	4.7520	0.4540	HC	1 POC	0.0001
94	H44	-0.2840	3.2590	1.0110	HC	1 POC	0.0001
95	H45	0.5930	2.8510	-1.3440	HC	1 POC	-0.0068
96	H46	-0.0830	4.4030	-1.8450	HC	1 POC	-0.0068
97	H47	1.3870	5.5830	-0.1540	HC	1 POC	-0.0063
98	H48	2.0160	4.0260	0.3820	HC	1 POC	-0.0063
99	H49	2.9490	3.6310	-1.9650	HC	1 POC	0.0044
100	H50	2.4060	5.2460	-2.4230	HC	1 POC	0.0044
101	H51	3.8760	6.2390	-0.6190	HC	1 POC	0.0171
102	H52	4.3900	4.6230	-0.1450	HC	1 POC	0.0171
103	H53	6.9770	4.7260	-2.9560	HA	1 POC	0.1468
104	H54	5.7800	3.0310	-0.6430	HC	1 POC	0.0077
105	H55	6.7560	2.4280	-1.9790	HC	1 POC	0.0077
106	H56	8.7530	3.7050	-1.1140	HC	1 POC	0.0148
107	H57	7.7750	4.3350	0.2120	HC	1 POC	0.0148
108	H58	7.3680	1.9700	1.0280	HC	1 POC	0.0067
109	H59	8.3600	1.3480	-0.2900	HC	1 POC	0.0067
110	H60	10.3680	2.5680	0.6410	HC	1 POC	-0.0122
111	H61	9.3710	3.2150	1.9460	HC	1 POC	-0.0122
112	H62	9.8930	0.2170	1.4490	HC	1 POC	0.0021
113	H63	8.9040	0.8670	2.7560	HC	1 POC	0.0021
114	H64	11.9170	1.4070	2.3900	HC	1 POC	0.0029
115	H65	10.9250	2.0570	3.6970	HC	1 POC	0.0029
116	H66	10.4300	-0.2890	4.4910	HC	1 POC	-0.0308
117	H67	11.4140	-0.9420	3.1820	HC	1 POC	-0.0308

```
     118 H68         12.7210   -0.9010    5.3270 HC        1 POC       0.0545
     119 H69         12.4780    0.8540    5.4710 HC        1 POC       0.0545
     120 H70         13.4700    0.2000    4.1500 HC        1 POC       0.0545
     121 H71        -10.4140    0.1900   -1.8820 H1        1 POC      -0.0146
     122 H72         -9.6820   -1.4090   -2.1510 H1        1 POC      -0.0146
     123 H73        -11.6250   -2.4740   -1.4840 HP        1 POC       0.0991
     124 C41        -13.0800   -0.5540   -2.6900 CT        1 POC      -0.4259
     125 C42        -13.9640   -1.6660   -0.6580 CT        1 POC      -0.4259
     126 H74        -13.9170    0.9980   -0.6480 HP        1 POC       0.1928
     127 H75        -12.1330    1.2330   -0.8330 HP        1 POC       0.1928
     128 H76        -12.7300    0.3920    0.5930 HP        1 POC       0.1928
     129 H77        -14.0590   -0.0890   -2.8370 HP        1 POC       0.1928
     130 H78        -13.0410   -1.5200   -3.2040 HP        1 POC       0.1928
     131 H79        -12.2980    0.1060   -3.0720 HP        1 POC       0.1928
     132 H80        -14.9280   -1.1740   -0.8200 HP        1 POC       0.1928
     133 H81        -13.7910   -1.8110    0.4110 HP        1 POC       0.1928
     134 H82        -13.9400   -2.6280   -1.1790 HP        1 POC       0.1928
@<TRIPOS>BOND
     1    1    2 1          20    7   63 1          39   13   76 1          58   21   88 1
     2    1   17 2          21    7   64 1          40   14   15 1          59   22   23 1
     3    1   18 1          22    8    9 1          41   14   77 1          60   23   24 1
     4    2    3 1          23    8   65 1          42   14   78 1          61   23   25 2
     5    2   53 1          24    8   66 1          43   15   16 1          62   24   26 1
     6    2   54 1          25    9   10 1          44   15   79 1          63   24   89 1
     7    3    4 1          26    9   67 1          45   15   80 1          64   24   90 1
     8    3   55 1          27    9   68 1          46   16   81 1          65   26   27 1
     9    3   56 1          28   10   11 1          47   16   82 1          66   26   91 1
    10    4    5 1          29   10   69 1          48   16   83 1          67   26   92 1
    11    4   57 1          30   10   70 1          49   18   19 1          68   27   28 1
    12    4   58 1          31   11   12 1          50   19   20 1          69   27   93 1
    13    5    6 1          32   11   71 1          51   19   84 1          70   27   94 1
    14    5   59 1          33   11   72 1          52   19   85 1          71   28   29 1
    15    5   60 1          34   12   13 1          53   20   21 1          72   28   95 1
    16    6    7 1          35   12   73 1          54   20   22 1          73   28   96 1
    17    6   61 1          36   12   74 1          55   20   86 1          74   29   30 1
    18    6   62 1          37   13   14 1          56   21   42 1          75   29   97 1
    19    7    8 1          38   13   75 1          57   21   87 1          76   29   98 1
```

```
    77   30   31 1          92   35  107 1          107   40  117 1          122   49   50 1
    78   30   99 1          93   36   37 1          108   41  118 1          123   49  124 1
    79   30  100 1          94   36  108 1          109   41  119 1          124   49  125 1
    80   31   32 1          95   36  109 1          110   41  120 1          125   50  126 1
    81   31  101 1          96   37   38 1          111   42   43 1          126   50  127 1
    82   31  102 1          97   37  110 1          112   43   44 1          127   50  128 1
    83   32   33 2          98   37  111 1          113   43   45 1          128  124  129 1
    84   32   52 1          99   38   39 1          114   43   46 1          129  124  130 1
    85   33   34 1         100   38  112 1          115   46   47 1          130  124  131 1
    86   33  103 1         101   38  113 1          116   47   48 1          131  125  132 1
    87   34   35 1         102   39   40 1          117   47  121 1          132  125  133 1
    88   34  104 1         103   39  114 1          118   47  122 1          133  125  134 1
    89   34  105 1         104   39  115 1          119   48   49 1
    90   35   36 1         105   40   41 1          120   48   51 1
    91   35  106 1         106   40  116 1          121   48  123 1
@<TRIPOS>SUBSTRUCTURE
     1 POC         1 TEMP              0 ****  ****    0 ROOT
```


.frcmod file:

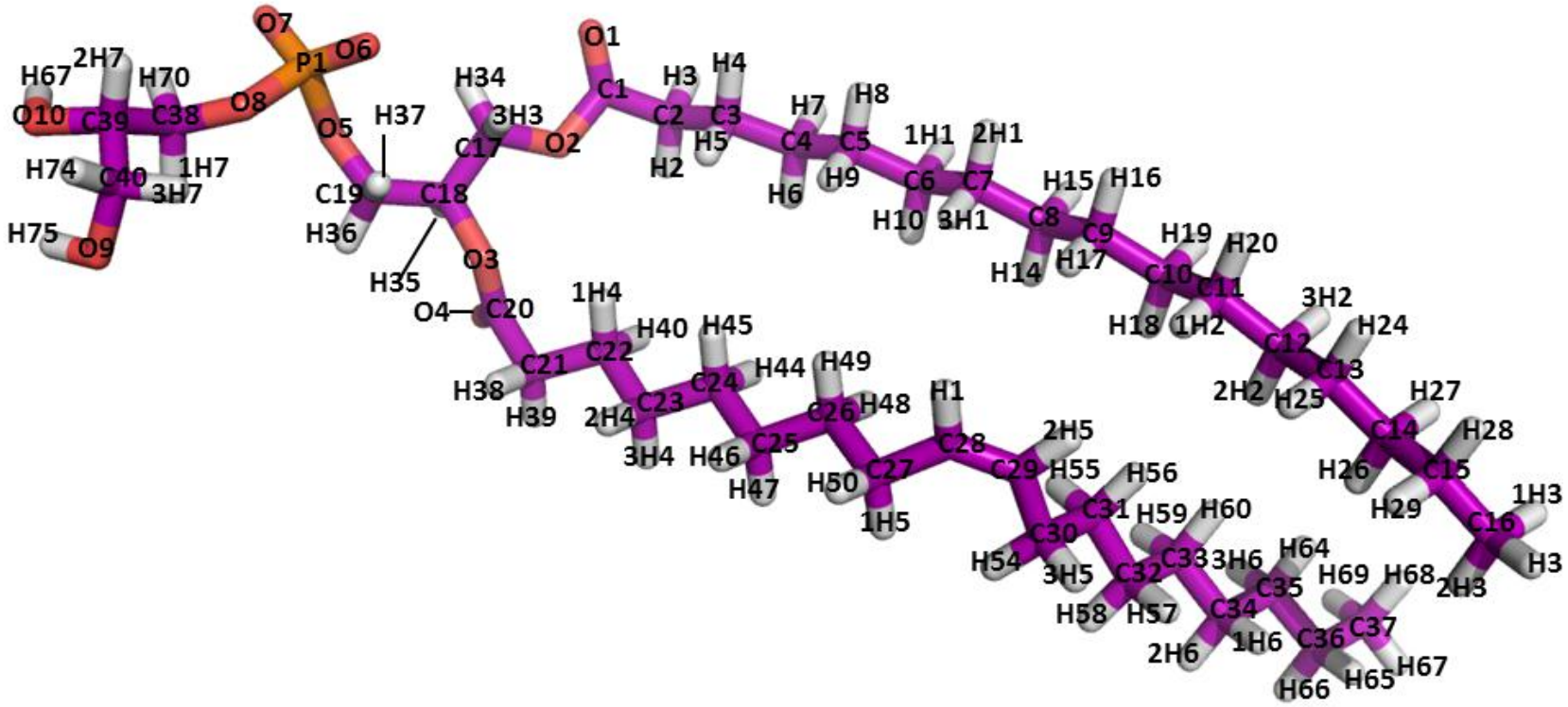
```
remark goes here
MASS
C  12.010        0.616              same as c
CT 12.010        0.878              same as c3
O  16.000        0.434              same as o
OS 16.000        0.465              same as os
CM 12.010        0.360              same as c2
P  30.970        1.538              same as p4
O2 16.000        0.434              same as o
N3 14.010        0.530              same as n4
HP 1.008         0.135              same as hc
HA 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc

BOND
C -CT  328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o
```

```
C -OS  411.30   1.343       same as c -os
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
OS-CT  301.50   1.439       same as c3-os
CT-H1  337.30   1.092       same as c3-hc
CT-CM  328.30   1.508       same as c2-c3
CM-CM  418.30   1.429       same as cc-cc
CM-HA  344.30   1.087       same as c2-hc
OS-P   311.60   1.636       same as os-p4
P -O2  456.40   1.503       same as o -p4
CT-N3  293.60   1.499       same as c3-n4
CT-HP  337.30   1.092       same as c3-hc

ANGLE
C -CT-CT   63.800     110.530   same as c -c3-c3
C -CT-HC   47.200     109.680   same as c -c3-hc
C -OS-CT   63.600     115.140   same as c -os-c3
CT-C -O    68.000     123.110   same as c3-c -o
CT-C -OS   69.300     111.960   same as c3-c -os
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
O -C -OS   76.200     122.430   same as o -c -os
OS-CT-CT   67.800     108.420   same as c3-c3-os
OS-CT-H1   50.900     108.700   same as hc-c3-os
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-OS-P    77.600     117.480   same as c3-os-p4
CT-CT-CM   63.700     110.960   same as c2-c3-c3
CT-CM-CM   64.300     123.420   same as c2-c2-c3
CT-CM-HA   45.100     120.000   same as c3-c2-hc
CM-CT-HC   47.000     110.490   same as c2-c3-hc
CM-CM-HA   50.300     119.700   same as c2-c2-hc
OS-P -O2   43.100     116.670   same as o -p4-os
OS-P -OS   44.700     100.340   same as os-p4-os
O2-P -O2   45.100     117.220   same as o -p4-o
CT-CT-N3   66.000     108.930   same as c3-c3-n4
CT-CT-HP   46.400     110.050   same as c3-c3-hc
CT-N3-CT   62.800     110.640   same as c3-n4-c3
```

```
N3-CT-HP   49.000     107.900   same as hc-c3-n4
HP-CT-HP   39.400     108.350   same as hc-c3-hc
HC-CT-HC   39.400     108.350   same as hc-c3-hc
H1-CT-H1   39.400     108.350   same as hc-c3-hc

DIHE
C -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
C -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
C -OS-CT-CT   1    0.383         0.000          3.000      same as X -c3-os-X
C -OS-CT-H1   1    0.383         0.000          3.000      same as X -c3-os-X
CT-C -OS-CT   1    2.700       180.000          2.000      same as X -c -os-X
CT-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-CT-C -OS   1    0.000       180.000          2.000      same as X -c -c3-X
O -C -CT-HC   1    0.000       180.000          2.000      same as X -c -c3-X
O -C -OS-CT   1    2.700       180.000          2.000      same as X -c -os-X
OS-C -CT-HC   1    0.000       180.000          2.000      same as X -c -c3-X
OS-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
OS-CT-CT-OS   1    0.156         0.000          3.000      same as X -c3-c3-X
OS-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-OS-P    1    0.383         0.000          3.000      same as X -c3-os-X
CT-OS-P -O2   1    1.050       180.000          2.000      same as X -os-p4-X
CT-OS-P -OS   1    1.050       180.000          2.000      same as X -os-p4-X
CT-CT-CT-CM   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CM-CM   1    0.000         0.000          2.000      same as X -c2-c3-X
CT-CT-CM-HA   1    0.000         0.000          2.000      same as X -c2-c3-X
CT-CM-CM-CT   1    6.650       180.000          2.000      same as X -c2-c2-X
CT-CM-CM-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
CM-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
CM-CM-CT-HC   1    0.000         0.000          2.000      same as X -c2-c3-X
P -OS-CT-H1   1    0.383         0.000          3.000      same as X -c3-os-X
OS-CT-CT-N3   1    0.156         0.000          3.000      same as X -c3-c3-X
OS-CT-CT-HP   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-N3-CT   1    0.156         0.000          3.000      same as X -c3-n4-X
CT-N3-CT-HP   1    0.156         0.000          3.000      same as X -c3-n4-X
```

```
N3-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HP-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HA-CM-CT-HC   1    0.000         0.000          2.000      same as X -c2-c3-X
HA-CM-CM-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
HC-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X

IMPROPER
CT-O -C -OS          1.1         180.0          2.0        Using default value
CM-CT-CM-HA          1.1         180.0          2.0        Using default value

NONBON
  C           1.9080  0.0860            same as c
  CT          1.9080  0.1094            same as c3
  O           1.6612  0.2100            same as o
  OS          1.6837  0.1700            same as os
  CM          1.9080  0.0860            same as cc
  P           2.1000  0.2000            same as p4
  O2          1.6612  0.2100            same as o
  N3          1.8240  0.1700            same as n4
  HP          1.4870  0.0157            same as hc
  HA          1.4870  0.0157            same as hc
  HC          1.4870  0.0157            same as hc
  H1          1.4870  0.0157            same as hc
```

POPG



.mol2 file:

```
@<TRIPOS>MOLECULE
POG
  127   126     1     0     0
SMALL
resp
@<TRIPOS>ATOM
      1 C1          4.5080    3.4210    0.9000 C         1 POG       0.8405
      2 C2          3.1430    3.7480    1.5140 CT        1 POG      -0.4121
      3 C3          2.0230    3.5210    0.4600 CT        1 POG       0.0549
```

4	C4	0.5880	3.7820	0.9770	CT	1 POG	-0.0511
5	C5	-0.4760	3.5290	-0.1200	CT	1 POG	-0.0190
6	C6	-1.9480	3.6370	0.3480	CT	1 POG	-0.0574
7	C7	-2.9540	3.2990	-0.7810	CT	1 POG	0.1014
8	C8	-4.4420	3.2870	-0.3540	CT	1 POG	-0.0547
9	C9	-5.3880	2.8220	-1.4880	CT	1 POG	0.0496
10	C10	-6.8750	2.6920	-1.0790	CT	1 POG	0.0042
11	C11	-7.7570	2.0510	-2.1800	CT	1 POG	-0.0262
12	C12	-9.2090	1.7500	-1.7350	CT	1 POG	0.1264
13	C13	-10.0210	0.9280	-2.7660	CT	1 POG	-0.0691
14	C14	-11.4110	0.4790	-2.2540	CT	1 POG	0.0010
15	C15	-12.1940	-0.4050	-3.2540	CT	1 POG	0.1184
16	C16	-13.5480	-0.9100	-2.7050	CT	1 POG	-0.2710
17	O1	5.0860	4.1480	0.0990	O	1 POG	-0.6061
18	O2	4.9560	2.1990	1.3180	OS	1 POG	-0.4641
19	C17	6.1670	1.6590	0.6510	CT	1 POG	0.0155
20	C18	6.6040	0.4190	1.4280	CT	1 POG	0.2668
21	C19	7.8240	-0.3090	0.8040	CT	1 POG	0.0580
22	O3	5.4940	-0.5740	1.3840	OS	1 POG	-0.5269
23	C20	4.9890	-1.0650	2.5540	C	1 POG	0.7444
24	C21	3.8790	-2.0810	2.2620	CT	1 POG	-0.2676
25	O4	5.3620	-0.7530	3.6820	O	1 POG	-0.5884
26	C22	2.6350	-1.4470	1.5860	CT	1 POG	0.0447
27	C23	1.5450	-2.4790	1.2130	CT	1 POG	-0.0388
28	C24	0.3410	-1.8390	0.4820	CT	1 POG	0.0075
29	C25	-0.7330	-2.8500	0.0170	CT	1 POG	-0.0396
30	C26	-1.8870	-2.1870	-0.7720	CT	1 POG	0.0440
31	C27	-2.9500	-3.1920	-1.3000	CT	1 POG	0.0760
32	C28	-4.0000	-2.5240	-2.1630	CM	1 POG	-0.2495
33	C29	-5.3360	-2.4570	-1.9790	CM	1 POG	-0.2929
34	C30	-6.1790	-3.0520	-0.8690	CT	1 POG	0.0552
35	C31	-6.9920	-1.9820	-0.0850	CT	1 POG	0.0249
36	C32	-8.1240	-2.5680	0.7910	CT	1 POG	-0.1027
37	C33	-8.9360	-1.4890	1.5500	CT	1 POG	0.1728
38	C34	-10.2050	-2.0280	2.2520	CT	1 POG	-0.0701
39	C35	-11.0660	-0.9240	2.9110	CT	1 POG	0.0314
40	C36	-12.3730	-1.4480	3.5520	CT	1 POG	0.1143

```
41 C37       -13.2570   -0.3340    4.1550 CT        1 POG      -0.2351
42 O5          9.0030    0.5060    0.8360 OS        1 POG      -0.5591
43 P1          9.6820    0.9650   -0.6610 P         1 POG       1.2657
44 O6          8.5640    1.3180   -1.6070 O2        1 POG      -0.7815
45 O7         10.8550    1.8360   -0.2980 O2        1 POG      -0.7815
46 O8         10.2430   -0.5380   -1.2190 OS        1 POG      -0.5584
47 C38        11.4590   -1.0710   -0.6690 CT        1 POG       0.2972
48 C39        12.1520   -1.9070   -1.7640 CT        1 POG       0.1672
49 C40        11.4450   -3.2240   -2.1290 CT        1 POG       0.1872
50 O9         11.4830   -4.2040   -1.0620 OH        1 POG      -0.6631
51 O10        13.5130   -2.2860   -1.3510 OH        1 POG      -0.6877
52 H1         -3.5980   -2.0260   -3.0530 HA        1 POG       0.1314
53 H2          2.9670    3.1290    2.4040 HC        1 POG       0.1140
54 H3          3.1600    4.8050    1.8160 HC        1 POG       0.1140
55 H4          2.2260    4.1710   -0.4040 HC        1 POG       0.0269
56 H5          2.0870    2.4840    0.0940 HC        1 POG       0.0269
57 H6          0.3840    3.1300    1.8440 HC        1 POG       0.0287
58 H7          0.5080    4.8200    1.3420 HC        1 POG       0.0287
59 H8         -0.3090    4.2330   -0.9520 HC        1 POG       0.0081
60 H9         -0.3160    2.5210   -0.5390 HC        1 POG       0.0081
61 H10        -2.1110    2.9470    1.1940 HC        1 POG       0.0083
62 H11        -2.1460    4.6510    0.7350 HC        1 POG       0.0083
63 H12        -2.8190    4.0100   -1.6140 HC        1 POG      -0.0223
64 H13        -2.7000    2.3050   -1.1860 HC        1 POG      -0.0223
65 H14        -4.5600    2.6060    0.5070 HC        1 POG       0.0083
66 H15        -4.7410    4.2870    0.0030 HC        1 POG       0.0083
67 H16        -5.2970    3.5060   -2.3490 HC        1 POG      -0.0157
68 H17        -5.0390    1.8380   -1.8450 HC        1 POG      -0.0157
69 H18        -6.9370    2.0640   -0.1740 HC        1 POG      -0.0112
70 H19        -7.2780    3.6790   -0.7970 HC        1 POG      -0.0112
71 H20        -7.7640    2.6910   -3.0780 HC        1 POG      -0.0007
72 H21        -7.2840    1.1030   -2.4870 HC        1 POG      -0.0007
73 H22        -9.1720    1.1840   -0.7900 HC        1 POG      -0.0350
74 H23        -9.7370    2.6910   -1.5050 HC        1 POG      -0.0350
75 H24       -10.1380    1.5060   -3.6980 HC        1 POG       0.0034
76 H25        -9.4400    0.0300   -3.0370 HC        1 POG       0.0034
77 H26       -11.2790   -0.0820   -1.3130 HC        1 POG       0.0060
```

78	H27	-12.0160	1.3660	-1.9990	HC	1 POG	0.0060
79	H28	-12.3650	0.1580	-4.1860	HC	1 POG	-0.0117
80	H29	-11.5720	-1.2730	-3.5310	HC	1 POG	-0.0117
81	H30	-14.0790	-1.5200	-3.4520	HC	1 POG	0.0627
82	H31	-14.2050	-0.0710	-2.4280	HC	1 POG	0.0627
83	H32	-13.4060	-1.5320	-1.8070	HC	1 POG	0.0627
84	H33	5.9210	1.4310	-0.3920	H1	1 POG	0.1386
85	H34	6.9670	2.4040	0.6550	H1	1 POG	0.1386
86	H35	6.7930	0.6680	2.4770	H1	1 POG	0.0878
87	H36	8.0060	-1.2150	1.4060	H1	1 POG	0.0894
88	H37	7.5720	-0.6140	-0.2210	H1	1 POG	0.0894
89	H38	4.2950	-2.8570	1.6010	HC	1 POG	0.0797
90	H39	3.6080	-2.5510	3.2160	HC	1 POG	0.0797
91	H40	2.2090	-0.6850	2.2590	HC	1 POG	0.0320
92	H41	2.9590	-0.9210	0.6780	HC	1 POG	0.0320
93	H42	1.9910	-3.2520	0.5640	HC	1 POG	0.0215
94	H43	1.1950	-3.0010	2.1200	HC	1 POG	0.0215
95	H44	-0.1250	-1.0840	1.1380	HC	1 POG	-0.0010
96	H45	0.7150	-1.2900	-0.3980	HC	1 POG	-0.0010
97	H46	-0.2570	-3.6170	-0.6170	HC	1 POG	-0.0027
98	H47	-1.1430	-3.3840	0.8920	HC	1 POG	-0.0027
99	H48	-2.3870	-1.4350	-0.1400	HC	1 POG	0.0083
100	H49	-1.4640	-1.6350	-1.6280	HC	1 POG	0.0083
101	H50	-2.4300	-3.9530	-1.9090	HC	1 POG	0.0192
102	H51	-3.4060	-3.7290	-0.4560	HC	1 POG	0.0192
103	H52	-5.9090	-1.9040	-2.7320	HA	1 POG	0.1476
104	H53	-6.8880	-3.7710	-1.3180	HC	1 POG	0.0279
105	H54	-5.5600	-3.6290	-0.1670	HC	1 POG	0.0279
106	H55	-6.2970	-1.3990	0.5410	HC	1 POG	0.0201
107	H56	-7.4290	-1.2680	-0.8010	HC	1 POG	0.0201
108	H57	-8.8080	-3.1470	0.1460	HC	1 POG	0.0091
109	H58	-7.7020	-3.2870	1.5140	HC	1 POG	0.0091
110	H59	-8.2860	-0.9930	2.2900	HC	1 POG	-0.0410
111	H60	-9.2380	-0.7050	0.8360	HC	1 POG	-0.0410
112	H61	-10.8190	-2.5680	1.5110	HC	1 POG	0.0047
113	H62	-9.9190	-2.7740	3.0140	HC	1 POG	0.0047
114	H63	-10.4680	-0.4030	3.6780	HC	1 POG	-0.0118

```
     115 H64        -11.3210   -0.1640    2.1520 HC        1 POG      -0.0118
     116 H65        -12.9520   -1.9970    2.7900 HC        1 POG      -0.0171
     117 H66        -12.1260   -2.1830    4.3360 HC        1 POG      -0.0171
     118 H67        -14.1740   -0.7500    4.6010 HC        1 POG       0.0535
     119 H68        -13.5600    0.3940    3.3870 HC        1 POG       0.0535
     120 H69        -12.7220    0.2180    4.9440 HC        1 POG       0.0535
     121 H70         12.1240   -0.2420   -0.3730 H1        1 POG       0.0074
     122 H71         11.2490   -1.6900    0.2190 H1        1 POG       0.0074
     123 H72         12.2070   -1.2890   -2.6760 H1        1 POG       0.0340
     124 H73         10.3890   -3.0160   -2.3380 H1        1 POG       0.0373
     125 H74         11.9130   -3.6420   -3.0390 H1        1 POG       0.0373
     126 H75         12.3940   -4.1550   -0.7100 HO        1 POG       0.3996
     127 H76         13.8780   -1.5260   -0.8600 HO        1 POG       0.3985
@<TRIPOS>BOND
     1    1    2 1            24    8   66 1            47   16   82 1            70   27   94 1
     2    1   17 2            25    9   10 1            48   16   83 1            71   28   29 1
     3    1   18 1            26    9   67 1            49   18   19 1            72   28   95 1
     4    2    3 1            27    9   68 1            50   19   20 1            73   28   96 1
     5    2   53 1            28   10   11 1            51   19   84 1            74   29   30 1
     6    2   54 1            29   10   69 1            52   19   85 1            75   29   97 1
     7    3    4 1            30   10   70 1            53   20   21 1            76   29   98 1
     8    3   55 1            31   11   12 1            54   20   22 1            77   30   31 1
     9    3   56 1            32   11   71 1            55   20   86 1            78   30   99 1
    10    4    5 1            33   11   72 1            56   21   42 1            79   30  100 1
    11    4   57 1            34   12   13 1            57   21   87 1            80   31   32 1
    12    4   58 1            35   12   73 1            58   21   88 1            81   31  101 1
    13    5    6 1            36   12   74 1            59   22   23 1            82   31  102 1
    14    5   59 1            37   13   14 1            60   23   24 1            83   32   33 2
    15    5   60 1            38   13   75 1            61   23   25 2            84   32   52 1
    16    6    7 1            39   13   76 1            62   24   26 1            85   33   34 1
    17    6   61 1            40   14   15 1            63   24   89 1            86   33  103 1
    18    6   62 1            41   14   77 1            64   24   90 1            87   34   35 1
    19    7    8 1            42   14   78 1            65   26   27 1            88   34  104 1
    20    7   63 1            43   15   16 1            66   26   91 1            89   34  105 1
    21    7   64 1            44   15   79 1            67   26   92 1            90   35   36 1
    22    8    9 1            45   15   80 1            68   27   28 1            91   35  106 1
    23    8   65 1            46   16   81 1            69   27   93 1            92   35  107 1
```

```
    93   36   37 1          102   39   40 1          111   42   43 1          120   48   51 1
    94   36  108 1          103   39  114 1          112   43   44 1          121   48  123 1
    95   36  109 1          104   39  115 1          113   43   45 1          122   49   50 1
    96   37   38 1          105   40   41 1          114   43   46 1          123   49  124 1
    97   37  110 1          106   40  116 1          115   46   47 1          124   49  125 1
    98   37  111 1          107   40  117 1          116   47   48 1          125   50  126 1
    99   38   39 1          108   41  118 1          117   47  121 1          126   51  127 1
   100   38  112 1          109   41  119 1          118   47  122 1
   101   38  113 1          110   41  120 1          119   48   49 1
@<TRIPOS>SUBSTRUCTURE
     1 POG         1 TEMP              0 ****  ****    0 ROOT
```

.frcmod file:

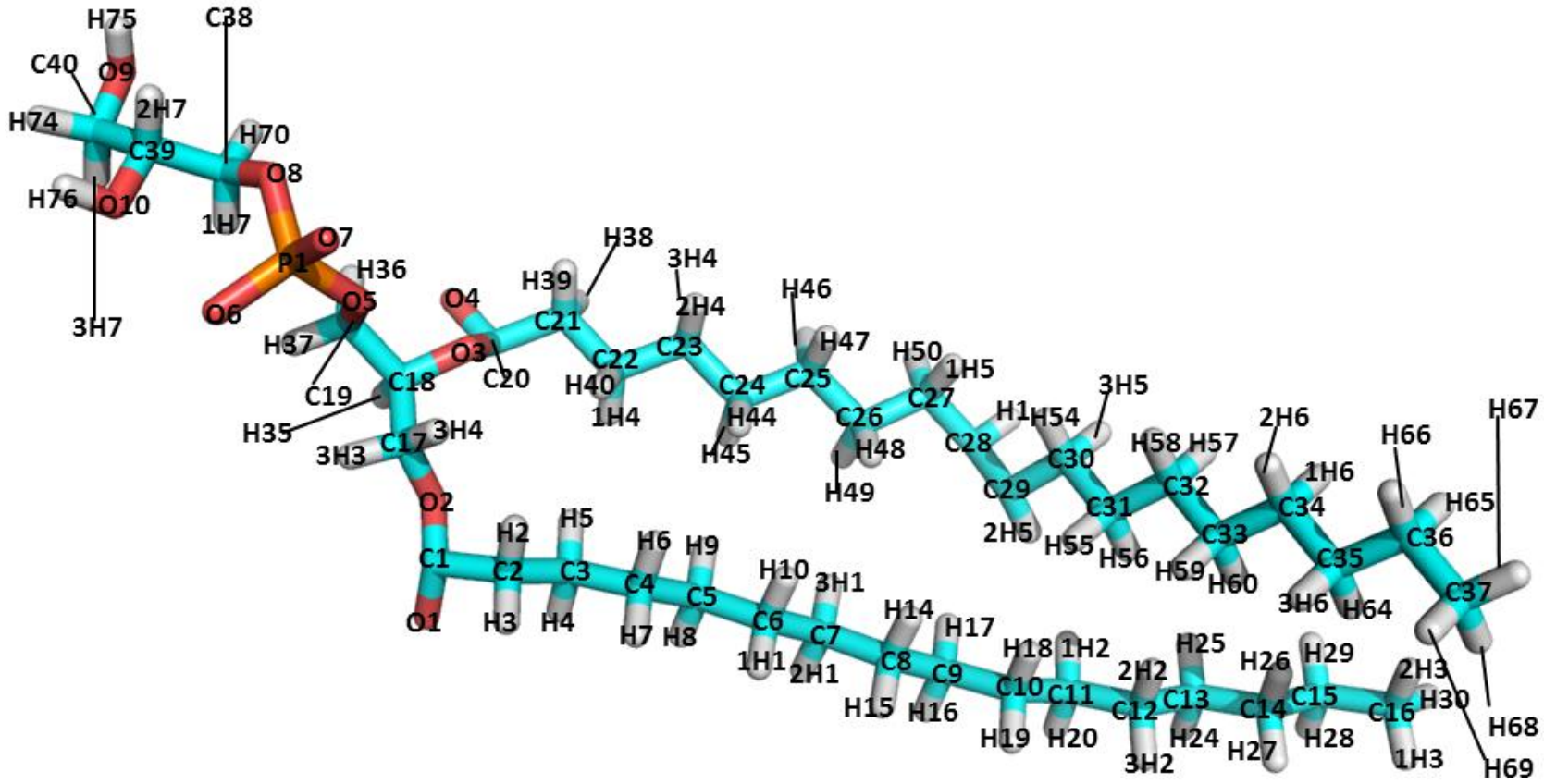
```
remark goes here
MASS
C  12.010        0.616              same as c
CT 12.010        0.878              same as c3
O  16.000        0.434              same as o
OS 16.000        0.465              same as os
CM 12.010        0.360              same as c2
P  30.970        1.538              same as p4
O2 16.000        0.434              same as o
OH 16.000        0.465              same as oh
HA 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
HO 1.008         0.135              same as ho

BOND
C -CT  328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o
C -OS  411.30   1.343       same as c -os
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
OS-CT  301.50   1.439       same as c3-os
CT-H1  337.30   1.092       same as c3-hc
```

```
CT-CM  328.30   1.508       same as c2-c3
CM-CM  418.30   1.429       same as cc-cc
CM-HA  344.30   1.087       same as c2-hc
OS-P   311.60   1.636       same as os-p4
P -O2  456.40   1.503       same as o -p4
CT-OH  314.10   1.426       same as c3-oh
OH-HO  369.60   0.974       same as ho-oh

ANGLE
C -CT-CT   63.800     110.530   same as c -c3-c3
C -CT-HC   47.200     109.680   same as c -c3-hc
C -OS-CT   63.600     115.140   same as c -os-c3
CT-C -O    68.000     123.110   same as c3-c -o
CT-C -OS   69.300     111.960   same as c3-c -os
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
O -C -OS   76.200     122.430   same as o -c -os
OS-CT-CT   67.800     108.420   same as c3-c3-os
OS-CT-H1   50.900     108.700   same as hc-c3-os
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-OS-P    77.600     117.480   same as c3-os-p4
CT-CT-CM   63.700     110.960   same as c2-c3-c3
CT-CM-CM   64.300     123.420   same as c2-c2-c3
CT-CM-HA   45.100     120.000   same as c3-c2-hc
CM-CT-HC   47.000     110.490   same as c2-c3-hc
CM-CM-HA   50.300     119.700   same as c2-c2-hc
OS-P -O2   43.100     116.670   same as o -p4-os
OS-P -OS   44.700     100.340   same as os-p4-os
O2-P -O2   45.100     117.220   same as o -p4-o
CT-CT-OH   67.700     109.430   same as c3-c3-oh
CT-OH-HO   47.100     108.160   same as c3-oh-ho
OH-CT-H1   51.100     109.500   same as hc-c3-oh
HC-CT-HC   39.400     108.350   same as hc-c3-hc
H1-CT-H1   39.400     108.350   same as hc-c3-hc

DIHE
C -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
```

C -CT-CT-HC	1	0.156	0.000	3.000	same as X -c3-c3-X
C -OS-CT-CT	1	0.383	0.000	3.000	same as X -c3-os-X
C -OS-CT-H1	1	0.383	0.000	3.000	same as X -c3-os-X
CT-C -OS-CT	1	2.700	180.000	2.000	same as X -c -os-X
CT-CT-CT-CT	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-CT-HC	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-C -O	1	0.000	180.000	2.000	same as X -c -c3-X
CT-CT-C -OS	1	0.000	180.000	2.000	same as X -c -c3-X
O -C -CT-HC	1	0.000	180.000	2.000	same as X -c -c3-X
O -C -OS-CT	1	2.700	180.000	2.000	same as X -c -os-X
OS-C -CT-HC	1	0.000	180.000	2.000	same as X -c -c3-X
OS-CT-CT-CT	1	0.156	0.000	3.000	same as X -c3-c3-X
OS-CT-CT-OS	1	0.156	0.000	3.000	same as X -c3-c3-X
OS-CT-CT-H1	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-CT-H1	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-OS-P	1	0.383	0.000	3.000	same as X -c3-os-X
CT-OS-P -O2	1	1.050	180.000	2.000	same as X -os-p4-X
CT-OS-P -OS	1	1.050	180.000	2.000	same as X -os-p4-X
CT-CT-CT-CM	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-CM-CM	1	0.000	0.000	2.000	same as X -c2-c3-X
CT-CT-CM-HA	1	0.000	0.000	2.000	same as X -c2-c3-X
CT-CM-CM-CT	1	6.650	180.000	2.000	same as X -c2-c2-X
CT-CM-CM-HA	1	6.650	180.000	2.000	same as X -c2-c2-X
CM-CT-CT-HC	1	0.156	0.000	3.000	same as X -c3-c3-X
CM-CM-CT-HC	1	0.000	0.000	2.000	same as X -c2-c3-X
P -OS-CT-H1	1	0.383	0.000	3.000	same as X -c3-os-X
OS-CT-CT-OH	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-CT-OH	1	0.156	0.000	3.000	same as X -c3-c3-X
CT-CT-OH-HO	1	0.167	0.000	3.000	same as X -c3-oh-X
OH-CT-CT-OH	1	0.156	0.000	3.000	same as X -c3-c3-X
OH-CT-CT-H1	1	0.156	0.000	3.000	same as X -c3-c3-X
HA-CM-CT-HC	1	0.000	0.000	2.000	same as X -c2-c3-X
HA-CM-CM-HA	1	6.650	180.000	2.000	same as X -c2-c2-X
HC-CT-CT-HC	1	0.156	0.000	3.000	same as X -c3-c3-X
H1-CT-CT-H1	1	0.156	0.000	3.000	same as X -c3-c3-X
H1-CT-OH-HO	1	0.167	0.000	3.000	same as X -c3-oh-X

```
IMPROPER
CT-O -C -OS         1.1          180.0         2.0          Using default value
CM-CT-CM-HA         1.1          180.0         2.0          Using default value

NONBON
  C           1.9080  0.0860            same as c
  CT          1.9080  0.1094            same as c3
  O           1.6612  0.2100            same as o
  OS          1.6837  0.1700            same as os
  CM          1.9080  0.0860            same as cc
  P           2.1000  0.2000            same as p4
  O2          1.6612  0.2100            same as o
  OH          1.7210  0.2104            same as oh
  HA          1.4870  0.0157            same as hc
  HC          1.4870  0.0157            same as hc
  H1          1.4870  0.0157            same as hc
  HO          0.0000  0.0000            same as ho
```

POPG2



.mol2 file:

```
PG2
  127   126     1     0     0
SMALL
resp

@<TRIPOS>ATOM
      1 C1          4.6850   -3.2340   -2.1100 C         1 PG2       0.5134
```

2	C2	3.3330	-2.5780	-2.4540	CT	1 PG2	-0.0300
3	C3	2.2460	-2.8670	-1.3990	CT	1 PG2	0.0266
4	C4	0.8740	-2.3210	-1.8330	CT	1 PG2	-0.1017
5	C5	-0.2780	-2.7510	-0.9110	CT	1 PG2	0.0344
6	C6	-1.6410	-2.3010	-1.4590	CT	1 PG2	0.0179
7	C7	-2.8310	-2.7670	-0.6070	CT	1 PG2	0.0765
8	C8	-4.1730	-2.3300	-1.2180	CT	1 PG2	-0.0897
9	C9	-5.3930	-2.7730	-0.3950	CT	1 PG2	0.0479
10	C10	-6.7070	-2.2280	-0.9780	CT	1 PG2	-0.0310
11	C11	-7.9140	-2.4500	-0.0540	CT	1 PG2	0.0828
12	C12	-9.1930	-1.7990	-0.6010	CT	1 PG2	0.0934
13	C13	-10.3480	-1.8290	0.4100	CT	1 PG2	-0.0295
14	C14	-11.6130	-1.1360	-0.1200	CT	1 PG2	0.0346
15	C15	-12.7490	-1.1150	0.9140	CT	1 PG2	0.1847
16	C16	-14.0020	-0.4010	0.3940	CT	1 PG2	-0.2761
17	O1	4.5300	-4.6260	-1.9900	O	1 PG2	-0.5306
18	O2	5.1820	-2.6970	-0.9300	OS	1 PG2	-0.3606
19	C17	6.0190	-1.5540	-1.1130	CT	1 PG2	0.0697
20	C18	6.5790	-1.1060	0.2440	CT	1 PG2	0.1684
21	C19	7.8460	-0.2390	0.1260	CT	1 PG2	0.1898
22	O3	5.5780	-0.3650	0.9350	OS	1 PG2	-0.2960
23	C20	5.2460	-0.8520	2.1910	C	1 PG2	0.5494
24	C21	4.0910	0.0080	2.7220	CT	1 PG2	-0.1930
25	O4	6.3350	-0.7610	3.0840	O	1 PG2	-0.5451
26	C22	2.7770	-0.3190	1.9930	CT	1 PG2	-0.0446
27	C23	1.6040	0.4980	2.5590	CT	1 PG2	0.0052
28	C24	0.2210	-0.0190	2.1310	CT	1 PG2	0.0535
29	C25	-0.9120	0.7400	2.8410	CT	1 PG2	-0.0997
30	C26	-2.2920	0.0860	2.6710	CT	1 PG2	0.0389
31	C27	-3.3640	0.7700	3.5370	CT	1 PG2	0.1552
32	C28	-4.6850	-0.0160	3.6320	CM	1 PG2	-0.3724
33	C29	-5.5330	-0.0160	2.3500	CM	1 PG2	-0.0937
34	C30	-6.2380	1.3220	2.0700	CT	1 PG2	-0.3867
35	C31	-7.2050	1.2190	0.8840	CT	1 PG2	0.0095
36	C32	-7.9460	2.5300	0.5940	CT	1 PG2	-0.0534
37	C33	-8.9330	2.3810	-0.5730	CT	1 PG2	0.1961
38	C34	-9.6890	3.6800	-0.8870	CT	1 PG2	-0.0575

39	C35	-10.6630	3.5170	-2.0620	CT	1 PG2	0.0115
40	C36	-11.4220	4.8120	-2.3820	CT	1 PG2	0.1702
41	C37	-12.3820	4.6510	-3.5670	CT	1 PG2	-0.2476
42	O5	7.6310	0.8500	-0.7510	OS	1 PG2	-0.4956
43	P1	8.9160	1.6660	-1.4710	P	1 PG2	1.1120
44	O6	9.8240	0.6540	-2.4440	O2	1 PG2	-0.7838
45	O7	8.2800	2.8470	-2.4660	O2	1 PG2	-0.7838
46	O8	9.8770	2.4930	-0.3610	OS	1 PG2	-0.5449
47	C38	10.8390	1.8070	0.4170	CT	1 PG2	0.3690
48	C39	12.2110	2.0780	-0.2250	CT	1 PG2	0.0480
49	C40	13.3910	1.5450	0.6090	CT	1 PG2	0.0598
50	O9	13.4950	2.2260	1.8410	OH	1 PG2	-0.6693
51	O10	12.2520	1.4250	-1.4780	OH	1 PG2	-0.5620
52	H1	-5.2810	0.3820	4.4540	HA	1 PG2	0.1832
53	H2	3.4760	-1.5030	-2.5690	HC	1 PG2	0.0216
54	H3	3.0200	-2.9610	-3.4250	HC	1 PG2	0.0216
55	H4	2.1620	-3.9430	-1.2450	HC	1 PG2	0.0492
56	H5	2.5350	-2.4410	-0.4390	HC	1 PG2	0.0492
57	H6	0.9160	-1.2330	-1.8870	HC	1 PG2	0.0329
58	H7	0.6620	-2.6710	-2.8420	HC	1 PG2	0.0329
59	H8	-0.2730	-3.8380	-0.8080	HC	1 PG2	-0.0171
60	H9	-0.1240	-2.3450	0.0890	HC	1 PG2	-0.0171
61	H10	-1.6550	-1.2120	-1.5390	HC	1 PG2	-0.0095
62	H11	-1.7610	-2.6820	-2.4740	HC	1 PG2	-0.0095
63	H12	-2.8120	-3.8530	-0.5120	HC	1 PG2	-0.0203
64	H13	-2.7380	-2.3640	0.4020	HC	1 PG2	-0.0203
65	H14	-4.1800	-1.2430	-1.3140	HC	1 PG2	0.0167
66	H15	-4.2590	-2.7270	-2.2310	HC	1 PG2	0.0167
67	H16	-5.4340	-3.8610	-0.3460	HC	1 PG2	-0.0146
68	H17	-5.2750	-2.4270	0.6310	HC	1 PG2	-0.0146
69	H18	-6.5970	-1.1600	-1.1640	HC	1 PG2	-0.0009
70	H19	-6.8950	-2.6880	-1.9490	HC	1 PG2	-0.0009
71	H20	-8.0760	-3.5170	0.0970	HC	1 PG2	-0.0346
72	H21	-7.6900	-2.0320	0.9280	HC	1 PG2	-0.0346
73	H22	-8.9860	-0.7620	-0.8700	HC	1 PG2	-0.0407
74	H23	-9.4940	-2.2990	-1.5220	HC	1 PG2	-0.0407
75	H24	-10.5790	-2.8620	0.6730	HC	1 PG2	-0.0130

76	H25	-10.0330	-1.3390	1.3320	HC	1 PG2	-0.0130
77	H26	-11.3740	-0.1130	-0.4110	HC	1 PG2	-0.0143
78	H27	-11.9530	-1.6410	-1.0250	HC	1 PG2	-0.0143
79	H28	-13.0070	-2.1360	1.1980	HC	1 PG2	-0.0355
80	H29	-12.4070	-0.6200	1.8240	HC	1 PG2	-0.0355
81	H30	-14.7900	-0.4010	1.1490	HC	1 PG2	0.0592
82	H31	-14.3970	-0.8920	-0.4960	HC	1 PG2	0.0592
83	H32	-13.7890	0.6370	0.1390	HC	1 PG2	0.0592
84	H33	6.8540	-1.8260	-1.7610	H1	1 PG2	0.0843
85	H34	5.4810	-0.7380	-1.5990	H1	1 PG2	0.0843
86	H35	6.8580	-2.0050	0.7970	H1	1 PG2	0.0374
87	H36	8.1510	0.1290	1.1050	H1	1 PG2	0.0458
88	H37	8.6650	-0.8490	-0.2580	H1	1 PG2	0.0458
89	H38	3.9710	-0.1830	3.7880	HC	1 PG2	0.1118
90	H39	4.3560	1.0590	2.6030	HC	1 PG2	0.1118
91	H40	2.8810	-0.1320	0.9220	HC	1 PG2	0.0129
92	H41	2.5630	-1.3840	2.0970	HC	1 PG2	0.0129
93	H42	1.6560	0.4740	3.6490	HC	1 PG2	0.0099
94	H43	1.7160	1.5450	2.2750	HC	1 PG2	0.0099
95	H44	0.1060	0.0580	1.0490	HC	1 PG2	0.0024
96	H45	0.1500	-1.0800	2.3740	HC	1 PG2	0.0024
97	H46	-0.6850	0.7900	3.9070	HC	1 PG2	0.0083
98	H47	-0.9450	1.7710	2.4860	HC	1 PG2	0.0083
99	H48	-2.5840	0.1000	1.6200	HC	1 PG2	0.0101
100	H49	-2.2220	-0.9670	2.9530	HC	1 PG2	0.0101
101	H50	-2.9740	0.8770	4.5500	HC	1 PG2	-0.0093
102	H51	-3.5470	1.7840	3.1820	HC	1 PG2	-0.0093
103	H52	-6.2920	-0.7940	2.4490	HA	1 PG2	0.1242
104	H53	-6.7870	1.6400	2.9550	HC	1 PG2	0.1653
105	H54	-5.4970	2.0960	1.8740	HC	1 PG2	0.1653
106	H55	-6.6500	0.9130	-0.0050	HC	1 PG2	0.0533
107	H56	-7.9350	0.4330	1.0830	HC	1 PG2	0.0533
108	H57	-8.4850	2.8480	1.4870	HC	1 PG2	-0.0001
109	H58	-7.2250	3.3170	0.3690	HC	1 PG2	-0.0001
110	H59	-8.3950	2.0490	-1.4610	HC	1 PG2	-0.0541
111	H60	-9.6520	1.5950	-0.3370	HC	1 PG2	-0.0541
112	H61	-10.2370	4.0050	-0.0000	HC	1 PG2	-0.0021

```
      113 H62         -8.9720    4.4700   -1.1150 HC        1 PG2       -0.0021
      114 H63        -10.1130    3.1910   -2.9470 HC        1 PG2       -0.0075
      115 H64        -11.3770    2.7230   -1.8360 HC        1 PG2       -0.0075
      116 H65        -11.9850    5.1330   -1.5050 HC        1 PG2       -0.0302
      117 H66        -10.7120    5.6100   -2.6000 HC        1 PG2       -0.0302
      118 H67        -12.9090    5.5840   -3.7700 HC        1 PG2        0.0522
      119 H68        -13.1310    3.8840   -3.3670 HC        1 PG2        0.0522
      120 H69        -11.8460    4.3690   -4.4730 HC        1 PG2        0.0522
      121 H70         10.7980    2.1760    1.4420 H1        1 PG2        0.0176
      122 H71         10.6230    0.7380    0.4580 H1        1 PG2        0.0176
      123 H72         12.3300    3.1520   -0.3860 H1        1 PG2        0.0756
      124 H73         13.2720    0.4760    0.7960 H1        1 PG2        0.0630
      125 H74         14.3290    1.6710    0.0650 H1        1 PG2        0.0630
      126 H75         13.7010    3.1360    1.6830 HO        1 PG2        0.4116
      127 H76         13.0430    1.6980   -1.9250 HO        1 PG2        0.3918
@<TRIPOS>BOND                    22    8    9 1          44   15   79 1          66   26   91 1
     1    1    2 1              23    8   65 1          45   15   80 1          67   26   92 1
     2    1   17 2              24    8   66 1          46   16   81 1          68   27   28 1
     3    1   18 1              25    9   10 1          47   16   82 1          69   27   93 1
     4    2    3 1              26    9   67 1          48   16   83 1          70   27   94 1
     5    2   53 1              27    9   68 1          49   18   19 1          71   28   29 1
     6    2   54 1              28   10   11 1          50   19   20 1          72   28   95 1
     7    3    4 1              29   10   69 1          51   19   84 1          73   28   96 1
     8    3   55 1              30   10   70 1          52   19   85 1          74   29   30 1
     9    3   56 1              31   11   12 1          53   20   21 1          75   29   97 1
    10    4    5 1              32   11   71 1          54   20   22 1          76   29   98 1
    11    4   57 1              33   11   72 1          55   20   86 1          77   30   31 1
    12    4   58 1              34   12   13 1          56   21   42 1          78   30   99 1
    13    5    6 1              35   12   73 1          57   21   87 1          79   30  100 1
    14    5   59 1              36   12   74 1          58   21   88 1          80   31   32 1
    15    5   60 1              37   13   14 1          59   22   23 1          81   31  101 1
    16    6    7 1              38   13   75 1          60   23   24 1          82   31  102 1
    17    6   61 1              39   13   76 1          61   23   25 2          83   32   33 2
    18    6   62 1              40   14   15 1          62   24   26 1          84   32   52 1
    19    7    8 1              41   14   77 1          63   24   89 1          85   33   34 1
    20    7   63 1              42   14   78 1          64   24   90 1          86   33  103 1
    21    7   64 1              43   15   16 1          65   26   27 1          87   34   35 1
```

```
 88   34  104 1
 89   34  105 1
 90   35   36 1
 91   35  106 1
 92   35  107 1
 93   36   37 1
 94   36  108 1
 95   36  109 1
 96   37   38 1
 97   37  110 1
 98   37  111 1
 99   38   39 1
100   38  112 1
101   38  113 1
102   39   40 1
103   39  114 1
104   39  115 1
105   40   41 1
106   40  116 1
107   40  117 1
108   41  118 1
109   41  119 1
110   41  120 1
111   42   43 1
112   43   44 1
113   43   45 1
114   43   46 1
115   46   47 1
116   47   48 1
117   47  121 1
118   47  122 1
119   48   49 1
120   48   51 1
121   48  123 1
122   49   50 1
123   49  124 1
124   49  125 1
125   50  126 1
126   51  127 1
```

.frcmod file:

```
PG2 .frcmod input
MASS
C  12.010        0.616              same as c
CT 12.010        0.878              same as c3
O  16.000        0.434              same as o
OS 16.000        0.465              same as os
CM 12.010        0.360              same as c2
P  30.970        1.538              same as p4
O2 16.000        0.434              same as o
OH 16.000        0.465              same as oh
HA 1.008         0.135              same as hc
HC 1.008         0.135              same as hc
H1 1.008         0.135              same as hc
HO 1.008         0.135              same as ho

BOND
C -CT  328.30   1.508       same as c -c3
C -O   648.00   1.214       same as c -o
C -OS  411.30   1.343       same as c -os
CT-CT  303.10   1.535       same as c3-c3
CT-HC  337.30   1.092       same as c3-hc
OS-CT  301.50   1.439       same as c3-os
CT-H1  337.30   1.092       same as c3-hc
CT-CM  328.30   1.508       same as c2-c3
CM-CM  418.30   1.429       same as cc-cc
CM-HA  344.30   1.087       same as c2-hc
OS-P   311.60   1.636       same as os-p4
P -O2  456.40   1.503       same as o -p4
CT-OH  314.10   1.426       same as c3-oh
OH-HO  369.60   0.974       same as ho-oh

ANGLE
C -CT-CT   63.800     110.530   same as c -c3-c3
C -CT-HC   47.200     109.680   same as c -c3-hc
C -OS-CT   63.600     115.140   same as c -os-c3
CT-C -O    68.000     123.110   same as c3-c -o
CT-C -OS   69.300     111.960   same as c3-c -os
CT-CT-CT   63.200     110.630   same as c3-c3-c3
CT-CT-HC   46.400     110.050   same as c3-c3-hc
O -C -OS   76.200     122.430   same as o -c -os
OS-CT-CT   67.800     108.420   same as c3-c3-os
OS-CT-H1   50.900     108.700   same as hc-c3-os
CT-CT-H1   46.400     110.050   same as c3-c3-hc
CT-OS-P    77.600     117.480   same as c3-os-p4
CT-CT-CM   63.700     110.960   same as c2-c3-c3
CT-CM-CM   64.300     123.420   same as c2-c2-c3
CT-CM-HA   45.100     120.000   same as c3-c2-hc
CM-CT-HC   47.000     110.490   same as c2-c3-hc
CM-CM-HA   50.300     119.700   same as c2-c2-hc
```

```
OS-P -O2   43.100     116.670   same as o -p4-os
OS-P -OS   44.700     100.340   same as os-p4-os
O2-P -O2   45.100     117.220   same as o -p4-o
CT-CT-OH   67.700     109.430   same as c3-c3-oh
CT-OH-HO   47.100     108.160   same as c3-oh-ho
OH-CT-H1   51.100     109.500   same as hc-c3-oh
HC-CT-HC   39.400     108.350   same as hc-c3-hc
H1-CT-H1   39.400     108.350   same as hc-c3-hc

DIHE
C -CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
C -CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
C -OS-CT-CT   1    0.383         0.000          3.000      same as X -c3-os-X
C -OS-CT-H1   1    0.383         0.000          3.000      same as X -c3-os-X
CT-C -OS-CT   1    2.700       180.000          2.000      same as X -c -os-X
CT-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-C -O    1    0.000       180.000          2.000      same as X -c -c3-X
CT-CT-C -OS   1    0.000       180.000          2.000      same as X -c -c3-X
O -C -CT-HC   1    0.000       180.000          2.000      same as X -c -c3-X
O -C -OS-CT   1    2.700       180.000          2.000      same as X -c -os-X
OS-C -CT-HC   1    0.000       180.000          2.000      same as X -c -c3-X
OS-CT-CT-CT   1    0.156         0.000          3.000      same as X -c3-c3-X
OS-CT-CT-OS   1    0.156         0.000          3.000      same as X -c3-c3-X
OS-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-OS-P    1    0.383         0.000          3.000      same as X -c3-os-X
CT-OS-P -O2   1    1.050       180.000          2.000      same as X -os-p4-X
CT-OS-P -OS   1    1.050       180.000          2.000      same as X -os-p4-X
CT-CT-CT-CM   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CM-CM   1    0.000         0.000          2.000      same as X -c2-c3-X
CT-CT-CM-HA   1    0.000         0.000          2.000      same as X -c2-c3-X
CT-CM-CM-CT   1    6.650       180.000          2.000      same as X -c2-c2-X
CT-CM-CM-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
CM-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
CM-CM-CT-HC   1    0.000         0.000          2.000      same as X -c2-c3-X
P -OS-CT-H1   1    0.383         0.000          3.000      same as X -c3-os-X
```

```
OS-CT-CT-OH   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-CT-OH   1    0.156         0.000          3.000      same as X -c3-c3-X
CT-CT-OH-HO   1    0.167         0.000          3.000      same as X -c3-oh-X
OH-CT-CT-OH   1    0.156         0.000          3.000      same as X -c3-c3-X
OH-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
HA-CM-CT-HC   1    0.000         0.000          2.000      same as X -c2-c3-X
HA-CM-CM-HA   1    6.650       180.000          2.000      same as X -c2-c2-X
HC-CT-CT-HC   1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-CT-H1   1    0.156         0.000          3.000      same as X -c3-c3-X
H1-CT-OH-HO   1    0.167         0.000          3.000      same as X -c3-oh-X

IMPROPER
CT-O -C -OS         1.1          180.0           2.0       Using default value
CM-CT-CM-HA         1.1          180.0           2.0       Using default value

NONBON
  C           1.9080  0.0860          same as c
  CT          1.9080  0.1094          same as c3
  O           1.6612  0.2100          same as o
  OS          1.6837  0.1700          same as os
  CM          1.9080  0.0860          same as cc
  P           2.1000  0.2000          same as p4
  O2          1.6612  0.2100          same as o
  OH          1.7210  0.2104          same as oh
  HA          1.4870  0.0157          same as hc
  HC          1.4870  0.0157          same as hc
  H1          1.4870  0.0157          same as hc
  HO          0.0000  0.0000          same as ho
```

APPENDIX B

AMBER9 Input Files

```
Minimization MIN0:
 &cntrl
   imin=1,
   maxcyc=5000,
   ntmin=1, ncyc=500,
   dielc=1.0, cut=10.0, nsnb=5,
   scee=1.2,
   ntpr=25,
   ntt=2,
   ntb=1, ntp=0,
   iwrap=1,
   nmropt=0,
   ntr=1,
   restraint_wt= 15.0
   restraintmask=':241-261',
 &end
END

Minimization MIN:
 &cntrl
   imin=1,
   maxcyc=3000,
   ntmin=1, ncyc=2500,
   dielc=1.0, cut=10.0, nsnb=5,
   scee=1.2,
   ntpr=50,
   ntt=2,
   ntb=1 ntp=0,
   nmropt=0,
   iwrap=1,
 &end


Equilibration (initial) Warming Step:
 &cntrl
   nstlim=21000,
   ntpr=500, ntwx=000, ntwv=000, ntwe=000,
   cut=10.0, nsnb=5,
   dt=0.0020, temp0=5.0,
   ntt=2,
   ntb=1, ntp=0,
   ntc=2, ntf=2,
   iwrap=1,
   nmropt=0,
 &end
 &wt type='TEMP0', istep1=0001,  istep2=1000,  value1=0.05,
value2=0.05, &end
 &wt type='TEMP0', istep1=1001,  istep2=3000,  value1=0.05,
value2=50.,  &end
```

```
 &wt type='TEMP0', istep1=3001,  istep2=5000,  value1=50.,
value2=100., &end
 &wt type='TEMP0', istep1=5001,  istep2=7000,  value1=100.,
value2=150., &end
 &wt type='TEMP0', istep1=7001,  istep2=9000,  value1=150.,
value2=200., &end
 &wt type='TEMP0', istep1=9001,  istep2=11000, value1=200.,
value2=250., &end
 &wt type='TEMP0', istep1=11001, istep2=16000, value1=250.,
value2=300., &end
 &wt type='TEMP0', istep1=16001, istep2=21000, value1=300.,
value2=310., &end
 &wt type='END'  &end
END
```

```
Md_npt:
Molecular dynamics
 &cntrl
   nstlim=5000000,
   irest=0, ntx=1,
   ntpr=500, ntwx=500, ntwv=0, ntwe=0,
   cut=10.0, nsnb=25,
   dt=0.0015, temp0=310.00, tempi=310.0,
   ntc=2, ntf=2,
   ntt=3, gamma_ln=1.0,
   ntb=2, ntp=2,
   pres0=1.0,
   taup=3.0,
   iwrap=1,
 &end
```

```
MD_nvt:
Molecular dynamics
 &cntrl
   nstlim=5000000,
   irest=1, ntx=5,
   ntpr=100, ntwx=1000, ntwv=0, ntwe=0,
   cut=10.0, nsnb=25,
   dt=0.0015, temp0=310.00, tempi=310.0,
   ntc=2, ntf=2,
   ntt=2,
   ntb=1, ntp=0,
   iwrap=1,
 &end
```

APPENDIX C

PHI/PSI/OMEGA CALCULATION SCRIPTS

THIS SCRIPT IS USED TO CALL THE SECOND SCRIPT- THE COMMAND LINE INPUT IS AS FOLLOWS:

./Script Name /directory path of pdb files /directory path of file with atom numbers.txt /Output directory and file name

```
#! /bin/sh
#tells you what the current directory is the directory from which the
script is started
currentdir=`pwd`

#this is where the pdb files are located it is the second entry of the
cmd. line
pdbdirectory=$1

#this is where the file with the atom numbers is located the third
entry of the cmd. line
coordinatefile=$2

#this is where the output will be located, the fourth entry of the
cmd. line
outputfile=$3
cd $currentdir

#print to the output file the following statement
echo "snapshot,phi,psi, omega1, omega2" >$outputfile
for lines in  `ls $pdbdirectory/*.pdb.*`;

#for lines in `ls /cygdrive/d/Scripts/prist/*.pdb.*`;
do
#     echo $coordinatefile
#     echo $lines
#     echo $currentdir
#     echo $pdbdirectory
#     echo $lines
      echo $lines" "$coordinatefile" "$currentdir" /asst"
#This tells you the number of the pdb it is taking the data from
      number=`echo $lines | awk -F '/' '{print $NF}' | awk -F '.'
'{print $NF}'`
#This is what you have to enter on the cmd line to get things started
      /usr/bin/python /cygdrive/d/Scripts/MagicScript $lines
$coordinatefile $currentdir/asst$number
$coordinatefile $currentdir/asst$number
#The value that is calculated is temporarily placed here
      value=`cat $currentdir/asst$number`
      echo $number","$value >> $outputfile

#This removes the temporary file
      rm $currentdir/asst$number
done
```

MAGICSCRIPT

```
import string
import math
import os
import sys

#The following lines have been included for debugging
#infile = "D:/Scripts/pro/noW.pdb.1"
#infile1 = "D:/Scripts/pro.txt"
#outfile = "D:/Scripts/pro/Angles.csv"
#will allow the submission of file names as arguments to read in a
sequence of files
# system argument (take something from the cmd. line for this could be
file name or variable #
infile=sys.argv[1]
# two input files
infile1=sys.argv[2]

#this is the file that will be read in
input = open(infile, 'r')
#this is the actual reading of the data from the file
#the whole pdb file is read in
text = input.read()
#splits at the end of the line and creates a list
line = text.split('\n')

input1 = open(infile1, 'r')
text1 = input1.read()
#split by the line
line1 = text1.split('\n')

outfile=sys.argv[3]
output = open(outfile, 'w')

# A function to remove duplicate bonds, angles and dihedrals

def duplicates(array):
    duplicate1 = 1
    while duplicate1 < len(array):
        duplicate2 = 0
        duplicate3 = (len(array[duplicate1]) - 1)
        duplicate4 = []
        while duplicate3 >= 0:
            duplicate4.append(array[duplicate1][duplicate3])
            duplicate3 = duplicate3 -1
        while duplicate2 < duplicate1:
            if duplicate4 == array[duplicate2]:
                del array[duplicate1]
                duplicate1 = duplicate1 -1
                break
```

```
            else:
                duplicate2 = duplicate2 + 1
        duplicate1 = duplicate1 + 1
    return array

#A function to calculate the difference between two vectors

def vectordiff(vector1,vector2):
    vectorpoint1 = 0
    vector3 = []
    while vectorpoint1 < 3:
        vector3.append(((vector1[vectorpoint1]) -
(vector2[vectorpoint1])))
        vectorpoint1 = vectorpoint1 + 1
    return vector3

#A function to calculate the cross product of two vectors

def crossproduct(vector4,vector5):
    vector6 = []
    vector6.append((vector4[1] * vector5[2]) - (vector5[1] *
vector4[2]))
    vector6.append(-((vector4[0] * vector5[2]) - (vector5[0] *
vector4[2])))
    vector6.append((vector4[0] * vector5[1]) - (vector5[0] *
vector4[1]))
    return vector6

#A function to calculate the angle between two vectors, the dihedral
angle

def costheta(vector7,vector8):
    costheta = ((vector7[0] * vector8[0]) + (vector7[1] * vector8[1])
+ (vector7[2] * vector8[2]))/((math.sqrt( pow(vector7[0],2) +
pow(vector7[1],2) + pow(vector7[2],2))) * (math.sqrt(
pow(vector8[0],2) + pow(vector8[1],2) + pow(vector8[2],2))))
    return costheta

#the split function is split by whatever you put in the ()  ex. , or a
space or a tab(5spaces)
intarray=line1[0].split()
phiarray=[]
phiarray.append(int(intarray[1]))
phiarray.append(int(intarray[2]))
phiarray.append(int(intarray[3]))
phiarray.append(int(intarray[4]))

intarray=line1[2].split()
psiarray=[]
psiarray.append(int(intarray[1]))
psiarray.append(int(intarray[2]))
psiarray.append(int(intarray[3]))
```

```
psiarray.append(int(intarray[4]))
psiarray.append(int(intarray[4]))

#print phiarray
#print psiarray
#print omega1array

intarray=line1[4].split()
omega1array=[]
omega1array.append(int(intarray[1]))
omega1array.append(int(intarray[2]))
omega1array.append(int(intarray[3]))
omega1array.append(int(intarray[4]))

intarray=line1[6].split()
#creates array
omega2array=[]
#appends data to array
omega2array.append(int(intarray[1]))
omega2array.append(int(intarray[2]))
omega2array.append(int(intarray[3]))
omega2array.append(int(intarray[4]))

#print phiarray
#print psiarray
#print omega1array
##print omega2array

#calculate the Phi Angle
c=1
coord=[]
while c < len(line)-1:
#while c < 5:
        coord1=[]
        coord2=[]
        coord1=line[c].rsplit()
        coord2.append(eval(coord1[5]))
        coord2.append(eval(coord1[6]))
        coord2.append(eval(coord1[7]))
        coord2.append(int(coord1[1]))
        coord.append(coord2)
        c = c + 1
#a=[40,46,48,50]
#phiarray=
#a=[15, 17, 27, 29]

d=0
vectorreference1=[]
while d< len(coord):
    coord4=[]
    e=0
    while e < len(phiarray):
```

```
        if phiarray[e] == coord[d][3]:
            coord5=[]
            coord5.append(coord[d][0])
            coord5.append(coord[d][1])
            coord5.append(coord[d][2])
            vectorreference1.append(coord5)
        e=e+1
    d=d+1


vectorij = vectordiff(vectorreference1[0],vectorreference1[1])
vectorjk = vectordiff(vectorreference1[1],vectorreference1[2])
vectorkl = vectordiff(vectorreference1[2],vectorreference1[3])
vectorm = crossproduct(vectorij,vectorjk)
vectorn = crossproduct(vectorjk,vectorkl)
phi = math.acos(costheta(vectorn,vectorm))* 180 * 7/22
sign = vectorm[0]*vectorkl[0] + vectorm[1]*vectorkl[1] +
vectorm[2]*vectorkl[2]
if sign > 0:
    phi = -1 *phi
#print phi



#calculate the Psi Angle
#a=[40,46,48,50]
#psiarray b=[17, 27, 29, 31]
d=0
vectorreference1=[]
while d< len(coord):
    coord4=[]
    e=0
    while e < len(psiarray):
        if psiarray[e] == coord[d][3]:
            coord5=[]
            coord5.append(coord[d][0])
            coord5.append(coord[d][1])
            coord5.append(coord[d][2])
            vectorreference1.append(coord5)
        e=e+1
    d=d+1


vectorij = vectordiff(vectorreference1[0],vectorreference1[1])
vectorjk = vectordiff(vectorreference1[1],vectorreference1[2])
vectorkl = vectordiff(vectorreference1[2],vectorreference1[3])
vectorm = crossproduct(vectorij,vectorjk)
vectorn = crossproduct(vectorjk,vectorkl)
Psi = math.acos(costheta(vectorn,vectorm))* 180 * 7/22
sign = vectorm[0]*vectorkl[0] + vectorm[1]*vectorkl[1] +
vectorm[2]*vectorkl[2]
if sign > 0:
```

```
    Psi = -1 *Psi
#print Psi

#calculate the OMEGA 1 Angle
#a=[9, 15, 17, 27]
#omega1array
d=0
vectorreference1=[]
while d< len(coord):
    coord4=[]
    e=0
    while e < len(omega1array):
        if omega1array[e] == coord[d][3]:
            coord5=[]
            coord5.append(coord[d][0])
            coord5.append(coord[d][1])
            coord5.append(coord[d][2])
            vectorreference1.append(coord5)
        e=e+1
    d=d+1


vectorij = vectordiff(vectorreference1[0],vectorreference1[1])
vectorjk = vectordiff(vectorreference1[1],vectorreference1[2])
vectorkl = vectordiff(vectorreference1[2],vectorreference1[3])
vectorm = crossproduct(vectorij,vectorjk)
vectorn = crossproduct(vectorjk,vectorkl)
omega1 = math.acos(costheta(vectorn,vectorm))* 180 * 7/22
sign = vectorm[0]*vectorkl[0] + vectorm[1]*vectorkl[1] +
vectorm[2]*vectorkl[2]
if sign > 0:
    omega1 = -1 *omega1
#print omega1


#calculate the OMEGA 2 Angle
#a=[40,46,48,50]
a=[27, 29, 31,33]
#omega2array
d=0
vectorreference1=[]
while d< len(coord):
    coord4=[]
    e=0
    while e < len(omega2array):
        if omega2array[e] == coord[d][3]:
            coord5=[]
            coord5.append(coord[d][0])
            coord5.append(coord[d][1])
            coord5.append(coord[d][2])
            vectorreference1.append(coord5)
        e=e+1
```

```
    d=d+1


vectorij = vectordiff(vectorreference1[0],vectorreference1[1])
vectorjk = vectordiff(vectorreference1[1],vectorreference1[2])
vectorkl = vectordiff(vectorreference1[2],vectorreference1[3])
vectorm = crossproduct(vectorij,vectorjk)
vectorn = crossproduct(vectorjk,vectorkl)
omega2 = math.acos(costheta(vectorn,vectorm))* 180 * 7/22
sign = vectorm[0]*vectorkl[0] + vectorm[1]*vectorkl[1] +
vectorm[2]*vectorkl[2]
if sign > 0:
    omega2= -1 *omega2
#print omega2

output.write("%f,%f,%f,%f\n" %(phi,Psi,omega1,omega2))
input.close()
input1.close()
output.close()
```

APPENDIX D
PEPTIDE 23 INSERTED INTO MODEL MEMBRANES

INTRODUCTION

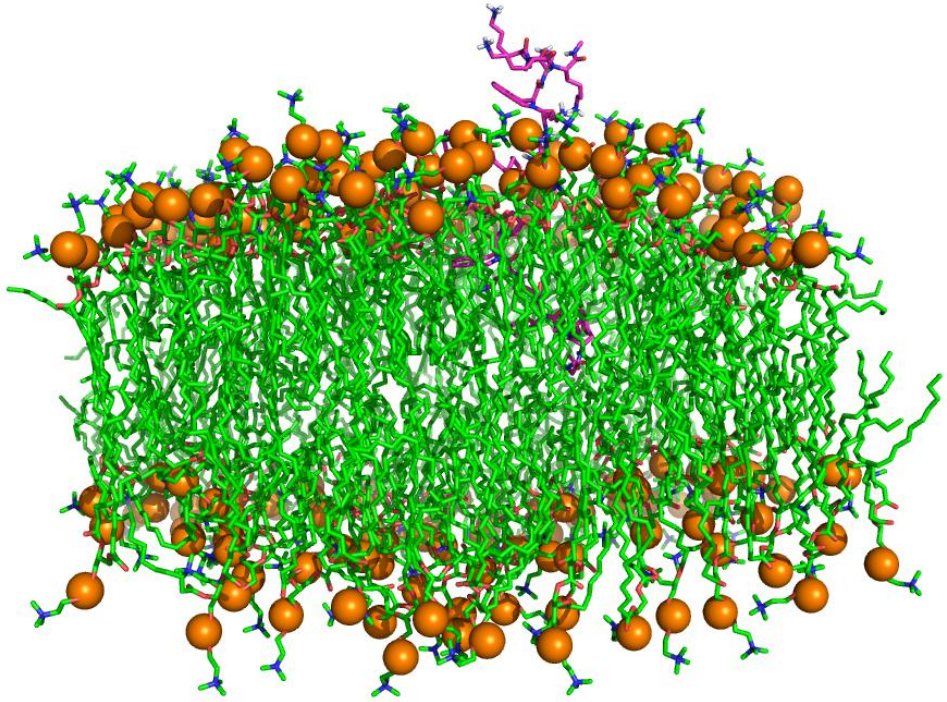
Peptide **23** was inserted into both the mixed and control bilayers. Two simulations with the mixed bilayer were conducted. The first had the N terminus inserted into the bilayer and the second had the C terminus inserted into the hydrophobic region of the bilayer. The following provides a brief perspective of the behavior of the peptide when placed in varying membrane environments.

CONTROL BILAYER PEPTIDE 23 INSERTED

Figures D.1 A, D.2 A & B, and D.3 show the starting orientation of peptide **23** when inserted into the control bilayer. The control bilayer was taken from the previous MD simulation of peptide **23** with the control bilayer (see Chapter 5) after 90 ns of the production simulation.

Figure D.1 B shows the position of peptide trajectory of the mixed bilayer simulation with peptide **23**. In the short 14 ns trajectory the system does not equilibrate. The interesting notation occurs in the conformation of the peptide relative to the bilayer. The polar C-terminus which had been placed at the bilayer surface moves away from the membrane surface allowing the more hydrophobic trajectory of the mixed bilayer simulation with peptide **23**. In the short 14 ns trajectory the system does not equilibrate. The interesting notation occurs in the conformation of the peptide relative to the bilayer. The polar C-terminus which had been placed at the bilayer surface moves away from the membrane surface allowing the more hydrophobic at the end of the simulation. Figure D.4 shows the trajectory of the control bilayer simulation with peptide **23** inserted into the membrane surface. In the short 14 ns trajectory the system did not equilibrate as evidenced by the average slope of the RMSD plot, Figure D.4. The interesting notation occurs in the conformation of the peptide relative to the bilayer surface. The polar C-terminus which had

A)



B)

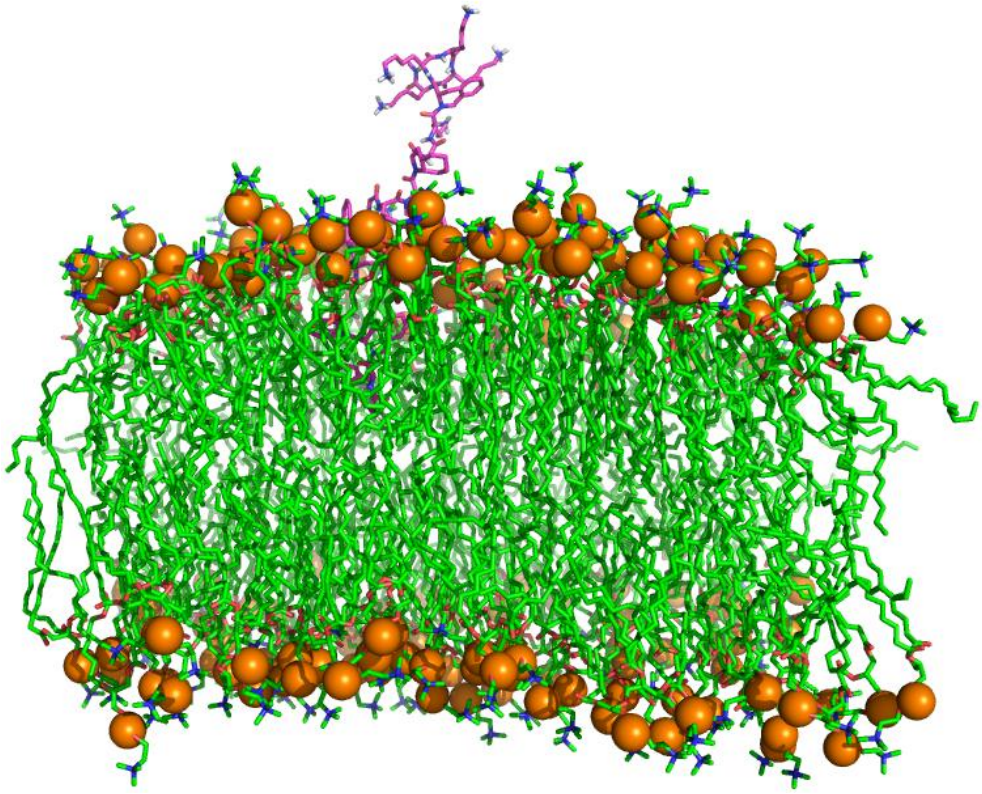


Figure D.1 A) Peptide **23** inserted into the control bilayer at the beginning of the simulation B) peptide 23 in the control bilayer at the end of the 14 ns trajectory

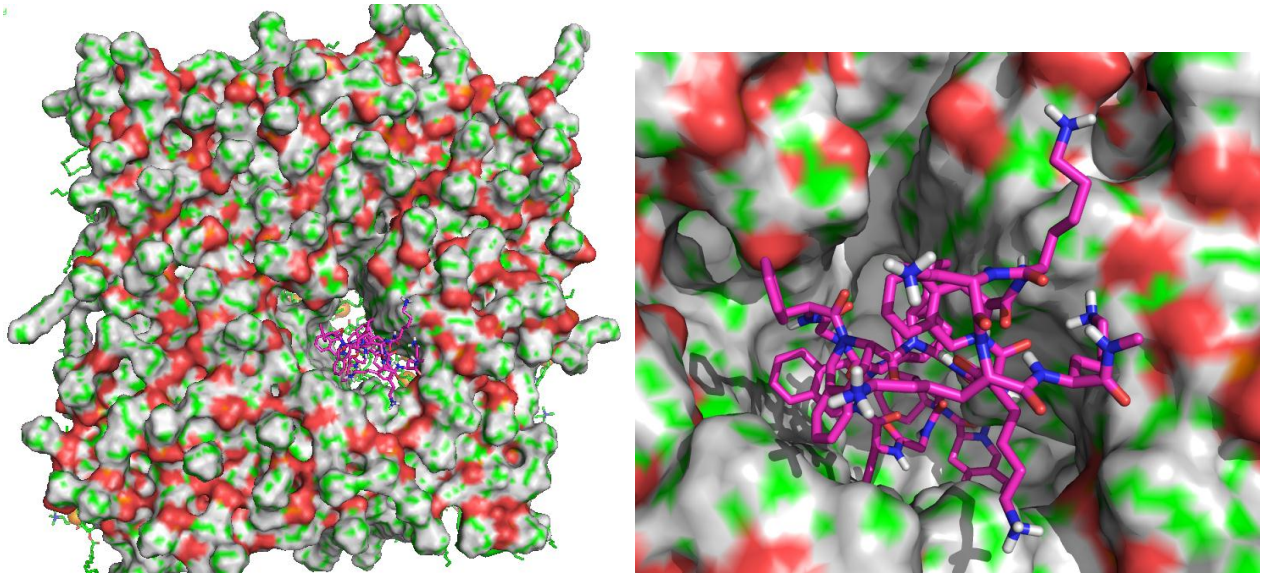


Figure D.2 Peptide **23** inserted into the control bilayer at the beginning of the simulation

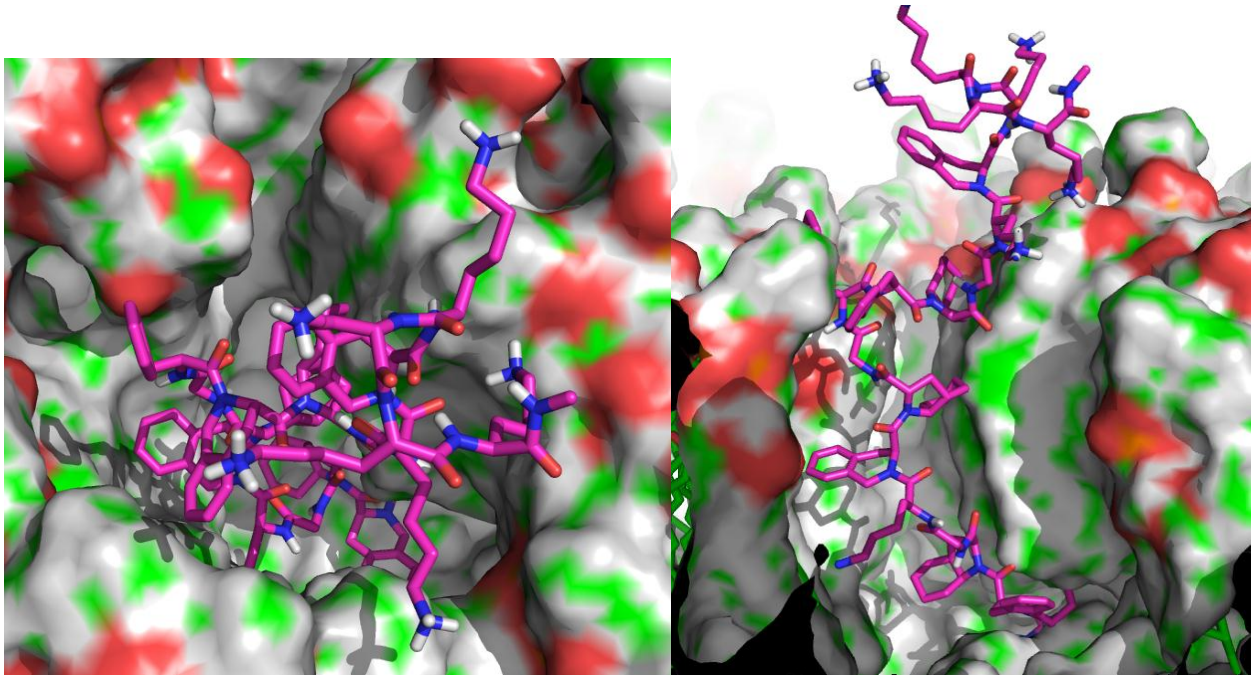


Figure D.3 Peptide **23** inserted into the control bilayer at the beginning of the simulation

been placed at the bilayer surface moves away from the membrane surface allowing the more hydrophobic rings near the N terminus to move out of the hydrophobic region of the bilayer D.5.

The short trajectory and lack of structure stabilization prevent conclusions from being drawn from the structures however it provides insight into how the peptide's residues behave when in a hydrophobic environment.

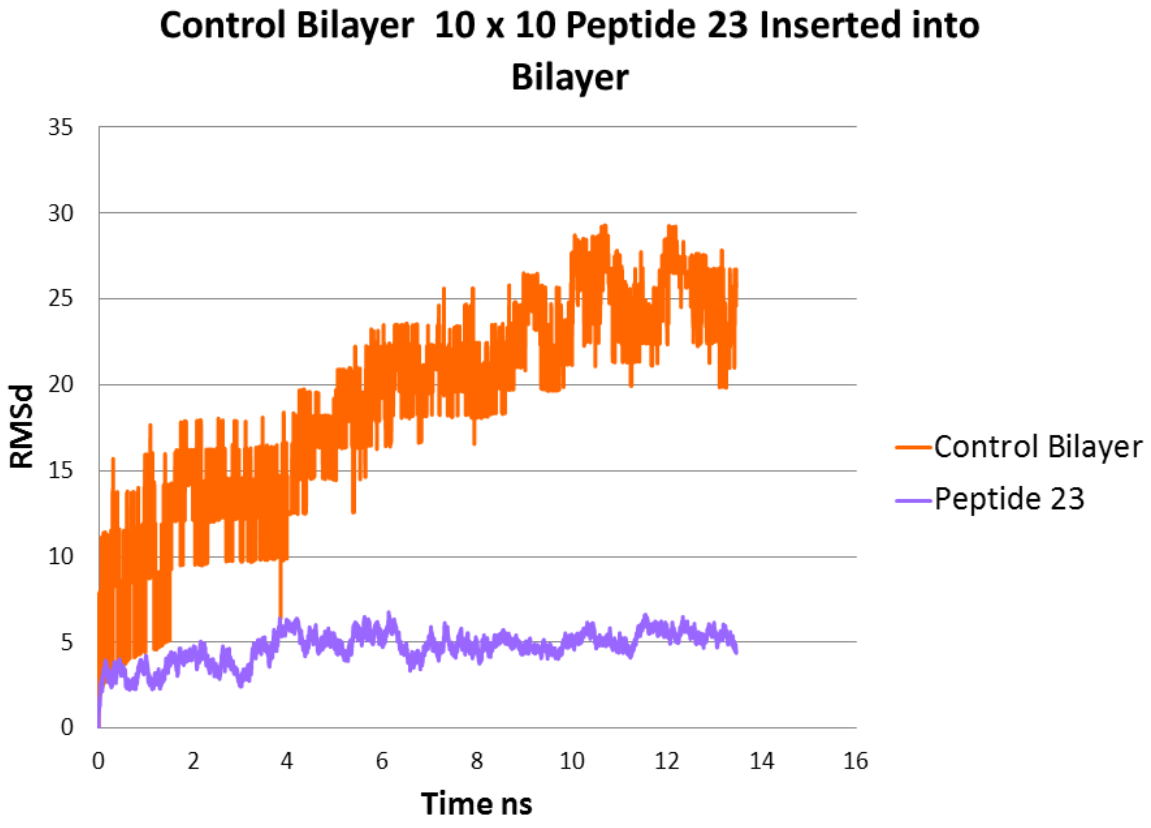


Figure D.4 Trajectory of peptide **23** inserted into the control bilayer

PEPTIDE 23 N TERMINUS INSERTED INTO MIXED BILAYER

Figure D.6 A and B show the starting orientation of peptide **23** when inserted into the mixed bilayer with the N terminus in the hydrophobic region of the bilayer. The mixed bilayer was taken from the previous MD simulation of peptide **23** with the mixed bilayer (see Chapter 5) after 105 ns of the production simulation. The simulation was continued for an additional ~ 30 ns without trajectory stabilization. The peptide was observed to move from the starting position relative to the bilayer surface.

Figure D.7 shows the trajectory of the mixed bilayer simulation with peptide **23** inserted into the hydrophobic region of the bilayer. In the 30 ns trajectory from the production simulation was sufficient for peptide **23** to equilibrate as evidenced by the average slope near 0. The short

Figure D.1 Atoms involved in the Omega 1 angle, outlined in blue. The Omega 2 angle is outlined in green. Phi angle, outlined in magenta and psi angle, outlined in yellow.

simulation time was not sufficient for the bilayer to equilibrate which is not surprising given the

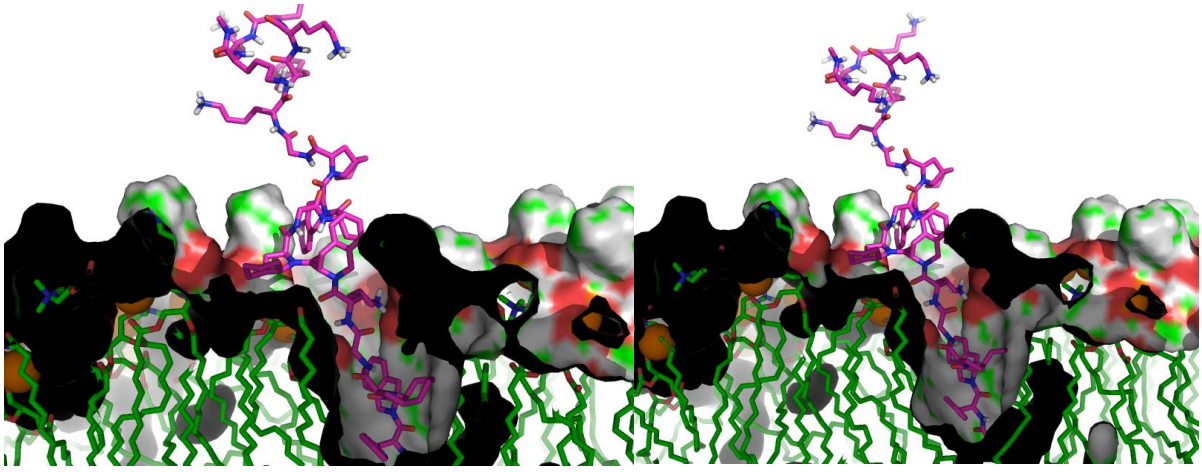


Figure D.5 Peptide **23** at the end of the ~14 ns MD trajectory

previous long trajectory attained without having the system equilibrate (Figure D.7).

The interesting observations are made with the peptide relative to both the starting position and the bilayer surface. At the beginning of the trajectory the N terminus was inserted into the bilayer with the C terminus, with the four polar lysine residues residing at the bilayer surface (Figure 5.6 A & B). The final orientation relative to the starting position shows the peptide is moving the peptide out of the hydrophobic region of the bilayer pushing all of the polar lysine residues to the bilayer surface (Figure D.8). It should be noted that the thinning of the membrane is a result of the previous simulation with peptide **23** at the surface (Chapter 5) and not a result of the peptide inserted into the bilayer. The N terminal hydrophobic rings of the Tic, Oic, and Phe residues are still partially buried at the within the hydrophobic region of the bilayer at the end of 30 ns (Figure D.8). The lysine residue closest to the N terminus have

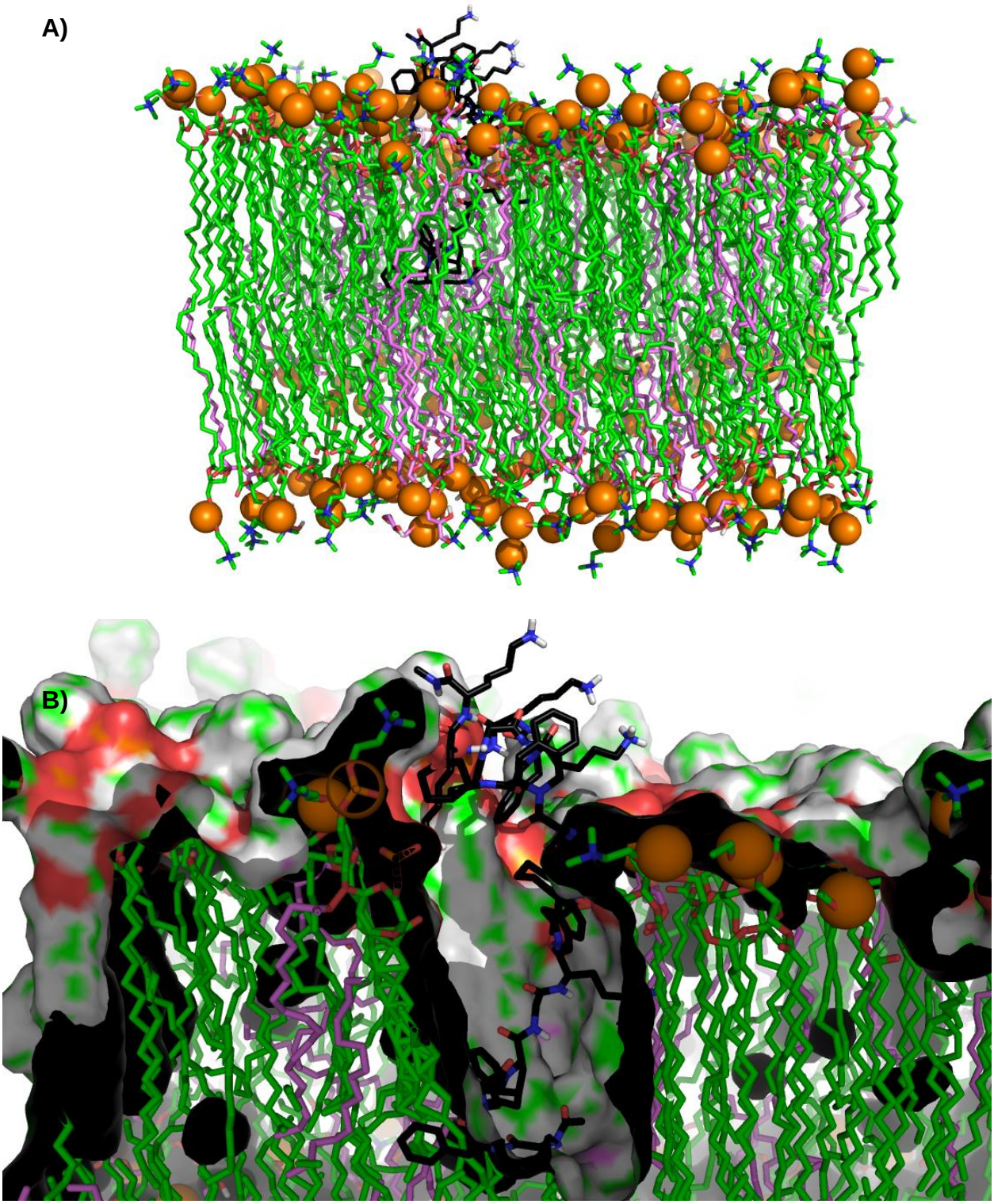


Figure D.6 Mixed Bilayer Peptide **23** inserted N terminus into the hydrophobic region of the bilayer.

moved to positions allowing interaction with the polar lipid head groups (Figure D.8). The position of the peptide should not be construed as the preferred orientation; the trajectory has not yet attained sufficient data to draw conclusions (Figure D.9).

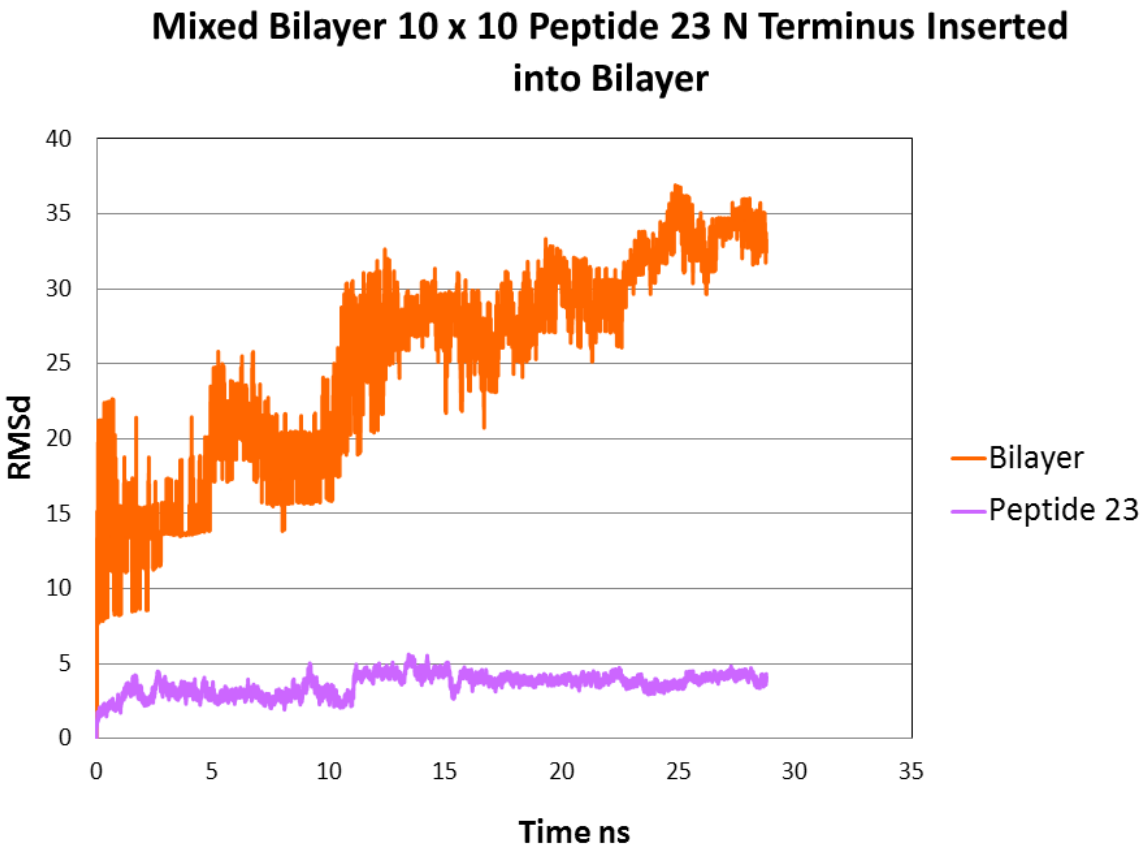


Figure D.7 Trajectory of peptide **23** inserted into the mixed bilayer

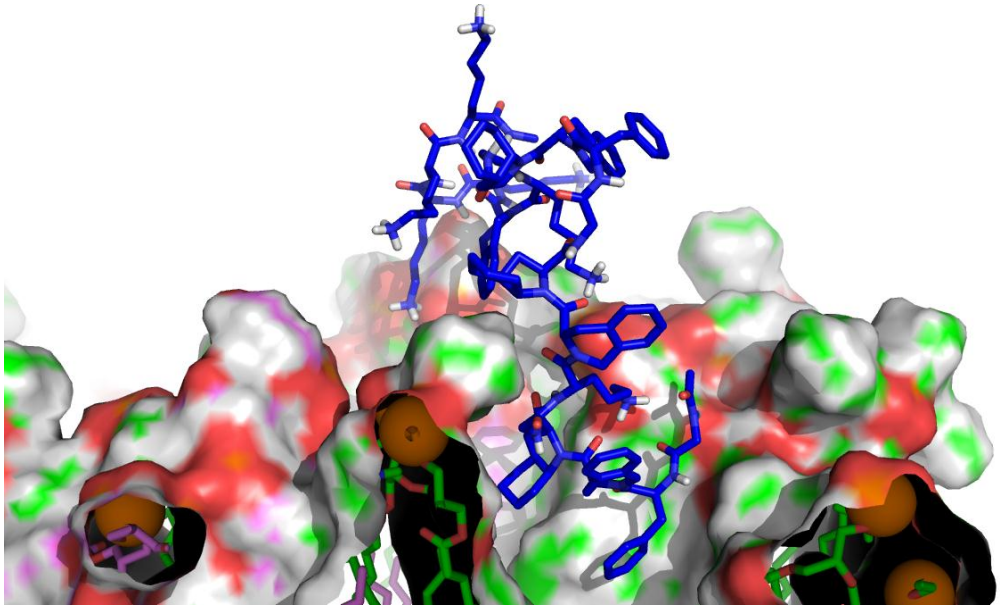


Figure D.8 Mixed bilayer with peptide **23** N terminus of the peptide inserted into the bilayer after 30 ns of MD simulation.

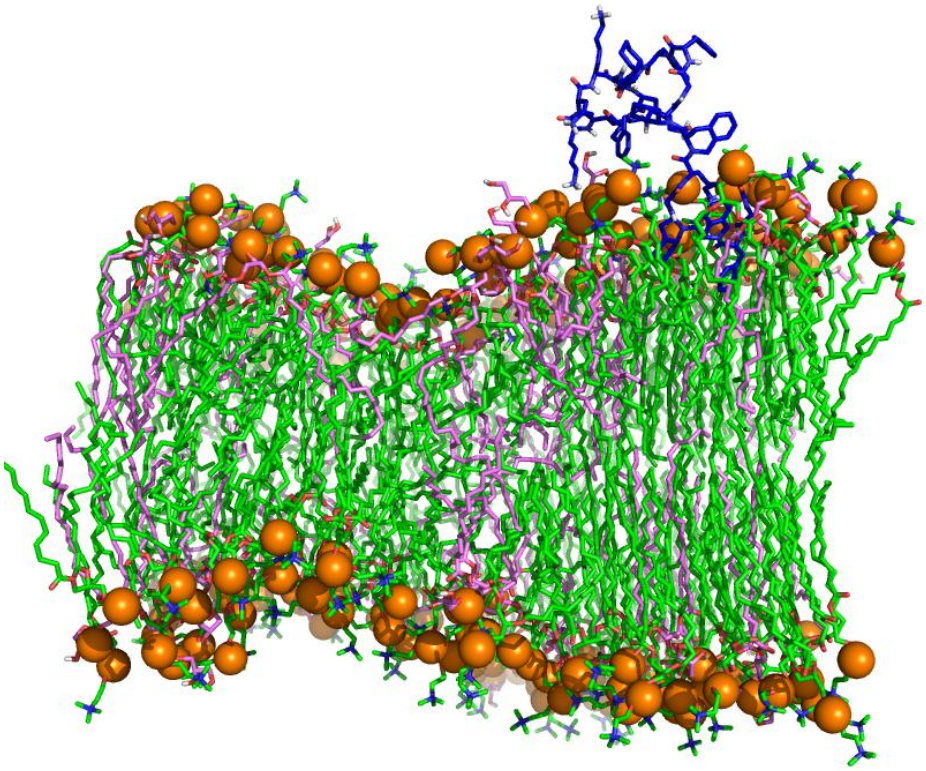


Figure D.9 Mixed bilayer with peptide **23** N terminus of the peptide inserted into the bilayer after 30 ns of MD simulation.

PEPTIDE 23 C TERMINUS INSERTED INTO MIXED BILAYER

Figure D.10 A and B show the starting orientation of peptide **23** inserted into the mixed bilayer with the C terminus inserted into the hydrophobic region of the bilayer. The mixed bilayer was taken from the previous MD simulation of peptide **23** with the mixed bilayer (see Chapter 5) after 105 ns of the production simulation. In Figure D.11 the starting position of the peptide relative to the two lipid types is depicted. The anionic POPG lipids are noted in lavender and the zwitterionic POPC lipids are noted in green. The peptide was placed near two anionic POPC lipids at the start of the simulation.

Figure D.12 shows the trajectory of the mixed bilayer simulation with peptide **23** inserted into the bilayer with the C terminus buried in the hydrophobic region of the bilayer. The bilayer appears to be stabilizing in the short 20 ns trajectory shown in Figure D.12. A longer simulation is necessary to draw any conclusions from the trajectory. The peptide inserted with the polar C terminus in the hydrophobic region of the bilayer underwent several conformational changes as evidenced by the slope of the RMSD plot in Figure D.12.

The interesting notation occurs in the conformation of the peptide relative to the bilayer (Figure D.13 and D.14). The less polar N terminus which had been placed at the bilayer surface moves away from the original position in the membrane surface allowing the polar lysine residues to move toward the polar head groups of the lipids forming hydrogen bonds, see figure D.14. Each of the lysine residues in peptide **23** has moved into reasonable proximity to form intermolecular hydrogen bonds with the surrounding lipids. There does not appear to be a preference for the charge associated with the lipid in the formation of these intermolecular interactions.

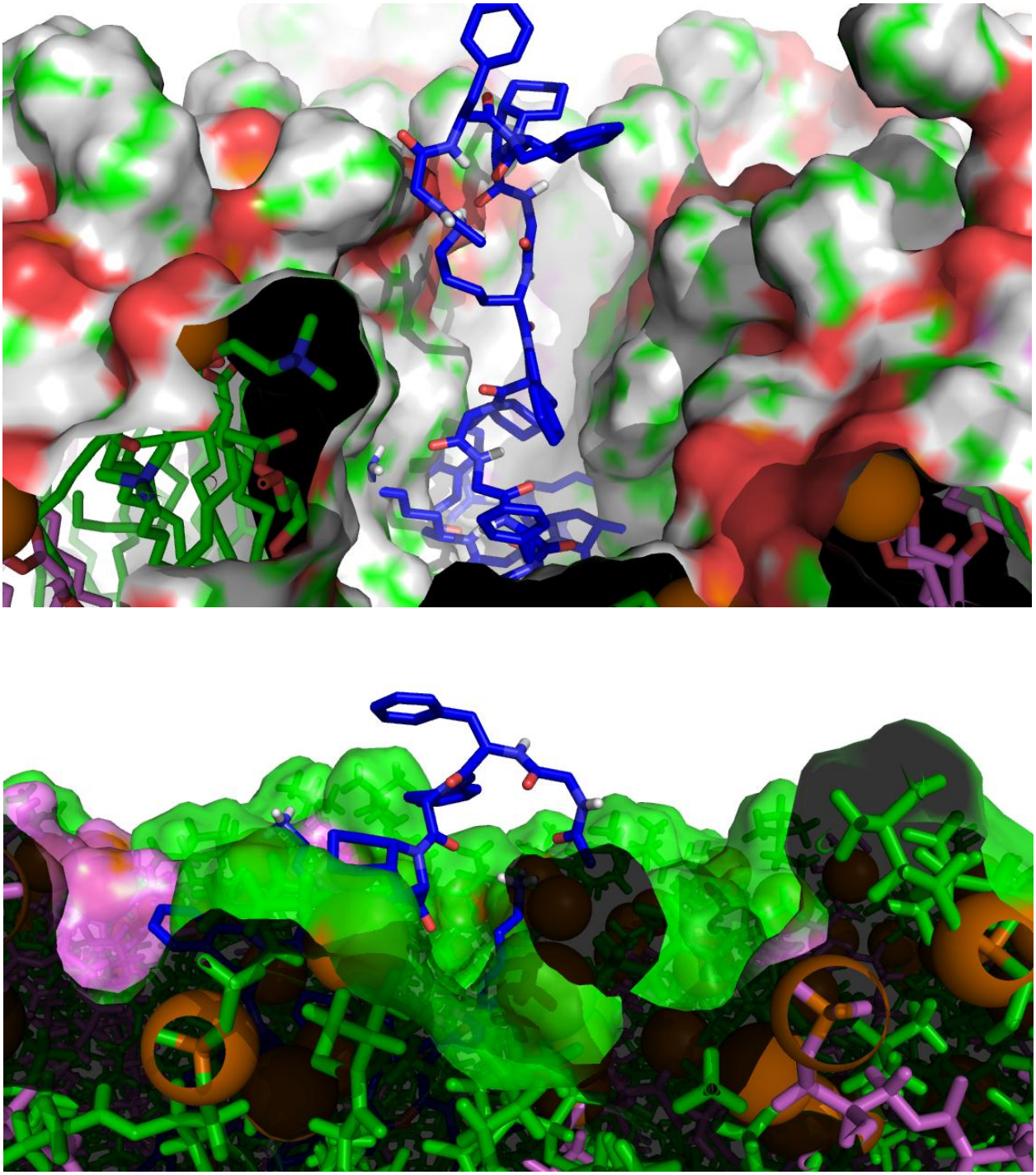


Figure D.10 Mixed bilayer with peptide **23** C terminus of the peptide inserted into the hydrophobic region of the bilayer.

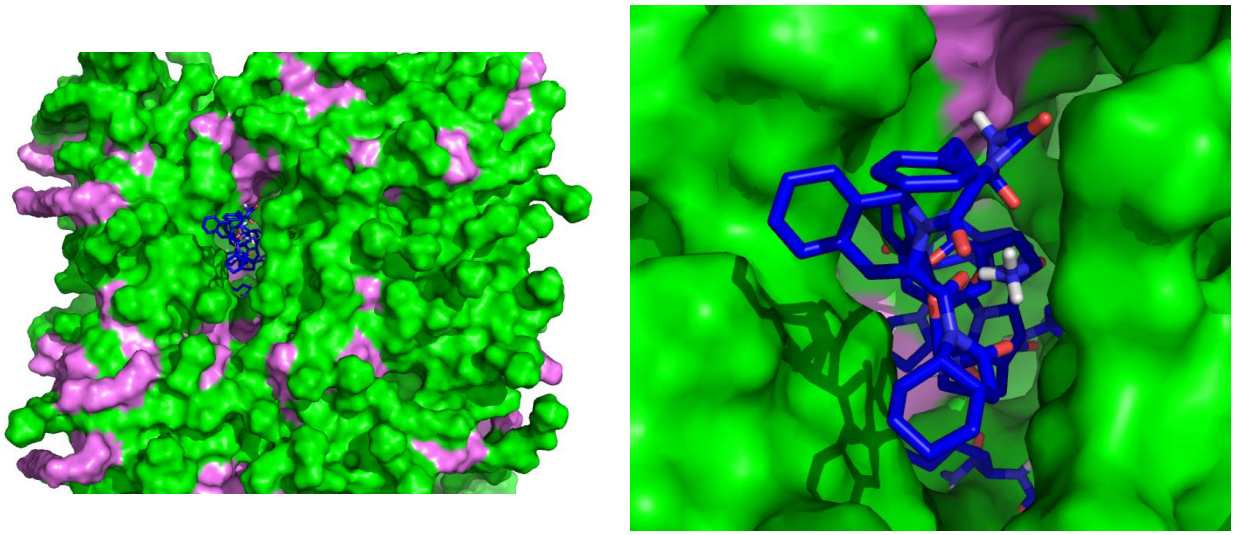


Figure D.11 Starting position of peptide **23** inserted with the C terminus in the hydrophobic region of the bilayer. Position of the peptide relative to the anionic POPG lipidsis noted in lavender and the zwitterionic POPG lipidsare noted in green.

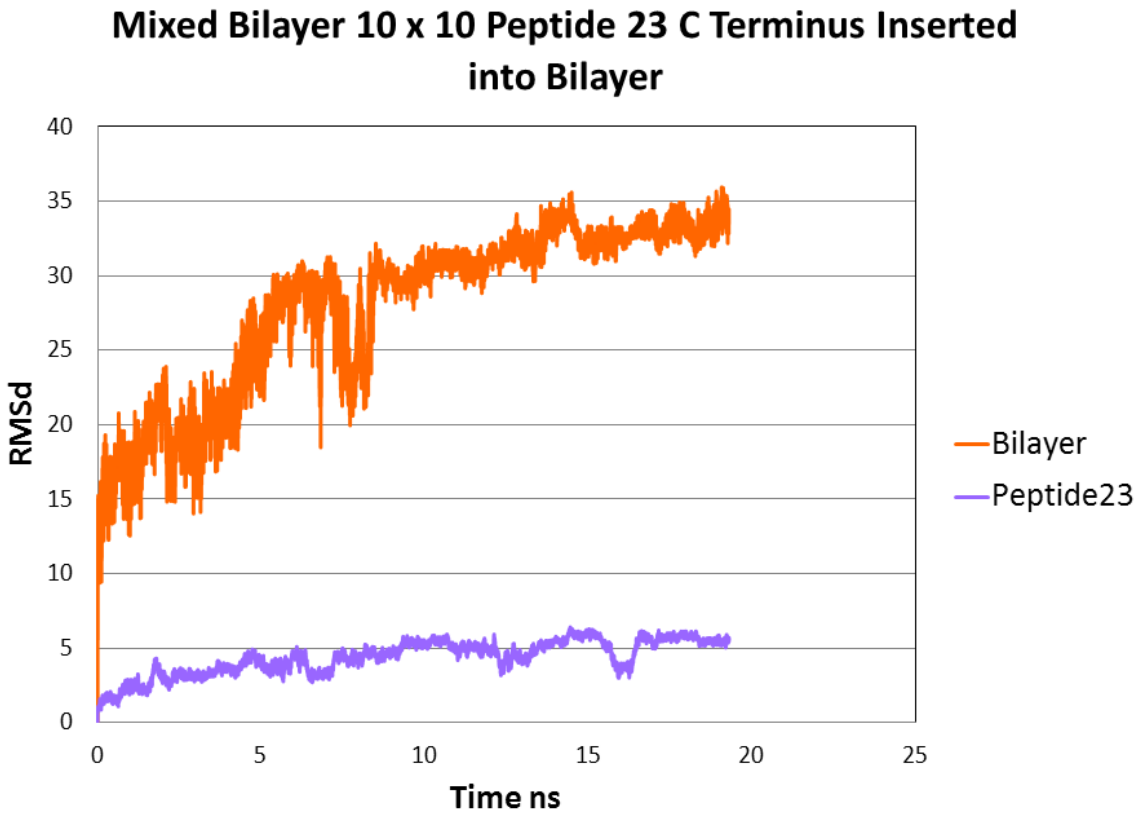


Figure D.12 Trajectory of peptide **23** with the C terminus inserted into the hydrophobic region of the mixed bilayer. Deviation is noted from the minimum energy structure at the beginning of the production run.

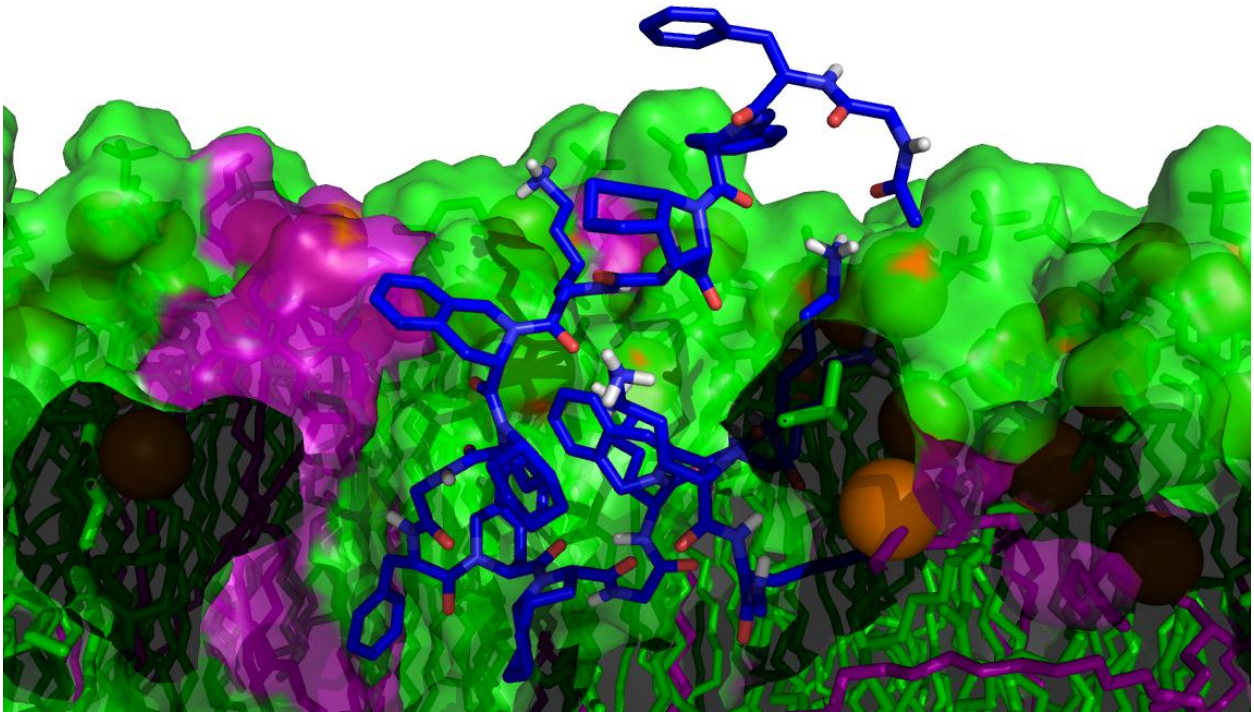


Figure D.13 Peptide **23** inserted with the C terminus in the hydrophobic region of the bilayer. Structure is from the end of a 20 ns production run.

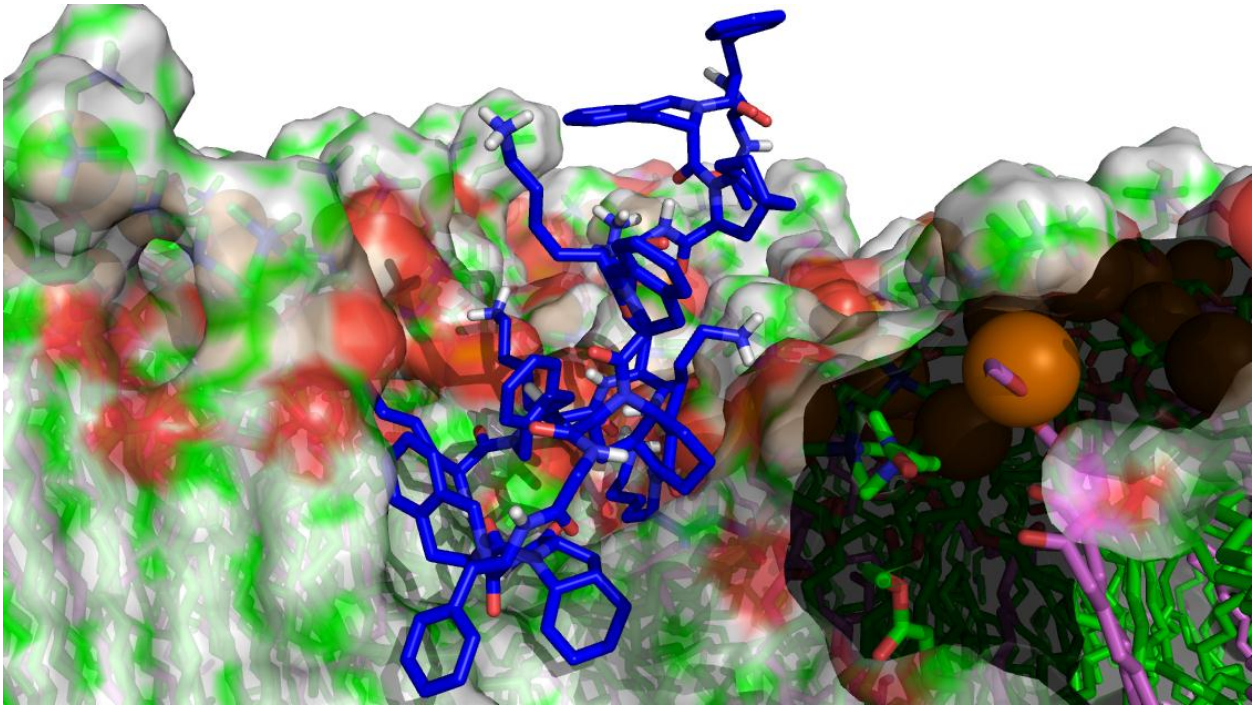


Figure D.14 Mixed bilayer with peptide **23** inserted with the C terminus inserted into the hydrophobic region of the bilayer. Structure was taken from the end of 20 ns of production simulation.