UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

Modeling Aspects of Proteins in Biologically Relevant Environments

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

Miranda Sheridan

Norman, Oklahoma

2024

Modeling Aspects of Proteins in Biologically Relevant Environments

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY

BY THE COMMITTEE CONSISTING OF

Dr. Ulrich Hansmann, Chair

Dr. Doerte Blume

Dr. Bayram Saparov

Dr. Yihan Shao

Contents

Abstract

Proteins and protein interactions are fundamental components of biological processes. Using molecular dynamics, the physical properties of proteins at various stages, and their association with functional or disease-causing complexes, can be probed under conditions that are difficult to investigate experimentally.

The first project presented in this work addresses the transition of the protein Serum amyloid A into a disease-causing conformation. Serum amyloid A is an acute phase protein and as such responds to trauma in the body. When incorrectly folded the protein may form amyloid fibrils which leads to severe diseases. Therefore, it is important to understand the transition from the normal to the pathological conformation. An advanced sampling technique, Replica Exchange with Tunneling (RET), was used for this purpose and it was established that there is a low energy barrier between the two forms of the protein and that the transition depends on the unfolding of the N-terminus of the protein.

The second project is motivated by the search for antibiotics. Because of its ability to form pores in membranes, nisin is pernicious to bacterial cells; however, the mechanism by which the pores are formed and are stabilized is unknown. Additionally, instability and insolubility at physiological conditions limit the broader use of the protein as an oral drug. To investigate the pore formation mechanism of nisin in bacterial cells and how the pores affect the integrity of the membranes, all atomistic molecular dynamics simulations were

performed. It was determined that residues 8-28 of nisin are important for the stability of the pore and that in the presence of lipid II, a component of the bilayer that is specific to bacteria cells, the pore is more stable. The integrity of the membrane is further reduced by local changes in thickness and viscosity.

In addition to having antimicrobial properties, nisin can inhibit cancer growth but does not appear to harm healthy human cells. As cancer cell membranes do not contain the bacteria-specific lipid II, there must be different mechanisms by which nisin induces cell death in cancer and bacteria cells. The third, and yet unfinished, project compares the interaction of nisin with different types of cell membranes – bacteria, cancer, and eukaryotic cells. Furthermore, simulating nisin in membrane environments allows for a more general inquiry into the not-well-understood physical dynamics of membranes. This is a promising area for future studies.

Chapter 1

Introduction

In this work, molecular dynamics is used to investigate the physical properties of proteins under conditions that are unfavorable to address experimentally, and to investigate the structures of proteins and their function in various environments. A concise rationale of why proteins should be studied, why the methods used in this work are appropriate for such studies, and a short prelude to each chapter of the dissertation are presented in the next sections of this introduction. As the topics discussed in this work, that is the proteins researched, differ strongly from each other, a more detailed introduction will be given to each topic in its respective chapter.

1.1 The Role of Proteins in Cell Biology

It has been well established that proteins mediate cellular activity; all processes that are associated with life ultimately are regulated and carried out by proteins on a cellular level. A detailed understanding of how biological proteins function is therefore needed for a more general understanding of the larger biological processes. In addition, biomedical research and treatment discoveries often depend on a detailed understanding of the working of

specific proteins. Pathologies, for example, are often a result of deviation from the normal functioning of these biological molecules. To treat diseases, then, a detailed knowledge of the normal functions and mechanisms of proteins is first required. The discovery of what we now know to be proteins, credited to Gerrit Jan Mulder in the 1830s, initiated the study of proteins. After heating various biological substances with NaOH, in the form of lye, he obtained a precipitate that contained approximately 16% nitrogen. From this he concluded that this precipitate was found in all living organisms; Berzelius later termed the substance 'proteine' – from Greek meaning primary. The idea of a common underlying substance among all life forms, e.g. proteins, thus emerged. In the mid-twentieth century, it became clear that each protein is comprised of a specific sequence of amino acids, called the primary structure, which ultimately determines the function of the protein. It has since remained a challenge and an objective to determine the function of a protein from its primary structure, that is to predict how a protein will behave from its amino acid sequence. This is known as the protein folding problem.

It is generally accepted that to carry out its normal function, a protein will assume a conformation that is the most thermodynamically stable state. After ribosomal production, the protein will fold into this so-called native structure. A theoretical description of how a protein folds into this state is provided by the free energy landscape of a protein. If every possible configuration that a protein can assume is plotted against the free energy associated with each configuration, the result would be a multidimensional graph referred to as the energy landscape. The native conformations are represented by a global minimum on this landscape.
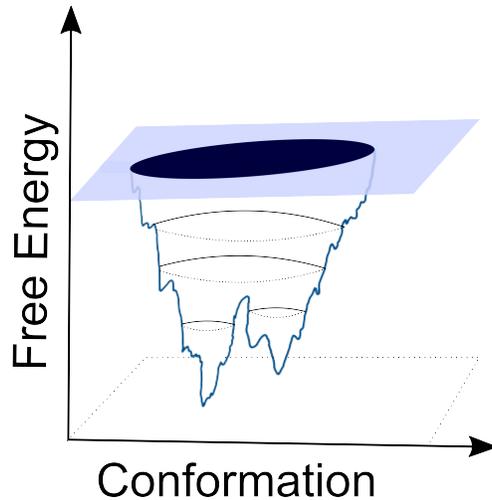
Figure 1.1: The possible configurations of a protein are projected onto a 3-dimensional Free Energy Landscap that dipicts how a protein folds.

Free energy landscapes depict the folding process; they show how the protein folds into

its native structure beginning from a disordered string of amino acids and ending in an

order stable state. This pathway of folding follows down the funnel-shaped landscape. [1]

The funnels describe the rate that a protein takes when folding into the native state. If the

funnel is smooth, folding is instant, and no alternative conformations are reached in the

folding process. If the energy landscape is rugged, then a protein will fluctuate confor-

mationally before arriving in its native state. Because this final structure determines the

protein function, it is important to both drug design and unveiling disease mechanisms to

understand the details of protein folding mechanisms.

The protein folding problem is a theoretical conceptualization of what information is needed

to understand how a protein works. The different configurations on the energy landscape

are dictated by physical characteristics: the solvent a protein is in, intermolecular contacts, electrostatics, hydrogen bonds, etc. Molecular dynamic (MD) simulations are a way to gain this type of structural details about a protein by sampling the protein's conformational landscape.

1.2   Tool-Set: Molecular Dynamics

Molecular dynamics (MD) relies on classical mechanics to predict the interactions between N particles in a system by integrating Newton's equations of motion

$$F = ma. \tag{1.1}$$

Because velocity is a derivative of position and acceleration is the derivative of velocity, the equation can be rewritten as

$$a = \frac{dv}{dt} = \frac{F(x)}{m} \tag{1.2}$$

where x is a vector that corresponds to the x, y, and z coordinates of the atoms, v is a vector that corresponds to the velocities for each atom, and m is the mass of each atom. A general from of the equation is

$$F_i = m_i \frac{d^2 r_i}{dt^2}, \tag{1.3}$$

where $\boldsymbol{F}_i$ is the force that is acting on a particle in the system, m_i is the mass of that par-

ticle, and r_i is the position of that particle. Numerical evaluation of the equation of motion leads to new positions and velocities, this is accomplished through integration algorithms as described below. [2]

The equations of motion can be approximated as a time series; a general form is below,

$$y(t + h) = y(t) + hy'(t) + \frac{1}{2}h^2 y''(t) + \frac{1}{6}h^3 y'''(t) + O(h^4) \qquad (1.4)$$

where h is the time step. If $-h$ is substituted in for h, the following equation is given

$$y(t - h) = y(t) - hy'(t) + \frac{1}{2}h^2 y''(t) - \frac{1}{6}h^3 h'''(t) + O(h^4). \qquad (1.5)$$

If 1.5 is subtracted from 1.4 the result is

$$y'(t) = \frac{y(t + h) - y(t - h)}{2h} + O(h^2) \qquad (1.6)$$

which is an approximation for the first derivative.

If 1.5 is added to 1.4 the result is

$$y''(t) = \frac{y(t + h) - 2y(t) + y(t - h)}{2h} + O(h^2) \qquad (1.7)$$

which is an approximation for the second derivative. Applying the position at time $t + h$ as

5

a function of the position at time t and $t - j$ gives

$$\boldsymbol{r}_i(t + h) = -\boldsymbol{r}_i(t - h) + 2\boldsymbol{r}_i(t) + \frac{h^2}{m_i}F_i(t) + O(h^4). \qquad (1.8)$$

Note, the accuracy of the equation depends on the size of the time step h, and the final values of the simulation are dependent initial values of $\boldsymbol{r}_i$; that is the initial configurations in MD simulation are important as the dynamics are an extension of them.

The velocity at time t can then be computed from

$$\boldsymbol{\nu}_i(t) \approx \frac{-\boldsymbol{r}_i(t + h) - \boldsymbol{r}_i(t - h)}{2h} \qquad (1.9)$$

For a more comprehensive review on time integrators see [3].

For an MD simulation, the initial positions and initial velocities of the particles are needed. The initial velocities, which are needed to determine the subsequent positions of particles are generated from a Maxwell-Boltzmann distribution at the specific temperature which is relevant to the system. Initial coordinates are usually taken from an x-ray, NMR, or Cryo-Em structure.

To simulate the interaction of particles, the bonds that make up a molecule, their respective angles, and interactions involving them or other nonbonded interactions also need to be accurately modeled. This information is included in various types of forcefields.

6

Equation 1.10 is a simplistic representation of the main two types of information that are needed

$$U(r) = U_{bonded} + U_{nonbonded}. \tag{1.10}$$

Associated with the first term are the energies that come from internal or intramolecular interactions, such as bonds stretching and the energy associated with the bond angles between the atoms. The unbonded terms describes interactions such as Van der Waal's and electrostatic interaction. An example is the CHARMM forcefield [4], a forcefield that is commonly used for protein simulations. It models the potential energy function in the following way:

$$U(\boldsymbol{r}) =$$

$$\sum_{bonds} K_b(b - b_0)^2 \tag{1.11}$$

$$+ \sum_{angles} K_\theta(\theta - \theta_0)^2 \tag{1.12}$$

$$+ \sum_{UB} K_{UB}(S - S_0)^2 \tag{1.13}$$

$$+ \sum_{dihedrals} K_\chi(1 + cos(n\chi - \delta)) \tag{1.14}$$

$$+ \sum_{impropers} K_{imp}(\varphi - \varphi_0)^2 \tag{1.15}$$

$$+ \sum_{\substack{nonbonded \\ i \neq j}} \left(\epsilon_{ij} \left[\left(\frac{R_{min_{ij}}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{min_{ij}}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon_l r_{ij}} \right) \tag{1.16}$$

The forcefield shown above is a potential energy function that represents the energy as-

sociated with bond movements (the term labeled as 1.11) , bond angles (1.12), cross-terms (1.13), dihedral angles (1.14), and angles that bonds cannot take (improper angles, 1.15), and the interaction between angles that are not bonded together (1.16). While the CHARMM additive forcefields is not the only way to model the interactions of a system, the above equation contains the terms that are common in forcefields. [4] Kukol et al. offer an analysis and comparison of forcefields for modeling proteins; see reference [5].

Because the MD simulations depend on the interaction between the particles present in a system, the resulting trajectories of a simulation yield detailed information about those particles. Intermolecular contacts, electrostatics, hydrogen bonds, etc. in a system can then be calculated and analyzed. These quantities are driving factors in protein folding and therefore should be studied in a manner that provides the most resolution possible.

As a type of physics-based modeling that allows the forces and dynamics of a protein to be understood, MD is particularly valuable in studying proteins that transition into or between conformations. Regarding these proteins, it is often not enough that the final structure is predicted. Gaining specific biochemical properties that drive the transition is a point of interest.

MD is also able to probe protein-protein and protein-lipid interactions. These interactions are involved in many vital biochemical processes; cell signaling pathways, structural organization in membranes, enzymatic activity, and the immune response system. Structures that are conformationally stable can interact with other structures; this can lead to

8

structural and functional changes in one or both molecules involved. MD is able to follow the dynamic changes and probe the atomic interactions that allow the changes to occur.

1.3   Overview of the Systems Studied

Serum amyloid A (SAA) is a protein that is associated with having both functional and pathological roles in physiology; it is known to mediate the innate immune system, regulates cholesterol, and cause amyloidosis which can lead to organ failure throughout the body. It is when the functional monomer protein misfolds that SAA becomes pathological; the misfolded monomer forms fibrils that aggregate on organs resulting in organ failure. The process of conversion from the functional form to a misfolded monomer is unknown. Experimental studies on monomer and aggregation processes are inhibited by the transient nature of the transformation. In this study advanced sampling techniques were used to probe the conversion of the protein. This advanced sampling technique makes use of multiple models simulated simultaneously and an exchange system, which allows configurations that are not possible to be skipped, to sample the energy landscape of a protein. The results of the study are presented, and expanded upon the work we previously published: *F. Yasar, M.S. Sheridan, and U.H.E. Hansmann,* Interconversion Between Serum Amyloid A Native and Fibril Conformations. *ACS Omega, 7 (2022), 12186* [6], in chapter 2.

Initially motivated by the search for novel antibiotics, the second pillar of my doctoral

research is a computational study into the workings of the antimicrobial peptide Nisin. Nisin is a small peptide that is naturally produced by certain bacteria. The function of the protein is to kill other non-host bacteria; this is accomplished when nisin forms pores in the membrane bilayer of bacteria cells. To form pores, it is thought that the N-terminus of nisin binds to a lipid II molecule, which is part of the bacteria membrane. This binding acts as an anchor and allows the C-terminus of nisin to be translocated into the membrane bilayer. Details about the structure of the pores in the membrane remain unknown.

Computational methods, specifically molecular dynamics (MD), are a viable approach to address this problem as MD affords specific atomic details. However, as the time scales of membrane formation are too long for brute force MD, I have taken a more restrictive approach. Experimentally there are 'clues' about the orientation of nisin proteins in a membrane and how the proteins are positioned concerning each other. These clues depict nisin proteins forming a ring structure (i.e. a pore) in the bilayer and provide an estimation of where the pore should reside in the membrane. In contrast, the only NMR structure of nisin binding to lipid II (resolved in an organic solvent) does not comport with other experimental data regarding the configuration of nisin in membranes. To bridge this gap, I started with the NMR structure of nisin and modified it to fit the collected experimental data, thereby creating a model that reflects the experimental knowledge. How the pore affects the integrity of the membranes is studied with MD and presented in chapter 3, which expands on our recently published article: *M.S. Sheridan, P. Pandey, and U.H.E. Hansmann,* In Bacterial Membranes Lipid II Changes the Stability of Pores Formed by the Antimicrobial Peptide Nisin, *J. Phys. Chem. B, 128(2024)* [7].

In addition, to being a natural antibacterial agent, nisin has also been shown to have inhibitory properties against cancer. In bacteria membranes the nisin molecules bind to lipid II, a component of the membrane, which gives the protein the stability needed for its C-terminus to move into the membrane. After this has been accomplished, a pore is formed with other nisin molecules. However, cancer membranes lack a lipid II molecule to bind to. For this reason, we are currently studying how nisin permeates into different membranes and why nisin is effective against bacterial and cancer cells but seems to be benign to healthy human cells.

Models that represent bacteria, cancer, and a normal eukaryotic cell are modeled with nisin to determine the mechanism that nisin uses to permeate into membranes. In addition, simulating nisin in membrane environments allows for a more general inquiry into the not-well-understood dynamics of membranes. This topic is addressed in chapter 4 and describes unfinished, and therefore yet unpublished, research.

In the final chapter 5, the research described in this dissertation is summarized. The conclusions are discussed and topics of future research that could grow out of the presented work are suggested.

Chapter 2

Conformational Transitions of Serum Amyloid A

Proteins can undergo phase transitions that can change the activity of the protein and induce other physical phenomenon such as crystallization, aggregation, or fibril formation. [8] While it is often assumed that a protein's amino acid sequence will determine the shape, and therefore the function that the protein will assume, proteins that are considered intrinsically disordered (assume no definite structure) do not find a single thermodynamically stable state but will assume multiple conformations and can transition between the conformations. It is worth noting that these intrinsically disordered proteins (IDPs), or regions of a protein (IDRs), are difficult to characterize experimentally – their lack of structure causes an electron density deficit, making them difficult to characterize by X-ray diffraction, and discourages crystals to be formed, making them difficult to character by cryo-EM. IDPs are found in almost all living organisms and have now gained recognition as a class of proteins that are known for their transitions on fast time scales. [9]

As a fundamental idea, protein transitions challenge what has been previously accepted by the scientific community on how proteins fold – that a sequence of amino acids leads

to a single functional structure. Proteins that transition prove that a sequence of amino acids can have multiple stable and functional structures. [10]

In addition to the intellectual challenge that transitional proteins present, proteins that misfold are also associated with common debilitating diseases that have no cure, e.g. Parkinson's, Alzheimer's disease, type 2 diabetes, Spongiform encephalopathies, and Huntington's disease. From a pathological standpoint, the transition of the proteins associated with such disease becomes of interest to understand the disease and the cure.

Implied in a protein transition is a dynamic process that results in the rearrangement of the interatomic forces. This chapter will focus on the conformational transition of the Serum Amyloid A monomer. Sections 2.2 - 2.5 are taken from our previously published work: *F. Yasar, M.S. Sheridan, and U.H.E. Hansmann,* Interconversion Between Serum Amyloid A Native and Fibril Conformations. *ACS Omega, 7 (2022), 12186* [6]

2.1 Serum Amyloid A

2.1.1 Introduction

Serum Amyloid A (SAA) proteins are a family of lipoproteins proteins that are commonly associated with inflammation and disease. SAA proteins are genetically encoded on chromosome 11 by 4 genes in mammal species and maintain homology across species in-

dicating a specific and conserved function in physiology. There are 4 subtypes of SAA proteins that are encoded by the four genes. SAA subtypes 1 and 2 (SAA1 and SAA2) are the proteins that are associated with the immune response, are produced overwhelmingly in the liver, and share essentially an identical sequence. Allelic variations and two polymorphisms account for the protein belonging to a class. [11] SAA3, in humans, has an early stop codon which causes the last 10 amino acids in the sequence to be discarded. In addition, SAA3 does not share the sequence identity that SAA1 and SAA2 share and is thought to be produced in tissues other than the liver. The functional purpose, and difference in function between SAA1/SAA2, of SAA3 is not well understood. SAA4 is not an acute phase protein but is regularly produced in the body. [12] For the purposes of this work, SAA will referer to subtypes 1 and 2.

Although SAA proteins were discovered along with the amyloid fibrils they can cause, and named and associated with these fibrils, they also have a positive role in physiology – they are an acute phase response protein. Any time the body experiences trauma caused by bacterial or viral infections, autoimmune disorders, and cancers [13], the concentration of SAA increases in the body as a response. Normal concentrations of SAA are under 3mg/L [11]; however, during the inflammatory response these levels can increase 1,000-fold compared to the normal levels of SAA. It is because of this clear change in protein concentration as a result of trauma that SAA is considered to be an acute phase protein. However, beyond this, the function and mechanism of SAA remains highly ambiguous. The purpose of the increased levels in the blood remains unknown, although it is thought that SAA helps regulate other pathways. By binding to receptors, located on different

cells in the body, SAA promotes chemotaxis. [12] Specifically SAA is thought to promote the inflammasome cascade and recruit various interleukins. These processes are then associated with various immune functions (e.g. produce helper T cells) which are responsible for alleviating the damage done by the trauma. SAA can promote interleukin-6 (IL-6), IL-1β, TNFα, and cause leukocyte migration, tissue infiltration, tissue infiltration, angiogenesis, cartilage destruction which is accomplished through multiple receptor activations. [11] Each of these pathways is highly complicated and still not well understood.

In addition to SAA being a part of the acute phase response, it is well established that SAA binds to high-density lipoprotein (HDL); however, it is not clear if this binding occurs because SAA functions as an HDL transporter, or because, as an insoluble protein, it needs to bind to another molecule, so that it can be transported in the serum – or some combination. It is worth noting that while circulating in the blood, 95% of SAA is an HDL lipoprotein (i.e. bound to HDL), however, during inflammation, SAA becomes an apolipoprotein to HDL (i.e. unbound to HDL). [14] In this unbound state, the SAA monomer is an intrinsically disordered protein. What causes SAA to dissociate from HDL is also unknown. It is clear that SAA plays a vital role in the immune defense response and can form fibrils that cause diseases. The role of SAA in the immune system seems not to be limited to one function. Because of the multipurpose function of SAA seen in the immune system, Wang et al. point out that the multipurpose SAA may be linked to multiple conformations. [15]

2.1.2 Historical Background

While it is becoming more apparent that SAA has a functional role in the body, historically SAA was associated with fibrils and the diseases that these fibrils cause. The first recorded reports of amyloidosis are from the mid-seventeenth century and in medical journals of the time such as Historiarum Anatomicarum Rariorum and Sepulchretum sive Anatomia Practica. [16] In the medical community cases continued to be noted which involved organs of abnormal size that were characterized by firm 'waxy' (Early descriptions of fibrils depict a gray fat-like substance covering the liver) like material. Historically, most amyloid cases were associated with the liver and a majority of the early cases were seen in patients who also had other diseases such as tuberculosis or bone tumors. In the mid-nineteenth century, experiments were being done to determine what amyloids were made up by staining fibrous tissue with iodine, methyl violet, and sulfuric acid (in 1922 Bennhold showed that Congo red bound readily to all amyloid types and is what is commonly used even today). It is interesting to note that the term amyloid was used when Rudolph Virchow was analyzing tissue from the nervous system by staining it with iodine (1838 AD [16]). From his experiments, he deduced that the tissues he was looking at were similar to the 'amyloid' deposits the botanists found on plants. When it was later determined that the waxy substances found on human organs and the amyloids found in plants were quite different, the term 'amyloid' had been in circulation to the extent where the name could not be changed. [16]

2.1.3   Relating Serum Amyloid A to Diseases

By definition, an amyloid is an aggregate of misfolded proteins that form a rigid fibril that has a distinct cross-beta motif. Amyloidosis is the disease that arises from these fibrillar depositions in various tissues of the body. To date, there are approximately 30 proteins that have amyloidogenic properties which enable them to form amyloids, and those 30 proteins account for around 50 amyloid disorders. [17, 10] It is estimated that these proteins will affect around 150 million people worldwide by 2050. [18] There is no cure for diseases associated with amyloid fibrils and treatment options are also limited or have unwanted side effects. For example, current treatment options for secondary amyloidosis suppress inflammation or use immune suppression drugs to keep SAA levels at a minimum. [17]

Amyloidosis can affect the body systemically or locally and can be related to genetic factors or can be acquired. The organs most affected by amyloids are the kidney, heart, liver, spleen, nerve, and blood vessels. Because amyloids are comprised of various types of proteins, the clinical manifestations of amyloidosis differ – this is how different diseases arise from the general concept of amyloid. The four main subtypes of systemic amyloidosis are primary Al amyloidosis, secondary amyloid A amyloidosis, Familial amyloidosis, and β2-microglobulin amyloidosis. Primary and Secondary amyloidosis, associated with the light chain protein and SAA respectively, are the most common amyloidosis worldwide. Allelic variants of the proteins that form the amyloids can change the severity and

toxicity of the disease. This is another factor that convolutes the understanding of amyloids and their formation.

Local amyloidosis manifests in clinical syndromes such as cardiomyopathy, autonomic dysfunction, neuropathy, renal failure, and corneal abnormalities. Generally, local amyloidosis is specific to one organ which has unique symptoms, this type of amyloidosis is distinct and unique. Proteins that are responsible for local amyloidosis, classified as amyloids, are distinct and different. Why these amyloids localize to one organ is unknown.

The diagnosis of amyloidosis is based on clinical symptoms and histological testing; however, diagnosing amyloidosis early and correctly can be challenging. Early clinical symptoms can bear resemblance to other disorders, in some cases, histological test will yield a negative result, but require deep tissue biopsy, and identification of the specific that constitute the fibrils becomes difficult after deposition has occurred. [19]

As mentioned before there are no treatment options for late cases of amyloidosis. Understanding the proteins involved in amyloids and how amyloids are formed is all the more important.

2.1.4   Formation of Amyloid Fibrils

The transition of the functional SAA monomer into the misfolded monomer in and of itself is not what causes diseases. The misfolded monomers aggregate and eventually form fibrils, and these fibrils are what build up in organ tissue. Formation of fibrils is important to detail for the broader picture of how the protein causes harm.

It is generally accepted that a protein's amino acids sequence determines its 3-dimensional shape, or its native structure, from which follows its function. However, there are proteins that do not fold into these functional stable states until they bind to a specific molecule. Usually, these proteins are associated with having a lack of structure and are termed intrinsically disordered proteins, and many are associated with misfolding diseases. It should be noted that proteins that eventually fold into an ordered globular state can also assume intermediate states and can also unfold and fold which leaves them in a partially unfolded state with the theoretical potential to also form aggregates. Partial unfolding or misfolding of proteins often leads to a final structure that is different from its normal form. Thus, the function of the protein will also be altered. The altered forms of proteins can cause fibril formation, ultimately caused by the abnormal proteins forming an ordered structure.

Amyloid fibrils are comprised of an assembly of one type of protein which self-assemble into a beta-sheet structure, and form a larger, more rigid, structure known as a fibril. For-

mally the fundamental structure that begins the fibril is a series of beta sheets. A long chain of these beta sheets forms a filament by orienting two hydrophobic cores from independent beta sheets together which are then stacked on top of other similar structures (approximately 1 μm in length). The fibrils gain their stability from the side-chain interaction and hydrogen bonds of the beta sheets that they are comprised from. The cross beta-sheet motif is a defining characteristic of amyloids; it should be noted that the 2-dimensional structure of the beta sheet folds can take on different structures, even from the same protein. This is what causes the amyloid polymorphisms. [20] A collection of filaments 'twist' together to make up a fibril bundle (approximately 10 μmm). The aggregation of these fibril bundles form the plaques that are known to accumulate on various tissues.

Although fibril formation is of interest and many studies seek to explain the phenomenon, the process is not well understood. There are three main fibril formation hypotheses:

Classical Nucleation Theory: This theory, when applied to proteins, poses that there is a specific concentration at which a protein will persist as a monomer in a solution, i.e. the saturated concentration. Once this saturated concentration is exceeded a phase change takes place and the monomers that are in excess of this saturated concentration form a cluster. [20] In this theory, there is one free energy barrier that a cluster has to overcome before a final phase is reached.

Non-classical Nucleation Theory: In this theory, the final phase is reached by passing

through intermediate phases, i.e. stages. [21] The theory is closely related to Oswald's rules of stages which states that a transition into a stable state occurs through intermediate states whose free energy is similar to the initial state; small changes in free energy that lead to more stable states is what ultimately achieves a globally stable state. [22]

Seeding Theory: The theory expresses the idea that an existing amyloid can cause monomers to assume fibril formation at a faster rate than classical nucleation. If an existing amyloid causes monomers of the same protein type to aggregate, is homologous and simply known as seeding. If an existing amyloid causes monomers of a different protein to aggregate it is known as cross-seeding, which is implicated in many neurodegenerative diseases. [22]

SAA fibrils are associated with constituent protein fragments that are approximately 76 amino acids in length (note not the full functional monomer length), which form ordered beta sheets that are predominately antiparallel and have around 10 angstroms of distance between the stacked sheets, making them thermodynamically stable. The fibrils formed are known to be insoluble and resistant to proteolytic degradation and assume a sold phase when associated with diseases, as opposed to the soluble forms which are more associated with biological functions. [13]

2.1.5   History of the Structural Characterization of Serum Amyloid A

While the sequence of SAA proteins is well known and studied, the 3-dimensional struc-
ture of SAA proteins are not well defined; this is attributed to the fact that SAA proteins are
not soluble, specifically when SAA is not bound to HDL. Consequently, the exact function
of SAA proteins is largely unknown nor is it clear how exactly SAA plays a role in the
diseases with which it is associated. The following is a proposed chronological timeline
of select structural milestones regarding the SAA monomer; the following papers not only
provide relevant information but also represent a general idea of the research that was
being done during that time. In no way is this an exhaustive list.

Secondary Structure Data, Sunde et al. 1987 —Using X-ray diffraction, Sunde et al.
show that six amyloid fibrils share similar chemical properties, even composed of different
proteins that do not share similar structures. [23]

Circular-dichroism (CD) Studies of SAA, McCubbin et al. 1998 — The secondary
structures of SAA proteins, and their isoforms, were actively studied in the late 1980s. In
their CD study, McCubbin et al. determined the secondary motif of the two murine SAA
proteins, SAA1 and SAA2. This study focused on the secondary structure in an attempt
to gain insight into how amyloidogenic proteins aggregate. They specifically quantified
the α-Helix, β-Sheet, and β-Turn in for each protein in different environments (various
concentrations of ions and in the presence and absence of heparin). In comparing the

two proteins they showed that SAA2 had around one-half of the α-Helix that SAA1 contained. In addition, they showed that in the presence of heparan, SAA2 has a higher β-Sheet content and there is no change in secondary structure for SAA1. They conclude that there are structural dissimilarities shown between the two proteins and that these are magnified in the presence of certain molecules which may favor amyloid formation (SAA2 as it was understood then, was more aggregation-prone). Note this experiment was still over murine SAA proteins. At this time human SAA structure was still being predicted. [24]

Crystal Resolution, Wang et al. 2002 — Wang et al. showed that lipid-free SAA2.2 has the ability to exist in both a monomeric state and a hexametric state in solution, with the monomeric state being the dominant state. Note, these experiments were performed over murine SAA2.2 which is more soluble and also known to be amyloid resistant. They also performed trypsin proteolysis from which they deduced that the residues 39-86 were resistant to protease degradation and located in flexible regions of the protein and that arginine 18 is cleaved rapidly. From bands on SDS/PAGE gels, the authors conclude that residues 1-18 may also interact highly with itself. In addition, they used circular-dichroism to find that around 50% of the residues keep a helical structure and around 10% assume beta-sheet motifs. They determined that regions with higher alpha-helical content, i.e. residues 13-29 and 43-86, where the regions where hexamerization could occur and that in regions with low alpha-helical, content hexamerization does not occur. [15]

Crystal structure of human SAA1, 2014 Lu et al. — Only recently was the crystal structure of SAA resolved. Lu et al. reported the crystal structure of human SAA1.1

23

(PDB-ID 4IP8) and also found that SAA can exist as a monomer, trimer, or a hexamer depending on the solution that the protein was in. They report that residues 1 and 3 are not exposed to solution when SAA exists as a hexamer which indicates that it is not an amyloidogenic in its native state. They determined that in its functional monomeric form, SAA is 104 amino acids in length, has 4 alpha helices that form a cone shape, and a disordered C-terminal that wraps around the 'cone'. [17]

2.2   Simulating Structural Transitions in SAA

Common computational techniques, such as molecular dynamics (MD), are limited as they often cannot cover the time scales on which biological processes occur. For instance, a computational simulation that accurately describes the folding of a protein requires information about the entire free energy landscape - if a protein is not simulated long enough, sufficient sampling of the landscape will not be achieved, nor will accurate configurations of the protein be found. Hence, there is a need for advanced simulation techniques that can sample the entire landscape of a protein.

One of the main challenges regarding MD is the energy barriers between configurations. This, and the use of small timesteps that are used in MD (femtosecond range) make sampling the conformational space of a protein computationally expensive. Advanced sampling techniques seek to sample the conformational landscape of a protein more efficiently. For a comprehensive review of advanced sampling techniques and description

see reference [25].

Replica Exchange with Tunneling (RET), an advanced sampling algorithm established by Hansmann et al. [26], was designed to study proteins that transition between distinct conformational states. Here, RET is used to study the conversion of the protein Serum amyloid A (SAA) from the state that is associated with normal physiological function to a misfolded state that results in amyloidosis, which can lead to organ failure throughout the body.

2.2.1 Replica-Exchange-with-Tunneling

Deriving the mechanism by which SAA assumes its fibril form through computer simulations requires an exhaustive and accurate sampling of the free energy landscape of the protein. This, however, is a difficult task as typical time scales in protein simulations are only of order μs. We have proposed in earlier work [27, 26, 28, 29] a variant of the Hamilton Replica Exchange method [30, 31] as a way to increase sampling for the special case of transitions between well-characterized states. Our approach relies on two ingredients. First, a ladder of replicas is set up, where on each replica a 'physical' model is coupled with a structure-based model. On one side of the ladder the structure-based model biases the physical system toward the native state of SAA, while on the other side the bias is toward the fibril. The strength of the coupling (biasing) on a replica is controlled by a parameter λ which is maximal at the two ends, and zero for the central replica where the

25

physical model is therefore not biased by one of the structure-based models. Exchange moves between neighboring replicas induce a random walk along the ladder by which the SAA conformation changes from one motif into the other. Accepting or rejecting these exchange moves with the criteria commonly used in Replica Exchange Sampling [32] guarantees that the correct and unbiased distribution of the (unbiased) physical model will be sampled at $\lambda = 0$. However, the acceptance probability becomes vanishingly small for large systems. In order to avoid this problem we conditionally accept the exchange and rescale the velocities of atoms in the two conformations A (moving from replica 1 to replica 2) and B (moving from replica 2 to replica 1):

$$v'_A = v_A \sqrt{\frac{E_2 - E_{phy}(q'_A) - E_{Go}(q_A) - \lambda_2 E_\lambda(q_A)}{E_{kin}(v_A)}} \quad , \qquad (2.1)$$

$$v'_B = v_B \sqrt{\frac{E_1 - E_{phy}(q'_B) - E_{Go}(q_B) - \lambda_1 E_\lambda(q_B)}{E_{kin}(v_B)}} \qquad (2.2)$$

such that the total energy at both replicas stays the same before and after the exchange. After the exchange, the systems on the two replicas evolve by microcanonical molecular dynamics, exchanging potential and kinetic energy. Once the velocity distribution for each replica approaches the one expected at the given temperature, the final configuration $\hat{B} = (\hat{q}_B, \hat{v}_B)$ on replica 1 has a comparable potential energy to the configuration A, while the potential energy of $\hat{A} = (\hat{q}_A, \hat{v}_A)$ on replica 2 will be close to that of the configuration B. Defining $\Delta E_{phy}^{(i)} = E_{phy}^i(\hat{q}_A) - E_{phy}^i(q_B)$, and accordingly $\Delta E_{Go}^{(i)}$ and $\Delta E_\lambda^{(i)}$, the *exchange*

is accepted with probability

$$\exp\left(-\beta_1\left(\Delta E_{phy}^{(1)} + \Delta E_{Go}^{(1)} + \lambda_1 \Delta E_\lambda^{(1)}\right) - \beta_2\left(\Delta E_{phy}^{(2)} + \Delta E_{Go}^{(2)} + \lambda_2 \Delta E_\lambda^{(2)}\right)\right).\qquad(2.3)$$

If rejected the simulation continues with configurations A (on replica 1) and B (on replica 2). We have shown that this *Replica-Exchange-with-Tunneling (RET)* procedure leads to the correct distribution for sufficient large systems and sufficiently long microcanonical segments, for detail see references [27, 26].

Currently, the largest limitation to the RET algorithm is the computational cost of the algorithm. In this experiment an average of 8 ns was achieved per day. Although there were 24 replicas and exchanges being evaluated, the computational cost is closely related to size of the proteins that can be simulated. The larger the proteins the greater the computational cost.

Regarding the computational speeds that currently limit the use of RET, there are two pragmatic actions that could be taken to increase the computational times; use GPUs with the code and move the algorithm to an alternative platform. The algorithm currently is programmed to run on CPU cards. It could be possible that configuring the algorithm to run on GPUs would increase the speed to be used for larger IDP proteins more efficiently.

RET simultaneously simulates multiple starting structures in the form of a replica and then evaluates the similarity of the replica in terms of energy. A large number of configurations

are generated in this process (the longer the simulation the more configurations are generated) and each of these configurations are physics based models. If more models were needed that were not similar, the user could use a large lambda value.

2.3   Materials and Methods

2.3.1   Simulation Set-up

In order to study the interconversion of Serum Amyloid A (SAA) between the native and fibril conformations, we describe in our RET simulations the "physical" model of the system by an all-atom energy function: the CHARMM36m force field [33] in conjunction with TIP3P [34] water molecules. The SAA_{1-76} fragment and 18,376 water molecules are put in a box of edge size 8.3 nm with periodic boundary conditions and the system neutralized with 2 sodium ions. The SAA fragment is capped with a acetyl group and a methylamine group on the N-terminus and the C-terminals, respectively. Start configurations are generated from the experimentally resolved structure of the full-size monomer as deposited in the Protein Data Bank (PDB-ID: 4IP8) [17] by discarding residues 77-104, and randomizing the resulting fragment SAA_{1-76} in a 1 ns simulation at 1500 K, followed by another simulation of 1 ns at 310 K. After visual inspection for loss of secondary structure, a final minimization of the resulting conformation follows. The biasing Go-model uses as target structures the fragment SAA_{1-76} as derived from the PDB-model 4IP8 after discarding residues 77 to 104, and the fibril model as deposited in the PDB under PDB-ID:

28

6MST [35]. As the fibril model is only for the fragment SAA_{2-56} we had to add an arginine at the first position, and the C-terminal residues 57-76 which was accomplished by using CHIMERA. [36] With the so-generated two target structures as input for the SMOG-Server [37] at http://smog-server.org we derive then expressions for the corresponding Go-model energies E_{Go}. Note that masses in the Go-models are scaled by a factor 14.21 to account for the lack of hydrogen atoms in the Go-models. Finally, we couple 'physical' and Go-models by an energy E_{λ} as define in references [38, 39], which quantifies the similarity between the two models. The strength of the coupling differs between the replicas. We used the following distribution of λ-parameters for the 24 replicas: $\lambda = 0.08, 0.07, 0.06,$ 0.05, 0.04, 0.03, 0.02, 0.01, 0.007,0.006, 0.005, 0, 0, 0.005, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.1. Note that for the replicas 0-10 the coupling is between physical model and a Go-model is defined by the native conformation, while for replicas 13-23 the physical model is coupled with a Go-model defined by the fibril model. In order to avoid an odd number of replicas we use two central replicas with $\lambda = 0$ (replica 11 and 12), where the physical model is not biased by a Go-model.

The above set-up is implemented in GROMACS 4.5.6 [40], used by us to compare our data with earlier work. Hydrogen bonds are constrained for both the physical and Go-models using the LINCS algorithm [41] leading to a time step of 2 fs. The Van Der Waals cutoffs are set to 1.2 nm, and the Velocity Verlet algorithm [42] is used to integrate the equations of motion. Temperature is controlled by the v-rescale thermostat [43]. Note that our implementation of the RET algorithm requires different temperatures for each replica, which therefore slightly changes between 310 K and 310.23 K (in steps of 0.01 K). The

length of the microcanonical segment is set to 1 ps in a RET move. Two independent trajectories of 100 ns and 85 ns were generated, taking measurements only from replica where $\lambda = 0$, i.e., where the physical model is not biased by any Go-term. Allowing for convergence of each trajectory we are left with a total of 125 ns for analyzing the free energy landscape of the system.

2.3.2 Analysis, Tools, and Protocols

We use GROMACS tools [40] and the MDTraj [44] software package in the analysis of the RET trajectories for measuring root-mean-square-deviation (RMSD), dihedral angles of residues, and number of contacts. Secondary structure analysis and evaluation of free energy landscapes is done by in-house-scripts. The transition pathway between native and fibril SAA conformations is derived from the free energy landscape projected on the RMSD to the respective structures. Using Dijkstra's algorithm [45] as implemented in the MEPSA software, [46] we construct for this purpose a minimum (free) energy pathway between the two corresponding minima, allowing us to identify the barriers among different minima along the pathway. While this pathway may not be the optimal (i.e., the physical) pathway, we found the method in past work [47] reliable and less costly than transition path sampling, [48, 49, 50] the string method, [51, 52] the kinetic network model, [53] traveling-salesman based automated path searching (TAPS), [54] or other competing approaches. For visualization, we use the PyMOL software. [55].

2.4 Results

We argue that the computational difficulties in simulating the conversion between native and fibril SAA conformations is alleviated in our variant of Replica Exchange Sampling. In order to support this assumption, we show in Figure 2.1a) the walk of a typical realization of our system along the ladder of replicas. At the start time (t = 0), the physical system sits on a replica, where it is biased with $\lambda = 0.1$ toward the folded structure but walks numerous times between this replica and the opposite endpoint of the ladder where the physical system is with λ max = 0.08, i.e. maximally biased toward the fibril. The average exchange rate between neighboring replicas along the ladder is approximately 25%. Monitoring the respective root mean square deviation (RMSD) toward both folded and fibril structure, we show in Figure 2.1b) that this walk through λ-space indeed induces interconversion between the two motifs.

The higher rate of transitions allows for a more exhaustive sampling of the free energy landscape. A measure for the efficiency of our method, and a lower limit on the number of independent configurations sampled at the $\lambda = 0$ replica, is the number of walks across the whole ladder, from the replica with maximal bias toward the folded structure to the one with maximal bias toward the fibril, and back. The number of such tunneling events is inverse to the average time needed to cross the ladder (termed by us the tunneling time). The higher the number of tunneling events, the shorter the tunneling time and the more efficient our approach will be. This convergence of the simulation is checked by

31
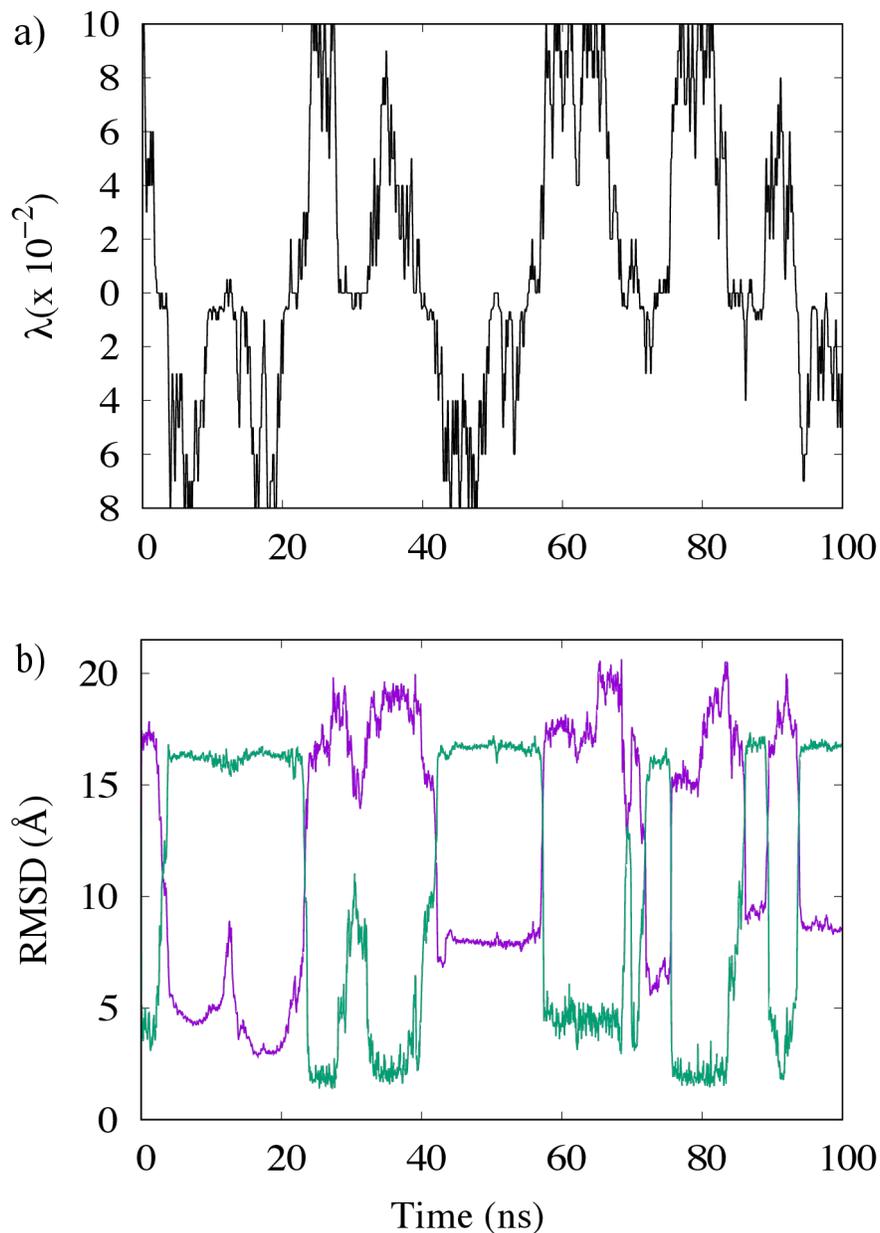
Figure 2.1: (a) A typical example of a replica walking through λ space starting from a replica, where the physical model is initially biased toward the folded SAA structure. While the system walks between replicas with bias toward the folded structure (upper half) and such with bias toward fibril structure (lower half), its configuration changes accordingly. This can be seen in b) where we show the corresponding time evolution of the RMSD to the native structure (in magenta) and the fibril structure (in green).
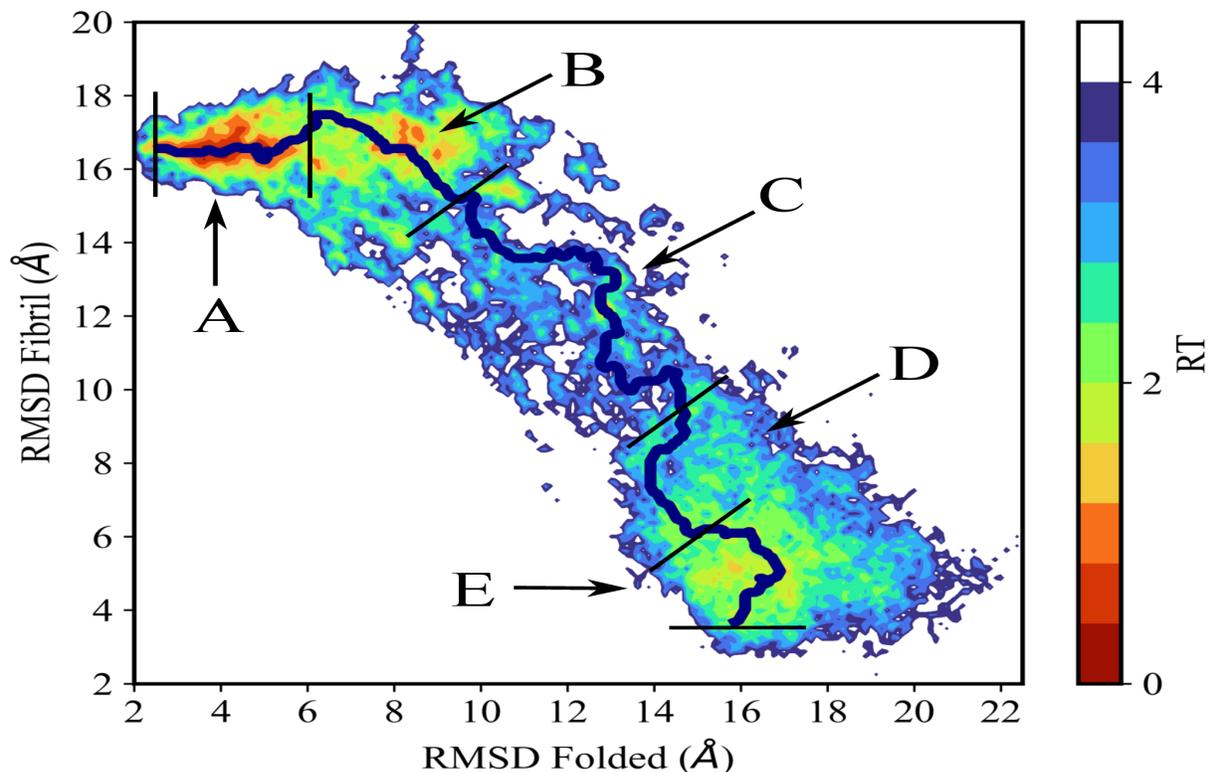
Figure 2.2: Free energy landscape as obtained from RET simulations, with data taken at replicas where the physical models are not biased by any Go-term. Energies are listed in units of RT. The prospective transition pathway is drawn in black, and the five regions crossed by this path marked in capital letters.

comparing for two distinct time intervals the free energy landscape projected on the two

distances introduced above. Resemblance of the data for these two intervals suggests

that the simulation has converged after 10 ns, therefore, we use the last 90 ns of this

specific trajectory for our analysis. In this time span, we find at least four tunneling events

with an average tunneling time of approximately 27 ns.

While small, the existence of at least a few tunneling events gives us confidence in the

reliability of our data at $\lambda = 0$, i.e., at a replica where the 'physical' model of our system

is not biased toward either structure. For instance, we show in Figure 2.2 the free energy

landscape of the SAA1–76 chain projected on the root mean square deviation (RMSD) to either folded or fibril conformation. This landscape is characterized by two prominent basins corresponding to either folded or fibril conformations, which are observed with frequencies of about 10% and 4%, respectively. The interconversion process corresponds to a path connecting the two basins. We have shown in earlier work [47] that the physical path does not necessarily correspond to the ones seen in tunneling events as RET relies on an unphysical dynamic. Hence, in order to find a more realistic pathway, we have used the MEPSA software [46] to determine the minimum energy pathway. This pathway, drawn as a black line in the landscape, proceeds through a series of basins and appears to be thermodynamically reasonable. Note that by construction this minimum energy pathway does not connect specific configurations but bins. Each bin contains a certain number of configurations sampled throughout the simulations. These configurations may differ in their secondary structure or the native contacts seen in either the folded or the fibril structure. The frequency with which the three quantities are observed allows one to identify five distinct regions along the pathway that correlate with the basins and barriers of the landscape. These regions are labeled as A to E in Figure 2.2, and characteristic conformations for these regions are shown in Figure 2.3, where in the case of region A (characterized by a high frequency of folded conformations) and region E (rich in fibril-like conformations) these characteristic conformations are superimposed on the respective reference structures. Region C appears to be a transition region spanning a large range of diverse conformations. Note, however, that the barrier height is only between two and three $k_B T$.
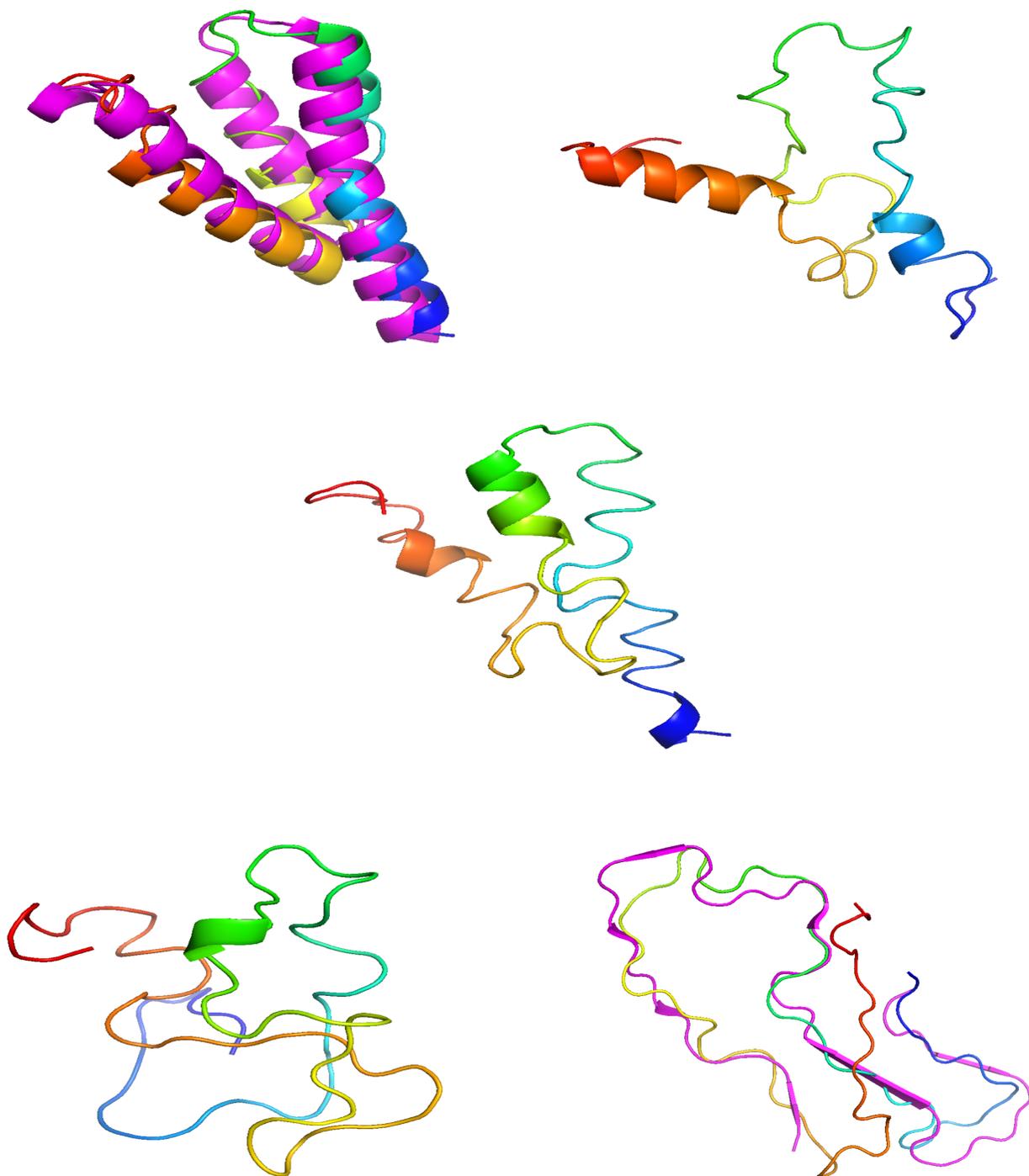
Figure 2.3: Characteristic conformations of SAA$_{176}$ as seen in each of the five regions identified on the proposed transition pathway. The N-terminus of the chains is colored in blue. For region A (dominated by folded-conformations) and region E (where fibril-like conformations are dominant) are these conformations superimposed on the respective reference structures.

In order to get insight into the structural changes along the proposed transition pathway, we augment the visual inspection of the dominant conformations by an analysis of the number n_{NS} of contacts found also in the folded structure and of the number n_{NS} of contacts also found in the fibril structure. Here, we define a contact by the condition that the distance between at least one pair of heavy atoms in residues i and j is smaller than the cutoff value of 4.5 Å. If, in addition, the two residues are separated by at least three other residues in the protein sequence, we call this a long-range contact. The two quantities, normalized to one, are shown in Figure 2.4a). As expected, the number n_{NS} of contacts found also in the folded structure decreases when going from region A, characterized by a low RMSD to the native structure, to region E (which has a low RMSD to the fibril structure). The opposite behavior is seen for the number n_{NS} of fibril-like contacts. However, while there is a steep increase in n_{NS} when going from region C to region E, the corresponding change for the number n_{NS} of contacts seen also in the folded structure appears to be more gradual when going from region A to region C. This picture changes when one considers only long-range contacts seen in the folded structure. This number n_{LR} is also shown in Figure 2.4a) and has a more pronounced behavior. While the number decreases rapidly going from region A to region B and from region C to region E, it changes little between region B and region C. This indicates that the conversion process from the helix bundle of the folded conformations starts with the decay of inter-helical contacts, while the transition region C is mainly characterized by the decay of intra-helical contacts, with the contacts found in strand-like conformations forming in region D. Note that we do not show separately the number of long-range fibril-like contacts as by definition contacts involving strands are long-range.

For a more detailed analysis, we have also shown in Figure 2.4b) the frequencies with which one of the three helices of the folded structure, or the two main strands of the fibril structure are found in the five regions. Values are again normalized to one. The three regions of residues 2–27, 32–47, and 50–69 are considered helical if at least 30% of residues have dihedral angles as seen in an α-helix, that is, if the dihedral angle pair (Φ, Ψ) takes values of between [-100°, -67 ^{o}] and [-40°, -7°]. Similarly, we define the 'strand-ness' of the two segments of residues 6–9 and 18–23 by the requirement that at least 20% of residues in the segment have dihedral angles as in a strand ([-160°, 160°] and ([100°, 100°]). The plots show that the N-terminal helix-I is the first to dissolve, with half of it already gone in region B. The central helix-II starts to decay later, but its propensity is also much decreased in the transition region C and neglectable in regions D and E. The C-terminal helix-III is the last one to decay and observed with substantial frequency even in region E. Note, however, that helix-III covers a region that was not resolved in the fibril structure and well may have transient helicity.

We observe a complementary picture for the two segments that are strand-like in the fibril. Again, we have initially a larger frequency for the N-terminal segment of residues 6–9, with only in the transition region C the second segment gaining higher propensity for strand formation. Both the early decay of the first helix and the higher propensity of the N-terminal segment to form strands are consistent with experimental observations and our earlier work that demonstrated the importance of the N-terminal residues 1–11 and the role of helix-I for fibril formation. [56, 57, 58, 59, 60] Interestingly, we observe in region

37
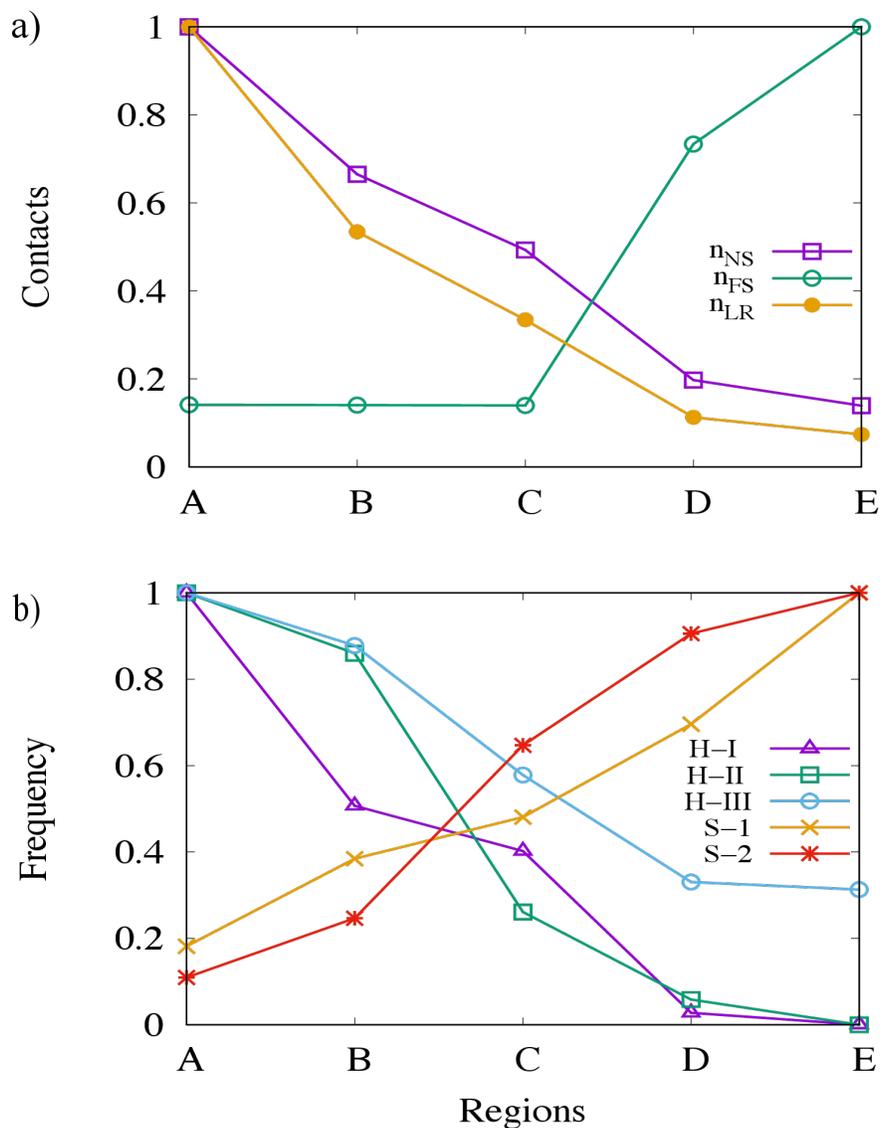
Figure 2.4: (a) The number of contacts (normalized to one) n_{NS} that are shared with the folded structure as measured in each of the five regions A to E of the transition pathway. The sub-set of long-range contacts n_{LR}, again normalized to one, is drawn separately. Shown are also the number n_{FS} of contacts also found in the fibril reference structure. In (b) and (c) we show the relative frequency with which one of the three characteristic helices of the folded structure, or the two main β-strands of the fibril structure, are observed.

B a mixture of the helix-broken and helix-weakened conformations, discussed in ref [56], while in region C the conformation either lost its helicity or resembled more the aggregation prone helix weakened conformations.

The above picture is also supported by the contact maps calculated for each of the five regions that are shown in Figure 2.5. Note that the coloring does not indicate frequency of contacts but the average distance between residue pairs. As discussed in [56], destabilization of the folded structure starts quickly after cleavage of the full-size monomer into a SAA1–76 fragment with a loss of contacts between residues 60–69 on one side and residues 70–76 on the other side. Already in region A are only about 30% of these contacts still existing, and the frequency decreases further to about 17% in regions B and C, before disappearing. As a result, helix-III is more flexible than in the full sized protein, allowing it to either break up (in helix-broken conformations) or partially dissolve (in helix-weakened conformations), thereby reducing contacts with helix-II. Overall is the loss of native contacts in regions A to C consistent with the process described in ref [56], where the data were derived with a different simulation technique. On the other hand, in region D starts the formation of fibril-stabilizing intrachain salt bridges and contacts between residues 26 and 34, 26 and 46, 29 and 33, and 35 and 43 that have been also described in reference [59]. Note, however, that the contact between 29 and 34 in the helix-I to helix-II linker region is already observed with substantial probability in region A and is therefore likely crucial for fibril formation of SAA. The salt bridge between residues 26 and 34 is also formed early, seen in all five regions with about 10–20% probability. Hence, this contact is therefore a bottleneck or a condensation point in fibril formation. Note that the contact

of 26–46, which competes with the 26–34 contact, is seen in region D with 5% frequency but only with about 1% in the fibril region E. These results are also consistent with earlier work [60] that emphasized the importance of the helix-I to helix-II linker region for fibril formation, albeit we do not see the reported local unfolding of the two helices around this region.

2.5 Conclusions

Using a variant of Replica Exchange with Tunneling (RET), we have probed the interconversion of SAA1–76 monomers between the folded structure and the one assumed in the experimentally resolved fibrils. In vivo, this conversion happens at high (micromolar) SAA concentrations and likely will therefore depend on the interaction with neighboring chains. However, modeling the conversion under these conditions is not feasible, and we therefore restrict ourselves to an isolated monomer. Our assumption is that both folded and fibril configurations are prominent local minima in the free energy landscape of the protein and connected by similar transition pathways as occurring under physiological conditions. Hence, the main limitation of our study is our conjecture that the energy landscape is not drastically altered by interactions with other chains changing barrier heights and local minima. The interaction with other chains becomes in this picture an external field that shifts the equilibrium toward the fibril structure. While this picture is an oversimplification, it allows us to propose a conversion mechanism for SAA fibril formation, identifying critical residues and intrachain interactions, that may guide future experiments. Specifically, we
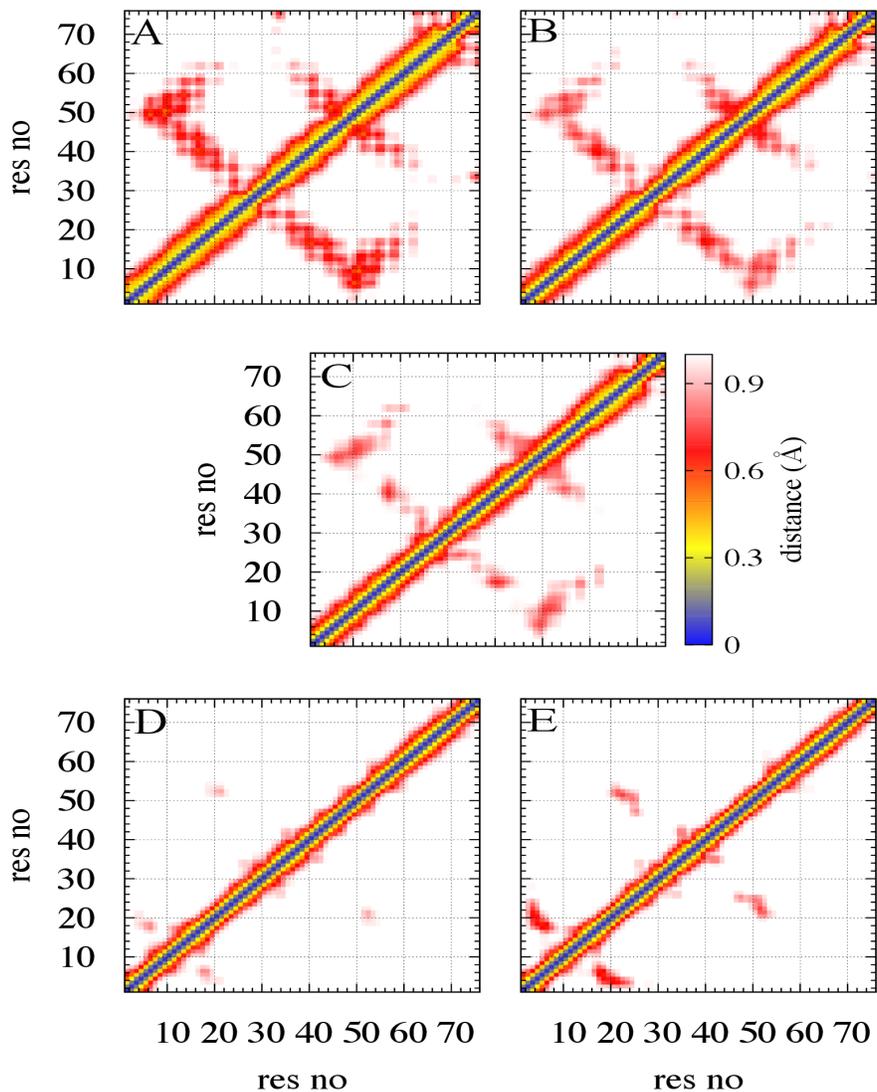
Figure 2.5: The residue-residue map of the average minimal distance between heavy atoms in a pair of residues, shown for each of the five regions. Residue pairs whose average contact distance is more than 10 Åare excluded. Numbers on the X and Y axis mark residues, and the average distance in Å is given by the color coding shown in the middle row.

find only small free energy barriers (of the order 2–$3k_BT$) separating the folded and fibril structures that can be crossed easily once a critical nucleus for fibril formation is formed. Consistent with our earlier work we find that the decay of the helix bundle of the folded structure progresses by a loss of interhelical contacts between helix-II and helix-III and helix-I and helix-II, leading to the helix-broken or helix-weakened conformations of [56]. Dominant on the path way are the more aggregation prone helix-weakened conformations where the C-terminus of helix-III interacts with the C-terminus of helix-I and increases the flexibility of helix-I. The resulting higher entropy leads to destabilization of helix I, causing a loss of helicity. Especially important is the release of the first 11 N-terminal residues from helix-I, which then can misfold into strand-like configurations. This strand segment is indeed the first one that appears in the conversion process seen in our RET simulation. We had shown in earlier work that the merging fibril structure is stabilized by intrachain contacts between residues 26 and 34, 29 and 33, and 35 and 43, connecting residues located on the helix-I to helix-II linker in the folded conformation. These contacts indeed appear once the barrier is crossed. Interestingly, the salt bridge between residues 26 and 34 is seen already transiently in early stages of the unfolding of the folded conformation and likely plays a crucial role for fibril formation. This could be tested by mutating one of these two residues to inhibit formation of this salt bridge.

Chapter 3

Understanding the Antimicrobial Effects of the Peptide Nisin

Nisin is a 34-residue protein that induces cell death through pore formation in the membranes of antibiotic resistant Gram-positive bacteria cells; however, the mechanism by which pores form and are stabilized remains unknown. For this reason, all-atom molecular dynamics is used to investigate the biophysical properties of nisin in a lipid environment, the dynamics of biological membranes, and how membrane dynamics affect the proteins in a membrane. Sections 3.7 - 3.9 are taken from and paraphrase our previously published work: *M.S. Sheridan, P. Pandey, and U.H.E. Hansmann,* In Bacterial Membranes Lipid II Changes the Stability of Pores Formed by the Antimicrobial Peptide Nisin, *J. Phys. Chem. B, 128(2024)* [7].

3.1   The Problem of Antibiotic Resistance Bacteria

Antibiotics, discovered at the beginning of the twentieth century, transformed medicine and improved the quality of life. The organisms that caused infectious diseases, for the first time, were specifically targeted; antibiotics cured many diseases by killing the small

bacteria that caused them. However, not quite a century later, the possibility of soon being without antibiotics is being realized.

The antibiotic resistance issue is especially concerning to the medical field as antibiotics play such a large role in treatment. Common medical procedures that rely on antibiotics include chemotherapy as a cancer treatment, surgeries, joint and hip replacements, organ transplants, dialysis, and care for premature babies. Without antibiotics, these common procedures will not be able to take place. [61]

Bacteria have self-defense mechanisms that promote and protect the life of the cell through various mechanisms such as mutations and horizontal gene transfer. Mutations allow the bacteria to impede the mechanism that an antibiotic uses. For example, if a bacteria changes the D-Ala-D-Ala amino acids sequence used in its cell wall synthesis vancomycin, which is designed to target that specific sequence, will no longer work. Because of horizontal gene transfer, and the innate way that bacteria replicate, that strain of bacteria will acquire the resistance very rapidly.

While new antibiotics that target bacteria in ways in which the bacteria have not become resistant to are needed, there is also a need to gain a more specific and detailed understanding of how bacteria become resistant, as well as explicitly how antibiotics work. In-depth knowledge of the mechanism that bacteria use to gain resistance is needed for recognize patterns of resistance and monitoring levels of resistance bacteria as well and choosing the most efficient drug for various strains of bacteria [62]; this is one way to

44

decrease levels of resistance. [63] There is also a possibility to exploit mechanisms from natural bacteriocins. However, the ability to use natural antibiotics with little or no modification, as in the case of penicillin, is unlikely. Understanding the proteins mechanism in its natural environment is needed to apply it to other environments, as in the case of nisin. To gain a detailed understanding of mechanisms, atomic details need to be addressed. For this, models are relevant and needed.

3.2   Introduction to the Lantibiotic Nisin

Nisin is widely utilized in the food industry, as it is highly effective, naturally produced, safe, and works well as a food preservative. It possesses strong antibacterial activity toward many gram-positive bacteria and therefore, it has the potential to be used as an antibiotic drug. Although nisin possesses promising antimicrobial activity, its susceptibility toward proteolytic degradation [64], poor solubility, and stability at physiological pH [65] limits its use. Its adaptation in drug design is further limited by a lack of understanding about the protein and the mechanism by which it is effectual. While likely related to its ability to form pores in membranes, the mechanism that nisin uses has to be distinct from other antibiotics. This indicative as nisin effective against methicillin-resistant bacteria (MERSA strains) [66] and shows anti-cancer properties [67, 68]. Hence, it is important to understand the pore formation mechanism of nisin in bacterial and cancerous cells, to identify the factors that affect the stability of the pore, and to probe the interaction between nisin and the membrane.

While it is known that nisin, a naturally occurring bactericide, is an effective agent against bacteria, the mechanism this protein uses remains elusive. The leading hypothesis is that a nisin peptide, external to the membrane, will first bind to the lipid II headgroup through hydrogen bonds between the amino protons of the nisin molecule and the pyrophosphate group of the lipid II molecule. Once bound, the lipid II molecule acts as an anchor for the peptide and the C-terminal end of nisin then penetrates into the bilayer where it forms a pore complex with 7 other nisin molecules and 4 lipid molecules. As experimental techniques are not sensitive enough to get information we turn to molecular dynamics simulations. To capture the atomic details needed to supplement experimental data, atomistic simulations are required.

The complexity of membranes makes it difficult to derive information about specific proteins embedded in a bilayer. As a result, there is little experimental data available about how nisin enters the membrane or forms a pore once it is in the bilayer. Computational studies are one way to augment experimental data as molecular dynamics simulations can yield atomic details about a protein's behavior that otherwise remains uncharacterizable. Hence, we will use large-scale molecular dynamics simulations to study the interaction of the protein nisin with lipid II molecules and how this interaction of nisin residues with membranes affects the integrity of the membranes.

3.2.1 Classification of Nisin

Nisin is produced by certain *Lactococcus lactis* (*L. lactis*) bacteria which have historically been used for fermentation of food products; the bacteria produces lactic acid which preserves food. *L. lactis* is a gram-positive bacteria that was originally found in plants. The bacteria was initially thought to be dormant but was found to become active in the gastrointestinal tracts of ruminant animals where the peptide displayed probiotics and antimicrobial properties.

Nisin is classified as a lantibiotics. Bacteria in this class are produced ribosomally by Gram-positive bacteria and undergo posttranslational modification; for lantibiotics posttranslational modification result in lanthionine rings. These rings results in connecting serine, threonine, and cysteine residues, through a monosulfide bond, to form a ring or rings in the peptide. These rings are resistant against certain types of enzymatic degradation.

In addition to having innate antibacterial properties, Song et al. point out that *L. lactis* is an important bacteria in terms of genetic engineering as it is a small bacteria and four *L. lactis* strains are fully sequenced. As a result many of its expression systems are well understood (including the genetically controlled expresses nisin); this makes it a good model to study experimentally and is the reason that nisin is the most studied of all the lantibiotics. [69]

3.2.2 A History of Nisin's Use

Interest in nisin was first described in 1928 when L. A. Rogers recorded that certain bacteria limited the growth of other species of bacteria. [70] Note, nisin was discovered the same year that penicillin was discovered. Since then, the role, interest, and significance of nisin in medicine has only increased.

Rogers et al. recorded that a group of bacteria, then called Group N Streptococcus and now known as *lactic streptococci*), inhibited bacterial growth in cheese cultures. This peptide responsible for inhibiting the bacterial growth was named nisin to reflect 'group **N** Streptococcus **I**nhibitory **S**ubstance' and where -in is the nomenclature to indicate antibacterial properties in 1949. [71] In the mid-twentieth century use of the protein in the food industry and in veterinary medicine was being explored. However, because of its low stability at physiological pH, the protein was not used in veterinary medicine; however, its importance in the food industry was being realized. In 1951 nisin was first used in the food industry [72] and since then has been approved for industry use by 80 countries including the FDA which has labeled nisin as GRAS (Generally Recognized As Safe). [71, 72]

In addition to being used to preserve cheese, nisin is now used in dairy and milk products, various canned goods, some meat, eggs, and cream products. [72] Because nisin is considered safe for human consumption large levels of nisin are added into the food products for preservation with no currently known side effects to humans. The harmless-

ness of nisin to humans and its success as an antimicrobial in the food industry is another reason why nisin's properties are being studied – with hopes that the unique antimicrobial properties can be used in medicine where new antimicrobial agents are needed.

3.2.3 Protein Synthesis

Nisin is produced ribosomally by *L. lactis*. The protein that is initially synthesized by the ribosome consists of 57 amino acids, known as pre-NisA. The first 23 amino acid of pre-NisA make up the N-terminal and the last 34 amino acids make up the C-terminal. [71] The N-terminal sequence is recognized by NisT which transports the protein (note, two enzymes, NisB and NisC, modify the protein before they can be recognized by NisT). [71] NisB, as Garg et al showed, dehydrates the serine and threonine amino acids on the C-terminal of pre-NisA in the presence of glutamate, ATP, and Mg^{2+}. [73] The dehydrated amino acids produced are then bound to cysteine molecules through cylase NisC. NisT then transport the protein, still 57 amino acids long, out of the cell, where the N-terminal is cleaved by NisP, leaving a 34 amino acid sequence. [74]

Once the protein is extracellular, *L. lactis* protects its-self from the destruction nisin causes with protein NisI and the protein complex NisFGE; NisI binds to nisin disallowing it to inter-act with the membrane and NisFGE detects nisin in the membrane bilayer and transports nisin out of the bilayer. It is thought that when nisin binds to NisK then NisK autophos-
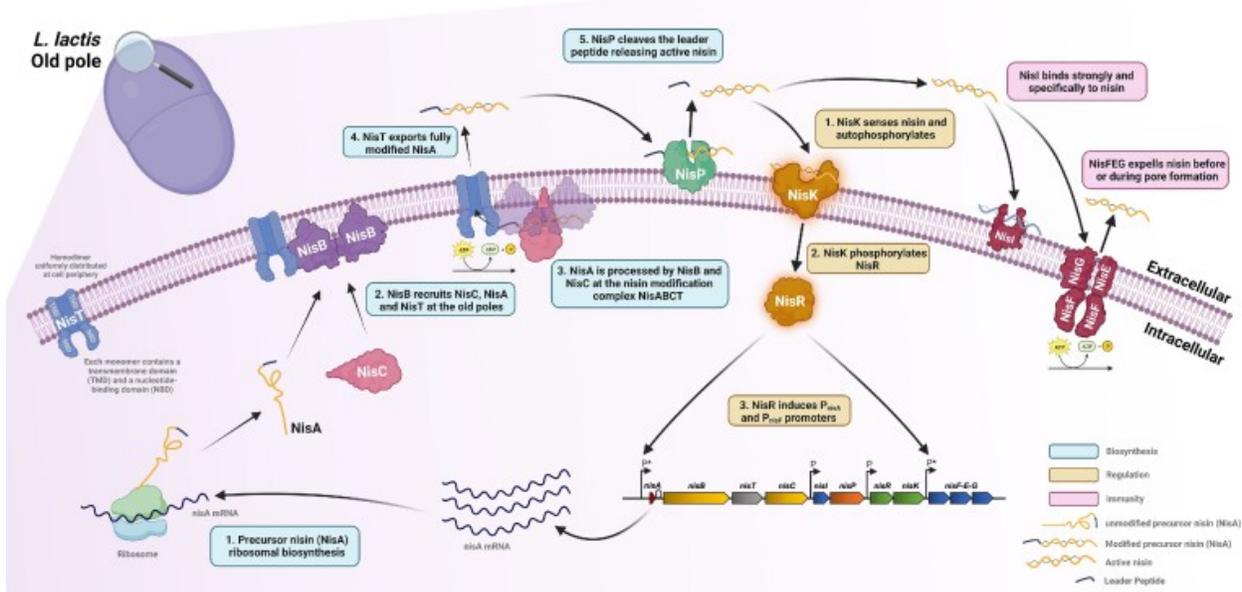
Figure 3.1: "Nisin biosynthesis, regulation, and immunity. Nisin is first synthesized as an unmodified precursor (NisA) consisting of a leader peptide and core peptide. NisA is processed by the dehydratase NisB and cyclase NisC and transported by NisT (NisABCT complex) at the old pole in L. lactis. The mature bioactive peptide is released by leader peptide cleavage performed by NisP. Immunity from the active peptide is provided by two distinct immunity systems composed of a lipoprotein NisI and an ABC transporter NisFEG. To ensure a proper balance between production and immunity, nisin expression is regulated via a two-component signal TCS composed of a sensor NisK, and response regulator NisR that activate the nisin promoter (P*)." Image and caption generated by Field et al.; see reference [71]

phorylates cause a phosphorous group to be transferred to NisR which induces gene expression of nisin. [74]

3.2.4 Biophysical Characteristics of Nisin

Some fundamental characteristics of nisin are worth noting. The activity of nisin is closely related to its structure which varies at different pH values. [75] Nisin is stable and soluble in very acid conditions and becomes unstable at high pH values. The stability of nisin

does not seem to be affected by temperature; it has been reported as stable at low temperatures [75] but can maintain activity at high temperatures and low pH [72], which is beneficial in terms of application.

There are 9 naturally occurring variants of nisin indicating that select mutations are allowed without impeding the overall function of the protein. [76] However, the variants may not be as potent as the wild-type. It has also been shown that mutating specific regions, specifically what is known as the hinge region of the protein or residues 19–21, decreases the effectiveness of the protein.

3.3   Lipid II

Lipid II, in Gram-positive bacteria, comprises less than 1 mol% of the lipids in the membrane and is responsible for the production of the bacterial cell wall. [77] As a result, it is estimated that a cycle of lipid II takes less than 1 second. [78] This is the reason lipid II is a target for antibiotics; There are, however, multiple ways in which lipid II can be targeted.

3.3.1 Biochemical Synthesis of Lipid II

The synthesis of lipid II is worth showing as it is fundamental to bacterial growth in both gram-positive and gram-negative bacteria. Because of this, it has also become a target for many antibiotics.

Intracellularly, precursors of what will eventually be N-acetylglucosamine (GlcNAc), N-acetylmuramic acid (MurNAc), and a pentapeptide – the components of the outer membranes of both gram-negative and gram-positive cells [79] – are bound to a polyisoprenoid lipid. The final structure, which is known as lipid II, consists of two sugar groups, a pentapeptide, two phosphate atoms, and a prenyl chain. lipid II is transported across the membrane bilayer and to the outside of the cell through a mechanism that is unknown where the intermediates in the lipid II structure will be transformed into part of the peptidoglycan wall. [80] The step-wise synthesis of lipid II is shown in figure 3.2 (for an in-depth discussion of synthesis steps see reference [81]).

Of the many intermediates in this pathway, undecaprenyl phosphate is especially interesting. lipid II is transported across the membrane by undecaprenyl phosphate (Und-P), which is a 55-carbon chain that plays a role in multiple biogenesis pathways that produce structures related to the membrane; specifically, Und-P acts as a transporter to carry structures that are synthesized in the cytoplasm to the outside of the cell. [82] The 55 carbons in the chain are highly conserved in bacteria. There has been debate about the

function of the long tail, which is longer than the average lipid in a membrane. It is worth noting that in our study, the C55 prenyl chain was conserved so that a model most reflecting nature was observed. The mechanisms and function of the transporter are also worth noting as they may directly relate to why bacteria have not acquired resistance against nisin. Nisin binds to lipid II, specifically the pyrophosphate group on lipid II. Because this moiety specifically functions to transport structures out of the cytosol, it has been suggested that this orientation cannot change without losing its function. This is a function that a bacteria cannot alter, and by doing so, becomes resistant to an antibacterial mechanism.[78]

3.3.2   Lipid II as an Antibacterial Target

Glycopeptides such as vancomycin, and its derivatives, target the D-Ala-D-Ala amino acids of the pentapeptide in the lipid II molecule, which prevent the penicillin-binding proteins (PBPs) to integrate the cell wall precursors into the cell wall, thus preventing the construction of the outer membrane and ultimately terminating the cell. Interestingly various glycolipids interact differently with those amino acids. The bacteria can alter the amino acid composition and by doing so become resistant to the drug. If the drug is altered to interact with the amino acid sequence in a new way that the bacteria does not 'recognize' a new drug is formed. [83]

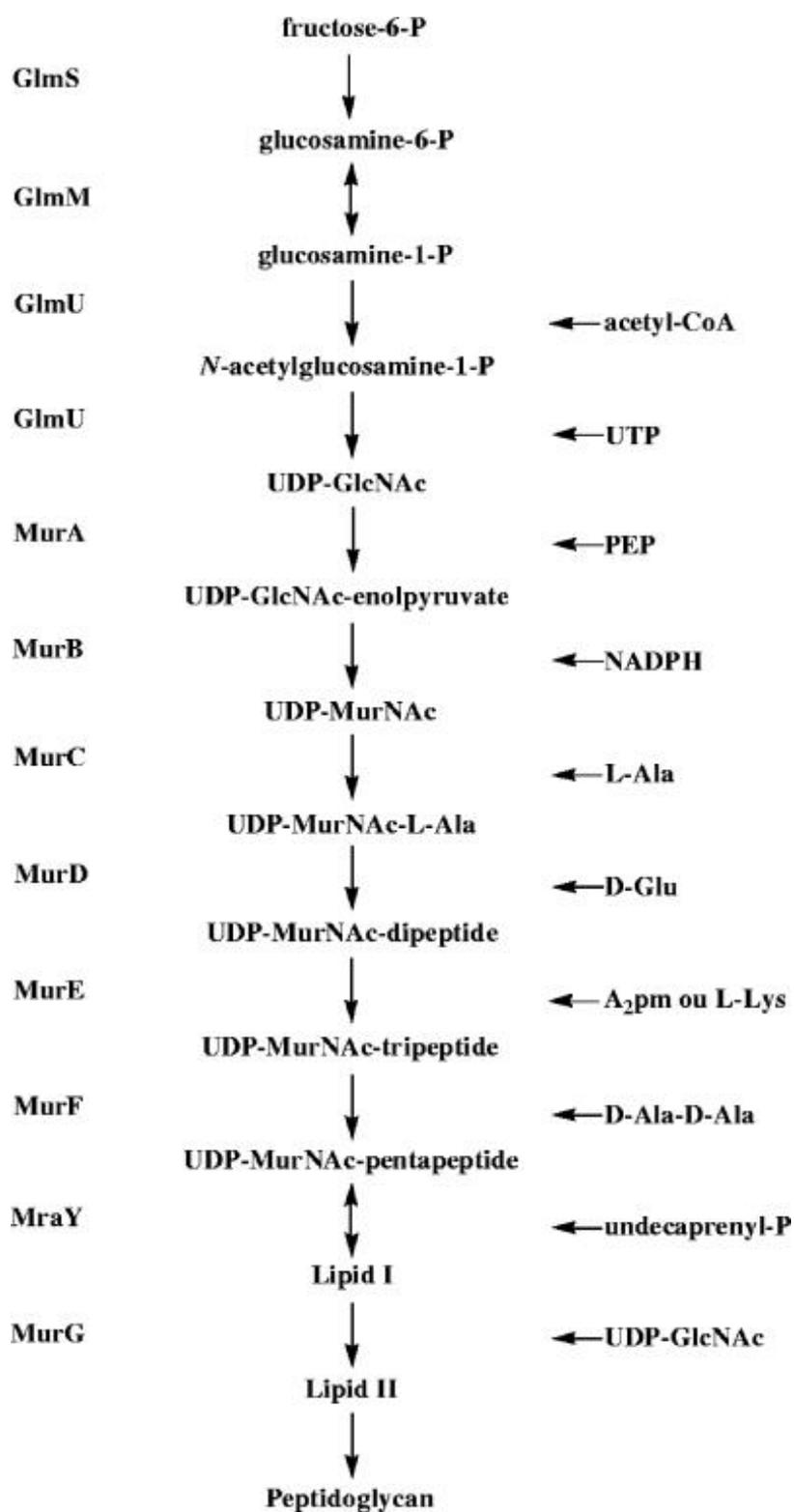Beta-lactams, such as penicillin, target transepetidases or peptidoglycan glycosyltrans-

Figure 3.2: [81]

ferases (PGTs) which are the enzymes that polymerize and integrate the lipid II cell wall. Specifically, Beta-lactams have a structure that is similar to the D-Ala-D-Ala terminus of the pentapeptide which allows them to bind to the transepetidases by which they become inactive. [84] Quio et al. point out that there are bacteria species where the specific mechanism of the transepetidases are unknown because the native lipid II of that species cannot be isolated in sufficient quantities and synthetic production is not efficient. [84] This is a limiting factor in research more beta-lactam derivatives.

Lantibiotics, e.g. nisin, target the pyrophosphate moiety of lipid II in the membrane of a cell. While nisin is also known to form pores in the cell membrane, restricting lipid II to the membrane also inhibits cell wall growth. In this way, nisin had multiple modes of action which may be the reason nisin has such high toxicity toward cells (nisin has been shown to be effective in nanomolar concentrations [85]).

A comparison of the different mechanisms that antibiotics use against lipid II highlights the fact that nisin binds to lipid II in a distinct manner. This unique interaction is of particular interest because bacteria have not yet become resistant to it.

3.4   Pore Formation — A Hypothesis

The exact mechanism of nisin remains unknown, but there is experimental evidence that supports two mechanisms. One is pore formation – nisin and lipid II molecules will form

pores in a membrane which will eventually kill the cell. The other is lipid II and nisin aggregation – these aggregates induce lipid budding which will terminate the cell.

However, the experimental support for the pore hypothesis seems to be stronger. Early studies showed that pores formed in membrane bilayers with lipid II were more stable and also that there was an increase in the affinity of nisin to lipid II. [86] In 2004 Hasper et al. performed fluorescence and circular dichroism experiments to determine the stereochemistry of the nisin lipid II pores in the membrane. From their results, they proposed a new model of the pore, one that had 4 lipid II molecules to 8 nisin molecules, and showed that this ratio provided stability to the pore which was quantified by the pore's lifetime. This new suggested stereochemistry contradicted the previously suggested stereochemistry of 1:1 [87]; this ratio was obtained from experiments that used micelles and not bilayers and this model did not explain certain characteristics of nisin. [88] Widemann et al. used single-channel experiments to determine that a nisin-lipid II has a pore diameter of around 2-2.5 nm [88] providing more evidence for the idea of pore formation.

3.5   Simulating the Nisin–Lipid II Pore Complex

The model which represents the interaction of nisin with a membrane and lipid II molecules will be large, especially when sufficient water molecules and ions are included. Simulating large systems on timescales that show accurate behavior is already a problem. Simulating the entire process of pore formation, which includes protein binding, permeation into

the membrane, the orientation of the proteins in the bilayer, and finally the assembly of the pore complex is not currently accessible.

It is for this reason that a preformed pore complex was modeled and simulated. The stability of the pore is not well understood; information regarding this is then relevant and needed. The preformed pore complex also allows MD to be used in a more prudent manner. The experiment was designed to answer questions about the stability of the pore and how the pore complex interacted with the membrane and vice versa. The questions were specific about the pore complex. The MD trajectories provided a meaningful and quantitative manner to answer the specific questions. Understanding how the nisin permeates into the membrane is important but outside the scope of this study.

Nisin is unique in that it is both an extracellular protein and an intracellular protein. Before it forms a pore complex it is found outside the cell and eventually moves into the bilayer. The structure of the protein may assume a different structure extracellularly compared to its form in the pore complex. Experimentally, there is no resolved structure of nisin in its natural environment extracellularly nor is there a structure of how it exists in a membrane. Structural identification of a protein generally involves first isolating the protein. This cannot be accomplished with transmembrane proteins as their structure necessary depends on the bilayer environment. Experimental techniques that are able to provide atomic structural details of the protein in the bilayer do not exist.

3.6   Limitations of the Model

To maintain an honest analysis of the results of our simulation it is important to outline the strengths and limitations of the model used in this study. This simulation relies on classical mechanics to create a trajectory of a nisin-lipid II pore complex in a membrane environment from an initial configuration; the trajectory results describe the initial config- uration. The model setup reflects all current data about the structure and location of this complex in a gram-positive bacteria membrane.

In this respect there is a highly theoretical component to the model as well. The model created relies on evidence from experimental studies. The studies performed to generate these structural clues were not solely carried out on the nisin pore complex as it exists in membranes; the technology does not exist to allow this. What the model does is take into account all that is known and synthesize it into one image, i.e. the starting configuration. In this way the model is theoretical, in this way, the model is a method that can be used to the pore complexes in a new way. Complete acceptance of the trajectory results would be imprudent, and a degree of skepticism is appropriately placed on this model.

It should be stressed that the importance of the model comes from the atomic description that it provides. Experimentally, the resolution is not specific enough to obtain information on an atomic scale about how nisin and lipid II interact. This makes MD an appropri- ate way to probe this system. Note, we make no claim that the results of the simulation

are necessarily what happens regarding the pore complex – we have not simulated the complete pore formation process. We make the claim that the MD trajectory is the dynamics of the model based on classical mechanics. Using the result of the simulation better questions can be asked. For example, the behavior of the middle region of nisin as it behaves in the membrane has remained elusive; the model, based on the initial setup, provides insight into what might be. Our trajectories show the middle rings interact with the bilayer of the membrane. This is new information about the dynamics of the protein in the membrane environment that would otherwise be inaccessible.

3.7   Materials and Methods

3.7.1   System Preparation

Our simulations, comparing the stability of a pore built from nisin-chains in the presence and absence of lipid II, rely on three elements: generation of a self-actualized membrane, the design and positioning of the nisin-pore complex into the membrane, and the addition and positioning of lipid II molecules.

Membrane Bilayer: As nisin is effective against gram-positive bacteria and therefore interacts with gram-positive membranes, we used the CHARMM-GUI Membrane builder [89] to generate a realistic model of a gram-positive bacteria cell wall with a 3:1 ratio of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG) with 1-palmitoyl-2-oleoyl-sn-

glycero-3-phosphoethanolamine (POPE)) [90, 91], that is 528 POPG and 177 POPE lipid molecules. Our membrane bilayer model was placed in a 150 x 150 x 230 Åbox and was solvated with 137,413 water molecules. Following the CHARMM-GUI membrane builder ion option, we neutralize the system with 0.15 mM KCl. The resulting membrane model was first minimized for 5000 steps and afterward simulated in short molecular dynamic simulations at 300 K and constant volume, gradually increasing step size and step time according to the standard CHARMM protocol. [89, 92, 93, 94] To equilibrize the lipid motions and minimize flip-flops of the bilayer lipid molecules, we simulated the membrane for an additional 1 microsecond at a constant temperature and pressure of 300 K and 1 bar. Finally, we removed the water and ions in this self-actualized membrane, leaving only the bilayer to build the systems considered for our simulations.

Nisin-Pore Complex: In the second step, we inserted a pore complex made of eight nisin peptides and four lipid II molecules in the so-generated self-actualized membrane. The pore was designed to have a diameter of 25 Å, enclosed by eight nisin molecules arranged in a circular fashion, with each lipid II molecule situated around two nisin molecules. The number of nisin proteins forming the pore and the pore size reflect the available experimental data. Breukink et al. used pyrene fluorescence and CD to show that the lipid II and nisin chains are arranged in the membrane bilayer in an 8:4 ratio. [87] Sahl et al. showed, through electrochemistry, that this ratio is constant and that the diameter is appropriately 20-25 Å. [88] As Medeiros et al. pointed out, residue 28 is the last residue in the membrane, [95] and the disordered C-terminus is outside of the membrane. Taking this into account, the nisin chains of our pore model must be long enough to span both layers of the

membrane. However, the sole model of nisin bound to lipid II, resolved by solution NMR and deposited in the Protein Data Bank (PDB) under identifier 1WCO, measures only 30 Åbetween the N-terminus and residue 28, while the bilayer has a thickness of about 43 Å. To obtain sufficient long nisin chains, we have simulated the nisin chains (starting from the 1WCO conformation) for one ns at 315 K, allowing them to unwind partially. The resulting conformation is still similar to the 1WCO model, but with a length of 40 Åbetween the N-terminus and residue 28 fits now to the thickness of the bilayer. This structure also has residue 17 sitting in the middle of the bilayer as indicated by both Medeiros-Silva [95] and Breukink. [86] Figure 3.3 shows the final models from a top-down and profile view.

Experimental data were also used to position the lipid II molecules. For instance, Ganchev et al. showed (using AFM imaging) that the pentapeptide located at the head group of lipid II is just exterior to the bilayer surface and points away from it, [96] and Chugnov et al. [90] showed that the pyrophosphate moiety of lipid II are similarly placed as the phosphate groups in the POPE and POPG headgroups in the bilayer. For this reason, we placed the lipid II head group in such a way that their phosphorus atoms are adjacent to the phosphorus groups of other lipids in the bilayer. As the Dodecaprenyl-C55 chain is highly conserved in bacteria, a chain of 55 carbons was used [95] as the acyl chains of lipid II. While these lipid II chains can adopt either an L, I, or V shape [90], only the I-shape allows them to form contacts with the nisin peptides. Therefore, we considered only this shape, a choice that is justified, as Chugnov et al. also showed that interactions between the head group and nisin are more important. [90] As the N-terminal residues are known to bind to the lipid II molecules through hydrogen bonding involving the amino hydrogens
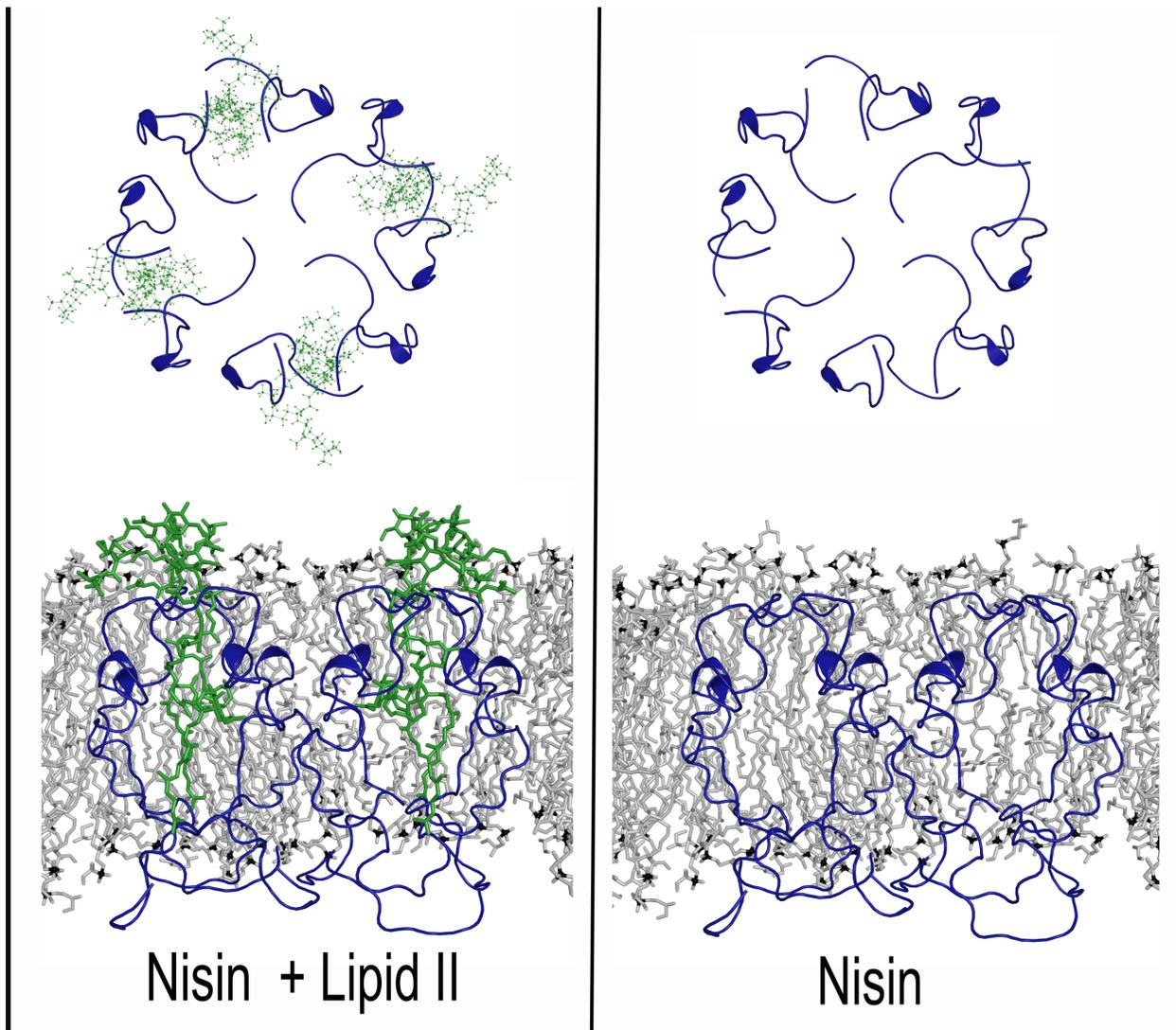
Figure 3.3: The final models of the pore complex represented first as a top-down view then as inserted into the membrane from a profile perspective. Nisin molecules are shown in blue and lipid II molecules are shown in green.
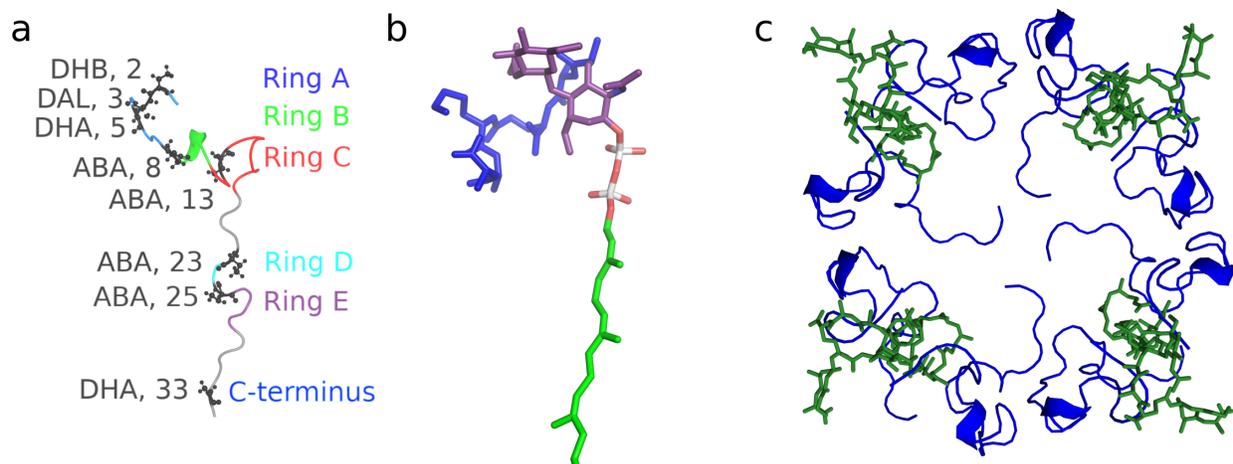
Figure 3.4: Sketch of (a) a nisin monomer, with the positions of rings A–E and the C-terminus color-coded. Listed are also the positions of the uncommon amino acids. In (b), we show the sketch of a lipid II chain, with the pyrophosphate moiety colored in red and white, the prenyl chain in green, and the additional pentapeptide and two sugars in blue and purple, respectively. In (c), we show our pore model formed by eight nisin chains (blue) interacting with four lipid II chains (green).

in the backbone of the PPI group [95], we added the lipid II molecules at a distance of around appropriately 3.5 Åto the first four residues of nisin.

Insertion of the Pore into the Membrane: The above-generated pore complex of nisin and lipid II molecules was then placed into the self-actualized bilayer in an opening of about 40 x 40 Å, which exceeded the pore by about two Åthat was created using the emulate function in Pymol. The resulting membrane-pore system was solvated in 40,000 TIP3P water molecules and 0.15 mM NaCl in a box 150 x 150 x 80 Å. The resulting model was again minimized over 5000 steps and simulated over 1.5 ns at a constant volume and temperature of 315 K. For equilibration, we carried our molecular dynamics runs at 315 K and one bar but restrained the pore in an initial step for five ns with a factor of 1000 kJ/mol. In three additional steps of five ns each, these constraints were gradually

removed in steps of 250 kJ/mol. This procedure leads to an equilibrated conformation that was used as the starting point for our simulations of nisin pore interacting with lipid II and the lipid bilayer in our model of a bacterial cell membrane. See Figure 3.4 for final configuration.

Using Pymol's atom selection feature to remove the lipid II molecules in the above model and simulating the resulting system with restrained positions of the nisin atoms for another one ns, we have generated a control model for the membrane-pore system. As in this model the lipid II molecules, this control allows us to pinpoint how lipid II affects the stability of the nisin pore and its interaction with the bilayer.

3.7.2   General Simulation Protocol

The stability of the above-designed nisin-pore-membrane systems in absence and presence of the bacterial membrane-specific lipid II are studied by molecular dynamic simulations utilizing the GROMACS 2022 and 2023 software packages. [97]

As nisin contains uncommon amino acids (dehydroalanine and dehydrobutyrine, and one lanthionine and four β-methyllanthionine), we had to modify the CHARMM36 all-atom forcefield [98] used by us to describe in combination with TIP3P water [34] inter and intramolecular interactions. Similarly, we had to derive topology and parameters for the lipid II molecules as shown in the Figure 3.4b. Eleven isoprene units, the first seven

units in the (Z)-configuration followed by three in the (E)-configuration, form a lipophilic anchor.[99, 78, 100, 101] It is connected by a pyrophosphate group to the polar headgroup made of a N-acetylglucosamine sugar (Glc) which is attached to a N-acetylmuramic acid sugar (Mur). Additionally linked to Glc is a peptide whose composition varies depending on bacterial species; commonly employed in research studies and also used by us in this study is the $_L$-$Ala - \gamma$-$_D$-$Glu - _L$-$Lys - _D$-$Ala - _D$-Ala pentapeptide. [102, 90] The saccharide MurNAc was created by combining the topologies of N-acetylglucosamine and 2-methoxy-propionate, both from the CHARMM36 carbohydrate force field. [103] Parameters for the non-proteinogenic peptide linkages in lipid II were taken from the CHARMM36 protein force field [98] and partial charges were redistributed among relevant atoms following a consistent trend among all peptide groups.

The membrane-pore complex was solvated by adding water such that the box is extended in z-direction on both sides by 20 Å. We then added 0.15 M NaCl ions to neutralize the system. Periodic boundary conditions are employed in all directions, and long-range electrostatic interactions are calculated with the Particle-Mesh Ewald (PME) technique using a real-space cutoff of 12 Å. Short-range van der Waal interactions are truncated at 12 Å with smoothing starting at 10.5 Å. Before the production run, each system was minimized with 50,000 steps with the steepest descent integrator in [97] and then equilibrated at constant volume for 750 ps at a temperature of 315 K with a leap-frog integrator [97] and a step size of 0.001 fs and then again for an additional 750 ps with a step size of 0.002 fs.

After equilibration, each system (the one which has lipid II interacting with the pore and

Table 3.1: System Setup

System	Number of Atoms	Water Molecules	Independent Trajectories	Trjectory Length (ns)	Total samplint (ns)
Control	183,347	30,968	3	200	600
Experiment	184,479	30,968	3	200	600

membrane is named by us experiment, and the other where lipid II is absent we termed

control) was followed in three independent trajectories that differ their initial distribution of

velocities. The lengths of the six trajectories are listed in Table 3.1, with in each simulation

the temperature and pressure set to 315 K and one bar, as controlled by a Nose-Hover

thermostat [104, 105] and Parrinello-Rahman barostat. [106] Using the SETTLE [107]

and LINCS [108] algorithms to control fast fluctuations and to restrain hydrogens allows

us to use a time step of 2 fs for integrating the equations of motion. PDB files of start and

final conformations are available as supplemental material.

3.7.3 Trajectory Analysis

The molecular dynamics trajectories are analyzed with the GROMACS tools [97], VMD

[109], MDAnalysis [110, 111], and Membrainy [112]. For visualization of trajectory con-

formations, we use VMD. Common quantities such as the root mean square deviation

(RMSD) or root mean square fluctuation (RMSF) are calculated using GROMACS tools,

using for the solvent-accessible surface area (SASA) a spherical probe of a 1.4 Å radius.

Contact frequencies are computed with VMD, where we define contacts by a cutoff of

4.5 Å in the closest distance between heavy atoms in a residue pair. Hydrogen bonds

are defined by a distance cutoff of 3.5 Å between the donor and acceptor atoms and an angle cutoff of $30°$. Membrane properties such as membrane thickness area per lipid, head group angles, and chain order parameters were calculated with Membrainy, with the respective molecules selected by MDAnalysis when calculated only for a shell at a certain distance to the pore.

3.8 Results and Discussion

3.8.1 Lipid II Modulated Nisin Pore stability

We start with looking into the stability of the pore, built from eight nisin peptides and an additional four lipid II molecules, inserted into a bilayer formed from POPE and POPG lipids that serves as a model of a bacterial membrane. Data from three independent trajectories (the experiments) covering 200 ns are compared with three trajectories of the control (where lipid II is absent) having the same length. In Figure 3.5, we contrast for each trajectory the final conformation with the start conformation. For both the experiment and the control systems, we see a slow decay of the pore. While we had 286 expected that in the presence of lipid II the pores (designed to 287 be consistent with the available experimental evidence) would be stable over the lifetime of our simulations, visual inspection shows already that the pore decay is smaller in the three trajectories where lipid II interacts with the nisin pore than in the control. We remark that our model differ from bacterial membranes in that there is no concentration gradient between the intracellular
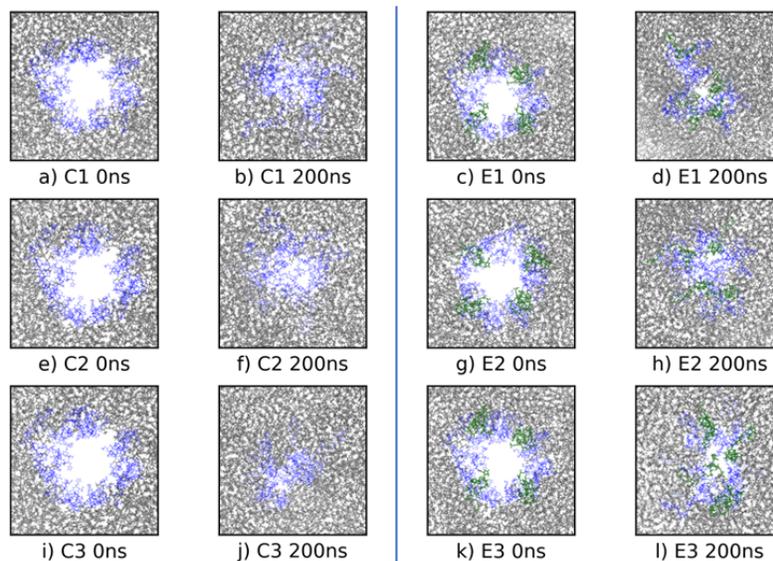
Figure 3.5: Start and final conformations of six trajectories of the nisin pore embedded in the membrane, either in absence (the control, C1 – C3), and presence of lipid II (the experiment, E1 – E3). Nisin chains are colored in blue while lipid II molecules are drawn in green, and POPE or POPG lipids in gray.

and extracellular environment, i.e., the pore cannot be kept open by dynamic pressure resulting from the flo3 of water through the membrane. [86, 87, 88, 95, 113]

To quantify the decay of the pore, one needs to define and measure its extension. Because of the irregular shape of the pore, we chose for this purpose the amount of water inside the pore. To account for the dynamically changing thickness of the membrane, we first calculated the center of mass of all eight nisin chains, taking into account only residues 8-28 located within the bilayer. We then measured the distance to the center of mass for each of the eight chains and chose the largest distance as the radius for a sphere around the center of mass. In this sphere, we deleted all water molecules that are outside the bilayer and the nisin assembly, counting only the remaining water molecules. The numbers obtained, as measured in the start and final conformations, are listed in Table 2. Note that in run E1, not only did one of the nisin chains detach from the pore,

68

Table 3.2: System Setup:Number of Water Molecules inside of the Nisin Pore Assembly for Both Control (Lipid II Missing) and Experiment (where Lipid II Interacts with the Nisin Chains); Data are Shown for the Start and the Final Frame and Their Differenc

Trjectory	Number of water molecules at t = 0	Number of water molecules at t = 200	Change in number of water molecules
C1	1359	523	836
C2	1365	556	809
C3	1335	497	838
Control Avg.	**1353**	**526**	**827**
E1	1340	398	951
E2	1338	796	542
E3	1326	694	632
Experiment Avg.	**1335**	**627**	**708**

but also one of the lipid II molecules lost contact with the pore and drifted away, which explains why this trajectory led to a more significant loss of water molecules and is indicative of the significance of the stereochemistry and presence of the lipid II molecules.

Comparing the initial and final conformations, we see that the number of water molecules at the start is similar in all six runs. However, at the end of the trajectories, the number of water molecules differs substantially between the control and experiment simulations. On average, one finds in the experiment trajectories approximately additional 100 water molecules within the pore, indicating a difference in the pore volume of about 1150 Å^3; at the end of the trajectory, the control has an average pore volume of 6000 Å^3, and the experiment has an average pore volume of 7150 Å^3. Assuming a membrane thickness of about 30 Å and cylinder geometry for the pore, this change in pore volume corresponds to a pore diameter that is around one Ålarger in diameter, i.e., the diameter of the pore in
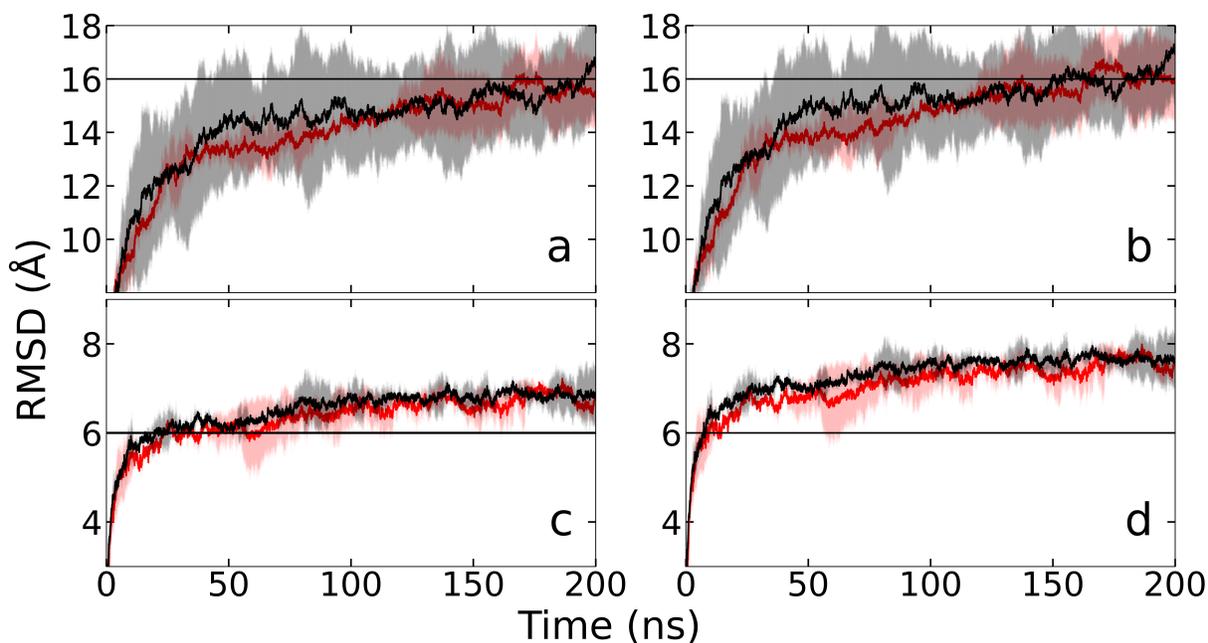
Figure 3.6: Global Root-mean-square deviation (RMSD) with respect to the start conformation as a function of time for the nisin assembly in the presence (Experiment, shown in red) and absence (Control, shown in black) of lipid II, calculated over Cα-atoms (a) and all heavy atoms (b). The corresponding plots for the chain RMSD are shown in (c) and (d). Horizontal lines are drawn to guide the eyes and to emphasize differences between the sub-figures.

the presence of lipid II is 9 Å compared to 8 Å in the control where lipid II is missing.

In order to describe the evolution of the pore with respect to time in more detail, we have

analyzed the average root-mean-square-deviation (RMSD) with respect to the start con-

formation as a function of time (see Figure 3.6). In Figures 3.6a) and 3.6b, the RMSD

is evaluated over the whole pore-forming assembly of nisin chains. We notice that this

global RMSD does not depend on whether we consider for the RMSD calculation only the

backbone Cα-atoms (3.6a) or all heavy atoms (3.6b) of the nisin chains, with differences

in the plots barely noticeable. This indicates that the overall change in the pore geom-

etry is driven by the re-arrangement of the complete nisin chains and not dominated by side-chain motion. When lipid II is present, the RMSD values are slightly lower, especially at the early stages of the simulation, and have a smaller standard deviation. However, after about 125 ns, the RMSD plots for both the control and the experiment plateau at similar values and with comparable standard deviations. The similarity in RMSD is reflected by similar values in the number of inter-chain contacts measured between the nisin molecules: 93(16) in the experiment (the bacterial membrane model containing lipid II) vs 118 (11) in the control. Note that none of these contacts is present in the start conformation; hence, while we find in both cases an increase of about 60 interchain contacts (72 in the presence of lipid II and 56 in the control), the larger number of contacts results from a decay of the pore that allows for movement of the nisin chains.

A different picture is seen in Figures 3.6c and 3.6d, where the RMSD is now evaluated separately for each of the eight chains, with the average over the eight chains (and the three trajectories) shown as a function of time. In these two sub-figures, we see that the chain RMSD values are larger when calculated over all heavy atoms (3.6d) instead of only $C\alpha$-atoms (3.6c), indicating that changes in the side chain orientations contribute strongly to the chain RMSD. Note that the standard deviations for both the lipid II system and the control are smaller than for the global RMSD plots in Figures 3.6a and 3.6b. Unlike the global RMSD, chain RMSD values for the experiment are lower than the control over most of the trajectory, with the difference appearing to be also more pronounced. We remark that while the number of intra-chain contacts (connecting residues within the same chain) stays for both control (about 23 contacts) and in the experiment (about 22 contacts) un-

changed throughout the simulation, this is not true for individual chains, where contacts are both lost and newly formed. Especially the large standard deviations in the chain RMSD for the lipid II system seen between 50 ns and 125 ns (and again 150 ns -175 ns) indicate significant variations in the chain geometries between the independent trajectories. However, all experiment trajectories still have smaller chain RMSD than seen in the trajectories of the control.

The visual inspection of the final conformations of our six trajectories and the estimated change in pore volume indicated that lipid II in the membrane stabilizes the pore-like nisin assembly. The four RMSD plots in Figure 3.6 not only support this observation but also suggest that lipid II stabilizes the nisin pore mainly by preserving the fold of the nisin peptides in the pore, thereby slowing down the decay of the pore. In order to establish the regions in the nisin chain conformations that are affected most by the presence of lipid II, we show in Figure 3.7 the root-means-square fluctuation (RMSF) of residues averaged over all eight chains, a quantity that allows one to identify the regions in the nisin chains that are most flexible and prone to change. The RMSF is evaluated with respect to the start conformations, but we remark that the overall picture does not change when other reference conformations (for instance, such as taken from the mid-points of the trajectories) are chosen (data not shown).

When evaluated over the first 50 ns (Figure 3.7a), the RMSF values in the absence and presence of lipid II differ little and agree within the error bars. This is different for the
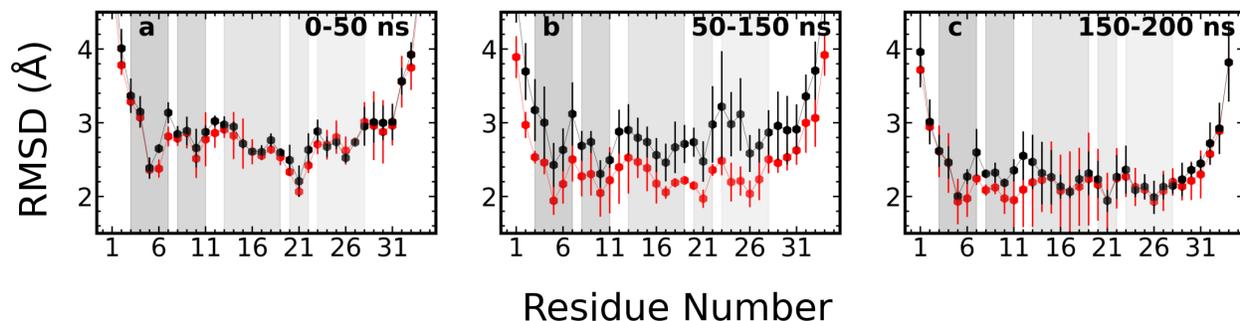
Figure 3.7: Root-mean-square fluctuation (RMSF) of nisin residues (averaged over all eight nisin chains in all three trajectories) in the presence (Experiment, drawn in red) and absence (Control, drawn in black) of lipid II. In (a), we show the values calculated over the first 50 ns; in (b), the values measured between 150-150 ns, and in (c), the ones obtained for the final 50 ns. Shaded areas mark the residue regions corresponding to the five rings A to E.

interval 50ns-150ns (Figure 3.7b), roughly corresponding to the time span, whereas in Figures 3.6c and 3.6d, the chain RMSD has a large standard deviation in the presence of lipid II. Here, we observe a pronounced lowering in the residue-wise fluctuations over the regions of residues 8-28 (which are embedded in the membrane) in the experiment (i.e., when lipid II is present), while there is little change in the values for the control. These differences start at the hinge region and ring D (residues 21-28) and around residues 8-11, i.e., ring B of the nisin chain (see Figure 1a), before later extending over all membrane-embedded residues. Over the last 50 ns (see Figure 3.7c), the differences between the control and the pore in the presence of lipid II decrease as the fluctuations in the control decrease, with the remaining differences again most pronounced for residues 8-13, i.e., ring B and residues 12 and 13 that lie between ring B and C and are important for the flexibility of the chain. Mutation in this residue has been shown to have enhanced antimicrobial activity. [114]

Table 3. Change in the SASA between the Last 50 ns and the Respective Start Conformation[a]

residues	ΔSASA		ΔWASA		ΔLASA		ΔΔSASA	ΔΔWASA
	E	C	E	C	E	C		
1−34	−1.3(1.1)	−2.6(4)	−4.2(4)	−4.2(5)	2.9(1.2)	1.5(9)	1.3(1.1)	−0.01(50)
1−7	0.6(6)	−0.8(2)	−1.6(1)	−1.8(1)	2.1(8)	1.0(5)	1.4(5)	0.2(2)
8−11	−0.1(2)	−0.3(1)	−0.2(3)	−0.2(2)	0.1(6)	−0.1(5)	0.1(2)	0.02(40)
8−28	−1.6(7)	−1.5(4)	−1.8(3)	−2.0(4)	0.2(1)	0.5(9)	−0.1(5)	0.2(1)
12−20	−1.1(4)	−0.78(3)	−1.0(1)	−0.8(1)	−0.1(7)	0.04(60)	−0.3(4)	−0.2(2)
21−28	−0.3(3)	−0.5(2)	−0.6(4)	−0.9(4)	0.3(8)	0.4(8)	0.2(1)	0.3(4)
29−34	−0.3(3)	−0.3(1)	−0.8(2)	−0.3(3)	0.5(7)	0.04(60)	0.02(20)	−0.5(4)

[a]If the area covered by POPE or POPG is excluded from the calculation, we call the resulting quantity the *water-accessible surface area* (WASA). In contrast, when the area exposed to water is excluded from the calculation, we call it the *lipid-accessible surface area* (LASA). Values are given in Å^2 and reported for both in the presence of lipid II (E) and in the absence of lipid II (C). We also list the differences between values measured in the presence and absence of lipid II (ΔΔSASA and ΔΔWASA). Our data are evaluated for different segments of the nisin chains, with residues 8−28 inside the bilayer and residues 1−7 and 29−34 outside the membrane. The standard deviation of the averages is given in parentheses, marking changes for the last digit.

Table 3.3

In order to understand what causes the observed differences in the mobility of the nisin residues when interacting with lipid II, we list in Table 3, the solvent-accessible surface area (SASA) measured for these segments both at the start of the trajectory and averaged over the last 50 ns of a trajectory. The listed SASA values describe the area of nisin chains in the pore that is exposed to the outside (i.e., the expanse of the pore) and not covered by intra-chain contacts, i.e., contacts with other nisin-chains or with lipid II molecules. Hence, it describes the area available for contact with either water molecules or POPE and POPG lipids in the membrane. By excluding the area covered through contact with POPE and POPG lipids, we arrive at the water-exposed area WASA, also listed in Table 3. The difference between the two quantities is the surface area covered by POPE or POPG, named by us LASA and again listed in Table 3.3. Note that we also list for all three quantities the difference between initial values and the averages over the final 50 ns, which quantifies for the segment change over the trajectory.

We see that independent of the absence or presence of lipid II, the pore-forming nisin

molecules lose accessible surface area over the course of the simulations; however, the loss of WASA is more significant than that of SASA, i.e., the surface area that at the start was exposed to water is now exposed to POPE or POPG, pointing to the bilayer. Hence, the difference between SASA and WASA describes the loss of pore geometry and the succeeding mixing of nisin molecules with the bilayer. Over the whole length of the nisin chain, the area covered by POPE and POPG in the presence of lipid II increases by 2.9(1.2) nm^2, while in the control by only 1.5(9) nm^2. The difference arises mainly from residues 1-7, which interact with the lipid II molecules: an additional area of 2.1(8) nm^2 versus 1.0(5) nm^2 in the control. For residues 8-28, which are embedded in the bilayer, the area that is exposed to the bilayer (i.e., POPE and POPG molecules) increases by 0.5(9) nm^2 in the control but only by 0.2(1) nm2 for the experiment, i.e., in presence of lipid II. This indicates that in presence of lipid II the side chains of nisin are more oriented into the pore and therefore interact less with the bilayer, stabilizing in this way the pore.

What causes the increased stability of the nisin pore in the presence of lipid II molecules? One factor may be the contacts between the nisin chains and the lipid II molecules and their reorganization along the trajectory. At the start, the eight nisin chains share a total of 45 contacts with the four lipid II molecules, with each lipid II forming contacts with two neighboring nisin chains, around 5-6 contacts with each chain. These contacts are primarily with N-terminal residues of the nisin chains, about three with the mostly extracellular residues 1-7 (mostly with the first three residues). About two contacts are with residues 8-28, which are located inside of the membrane bilayer. At the end of the trajectory, the four lipid II molecules still form about 38 contacts with the nisin chains; however, each

75

lipid II has now, on average, only about two contacts with the exterior segment (residues 1-7) of a nisin chain but has increased the number of contacts with the interior residues 8-28 of a nisin chain by around 2-3 contacts. These contacts, formed mainly by the head group atoms of lipid II with adjacent nisin molecules, provide additional stabilization to the pore geometry. Note that these contacts involve the pyrophosphate moiety but not the pentapeptide of lipid II; i.e., binding of nisin to lipid II cannot be evaded by mutations in the pentapeptide. This is similar to what is observed for another antimicrobial peptide, teixobactin, [115] and may explain why resistance to nisin is rare. Note also that these contacts coordinate the lipid II head groups to nisin chains. Visual inspection of frames along the trajectories E1, E2, and E3 indicates that the proximity of the flexible and moving lipid II tails constraints the movement of the nisin molecules within the membrane, therefore adding to the stability of the pore. However, as discussed by us later in the following section, this movement of the lipid II tails (with a RMSF of about 6 Å for the final ten carbon atoms) also disturbs and weakens the membrane directly. Interestingly, we did not see that lipid II changes the interaction of Na+ ions with nisin and the lipid bilayer as was observed in previous work. [116] Specifically we did not see any specific change in the binding pattern of the ions between the control and experiment and, therefore, do not believe that such ion-mediated interactions contribute to the stability of the pore.

3.8.2 Nisin - Lipid II Mediated Membrane Integrity

What are the effects of the nisin pore complex on membrane integrity? The change in the surface area of the nisin chains that are exposed to either water or POPE/POPG lipids leads to a change in the number of contacts and hydrogen bonds between water or POPE/POPG molecules with nisin chains. We show this for water molecules in Figures 5a and 5c. For residues 8-28 (i.e., the ones embedded in the interior of the membrane), the number of hydrogen bonds is reduced by 40 in the control (i.e., about five hydrogen bonds per nisin chain) but only by around 30 in the experiment, in presence of lipid II (about four hydrogen bonds per chain). Approximately 16 hydrogen bonds (two per chain) are lost for both systems in the region of residues 12-20. On the other hand, the total number of contacts between water and nisin chains is for residues 8-28 about 40 contacts lower in the presence of lipid II than in the control, with the difference mainly resulting from residues 12-20. This is because in the control, some nisin residues no longer form hydrogen bonds with water molecules, but they stay close enough to the water molecules to form contacts, i.e., the lost hydrogen bonds are now counted as contacts. Nevertheless, contact and hydrogen bond numbers together indicate that the presence of lipid II stabilizes the hydrogen bonds between water molecules and the nisin residues 8-11 and 21-28, which corresponds to ring B and ring D. As a result, the two rings are more likely oriented towards the water, pointing towards the pore interior thus constraining and stabilizing the pore geometry.

Residues in the nisin chains can also form contacts with the POPE and POPG lipids that make up the membrane bilayer. The time evolution of these contacts is shown in Figures 3.8c and fig:hbondsd. The number of contacts between residues 8-28, embedded inside of the membrane, and POPE lipids on average decreased by around 20 contacts for the control group: at the beginning of the trajectory, there were 121 contacts which fell to an average of 104(14) contacts in the last 50 ns of the trajectory. However, in presence of lipid II, the number of contacts between residues 8-28 and POPE lipids on average increased by around 40 contacts; at the beginning of the trajectory there were 93 contacts which increased to an average of 125(4) contacts in the last 50 ns, with almost 30 contacts formed by residues 8-11 (ring B) and 21-28 (ring D), see Figure 1. The corresponding number of contacts between the nisin chains and the more frequent POPG molecules increases for the experiment, i.e., in the presence of lipid II, by about 170 contacts, beginning with 394 contacts at the start of the trajectory and increasing to an average of 568(72) contacts in the last 50 ns. However, the increase is larger in the control, where the number of contacts grows by about 190 contacts, initially having 412 contacts, which increased to an average of 600(60) contacts at the end of the trajectory. Again, most of the newly formed contacts of POPG are with nisin residues 8-11 (ring B) and 21-28 (ring D): about 90 in the experiment and 100 in the control.

Hydrogen bonds between the nisin chains and POPE are rare and do not change in frequency; on average, we observe, in the last 50 ns, 3(1) hydrogen bonds in the control and 4(2) hydrogen bonds in the presence of lipid II, i.e., in the experiment. On the other hand, hydrogen bonds are not only seen, but increase, between the nisin chains and

78
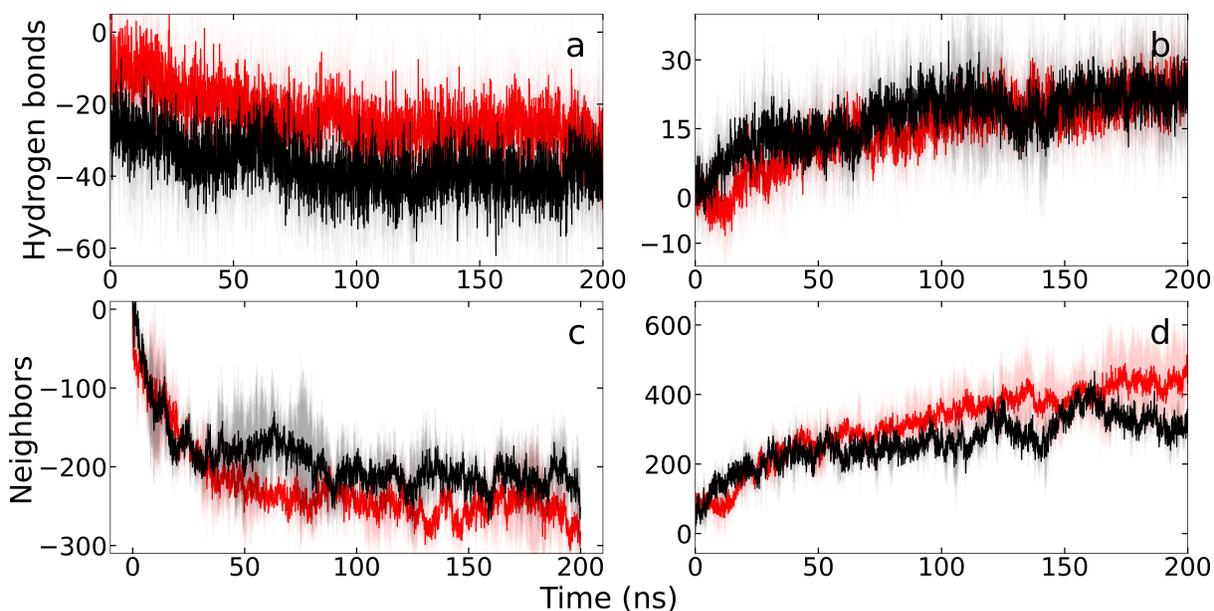
Figure 3.8: The number of hydrogen bonds formed between nisin chains and water molecules located within the pore are shown in (a). Only the residues 8-28, located in the interior of the membrane, are considered. The plots show the difference between the average, taken over the last 50 ns and three trajectories, and the initial values, i.e., a negative value indicates a loss of hydrogen bonds. Similarly, we show in (b) the corresponding number of hydrogen bonds formed with either POPE or POPG lipids. The number of water molecules that do not form hydrogen bonds with nisin residues but are within 4.5 Å to nisin residues is shown in (c), and the corresponding number for contacts between nisin chains and POPE or POPG is shown in (d). Red indicates the experiment, and black the control.

the more numerous POPG: there was an increase of 13 hydrogen bonds for the control (beginning with 13 hydrogen bonds and increasing to 26(5) at end of the trajectory) and an increase of around 9 hydrogen bonds for the experiment (beginning with 17 hydrogen bonds and increasing to 26(4) at the end of the trajectory). Note that about 66% of the new hydrogen bonds are formed between POPG and residues 8-11 or 21-28: 9 in the control and 6 in the experiment.

Hence, both the number of hydrogen bonds and the total number of contacts between the nisin chains and the POPE/POPG lipids increase over the lengths of the trajectories, but less so in presence of lipid II. This is consistent with the stronger decay of the pore in the control, leading to more nisin residues coming in contact with the membrane lipids while reducing the number of hydrogen bonds formed with water. Most of this effect comes from rings B (residues 8-11) and D (residues 21-28). The contact of these residues in the nisin chains with lipid II encourages an orientation of these rings where they face toward the pore, which stabilizes hydrogen bonds with water molecules and reduces the chance for contacts or hydrogen bonds to form with the membrane constituents, i.e., the POPG/POPE lipids. This explains the differences between our experiments and the control runs.

This growing interaction of the nisin chains with POPE and POPG molecules, as seen throughout the respective trajectories, disturbs the membrane. At the start of the simulations, there are no significant differences in the distribution of lipids around the pore between the control and the experiment. This is different at the end of the simulations,

when, in the vicinity of the pore (defined as the shell with a radius between 25 Å and 40 Å from the center of the pore), the number of POPE and POPG molecules increases from 113 to 141 molecules in the control, but only from 109 to 134 molecules in the experiment, i.e., in presence of lipid II. Hence, while at the start, the density of lipids was similar in the control and experiment, it became lower in the experiment than in the control as the simulation evolved. This is because the pore decays in both experiment and control over the course of the trajectories, allowing for an inflow of POPG and POPE lipids. This influx is more pronounced in the control than in the experiment where presence of lipid II limits this effect. Correspondingly, there is a difference in the thickness of the membrane in the vicinity of the nisin assembly: a membrane thickness of 31.8(8) Å is measured in the control but only 27.5(1) Å in the experiment. For distances larger than 40 Å the densities and thickness of the bilayer (on average 30.4(1) Å over the last 50 ns in both systems) are again similar in experiment and control and changed little between the start and finish of the trajectories. This indicates that the presence of lipid II not only stabilizes the nisin pore but also leads locally to a thinning of the bilayer.

One way to characterize the behavior of the membrane is by the area per lipid (APL), This quantity is defined as the area of a selected segment ("box") of the membrane divided by the number of lipid molecules in the box. As this quantity changes in response to the presence of proteins and other non-lipids, we have measured it for both the control and the experiment trajectories as a function of distance to the center of the pore. The APL was evaluated separately for the two layers of the membrane as the lipid II molecules are located in the layer close to the extracellular space and the N-terminus of the nisin

chains. In our notation, this is layer 1. In bulk, the difference between the two layers is neglectable in both control and experiment; however, differences are seen near the pore. Comparing the averages taken over the last 50ns and all three trajectories of each system, we notice that in the shell with a radius between 25 Å and 40 Å from the center of the pore, the APL is in both layers larger in the presence of lipid II than in control, and in both cases are the APL values larger in layer 1 than in layer 2: in layer 1, we have in the control a value of 54.1(7) $Å^2$ vs 56.0(6) $Å^2$ in the experiment, while in layer 2 the corresponding numbers are 47.2(6) $Å^2$ for the control and 50.5(6) $Å^2$ in the experiment. The APL values for layer 1, in both the control and the experiment, remain essentially unchanged from the beginning to the end of the simulation, 53.7 $Å^2$ and 55.6 $Å^2$, respectively. However, the APL values decreased for layer 2 in both absence and presence of lipid II: the start value for the control is 54.7 $Å^2$ and 56.6 $Å^2$ in the experiment. These differences of the time evolution of APL values show that the presence of the nisin chains distorts layer 2 more than layer 1 but that this effect is lesser in experiment, i.e., is restrained by presence of lipid II; in the control is in layer 2 the APL by 6.9(7) $Å^2$ smaller than in layer 1, but the difference is in the experiment only -5.5(6) $Å^2$. Note that at the start, the corresponding differences are about 1 $Å^2$ in both the control and in the presence of lipid II.

Another way to characterize the membrane behavior is through the head group angles of the lipids in the membrane. For the head group atoms, this orientation can be described by head group angles defined such that for zero angle, the lipid head is parallel to the membrane surface, and for positive angles points away from the membrane. For the less frequent POPE in the vicinity of the pore (in a shell with a radius between 25 Åand

82

40 Å), we find that there are no significant differences in the distribution of head group angles when comparing the control to the experiment. Regarding the POPG headgroup angles, there is only a slight shift toward more negative headgroup angles. In all cases, the averages are compatible with zero, which suggests that the changes in membrane properties more likely result from the orientations in the lipid tails. A measure for this quantity is the order parameter, -SCD, which characterizes the distribution of angles for the carbon atoms in the POPE and POPG lipids, i.e., it describes the fluidity of the lipids in a membrane segment. By its definition, this quantity takes values between -1 (disorder) and 0.5 (completely ordered). [112] Values of this quantity, as measured and averaged over the final 50 ns, differ in both layers little for POPG and POPE lipids in bulk or far away from the nisin assembly. However, when measured in a shell of radius between 25-40 Å i.e., close to the pore, the order parameter values are for layer 2 in the presence of lipid II higher than in the control. However, for layer 1, where the peptide chains of lipid II reside, the values are lower than in the control. This can be seen in Figure 3.9, where we show either for each layer the differences between order parameters measured in experiment trajectories and the ones measured in control runs (in Figure 6c for POPE and in 3.9d for POPG), or the differences between order parameter values measured in the two layers (in Figure 3.9a for POPE and in 3.9b for POPG) in either experiment or control. Comparing the order parameter difference between the two layers, we find that in the experiment, the order parameter values are for both POPE and POPG higher in layer 1 than in layer 2, while the opposite behavior is seen for the control. This indicates that the presence of lipid II leads to more disordered chains of the bilayer lipids POPE and POPG in layer 2. In contrast, its lipid tails cause more ordering in the POPE and POPG lipid tails of layer 1,
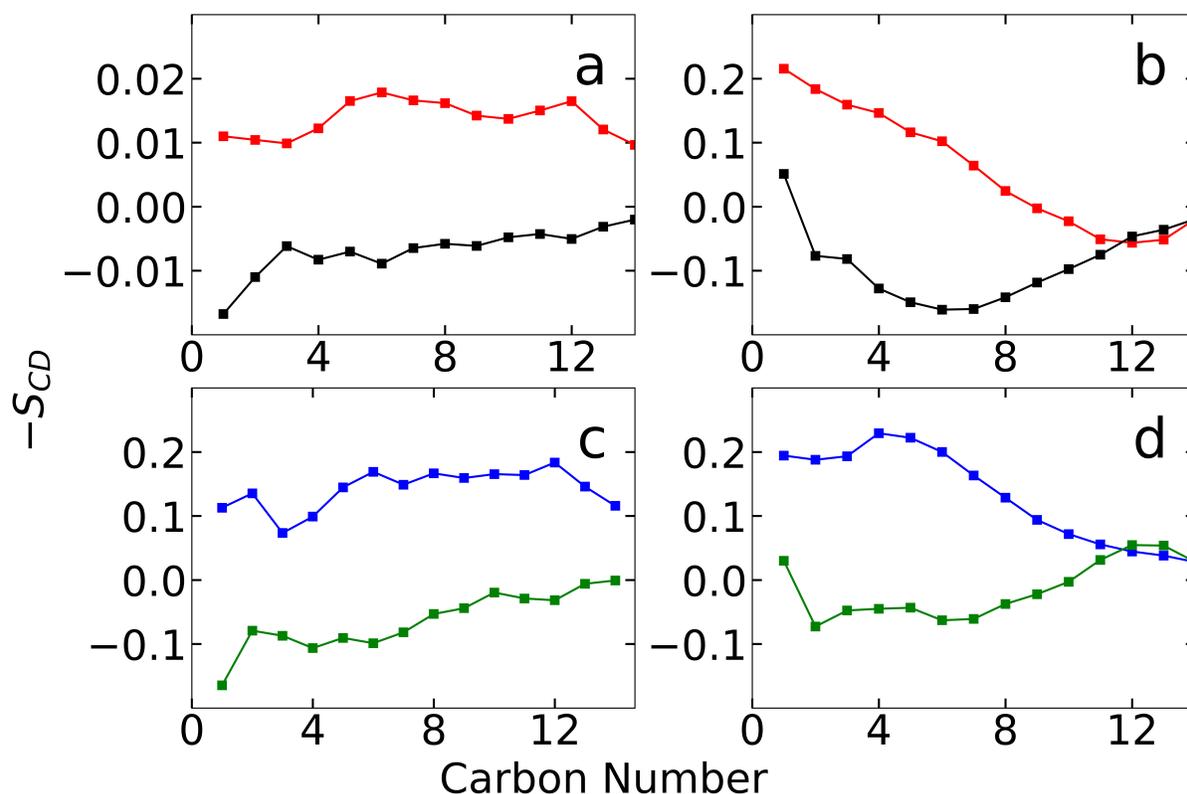
83

Figure 3.9: Tail group order parameter $-S_{CD}$, averaged over the final 50 ns of three trajectories and measured separately for POPE and POPG lipids and both layers of the membrane. In (a), we show the differences between the values measured for POPE in the two layers, with the results from simulations in the presence of lipid II drawn in red and that of the control in black. The corresponding plots for POPG are shown in (b). On the other hand, the differences between the order parameter values measured in the experiment and control are shown in (c) for POPE and in (d) for POPG, where we plot in blue the results for layer 1, and in green the results for layer 2.

i.e., the presence of lipid II leads to a more fluid layer 2 and a more viscous, gel-like layer 1. Such altering of the local membrane environment by the presence of lipid II has also been observed in earlier work [116], but in our case, is connected with the presence of nisin chains. Hence, similar to what is seen for teixobactin, [115] the interactions between nisin and lipid II reduce the integrity of the bilayer.

3.9 Conclusion

Although nisin is non-toxic to humans and has broad and strong antimicrobial activity, its poor solubility and stability limit its use as an antibiotic. Considerable efforts have been put into deriving nisin analogs that overcome these limitations, but equally crucial for the development of future antibiotics is obtaining an in-depth understanding of the mechanisms for nisin's antimicrobial effectiveness. In order to gain such an understanding, we have conducted large-scale molecular dynamics simulations of a nisin assembly, designed to be consistent with experimental data, that was inserted in a gram-positive bacterial cell membrane model. As a control, we used the same model of a nisin pore in a bilayer but removed the lipid II molecules. The purpose of our simulations was to study the effect of the bacteria-specific lipid II on the nisin pore stability, and to probe how in the absence and presence of lipid II nisin chains change the bilayer in our model of a bacterial membrane.

We find that contacts formed between the lipid II head group atoms and nisin residues stabilize the geometry of the nisin chains and, as a consequence, the integrity of the pore assembly. Additional stabilization is provided by the proximity of the flexible and moving lipid II tails that constrain the movement of the nisin molecules within the membrane. The pore decays in both absence and presence of lipid II, but more so in the control where lipid II is not present. It should be noted that the lifetime of a pore formed by nisin chains will likely be larger in a bacterial cell than seen in our model which did not account for a

85

flow of water molecules through the pore that would keep the pore open.

The assembly of nisin chains interacting with lipid II chains also changes the property and behavior of the membrane, leading in the neighborhood of the nisin chains to more disordered lipid chains in layer 2. In contrast, its lipid tails cause more ordering in the POPE and POPG lipid tails of layer 1, where the lipid II chains reside. Of critical importance for these processes are residues 8-11 and 21-28, i.e., rings B and D of the nisin chains. The presence of lipid II leads in both layers locally to an increase in the Area per Lipid (APL), i.e., the density of lipids in the bilayer, with the difference between the two layers much more prominent than in the control. The effect of the increased density is twofold: a local thinning of the bilayer and a spread in the viscosity of the two layers, with layer 2 becoming more fluid and layer 1 (where the lipid II are located) more gel-like. Both effects reduce the stability of the bacterial membrane.

In conclusion, our computational study shows that interaction with lipid II enhances the stability of pores formed by nisin chains in the membranes of gram-positive bacteria, thus increasing the chance for cell leakage. Interactions between nisin chains and lipid II molecules also cause local changes in membrane viscosity and thickness that further destabilize the bacterial membrane. We propose that it is the combination of the two effects that makes nisin an effective antimicrobial agent.

Chapter 4

Permeation of Nisin
into Biological Membranes

In addition to bacterial cells, nisin has also been shown to have inhibitory properties against cancer (Table 4.1). In contrast to bacteria, cancer cells lack lipid II molecules. This suggests that in a cancer cells nisin binds to other, non-lipid II, molecules can interact with the membrane in a manner that compromises the integrity of the membrane. Here the possible mechanisms by which nisin can cause the death of a non-bacterial-membranes are addressed. In addition, simulating nisin in membrane environments allows for a more general inquiry into the physical dynamics of membranes. This is not well understood as of now but is a promising area for future study. This chapter describes ongoing and unfinished research. A research plan will be presented, and while some simulations have been completed our analysis has just started.

4.1 Introduction

Cells function to separate the internal components of the cell from that of its surrounding environments.This organization and serperation allow the processes of a cell to be car-

ried out. Proteins that cross membrane bilayers play a role in the homeostasis of cells; from cell signaling to membrane dynamics, they help regulate the membrane function and cellular processes. Conversely, permeation into cell membranes is often a needed step for bioactive molecules as their targets are intracellular. Investigation into how proteins or small molecules, natural or synthetic, permeate into membranes is considered necessary for further drug development.

Nisin is produced intracellularly, then moves outside of the host membrane, before re-entering other bilayers of different cells. The protein is capable of crossing membranes, yet the same protein behavior seems to differ depending on the membrane environment. Nisin is highly effective against bacterial cells, it shows inhibitory properties at the nanomolar concentrations. In cancer membranes, the protein can inhibit growth, but the effectiveness is not as pronounced by itself (combined with other drugs, the effect on cancer cells is again more pronounced). It is therefore of interest how nisin interacts with different membranes.

In bacteria membranes the nisin molecules bind to lipid II which gives the protein the stability needed for its C-terminus to move into the membrane. After assembly with other nisin molecules, a pore is then formed. While this mechanism is effective at low concentrations, indicating that it works very well, cancer cells lack a lipid II molecule to bind to. In addition to not having lipid II molecules, it is not indicated in literature that pores are formed in cancer membranes. However, there is sufficient evidence that nisin interacts with different cancer types and inhibits cellular growth. The question is how.

Were et al., in a study looking at antimicrobial behavior, show that nisin was more effective at a certain ratio of lipids. [117] This suggests that nisin shows preferential interaction with lipids in cells and that the components of the bilayer impact the behavior of the protein. The lipid types and specific lipid ratios are relevant.

Sadri et al. point out that nisin induces apoptosis in cancer cells by first initiating mitochondrial dysfunction and oxidative stress. [118] Similarly, Lewies et al. also indicate that nisin perturbs the metabolic pathways of melanoma cells which also results in apoptosis. [119] Nisin may perturb the metabolic pathways of cancer cells.

While mitochondrial and cancer membranes vary in composition, for nisin to perturb the mitochondrial membrane directly, it first needs to enter the cell. Probing how nisin interacts with membranes, regardless, is still a relevant question. Table 4.1 indicates other studies where nisin has been shown to impede tumor growth and highlights the potential of the protein.

Nisin is considered safe for human consumption. It is curious that this protein has the ability to terminate bacterial and cancer cells yet seems to have no negative impact on healthy human cells. This is also worth looking into.

Table 4.1: Types of Cancers to which Nisin has Inhibitory Properties

Cancer Type	Reference
Lung	[120]
Melanoma	[119]
Head and Neck Squamous Sarcinoma	[121]
Breast	[67, 122]
Colon	[67]
Colorectal	[123]
Liver	[124, 125]
Cervical	[118]

4.2 Questions of Interest

Mechanism of Protein Permeation: The exact mechanism of how nisin enters a membrane is only surmised. The current hypothesis is that nisin interacts with the surface of the membrane which allows binding to an 'anchor' to occur. The peptide uses the anchor to pull itself into the membrane. In the case of bacterial members, the anchor is a lipid II molecule and in the case of cancer, it is speculated that it is another type of lipid molecule. The importance of this step is significant as the peptide has to enter the bilayer before a pore can be formed and the pore formation is what leads to the termination of the toxic cell.

Membrane Dynamics: Membrane proteins, those that permanently reside in the bilayer or those that are temporarily associated with the bilayer, are both directly affected by the dynamics of the membrane as the membrane is their immediate environment. Recently, computational advancements in terms of simulation time have made membrane simulations possible; however, the physical dynamics and movements of lipids in the membranes

themselves have not been widely studied and therefore are largely not well understood. The influence of the dynamics of the membrane on proteins comes in the form of multidimensional movements acting on the protein. The independent dynamics of membranes should be investigated to properly understand the behavior of proteins in a membrane. To gain more understanding of the membrane behavior, the most fundamental properties cannot be overlooked, but rather need to be addressed directly. Questions such as what effect on dynamics does the size of the system have and how, in turn, does this affect the proteins in the bilayer?

4.3   Guiding Hypothesis

It is expected that the N-terminus of the peptide will bind to a lipid in the membrane bilayer: lipid II for the bacterial membrane and DOPE/DOPS for the cancer membrane, and not all for the eukaryotic membrane. The attraction to the lipid types is expected to be a result of charge; the positive residues in the N-terminus of the nisin protein should be attracted to the negative charges of the lipid II in bacterial cells, or DOPE or DOPS in cancer cells. When the N-terminus of the protein is bound to a lipid molecule, it is expected that the protein will lay vertically across the membrane allowing the C-terminus of the peptide to move into the bilayer. Healthy human cells do not have, in their extracellular layer, the same abundance of negatively charged lipids, we do not expect to see this in our model of healthy cells.

4.4 Approach

4.4.1 The Bilayer Models

This study will focus on three types of bilayers: a bacterial bilayer, a cancer bilayer, and a eukaryotic (or healthy) bilayer. All bilayers were generated with GARMM-GUI Membrane Builder.

The Bacterial Bilayer — The bacterial membrane is modeled by a 3:1 ratio of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE). [90, 91] This ratio is chosen as it models a Gram-positive bacteria bilayer.

The Cancer Bilayer — The cancer will be modeled with sphingomyelin (SSM), 3-palmitoyl-2-oleyl-D-glycero-1-phosphatidylcholine (POPC), 1,2-dioleoyl-sn-glycero-3-phosphoethanolamine (DOPE) 1,2-dioleoyl-sn-glycero-3-phospho-L-serine (DOPS) and Cholesterol (CHL1) lipid types. The ratios of these lipids are presented in table 4.2. The lipid types and ratios are given by Rivel et al. and are derived from experimental data. [126]

The Healthy Cell Bilayer — Rivel et al. also suggests a normal membrane which takes into account the asymmetry that is present in them. [126] The same types of lipids used to model the cancer bilayer will again be used but the ratios of the inner and outer leaflet

Table 4.2: Ratio of Lipids in a Cancer Membrane

Lipid type	Percentage in Outer Bilayer	Percentage in Inner Bilayer
SSM	18	18
POPC	20	20
DOPE	20	20
DOPS	10	10
CHL1	32	32

 Lipid Nomenclature
SSM-sphingomyelin
POPC-3-palmitoyl-2-oleyl-D-glycero-1-phosphatidylcholine
DOPE-1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
DOPS-1,2-dioleoyl-sn-glycero-3-phospho-L-serine
CHL1-Cholesterol

Table 4.3: Ratio of Lipids in a Eukaryotic Membrane

Lipid type	Percentage in Outer Bilayer	Percentage in Inner Bilayer
SSM	28	8
POPC	30	9
DOPE	9	30
DOPS	0	20
CHL1	33	33

 Lipid Nomenclature
SSM-sphingomyelin
POPC-3-palmitoyl-2-oleyl-D-glycero-1-phosphatidylcholine
DOPE-1,2-dioleoyl-sn-glycero-3-phosphoethanolamine
DOPS-1,2-dioleoyl-sn-glycero-3-phospho-L-serine
CHL1-Cholesterol

will change. This is shown in 4.3

4.4.2   Lipid Preference

To determine what lipid molecules the protein has preference for in the cancer and normal

membrane a 200 ns simulation was carried out. Here, the protein was placed above the

bilayer and allowed to interact freely. From the resulting trajectory the contact analysis

between the N-terminus of the protein and bilayer was analyzed. It was assumed that the lipid that has the most interactions with the N-terminus of nisin is the preferential lipid.

The N-terminal residues (1-10) of nisin interacted more with the DOPE and DOPS lipid molecules. Specifically, residue 8 had the most interaction with DOPE, and residues 1,3-4,6 interacted the most with DOPS. To see if the presence of ions had an effect on the number of interactions, a similar simulation was carried out (this time for 10 ns), however, the system was neutralized. The ratio of interactions was the same and therefore it was concluded that the preference of nisin was to DOPE and DOPS molecules. Not these comprise 20 and 10% of the cancer bilayer respectively.

4.4.3 Simulations

Bilayer Only

Each membrane bilayer will be simulated independently for 1 μs. This long equilibrated simulation of the membrane is particularly important for lipid motion as it provides a membrane structure where flip-flops of the bilayer lipids are less likely to occur. Flip-flops in the lipids can introduce pore formation effects and bias the simulation. Explicitly, the membrane characteristics are important to this study because the central question of how nisin interacts with the membrane and therefore a well-equilibrated membrane is significantly important for reliable results.
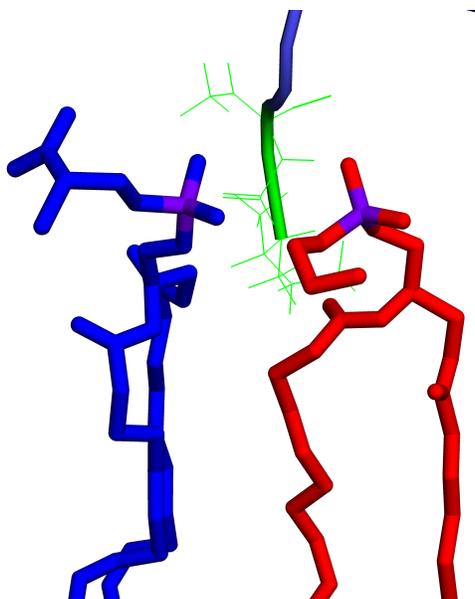
Figure 4.1: The first two residues of nisin (shown in green) have been placed next to a DOPS (colored in blue) and DOPE (colored in red).

Bilayer and Protein Because of time scales, nisin will be placed next to the lipids which it displays a preference for. It is assumed that the first residues of the protein are what initially bind to the lipid molecules. To mimic this, nisin will be placed next to the preferential lipid at around 2 Å deep into the membrane; for the cancer simulation nisin will be placed in between a DOPS and a DOPE molecule (see figure4.1) and in the bacterial membrane nisin will be placed next to a lipid II molecule.

It is thought that the configuration of nisin, with respect to the membrane, just before entry is horizontal on the bilayer surface and bound to an anchor lipid. A similar result, the formation of a pore, occurs from nisin interacting with a bacterial and cancer membrane respectively. Therefore, both will be independently simulated in this aim.

The orientation of the nisin in the simulation will determine if the simulation was suc-

cessful regarding the initial question. If the peptide moves into the bilayer, analysis can be performed on this nisin and membrane interaction.

4.5 Expected Results

It is expected that the N-terminus of the peptide will interact with the respective lipid causing the body of the protein to move into the bilayer. From the trajectory of the simulation the following information with be obtained and will be analyzed through GROMACS, Membrainy, and MDAnalysis:

- Reside orientation of the peptide as it enters the bilayer: this gives information as to what residues of the protein are important and why.

- Orientation of lipid anchor with respect to nisin molecules and membrane lipids: this gives information about the binding mechanism and what is unique about it.

- Acyl chain order parameters of the membrane: Shows how the respective membranes interact with a nisin lipid anchor complex.

- Membrane Dynamics: this will shed light on how the protein changes the behavior of respective membranes.

4.6   Current State of Simulations and Results

The lipid types for which nisin has a preference in the cancer bilayer have been determined. Initial simulation of this configuration has been carried out to 70ns where visual inspection shows that the peptide is integrating into the membrane (see figure 4.2 and 4.3). The protiins movement into the membrane is still being simulated. Further simulations, analysis, and comparison still need to be completed.

Because of the asymmetry of the eukaryotic membrane flat bottom potential needs to be applied to the membrane for equilibrations. This is being carried out and a 1 microsecond trajectory will be the result.
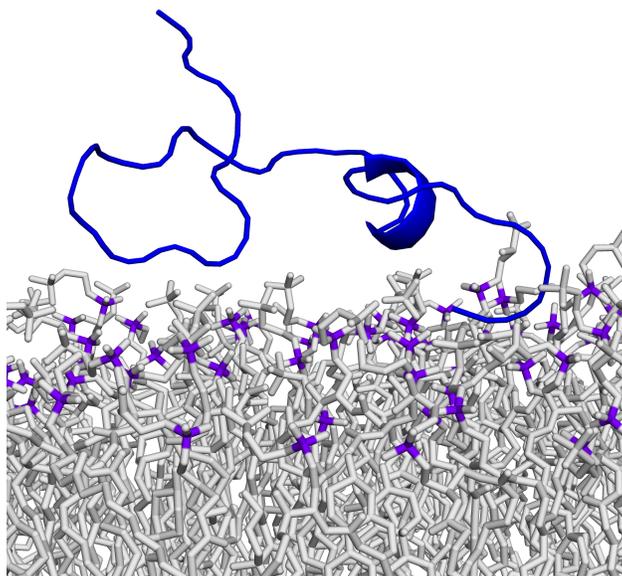
Figure 4.2: The starting configuration of nisin. The first two amino acids have been placed in close proximity to a DOPS and an DOPE lipid molecule.
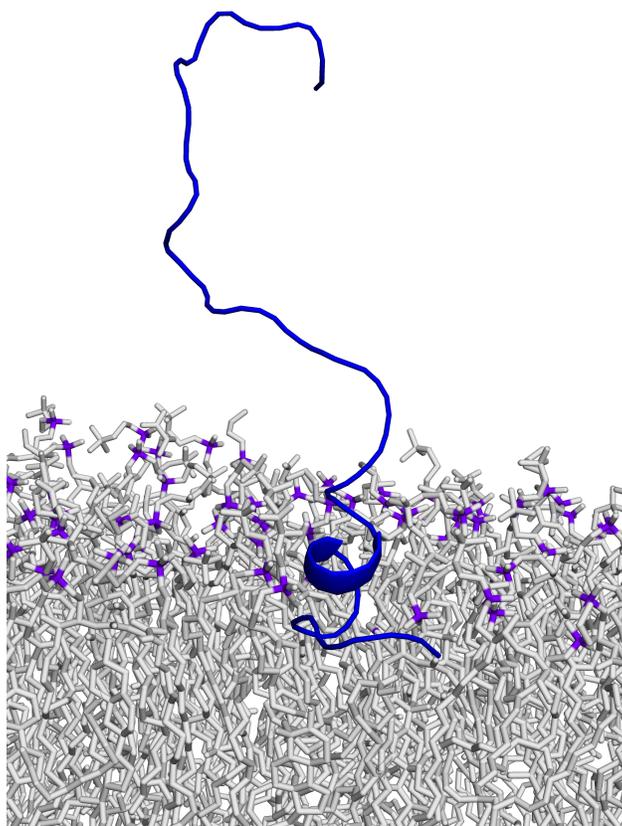

Figure 4.3: The permeation of nisin into the cancer membrane is seen. At 70 ns the first 10 residues of the N-terminus have permeated into the bilayer.

Chapter 5

Summary and Outlook

This final chapter summarizes the research presented in this dissertation, outlines possible extensions, and discusses the significance of potential applications.

In my doctoral research, I have used computer simulations to study transitions and interactions of proteins. While the considered proteins are of biological and medical relevance, the focus was on fundamental questions. For the first system, I used an advanced sampling technique, Replica Exchange with Tunneling, to study the conversion of the protein Serum amyloid A (SAA) from the state that is associated with normal physiological function to a misfolded state that results in amyloidosis. From the simulation a Free Energy Landscape was derived, and in turn, a possible pathway between the two conformational states was produced. From this pathway it became clear that transition depends on the unfolding of the N-terminus of the protein. The results are published in: *F. Yasar, M.S. Sheridan, and U.H.E. Hansmann,* Interconversion Between Serum Amyloid A Native and Fibril Conformations. *ACS Omega, 7 (2022), 12186* [6]

Two possible future extensions of this work come to mind. First, while RET is limited

by the size of the system that can be studied, there is a need, particularly in the pharmaceutical sciences, to generate possible configurations of small molecules. Because RET is designed to sample the landscape of proteins exhaustively, RET is well suited for this purpose, generating many diverse configurations. These have the additional advantage of being physics-based models. Secondly, the function of SAA is often regulated by post-translational modifications. Such modifications of SAA have been not yet studied. RET may also allow simulating SAA in different solvent environments, say in higher concentrations of Ca^{2+} where it is known that this environment changes the secondary structure.

As a second system, I have studied nisin, a naturally occurring antimicrobial peptide that is produced by certain bacteria to kill other non-host bacteria. Because it is effective at micromolar concentrations and safe for human consumption, and is effective against antibiotic-resistance gram-positive bacteria, nisin is considered a potential candidate for future antimicrobial drugs.

The likely mechanism of nisin's antimicrobial avtivity is forming pores through the cell membrane, leading to cell death. In this study a preformed pore complex, consisting of lipid II and nisin molecules and reflecting all known experimental structural data was created and simulated. Analysis of the data indicates that residues 8-28 of nisin are important for the stability of the pore, that in the presence of lipid II the pore is more stable, and that the integrity of the membrane is further reduced by local changes in thickness and viscosity. Our results are published in: *M.S. Sheridan, P. Pandey, and U.H.E. Hansmann,* In Bacterial Membranes Lipid II Changes the Stability of Pores Formed by the Antimicrobial

Peptide Nisin, *J. Phys. Chem. B, 128(2024)* [7].

Addressing how the pore complex behaves in a bilayer is only the initial step in understanding how nisin induces cell death. Information about how the proteins initially form a pore as well as how the protein enters the membrane is also needed. These questions are addressed by us in ongoing work where we compare the mechanisms of pore formations in bacterial cells with that of cancer cells, and also ask why nisin is not negatively affecting healthy human cells. We have been able to identify what lipid molecules in the bilayer of a cancer cell the protein nisin is most attracted to. From this, a simulation has been created where the permeation of the protein can be seen. Simulations with a similar initial configuration, nisin close to a preferred lipid in a membrane, remain to be repeated in bacteria and eukaryotic bilayers (this is ongoing). The simulations will seek to compare how nisin permeates into different membranes and the membrane dynamics of the respective systems.

Future work may also extend to the analysis other lantibiotics or the pore formation by A-β peptides. Pore formation may be one of the reasons for the cell toxicity of A-β peptides that contributes to Alzheimer's disease. Another possible extension of my work would be a more detailed investigation into the dynamics of membranes and membrane composition effects membrane dynamics.

Bibliography

[1] Ken A Dill, S Banu Ozkan, M Scott Shell, and Thomas R Weikl. The protein folding problem. *Annu. Rev. Biophys.*, 37:289–316, 2008.

[2] Miguel A González. Force fields and molecular dynamics simulations. *École thématique de la Société Française de la Neutronique*, 12:169–200, 2011.

[3] Nawaf Bou-Rabee. Time integrators for molecular dynamics. *Entropy*, 16(1):138–162, 2013.

[4] Xiao Zhu, Pedro EM Lopes, and Alexander D MacKerell Jr. Recent developments and applications of the charmm force fields. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2(1):167–185, 2012.

[5] Olgun Guvench and Alexander D MacKerell. Comparison of protein force fields for molecular dynamics simulations. *Molecular modeling of proteins*, pages 63–88, 2008.

[6] Fatih Yasar, Miranda S Sheridan, and Ulrich HE Hansmann. Interconversion between serum amyloid a native and fibril conformations. *ACS omega*, 7(14):12186–12192, 2022.

[7] Miranda S Sheridan, Preeti Pandey, and Ulrich HE Hansmann. In bacterial membranes lipid ii changes the stability of pores formed by the antimicrobial peptide nisin. *The Journal of Physical Chemistry B*, 2024.

[8] Vladimir N Uversky and Alexei V Finkelstein. Life in phases: Intra-and intermolecular phase transitions in protein solutions. *Biomolecules*, 9/Doc(12):842, 2019.

[9] Rakesh Trivedi and Hampapathalu Adimurthy Nagarajaram. Intrinsically disordered proteins: An overview. *International journal of molecular sciences*, 23(22):14050, 2022.

[10] Tuomas PJ Knowles, Michele Vendruscolo, and Christopher M Dobson. The amyloid state and its association with protein misfolding diseases. *Nature reviews Molecular cell biology*, 15(6):384–396, 2014.

[11] Iva Sorić Hosman, Ivanka Kos, and Lovro Lamot. Serum amyloid a in inflammatory rheumatic diseases: a compendious review of a renowned biomarker. *Frontiers in Immunology*, 11:631299, 2021.

[12] Laura J den Hartigh, Karolline S May, Xue-Song Zhang, Alan Chait, and Martin J Blaser. Serum amyloid a and metabolic disease: evidence for a critical role in chronic inflammatory conditions. *Frontiers in Cardiovascular Medicine*, 10:1197432, 2023.

[13] Olga Gursky. Structural basis for vital function and malfunction of serum amyloid a: an acute-phase protein that wears hydrophobicity on its sleeve. *Current atherosclerosis reports*, 22:1–11, 2020.

[14] Nancy R Webb. High-density lipoproteins and serum amyloid a (saa). *Current Atherosclerosis Reports*, 23:1–8, 2021.

[15] Limin Wang, Hilal A Lashuel, Thomas Walz, and Wilfredo Colón. Murine apolipoprotein serum amyloid a in solution forms a hexamer containing a central channel. *Proceedings of the National Academy of Sciences*, 99(25):15947–15952, 2002.

[16] Robert A Kyle. Amyloidosis: a convoluted story. *British journal of haematology*, 114(3):529–538, 2001.

[17] Jinghua Lu, Yadong Yu, Iowis Zhu, Yifan Cheng, and Peter D Sun. Structural mechanism of serum amyloid a-mediated inflammatory amyloidosis. *Proceedings of the National Academy of Sciences*, 111(14):5189–5194, 2014.

[18] Zaida L Almeida and Rui MM Brito. Structure and aggregation mechanisms in amyloids. *Molecules*, 25(5):1195, 2020.

[19] Diego Real de Asúa, Ramón Costa, Jose María Galván, María Teresa Filigheddu, Davinia Trujillo, and Julen Cadiñanos. Systemic aa amyloidosis: epidemiology, diagnosis, and management. *Clinical epidemiology*, pages 369–377, 2014.

[20] Thomas CT Michaels, Daoyuan Qian, Anđela Šarić, Michele Vendruscolo, Sara Linse, and Tuomas PJ Knowles. Amyloid formation as a protein phase transition. *Nature Reviews Physics*, 5(7):379–397, 2023.

[21] S Karthika, TK Radhakrishnan, and P Kalaichelvi. A review of classical and non-classical nucleation theories. *Crystal Growth & Design*, 16(11):6663–6681, 2016.

[22] Sushma Subedi, Santanu Sasidharan, Niharika Nag, Prakash Saudagar, and Timir Tripathi. Amyloid cross-seeding: Mechanism, implication, and inhibition. *Molecules*, 27(6):1776, 2022.

[23] Margaret Sunde, Louise C Serpell, Mark Bartlam, Paul E Fraser, Mark B Pepys, and Colin CF Blake. Common core structure of amyloid fibrils by synchrotron x-ray diffraction. *Journal of molecular biology*, 273(3):729–739, 1997.

[24] William D McCubbin, CM Kay, S Narindrasorasak, and R Kisilevsky. Circular-dichroism studies on two murine serum amyloid a proteins. *Biochemical Journal*, 256(3):775–783, 1988.

[25] Yi Isaac Yang, Qiang Shao, Jun Zhang, Lijiang Yang, and Yi Qin Gao. Enhanced sampling in molecular dynamics. *The Journal of Chemical Physics*, 151(7), 2019.

[26] Nathan A. Bernhardt, Wenhui Xi, Wei Wang, and Ulrich H. E. Hansmann. Simulating protein fold switching by replica exchange with tunneling. *Journal of Chemical Theory and Computation*, 12(11):5656–5666, NOV 2016.

[27] Fatih Yasar, Nathan A. Bernhardt, and Ulrich H. E. Hansmann. Replica-exchange-with-tunneling for fast exploration of protein landscapes. *Journal of Chemical Physics*, 143(22), DEC 14 2015.

[28] Huiling Zhang, Wenhui Xi, Ulrich H. E. Hansmann, and Yanjie Wei. Fibril-barrel transitions in cylindrin amyloids. *Journal of Chemical Theory and Computation*, 13(8):3936–3944, AUG 2017.

[29] Nathan A. Bernhardt and Ulrich H. E. Hansmann. Multifunnel landscape of the fold-switching protein rfah-ctd. *Journal of Physcial Chemistry B*, 122(5):1600–1607, FEB 8 2018.

[30] H Fukunishi, O Watanabe, and S Takada. On the hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *Journal of Chemical Physics*, 116(20):9058–9067, MAY 22 2002.

[31] W Kwak and UHE Hansmann. Efficient sampling of protein structures by model hopping. *PHYSICAL REVIEW LETTERS*, 95(13), SEP 23 2005.

[32] UHE Hansmann. Parallel tempering algorithm for conformational studies of biological molecules. *CHEMICAL PHYSICS LETTERS*, 281(1-3):140–150, DEC 19 1997.

[33] Robert B. Best, Xiao Zhu, Jihyun Shim, Pedro E. M. Lopes, Jeetain Mittal, Michael Feig, and Alexander D. MacKerell, Jr. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ÏȨ, ÏĹ and side-chain ÏǦ$_1$ and ÏǦ$_2$ dihedral angles. *Journal of Chemical Theory and Computation*, 8(9):3257–3273, SEP 2012.

[34] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.

[35] Falk Liberta, Sarah Loerch, Matthies Rennegarbege, Angelika Schierhorn, Per Westermark, Gunilla T. Westermark, Bouke P. C. Hazenberg, Nikolaus Grigorieff, Marcus Faendrich, and Matthias Schmidt. Cryo-em fibril structures from systemic aa amyloidosis reveal the species complementarity of pathological amyloids. *Nature communications*, 10, MAR 7 2019.

[36] EF Pettersen, TD Goddard, CC Huang, GS Couch, DM Greenblatt, EC Meng, and TE Ferrin. Ucsf chimera - a visualization system for exploratory research and analysis. *JOURNAL OF COMPUTATIONAL CHEMISTRY*, 25(13):1605–1612, OCT 2004.

[37] Jeffrey K. Noel, Paul C. Whitford, Karissa Y. Sanbonmatsu, and Jose N. Onuchic. Smog@ctbp: simplified deployment of structure-based models in gromacs. *NUCLEIC ACIDS RESEARCH*, 38(2):W657–W661, JUL 2010.

[38] Weihong Zhang and Jianhan Chen. Accelerate sampling in atomistic energy landscapes using topology-based coarse-grained models. *Journal of Chemical Theory and Computation*, 10(3):918–923, MAR 2014.

[39] Kei Moritsugu, Tohru Terada, and Akinori Kidera. Scalable free energy calculation of proteins via multiscale essential sampling. *Journal of Chemical Physics*, 133(22), DEC 14 2010.

[40] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. Gromacs 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, MAR 2008.

[41] Berk Hess. P-lincs: A parallel linear constraint solver for molecular simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, JAN 2008.

[42] William C Swope, Hans C Andersen, Peter H Berens, and Kent R Wilson. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of chemical physics*, 76(1):637–649, 1982.

[43] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *Journal of Chemical Physics*, 126(1), JAN 7 2007.

[44] Robert T. McGibbon, Kyle A. Beauchamp, Matthew P. Harrigan, Christoph Klein, Jason M. Swails, Carlos X. Hernandez, Christian R. Schwantes, Lee-Ping Wang, Thomas J. Lane, and Vijay S. Pande. Mdtraj: A modern open library for the analysis of molecular dynamics trajectories. *BIOPHYSICAL JOURNAL*, 109(8):1528–1532, OCT 20 2015.

[45] E Dijkstra. Anote on two problems in connection with graphs. *Numer Math*, 1:101–18, 1959.

[46] Inigo Marcos-Alcalde, Javier Setoain, Jesus I. Mendieta-Moreno, Jesus Mendieta, and Paulino Gomez-Puertas. Mepsa: minimum energy pathway analysis for energy landscapes. *BIOINFORMATICS*, 31(23):3853–3855, DEC 1 2015.

[47] Prabir Khatua, Alan J. Ray, and Ulrich H. E. Hansmann. Bifurcated hydrogen bonds and the fold switching of lymphotactin. *Journal of Physcial Chemistry B*, 124(30):6555–6564, JUL 30 2020.

[48] PG Bolhuis, D Chandler, C Dellago, and PL Geissler. Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *ANNUAL REVIEW OF PHYSICAL CHEMISTRY*, 53:291–318, 2002.

[49] C Dellago, PG Bolhuis, FS Csajka, and D Chandler. Transition path sampling and the calculation of rate constants. *Journal of Chemical Physics*, 108(5):1964–1977, FEB 1 1998.

[50] C Dellago, PG Bolhuis, and D Chandler. Efficient transition path sampling: Application to lennard-jones cluster rearrangements. *Journal of Chemical Physics*, 108(22):9236–9245, JUN 8 1998.

[51] E Weinan, WQ Ren, and E Vanden-Eijnden. String method for the study of rare events. *PHYSICAL REVIEW B*, 66(5), AUG 1 2002.

[52] Luca Maragliano, Alexander Fischer, Eric Vanden-Eijnden, and Giovanni Ciccotti. String method in collective variables: Minimum free energy paths and isocommittor surfaces. *Journal of Chemical Physics*, 125(2), JUL 14 2006.

[53] Ilona Christy Unarta, Lizhe Zhu, Carmen Ka Man Tse, Peter Pak-Hang Cheung, Jin Yu, and Xuhui Huang. Molecular mechanisms of rna polymerase ii transcription elongation elucidated by kinetic network models. *Current Opinion in Structural Biology*, 49:54–62, 2018.

[54] Lizhe Zhu, Fu Kit Sheong, Siqin Cao, Song Liu, Ilona C. Unarta, and Xuhui Huang. Taps: A traveling-salesman based automated path searching method for functional conformational changes of biological macromolecules. *Journal of Chemical Physics*, 150(12), MAR 28 2019.

[55] LLC Schrodinger. The pymol molecular graphics system. *Version*, 1:8, 2015.

[56] Wenhua Wang, Prabir Khatua, and Ulrich H. E. Hansmann. Cleavage, downregulation, and aggregation of serum amyloid a. *Journal of Physcial Chemistry B*, 124(6):1009–1019, FEB 13 2020.

[57] Gunilla T Westermark, Ulla Engström, and Per Westermark. The n-terminal segment of protein aa determines its fibrillogenic property. *Biochemical and biophysical research communications*, 182(1):27–33, 1992.

[58] Hussein N. Yassine, Olgica Trenchevska, Huijuan He, Chad R. Borges, Dobrin Nedelkov, Wendy Mack, Naoko Kono, Juraj Koska, Peter D. Reaven, and Randall W. Nelson. Serum amyloid a truncations in type 2 diabetes mellitus. *PLOS ONE*, 10(1), JAN 21 2015.

[59] Wenhua Wang and Ulrich H. E. Hansmann. Stability of human serum amyloid a fibrils. *Journal of Physcial Chemistry B*, 124(47):10708–10717, NOV 25 2020.

[60] Nicholas M. Frame, Meera Kumanan, Thomas E. Wales, Asanga Bandara, Marcus Fandrich, John E. Straub, John R. Engen, and Olga Gursky. Structural basis for lipid binding and function by an evolutionarily conserved protein, serum amyloid a. *JOURNAL OF MOLECULAR BIOLOGY*, 432(7):1978–1995, MAR 27 2020.

[61] Antimicrobial resistance. https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance, 2023. World Health Organization (WHO).

[62] Elizabeth M Darby, Eleftheria Trampari, Pauline Siasat, Maria Solsona Gaya, Ilyas Alav, Mark A Webber, and Jessica MA Blair. Molecular mechanisms of antibiotic resistance revisited. *Nature Reviews Microbiology*, 21(5):280–295, 2023.

[63] U.S. Department of Health and Human Services. National action plan for combating antibiotic-resistant bacteria (carb), 2020-2025. https://www.hhs.gov/sites/default/files/carb-national-action-plan-2020-2025.pdf, 2020.

[64] J Yuan, Z-Z Zhang, X-Z Chen, W Yang, and L-D Huan. Site-directed mutagenesis of the hinge region of nisinz and properties of nisinz mutants. *Applied microbiology and biotechnology*, 64:806–815, 2004.

[65] Jae M Shin, Ji Won Gwak, Pachiyappan Kamarajan, J Christopher Fenno, Alexander H Rickard, and Yvonne L Kapila. Biomedical applications of nisin. *Journal of applied microbiology*, 120(6):1449–1465, 2016.

[66] Shiva Ahmadi, Marzieh Ghollasi, and Hamideh Mahmoodzadeh Hosseini. The apoptotic impact of nisin as a potent bacteriocin on the colon cancer cells. *Microbial pathogenesis*, 111:193–197, 2017.

[67] Yousef Khazaei Monfared, Mohammad Mahmoudian, Claudio Cecone, Fabrizio Caldera, Parvin Zakeri-Milani, Adrián Matencio, and Francesco Trotta. Stabilization and anticancer enhancing activity of the peptide nisin by cyclodextrin-based nanosponges against colon and breast cancer cells. *Polymers*, 14(3):594, 2022.

[68] César Matos Ribeiro da Silva, Waleska da Silva Albuquerque, Jucyara Natállia Araujo de Oliveira, Maria Regina Pires Carneiro, Sona Jain, Silvio Santana Dolabella, and Ana Andréa Teixeira Barbosa. Nisin activity against methicillin-resistant and methicillin-sensitive staphylococcus aureus and risk of resistance acquisition. *Research, Society and Development*, 10(7):e4610716178–e4610716178, 2021.

[69] Adelene Ai-Lian Song, Lionel LA In, Swee Hua Erin Lim, and Raha Abdul Rahim. A review on lactococcus lactis: from food to factory. *Microbial cell factories*, 16:1–15, 2017.

[70] LA375033 Rogers. The inhibiting effect of streptococcus lactis on lactobacillus bulgaricus. *Journal of bacteriology*, 16(5):321–325, 1928.

[71] Des Field, Miguel Fernandez de Ullivarri, R Paul Ross, and Colin Hill. After a century of nisin research-where are we now? *FEMS microbiology reviews*, 47(3):fuad023, 2023.

[72] K Müller-Auffermann, Felipe Grijalva, Fritz Jacob, and Mathias Hutzler. Nisin and its usage in breweries: a review and discussion. *Journal of the Institute of Brewing*, 121(3):309–319, 2015.

[73] Neha Garg, Luis MA Salazar-Ocampo, and Wilfred A Van Der Donk. In vitro activity of the nisin dehydratase nisb. *Proceedings of the national academy of sciences*, 110(18):7258–7263, 2013.

[74] Xu Xia Zhou, Wei Fen Li, Guo Xia Ma, and Yuan Jiang Pan. The nisin-controlled gene expression system: construction, application and improvements. *Biotechnology Advances*, 24(3):285–295, 2006.

[75] Enriqueta Garcia-Gutierrez, Paula M OâĂŹConnor, Gerhard Saalbach, Calum J Walsh, James W Hegarty, Caitriona M Guinane, Melinda J Mayer, Arjan Narbad, and Paul D Cotter. First evidence of production of the lantibiotic nisin p. *Scientific reports*, 10(1):3738, 2020.

[76] Yue Zheng, Yuhui Du, Zekai Qiu, Ziming Liu, Jianjun Qiao, Yanni Li, and Qinggele Caiyin. Nisin variants generated by protein engineering and their properties. *Bioengineering*, 9(6):251, 2022.

[77] Naomi E Kramer, Eddy J Smid, Jan Kok, Ben de Kruijff, Oscar P Kuipers, and Eefjan Breukink. Resistance of gram-positive bacteria to nisin is not determined by lipid ii levels. *FEMS microbiology letters*, 239(1):157–161, 2004.

[78] Ben de Kruijff, Vincent van Dam, and Eefjan Breukink. Lipid ii: a central component in bacterial cell wall synthesis and a target for antibiotics. *Prostaglandins, Leukotrienes and Essential Fatty Acids*, 79(3-5):117–121, 2008.

[79] Benjamin Schwartz, Jay A Markwalder, and Yi Wang. Lipid ii: Total synthesis of the bacterial cell wall precursor and utilization as a substrate for glycosyltransfer and transpeptidation by penicillin binding protein (pbp) 1b of eschericia c oli. *Journal of the American Chemical Society*, 123(47):11638–11643, 2001.

[80] Miriam Wilmes, Bruno PA Cammue, Hans-Georg Sahl, and Karin Thevissen. Antibiotic activities of host defense peptides: more to it than lipid bilayer perturbation. *Natural product reports*, 28(8):1350–1358, 2011.

[81] Jean van Heijenoort. Lipid intermediates in the biosynthesis of bacterial peptidoglycan. *Microbiology and Molecular Biology Reviews*, 71(4):620–635, 2007.

[82] Sujeet Kumar, Aurelio Mollo, Daniel Kahne, and Natividad Ruiz. The bacterial cell wall: from lipid ii flipping to polymerization. *Chemical reviews*, 122(9):8884–8910, 2022.

[83] Hester E Hasper, Naomi E Kramer, James L Smith, JD Hillman, Cherian Zachariah, Oscar P Kuipers, Ben De Kruijff, and Eefjan Breukink. An alternative bactericidal mechanism of action for lantibiotic peptides that target lipid ii. *Science*, 313(5793):1636–1637, 2006.

[84] Yuan Qiao, Veerasak Srisuknimit, Frederick Rubino, Kaitlin Schaefer, Natividad Ruiz, Suzanne Walker, and Daniel Kahne. Lipid ii overproduction allows direct assay of transpeptidase inhibition by β-lactams. *Nature chemical biology*, 13(7):793–798, 2017.

[85] Ka Teng Chan, Xin Song, Leyao Shen, Nian Liu, Xuedong Zhou, Lei Cheng, and Jing Chen. Nisin and its application in oral diseases. *Journal of Functional Foods*, 105:105559, 2023.

[86] Hester Emilie van Heusden, Ben De Kruijff, and Eefjan Breukink. Lipid ii induces a transmembrane orientation of the pore-forming peptide lantibiotic nisin. *Biochemistry*, 41(40):12171–12178, 2002.

[87] Hester Emilie Hasper, Ben de Kruijff, and Eefjan Breukink. Assembly and stability of nisin- lipid ii pores. *Biochemistry*, 43(36):11567–11575, 2004.

[88] Imke Wiedemann, Roland Benz, and Hans-Georg Sahl. Lipid ii-mediated pore formation by the peptide antibiotic nisin: a black lipid membrane study. *Journal of Bacteriology*, 186(10):3259–3261, 2004.

[89] Sunhwan Jo, Joseph B Lim, Jeffery B Klauda, and Wonpil Im. Charmm-gui membrane builder for mixed bilayers and its application to yeast membranes. *Biophysical journal*, 97(1):50–58, 2009.

[90] Anton Chugunov, Darya Pyrkova, Dmitry Nolde, Anton Polyansky, Vladimir Pentkovsky, and Roman Efremov. Lipid-ii forms potential âĂIJlanding terrainâĂİ for lantibiotics in simulated bacterial membrane. *Scientific reports*, 3(1):1–11, 2013.

[91] Florian Prossnigg, Andrea Hickel, Georg Pabst, and Karl Lohner. Packing behaviour of two predominant anionic phospholipids of bacterial cytoplasmic membranes. *Biophysical chemistry*, 150(1-3):129–135, 2010.

[92] Sunhwan Jo, Taehoon Kim, Vidyashankara G Iyer, and Wonpil Im. Charmm-gui: a web-based graphical user interface for charmm. *Journal of computational chemistry*, 29(11):1859–1865, 2008.

[93] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. Charmm: the biomolecular simulation program. *Journal of computational chemistry*, 30(10):1545–1614, 2009.

[94] Jumin Lee, Xi Cheng, Sunhwan Jo, Alexander D MacKerell, Jeffery B Klauda, and Wonpil Im. Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Biophysical journal*, 110(3):641a, 2016.

[95] João Medeiros-Silva, Shehrazade Jekhmane, Alessandra Lucini Paioni, Katarzyna Gawarecka, Marc Baldus, Ewa Swiezewska, Eefjan Breukink, and Markus Weingarth. High-resolution nmr studies of antibiotics in cellular membranes. *Nature communications*, 9(1):1–10, 2018.

[96] DN Ganchev, HE Hasper, E Breukink, and B de Kruijff. Size and orientation of the lipid ii headgroup as revealed by afm imaging. *Biochemistry*, 45(19):6195–6202, 2006.

[97] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C Smith, Berk Hess, and Erik Lindahl. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1:19–25, 2015.

[98] Robert B Best, Xiao Zhu, Jihyun Shim, Pedro EM Lopes, Jeetain Mittal, Michael Feig, and Alexander D MacKerell Jr. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone ϕ, ψ and side-chain χ1 and χ2 dihedral angles. *Journal of Chemical Theory and Computation*, 8(9):3257–3273, 2012.

[99] Tanja Schneider and Hans-Georg Sahl. An oldie but a goodie–cell wall biosynthesis as antibiotic target pathway. *International Journal of Medical Microbiology*, 300(2-3):161–169, 2010.

[100] Shang-Te Hsu, Eefjan Breukink, Ben de Kruijff, Robert Kaptein, Alexandre MJJ Bonvin, and Nico AJ van Nuland. Mapping the targeted membrane pore formation mechanism by solution nmr: The nisin z and lipid ii interaction in sds micelles. *Biochemistry*, 41(24):7670–7676, 2002.

[101] Eefjan Breukink and Ben de Kruijff. Lipid ii as a target for antibiotics. *Nature reviews Drug discovery*, 5(4):321–323, 2006.

[102] ZhiGuang Jia, Megan L O'Mara, Johannes Zuegg, Matthew A Cooper, and Alan E Mark. The effect of environment on the recognition and binding of vancomycin to native and resistant forms of lipid ii. *Biophysical journal*, 101(11):2684–2692, 2011.

[103] Olgun Guvench, Sairam S Mallajosyula, E Prabhu Raman, Elizabeth Hatcher, Kenno Vanommeslaeghe, Theresa J Foster, Francis W Jamison, and Alexander D MacKerell Jr. Charmm additive all-atom force field for carbohydrate derivatives and its utility in polysaccharide and carbohydrate–protein modeling. *Journal of Chemical Theory and Computation*, 7(10):3162–3180, 2011.

[104] Shuichi Nosé and ML Klein. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 1983.

[105] William G Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.

[106] Michele Parrinello and Aneesur Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics*, 52(12):7182–7190, 1981.

[107] Shuichi Miyamoto and Peter A Kollman. Settle: An analytical version of the shake and rattle algorithm for rigid water models. *Journal of computational chemistry*, 13(8):952–962, 1992.

[108] Berk Hess, Henk Bekker, Herman JC Berendsen, and Johannes GEM Fraaije. Lincs: A linear constraint solver for molecular simulations. *Journal of computational chemistry*, 18(12):1463–1472, 1997.

[109] William Humphrey, Andrew Dalke, and Klaus Schulten. Vmd: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

[110] Richard J Gowers, Max Linke, Jonathan Barnoud, Tyler JE Reddy, Manuel N Melo, Sean L Seyler, Jan Domanski, David L Dotson, Sébastien Buchoux, Ian M Kenney, et al. Mdanalysis: a python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th python in science conference*, volume 98, page 105. SciPy Austin, TX, 2016.

[111] Naveen Michaud-Agrawal, Elizabeth J Denning, Thomas B Woolf, and Oliver Beckstein. Mdanalysis: a toolkit for the analysis of molecular dynamics simulations. *Journal of computational chemistry*, 32(10):2319–2327, 2011.

[112] Matthew Carr and Cait E MacPhee. Membrainy: a âĂŸsmartâĂŹ, unified membrane analysis tool. *Source code for biology and medicine*, 10(1):1–10, 2015.

[113] Imke Wiedemann, Eefjan Breukink, Cindy Van Kraaij, Oscar P Kuipers, Gabriele Bierbaum, Ben De Kruijff, and Hans-Georg Sahl. Specific binding of nisin to the peptidoglycan precursor lipid ii combines pore formation and inhibition of cell wall biosynthesis for potent antibiotic activity. *Journal of Biological Chemistry*, 276(3):1772–1779, 2001.

[114] Evelyn M Molloy, Des Field, Paula M O' Connor, Paul D Cotter, Colin Hill, and R Paul Ross. Saturation mutagenesis of lysine 12 leads to the identification of derivatives of nisin a with enhanced antimicrobial activity. *PloS one*, 8(3):e58530, 2013.

[115] Rhythm Shukla, Francesca Lavore, Sourav Maity, Maik GN Derks, Chelsea R Jones, Bram JA Vermeulen, Adéla Melcrová, Michael A Morris, Lea Marie Becker, Xiaoqi Wang, et al. Teixobactin kills bacteria by a two-pronged attack on the cell envelope. *Nature*, 608(7922):390–396, 2022.

[116] Sarah Witzke, Michael Petersen, Timothy S Carpenter, and Syma Khalid. Molecular dynamics simulations reveal the conformational flexibility of lipid ii and its loose association with the defensin plectasin in the staphylococcus aureus membrane. *Biochemistry*, 55(23):3303–3314, 2016.

[117] Lilian M Were, Barry D Bruce, P Michael Davidson, and Jochen Weiss. Size, stability, and entrapment efficiency of phospholipid nanocapsules containing polypeptide antimicrobials. *Journal of agricultural and food chemistry*, 51(27):8073–8079, 2003.

[118] Houri Sadri, Mahmoud Aghaei, and Vajihe Akbari. Nisin induces apoptosis in cervical cancer cells via reactive oxygen species generation and mitochondrial membrane potential changes. *Biochemistry and Cell Biology*, 100(2):136–141, 2022.

[119] Angélique Lewies, Johannes Frederik Wentzel, Hayley Christy Miller, and Lissinda Hester Du Plessis. The antimicrobial peptide nisin z induces selective toxicity and apoptotic cell death in cultured melanoma cells. *Biochimie*, 144:28–40, 2018.

[120] Suyash M Patil and Nitesh K Kunda. Nisin zp, an antimicrobial peptide, induces cell death and inhibits non-small cell lung cancer (nsclc) progression in vitro in 2d and 3d cell culture. *Pharmaceutical Research*, 39(11):2859–2870, 2022.

[121] Pachiyappan Kamarajan, Takayuki Hayami, Bibiana Matte, Yang Liu, Theodora Danciu, Ayyalusamy Ramamoorthy, Francis Worden, Sunil Kapila, and Yvonne Kapila. Nisin zp, a bacteriocin and food preservative, inhibits head and neck cancer tumorigenesis and prolongs survival. *PloS one*, 10(7):e0131008, 2015.

[122] Kayla Jaye, Dennis Chang, Chun Guang Li, and Deep Jyoti Bhuyan. Gut metabolites and breast cancer: the continuum of dysbiosis, breast cancer risk, and potential breast cancer therapy. *International Journal of Molecular Sciences*, 23(16):9490, 2022.

[123] Zohreh Norouzi, Ali Salimi, Raheleh Halabian, and Hossein Fahimi. Nisin, a potent bacteriocin and anti-bacterial peptide, attenuates expression of metastatic genes in colorectal cancer cell lines. *Microbial pathogenesis*, 123:183–189, 2018.

[124] Pelin Balcik-Ercin and Belgin Sever. An investigation of bacteriocin nisin anti-cancer effects and fzd7 protein interactions in liver cancer cells. *Chemico-Biological Interactions*, 366:110152, 2022.

[125] Noha El-Sayed Ibrahim, Heba Morsy, and Marwa Abdelgwad. The comparative effect of nisin and thioridazine as potential anticancer agents on hepatocellular carcinoma. *Reports of biochemistry & molecular biology*, 9(4):452, 2021.

[126] Timothée Rivel, Christophe Ramseyer, and Semen Yesylevskyy. The asymmetry of plasma membranes and their cholesterol content influence the uptake of cisplatin. *Scientific reports*, 9(1):5627, 2019.